The HMGB1-CXCL12 heterocomplex and its interaction with CXCR4 as a target for tissue repair



Álvaro Viñals Guitart
Saint Hilda's College

In partial fulfilment of the requirements for the title of DPhil in Molecular and Cellular Medicine

Supervised by Prof. Jagdeep Nanchahal, Dr. Nicola Burgess-Brown, Dr. Katharina Duerr

Acknowledgements

“Well, the project it´s like science fiction! And I get to play with fancy machines!”

Despite this enthusiastic, but not exactly realistic, motivation, Professor Jagdeep Nanchahal gave me the opportunity of joining his lab for a DPhil in tissue repair. I took the scenic route to this place: majoring in Biochemistry, then three years as a research assistant in labs working in biomaterials and tissue engineering. And now I needed to become a structural biologist, animal surgeon, AND understand the medical aspect of the project? Jagdeep nevertheless trusted me as capable to do so and has ensured throughout this DPhil that I didn´t lose the focus in the important issues, or the grip whatever passed for my sanity. He has also gone above and beyond a supervisor´s call: not only did he help me with my scientific issues, but with my personal and mental issues, including when I almost gave up in more than one sense. He also acknowledged from the get-go that I would need to learn skills outside his expertise which I didn´t even know I lacked and ensured that I came directly into the proper environment to learn.

Enter the Structural Genomics Consortium: Dr. Nicola Burgess Brown (Biotechnology group) and Dr. Katharina Düerr (IMP2 group) were my direct co-supervisors and taught me from scratch how to coax pesky proteins into existence despite their best efforts in the contrary. In addition, and despite not being acknowledged in the paperwork as such, Dr. Wyatt Yue has effectively acted as a fourth co-supervisor and it is thanks to him that I was able to turn a hodgepodge collection of biophysics experiments into a coherent body of work.

I would not be submitting this work either without having the help of friends. In the Nanchahal Group, Doctor Ana Isabel Espírito Santo has been like a left leg in this project; from teaching me a good chunk of the biology techniques I would need to learn, to the *fun* evenings in the animal unit at the Kennedy, and the occasional whack in the head when I was about to mess up badly. Carl Lee, my fellow DPhil student (you are next, remember!) ensured that I did not utterly eradicate any cells I worked with thanks to his status as our resident Stem Cell God. I would also like to thank Dr. Geoffrey Lee, my predecessor as DPhil student, without whom I would not have a project to work on! Special thanks also must go to DPhil student Anna Marzęda in the Kennedy for being the resident immunologist and as the third member of our office´s Kennedy Misery Trio (as Polish Misery, with me being Spanish Misery and Ana as Portuguese Misery).

I spent a substantial amount of time at the SGC, adopted by the Biotechnology and IMP2 groups. I am thankful in particular to Tina Bohstedt in the IMP group for being able to set up a CXCR4 expression system, and for being my fellow dark humour partner at very late lab nights where sanity ebbs, unholy entities sabotage our experiments, and we dread existence. I wouldn´t have made it through those on my own. I also want to thank Dr. Patrizia Abrusci and Cassandra Dix, who have also assisted in this project and are now in the next steps of their respective careers. In Biotechnology, special thanks go to Dr. Alejandra Fernández-Cid Fernández-Canteli has been a huge well of knowledge; Dr. Rod Chalk and Dr. Tiago Moreira taught me all I needed to know and more about mass spectrometry, and DPhil student Aleksandra Szykowska (also known as Scary Aleks) who showed me how to use FPLC systems without destroying them. I also want to acknowledge the whole of the SGC, as an open-source institute where knowledge is shared freely, and which I was able to access and adapt for my own purposes, such as several of the vectors present in the SGC repositories. Dr Michael Fairhead developed the high-throughput large scale protein production protocols I adapted; DPhil student Jesse Coker (MOB group) helped me with the validation of my disulfide quantitation methods, and together with Dr. Daniel Pinkas helped me better understand kinetic analyses.

I also thank my parents for their unconditional support in this crazy idea of being a scientist very far away from them with all the emotional cost it took. Finally, I also must thank Isis Cortés Botella, my partner and bride, without whom I would likely not be alive to write this after four years of DPhil. She has almost saintly patience with my near-incessant rambling, loud cursing at data, and very geek day-to-day habits; somehow managed to turn me into a better person throughout this process.

During this project, I tried to keep close to heart the words of the immortal Douglas Adams in his *Hitchhiker Guide to the Galaxy* series:

Don´t panic.

(The following was more representative of reality; I must have misplaced my towel).

Can I panic now, *please*?

Álvaro Viñals Guitart, Oxford, 18th Oct 2020

Abstract

Reduced High Mobility Group Box 1 (HMGB1) protein binds to CXC Ligand 12 (CXCL12), signals through CXC Receptor 4 (CXCR4) to promote tissue regeneration and accelerates repair by transitioning stem and progenitor cells to G_{Alert}. Therefore, administration of fully reduced HMGB1 (FR-HMGB1) could prove beneficial in clinical practice. However, local conversion of FR-HMGB1 to the disulfide form (DS-HMGB1) may result in deleterious inflammation through signalling via Toll-Like Receptors 2 and 4, and the Receptor for Advanced Glycation End Products (RAGE).

Engineering HMGB1 to eliminate these potentially deleterious properties requires knowledge on the mechanism of its interaction with CXCL12, for which there is sparse literature. I identified the residues involved in formation of the heterocomplex using a combination of peptide arrays, biolayer interferometry and nuclear magnetic resonance. With these, and the published literature on the interaction of HMGB1 with its proinflammatory receptors, I designed a panel of HMGB1 derivatives aimed at removing the deleterious activities. I identified a construct comprising two HMG B Boxes in tandem (DbB-HMGB1) which is unable to signal through either TLR-2 or TLR-4 and has greatly decreased RAGE binding but retains regenerative activity equivalent to FR-HMGB1 *in vivo* and *in vitro.*

Table of contents

List of figures v
List of tables vii
List of abbreviations and acronyms ix
Chapter 1 Introduction 11
1.1 Incomplete tissue repair after injuries leads to loss of function 12
1.2 High-Mobility Group Box Protein 1 (HMGB1) 14
1.3 The CXCL12-CXCR4 axis 27
1.4 HMGB1 breaks the concentration-dependence of CXCL12 signalling 31
1.5 Aim: improving HMGB1 as a tissue repair therapeutic 33
1.6 Specific objectives 34
Chapter 2 Materials and Methods 35
2.1 Nomenclature and notes 35
2.2 Statistical analysis 36
2.3 In-house enzyme production 38
2.4 Protein purification resins 38
2.5 *E. coli* culture media preparation 39
2.6 Electrophoresis 40
2.7 Mass spectrometry (MS) 40
2.8 Cloning and test expression from *E. coli* 43
2.9 Test purification from *E. coli* 48
2.10 Large scale expression from *E. coli* 48
2.11 Purification of HMGB1 from *E. coli* cell pellets 49
2.12 Purification of CXCL12 from *E. coli* periplasm 51
2.13 Endotoxin-free protein production 52
2.14 Differential scanning fluorimetry (DSF) 55
2.15 Peptide arrays (SPOT array) 56
2.16 Crystallography (hanging drop) 56
2.17 Biolayer interferometry (BLI) 57
2.18 Nuclear magnetic resonance (NMR) 58
2.19 *In vivo* muscle injury model 59
2.20 *In vitro* transwell migration assay [31] 60
2.21 Gene design and ordering 60
2.22 HMGB1 *in vitro* H9 differentiation assay 61
2.23 C2C12 differentiation assay 62
2.24 TLR-4 and TLR-2 HEK reporter assay 63
2.25 HMGB1 RAGE binding enzyme-linked immunosorbent assay (ELISA) 63
2.26 Experiments done by collaborators 64
Chapter 3 Purification of recombinant human HMGB1 and CXCL12 67
3.1 Chapter introduction: design considerations 68
3.2 Generation of entry clones 71
3.3 Disulfide bond detection 74
3.4 Expression and purification of HMGB1 77
3.5 Expression, and purification of CXCL12. 112
3.6 Final remarks 131

Chapter 4 The interaction between CXCL12 and HMGB1 133
4.1 Chapter introduction 134
4.2 Biology of HMGB1 as a potential therapeutic 135
4.3 Characterisation of the heterocomplex 138
4.4 Sequence-specificity of the HMGB1-CXCL12 interaction 146
4.5 Conclusion: defined, homologous regions within each HMG Box domain bind CXCL12. 162
Chapter 5 The HMGB1 construct panel 167
5.1 Chapter introduction 168
5.2 Data used in construct design 169
5.3 HMGB1-C1 construct panel 175
5.4 RAGE-binding screen 184
5.5 TLR-4 and TLR-2 activity screen 193
5.6 Regenerative effect of Double box B (DbB) construct -HMGB1A-C1-c048 196
5.7 Stability of DbB-HMGB1 200
5.8 Conclusions 211
Chapter 6 Future work and conclusions 213
Chapter 7 Bibliography 219
Chapter 8 Appendices 233
Appendix A1 Supplemental Methods ..... - 3 -
Chapter 2 Reagents and primers ..... - 3 -
Chapter 3 Designed constructs and sequences ..... - 9 -
Chapter 4 ..... - 27 -
Chapter 5 Sequences ..... - 27 -
Appendix A2 Mass spectrometry data ..... - 51 -
Chapter 3 MS data ..... - 51 -
Chapter 4 MS data ..... - 69 -
Chapter 5 MS data ..... - 70 -
Appendix A3 Supplemental data ..... - 81 -
Chapter 3 Figure supplements ..... - 81 -
Chapter 4 Figure supplements ..... - 86 -
Appendix A4 NMR tables ..... - 92 -
Chapter 4 NMR spectra (Figure 4.8/Figure 4.9) ..... - 92 -
Appendix A5 Statistics ..... - 124 -
Chapter 3 Data analysis ..... - 124 -
Chapter 4 Data analysis ..... - 132 -
Chapter 5 Data analysis ..... - 137 -

List of figures

Figure 1.1: Conservation of HMGB1 sequences across different vertebrates 15
Figure 1.2: Structure of the HMG Box domains in NHP6A (blue, PDB 1CG7) and HMGB1 (grey, PDB 2YRQ) 17
Figure 1.3: Interactions of HMGB1 18
Figure 1.4: CXCL12 monomer structure (PDB 2kec) 27
Figure 1.5: Structure of CXCR4 (PDB 3ODU) 28
Figure 2.1: Elements of an SGC identifier 35
Figure 2.2: Statistical test selection criteria 37
Figure 2.3: Correlation between m/z and surface area 42
Figure 2.4: Calculation of safe pyrogen levels for recombinant protein injection into mice 52
Figure 2.5: Identification of PEG and Triton contamination in protein samples after LPS removal 54
Figure 3.1: Disulfide bridges of HMGB1 and CXCL12 71
Figure 3.2: Mutagenesis of HMGB1A-s002 72
Figure 3.3: ESI-QTOF/MS reduction state measurements 76
Figure 3.4: Initial HMGB1 constructs 78
Figure 3.5: Test expression of initial HMGB1 panel 79
Figure 3.6: Removal of the Gly1-Pro8 peptide results in loss of enhanced tissue repair and chemotaxis 80
Figure 3.7: DNA contamination of HMGB1 after size exclusion chromatography (SEC). 81
Figure 3.8: 1 M NaCl purification of HMGB1A-c001 82
Figure 3.9 : DNA contamination of high-throughput HMGB1 preparations 84
Figure 3.11: Biotin-affinity purification of HMGB1 constructs 85
Figure 3.10: Vector map of pNIC-Bio3 85
Figure 3.12: PTM analysis of FR-HMGB1 constructs (initial panel) 86
Figure 3.13: Deficient N-terminal TEV cleavage in constructs with proline in P3´ 87
Figure 3.14: "Locking" region in both Box A and Box B 89
Figure 3.15: Biological activity of HMGB1A-c031 (1-164) 90
Figure 3.16: SEC profile of different HMGB1 constructs 92
Figure 3.17: Surface representations of hydrophobicity and polarity in HMGB1 94
Figure 3.18: Native ESI/QTOF-MS of HMGB1 in 0.2 M ammonium acetate, pH 6.8 96
Figure 3.19: Native ESI/QTOF-MS of HMGB1 in 50 mM ammonium acetate, pH 6.8) 98
Figure 3.20: Denaturing LC-ESI-QTOF/MS of HMGB1 in 0.1% formic acid, pH 2.2 100
Figure 3.21: NMR conformers of HMGB1 in PDB 2YRQ 103
Figure 3.22: Hydrophobicity colouring of HMGB1 1-164 (PDB 2YRQ) 105
Figure 3.23: DSF melting curve profiles and Tm_{50} analysis of HMGB1A constructs 108
Figure 3.24: Initial CXCL12 constructs and test expression 112
Figure 3.25: Expression of CXCL12 in disulfide-competent strains 114
Figure 3.26: Periplasmic export of CXCL12A-c003 and c006 with N-terminal SUMO tag fusion 115

Figure 3.27: Time optimisation of CXCL12A-c027 (TAT export) expression and 24h test expression of CXCL12A-c021/25 (pDsbC fusions) 117
Figure 3.28: Degradation sub-products of CXCL12 constructs 119
Figure 3.29: Optimisation of CXCL12 expression in pDsbC constructs (CXCL12A-c021) 121
Figure 3.30: Comparison of cell lysis methods for CXCL12A-c021 purification 123
Figure 3.31: SEC chromatography of CXCL12 124
Figure 3.32: Hydrophobicity colouring of CXCL12 125
Figure 3.33: ESI-QTOF/MS and SEC of CXCL12A-c021 (cut) 127
Figure 3.34: Interaction of CXCL12A-c021 with CXCR4-GFP 129
Figure 4.1: Optimal parameters for HMGB1 treatment in $BaCl_2$ muscle injury model 135
Figure 4.2: Effect of CXCL12 co-administration in the regenerative effects of HMGB1 137
Figure 4.3: SEC chromatography of the HMGB1-CXCL12 heterocomplex 139
Figure 4.4: Native ESI-QTOF/MS of HMGB1A-c001 (FR, full length), CXCL12A-c021 (wild type) and heterocomplex, in 0.2 M ammonium acetate, pH 6.8 141
Figure 4.5: BLI kinetics of CXCL12 binding to full length FR and 3S-HMGB1 144
Figure 4.6: Alanine scan of HMGB1 peptides interacting with CXCL12 across Box A and Box B 148
Figure 4.7: Peptide array of CXCL12 peptides interacting with HMGB1 149
Figure 4.8: CSP and intensity change in HMGB1 Box B constructs with and without flanking regions upon CXCL12 binding 152
Figure 4.9: Figure 4.9: CSP changes in HMGB1 1-184 (3S) upon CXCL12 binding 156
Figure 4.10: Summary of peptide array and NMR data 157
Figure 4.11: Interferograms in BLI of HMGB1 constructs' binding to CXCL12 159
Figure 4.12: Kinetics of CXCL12 binding to different HMGB1 constructs (Figure 4.11) 161
Figure 5.1: Alignment of RAGE binding sites in HMGB1 Box A (C22-K49), Box B (K149-K182), and RAP peptides from S100A/S100A4 174
Figure 5.2: Types of constructs in the HMGB1 construct panel (HMGB1-C1) 176
Figure 5.3: Potential TLR-4 binding hydrophobic pocket in Box B 177
Figure 5.4: Schematic of HMGB1-C1-c048 178
Figure 5.5: Test expression of HMGB1 construct panel C1 181
Figure 5.6: Large scale purification of HMGB1-C1 constructs 183
Figure 5.7: Adaptation of RAGE ELISA to 384-well format 185
Figure 5.8: RAGE-Fc binding in ELISA of HMGB1-FL (FR, DS and 3S) and full HMG Box constructs 185
Figure 5.9: RAGE-Fc binding in ELISA of HMGB1 linker and RAGE-site mutated constructs 186
Figure 5.10: RAGE-Fc binding in ELISA to tail-truncations of HMGB1 and Double Box B HMGB1 187
Figure 5.11: Interferograms in BLI of HMGB1 constructs' binding to RAGE 189
Figure 5.12: BLI validation of RAGE binding to HMGB1 constructs 190
Figure 5.13: TLR signalling in response to recombinant FR-HMGB1-C1-c001, HMGB1-C1-c048 and commercial DS-HMGB1 194

Figure 5.14: Response of C2C12 differentiation to HMGB1 **196**
Figure 5.15: H9 differentiation assay with FR-HMGB1-C1-c001 (1-214) and DbB-HMGB1A-C1-c048 **197**
Figure 5.16: Time course of muscle cell regeneration induced by FR-HMGB1A-c001 and DbB-HMGB1 (HMGB1A-C1-c048) **199**
Figure 5.17: SEC of HMGB1 constructs in comparison with DbB-HMGB1 **201**
Figure 5.18: Native ESI-QTOF/MS SASA plots for HMGB1 constructs and DbB-HMGB1 in 0.2 M ammonium acetate, pH 6.8 (Figure 5.19) **203**
Figure 5.19: Denaturing ESI-QTOF/MS SASA plots of HMGB1 constructs and DbB-HMGB1 in 0.1% formic acid, pH 2.2 **204**
Figure 5.20: DSF melting curve profiles and Tm_{50} (°C) of DbB HMGB1-C1-c048 and comparison with FR-HMGB1 1-214 **206**
Figure 5.21: Solution stability of FR-HMGB1 FL and DbB **209**
Figure 5.22: Time-of-flight mass spectrometry (ESI/QTOF-MS) of the samples analysed in Figure 5.21 **210**
Figure 8.1: Function test of Superdex 75 Increase, 10/300 **- 81 -**
Figure 8.2: Conductivity traces and sample exclusion criteria **- 82 -**
Figure 8.3: Nonreducing denaturing SDS-PAGE of HMGB1 constructs in pH 6.8 ammonium acetate (Figure 3.19) **- 83 -**
Figure 8.4: Preparative SEC of HMGB1 constructs c001, c002, c031, c006 (no TCEP) **- 84 -**
Figure 8.5: DSF melting curve profiles and Tm_{50} analysis of HMGB1A constructs **- 85 -**

List of tables

Table 2.1: Enzymes and cell lines used in cloning **43**
Table 2.2: Purification tags and uses of SGC *E. coli* expression vectors present in this work **44**
Table 2.3: Contents of custom DSF screen **55**
Table 3.1: Summary of retention volumes from HMGB1 constructs in Superdex 75 Increase, 10/300 **93**
Table 3.2: SASA calculations from ESI/QTOF-MS experiments **101**
Table 3.3: Summary Tm_{50} of HMGB1 constructs with different and redox states **106**
Table 5.1: Retention volumes for FR-HMGB1 constructs and DbB-HMGB1 **200**
Table 5.2: SASA calculations from ESI/QTOF-MS experiments for DbB- and FR-HMGB1 constructs **202**
Table 8.1: List of primers used in generation of PCR products (alleles) **- 7 -**
Table 8.2: Primer and vector combinations for allele and construct generation **- 8 -**
Table 8.3: Designed HMGB1 constructs for characterisation of the HMGB1-CXCL12 interaction **- 12 -**
Table 8.4: Designed CXCL12 constructs for characterisation of the HMGB1-CXCL12 interaction. **- 24 -**
Table 8.5: HMGB1-C1 construct panel **- 28 -**

List of abbreviations and acronyms

Acronym	Definition
AHC	Anti IgG Capture Biosensor
AICc	Aikaike´s Information Criterion (Corrected)
AIM-TB	Autoinduction Medium Terrific Broth
ANOVA	ANalysis Of VAriance
ASPA	Animal Scientific Procedures Act
BB	Base buffer
BLI	Biolayer Interferometry
BMRB	Biological Magnetic Resonance Bank
BSA	Bovine Serum Albumin
CD14	Cluster of differentiation 14
Cryo-EM	Cryo-Electron Microscopy
CSA	Cross Sectional Area
CSP	Chemical Shift Perturbation
CV	Column Volume
CXCL12	CXC Chemokine Ligand 12 (also known as SDF-1)
CXCR4	CXC Chemokine Receptor 4
DAMP	Damage Associated Molecular Pattern
DBB-HMGB1	Double Box B HMGB1
DDM	N-Dodecyl-Beta-D-Maltoside
DMEM	Dubecco´s Modified Eagle Medium
DSF	Differential Scanning Fluorimetry
DS-HMGB1	Disulfide HMGB1
DTT	Dithiothreitol
DVT	Deep Venous Thrombosis
EB	Elution Buffer
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray Ionisation
EU	Endotoxin Activity Units
FA	Formic Acid
FL-HMGB1	Full-Length HMGB1
FPLC	Fast Pressure Liquid Chromatography
FR-HMGB1	Fully Reduced HMGB1
GPCR	G-Protein Coupled Receptor
HEK-293	Human Embryonic Kidney Cells 293
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HI-FBS	(Heat Inactivated) Foetal Bovine Serum
HIV	Human Immunodeficiency Virus
HMG(XY)	High Mobility Group (Protein), e.g. HMGB1 = B1, HMGN = N, HMGA=A)
HPLC	High Pressure Liquid Chromatography
HRP	Horse Radish Peroxidase
HSQC	Heteronuclear Single Quantum Correlation experiment
IA	Iodoacetamide
IDR	Intrinsically Disordered Region
IMAC	Ion-Metal Affinity Chromatography
IPTG	Isopropyl β-d-1-thiogalactopyranoside
KB	Kinetics buffer
K_d	Equilibrium dissociation constant (M)
KO	Knock-Out
K_{off}	Dissociation rate constant (1/s)
k_{on}	Association rate constant (M/s)
LAL	*Limulus* amoebocyte lysate
LB	Luria-Bertani (medium/agar)
LBP	LPS binding protein
LC	Liquid Chromatography
LDS	Lauryl Dodecyl Sulphate
LIC	Ligation Independent Cloning
LPS	Lipopolysaccharide (endotoxin)
LTA	Lipoteichoic acid
MAPK	Mitogen Activated Protein Kinase
MD-2	Myeloid Differentiation Factor 2 (also known as LY96)
MES	2-ethanesulfonic acid
MGC	Mammalian Gene Collection
MI	Myocardial Infarction
MOPS	3-(N-morpholino)propanesulfonic acid
MS	Mass spectrometry
MS/MS	Tandem Mass Spectrometry
MYD-88	Myeloid Differentiation Factor 88

Acronym	Definition
NACWO	Named Animal Care and Welfare Officer
NCBI	National Centre for Biotechnology Information
NEM	N-Ethyl Maleimide
NET	Neutrophil Extracellular Trap
NHS	Esther reactive group for amine crosslinking
NLLS	Nonlinear Least Squares
NLS	Nuclear Localisation Signal
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
NVS	Named Veterinary Surgeon
o.n.	Over-Night
OD_X	Optical density (absorbance unit) at wavelength X
PAGE	Poly-Acrylamide Gel Electrophoresis
PAMP	Pathogen Associated Molecular Pattern
PBS (T)	Phosphate Buffer Saline (Tween)
PCR	Polymerase Chain Reaction
PDI	Protein Disulfide Isomerase
PEG	Polyethylene Glycol
PEI	Polyethyleneimine
PIL	Personal License (under ASPA 1996 Regulations)
PTM	Post Translational Modification
QQ	Quantile-Quantile
QTOF	Quadrupole Time of Flight (mass spectrometry)
RAGE	Receptor for Advanced Glycation End-products (AGE)
RAP	RAGE Antagonist Peptide
RCF (*xg)*	Relative Centrifugal Force
REML	Restricted Maximum Likelihood
RFU	Relative Fluorescence Units
RMSD	Root Mean Square Deviation (of atomic positions)
ROCK	Rho Associated Protein Kinase
RPM	Revolutions Per Minute
RT	Room Temperature
RV	Retention Volume
SASA	Solvent Accessible Surface Area
SD	Standard Deviation (of the Mean)
SDF-1	Stromal-Derived Factor 1 (see: CXCL12)
SDS	Sodium Dodecyl Sulphate
SEAP	Secreted Embryonic Alkaline Phosphatase
SEC	Size Exclusion Chromatography
Sec (italics)	Secretory pathway for unfolded peptides
SEM	Standard Error (of the Mean)
SGC	Structural Genomics Consortium
SNP	Single Nucleotide Polymorphism
SOB	Super Optimal Broth
SOC	Super Optimal Broth with Catabolite repression
SPG	Succinic Acid, Sodium Dihydrogen Phosphate and Glycine buffer
SPR	Surface Plasmon Resonance
SUMO	Small Ubiquitin-like Modifier
TA	Tibialis Anterioris muscle
TAE	Tris Acetate EDTA buffer
TAT	Twin Arginine Transporter
TB	Terrific Broth
TCEP	Tris(2-carboxyethyl)phosphine
TEV	Tobacco Etch Virus
TGF	Transforming Growth Factor
TIC	Total Ion Count chromatogram
TLR-(X)	Toll-Like Receptor (X)
TMB	3,3′,5,5′-Tetramethylbenzidine
TNF	Tumour Necrosis Factor
TOCSY	Total Correlated Spectroscopy
TOF	Time-Of-Flight
TROSY	Transverse Relaxation Optimised Spectroscopy
UBMTA	Uniform Biological Material Transfer Agreement
Ulp-1	Ubl-specific protease 1 (SUMO protease)
YFP	Yellow Fluorescent Protein

Chapter 1 Introduction

1.1 Incomplete tissue repair after injuries leads to loss of function 12
1.2 High-Mobility Group Box Protein 1 (HMGB1) 14
1.2.I Characteristics of the high-mobility group (HMG) family 16
1.2.II Structure and functions of the HMG Boxes within HMGB1 17
1.2.III Subcellular location and PTMs: the functional redox switch in HMGB1 19
1.2.IV DS-HMGB1 is both a damage- (alarmin) and pathogen-associated molecular pattern 22
1.2.V FR-HMGB1 is not only a chemotactic mediator but a strong pro-regenerative molecule 25
1.3 The CXCL12-CXCR4 axis 27
1.3.I Structures of CXCL12 and CXCR4 27
1.3.II Biphasic signalling mechanism of CXCL12 29
1.4 HMGB1 breaks the concentration-dependence of CXCL12 signalling 31
1.5 Aim: improving HMGB1 as a tissue repair therapeutic 33
1.6 Specific objectives 34

1.1 Incomplete tissue repair after injuries leads to loss of function

Most tissues have a resident population of stem and progenitor cells which were important during development and subsequently replenish the differentiated cell populations for homeostatic maintenance and repair following injury [1]. Upon injury, physiological mechanisms are triggered to promote repair by activation of the immune system to remove damaged tissue and activation of these resident progenitors. Whereas for some tissues this mechanism achieves near-total restoration of function (e.g. bone and liver), in other cases this process is imperfect and the tissue is instead replaced by fibrotic scars (e.g. cartilage, skin, skeletal muscle [2][3]).

Stem cells reside in their niches and receive signals both from neighbouring cells through e.g. cell surface embedded molecules, extracellular matrix substances, and growth factors secreted by other cells (paracrine or endocrine signalling) [4]. Any alteration of these signalling mechanisms can influence the activity of stem cells [5]. Some tissues have limited intrinsic repair capabilities due to their structure or differentiation process [6]. Ultimately, following incomplete restoration of native architecture, function of damaged tissue is impaired, increasing the risk for further injury.

As the resident progenitor cell populations are instrumental for the regenerative process, it would appear logical to target them to promote repair. The simplest approach is to administer exogenous stem cell and the greatest success has been for some haematological disorders, where stem cell therapy has become standard of care [7]. However, for solid organs no approaches have moved beyond small-scale clinical trials [8]. This is due to several factors. Exogenous stem cell culture is typically complex and requires significant investment [9]. Furthermore, in contrast to haematological malignancies where the stem cell niche is maintained, following injury to solid organs, the niche is damaged. This adversely affects the ability of the exogenously administered stem cells to engraft into the inflamed and scarred host bed [10]. Furthermore, stem cells cultured *ex vivo* are prone to change their phenotype as the environment is seldom identical to the native stem cell niches due to their complexity [5,9,11].

The inflammation following injury is crucial for modulating downstream repair processes [12]. This initial inflammatory stage is responsible not only for the removal of damaged tissues, but also for the recruitment of progenitor cells to injury sites via chemokines such as CXC Chemokine Ligand 12 (CXCL12) [13,14]. The severity and duration of inflammation following injury is a key determinant of subsequent outcomes. Whilst low levels of inflammation in the early stages of repair are necessary for healing to occur [12,15], this response needs to be brief and give way to M2 macrophage-specific activity. Prolongated M1-macrophage activity upon injury is detrimental to successful regeneration [16,17]. This occurs when the acute inflammatory response is excessive, for example following myocardial infarction (MI) and reperfusion injury [18] or when other factors result in chronic low-intensity inflammation, such as type 1 diabetes mellitus [16].

The first mediators to be released following injury are a subset of endogenous molecules called the alarmins, a type of damage-associated molecular pattern (DAMP) [19]. These are constitutive components of cells which typically perform non-immune related functions. However, they are released upon non-programmed (e.g. necrotic) cell death and activate the immune system [20,21]. Alarmins are a heterogeneous group of molecules and include proteins, peptides, and metabolites. Given that they are the earliest mediators to be released, modulation of alarmin activity can alter the inflammatory landscape at the site of injury and hence targeted to enhance endogenous tissue repair.

A prototypical alarmin is High Mobility Group Box 1 (HMGB1), which forms the core of my DPhil project. This protein was identified in 1988 as a DNA-binding chromatin component [22] and transcriptional regulator. It was found later to also be a mediator of the immune response to bacterial infection [23], and found to be capable of modulating immune activity and chemotaxis in 2012 [24] which cemented its classification as an alarmin. In 2018, another function was identified by my group for this protein, namely as a potent mediator of stem-cell directed tissue repair [25]. Therefore, HMGB1 is a multifunctional protein. This dual activity of HMGB1 in injury, as both an alarmin that promotes inflammation and as a mediator of tissue repair, makes it an interesting therapeutic target.

1.2 High-Mobility Group Box Protein 1 (HMGB1)

Formerly known as HMG-1, HMGB1 is essential for development. *HMGB1*$^{-/-}$ animals are not viable and die due to hypoglycaemia and down-regulation of the glucocorticoid receptor shortly after birth [26,27]. HMGB1 has multiple functions. It is involved in transcription regulation [28], nucleosome function and genome stability. Extracellular HMGB1 is associated with multiple immune processes, including late-onset endotoxin response [23,29], thrombus formation [30], chemotaxis [31,32], as well as tissue repair [25], amongst others. It also has powerful antibacterial effects [33]. Thus, HMGB1 has a crucial role in human health and disease.

HMGB1 is also extremely conserved, with 99% identity between the amino acid chains of human, mouse and rat, and has high conservation across other vertebrates (see Figure 1.1), as well as homologues in *D. melanogaster* [34,35] and yeast [36]. The protein family to which HMGB1 belongs, the HMG Box group, also possess high sequence identity to HMGB1. There is 80% amino acid conservation between HMGB1 and HMGB2/3, despite each gene having clearly identifiable knockout (KO) phenotypes [37]. This is striking when the fact that HMGB1 is a particularly ancient gene is taken into consideration, with the estimated origin of the HMG Box family being dated as far back as 525 million years ago [38].

HMGB1 is constitutively expressed in almost all cell types. Whilst in most cases the protein accumulates in the nucleus, even non-nucleated cells such as platelets [39] express large amounts of HMGB1. Additional to constitutive expression, several cell types also have specifically regulated pathways for HMGB1 expression [40]. HMGB1 expression varies as much as a hundred-fold between undifferentiated cells (high level) and differentiated cells (low level). Cancer cells also increase HMGB1 expression, resulting in enhanced angiogenesis, tumour survival and dispersion, and excessive proliferation of these cells [41].

Within the cell, HMGB1 is primarily a chromatin protein, being involved in DNA repair [36,42] and transcriptomic regulation [43]. However, since HMGB1 is an alarmin, it can also be released to the extracellular medium after tissue damage or infection. This broad variety of extracellular functions result in a protein with strong evolutionary constraints which explain the high degree of conservation of the sequence and structure of HMGB1.

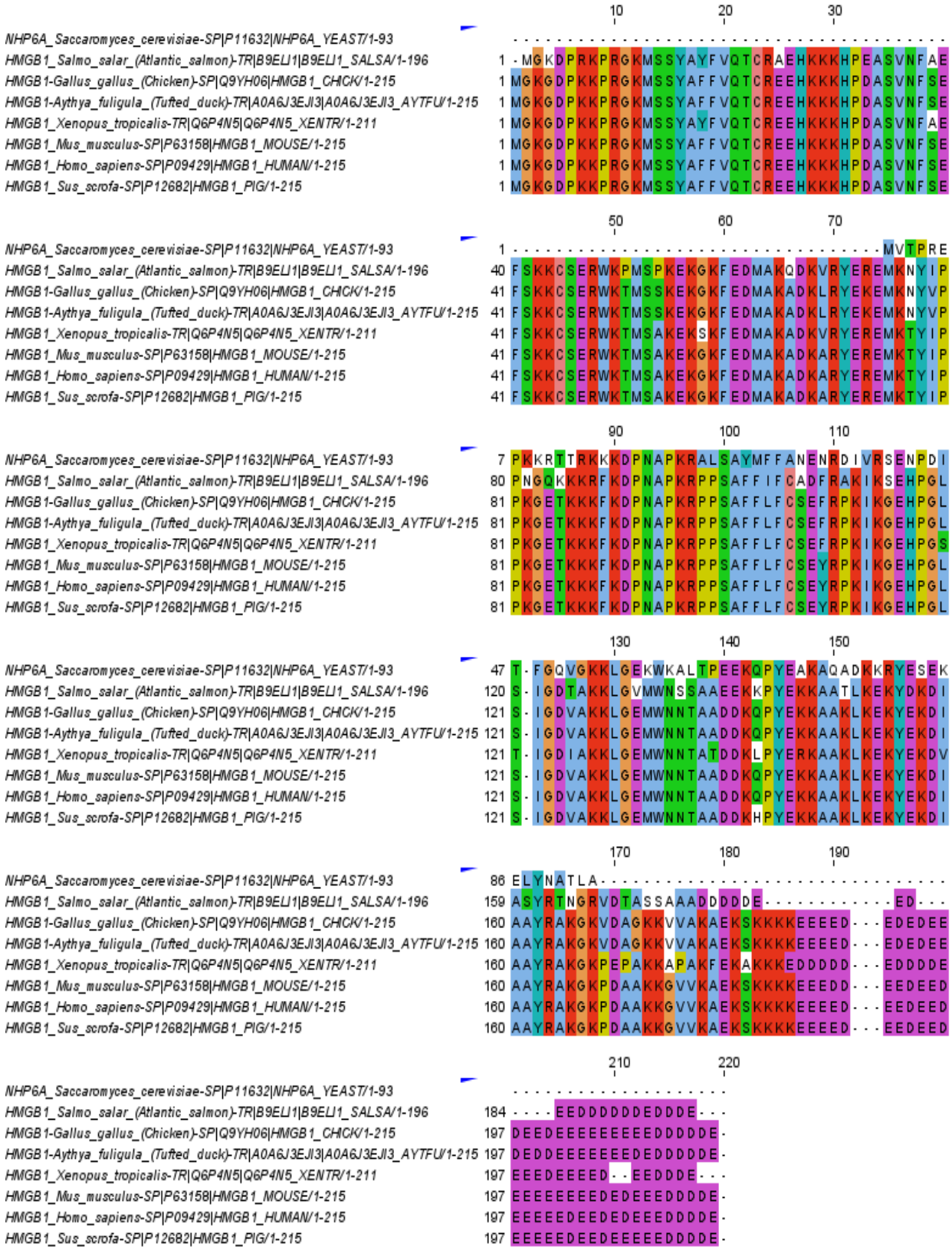


Figure 1.1: Conservation of HMGB1 sequences across different vertebrates

Alignment of known HMGB1 homologues in vertebrates and also yeast HMG-like protein NHP6A, as retrieved from the Uniprot database [293]. Alignment performed in the Clustal Omega EMBL server [294], with default parameters. Image generated via Jalview [295]. Colour code: hydrophobic (blue), positive charge (red), negative charge (lilac), polar (green), cysteine (orange), proline (yellow), aromatic residues (cyan), not conserved (white). Uniprot IDs and common names provided in the left.

1.2.I Characteristics of the high-mobility group (HMG) family

All of the known HMG class proteins have a series of defining characteristics [44]:

- Extractable from chromatin via 0.35 M sodium chloride
- Solubility in 5% perchloric acid or trichloroacetic acid
- Low molecular weight of <30 kilodalton (kDa)
- High content of charged amino acids and fast migration in sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis (PAGE)
- Presence of extensive post-translational modifications which affect its activity
- Tissue and development-dependent regulation of expression

There are several families of HMG proteins [37]: HMGA proteins, whose key structural feature is the presence of a triplet of AT-hook motifs that are functionally independent and act as transcription regulators; HMGN proteins, a group found exclusively in vertebrates, that bind nucleosomes and regulate chromatin structure and function. Finally, there are the HMGB proteins, which are the most abundantly represented across organisms with homologues in vertebrates but also plants such as *Arabidopsis thaliana*. The key feature of the HMGB group is the presence of the so-called HMG-box domains, which bind DNA without sequence specificity to induce structural changes [28].

1.2.II Structure and functions of the HMG Boxes within HMGB1

The structure of HMGB1 has been elucidated, both via solution NMR (Protein Data Bank - PDB - 2YRQ) and in a complex with deoxyribonucleic acid (DNA) and RAG 1/2 recombinase via cryo-electron microscopy and X-ray crystallography (PDB 6CG0/6CIK). The main structural element in HMGB1 is the HMG Box domain, an archetypical L-shaped structure common to the HMG family, and analogous to proteins such as yeast non-histone chromosomal protein 6A (NHP6A, Figure 1.2) which performs similar transcriptomic functions [36].

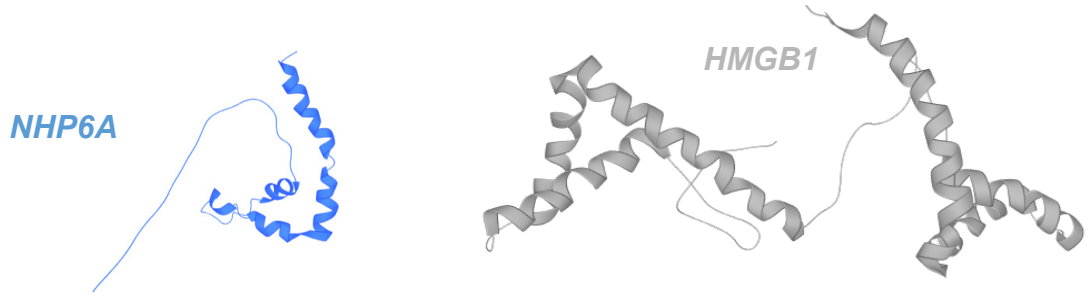


Figure 1.2: Structure of the HMG Box domains in NHP6A (blue, PDB 1CG7) and HMGB1 (grey, PDB 2YRQ)

The structure of these is rigid and conserved across HMG Box proteins. HMG Box domains act typically as genomic stabilising agents, protecting against e.g. DNA damage [36]. This function is mediated by sequence-independent DNA binding of the HMG Box proteins [45,46]. HMGB1 is known to also regulate transcription via interaction with transcription factors such as p53 [47]. Each HMG Box contains a nuclear localisation signal responsible of nuclear import [37].

HMGB1 contains two such domains (Box A and Box B) with a high degree of sequence conservation to each other, which suggests a tandem duplication event from a single-box HMG protein. Box A can form an intramolecular disulfide bridge (Cys 22-Cys 44) which modulates several of its activities; this is absent in Box B where only one cysteine (Cys 105) is present. The HMG Box domains are intrinsically charged due to the high proportion of amine side chains (pI of the Box domain > 9.5). In addition, HMGB1 has a carboxylic-acid tail at the C-terminus that regulates its DNA bending activity [48] and antibacterial effects [33], as well as modulating proinflammatory properties [49]. Interaction of the tail with the Box domains results in several different conformations [50,51] due to the opposite charges. Different functions have been mapped within each HMG Box (Figure 1.3). These involve interactions with compounds such as lipopolysaccharide (LPS) [52], heparin (residues 1-9 [53]) or pathogen and damage associated molecular pattern (PAMP/DAMP) receptors such as Toll-like receptors 2 and 4 (TLR-2, TLR-4) or the receptor for advanced glycation end-products (RAGE).

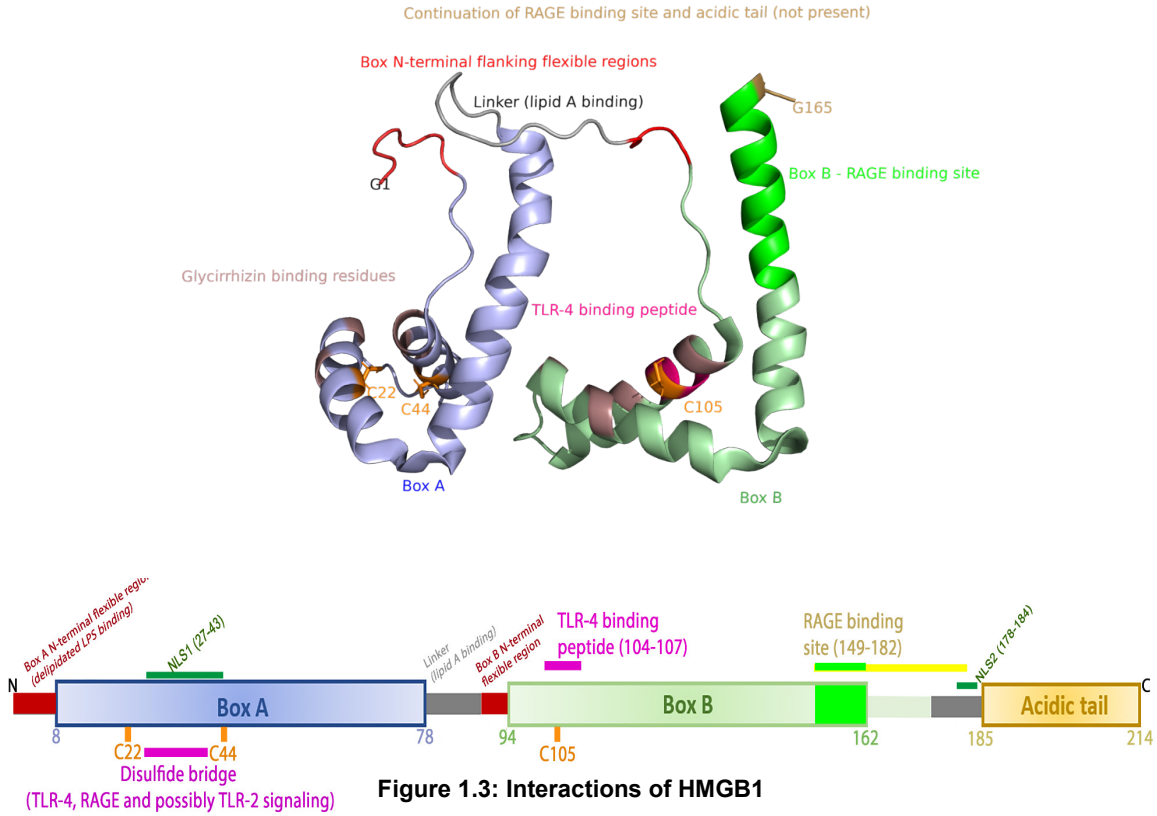


Figure 1.3: Interactions of HMGB1

Structure of HMGB1 (PDB 2YRQ, conformer 1), coloured according to known interactions with PAMPs and molecules. The acidic tail (transcriptional modulation and bactericidal activities) is not contained within the structure and therefore is not displayed; a schematic is provided below. Heparin binds the N-terminal flanking flexible region before Box A.

1.2.III Subcellular location and PTMs: the functional redox switch in HMGB1

As with many other proteins, post-translational modifications (PTMs) significantly modulate the activities of HMGB1. Known modifications include acetylation [54,55], phosphorylation [56], glycosylation [57] and oxidation of cysteines, resulting in a disulfide bond between cysteines 22 and 44 (23 and 45 if including the N-terminal Met, which is typically lost *in vivo* after translation) [24].

The cysteine bond between C22 and C44 is one of the key determinators for the activity of HMGB1. Fully-reduced HMGB1 (FR-HMGB1) has been described as a chemoattractant [24], capable of recruiting immune cells to injury sites, an activity which is inhibited by the alkaloid glycyrrhizin [58] that binds in the concave surface between helices I and II. Oxidation of C22-C44 into cystine results in disulfide HMGB1 (DS-HMGB1), which is a prototypical DAMP, which binds several proinflammatory receptors including TLR-2, TLR-4, and RAGE. It can also bind PAMPs (such as endotoxin [29] or lipotheicoic acid [59]) to induce inflammatory responses to infection through these same receptors. In the nucleus, HMGB1 exists in the fully-reduced state: due to the glutaredoxin system [60] preserving the free thiols.

Lacking classical secretion signals, HMGB1 cannot be secreted actively via the endomembrane system. Bianchi and colleagues [55] first described the acetylation of the nuclear localisation signals (NLS) of HMGB1 as a novel mechanism for its active export via the nuclear exportin CRM1 into the extracellular medium upon endotoxin challenge. This process was later shown to involve the N-glycosylation of N36/133/134 as a requirement for CRM1-mediated export [61]. Other pathways have since been described, including other post-translational modifications such as phosphorylation via protein kinase C and NF-κβ/Mitogen associated protein Kinase (MAPK)-dependent export. These are triggered by innate immune responses, such as monocyte activation [55,56,62] in response to bacterial challenges [56], inflammasome-dependent release in autoimmune processes [29,63], as well as many others [37], highlighting the role of HMGB1 in activating the innate immune system. Certain non-immune cells are also capable of HMGB1 export. For example, platelets which express HMGB1 despite lacking a nucleus, oxidise and translocate it from the cytoplasm to the membrane surface upon activation [64].

Ultimately, phosphorylation of HMGB1 and acetylation of the NLS sequences prevents nuclear import, resulting in accumulation of HMGB1 in the cytoplasm. Cysteine 105 (106 with N-terminal Met) is also involved in this process, as C105S mutants of HMGB1 constitutively localise to the cytoplasm [60]. Cytoplasmic HMGB1 then accumulates in secretory lysosomes by default [65], where the oxidising environment converts the protein to the disulfide form [24]. Upon stimulus such as monocyte activation, these lysosomes are secreted to the extracellular medium, releasing DS-HMGB1. In the final stages of inflammation, the thiol groups of extracellular HMGB1 are terminally oxidised into sulfonates, resulting in its inactivation [66].

Reduced HMGB1 (FR-HMGB1) can also be released passively. The archetypical scenario for FR-HMGB1 release is non-programmed cell death (e.g., necrosis). FR-HMGB1 is loosely bound to chromatin under normal conditions and, therefore, if the cells become permeable, FR-HMGB1 is released [67]. During programmed cell death (e.g. apoptosis), chromatin, including HMGB1, is hypoacetylated [56,68]; hypoacetylated HMGB1 accumulates in the nucleus as the nuclear localisation signals (NLS) are exposed and is also more tightly bound to DNA [55,69]. FR-HMGB1 in the extracellular medium acts as a chemoattractant [24,32,70], recruiting immune cells to the damaged tissue sites. What occurs after this recruitment is controversial.

It was originally proposed that FR-HMGB1, once sufficient immune cells have been recruited to the site of injury, oxidises due to the presence of reactive oxygen species and is converted to the proinflammatory disulfide form [24]. However, the half-life of HMGB1 in the extracellular medium is extremely short. HMGB1, whether FR or DS, is cleared from the bloodstream with a half-life of 26 minutes [71]. HMGB1 is removed from circulation by binding to circulating protein haptoglobin-β (preferentially bound by DS-HMGB1 [72]). HMGB1-haptoglobin complexes are subsequently removed from the circulation by hepatic macrophages. Binding of the HMGB1-haptoglobin complex to CD-163 in the cell surface also triggers production of anti-inflammatory enzymes (heme oxygenase) and cytokines (IL-10) [73]. Since oxidation half-life of HMGB1 in serum is in the same order of magnitude as clearance (V_{50} = 17 minutes [74]), the initial rapid clearance of DS-HMGB1 would suggest that the systemic proinflammatory effects of physiological levels of FR-HMGB1 released by necrotic cells are negligible, as the oxidised form is cleared faster than it is produced.

It has also been reported recently that there are two peaks in HMGB1 plasma concentration following trauma, corresponding to separate release events [71]. The first peak occurs immediately after the initial trauma, decays exponentially, and is not correlated with clinical outcomes. The second peak occurs after the initial peak has subsided at 3-6 hours, and patients with high levels of HMGB1 in the second release episode typically experience worse outcomes. Based on the clinical outcomes, the authors proposed that the two peaks related to different oxidation states of HMGB1. Upon trauma, necrotic cells release FR-HMGB1, which recruits immune cells and is promptly cleared from the circulation. Then, these immune cells release DS-HMGB1, triggering inflammation and explaining the association to poor clinical outcomes. It should be noted that the authors did not assess the oxidation or acetylation status of this circulating HMGB1: however, two successive cell death episodes after a single trauma event are unlikely [71].

These arguments only discount systemic effects of the oxidation of passively released FR-HMGB1, as the local clearance might be slower. The short half-life has also been measured internally in the Nanchahal group with exogenously administered HMGB1 (unpublished data by Geoffrey Lee and Ana Espirito Santo) but two clearance events, with half-lives of 11 and 120 minutes, were observed. This could represent separate clearance events and slower local elimination, where HMGB1 has enough time to locally oxidise and trigger inflammation before it is fully cleared.

1.2.IV DS-HMGB1 is both a damage- (alarmin) and pathogen-associated molecular pattern

DS-HMGB1 can induce production of tumour necrosis factor (TNF) several interleukins (IL; IL-1α, IL-1β, IL-1RA, IL-6, IL-8), macrophage inflammatory protein (MIP)-1α, and MIP-1β [75] by macrophages. It also can affect endothelial cells, inducing both expression of adhesion molecules, RAGE, and cytokines IL-8 and TNF, influence dendritic cell maturation, and stimulate production of TNF, IL-1β, and IL-8 by neutrophils [76]. Furthermore, it can increase the permeability of epithelial cells. All these functions result in DS-HMGB1 acting as a key mediator of inflammation and sepsis in the extracellular medium. It is worth mentioning that, until 2012, the literature did not distinguish between the different redox states of HMGB1 as their significance was unknown, leading to conflicting publications due to the lack of control for the redox state of the molecule in the experiments. The first studies in the matter highlighting the immunomodulatory roles of HMGB1 described its enhancement of responses to endotoxemia during bacterial sepsis [23]. Antibody neutralisation of HMGB1 activity in mice improved survival upon lethal endotoxemia; conversely, administration of HMGB1 increased mortality.

HMGB1 was promptly identified as a key mediator of sterile inflammation in response to injury (DAMP), where no exogenous components such as bacteria are present, and the inflammatory response is triggered by e.g., tissue damage. In several models, such as drug-induced liver injury [77], ischemia and reperfusion [78], hypoxia [79], thrombosis [64,80] and the aforementioned case of trauma [71], HMGB1 is critical for the inflammatory response to tissue damage. HMGB1 can perform these activities on its own or operate in concert with a plethora of both host and exogenous molecules. These are in addition to the co-factors or co-receptors required by the receptors themselves: cluster of differentiation 14 (CD14) and myeloid differentiation factor 2 (MD-2) [81] for TLR-4, and heparan sulphate for RAGE [82]. DAMP signalling of HMGB1 is mainly effected via the receptors TLR-4, TLR-2, or RAGE, and results in activation of the NF-κβ pathway, whether directly in the case of TLRs or indirectly through MAPK pathways in the case of RAGE [83]. All three pathways converge in signalling through myeloid differentiation factor 88 (MyD88) to induce NF-κβ activity [67,84].

TLR-4/2-HMGB1 signalling has been associated with the production of proinflammatory cytokines and reactive oxygen species in neutrophils [85]. Blockade of HMGB1 proved beneficial in several diseases, such as paracetamol-induced hepatic injury, sepsis, or arthritis. Conversely, exogenous addition of HMGB1 was found to be deleterious in the aforementioned models, owing to its proinflammatory activities [67]. Sequestration of HMGB1 by haptoglobin also confers protection against lethal sepsis and endotoxemia in injury [72,73], via its interaction with CD163. Both TLR-4 and TLR-2 are key for the activation of macrophages by the disulfide form of HMGB1, although TLR-2 signalling is not restricted to DS-HMGB1 [49]. Furthermore, they can act synergistically with each other. Co-transfection in NF-κβ reporter HEK cells with TLR-2 and TLR-4 resulted in higher reporter activity than cells co-transfected with either TLR-2 and RAGE or TLR-4 and RAGE [86], owing to their shared signalling pathway [83].

DS-HMGB1 can induce both LPS-independent and dependent (replacing LPS binding protein LBP [52]) TLR-4 signalling on its own [87,88], but whether a ligand is required for TLR-2 mediated HMGB1 signalling is controversial [49,59]. HMGB1 forms complexes with exogenous TLR-2 [59] (lipotheicoic acid and peptidoglycan) and TLR-4 (endotoxin) [29] ligands, and synergistically enhances the response of cells expressing these receptors to their respective ligands [89]. These interactions result in a quantitative increase in receptor activity, which does not involve additional co-receptors or different signalling pathways compared to the ligand alone. More interestingly, these interactions with the TLR ligands have been reported to occur with the fully reduced form, although the contribution of these complexes to signalling has not been described. Whilst the MD-2/TLR-4 interaction occurs preferentially with DS-HMGB1 [88], FR-HMGB1 is still capable of complexing with Pam_3CSK_4 or lipoteichoic acid (LTA) [59]. It remains unknown whether endogenous molecules can also act synergistically with DS-HMGB1 to enhance sterile inflammation; in the absence of serum, no TLR-2 signalling is induced by HMGB1 alone in macrophages, whereas in whole blood cells or HEK reporter cells (TLR-1/2 dimers) grown with serum a response to DS-HMGB1 alone was observed [49], which has been shown in other studies as well [90]. It could be possible that the presence of a co-ligand enhances basal TLR-2 agonism of DS-HMGB1, enabling signalling at much lower concentrations.

RAGE can mediate chemotaxis via interaction with FR-HMGB1 [31,91]. This interaction induces CXCL12 expression through NF-κβ signalling [70,92]. However, in the context of injury and inflammation, there are two DS-HMGB1 mediated RAGE-mediated pathways of special relevance, Neutrophil Extracellular Trap (NET) formation and subsequent deep venous thrombosis (DVT) [64].

NETs are extracellular structures composed of DNA, histones, and several granular proteins, such as neutrophil elastase, cathepsin G, and myeloperoxidase. They function to kill bacteria and prevent their spreading [93]. They also possess significant prothrombotic activity [94]. Platelets have been known to promote to NET generation by neutrophils [80] when challenged with microbial stimuli. When activated, platelets translocate their HMGB1 from the cytoplasm and the surface in the disulfide form, presenting it to neutrophils. Blockade of RAGE or knock-out of HMGB1 in neutrophils (chimeric animals) significantly decreases the capacity of platelets to induce NET production by neutrophils [64].

Clot formation in DVT occurs under sterile injury conditions and is dependent on the innate immune system [95]. It has been shown that thrombi accumulate HMGB1 over time and that HMGB1$^{-/-}$ chimeric animals show reduced thrombus weight in a murine thrombus model [64]. Thrombus formation is restored when these animals are injected with wild-type platelets: activated platelets translocate their HMGB1, oxidising it in the process, which in turn triggers NETosis and starts the formation of thrombi.

However, TLR-2 also cooperates with RAGE on thrombosis. Individual knock-out of TLR-2, TLR-4, or RAGE does not affect thrombus weight, but knockout of the common adaptor protein, Myd88, results in a decrease. Knock-out of TLR-2, RAGE or Myd88, but not TLR-4, reduces monocyte recruitment to the clot site, and only RAGE or Myd88 knockout has an effect in NET reduction, suggesting that despite both receptors being redundant to each other, RAGE signalling plays a greater role than TLR-2 in thrombosis. This redundancy is consistent with the convergence of signalling cascades through all three receptors into the NF-κβ pathway [83].

1.2.V FR-HMGB1 is not only a chemotactic mediator but a strong pro-regenerative molecule

In most clinical publications, HMGB1 inhibition has been shown as beneficial in several contexts due to its deleterious proinflammatory activities. This suggests that HMGB1 activity outside the cells is detrimental for the organism. However, this is only half of the story. Whilst DS-HMGB1 appears to be highly detrimental in most cases, the reduced form of HMGB1 is beneficial for wound healing. The redox switch allows HMGB1 to orchestrate both the initial phase, recruiting immune cells to damage sites, and the subsequent inflammatory phase, is a unique trait which is not shared with other DAMPs and sets HMGB1 apart from them.

FR-HMGB1 acts by forming a complex with the CXC chemokine ligand 12 (CXCL12), also known as stromal derived factor 1 (SDF-1). This complex signals through CXCR4, one of the cognate receptors for CXCL12, to induce chemotaxis. Initially this was presumed to be part of the inflammatory response alone, where FR-HMGB1 released by necrotic cells recruits inflammatory cells to the tissue damage site [24,32,96]. Interest on this pathway was first sparked by the discovery of the interaction between fully reduced HMGB1 and CXCL12 [24,96,97]. HMGB1 was found to enhance the chemotactic activity of CXCL12 in immune cell types, allowing for chemotaxis even at suboptimal concentrations of CXCL12, and inducing a different conformation of CXCR4 in the plasma membranes of CXCR4-expressing cells. This interaction was shown to be sensitive to glycyrrhizin, an alkaloid which inhibits the chemotactic activities of HMGB1, and offered a possible explanation for its mechanism of action by which HMGB1 modulates CXCL12 signalling.

The activity of FR-HMGB1 in injury is not limited to cell recruitment. FR-HMGB1 is also a key initiator of the tissue repair response. Upon injury, many tissues, including bone, skeletal muscle, and blood, rely on endogenous populations of stem and/or progenitor cells. These cell populations are typically quiescent in mature tissue and on tissue damage they enter the cell cycle to proliferate and subsequently differentiate to repair the damage. FR-HMGB1, by the signalling through the CXCL12-CXCR4 axis, has been shown to be key for tissue healing [25].

In 2014, Rodgers and his colleagues [98] explained a curious finding first reported in 1970: a priming injury at a site distant to subsequent trauma resulted in accelerated healing in rabbits [99]. This study showed that the priming injury resulted in the release of a circulating factor which forced stem cells to enter in "alert" state (G_{Alert}) via mTORC1 signalling. G_{Alert} is a state intermediate between the quiescent (G_0) and fully activated (G_1) states, and is characterised by an increased number of mitochondria, adenosyl triphosphate (ATP) levels and cell size compared to quiescent cells but lower than in activated ones. Cells in G_{Alert} enter G_1 much faster, compared to stem cells in G_0, upon exposure to the appropriate activating factor. However, if not activated, the cells in G_{Alert} revert to G_0 after 3-4 weeks. They went on to show that hepatocyte growth factor activator (HGFA) could induce G_{Alert} and accelerate healing [100]; however, HGFA is also an activator of cell division (induces G_1 entry), bypassing G_{Alert}.

FR-HMGB1 is capable of inducing the G_{Alert} state in quiescent stem cell populations via signalling through the CXCL12-CXCR4 axis, effectively accelerating the entry of stem and progenitor cells into the cell cycle [25]. Inhibition of the HMGB1-CXCL12 complex by glycyrrhizin, global conditional knockouts, or inhibition of the CXCL12-CXCR4 interaction greatly impaired bone healing. Administration of exogenous HMGB1 up to two weeks prior or just after injury enhanced bone regeneration and rescued the poor healing in conditional HMGB1 knockout animals. CXCL12 administration alone was unable to produce these effects, instead resulting in disordered fracture callus. Stem cells isolated from tissues exposed to HMGB1 showed all the features of G_{Alert}, and when the model was extended to other tissues such as blood (chemotherapeutic depletion of white blood cells) or muscle (barium chloride injury), accelerated healing was observed. Therefore, agonism of the FR-HMGB1/CXCL12/CXCR4 pathway could be a therapeutic strategy to promote functional recovery of tissues. However, the multiple deleterious functions of DS-HMGB1 preclude this. Whilst the short half-life of FR-HMGB1 did not lead to a significant increase in proinflammatory cytokines in the circulation in normal mice [25], local conversion to DS-HGMB1 around the site of injury is likely lead to deleterious effects e.g. in the coronary circulation following MI, such as thrombosis. A detailed understanding of the interaction of HMGB1 with CXCL12 is necessary to engineer variants of HMGB1 capable of promoting repair whilst eliminating potential deleterious proinflammatory signalling.

1.3 The CXCL12-CXCR4 axis

CXCL12 is the sole known chemokine ligand of CXC Chemokine Receptor 4 (CXCR4), a G-protein coupled receptor (GPCR) with a single seven transmembrane helix bundle. The CXC designation corresponds to a specific Cys-X-Cys motif in CXCL12, where X can be any residue. This axis has been extensively studied in the literature due to its critical function in stem [13,14,101] and immune cell mobilisation, human immunodeficiency virus (HIV) entry [102–105], regulation of cell fate and tissue development [106–108]. Knockouts of this signalling axis are typically lethal due to abnormal organ formation [109,110]. This combination of essential functions means that CXCR4 is of great interest.

1.3.I Structures of CXCL12 and CXCR4

The structure of monomeric CXCL12 has been available for quite some time [111], although refinements to the initial model have been made [112] (Figure 1.4) and dimer structures are also available [113]. It consists of a flexible N-terminus which is linked to a β-sheet by the RFFESH loop, critical in CXCR4 signalling. The C-terminus of the molecule ends in an α-helix. CXCL12 is secreted as a mature protein, but the nascent peptide possesses a *Sec*-type secretion signal that is lost upon processing. Despite being highly positively charged (pI > 9), CXCL12 forms homodimers with weak affinity (K_d > 100 μM). However, dimerisation only occurs at pH > 5.5 and requires coordination of multivalent anions (e.g. sulphate or phosphate) to neutralise the positive charge [114]. Low pH leads to His25 protonation and electrostatic repulsion with Lys 27 in an adjacent monomer [114]; mutation to a cationic residue (e.g. Arg, H25R) blocks dimerisation in a similar manner to histidine 25 protonation [112,114].

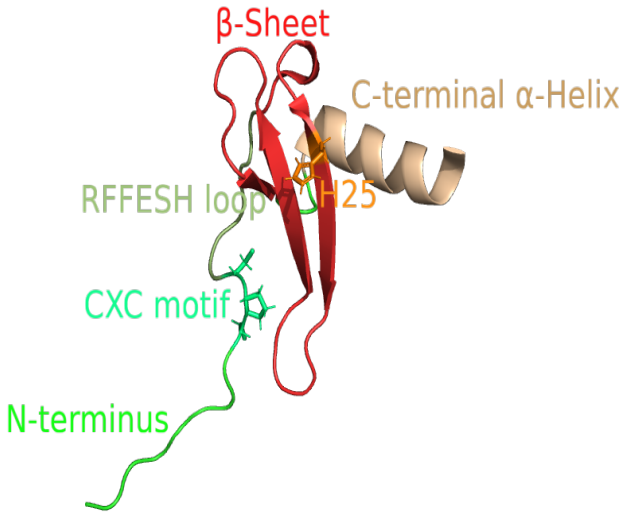


Figure 1.4: CXCL12 monomer structure (PDB 2kec)

Elucidation of the structure for CXCR4 is far more recent. In 2009, Raymond Stevens´ laboratory crystallised this GPCR in the inactive conformation, aided by fusion of T4 lysozyme within the intracellular loops [115] and thermostabilising mutations L125W and T240P, the latter which distorts one of the alpha helices and inactivates signalling via the receptor. A second structure was published in 2015 in a complex with a viral protein [116]; again, this was an inactive form of the receptor due to the T240P mutation. There is no structure of the active conformation of CXCR4 in complex with CXCL12, although computational models exist [117]. This hampers computational modelling of the HMGB1-CXCL12-CXCR4 ternary complex, which likely signals through the activated CXCR4 conformation. This is due to the highly dynamic nature of GPCRs which makes formation of homogeneous crystals challenging.

The published structure (Figure 1.5) shows a core domain with seven-transmembrane α-helices (red), joined by 3 extracellular (bright green) and 3 intracellular loops (matte green), a flexible N-terminus (cyan), and a shorter eight alpha helix in the C-terminus. The extracellular loops and N-terminus of CXCR4 are critical for CXCL12 binding [113,117] and are used as epitopes for CXCR4-neutralising and detection antibodies, whereas the C-terminal alpha helix is involved in regulation of CXCR4 levels in the membrane and on-agonist inactivation [118,119]. In the cell surface membrane, it forms either homo- or heterodimers with CXCR7, a CXCL12-scavenging receptor which antagonises CXCR4 signalling by ligand depletion [108] and p38-mediated CXCR4 inhibition [120].

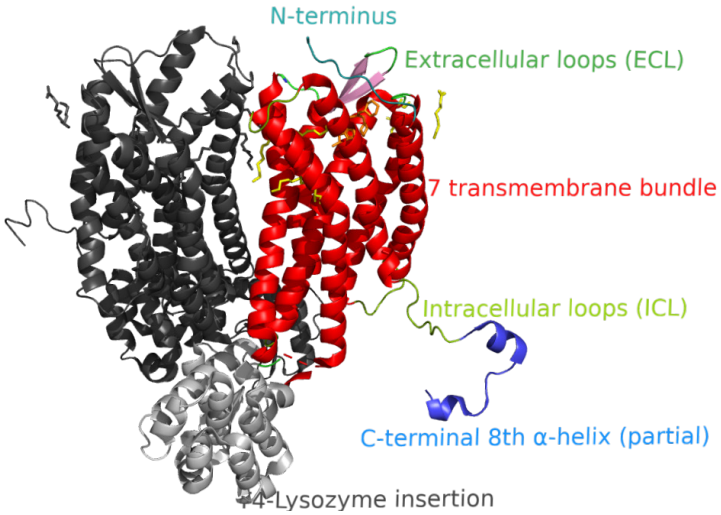


Figure 1.5: Structure of CXCR4 (PDB 3ODU)

Structure of a CXCR4 homodimer, colored according to the protein domains. Lipids are colored yellow, and the compound present in the crystal (inhibitor ITD) in orange; the second CXCR4 monomer is colored dark grey.

1.3.II Biphasic signalling mechanism of CXCL12

The interaction between CXCL12 and CXCR4 has been known almost from the discovery of both proteins [105,121] to be highly specific. CXCL12 interacts strongly with CXCR4 receptor [102] (3.6 nM [122]), although its interaction with CXCR7 is ten-fold stronger [123]. Despite intensive investigation, the structure of the complex formed between CXCL12 and CXCR4 remains unclear. This is largely as a result of the inherent challenges of working with activated GPCRs [124] due to their conformation shift between activated and inactive being fluid in the absence of additional components [125]. The first proposed model described two independent binding events in binding of a CXCL12 monomer to a CXCR4 monomer [126]. This study also noted that the interface in Site 1 overlaps with the CXCL12 dimerisation interface, highlighting a possible cause for the different signalling pathways elicited:

- Site 1: N-terminal RFFESH loop, β-sheet surface, and the loop between this and the C-terminal α-helix (50s loop) in a CXCL12 monomer (Figure 1.4) bind the extracellular loops of a CXCR4 monomer. This site likely positions CXCL12 in the correct conformation for signalling through Site 2 to occur.
- Site 2: The N-terminal flexible peptide in CXCL12 burrows within the GPCR core and triggers signalling. Addition, mutation or removal of single amino acids within the N-terminus of CXCL12 greatly affects its activity [111,127,128]. N-terminal truncated forms exist *in vivo* [128] and modulate CXCL12-CXCR4 signalling.

A coetaneous study identified sulfation of three tyrosine residues (7, 12, 21) in CXCR4 as critical for dimeric CXCL12 signalling [113]. These sulphated residues stabilise the formation of a CXCL12 dimer within the CXCR4 binding interface, provided there is sufficient CXCL12 to saturate both receptors as a homodimer, triggering the pro-quiescence dimeric signalling. In turn, this modifies the presentation of the CXCL12 N-terminus within the CXCR4 transmembrane bundles Therefore, both CXCL12 concentration and receptor post-translational modifications (PTMs) modulate signalling.

Further work has proven that the influence of CXCR4 on CXCL12 dimerisation is key for the biphasic signalling of CXCL12 [117]. Upon dimerisation of CXCL12, the specific contacts for the monomer form are lost as they become trapped in the dimer interface. At any given concentration of CXCL12, where both CXCR4 monomers in a homodimer will be occupied, the signalling will be biased towards the dimer-associated pathway rather than two independent monomer-associated signalling events [117]. Both the monomer and the dimer activate CXCR4 signalling and calcium mobilisation within the cell, but with different downstream signalling cascades and effects in the cell:

- CXCL12 monomers preferentially bind CXCR4 monomers (within a dimer) when Tyr7 and Tyr 12 are not sulfonated [113], as well as CXCR7. Signalling through CXCR4 of monomeric CXCL12 leads to chemotactic activity [32] and pro-differentiative pathways [129], with downstream GPCR signalling cascade through both β-arrestin and G-protein pathways.
- CXCL12 dimers preferentially bind sulfonated CXCR4 dimers and their signalling is G-protein biased, and thus do not induce chemotaxis, but instead enforce stem cell quiescence. NMR analysis showed that monomer-specific contacts between CXCL12 and CXCR4 are lost upon dimerisation of the former [130], and that binding of two independent CXCL12 monomers to each of the subunits CXCR4 dimer promotes dimeric-associated signalling [117].

This signalling shows a biphasic dose response due to the interaction between the CXCL12 binding equilibrium and CXCR4 binding. At low CXCL12 concentrations, due to the weak dimerisation affinity, the predominant form is the monomer, inducing chemotaxis and cell differentiation [114]. Once CXCL12 concentration is sufficiently high in solution it dimerises and enforces stem cell quiescence, antagonising its own chemotactic activity, a process enhanced by dimerisation of CXCL12 in the receptor surface [117]. This process allows for a CXCL12 gradient to promote progenitor cells to home towards their niche, where the high local CXCL12 concentration induces quiescence and preserves a resting stem cell population in the bone marrow. The CXCL12 antagonist, AMD 3100, targets this pathway to mobilise bone marrow stem cells after chemotherapeutic depletion.

1.4 HMGB1 breaks the concentration-dependence of CXCL12 signalling

HMGB1 was initially described in relation to the CXCL12-CXCR4 axis as an enhancer of the chemotactic activities of CXCL12. At concentrations of CXCL12 below the threshold required for chemotaxis, co-addition of HMGB1 resulted in chemotaxis of previously unresponsive cells [70,91,96], which was lost upon neutralisation of CXCL12 activity or increase of CXCL12 concentration beyond the dimer formation threshold. This suggests that HMGB1 enhances monomer-associated signalling pathway but cannot induce chemotactic activity on its own.

However, the effects of HMGB1 in tissue repair mediated by this pathway imply the existence of a third signalling pathway in addition to the biphasic CXCL12-CXCR4 axis. Besides the monomeric (pro-chemotactic) and the dimeric (pro-quiescent) signals, HMGB1 triggers a pro-regenerative signal which is distinct from the other functions of CXCL12-CXCR4 alone. Mechanistic information on the HMGB1-CXCL12 interaction was extremely limited at the start of my DPhil and remains so to date. Only one publication described the biophysics of the interaction between both [96].[1] However, this study omitted several critical pieces of evidence, in particular, the NMR assignment tables corresponding to these. Furthermore, several key heteronuclear single quantum correlation experiments (HSQC) spectra cannot be found in either the BMRB or the associated DPhil thesis [131], which makes interpretation difficult.

[1] Near the end of my DPhil further publications from the same group followed [182,183]. The implications of these are discussed in Chapter 4 as they did not influence initial planning of my DPhil.

Nonetheless, this publication still offered key pieces of information which influenced my work plan:

- A known inhibitor of HMGB1-mediated inflammation, glycyrrhizin [132], was shown to specifically inhibit HMGB1-CXCL12 interaction. Glycyrrhizin binds to each HMG Box domain independently, which suggests that two independent binding events, each for one CXCL12 monomer to one Box. This was validated by the Box domains alone inducing different chemical shift perturbations.
- Binding of CXCL12 to a Box domain alone induces different changes in the ^{1}H-^{15}N-HSQC spectrum of CXCL12 compared to binding of CXCL12 to a Box domain within the full-length HMGB1. This suggests that binding of CXCL12 to each HMG Box domain can either be co-operative or position each CXCL12 monomer relative to the other differently in the full-length molecule than in a single Box domain.
- When CXCR4 containing C-terminal fusion cyan and yellow fluorescent protein (CFP/YFP) is transfected in cells, bioluminescence resonance energy transfer (FRET) can be observed as heterodimers of CXCR4-GFP and CXCR4-YFP are formed, and the efficiency of FRET depends on the relative conformation of each fusion protein within the CXCR4 dimer. HMGB1-CXCL12 induces a different FRET signal pattern within a CXCR4 dimer compared to CXCL12 alone, suggesting a different conformation which could lead to different downstream signalling pathways [96].

Based on this information, I hypothesised that binding of HMGB1 to CXCL12 might effectively “break” the propensity of CXCL12 to form dimers in the surface of CXCR4, leading to a distinct signalling pathway where both CXCR4 monomers within a homodimer induce the β-arrestin and G-protein signalling pathways. Recent publications would support this hypothesis [133], albeit with different activation of β-arrestin 1 and 2 compared to monomeric CXCL12 alone, which prevents on-agonist internalisation of CXCR4. This ultimately leads to induction of G_{Alert} and enhanced tissue repair through mTORC1 signalling [25]. Formation of a different downstream signalling complex would provide a mechanistic explanation for this novel signalling pathway within the CXCL12-CXCR4 axis, which cannot be reduced to either the monomer- or dimer-mediated CXCL12 exclusive signalling.

1.5 Aim: improving HMGB1 as a tissue repair therapeutic

Enhancing FR-HMGB1-CXCL12-CXCR4 activity on injured areas improves tissue repair in multiple models [25]. To develop FR-HMGB1 as a therapeutic to promote tissue regeneration, is important to minimise the potential for deleterious signalling following local conversion *in vivo* of exogenous FR-HMGB1 into DS-HMGB1. To remove TLR-2/4 or RAGE signalling whilst retaining full regenerative activity and CXCL12 binding, the residues involved in binding need to be known. The ideal case scenario would be to elucidate the structure of the HMGB1-CXCL12-CXCR4 ternary complex. Protein structures of HMGB1 (PDB 2YRQ), CXCL12 [112] and an inactive CXCR4 mutant [134] alone are already known. Furthermore, several of the proinflammatory functions of HMGB1 have been at least partially mapped to its sequence (at broad peptide level) [87,135], which in combination with the aforementioned structures could allow to compare the distribution of these epitopes over folded HMGB1. However, it was not realistic for me to work on all three components of the ternary complex. Therefore, recombinant expression and optimisation of CXCR4 was undertaken as an independent project by other researchers, whilst I focused solely on the FR-HMGB1-CXCL12 binary complex. The plan was that both projects would eventually converge to elucidate how FR-HMGB1-CXCL12-mediated signalling via CXCR4 differs from the CXCL12-exclusive pathways.

Knowledge of the binary structure would inform the regions of HMGB1 involved in CXCL12 binding. With this knowledge, specific mutations aimed at eliminating proinflammatory activities of HMGB1 could be designed. The potential mutations to test, which can be based only on the available literature on TLR-2, TLR-4, and RAGE HMGB1 signalling would be too numerous. By identifying residues or motifs that should be preserved or minimally modified, the size of theoretical construct panels decreases by at least one order of magnitude from several hundred constructs to less than a hundred. Furthermore, if the mutations are designed rationally to preserve CXCL12 binding capacity, the chances of retaining pro-regenerative activity of a designed HMGB1 construct increase.

1.6 Specific objectives

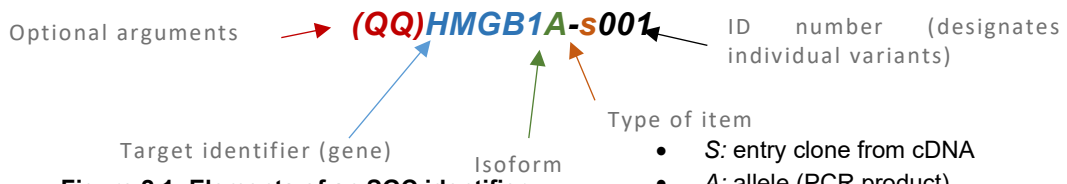
1. **Express and purify sufficient HMGB1 and CXCL12 to support structural studies:** Whilst commercial sources for both proteins are available, protein crystallisation and *in vivo* studies are destructive assays. Several hundred micrograms of protein are typically consumed in a single of these experiments. Furthermore, if the protocols described in the literature are followed, CXCL12 purification is particularly laborious. My first objective was to establish protocols for purification of both HMGB1 and CXCL12 in supra-milligram amounts. Repeated commercial purchase of such amounts is inviable, and it would greatly assist in the expression and purification of different constructs or mutants in objective 3. Characterisation of these proteins would also assist downstream studies, as information on the biophysical properties of both proteins is limited in the literature.
2. **Solve the structure of the FR-HMGB1-CXCL12 binary heterocomplex to identify binding contacts**: This objective was eventually superseded by the following due to technical limitations: **Identification of the CXCL12-binding residues of HMGB1.** As I advanced in my understanding of this interaction, it became clear that crystallising the complex was not feasible due to its molecular properties. Therefore, I decided to identify the residues involved in the interaction by techniques which did not require structural resolution, as well as attempt to elucidate the requirements for an HMG Box containing protein to emulate the pro-regenerative properties of HMGB1.
3. **Design of a panel of HMGB1 constructs which preserve CXCL12 binding but have reduced or absent proinflammatory activities.** With knowledge about which residues or motifs within HMGB1 bind CXCL12 and must be preserved, and the available literature on TLR-2, TLR-4 and RAGE, a panel of HMGB1 constructs can be designed. I would also need a suitable set of *in vitro* screening assays to select from the panel the final candidate for further refinement and *in vivo* analysis.

I have described the results of each of these objectives in an independent results chapter, although the process was often performed in parallel.

Chapter 2 Materials and Methods

2.1 Nomenclature and notes

- DNA sequences and recombinant proteins are identified using SGC nomenclature.



- *S:* entry clone from cDNA
- *A:* allele (PCR product)
- *C:* construct (allele cloned into a vector)
- *K:* clone (transformant in *E.coli*)
- *E:* Expression experiment
- *P:* Purified protein

Figure 2.1: Elements of an SGC identifier
HMGB1A-s001 designates the wild-type sequence of the target HMGB1 from the Mammalian Gene Collection (MGC), is the first isoform of this target (*A*) directly derived from the MGC (entry clone; *s*-), and also the first entry clone generated (*001*). The lack of optional arguments indicates no mutations from the MGC recorded sequence

- All kits and reagents were used as per manufacturer brochure unless otherwise indicated. Most chemical reagents can be substituted with equivalents, excepting antibodies and cell lines.
- This section is written with the goal of providing the reader with enough information to reproduce my results. Footnotes have been added to reference steps which were subjected to optimisation to facilitate comprehension; these are described in respective Results chapters.
- Academic citations follow Current Opinion in Cell Biology format (in-text number citation). Footnotes appear as superscript numbers after the relevant main text. Cross-references in this thesis follow this format: the first letter corresponds to the chapter (e.g., 3.3.VI stands for Chapter 3, section 3, subsection VI).
- Recombinant protein batches were used if meeting the following criteria:
    - *Purity:* Monodisperse SEC profile, single species in ESI-QTOF/MS (2.7.IV) with exact mass match, correct reduction state and biotinylation if present, only faint lower MW bands in SDS-PAGE (2.6.II). For *in vivo* and *in vitro* work, protein was assessed for endotoxin contents under < 4 EU/mg.
    - *DNA contamination, protocol 2.6.II:* $OD_{260/280}$ < 0.57, no extra bands visible on SDS/PAGE stained with ethidium bromide.
- HMGB1 residues are numbered after the N-terminal methionine (residue 0), since it is lost in HMGB1, and to keep consistency with published NMR spectra (BMRB 15148, 15149). CXCL12 residues are numbered from the mature protein (no signal peptide).

2.2 Statistical analysis

All statistical analysis was performed in GraphPad Prism (v. 8.31+). For single-variable analysis, either one-way Analysis of Variance (ANOVA) with results shown as mean±standard error of the mean (SEM) or Kruskal-Wallis (box-whiskers plot with Tukey whiskers) tests were applied, depending on homogeneity of variances[2]. When two independent variables were present, a two-way ANOVA was used. In animal models, a nested ANOVA (mixed-effects model ANOVA) with restricted likelihood estimation was applied if the assumptions were met. In these, each animal (n1= number of animals) became a subgroup (n2= repeated measurements) in the nested analysis, and all animals in each condition were considered as the replicates to avoid error deflation, and to account for biological variability separate from treatment effect. This approach also allows for unequal N numbers in the subgroups. If these were not met, a Kruskal-Wallis was applied, randomly selecting equal numbers of samples for each group to avoid skewing of data towards individual animals.

Kinetics (XY) data were fitted via nonlinear least squares (NLLS) method, rather than linearisation into a Scatchard plot, as heteroskedasticity is rarely absent and NLSS is less sensitive to it [136]. General comparisons between curves were performed with Aikaike Information Criterion, corrected (AICc) to avoid bias towards the simplest (null) hypothesis; multiple comparisons of best-fit values were performed via horizontal averaging of the replicates (NLLS fitting) and these were compared via either 2-way ANOVA or Brown-Forsythe/Welch one-way ANOVA (Dunnett T3 post-hoc test). When performing these, n=degrees of freedom+1 for each best-fit value; best-fit parameters were assumed to be representative of a Gaussian distribution.

Post-hoc comparisons were corrected via Holms-Sidak correction to decrease Type II error [137] for parametric tests, and Dunnett´s T3 for non-parametric. When statistical summaries for *p values* are provided, they follow the default GraphPad Prism format. Unless otherwise indicated, α= 0.05 (two-tailed test). Significance levels: $p > 0.05$ = n.s, $0.05 > p > 0.01$ = *, $0.01 > p > 0.001$ = **, $0.001 > p > 0.0001$ = ***, $0.0001 > p$ = ****. Some p values < 0.1 have been reported as numbers if appropriate (α= 0.1). Criteria for test selection are summarised in Figure 2.2. Exact technical/biological replicate numbers and tests used are indicated in the relevant figures. Data selection criteria and statistic test outputs are collected in in Appendix A5.

[2] ANOVA is robust against moderate departure from normality, provided that the variances are equal [282], and no non-parametric alternatives exist for e.g. 2-way or nested ANOVA. Tests for Gaussian residual distribution applied: Anderson-Darling (A2*), D'Agostino-Pearson Omnibus K2, Saphiro-Wilk, Kolmogorov-Smirnov. Distributions analysed: quantile/quantile (Q/Q) plot and heteroskedasticity plot.

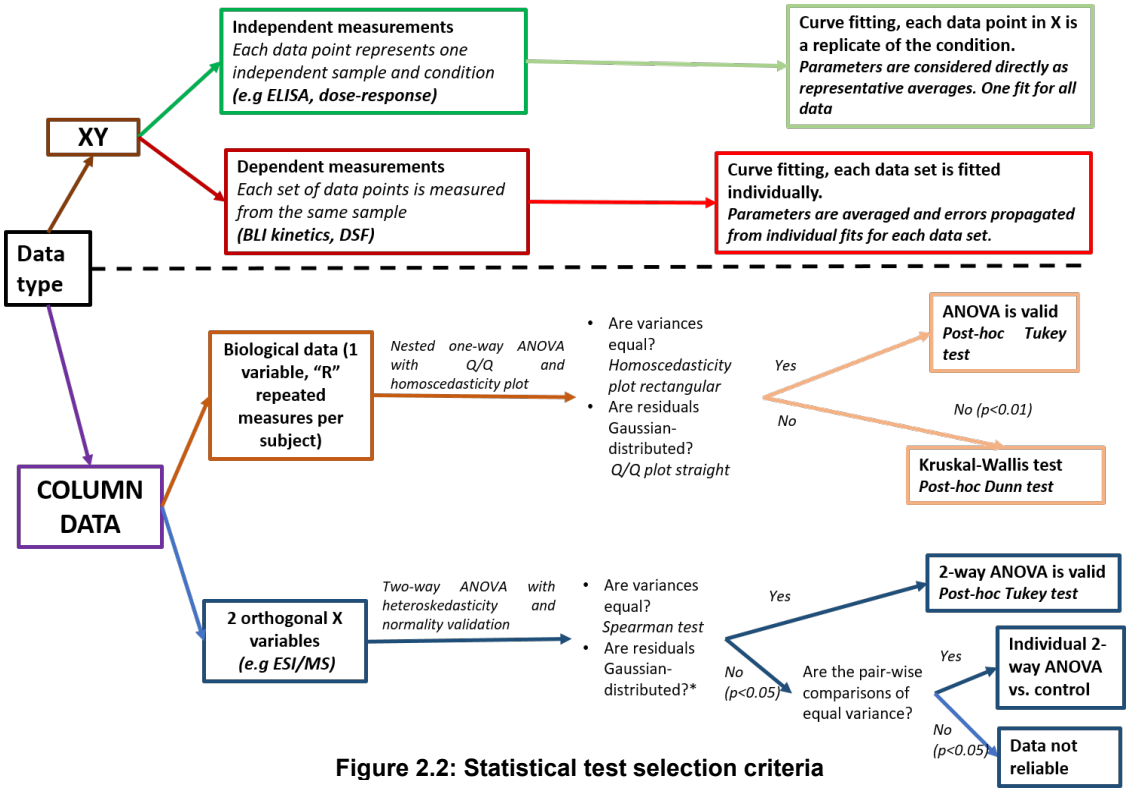


Figure 2.2: Statistical test selection criteria

2.2.I Free software

Tool/Database	URL	Usage case
CCPNMR Analysis suite	https://www.ccpn.ac.uk/v3-software	NMR data analysis. Version 3.02
EuGene [161]	http://bioinformatics.ua.pt/eugene/	Optimisation and redesign of protein sequences
EXPASY ProtParam [296]	https://web.expasy.org/protparam	Theoretical calculations of isoelectric point, extinction coefficient, mW, and aliphatic contents of proteins
PredTAT [297]	http://www.compgen.org/tools/PRED-TAT	Prediction of twin-arginine transporter signal peptides
ProtSA [145]	http://webapps.bifi.es/protsa	Theoretical calculation of surface areas for unfolded peptides and PDB ensembles
PyMol	https://pymolwiki.org/index.php/Windows_Install#Open-Source_PyMOL	Visualisation of protein structures and preparation of figures. Scripts for colouring residues based on hydrophobicity and charge were derived from the literature or the database.
SATurn [298]	N/A (internal SGC resource)	SGC database for sample organisation and ease of access to data within the Institute
SignalP [299,300]	http://www.cbs.dtu.dk/services/SignalP/	Prediction of *Sec* signal peptides
UNIPROT database [293]	https://www.uniprot.org/	Sequences and deposited structures for various proteins
Unimod database	http://www.unimod.org/modifications_list.php	Annotation of adducts in ESI/MS deconvoluted spectra of unknown identity
PAWS	http://prowl. rockefeller.edu	Mapping of peptide sequences to a given MW species

2.3 In-house enzyme production

Benzonase-Maltose Binding Protein (MBP), Ulp-1 protease, and glutathione-S transferase tagged (GST)-TEV protease were produced in-house from vectors available at the Structural Genomics Consortium (SGC). Expression was performed as described in 2.10 (ampicillin resistance). For Ulp-1 and GST-TEV proteases, the protocol was the same as described in Section 2.11, IMAC section. For Benzonase-MBP construct, outer membrane lysis was performed (see Section 2.12) and the protein was purified with amylose resin (NEB, E0821) according to manufacturer instructions.

In both cases, the resulting proteins were concentrated to 10 mg/mL in 50 mM HEPES pH 7.5, 0.3 M NaCl, 10% glycerol. GST-TEV protease and Ulp-1 were flash-frozen with liquid nitrogen and supplemented with 0.5 mM TCEP during purification; Benzonase-MBP was supplemented with 50% glycerol and 2 mM $MgCl_2$ and stored at -20ºC.

Lysozyme was purchased commercially (Sigma-Aldrich 10837059001).

2.4 Protein purification resins

Name	Format	Supplier	Item	Purification method	Protocol	Regeneration/CIP and storage
Ni-Sepharose 6, FF	50% slurry	GE Healthcare	17531	IMAC (nickel affinity)	2.3, 2.9, 2.11	Regeneration: 0.1 M EDTA in 20% ethanol (10 column volume/CV), 1 M NaOH (10 CV), 1 M acetic acid with 1% TX-100 (10 CV), 100 mM $NiSO_4$ + 20 mM Tris-HCl pH 8.0 (2 CV). 40 CV of water was applied between steps. Storage: 20% ethanol.
	GraviTrap 1 mL column		11-0033-99			
Ni-NTA 5 mL Superflow	5 mL prepacked cartridge	Quiagen	30761		2.12	
HiScreen CaptoS ImpAct	7.8/100 Tricorn column (4.7 mL)	GE Healthcare	17-3717-47	Cation exchange	2.12	Regeneration: 1.5 M NaCl in 20 mM HEPES. Storage: 0.2 M Sodium acetate in 20% ethanol.
Monomeric Avidin/Agarose resin	50% slurry	Thermo Scientific	20228	Biotin affinity	2.11	Regeneration: 0.1 M Glycine, pH 2.8 (40 CV), water (20 CV). Storage: PBS 1X with 0.02% sodium azide.
Strep-Tactin XT	50% slurry	IBA Bioscience	2-4010	Biotin affinity	2.11	10 mM NaOH (40 CV), water (40 CV) Storage: 100 mM Tris-HCl pH 8.0, 1.5 M (NH4)2SO4, 150 mM NaCl, 1 mM EDTA, 0.02% sodium azide.
Superdex 75 pg (Prepacked Tricorn columns)	10/300 (24 mL)	GE Healthcare	29148721	Size exclusion	2.3, 2.9, 2.11, 2.12	CIP: 0.5 M NaOH (2 CV), reverse flow at 27 cm/h Storage: 0.2 M Sodium acetate in 20% ethanol.
	16/600 (120 mL)		28989333			

2.5 *E. coli* culture media preparation

All media were prepared with 18.2 mΩ water and autoclaved. Stocks were either autoclaved or filtered through a 0.22 µM filter.

- Luria-Bertani (LB) broth, 16X: 160 g/L tryptone, 80 g/L yeast extract, 160 g/L NaCl, pH 7.2. Diluted to either 2X or 1X before use. For LB agar, 2% w/v agar powder was added before autoclaving.
- Terrific Broth (TB): The below solutions were mixed to 1X concentration, plus 0.01% antifoam-204 solution and 1X trace metal solution for supplemented TB media.
    - Base, 10X: 120 g/L tryptone, 240 g/L yeast extract, 40 g/L glycerol.
    - TB salts, 10X: 125.13 g/L K_2HPO_4, 23.5 g/L KH_2PO_4.
    - TB+ supplement, 10X (filtered): 16% w/v glycerol, 10% glucose, 250 mM $(NH_4)_2SO_4$, 10 mM $MgSO_4$, 10X trace metals. Sterile filter in a sterile container.
- Trace metal solution, 1000X (filtered): 50mM $FeCl_3$ (13.5 g/L), 20mM $CaCl_2$ (2.94 g/L), 10mM $MnCl_2$ (1.96 g/L), 10mM $ZnSO_4$ (2.88 g/L), 2mM $CoCl_2$ (0.48 g/L), 2mM $CuCl_2$ (0.34 g/L), and 2mM $NiCl_2$ (0.48 g/L), in 0.1M HCl. It is recommended to first dissolve the $FeCl_3$ in 0.5X final volume of 0.2M HCl and adjust volume at the end. Sterile filter in a sterile container.
- Super-Optimal Broth (SOB)*:* 20 g/L tryptone, 5 g/L yeast extract, 0.5 g/L NaCl, 0.1862g/L KCl. 4.132 g/L $MgCl_2$ were added after autoclaving (from 2M sterile-filtered stock). 20 mM glucose was added for SOC (catabolite repressed) media.
- M9 medium:
    - 1X M9 salts (from 5X stock: 80 g/L Na_2HPO_4 (dihydrate), 20 g/L K_2HPO_4 (anhydrous), 5 g/L NaCl, pH 7.2-7.3)
    - 1mM $MgSO_4$
    - 1X mL M9 micronutrients (from 100X stock: 0.5 g/L $FeSO_4$, 5 mg/L $ZnCl_2$, 1 mg/L $CuSO_4$, 5 g/L EDTA)
    - 0.2 X Trace metals mixture
    - 8 g/L glucose; U-^{13}C if labelling is necessary (filtered)
    - 1 g/L NH_4Cl; U-^{15}N if labelling is necessary (filtered).
    - 0.3 mM $CaCl_2$ (filtered)
    - 10 mL BME essential vitamins (Sigma B6891).

2.6 Electrophoresis

2.6.I Agarose DNA electrophoresis

DNA samples were separated in agarose gels (1.5% w/w or 0.75% for bands expected to be > 3 kilobases) in TAE buffer (50X stock, 242g/L Tris-HCL, 57.1 mL/L glacial acetic acid, 50 mM EDTA) and stained with 1:15000 PeqGreen dye for UV visualisation, according to published protocols [138]

2.6.II SDS-PAGE protein electrophoresis

All proteins purified by chromatography were assessed by SDS-PAGE (4-12% acrylamide gradient precast gel) in MES/SDS buffer [139]. Protein bands were stained with Colloidal Coomassie stain ([140]), and further identified by tryptic digest MS/MS (2.7.IV).

2.7 Mass spectrometry (MS)[3]

2.7.I Instruments

- High-pressure liquid chromatography (HPLC) stack (Agilent 1100 HPLC).
- Electrospray-ionisation quadrupole time-of flight (ESI/QTOF) spectroscope (Agilent Q-TOF 6545), in positive ion mode.
- Gas-tight glass syringe and sample pump (e.g., Sigma Z401366) for direct injection.

2.7.II Solutions, all liquid chromatography (LC)/MS grade

- Buffer A: 0.1% formic acid (FA), in water.
- Buffer B: 0.1% FA, in methanol.
- Buffer C: 0.1% FA, in acetonitrile.
- Volatile sample buffer: 50/200 mM ammonium acetate, pH 6.8.

2.7.III Tryptic digest tandem mass spectrometry (MS/MS)

Bands from SDS-PAGE gels of suspect identity were excised and submitted to the SGC open access MS platform for protein identification, performed by Dr. Rod Chalk, Dr. Tiago Moreira and Oktawia Borkowska, as published [141,142]. Data analysis (peptide mapping) to annotate protein identity was performed with and the MASCOT search engine against the Uniprot and SGC databases.

[3] I would like to acknowledge Dr. Rod Chalk and Dr. Tiago Moreira for both devising these methods as well as the extensive assistance they offered me in learning them.

2.7.IV LC-ESI-QTOF/MS

1. Samples were diluted to 16 ng/mL (1 µg) in buffer A: blanks (1% formic acid) were run between injections.
2. Instrument mass calibration range: 600-3000 m/z. See [141] (pp.378, table 2) for detailed instrument settings. Data were acquired for the entirety of the chromatographic run.
3. Chromatographic separation took place at 40°C in either:
    a. 2.1 mm × 12.5 mm Zorbax 5 µm 300SB-C3 guard column and column holder (Agilent), for soluble proteins. LC settings; see Table 1 (pp.378) in [141]
    b. RP-18e Chromolith 25 × 2 mm (Merck), for separation and detection of detergent in LPS-free samples. LC settings; see Table 9 (pp.378) in [141].

2.7.V Native ESI-QTOF/MS

1. Samples were desalted via size-exclusion spin columns (MicroSpin 6, BioRad7326225) for a minimum of 3 times in volatile buffers, then concentrated to 1-3 mg/mL (optimal concentration for injection varied depending on the protein).
2. Instrument mass calibration range: 1000-20000 m/z. See [141] (pp. 388-390, Table 8) for detailed instrument settings.
3. A gas-tight syringe and tubing were washed 3X with 1.5 mL of acetonitrile, then were used to inject acetonitrile directly into the nebuliser at 660 µL/h until a stable ion stream was observed in the total ion chromatogram (TIC). The process was repeated with volatile buffer and data from 10'' in the stable phase were collected as a blank.
4. Protein samples were injected at 360 µL/h: data was collected as above provided sufficient signal-to-noise ratio was observed in the live spectra for at least 30''.

2.7.VI Quantification of disulfide bonds

1. Protein samples were diluted to 1 mg/mL in 0.1 M ammonium bicarbonate, pH 8.0 and 8 M urea.
2. Parallel samples for each protein were treated with iodoacetamide (10 mM) for 45 minutes in the dark, with or without addition of 2 mM TCEP.
3. Alkylated samples were measured via denaturing ESI-QTOF/MS. The presence of a disulfide bridge was characterized by a mass difference of 114 Da (two iodoacetamide adducts) between the samples treated with iodoacetamide and TCEP, absent in the sample treated with iodoacetamide alone. If both masses were equal, the protein was assumed to be fully reduced.

2.7.VII ESI/MS data analysis

M/Z spectrum extraction from the TIC and deconvolution were performed with Agilent MassHunter Workstation software (v 7.0). Ambiguous post-translational modifications were annotated with the information in the Unimod database. Sodium adducts (22*n series) were prevalent in denaturing ESI/MS samples, as well as deconvolution artifacts (sodium series either +11*n; ½x mass artifact or +44*n, 2x mass artifact). In native ESI/MS (no sodium present), ammonium adducts (+18*n) could be observed instead. Charge states were assigned to each species in the M/Z spectrum according to the following formula, assuming that each charge acquired in positive mode comes from protonation of e.g., amine groups:

$$protein\ mass = \left(\left(\frac{protein\ mass}{z}\right) - H^+\right) * z$$

***H*⁺** = proton mass (1.07 Da)
MW = deconvoluted molecular weight
Z = charge state

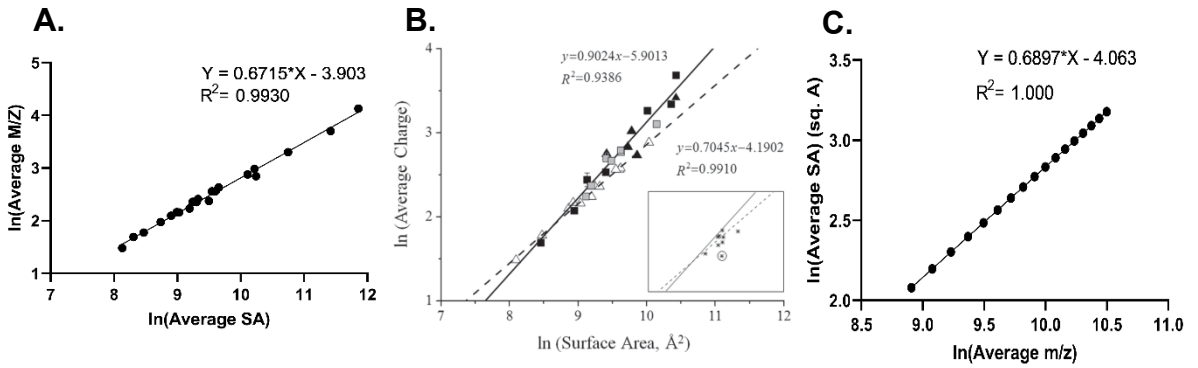


Figure 2.3: Correlation between m/z and surface area

A. Standard curve for ESI/MS charge states under nondenaturing conditions for several proteins, fitted from the data reported in Kaltashov et. al (2005). **B.** Standard curve for ESI/MS charge states of non-intrinsically disordered proteins under nondenaturing conditions (dashed line, white triangles) and denaturing conditions (continuous line, black triangles). Intrinsically disordered proteins (gray square: nondenaturing, black square: denaturing) are also shown. Image extracted from Testa et. al, 2011 (Figure 3 in the original publication). **C:** combined equation from both sets of data which I used for my calculations for globular protein denaturing surface area.

Oligomers were separated by the continuity of the series. Some M/Z states are identical across oligomers, but a discontinuity in the M/Z series invariably represents distinct oligomeric states. For example, monomer Z= +8 has identical m/z to a dimer with Z=+16; if a +7-+8/+16-+9 series is observed, the middle state must be +8 as with the dimer +15 and +17 would be observed instead. Individual charge states were assigned a SASA according to the formulas derived from the literature (Figure 2.3, [143,144]). Typically, a given conformation will follow a unimodal distribution. Thus, a distribution with multiple local maxima implies multiple conformational states being present. Relative intensity for each peak was extracted as the "abundance percentage" readout in the software, with the baseline adjusted to the lowest point in the TIC. The solvent accessible surface area (SASA) for each charge state distribution was then averaged to estimate the SASA of each oligomer:

$$SASA\,(\text{Å}) = \sum\left(\frac{I_{\frac{m}{z}}}{\sum I_{\frac{m}{z}}} * SA_{\frac{m}{z}}\right)$$

$\boldsymbol{I_{\frac{m}{z}}}$**=** relative abundance of a given M/Z state

$\boldsymbol{SA_{\frac{m}{z}}}$ = surface area of a given M/Z state

HMGB1 possesses several intrinsically disordered regions, rather than being perfectly globular. I chose to use this approach as the models have been validated also with non-spherical globular proteins, and therefore would be better applicable than estimation based on the radius. The studies also showed that ionisation is not be affected by the isoelectric point [143,144], as the backbone amines will acquire proton charges as well, a critical parameter for highly charged molecules such as HMGB1 and CXCL12. These data were compared with SA data from PDB-available structures for HMGB1 and CXCL12. Both calculations with PyMol **get_area** function (**dot_density** = 4, **dot_solvent** = 1 for folded SASA, and **dot_solvent** = 0 for total unfolded SASA), and from the server ProtSA [145] were considered. The ProtSA server was also used to validate surface area calculations from unfolded proteins without a PDB structure available (e.g., HMGB1 constructs).

2.8 Cloning and test expression from *E. coli*

2.8.I Cell lines and enzymes

All these cells are routinely expanded in-house [146] and frozen as chemically competent aliquots in liquid nitrogen, stored at -80ºC before use as part of the Structural Genomics Consortium (SGC) pipeline. Inducible strains were grown in 1% glucose to repress background *lac* activity.

CELL LINES	Inducible?	Antibiotics (µg/mL)	Growth temp. (°C)	Induction temp. (°C)	Features
Mach-1 T1R (Invitrogen, C862003)	No	None	37	N/A	Fast-growth DNA storage strain
BL21(DE3)-R3-pRARE2	*lac* operon expressing T7 polymerase	Chloramphenicol (34)		18	pRARE2 plasmid (human tRNAs)
BL21(DE3)-R3-pRARE2-BirA		Chloramphenicol (34), Streptomycin (50)			pRARE2 plasmid (human tRNAs), Constitutive BirA expression (overbiotinylation)
BL21(DE3)-pLysS-CyDisCo PMJ226 cells ***Gift from DPhil Jesse A. Coker, who generated in-house competent cells in pLysS background***		Chloramphenicol (34)	30 due to *Erv1p* toxicity [210,212]		pRARE2 plasmid (human tRNAs), cytoplasmic disulfide-competent [301]

ENZYMES	Supplier	Catalogue Nº	Assay buffer (at 1X)
BfuAI restriction enzyme	New England Biolabs	R0701	NEB 3.1
BsaI restriction enzyme	New England Biolabs	R0535	CutSmart
DpnI restriction enzyme	New England Biolabs	R0176	CutSmart
Herculase II Fusion DNA polymerase	Agilent Technologies	600679	Herculase II buffer
MyTaq polymerase	Bioline	BIO-2110	MyTaq Red Reaction buffer
T4 polymerase	New England Biolabs	M0203	CutSmart

Table 2.1: Enzymes and cell lines used in cloning

NEB CutSmart buffer (from 10X stock): 50 mM potassium acetate, 20 mM Tris acetate, 10 mM magnesium acetate, 0.1 mg/mL bovine serum albumin (BSA), pH 7.9 at 25ºC. **NEB 3.1 buffer (from 10X stock):** 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 0.1 mg/mL BSA, pH 7.9 at 25ºC. Herculase II buffer and MyTaq Red buffers are proprietary formulas supplied with the enzymes; Herculase II also is supplied with 100 µM dNTP

2.8.II Plasmids

Vectors pFB-CT10HF, pNIC-CTHF, pNIC-SufI-NHT-cTStII, pNIC-OmpA-NHT-cTStII, pNIC-TolB-NHTcTStII, pDSBC-HT-Bio, pNIC-Bio3 and pSUMO-LIC were sourced from the Structural Genomics Consortium (SGC) at Oxford. pFB vectors were selected with ampicillin (200 µg/mL) or carbenicillin (50 µg/mL). pNIC vectors were selected with 50 µg/mL kanamycin instead, in addition to antibiotics for the specific *E. coli* strains. DNA for the human HMGB1 and CXCL12 genes was sourced from the Mammalian Gene Collection repository present at the SGC [147]. All vectors constitutively express levansucrase when stored in *E. coli* (SacB) for negative selection in the presence of sucrose. This coding sequence can be excised by a specific restriction enzyme generating the LIC cloning site. They are all covered by open access (UBMTA) licenses.

Except for pFB-CT10HF-LIC (ampicillin/gentamycin resistance, pUC/F1 origin; *E. coli/*baculovirus shuttle vector), which was only used in *E. coli* for DNA storage of entry clones, all of these vectors were exclusive for bacteria and expression-capable. Common properties of these bacterial vectors were:

- Replication origins for *E. coli* (pBR322 and phage F1)
- Kanamycin resistance.
- Expression regulated via the lactose operon (*lacI)* controlling expression of T7 polymerase, which in turn induced expression of the inserted genes in the presence of IPTG/lactose.
- Specific restriction enzyme for LIC cloning, tags and applications are summarized in Table 2.2

Name	N-terminal tag	C-terminal tag	Restriction enzyme for LIC	Signal peptide	Use
pNIC-CTHF	None	(TEV site)-6X His-FLAG	BfuAI	None	Cytoplasmic expression of protein
pNIC-SufI-NHT-cTStII	6X His-(TEV site)	None	BsaI	SufI (Twin arginine translocase path)	Periplasmic export of recombinant protein
pNIC-OmpA-NHT-cTStII	6X His-(TEV site)	None	BsaI	OmpA (*sec SpI* system, unfolded translocation into periplasm)	Periplasmic export of recombinant protein
pNIC-TolB-NHTcTStII	6X His-(TEV site)	None	BsaI	OmpA (*sec SpII* system, unfolded translocation into membrane)	Periplasmic export of recombinant protein
pNIC-Bio3	6X His (TEV cleavable)	Biotinylation epitope (Avi tag)	BsaI	None	Cytoplasmic expression of biotinylated protein ( )
pDSBC-HT-Bio	DsbC-6X His-(TEV site)	Biotinylation epitope (Avi tag)	BsaI	OmpA (*sec SpI* system, unfolded translocon into periplasm)	Periplasmic export of recombinant protein with a chaperone to assist in disulfide formation. A SUMO site was inserted between the TEV site and the protein coding sequence.
pSUMO-LIC	His6-(SUMO motif)	His6	BsaI	None	Template to insert SUMO site in pDsbC vector

Table 2.2: Purification tags and uses of SGC *E. coli* expression vectors present in this work

2.8.III Methods

Primer pairs for all constructs designed are collected in Suppl. data A1-3-B. For a list of the specific constructs designed, see 2.11 (HMGB1), or 2.12 (CXCL12). PCR reactions were all performed according to the following scheme: variables are defined for each polymerase in the relevant protocol. Primers were designed to encompass both the first 8-10 residues of the protein adjacent to the ligation site and the designed Ligation Independent Cloning site present in the vectors, with a Tm_{50} between 55 and 60ºC.

1. 95ºC, 10'.
2. 25X cycles:
   a. 95ºC, 30 sec.
   b. [Primer Tm_{50}]ºC, 1' .
   c. [Extension time]' at [Extension temperature]ºC
3. [Extension temperature]ºC, 10'.

2.8.III.i Herculase II polymerase chain reaction (PCR)

Complementary DNA (cDNA) for HMGB1 and CXCL12 was sourced from the Mammalian Gene Collection at the Structural Genomics Consortium. Sequences directly amplified from the MGC were sequenced to verify identity. All reactions were made up to final volumes in nuclease free water. [Primer Tm] = variable. [Extension time] = 1 minute per kilobase of product. [Extension temperature] = 68ºC

Reaction mix (25 µL)

- 5 µL Herculase 5X reaction buffer
- 2.5 µL primers (from 10 µM stock) and 1.5 µL template (from 100 µg/mL stock)
  - For megaprimer PCR, 5 µL of each PCR product were used instead of the primer mixtures and templates.
- 3.75 µL dNTP mixture (10 µM stock, from 100 µM stocks of each)
- 0.2 µL Herculase II polymerase

The resulting PCR products were kept at 4ºC for immediate use or frozen at -20ºC for long-term storage. All reactions were designed so that the PCR product sizes were over 100 kilobases (kB). When the antibiotic resistance of the entry clone template matched that of the target vector for cloning, the PCR products were treated with Dpn1 enzyme to degrade the methylated DNA template (1 unit per PCR product, 60' incubation at 37ºC). All PCR products were purified using PureLink column purification kits (ThermoFisher K310001) and kept either at 4ºC for short-term or at -20ºC for long-term storage.

2.8.III.ii Ligation Independent cloning (LIC)

1. Insert preparation

PCR products were treated with T4 polymerase to generate overhang fragments via addition of a single dNTP to the insert and the complementary base to the vector reactions to promote T4 polymerase exonuclease activity. The reactions were incubated at 22ºC for 30' followed by incubation at 75ºC for 20' to inactivate the enzyme. Products were used without purification.

Reaction mix (10 µL)

- 5 µL purified PCR product
- 1 µL 25 mM dNTP
- 1 µL CutSmart/NEB Buffer 2.1
- 0.1 µL BSA, molecular biology grade
- 0.25 µL DTT

2. Vector preparation

Vectors were digested with the appropriate restriction enzyme to remove the SacB "stuffer" fragment, which results in the apparition of a 2 kB fragment in agarose gel electrophoresis.

- *pNIC-CTHF-LIC, pNIC-CTH0-LIC, pFB-CT10-StII-LIC*: BfuAI digestion (3 hours at 60ºC), vector treated with dGTP, insert treated with dCTP
- *All other plasmids:* BsaI digestion (2 hours at 37ºC), vector treated with dCTP, insert treated with dGTP.

Vector digestion reaction mix:

- 2 µL BfuAI (5 units) / BsaI (10 units)
- 5 µg plasmid DNA
- 5 µL NEB Buffer 3.1 (for BfuAI) / CutSmart (BsaI)
- Final volume 50 µL with nuclease-free water.

The resulting digests were purified with PureLink PCR purification kits and treated with T4 polymerase in the same way as the inserts with the corresponding complementary dNTP, scaling the final volume to 100 µL.

3. Cloning

a. 1 µL of T4-treated vector was mixed with 2 µL T4 treated insert. Annealing was performed by 30' incubation at room temperature in a sterile container e.g., PCR tube.
b. Mach-1 Competent cells were freshly thawed on ice and 40 µL[4] added to the DNA tube. This mixture was equilibrated for 30' on ice.
c. Cells were heat-shocked for 45'' in a 42ºC water bath, then chilled on ice for 1'.
d. 100 µL of SOC medium was added to the mixture and incubated for 1 hour at 37ºC.
e. Transformants were selected in plates with the appropriate antibiotics and 5% sucrose to negatively select SacB containing cells (intact vector). Plates were air-dried and incubated overnight to pick colonies.

[4] This amount can be adjusted depending on the competence of the cells. Same applies for the plated volumes.

4. Colony PCR

a. A single colony was inoculated first in 1 mL of selective 2xLB medium, on a 2 mL 96-well block, and cultured overnight.
b. The inoculation loop was dipped in 20 µL of MyTaq polymerase master mix (according to manufacturer instructions).

Vector	Primer name	Sequence
pFBac-CT10HF	FBAC-3	TTAAAATGATAACCATCTCG
	FBAC-2	GTGTGGGAGGTTTTTTAAAGCAAGTAAA
All other vectors	pLIC-F	TGTGAGCGGATAACAATTCC
	pLIC-R	AGCAGCCAACTCAGCTTCC

c. [Primer Tm] = 52ºC. [Extension time] = 30´´ per kilobase of product. [Extension temperature] = 72ºC. Sequencing primers are shown above.
d. Correct clones were verified by detection of a PCR product the same size of the insert + 200 bp product of the colony PCR primer amplification (1000 bp for pDsbC vector) in agarose gel electrophoresis.

5. Miniprep and storage

a. 120 µL of overnight culture for positive colonies were mixed with 30 µL of 60% in a 96-well plate (both sterile) at -80ºC.
b. Remaining culture was centrifugated 3000 RCF for 10' to harvest cell pellets.
c. Plasmid DNA (construct) was purified via commercial miniprep kits (e.g., Macherey-Nagel NucleoSpin, 740952) and stored until further use.

2.8.III.iii Test expression

1. 3 µL of construct DNA were added to 20 µL of the desired expression strain. Steps 2.8.III.ii 3b-d were followed.
2. Selection was performed in LB agar plates with antibiotics appropriate for the vector and strain, and positive colonies grown overnight in 2X LB.
3. Glycerol stocks were made as in 2.8.III.ii1.a
4. 20 µL of overnight culture were inoculated in 1 mL of TB medium with antibiotics,[5] in a 96-well 2 mL block, with 700 revolutions per minute (rpm) orbital shaking.
5. Grow at 37ºC until OD_{600} 2-3 is reached.
6. Cool the cultures at 18ºC.
7. Induce cells with 0.1 mM Isopropyl β-d-1-thiogalactopyranoside (IPTG, lactose analogue) and culture overnight at 18ºC
8. Harvest at 3000 RCF, 20' and discard supernatant.

[5] For CyDisCo cells, TB+ was used instead.

2.9 Test purification from *E. coli*

2.9.I Solutions and enzymes

To purify reduced HMGB1 constructs, add 0.5 mM TCEP to all buffers prior to use.

- Lysis buffer (LBt): 0.5 M NaCl, 0.1 M HEPES pH 8.0, 10% glycerol, 10 mM imidazole, 1 mM $MgSO_4$.
- Before purification, the following were added: 3 µg/mL Benzonase-MBP, 0.5 mg/mL lysozyme, 1:1000 protease inhibitor cocktail (Calbiochem, set III, EDTA-free), 0.1% DDM.
- Wash buffer (WBt): 0.5 M NaCl, 50 mM HEPES pH 8.0, 5% glycerol, 25 mM imidazole.
- Elution buffer (EBt): 0.5 M NaCl, 50 mM HEPES pH 8.0, 5% glycerol, 500 mM imidazole.

2.9.II Method

1. Cell pellets were resuspended in 200 µL of lysis buffer with a multichannel and orbital shaker at 1500 RCF for 15', then fully frozen at -80ºC.
2. Pellets were then thawed at 22°C for > 20' and resuspension at 1500 RCF, prior to centrifugation at 3000 RCF for 10'.
3. 200 µL of each supernatant was transferred to a MultiScreen filter plate (Merck MSBVN1210) with 50 µL of resin suspension/well (Ni-Sepharose 6, FF). Batch-binding was performed at 18ºC for 1 hour, shaking at 0.0025 RCF.
4. After binding, lysate was filtered through a vacuum pump with an appropriate adaptor.
5. Wells were washed 4 times with 200 µL WBt, and the plate was further spun for 3' at 300 RCF to fully remove buffer.
6. Proteins were eluted by incubation for 20' of the wells in 40 µL EBt over a 96-well V-bottom plate, then collected by centrifugation at 300 RCF, 3'.
7. Bands of interest in SDS-PAGE were submitted for tryptic digest MS/MS analysis (2.7.III)

2.10 Large scale expression from *E. coli*

If producing biotinylated proteins, 10 mL D-biotin solution in sterile phosphate buffer saline (PBS) was added twice at steps 4 and 6. Orbital shaking (0.45 RCF) was used in all cases.

1. Overnight precultures were grown in 2X LB, from a freshly streaked agar plate with antibiotics.
2. Media was dispensed in flasks sterilised by decontamination in Virkon 2% (30´) and autoclave (121°C, 15 min).
3. 1:50 preculture/media ratios were used for inoculation. Cultures were grown until OD_{600} 1.5-2 (TB, TB+) or 0.7 (M9) is reached.
4. Cultures were cooled at 18ºC for 1 hour.
5. Expression was induced with 0.5 mM IPTG (0.25 for CXCL12, 0.1 for M9 media) over 16h.
6. Pellets were harvested by centrifugation at 4000 RCF, 20' for future purification.

2.11 Purification of HMGB1 from *E. coli* cell pellets

A full list of all HMGB1 constructs designed during this project, and the specific applications can be found in Appendix A1. DNA and coding sequences are collected in either Suppl. data A1-3-B to D (constructs used in Chapters 3 and 4) or Suppl. data A1-5-A to C (HMGB1 constructs used in Chapter 5). Purification was preferably performed at 4°C after lysis.

2.11.I Buffers

- 4X Buffer A: 4 M NaCl, 0.2 M HEPES pH 7.5
- 5X Buffer B: 2.5 M NaCl, 0.25 M HEPES pH 7.5, 25% glycerol
- SEC storage buffer: 0.2M sodium acetate in 20% ethanol
- Strep-Tactin base buffer: 50 mM HEPES pH 7.5, 750 mM NaCl.

2.11.II Methods

1. **Cell lysis**
    a. Induced cell pellets (from 2.10) were fully resuspended (7 mL/g) in 1X Buffer A + 10 mM imidazole + 5% glycerol, supplemented with 3 µg/mL Benzonase-MBP, 0.5 mg/mL lysozyme, 1:1000 protease inhibitor cocktail and 1 mM $MgSO_4$. The solution was aliquoted in 50 mL tubes (30 mL each), plus 0.1 mM D-biotin for biotinylated targets.
    b. To lyse the cells, 0.5% TX-110 was added before freezing overnight at -80ºC, followed by thawing at 22°C for 1 hour.
    c. Cell pellets were then removed by centrifugation at 6500 RCF, 45', and filtered through 5 µM membranes if needed.
2. **First affinity step (IMAC)**
    a. Supernatant was flown through in a pre-equilibrated in 1X buffer A His GraviTrap column fitted with LabMate extender.
    b. Remaining lysate was removed with 10 mL of 1X Buffer A.
    c. Pipette 2 mL of 0.4X Buffer A + 1 mM $MgSO_4$ + 3 µg/mL Benzonase-MBP and incubate for 30'.
    d. Contaminants were removed with 15 mL of 1X Buffer B + 30 mM Imidazole.
    e. IMAC elution was performed directly into a PD-10 column, equilibrated in 1X Buffer B + 20 mM imidazole, with 2.5 mL 1X Buffer B + 500 mM imidazole.
    f. Protein was recovered from PD-10 with 3.5 mL of 1X Buffer B + 20 mM imidazole.
    g. 1 OD_{280} GST-TEV for every 20 OD_{280} was added to pooled fraction and incubated overnight.
    h. GraviTrap columns were equilibrated in 1X Buffer B + 20 mM imidazole (if performing reverse IMAC).

3. Optional post-IMAC affinity steps (if required)

Reverse IMAC (if cleaving tag)

a. TEV was removed by flow back through the previously used GraviTrap column in IMAC.
b. The column was washed with further 2.5 mL 1X Buffer B + 20 mM imidazole to maximize recovery

Biotin-affinity (for pNIC-Bio3)

a. Sample was diluted to 5 CV and flown through Strep-Tactin XT resin (1 mL for every 2 OD) or Monomeric Agarose resin (2 mL for every 0.5 OD)
b. Column was washed with 30 CV Strep-Tactin base buffer, then with 1 CV 100 mM Biotin in 0.2X Strep-Tactin base buffer
c. Elution was done by incubating for 1h with 3 CV 100 mM Biotin in 0.2X Strep-Tactin base buffer.

4. Size Exclusion Chromatography (SEC)

This protocol was also applied for analytical SEC experiments, such as those described in Figure 3.16.

a. Samples were concentrated to the appropriate injection volume (e.g., 0.33 mL)
b. When required, endotoxin removal from samples and chromatography system was performed (see 2.13)
c. The chromatography system was equilibrated in either sterile PBS (for biological work) or 0.2X StrepTactin base buffer (for NMR)
d. Samples were injected in either a Superdex 75 pg 10/300 (0.3 mL) or 16/600 (5 mL), and eluted at 27 cm/h.
e. Pooled fractions were concentrated and snap frozen. 1 mM TCEP was added if reduction was desired; samples were incubated for 1 h before analysis of reduction state.

2.12 Purification of CXCL12 from *E. coli* periplasm

A full list of all CXCL12 constructs designed during this project, and the specific applications can be found in Suppl. data A1-3-E. Primer pairs for all constructs are collected in Suppl. data A1-3-B; DNA and coding sequences are collected in Suppl. data A1-3-F/G.

2.12.I Buffers

- 5X Buffer A.b: 0.2M HEPES pH 7.5, 1 M sucrose, 1 mM EDTA, 50 mM imidazole
  - *Added before use: 1 mg/mL lysozyme, 2X protease inhibitor tablets, 3 µg/mL benzonase.*
- 5X Buffer B.b: 2.5 M NaCl, 0.25 M HEPES pH 7.5, 25% glycerol
- IMAC column regeneration buffers (from 2.11)
- Cation Exchange buffer A.c: 0.2 M NaCl, 20 mM HEPES pH 8.0
- Cation Exchange buffer B.c: 1.5 M NaCl, 20 mM HEPES pH 8.0
- SEC buffer: 10 mM HEPES, 150 mM NaCl, pH 7.5

2.12.II Methods

The method described here pertain to outer membrane lysis (pDsbC-expressed proteins). For purification from CyDisCo strains, replace crude IMAC purification with that described in 2.11.

1. Induced cell pellets (from 2.10) were resuspended in 2.5 mL/g Buffer A.b 5X at room temperature and stir for <45': an acrid smell indicated full outer membrane lysis, then diluted with 4 volumes of 18.2 mΩ water, and stirred for 10'. EDTA was neutralised afterwards by addition of 1 mM $MgSO_4$.
2. Supernatant was collected by centrifugation (16000 RCF, 1 h), and loaded at 10 mL/min in serial 5 mL Ni-NTA columns (1 column/6 L cell culture) in a fast-pressure liquid chromatography (FPLC) system (GE Healthcare Äkta Xpress).
3. Gradient elution was done stepwise (buffers A.b/B.b): 10 mM-25 mM, 10 CV; 25-500 mM 8 CV. Protein identity was verified by SDS-PAGE and LC-ESI-QTOF/MS.
4. 1 OD_{280} (absorbance unit at 280 nm) Ulp-1 protease was added to the pooled fractions for every 20 OD_{280} CXCL12, and the sample was diluted buffer A.c to < 3 OD_{280}/mL to prevent precipitation. Enzymatic cleavage was done in dialysis against > 50X volumes of Buffer A.c overnight and verified as in the previous step.
5. Cleaved CXCL12 was loaded in serial CaptoS HiScreen columns (1 column ≈ 150 OD_{280}). Gradient elution profile: 0.2-0.4 M, 8 CV; 0.4-1.5 M, 6 CV, 1.5 M, 4 CV.
6. Pooled fractions were concentrated to 5 mL and further purified by SEC (as per 2.11), in SEC buffer.

2.13 Endotoxin-free protein production

For non-intrathecal injection of substances in animals and humans, the USA Food and Drug Administration (FDA) establishes a limit of 5 endotoxin units (EU)/kg of body weight due to its toxic effects *in vivo*. EU are typically monitored by *Limulus* amoebocyte lysate (LAL) assays or their derivatives and threshold depends on the application (Figure 2.4). In cell culture assays, 0.1 EU/mL is assumed to be safe as it is below the threshold for immune cell activation [148].

$$5\frac{EU}{kg} * 0.025\ kg \over \mu g\ protein = \frac{0.125\ EU}{\mu g\ protein}$$

Figure 2.4: Calculation of safe pyrogen levels for recombinant protein injection into mice

Whilst there are *E. coli* strains which express a modified LPS which has little to no pyrogenic activity in human immune cells as it is not recognised by human TLR-4 [149], this endotoxin is recognised by murine TLR-4 and induces LPS responses. As several of my experiments involved mouse models, or human cells expressing murine TLR-4, this was not suitable. Whilst LPS removal media exists, they are either ineffective for specific LPS-binding proteins such as HMGB1 (e.g. polymyxin-B and positively charged resins) [150] or require addition of calcium to the samples (e.g. Hyglos EndoTrap). Both resins also have a limited number of uses which leads to increased costs for high-throughput applications, and can leech ligand into the sample leading to toxicity *in vivo* [151][6].

Instead, I removed endotoxins post-purification via phase separation in Triton-X114 [150,152]. Triton-X114 is known to decrease the viability of cells, absorbs in the 280 nm range ($\varepsilon 1\%_{280}$ = 20.55) and prevents the LAL reaction from happening [153]. It also invariably contains low molecular weight polyethylene glycol (PEG) which can have anti-inflammatory properties [154] and could mask the proinflammatory effects of HMGB1. Therefore, it was removed post-purification via tandem SEC and adsorbent polymers[7]. Only sterile plasticware was used in this protocol (factory pyrogen-free certified). Glassware was heat-inactivated at 180°C for 45 minutes. All solutions used in this protocol were prepared with endotoxin-free 18.2 mΩ water (validated < 0.03 EU/mL in house). TCEP was observed to inhibit LAL activity at assay concentrations over 0.5 mM, but not below (clotting observed in samples spiked with LPS).

[6] In my experiments, I did not observe differences in endotoxin levels with the described protocol with or without the addition of a 5 mL EndoTrap HD FPLC column in-line with the SEC column. In both cases, if the sample pool after SEC was > 2 mg/mL, endotoxin levels ranged between 1-4 EU/mg, below the maximum threshold for my assays. For simplicity, I prescinded of these columns.

[7] Whilst adsorbent polymers can partially remove PEG contamination [283], addition of an amount of polymer sufficient for full PEG removal resulted in >80% protein losses due to nonspecific protein binding to the polymers. Therefore, the LPS removal process was performed prior to SEC, which would remove both detergent micelles (high mW) and PEG in solution (< 1500 Da).

2.13.I Protocol

SM-2 BioBeads preparation

A 10% v/v slurry of BioBeads (Bio Rad 152-8920) was prepared in 2% TX-114 and water. This slurry was incubated with gentle rotation at 4°C to produce a homogeneous solution, then drained and washed with 20 CV methanol, 30 CV endotoxin-free water, and equilibrated with 20 CV of sterile PBS before usage.

Chromatography system sanitisation

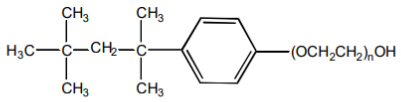
The entire system (fraction collector, columns, sample loops, flow path) was cleaned with water, then incubated in 0.5 M NaOH for 16h to remove protein traces and endotoxin. A second wash with 0.2 M acetic acid in 20% ethanol was applied to further sanitise the system before equilibration in water and PBS buffer. All columns were stored in PBS + 2 mM EDTA to prevent bacterial growth without resorting to toxic agents (e.g., sodium azide).

1. Final 1% TX-114 was added to protein samples concentrated to 15-10 mg/mL in no more than 1 mL, and homogenised at 4°C for 20'[8] before heating for 5 minutes at 37°C.
2. Detergent phase was removed by centrifugation at > 25°C (Triton cloud point), 8000 RCF, 10'
3. 5% v/v of freshly regenerated SM-2 Biobeads were added to the aqueous phase, with wide-bore pipette tips. The suspension was incubated for 2h at room temperature with 1200 orbital shaking to allow for detergent adsorption.
4. The sample was diluted to 0.5 mL (if needed) and purified via SEC as per 2.11, in a 10/300 Superdex 75 pg Increase column. Injection volumes were between 0.5-1 mL.
5. TX-114 or PEG contamination was evaluated by LC-ESI-QTOF/MS (,Figure 2.5). If TX-114 is present the samples were processed again as in Step 3. Samples which could not pass both the LAL test or where polymer contamination was unavoidable **were discarded**. [148]
7. Endotoxin levels were measured with a commercial endotoxin kit (GenScript ToxinSensor L00350) [155][9]. Sample concentrations were adjusted with a protein-free sample treated in the same manner to correct for potential Triton absorbance. If samples are contaminated the protocol was repeated from Step 1.
8. Samples were sterile filtered with a 0.22 µM filter and aliquoted at the desired concentration prior to snap-freezing.

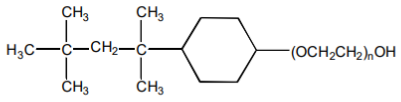
[8] Protein concentrations over 20 mg/mL can lead to deficient phase separation due to high viscosity in the aqueous phase. Crowding agents (e.g. sucrose, glycerol) were also avoided due to this. However, too dilute samples will fail the endotoxin tests as there will always be a volume-dependent residual endotoxin presence. I adjusted the scale of my experiments to ensure the yields would allow for this.

[9] Whilst more reliable and ethical alternatives to the LAL assay (which stresses dwindling populations of horseshoe crab) exist such as recombinant Factor C [155], I was unable to observe reliable results with these in my experiments. Therefore, classical LAL-based kits were used [284]

A.



Triton X-100



Triton X-100R (reduced)

Table B Peaks that may be Observed in ESI Spectra if Triton X-100, X-114, X-405, or X-45 are Present

# of repeat units (n)	Exact Mass (u)		
	H^+	Na^+	Reduced Form Na^+
1	251.2006	273.1825	279.2294
2	295.2268	317.2087	323.2557
3	339.2530	361.2349	367.2819
4	383.2792	405.2611	411.3081
5	427.3054	449.2874	455.3343
6	471.3316	493.3136	499.3605
7	515.3578	537.3398	543.3867
8	559.3841	581.3660	587.4130
9	603.4103	625.3922	631.4392
10	647.4365	669.4184	675.4654
11	691.4627	713.4446	719.4916
12	735.4889	757.4709	763.5178
13	779.5151	801.4971	807.5440
14	823.5414	845.5233	851.5702
15	867.5676	889.5495	895.5965
16	911.5938	933.5757	939.6227
17	955.6200	977.6019	983.6489
18	999.6462	1021.6282	1027.6751
19	1043.6724	1065.6544	1071.7013
20	1087.6986	1109.6806	1115.7275
¾ etc. ¾			

Note: X-100 has on average 9 to 10 repeat units (also reduced form)
X-114 has on average 7 to 8 repeat units (also reduced form)
X-405 has on average 40 repeat units (also reduced form)
X-45 has on average 2 repeat units (only reduced form)
Na^+ cationised form is observed most intensely, H^+ weakly (if at all)
K^+ cationised form also possible
a by-product of Triton production is polyethylene glycol

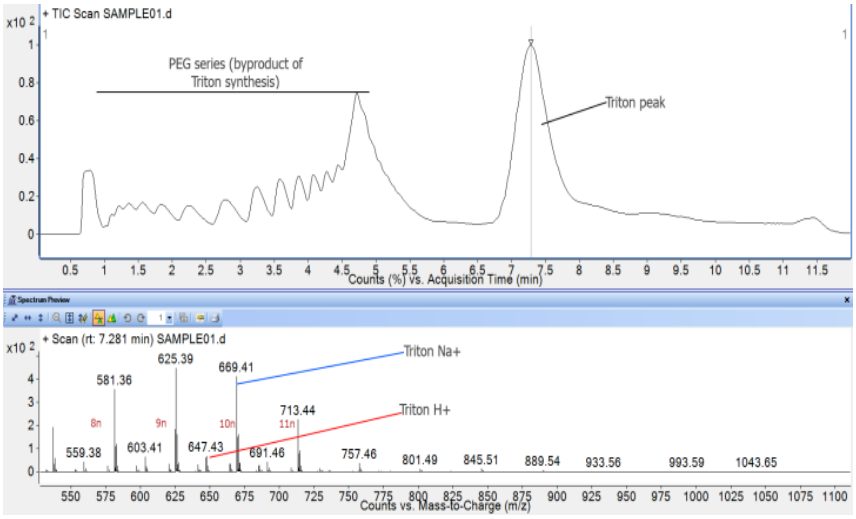
$H–(OCH_2CH_2)_nOH$

Table D Peaks that may be Observed in ESI Spectra if Polyethylene Glyco Present

# of repeat units (n)	Exact Mass (u)		
	H^+	Na^+	K^+
1	63.0441	85.0260	100.9999
2	107.0703	129.0522	145.0262
3	151.0965	173.0784	189.0524
4	195.1227	217.1046	233.0786
5	239.1489	261.1309	277.1048
6	283.1751	305.1571	321.1310
7	327.2013	349.1833	365.1572
8	371.2276	393.2095	409.1834
9	415.2538	437.2357	453.2097
10	459.2800	481.2619	497.2359
11	503.3062	525.2881	541.2621
12	547.3324	569.3144	585.2883
13	591.3586	613.3406	629.3145
14	635.3849	657.3668	673.3407
15	679.4111	701.3930	717.3669
16	723.4373	745.4192	761.3932
17	767.4635	789.4454	805.4194
18	811.4897	833.4717	849.4456
19	855.5159	877.4979	893.4718
20	899.5421	921.5241	937.4980
¾ etc. ¾			

Note: Na^+ cationised form is observed most intensely, H^+ weakly (if at all)

B.



C.

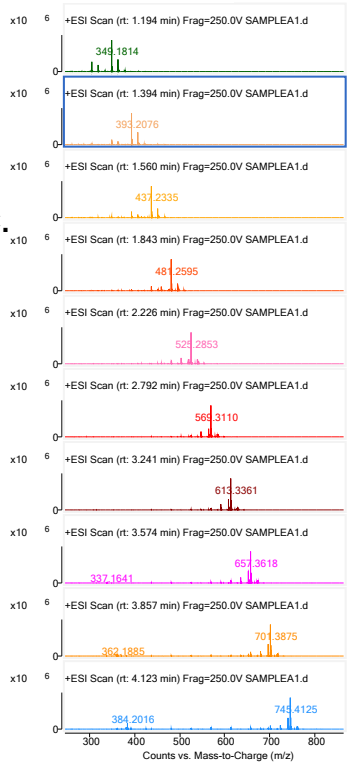


Figure 2.5: Identification of PEG and Triton contamination in protein samples after LPS removal

A. ESI/MS ions visible when either Triton detergents or PEG are present Tables obtained from the University of Alberta Mass Spectrometry facility database (https://www.chem.ualberta.ca/~massspec/es_ions.pdf), accessed 28.12.19. Series for TX-114 is centred at the ions corresponding to 7/8 repeats. PEG will also produce a distinctive pattern on LC
B. Example from LC-ESI-QTOF/MS of a protein sample contaminated with Triton-X114 and PEG. LC chromatogram (top) and M/Z ESI spectra of the peak at 7.2 min (Triton peak; specific ions have been annotated). Note the distinctive PEG elution profile between 1 and 5 min; each peak corresponds to different repeat lengths. **C.** PEG ion series M/Z spectra from each of the peaks shown in (b).

2.14 Differential scanning fluorimetry (DSF)

1. 16 µL of each buffer condition were mixed with 4 µL of 8 µM protein solution containing 5X SyPro Orange (Invitrogen S6650), in a white qPCR-compatible 96 well plate (ThermoFisher AB700W) and sealed with fluorescence-compatible film (BioRad MSB1001), allowing for space between wells to avoid detachment of the film. Conditions tested are summarised in Table
2. The plate was spun down (2500 RCF, 1') and allowed to equilibrate for 10' in the dark at 22°C.
3. Unfolding was measured by heating the plate in increments of 1°C for 1' (from 25 to 95°C) and recording SyPro fluorescence at endpoints in a ThermoFisher MX3005P qPCR thermocycler.
4. Data were exported as the multicomponent view of response in .XSLX format (horizontal grouping). The sigmoidal segment of the melting profile was extracted from the dataset (I used a MS Excel macro available at the SGC [156]) and fitted to a Boltzmann sigmoidal to extrapolate Tm_{50}
5. To calculate kD from nanoDSF data, a modified Langmuir binding isotherm equation was applied [157], assuming CXCR4 monomers and thus a 1:1 ratio due to the absence of agents promoting CXCL12 dimerisation, e.g. divalent anions. Tm_0 = Tm_{50} of CXCR4 without CXCL12, Tm_{max} = steady-state Tm50, *CXCR4/ CXCL12* = concentration of each protein, in molar units.

$$f'\left(\frac{350}{330},T\right) = Tm_0(Tm_{\max} - Tm_0) * \frac{1 - \left(CXCR4 - kD - CXCl12 + \sqrt{([\![CXCR4 + CXCL12 + kD)]\!]^2 - 4 * CXCL12 * CXCR4}\right)}{2 * CXCR4}$$

	1	2	3	4	5	6	7	8	9	10	11	12	
A	De-mineralised water			Saline (150 mM NaCl in DI water)			50 mM HEPES, pH 7.5, 50 mM NaCl			50 mM HEPES, pH 7.5, 150 mM NaCl			Salt screen
B	50 mM HEPES, pH 7.5, 250 mM NaCl			50 mM HEPES, pH 7.5, 500 mM NaCl			50 mM HEPES, pH 7.5, 750 mM NaCl			50 mM HEPES, pH 7.5, 1 M NaCl			
C	50 mM HEPES, pH 7.0			50 mM HEPES, pH 7.5			50 mM HEPES, pH 8.0			50 mM Tris, pH 7.0			Buffer screen
D	50 mM Tris, pH 7.5			50 mM Tris, pH 8.0			50 mM $NH_4CH_3CO_2$, pH 6.8			100 mM $NH_4CH_3CO_2$, pH 6.8			
E	200 mM $NH_4CH_3CO_2$, pH 6.8			0.33X PBS, pH 7.4			1X PBS, pH 7.4			2X PBS, pH 7.4			
F	0.1% Formic acid			0.1 M SPG, pH 4.5			0.1 M SPG, pH 5.0			0.1 M SPG, pH 5.5			pH screen
G	0.1 M SPG, pH 6.0			0.1 M SPG, pH 6.5			0.1 M SPG, pH 7.0			0.1 M SPG, pH 7.5			
H	0.1 M SPG, pH 8.0			0.1 M SPG, pH 8.5			0.1 M SPG, pH 9.0			0.1 M SPG, pH 10.0			

Table 2.3: Contents of custom DSF screen

To assess the suitability of HMGB1 constructs as a therapeutic target, I introduced buffer conditions which would be present in standard clinical practice (saline solution and sterile water), in addition to pH and salt ranges present in my purification conditions.
Concentrations shown as in the final DSF assay (stocks are at 1.25X to account dilution product of adding protein and SyPro Orange stock). This screen is logged in the SGC database as "AVG_DSF_followup_z001". SPG: Succinic acid, sodium phosphate and glycine. Buffer screen contains all buffers which I applied in the experiments described in this work, plus a pH and salt range in Tris/HEPES.. Darker hues represent either increasing pH (Tris, HEPES and SPG pH screens) or salt concentration (ammonium acetate, PBS and NaCl in HEPES buffer screens).

2.15 Peptide arrays (SPOT array)

1. Nitrocellulose membranes were printed with either overlapping 15-mers (2 residues shift) from the sequences of HMGB1 (Uniprot P09429) or CXCL12 (Uniprot P48061) via Fluorenylmethoxycarbonyl (FMOC)-coupled synthesis by Dr. Sarah Picaud at the SGC [158], or hit peptides with serial single residue mutation to alanine for alanine scanning; 10x His peptides were incorporated as positive controls.
2. Membranes were rinsed 3X with ethanol (1´), 3X PBST[10], and blocked for 8h in 10% BSA/PBST.
3. BSA was removed with 2X PBST washes and 1 PBS wash.
4. 1 µM 6-His tagged FR-HMGB1 (+ 0.5 mM TCEP) or CXCL12 in PBS was added and incubated with gentle rocking for 16h.
5. Unbound proteins were washed 3X with PBST.
6. Bound protein was detected with addition of anti-his horse radish peroxidase (HRP) conjugated antibodies (1:3000 dilution in 2.5% BSA/PBST, e.g., Quiagen 34460) for 1h. Excess antibody was removed with 3X washes in PBST (20')
7. Detection of HRP was done via chemiluminescence (Pierce ECL kit, 32109). Chemiluminescence was registered in incremental images with a 5-minute interval until sufficient development occurred in an ImageQuant LAS 4000 camera.
8. Images were analysed with Fiji ImageJ and the intensity of spots was normalised to either the 10X histidine control (screening array) or the unmodified hit peptide (alanine scan). For the alanine scan, hits were considered for peptides with signal variation higher than peptides mutated at Ala positions (no mutation; measures experimental noise)

2.16 Crystallography (hanging drop)

1. HMGB1 and CXCL12 solutions were incubated with 1:1000 Alexa amine-reactive dyes in 10 mM HEPES 150 mM NaCl (Alexa488-NHS and 555-NHS; one for each protein; ThermoFisher A20009/A20000) on ice for 30'. Final 1 mM Tris-HCl was added to neutralise excess dye afterwards.
2. Meanwhile, 20 µL of a 96-well crystallisation screen were transferred to the reservoirs of a SwissCi 3-well plate (3W96T).
3. From labelled proteins 1:1 HMGB1 Box/CXCL12 mixture was prepared and concentrated to 25 µL; molar ratios were verified by the relative fluorescence of each dye in the mix.
4. With a ttpLabtech MosquitoCrystal, 50, 100 and 150 nL of protein solution was transferred to each sub-well in the plate (a-c-d), then 150, 100 or 50 nL of screen mixed with each.
5. Plates were sealed with fluorescence-compatible seals (Molecular Dimensions Clear Vue, MD6-01S) and incubated at 4ºC with imaging in fluorescent and brightfield channels at 0.5, 1, 3, 7, 14, 28 and 56 days in a temperature-controlled incubator (Formulatrix RockImager 1000).

[10] Unless stated otherwise, step duration was 10' and temperature 22°C.

2.17 Biolayer interferometry (BLI)

BLI was performed with the OctetRed 384 system (Molecular Devices). All experiments were performed with 1000 rpm orbital shaking, 5 Hz acquisition rate, 20 averages, at 25 °C.

2.17.I Buffers

- Base buffer (BB): 10 mM HEPES, 150 mM NaCl, pH 7.5, 0.5 mM TCEP.
- Kinetics buffer (KB): Base buffer with 0.05% Tween-20, 1% BSA, 0.0025% sodium azide and no TCEP for CXCL12 binding; PBS, 0.05% Tween-20 and 0.1% BSA for RAGE binding
- Reduction buffer (RB): Base buffer + 2 mM TCEP

2.17.II HMGB1/CXCL12 interaction

1. Streptavidin sensors (FortéBio, 18-5019, pre-wetted in BB for 10') were coated with 3.33 µM solutions of biotinylated HMGB1 constructs in BB for 60s and blocked for 180s in KB before dissociation of excess HMGB1 for 600s. Sensors were kept wet at all times.
2. A concentration gradient of CXCL12 in KB, from 0 to 150 µM with a 1:2 dilution factor was set up in a black tilted 384-well plate (FortéBio, 18-5080). Adjacent wells were filled with either KB or RB.
3. Kinetic measurements were taken with 60s baseline in, 500s association, 420s dissociation (all in KB) and 180s reduction (in RB). Each replicate was done with one sensor.

2.17.III RAGE-HMGB1 interaction

1. Anti-human IgG Capture (AHC, FortéBio, 18-5060, pre-wetted in KB for 10') biosensors were coated with 15 µg/mL of RAGE-Fc and blocked for 180s before dissociation of excess HMGB1 for 120s, all in KB.
2. A concentration gradient of HMGB1 constructs in KB, from 0 to 25 µM over 7 steps, or 0 to 22.22 µM over 10, all with 1:3 dilution factor was set up, plus and adjacent well with KB
3. Kinetic measurements were taken with 60s baseline, 200s association, 200s dissociation. Each sensor was run three times over the solutions.

2.17.IV Analysis

Data were extracted with FortéBio Data Analysis (v 11.0) and the equilibrium response (R_{eq}) at each step was fitted in a saturation binding fit versus [ligand] according to a single-site, single-ligand model, subtracting both the blank ligand and blank chip (double reference). Further processing was done in GraphPad Prism (v. 8.31):

- *kD (dissociation constant, M)* and maximum bound protein in the case of CXCL12 (B_{max}) were derived from the initial saturation fits for each replicate.
- The apparent first-order rate constant *kObs* [11] was fitted as a linear equation to [CXCL12] to compute the *kOn (association rate, $M^{-1}s^{-1}$)* as the line slope, and *kOff (dissociation rate, s^{-1})* as the Y-intercept [159], as well as derived from the horizontal average of the directly measured values in the instrument for each concentration point.

[11] Ligand concentrations 10-fold different from kD had to be excluded due to the plot approaching a second order reaction even for first order kinetics [285].

2.18 Nuclear magnetic resonance (NMR)

Cross-peaks in the NOESY spectrum (inter-residue interactions; adjacent residues in the chain) were cross-compared with the TOCSY spectra (intra-residue interactions; provides information on side chain type) to assign residues in the HSQC spectra, together with deposited BMRB data by other groups on HMGB1 (15148, 15419) [58]. The full HSQC spectrum of an HMGB1 fragment 1-184 is largely the sum of individual Box A and Box B spectra [160]; therefore, published assignments were used to validate from NOESY/TOCSY. For CSP/intensity change tracking, parallel HMGB1 samples without CXCL12 added, stored in the same conditions as the titrated samples, were included to account for chemical shift drifting during the experiments (parallel control). Median CSP was used instead of the mean as the distribution of CSP values is not Gaussian in my data; I/I0 changes did and mean was used. Sample dilution was necessary in several experiments; CCPNMR Analysis 3.0 weights the average intensity/CSP change of the whole spectrum in its calculations for variations in individual residues, and therefore, N15 labelled protein dilution should not be a factor. pH and volume in the NMR tube were also kept constant between samples to avoid shim changes (tube position)

Full method files and raw data can be provided upon request. For all experiments, data were collected with 25% non-uniform sampling (unweighted) Pulse programs were based upon those in use by Professor Christina Redfield at Oxford Biochemistry Dept., Rex Richards building, without whose expertise I would have not been able to perform any NMR experiments.

1. ^{15}N-labelled recombinant HMGB1, at 2X working concentration in 10 mM HEPES 150 mM NaCl pH 7.5, was mixed to 1X with either buffer or CXCL12, plus 5 v/v D_2O to the protein sample and pipetted with a glass Pasteur pipette into a 5 mm Shigeimi tube (Final volume > 330 µL).
2. After embolus insertion and sealing with paraffin, the tube was placed into the spectrometer sample holder.
3. Deuterium signal was locked on until the baselines overlapped and maxima of each of the two lock-on signals were symmetrical to each other.
4. Optimal shimming (tube position) was manually adjusted for highest possible signal homogeneity, varying the gain if required (X, Y, XY, (X^2-Y^2, XZ/YZ parameters).
5. The ^{15}N and ^{1}H nuclei were automatically tuned and matched (*atma* command*)*, then manually refined (*atmm)* so the frequency valley in the plot matched the reference nucleus frequence (tuning) and the reflected signal was minimal (matching).
6. Shimming was re-refined automatically (*topshim)*
7. Water signal was supressed as follows: intensity was estimated via the *pulsecal* command, and a ^{1}H spectra was acquired (*zg:ft:apk)* with the power level 1 set to the *pulsecal* value. If multiple peaks appeared, shimming was repeated: otherwise, the process was repeated with p1= 4 times the initial value. If the new spectra did not display a symmetric positive and negative peak, p1 was modified (higher if negative skew or vice-versa) until peaks were symmetrical. This p1 value was utilised for all further experiments.
8. NMR experiments were performed after these calibration steps. ^{1}H-NMR, ^{15}N-HSQC, ^{15}N-NOESY-HSQC, ^{15}N-TOCSY-HSQC; full log files for settings are available if requested. Specifics for frequencies and pulse width settings can be found in Appendix A4.

2.19 *In vivo* muscle injury model

2.19.I Animals

C57BL/6 inbred mouse strain, females of 11-12 weeks of age were procured from Charles River, UK, and housed in the local Biological Safety Unit (BSU) at the Kennedy Institute. Acclimatisation period was 1-2 weeks.

All protocols performed on live animals have been approved by the UK Home Office (Project licenses 30/3330 and), the local Named Animal Welfare and Competence Officer (NACWO), Named Veterinary Surgeon (NVS) and Animal Welfare and Ethical Review Body (AWERB) in compliance with the Animal Scientific Procedures Act (ASPA) of 1996. I held a Personal License (PIL number IFA8DA776) throughout this entire work and have been assessed by the NVS and NACWO of the BSU as fully competent for these procedures, as well as for Schedule 1. Full tracking and procedure data are stored in the local database (Sanger) and are available to accredited Home Office inspectors if required. Individual animals were tracked via ear clip and cages via the Sanger database system at Oxford University.

All consumables were surgically certified; recombinant proteins were made endotoxin-free (2.13). Surgeries were done in a clean environment separate from culling facilities. All animals were monitored for 1h post-operatively, and daily for the following 3 days: monitoring was then transferred to the NVS/NACWO. Isoflurane and buprenorphine were provided by the local NVS.

2.19.II Method [98]

1. Animals were anaesthetised via inhalation of 2-3% v/v isoflurane, and 150-250 µL of buprenorphine (Vetergesic, diluted to 15 µg/mL) were injected subcutaneously (scruff; final 0.125 mg/kg) as an analgesic.
2. Right hindleg tibial area was disinfected with iodine and the tail (if performing i.v. injection) with ethanol.
3. Proteins were injected either intramuscularly (tibialis anterioris; TA muscle) or intravenously (tail vein) with FR-HMGB1 or CXCL12 constructs; doses are listed in the relevant experiments. Recombinant HMGB1 injection (0.075 to 6 mg/kg) did not induce additional stress in the animals.
4. 50 µL 1.2% BaCl2 was injected in the TA muscle to induce cell death.
5. At the desired time points, mice were sacrificed via rising CO_2 concentration before dissecting the right hindleg for fixation in 3.7% formalin (24h). After this period, the TA muscle was dissected and fixed for another 24h before transferring to PBS. Tissues were embedded in paraffin blocks at the Kennedy Institute Histological facilities.
6. 5 µm thickness sections were cut and stained with haematoxylin-eosin at the Histology facility and imaged at 10X ocular/40X lens magnification in a bright-field Olympus BX51 microscope, with at least 4 images per animal are recommended.
7. Muscle cross-sectional area (CSA) of injured muscle fibres was measured in ImageJ and analysed as described in Prot. 2.2.

2.20 *In vitro* transwell migration assay [31]

1. 3T3 fibroblasts (ATCC CRL-1658) were grown in 10% v/v heat-inactivated foetal bovine serum (HI-FBS) supplemented with penicillin/streptomycin 1X up to confluence 70-80% and serum starved in 0.5% HI-FBS without antibiotics, in an incubator at 37ºC, 5% CO_2. Frozen cell stocks were prepared in growth medium without antibiotics and 5% v/v DMSO.
2. Cells were washed with sterile DPBS, lifted with 0.1X Trypsin for 5' at 37ºC, and resuspended in media containing 0.5% HI-FBS to 2.5 x10^5 cells/mL.
3. 5000 cells were pipetted in each upper chamber of a Boyden transwell plate (Corning CLS3422-48EA), with the bottom wells containing either 0.5%/20% HI-FBS, 1 nM HMGB1 constructs in 0.5% HI-FBS, or commercial CXCL12 (PeproTech, 300-28A).
4. Migration was allowed for 6h in the incubator, and the upper chamber swabbed to remove non-migrating cells with a cotton swab.
5. Transwells were stained with Giemsa solution and 10 random fields were counted per chamber to quantify migrating cells

2.21 Gene design and ordering

Optimisation of HMGB1 sequences, when necessary, was performed with the software EuGene [161]. All genomes were sourced from the National Centre for Biotechnology (NCBI) repositories; for the purposes of optimisation, only cDNA sequences were loaded for each genome.

- Origin genome: HMGB1 human sequence, loaded as mRNA in the human genome (*Homo sapiens* assembly GRCh38, *ab initio* predictions)
- Destination genome: *E. coli* BL21-DE3 genome (assembly ASM956v1)
- Optimisation settings:
    - Harmonise relative synonymous codon usage.
    - Keep rare codons: ON
    - Maximise codon correlation.
    - Remove repeated nucleotides (>5)
    - Iterative ("slow") protocol. *A series of fits is produced, and one amongst them was manually chosen for synthesis.*

Customised DNA sequences were purchased as cloned fragments in pNIC-CTH0 from Twist Biosciences (San Francisco, USA). If necessary, sequences were optimised to maximise success chances; the supplier also filters input sequences and returns an error if they are non-feasible. Purchased constructs were processed as described in previous sections.

2.22 HMGB1 *in vitro* H9 differentiation assay[12]

2.22.I Consumables and cell lines

- Irradiated mouse embryonic fibroblasts (MEF-KO) [StemCell Technologies CF-1 MEF IRR 4M (A34181)].
- H9 Human Embryonic Stem Cells (WiCell, WA09 line)
- Stem Cell Cardiomyocyte Maintenance Kit (StemCell Technologies, 05020).
- StemDiff Stem Cell Cardiomyocyte Differentiation Kit (StemCell Technologies, 05010).
- Matrigel (Corning 354277)
- Gelatin (0.1% solution, StemCell Technologies 07903)
- Rho-associated kinase (ROCK) inhibitors [162,163] (Y27632; Sigma SCM075)
- Gentle Cell Dissociation reagent (StemCell Technologies 07174).

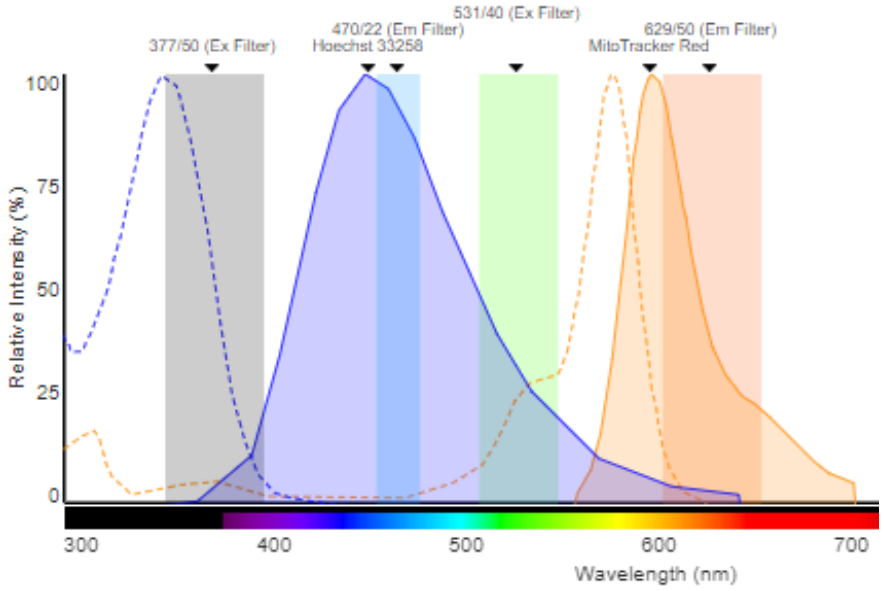
2.22.II Protocol

1. 60 mm glass Petri dishes were coated with gelatin (0.1%) at room temperature (RT) for 2h, then washed with DPBS.
2. $2x10^6$ MEF-KO cells (≈50% confluence) were seeded in a homogeneous layer in each plate in 10% HI-FBS/DMEM without antibiotics (from frozen stock) and cultured for 1 day at 37°C, 5% CO_2. These plates were used within a maximum of 14 days.
3. An H9 frozen vial was thawed and gently resuspended in 4 mL of mTESR+ media with 1 μM Y27632, taking care to not fully break the stem cell colonies (colloidal suspension of 20-40 cells per colony).
4. The confluent MEF-KO plates were washed in DPBS prior to addition of 10^6 H9 stem cells in suspension and filled up to 3 mL of mTESR+ media with Y27632
5. On the next day, the media was replaced with regular mTESR+. Plates were grown for 4-6 days, removing colonies over 200 cells (visible white opaque colour with thicker cell wall). On this point, cells were either frozen or subcultured into Matrigel (next step).
6. 60 mm glass Petri dishes were coated with 2 mL of 1:60 dilution of Matrigel stock with ice-cold mTESR for 30'.
7. Media in the H9 cell plates was replaced with 1 mL of Gentle Cell Dissociation medium to release pluripotent cells. This was then aspirated to leave a thin layer and plates were incubated for 1' at 37ºC.
8. The cells were then gently resuspended with 4 mL of mTESR+ medium with Y27632 to a colony size of 40-50 cells approximately and distributed into Matrigel-coated plates (1:2 dilution factor).
9. The next day, media was replaced with regular mTESR+; plates at this point were cultured until 70-80% confluence prior to differentiation.
10. At 50% confluence 100 nM/mL of each HMGB1 construct was added to the media for 24 hours.
11. Differentiation was induced with the StemDiff Cardiomyocyte differentiation kit:
    a. 2 days culture in Differentiation medium A + 10 μM Chir 99021 (StemCell Technologies)
    b. 2 days culture in Differentiation medium B
    c. 4 days culture in Differentiation medium C (2 changes). Cells were maintained from this point on in StemDiff Maintenance media with changes every 48h.
12. Differentiation was assessed by cardiac troponin content via commercial kits (Human Cardiomyocyte Immunocytochemistry Kit, Thermo Fisher A25973).
13. Contractility measurements were done via the MyoCyter macro for Fiji. Frequency of motion in cardiomyocyte clumps was measured and averaged across 3 biological replicates, at day 17.

[12] I would like to acknowledge DPhil student Tzung Yuan (Carl) Lee, MSc in our group for developing and optimising this assay.

2.23 C2C12 differentiation assay

C2C12 cells (ATCC CRL 1722) were used to analyse myoblast formation via measurement of mitochondrial activity [164]. Cells were maintained in 10% HI-FBS-High-glucose DMEM with glutamine, 1X penicillin/streptomycin and non-essential amino acids (Gibco 11140050), without allowing the cells to exceed 70% confluence to prevent differentiation; frozen stocks were prepared with growth medium without antibiotics and 5% v/v DMSO under standard tissue culture practices. Dual *in vivo* staining was performed with MitoTracker Red (ThermoFisher M7512; $577_{ex}/599_{em}$) and Hoescht 33258 (ThermoFisher H3569; $352_{ex}/461_{em}$) dyes. A schematic of the fluorophores and filters used is shown below (no filter combination more suitable for MitoTracker was available):



1. C2C12 cells were passaged to 25% confluence in a 96-well plate.
2. The next day, the cells were supplemented with growth medium and 1 nM of each HMGB1 construct tested (if appropriate)
3. Upon reaching confluence (2 days), media was exchanged to 2% HI-FBS growth media to start the differentiation process, except on control wells kept in regular growth medium.
4. At the designated time points, media was replaced with differentiation medium containing 25 ng/mL MitoTracker and 1 nM Hoescht dye and incubated for 30 minutes to allow dye permeation before replacing the medium.
5. Fluorescence was measured in the red channel (MitoTracker, $531/40_{ex}$-$629/53_{em}$) and blue channel (Hoescht, $377/50_{ex}$-$470/22_{Em}$), in a CeligoS instrument. The measurements were kindly performed by Ms. Marisa Cabrita in the Nanchahal Group. To calculate relative mitochondrial activity at each timepoint and therefore progress of differentiation, the fluorescence on the red channel was normalised by that on the blue channel (Target 2% parameter).

2.24 TLR-4 and TLR-2 HEK reporter assay

TLR activity was evaluated via commercial HEK-Dual NF-κβ reporter cells (Invivogen: hTLR-2; hkd-htlr2ni, mTLR-4; hkd-mTLR-4ni, Null; hkd-nullni), maintained in 10% HI-FBS-High-glucose DMEM with glutamine, plus penicillin/streptomycin 1X and selection antibiotics (Zeocin 50 µg/mL; Hygromycin B 100 µg/mL) added after 4 days of subculture. Cells were frozen in 20% HI-FBS/10% DMSO (both v/v) in DMEM. These cells express secreted embryonic alkaline phosphatase under NF-κβ control and the respective TLR or coreceptors.

1. A confluent cell layer was subcultured into cell-culture 96-well plates at 25000 cells/well and cultured for 24h at 37°C, 5% CO_2. A Null cell line row was added as a control (blank).
2. Medium was replaced with 200 µL of media supplemented with either HMGB1 constructs at 400 nM, positive controls: 0.2 ng/mL LPS-EB Ultrapure (Invivogen **tlrl-3pelps)** and 1 ng/mL LTA-SA Ultrapure (Invivogen **tlrl-pslta) or media (baseline)**: both controls were added in each case to evaluate cross-contamination. Cells were incubated for 24 hours to allow SEAP secretion.
3. 20 µL of medium were mixed into a clear bottom, tissue-culture treated black-well 96 well plate (Costar 3603) with 180 µL of Quanti-Blue reagent and incubated at 37°C for 1-3h (based on the colour development)
4. OD_{650} data were transformed following this formula: $R = \frac{Signal - blank}{Baseline - blank}$ (fold-change from Null control). TLR agonism will equal higher SEAP expression, and thus colour development.

2.25 HMGB1 RAGE binding enzyme-linked immunosorbent assay (ELISA)

1. 384-well high-binding clear flat bottom plates (SantaCruz Biotechnology, sc-206072) were coated with 50 µL of 40 nM HMGB1 constructs overnight at 4ºC: 2 µM TCEP was added for FR-HGMB1 forms in all steps up to RAGE-Fc incubation. Each plate included a blank row and a HMGB1-FL control in each plate to normalise the data. Each well was assayed in quadruplicate.
2. Plates were washed 3x with PBST (100 µL). This was repeated after steps 3-5.
3. Nonspecific binding was blocked with 100 µL of 10% BSA/1X PBST, for 2 h at 22°C (+ TCEP as per Step 1).
4. 20 µL of RAGE-Fc chimera in 10% BSA/PBS (BioTechne, 1145-RG) were added to the plates and these were incubated 2h at 4ºC.
5. Plates were then treated with 50 µL of 1:10000 dilution of anti-hIgG-HRP (in 10% BSA/PBS), for 2h at RT.
6. Bound secondary antibody (Agilent Dako P021402-2, in 1:1 ratio to RAGE-Fc) was measured by addition of 25 µL TMB substrate (ThermoFisher N301) and incubation until a clear colour gradient is visible for DS-HMGB1-FL (positive control)
7. The reaction was stopped with 25 µL 1 M H_2SO_4., and OD_{450} quantified in a plate reader.
8. To compare data from different plates, values were normalised to the DS-HMGB1 controls in each plate. RAGE concentration used in calculations was double the concentration of RAGE-Fc (each molecule contains 2 RAGE monomers, bound by the IgG fraction). Curves were normalised to a saturation binding fit (Michaelis-Menten equation).

2.26 Experiments done by collaborators

2.26.I nanoDSF and SEC of CXCL12 and CXCR4 binding

CXCR4 purification, nanoDSF and SEC were all done by MSc Cassandra Dix, SGC. I did the plasmid design and generation of CXCR4A-c038 and CXCL12 as well as DSF/SEC data analysis. Original optimisation of CXCR4 expression is the work of Patrizia Abrusci, PhD, and Tina Bohstedt, MSc, at the SGC.

CXCR4, isoform I (MEGIS) was expressed in EXPI-293 cells transfected with a pHTBV1.1-CTGFP-SIII-10H plasmid (Ampicillin resistance, BacMam system) encoding CXCR4 1-325 (CXCR4A-c038) with the thermostabilising mutations L125W and T240P with a TEV site, His10-Twin Strep tag and GFP on the C-terminus.

1. Cells were lysed via a Dounce Homogeniser in hypotonic buffer (20 mM KCl, 10 mM $MgCl_2$, 10 mM HEPES pH 7.5) and pelleted for 40 ´ at 50000 RCF.
2. The process was repeated twice in hypertonic buffer (hypotonic buffer + 1 M NaCl)
3. Samples were resuspended with 50 mM HEPES pH 7.5, 5% glycerol, 0.1 M NaCl. For SEC, CXCL12 was added at this stage (2 µM final)
4. Once resuspended, CXCR4 was solubilised with 1% DDM/0.003% cholesterol for 2 h at 4°C. Solubilised samples were incubated bind to Strep-Tactin XT resin overnight.
5. Columns were washed with base buffer with 0.1%DDM/0.001% cholesterol (30 CV) and eluted with 3 CV of the same buffer with 100 mM d-biotin for 30 minutes before concentration.
6. SEC was performed as described in other sections of this work, in a Superose S200 column, in 50 mM HEPES pH 7.5, 0.1 M NaCl and 0.1%DDM/0.001% CHS.
7. For nanoDSF, purified CXCR4 (without CXCL12) and CXCL12 were premixed at final 0.66 µM CXCR4 and 1.1, 2.2 ,6 and 11.8 µM CXCL12, all in SEC buffer; final volume was 50 µL. Duplicate samples were set for CXCR4/CXCL12, CXCR4 alone and CXCL12 alone at each concentration.
8. Samples were equilibrated at RT for 1 hour before loading in a nanoDSF (Prometheus, NanoTemper) capillary tube. Melting was assessed by intrinsic fluorescence heating from 20-95°C (1°C intervals) and analysed via Pr.ThermControl software (v 2.1.2). Computed Tm_{50} values were fitted to a modified Langmuir isotherm [157] to account for the large amounts of CXCR4 required in the assay to estimate K_d.

2.26.II Monocyte TNF production assay

This experiment was entirely performed by DPhil student Tzung Yuan (Carl) Lee in the Nanchahal Group on my behalf; I only assisted in design of the conditions and data analysis.

I was also unable to produce LPS-free DS-HMGB1, full length for this assay due to the laboratory lockdown because of the Covid-19 pandemic[13][14]. Therefore, it was purchased from a commercial supplier (HMG Biotech, S.R.L, HM-122). The sequence of the commercial protein is identical to my recombinant HMGB1-c001, but it lacks the AENLFYQ TEV protease site at the C-terminus.

Human peripheral blood monocytes (StemCell Technologies, 70034) were maintained in DMEM, 4.5 g/L glucose supplemented with 10 % HI-FBS (Gibco). 10^5 human monocytes were plated (in triplicate for each donor) into wells of a 96 well plate and stimulated with 10 µg/mL HMGB1 constructs and/or with 50 ng/mL LPS (TLR-4 agonist) or 10 ng/mL LTA (TLR-2 agonist), plus a control of proteinase K digested FR-HMGB1-C1-c001, full length (following manufacturer protocols and heat inactivation at 95°C for 30 minutes). All conditions were premixed and incubated in media for 2 hours before stimulating cells. 24 hours after stimulation, TNF levels were determined by anti-TNF ELISA kits (Abcam ab181421).

Data were analyzed via nested ANOVA considering each donor as a subgroup and the technical replicates as repeated measurements; TNF secretion was normalized to secretion in untreated cells. Conditions pertaining to TLR-2 and TLR-4 were analyzed separately to prevent violations of heteroskedasticity.

[13] The equipment I needed to produce LPS free proteins was divided between the Kennedy Institute (LPS-free station) and the Structural Genomics Consortium (Protein purification equipment) This would have required me to access multiple buildings on the same day, which would be a direct violation of the current Safe Working protocols in place at both institutes.

Chapter 3 Purification of recombinant human HMGB1 and CXCL12

3.1 Chapter introduction: design considerations **68**
3.2 Generation of entry clones **71**
3.3 Disulfide bond detection **74**
3.4 Expression and purification of HMGB1 ..... **77**
3.4.I Initial construct panel 77
3.4.II Large scale expression and purification of HMGB1 81
3.4.II.i Ion-metal affinity chromatography (IMAC) (pNIC-CTHF constructs) 81
3.4.II.ii High-throughput IMAC 83
3.4.II.iii Biotinylated HMGB1 constructs 85
3.4.III PTMs of purified HMGB1 86
3.4.I Characterisation of biologically active HMGB1 constructs 88
3.4.I.i SEC profiles of different HMGB1 constructs 91
3.4.I.ii ESI-MS analysis of HMGB1 conformations 95
3.4.I.iii Melting behaviour of different HMGB1 redox states 105
3.4.I.iv Summary of characterisation experiments 109
3.4.II Conclusions – HMGB1 110
3.5 Expression, and purification of CXCL12. ..... **112**
3.5.I First iteration of CXCL12 constructs 112
3.5.II CXCL12 production in cytoplasmic disulfide-competent strains is inefficient 113
3.5.III Construct design for periplasmic export 114
3.5.IV Optimisation of large-scale expression conditions 120
3.5.V Optimisation of cell lysis and downstream purification 122
3.5.VI Characterisation of CXCL12 125
3.5.VI.i Validation of CXCL12 binding to CXCR4 128
3.5.VII Conclusions – CXCL12 130
3.6 Final remarks **131**

3.1 Chapter introduction: design considerations

As the first step for this project, I had to produce enough HMGB1 and CXCL12 to perform structural and biological characterisation. This chapter consolidates all the cloning and protein production experiments to this end, as well as the characterisation of the individual proteins, whilst data pertaining to their interaction will be discussed in Chapter 4. Commercial acquisition of these proteins was not an option due to the large costs associated – a single NMR experiment requires approximately 3-4 mg of HMGB1 and 2-3 mg of CXCL12. Even without accounting for isotopic labelling, the ballpark estimate would encroach £10-15k, so instead I opted for recombinant expression.

HMGB1 has been recombinantly produced in *E. coli* [152] or mammalian cells [152,165], expressed in cell-free systems [166], and even derived from raw tissue by heparin affinity chromatography [39,167]. The main difference between hosts lies in costs and post-translational modifications (PTMs) present - eukaryotic hosts show varying degrees of glycosylation, which is not required for either the cytokine-inducing or regenerative activities of HMGB1. Ion exchange chromatography (IEX), size exclusion chromatography (SEC), and affinity chromatography (GST/Histidine) have all been applied in these studies. These methods were largely optimised for purification of single constructs: FL-HMGB1 or the 1-184 HMGB1 tail-less variant for *in vitro* or *in vivo* assays, where large amounts of protein are not essential, rather than the high-throughput approaches I would need.

CXCL12 has been produced solely by protein refolding [70,111,130]. This approach, besides being extremely laborious, requires preparative-grade high-pressure liquid chromatography (HPLC) [121] to separate the various disulfide-bridged states of CXCL12, an equipment to which I lacked access, and therefore I could not use this protocol despite it having been optimised in the literature up to the point of protein structures being available [112]. CXCL12 is secreted *in vivo*; however, the level of secretion and the tight regulation its signalling is subject to (leading to low expression levels) are likely the reason why no published studies with recombinantly expressed, secreted CXCL12 can be found in the literature.

Therefore, I decided to design my own methodologies, based on the frameworks present at the Structural Genomics Consortium (SGC) [146], including the Mammalian Gene Collection repository (MGC) from where I amplified the sequences for wild-type human HMGB1 (entry clone HMGB1A-s001) and wild-type human CXCL12 (entry clone CXCL12A-s001) complementary DNA (cDNA).

When expressing recombinant proteins, the host organism must be carefully considered. Whereas the genetic code is universal, codon bias differences, [168–170], the ability of the selected host to produce specific PTMs [171], and even the membrane structure and available protein folding machinery can influence not only expression, but also the activity of the recombinant proteins. Given than only redox-state modifications are required on both proteins for their functional activity [111,152], *E. coli* was selected as the expression host.

- CXCL12 has few PTMs, the most relevant being two disulfide bridges which are key in its binding to CXCR4.
- HMGB1 has a much broader spectrum of PTMs, but for its interaction with CXCL12 and signalling via CXCR4, only one has key relevance: the disulfide bridge between C23 and C45.
- In addition, the use of NMR in Chapter 4 requires a host able to express protein in minimal media conditions with labelled isotopes (^{15}N for the purposes of this project), a requirement that can be easily met by *E. coli* [172][15].

Expression was regulated via inducible T7 polymerase under the control of the lac operon [173]. Detailed description of primer, entry clone, and vector combinations for each construct can be found in Suppl. data A1-2-B.

[15] Yeast was avoided to prevent contamination issues in the laboratory, and due to the much larger *E. coli* framework of vectors and protocols available in-house.

3.2 Generation of entry clones

As mentioned in the previous section, fully reduced HMGB1 exerts its regenerative activities via CXCL12, which itself has two disulfide bonds critical for CXCR4 binding as they position the N-terminus in the correct position for proper CXCR4 interaction. This complicates long-term structural studies, as conditions preserving the reduced state of HMGB1 will necessarily reduce the disulfide bridges in CXCL12. There is no information in the literature whether the redox state of CXCL12 is relevant for HMGB1 binding; however, I decided against trying to optimise a set of conditions whose redox potential could preserve the disulfide bridges in CXCL12 whilst breaking those of HMGB1 due to the different solvent exposure of these (Figure 3.1). Both disulfide bridges in CXCL12 are exposed to the solvent and consequently are highly accessible, whereas the disulfide bridge in HMGB1 is buried within two alpha helices in Box A. Therefore, I assumed that any condition capable of breaking the disulfide in HMGB1 would also reduce those in CXCL12. This was observed experimentally during my DPhil (Section 3.3).

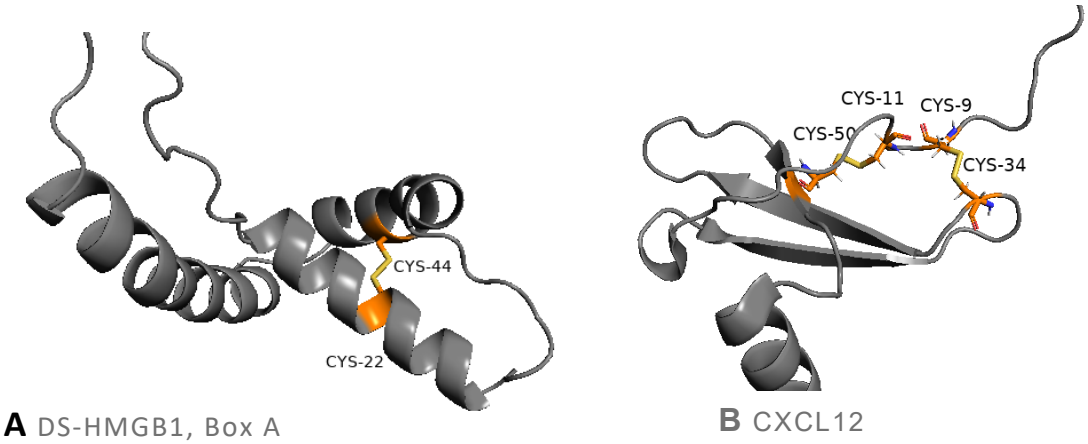


Figure 3.1: Disulfide bridges of HMGB1 and CXCL12

A: Structure of oxidised HMGB1 Box A (PDB 2rtu). Main chain displayed in grey. Sulphur residues in cysteine are displayed in yellow, cystine backbone in orange. **B.** Structure of a CXCL12 dimer (PDB 2kec), same colour scheme. Residue numbers correspond to the unmodified peptide chain, with N-terminal methionine and signal peptide for CXCL12.

3S-HMGB1 is a synthetic mutant where Cys22, Cys44 and Cys105 are substituted by serine residues. This mutant preserves the chemotactic and G_{Alert} mediated regenerative activities of HMGB1, but is not susceptible to oxidation [24,25] and therefore more suitable for long-term solution structural studies such as NMR. However, it has also been shown to be detrimental for the regeneration of some tissues [173]; therefore, I reserved this form only for biophysical assays, and did not use it instead of FR-HMGB1 for *in vivo* work beyond validating the regenerative effects.

To generate HMGB1A-s002 (containing the mutations C22S, C44S, and C105S), fragments were generated with the primer pairs HMGB1WTLICF/HMGB1C106S and HMGB1C45S/HMGB1WTLICR using PCR (Figure 3.2.1). These two fragments were used as megaprimers for a partial amplification over HMGB1A-c001, generating a variant of HMGB1 with C44S and C105S mutations. After cloning and sequencing this temporary construct, it was amplified (Figure 3.2.2), with the primer pair HMGB1C23S/HMGB1WTLICR to generate a fragment containing all three mutations. This fragment became the reverse primer of a reaction with HMGB1A-s001 as a template and HMGB1WTLICF, yielding HMGB1A-s002 as a result.

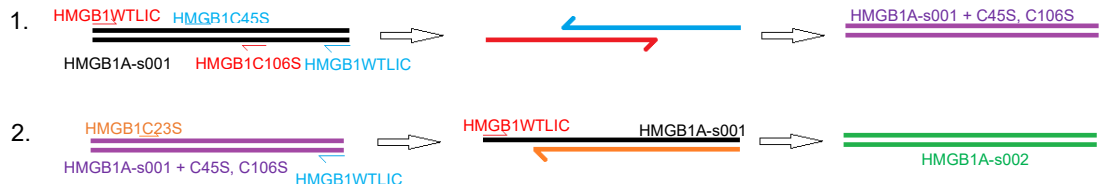


Figure 3.2: Mutagenesis of HMGB1A-s002

Schematics of mutagenesis process for HMGB1A-s002. Arrows in the DNA fragments are directed in the 5'->3' direction. Each outlined arrow represents a single step of PCR reaction with Herculase II and LIC cloning.

CXCL12 is a very small protein (68 amino acids when fully processed, after signalling peptide removal), but two mutants are described in the literature which I thought could prove useful to have:

- Val 3[16] Ile (entry clone CXCL12A-s002) [111]. This mutation binds CXCR4 with equivalent affinity to that of wild-type CXCL12, but is functionally much more active, with increased Ca^{2+} signalling and a lower EC50 for the inhibition of HIV infectivity.
- His 25 Arg (entry clones CXCL12A-s004 -wt - and s005 - V24I) [130]: Histidine 25 is critical for the self-dimerisation of CXCL12 [114]; protonation blocks dimerisation by repulsion with the adjacent monomer. It was postulated that dimerisation hampers resolution in NMR experiments of CXCL12 in complex with HMGB1 [96].

Mutagenesis of CXCL12-s002 was performed following the same procedure as the second step of HMGB1A-s002 mutagenesis. CXCL12A-s001 was amplified with the primer pair CXCL12V24IF/CXCL12LICR. The subsequent PCR product was used as a reverse primer, with CXCL12LICF as the forward primer over CXCL12A-s001 template to extend the construct into the full-length CXCL12A-s002 entry clone. To generate CXCL12A-s004 and s005, the entry clones s001 and s002 were subjected to this same PCR scheme replacing CXCL12V24IF for CXCL12H25RF.

[16] Val 24 Ile if including N-terminal secretion sequence (numbering corresponds to the mature CXCL12 protein)

3.3 Disulfide bond detection

Note that this methodology was optimised after both purification protocols for HMGB1 (Section 3.4) and CXCL12 (Section 3.5) but is described before them as it was applied to both proteins.

Nonreducing, denaturing SDS-PAGE, can provide initial insight on this [174]: disulfide forms do not fully unfold, resulting in a smaller profile and a shift in SDS-PAGE of -1 kDa (faster migration; e.g. Figure 8.3). However, this method has several shortcomings as the potential for interference from e.g., gel quality or position will mask this effect. Furthermore, if the sample contains a mixture of oxidised and reduced HMGB1, this shift is barely appreciable.

Whilst ESI-QTOF/MS has enough accuracy to detect disulfides by MW measurements alone (-2 Da mass shift from loss of 2 protons), there is a margin of error of 50 ppm (1.3 Da for a 26 kDa protein).Direct MS/MS peptide quantitation methods exist [175] and have been used to determine HMGB1 redox states [24,54,72,87,176,177] by alkylation of the free cysteines with iodoacetamide (IA), dialysis in DTT to reduce disulfide bonds, and further alkylation of the newly reduced cysteines in N-ethyl maleimide (NEM); the different alkyl groups have vastly different masses (+57 IA, +125 NEM). These publications describe cleaving of the proteins and quantitation of peptides with different reduction states, and are all authored by a single group, formerly led by Daniel J. Antoine, which performed these measurements externally for these studies.

However, this protocol would require external submission of the samples to a specific research group, which is incompatible with a routine screening tool which I performed for every protein batch and in large numbers: this would severely bottleneck my output. Instead, I tried to replicate it in-house replacing peptide quantitation with ESI-QTOF/MS measurement of the number of adducts added after iodoacetamide, reduction, and NEM treatment. As my samples are recombinant, this is feasible since the PTM landscape should be homogeneous, unlike *in vivo* derived HMGB1 samples which display a wide PTM variety. This would not allow to detect specific peptides but given a homogeneous sample, the number of adducts can be equated to the number of free cysteines in the protein.

This protocol, however, resulted in insufficient alkylation or maleimidation for several targets, including HMGB1[17]. In other cases, such as CXCL12, I detected 4 iodoacetamide adducts which would imply full reduction, despite it showing biological activity and having a -4 Da adduct when untreated (Figure 3.3; full validation data can be found in Suppl. data A2-3-A). As several of the tested proteins used to validate this method possess highly compact structures, I suspected that the disulfides would be not accessible in solution; conversely, CXCL12 has highly exposed disulfides (Figure 1.4) and therefore could still be attacked by the iodoacetamide before it is fully neutralised by DTT. This was striking as this protocol has been cited in several publications; however, due to the recent retraction of several publications involving this protocol, I surmised that it might be an experimental dead end.

To bypass this issue, chemical unfolding was induced: since iodoacetamide can non-specifically modify amines too at pH outside 8-9 [178], and detergents are ion suppressants and thus MS-incompatible, I denatured my samples with 8M urea[18], buffered in 100 mM ammonium carbonate pH 8.0, and in the presence or absence of 2 mM TCEP, which does not neutralise iodoacetamide. With this method, samples with disulfide bonds would have two extra iodoacetamide adducts per disulfide when reducing agents are present, whereas initially reduced samples would not change the number of adducts.

This new iteration of the protocol resulted in consistent results with the published structures and expected reduction states, as shown in Figure 3.3 (excepting tenascin C, which has a particularly stable fold and could not fully reduce). As the samples are diluted 1:60 prior to injection, the residual urea concentration is negligible and does not hamper analysis. This method allowed me to perform high-throughput, accurate characterisation of the reduction state of my samples.

[17] I would like to acknowledge DPhil student Jesse A. Coker for providing me with samples of several of his TLR-4 disulfide containing ligands. These were instrumental for my setting up of this methodology as they I had enough different proteins to test whether my method was robust.

[18] Whilst some publications recommend heating for reduction reactions, this is inadvisable with urea as it could cause decomposition and carbamylation of the proteins due to the decomposition of urea to ammonium cyanate when heated. As alkylation blocks the cysteine residues, oxidation during the reaction was considered to not be a factor.

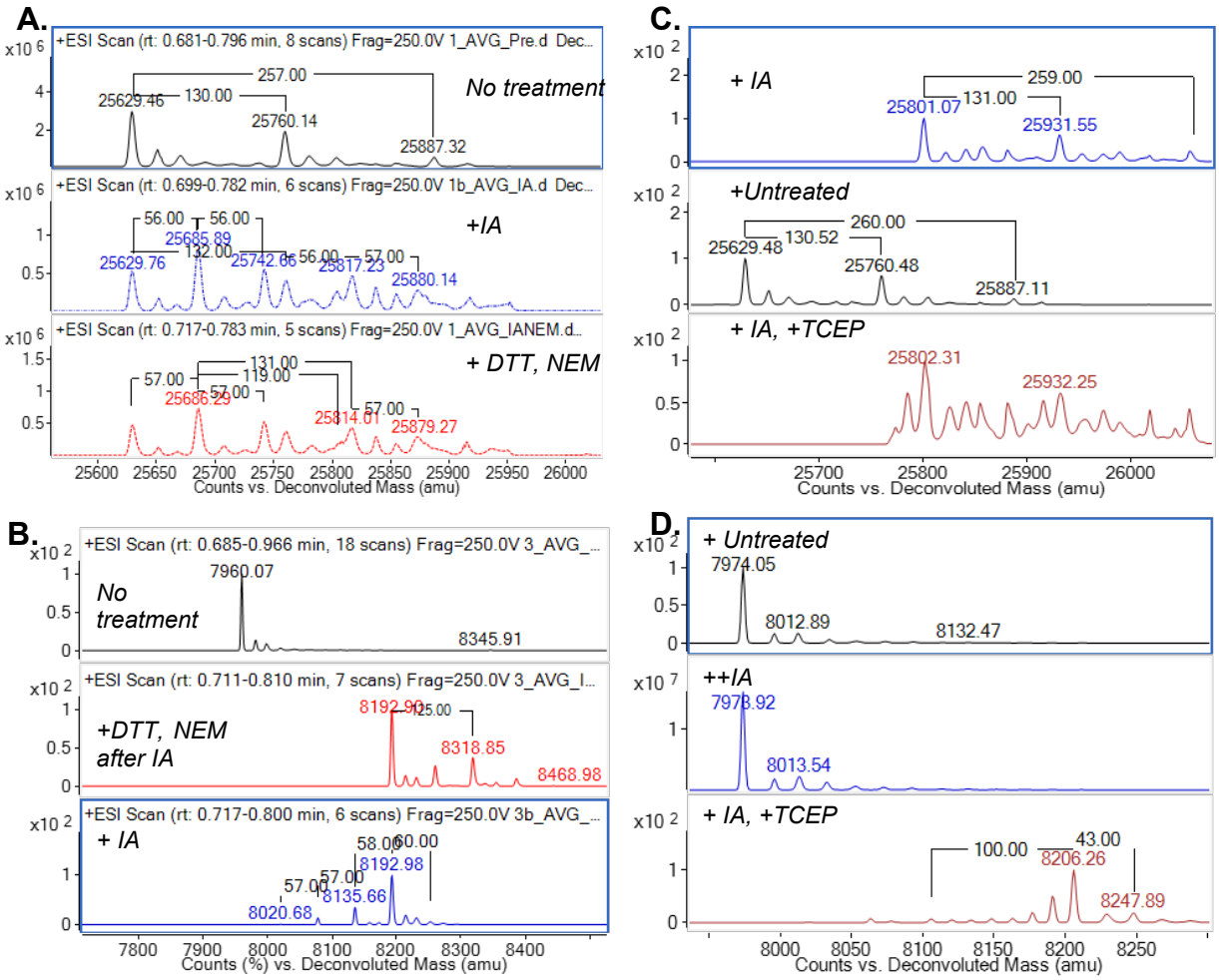


Figure 3.3: ESI-QTOF/MS reduction state measurements

A. ESI-QTOF/MS of reduced HMGB1 (black), treated with iodoacetamide under nondenaturing conditions (blue), and then dialysed into DTT followed by N-ethyl maleimide treatment (bright red). Only one cysteine was fully alkylated; partial IA adducts are visible **B.** Same experiment with CXCL12 V3I (mW 7978 Da); only one n-ethyl maleimide (NEM) adduct is visible and all 4 cysteines were alkylated by IA, which would suggest reduced protein despite -4 Da mW. **C.** ESI-QTOF/MS of reduced HMGB1 (black) in pH 8.0 8M Urea, treated with iodoacetamide in the presence (red) or absence (blue) of 2 mM TCEP. In both cases, all cysteines alkylate (reduced) **D:** Same experiment with oxidised CXCL12. Only when the protein is treated with TCEP the cysteines alkylate (oxidised)

Each free cysteine treated with iodoacetamide results in a +57-mass shift in both the presence and absence of TCEP; reduction of the protein adds a further +116 mass shift (114 from 2 alkylation + 2 protons) for each broken disulfide (e.g., +232 Da for CXCL12) if treated with iodoacetamide, and +250 (2 125 maleimide adducts).

3.4 Expression and purification of HMGB1

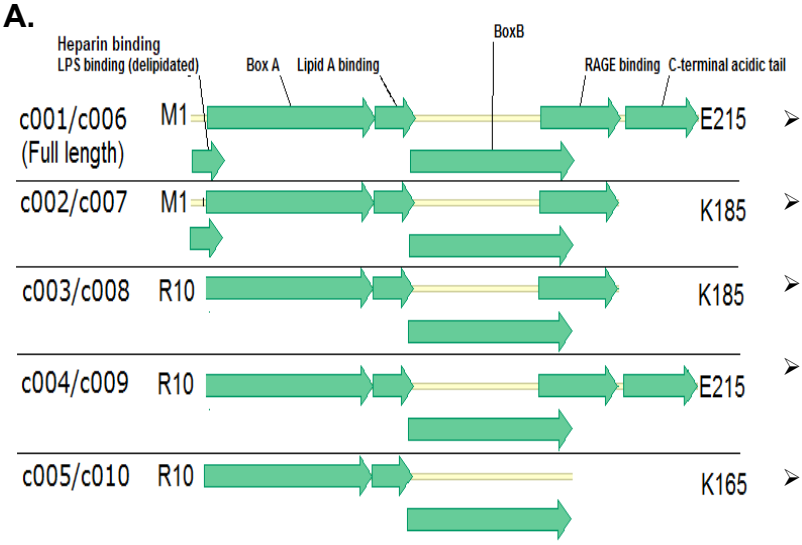
3.4.I Initial construct panel

HMGB1 has multiple unstructured regions, which would hamper crystallisation. These include the canonical RAGE binding site (149-182) [135,179], as well as the LPS binding regions (1-9, 79-95) [52], which are involved in proinflammatory activities. HMGB1 also undergoes C-terminal degradation due to the toxicity of the 200-204 peptide [33,152] when expressed recombinantly in *E. coli.* Truncation of the acidic tail markedly changes the (pI) and charge distribution of HMGB1, due to the different pI for HMG Boxes and the acidic tail. Ion exchange chromatography has been applied for separation of these truncated forms from full-length protein [152] when the protein is tagged at the N-terminus. However, adapting this methodology to constructs with different boundaries and therefore pI would require a high-throughput automated chromatography system and extensive optimisation.

Instead, I purified HMGB1 via a C-terminal His tag, which would readily eliminate the C-terminal truncated products as they would also lose the affinity epitope. Cytoplasmic expression was preferred as I needed both FR and DS-HMGB1. This avoids complications associated to periplasmic expression e.g., inefficient secretion or proteolysis, as the protein could be oxidised once purified.

This initial panel was designed conservatively, to produce a basic set of constructs which could both serve for future crystallography experiments (Figure 3.4) whilst removing regions associated to known proinflammatory functions such as binding to the glycan portion of LPS [52], mimicking thrombin cleavage (cleaves after R9, inactivating several proinflammatory functions), truncating the known RAGE binding site [180,181], or minimising bactericidal activity to increase yield in *E. coli* [33] by removing the acidic tail.

A.



Rationale

- Positive control (full-length protein)
- Removal of the bactericidal C-terminal acid tail, from E186(185)-E215(214)
- Removal of C-terminal acid tail and Gly2(1)-Pro9(8) peptide which binds delipidated LPS
- Removal of Gly2(1)-Pro9(8) peptide alone
- Minimal HMGB1 Box construct R10(9)-K165(164) with the RAGE binding peptide removed (166-185)

B.

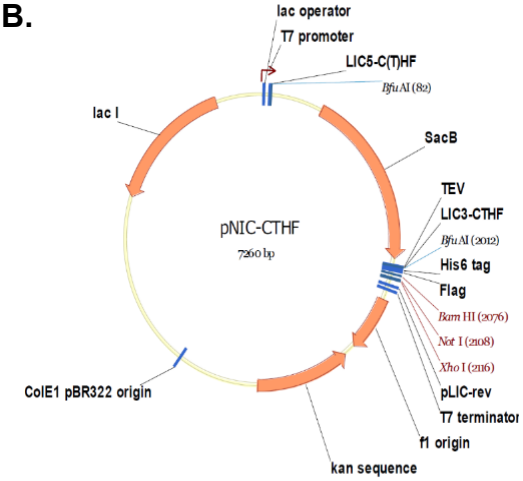


Figure 3.4: Initial HMGB1 constructs

A. Initial HMGB1A constructs. Two construct IDs are supplied: IDs 1-5 represent constructs derived from HMGB1A-s001 (wt) whereas IDs 6-10 represent constructs derived from HMGB1A-s002 (3S). Residue numbering corresponds to the polypeptide with the N-terminal methionine, which is lost post-translationally. Construct boundaries have the numbers excluding N-terminal Met in parentheses (**B**).Schematic of the expression vector pNIC-CTHF.

As seen in Figure 3.5, most constructs expressed and the purification tags were accessible for cleaving.[19] However, HMGB1A-c003 (9-184) underwent extensive degradation. It is interesting that other further C-terminally truncated constructs (HMGB1A-c005, 9-164, and c002, 1-184) or HMGB1 with the full C-terminal region (HMGB1A-c004, 9-214), do not suffer this degradation. which means that the lack of expression is unlikely to be solely dependent on the Gly1-Pro8 peptide being present. Furthermore, HMGB1A-c008 (the equivalent 3S-HMGB1 construct to HMGB1A-c003) was expressed without evident degradation. As expected, HMGB1A-c002/7 (no C-terminal tail) expresses to higher levels owing to lack of toxicity in *E. coli.*

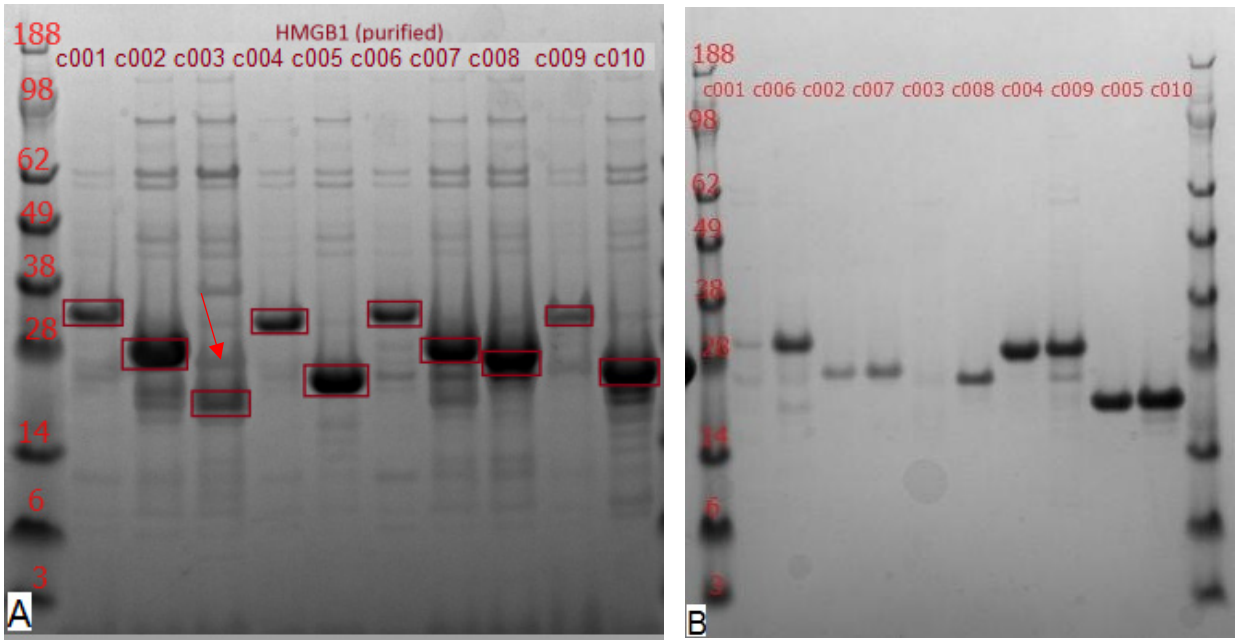


Figure 3.5: Test expression of initial HMGB1 panel

Test expression of HMGB1 constructs. Protein was purified using IMAC and cleaved with TEV protease, then analysed by SDS-PAGE. Bands analysed by MS/MS are highlighted; all identifications matched the target. The full-length putative HMGB1A-c003 band is designated by an arrow; note that the highest intensity band has a lower MW than this.
A: Soluble expression of HMGB1 constructs in 1 ml of BL21(DE3)-R3-pRARE2.
B: Change in MW of cleaved HMGB1 after tag removal and reverse IMAC. Scale: 50 mL

To see whether the loss of certain construct boundaries and lack of expression was related to protein activity, I analysed the in vivo activity of the four construct boundaries which expressed for wt HMGB1, plus 3S-HMGB1 (HMGB1A-c006).

[19] It should be noted that HMGB1 appears to migrate higher than the expected MW [152], due to the high density of positively charged groups, as do many other HMG proteins [37].

I used both an established muscle injury regeneration model, where HMGB1 is known to enhance tissue repair [25], and a chemotaxis assay, as HMGB1 can enhance the chemotactic activities of CXCL12 [96]; for the latter I used only 3S-HMGB1 as it had already been validated in this model and would not be subject to oxidisation throughout the test (>6 h at 37ºC).

Constructs truncated after Pro 8 did not show improved regenerative potential (Figure 3.6A) or chemotactic activity (Figure 3.6B) when compared to controls without HMGB1 treatment, whereas full-length FR (c001) or 3S (c006) HMGB1 did improve tissue repair. The loss of the C-terminal acidic tail (HMGB1A-c002) does not impair tissue healing.

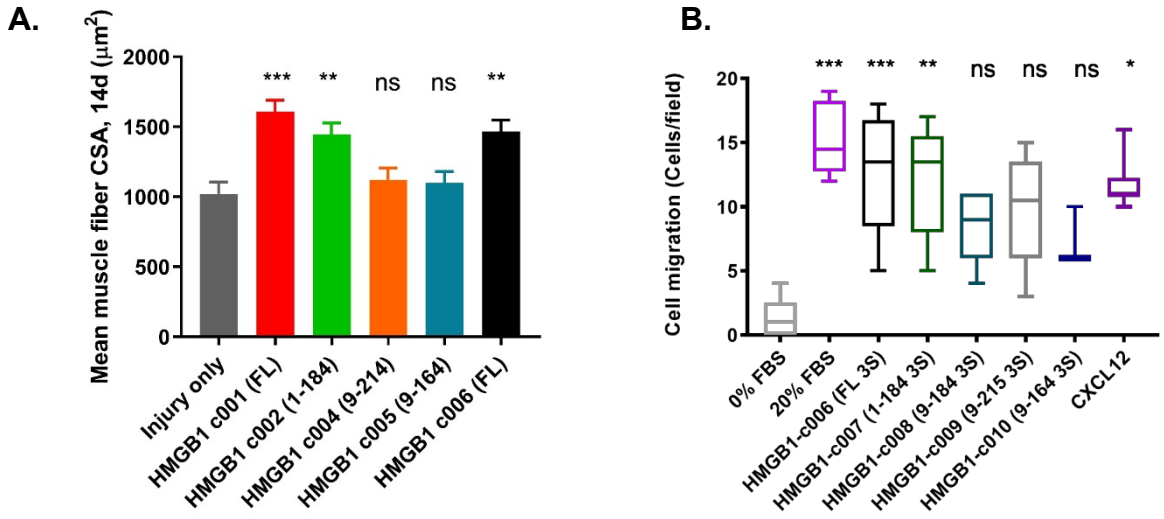


Figure 3.6: Removal of the Gly1-Pro8 peptide results in loss of enhanced tissue repair and chemotaxis

A. Results from barium chloride muscle injury model after local injection of 0.6 mg/kg of HMGB1 constructs (mean muscle fibre cross-sectional area (CSA) at 14 days after injury). A higher value indicates faster tissue repair; significance values with respect to control animals (Gaussian, non-heteroskedastic dataset; 1-way nested ANOVA test with Holm-Sidak correction for multiple comparisons). N1= 5 animals per group, n2= 10 cells per animal **B.** Boyden chamber transwell migration assay**.** 3T3 mouse fibroblasts were exposed to 1 nM of 3S-HMGB1 constructs or commercially purchased CXCL12 (as I had yet to optimise expression and purification) and fixed after 6 hours. 6 image fields were counted per condition. Statistical comparison with respect to 0% HI-FBS control (non-Gaussian, heteroskedastic dataset; 1-way Kruskal-Wallis test with Dunn´s multiple comparison). Plots represent, top to bottom 1st, 2nd (median) and 3rd quartiles. Whiskers set to min-max range. n= at least three fields counted per condition (migrated cells per field).

One possible hypothesis to explain the extensive degradation in HMGB1A-c003 and the activity loss of HMGB1A-c004 and c005 (which are all truncated after Pro 8) is that either loss of some residues in the N-terminal domain induces partial unfolding of Box A during translation leading to instability, or that this domain contains residues critical for CXCL12 interaction, despite not having been described as such before [131,182,183].

3.4.II Large scale expression and purification of HMGB1[20]

3.4.II.i Ion-metal affinity chromatography (IMAC) (pNIC-CTHF constructs)

HMGB1 is a highly effective DNA binding protein, forming large cruciform structures in its presence [22], and this was reflected by large DNA carryover during purification. This occurred despite the combined use of benzonase (w/ 1 mM $MgSO_4$), moderate ionic strength to dissociate the DNA from the protein (0.5 M NaCl) and polyethyleneimine (PEI) precipitation of nucleic acids (0.15% w/v final, for 30 minutes at 4 °C). HMGB1 co-eluted with this residual DNA in the void volume of SEC, which resulted in large losses due to the DNA-contaminated fractions being unsuitable for biological work (Figure 3.7). The main protein elution peak had still substantial DNA content, with A260/280 values for purified protein > 1.2 [184–186].

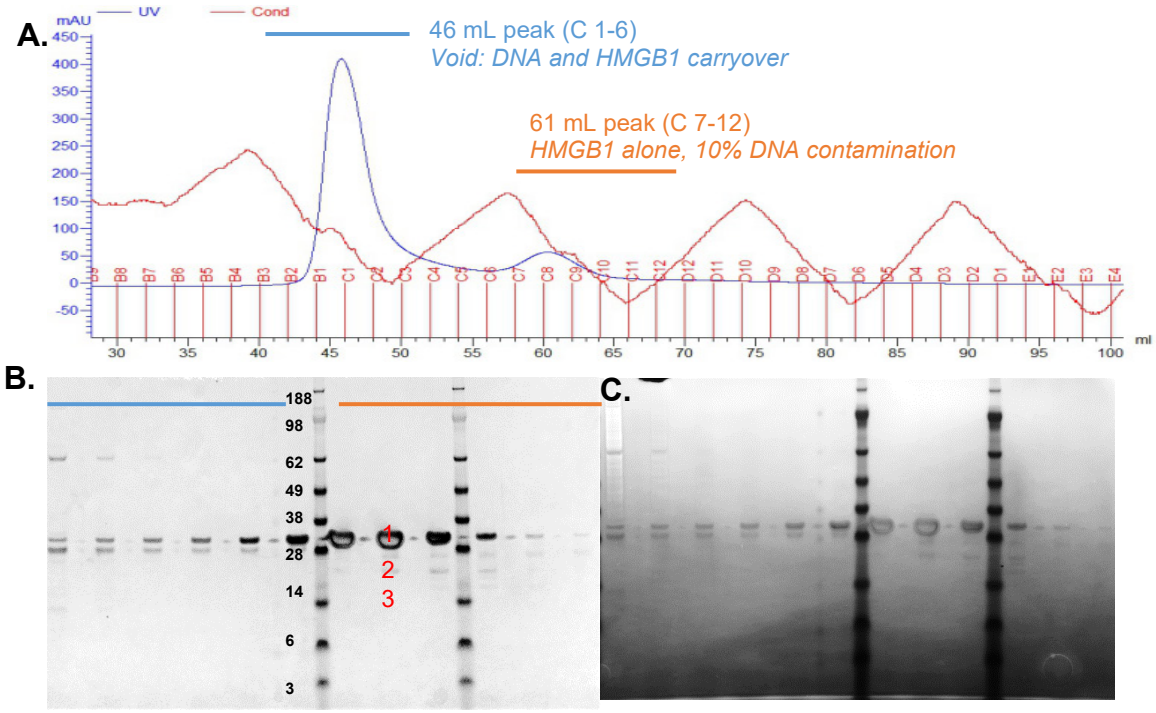


Figure 3.7: DNA contamination of HMGB1 after size exclusion chromatography (SEC).

HMGB1A-c006 was injected into a Superdex 75pg 16/600 column and eluted at 1.2 mL/min in 0.3M NaCl 5% Glycerol 50 mM HEPES pH 7.5. **A.** Chromatogram of the experiment. Two main peaks appear: one at 46 mL and another at 61 mL. **B.** SDS-PAGE gel (Coomassie stain) of peak fractions. Bands identified by tryptic digest MS-MS (Suppl. data A1-3-A): 1 = FL HMGB1A-c006; 2, 3 = C-terminal truncated fragments (see 3.3.IV). **C.** Gel from (A), stained with ethidium bromide and imaged under UV light. The first peak fraction contains HMGB1, TEV and large amounts of DNA; pooled fractions C7-C9 had approximately 5-10% DNA content.

[20] Since HEPES does not allow CXCL12 dimerisation [114], nor does it suffer significant pH alteration with temperature as Tris, it became the buffer of choice for my purification protocols.

I hypothesised that the binding of HMGB1 to DNA could be protecting it from both endonuclease degradation and preventing the formation of complexes with PEI, which could be separated from the protein fraction by centrifugation. HMGB1 is extractable from chromatin with 0.35 M NaCl [44], but since 0.5 M NaCl proved insufficient to separate DNA during lysis, I increased the concentration to 1 M during cell lysis. This resulted in recombinant proteins with A260/A280 ratio <0.58[21]. The protein remained soluble throughout the process up to and after SEC and retained biological activity (Figure 3.6 and Figure 3.15). However, a significant fraction of total protein (≈50% of total yield) was lost during DNA removal with 1 M NaCl, of equivalent purity to the elution fractions (Figure 3.8). Nevertheless, protein yields for the initial construct panel (HMGB1A-c001 to HMGB1A-c010) ranged between 5 and 10 mg/L of TB+ or AIM-TB for full-length constructs. Removal of the C-terminal acidic tail (toxic to *E. coli* [33]) vastly increases yields as expected. This became my IMAC purification method through the first years of my DPhil[22].

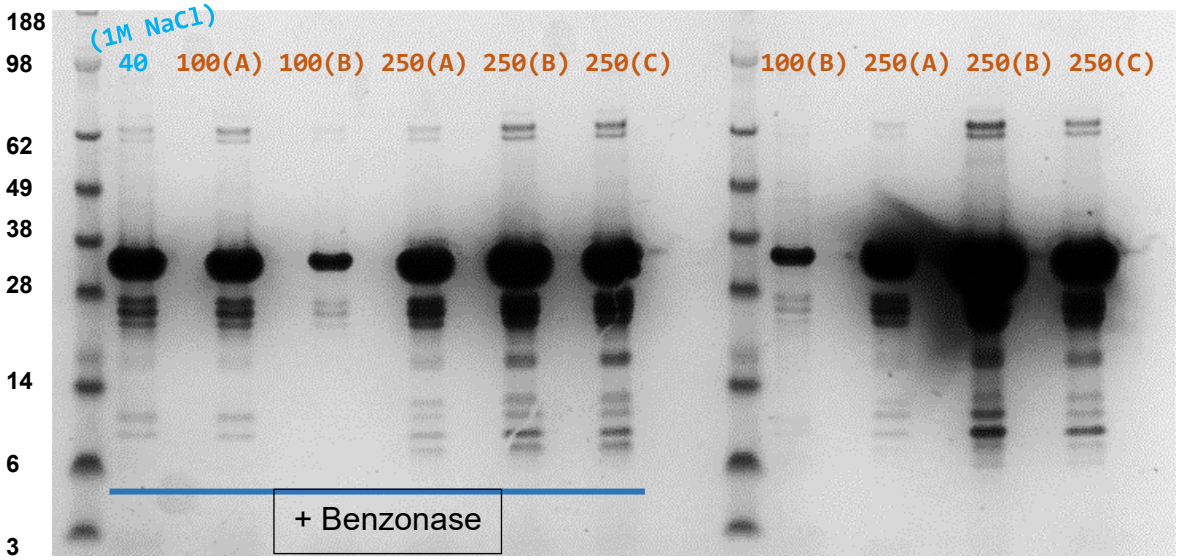


Figure 3.8: 1 M NaCl purification of HMGB1A-c001

IMAC purification fractions of HMGB1A-c001. The imidazole concentration (in mM) is indicated on top of each lane. Fractions coloured blue had 1 M NaCl: in brown, 0.5 M NaCl. The 40 mM imidazole, 1 M NaCl wash fraction (40 mL) has similar protein purity and concentration to the 100 mM imidazole elution fraction (0.5 M NaCl) (5 mL) but much higher A260/A280 ratio (1.12).

[21] Benzonase did not affect DNA in this protocol despite a visible loss in viscosity being observed in lysates upon its addition (DNA lysis) as PEI precipitated all the DNA dissociated from HMGB1.

[22] DSF results validated that these conditions were tolerable by the protein (Section 3.4.I.iii).

3.4.II.ii High-throughput IMAC

However, at a later stage of my DPhil (see Chapter 5, section 5.3), I encountered the need to purify large numbers of constructs simultaneously (> 24) in sufficient quantities for *in vivo* and *in vitro* experiments (1-2 mg). Whilst the base protocol yielded good results, several steps were unsuitable for high-throughput production:

- Different HMGB1 constructs, particularly those that include the C-terminal acidic tail, adversely affect E. coli cell viability. Therefore, growth rates and induction times are **difficult to synchronise for multiple samples**.
- **Lysis via sonication** (or other mechanical protocols) **is not readily performed for several large samples**. Due to the necessity to remove LPS and the subsequently large expression experiments (see 2.13), at least 30-40 g wet weight of cells need to be processed for each sample and could not be done in parallel with the available equipment.
- In addition, PEI precipitation requires high-speed capable centrifugation, > 16000 RCF; much as with sonication, equipment limitations prevent up-scaling of this process.

I hypothesised that protocols using auto-induction media available at the SGC for high-throughput production of proteins for crystallography[23] could be adapted for DNA-binding proteins like HMGB1 to facilitate screening protocols. PEI precipitation is substituted by a higher concentration of benzonase, and chemical lysis replaces sonication [146]. Use of AIM-TB eliminates the need for individual OD monitoring and allows for higher final growth densities after induction, reducing culture volumes, and IMAC chromatography is readily scalable in parallel, with the only limit being sample tracking.

However, this failed to yield DNA-free HMGB1, and the quality of the preparations was inconsistent (Figure 3.9); furthermore, yields for full-length HMGB1 constructs were too low with that scale of preparation (100 mL in autoinduction medium) for LPS removal. Increasing the scale would run again into issues with sample throughput. Since HMGB1 is toxic to *E. coli* when expressed in full-length form, I decided to exploit this property to solve this problem by increasing induction strength; higher induction of

[23] I would like to acknowledge Dr. Michael Fairhead at the SGC for devising of the original protocol which I adapted.

HMGB1 expression would theoretically hamper cell growth, leading to smaller pellets, but potentially similar protein yield/L. I added as well 1% glucose to agar plate, preculture and expression culture to fully repress background *lac* activity [187], to synchronise growth rates for all constructs. As IPTG is not reliant on the lactose transporter to enter the cells, it can induce expression in the presence of high glucose concentrations [187,188]. This led to a large reduction in cell mass (from 40-50 g/L culture to 10-15 g/L) but no decrease in yield of protein (mg/L culture) or changes in post-translational modifications (see 3.4.III).

To fully eliminate DNA contamination without excessive losses from high salt concentrations in combination with imidazole, as in my previous protocol (3.4.II.i), I added on-column digestion of DNA with benzonase after a high-salt (1 M) wash in the absence of imidazole (to dissociate DNA from HMGB1); either step alone was insufficient to fully remove DNA. This eliminated DNA contamination whilst achieving comparable yields/L to the original high-throughput protocol (Figure 3.9). As the cell mass is decreased by a large factor, the scale of culture could be increased up to 0.5-1L whilst allowing for quick and straightforward processing besides the initial centrifugation. This became my final IMAC-based HMGB1 purification.

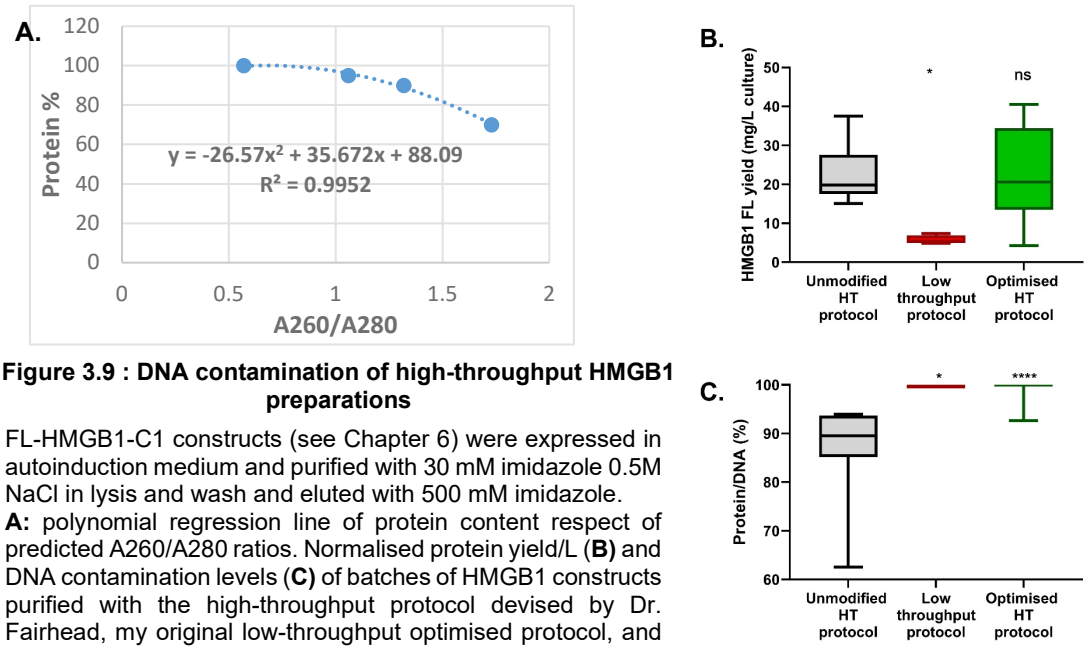


Figure 3.9 : DNA contamination of high-throughput HMGB1 preparations

FL-HMGB1-C1 constructs (see Chapter 6) were expressed in autoinduction medium and purified with 30 mM imidazole 0.5M NaCl in lysis and wash and eluted with 500 mM imidazole.
A: polynomial regression line of protein content respect of predicted A260/A280 ratios. Normalised protein yield/L (**B)** and DNA contamination levels (**C)** of batches of HMGB1 constructs purified with the high-throughput protocol devised by Dr. Fairhead, my original low-throughput optimised protocol, and optimised high-throughput protocol with more stringent induction conditions. Data analysed by Kruskal-Wallis test with Dunn´s post-hoc analysis. Unmodified HT protocol: n=17 samples, 50 mL, AIM-TB. Low throughput protocol: n=4 samples, 1 L of TB+ with 0.5 mM IPTG induction. Optimised HT protocol: n= 24 samples, 0.2 L of TB+ with 0.5 mM IPTG induction, with wash at 1 M NaCl.

3.4.II.iii Biotinylated HMGB1 constructs

I designed several constructs with the intent to couple them to streptavidin chips (see Chapter 4) and measure the binding affinity of CXCL12 to diverse HMGB1 constructs. These were all expressed in the pNIC-Bio3 vector, which has a histidine-tag in the N-terminus, and a biotinylation site at the C-terminus (Figure 3.10). This allowed me to purify C-terminal degradation products via affinity to streptavidin and avidin resins [152] in addition to the IMAC and reverse IMAC steps. Of these, streptavidin proved to be the better choice due to higher capacity but similar specificity (Figure 3.11). As HMGB1 is able to dimerise and form higher order oligomers under different conditions [51], some carryover of C-terminally degraded forms was observed Increasing salt concentration during binding and wash to 750 mM NaCl allowed for elimination of these.

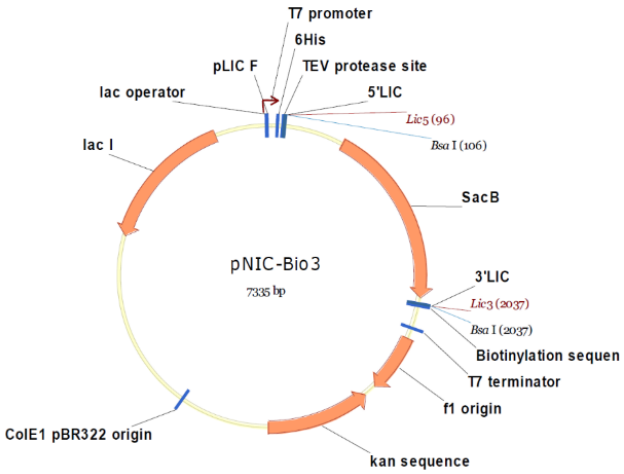


Figure 3.11: Vector map of pNIC-Bio3

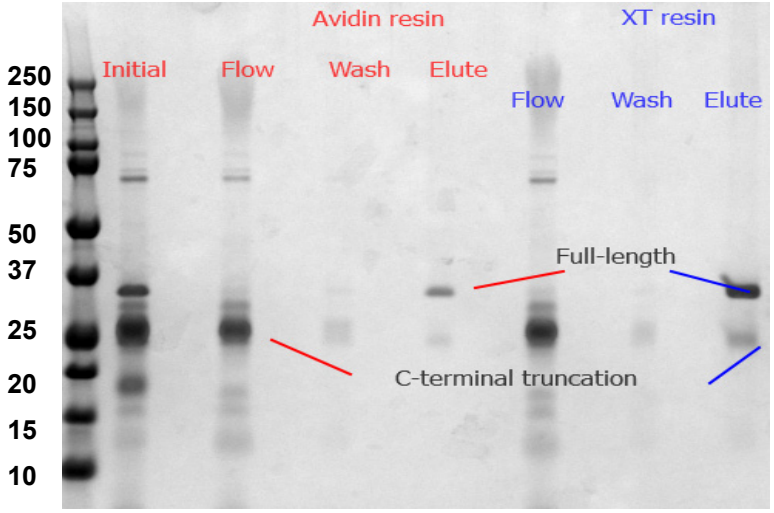


Figure 3.10: Biotin-affinity purification of HMGB1 constructs

Purification of HMGB1A-c011 (full-length FR-HMGB1) from both monomeric agarose resin (1 mL per OD unit) and Strep-Tactin XT resin (0.2 mL per OD unit). Protein was allowed to bind to the resin for 1 hour and washed with 30 CV of 10 mM HEPES, 150 mM NaCl, followed by elution with 3 CV of 2 mM biotin (Pierce monomeric agarose resin) or 50 mM (Strep-Tactin XT resin). Both resins seem to be capable of retaining most of the biotinylated HMGB1 in the sample, but the use of avidin resin led to much larger sample dilution than with the Strep-Tactin XT due to the much lower dynamic binding capacity of the former (0.2 mg/mL for avidin and 2-3 mg/mL for Strep-Tactin XT) Therefore this was the medium of choice, despite the less reversible binding of biotin to Strep-Tactin XT requiring higher biotin concentration for elution and longer elution times.

3.4.III PTMs of purified HMGB1

When purified from *E. coli*, HMGB1 only experiences minor PTMs. N-terminal methionine loss, gluconoylation/phosphogluconoylation and oxidation (when prepared in the absence of reducing agents) could be observed (Figure 3.12, right). All these modifications follow the same charge state distribution on the M/Z spectrum (Figure 3.12, left), which suggests that they have no overall effect on protein unfolding.

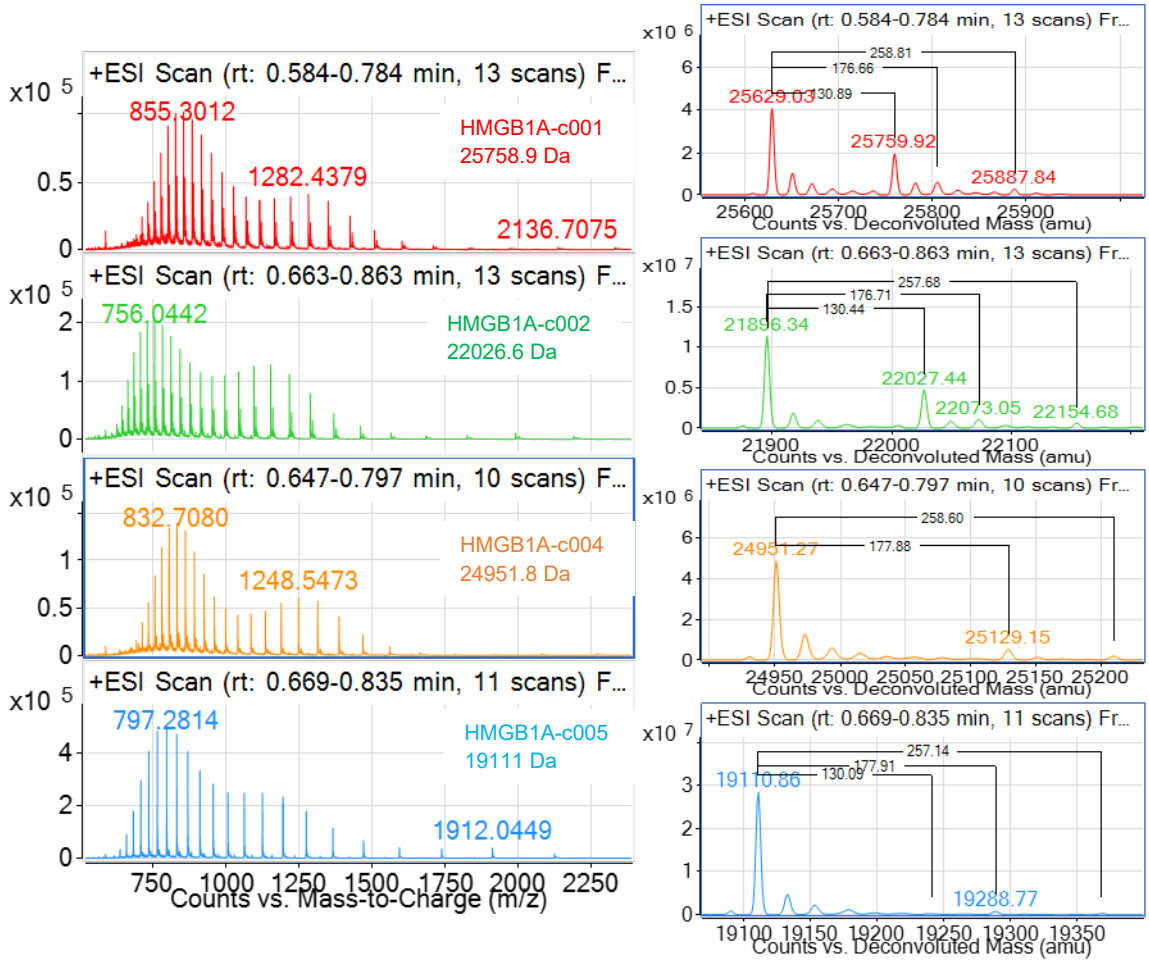
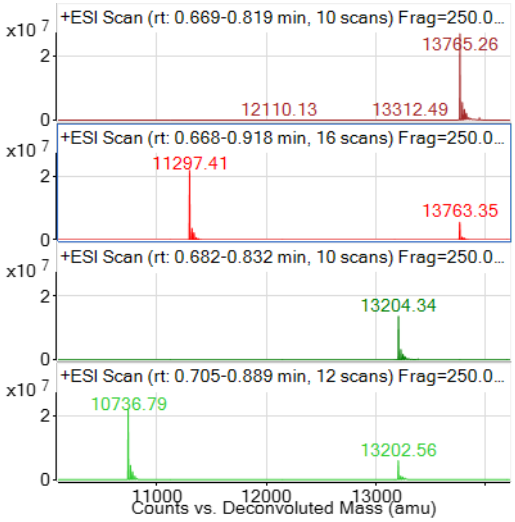


Figure 3.12: PTM analysis of FR-HMGB1 constructs (initial panel)

Q-TOF spectra of the different FR-HMGB1 constructs from the initial panel (HMGB1A-c001/c002/c004/c005). Observed PTMs: -131 Da (N-terminal methionine loss), +258 (phosphogluconoylation), +176 (gluconoylation). Theoretical average MW of the constructs (reduced, cleaved, without any PTM) is annotated in each M/Z spectrum. Unannotated peaks correspond to +21/+22 Na adducts (with or without proton loss). 2X and ½ X mass deconvolution artefacts were present (characterised by a +11 or +42-44 Na adduct series rather than a +21 adduct)

Constructs truncated after Pro 8 experience no N-terminal methionine removal. Arg 9 has a long side chain, a factor known to affect excision by methionine aminopeptidase [189]. Later constructs starting at Pro8 or Pro94 (in the case of HMG Box B constructs) did not experience methionine loss and had deficient TEV cleavage (≈60% efficiency) of N-terminal tags (e.g., pNIC-Bio3). Proline in position P1 is known to block TEV activity, likely due to the loss of flexibility in the protease cleavage site resulting in it being too close to the main protein chain for cleaving to properly occur. [190]. Whilst the TEV site in my constructs has a Ser residue on that position (ENLYFQ/SM), having a Pro residue in the P3´ position is likely to cause the same loss of flexibility, to some extent, and decrease cleavage efficiency.



	c027	c028
Tagged	13536.3	12975.6
Tagged, biotinylated	13762.3	13179
Cleaved	11070.6	10510
Cleaved, biotinylated	11296.6	10736

Figure 3.13: Deficient N-terminal TEV cleavage in constructs with proline in P3´

Intact mass analysis of HMGB1-c027 and HMGB1-c028 (both biotinylated) before and after TEV treatment. Theoretical average MW (in Da) of the different species is annotated on the table to the side. Spectra of the forms without TEV treatment has a systematic +2 Da error due to instrument calibration issues at the time of the experiment

3.4.I Characterisation of biologically active HMGB1 constructs

Whilst the core of Box A does not start until Ser 14, it could be is possible that the flexible loop adjacent to the N-terminus of HMGB1 has a role in stabilising the Box domain. The published NMR structure of HMGB1 1-164 (PDB 2YRQ) shows that Proline 8 in Box A, and Proline 94 in Box B, seemingly configure a molecular lock which could help stabilise the Box structure (Figure 3.14A). The peptide bond of each proline with residue n+1 is in proximity (of the hydroxyl group of a tyrosine ring (Tyr 70 in Box A and 154 in Box B), and an apolar side chain (Met 74/Ile 158).

These three residues have similar spatial arrangement across the conformers shown in the NMR structure, particularly in Box B (Figure 3.14B); superposition of all 20 conformers according to the structure of only Box B (88-162) results in a near perfect match. Additionally, if the sequences of Box A and B are aligned, these residue pairs are in conserved positions across each HMG Box (Figure 3.14C). In Box A, the OH group of the tyrosine preferentially occupies positions across the conformers which allow a hydrogen bond with the backbone oxygen of Pro 8 (average distance across conformers 1-10, 2.8±0.2 Å; hydrogen bonds can occur up to 3 Å [191,192]), but it has a much wider movement range than Tyr 154, which suggests more conformational flexibility for Box A than Box B. In Box B, however, this hydrogen bond seems less likely, as Tyr 154 is too far from Pro 94 in all conformers (> 5 Å)

Whilst this was a purely speculative hypothesis, I decided to test whether restoration of Proline 8 would allow expression of HMGB1A-c003 (9-185), and potentially restore biological activity. I designed a second set of HMGB1 constructs with Pro 8 restored at the apex of Box A in both wild-type (derived from HMGB1A-s001; c012 to c014) and 3S (derived from HMGB1A-s002; c015-c017). These correspond to the C-terminal truncations at positions 184 (c012/15), none (c013/16) and 164 (c014/c017). I also designed a variant of HMGB1A-c005 with an intact N-terminus but truncated at position 164 (HMGB1A-c031). All these constructs expressed at levels comparable to wt-HMGB1 and 1-184 HMGB1A (Figure 3.14D). I did not test, however, the *in vivo* activity of these truncations as later experiments (Section 4.4) proved that, whilst resulting in stable protein, a truncation at this position removes residues critical for CXCL12 interaction, which explains the loss of biological activity in constructs containing this truncation. This, however, reinforced the hypothesis that indeed Proline 8 is involved in the correct folding and stabilisation of HMGB1 Box A.

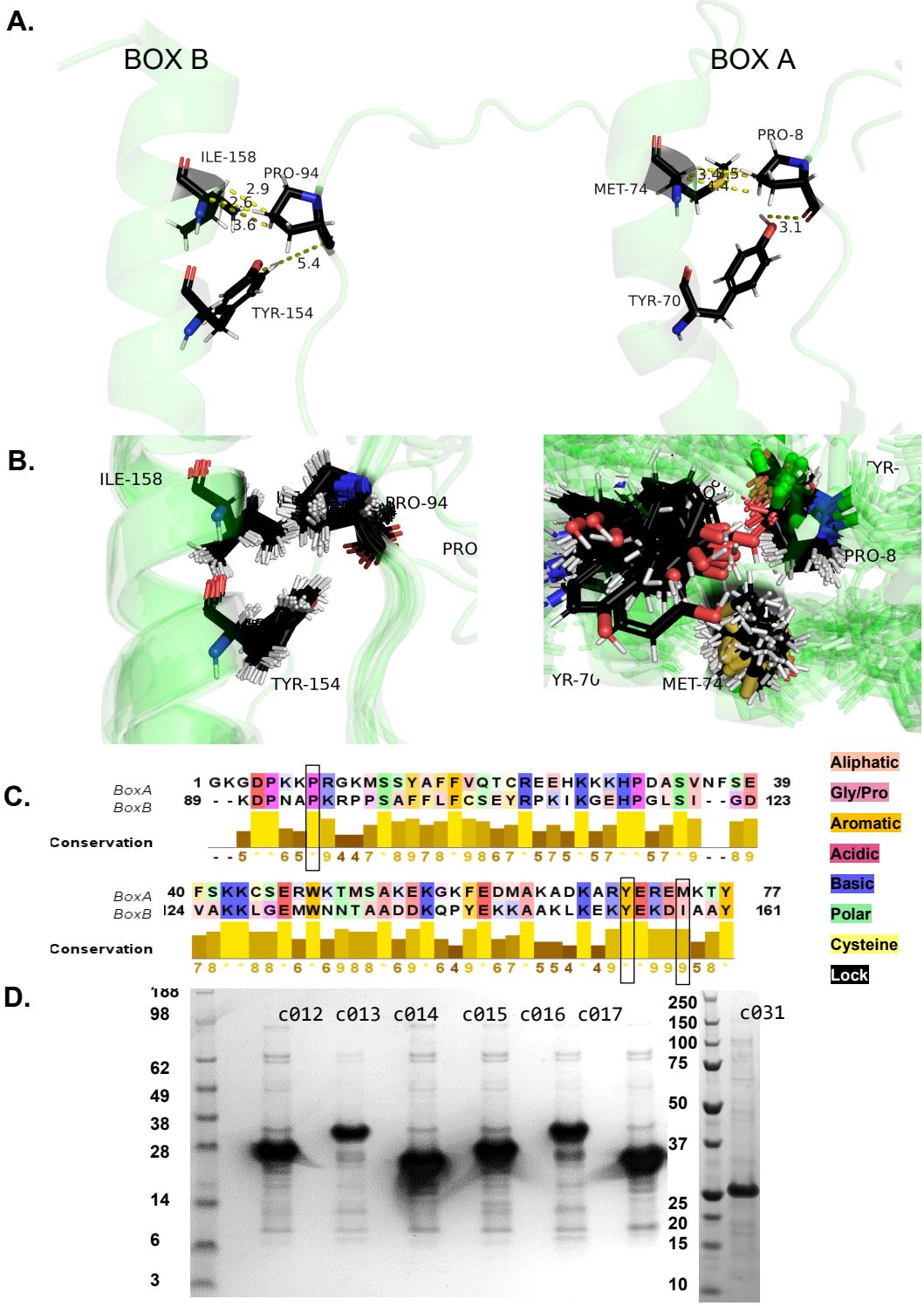


Figure 3.14: "Locking" region in both Box A and Box B

A: "Locking" mechanism at the apex of the HMG Box domains A and B**.** Distances between the closest C-H groups of the proline and the tyrosine and aliphatic chains are detailed (in Å). The residues involved in the lock have black backbone (Pro8, Tyr70 and Met74 in Box A; Pro94, Tyr154 and Ile 158 in Box B.
B: Superposition of all 20 conformers of HMGB1 contained in PDB 2YRQ according to only the sequences of Box B (Lys89-Tyr161), in the left, or Box A (Gly1-Tyr77) in the right. Pro8 and Pro 94 do not deviate between conformers from the relative positions, whereas Tyr 70 and to a lesser extent Met 74 do. The whole of Box B is much more rigid than Box A, with little to no variation across conformers. **C:** Sequence alignment between residues of Box A (Gly1-Tyr77) and Box B (Lys89-Tyr161). Residues involved in the proposed locking mechanism are surrounded by a black square; conservation histogram is shown ("+" = no conservation, "*" = full identity. All three positions (Pro8-Pro94, Tyr70-Tyr154 and Met74-Ile158) are highly conserved. D. SDS-PAGE gel from test expression (1 mL) of HMGB1 constructs with Pro 8 restored at the apex of Box A (c012-c017) and 1-164 truncate (c031). Specific construct identifiers have been annotated.

I did, however, test the biological activity of construct HMGB1A-c031 (G1-K164 truncation; Figure 3.15) in comparison to the other active HMGB1 constructs. Regenerative activity was partially restored when the N-terminal peptide was present again, potentially confirming the results where truncation after Pro 8 ablates HMGB1-enhanced tissue repair. This is consistent with findings that report HMGB1 losing its proinflammatory activities upon thrombin cleavage after Arg 9 [193], and would suggest that this loss of activity is due to a change in the structure in Box A, rather than to removal of specific functional peptides.

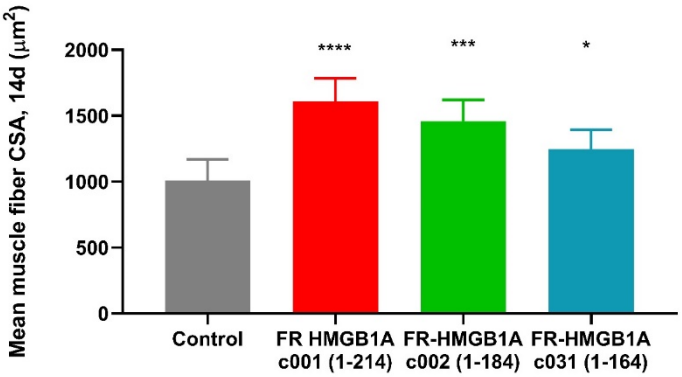


Figure 3.15: Biological activity of HMGB1A-c031 (1-164)

Results from barium chloride muscle injury (mean muscle fibre cross-sectional area at 14 days after injury from experiments with FR-HMGB1A-c001, c002 and c031; significance values with respect to control group (nested ANOVA with Holm-Sidak post-hoc correction). N1= 5 animals per group, N2= at least 40 cells per animal. Proteins were free of LPS. The differences with the respective control seem to decrease with further truncation. As c001/c002 and c031 were tested in different batches, the control groups were compared; no difference could be found between controls. However, the c001 group (positive control) assayed with c031 did not show enhanced regeneration.

Removal of Gly 1 – Pro 8 appears to influence protein stability, and this could be due to a loss of stabilising interactions in Box A with Pro 8. However, this alone is not sufficient to fully destabilise the protein; it is only apparent when it is combined with the 165-184 fragment being present and absence of acidic tail, with complete loss of expression. It is known that removal of the acidic tail alone decreases the thermal stability of HMGB1 [48,50] due to of the loss of stabilising interactions with the HMG Box core domains, and conversely, mutation of the three cysteines to serine can also counteract this stability.

I assessed the stability of all the biologically active HMGB1 constructs to attempt elucidation of the factors influencing HMGB1 stability: HMGB1 full length (c001; 1-214), tail-less (c002; 1-184) and a minimal HMG Box construct (c031 1-164). All of these were tested in both oxidised and reduced forms, as well as full length HMGB1-3S (c006; 1-214). All proteins contained a TEV scar on the C-terminus from purification.

3.4.I.i SEC profiles of different HMGB1 constructs

I first analysed the behaviour in SEC of these different FR-HMGB1 constructs. Lower retention volumes (RV) generally can be associated to a larger hydrodynamic radius and it can be approximated to an estimate of the solvent accessible surface area (SASA) and molecular weight (MW) (Suppl. data A3-1-A, Figure 8.1, [194]), although this can be affected by other factors such as charge and column interaction and is not a direct correlation. I performed the assay in two types of buffer at different concentrations to assess the effect of pH and ionic strength (the latter indirectly measured by conductivity): either PBS pH 7.4 (carrier buffer for *in vivo* experiments) or ammonium acetate pH 6.5 (for native mass spectrometry).

The behaviour of HMGB1 constructs in 1X or 2X PBS followed a predictable pattern. The constructs elute in reverse molecular weight order (higher to lower size) and in monodisperse, single peaks, with no difference in RV for disulfide constructs. (Figure 3.16). RV ranged between 10.4 to 11.5 mL in 1X/2X PBS across constructs. HMGB1 seems to elute much earlier than a globular protein standard with comparable mW. Ovalbumin (mW 43 kDa, SASA 15586 Å) had a retention time of 10.8 mL in 1X PBS (Suppl. data A3-1-A); FL-HMGB1 eluted at earlier volume despite being 25 kDa in size, and truncated constructs . This would suggest that the different forms of HMGB1 all have much larger SASA than what could be predicted by their molecular weights alone.

Full-length HMGB1 forms were comparatively much less affected by ionic strength than tail-less constructs, which exponentially decrease retention volume with increasing conductivity at < 6 mS/cm due to likely nonspecific column interaction. Disulfide forms were much more affected by this, with DS-HMGB1 1-184 even failing to elute in this peak in all cases (Figure 3.16). This was observable by the large amount of sample carryover between runs for these constructs, much higher than that observed for FL-HMGB1 proteins (Figure 8.3)[24], and the apparition of secondary peaks at 19-20 mL which likely represent elution of protein adsorbed to the column. Back-pressure also increased significantly after the experiments at low ionic strength due to the protein accumulating in the column. Retention volumes are summarised in Table 3.1.

[24] Whilst a more complete experiment would also have included at least the 3S version of HMGB1A-c002 (HMGB1A-c007), and/or potentially any of the R9 truncations, I would not have been able to perform all of the chromatography runs together, introducing potential technical variation.

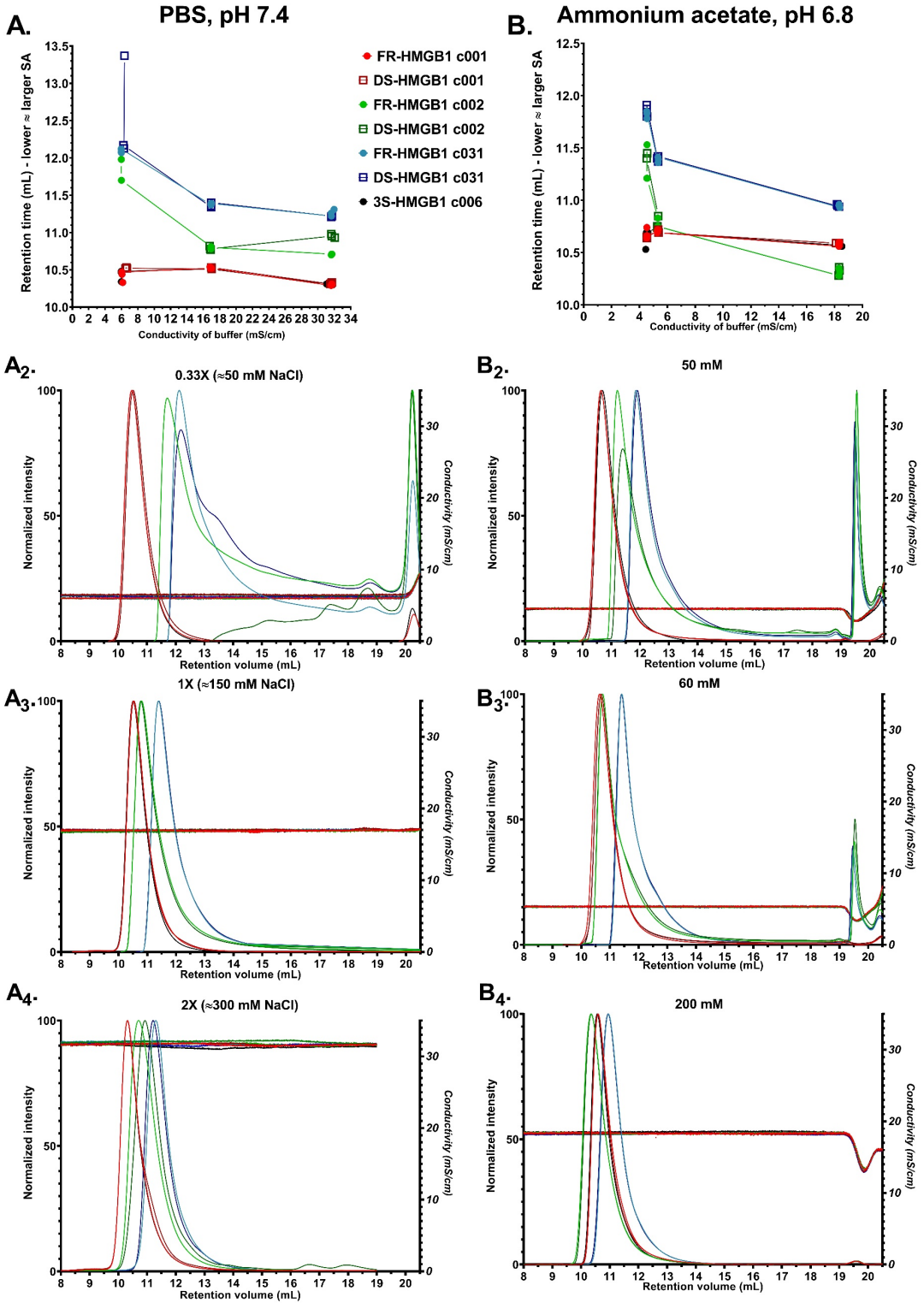


Figure 3.16: SEC profile of different HMGB1 constructs

Retention volume analysis of HMGB1A-c001 (FL, red), c002 (1-184, green), c031 (1-164, blue), both reduced and oxidised (darker hue), and HMGB1-c006 (3S, FL, black). 0.3 mL of 1.25 mg/mL samples of each construct were injected into a Superdex 10/300 75 Increase column pre-equilibrated in buffer and eluted at 0.5 mL/min. Each RV is the average of three independent injections. **A.** Plot of retention volume versus average conductivity of each run (8-18 mL) in PBS buffers. **A2, A3, A4.** Representative SEC profiles of 0.33X, 1X, 2X PBS (pH 7.4) respectively, with conductivity plots (right axis) and normalised UV trace (left axis). **B.** Plot of retention volume versus average conductivity of the run (8-18 mL) in ammonium acetate buffers. **B2, B3, B4.** Representative SEC profiles of 50, 60, 200 mM ammonium acetate (pH 6.8), with conductivity plots (right axis) and normalised UV trace (left axis). Conductivity traces are shown in Suppl. data A3-1-B (Figure 8.2).

A.

		Conductivity (mS/cm)	FR-HMGB1A, c001, FL	DS-HMGB1A, c001, FL	FR-HMGB1A, c002, 1-184	DS-HMGB1A, c002, 1-184	FR-HMGB1A, c031, 1-164	DS-HMGB1A, c031, 1-164	3S-HMGB1A, c006, FL
Amm. Ac, pH 6.8	50 mM	4.53±0.01	10.7±0.1	10.7±0.1	11.3±0.2	11.5±0.1	11.9±0.1	11.9±0.1	10.7±0.1
	60 mM	5.33±0.01	10.8±0.1	10.7±0.1	10.8±0.1	10.9±0.1	11.4±0.1	11.5±0.1	10.7±0.1
	200 mM	18.31±0.02	10.6±0.1	10.7±0.1	10.4±0.1	10.4±0.1	11±0.1	11±0.1	10.6±0.1
PBS, pH 7.4	0.33X	6.21±0.24	10.5±0.2	10.6±0.1	11.9±0.2	N/A	12.3±0.1	N/A	10.5±0.1
	1X	16.94±0.01	10.6±0.1	10.5±0.1	10.8±0.1	10.9±0.1	11.4±0.1	11.4±0.1	10.6±0.1
	2X	31.65±0.04	10.3±0.1	10.4±0.1	10.7±0.1	10.9±0.1	11.3±0.1	11.2±0.1	10.5±0.1

Table 3.1: Summary of retention volumes from HMGB1 constructs in Superdex 75 Increase, 10/300

A. Peak retention volume (in mL, mean±SEM) of each HMGB1 construct in size exclusion experiments shown in Figure 3.19. All data were normalised by the ratio between conductivity of the run/average conductivity of all runs in that buffer. Lower RV (≈ higher SASA) in purple; higher RV (≈ lower SASA) in orange. Midpoint in green. Retention volumes averaged over three independent runs. HMGB1A c002 and c031 have large increases in retention volume at lower ionic strength, due to likely column interaction. HMGB1A-c002, and oxidised constructs, appear to be particularly affected by this nonspecific binding as a larger proportion elutes in a secondary peak after 21 mL (Figure 3.18). A comparison between the proteins present in the elution peaks in 50 mM ammonium acetate and 200 mM is shown in Suppl. data A3-1-B (Figure 8.3).

Buffer pH also seemed to have a role in retention volume. The RV increase for tail-less forms of HMGB1 at low ionic strength was much larger in 0.33X PBS (pH 7.4, 6.2 mΩ) than in 60 mM ammonium acetate (pH 6.8, 5.33 6.2 mΩ), despite the higher conductivity. Tail-less constructs (pI>9.6) decreased retention volumes with decreasing pH in the 4.5-18 mS/cm conductivity range (100-330 mΩ ionic strength; equivalent to 40-165 mM NaCl) or even fail to elute when oxidised (increased column binding), whereas full-length HMGB1 constructs (pI≈5.6) remain relatively unaffected by pH changes. At higher ionic strength but lower pH, HMGB1A-c002 elutes before HMGB1A-c001/c006 (higher SASA). The increased binding at low ionic strength being affected by pH and reduction state could be explained by the different isoelectric points of the constructs and the interaction of the tail with the HMG Box domains.

HMGB1 Box B contains fewer charged groups than Box A, particularly at the bottom and frontal surfaces (Figure 3.17; Yellow/Red/Blue-YRB colouring [195]). Box B has less charged and more apolar side chains than Box A, but similar pI (Box B: pI 9.42, aliphatic index 55.3, 36% charged side chains, and Box A: pI 9.59; aliphatic index 22.14, 41% charged side chains). As pH decreases, the amine groups would be more likely to be protonated and the charges would repel better the column. In full-length HMGB1, the acidic tail decreases nonspecific interaction further as it increases the net amount of charged groups (the acidic residues would be deprotonated at pH > 5) in addition to the possible effect of the conformational changes induced by it [48,196] but not in tail-less HMGB1.

It is known that the interaction of the acidic tail with the HMG Boxes is modulated by the oxidation state: both TLR-2 [49] and RAGE [196] interaction are preferential for oxidised HMGB1, with the tail acting as a negative regulator which can be displaced by e.g. ligands. The acidic tail binds to the HMG Boxes [50], in both Box A [160] and the C-terminal segment of Box B (I158, R162, K164 [196][196]. Oxidation of Box A might result in dissociation of the acidic tail with the Box A core, exposing more charged residues and increasing nonspecific electrostatic binding. This could potentially explain why oxidised forms of HMGB1 have more nonspecific column interaction than the reduced equivalents.

HMGB1-c002 (1-184) binding more strongly to the column than HMGB1-c031 (1-164) is however not explainable by these data, although the presence of the 165-184 peptide could be modifying the conformation of the protein. This peptide is highly positively charged (45% Lys) and therefore would be subject to the same electrostatic effects as the HMG Box core.

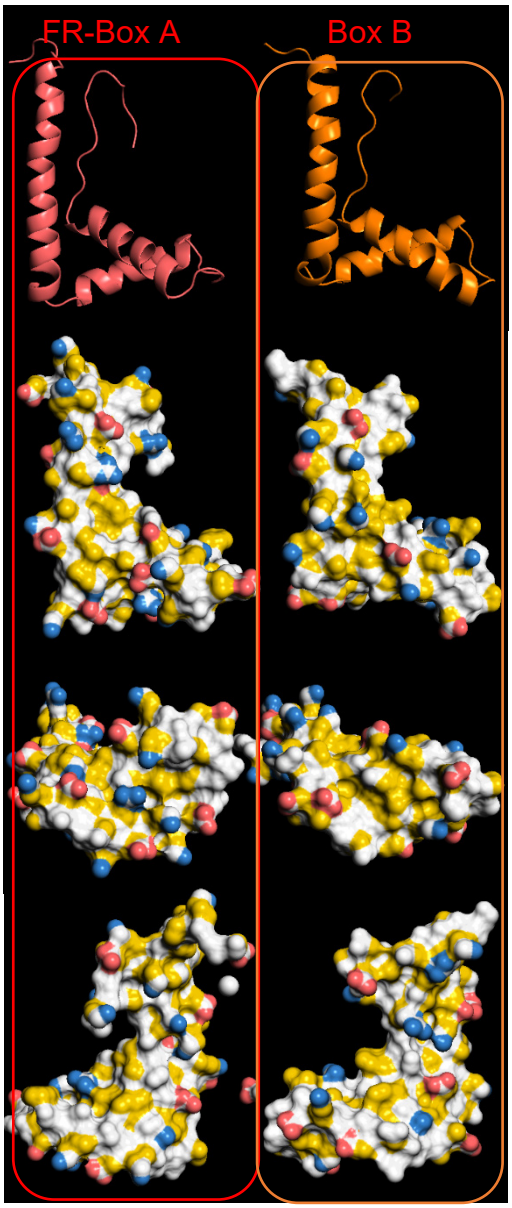


Figure 3.17: Surface representations of hydrophobicity and polarity in HMGB1

YRB colouring of PDB 2YRQ. **Yellow** = aromatic and aliphatic side chains. **Red** = carboxyl side chains. **Blue** = amine side chains. Polar groups and main chains are coloured white. HMGB1 FR Box A vistas are surrounded by a red square, Box B by an orange square. Box B has a much higher concentration of hydrophobic patches and lower concentration of charged residues, particularly in the bottom segment.

3.4.I.ii ESI-MS analysis of HMGB1 conformations.

The samples processed by SEC were then analysed in ESI-MS to calculate SASA by surface charge[25]. Protein surface charge acquisition is proportional to SASA [143,144], albeit unfolded proteins typically reach higher charge densities as the backbone charged groups are more accessible (e.g. amine in positive ion mode). Nevertheless, the maximum possible charge acquisition (I/Z, residues/charge) for a fully unfolded (linearised) protein is 1 charge per every 6 residues (I/Z = 6 [197]) due to electrostatic repulsion. The distribution of individual charge states for a given protein conformation and oligomeric state should follow a unimodal distribution, without discontinuities in the charge series. Multiple continuous distributions, each with its own mode, on the M/Z spectrum represent different charge envelopes, with the protein shifting between all these dynamically. These can happen either in a continuous charge series (different conformations of the same oligomerisation state) or discontinuous (distinct oligomerisation states).

I analysed each construct under native ESI-QTOF/MS in the two non-denaturing (native) conditions I used in size exclusion chromatography (50/200 mM ammonium acetate, pH 6.5) as well as under denaturing conditions (mobile phase in methanol with 0.1% formic acid) to analyse the conformation landscape of folded and unfolded HMGB1 constructs, and whether folding varied depending on buffer conditions or construct.

Several distinct charge envelopes of HMGB1 were observed in native solution ESI/MS at both near physiological (200 mM ammonium acetate, pH 6.8 - Figure 3.18) and infra physiological (50 mM - Figure 3.19) ionic strength for all constructs tested: two charge state distributions corresponding to monomers, and higher state oligomers, mainly dimers. Under denaturing conditions, three charge state distributions corresponding to monomers were observed, two of which overlapped with those observed under native conditions. Whilst HMGB1 has been reported to homo-oligomerise [51,198], a process enhanced by divalent cations and low ionic strength, native MS can sometimes generate artificial oligomerisation states [199–201].

[25] Salt adducts (sodium) could be observed despite size exclusion and manual desalting

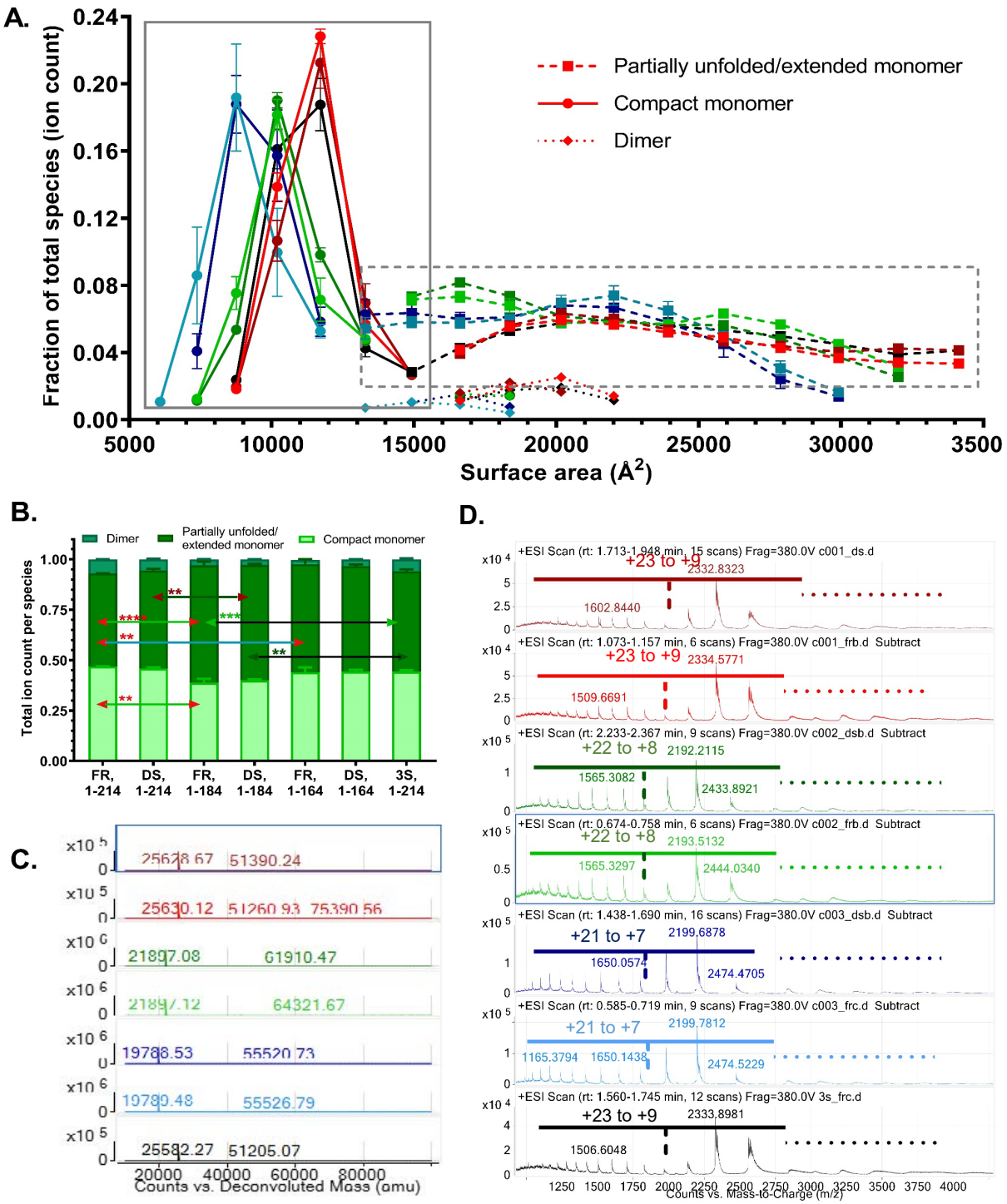


Figure 3.18: Native ESI/QTOF-MS of HMGB1 in 0.2 M ammonium acetate, pH 6.8

A. Average solvent accessible surface area vs. charge state abundance of HMGB1A-c001, c002, c031 (FR/DS) and HMGB1A-c006 (3S) charge state distribution. Data result of 3 independent injections, each with at least 10 counts of acquisition. Continuous lines: monomer, high M/Z monomer species. Dashed lines: low M/Z monomer species. Dotted lines: dimer and other multimers. **B.** Total ion counts for the folded monomer, unfolded monomer, and dimer states (2-way ANOVA; arrowhead and significance value colour designate parent construct, line colour indicates what construct is it compared to). **C.** Deconvolution of the native spectra in (A). Disulfide forms have a -1/-1.5 Da mass shift due to proton loss; DS-HMGB1 c002 and c003 are not fully oxidised as this shift is not apparent. Results representative of 3 runs; full spectra available in Suppl. data A1-3-C. **D.** Example M/Z spectra of HMGB1A-c001 (red), c002 (green), c031 (blue), in both FR and DS forms (darker hue) and HMGB1A-c006 (3S) (3 mg/mL, direct infusion into Q-TOF system). Horizontal, continuous lines: monomer. Horizontal, dotted lines: dimer and other multimers. Charge states belonging to the monomer are noted in the image. Vertical, dashed lines indicate what states were considered the threshold between compact monomer envelope (right) and the extended (left); +13 for c001/c006, +12 for c002 and +11 for c031.

The three monomer charge envelopes would be a compact monomer state (charge states +9 to +13 for FL-HMGB1) and an extended or partially unfolded monomer (+14 to +23; these values shift by -1 charge for c002 and -2 for c003). The charge state distribution for the larger monomer species in HMGB1A-c002 (1-184) and HMGB1A-c031 (1-164) seems to be bimodal and resolves into two overlapping distributions, unlike in FL-HMGB1 (c001 and c006) constructs which have a unimodal distribution in this species (Figure 3.18A). I could not separate this intermediate distribution from the extended monomer in this experiment, but this was feasible on the denaturing MS experiments where both these distributions are the major species. These likely represent fully unfolded and partially folded monomers. This partially extended/unfolded charge state is also more prevalent in HMGB1 1-184 (c002) than in full-length (1-214) or further C-terminally truncated (1-164) constructs (Figure 3.18-B). The bimodality of the extended monomer distribution is also more evident for HMGB1 1-184 than it is in 1-164. These would suggest that tail-less HMGB1 with the RAGE binding tail (1-184) might be more prone to partially unfolding than HMGB1 with the HMG Box domains only (1-164), matching the test expression results where HMGB1 9-184 did not express but HMGB1 9-164 did. Whilst for a construct with intact N-termini (HMGB1 1-184) this does not destabilise the protein enough to prevent expression or loss of biological activity, when combined with other factors it could be enough to render the protein unstable.

Redox state of the protein (FR, DS and 3S) does not lead to appreciable variations in charge envelope, reflecting the SEC results where disulfide and reduced HMGB1 constructs differed only in the propensity for column interaction (affected by surface charge profile) but not in retention volume (RV) (affected by SASA and radius of the protein)[26]. Regardless of the buffer ionic strength, the effect of construct boundaries and redox state in the proportion of monomer in the more compact conformation appears to be very similar, with the only difference being that at lower ionic strength, the compact monomer conformation appears to be more prevalent, even in the absence of the acidic tail.

[26] I excluded DS-HMGB1A c002 and c031 in the experiments with 50 mM ammonium acetate buffer due to limited equipment access and these results.

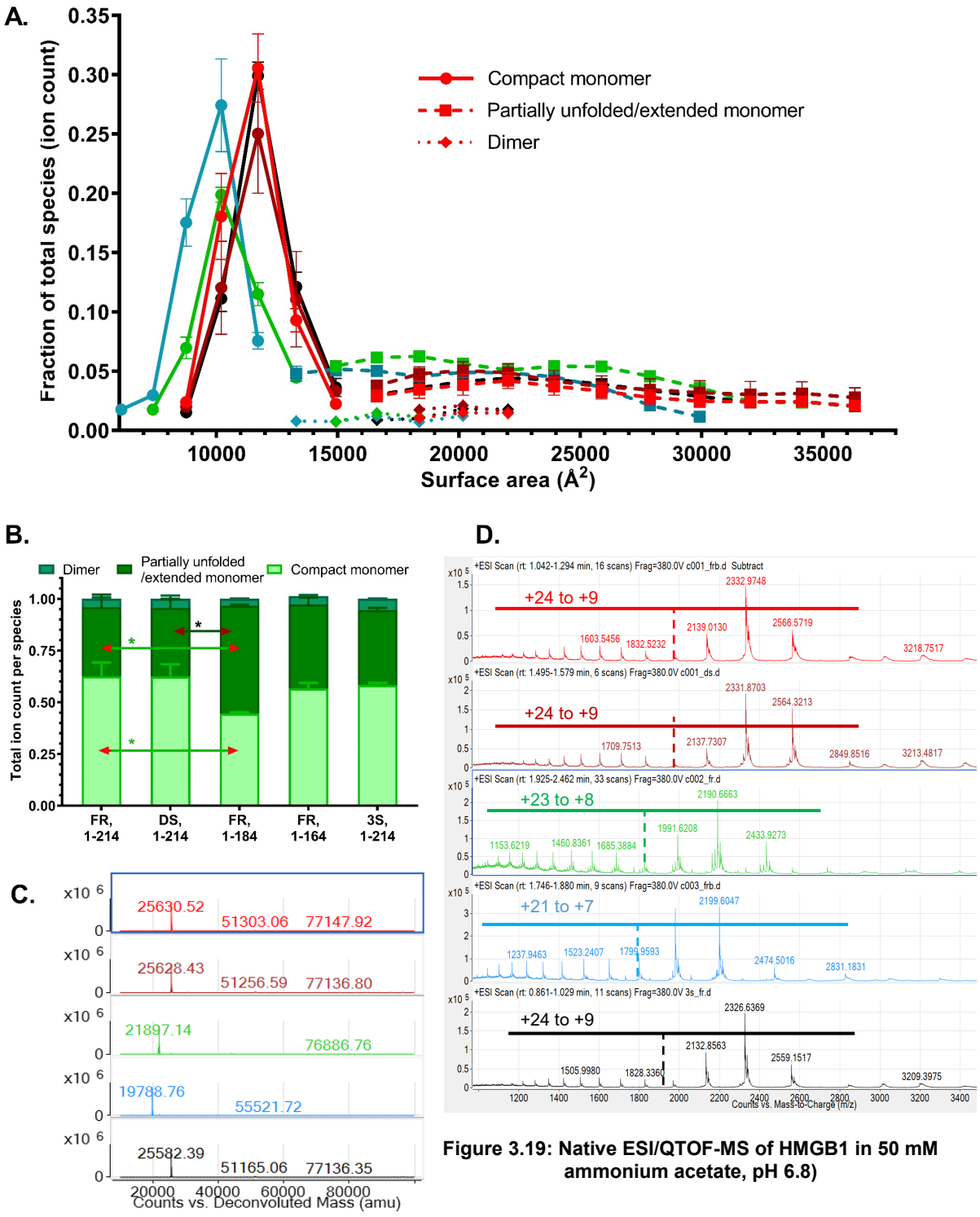


Figure 3.19: Native ESI/QTOF-MS of HMGB1 in 50 mM ammonium acetate, pH 6.8)

A. Average solvent accessible surface area vs. charge state proportion of HMGB1A-c001, c002, c003 (FR), HMGB1A-c001 (DS) and HMGB1A-c006 (3S). Data result of 3 independent injections, each with at least 10 counts of acquisition. Continuous lines: monomer, high M/Z monomer species. Dashed lines: low M/Z monomer species. Dotted lines: dimer and other multimers. **B** Total ion counts for the compact monomer, partially unfolded/extended monomer, and dimer states (2-way ANOVA; arrowhead and significance value colour indicate parent construct, line colour indicates what construct is it compared to). **C.** Ratios of folded monomer to unfolded monomer per construct. **C:** Deconvolution of the native spectra in (A). Results representative of 3 runs; full spectra available in Suppl. data A1-3-D. DS-HMGB1-c002 and c003 were not tested in this experiment due to restrictions in equipment and user time. **D.** Example M/Z spectra of HMGB1A-c001 (bright red-FR, dark red-DS), c002 (green), c031 (blue) and HMGB1A-c006 (3S) (3 mg/mL, direct infusion into Q-TOF system). Horizontal, continuous lines: monomer. Charge states belonging to the monomer are noted in the image. Vertical, dashed lines indicate what states were considered the threshold between compact monomer envelope (right) and the extended (left); +13 for c001/c006, +12 for c002 and +11 for c031.

Under denaturing conditions (0.1% formic acid in methanol mobile phase) the effect of redox state in protein conformation changes from what I observed in native MS (Figure 3.20A). I still observed in denaturing MS the two monomer conformations which appeared in native MS, but in different proportions of the total protein, and could clearly resolve the third conformation at very high (> +24 in FL-HMGB1) charge state which reaches maximum I/Z = +6 (linearised peptide). The fully folded monomer was now a minor species but could still be observed, and the proportion of fully unfolded peptide was inversely proportional to construct length, unlike in native MS.

All observed monomer surface charge distributions had comparable charge states profiles as observed in native MS. This further suggests that the surface charge profile of these conformations does not depend on the buffer, but the proportion of monomer in each state is. All fully unfolded peptides reached maximum I/Z = 5-6, indicating that the protein reaches complete unfolding in these experiments, despite the presence of both folded monomer species in denaturing ESI/MS.

However, the fully denatured distribution across different oxidation states differs in both the average and most abundant charge states. Under denaturing conditions, the disulfide states of HMGB1 are more prone to reach their fully unfolded states, but the charge state envelopes appear smaller, with -1 charge shift in the fully unfolded state. This would represent a lower degree of linearisation over 3S- or FR-HMGB1 which favour the extended/partially unfolded monomer under denaturing conditions (Figure 3.20B) but reach higher levels of unfolding. This would suggest that disulfide conformations of HMGB1 have a smaller energetic cost associated for the transition between the fully unfolded and extended monomer state compared to reduced protein. Disulfide bridges would limit the maximum conformational change of the molecule from folded to fully unfolded state (by not being labile in acidic conditions), and therefore the acid-denaturation associated unfolding ΔG would be smaller. 3S-HMGB1 is even more prone to retaining partial folding under denaturing conditions than FR-HMGB1, and therefore would appear to be more resilient to full denaturation. This could explain why the same construct boundaries express as FR-HMGB1 but not as 3S-HMGB1 (9-184).

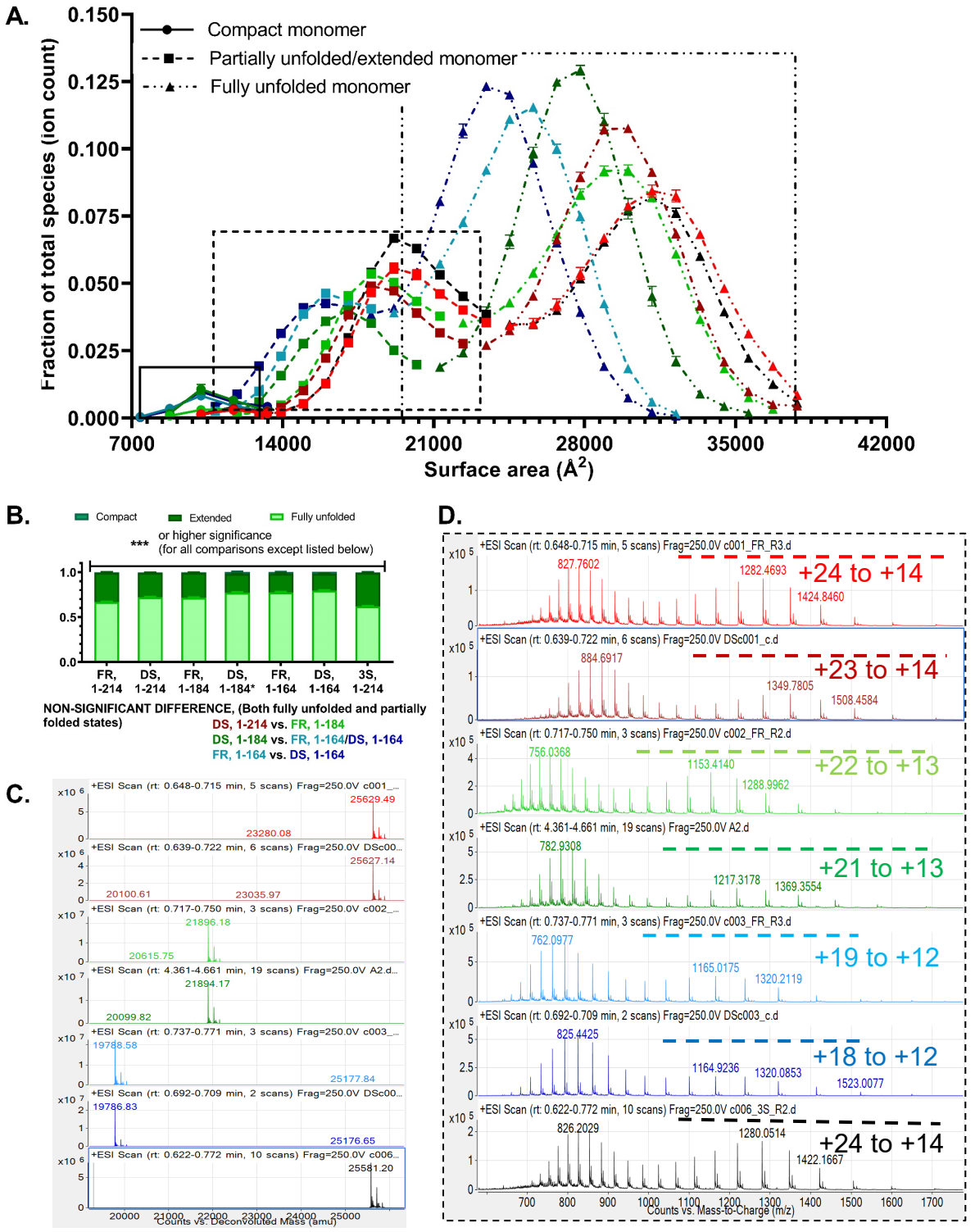


Figure 3.20: Denaturing LC-ESI-QTOF/MS of HMGB1 in 0.1% formic acid, pH 2.2

A. Average solvent accessible surface area vs. charge state abundance of HMGB1A-c001, c002, c031 (FR/DS) and HMGB1A-c006 (3S) charge state distribution. Data result of 3 independent injections, each with at least 10 counts of acquisition. **B.** Total ion counts for the fully unfolded monomer, partially unfolded monomer, and dimer states (2-way ANOVA; arrowhead and significance value colour indicate parent construct, line colour indicates what construct is it compared to). **C.** M/Z spectra of HMGB1A-c001 (red), c002 (green), c031 (blue), in both FR and DS forms (darker hue) and HMGB1A-c006 (3S) (3 mg/mL, direct infusion into Q-TOF system). Both curves correspond to different unfolded conformations of HMGB1; a low M/Z (high charge) component representing fully unfolded protein, and a high M/Z (lower charge) component corresponding to partially folded protein (e.g. core HMG Box folded and flexible regions extended) with identical M/Z values to the second distribution observed under native conditions (50 mM / 0.2 M ammonium acetate pH 6-8, dashed line). Highest charge states observed were +37 (c001/c006; 221 residues), +34 (c002; 191 residues), +30 (c031, 171 residues). Some charge states corresponding to the first distribution were also observed (not shown due to scale). Full spectra available in Suppl. data A1-3-E. **D.** Deconvolution of the denaturing spectra.

I calculated the mean Δfolding SASA (5048 Å^2) from the published NMR structure (PDB 2YRQ; folded *average* SASA = 14251 Å^2, unfolded *average SASA* of 19299 Å^2). I also obtained unfolded theoretical SASA calculations from the server ProtSA [145], to contrast with the theoretical SASA which can be derived from the charge envelopes. SASA calculated from native MS of HMGB1A-c001, c002, and c031 monomers ranges between 17000-15000 Å^2, much larger than expected for a protein of 25 kDa spherical protein [202]. Compared to e.g., ovalbumin, which eluted at similar volumes from the same column despite being twice as large in MW (44 kDa), the SASA is comparable between both proteins (15000 Å^2; Suppl. data A3-1-A, Figure 8.1), which would support these calculations[27]. The decrease in average monomer SASA across buffers would coincide with the proteins shifting their conformation equilibrium towards the more compact monomer conformation, accruing to the size reduction in ammonium acetate 50 mM compared to 200 mM.

All values in Å^2		**FR-HMGB1A, c001, FL** *26.3 kDa*	**DS-HMGB1A, c001, FL**	**FR-HMGB1A, c002, 1-184** *21.9 kDa*	**DS-HMGB1A, c002, 1-184**	**FR-HMGB1A, c031, 1-164** *19.8 kDa*	**DS-HMGB1A, c031, 1-164**	**3S-HMGB1A, c006, FL** *25.6 kDa*
Unfolded (ProtSA)		26337		22101		20691		26300
Unfolded (ESI/MS)		26945±41	26334±50	25487±92	24748±179	22488±64	21608±38	26237±100
Potential IDR length (including TEV scar)		9818	9818	5582	5582	4172	4172	9818
2YRQ (1-164) folded/unfolded		N/A	N/A	N/A	N/A	14251/19299	N/A	N/A
200 mM Amm. Ac	*Average monomer*	17839±17	18782±456	17502±219	17154±57	15248±304	15109±295	18394±157
	Dimer	19466±24	18403±51	17454±40	17531±18	14199±116	16501±125	19330±78
50 mM Amm. Ac	*Average monomer*	16267±880	17640±1500	17096±81	N/A	13961±315	N/A	16939±295
	Dimer	20560±161	20080±142	17243±88	N/A	16777±249	N/A	20783±218
Δ FR Monomer SASA between 0.2 M and 50 mM Amm. ac		1572±880		406±233		1286±438		1454±334

Table 3.2: SASA calculations from ESI/QTOF-MS experiments

Average SASA of the HMGB1 constructs tested, estimated from ESI/QTOF-MS in native and denaturing conditions (average of three injection events) from Figure 3.18/Figure 3.19/Figure 3.20. Differences of each average monomer SASA from either 50 mM or 0.2 M ammonium acetate (native conditions) with the average monomer SASA under denaturing conditions are also shown (approximate denaturation SASA) Unfolded SASA of non-structured regions of each construct: Box A N-terminus (1031 Å2), linker between HMG Boxes (2034 Å2), TEV scar AENLFYQ (1106 Å2). Regions absent in c031: RAGE binding peptide 165-184 (1410 Å2). Regions absent in both c002 and c031: Acidic tail 185-214 (4236 Å2). Amm. Ac= ammonium acetate pH 6.8. Errors in the "Δ Monomer SASA…" row were propagated according to $SE_{\Delta SASA} = \sqrt{\frac{SD1^2}{n1} + \frac{SD2^2}{n2}}$

[27] Anecdotally, when concentrating the protein prior to SEC, HMGB1A-c001, c002 and c006 did not penetrate a membrane with molecular weight cut-off (MWCO) > 30000 kDa; only HMGB1A-c031 permeated to some extent but even then 90% of the protein could be recovered.

The SASA of FL-HMGB1 1-214 and HMGB1 1-184 quite similar (n.s; Figure 3.18C), despite the difference of MW of 15% and the conformational states of HMGB1 1-184 having lower M/Z values than those of FL-HMGB1 (smaller SASA). Retention volumes in SEC were also much more similar between these truncations despite the MW difference, and in 200 mM ammonium acetate, HMGB1 1-184 eluted before the full-length protein. Furthermore, when the ionic strength is decreased in native MS, this construct has the smallest decrease in SASA amongst all three reduced constructs and 3S when ionic strength is decreased (roughly a third of the decrease observed in either FR 1-214, 1-164 or 3S 1-214), and less propensity for adopting the compact monomer conformation under low ionic strength, and overall is more propense to adopting the extended monomer conformation under native conditions than other HMGB1 constructs.

This suggests that removal of the carboxylic acid tail shifts the solution equilibrium between monomer conformations of HMGB1 towards a higher occupancy of these “extended” conformations, consistent with its role in protein stability and folding [48,50,196]: a more open protein conformation would have a lower free energy for its transition into unfolded states. Further truncation of the 165-184 peptide partially reverses this shift in the equilibrium back towards the more compact monomer conformations.

However, both the extended and compact monomer charge envelopes observed in native MS for HMGB1 constructs are likely “native” and exist in an equilibrium, rather than the extended conformation representing partially denatured protein. HMGB1 has several intrinsically disordered regions: 185-214 acidic tail, 165-184 RAGE binding tail, 79-88 linker region, N-terminal Gly1-Pro8 region, and a TEV scar in the case of my recombinant proteins. These can exist in an unfolded conformation or coil around the molecule in the native folded conformation of HMGB1 [46,48,50], resulting in theoretically great conformational flexibility, and which could allow for the presence of all these different packing states of HMGB1 monomers natively in solution.

In the performed SEC experiments in this section, whether analytical or preparative (≈ up to 2-3 mM injection concentration, see Suppl. data A3-1-C) the peaks were monodisperse with no void (aggregate) peak; SDS-PAGE also shows a monodisperse concentration gradient across the elution, which would imply that the protein interchanges between both monomer conformations but is otherwise still in solution, rather than denatured.

Furthermore, the average SASA of all observed monomer charge states agrees with both the published HMGB1 structure and my SEC data suggesting retention times similar to ovalbumin, which has equivalent SASA. Furthermore, all of these construct boundaries (1-214, 1-184 and 1-164) are biologically active at previously described doses with commercial FR-HMGB1 [25], and express reliably in *E. coli*. A published NMR structure contains several viable HMGB1 conformers (Figure 3.21) which highlights this conformational flexibility, with very similar structures for the HMG Box domains, but drastic changes in the relative position of Box A with respect to Box B across conformers.

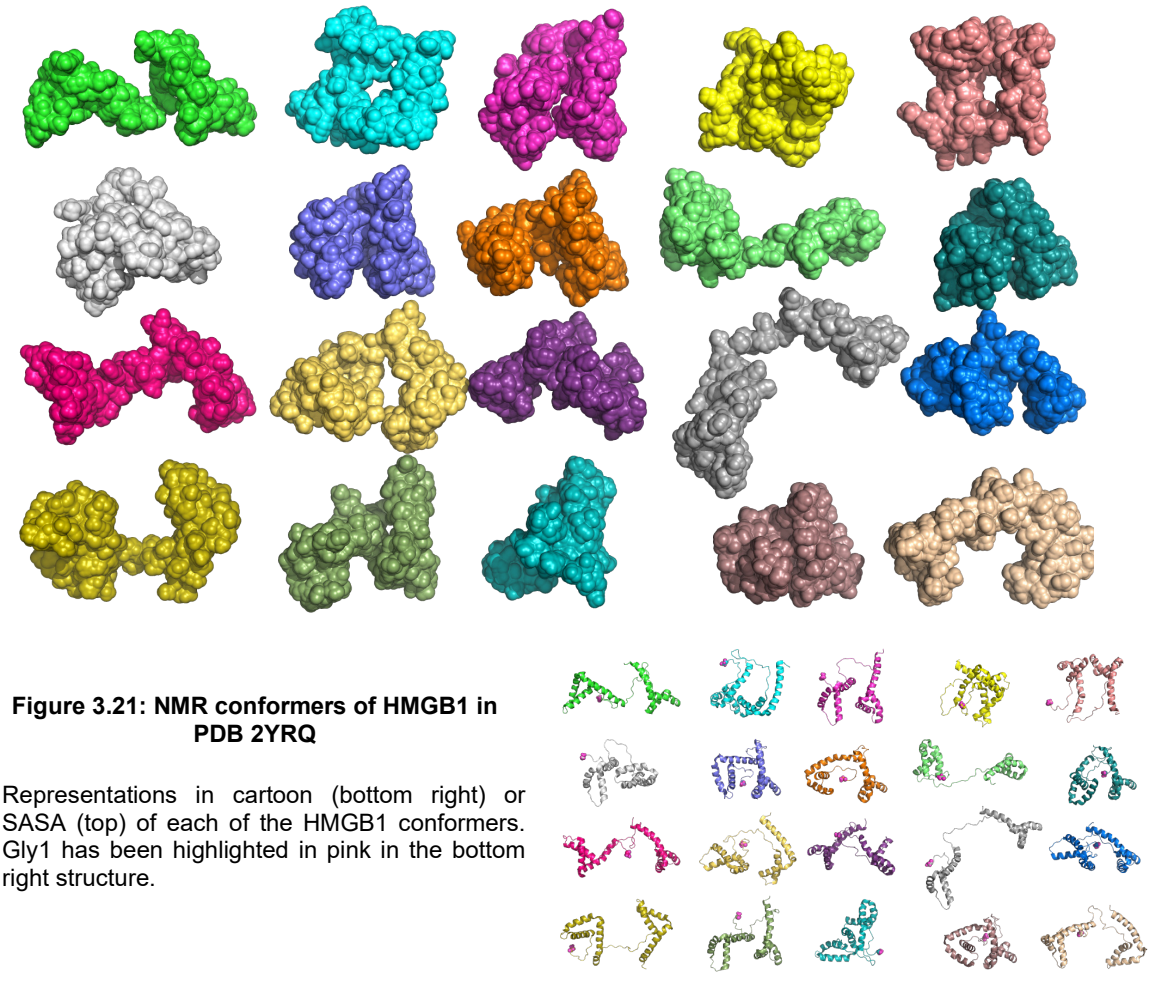


Figure 3.21: NMR conformers of HMGB1 in PDB 2YRQ

Representations in cartoon (bottom right) or SASA (top) of each of the HMGB1 conformers. Gly1 has been highlighted in pink in the bottom right structure.

However, a recent study [203] modelled the SASA of FL FR HMGB1 as ≈12000 Å^2, and FL DS HMGB1 at 11500, significantly lower than the average SASA for all conformational states pooled together. Calculations in this study are product of molecular dynamics (MD) assuming water as a solvent and the area of the tail was extrapolated. The SASA in the published NMR structure for HMGB1 1-164 (PDB 2YRQ) matches my average SASA calculations for these boundaries. My data showed that at higher ionic strength, HMGB1 is more propense to adopt conformations with higher SASA. Should the sample exist in a salt-free environment such as water, it would be reasonable to expect the conformation landscape of HMGB1 to be highly skewed towards the compact monomer conformation, reconciliating my results with both pieces of published information. A compact FL-HMGB1 monomer (average charge state approximately +11; approximate SASA 10872 Å^2) would concur with these published MD calculations of the HMGB1 SASA in water.

The sum of these native MS experiments could explain why HMGB1 c003 (FR, 8-184) failed to express, but c004 (acid tail present, 8-214), c005 (FR, 8-164) or c008 (3S, 8-184) expressed without significant aggregation or degradation. The absence of the acidic tail destabilises the protein [48], due to a decrease in the net helical content (folded protein) [196], indicative of partial unfolding. This effect appears to be more appreciable if the 165-184 peptide is present, with the protein monomer conformation being skewed towards more open monomers in HMGB1 1-184 than in HMGB1 1-164. If other destabilising factors are added (e.g., the interaction between Pro8, Tyr 70 and Met 94 is lost), the protein becomes further unstable. Fully reduced forms seemingly have a higher energy cost associated to the transition between partially folded, and fully unfolded states, and lower between fully unfolder and partially folded, compared to 3S-HMGB1. [28]

[28] Whilst the QTOF experiments do not involve SEC chromatography, extensive washes on the ion source with acetonitrile and running buffer were needed to fully baseline the signal due to nonspecific adsorption.

3.4.I.iii Melting behaviour of different HMGB1 redox states

To validate the conclusions derived from native and denaturing ESI-QTOF/MS, I analysed HMGB1A-c001 (FR and DS), HMGB1A-c002/c031 (FR), and HMGB1A-c006 (3S) in differential scanning fluorimetry (DSF), to assess protein stability (indirectly, via the melting behaviour) in multiple conditions. HMGB1 has a hydrophobic pocket buried within each HMG box domain, and therefore should be suitable for DSF (Figure 3.22).

However, the presence of several exposed hydrophobic residues resulted in high levels of initial fluorescence in most cases (10-20K RFU), and I could not analyse HMGB1-c002 or c031 as these hardly followed a one-step transition (Suppl. data A3-1-D). I however able to analyse FR, DS and 3S-full-length HMGB1, particularly regarding ionic strength and pH, which would be the most differential as the physicochemical properties of cysteine and serine are quite different – let alone the presence of a disulfide bridge. Calculated Tm_{50} values from Boltzmann fits are collected in Table 3.3, and analysis of the fits, plus the raw data curves, are shown in Figure 3.23.

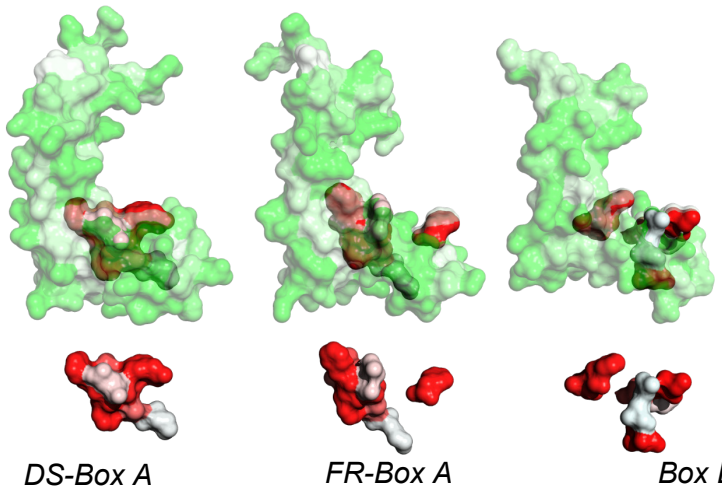


Figure 3.22: Hydrophobicity colouring of HMGB1 1-164 (PDB 2YRQ)
Eisenberg hydrophobicity colouring (white = most hydrophobic, green = most hydrophilic) of DS-Box A (PDB 2RTU), FR-Box A, and Box B (PDB 2YRQ). The SASA of the glycyrrhizin binding site (buried hydrophobic core) residues has been separately coloured with hydrophobic residues being coloured in red and hydrophilic in white; note how most of these are buried within each Box domain and not solvent-accessible.

	FR-HMGB1A-c001, 1-214	DS-HMGB1A-c001, 1-214	3S-HMGB1A-c006, 1-214
Water	49.6±0.6	51.2±0.9	46.1±0.3
150 mM NaCl	48.3±0.4	48.3±1	43.7±1.5
50mM HEPES pH 7.5, 50 mM NaCl	49.5±0.4	51.4±0.3	46.1±0.6
50mM HEPES pH 7.5, 150 mM NaCl	48.1±0.7	49.3±1	45.4±2
50mM HEPES pH 7.5, 250 mM NaCl	47.7±0.5	47.4±0.7	42.9±0.7
50mM HEPES pH 7.5, 500 mM NaCl	47.6±0.3	47.2±0.7	44.2±1.3
50mM HEPES pH 7.5, 750 mM NaCl	47.8±0.8	47.9±1	42.6±1.4
50mM HEPES pH 7.5, 1000 mM NaCl	49±0.5	48.9±0.3	44.2±0.3
50mM HEPES, pH 7.0	46.8±0.7	47.1±0.5	42.6±0.7
50mM HEPES, pH 7.5	48.3±0.6	48.6±1.2	45.7±0.7
50mM HEPES, pH 8.0	51.5±1.4	53.5±0.2	45.7±0.7
50mM Tris, pH 7.0	45.2±0.6	46±1.4	41.4±1.4
50mM Tris, pH 7.5	47.4±0.8	47.8±1.2	45.5±1.4
50mM Tris, pH 8.0	50±0.3	48.6±0.6	45±1.1
50mM Ammonium acetate, pH 6.8	45.1±1	43.7±1	41±0.3
100mM Ammonium acetate, pH 6.8	47.1±0.4	46.9±0.3	42.6±0.4
200mM Ammonium acetate, pH 6.8	45.6±0.6	45.7±0.9	40.4±0.4
0.33X PBS, pH 7.4	47.9±0.3	47.2±0.4	42.9±0.2
1X PBS, pH 7.4	50.8±0.1	52.1±0.5	44.9±0.5
2X PBS, pH 7.4	41±5.5	49±1.8	39.9±2
0.1 M SPG, pH 4.5	35.6±0.9	36.6±0	34.9±0.3
0.1 M SPG, pH 5.0	35.4±0.1	34.2±0	34.9±0.5
0.1 M SPG, pH 5.5	37±1.3	37.4±0.9	36.3±1.5
0.1 M SPG, pH 6.0	41.4±3.2	42±4.1	37.4±1.7
0.1 M SPG, pH 6.5	45.6±2.2	43.6±1	35.6±2.1
0.1 M SPG, pH 7.0	50±1	51.7±1.4	44.3±1.4
0.1 M SPG, pH 7.5	48.5±2.2	51.7±3.3	43.8±4.2
0.1 M SPG, pH 8.0	45.4±5.3	52±2.7	42.7±4.4
0.1 M SPG, pH 8.5	49.8±1.6	49.9±2.4	45±2.1
0.1 M SPG, pH 9.0	50±0.2	50.9±0.8	46.1±0.7
0.1 M SPG, pH 10.0	53.2±0.7	53±0.7	41.9±4.8

Table 3.3: Summary Tm_{50} of HMGB1 constructs with different and redox states

Calculated Tm_{50} values (in ºC) for HMGB1A-c001 (bright red: FR, dark red: DS), 3S-HMGB1A-c006 (black). Colour gradients for the Tm_{50} values indicate highest (green) and lowest (red) values within the global dataset; individual condition identifiers have been coloured as per Table 3.4 and each set of screening conditions is separated by a double-space line. Data result of three melting experiments, except for conditions with an SEM of 0 (DS-HMGB1, SPG pH 4.5 and 5.0) where several melting curves melts reached fluorescence signal saturation and were discarded since the computed Tm_{50} would be artificially low. A sample of curves for these conditions is shown in Figure 3.23.

Whilst HMGB1 seems to be moderately more stable at pH 8.0 than it is at pH 7.5 in HEPES, pH 8.0 would be too much of a departure from physiological range for my biochemical studies, whereas protein used in biological studies would be diluted in PBS (pH 7.4), and I did not vary the pH of my purification buffers. FR-HMGB1A-c002 and c031 melting profiles were too wide to estimate realistic Tm_{50} values and have been excluded; Curves for these constructs are shown in Suppl. data A3-1-D for evaluation purposes only.

Interestingly, 3S-HMGB1A-c006 is overall slightly less thermostable than both DS and FR-HMGB1A-c001. This is in apparent contradiction with both my test expression results (Figure 3.5), where the same protein boundaries of HMGB1 express as 3S but not as FR, and my MS data where 3S-HMGB1 favours more compact conformations. However, if this reflects a lower energetic cost for the folding/unfolding process, it could be argued that 3S HMGB1 can tolerate the presence of more destabilising truncations than FR-HMGB1 since shifting between states requires less free energy change and therefore interchange between these can occur more easily. 3S-HMGB1A-c006 is also more sensitive to high salt concentration than FR- HMGB1A-c001; increasing concentrations of salt destabilise 3S-HMGB1-c006 slightly (ΔTm_{50} between 50 and 1000 mM ≈ -2ºC, 0.5ºC for FR-HMGB1A-c001), a trait apparently shared with DS-HMGB1-c001 (ΔTm_{50} ≈ -3ºC). However, DS-HMGB1A-c001 has apparent "destabilisation" due to it being much more stable than FR-HMGB1A-c001 under low salt conditions, and as stable as it under higher salt concentrations.

All 3 redox isoforms have similar ionic strength and pH tolerances. NaCl concentration in the range of 150-500 mM is suboptimal when compared to values outside this range. This was reproducible in ammonium acetate, HEPES+NaCl and PBS buffer, which rules out buffer-specific effect. This decreased thermostability compared to FR-HMGB1 might be due to the structure of 3S-HMGB1A-c006 being more influenced by the hydrogen bond between serines, whereas the contribution of a hydrogen bond between Cys 22 and Cys 44 in FR-HMGB1 would be negligible due to the low differences in electronegativity between sulphur and hydrogen.

All FL-HMGB1 forms are more unstable at pH 5.5 or less, and pH increases beyond this point lead to increasing Tm_{50}. The pI of FL-HMGB1 is 5.6 due to its high content of acidic residues in the tail offsetting the lysine- and arginine-rich HMG Boxes. When >1 pH unit below the pI, the side chain groups in the acidic tail (aspartic/glutamic acid) begin to protonate, dissociating it from the HMG box core and destabilising the protein, offsetting the increased solubility from departing the pI, because all the amine groups it interacts with would be protonated as well. 3S-HMGB1A-c006 is less stabilised by an increase in pH than wild-type HMGB1 as well; whilst its Tm_{50} at pH values under its pI is equivalent to that of FR and DS-HMGB1A-c001, it is markedly less thermostable above this value in Tris, HEPES and SPG buffers, with a ΔTm_{50} in the pH 4-9 range of ≈11ºC, compared with ΔTm_{50} for FR-HMGB1 ≈14ºC.

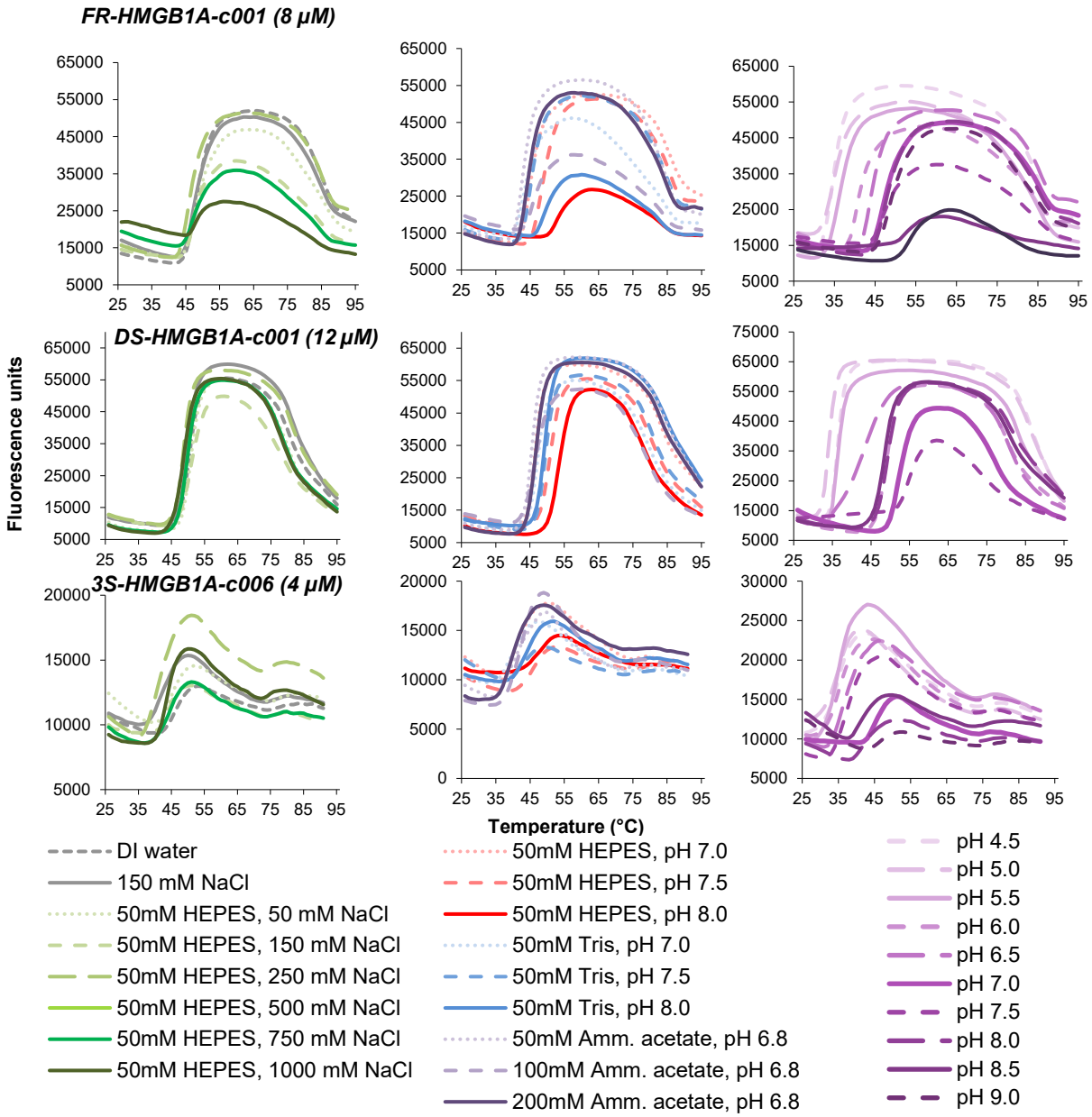


Figure 3.23: DSF melting curve profiles and Tm_{50} analysis of HMGB1A constructs

Thermal denaturation profiles for HMGB1A-c001, c002 and c031 (FR) in my custom screen; conditions have been coloured in the same code as shown in Table 2.3. Each row of graphs corresponds to the listed construct; all experiments performed with 8 µM protein (desalted in 10 mM HEPES pH 7.5) and 1X SyPro Orange. I did not dare to fit equations to the plots in (A) and (B) as I could not deduce a suitable mathematical model. Profiles representative of three replicates. The concentration of protein in the final experiment is shown. In green: Ionic strength conditions. In red and blue: HEPES, Tris and ammonium acetate screen. In purple: SPG pH screen.

29

[29] Unfortunately, the melting profiles in pH 2.2 0.1% formic acid could not be analysed as these conditions had little to no transition and initial relative fluorescence units (RFU) comparable to other runs. I would presume that the fluorescence of SyPro is quenched at this pH.

3.4.I.iv Summary of characterisation experiments

In both ESI-QTOF/MS, SEC, and DSF, I observed that a construct encompassing the HMG Box core and the RAGE binding peptide (1-184) had a smoother transition between the fully folded and the partially folded/extended monomer states than both full-length HMGB1 and a construct containing the HMG Box core alone. In ESI-QTOF/MS, HMGB1 1-184 had a higher proportion of this conformational state than FL constructs, and its SASA was comparable to that of FL-HMGB1 despite having 20% less mW. Recent studies have also shown that tail-less HMGB1 constructs (such as 1-184 or 1-164) have a lower net helical content than full-length protein [196], which would match my ESI/MS results as these constructs are more prone to partial unfolding.

HMGB1 1-184 is likely prone to adopt these partially unfolded conformations during translation. This increases the chances of the protein being recognised by the protein quality control mechanisms as misfolded. If other destabilising factors such as truncation of the molecular lock in Box A (9-78) are present, this results in the lack of expression I saw for HMGB1 9-184 from protein degradation.

Substitution of cysteine to serine in 3S-HMGB1 in HMGB1 9-184 results in expression due to the different effects of the -SH and the -OH hydrogen bonds. FL-3S-HMGB1 is more prone than FL-FR-HMGB1 to retain partial folding in denaturing conditions, yet its melting transition temperatures in non-denaturing conditions are lower (less stable). This suggests that this mutant is more stable in the partially unfolded state (lower energy cost for full unfolding, but higher probability to retain folded state) than FR-HMGB1; both adopt identical conformations under non-denaturing conditions. The 9-184 constructs are more prone to partial unfolding on their own; in this case, 3S-HMGB1 9-184 will be more stable and retain this partial unfolding than FR-HMGB1 9-184, which is likely to follow through to a fully unfolded state and degradation.

3.4.II Conclusions – HMGB1

Whilst expression of HMGB1 has been described in several publications prior to my work, these all had shortcomings which I have overcome to scale these for high throughput. Different HMGB1 constructs possess varying degrees of bacterial toxicity and thus growth rate of expressing cells differs, and the described purification strategies required individual optimisation for each construct which was not viable. I now had access to all the HMGB1 constructs I would require and could purify them in milligram quantities reproducibly.

In the process of optimising these, I identified additional factors affecting HMGB1 stability, observable after removing the contribution of the carboxylic acid tail [50].

- Within each HMG Box there is a molecular lock between a proline (Pro 8/94) , tyrosine (Tyr 71/Tyr 154) and an apolar (Met 74/Ile 158) residues which helps stabilize its structure, even though in Box B some contacts in this lock are absent and overall it adopts a much more rigid structure in published NMR structures (little to no variation across conformers).
    - Removal of Pro 8 is a destabilizing factor which results in no expression when combined with other modifications such as removal of the acidic tail but not of the RAGE binding peptide (165-184). This can be observed by an increase in the proportion of monomers in the extended conformation in ESI/MS for 1-184 constructs, as well as higher retention times in SEC than FL-HMGB1 despite lower MW.
    - I did not attempt mutation of either the Tyrosine or the apolar residues; this would confirm or deny this molecular lock hypothesis.
- The redox state of the protein affects both its stability and its surface charge conformation.
    - All three redox forms have similar thermostability trends depending on buffer conditions and adopt similar conformation distribution in solution, but 3S-HMGB1 has lower energy cost for unfolding (thermostability decreased) and a higher propensity to remain in the extended monomer conformation under denaturing conditions rather than fully unfolded.

- Both 3S-HMGB1 and FR-HMGB1 have equivalent regenerative activity [25]. Combined with the similar thermostability trends, this allowed me to use 3S-HMGB1 as a substitute for FR-HMGB1 in assays where protein oxidation could interfere, such as NMR in the presence of CXCL12, despite the apparent lower thermostability. However, the slight differences between these could also imply functional differences. I therefore needed to validate that the binding properties for both constructs regarding CXCL12 were equivalent. In Chapter 5, I did observe differences in their kinetics for RAGE binding.

These factors do not fully explain the requirements for HMGB1/CXCL12-mediated tissue repair. Despite expressing, a construct without the 165-184 peptide (HMGB1A-c005, 9-164) does not have pro-regenerative activity. Conversely, a construct without the acidic tail but with the 165-184 peptide (HMGB1A-c002) retains biological activity. This could be explained if the 1-8 peptide contains residues critical for CXCL12 interaction or ternary complex signalling; this hypothesis will be explored in Chapter 4.

3.5 Expression, and purification of CXCL12.

3.5.I First iteration of CXCL12 constructs

In eumetazoans, CXCL12 is secreted through the *Sec* pathway. The *Sec* system is fairly conserved [204], and out of curiosity I designed constructs with and without the peptide present, for wild-type and the Val24-Ile mutant, to assess if this native *Sec* peptide could induce secretion of recombinant CXCL12 in *E. coli* to the periplasm, in addition to cytoplasmic expression. The disulfide bonds of CXCL12 are required for CXCR4 signalling [126,205–207], but it is not known if they are required for HMGB1 binding or folding. I also designed a truncation of the C-terminal alpha helix of CXCL12; however, as the alpha helix of CXCL12 is involved in dimerisation [114], and a later study [131] suggested a role in HMGB1 interaction, that truncation was abandoned (Figure 3.24). This first set of CXCL12 constructs was initially cloned into pNIC-CTHF (cytoplasmic expression unless the secretion peptide was active). Disappointingly, all these CXCL12 constructs expressed at minuscule levels (Figure 3.24). Constructs without secretion peptide were not soluble, despite abundant protein content in the total fraction (unclarified lysate), likely forming inclusion bodies. Constructs containing the secretion sequence were absent even in the total fractions, failing to express.

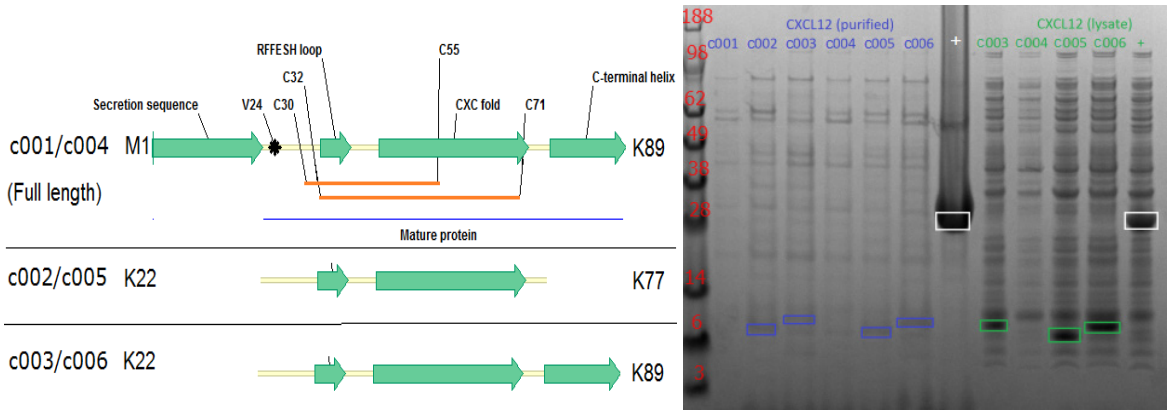


Figure 3.24: Initial CXCL12 constructs and test expression

A: Schematic representation of initial construct boundaries. Two construct IDs are supplied: IDs 1-3 represent constructs derived from CXCL12A-s001 (wt) whereas IDs 4-6 represent constructs derived from CXCL12A-s001 (V24I). Residue numbering corresponds to the polypeptide with the N-terminal methionine and signal peptide. **B:** Test expression of CXCL12 constructs in BL21(-DE3)-(R3)-pRARE2. Protein was purified using IMAC, then analysed by SDS-PAGE. Bands analysed by MS/MS are highlighted. Blue: supernatant fraction. Green: total fractions (c001 and c002 totals were lost). +: positive control (NANPA-c000).

3.5.II CXCL12 production in cytoplasmic disulfide-competent strains is inefficient

Refolding CXCL12 from *E. coli* is a costly and time-consuming strategy, which requires either very lengthy and complex dialysis setups or equipment to which I had no access (preparative-grade RP-HPLC) [112,208], as otherwise a mixed population of CXCL12 with varying disulfide bonds is produced (e.g. C32-C55/C30-C71), I instead set to optimise soluble expression of CXCL12 in *E. coli* through other means[30].

There are several strains of *E. coli* capable of producing proteins in the cytoplasm in their oxidised state. BL21-DE3- pRARE2 derivatives such as Rosetta-Gami2, combine double deletion of glutathione-oxide and thioredoxin B reductases (*Δgor/ΔtrxB).* Shuffle cells enhance these with cytoplasmic overexpression of disulfide isomerase C (DsbC) to assist in the folding of cysteine-containing proteins. However, yields for my CXCL12 constructs without *Sec* signals in Rosetta-Gami2 cells were as low as with regular BL21-(DE3)-R3-pRARE2. Due to the metabolic stress from these genomic deletions [209,210], these cells also had poor viability from glycerol stocks, which led me to not attempt testing similar systems such as Shuffle cells.

Expression in BL21(DE3)-pLysS cells, augmented with the CyDisCo pMJS226 system, whilst more productive than Rosetta-Gami2 (Figure 3.25B) still had very low yield of soluble protein. Instead of *Δgor/ΔtrxB* deletion, these overexpress human protein-disulfide isomerase (PDI) and yeast mitochondrial sulfhydryl oxidase (*erv1p)* to assist in the folding of disulfide-rich proteins [210–212]. Expression in CyDisCo cells provided yields sufficient for some of my experiments (e.g., peptide arrays, where ≈10 µmol/100 µg CXCL12 are needed), but insufficient for kinetics or *in vivo* assays. Furthermore, cytoplasmic lysis of *E. coli* results in co-purification of contaminants with similar pI and MW to CXCL12. This led me to conclude that cytoplasmic expression of CXCL12 is not efficient. Instead, I focussed on periplasmic expression of this protein.

[30] To my understanding, there is no instance of CXCL12 being expressed in insect or mammalian cells in the literature, and I was restricted to *E. coli* due to NMR labelling requirements; yeast culture would have introduced the risk of contamination for other researchers in the lab. In-house experiments performed by collaborators, where CXCL12 was either secreted or intracellularly expressed in *Sf9* insect cells or EXPI-293 mammalian cells failed to yield even insoluble protein, likely owing to the extreme regulation CXCL12 is subjected to in eumetazoans [13,108,286,287].

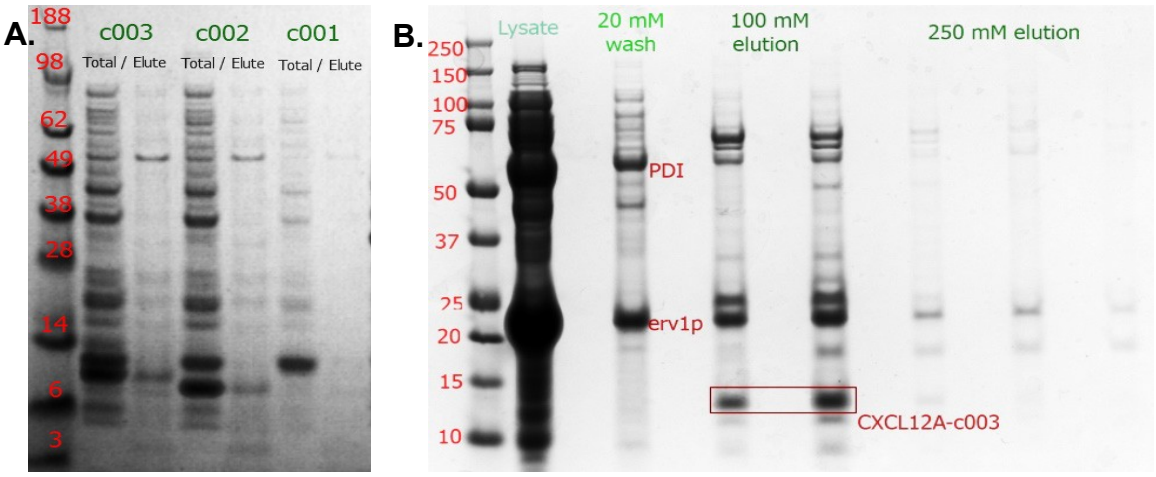


Figure 3.25: Expression of CXCL12 in disulfide-competent strains

A. SDS-PAGE of 1 mL test expression of CXCL12 constructs in Rosetta-Gami2, purified using IMAC **B**. 1L expression of CXCL12A-c003 in BL21-DE3-pLysS CyDisCo pMJS226 cells; concentrations of imidazole are indicated in the image. Bands corresponding to PDI, erv1p and CXCL12 are highlighted.

3.5.III Construct design for periplasmic export

Use of the native *Sec* peptide present in CXCL12 led to no expression, despite the conservation of this system across organisms, and therefore I attempted secretion to the periplasm via native *E. coli Sec* signal peptides for both the unfolded translocon (*SPI)* and membrane translocon (*SPII)* pathways. *E. coli* has also a folded protein translocon system (TAT, twin-arginine transporter) [213]. I tested expression of CXCL12 through all three export pathways in *E. coli*, as vectors were available in-house: pNIC-OmpA-NHT-cTStII for *Sec-SPI (*unfolded translocation), pNIC-Suf1-NHT-cTStII for the TAT pathway (folded translocation), and pNIC-TolB-NHT-cTStII for *Sec-SPII* (likely to fail; membrane embedding translocation/atypical export). For these experiments, I only amplified CXCL12A-c003 (CXCL12A-c022 to 24) and CXCL12A-c006 (CXCL12A-c026 to 28) as the native secretion peptide was unneeded and the C-terminal alpha helix truncation had been discarded by this point.

However, expression of CXCL12 in these vectors would require addition of residues to its N-terminus as they have all N-terminal tags. Addition of single residues to the N-terminus of CXCL12 is known to dramatically alter its biological activity [103], as the flexible N-terminal loop is critical in receptor interaction and activation [117,207]. Expression of CXCL12 with N-terminal tags could result in dysfunctional CXCL12, and I had to adapt the system to prevent this. Ubiquitin-like specific protease, also known as sentrin-specific protease (Ulp-1/SenP1*)* is unique since it does not recognise a specific peptide motif. Rather, it will target the first peptide bond C-terminal to a

structural motif (14 kDa) known as Small Ubiquitin-like Modifier (SUMO) tag. This protease will leave no trace residues after cleaving, resulting in intact proteins at their N-terminus [214]. Therefore, Ulp1 sites would be ideally suited for CXCL12 production in N-terminal tagged vectors. Addition of SUMO to the N-terminus of CXCL12A constructs was readily done via PCR with overhang fragments at the C-terminus of the SUMO site matching the N-terminus of CXCL12 constructs, and vice-versa.[31]

Initial test expression of SUMO-CXCL12 yielded appreciable amounts of protein (Figure 3.26), but only when through the TAT pathway which was striking, given that CXCL12 is likely unfolded in the cytoplasm. I suspected that this was due to the presence of a SUMO tag, which is folded, and therefore would hamper *Sec* translocation, particularly since it comprises most of the recombinant protein (14 kDa for the SUMO tag, compared to 8 kDa for CXCL12). However, even with this method, a significant amount of the protein was degraded during expression (Figure 3.26): whilst the bands for CXCL12A-c027 were identified as SUMO and CXCL12, the full mass of the protein was not detectable. Lower MW bands were identified as SUMO alone; degradation was likely C-terminal to SUMO and CXCL12 is lost, since the tag is N-terminal to SUMO.

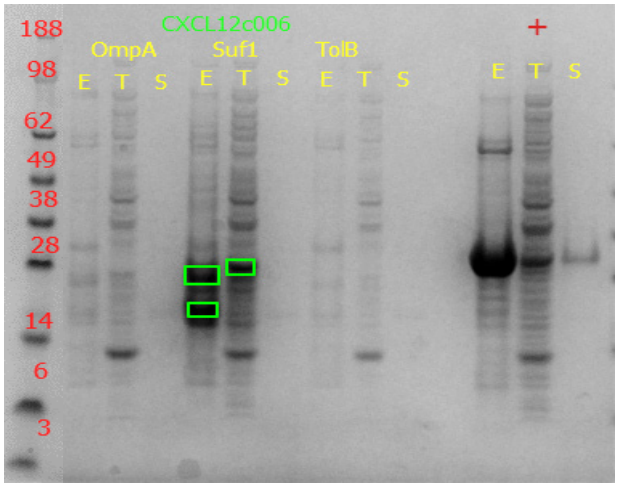


Figure 3.26: Periplasmic export of CXCL12A-c003 and c006 with N-terminal SUMO tag fusion

Test expression of CXCL12 constructs in BL21-DE3(R3)-pRARE2 and periplasmic expression vectors. OmpA: *Sec-SPI* pathway. Suf1: *TAT* pathway. *TolB: Sec-SPII pathway.* Protein was purified using IMAC, then analysed by SDS-PAGE. Bands corresponding to CXCL12 are highlighted. S: supernatant fraction. T: total fractions. E: elution from IMAC. +: positive control (NANPA-c000). Only constructs corresponding to CXCL12 V4I (derived from CXCL12-c006) are shown; cloning of the SUMO tag in CXCL12A-c003-derived constructs introduced a frameshift mutation and they had to be repeated in another experiment (data not shown). MS/MS data can be found in Suppl. data A2-3-I

[31] Whilst Ulp1 is a cysteine protease and would require reducing agents, a 1:10 protease OD_{280} /target OD_{280} ratio was sufficient to fully cleave SUMO-CXCL12 constructs in the absence of TCEP, preserving the oxidised state of purified CXCL12. This has been shown for other Cys-proteases such as TEV [217].

I performed experiments to optimise CXCL12 TAT-mediated periplasmic expression, showing that at 4h after induction the ratio of intact to degraded protein is highest (Figure 3.27A). Induction with 50 μM IPTG is slightly less productive than induction with 100 μM; this led me to believe that the *in vivo* degradation of CXCL12 occurs at a constant rate during expression, and that can be overcome by refining the expression methodology to increase raw protein yields. In parallel, I had also designed an alternative CXCL12 construct, based on the spectacular results achieved via N-terminal fusion of disulfide-rich snake venom proteins to DsbC [215–217]. The SGC library had a vector for expression via this system (pDsbC-HT-CBio), and my SUMO-CXCL12 construct could easily be inserted in this vector without affecting the final protein as it had its own N-terminal cleavage site and C-terminal stop codon would remove any further epitopes. Whilst it would export through the *Sec-SPI* pathway instead of the TAT pathway - required for DsbC expression - the combined mass of unfolded protein (26.3 kDa for DsbC + 8 kDa for CXCL12) was larger than the folded mass (14 kDa for SUMO) so I decided to attempt expression in this system.

These constructs had much higher expression when compared to those in pSuf1 (Figure 3.27A), and the main species in intact mass was the full target protein. Therefore, I abandoned the TAT-exported constructs in favour of the pDsbC fusion for my work, and purified protein from CyDisCo cells for applications which required histidine-tagged CXCL12 (peptide arrays), as the SUMO+DsbC tag would render the protein too large to be used as a ligand on these [158].

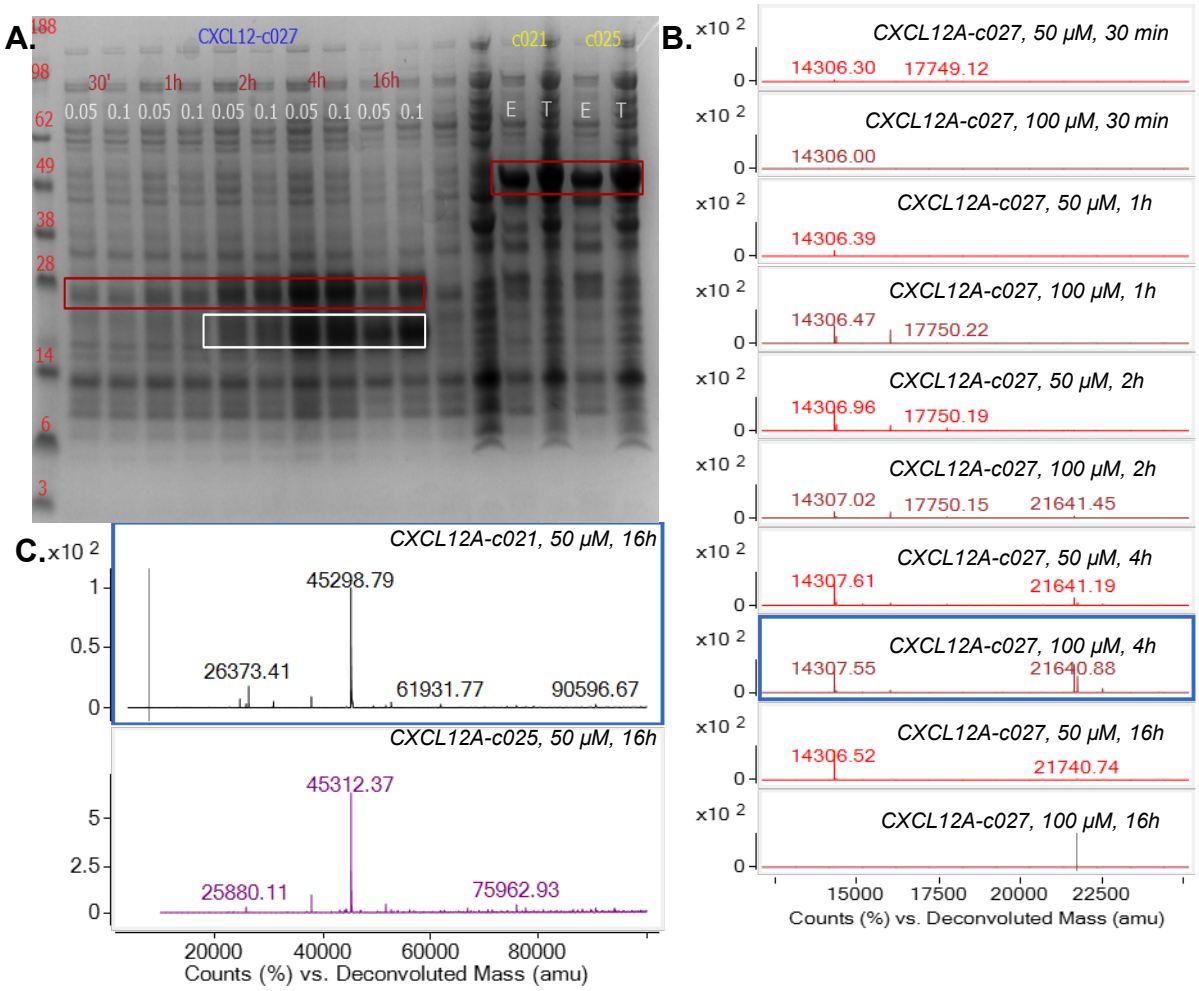


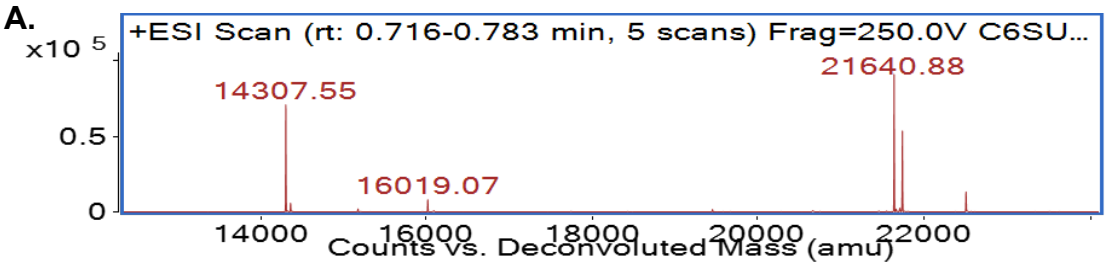
Figure 3.27: Time optimisation of CXCL12A-c027 (TAT export) expression and 24h test expression of CXCL12A-c021/25 (pDsbC fusions)

1 mL of BL21(DE3)-R3-pRARE2 cells were induced with either 50 or 100 µM IPTG and harvested at the indicated timepoints. In parallel, cells transformed with CXCL12A-c021 and c025 were induced and harvested after 16 hours (overnight). All were purified by IMAC and the elutions analysed by intact mass spectrometry. **A.** SDS-PAGE gel. **B.** LC-ESI-QTOF/MS deconvolution of time course. Dark red: 100 µM IPTG, bright red: 50 µM. MW of CXCL12A-c027: 21644.70 Da. Mass of the N-terminal SUMO fusion: **C.** LC-ESI-QTOF/MS deconvolution of pDsbC constructs. MW of CXCL12A-c021: 45303 Da. MW of CXCL12A-c025: 45316 Da. All masses assume reduction of all cysteines and loss of the signal peptide; a -4 Da adduct signals the formation of two disulfide bonds in CXCL12.

Degradation of SUMO-CXCL12 constructs in *E. coli* appears to be specific to certain sequences within CXCL12 (Figure 3.28). The cleaving sites seem to be identical for both pSuf1 and pDsbC constructs and the MW of the observed degradation products was consistent across multiple instances of protein expression. In the pSuf1 construct (CXCL12A-c027), a degraded fragment confirmed by MS/MS to be SUMO can be seen, with roughly 21 kDa apparent MW in SDS compared to 28 kDa for the full SUMO+CXCL12 constructs (actual MW 21644, oxidised 21641). A peptide in ESI-QTOF/MS with MW 14307 can be seen after induction by IPTG. This MW matches that of a SUMO tag with N-terminal Met-His6 and the first 6 residues of CXCL12 (KPISLS) C-terminal (4 Da error). The pDsbC construct produces 2 additional species detectable in ESI-QTOF/MS. One of them, MW 26373 Da, matches N-His6-DsbC cleaved 2 positions after TEV, despite no TEV being used at any point of this protocol. The second species with MW 39782.13 Da would match N-His6-pDsbC-SUMO-KPISLS. Again, I can detect the exact same cleaving site, which led me to assume the presence of either a protease recognising that site or that the SUMO-CXCL12 was somehow toxic to the cells. This was not unexpected as both constructs comprise a folded and unfolded peptide, which then is exported through either a folded protein-specific translocon (TAT system) or unfolded-specific (SEC system). This is consistent with the fact that the cells also cleave the protein ahead of the SUMO tag in the pDsbC construct. This behaviour can be seen with other toxic proteins – e.g., FL-HMGB1 (in particular, the 201-205 peptide) - where cells cleave the toxic fragments of proteins.

Modification of the N-terminus of CXCL12 to alleviate this degradation is not ideal given the relevance of the first 4 N-terminal residues in CXCR4 activation [116], where even single mutations in this region turn CXCL12 from agonist to a highly potent antagonist of CXCR4; therefore, I considered prudent to refrain from further modifications, given that regardless of degradation, sufficient quantities of intact protein could be obtained from this system. The next step was to test the pDsbC system in large scale.

A.



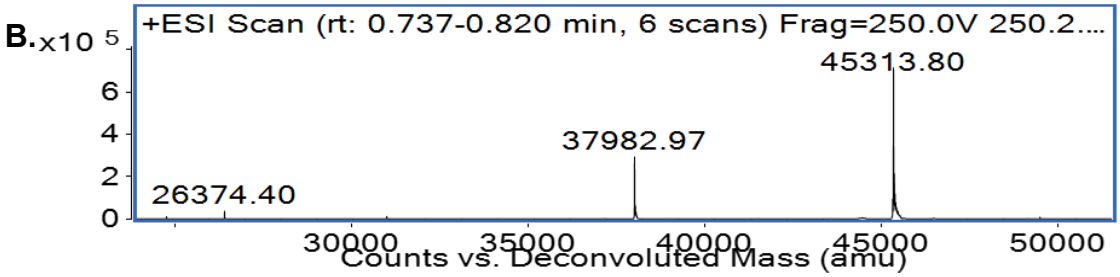
A1. SUMO fragment

```
        10         20         30         40         50         60
MHHHHHHSSG VDLGTENLYF QSMSDQEAKP STEDLGDKKE GEYIKLKVIG QDSSEIHFKV
        70         80         90        100        110        120
KMTTHLKKLK ESYCQRQGVP MNSLRFLFEG QRIADNHTPK ELGMEEEDVI EVYQEQTGGK

PISLS
```

Molecular weight: 14311.02
Theoretical pI: 5.81

B.



B1.- Intact protein

```
        10         20         30         40         50         60
DDAAIQQTLA KMGIKSSDIQ PAPVAGMKTV LTNSGVLYIT DDGKHIIQGP MYDVSGTAPV
        70         80         90        100        110        120
NVTNKMLLKQ LNALEKEMIV YKAPQEKHVI TVFTDITCGY CHKLHEQMAD YNALGITVRY
       130        140        150        160        170        180
LAFPRQGLDS DAEKEMKAIW CAKDKNKAFD DVMAGKSVAP ASCDVDIADH YALGVQLGVS
       190        200        210        220        230        240
GTPAVVLSNG TLVPGYQPPK EMKEFLDEHQ KMTSGKGSSM HHHHHHSSGV DLGTENLYFQ
       250        260        270        280        290        300
SMSDQEAKPS TEDLGDKKEG EYIKLKVIGQ DSSEIHFKVK MTTHLKKLKE SYCQRQGVPM
       310        320        330        340
NSLRFLFEGQ RIADNHTPKE LGMEEEDVIE VYQEQTGGKP ISLS
```

Molecular weight(reduced): 37984.22
Theoretical pI: 5.84

B2.-DsbC fragment

```
        10         20         30         40         50         60
DDAAIQQTLA KMGIKSSDIQ PAPVAGMKTV LTNSGVLYIT DDGKHIIQGP MYDVSGTAPV
        70         80         90        100        110        120
NVTNKMLLKQ LNALEKEMIV YKAPQEKHVI TVFTDITCGY CHKLHEQMAD YNALGITVRY
       130        140        150        160        170        180
LAFPRQGLDS DAEKEMKAIW CAKDKNKAFD DVMAGKSVAP ASCDVDIADH YALGVQLGVS
       190        200        210        220        230        240
GTPAVVLSNG TLVPGYQPPK EMKEFLDEHQ KMTSGKGSSM HHHHHHSSGV DLGTENLYFQ

SM
```

Molecular weight(reduced): 26375.13
Theoretical pI: 6.04

Figure 3.28: Degradation sub-products of CXCL12 constructs

ESI-QTOF/MS spectra of IMAC-purified fractions of V3I mutant CXCL12 in (**A)** CXCL12A-c027 (pSuf1) and (**B)** CXCL12A-c025 (pDsbC). Sequences of the Met-His6-SUMO-KPISLS (**A1**), DsbC-His6 (**B1)** and DsbC-His6-SUMO-KPISLS (**B2)** fragments with MW and pI calculations are also shown; MW corresponds to reduced mass. -2 Da correspond to oxidation of DsbC in (B1, B2). Peptides from CXCL12 in green; peptides from SUMO in red. TEV cleavage site in blue.

3.5.IV Optimisation of large-scale expression conditions

Despite the initial success in test expression scale, large scale expression yielded inconsistent results, and intact protein yield greatly varied between batches. These were apparently related to cell stress: where degradation was extensive reach lower cell densities after induction were achieved (16 g/L; successful batches yield 27 g/L). The degradation products having consistent MW across experiments would hint to specific stress response e.g., protease activity. One of the main causes appeared to be the age of the streaked plate – colonies kept at 4ºC yielded little to no protein, whereas freshly streaked plates produced mainly intact fusion protein CXCL12. Use of fresh colonies (from glycerol stock) solved the inconsistency problem but *in vivo* truncation of the DsbC-SUMO-CXCL12 construct was still significant.[32]

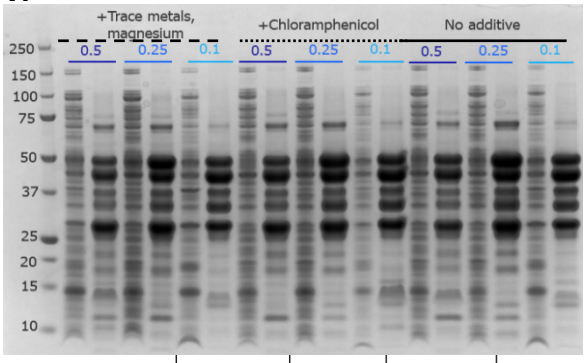
Sec pathway-mediated export is adversely affected by low temperature [218] as several of the chaperones involved belong to the heat-shock family, but higher induction temperatures could increase protease activity. Furthermore, the periplasmic space is rich in metal-dependent proteases: my use of rich media (TB+) containing metal cations could increase their activity. The presence of additional antibiotics (such as chloramphenicol) could also be inducing additional metabolic stress to the cells. I assessed yields modifying these parameters as well as culture conditions (aeration, and additional carbon/nitrogen sources), plus *lac* inducer concentration (Figure 3.29). Both addition of metal micronutrients, or temperature increase over 18°C, resulted in increased truncation of CXCL12 and lower protein yields, potentially due to higher metalloprotease activity in the periplasm. There was also an optimal point for media richness; whilst 2XLB yields almost no intact protein compared to TB, additional carbon and nitrogen sources in TB were detrimental for expression of full-length CXCL12. Harsher induction was also beneficial, up to a limit (0.25 mM IPTG being the optimal concentration). Addition of antifoaming detergents had no beneficial effect, whereas increased agitation decreased yield.[33] I therefore performed my expressions of CXCL12 constructs in pDsbC-HT-CBio in TB without additional nutrients or detergents, inducing at 0.25 mM IPTG (16h/ 18ºC), and 180 rpm agitation (0.45 RCF).

---

[32] Addition of glucose to the transformant plates and preculture did not improve this degradation, ruling out *lac* expression leakage. To date, I am unable to explain this colony age-related toxicity.

[33] Different induction times at higher temperature were not tested as even if they were optimal, this would result in a final purification process requiring more than 24 hours of continuous work. Freezing cells prior to outer membrane lysis (pause point) could also lyse the cytosol leading to contaminants (see 3.4.V)

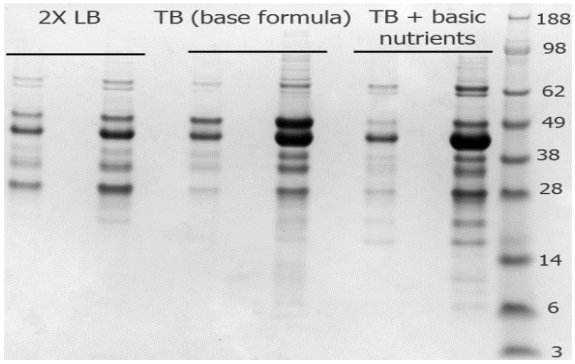
A1



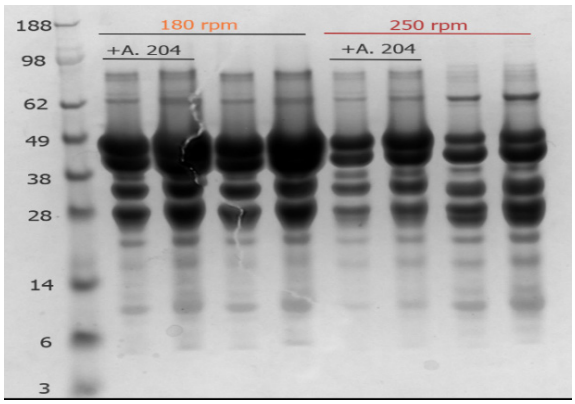
A2.

Temperature (ºC)	Additive	IPTG (mM)	Cell mass (g)	Yield (mg/L)
37	None	0.5	3.85	0.5
		0.25	5	2.3
		0.1	3	0.5
25		0.5	3.75	11.0
		0.25	3.9	10.3
		0.1	5.2	12.8
18		0.5	3.3	14.8
		0.25	4.8	33.0
		0.1	5.5	13.0
	Trace Metals, MgSO4	0.5	3.55	9.1
		0.25	4.2	24.1
		0.1	5.63	11.2
	Chloram-phenicol	0.5	3	15.6
		0.25	4.45	17.6
		0.1	5.2	17.9

B.



C.



A3.

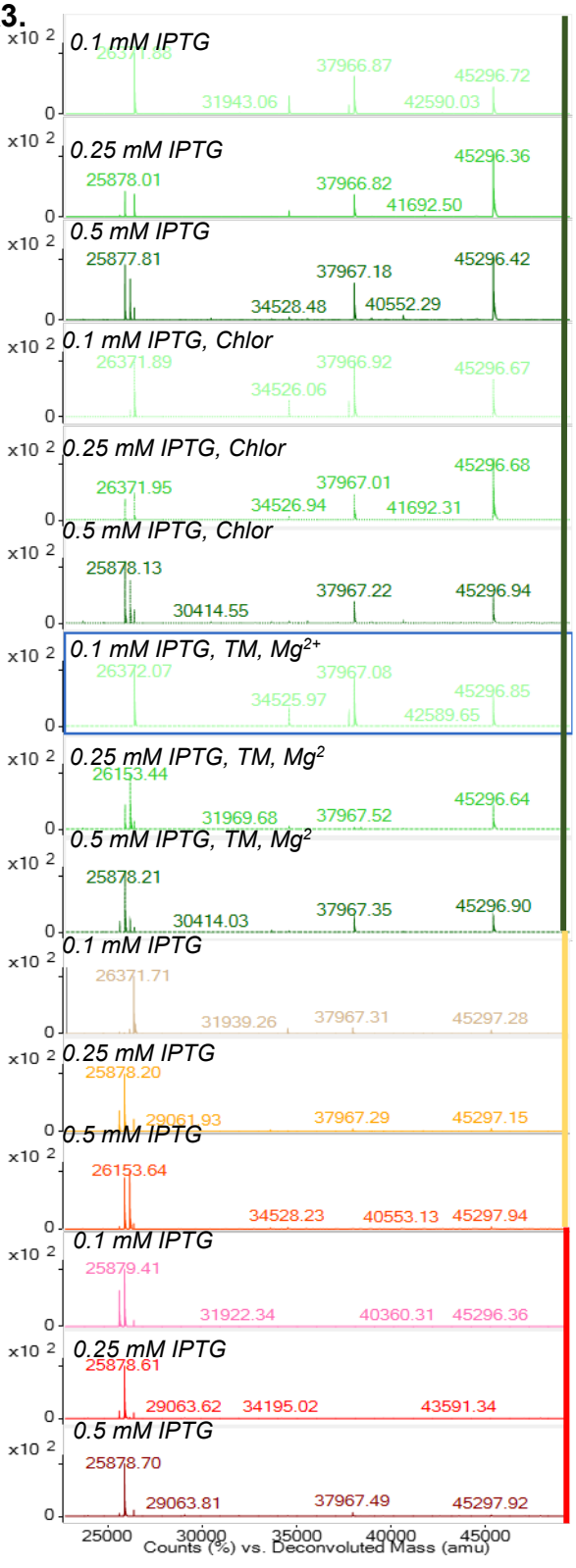


Figure 3.29: Optimisation of CXCL12 expression in pDsbC constructs (CXCL12A-c021)

A1. SDS-PAGE of CXCL12A-c021 (whole cell lysis) expressed with the addition of 1X trace metal mixture and 1 mM $MgSO_4$, chloramphenicol, and no additives, in TB with 1% glucose, 25 mM $(NH4)_2SO_4$ and 1.6% glycerol. For each condition, left lane is lysate, right lane elution. **A2.** Quantification of yields and cell mass with different temperature ranges and IPTG concentration for induction from (A). **A3.** ESI-QTOF/MS of the protein samples in (B). Note the increase in degradation with temperature (green = 18 ºC, yellow = 25 ºC, red = 37 ºC),or supplementation of media. **B.** Comparison between 2X LB, TB without additives, and TB with 1% glucose, 25 mM $(NH4)_2SO_4$ and 1.6% glycerol. **C.** Comparison between agitation speed and addition of antifoaming agent (0.01% Antifoam-204).

3.5.V Optimisation of cell lysis and downstream purification

Due to the large difference in pI between cleaved CXCL12 (pI 9.9), DsbC or DsbC-SUMO fragments (pI 6.04 and 5.84, respectively; see Figure 3.28) and Ulp1 (pI 8.64), cation exchange chromatography would theoretically be ideal to remove these. However, some of these contaminants apparently had similar MW and pI to CXCL12 (Figure 3.30A) which made them impossible to remove by cation exchange or further SEC without losses[34].

However, as the construct was being exported to the periplasm, I tested the use of outer membrane lysis, a protocol with several variants [219]. I used EDTA (removes the cations holding the outer membrane together) and lysozyme (digests the peptidoglycan in the membrane) A high-osmotic environment (0.2M buffer + 1M sucrose) prevents the outer membrane from fully lysing and thus exposing the inner membrane for lysis. To counteract metal ion stripping and permit benzonase activity (inhibited by EDTA), $MgCl_2$ was added in 5-fold molar excess to EDTA, prior to osmotic shock in a large excess of water to burst the outer membrane and leave spheroplasts containing the cytoplasmic fraction to be pelleted.

With outer membrane lysis, all purified protein bands from IMAC corresponded to DsbC-SUMO-CXCL12 fusion protein and its varying truncations (Figure 3.30B) as a large fraction of the cytoplasm and its contaminants is discarded with the spheroplasts, resulting in much higher purity from cation exchange. To better resolve CXCL12 from the cytoplasmic contaminant, particularly when using the CyDisCo expression system (where outer membrane lysis is not an option), I also switched to 7.7mm*100 mM CaptoS Tricorn column format (from HiTrap HP-SP 5 mL). This required a decrease in flow rates, but as sample volumes decreased significantly due to the lower protein content (allowing for higher concentration in the loaded sample) it was feasible.

[34] Reverse IMAC was not feasible as I found out that tag-less CXCL12 non-specifically bound to agarose resin, even at high ionic strength (> 500 mM NaCl). CXCL12 also binds non-specifically to other resins such as Strep-Tactin XT, and could be due to CXCL12´s propensity to bind glycans electrostatically [114,288].

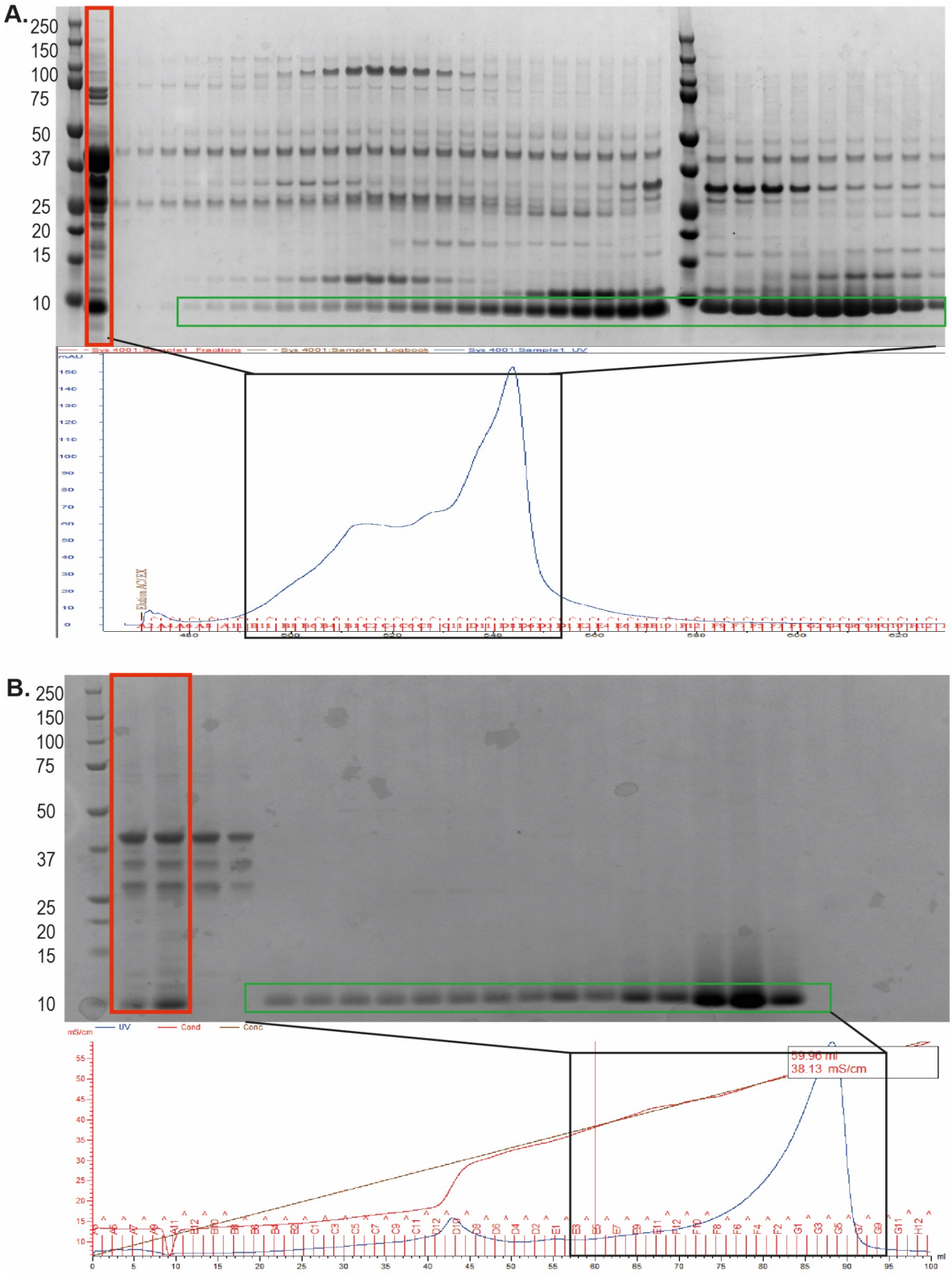


Figure 3.30: Comparison of cell lysis methods for CXCL12A-c021 purification

SDS-PAGE of CXCL12A-c021 purified by IMAC and cation exchange chromatography. Red rectangles show the IMAC pool; green rectangles the cleaved CXCL12 in the IEX run. Black rectangles with lines show the chromatogram fractions loaded.
A. Cells lysed by sonication, eluted by 5 mL HiTrap SP-HP column. **B.** Cells lysed by outer membrane lysis. The lower protein concentration allowed use of a higher resolution 4.7 mL CaptoS Tricorn column due to lower sample volumes.

Final protein polishing via SEC (Figure 3.31) yielded nearly 100% pure protein in SDS-PAGE, and ESI-QTOF/MS. Yields, whilst not as large as those of HMGB1 constructs, were sufficient to remain within the realm of practicality especially when taking into account the MW of this protein. 1 mg (150 nmol) can be purified from a litre of culture with these methods, an amount sufficient even for NMR since a 7 mg/mL solution is already at 1 mM.

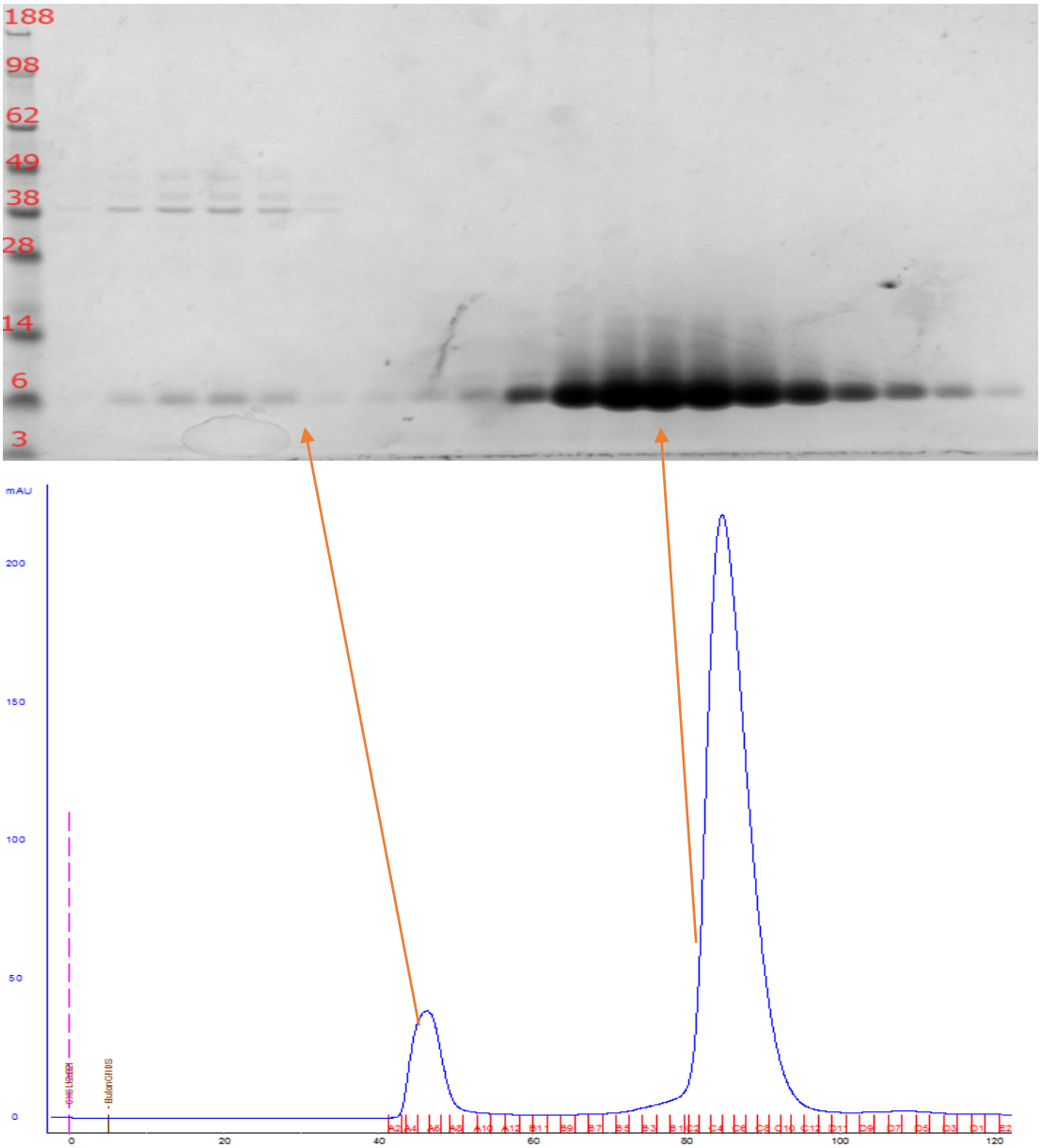


Figure 3.31: SEC chromatography of CXCL12

Superdex 75 PG 16/600 elution profile of CXCL12A-c021 (after cleaving and cation exchange). The purified protein has a monodisperse elution profile, whilst the void peak contains some CXCL12 and the remainder of the DsbC-SUMO tags.

3.5.VI Characterisation of CXCL12

I was regrettably unable to obtain reproducible thermal melting profiles for CXCL12. CXCL12 has several surface-accessible hydrophobic patches and no buried core to speak of (Figure 3.32) [220]. As SyPro Orange dye binds hydrophobic residues which are exposed to solvent upon melting to produce fluorescence, it could also bind these, particularly in those conditions where CXCL12 would be monomeric, e.g., HEPES or pH < 5.5, although this should give large initial fluorescence values I did not detect. The combination of a lack of a buried hydrophobic core, and significant presence of hydrophobic residues in the solvent-accessible surface would suggest that direct hydrophobicity monitoring is not adequate to characterise CXCL12 unfolding and would explain why without background artefacts no proper melting curves could be observed.

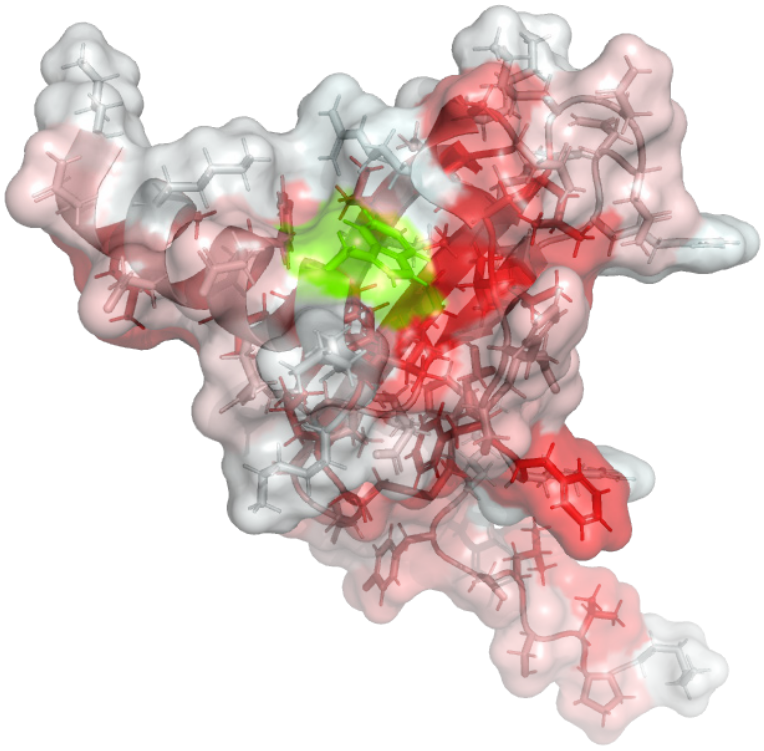


Figure 3.32: Hydrophobicity colouring of CXCL12

Eisenberg hydrophobicity colouring (bright red = most hydrophobic, white = most hydrophilic) in PyMol of **A.** CXCL12 monomer (PDB 2KED) with Trp 57 highlighted (in green). CXCL12 has large hydrophobic patches which are solvent-accessible (red areas; side chains of hydrophobic residues) and would bind Sypro Orange. The side chain of Trp 57 is partially buried and could be used for nanoDSF (tryptophan fluorescence) thermal unfolding measurements.

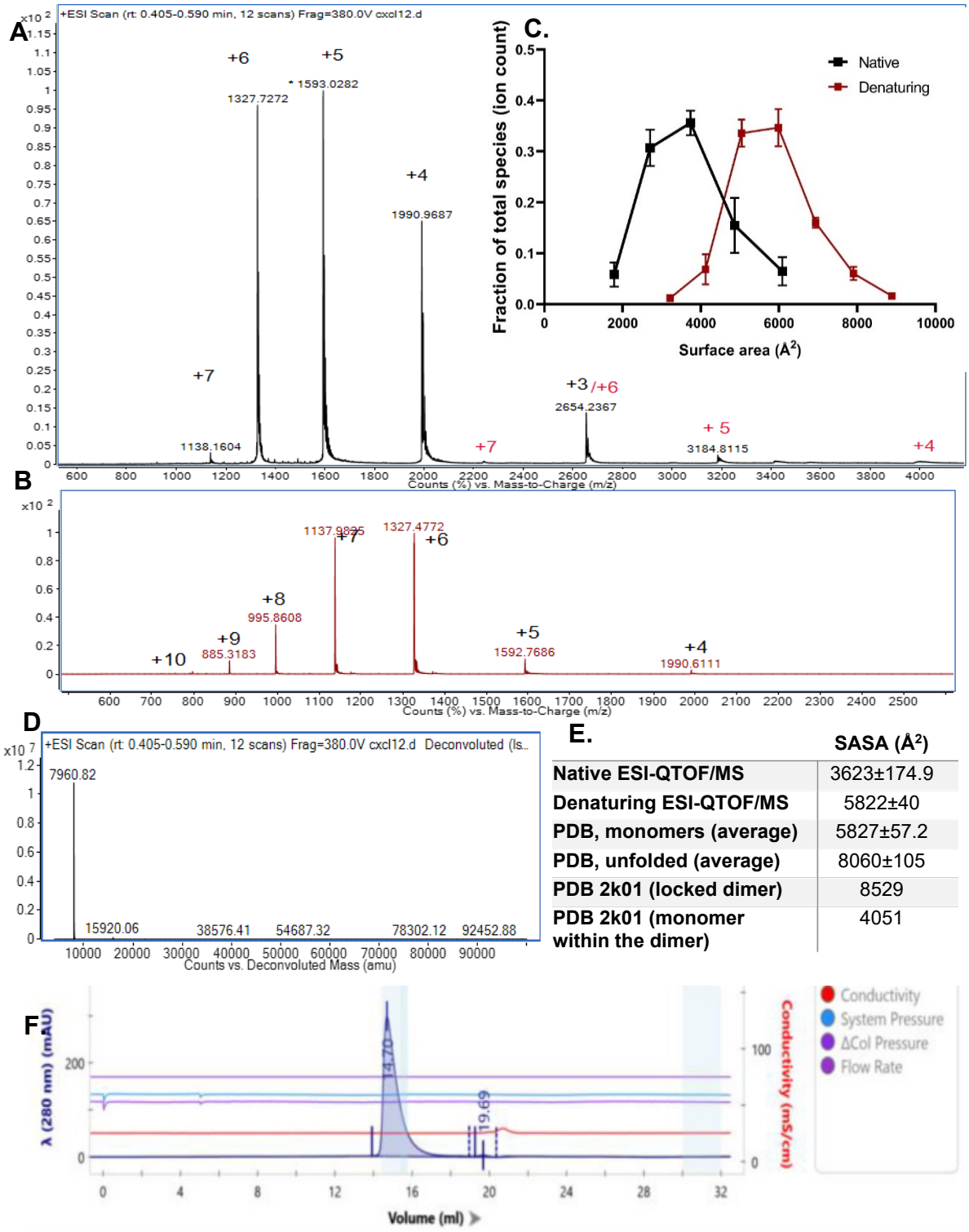
One partially buried hydrophobic residue within CXCL12 is Trp 57 (Figure 3.32A), which lies beneath the loop between the α-helix and the β-strand of CXCL12. Therefore, nanoDSF[35] (Monolith Prometheus platform), which relies on tryptophan fluorescence alone, could be a more accurate measurement of unfolding. A pilot experiment (Figure 3.34) measured apparent Tm_{50} > 60 ºC, but which increased with CXCL12 concentration. I have been unable to pursue this approach further due to the Covid-19 lockdown. Therefore, I could not use thermal melting to evaluate thermostability of CXCL12.

Next, I performed denaturing and native ESI-QTOF/MS of CXCL12. Whilst I grossly underestimated the ionizability of CXCL12, which led to saturation of the detector, I was able to identify a major species consisting of a folded monomer with SASA 3623 $Å^2$, and a potential dimer. The SASA of these dimers is however not feasible: the observed charge states are the same as I see in CXCL12 monomers (+7 to +4) and global theoretical SASA is therefore the same as for the monomer. In contrast, the SASA of available CXCL12[36] dimers (2K01 [221], locked dimer) is around 8529 $Å^2$, and I suspect these to be an artifact product of ionisation [199].

The SASA average of published CXCL12 monomer structures (PDB 2SDF [111], 2KEC, 2KE - pH < 5.5 precludes dimerisation, 2KEE [112] - no divalent anions are present) averages 5827 $Å^2$. This is better matched by my denaturing ESI-QTOF/MS measurements at pH 2.0 in methanol/formic acid (5822 $Å^2$; H25 protonated due to low pH [114] and therefore monomeric), rather than those from native ESI-QTOF/MS (3623 $Å^2$). It is also be possible that, due to the compact chemokine fold structure and high percentage of positive residues in CXCL12, charge acquisition in native MS is hampered [143,144,197] whereas on denaturing this alleviated by the unfolding and the extension of the N-terminus, despite published data suggesting this unlikely [143]. Further elucidation of this phenomenon would require a full NMR analysis in the same buffer conditions (pH 6.8 ammonium acetate, 0.2 M), coupled with other measurements of protein folding (e.g., circular dichroism or Fourier transform infrared spectroscopy).

[35] Performed by Cassandra Dix as part of Figure 3.51B; CXCL12 sample contained detergent.

[36] Given that a large proportion of CXCL12 is comprised of flexible regions, I only analysed NMR structures (which contain multiple conformations of the protein), rather than crystal structures where only one static conformation is represented.



	SASA (Å²)
Native ESI-QTOF/MS	3623±174.9
Denaturing ESI-QTOF/MS	5822±40
PDB, monomers (average)	5827±57.2
PDB, unfolded (average)	8060±105
PDB 2k01 (locked dimer)	8529
PDB 2k01 (monomer within the dimer)	4051

Figure 3.33: ESI-QTOF/MS and SEC of CXCL12A-c021 (cut)

A. Native ESI-QTOF/MS of CXCL12A-c021 M/Z spectra, in 50 mM ammonium acetate pH 6.8; charge states and conformations (charge state distributions) have been annotated. Asterisks indicate detector signal saturation; full spectra located in Suppl. data A2-3-J. **B.** Denaturing ESI-QTOF/MS CXCL12A-c021 M/Z spectra; charge states have been annotated **C.** Ion count vs. SASA for each charge state in (A) and (B). **D.** Deconvolution of (A). **E.** Solvent accessible surface area calculations for (C) and several PDB structures. **F.** SEC elution profile of CXCL12 in 50 mM HEPES pH 7.5, 300 mM NaCl, 5% glycerol.
For MS experiments, Ion counts and SASA calculations averaged over three spectra, with at least 10 counts each.

3.5.VI.i Validation of CXCL12 binding to CXCR4[37]

Whilst originally it was also part of my DPhil project, I was unable to devote time to expression and purification of CXCR4 beyond the initial construct design. However, other researchers at the SGC were able to pursue the project based on my initial constructs, and at the point of me having characterised CXCL12, CXCR4 was also available in our labs. Given the reported difficulties in producing CXCL12 - even requiring HPLC to discriminate disulfide bridge pairs [126,222], and my inconsistent DSF experiments, the question arose – was my CXCL12 properly folded? A robust way to be certain of this is functional characterisation of its binding to the native receptor – CXCR4.

On SEC (Figure 3.34A), CXCL12 co-elutes with CXCR4, despite the vast difference in molecular weight (128 kDa for a CXCR4-GFP fusion construct, 8 kDa for CXCL12) in a Superdex 200 column (Mr 600-10 kDa), eluting significantly earlier than when alone. Interestingly, it preferentially bound to the peak corresponding to CXCR4 monomers, rather than to the CXCR4 dimer peak. The monomer peak was increased in abundance with respect to the dimer peak when compared to the elution profile when CXCL12 was absent. Furthermore, a titration curve in nanoDSF (Figure 3.34B), showed interaction of CXCL12 with CXCR4 (K_d = 1.9 μM). Whilst a thousand-fold lower than the consistently reported values in the literature of 1-10 nM range [102,123,223], and therefore potentially representing incorrect folding, the strong dose-dependency of the effect, and the disparity of these experimental conditions to those published with either membrane-embedded CXCR4 [111,122] or with a heavily mutated CXCR4, stable without detergents, via microscale thermophoresis (MST) [224]; all of these were performed in the absence of detergents which will necessarily weaken the interaction between macromolecules.

[37] Production of CXCR4 and the experiments described in Figure 3.34 were performed with my recombinant CXCL12 by Cassandra Dix, the Research Assistant collaborating with me in elucidating the ternary structure of the heterocomplex; I produced CXCL12, designed the experiments, and analysed the data. I would also like to acknowledge the contribution of Patrizia Abrusci and Tina Bohstedt for optimising CXCR4 expression based on the CXCR4-GFP (L125W, T240P) thermostabilised constructs I designed at the beginning of my DPhil.

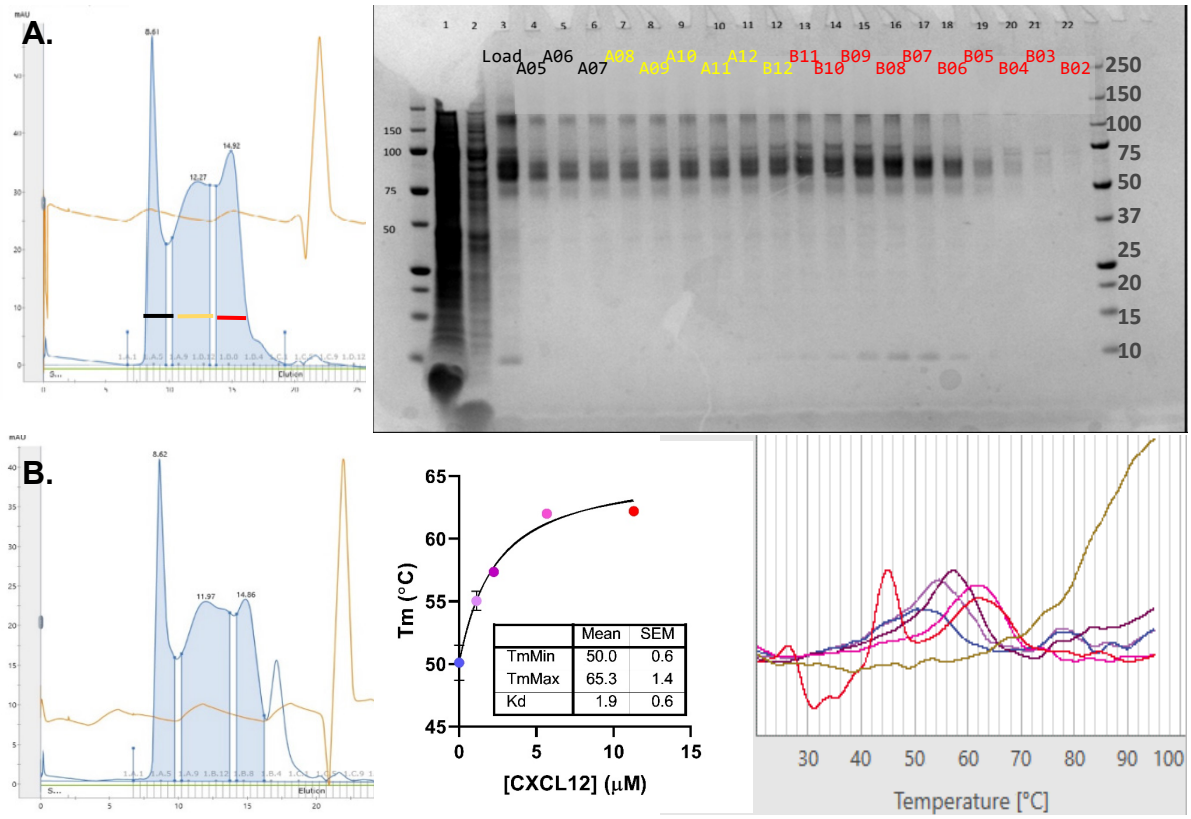


Figure 3.34: Interaction of CXCL12A-c021 with CXCR4-GFP

A. CXCR4-GFP construct, purified in 50 mM HEPES pH 7.5, 100 mM NaCl, 0.01% DDM/0.001% CHS, in the presence of 2 μM CXCL12A-c021 via streptavidin affinity, and injected in a 10/300 Superdex 200 Increase column. Peak fractions were loaded in SDS-PAGE for qualitative analysis. In black: void peak (8.6 mL). In yellow (12.27 mL): Potential CXCR4 dimers (estimated mW, 160 kDa). In red: Potential CXCR4 monomer; CXCL12 elutes in this fraction (14.9 mL) **B.** nanoDSF (Prometheus, NanoTemper) of CXCL12 alone (brown), CXCR4 alone (blue) and CXCL12-CXCR4 in different ratios (red and violet lines). The Tm_{50} of each condition was fitted to a modified Langmuir binding isotherm to calculate the approximated kD, with two replicates for each CXCL12 concentration. Note the second melting transition in CXCR4 melting: this biphasic melt profile has been characterised in the literature [224]and is not present with CXCL12 alone.

Whilst no studies of CXCL12 binding to solubilised CXCR4 in the presence of detergents are reported in the literature, a comparative study on *Sf9* insect cell-expressed CXCR7 (ACKR3) [123] shows a sharp increase in off rates of detergent-solubilised receptor in SPR compared to those measured from receptor in the cell membrane. The K_d computed in this study is in the 400 nM range, more than two orders of magnitude higher than the average K_d for membrane-embedded CXCR7 and CXCL12 (1.21 nM [123]). As CXCR4 and CXCR7 are remarkably similar, it would be reasonable to expect this phenomenon to also apply, and K_d obtained from cell-free CXCR4 to be much weaker than those values derived from membrane-embedded CXCR4. Furth work with inhibitors (e.g., AMD-3100) or neutralising antibodies will be needed to assess the biological significance of these results and comparing to published data from binding to membrane CXCR4.

3.5.VII Conclusions – CXCL12

Whilst HMGB1 expression was relatively straightforward, CXCL12 proved to be a more complex target, and most techniques I attempted to express it with failed. Whilst it has been expressed via refolding and successfully crystallised, I lacked the access to the equipment required (large-scale RP HPLC). Even in my final protocol, DsbC-SUMO-CXCL12 fusion protein is truncated extensively in *E. coli* in a specific sequence within the N-terminus of CXCL12. Mutagenesis of those residues is unfeasible, as the N-terminus of CXCL12 is critical for biological activity and any modification would drastically alter its activity [103,111]. However, this system is likely adaptable to other disulfide-containing molecules with critical N- or C-terminal residues for activity, when tag-free production is needed, such as most CXC/CC chemokines and growth factors such as transforming growth factor beta (TGF-β), since the use of SUMO tag ensures no alteration of the final protein.[38]

On the other hand, its small size does not leave much room for unusual behaviour, although it was striking to find that the monomeric conformations published in the PDB all seem to match the acid-shaped form of CXCL12 with larger SASA than my ESI-QTOF/MS and SEC measurements. However, this is to be expected as CXCL12 seems to be overall more stable in acid pH and therefore, amenable to either long-term NMR spectroscopy or crystallography; most available structures have been obtained in such buffer conditions. CXCL12 is remarkably stable; denaturing ESI-QTOF/MS cannot fully unfold it due to a combination of disulfide bridges and high pI, and the protein was stable for extended periods of time (>1 week at RT). Ironically, this is likely the reason expression of the recombinant protein is so complex; a high free folding energy would hamper the transition between the nascent peptide (unfolded) and the folded protein, but this would work both ways and thus increase the free energy associated with protein denaturation.

[38] I attempted modification of the pDsbC-HT-CBio vector to include an in-frame SUMO tag for this purpose, but I was regrettably unable to complete this due to time constraints. This was motivated by the presence of a poly-adenine stretch 3´terminal in SUMO which leads to frameshift mutations as a result when merging with CXCL12 DNA fragments (such as those occurring in CXCL12-c022/23/24 from the experiment in Figure 3.39).

3.6 Final remarks

With both proteins in hand, and confident that my protocols were robust, reproducible, and scalable, as well as knowing that in both cases the result was biologically active proteins, I was in a good position to perform all the experiments required to probe the HMGB1-CXCL12 interaction. Furthermore, it would also mean that any unexpected results downstream would not be a product of inconsistent proteins or batch-to-batch variation. As I will describe in Chapter 4, this was vital since the results from most experiments I performed on the heterocomplex would not agree with each other when interpreted in isolation; only by analysing the whole sum of those data I could derive reasonable conclusions. In the next chapter I will describe the structural determination of the HMGB1-CXCL12 heterocomplex.

Chapter 4 The interaction between CXCL12 and HMGB1

4.1 Chapter introduction 134
4.2 Biology of HMGB1 as a potential therapeutic 135
4.3 Characterisation of the heterocomplex 138
4.3.I SEC of the HMGB1-CXCL12 complex 138
4.3.II Native ESI-QTOF mass spectrometry 140
4.3.III Screening of FL-HMGB1 and CXCL12 interaction 142
4.3.IV Crystallisation trials 145
4.4 Sequence-specificity of the HMGB1-CXCL12 interaction 146
4.4.I Peptide arrays 147
4.4.II Nuclear Magnetic Resonance (NMR) 150
4.4.III Kinetic validation of CXCL12 binding to different HMG Box domains and redox states 158
4.5 Conclusion: defined, homologous regions within each HMG Box domain bind CXCL12 162

4.1 Chapter introduction

Following optimising expression and purification of the proteins described in Chapter 3, I had now enough starting material to analyse the interaction between HMGB1 and CXCL12. But where to start?

The interaction between these two molecules has been identified fairly recently, with the first reports in the literature in 2007 [32,225]. Most publications in the subject study the biological effects of the heterocomplex, such as chemotaxis [96,226,227] and tissue repair enhancement [25,173,228]. Only one study published in 2012 analysed the interaction from a structural point of view [96] prior to my DPhil. In this publication, HMGB1 was shown by SPR to form a complex with CXCL12, inhibited in a dose-dependent manner with glycyrrhizin. This study identified that some residues in both CXCL12 and HMGB1 experienced chemical shift perturbation (CSP) upon binding. However, the data from this study have never been made fully available to the academic community. The complete SPR data were only published 5 years later in a thesis [131], but the NMR spectra which describe the interaction of CXCL12 with HMGB1 have not been deposited in any database.

Therefore, when I commenced this part of my project, only very limited information was available on how to proceed. Spectra for HMGB1 interacting with CXCL12 were only made available by the same group in a later publication (2019) [183], where the residues involved in CXCL12 binding were now identified by intensity variation rather than CSP. This will be discussed in detail in the relevant section as, whilst having limited influence in my experimental design, these studies could be contrasted with my own data which I describe in this chapter.

4.2 Biology of HMGB1 as a potential therapeutic

The biological effects of the interaction of the FR-HMGB1-CXCL12 heterocomplex with CXCR4 on tissue repair are well described [25]. FR-HMGB1 can trigger [25] the G_{Alert} state [98] in resident stem cells, enhancing tissue regeneration, even when the injury and the HMGB1 stimulation are temporally separated by long periods (>2 weeks). However, it was not known whether FR-HMGB1 would have the same regenerative effect if it were administered after the injury, a situation much more typical in clinical practice (except for planned surgery), and no detailed dose-response curves were available [25]. I used the $BaCl_2$ muscle injury model [98] to assess these parameters, with FR-HMGB1 being injected intravenously (i.v.) instead of locally as local administration is not feasible in many clinical settings (e.g. myocardial infarction).

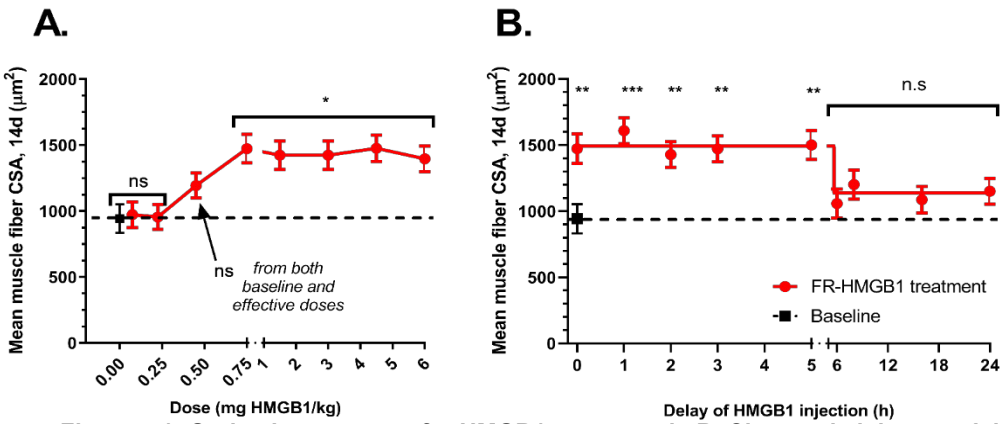


Figure 4.1: Optimal parameters for HMGB1 treatment in $BaCl_2$ muscle injury model

A. Dose-response curve of FR-HMGB1, LPS-free, injected i.v. immediately prior to injury. A dose of 0.75 mg/kg was proven sufficient to achieve maximal regeneration enhancement. A dose of 0.45 mg/kg seems to be unable to reach full regeneration but achieves intermediate regeneration (n.s. with both control and optimal dose). To account for variations in e.g., administration route, 1.2 mg/kg (30 µg in a 25g animal) was used henceforth. **B.** Injury timing experiment: HMGB1 was injected at the indicated time after injury, and regeneration was assessed at 14 days.
Both datasets analysed via nested one-way ANOVA. N1=5 animals per dose or timepoint, and at least 40 cells assessed over 4 fields (n2 > 40)

Regenerative effect was maximal at doses of 0.75 mg/kg (29 nM/kg) and above: there was no further improvement in efficacy beyond this limit. Therefore, I selected 1.5 mg/kg as the optimal dose (Figure 4.1A) to ensure maximal regenerative effect whilst accounting for potential future variations, particularly related to animal batches. Intravenous administration of HMGB1 was effective in promoting regeneration up to and including 5 h post-injury (Figure 4.1B).

It has been shown in *Danio rerio* (zebrafish) [229] that muscle cell regeneration starts fairly quickly after injury, with activation and migration of *myf5*-positive satellite muscle cells to the injured area occurring between 6 and 9 h post injury. These cells comprise the resident stem cell population which give rise to new muscle fibres. My data agree with those published using the zebrafish model: the lack of effect after a similar interval of 5-6 h is due to the muscle satellite cells responsible for regeneration being already activated by the injury [229]. Once activated, the stem cells cannot enter G_{Alert} as they have entered the G_1 phase of the cell cycle and hence HMGB1 would not be expected to have any effect at this stage.

The hypothesis of stem cell activation occurring after this period of 5 h would also be in line with the observation of Rando and his colleagues [230], who studied mTORC1 activation and hence induction of the G_{Alert} state in a gut injury model using bacterial challenge in *Drosophila melanogaster*. Upon bacterial challenge, the levels of phosphor-eIF4E binding protein (4-EBP), which is phosphorylated upon mTORC1 activation, increased after 4 h post-injury, and steadily decay afterwards. The latter corresponded to entry of the resident stem cells into cell cycle prior to tissue regeneration after 4 h. There was a marked increase in cell size (entry into G_1) at 17 h (1020 min), which coincided with the timing in the aforementioned zebrafish model [229], where asymmetric stem cell division occurs between 750 min (12h) and 1020 min (values approximated from Fig. 1-I in the publication).

Exogenous FR-HMGB1 promotes tissue regeneration [25], but it is not known if or if heterocomplex formation is limited by the CXCL12 available endogenously, to achieve maximum regenerative effect. I tested whether additional CXCL12 would boost regeneration if this were the case. As I also had access to the V3I mutant of CXCL12, which has been described as leading enhanced signalling as assessed by calcium influx compared to wild-type CXCL12 on binding CXCR4, despite the binding affinity for both isoforms to CXCR4 being equivalent [111], I also tested this construct to assess if more effective CXCL12-CXCR4 signalling would also boost tissue repair.

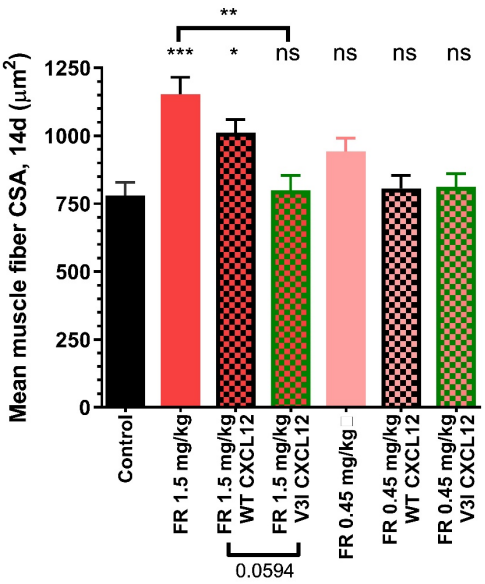


Figure 4.2: Effect of CXCL12 co-administration in the regenerative effects of HMGB1

CXCL12 was locally injected immediately prior to HMGB1 injection and injury, at 0.5 mg/kg; HMGB1 was injected at optimal dose (bright red; 1.5 mg/kg) and suboptimal (pale red; 0.45 mg/kg). Black chequers: wild-type CXCL12, green chequers: V3I mutant CXCL12 with enhanced CXCR4 activation. In this experiment, both control and treated animals displayed lower average mean muscle fibre CSA than the usual values (≈1500 μm^2 for subjects with enhanced regeneration and ≈1000 μm^2 for control subjects); this was also visible upon inspection of the slides.
All data analysed via nested-one-way ANOVA; post-hoc comparisons corrected via Holm-Sidak method. Statistical significance values shown correspond all to control PBS injection, represented by black lines in (A) and (B), unless otherwise indicated. N=5 animals per dose or timepoint, and at least 40 cells assessed over 4 fields (n2 > 40)

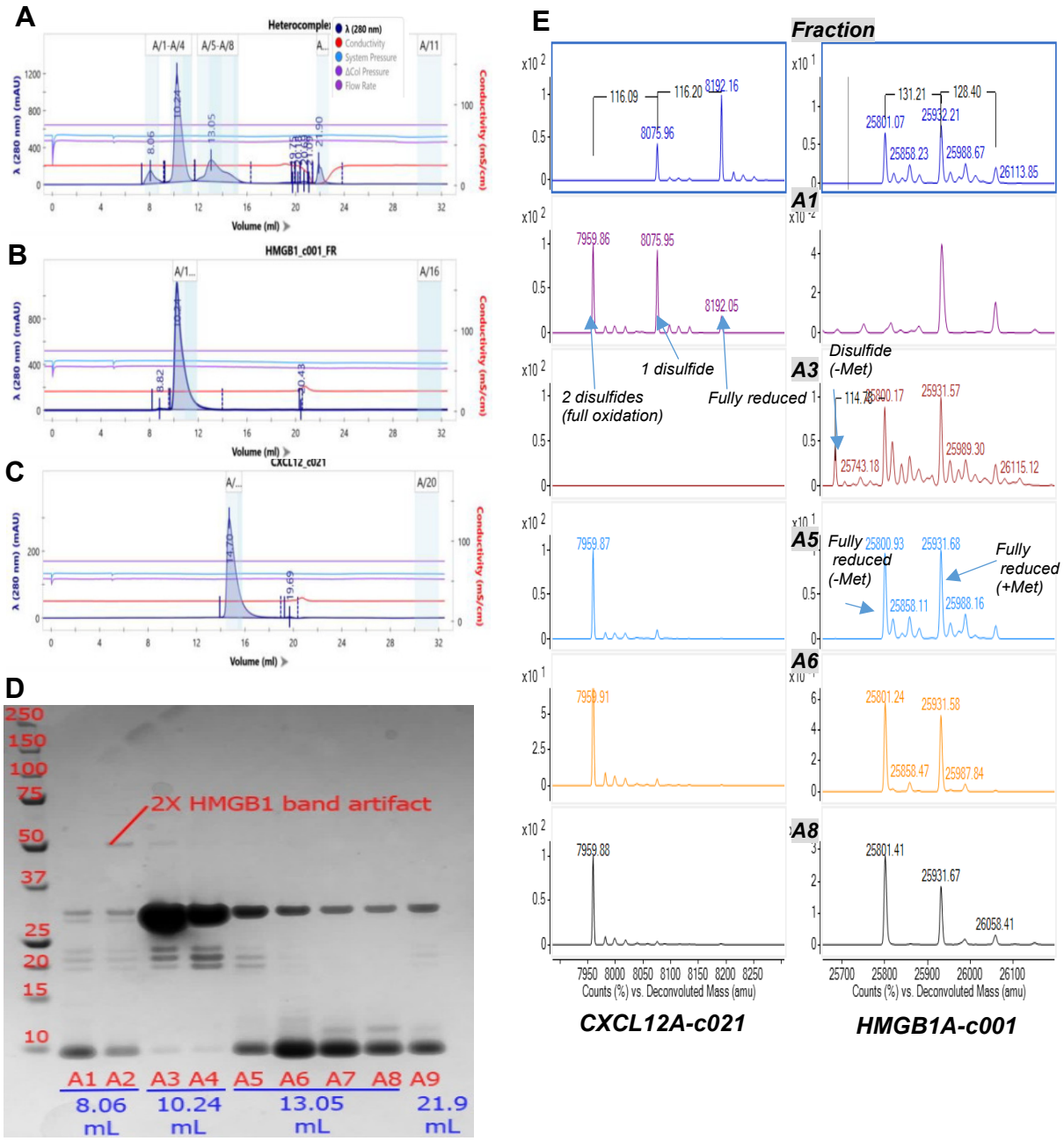
However, co-administration of exogenous wild-type CXCL12 did not enhance the effects of FR-HMGB1 (Figure 4.2), and in fact appeared to reduce the efficacy of FR-HMGB1, though this did not reach statistical significance at the doses tested. Co-administration of V3I mutant abrogated the effect of FR-HMGB1, even at the optimal dose of FR-HMGB1, completely inhibited the regenerative effect of FR-HMGB1. Given that both forms of CXCL12 forms bind CXCR4 with equal affinity, their differing effect is most likely due to their differing CXCR4 agonism. On activation by CXCL12 alone, CXCR4 is internalised [133], resulting in desensitisation to CXCL12, and consequently, to the CXCL12-HMGB1 heterocomplex. A stronger CXCL12-like agonist such as the V3I mutant would, therefore, be expected to have a greater effect on inhibiting the action of the FR-HMGB1-CXCL12 heterocomplex.

4.3 Characterisation of the heterocomplex

4.3.I SEC of the HMGB1-CXCL12 complex

Whereas on their own FL-FR HMGB1 (RV 10.24 mL, SASA from native MS ≈ 17000Å^2) and CXCL12 (RV 14.70 mL, SASA ≈ 4000 Å^2) elute in single peaks in SEC, additional peaks are observed when a 1:2 mixture of both is eluted (Figure 4.3). Both the main HMGB1 and CXCL12 peaks appear at the same RV, with mixtures of both proteins in the additional peaks. Analysis of these would indicate that the complex is very unstable, which has been theorised before [96].

- RV 8.06 mL (before void volume for thyroglobulin at 670 kDa; 8.71 mL), containing HMGB1 and CXCL12, apparently in a 1:2 ratio. It is not clear whether this represents the complex or aggregation. However, this peak is beyond the exclusion limit and thus identification is unreliable.
    - C-terminally degraded HMGB1 fragments are present in similar proportion to FL-HMGB1, unlike in the main HMGB1 peak where they are a minor species. These would contain Box A and Box B, both of which should be able to still bind CXCL12 based on my biological activity data. The individual proteins did not elute in the void peak, however.
- Two overlapping peaks at RV 13.05 and 14.70 mL, with CXCL12 and HMGB1 in uneven concentration gradients. These could represent dimeric and monomeric CXCL12, but as no negative divalent anions are present the dimerisation could be induced by e.g., binding to HMGB1 in a 2:1 manner followed by dissociation of CXCL12 after SEC.
    - Binding to FR-HMGB1 will change the position of the N-termini of CXCL12 into a conformation more favorable for CXCR4 binding [203]. I hypothesize that, whilst the complex itself is only observable in small quantities and elutes before the void volume, the surface conformation of CXCL12 is changed during its association and dissociation from HMGB1 before entering SEC medium.
- RV 21.9 mL, containing both proteins with CXCL12 being more abundant, likely a result of non-specific column interaction (as shown in Figure 3.16) due to the high load, and is too delayed to represent size-exclusion related retention.



F

	Fully oxidised		Partially oxidised		Fully reduced	
	Untreated MW	*+IA MW*	*Untreated mW*	*+IA MW*	*Untreated mW*	*+IA MW*
HMGB1 *(no Met, 3 cysteines)*	25626.8	25683 (Cys 105 free)	N/A	N/A	25628.8	25801 (3 free cysteines)
CXCL12 *(4 cysteines)*	7959.23	7959.9 (no free cysteines)	-	8075.96 (2 free cysteines)	7963.43	8192.16 (4 free cysteines)

Figure 4.3: SEC chromatography of the HMGB1-CXCL12 heterocomplex

FR-HMGB1A-c001 was mixed with CXCL12, at final 0.5 mM HMGB1 and 0.8 mM CXCL12A-c021 (wt), in 20 mM HEPES pH 7.5, 250 mM NaCl, 1 mM TCEP, and loaded in a Superdex 75 pg 10/300 column, eluted in 50 mM HEPES, pH 7.5 300 mM NaCl, 5% glycerol at 0.4 mL/min (**A**). Control samples with HMGB1 (**B**) and CXCL12 (**C**) at the same concentrations. **D**: SDS-PAGE gel of the fractions in (A). Retention volume of the peaks listed. **E:** Close-up of the spectra in LC-ESI/QTOF-MS, at 7900-8300 Da (CXCL12) or 25700-26200 Da (HMGB1). Peak series at +22n corresponds to sodium adducts. **F. MW** of redox states of both proteins, before and after (IA) reaction; carboxyamidomethyl adduct = +57 Da. MW change from reduction = +2 Da (2 protons). +131 adduct in HMGB1 corresponds to the N-terminal Methionine. The untreated MW for HMGB1 is not observable as all the protein is alkylated in this experiment since at least 1 cysteine (105) is reduced in the oxidised protein.

4.3.II Native ESI-QTOF mass spectrometry

When I analysed the heterocomplex in native mass spectrometry I was not able to detect the presence of a heterocomplex (Figure 4.4). Injection of a mixture of HMGB1 and CXCL12 in the nebuliser does not result in the detection of a charge state envelope corresponding to a species matching the theoretical MW (41.5 kDa) of the heterocomplex. I observed however changes in the charge state envelope for both proteins, which would imply that there was a change in conformation associated with the mixture even though direct evidence of a complex was absent.

CXCL12 shifted towards higher charge states (higher SASA) with the main species shifting to charge state +6, from +5 for CXCL12 alone, as in SEC where elution is faster when HMGB1 is present (higher protein size). This is consistent with published data which show that the formation of a complex between CXCL12 and HMGB1 affects the position of the CXCL12 N-terminus [131] and could potentially acquire higher charge if it is more exposed after binding to HMGB1. FR-HMGB1 also slightly shifted towards the compact monomer state (higher abundance of the charge states +13 to +9), but this was minor compared to the shift in CXCL12 charge distribution and could be due to experimental noise.

There is clear biological evidence of the existence of the heterocomplex: I observed regenerative activity *in vivo* with my recombinant HMGB1 preparations in multiple experiments, and my recombinant CXCL12 was capable of binding CXCR4. However, biophysical evidence that the recombinant proteins form a complex appears to be weak and circumstantial. SEC and native MS potentially show changes in the SASA of CXCL12 upon mixture but a small proportion of complex formation which could be artificial. Therefore, I attempted to measure the K_d of this complex directly with quantitative techniques, compared to the qualitative data SEC or MS provide. A very weak interaction could potentially explain these data, where only indirect effects of complex formation can be observed.

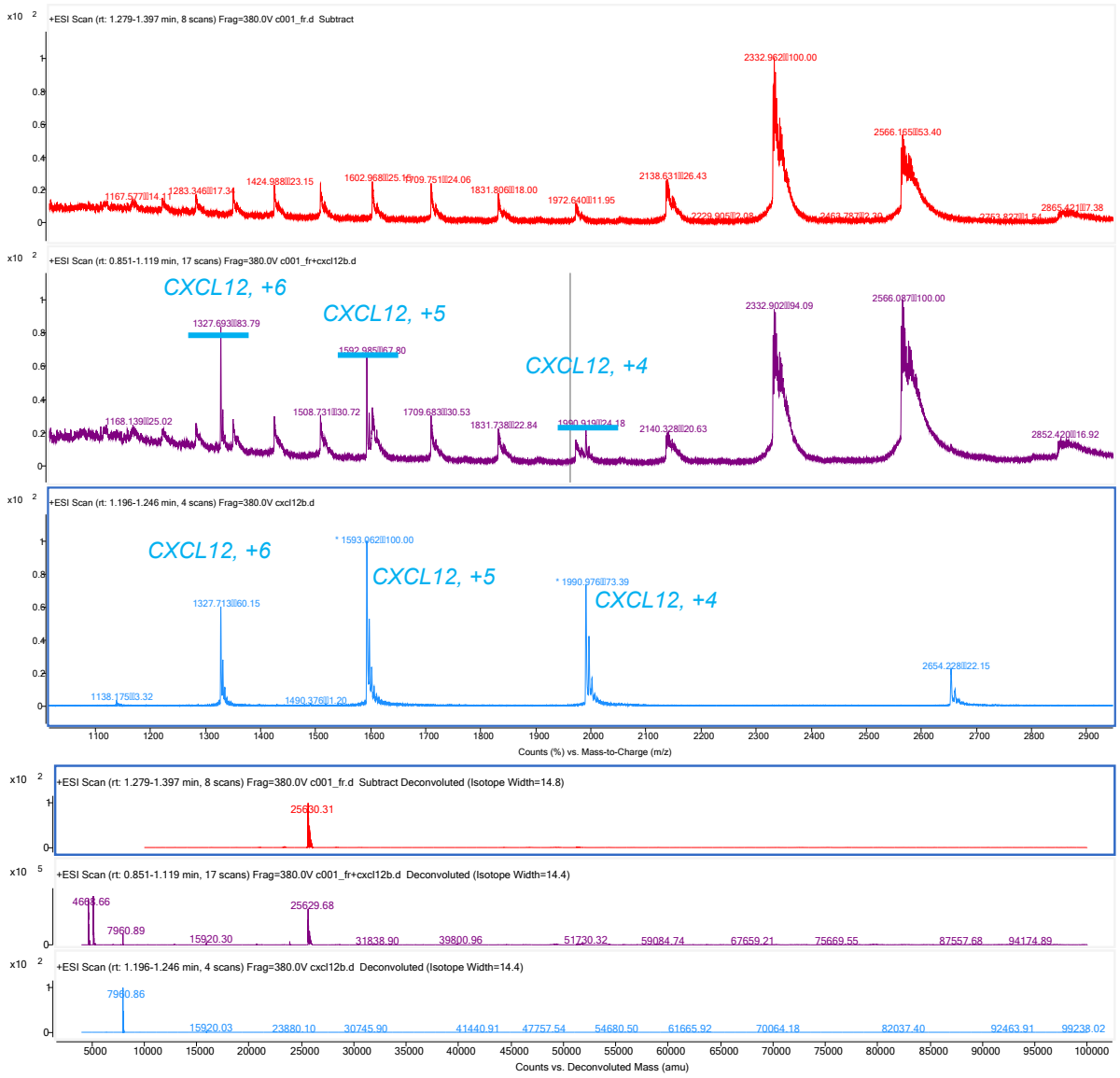


Figure 4.4: Native ESI-QTOF/MS of HMGB1A-c001 (FR, full length), CXCL12A-c021 (wild type) and heterocomplex, in 0.2 M ammonium acetate, pH 6.8

Native M/Z spectra of FR-HMGB1A-c001 (red, 3 mg/mL), heterocomplex (purple, 2 mg/mL HMGB1+ 0.2 mg/mL CXCL12) and CXCL12A-c021 (blue, 0.4 mg/mL). CXCL12 shifts towards higher charge states (larger SASA) when mixed with HMGB1, but no charge states corresponding to the heterocomplex can be observed nor is it visible in the deconvoluted spectra (41549 Da)

4.3.III Screening of FL-HMGB1 and CXCL12 interaction

There are several techniques available to study protein-protein interactions which I had access to. I considered the suitability of each to the analysis of the HMGB1-CXCL12 interaction to confirm whether my recombinant proteins were capable of association in the heterocomplex:

- **Endpoint assays:** Besides the technical drawbacks, these only allow for K_d estimation, although some approaches exist to derive kinetic rates from ITC [231].
    - *Calorimetry*: The large amounts of protein required for isothermal calorimetry (**ITC**), particularly CXCL12, was a significant drawback. Microscale thermophoresis (**MST)** has been applied to the HMGB1-CXCL12 [232] and requires much less protein, but the tagging method was not straightforward for several constructs, particularly as the interaction regions were not well described at the time.
    - *Fluorimetry:* **DSF** was not an option I was unable to get reliable melting curves for CXCL12 (3.5.VI) due to potential incompatibility of this target.
    - *ELISA assays:* These would require specific labelled antibodies which were not readily available. Kinetics from **ELISA** are also biased towards dissociation rate due to the methodology requiring extensive washes (dissociation steps).
    - *Radioligand displacement* was not an option as I did not have access to a laboratory with appropriate facilities. Furthermore, is mostly suited to high-affinity interactions [233] and would add interference from the CXCL12-CXCR4 binding equilibrium.
- **NMR** would be feasible and was actually used later on for in-depth analysis of the interaction (Section 4.4.II), but not for screening due to costs (isotopes) and equipment access limitations (high-field NMR spectroscope access time).
- **Real-time kinetics:** Surface plasmon resonance **(SPR)** has been applied to study the HMGB1-CXCL12 interaction [131]. The only limitation would be sample throughput [234], unless very specialized equipment is available. A similar technique, biolayer interferometry **(BLI)** was also available in-house and allowed for easier set-up of multiple samples in parallel due to lower fluidics complexity and cost I was my method of choice.

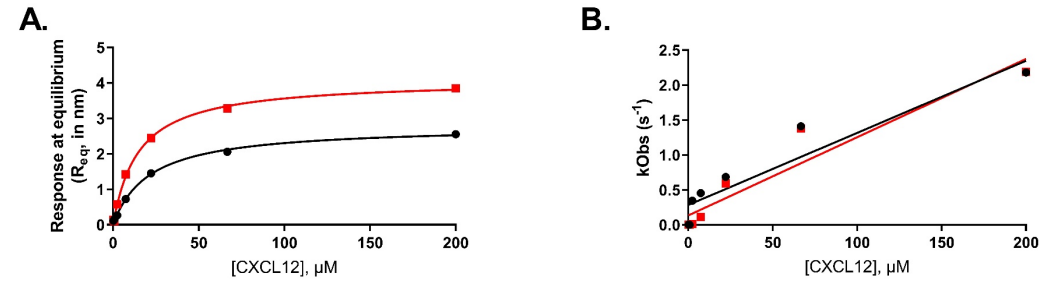
To simulate physiological conditions, but without interference from CXCL12 dimers, I performed these assays in HEPES 10 mM, 150 mM NaCl (isosmotic to physiological) at pH 7.5, to prevent CXCL12 dimerisation[39]. CXCL12 has been reported in these studies to bind non-specifically to several surface chemistries, particularly amine-reactive (CN-), precluding reliable K_d estimation, but other chemistries can be used, and I immobilised HMGB1 via biotinylation to streptavidin (SA) biosensors. These biotinylated constructs (3.4.II.iii) were derived from my initial panel: FL FR-HMGB1 (c011) and 3S (c022). Both constructs were capable of binding CXCL12 (Figure 4.5), and ruling out the CXCL12 binding properties having been lost during purification due to e.g. denaturation, coincident with the SASA values in line with published structures for HMGB1 (3.4.I.ii). Later during my DPhil, two publications reported K_d values for the heterocomplex (via MST) of 4 μM [183] and 77 μM [182]: these encompass the range of K_d in which my data for FL-HMGB1 lie and suggest that native binding properties are indeed preserved.[40]

Observed apparent affinities of FR/3S FL-HMGB1 for CXCL12 were in the low micromolar range, suggesting a relatively fluid interaction, rather than a stable complex, which matches my observations in SEC (Figure 4.3) and native MS (Figure 4.4). However, a cause for concern is that I was fits the kinetic parameters in these experiments gave inconsistent results and had very large errors, both via a horizontal fit of measured k_{on}/k_{off} or a linear fit of the first rate apparent constant k_{obs} to the concentration (non-linear curve), which indicates an incorrect model due to e.g. nonspecific binding preventing this curve from being linear [159] or experimental noise due to each concentration point being measured with a different sensor.

Nevertheless, this was proof that the recombinant proteins had the capacity to bind CXCL12, as in all cases a concentration-dependent relationship could be observed, and measured affinities were in the range of published data. However, the affinity appeared to be very weak, which greatly complicated structural determination via crystallography.

[39] Pilot experiments with 50 mM phosphate buffer added showed no significant alteration in CXCL12 binding to HMGB1.

[40] It is common that different techniques, or even the same technique in different setups, report slightly different measurements in the same order of magnitude for a given interaction.



C.

	Calculation	K_d (µM) Mean	K_d (µM) Individual	k_{On} ($s^{-1}µM^{-1}$)	k_{Off} (s^{-1})	B_{max} (nm)
c011 (FL, FR)	*Sat. fit*	10.1±3.4	14.8±1.1	-	-	4.1±0.1
	kObs fit		11.9±12.2	0.13±0.13	0.01±0.002	-
	Direct rates.		3.6±3.4	0.07±0.07	0.02±0.005	-
c022 (FL, 3S)	*Sat. fit*	16.7±7.8	21.2±1.8	-	-	2.8±0.1
	kObs fit		27.5±15.4	0.1±0.04	0.01±0.002	-
	Direct rates		1.5±0.9	0.28±0.13	0.07±0.031	-

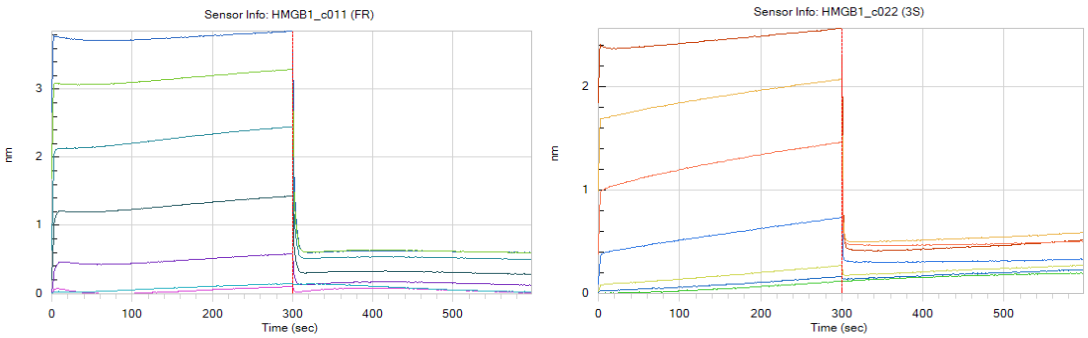


Figure 4.5: BLI kinetics of CXCL12 binding to full length FR and 3S-HMGB1

Biotinylated HMGB1 constructs were immobilised in streptavidin-coated Octet biosensors and dipped in different concentrations of CXCL12 (0-150 for FL-HMGB1, in 1:3 dilutions; 8 sensors were used per construct and were used as independent points in a least squares fit. **A.** Saturation binding plot to equilibrium response (Michaelis-Menten plot), and calculated maximum binding (Bmax, in nm) and Kd. Kd for FR-HMGB1 (**red**) and 3S-HMGB1 (**black**) was equivalent. **B.** Linear least squares fit of the apparent first order rate constant to CXCL12 concentration. **C.** Kinetics summary table; all three calculations do NOT produce equivalent results, indication of potential errors in the methodology: errors in the estimation of kinetic constants are also unacceptably large (> 1 mean) **D.** Interferogram for FL FR-HMGB1A-c011; red fits in A/B. **E.** Interferogram for FL 3S-HMGB1A-c022; black fits in A/B. Sensor coating curves are shown in Suppl. data A3-1-D.
This experiment was also performed with HMGB1A-c027 (9-78) and c028 (94-162), but the low quality of the fits and presence of baseline drift resulted in those data not being evaluated. See Figure 4.12 *for a repeat of these experiments with a more correct methodology.*

4.3.IV Crystallisation trials

Despite my data indicating that the complex was highly unstable with a very weak affinity, I nevertheless attempted to crystallise several HMGB1 constructs in combination with CXCL12 in parallel to the experiments described in Section 4.4. To differentiate potential complex crystals from single-protein crystals, each protein was stained with amine-reactive dyes Alexa-495 and Alexa-555 at a 1:1000 ratio. These dyes stain the primary amines residues in each protein (e.g., lysine side groups, N-termini) and can be imaged in an automated manner.

However, despite my best efforts, no HMGB1 construct crystallised in combination with CXCL12. This is to be expected: the low affinity of the interaction, coupled with the high structural flexibility of HMGB1, and the high solubility of both molecules. CXCL12 remains stable at > 1 mM for extended periods of time, and I was unable to determine an upper solubility limit for HMGB1 experimentally, with solutions in excess of 90 mg/mL (>4 mM) aggregate in SEC or even precipitate. At best, I could observe spherulites stained with both dyes. More typically, either clear drops, separate phases of protein and solvent, granular precipitates, or extensive nucleation of small crystals which failed to increase in size were observed.

Ultimately, the lack of success, combined with the strong evidence that the interaction between HMGB1 and CXCL12 was very weak and the costs of producing CXCL12 in sufficient amounts for crystallography, led me to conclude that this approach was not worth pursuing further. Crystallisation of proteins relies on very strong contacts between regular subunits. HMGB1 is already intrinsically flexible, and the weak affinity of the interaction with CXCL12 would result in a high variability between multimers within the crystallisation solution.

4.4 Sequence-specificity of the HMGB1-CXCL12 interaction

Given my lack of success of crystallising the heterocomplex, I decided to explore alternative approaches. Cryo-electron microscopy (cryo-EM) is better suited to weak complexes due to the ability to identify individual particles/conformation states and providing the option of crosslinking the two components. However, the molecular weight of the final complex – approximately 40 kDa – is below the threshold for current cryo-electron microscopes (> 80 kDa) [235], although smaller structures exist [236], of either single proteins or stable complexes structures exist in this range [237,238]. Unfortunately, HMGB1 is structurally heterogeneous, and together with the weak affinity of the HMGB1-CXCL12 complex this would introduce a great amount of particle heterogeneity and therefore sample complexity. Furthermore, the low molecular weight of the complex precludes the use of gradient molecular crosslinking techniques to overcome the low affinity [239] as is not straightforward to construct a sucrose gradient of sufficient resolution to separate monomeric HMGB1 and CXCL12 from the heterocomplex or CXCL12 dimers.

Rather than trying to obtain the structure of the heterocomplex, and from it deduce the residues that are critical for CXCL12 interaction, my new objective would be to identify the residues and motifs involved in CXCL12 binding and hence for tissue regeneration. HMGB1 has two Box domains, and each has been described as involved in CXCL12 binding. However, the actual binding mechanism is not entirely understood, with two possible binding models:

- The binding of CXCL12 to HMGB1 is cooperative between Box A or Box B, e.g., binding of one CXCL12 monomer to one HMG Box enhances binding of a second CXCL12 monomer to the other.
- The described binding interface for HMGB1 Box A binding to CXCL12 (Ala 16-Trp 38) and Box B (Phe 101-Lys 140) is incomplete and Box A 8-78, or Box B 94-162, are missing residues critical for CXCL12 binding leading to their decreased capacity to bind CXCL12. This interface was described by NMR [131,232], but no other orthogonal experimental validation exists in the literature besides computational modelling.

4.4.I Peptide arrays

Peptide array, also known as SPOT arrays [158], rely on the capacity of isolated amino acid motifs of a protein to recruit the ligand and can detect peptides binding with K_d up to 100 μM. Application to the HMGB1-CXCL12 complex revealed a mirrored pattern between HMG Box domains, albeit with weaker intensity in Box B (Figure 4.6 A, B). Within both Box A and Box B there are two groups of peptides which bind CXCL12, found in similar positions within each Box. The first of these peptides covers approximately residues Asp4-Ser41 within Box A and Asp90-Gly122 in Box B; in both HMG Boxes this overlaps the binding site for glycyrrhizin, which inhibits the regenerative effects of the HMGB1-CXCL12 heterocomplex [25]. Interestingly, this region also extends beyond the described structural domain for Box A (Pro8-Ile78) and Box B (Pro94-Arg162) in the Uniprot database and includes a portion of the N-terminus which was not included in my initial Box A and Box B constructs.

A second peptide is present in the C-terminal region of the HMG Box (Lys65-Gly82 in Box A, Lys151-Asp168 in Box B), the signal of which is much weaker in Box B compared with A. Both CXCL12 binding peptides in each HMG Box extend beyond the described interface which only encompasses the helical segments of both HMG Boxes [96,182,232]. Assuming that both Box A and Box B are capable of independently binding CXCL12, it follows that the interaction regions are likely to be similar between the two. The weaker binding of CXCL12 to peptides within Box B than to those within Box A would match my BLI data where Box B had decreased maximum binding to CXCL12 compared to Box A, and potentially lower affinity. Other possibilities include a faster association and dissociation of CXCL12 to Box B than Box A, or the binding to Box B being more reliant in the structure than the peptide sequences alone. From these peptides, I performed an in-depth scan to identify key positions in CXCL12 binding (Figure 4.6 C, D). Each residue within a target peptide sequence was mutated to alanine, a neutral side chain. If the mutated residue were involved in CXCL12 binding, the intensity of the peptide spot containing the mutation would be expected to decrease. Both Box A and Box B involve residues in their flexible N-termini (D-P-X-X) and C-termini. Several of the residues in Box B appeared to show higher binding upon mutation to alanine, although this is may have been due the low initial signal. Nevertheless, I considered them as potential hits.

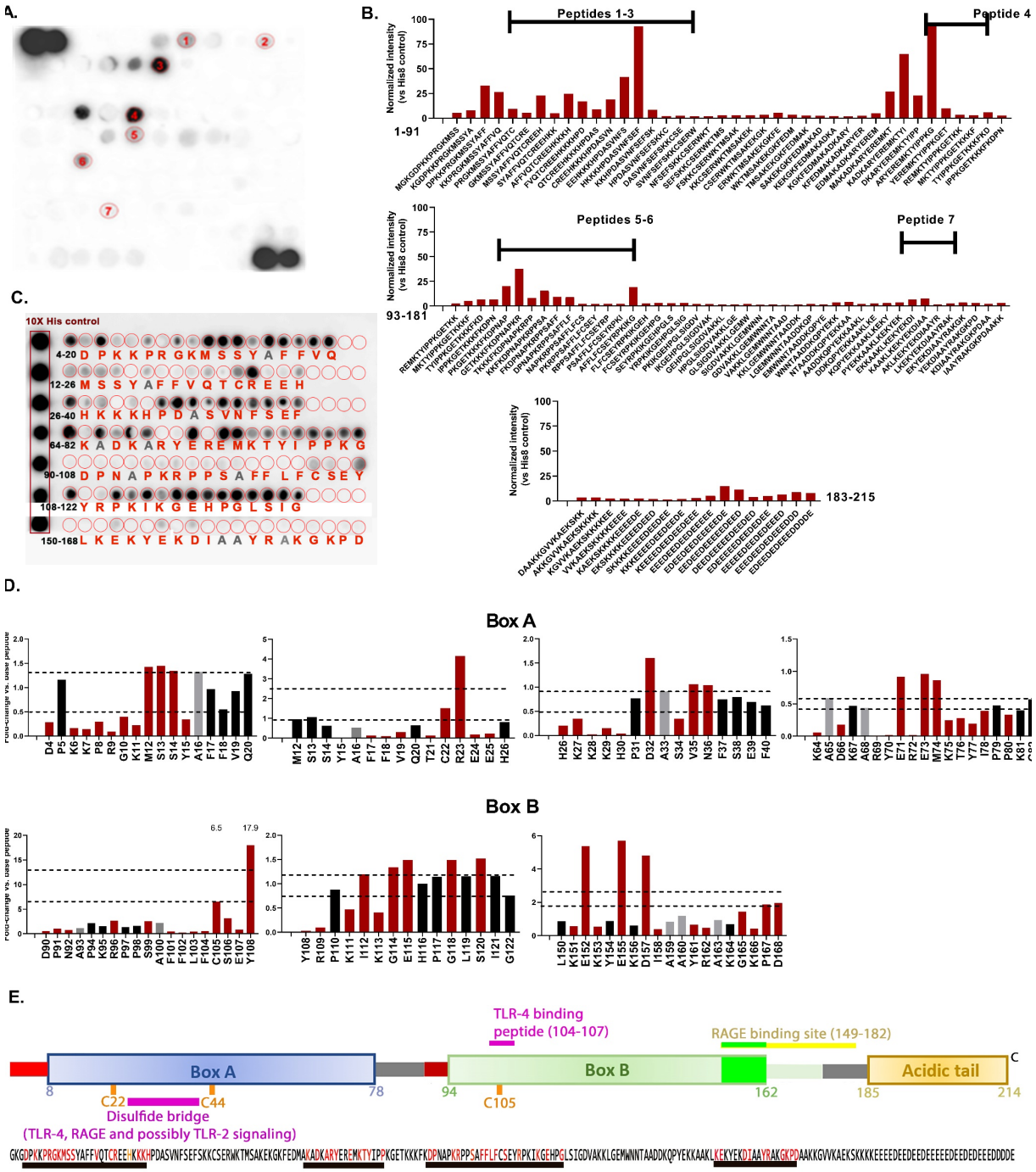


Figure 4.6: Alanine scan of HMGB1 peptides interacting with CXCL12 across Box A and Box B

A. Peptide array of full-length HMGB1. "+" positions correspond to positive control 10-His peptides; the rest of the peptides comprise HMGB1 15-mers shifted 2 residues in succession towards the C-terminus. The membrane was exposed to 1 μM CXCL12-His6 for 24h. Protein binding was detected by incubation with anti-His-HRP conjugate chemiluminescence; this experiment was performed twice. High intensity spots represent bound CXCL12 (HRP activity). **B.** Image analysis of spot intensity compared to control spot intensity (+); peptides selected for further analysis have been highlighted. The numbers indicate roughly the HMGB1 domain represented in each graph; box A (1-89), box B (90-175), or the acidic tail (185-215), numbered from Met 1. **C, D:** From each peptide identified in (A) , single positions were mutated to alanine and the changes in intensity with respect to the control were analysed with Fiji ImageJ. Mutations of alanine positions (identical to control peptide) nevertheless had signal variation; these were used to define a threshold to consider a position relevant for CXCL12 binding (random technical variation). Residues with changes in intensity of bound CXCL12 upon alanine mutation beyond those seen in alanine residues are shown in red. As the signal in peptides 5 and 7 (box B) was quite weak, it is possible that false positives/negatives exist amongst these. **E.** Schematic of FL-HMGB1 with sequence. CXCL12 binding peptides are underlined in black, and key residues observed in the alanine scan are highlighted in red.

A reverse peptide array was therefore performed to validate these results (Figure 4.7). As the sequence specificity of CXCL12 was not of concern to these studies, this experiment was designed as qualitative analysis of the effect of the flanking sequences in each Box domain, to assess whether the Box domains alone could interact with the same peptides as a full-length HMGB1 protein.

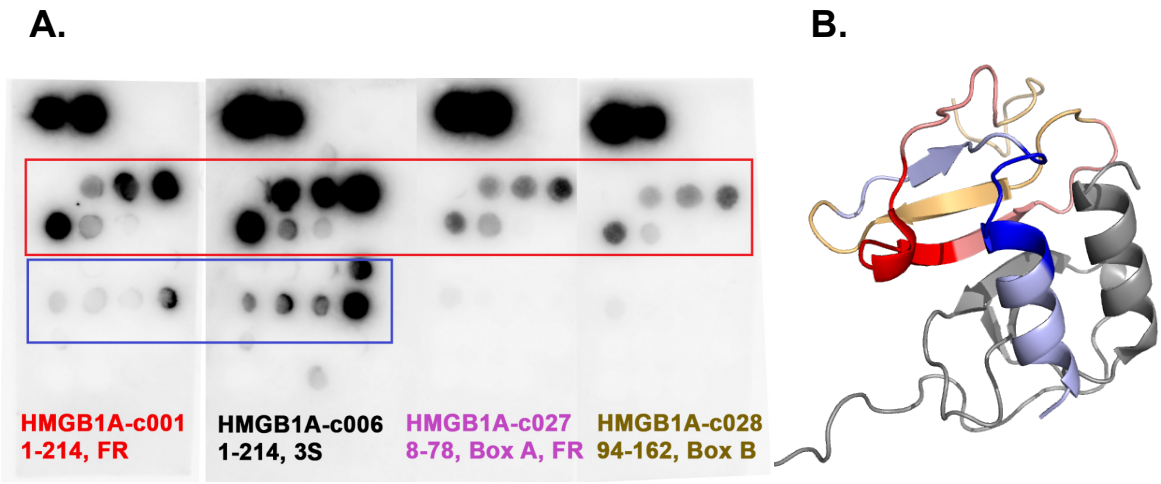


Figure 4.7: Peptide array of CXCL12 peptides interacting with HMGB1

A. Peptide array of full-length CXCL12. "+" positions correspond to positive control 10-His peptides; the rest of the peptides comprise CXCL12 15-mers shifted 2 residues in succession towards the C-terminus. The membrane was exposed to 1 μM HMGB1(FR or 3S)-His6 (1-214), BoxA-His6 (8-78 and BoxB-His6 (94-162) for 24h . Bound protein was detected by anti-His-HRP conjugate chemiluminescence. A peptide of CXCL12 interacting with full length HMGB1 cannot interact with either Box A or Box B alone (in **blue**), confirming the requirement of the N-terminal segment of each Box domain (particularly, D4 in box A/D90 in box B): intensity of the spots pertaining to the common CXCL12 peptide (red) is also markedly decreased upon binding to the Box domains alone when compared to FL-HMGB1. *Binding to 3S seems to be of higher intensity than that to FR; this is likely due to protein oxidation during the assay, although this was not quantified due to the low concentration of protein used being unsuitable for ESI/TOF MS. BLI data from* Figure 4.13, *however, do suggest a lower off rate of CXCL12 from 3S than from FR-HMGB1.***B.** CXCL12 dimer (PDB 2J7Z) with the regions binding HMGB1 highlighted. In red: shared binding region; darker red indicates the common sequence across all spots. In blue: non-shared binding region; darker blue indicates the common sequence across all spots. In grey: second monomer in the structure.

FR and 3S- HMGB1 bind CXCL12 peptides over the full β-sheet of CXCL12, whereas HMGB1 8-78 (Box A alone) or HMGB1 94-162 (Box B alone), the minimal HMG Box constructs [240] bound to the N-terminal portion of this region but not to the C-terminal segment of the β-sheet. The lack of this interaction could potentially destabilise the binding of HMGB1 to CXCL12. This does not completely disprove the hypothesis of the binding to CXCL12 being cooperative between Box A and Box B, rather than independent for each Box domain, as the affinity is lower for each Box domain alone and this could be due to either these missing regions or a cooperative binding mechanism requiring two Box domains. However, the binding regions being mirrored between boxes would suggest the former.

4.4.II Nuclear Magnetic Resonance (NMR)

Peptide arrays are restricted to sequence-specific, structure-independent effects as the peptides in the membrane are 15-22-mers and at most can only adopt short secondary structures. The short α-helices present in HMGB1 (11/10/22 residues), which turns every 4 residues, could potentially form with the peptides in the membrane, but I felt that additional experiments which specifically involved the contribution of the tertiary protein structure to binding would be needed to confirm these results.

Initially, my intention was to use NMR to resolve a full structure of HMGB1-CXCL12 complex. NMR experiments have been performed for several HMGB1 complexes in the literature [96,131,183]. NMR has been a pivotal technique in the HMGB1 literature for identifying both its intra- [46,50,74,241] and inter-molecular interactions [96,131,132,232]. However, full structural determination would require lengthy experiments at high magnetic field such as NOESY (Nuclear OVerhauser SpectroscopY) and TOCSY (TOtal Correlation SpectroscopY). This could lead to HMGB1 oxidising during the assay and affecting the data. Furthermore, the molecular weight of the heterocomplex makes it too large for accurate structural determination. Whilst NMR size limits have greatly increased in the past decade [242,243], thanks to the advent of techniques such as high-field TROSY [244], these have been applied to stable complexes or single proteins, rather than to fast exchange rate interactions such as HMGB1-CXCL12.

A third problem lies in the dynamic nature of the complex. The original studies in the heterocomplex describe CSPs in both CXCL12 and HMGB1 spectra, whereas the 2019 study measured peak height (I/I0) changes; both papers show significant peak broadening in either CXCL12 or HMGB1 spectra due to the high exchange rate [96,131][41]. Together with my affinity measurements (μM range), I concluded that determination of a full structure by NMR would be too costly in both time and resources, and not feasible during my DPhil.

[41] None of these experiments have been deposited in the BMRB prior to publication, nor are the chemical shift tables are shown or referenced in any source. Hence, I had to perform my own data collections. Furthermore, spectra for ^{15}N labelled HMGB1 bound to CXCL12 are mentioned in both 2012 [96] and 2017 [131], but never shown.

Instead, I decided to verify my peptide array data using NMR: the residues within HMGB1 which I found to interact with CXCL12 in the peptide array should also experience changes upon CXCL12 addition. If the D-P-X-X peptide is involved in binding to CXCL12, the changes in the NMR spectra of an HMG Box domain with that peptide present should be substantially different than those observed in an HMG Box without, and potentially changes in the signal residues within it should also appear.[42]

I used HMGB1A-c028 (FR, Box B 94-162) and c038 (FR, Box B 89-174) to represent native HMG Box binding domains with and without the flanking regions via $^{1}H^{15}N$-HSQC (Figure 4.8). Using Box B alone would provide information on the binding to a native HMG Box without need for reducing agents, assuming CXCL12 binding to each HMG box is independent. This experiment was performed in identical buffer conditions to the BLI experiments in Figure 4.5.

I could observe CXCL12-concentration dependent CSP and I/I0 in several residues identified in the peptide arrays, changes absent in the control samples (Figure 4.8). The median CSP in Box B constructs are much greater when a full HMG Box is present (0.005/0.017) and several residues in the flanking regions, particularly the C-terminus, showed significant CSP or intensity changes upon CXCL12 titration in HMGB1-c038, reflecting their involvement in CXCL12 binding. Other residues flagged in the peptide array, such as F101, D157, I158, show CSP changes in the full HMG Box but not in the helical-only construct. This would suggest that the presence of the flanking regions increases the affinity of Box B for CXCL12 and allows additional contacts to be formed between both proteins. Several residues identified as involved in CXCL12 binding, in both peptide arrays (Figure 4.6) and published literature [182,232] show no changes in either construct when titrated with CXCL12. (e.g., A100, I112, L119, A136, Y154). In particular, the signal of A100 has been shown to disappear after CXCL12 addition [232], although previous publications from the same group do not mention this phenomenon [96,131]. [43]

[42] The optimisation of HMGB1 for NMR experiments was largely trial and error, and therefore will not be listed. Only experiments which yielded interpretable data are shown, with the parameters and resonance tables being collected in Appendix A4.

[43] These experiments were performed in infra-physiological pH and ionic strength (20 mM sodium phosphate buffer, 20 mM NaCl, pH 6.0), which improves both NMR signal and protein binding. However, no indication of pH monitoring is shown in the study reporting the signal from Ala 100 to disappear, where HMGB1 and CXCL12 were in different pH and salt concentration buffers. Due to these differences in methodology and controls, it is likely that they caused these differences in detected CSP changes.

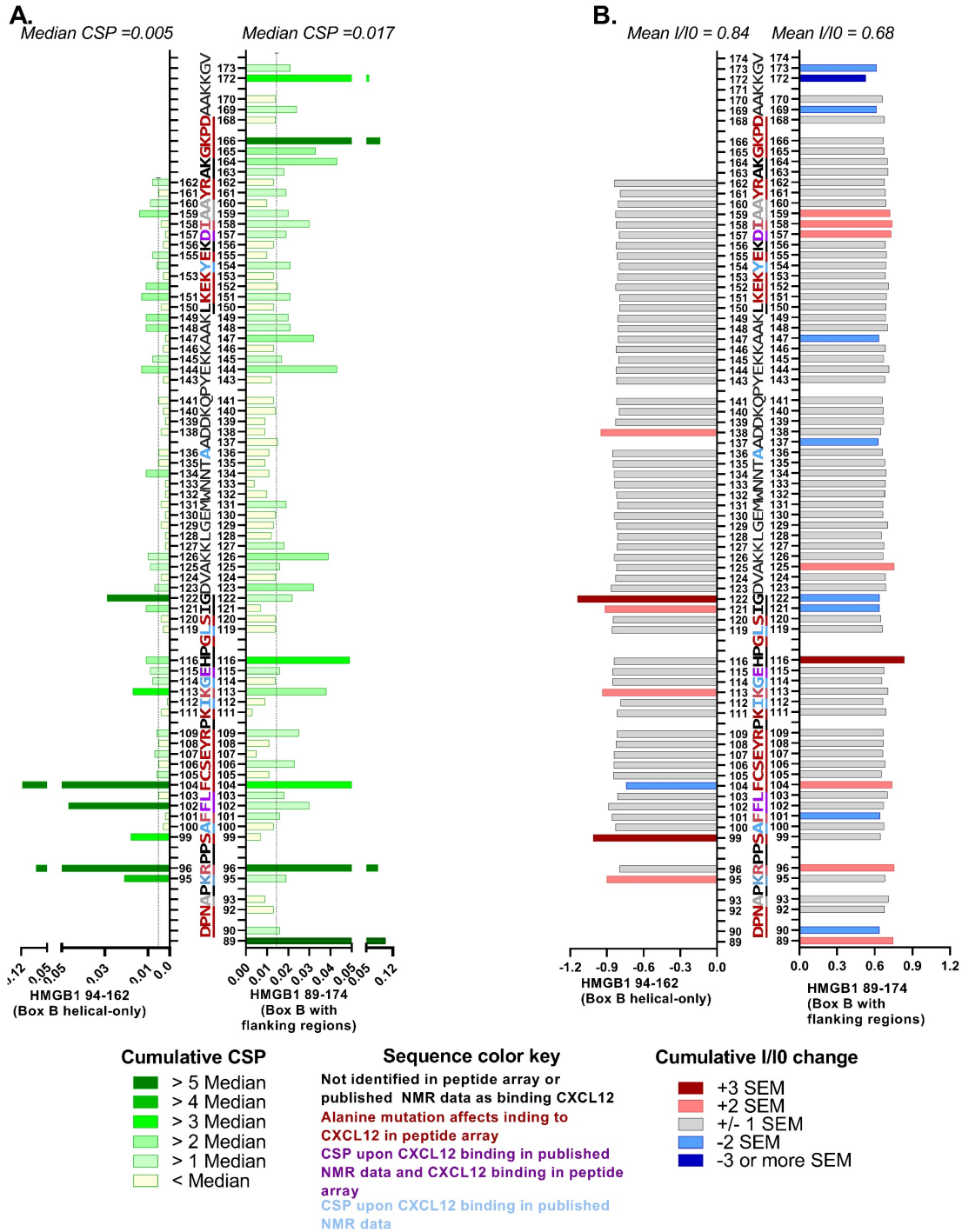
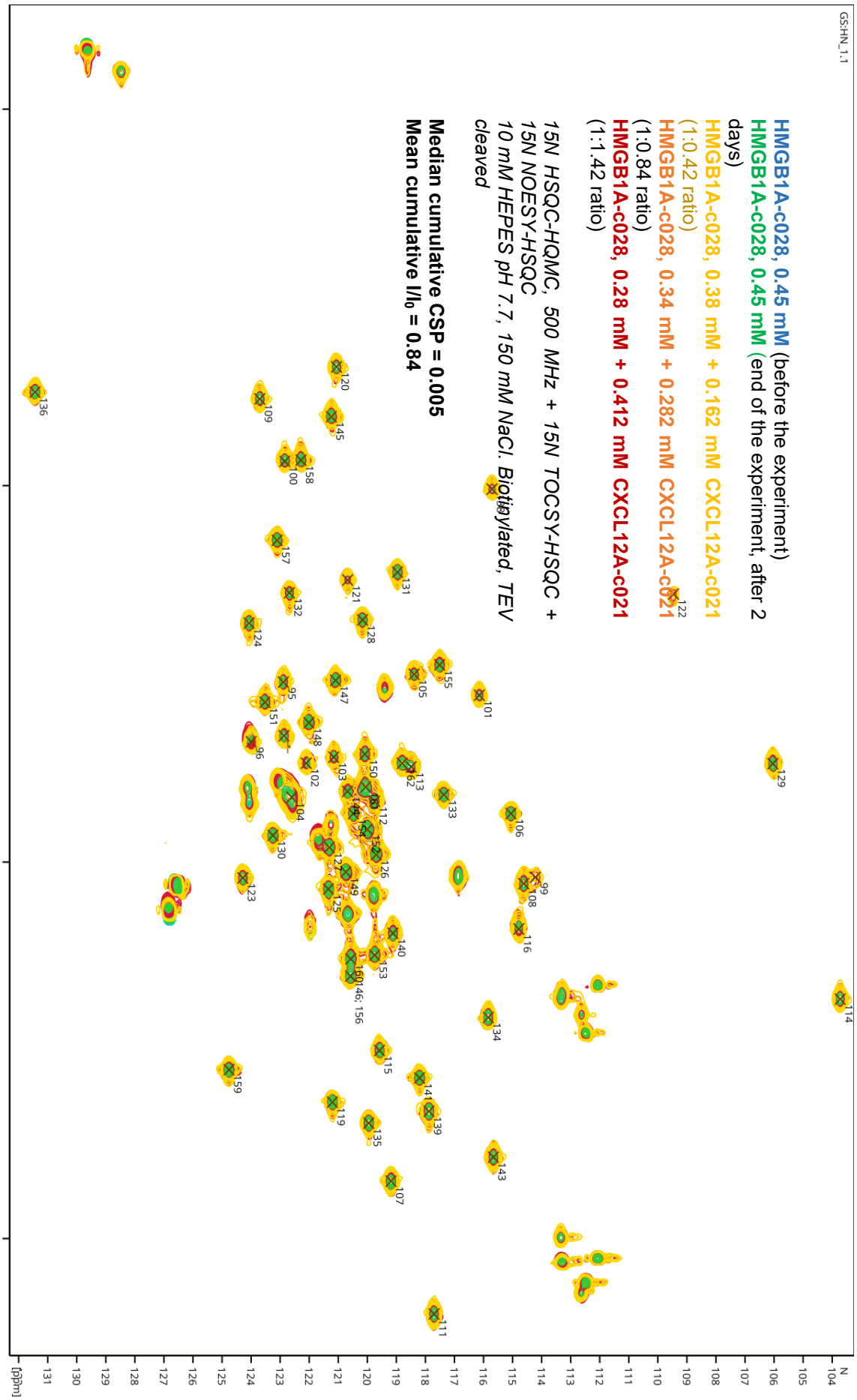


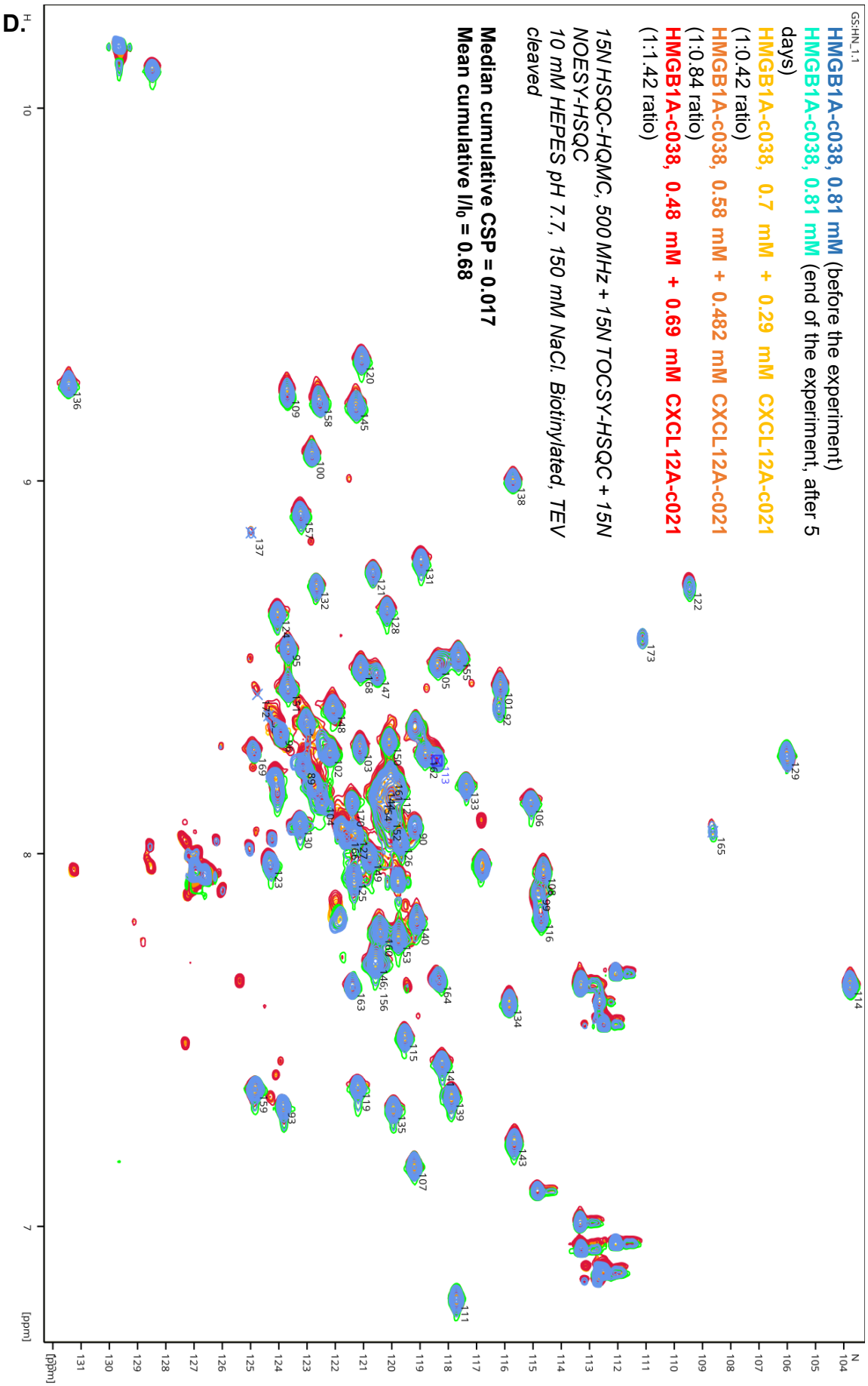
Figure 4.8: CSP and intensity change in HMGB1 Box B constructs with and without flanking regions upon CXCL12 binding

A. Cumulative CSP of either helical-only biotinylated Box B (94-162, HMGB1A-c028) or complete Box B (89-174, HMGB1A-c038) after titration with CXCL12 (0.42, 0.84 and 1.42 molar equivalents), calculated over several HSQC spectra. A sample with no CXCL12 measured at the last timepoint was used to account for random CSP and intensity drift during the experiments in the curve fitting on CCPNMR software. Intensity of the green colour in the graph indicates higher relative CSP. The sequence of each HMGB1 construct has been overlaid with the residue number; an empty column (no number) represents residues which could not be mapped in the parallel 3D 1H-15N HSQC/NOE/TOCSY experiments. The sequence of the residues corresponding to each box is in the middle of each condition, coloured as per: Light blue, residues previously reported in the literature as involved in CXCL12 binding; **red**, residues weakly involved in the peptide array; **purple**, residues involved according to both published literature and peptide array. **B.** Cumulative peak height change of (A); red indicates I/I0 change over 1 SEM of all residues, blue decreases over -1 SEM. . **C, D (next 2 pages):** 15N HSQC-HQMC peak diagrams for all samples, in 10 mM HEPES 150 mM NaCl pH 7.5 buffer Protein concentrations are indicated in the spectra overlay. **C:** HMGB1A-c028 (2 day experiment), full details in Suppl. data A4-1-A; **D:** HMGB1A-c038 (6 day experiment; minor degradation occurs after day 4), full details in Suppl. data A4-1-C.

C.


HMGB1A-c028, 0.45 mM (before the experiment)
HMGB1A-c028, 0.45 mM (end of the experiment, after 2 days)
HMGB1A-c028, 0.38 mM + 0.162 mM CXCL12A-c021 (1:0.42 ratio)
HMGB1A-c028, 0.34 mM + 0.282 mM CXCL12A-c021 (1:0.84 ratio)
HMGB1A-c028, 0.28 mM + 0.412 mM CXCL12A-c021 (1:1.42 ratio)
15N HSQC-HQMC, 500 MHz + 15N TOCSY-HSQC + 15N NOESY-HSQC
10 mM HEPES pH 7.7, 150 mM NaCl. Biotinylated, TEV cleaved
Median cumulative CSP = 0.005
Mean cumulative I/I0 = 0.84
GS:HN_1.1
H [ppm]
N [ppm]

D.


GS:HN_1.1
HMGB1A-c038, 0.81 mM (before the experiment)
HMGB1A-c038, 0.81 mM (end of the experiment, after 5 days)
HMGB1A-c038, 0.7 mM + 0.29 mM CXCL12A-c021 (1:0.42 ratio)
HMGB1A-c038, 0.58 mM + 0.482 mM CXCL12A-c021 (1:0.84 ratio)
HMGB1A-c038, 0.48 mM + 0.69 mM CXCL12A-c021 (1:1.42 ratio)
15N HSQC-HQMC, 500 MHz + 15N TOCSY-HSQC + 15N NOESY-HSQC
10 mM HEPES pH 7.7, 150 mM NaCl. Biotinylated, TEV cleaved
Median cumulative CSP = 0.017
Mean cumulative I/I0 = 0.68
H [ppm]
N [ppm]

I also evaluated the same interaction with a full-length HMGB1 molecule, to compare Box A and Box B within the same molecule to determine whether they have equivalent CXCL12 binding properties. Due to the long times required for this experiment (> 5 days) I selected a 3S-HMGB1 construct to prevent sample oxidation in the NMR tube. I used a tail-less HMGB1 construct (HMGB1-c007, 1-184) to ensure I would have sufficient protein to achieve good signal-to-noise ratios[44].

When this experiment was repeated with 3S-HMGB1 1-184, CSP changes were very modest due to limitations in the signal/noise ratio and the solubility of CXCL12 precluding concentrations over 1:2 molar ratio. Overall, HMG Box B showed a weaker CSP change for its residues involved in the interaction than Box A, which could be representative of a more fluid equilibrium of CXCL12 binding in this domain (faster association and dissociation rates, not necessarily different affinity). Most of the residues I observed as having significant chemical shift changes upon CXCL12 addition within Box A were also contained in the regions within it identified in the spot array, reinforcing the hypothesis of a mirrored pattern between Box A and Box B.

Based on these NMR data, I decided to reclassify several residues weakly flagged in the alanine scan as CXCL12 binders due to the lack of response in NMR. Whereas a residue indirectly involved in binding could not be flagged in the alanine scan (sequence-specificity) but be involved according to NMR (structure and environment effects), the opposite is unlikely. Residues involved only in NMR but not in the peptide array, such as A136 in the published experiments, or K127, G129 and M131, are likely experiencing changes due to either the perturbation of neighbouring residues or interaction with CXCL12 via the main chains (sequence-independent).

[44] 20+ litres of FL-HMGB1 culture in M9 media would be required to purify enough protein to reach 0.2-0.3 mM in the lumen of a 5mm Shigeimi tube (300 µL); 4-5 L yield 0.5 mM for a tail-less construct. I was also limited to a single titration point due to magnet time availability, as I required a higher field (750 MHz) spectrometer. I was also limited by CXCL12 concentration, due to the necessity of reaching a 1:2 HMGB1/CXCL12 ratio to have 1:1 HMG Box domain/CXCL12 ratio; I was unable to concentrate CXCL12 over 1.8 mM.

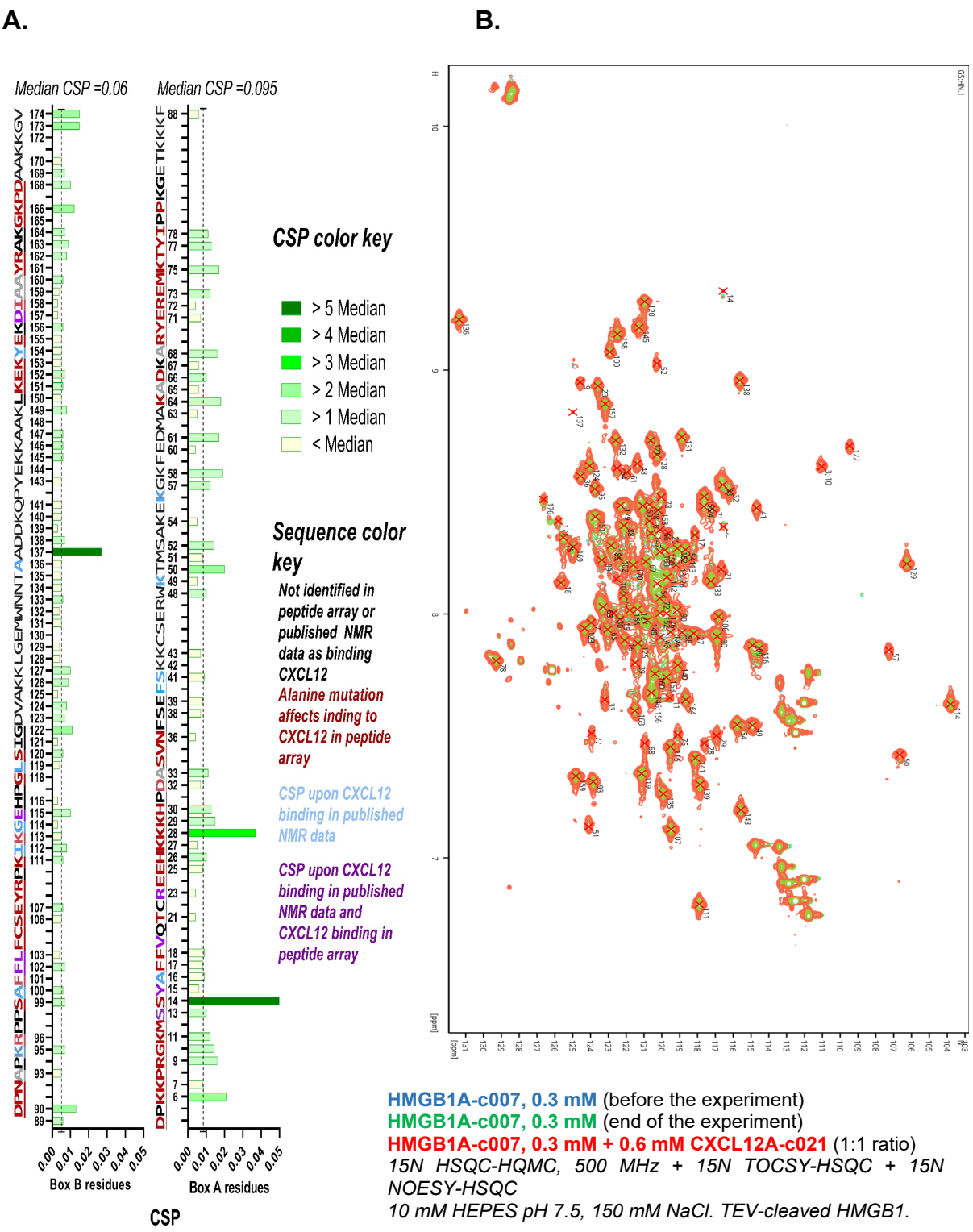


Figure 4.9: Figure 4.9: CSP changes in HMGB1 1-184 (3S) upon CXCL12 binding

A. Cumulative CSP of HMGB1 3S 1-184 (HMGB1A-c007) upon addition of 1:2 molar equivalents of CXCL12 in one step (1:1 HMG Box to CXCL12 ratio. Box A and Box B residues have been considered separate molecules for the purposes of median CSP calculation. Intensity of the green colour in the graph indicates higher relative CSP. The sequence of each HMGB1 construct has been overlaid with the residue number; an empty column (no number) represents residues which could not be mapped in the parallel 3D 1H-15N HSQC/NOE/TOCSY experiments. The sequence of the residues corresponding to each box is in the middle of each condition, coloured as per: **Light blue**, residues previously reported in the literature as involved in CXCL12 binding; **red**, residues weakly involved in the peptide array; **purple**, residues involved according to both published literature and peptide array. **B.** 15N HSQC-HQMC peak diagrams for all samples, in 10 mM HEPES 150 mM NaCl pH 7.5 buffer Protein concentrations are indicated in the spectra overlay; full details in Suppl. data A4-1-B.

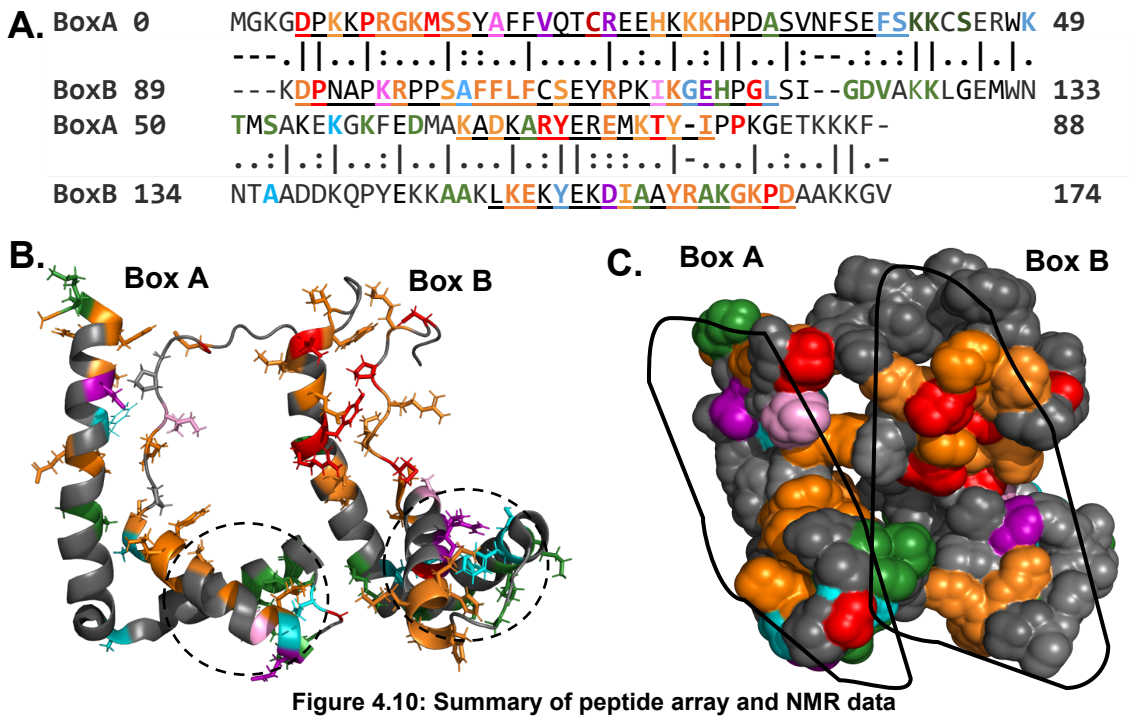


Figure 4.10: Summary of peptide array and NMR data

A. Alignment of Box A + linkers (1-88) with Box B + linkers (89-174). Vertical lines designate conserved positions, and double dots similar substitutions. Underlined: peptide regions binding CXCL12 from the coarse peptide array. In **red**: residues flagged in the alanine scan as involved in CXCL12 binding which could not be checked in NMR. In **orange**: residues flagged in the alanine scan and showing CSP changes upon CXCL12 binding in NMR. In **cyan**: residues not flagged in my experiments but described in the NMR literature as contributing to CXCL12 binding [182]. In purple: residues flagged in both my peptide array experiments and the published NMR data, but not my own NMR experiments. In **green**: residues flagged only on NMR experiments, which can either directly bind CXCL12 or be affected by binding to nearby residues. In pink: residues flagged in both my NMR experiments and the published NMR data. Proposed CXCL12 binding peptides are underlined. Alanine residues can only be identified via NMR. **B.** Surface representation of the structure of FR-HMGB1 1-166 (2YRQ), with the residues on (A) highlighted, following the same colour code. Side chains of all involved residues have been shown. **Black dashed circles** surround the regions where glycyrrhizin binds each HMG Box. **C.** Surface representation of (B)., highlighting the surfaces defined by residues involved in CXCL12 interaction in the concave region of each HMG Box (black full lines)

When the peptides defined from the SPOT array and corrected by NMR, as well as the residues identified only in NMR as affected by CXCL12 binding are overlaid over the structure of HMGB1 1-164, Box A + Box B (PDB 2YRQ; Figure 4.10), and the aligned Box A and Box B sequences, it is apparent that the interaction surface defined is homologous between both. Two interaction sites exist within each Box domain: the C-terminal segment of the third alpha helix and the loop between the first and second, which also includes the glycyrrhizin binding site [132], an inhibitor of HMGB1-CXCL12-CXCR4 mediated tissue repair [25]. Specific residues loosely homologous between each HMG Box and following a rough i+3/4 series in each Box, resulting in their side chains aligning across the α-helices. The side chains of the residues flagged in both peptide array and NMR all localise to the concave surface of the HMG Box domain; conversely, residues identified only via NMR typically arrange their side chains outside of this surface.

4.4.III Kinetic validation of CXCL12 binding to different HMG Box domains and redox states

I observed a lower binding response to CXCL12 binding in Box B constructs with only the core helical domain in BLI when compared to FL-HMGB1 or a full-length Box B construct, but I had been unable to calculate valid kinetic rates to confirm this. If my hypothesis were correct, addition of the peptides would restore native binding capacity. It had been brought to my attention that my previous experimental setup (Section 4.3.III) could be improved to yield more reliable data. The use of one sensor per concentration of ligand could result in more noise. As I could not calculate reliable kinetic rates with the previous methodology, I modified my setup and performed these assays with a single sensor per HMGB1 construct (serial-dip; Figure 4.11), which then would be exposed to a rising concentration of CXCL12, with replicates performed over the same sensor in cycles.

This method introduced other complications. I observed CXCL12 saturation of the sensor due to nonspecific (irreversible) binding to the uncoated sensors, an effect which I could not reproduce with the older batch of sensors used in 4.3.III. Constructs with low binding (e.g., Box B 94-162) would thus have an apparent increase in interference (nm) upon CXCL12 dissociation after subtracting the baseline sensors, which is not possible. After optimisation of the experimental parameters, 1% BSA, 0.05% Tween-20, 0.02% sodium azide were added to CXCL12 assay buffers to minimise nonspecific binding; coated sensors were equilibrated in this buffer for 20 min. Additionally, assay time was increased in order to allow for reaching of steady-state and more reliable kinetic rate measurements matching the saturation fits. In addition, serial-dip methods extend assay times, potentially leading to HMGB1 oxidation; however, TCEP is incompatible with CXCL12.[45] To better improve reliability, I shifted to SAX (High-precision streptavidin) sensors, which are rated for quantitation and are more strictly controlled than regular SA sensors, to be able to compare B_{max} values.

[45] HMGB1 remains reduced in neutral solutions for at least a few hours [74]; oxidation in physiological fluids or cell culture supernatant is much faster. In my experience however, samples in test tubes can take up to > 1 week to oxidise at RT without oxidising agents, albeit this timeframe was not consistent.

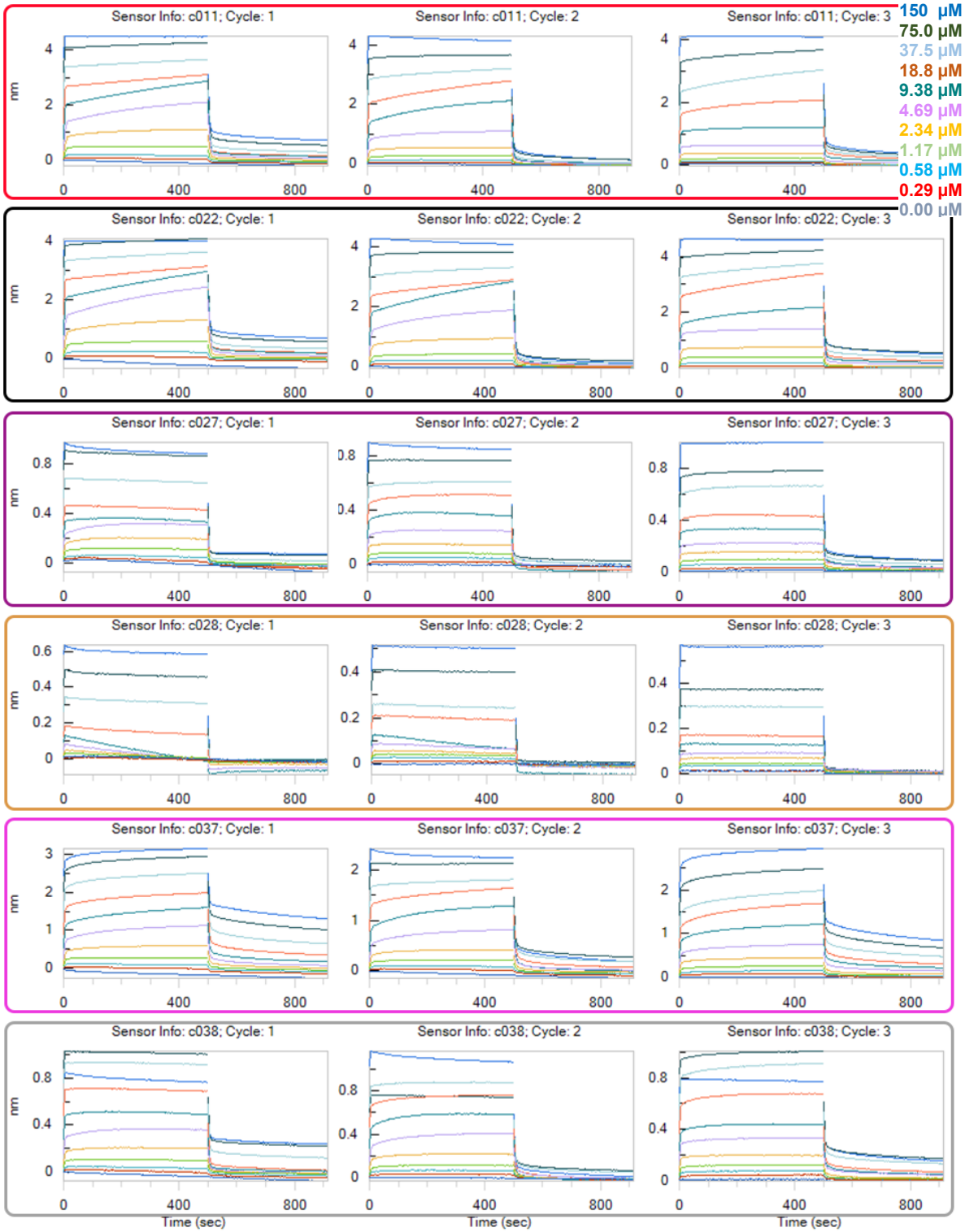


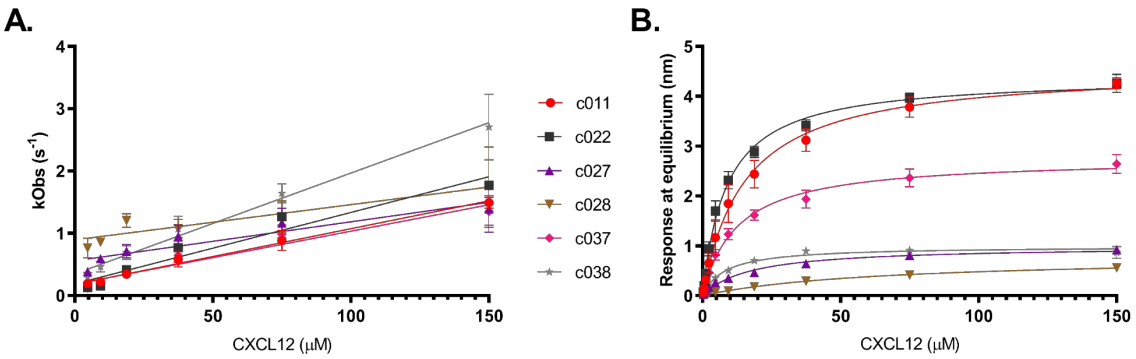
Figure 4.11: Interferograms in BLI of HMGB1 constructs' binding to CXCL12

Biotinylated HMGB1 constructs were immobilised on streptavidin-coated Octet biosensors and dipped in rising concentration of CXCL12. Interferograms are coloured according to CXCL12 concentration (key in top right). Each set of three replicates (cycles) for a given sensor is surrounded by a coloured overlay according to construct. Binding curves to sensor of HMGB1 shown in Suppl. data A3-1-C.
Red: FR FL-HMGB1 (c011), **black**: 3S-FL HMGB1 (c022), **purple:** FR-HMGB1 Box A 8-78 (c027), **brown:** FR-HMGB1 Box B 94-162 (c028), **pink:** FR-HMGB1 Box A 1-88 (c037), **grey**: FR-HMGB1 Box B 89-174 (c038).

Kinetic analysis shows that the presence of the flanking flexible regions in each box greatly increases the binding affinity of HMG Box domains to CXCL12 (Figure 4.12). HMGB1-c027 Box A (8-78) and HMGB1-c028 Box B (94-162) constructs had both decreased binding capacity (B_{max}) and weaker affinity (increased K_d) than constructs HMGB1-c037 Box A (1-88) and HMGB1-c038 Box B (89-174), in which the peptides are present: these two bound CXCL12 with equivalent affinities to FL-HMGB1 3S (c022) and FR (c011), which both had similar affinity for CXCL12. The relationship between K_d values was the same whether these were derived from the saturation fit to concentration, the first order apparent constant (k_{obs}) fit to concentration, or derivation from association and dissociation rates directly measured by the instrument.

The differences in relative affinity (K_d) between the HMG Box constructs without flanking regions and all others was due to a much faster dissociation rates (k_{Off}) of CXCL12 in HMGB1A-c027 and c028 compared to all other constructs, whereas association rates (k_{On}) remained constant across constructs excepting HMGB1A-c038 which had a higher association rate. However, maximum binding response (B_{max}) of Box A c037 with the peptide was significantly higher than that of Box B c038, and the K_d values between both were similar[46]. This is due to the dissociation rate from full-length Box B being also higher than from Box A. This is likely to represent a much more fluid binding, where CXCL12 is recruited more quickly to Box B, but is also less prone to remain bound to it than to Box A. This is consistent with both my NMR data where CSP for CXCL12 binding residues in Box B was of a lower magnitude than those in Box A, and the published data showing that CXCL12 is likely interchanging between Boxes [96], likely due to the different kinetic rates leading to the exchange. A more fluid binding would result in a less defined change in the chemical environment as it switches more often between both bound and unbound state but result in an overall equivalent affinity as the changes between Box A and B are within the same order of magnitude. The functional significance of this is unclear, but this is likely a consequence of multiple functions for HMGB1 leading to different “compromises” in the sequence of each HMG Box, affecting CXCL12-binding properties.

[46] Response during sensor coating was 2-fold lower for Box B when compared to Box A or FL-HMGB1 constructs (Appendix A3) suggesting lower protein density on the sensor surface for Box B than for Box A as the MW is nearly identical. This could be due to their different surface area charge profiles (Figure 3.21) affecting solvation and allowing for more compact packing of Box A in the sensor surface, i.e., there are more CXCL12 binding sites in the sensors coated with Box A than those with Box B.



C.

	Calculation	K_d *Mean*	K_d *Individual*	k_{On} ($s^{-1}\mu M^{-1}$)	k_{Off} (s^{-1})	B_{max} *(nm)*
c011 (FL, FR)	*Sat. fit*	19.1±2.1	17.8±1.7	-	-	4.67±0.07
	kObs fit		19.9±5.9	0.009±0.001	0.18±0.05	-
	Direct rates.		20.9±4	0.009±0.002	0.18±0.01	-
c022 (FL, 3S)	*Sat. fit*	11.4±1.6	9.2±0.7	-	-	4.43±0.09
	kObs fit		16.3±5.2	0.012±0.001	0.19±0.06	-
	Direct rates		11.7±1.5	0.013±0.002	0.15±0.02	-
c027 (FR, 8-78)	*Sat. fit*	38±6.9	18±1.9	-	-	1±0.03
	kObs fit		88.9±19.3	0.007±0.002	0.56±0.06	-
	Direct rates		18.7±4.6	0.017±0.004	0.32±0.04	-
c028 (FR, 94-162)	*Sat. fit*	91.7±19.5	64.3±7.7	-	-	0.79±0.05
	kObs fit		200±69.9	0.005±0.002	0.92±0.07	-
	Direct rates		39±10.7	0.017±0.005	0.64±0.07	-
c037 (FR, 1-88)	*Sat. fit*	13.6±1.9	13±1	-	-	2.78±0.07
	kObs fit		21.7±5.9	0.009±0.001	0.19±0.05	-
	Direct rates		6.1±1.4	0.014±0.002	0.08±0.02	-
c038 (FR, 90-162)	*Sat. fit*	11.4±1.5	8.2±1.2	-	-	0.99±0.04
	kObs fit		21.5±3.8	0.017±0.002	0.35±0.06	-
	Direct rates		7.1±1.4	0.023±0.003	0.16±0.03	-

Figure 4.12: Kinetics of CXCL12 binding to different HMGB1 constructs (Figure 4.11)

A. Linear least squares fit of the apparent first order rate constant to CXCL12 concentration (global fit of 3 replicates). **Red**: FR FL-HMGB1 (c011), **black**: 3S-FL HMGB1 (c022), **purple:** FR-HMGB1 Box A 8-78 (c027), **brown:** FR-HMGB1 Box B 94-162 (c028), **pink:** FR-HMGB1 Box A 1-88 (c037), **grey**: FR-HMGB1 Box B 89-174 (c038). **B.** Saturation Michaelis-Menten fits of response at equilibrium to CXCL12 concentration (global fit of 3 replicates). Red: FR FL-HMGB1 (c011), black: 3S-FL HMGB1 (c022), purple: FR-HMGB1 Box A 8-78 (c027), brown: FR-HMGB1 Box B 94-162 (c028), pink: FR-HMGB1 Box A 1-88 (c037), grey: FR-HMGB1 Box B 89-174 (c038). **C.** Summary table of K_d values, kinetic constants (derived from both *kObs* fitted to concentration and raw measurements in the instrument – direct rates) and maximum binding (response in BLI) for each construct. Averaged kinetics parameters performed by fitting to a horizontal line representing the mean. Initial comparisons between kinetic parameters in fits (k_{On}, k_{Off}, K_d in both equations) performed via AICc without other parameters constrained. As in all cases differences were observed in the in-fit AICc test, post-hoc comparisons were performed via 2-way ANOVA, accounting all calculated values. Errors were propagated throughout via fitting the replicates to a horizontal line before analysis, with n= df+1.

Statistical summary:

- K_d for c027 was significantly lower than those calculated for all other constructs save FL-HMGB1 c011 ($p< 0.1$). K_d for c028 was significantly lower than those calculated for all other constructs ($p< 0.0001$).
- k_{On} was only different for c038, which had significantly higher association rate than the rest ($p< 0.01$).
- k_{Off} was significantly higher for both c027 and c028 when compared to all other constructs ($p < 0.0005$)
- All other comparisons were non-significant ($p>0.25$ or above)
- B_{max} was not formally compared as the equal variance hypothesis could not be validated, but the AICc test reported differences (99.99% probability)
- Pairwise comparisons for each parameter between calculations only highlighted differences between the first order fit for c027/c028 and the other calculations; this could be due to experimental noise from low binding.

4.5 Conclusion: defined, homologous regions within each HMG Box domain bind CXCL12.

One of my original aims was to determine the structure of the HMGB1-CXCL12 complex to better design HMGB1 derivatives which can induce G_{Alert} and enhance tissue repair, whilst eliminating potential deleterious proinflammatory signalling via TLR-2, TLR-4 and RAGE. I attempted both crystallisation of the heterocomplex and NMR structural determination. However, it quickly became apparent that elucidating the structure of the heterocomplex was a much more difficult prospect than originally anticipated. The affinity of HMGB1 for CXCL12 is very weak, in the micromolar range. Furthermore, the high solubility of both proteins in a wide range of buffers, and the large degree of flexibility HMGB1 has in solution, resulted in my abandoning this approach and instead focusing on indirectly determining the critical elements of HMGB1 for CXCL12 binding. NMR, whilst failing to help me determine the structure of the heterocomplex was nevertheless informative such that when combined with other techniques such as BLI and peptide arrays, and I was able to better understand the interaction of HMGB1 with CXCL12.

The affinity of the interaction of HMGB1 with CXCL12, despite being weak, is much higher than the reported affinity of CXCL12 for itself (although with a different technique), which is in the high micromolar range – one to two orders of magnitude weaker. Each HMG Box domain seemingly binds CXCL12 independently of whether the domain exists on its own or as part of a full-length HMGB1 protein. A non-cooperative binding mechanism for each HMGB1 Box to a CXCL12 monomer would also support recent data showing that signalling by the heterocomplex occurs through balanced GPCR signalling, much like monomeric CXCL12 alone [133], but with a different downstream profile.

This supports my initial hypothesis where HMGB1 binds two CXCL12 monomers, allowing each to simultaneously signal over the same CXCR4 homodimer in the monomeric signalling pathway, rather than the CXCL12 molecules forming a dimer and inducing persistence of G_0. However, there is still no direct evidence of the 1:2 stoichiometry for the HMGB1-CXCL12 heterocomplex. Whilst my observations of CXCL12 binding independently to both Box A and B Boxes supports this, direct stoichiometry measurement via e.g. ITC [245] would fully clarify this model.

Therefore, CXCL12 would have three different signalling pathways through CXCR4 depending on the binding mechanism:

1. A single CXCL12 monomer bound to one subunit within a CXCR4 homodimer (pro-chemotactic), inducing balanced CXCR4 signalling.
2. Two CXCL12 monomer bound each to one subunit within a CXCR4 homodimer, kept apart from dimerising by HMGB1 (pro-chemotactic, pro-regenerative) as the affinity of CXCL12 for HMGB1 (≈15 μM) is greater than to itself (100 μM, [114]), promoting G_{Alert} through balanced CXCR4 agonism but with a different arrestin profile, resulting in no CXCR4 internalisation [133].
3. A CXCL12 dimer bound to a CXCR4 homodimer, each CXCL12 subunit interacting with both CXCR4 subunits (pro-quiescence), inducing G-protein signalling only.

The results described in this chapter expand upon the existing literature, where NMR data have been used to speculate on which residues of HMGB1 are involved in CXCL12 binding [183]. However, a key limitation of only using NMR in this type of analysis is that it relies on changes of the local microenvironment of each residue, which translate in either perturbations of chemical shift, or decreases in intensity due to peak broadening in the case of weak interactions. A protein-protein interaction will necessarily affect the microenvironment of residues not involved in the interaction, simply due to the presence of another large macromolecule and, therefore, would present “false positives” as these residues show changes in NMR signal. This is also true for residues which are affected elsewhere in the protein, e.g., by a binding-induced conformational change, and it also cannot distinguish between a main-chain driven interaction and a side chain-driven interaction unless all the atoms in the chain are mapped via $^{13}C/^{1}H/^{15}N$ experiments. This is the case of, for instance, the N-terminus of CXCL12, which is profoundly affected by HMGB1 binding despite not being involved in the interaction [96,131]. It also behaves differently upon binding to FL-HMGB1 compared to a Box domain alone despite this binding likely being uncooperative; this is due to a change in the relative position of the CXCL12 monomers in the heterocomplex, compared to the dimer or as free monomers.

Since NMR data alone are not sufficient, other techniques are needed to confirm residues identified in these assays, as well as kinetics data to validate the proposed domains of interaction. My peptide array experiment shows a much larger number of residues being involved in the CXCL12 interaction, all of which localise in defined peptides conforming a binding pocket in the concave surface of each HMG Box where their side chains align, whereas some of the residues proposed in the existing literature and identified via NMR are positioned in regions well away from these pockets and do not align their side chains within them, suggesting an indirect contribution or response to CXCL12 binding. This is consistent with computational models published towards the end of my DPhil [182,183] based on published NMR data [96,131,232] describing this same interface, but which do not mention the mirrored pattern between Boxes or the contribution of the flanking sequences.

I had now a clear picture of the regions of HMGB1 required for CXCL12 binding. Whilst there are key sequence-specificities and several of the positions are conserved between the Boxes, both regions define the same binding interface within each Box domain structure. Furthermore, the tail domain is dispensable in the interaction, confirming my results in Chapter 3, where constructs truncated at 184- (no acid tail) or 164- retain regenerative activity, although some decrease was observed for HMGB1 1-164 which could be attributed to the lack of the 165-168 region which was shown to be strongly influenced by CXCL12 binding in both NMR and peptide arrays. Furthermore, binding to each HMG Box seems to be non-cooperative between Boxes. Isolated HMG Box constructs are capable of binding CXCL12 with equivalent affinity to FL-HMGB1. As such, a key conclusion of this chapter is that each specific Box is “redundant”, i.e., a construct with two Box A domains, or two Box B domains, would be as capable of binding CXCL12 to the same extent as the full-length molecule. Further validation of the interacting sequences of HMGB1 and CXCL12 can be done via techniques such as crosslinking mass spectrometry [246] to isolate peptides from both proteins in close proximity. These peptides would allow identification of the regions of CXCL12 which bind to the identified peptides within the HMG Boxes, which likely are, at least partially, those identified in the reverse peptide arrays. These experiments are currently in process.

Despite having failed to determine full structure of the heterocomplex, these data were sufficient to progress with my DPhil. Future research could attempt to approach the full ternary structure of the complex with CXCR4 with cryo-EM, as the presence of the GPCR dimers would add sufficient mass to allow structural determination via cryo-EM. Several techniques can be applied to stabilise the likely active conformation of the GPCR in the complex, such as antibodies against the HMGB1-CXCL12 complex or recombinant mini G-protein [125] C-terminal fusion of proteins to stabilise crystal contacts (e.g. chimeric lysozyme [134]) is also feasible but could interfere with the conformational changes associated to the ternary complex. Gradient crosslinking [239] also is an option due to the higher molecular weight of the full complex, compared to HMGB1-CXCL12 complexes on their own.

Having determined which regions are required for CXCL12 binding, as well as what residues within them are critical, mutagenesis of those involved in the proinflammatory and deleterious activities of HMGB1 could be planned in a rational manner. I could keep the positions involved in CXCL12 binding as free of mutation as possible, or only modified with conservative substitutions. This will be discussed in the next chapter, where I describe how I modified HMGB1 to retain CXCL12 binding whilst removing interactions with DAMP receptors potentially deleterious in a therapeutic molecule.

Chapter 5 The HMGB1 construct panel

5.1 Chapter introduction **168**
5.2 Data used in construct design **169**
5.2.I Results on HMGB1-CXCL12 interaction 169
5.2.II The interaction of HMGB1 with TLR-4, TLR-2 and RAGE as described in the literature 170
5.2.II.i Toll-like receptor 4 (TLR-4) 170
5.2.II.ii Toll-like receptor 2 TLR-2 171
5.2.II.iii Receptor for advanced glycation end products (RAGE) 173
5.3 HMGB1-C1 construct panel **175**
5.3.I Construct design 175
5.3.II DNA design and protein expression 180
5.4 RAGE-binding screen **184**
5.4.I Assay optimisation 184
5.4.II Screen of first batch of HMGB1 constructs 186
5.4.III Selected RAGE-deficient constructs 192
5.5 TLR-4 and TLR-2 activity screen **193**
5.6 Regenerative effect of Double box B (DbB) construct -HMGB1A-C1-c048 ..... **196**
5.7 Stability of DbB-HMGB1 ..... **200**
5.7.I SEC profile of DbB compared to other HMGB1 constructs 200
5.7.II ESI-QTOF/MS 202
5.7.III Differential scanning fluorimetry (DSF) 205
5.7.IV Long-term solution stability at room temperature of FR-HMGB1 FL, 1-164, and DbB 207
5.8 Conclusions **211**

5.1 Chapter introduction

Following completion of the work described in the preceding two chapters, I had sufficient information on the interaction between HMGB1 and CXCL12 to deduce which residues would be suitable for modification, whilst preserving the regenerative potential of HMGB1.

HMGB1 can bind multiple PAMP and DAMP receptors. TLR-4 and TLR-2 have been reported to promote excessive inflammatory signalling in response to DS-HMGB1 and RAGE is known to be a key mediator of prothrombotic signalling [64,247]. These interactions would be highly undesirable for the use of HMGB1 as a therapeutic. For instance, RAGE interaction and the associated thrombosis is deleterious in the context of cardiac repair following myocardial infarction. Whilst it has been argued that substitution of cysteines with serines in 3S-HMGB1 would circumvent this, 3S-HMGB1 was found to be deleterious for cardiac tissue repair, leading to excessive fibrosis (aberrant healing) [173], whilst FR-HMGB1 promoted cardiac regeneration. Therefore, a variant of HMGB1 other than 3S-HMGB1 is required if HMGB1 is to be used as a therapeutic.

Whilst the downstream biological effects of HMGB1 interaction with TLR-2/4 and RAGE have been studied extensively, there is limited information regarding the molecular basis interaction of HMGB1 with these receptors, although TLR-4 and RAGE are partially described[47]. To further complicate matters, signalling via all three receptors converges through Myd88 and NF-κβ [83], greatly hampering functional studies. Nevertheless, based on the data I gained on the residues and motifs involved in CXCL12 interaction, I was well-placed to design my first panel of constructs, by modifying residues known or likely to be involved in TLR-2/4 signalling and RAGE binding but are not required for CXCL12 interaction. In this chapter I will summarise the information which I used to design my panel based on data from my experiments and the literature, the basis of my design and the assays I used to assess the effects of my truncations and substitutions.

[47] Whilst peptide arrays, such as those described in Chapter 4, would have been helpful, the difficulty in producing in-house large batches of His-tagged TLR-4 and TLR-2, precluded my using this technique to identify.
Although the soluble, HMGB1-binding domain of RAGE has been expressed, the MGC collection lacked the full RAGE sequence. Consequently, I used commercially available protein.

5.2 Data used in construct design

5.2.I Results on HMGB1-CXCL12 interaction

In Chapter 4, I found by peptide array (Section 4.4.I) and confirmed by NMR (Section 4.4.II) that there is a repeated pattern within HMGB1 critical for interaction with CXCL12. Residues involved in this pattern are aligned across the same sides of the α-helices in HMGB1, with their side chains conforming a pocket in the concave surface of each HMG Box. Flanking regions at the N- and C-terminus of each HMG Box domain, outside the helical components, are also involved in CXCL12 binding, dramatically stabilising the interaction. This was apparent from the data from my NMR and BLI experiments (Section 4.4) where removal of this peptide decreased affinity but did not completely abolish interaction.

Whilst Cys 22 and Cys 44 do not seem to be directly involved in the interaction with CXCL12, the presence of a disulfide bridge flips one of the α-helices in Box A [241], thereby hampering binding to CXCL12 [203]. The binding of CXCL12 to HMGB1 appears to be Box-dependent, rather than requiring two HMG Box domains acting in a cooperative fashion. HMGB1 is likely allowing two CXCL12 monomers to bind the two monomers of CXCR4 within a CXCR4 dimer independently, rather than a CXCL12 dimer binding a CXCR4 dimer, typical at high CXCL12 concentrations [117]. In fact, binding of CXCL12 monomers to a CXCR4 dimer promotes CXCL12 dimerisation over the dimer surface [117]. As shown by my BLI experiments, both Box A and Box B can recruit CXCL12 with similar affinity to FL-HMGB1, and in comparable amounts.

5.2.II The interaction of HMGB1 with TLR-4, TLR-2 and RAGE as described in the literature

5.2.II.i Toll-like receptor 4 (TLR-4)

TLR-4 is one of the main pattern recognition receptors for immune signalling, particularly for pathogen-associated patterns (PAMPs, e.g. LPS) but also damage-associated patterns (DAMPs [248,249]). DS-HMGB1, but not FR-HMGB1, binds TLR-4/MD-2 complexes to mediate inflammatory signalling: in the absence of MD-2 signalling cannot occur [87]. Whilst the pro-inflammatory signalling pathways of DS-HMGB1 through TLR-4/MD-2 have been extensively studied [86,96,97,250,251], studies on the specific epitopes or kinetics of the interaction are more recent.

TLR-4 signalling epitopes were initially mapped within HMGB1-Box B [252], residues 88-107 (89-108 if the N-terminal Met is counted), despite the apparent contradiction with the vast majority of data which indicate that the disulfide in Box A is critical for the proinflammatory activities. The FCSE motif (104-107) in box B is critical for HMGB1 interaction with TLR-4/MD-2 [87]; a small peptide with the sequence FSSE competitively inhibits signalling by binding to MD-2 and displacing HMGB1. This results in suppression of DS-HMGB1 dependent signalling, but does not modify HMGB1 enhancement of TLR-4 responses to LPS [87]. HMGB1 can bind endotoxin via two separate peptides in Box A, residues 2-14 (binds the delipidated component of LPS) and the linker, residues 79-95 (binds to Lipid A). It is hypothesised that this binding is necessary for the transfer of LPS from HMGB1 to CD14 and to TLR-4/MD-2 [52], where HMGB1 can replace the function of LPS-binding protein (LBP) .

A recent study [88] described the full binding mechanism for the TLR-4/MD-2/HMGB1 ternary complex in the absence of LPS (DAMP signalling). Disulfide HMGB1 binds TLR-4/MD-2 complexes with higher affinity (10-fold faster association rate) than FR or 3S-HMGB1. An initial interaction of DS-Box A with TLR-4 with high affinity is fully stabilised by Box B binding to MD-2, with low affinity but extremely low k_{off}. Their results also show that DS and FR-HMGB1 can bind TLR-4 in the absence of MD-2 with greater affinity than 3S-HMGB1. This interaction is likely devoid of biological significance, as MD-2 is required for downstream inflammatory signalling [87]. Binding of HMGB1 to the TLR-4/MD-2 complex triggers HMGB1-TLR-4-MD-2 complex dimerisation, with HMGB1 performing a similar role to that of LPS in the TLR-4/MD2 complex.

5.2.II.ii Toll-like receptor 2 TLR-2

Unlike TLR-4, the biophysics of the interaction of HMGB1 with TLR-2 have been much less studied, although there are several publications analysing the biological consequences of this interaction [83,253]. The binding mechanism for TLR-2-mediated HMGB1 signalling is ill-understood, with conflicting evidence available in the literature. Whereas several studies report direct TLR-2 agonism by HMGB1 in either TLR-2 reporter HEK cells or macrophages [90,254–256], more recent publications support a model where HMGB1 requires binding of TLR-2 co-ligands to signal through TLR-2 [49,59]. Unlike with TLR-4, TLR-2 signalling might not be restricted to the DS-HMGB1 since both full-length FR and DS forms cannot bind on their own, albeit this study showed tail-less HMGB1 (natively binding TLR-2) to decrease its binding upon reduction. Publications prior to 2012 [83,254] do not differentiate FR and DS HMGB1 since this redox switch was unknown before then [24]: however, studies after this date still do not report the oxidation status of HMGB1 in their experiments [255,256], even after the redox switch was identified. Based on the few biophysical studies on the TLR-2/HMGB1 interaction, it involves at least one HMG Box domain and the acidic tail.

Involvement of the acid tail became apparent when studying how HMGB1 assists in the transfer of lipotheicoic acid (LTA, a peptidoglycan) [59] to TLR-2: the C-terminal acidic tail seems to modulate this response [49]. Tail-truncated forms (1-184, -Δ30; 1-196, -Δ18) of HMGB1 are capable of binding TLR-2 directly, unlike DS-HMGB1, but they do not signal through this receptor unless a native TLR-2 ligand such as LTA is also present. This would suggest that the tail is masking the TLR-2 binding epitopes, perhaps by binding the HMG Box core. Other activities of HMGB1 through TLR-2 have been reported in sterile inflammation (e.g. thrombosis [64] or autoimmune disorders [255]). As in sterile injury PAMPs should not be present, other endogenous co-ligands are likely present *in vivo* to allow for active signalling of HMGB1 through binding to TLR-2[48], or post-translational modifications such as lysine acetylation [54,55] by modifying the electrostatic interaction of the acid tail with the core domains of HMGB1 [50].

[48] This publication does not specify the redox state of HMGB1 1-184, although the lack of reducing agents suggests that the of DS form was used. Additionally, neither study mentions removal of PEG or TX-114 from their preparations. This could have masked pro-inflammatory effects due to polymer contamination [148] and LPS levels/protein concentration due to TX-114 interfering with both assays.

This would also explain the results published where DS-HMGB1 alone is capable to induce TLR-2 agonism, as they were all performed in the presence of serum which could contain such ligands. In all scenarios, the currently unknown TLR-2 binding sites within HMGB1 end exposed to solvent.

A study on the effects of HMGB1 in autoimmune thyroid disease also provided indirect information on the requirement of an HMG Box. Glycyrrhizin, an established HMGB1-CXCL12 inhibitor, also inhibits its response in a sodium iodide model, where administration of glycyrrhizin decreased the expression of HMGB1, TLR-2, Myd88 and NF-κβ in the thyroid gland [255]. Glycyrrhizin binds to each HMG Box domain, and the binding site is known to be distributed in the segment between α-Helices I and II of HMGB1. I hypothesised that either the residues involved in this interaction, or neighbouring ones, could be involved in TLR-2 interaction. It could be possible to elucidate these via e.g., peptide arrays, but time and equipment constraints precluded these experiments.

5.2.II.iii Receptor for advanced glycation end products (RAGE)

There is a confirmed RAGE-binding epitope mapped to residues 149-182 of HMGB1, with significant sequence and structural conservation when compared to other RAGE ligands of the S100 family [135]; it has been proposed that this interaction is not sequence-specific but rather pattern-specific, given RAGE´s status as a DAMP receptor [179]. This site was later associated to inducing survival and outgrowth of neurites expressing RAGE [166], before amphoterin was renamed to HMGB1. A peptide with this sequence is an efficient RAGE antagonist [257]. Peptides were devised [179] from the S100P and S100A4 sequences as well; peptides homologous to positions 149-158, and 156-169, of HMGB1 are commercialised as RAGE inhibitor peptides (RAPs)

It is likely that most of the signalling pathways described for RAGE which involve HMGB1 occur through the interaction with this domain. However, Box A is clearly also involved in this interaction as DS-HMGB1 has been shown to be a critical mediator for RAGE-induced thrombosis [64,80,247]. Stark et. al showed that disulfide, and not sulfonyl or 3S-HMGB1, is capable of recapitulating prothrombotic and proinflammatory effects that are RAGE-dependent. Inexplicably, the authors omitted FR-HMGB1 controls from this analysis. However, both 3S-HMGB1 and sulfonyl-HMGB1 were reported in this study to bind RAGE with high affinity in almost irreversible manner. Whether the prothrombotic activities of HMGB1 require binding of RAGE to peptide sequences within Box A specifically, or Box A oxidation is inducing a conformational change in HMGB1 allowing for better RAGE interaction via the C-terminal site, is unknown. However, Box A and derived peptides have been reported to be capable of inhibiting RAGE-mediated HMGB1 signalling in several contexts [258,259] which supports direct implication in RAGE binding.

A second RAGE binding site has also been described in Box A [260]. However, this site is initially inactive and only becomes accessible through Caspase-1 cleavage of Asp 66 and is therefore unlikely to be involved in the prothrombotic pathways. Once cleaved, this site can signal through RAGE and trigger reversal of immune tolerance to apoptotic cells and secondary infection [260], improving survival of animals upon secondary septic infection as immune cells are primed. Should a Caspase-1 site within Box A be mutated (Asp 66-Ala), this effect is lost. However, this site has not been directly implicated in any other activities, particularly prothrombotic ones.

Other publications report pro-differentiating [135,167] or chemotactic [31,91,261] properties of the HMGB1-RAGE interaction, though these were published before the description of the multiple redox states of HMGB1 [262]. Based on the evidence of the dynamic interchange between reduced, oxidised and sulfonyl HMGB1 [74], I would expect they pertained to the reduced form, which was shown to be chemoattractant in later studies, both with respect to RAGE [24] and to CXCL12-CXCR4 signalling [96]. It is important to note that following injury, the local proinflammatory conditions could favour conversion of FR-HMGB1 to DS-HMGB1, shifting its activity towards triggering of detrimental prothrombotic pathways in injury. A summary of the RAGE binding peptides is shown in Figure 5.1.

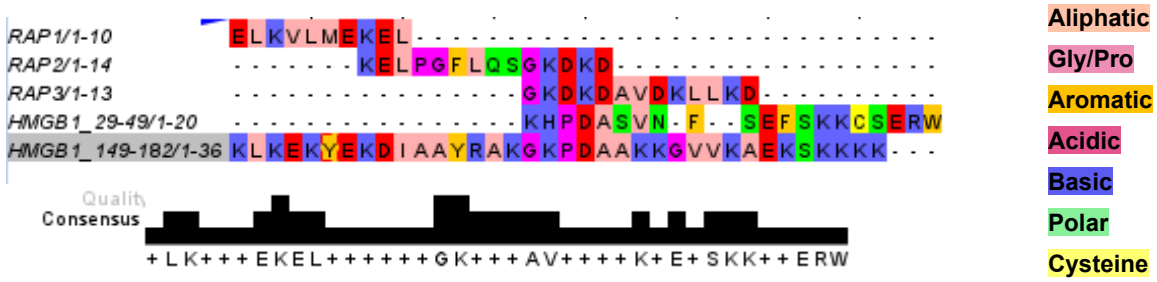


Figure 5.1: Alignment of RAGE binding sites in HMGB1 Box A (C22-K49), Box B (K149-K182), and RAP peptides from S100A/S100A4

Colour key is shown in the right-hand side; conservation score (ClustalOmega) below. Residue positions have been taken as starting at -1 from the N-terminal methionine. RAP1: 149-158. RAP2: 156-169. RAP3: 163-176. RAP3 was shown in the literature to be the least effective in RAGE inhibition

There is also conflicting evidence in the literature with regards to the exact affinity of the interaction between HMGB1 and RAGE. The original study which identified RAGE binding to HMGB1 reported, via radiolabelling of HMGB1 and adsorbed RAGE, a dissociation constant (K_d) of approximately 6 nM [167], but further work has yielded affinity constants in the range of 100-650 nM by SPR [53,64,263]. It is also unclear how many binding events are involved in the interaction, whether it is one or two. The differences in affinity could be attributed to either the different sensitivity of the techniques, or steric effects of the immobilisation of either HMGB1 or sRAGE. However, two of the studies coincided in reporting that heparin, a known ligand of HMGB1, inhibits the interaction between HMGB1 and RAGE. This is likely mediated through direct binding of heparin to HMGB1, rather than to RAGE, as heparin and sRAGE showed little to no affinity for each other in the SPR model. This was supported by the loss of proinflammatory effects of HMGB1 in RAW264.7 macrophages and human vascular endothelial cells (HUVEC) [53] upon exposure to heparin.

5.3 HMGB1-C1 construct panel

5.3.I Construct design

Due to the multiple signalling pathways associated with HMGB1, designing a potential therapeutic that retains regenerative activity whilst eliminating potential deleterious proinflammatory effects is challenging. The sequence of HMGB1 is highly conserved (>99% across vertebrates), likely due to its multiple biological functions restricting the number of possible mutations without deleterious effects (negative selection against mutants). My HMGB1 construct panel (henceforth referred to as HMGB1-C1) was based on the following criteria:

- Preservation of residues and peptide patterns which I found to be critical for CXCL12 binding (5.2.I)
- Removal of residues known to be important for binding to TLR-2, TLR-4, and RAGE.

First, I screened those constructs that expressed successfully with yields comparable to those of FL FR-HMGB1 for decreased TLR-4/2 signalling and RAGE binding (item 2). Constructs fulfilling these criteria would be assessed *in vitro* and *in vivo* for regenerative activity (item 1). Promising constructs would be then evaluated for stability. The designed constructs can be classified in the following groups, depending on which regions of HMGB1 are modified (Figure 5.2 and Table 8.5; sequence details in Appendix A1Chapter 4):

- **Control constructs (C1-c001 to c003):** Re-cloned versions of HMGB1A-c001, c002 and c031. These constructs cover all three of the truncations which I found had regenerative activity and are described in detail in Chapter 3. They fulfilled the dual function of positive expression and regenerative activity controls.
- **Pro-apoptotic RAGE-binding site (C1-c004 to c011):** Whilst it is not known if this site affects wound healing and injury, it is established that removal of the Caspase-1 sites in Box A disables this interaction [260]. These constructs contain mutations of Asp66, 157 and 168 (the critical Caspase-1 residues) to either Alanine (published mutation; likely inactive based on my peptide array results), Glu (different side chain length, same charge) or Asn (same length, different charge).

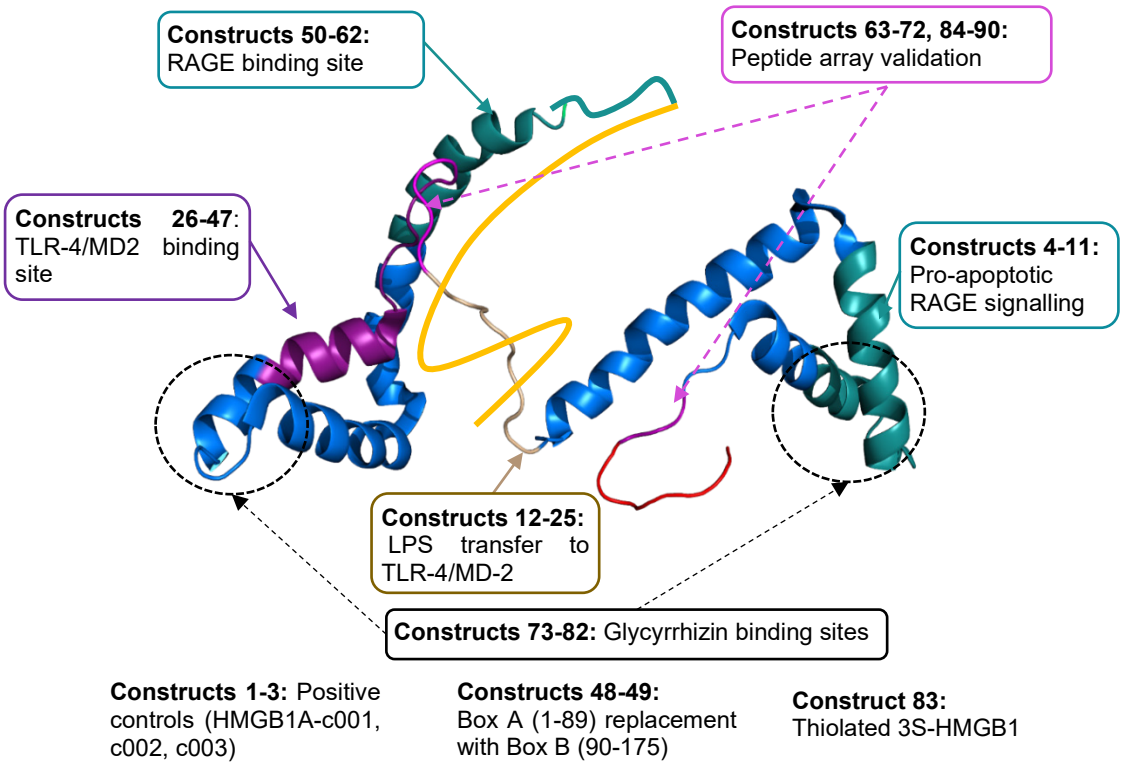


Figure 5.2: Types of constructs in the HMGB1 construct panel (HMGB1-C1)

Designed HMGB1 constructs, grouped by type of single amino acid mutations, and highlighted in the structure of HMGB1 (PDB 2YRQ the unstructured 165-214 peptide has been drawn overlaid as it is absent in said structure). Constructs not highlighted over the structure correspond to broader truncations of the full-length molecule.

- **Flexible LPS-binding linker (C1-c012 to c025):** Whilst most internal tissue injuries occur in a sterile setting, e.g., myocardial infarction, it is conceivable that the patient may have a concurrent infection. Open wounds such as fractures or surgery may also be contaminated. Therefore, it is desirable to reduce the potential binding of LPS to the optimised HMGB1 construct. Whilst the first 4 residues present before each HMG box are involved in CXCL12 interaction, the remaining linker residues (79-89) were deleted in serial pairs, from both N- and C-termini. If these mutations do not negatively impair tissue repair, the effect on LPS binding could be evaluated *in vitro* via binding assays such as BLI. These would also inform on whether the flexibility of the region within the HMG Boxes is relevant for optimal CXCL12 positioning. I also replaced pentamers within this region with GSGSG linkers, including the N-terminal region ahead of Box B required for CXCL12 binding (potential negative control).

- **Modification of the TLR-4 and MD-2 binding site (C1-c026 to c047):** The 88-109 peptide is well known as the TLR-4-MD-2 signalling site. Cys105 has been identified as critical. However, modification of this residue is risky as e.g. 3S-HMGB1 (Cys 22, 44, 105 -> Ser), despite not inducing thrombosis in a DVT model [63] and allegedly losing TLR-4 binding capabilities [264], has been reported to causes excessive fibrosis in an MI model [174].

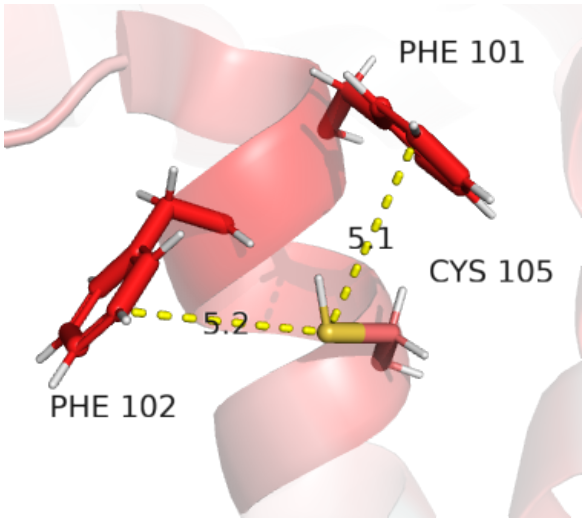


Figure 5.3: Potential TLR-4 binding hydrophobic pocket in Box B

Eisenberg hydrophobicity colouring of HMGB1 residues (Ala100-Lys111). Distance of the thiol in Cys105 to the sidechains of Phe101/102 (Å) are shown for Conformer 1. Red=apolar, white=polar. All conformers in Box B align nearly superimposed (see Figure 3.14)

Instead, I took an alternative approach and modified the neighbouring residues. The rings of Phe101 and 102 are positioned in both sides of the sidechain of Cys105 (Figure 5.3). It is known that both C105->S/E suppresses HMGB1-mediated TLR-4 signalling [265]. I hypothesised that the addition of a more polar group to the position 105 (e.g., Ser/Glu, when compared to the low polarity of thiols) between both Phe rings could disrupt a local hydrophobic pocket involved in TLR-4 binding.

I attempted to disrupt this pocket by modifying both Phe residues to other aromatic groups (Tyr; polar group added, His/Trp; different aromatic group altogether), or to glutamic acid. This could affect regenerative activity as Phe 101/102 appear to be involved in CXCL12 binding. In addition, the FSSE tetrapeptide (where C106->S) has been shown as being a potent inhibitor of the interaction with the TLR-4/MD-2 complex by binding MD-2 instead [87] by having a higher affinity than the FCSE motif present in Box B. Interestingly, 3S-HMGB1 has this same tetrapeptide sequence. Therefore, I also designed constructs with this peptide present in the flexible, unstructured regions of HMGB1 to act as a potential “decoy”, binding MD-2 in an inappropriate position for activation of TLR-4 signalling.

- **Double Box B constructs (C1-c048, c049) (**Figure 5.4**):** Whereas Cys105 is critical for TLR-4/MD-2 signalling, it is interaction with disulfide Box A that initiates the binding of this complex to HMGB1 [88]. As the pattern of peptides and residues binding CXCL12 is mirrored between Box A and Box B, and both boxes are similar in sequence and nearly identical in structure, substitution of Box A for the Box B sequence could still preserve regenerative activity via CXCL12 binding and eliminate DAMP signalling. The duplicated Box B domains include the N-terminal CXCL12 binding peptide, and positions 165-174 C-terminal to Box B replace the flexible LPS binding linker; these include the flexible regions involved in CXCL12 binding. c049 is identical to c048 but has an additional GSGSGSG linker between each Box B repeat to add extra flexibility of the linker domain.

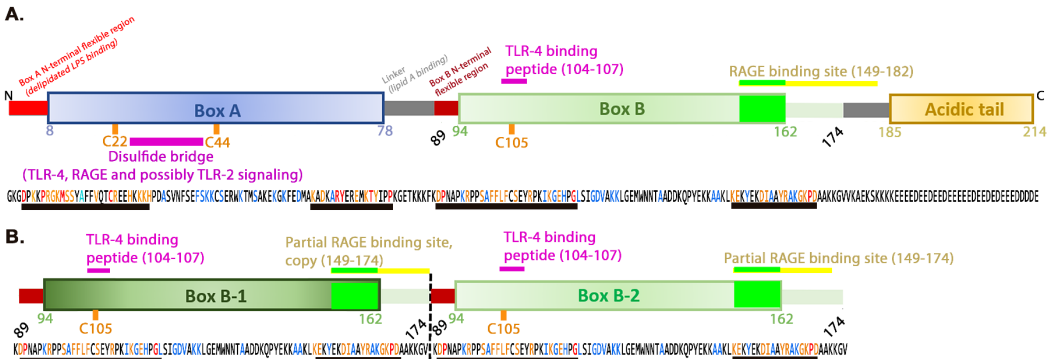


Figure 5.4: Schematic of HMGB1-C1-c048

A. Schematic representation of HMGB1 domains overlaid with the sequence (from Figure 1.3). Residues in red: identified in peptide array only. In orange: identified in peptide array and NMR. Blue: identified in NMR (my own, published data or both; see Figure 4.11 for details). **B.** Representation of HMGB1-C1-c048 (double Box B)

- **Box B RAGE binding site (C1-c050 to c062):** As the conservation of RAGE-binding motifs has been studied across several of its ligands, including HMGB1 [135], in an attempt to disrupt RAGE interaction I used this knowledge to modify the residues conserved in these positions to different side chain groups. This region (149-182) overlaps with the CXCL12-binding peptide B3 in Box B so several of these were not modified as they were flagged in the alanine scan.

- **Glycyrrhizin binding site mutations (C1-c073 to c082):** Whilst I was unable to find any information on specific residues involved in TLR-2 binding**,** glycyrrhizin can inhibit TLR-2 interaction as well as CXCL12 interaction with HMGB1. I hypothesised that modifications to the conserved glycyrrhizin binding residues (Arg24/Lys42 in Box A, Arg109/Lys126 in B) could negatively influence TLR-2 signalling, although as these are contained in the CXCL12 binding regions it could affect regenerative activity as well. These mutations were Arg->Ala/His/Lys and Lys->Met/Gln.

I also designed several constructs without the intent of testing them for pro-inflammatory properties; rather, these were devised as potentially useful tools for validating the results from Chapter 4.

- **Negative control constructs, from alanine scanning (C1-c063 to c072):** Modification of the residues marked as critical for CXCL12 binding in the N-terminal regions of each Box domain, to alanine, in order to confirm their importance, as highlighted by my BLI and NMR experiments. These were not part of the proinflammatory screening.

- **Thiolated 3S-HMGB1:** As 3S-HMGB1 can be used to replicate the regenerative properties of FR-HMGB1, I modified Lys 87 in the linker to cysteine. *in vivo* tracking of HMGB1 were required in the future, maleimide or Click chemistry could be applied to label HMGB1 in a biologically unimportant position.

- **N-terminal scanning of Box A (C1-c084 to c090):** To rule out any other residues in Box A being part of the molecular lock described in Chapter 3, I serially deleted residues 1 to 7 from the N-terminus to confirm that all of these constructs would express. Some of these forms would be inactive due to truncation of the 4-8 N-terminal peptide, but as with the alanine scan mutants, these were not part of the proinflammatory screening.

5.3.II DNA design and protein expression

Due to the sheer volume of work and complexity necessary for the expression of all these constructs, DNA sequences were designed *in silico* and purchased directly cloned into pNIC-CTH0 from a commercial synthetic gene supplier (Twist Biosciences), rather than cloned in-house as for those constructs described in Chapter 3. This vector adds a minimal tag (C-terminal TEV-His6) to the recombinant protein.

Whilst most modifications were either short deletions or SNP (single nucleotide polymorphism) mutations to preserve as many base pairs in each triplet identical to the wild type sequence as possible, the double Box B constructs (c048 and c049) required special consideration as synthesis of a 300 bp repeat sequence is almost impossible without optimising the sequence first. There is ample literature on gene optimisation [161,266–270] and user-available software such as EuGene [161]. However, optimisation of the native Double Box B sequence would lead to similar problems with the repeats. Therefore, I only optimised the first Box B repeat substituting Box A, and preserved the native Box B sequence, to generate synonymous DNA sequences with different nucleotide sequences.

All sequences barring HMGB1-C1-c049 were successfully produced by the supplier. A second attempt resulted in the same construct failing and, therefore, it was abandoned. As the sequences were synthesised *in silico* and sequenced by the manufacturer, this also bypassed the need for in-house verification of the DNA sequences. They were transformed into the same DNA storage and expression strains as for my initially designed HMGB1 constructs. However, an incident at the manufacturer site resulted in constructs arriving in several batches. Constructs c014-c016, c019, c022, c029-c041, c047-c055 and c063 were significantly delayed by >2 months and I proceeded with those I had at the time. This ultimately led to several constructs being excluded *a priori* from further study due to identification of a promising candidate early in the screens and the necessity to reduce on-site work due to the Covid-19 pandemic.

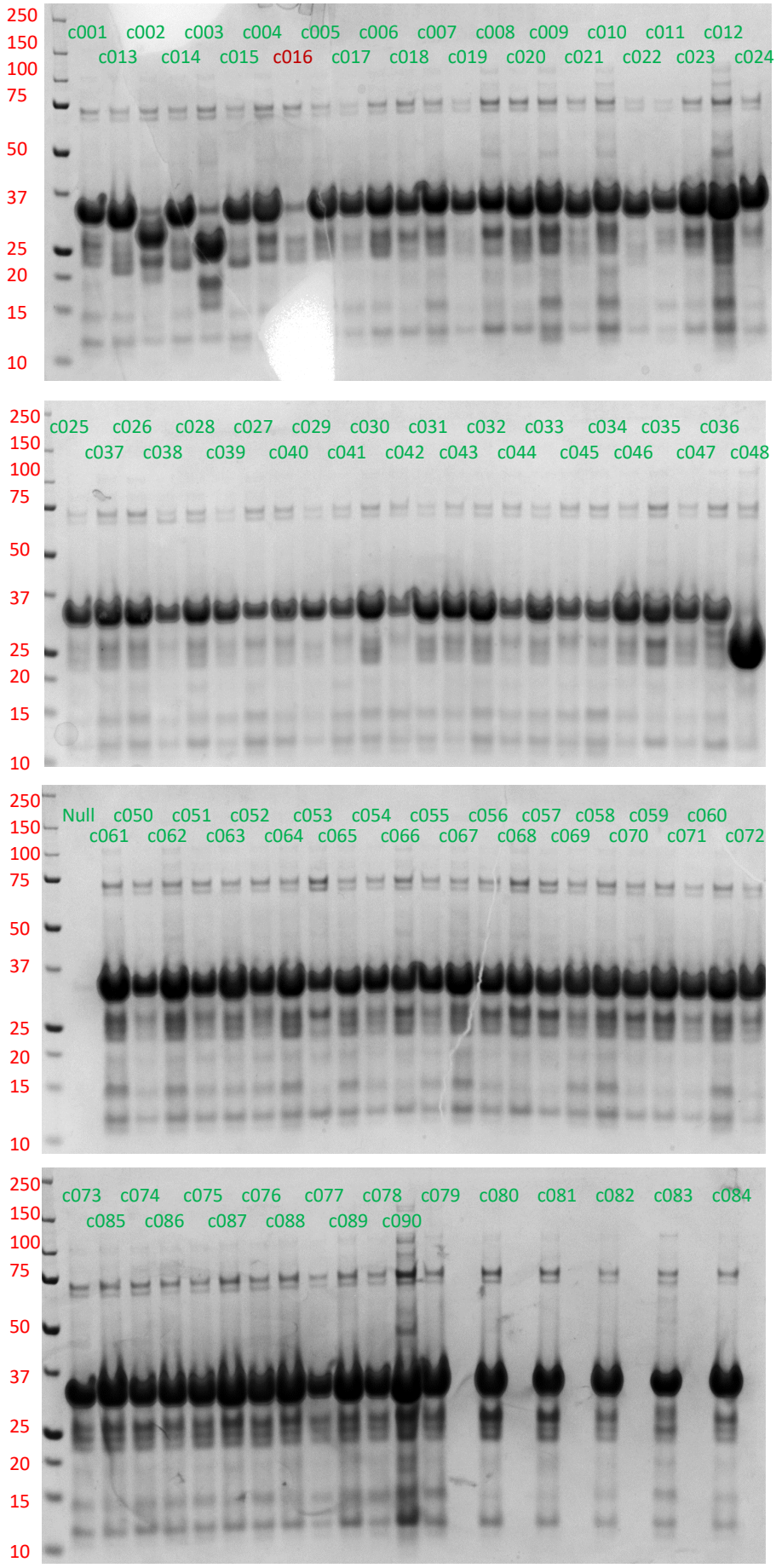


Figure 5.5: Test expression of HMGB1 construct panel C1

Test expression of HMGB1 constructs in 1 ml of BL21(DE3)-R3-pRARE2. Molecular weights of samples were verified by denaturing LC-ESI-TOF/MS to confirm identity (see Suppl. data A2-5-A). A detailed list of the mutations present in each construct can be found in Suppl. data A1-5-A.

When cloned into expression strains, expression was observed for nearly all constructs (Figure 5.5); only HMGB1A-C1-c016 (**PP**KGETKKKFKDPNA->**--**KGETKKKFKDPNA) failed to express. It is possible that removal of the two highlighted Proline residues destabilises the linker, although I was not able to perform the same in-depth validation as described in Chapter 3 since it would defeat the purpose of a screening panel. Interestingly, the double Box B construct which lacks the full linker, all other linker truncations, and constructs with serial truncation of the 1-7 fragment in Box A did not have decreased expression, which would rule out instability of the HMG Box domain itself. As proline introduces sharp turns in peptides, it is likely that this results in incompatible folding, but the turn itself is not needed with other sequences.

Several of the Phe101/102 mutants had noticeably lower yield in test expression, suggesting that the modification of these two phenylalanine side chains results in protein instability. This is better seen when upscaling expression (0.5-1 L), although small amounts of soluble protein could be isolated for most constructs[49]. In these experiments, often a white, opaque cell pellet was observed in low-expressing constructs, potentially indicating the formation of inclusion bodies, and therefore identifying those constructs as potentially unstable. Consequently, I excluded them temporarily from further testing; if no other suitable construct candidates were identified in the panel screening, I could revisit those in future studies. A summary of the results from large-scale expression of the constructs used in downstream tests can be found in Figure 5.6; these the constructs used for the optimisation experiments described in 3.4.II.ii.

[49] All these experiments were performed parallel to the optimisation described in Chapter 3.

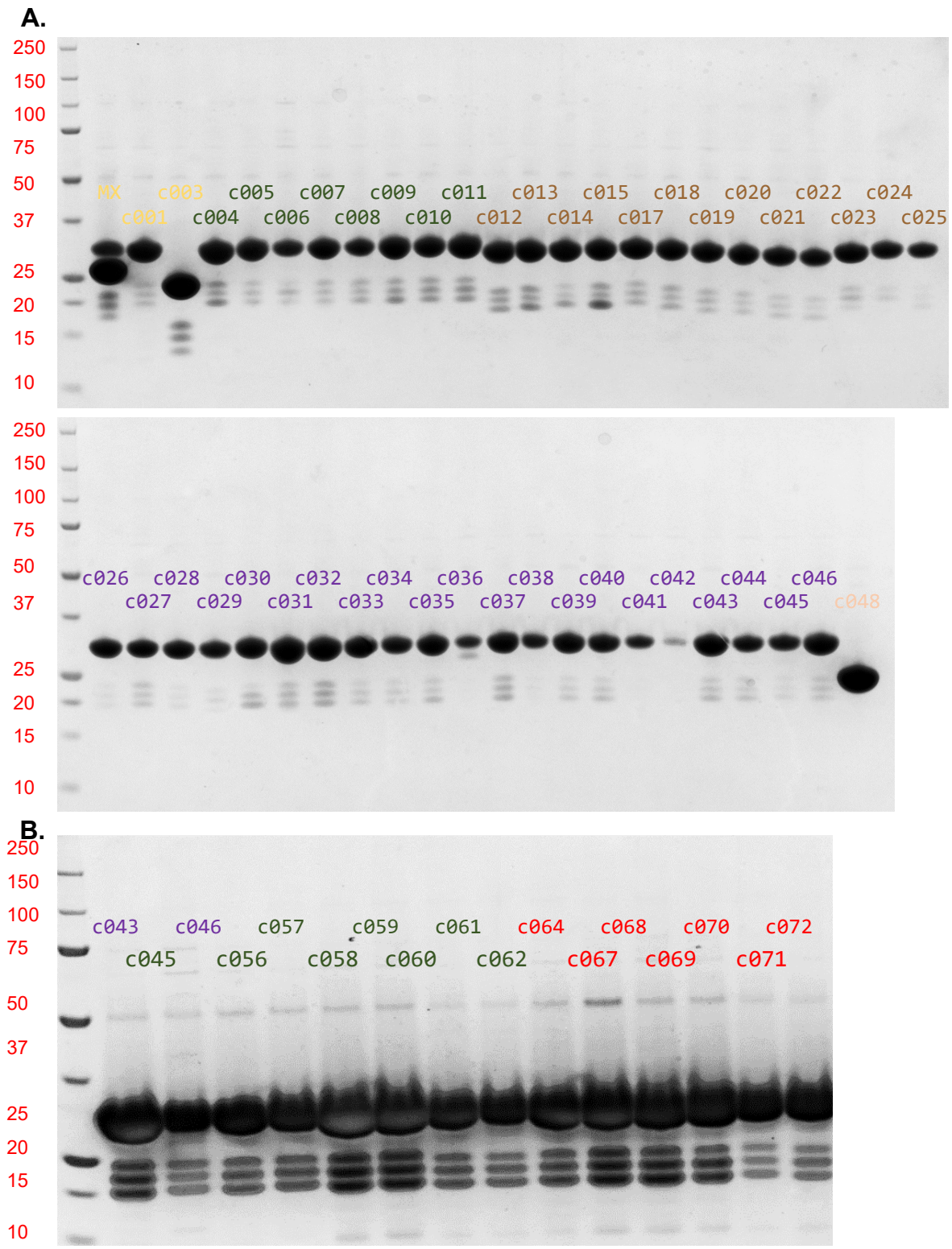


Figure 5.6: Large scale purification of HMGB1-C1 constructs

SDS-PAGE gel of purified HMGB1-C1 constructs (0.5L of BL21(DE3)-R3-pRARE2, induced at OD 2 with 0.5 mM IPTG). Lysate was purified by IMAC, followed by TEV cleavage and reverse-IMAC to remove contaminants; these samples were further purified by SEC chromatography. **A: Batch 1, HMGB1-C1-c001 to c048.** 1 µL of sample was diluted in 59 µL water and 20 µL 4X loading dye, supplemented with DTT. MX = cross-contaminated sample with c001 and c002 constructs (this sample was discarded and re-expressed in a separate experiment). Constructs have been coloured as per Table 5.1. **B: Batch 2, HMGB1-C1-c043 to c072.** 1 µL of sample was diluted in 29 µL water and 10 µL 4X loading dye, supplemented with DTT. Protein concentration in this batch was higher resulting in gel overloading. All protein samples correspond to those purified with the final high-throughput IMAC protocol described in 3.4.II.ii.

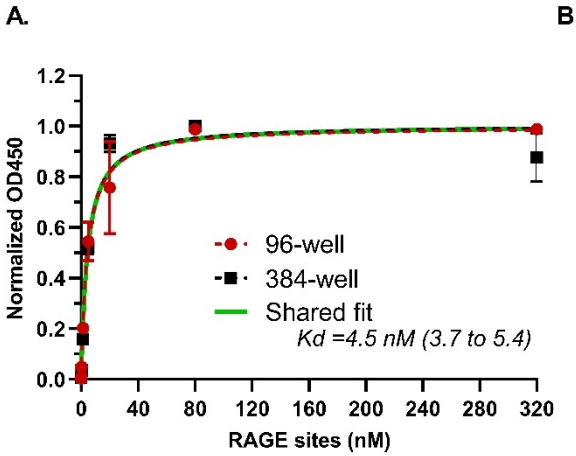
5.4 RAGE-binding screen

5.4.I Assay optimisation[50]

Whilst RAGE-mRNA quantitation has been used to measure HMGB1 agonism to this receptor [91,167], this is an indirect measurement of positive feedback signalling, whereby HMGB1 stimulation leads to increased RAGE expression, rather than pro-inflammatory response. To my knowledge, there is no RAGE-reporter cell line described in literature. It would be possible to construct a signalling assay with macrophages, for which there is a reporter cell line (THP-1 Dual NF-κβ, Invivogen). However, use of this line to assess RAGE signalling would require TLR-4 and TLR-2 inactivation as they express all three receptors, which converge in NF-κβ signalling and therefore the specific signal to each cannot be resolved from each other. TLR-2 inhibition is also challenging and costly for a screening panel, requiring siRNA or inactivating antibodies to prevent HMGB1-induced RAGE responses from being masked by TLR signalling. Therefore, I screened the constructs via their binding affinity to RAGE.

I adapted the methodology described for TLR-2 affinity measurements (ELISA) [49] to the existing RAGE-HMGB1 adsorption assays [180] (IMAC-agarose-HRP beads binding His-tagged RAGE), switching them to 384-well format. Otherwise, the DS-HMGB1 and blank controls would allow only for 2 constructs to be tested per plate, greatly decreasing throughput (Figure 5.7). Due to the nature of ELISA assays, where dissociation is greatly overrepresented (as only one step occurs in the presence of ligand excess, compared to several wash and secondary antibody steps which are effectively dissociation steps), the kinetic parameters derived from these would be more representative of dissociation rates. Nevertheless, it would allow for initial screening of multiple constructs in a much faster way than BLI or SPR.

[50] In-house expression of RAGE would have required a significant time commitment. Therefore, I purchased recombinant protein conjugated to IgG (Fc fragment), to develop this method.



B.

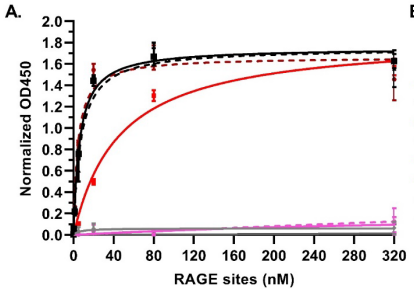
Goodness of Fit			
Degrees of Freedom			45
R squared	0.98	0.97	0.97
Adjusted R squared			0.97
Sum of Squares	0.052	0.16	0.22
RMSE			0.068

	Global (shared)
Comparison of Fits	
Simpler model	Kd same for all data sets
Probability it is correct	75%
Alternative model	Kd different for each data set
Probability it is correct	25%
Ratio of probabilities	3.0
Preferred model	Kd same for all data sets
Difference in AICc	-2.2

Figure 5.7: Adaptation of RAGE ELISA to 384-well format

A. Michaelis-Menten plot of ELISA measuring RAGE-Fc, bound to plates coated with HMGB1-C1-c001-DS (40 nM). Red circles: 96-well plate (3 replicate wells per concentration). Black squares: 384-well plate (4 replicate wells per concentration). Fits for both shown as discontinuous lines. Green line: shared fitted curve between both. Data points: mean ± SD (as otherwise error bars would not be visible) **B.** Statistical comparison between fits and probability of shared fit being correct (AICc method). RAGE-sites = 2x RAGE-Fc concentration .Data have been normalised for display purposes as absolute response values differ between

I then assessed binding of RAGE to HMGB1 full-length, Box A and Box B alone (full-length) in both reduced and oxidised states (Figure 5.8). In the ELISA assays, 3S-HMGB1 (K_d = 6.1-7.2 nM) bound equivalent amounts of RAGE and with similar affinity to DS-HMGB1 (4.4 nM), regardless of reducing or oxidising conditions. Published SPR data show high-affinity binding of 3S to RAGE, but with no dissociation [64]. FR-HMGB1 binds the same RAGE amount that both DS- and 3S-HMGB1 but with weaker affinity at equilibrium (45.7 nM) and not reaching full curve saturation. Box A and Box B were incapable of binding RAGE on their own, resulting in invalid fits. The values for DS-HMGB1 RAGE binding affinity would be in agreement with the results obtained by radioligand displacement (K_d= 6 nM) with RAGE adsorbed to a microtiter plate (in a very similar setup to ELISA) [167] but not with those from SPR (K_d= 650 nM).



B.

	Bmax		Kd		
	Value	SEM	Value	SEM	
DS-HMGB1A-C1-c001 (1-214)	1.7	4.9e-002	4.4	0.6	
DS-HMGB1A-c037 (1-88)	1.1	8.1	2553.3	20577.8	
"DS"-HMGB1A-c038 (89-174)	2.3e+013	4.4e+028	5.6e+017	1.1e+033	R^2=0.4
"DS" 3S-HMGB1A-c006 (1-214)	1.7	0.1	7.2	1.0	R^2=0.02
FR-HMGB1A-C1-c001 (1-214)	1.9	0.1	45.7	4.3	
FR-HMGB1A-c037 (1-88)	0.2	0.2	407.3	548.5	R^2=0.61
"FR"-HMGB1A-c038 (89-174)	0.1	1.3e-002	5.8	5.5	R^2=0.4
"FR" 3S-HMGB1A-c006 (1-214)	1.8	2.9e-002	6.1	0.5	

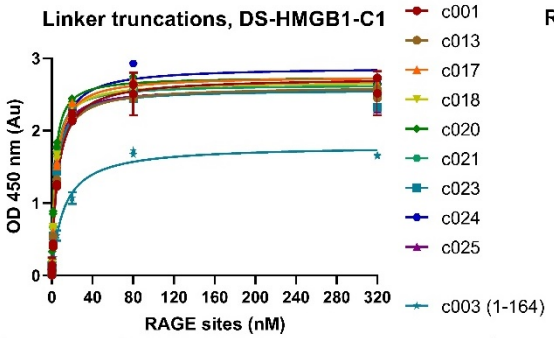
Figure 5.8: RAGE-Fc binding in ELISA of HMGB1-FL (FR, DS and 3S) and full HMG Box constructs

A. Michaelis-Menten plots. "FR" or "DS" designates reduction state. Box B (no disulfide possibility) was also assayed in reducing (+2 mM TCEP) or oxidising conditions. 4 replicate wells per concentration and construct. **B.** Fitted K_d constants (nM) for RAGE binding to different redox forms of HMGB1. Maximum binding (B_{max}) was roughly equivalent for all fitted curves (50% chance of difference); only FR-HMGB1A-C1-c001 had a different K_d from DS-HMGB1). Flagged: Adjusted R^2 < 0.65; all other fits, > 0.95. **Red asterisk**: The Michaelis-Menten plot for FR-HMGB1-C1-c001 shows that binding did not reach full saturation (plateau).

5.4.II Screen of first batch of HMGB1 constructs

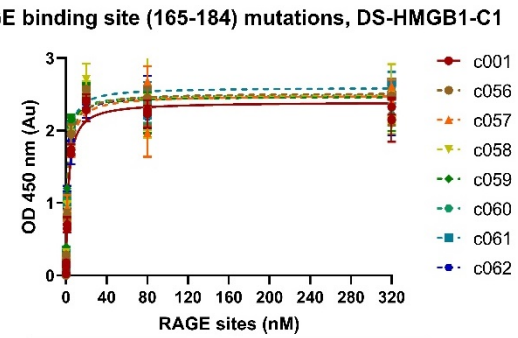
As explained above, some of the RAGE-deficient and varying linker-length constructs had not arrived by the time I conducted these experiments. Therefore, my initial screen was incomplete as it only included the ones to which I had access at the time (Figure 5.9). Neither the pair-wise linker truncations I had available for this experiment (C1-c013, c017, c018, c020, c021) nor replacement of segments of the linker by GSGSG (C1-c023-c025) influenced RAGE binding. The seven RAGE binding site mutants I had at my disposal (C1-c056-c062) had no effect either. As expected, HMGB1 1-164 (C1-c003) exhibited decreased maximum RAGE binding and affinity for RAGE due to the truncation of the 149-182 RAGE binding region, although . As I already knew that this construct had regenerative activity (Figure 3.6), it validated the hypothesis of G_{Alert} induction via HMGB1 being RAGE-independent, despite RAGE-HMGB1 signalling resulting in CXCL12 production [97].

A.



	Bmax			Kd (nM)		
	Value	+Error	-Error	Value	+Error	-Error
c001	2.73	0.09	0.09	5.90	0.93	0.81
c013	2.61	0.08	0.08	4.92	0.73	0.64
c017	2.75	0.08	0.07	4.12	0.58	0.51
c018	2.68	0.08	0.08	3.54	0.56	0.49
c020	2.74	0.05	0.05	2.67	0.26	0.24
c021	2.64	0.08	0.08	3.07	0.49	0.43
c023	2.57	0.10	0.10	3.90	0.82	0.68
c024	2.88	0.20	0.20	4.93	1.92	1.40
c025	2.58	0.10	0.10	3.62	0.73	0.61
c003 (1-164)	1.79	0.07	0.07	11.47	2.08	1.79

B.



	Bmax			Kd (nM)		
	Value	+Error	-Error	Value	+Error	-Error
c001	2.40	0.11	0.10	2.49	0.59	0.48
c056	2.52	0.12	0.12	2.01	0.52	0.42
c057	2.49	0.12	0.12	2.27	0.64	0.51
c058	2.50	0.16	0.15	1.73	0.63	0.47
c059	2.47	0.12	0.11	1.27	0.35	0.28
c060	2.50	0.10	0.10	1.42	0.32	0.27
c061	2.59	0.09	0.09	1.68	0.34	0.28
c062	2.48	0.11	0.11	1.60	0.40	0.33

C.

	Global (shared)
Comparison of Fits	
Simpler model	Kd same for all data sets
Probability it is correct	<0.01%
Alternative model	Kd different for each data set
Probability it is correct	>99.99%
Ratio of probabilities	
Preferred model	Kd different for each data set
Difference in AICc	23

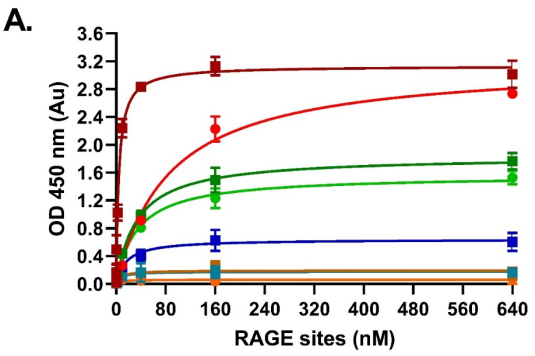
	Global (shared)
Comparison of Fits	
Simpler model	Bmax same for all data sets
Probability it is correct	<0.01%
Alternative model	Bmax different for each data set
Probability it is correct	>99.99%
Ratio of probabilities	
Preferred model	Bmax different for each data set
Difference in AICc	99

Figure 5.9: RAGE-Fc binding in ELISA of HMGB1 linker and RAGE-site mutated constructs

Michaelis-Menten plots for binding of RAGE-Fc to ELISA plates coated with 40 nM of HMGB1-C1-DS constructs (mean ± SD, baseline-corrected), and computed maximum binding/Kd constants (mean with asymmetric confidence intervals). 4 replicate wells per concentration and construct. **A.** Linker truncations and control constructs. **B.** RAGE site mutants: all Kd values seem to be off by a factor of approximately two (seen in the c001 control construct), hence the need for inclusion of a wild-type control in all experiments. **C.** Statistical comparison (AICc) between c001 (full-length wild type) and c003 (1-164 truncation).

Despite the main RAGE binding site being in the 149-182 fragment, DS-HMGB1 is the isoform assumed to effect prothrombotic signalling through RAGE. This led me to consider whether 1) HMGB1 1-184 (C1-c002) would bind RAGE any differently from 1-164 (C1-c003) due to the presence of the full RAGE binding epitope and 2) what would happen with a construct with two HMG Box domains but Box A replaced by Box B (C1-c048), and therefore unable to have this disulfide, as the C-terminal tail of HMGB1 is known to modulate the binding to other PAMPs such as TLR-2 [49] by preventing it in the absence of ligand. For this assay, I increased the concentration of ligand up to 640 nM to improve fit quality for FR-HMGB1 by reaching binding saturation.

The disulfide bridge in Box A seems to be key for stabilising the interaction between RAGE and HMGB1 (Figure 5.10). Whilst FL-FR-HMGB1 bound similar amounts of RAGE as FL-DS-HMGB1 at saturation, the affinity of the reduced form is markedly lower. Even though total saturation was still not completely achieved (at 640 nM ligand), affinity upon oxidation is greatly increased (K_d decreased from 86.11 to 4.35 nM [64]), which would likely explain the loss of prothrombotic activity with the reduced protein. HMGB1 1-164 binding to RAGE is affected by oxidation to a lesser extent - whereas the behaviour of the oxidised form could be fitted to a Michaelis-Menten model, bound RAGE at equilibrium for FR-HMGB1 1-164 was almost nil.



	Bmax		Kd (nM)	
	Value	SEM	Value	SEM
DS HMGB1-C1-c001 (1-214)	3.13	0.05	4.35	0.35
DS HMGB1-C1-c002 (1-184)	1.84	0.05	33.87	3.54
DS HMGB1-C1-c003 (1-164)	0.64	0.04	16.56	4.24
"DS" HMGB1-C1-c048 (DbB)	0.19	0.02	7.98	4.37
FR HMGB1-C1-c001 (1-214)	3.19	0.08	86.11	7.06
FR HMGB1-C1-c002 (1-184)	1.56	0.04	34.37	3.21
FR HMGB1-C1-c003 (1-164)	0.18	0.02	7.71	4.57
"FR" HMGB1-C1-c048 (DbB)	0.06	0.01	6.66	7.44

Figure 5.10: RAGE-Fc binding in ELISA to tail-truncations of HMGB1 and Double Box B HMGB1

Michaelis-Menten plots of RAGE-Fc binding to HMGB1 full-length, 1-184 and 1-164 (both FR and DS), plus double box B construct (c048; with and without 2 mM TCEP) (mean ± SD, baseline-corrected), and computed B_{max} and K_d constants (mean with asymmetric confidence intervals). Whilst the quality of the fit for FR-HMGB1-FL improved, saturation was still not reached. c003-FR and both c048 conditions did not bind RAGE-Fc, resulting in flagged (incomparable) fits. 2-way ANOVA post-hoc analysis reflected this as well. 4 replicate wells per concentration and construct.

HMGB1 1-164 (DS) displayed much lower B_{max} in this assay (Figure 5.10) than it did in the previous one (Figure 5.9), although K_d values were comparable (16.5 vs 11.5 nM), which might be due to nonspecific binding in the first assay, or technical variation (e.g. washing).[51] HMGB1 1-184, in contrast, bound RAGE to a much greater extent than 1-164 but still less than full-length constructs. Unlike full-length HMGB1, both the reduced and oxidised forms of HMGB1 1-184 were able to bind RAGE with equal K_d, and similar B_{max}, and this binding is stronger in both cases than the one to FR-HMGB1 full-length. This is probably similar to the phenomenon described for HMGB1 1-184 and TLR-2-Fc binding [49], where tail-truncated HMGB1 could bind, but not signal, through TLR-2 in the absence of ligand, whereas FL-HMGB1 (either FR or DS) could only bind in the presence of TLR-2 ligand. This could suggest that the effect of the disulfide bridge in RAGE binding is related, at least partially, to the interaction of the acidic tail and Box A when oxidised.

HMGB1A-C1-c048 (double Box B) behaved in the same manner as would be expected of Box B alone. This construct lacks both part of the RAGE-binding site and the disulfide bridge of Box A and therefore RAGE binding should be greatly decreased. Accordingly, the curves for this construct could not be fitted to a Michaelis-Menten model whether TCEP was present or not. The lack of effect of TCEP on RAGE binding is logical as this construct has no disulfide bridges and, therefore, it would not be expected to be sensitive to reducing conditions.

To validate these data, I analysed the binding of my two RAGE-deficient constructs (DS-HMGB1 1-164, HMGB1-C1-c048), both redox states of HMGB1, 3S-HMGB1, and disulfide Box A and HMGB1 1-184 via BLI (Figure 5.11). To do this, I used the same RAGE-Fc construct applied in my ELISA, but this time RAGE-Fc was immobilised via anti-IgG sensors. This was suboptimal as the mass difference for the ligand and RAGE would be very small (25 kDa vs. 120+ kDa) and therefore signal-to-noise ratio would be reduced, but RAGE-Fc had too much nonspecific binding to SA sensors in pilot experiments to utilise it as a soluble ligand. I also included 3S-HMGB1 (HMGB1A-c006) as my ELISA data did not agree with the published SPR data. I also included DS Box A (1-88) as a negative control based on my ELISA data.

[51] When 2 mM TCEP was present in the blocking and coating steps, absorbance of the blank (no HMGB1) was slightly lower than when TCEP was absent (0.15/0.3 Au); this does not account for the differences except with c048 and 1-164 constructs.

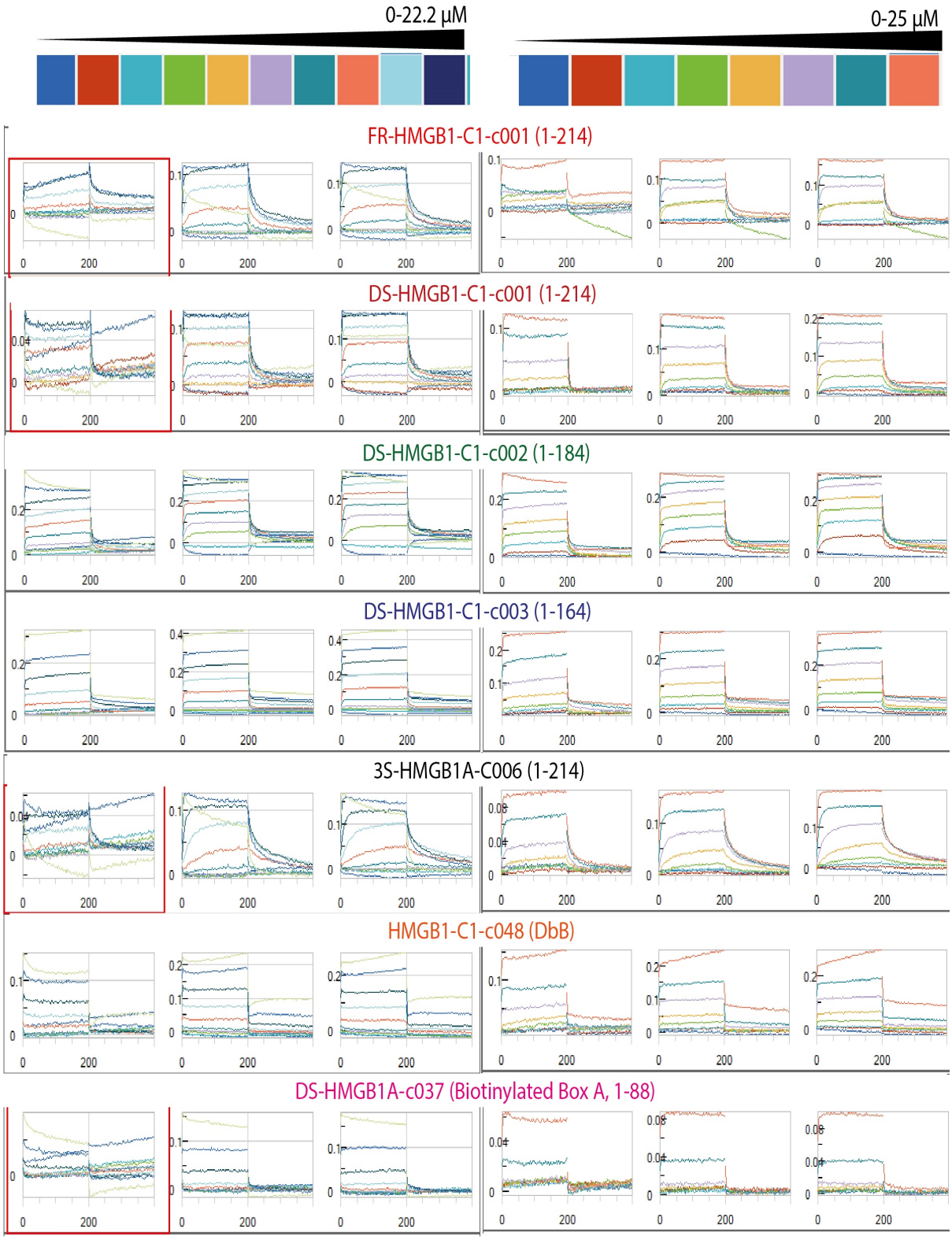
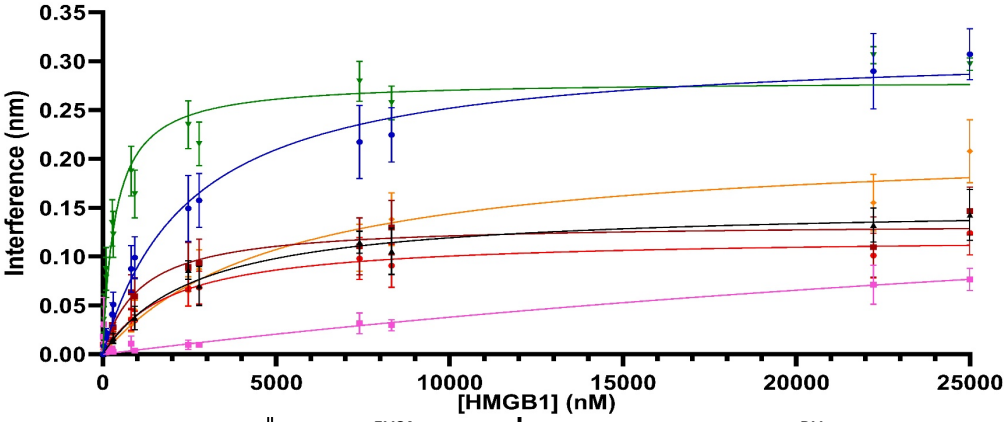


Figure 5.11: Interferograms in BLI of HMGB1 constructs' binding to RAGE

RAGE-Fc was immobilised in the surface of AHC sensors and dipped in rising concentration of different HMGB1 constructs. Two experiments were run with different concentration ranges: the three columns of graphs in the left, 0 to 22.22 µM HMGB1 over 9 steps; on the right, 0 to 25 µM over 7 steps. Colours indicate the specific construct concentration. Each line of graphs corresponds to all replicates for a given HMGB1 construct: red, FR FL-HMGB1; dark red, DS FL-HMGB1; dark green, DS HMGB1 1-184; dark blue, DS-HMGB1 1-164; black, 3S-FL HMGB1; orange, HMGB1-C1-c048; pink, FR-HMGB1 Box A 1-88 (c037). *Interferograms surrounded by a red rectangle had data points excluded due to poor quality (e.g. drift)*

Binding of RAGE to all DS-HMGB1 constructs (Figure 5.12) is of higher affinity due to a faster association rate for RAGE. DS-HMGB1 1-164 had a decreased affinity for RAGE compared to DS-FL-HMGB1, and HMGB1-C1-c048 had the least RAGE affinity of all tested HMGB1 constructs with two HMG Box domains. This is due to both constructs having increased k_{off} compared to the rest; despite the also increased k_{on}, this leads to a large reduction in RAGE affinity as the proportional increase in dissociation rate is larger. FL 3S-HMGB1, despite having equivalent K_d for RAGE than FL FR-HMGB1 FL, has a much slower RAGE binding equilibrium (large reduction in both k_{off} and k_{on} by an equivalent factor). The K_d for RAGE binding to FL-DS-HMGB1 I measure (0.2 μM from kinetics rates and 1.3 μM from saturation fits) are in the range of the previously reported values via SPR (0.1 from kinetics rates [180] and 0.65 μM from saturation fits [64]).



	Calculation method	ELISA B_{max} (nM)	ELISA K_d (nM)	K_d (μM) Saturation fit	K_d (μM) Kinetics plot	BLI k_{On} ($\mu M^{-1}s^{-1}$) Kinetics plot	k_{Off} (s^{-1}) Kinetics plot
	DS-HMGB1-C1-C001	1±0.02	4.35±0.35	1.3±0.2	0.2±0.04	0.215±0.038	0.043±0.003
	FR-HMGB1-C1-C001	1±0.02	86.11±7.06	2.3±0.4	0.8±0.27	0.051±0.011	0.041±0.01
	HMGB1-C1-C048	0.06±0.01	**Invalid fit**	5.5±0.7	1.3±0.33	0.164±0.034	0.206±0.033
	DS-HMGB1A-C037	0.08±0.02	**Invalid fit**	53.4±8.5	15.6±5.87	0.028±0.009	0.439±0.089
	3S-HMGB1A-C006	1.05±0.03	7.22±0.97	3±0.5	0.9±0.11	0.017±0.002	0.015±0.001
	DS-HMGB1-C1-C002	0.59±0.02	33.87±3.54	0.5±0.1	0.2±0.04	0.349±0.059	0.067±0.01
	DS-HMGB1-C1-C003	0.2±0.01	**Invalid fit**	3.1±0.3	0.9±0.26	0.178±0.037	0.152±0.035

Figure 5.12: BLI validation of RAGE binding to HMGB1 constructs

Biolayer interferometry kinetic plots of different HMGB1 constructs binding to RAGE-Fc. K_d constants derived from the saturation fit (Michaelis-Menten equilibrium binding, nonlinear least squares) or from the direct measurement of kinetics in the instrument (k_{Off}/k_{On}); all values in mean±SEM. Values from ELISA are also shown. AICc informed of differences for all parameters between constructs; B_{Max} excluded as AHC sensors are not certified for quantitation. Post-hoc comparisons, after excluding data for HMGB1A-c037 (SD too high): 6 sensors used per construct across two concentration sets (Figure 5.11); some had to be excluded due to drift

- ***K_d*, 2-way ANOVA:** Higher affinity for DS-HMGB1-c1-c001 and all other constructs save DS-HMGB1-C1-c002 ($p < 0.01$ or lower). HMGB1-C1- c048 and DS HMGB1-C1-c003 had lower affinity for RAGE ($p < 0.0001$). FR and 3S HMGB1 (c001/c006) had equivalent RAGE-Fc affinity.
- ***k_{On}*, Brown-Forsythe/Welch ANOVA test:** Association rates for DS-HMGB1-C1-c001/c002/c003 and HMGB1A-c048 are equivalent and different ($p< 0.05$) from those for FR/3S HMGB1 FL.
- ***k_{Off}*, Brown-Forsythe/Welch ANOVA test:** Dissociation rates for FR/DS-HMGB1-C1-c001 were equivalent and slower than those for DS-HMGB1 c002, c003, and c048; the latter two were equivalent between each other. 3S-HMGB1 had significantly slower dissociation rate than all other constructs.

These results confirm the observations from ELISA; the only discrepancies would be FL-3S-HMGB1 and DS HMGB1 1-184, due to the different methodologies for quantification of binding in each assay. BLI is a “true“ equilibrium, where association and dissociation steps are measured in real-time and equally weighted in computing affinity. In ELISA, the effect of dissociation rates is more prevalent because of the extensive dissociation steps required after the association step. In fact, the B_{max} values from ELISA follow the exact same pattern as the BLI k_{off} measurements. In BLI, the effect of the different association and dissociation rates can be deconvoluted from each other and individual effects of each in global affinity become observable.

FL 3S-HMGB1 bound RAGE in BLI with equivalent K_d to FL FR-HMGB1, rather than to FL DS-HMGB1 as in ELISA, but with much slower kinetic rates (slower equilibrium). Assuming all constructs reach saturation equilibrium during the ELISA association step (equal B_{max} after association) and that all redox HMGB1 forms bind equivalent amounts of RAGE at equilibrium if fully saturated (before dissociation), the decreased k_{off} of FL-3S-HMGB1 allows for better RAGE retention than either FR or DS FL-HMGB1 (dissociation during the experiment is slower). This inflates the ELISA measured K_d values approach those for DS-HMGB1. This apparent high-affinity, low dissociation rate binding was also reported for 3S-HMGB1 in the published SPR assays [271]. Conversely, DS HMGB1 1-184 bound RAGE in BLI with even greater affinity than full-length DS-HMGB1, opposite to the ELISA results. This is due to a large increase in k_{on} over FL DS-HMGB1, but it also has slightly higher k_{off} than FL DS-HMGB1, which decreases the amount of remaining bound RAGE by the time it´s measured, apparently decreasing its affinity for it in ELISA. Therefore, loss of the C-terminal RAGE binding site drastically increases dissociation of RAGE from HMGB1, whereas the disulfide bridge in Box A increases the association of both proteins. However, disulfide Box A alone is incapable of effectively binding RAGE, having the lowest overall affinity, fastest dissociation rate, and second slowest association rate amongst all constructs, not even reaching equilibrium in the experiment (calculated K_d is artificially lower than reality as a result). This could confirm my hypothesis of the acidic tail in FL-HMGB1 interfering with RAGE binding as part of a regulatory mechanism in concert with the oxidation state of the protein.[52]

[52] Concentrations above 25 µM resulted in excessive nonspecific binding to the sensor surface and had to be excluded.

5.4.III Selected RAGE-deficient constructs

Whilst all the mutations and designed truncations of the linker failed to modify RAGE binding of HMGB1 in my first batch of tests, I could identify HMGB1A-C1-c048 (DbB) as a construct with significantly impaired RAGE binding for further testing. These experiments also provide further insight on the mechanism of HMGB1/RAGE :

- Optimal binding requires a full-length HMGB1-like protein structure with two HMG boxes: Box A and Box B alone have very decreased RAGE binding capacity when compared to FL-HMGB1. Whilst Box A binding to RAGE has been shown *in vitro*, and Box A can inhibit HMGB1/RAGE mediated signalling [258], in both cases a significant excess of Box A is required compared to HMGB1, representing this weakened binding.
- The disulfide bridge in Box A is also essential for this interaction, greatly increasing the association rate of HMGB1 for RAGE when present and explaining why DS-HMGB1 is prothrombotic whereas FR or 3S are not [64]. Reduction of Box A in FL-HMGB1 decreases its affinity for RAGE (lower k_{off} and higher K_d), but it does not decrease total binding capacity (B_{max}) at saturation in the ELISA experiments as the dissociation rates remain unaffected by oxidation. It could be that the presence of a second HMG Box is needed in combination with oxidation to expose the sequences in Box A for RAGE binding, as peptides within Box A have been shown to bind RAGE [259,260]. In a construct with two tandem B-boxes (HMGB1-C1-c048) with RAGE binding peptides, association rate is increased but RAGE cannot remain bound to this construct due to the lack of interaction with DS-Box A.
- Removal of the C-terminal acidic tail (HMGB1 1-184) leads to enhanced binding to RAGE. It is possible that the binding of the acidic tail to the Box A core is involved in modulation of RAGE binding. In BLI, a tail-less oxidised HMGB1 construct has even greater affinity for RAGE than FL-HMGB1, unless the RAGE binding motif itself is truncated, due to much faster HMGB1-RAGE association rate. This has been proposed before [196], as some of the residues involved in acid tail interaction (e.g. R162) [160] are also involved in RAGE binding [39]. A construct with both Box A and B, but no 165-184 peptide has greatly reduced RAGE binding due to increased dissociation rate, as the key RAGE binding epitope is now absent.

5.5 TLR-4 and TLR-2 activity screen

HMGB1 binds TLR-4, TLR-2 and RAGE; all three routes ultimately converge in the NF-κβ pathway [64,83,84], and consequently it is difficult to attribute downstream signalling to a given receptor amongst them. Therefore, I first studied TLR activity via reporter lines which expressed only one of the TLRs and cofactors [153] to identify receptor-specific signalling, followed by a more representative model in monocytes, capable of expressing and responding through all three receptors[53].

FR-HMGB1 and HMGB1A-C1-c048 (DbB-HMGB1) did not induce NF-κβ proinflammatory signalling via TLR-2 (Figure 5.13A) or TLR-4 (Figure 5.13B) in reporter cells (via NF-κβ activation), as well as TNF production in macrophages (Figure 5.13C), DS-HMGB1 induced significant responses via these receptors in both models. The proinflammatory properties of HMGB1A-C1-c048 are equivalent to those of reduced HMGB1 (i.e., nil) whilst being incapable of shifting to a pro-inflammatory phenotype due to the lack of a disulfide bond within Box A. Observing TLR-2 signalling from DS-HMGB1 alone was somewhat puzzling, given the contradictory evidence in the literature for [254,255,271] and against [49,59] the requirement of a co-ligand. It is possible that within HI-FBS there are proteins capable of substituting LTA as a TLR-2 co-ligand for HMGB1.

DS-HMGB1 did also induce proinflammatory signalling in monocytes. Addition of suboptimal amounts of LTA to DS-HMGB1 and preincubating to allow complex formation resulted in a synergistic effect, producing more TNF than the sum of the levels reached with either treatment alone, agreeing with the published data on the synergy between these molecules [272]; FR-HMGB1 and DbB-HMGB1 did not induce TNF secretion over baseline, whether alone or synergistically with LTA. Interestingly, FR-HMGB1 and DbB-HMGB1 inhibited the response to LPS when macrophages were stimulated pre-incubated HMGB1/LPS complexes, whereas DS-HMGB1-LPS complexes induced a similar response to LPS alone.

[53] Due to limited laboratory access because of the Covid-19 pandemic, I was unable become proficient in monocyte culture techniques, and had to greatly reduce the number of tested HMGB1-C1 constructs. Experiments with these cells were the work of fellow DPhil student, Tzung Yuan (Carl) Lee, at the Nanchahal Group using my recombinant HMGB1 proteins and are reproduced with his permission. DS-HMGB1 had to also be purchased commercially due to the same restrictions for these assays.

This is likely due to the function of HMGB1 as an LPS co-receptor, substituting LPS binding protein (LBP) [23,52] which is likely present in the serum from the media [273]. Whereas DS-HMGB1 can induce proinflammatory signalling, and thus LPS-DS HMGB1 complexes are still active, LPS bound to either FR or DbB-HMGB1 is likely inactivated as neither has the correct conformation for TLR-4 binding, since an oxidised box A is missing, and binding cannot be initiated. It is not known whether LPS binding to HMGB1 is reliant on the redox state; however, biotinylated LPS is commercially available and therefore binding kinetics of LPS to DS-HMGB1, FR-HMGB1 and DbB-HMGB1 could be analysed via BLI to assess whether these constructs are indeed capable of specifically binding LPS, despite the absence of canonical LPS binding sites within DbB-HMGB1.

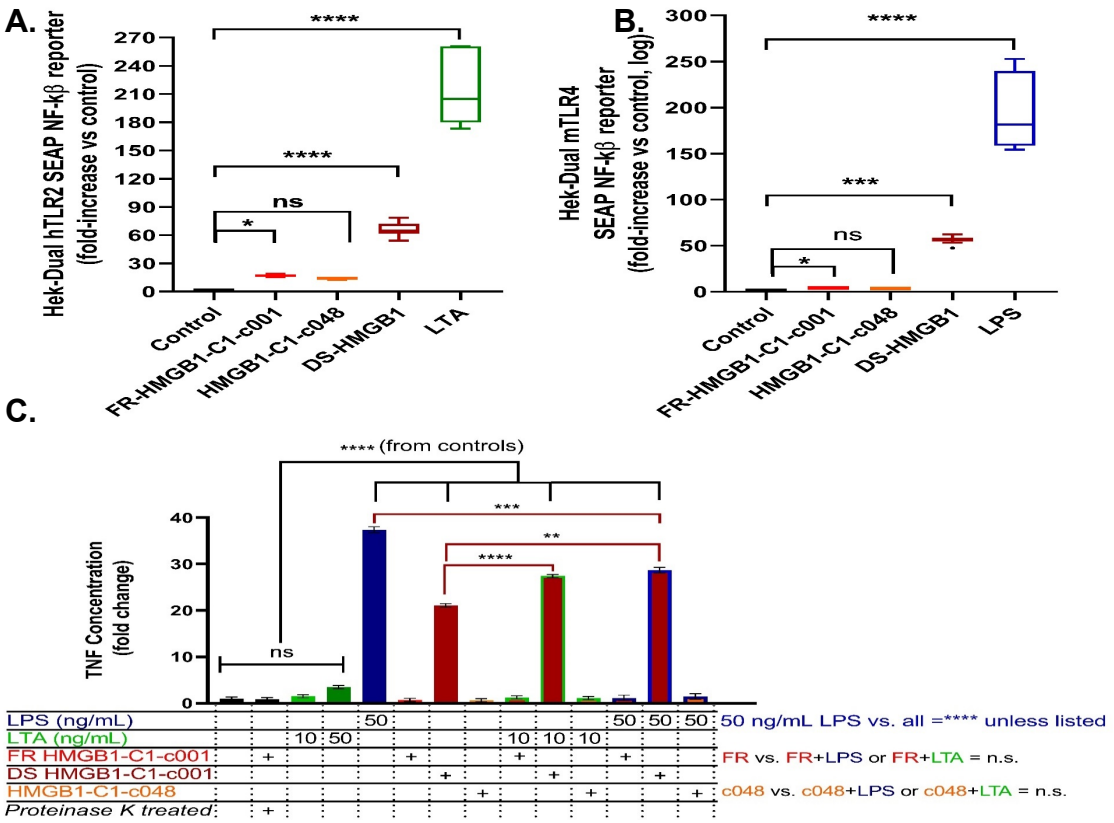


Figure 5.13: TLR signalling in response to recombinant FR-HMGB1-C1-c001, HMGB1-C1-c048 and commercial DS-HMGB1

(**A, B)** DS-HMGB1 induces NF-κβ activity in reporter HEK-Dual cells expressing human TLR-2 and CD14 (A) or murine TLR-4, MD-2 and CD14 (B). DbB-HMGB1 did not induce significant amounts of signalling, whereas FR-HMGB1 only induced minor NF-κβ activation in both cell lines, potentially due to minor oxidation during preincubation or assay. Values are shown as median fold change compared to control (media alone); Kruskal-Wallis test. N= 8 wells per condition across two experiments. **(C)** Disulfide HMGB1 (DS-HMGB1) increases TNF production in monocytes, which is further enhanced by the presence of suboptimal amounts of LTA, but not of LPS. Both FR-HMGB1 and DbB-HMGB1 were unable to elicit monocyte TNF secretion, even when pre-incubated for 24h with LPS or LTA; response to LPS pre-incubated with these constructs was also significantly reduced. n= 3 donors with 3 technical replicates each. Data analysed by nested ANOVA. All HMGB1 constructs were added at a dose equivalent to 10 µg/ml of FL-HMGB1 (0.39 nM).

An alternative hypothesis for the decrease in LPS-mediated signalling seen with my HMGB1 preparations (FR and dBB) may be the presence of TLR inhibitors in the buffer components, such as PEG contamination as a residue from LPS removal [154], although this is unlikely as I verified their absence using ESI/MS (2.13). Another potential explanation is that the presence of the C-terminal TEV scar in my protein, which was not present in the commercially available DS-HMGB1, interferes with TLR-4 binding. The C-terminus of the acidic tail has however not been described previously as involved in this interaction, and the acidic tail is completely absent in DbB-HMGB1, which would make this unlikely to be the cause. During optimisation of the HEK experiments, direct addition of my recombinant HMGB1 in combination with LPS or LTA without preincubation did not lead to inhibition of either PAMP, which would rule out direct TLR-4 or TLR-2 inhibition by a buffer component. A control with LPS pre-incubated with proteinase K treated HMGB1 would confirm whether any buffer component is inhibiting LPS-mediated TLR-4 signalling.

Based on my findings that the acid tail is dispensable for the regenerative activity of HMGB1 (Figure 3.15) and that each of the Box B domains in DbB-HMGB1 are able to bind CXCL12 independently (Figure 4.12), I proceeded to test this construct for regenerative activity rather than completing the TLR screening of all my constructs.

5.6 Regenerative effect of Double box B (DbB) construct -HMGB1A-C1-c048

I had now identified two HMGB1 constructs with reduced RAGE signalling (HMGB1A-C1-c003, 1-164, and HMGB1A-c1-c048, double box B), one of which (c048) also had reduced TLR-2 and TLR-4 signalling. Whilst FR-HMGB1A-c1-c003 was known to be biologically active (as it is identical to FR-HMGB1A-c031), HMGB1A-c1-c048 had to be screened, which led me to re-evaluate my methodology.

Alone, the *in vivo* muscle regeneration suffers critical limitations for screening purposes as it is not suitable for high-throughput and does not comply with the NC3R objectives for animal welfare [274]. Whilst an *in vitro* models for skeletal muscle regeneration models have been described via measurement of mitochondrial activity [275] or direct myoblast fusion via split-GFP technology [164], I found that they are inadequate for HMGB1 early in my DPhil (Figure 5.14). When pre-treating the cells before addition of differentiation medium with active (FR-HMGB1 1-214, 1-184, and 1-164) and inactive (FR-HMGB1 9-214) constructs, neither resulted in increased differentiation. Furthermore, even control cells underwent spontaneous differentiation to the same extent as stimulated cells. This would suggest that this cell line is already activated in the proliferative state and committed to myogenesis, and not a suitable model.

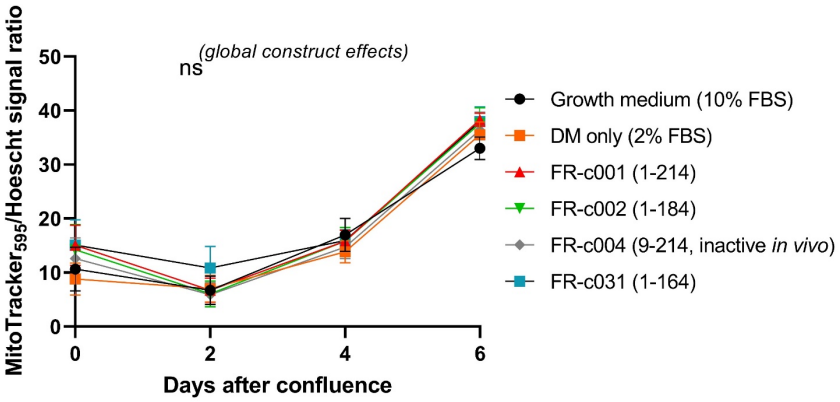


Figure 5.14: Response of C2C12 differentiation to HMGB1

A. Effect of the different HMGB1 constructs, administered 1 day before full confluence and stimulation with differentiation medium, in myocyte C2C12 differentiation; neither the differentiation medium nor HMGB1 increased differentiation into fused, tubular myocytes compared to basic growth medium. Cells untreated with differentiation media experienced the same increase in mitochondrial activity (ratio of MitoTracker to Hoescht DNA fluorescence stain)

Progenitor cells have to be alerted by HMGB1 prior to activation and cell cycle entry in order to enhance tissue repair [25]. This would explain the lack of effect, reproducing the effects observed in my time course experiment (Figure 4.1) where after 5 hours post-injury FR-HMGB1 injection did not enhance regeneration. Furthermore, it only reflects one of the multiple tissues where HMGB1 can induce enhanced regeneration [25]. Therefore, an alternative was needed.

Stem cells are the "gold standard" for tissue differentiation models, albeit extremely costly and complex to work with.[54] However, one key advantage of established pluripotent human embryonic stem cell lines such as H9 is that they can reliably differentiate into a wide variety of tissues. Cardiomyocyte differentiation was chosen as a model given its clinical significance, whilst being related to our *in vivo* skeletal muscle injury model.

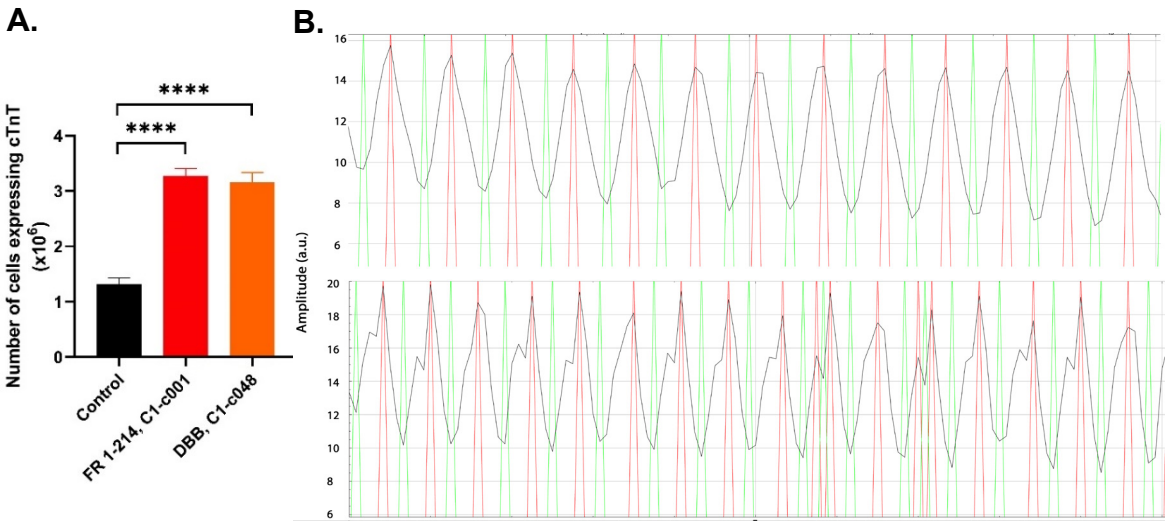


Figure 5.15: H9 differentiation assay with FR-HMGB1-C1-c001 (1-214) and DbB-HMGB1A-C1-c048

H9 embryonic stem cells were stimulated with 100 nM FR-HMGB1 full-length (2.5 µg/mL) or DbB-HMGB1 (2 µg/mL) for 24 hours prior to the start of the differentiation process with StemDiff Cardiomyocyte Kit. **A.** At day 8, the cells were lifted and cardiotroponin expression (cTnT) was analysed (ELISA). **B.** Contractile potential was measured using Myocyter (ImageJ plugin) on raw video of the cells in differentiation: FR-HMGB1 treated cells contract at 2.6±0.22 Hz (average of 62 beats) whereas control (untreated) cells contract at 1.5±0.27 Hz. DbB could not be imaged due to the Covid-19 lockdown.

[54] I would like to acknowledge the essential contribution of fellow DPhil student Tzung Yuan (Carl) Lee, who set up the model in our lab and provided invaluable technical help with performing this assay, as well as the FACS analysis. He, and not I, adapted the protocols existing in the literature [289–291] to the needs in our group.

When resting H9 stem cells are stimulated with either FR-HMGB1 1-214, or DbB-HMGB1, prior to the induction of cardiomyogenesis, two pronounced effects can be seen with both constructs (Figure 5.15). Firstly, expression of specific cardiac cell markers greatly increases; cardiac troponin-expressing cells almost double at day 6 upon HMGB1 treatment, where the cardiomyocytes begin to mature. Additionally, HMGB1-treated cells begin contracting much earlier and at higher frequency (≈1 Hz in treated cells vs. ≈0.5 Hz in untreated) [276], indicating not only a biochemical difference in the differentiation process, but also a functional effect on the contractile potential of the cardiomyocytes.

Whilst this is not entirely representative of mature human heart - regeneration of which is mediated by mature cardiomyocytes, rather than stem cells [277] – other researchers in our group are in the process of analysing this effect in both a murine MI model, as well as reproduced this effect in mature cardiomyocytes which are subjected to damage via doxycycline, where FR-HMGB1 has a protective and tissue repair enhancement effect, with HMGB1A-C1-c048.

Having found that the DbB construct promoted cardiomyocyte proliferation and differentiation *in vitro*, I proceed to test it *in vivo*. To better assess whether DbB-HMGB1 is comparable to FR-HMGB1in promoting regeneration , I reproduced the time course experiments previously published from my group [25][55] (Figure 5.16). In this study, HMGB1 was shown to accelerate skeletal muscle regeneration from day 14 onwards, with a more modest increase at the first week, and a non-significant difference with the control group at day 28, when the process has plateaued . Animals were injured and the muscle cell cross-section area was quantified at 6, 14, and 28 days post-injury. The results are summarised in Figure 5.16, which shows that FR-HMGB1A-c001 (full-length wild type protein) and HMGB1A-C1-c048 (DbB) accelerated repair at day 14 to equivalent levels.

[55] C57BL/6 mice have a propensity for overgrooming to the point that they can cause injuries in the skin. These could act as priming injuries [99], priming cells in the alert phase and masking the effects of HMGB1 as a result. As a result, animals had to be routinely excluded from these analyses due to the presence of external injuries.

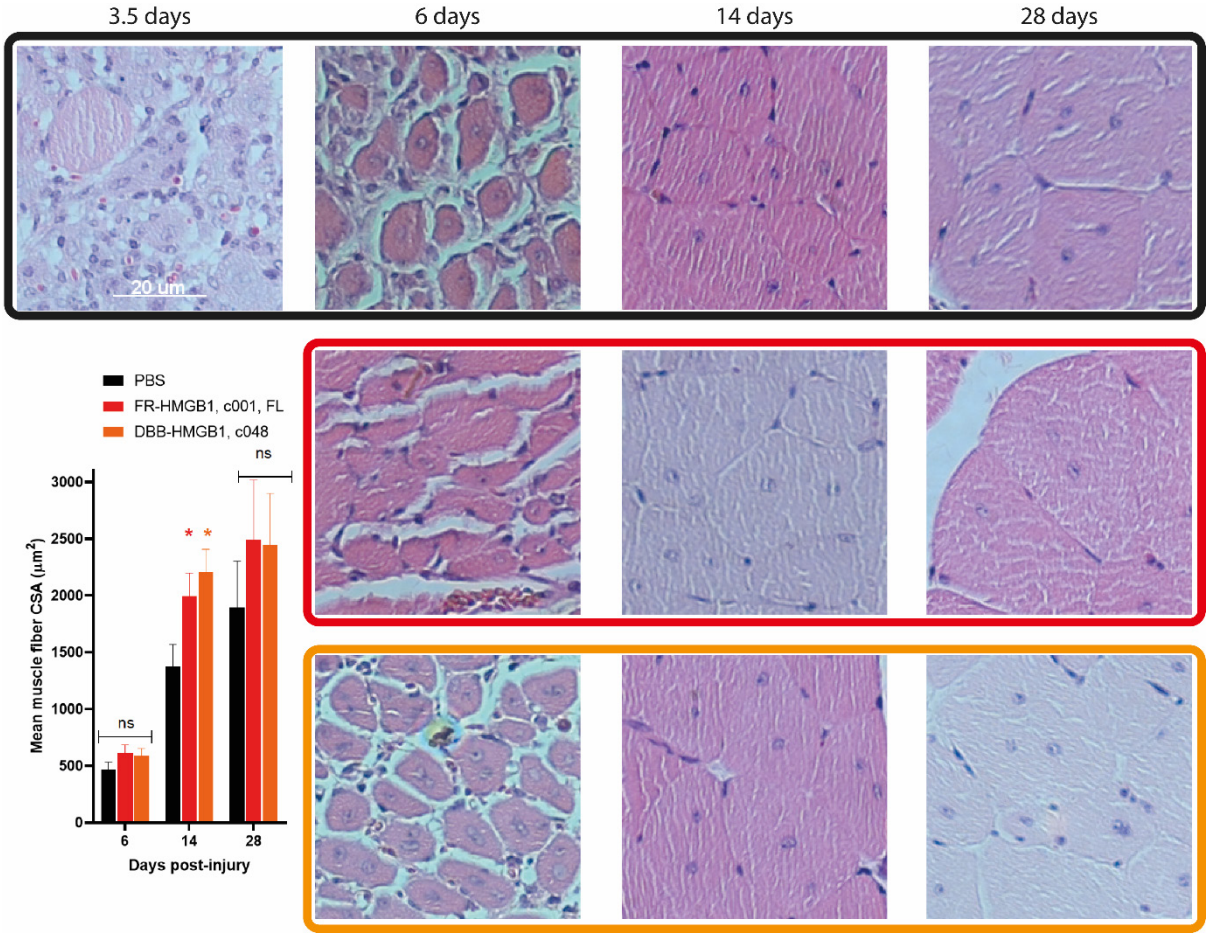


Figure 5.16: Time course of muscle cell regeneration induced by FR-HMGB1A-c001 and DbB-HMGB1 (HMGB1A-C1-c048)

Mean muscle cross-section area in animals injected intramuscularly with either PBS (black) 30 µg of FR-HMGB1A-c001 (red) or HMGB1A-C1-c048 (orange) and injured using intramuscular injection of barium chloride. 5 animals in each group were culled at 6, 14, and 28 days to measure muscle cross-sectional area; Data analysed by nested one-way ANOVA (Holm-Sidak post-hoc correction). A representative image of each point, as well as a control group harvested at day 3.5 (too early to show regeneration) is shown.

5.7 Stability of DbB-HMGB1

In Chapter 3, I analysed the surface area profiles and melting profiles of FL HMGB1 (c001), 1-184 (c002) and 1-164 (c031). I had also tested HMGB1A-C1-c048 (double box B) in parallel; here I summarise these results. [56]

5.7.I SEC profile of DbB compared to other HMGB1 constructs

The behaviour of HMGB1A-C1-c048 is like that of other tail-less HMGB1 constructs (HMGB1A-c002, c031; Figure 5.17 and Table 5.1). It eluted between c002 and c031 except at very low ionic strengths, which is to be expected given their molecular weights (1-184, 21.8 kDa; 1-164, 19.7 kDa; DbB, 20.1 kDa). Its behaviour in low salt concentrations was also like that of tail-less constructs, with retention volumes increasing much more than those for FL-HMGB1 at decreasing ionic strength, non-specific column binding at low salt concentrations, and proportionally larger SASA than HMGB1 1-164 at lower pH (ammonium acetate). However, this binding was much lower than that of 1-164 or 1-184 HMGB1, resulting in earlier elution at low ionic strength. For oxidisable constructs, oxidation of Box A appeared to increase nonspecific column binding (Figure 3.16 and Table 3.1); since DbB completely lacks Box A it would be reasonable to expect lower nonspecific column binding, even lower than either of the other HMGB1 tail truncations.

		Conductivity (mS/cm)	FR-HMGB1A, c001, FL	FR-HMGB1A, c002, 1-184	FR-HMGB1A, c031, 1-164	DbB-HMGB1A-C1-c048
			RV (mL)	RV (mL)	RV (mL)	RV (mL)
Amm. Ac, pH 6.8	50 mM	4.53±0.01	10.69±0.03	11.29±0.11	11.82±0.04	11.45±0.04
	60 mM	5.33±0.01	10.77±0.03	10.77±0.04	11.36±0.02	11.06±0.01
	200 mM	18.31±0.02	10.58±0.01	10.31±0.02	10.93±0.02	10.8±0.05
PBS, pH 7.4	0.33X	6.21±0.24	10.41±0.11	11.85±0.15	12.09±0.03	11.48±0.02
	1X	16.94±0.01	10.56±0.04	10.79±0.01	11.36±0.01	11.05±0.04
	2X	31.65±0.04	10.29±0.01	10.7±0.01	11.23±0.02	10.99±0.06

Table 5.1: Retention volumes for FR-HMGB1 constructs and DbB-HMGB1

Peak retention volume (in mL, mean±SEM) of each HMGB1 construct in size exclusion experiments shown in Figure 5.18. Each buffer was treated as a categorical variable due to small variations in conductivity. All other data were normalised by the ratio between conductivity of the run/average conductivity of all runs in that buffer. higher RV (≈ lower SASA) in purple; shorter RV (≈ larger SASA) in orange. Midpoint in green. Retention volumes averaged over three independent runs.

[56] For convenience, only data for FR-HMGB1 FL, 1-184 and 1-164 are shown in addition to DbB, even though the experiment shown in Figure 3.16 was performed in parallel.

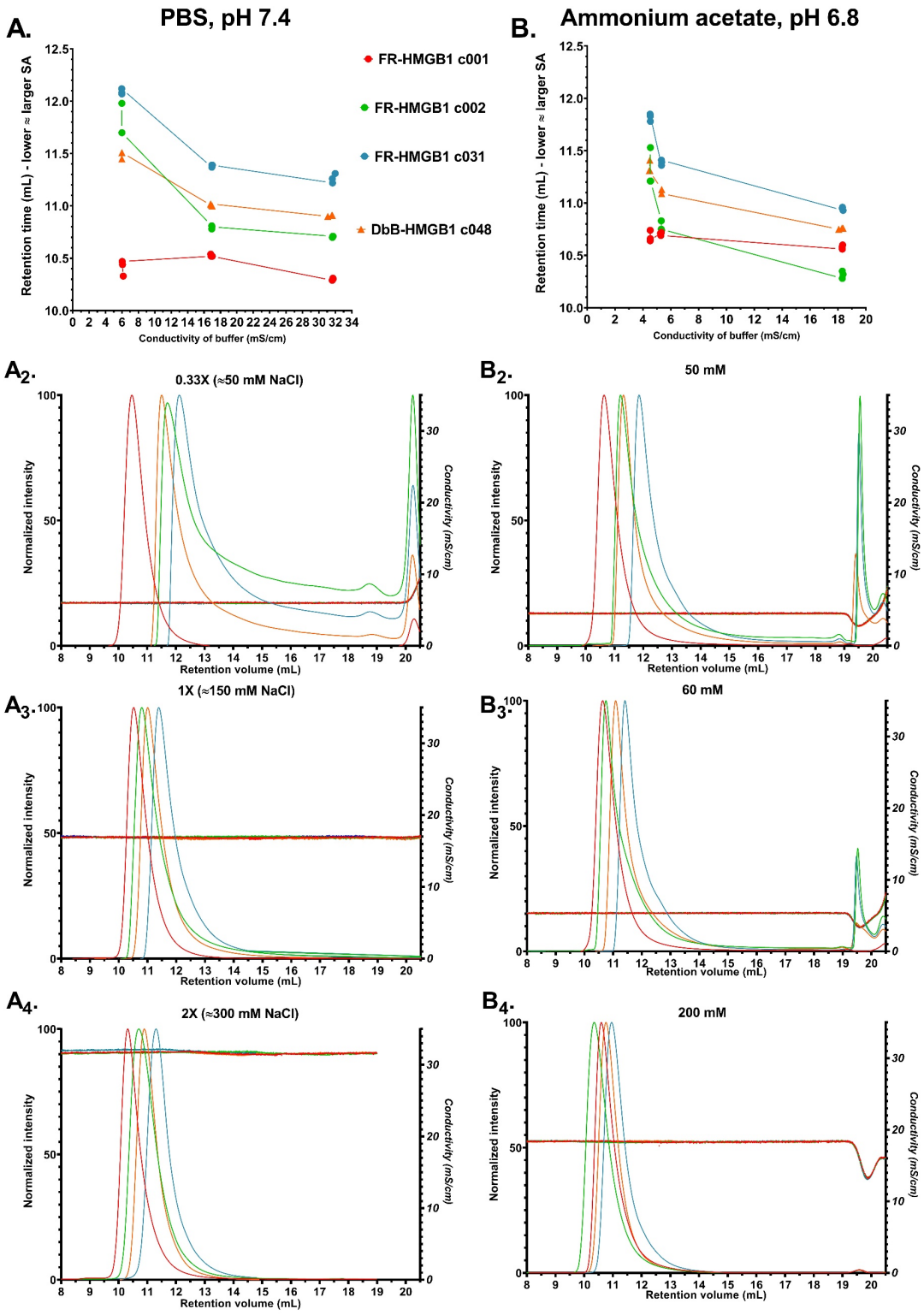


Figure 5.17: SEC of HMGB1 constructs in comparison with DbB-HMGB1

Retention volume analysis of HMGB1A-c001 (FL, red), c002 (1-184, green), c031 (1-164, blue), all reduced and HMGB1A-C1-c048 (DbB, orange). 0.3 mL of 1.25 mg/mL samples of each construct protein were injected into a Superdex 10/300 75 Increase column pre-equilibrated in buffer and eluted at 0.5 mL/min. **A.** Plot of retention volume versus average conductivity of each run (8-18 mL) in PBS buffers. **A2, A3, A4.** Representative SEC profiles of 0.33X, 1X, 2X PBS (pH 7.4) respectively, with conductivity plots (right axis) and normalised UV trace (left axis). **B.** Plot of retention volume versus average conductivity of the run (8-18 mL) in ammonium acetate buffers. **B2, B3, B4.** Representative SEC profiles of 50, 60, 200 mM ammonium acetate (pH 6.8), with conductivity plots (right axis) and normalised UV trace (left axis).

5.7.II ESI-QTOF/MS

In native ESI (0.2 M ammonium acetate; Figure 5.18), DbB-HMGB1 acquires surface charges in an analogous distribution to that seen in FR-HMGB1 1-164 (C1-c003/c031). Charge states for DbB-HMGB1 monomers are equivalent to that of HMGB1 1-164 monomers, but the relative abundance is different, with DbB-HMGB1 overall shifting towards higher charge states on the high M/Z envelope and lower charge states on the low M/Z envelope than HMGB1A 1-164, resulting in less variation between the potential SASA of each charge state envelope. Given the apparent higher rigidity of Box B compared to Box A in the published NMR structure (Figure 3.14), this is to be expected.

Whilst the average monomer SASA of a construct with a Box A and Box B is equivalent to that of a construct with two Box B domains, there is much less variation of charge state distributions from the average in DbB-HMGB1, potentially suggesting more rigidity in its folding states. Both HMGB1 1-164 and DbB-HMGB1 are more predisposed towards the compact monomer charge state envelope than 1-184, which remains the more predisposed towards the extended monomer conformations amongst all tested constructs. Denaturing ESI-MS (Figure 5.19) also provided results consistent with the observations in native MS. Under denaturing conditions, the average unfolded monomer SASA for DbB was equal to that of HMGB1 1-164, and both constructs adopt similar charge distributions as well.

From these data, it can be deduced that the solvent exposed surface of DbB is comparable to a reduced HMGB1 construct with a Box A and Box B domains alone, such as FR HMGB1 1-164, but with potentially less flexibility in its conformational equilibrium, as the difference between the high M/Z and the low M/Z native monomers is much smaller for the double box B construct.

Surface area w/ SEM (Å^2)		**FR-HMGB1A, c001, FL** *25.6 kDa*	**FR-HMGB1A, c002, 1-184** *21.9 kDa*	**FR-HMGB1A, c031, 1-164** *19.8 kDa*	**HMGB1A-C1, c048, DbB** *20.1 kDa*
Unfolded (ProtSA)		26337	22101	20691	21075
Unfolded (ESI/MS)		26945±41	25487±92	22488±64	22860±217
Potential IDR length		9819	5583	4173	4,717.89
2YRQ (1-164) folded/unfolded		N/A	N/A	14251/19299	N/A
200 mM Amm. Ac	*Average monomer*	17839±17	17502±219	15248±304	15377±203
	Dimer	19466±24	17454±40	14199±116	16765±36

Table 5.2: SASA calculations from ESI/QTOF-MS experiments for DbB- and FR-HMGB1 constructs

Average SASA of the HMGB1 constructs tested, estimated from ESI/QTOF-MS in native and denaturing conditions (average of three injection events) from Figure 5.18 and Figure 5.19.

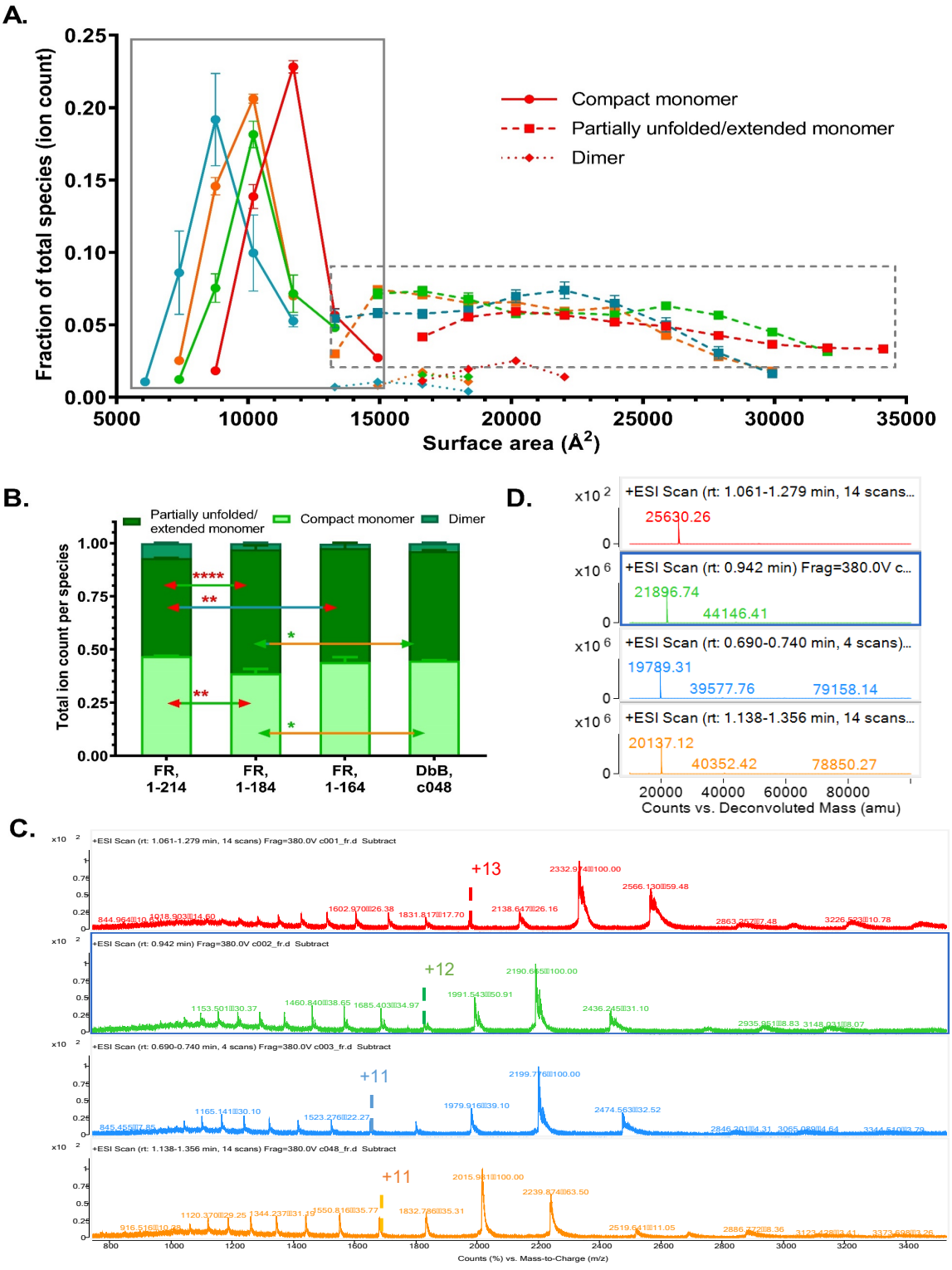


Figure 5.18: Native ESI-QTOF/MS SASA plots for HMGB1 constructs and DbB-HMGB1 in 0.2 M ammonium acetate, pH 6.8 (Figure 5.19)

A. Average solvent accessible surface area vs. charge state abundance of HMGB1A-c001, c002, c031 (FR) and HMGB1A-c048 (DbB) charge state distribution. Data average of three injections, each acquired for at least 10 counts. Continuous lines: monomer, high M/Z monomer species. Dashed lines: low M/Z monomer species. Dotted lines: dimer and other multimers. **B.** Total ion counts for the compact monomer, unfolded monomer, and dimer states (2-way ANOVA; arrowhead and significance value colour indicate parent construct, line colour indicates what construct is it compared to). **C.** Example M/Z spectra of HMGB1A-c001 (red), c002 (green), c031 (blue), and HMGB1A-c048 (orange, DbB) (3 mg/mL, direct infusion into Q-TOF system). States which were considered the threshold between compact monomer envelope (right) and the extended (left); +13 for c001, +12 for c002, +11 for c031 and c048. **D.** Deconvoluted spectra from (C)

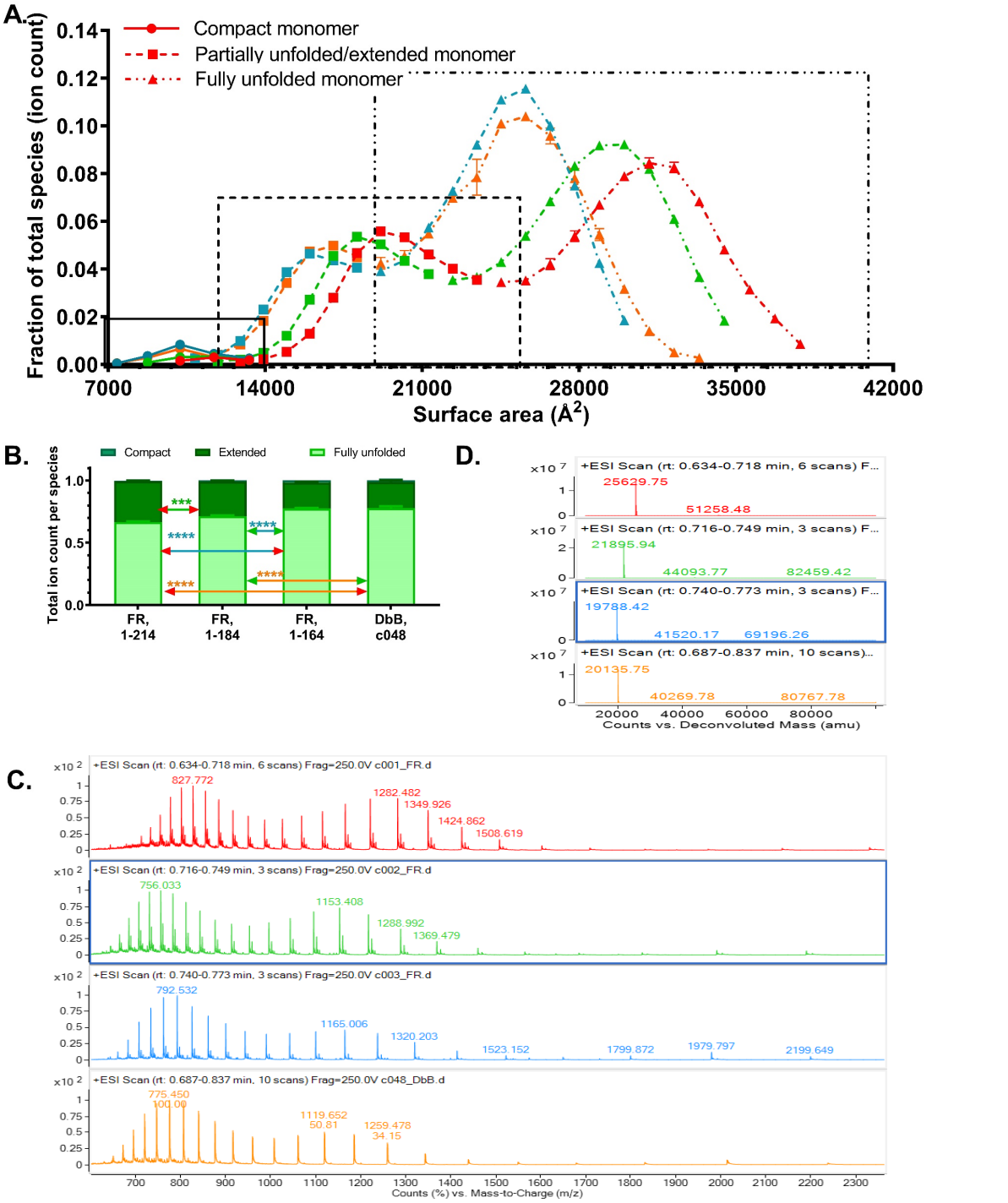


Figure 5.19: Denaturing ESI-QTOF/MS SASA plots of HMGB1 constructs and DbB-HMGB1 in 0.1% formic acid, pH 2.2

A. Average solvent accessible surface area vs. charge state abundance of HMGB1A-c001, c002, c031 (FR) and HMGB1A-c048 (DbB) charge state distribution. Data average of three injections, each acquired for at least 10 counts. Continuous lines: monomer, high M/Z monomer species. Dashed lines: low M/Z monomer species. Dotted lines: dimer and other multimers. **B.** Total ion counts for the compact monomer, unfolded monomer, and dimer states (2-way ANOVA; arrowhead and significance value colour indicate parent construct, line colour indicates what construct is it compared to). **C.** Example M/Z spectra of HMGB1A-c001 (red), c002 (green), c031 (blue), and HMGB1A-c048 (orange, DbB) (3 mg/mL, direct infusion into Q-TOF system). **D.** Deconvoluted spectra from (C)

5.7.III Differential scanning fluorimetry (DSF)

DSF highlighted significant differences between DbB-HMGB1 and FL-HMGB1 (Figure 5.20). Full-length FRHMGB1 is the most stable if NaCl concentration is outside 150-500 mM, whereas ionic strength within this range is optimal for DbB-HMGB1. However, pH behaviour of DbB-HMGB1 differs, with optimal pH ranges being which is to be expected as the acidic tail of HMGB1 is the key modulator of the pH-dependant conformational changes in HMGB1, rather than the Boxes. As pH approaches 10 (pI of DbB: 9.6), the melting temperature starts to drop, as solubility decreases due to approaching isoelectric point, and at pH 10, the curves could not be fitted as no sharp transition could be observed.

At increasing ionic strength concentrations, the melting behaviour of DbB-HMGB1 goes from monophasic to biphasic melting, and a second peak occurs at the same temperature as in deionised water. This would indicate that the protein is salting out of solution. At 1000 mM, this second peak is less prominent, but initial fluorescence is extremely high, which would indicate denatured protein. This low melting point species could also be a dimer. However, since it also appears at high salt concentrations which prevent HMGB1 oligomerisation [51], I consider this unlikely.[57] Despite this behaviour, DbB-HMGB1 can tolerate 1 M NaCl concentrations during purification, at least temporarily, whilst retaining tissue repair enhancement activity *in vitro* and *in vivo*. Some protein might aggregate under these conditions and this first, low melting temperature species, could represent the aggregates dissociating.

DbB-HMGB1 has very different ionic strength and pH tolerances to FR-HMGB1, which is to be expected due to the removal of the acidic tail and the exchange of Box A for Box B. Unlike other tail-less HMGB1 truncations (1-184/164) with reduced Box A and Box B, it does show proper two-phase transition in DSF and likely has a more stable hydrophobic core within Box B, which is more rigid than Box A. Whilst if formulated for clinical use the protein would be reconstituted in saline solution or freeze-dried with an optimal buffer for reconstitution with water, there could be potential for improving the stability of DbB-HMGB1, either by modification of the construct or the purification protocol.

[57] At low salt concentration, it is also possible than Box B binds to the plate. HMGB1 is known to have a high affinity for plasticware [292]

	FR-HMGB1A-c001, 1-214	DbB-HMGB1A-C1-c048
Water	49.6±0.6	32.1±0.3
150 mM NaCl	48.3±0.4	50.4±1
50 mM HEPES pH 7.5, 50 mM NaCl	49.5±0.4	48.8±0.2
50 mM HEPES pH 7.5, 150 mM NaCl	48.1±0.7	51.8±0.3
50 mM HEPES pH 7.5, 250 mM NaCl	47.7±0.5	52.7±0.2
50 mM HEPES pH 7.5, 500 mM NaCl (1)		32.8±0.1
50 mM HEPES pH 7.5, 500 mM NaCl (2)	47.6±0.3	50.8±0.4
50 mM HEPES pH 7.5, 750 mM NaCl (1)		30.2±0.2
50 mM HEPES pH 7.5, 750 mM NaCl (2)	47.8±0.8	51.4±0.2
50 mM HEPES pH 7.5, 1000 mM NaCl	49±0.5	46.2±0.2
50 mM HEPES, pH 7.0	46.8±0.7	39.3±1.2
50 mM HEPES, pH 7.5	48.3±0.6	45.6±0.9
50 mM HEPES, pH 8.0	51.5±1.4	39±4.2
50 mM Tris, pH 7.0	45.2±0.6	41.1±2.1
50 mM Tris, pH 7.5	47.4±0.8	42.4±0.2
50 mM Tris, pH 8.0	50±0.3	41.7±0.2
50 mM Ammonium acetate, pH 6.8	45.1±1	46.3±0.7
100 mM Ammonium acetate, pH 6.8	47.1±0.4	47.1±1.4
200 mM Ammonium acetate, pH 6.8	45.6±0.6	41.1±2.4
0.33X PBS, pH 7.4	47.9±0.3	50.6±0.3
1X PBS, pH 7.4	50.8±0.1	49.1±1.4
2X PBS, pH 7.4	41±5.5	51.1±0.5
0.1 M SPG, pH 4.5	35.6±0.9	42.9±1.3
0.1 M SPG, pH 5.0	35.4±0.1	49.8±0.4
0.1 M SPG, pH 5.5	37±1.3	51.3±0.7
0.1 M SPG, pH 6.0	41.4±3.2	50.8±0.7
0.1 M SPG, pH 6.5	45.6±2.2	50.1±0.3
0.1 M SPG, pH 7.0	50±1	49.9±0.3
0.1 M SPG, pH 7.5	48.5±2.2	50±0.3
0.1 M SPG, pH 8.0	45.4±5.3	51.8±0.9
0.1 M SPG, pH 8.5	49.8±1.6	49.8±0.2
0.1 M SPG, pH 9.0	50±0.2	49.8±0.1
0.1 M SPG, pH 10.0	53.2±0.7	N/A

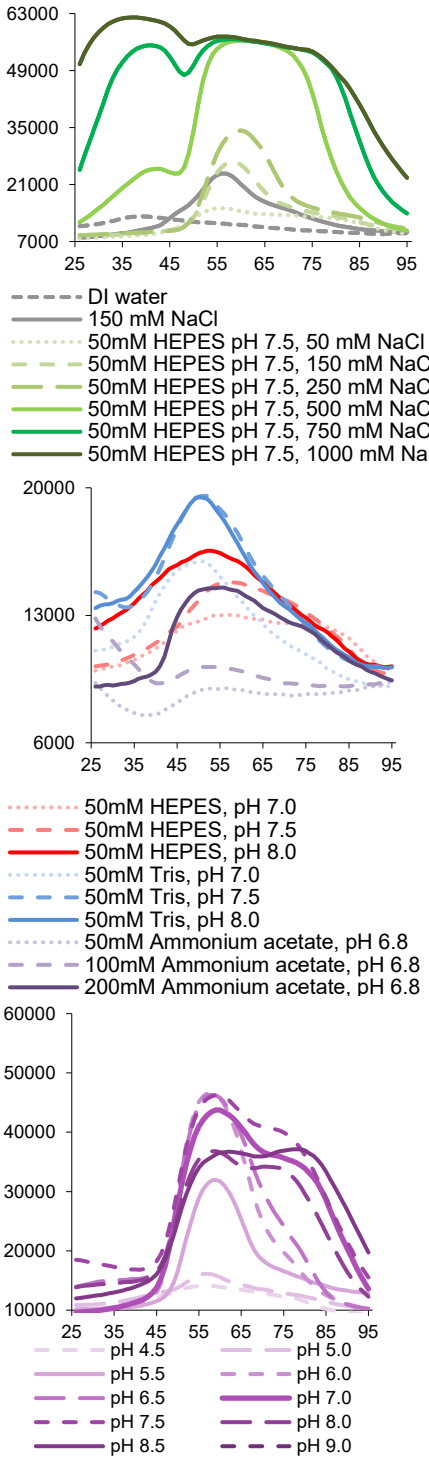


Figure 5.20: DSF melting curve profiles and Tm_{50} (°C) of DbB HMGB1-C1-c048 and comparison with FR-HMGB1 1-214

A: Calculated Tm_{50} values (in ºC) for FR-HMGB1A-c001 (bright red) and DbB-HMGB1-C1-c048 (orange) Colour gradients for the Tm_{50} values indicate highest (green) and lowest (red) values within the global dataset for all constructs; individual condition identifiers have been coloured as per Table 3.4 and each set of screening conditions is separated by a double-space line. N/A: curve too unreliable to fit. N/P: not present. **B:** Thermal denaturation profiles for HMGB1A-C1-c048 in my custom screen; conditions have been coloured in the same code as shown in Figure 3.23. Each row of graphs corresponds to the listed construct; all experiments performed with 8 µM protein (desalted in 10 mM HEPES pH 7.5) and 1X SyPro Orange.

5.7.IV Long-term solution stability at room temperature of FR-HMGB1 FL, 1-164, and DbB

I have not considered whether under prolonged storage these constructs suffer any degradation or aggregation. In a laboratory setting, recombinant proteins are either produced on-demand or flash-frozen at -80°C until they are needed. However, both these approaches are incompatible with clinical practice.

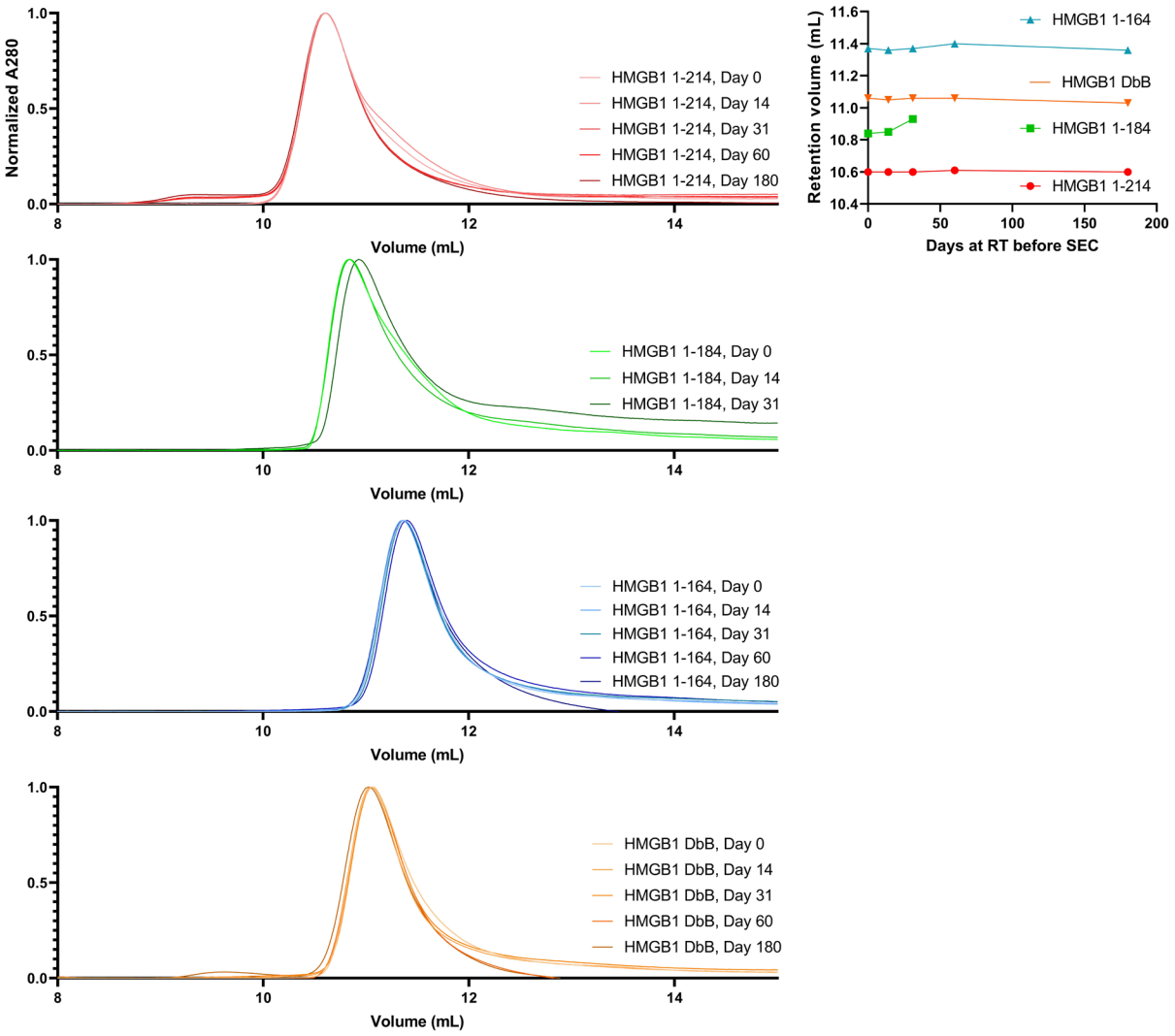
I already knew from my NMR experiments that, as there were no fine structural changes in the HSQC spectra of 1-184 3S HMGB1 and full-length Box B (89-174) HMGB1 after serial NOE/TOCSY experiments, the chemical environment of each residue and thus protein structure remains constant in solution at room temperature for moderate amounts of time (4-5 days). I performed a proof-of-concept experiment to study the stability of DbB-HMGB1 in solution after my standard purification protocol by incubating long-term at room temperature. These conditions are harsher than those likely to be encountered in clinical practice, where the protein could be kept refrigerated or freeze-dried. To cross-validate my results from Chapter 3, I also included HMGB1 1-214, HMGB1 1-184, and HMGB1 1-164 as controls (Figure 5.21). The samples were stored in sterile pH 6.5 ammonium acetate, 0.2 M, as a preparation for non-denaturing ESI-QTOF/MS.[58] In this experiment, degradation would be mediated only by residual proteases or e.g., microbial growth, although I did not observe any indication of the latter.

[58] Whilst the time interval between 60 and 180-day timepoints would appear excessive, this was due to the Covid-19 lockdown preventing me from accessing the lab in this period. I was also unable to perform the planned native MS experiments due to this. The sample for HMGB1 1-184 was used up by day 31.

No significant signs of aggregation could be observed foe any construct. When the samples were loaded in SEC without filtering or centrifugation (to avoid removing aggregates) an increase of a peak at ≈ 9 mL was observed for both HMGB1 1-214 and DbB, which could indicate a modicum of aggregation, after 60 days of storage. Nevertheless, this seems to comprise a small proportion of the total protein sample which elutes at the same retention volume; the main peak for HMGB1 1-214, 1-164, and DbB elutes almost exactly at the same retention volume from day 0 to day 180 (Figure 5.21A). However, HMGB1 1-184 does experience a larger increase in retention volume compared to all other constructs, due to the presence of degradation.

Degradation was minor for all constructs save HMGB1 1-184, with a gradual increase in relative abundance of lower MW species already present in the initial samples but even at 180 days these comprised a minor fraction of the total protein. New lower MW bands appear only after day 31/60. As the degradation products were present to some extent in the initial samples (*in vivo* degradation) they are likely to be a product of trace protease activity from the purification. The only exception to this is HMGB1 1-184, for which several lower MW species appear at day 7 which were not present in the day 0 samples (Figure 5.21B), and which would explain the increase in retention volume; it is noteworthy, however, these were not visible in ESI-QTOF/MS. This would be coincident with my observations in Chapter 3 where this construct was shown to be likely more unstable than both 1-214 and 1-164 constructs.

Full protein oxidation for HMGB1 1-214 only occurred after 14 days with 0.5 mM TCEP, whereas all of the HMGB1 1-184 and 1-164 oxidised after just 7 days; as the acidic tail predisposes the protein towards more compact conformations, this is to be expected due to higher cysteine exposure. As DbB lacks Box A, it cannot undergo intramolecular oxidation. These results suggest that, whilst the protein is quite stable in solution, freeze-drying might be preferable in a clinical setting to prevent potential oxidation of intermolecular disulfide bonds between Box B domains in DbB or extreme long-term (> months) degradation.



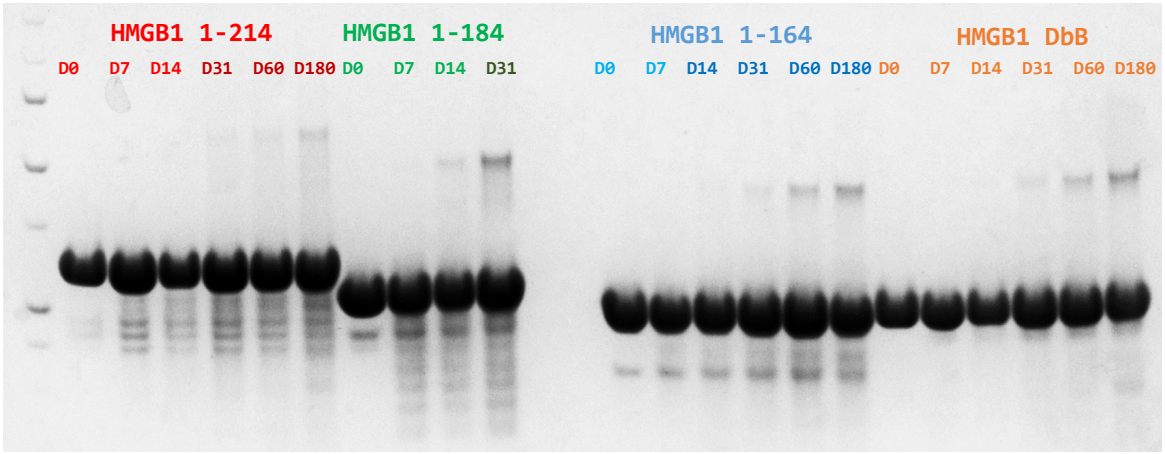


Figure 5.21: Solution stability of FR-HMGB1 FL and DbB

Samples of FR-HMGB1 were subjected to three heat-shock phase separation cycles (addition of 2 % TX-114, equilibrium at 4 °C, equilibrium at 37 °C, and centrifugation) and dialysed into 0.2 M ammonium acetate (pH 6.5)+0.5 mM TCEP. At 0, 7, 14, 31, 60 and 180 days, 100 µL of each protein was frozen at -80 °C without cryoprotectants until required.

A. SEC profiles of the samples (Superdex 75 pg, 10/300, 0.35 mL/min, in 1X PBS). **B.** Reducing SDS-PAGE gel of the samples used in this experiment. Note the higher MW bands which are 2X the MW of the original construct (visible after day 14 for 1-184, and day 31 for all other constructs) and the apparition of new lower MW bands for 1-184 after day 7; for all other constructs this only occurred after day 60. Higher MW bands would correspond to dimers of each construct, which would need to be covalently linked or aggregated for them to appear as a single band in SDS-PAGE. Whilst 2X MW species were observed in ESI, the adduct series (+36n) suggests a deconvolution artifact from 18n (ammonium) series. In SEC, these could correspond to the small increases in void peak observed.

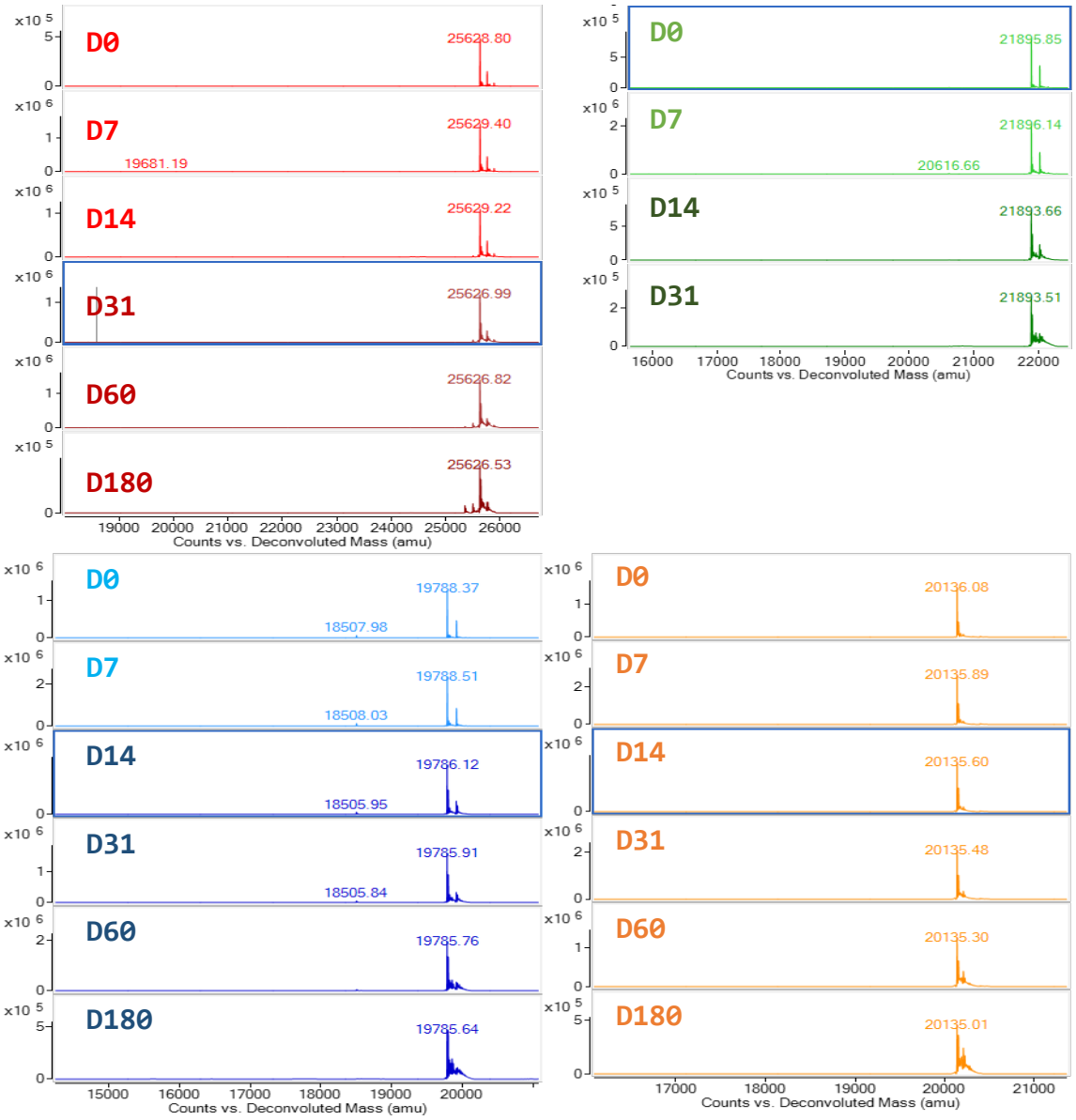


Figure 5.22: Time-of-flight mass spectrometry (ESI/QTOF-MS) of the samples analysed in Figure 5.21

No decrease in molecular weight or appearance of lower MW species (indicative of degradation) were observed until past day 60 where smaller MW species appear in the ESI/MS spectra. The adduct series is +18n (ammonia) rather than +22n (sodium). The amount of ammonium adducts increases with time spent in the solution. FL-FR-HMGB1 remained in the reduced state in the presence of 0.5 mM TCEP for at least 14 days at RT (The -2 Da mass change indicative of oxidation occurs only after Day 14); tail-less forms oxidised earlier due to their more open solution conformation. Complete deconvoluted spectra are shown in Suppl. data A1-5-D; these are zoomed in to show specific peaks for HMGB1 constructs.

5.8 Conclusions

In this chapter, I applied the findings on the molecular basis of the interaction between HMGB1 and CXCL12. Together with information available in the literature on the interaction of HMGB1 with PAMP receptors, I was able to design a panel of constructs which potentially absent TLR-2/4 or RAGE signalling. From this panel, I identified both DbB-HMGB1 and FR HMGB1 1-164 as having reduced RAGE binding, and DbB-HMGB1 being additionally incapable of signalling through TLR-2 or TLR-4 beyond the levels induced by FR-HMGB1. DbB-HMGB1 is however still capable of stimulating hESC differentiation into cardiomyocytes and enhance skeletal muscle tissue repair in a mouse model. Regrettably, I was unable to complete the testing of the panel for proinflammatory activities due to the restrictions imposed by Covid-19, which could have allowed further information on the specific residues involved in TLR-2, TLR-4 and RAGE signalling via HMGB1 and therefore mechanistic insight on why DbB-HMGB1 loses these activities.

I had already analysed the biophysical properties of FR-HMGB1 1-214, 1-184 and-1-164 in Chapter 3. In this chapter, I analysed DbB using the same techniques, confirming that its behaviour and surface conformation is like that of other HMGB1 constructs. As the structure of DbB is not available, this would be the next logical step to assess whether they are entirely similar. CXCL12 binding properties of DbB are likely to be identical to those of the Box B alone construct (89-174), of which DbB contains two copies in tandem.

There is still likely room for further refinement of DbB-HMGB1, despite all the objectives having been met at face value. The proinflammatory effect assays I have performed are *in vitro*. However, *in vivo*, HMGB1 may form complexes with other DAMPs/PAMPs acting as a coreceptor, such as with LPS [29,52], peptidoglycan and LTA [49,59]. RAGE-HMGB1 signalling pathways could also involve cofactors as TLR-2 and 4 do, a hypothesis further reinforced by the convergence of TLR-2, TLR-4, and RAGE on sterile injury [64,83,253] being involved in several deleterious processes in vascular injury such as thrombosis. Heparan sulphate binds both RAGE and HMGB1 [180] and is required for signalling to occur, despite both molecules interacting in its absence in my experiments; heparan sulfation of RAGE seems to however a requirement of RAGE signalling itself, rather than specific to HMGB1, much as MD-2 is for TLR-4 [82].

Chapter 6 Future work and conclusions

My initial plan was to elucidate the 3D structure of the FR-HMGB1-CXCL12 heterocomplex to determine the requirements for complex formation, and use this information to design HMGB1 constructs engineered to eliminate potential proinflammatory signalling via TLR-2, TLR-4, and RAGE.

I started by optimising protocols for large scale production of both HMGB1 and CXCL12 in quantities required for high-volume structural studies and of sufficient purity for *in vivo* work. Whilst in the case of HMGB1 I was able to adapt published protocols, CXCL12 production was particularly challenging as most current protocols require complex refolding and chromatography approaches. The DsbC-SUMO fusion technique I adapted [215] is likely to be translatable to other CXC/CC chemokines that are of small size, include disulfide bonds and have strict modification requirements that make it difficult to express them in eukaryotes due to tight regulation of their signalling. Expression in *E. coli* allowed me to bypass most of these pathways as these do not exist in unicellular prokaryotes. However, determining the structure of the heterocomplex proved to be challenging due to several factors. The conformational flexibility of HMGB1 itself (3.4.I.ii) and the weak, highly dynamic binding equilibrium between HMGB1 and CXCL12 (4.4.III) meant that isolation of stable and homogeneous complexes in solution was not feasible.

I have nevertheless been able to identify the specific residues in HMGB1 responsible for this interaction and to discern that they form a repeated pattern that is conserved between HMGB1 Boxes. I also found that each of the Boxes binds CXCL12 independently of the other with equivalent affinity (4.4.III), but with different kinetic rates. The CXCL12 binding region comprises the first and a half α-helices of an HMG Box, including the flexible N-terminal segment and the C-terminal half of the third α-helix. I found that these flanking flexible regions are critical for stabilising the interaction as their removal resulted in dramatically increased dissociation rates of CXCL12 from the Box domains, as well as a much-reduced signal alteration in NMR in response to CXCL12. An HMG Box with the flanking regions demonstrated increased CSP in several residues over the median value, with additional residues becoming involved in the interaction, compared to an HMG Box without the flanking regions.

Identification of this pattern allowed me to design a construct panel with mutations to disrupt the interaction of HMGB1 with proinflammatory receptors as the regions responsible for latter are partially described. Amongst this panel, a tail-less construct where Box A is replaced with a copy of Box B retained both *in vitro* and *in vivo* regenerative activity (Sec. 5.6) owing to the presence of two full-length HMG Box domains capable of binding CXCL12, whilst eliminating proinflammatory activities:

- Loss of TLR-4 signalling (Sec. 5.5) achieved by removal of Box A, resulting in a construct unable to initiate TLR-4 binding, despite possessing the intact peptide in Box B associated with TLR-4 signalling by binding MD2(89-106) [278].
- Whilst the interaction of TLR-2 with HMGB1 is less well described, it is known that it requires both a co-ligand (such as LTA, or other proteins present in the serum) to compete with the acidic tail [49] and also involves the HMG Box domains near the glycyrrhizin binding site [255]. Signalling via TLR-2 also seems to require the disulfide form of HMGB1 as reduced FL-HMGB1 or my construct lacking Box A and the acidic tail failed to elicit TLR-2-mediated cytokine expression *in vitro*.
- RAGE binding (Sec. 5.4) was abrogated by truncating the RAGE binding peptide (149-182) and the removal of the disulfide bridge in Box A by replacing this Box with another Box B. The DbB-HMGB1 construct failed to effectively bind RAGE despite removal of the acidic tail. The disulfide form of HMGB1 has been shown to be critical for prothrombotic activity via RAGE [64] both *in vitro* and *in vivo*. The biophysical data presented in this thesis show that it is required for binding to the receptor, despite Box A being unable to bind RAGE on its own, suggesting a cooperative mechanism between Box A and Box B.

I assessed *in vitro* signalling for TLR-2 and TLR-4 but not for RAGE. The levels of RAGE binding shown with DbB-HMGB1 were significantly lower than those seen within FR-HMGB1, which cannot signal through the proinflammatory RAGE pathways [63], and in cells expressing TLR-2, TLR-4 and RAGE, no TNF was produced in response to DbB-HMGB1. This would suggest no RAGE agonism by this construct, although full validation *in vitro* is required to confirm this.

Cells that exclusively respond to RAGE have not been described. Monocytes respond to HMGB1 via TLRs, and as all three pathways ultimately converge to the NF-κβ, it would be difficult to isolate the effects due to RAGE signalling alone. Development of a RAGE reporter assay was outside the scope of my DPhil project and it would require dual inhibition of TLR-2 and TLR-4 signalling using e.g. a combination of small molecule inhibitors of TLR-4 such as TAK-242 [279] and neutralising antibodies for TLR-2 [254], as well as synthetic RAGE inhibitors such as RAGE antagonistic peptide (RAP) [179,257]. The alternative approach of using monocytes from dual TLR-2 and TLR-4 knockout animals would be extremely laborious and time consuming. Developing an *in vitro* RAGE reporting assay would be one of the next goals of this project, likely based on the monocyte model described in Section 5.5 in combination with inhibitors of TLR-2 (anti-TLR-2 antibodies), TLR-4 (TAK-242 or TLR-2/4 blockade via TIRAP inhibition [280]) and RAGE (RAP).

The pathways by which the HMGB1-CXCL12 heterocomplex signals downstream of binding to CXCR4 remain to be defined. The biological effect, namely transition of stem cells to G_{Alert}, has been described by the Nanchahal group previously [25]. However, it is unclear how the intracellular pathways triggered by HMGB1-CXCL12 binding to CXCR4 differ from those of CXCL12-induced signalling, although it is known that the ternary complex adopts a different conformation compared to the binary complex [96]. Independent HMGB Box binding to the monomeric form of CXCL12 supports the monomer-related signalling hypothesis, and recent evidence shows that the heterocomplex acts as a balanced CXCR4 agonist with both G protein and arrestin signalling much like monomeric CXCL12 [133], but altered in an unknown manner which results in no internalisation of CXCR4 and increased signalling compared to monomeric CXCL12. Potentially, the heterocomplex also allows this monomer-like signalling to occur even at CXCL12 concentrations where physiologically it would form dimers in the CXCR4 surface and induce G-protein mediated signalling and quiescence alone.

Elucidation of the structure of the full ternary complex with FR-HMGB1 and DbB-HMGB1, compared with the CXCL12-CXCR4 complex alone would provide information on the conformational changes induced by HMGB1 in CXCR4. This structure would also greatly assist in confirming the residues and binding interface involved in the HMGB1-CXCL12 interaction, as well as whether direct contact between HMGB1 and CXCR4 occurs in this complex, as has been previously suggested for 3S-HMGB1 but not FR-HMGB1 [173]. The CXCL12-CXCR4 structure has proved remarkably elusive due to the inherent challenges of GPCR work [115,116,281], particularly in stabilising the active conformation of a GPCR. However, collaborators at the SGC have optimised protocols for expression of stable CXCR4 in large quantities in parallel to my DPhil, solving one of the main issues of this approach.

I observed that certain truncations of the HMGB1 linker in the HMGB1-C1 panel resulted in loss of expression or construct destabilisation. As the thermal melting properties of DbB-HMGB1 are slightly different to those of FR-HMGB1 (Sec. 5.7.III) and it lacks the stabilising effect of the acidic tail, there could also be room for further thermostabilising this construct by modifying linker length or sequence. Furthermore, the current linker in DbB-HMGB1 replaces part of the Lipid A binding region (residues 80-89) [52] with the C-terminal flexible Box B tail (163-174). This results in a partial duplication of the native RAGE binding sequence in Box B (149-174); this is similar in sequence to the RAGE antagonist peptides used in the literature [179]. Whilst the overall binding to RAGE is greatly hampered by the lack of Box A and a high dissociation rate, DbB-HMGB1 has an increased association rate compared to FR-HMGB1, likely due to these two copies of the RAGE binding site. However, I found that several of these residues are involved in CXCL12 binding and, therefore, scope for modification is likely to be limited without further details on specific RAGE binding residues.

This could be achieved via peptide arrays, for both RAGE and TLR-2/4. Whilst recombinant production of both these DAMP receptors is challenging, the available fusions with IgG, such as that used for RAGE binding characterisation, could be used to perform these assays by relying on the detection of IgG rather than His tag, albeit the size limit for the bound ligand is significantly lower than the mass of these proteins [158].[59] These peptide arrays could inform of the presence of unmodified DAMP binding regions within this construct, as well as provide a sequence-based explanation for the loss of RAGE binding as well as TLR-2 and TLR-4 signalling in double box B HMGB1. The TLR mutants I designed in my HMGB1-C1 panel could also be used to validate this if the mutations in these are identified in the peptide array. The methodology adapted to the BLI with RAGE-Fc could be adapted for TLR-2 and TLR-4, although due to the requirements of co-ligands for TLR-2 [49] and MD-2 for TLR-4 [87], this would require extensive optimisation as the mass of the complexes bound to the sensor greatly exceeds that of HMGB1, likely resulting in very low signals.

Determination of the optimal linker length of HMGB1 for preserving pro-regenerative activities would also inform the role of molecular flexibility in DbB-HMGB1 and ternary complex signalling, as the relative position of each HMG Box with respect to each other is relevant for successful CXCL12/CXCR4 mediated signalling via HMGB1 [182]. I am currently analysing thermostability of the linker length modification constructs from the HMGB1-C1 panel. Those which do not demonstrate a stability decrease compared to FL-HMGB1 will then be tested for regeneration *in vivo* in attempt to determine whether, or how much of, the linker is necessary for HMGB1-CXCL12/CXCR4 regenerative signalling in preparation for further modification of DbB-HMGB1. This could further decrease RAGE binding for DbB, as part of the native RAP peptide (149-174) is still present in DbB-HMGB1.

[59] Currently, this was not possible as I lacked the capacity to both access training in the peptide synthetiser equipment or to request the membranes being made by a third party, due to Covid-19 restrictions. The amount of optimisation needed was also incompatible with the laboratory access restrictions.

In summary, the novel findings described in my thesis are as follows:

I. The interaction of HMGB1 and CXCL12 is dynamic and relies on a conserved pattern mirrored across each HMG Box domain, defining a CXCL12-binding pocket in the concave surface of each HMG Box. This pattern involves not only the helical components but also the flexible flanking regions of each HMG Box which stabilise the binding.
II. Binding of CXCL12 to HMGB1 is non-cooperative, with each HMG Box independently recruiting a CXCL12 monomer. Affinity of CXCL12 for the individual HMG Boxes is like that observed for FL-HMGB1.
III. Interaction of HMGB1 with RAGE requires an oxidised Box A domain and a Box B domain containing at least the 149-182 peptide. RAGE binding is competitively inhibited by the presence of the acidic tail, with Box B being responsible for recruiting RAGE and Box A inducing structural changes which ultimately stabilise the interaction between both.
IV. A construct containing two copies of an HMG Box B domain, but no acidic tail or RAGE binding domain, fails to elicit TLR-2 and TLR-4 signalling or bind RAGE but retains full regenerative activity equivalent to full length *in vivo* in a mouse model of skeletal muscle injury. Further optimisation of this construct is feasible, but it will require more information on the function of the HMGB1 linker (if any) and the specific residues or patterns of interaction of HMGB1 with TLR-2 and RAGE.

Chapter 7 Bibliography

1. Weissman IL: **Stem cells: Units of development, units of regeneration, and units in evolution**. *Cell* 2000, **100**:157–168.
2. Brittberg M: **Clinical articular cartilage repair—an up to date review**. *Ann Jt* 2018, **3**:94–94.
3. Uygur A, Lee RT: **Mechanisms of Cardiac Regeneration**. *Dev Cell* 2016, **36**:362–374.
4. Watt FM, Huck WTS: **Role of the extracellular matrix in regulating stem cell fate**. *Nat Rev Mol Cell Biol* 2013, **14**:467–473.
5. Lane SW, Williams DA, Watt FM: **Modulating the stem cell niche for tissue regeneration**. *Nat Biotechnol* 2014, **32**:795–803.
6. Medvedeva E V., Grebenik EA, Gornostaeva SN, Telpuhov VI, Lychagin A V., Timashev PS, Chagin AS: **Repair of damaged articular cartilage: Current approaches and future directions**. *Int J Mol Sci* 2018, **19**.
7. Gratwohl A, Pasquini MC, Aljurf M, Atsuta Y, Baldomero H, Foeken L, Gratwohl M, Bouzas LF, Confer D, Frauendorfer K, et al.: **One million haemopoietic stem-cell transplants: a retrospective observational study**. *Lancet Haematol* 2015, **2**:e91–e100.
8. Trainor N, Pietak A, Smith T: **Rethinking clinical delivery of adult stem cell therapies**. *Nat Biotechnol* 2014, **32**:729–735.
9. Watt FM, Hogan BLM: **Out of eden: Stem cells and their niches**. *Science* 2000, **287**:1427–1430.
10. Domb AJ, Nudelman R: **In vivo and in vitro elimination of aliphatic polyanhydrides**. *Biomaterials* 1995, **16**:319–23.
11. Ghafar-Zadeh E, Waldeisen JR, Lee LP: **Engineered approaches to the stem cell microenvironment for cardiac tissue regeneration.** *Lab Chip* 2011, **11**:3031–48.
12. Chan JK, Glass GE, Ersek A, Freidin A, Williams GA, Gowers K, Espirito Santo AI, Jeffery R, Otto WR, Poulsom R, et al.: **Low-dose TNF augments fracture healing in normal and osteoporotic bone by up-regulating the innate immune response.** *EMBO Mol Med* 2015, **7**:547–61.
13. Ceradini DJ, Kulkarni AR, Callaghan MJ, Tepper OM, Bastidas N, Kleinman ME, Capla JM, Galiano RD, Levine JP, Gurtner GC: **Progenitor cell trafficking is regulated by hypoxic gradients through HIF-1 induction of SDF-1**. *Nat Med* 2004, **10**:858–864.
14. Kitaori T, Ito H, Schwarz EM, Tsutsumi R, Yoshitomi H, Oishi S, Nakano M, Fujii N, Nagasawa T, Nakamura T: **Stromal cell-derived factor 1/CXCR4 signaling is critical for the recruitment of mesenchymal stem cells to the fracture site during skeletal repair in a mouse model**. *Arthritis Rheum* 2009, **60**:813–823.
15. Ding J, Tredget EE: **The Role of Chemokines in Fibrotic Wound Healing**. *Adv Wound Care* 2015, **4**:673–686.
16. Dasu MR, Thangappan RK, Bourgette A, Dipietro LA, Isseroff R, Jialal I: **TLR2 expression and signaling-dependent inflammation impair wound healing in diabetic mice**. *Lab Investig* 2010, **90**:1628–1636.
17. Deonarine K, Panelli MC, Stashower ME, Jin P, Smith K, Slade HB, Norwood C, Wang E, Marincola FM, Stroncek DF: **Gene expression profiling of cutaneous wound healing.** *J Transl Med* 2007, **5**:11.
18. Liu J, Wang H, Li J: **Inflammation and inflammatory cells in myocardial infarction and reperfusion injury: A double-edged sword**. *Clin Med Insights Cardiol* 2016, **10**:79–84.
19. Chan JK, Roth J, Oppenheim JJ, Tracey KJ, Vogl T, Feldmann M, Horwood N, Nanchahal J: **Alarmins: awaiting a clinical response.** *J Clin Invest* 2012, **122**:2711–9.
20. Yang D, Rosa G De, Tewary P, Oppenheim JJ: **Alarmins Link Neutrophils and Dendritic Cells**. *Trends Inmunol* 2009, **30**:531–537.
21. Oppenheim JJ, Yang D: **Alarmins: Chemotactic activators of immune responses**. *Curr Opin Immunol* 2005, **17**:359–365.
22. Beltrame M, Bianchi ME: **Specific recognition of cruciform DNA by a rat liver protein**. *Atti Assoc Genet Ital* 1988, **34**:23–24.
23. Wang H, Bloom O, Zhang M, Vishnubhakat JM, Ombrellino M, Che J, Frazier A, Yang H, Ivanova S, Borovikova L, et al.: **HMG-1 as a Late Mediator of Endotoxin Lethality in Mice**. *Science* 1999, **285**:248–251.

24. Venereau E, Casalgrandi M, Schiraldi M, Antoine DJ, Cattaneo A, De Marchis F, Liu J, Antonelli A, Preti A, Raeli L, et al.: **Mutually exclusive redox forms of HMGB1 promote cell recruitment or proinflammatory cytokine release.** *J Exp Med* 2012, **209**:1519–28.

25. Lee G*, Espirito Santo, A.I.*, Zwingenberger Cai L, Vogl T, Feldmann M, Horwood N, J.K-K C, Nanchahal J: **HMGB1 accelerates the regeneration of multiple tissues by transitioning stem cells to G(Alert)**. *Proc Natl Acad Sci U S A (*authors Contrib Equal to this Proj* 2018, doi:10.1073/pnas.1802893115.

26. Tang D, Kang R, Van Houten B, Zeh HJ, Billiar TR, Lotze MT: **High mobility group box 1 (HMGB1) phenotypic role revealed with stress.** *Mol Med* 2014, **20**:359–62.

27. Calogero S, Grassi F, Aguzzi A, Voigtländer T, Ferrier P, Ferrari S, Bianchi ME: **The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice**. *Nat Genet* 1999, **22**:276–280.

28. Bianchi ME, Agresti A: **HMG proteins: Dynamic players in gene regulation and differentiation**. *Curr Opin Genet Dev* 2005, **15**:496–506.

29. Deng M, Tang Y, Li W, Wang X, Zhang R, Zhang X, Zhao X, Liu J, Tang C, Liu Z, et al.: **The Endotoxin Delivery Protein HMGB1 Mediates Caspase-11-Dependent Lethality in Sepsis**. *Immunity* 2018, **49**:740-753.e7.

30. Vogel S, Bodenstein R, Chen Q, Feil S, Feil R, Rheinlaender J, Schäffer TE, Bohn E, Frick JS, Borst O, et al.: **Platelet-derived HMGB1 is a critical mediator of thrombosis**. *J Clin Invest* 2015, **125**:4638–4654.

31. Palumbo R, De Marchis F, Pusterla T, Conti A, Alessio M, Bianchi ME: **Src family kinases are necessary for cell migration induced by extracellular HMGB1.** *J Leukoc Biol* 2009, **86**:617–23.

32. Palumbo R, Galvez BG, Pusterla T, De Marchis F, Cossu G, Marcu KB, Bianchi ME: **Cells migrating to sites of tissue damage in response to the danger signal HMGB1 require NF-??B activation**. *J Cell Biol* 2007, **179**:33–40.

33. Gong W, Li Y, Chao F, Huang G, He F: **Amino acid residues 201-205 in C-terminal acidic tail region plays a crucial role in antibacterial activity of HMGB1**. *J Biomed Sci* 2009, **16**:83.

34. Ner SS, Churchill EA, Alexandra M, Travers AA: **dHMG-Z, a second HMG-1-related protein in Drosophila melanogaster**. *Nucleic Acids Res* 1993, **21**:4369–4371.

35. Wagner CR, Hamana K, Elgin SC: **A high-mobility-group protein and its cDNAs from Drosophila melanogaster.** *Mol Cell Biol* 1992, **12**:1915–23.

36. Giavara S, Kosmidou E, Prakash Hande M, Bianchi ME, Morgan A, di Fagagna A, Jackson SP: **Yeast Nhp6A/B and Mammalian Hmgb1 Facilitate the Maintenance of Genome Stability indicate that Nhp6A/B and Hmgb1 protect DNA from damaging agents and thus guard against the genera-tion of genomic aberrations**. *Curr Biol* 2005, **15**:68–72.

37. Kang R, Chen R, Zhang Q, Hou W, Wu S, Cao L, Huang J, Yu Y, Fan XG, Yan Z, et al.: **HMGB1 in health and disease**. *Mol Aspects Med* 2014, **40**:1–116.

38. Sharman AC, Hay-Schmidt A, Holland PW.: **Cloning and analysis of an HMG gene from the lamprey Lampetra fluviatilis: gene duplication in vertebrate evolution**. *Gene* 1997, **184**:99–105.

39. Rauvala H, Rouhiainen A: **Physiological and pathophysiological outcomes of the interactions of HMGB1 with cell surface receptors**. *Biochim Biophys Acta - Gene Regul Mech* 2010, **1799**:164–170.

40. Muller S, Ronfani L, Bianchi ME: **Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function**. *J Intern Med* 2004, **255**:332–343.

41. Tang D, Kang R, Zeh HJ, Lotze MT: **High-mobility group box 1 and cancer**. *Biochim Biophys Acta - Gene Regul Mech* 2010, **1799**:131–140.

42. Joshi SR, Sarpong YC, Peterson RC, Scovell WM: **Nucleosome dynamics: HMGB1 relaxes canonical nucleosome structure to facilitate estrogen receptor binding**. *Nucleic Acids Res* 2012, **40**:10161–10171.

43. Itou J, Taniguchi N, Oishi I, Kawakami H, Lotz M, Kawakami Y: **HMGB factors are required for posterior digit development through integrating signaling pathway activities**. *Dev Dyn* 2011, **240**:1151–1162.

44. Bustin M: **Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins.** *Mol Cell Biol* 1999, **19**:5237–46.

45. Allain FH-T, Yen YM, Masse JE, Schultze P, Dieckmann T, Johnson RC, Feigon J: **Solution structure of the HMG protein NHP6A and its interaction with DNA reveals the structural determinants for non-sequence-specific binding**. *EMBO J* 1999, **18**:2563–2579.

46. For TEM: **Structure of the HMG box motif in the B-domain of HMG1 HMGB1**. *Hazel MWeir, Per JKraulis, Caroline SHill1, Andrew RCRaine, Ernest DLaue Jean OThomas2 Cent* 1993, **12**:1311–1319.

47. Dintilhac A, Bernués J: **HMGB1 interacts with many apparently unrelated proteins by recognizing short amino acid sequences**. *J Biol Chem* 2002, **277**:7021–7028.

48. Belgrano FS, De Abreu Da Silva IC, Bastos De Oliveira FM, Fantappié MR, Mohana-Borges R: **Role of the acidic tail of high mobility group protein B1 (HMGB1) in protein stability and DNA bending**. *PLoS One* 2013, **8**:1–12.

49. Aucott H, Sowinska A, Harris HE, Lundback P: **Ligation of free HMGB1 to TLR2 in the absence of ligand is negatively regulated by the C-terminal tail domain**. *Mol Med* 2018, **24**:19.

50. Watson M, Stott K, Thomas JO: **Mapping Intramolecular Interactions between Domains in HMGB1 using a Tail-truncation Approach**. *J Mol Biol* 2007, **374**:1286–1297.

51. Anggayasti WL, Mancera RL, Bottomley S, Helmerhorst E: **The self-association of HMGB1 and its possible role in the binding to DNA and cell membrane receptors**. *FEBS Lett* 2017, **591**:282–294.

52. Youn JH, Kwak MS, Wu J, Kim ES, Ji Y, Min HJ, Yoo JH, Choi JE, Cho HS, Shin JS: **Identification of lipopolysaccharide-binding peptide regions within HMGB1 and their effects on subclinical endotoxemia in a mouse model**. *Eur J Immunol* 2011, **41**:2753–2762.

53. Ling Y, Yang Z, Yin T, Li L, Yuan W, Wu H, Wang C: **Heparin changes the conformation of high-mobility group protein 1 and decreases its affinity toward receptor for advanced glycation endproducts in vitro**. *Int Immunopharmacol* 2011, **11**:187–193.

54. Lu B, Antoine DJ, Kwan K, Lundbäck P, Wähämaa H, Schierbeck H, Robinson M, Van Zoelen MAD, Yang H, Li J, et al.: **JAK/STAT1 signaling promotes HMGB1 hyperacetylation and nuclear translocation.** *Proc Natl Acad Sci U S A* 2014, **111**:3068–73.

55. Bonaldi T, Talamo F, Scaffidi P, Ferrera D, Porto A, Bachi A, Rubartelli A, Agresti A, Bianchi ME: **Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion**. *EMBO J* 2003, **22**:5551–5560.

56. Youn JH, Shin J-S: **Nucleocytoplasmic shuttling of HMGB1 is regulated by phosphorylation that redirects it toward secretion.** *J Immunol* 2006, **177**:7889–7897.

57. Bernice Chao Y, Betty Yan S, Scovell WM: **High mobility group protein, hmg-1, contains insignificant glycosyl modification**. *Protein Sci* 1994, **3**:2452–2454.

58. Mollica L, De Marchis F, Spitaleri A, Dallacosta C, Pennacchini D, Zamai M, Agresti A, Trisciuoglio L, Musco G, Bianchi ME: **Glycyrrhizin Binds to High-Mobility Group Box 1 Protein and Inhibits Its Cytokine Activities**. *Chem Biol* 2007, **14**:431–441.

59. Kwak MS, Lim M, Lee YJ, Lee HS, Kim YH, Ho Youn J, Choi JE, Shin J-S: **HMGB1 Binds to Lipoteichoic Acid and Enhances TNF-α and IL-6 Production through HMGB1-Mediated Transfer of Lipoteichoic Acid to CD14 and TLR2**. *J Innate Immun* 2015, **7**:405–416.

60. Hoppe G, Talcott KE, Bhattacharya SK, Crabb JW, Sears JE: **Molecular basis for the redox control of nuclear transport of the structural chromatin protein Hmgb1**. *Exp Cell Res* 2006, **312**:3526–3538.

61. Kim YH, Kwak MS, Park JB, Lee SA, Choi JE, Cho HS, Shin JS: **N-linked glycosylation plays a crucial role in the secretion of HMGB1**. *J Cell Sci* 2016, **129**:29–38.

62. Cecchinato V, D'Agostino G, Raeli L, Nerviani A, Schiraldi M, Danelon G, Manzo A, Thelen M, Ciurea A, Bianchi ME, et al.: **Redox-Mediated Mechanisms Fuel Monocyte Responses to CXCL12/HMGB1 in Active Rheumatoid Arthritis**. *Front Immunol* 2018, **9**:1–12.

63. Lu B, Nakamura T, Inouye K, Li J, Tang Y, Lundback P, Valdes-Ferrer SI, Olofsson PS, Kalb T, Roth J, et al.: **Novel role of PKR in inflammasome activation and HMGB1 release**. *Nature* 2012, **488**:670–674.

64. Stark K, Philippi V, Stockhausen S, Busse J, Antonelli A, Miller M, Schubert I, Hoseinpour P, Chandraratne S, Von Bruhl ML, et al.: **Disulfide HMGB1 derived from platelets coordinates venous thrombosis in mice**. *Blood* 2016, **128**:2435–2449.

65. Gardella S, Andrei C, Ferrera D, Lotti L V., Torrisi MR, Bianchi ME, Rubartelli A: **The nuclear protein HMGB1 is secreted by monocytes via a non-classical, vesicle-mediated secretory pathway**. *EMBO Rep* 2002, **3**:995–1001.

66. Lu B, Wang C, Wang M, Li W, Chen F, Tracey KJ, Wang H: **Molecular Mechanism and Therapeutic Modulation of HMGB1 Release and Action: An Updated Review**. *Expert Rev Clin Immunol* 2014, **10**:713–727.

67. Lotze MT, Tracey KJ: **High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal.** *Nat Rev Immunol* 2005, **5**:331–42.

68. Scaffidi P, Misteli T, Bianchi ME: **Release of chromatin protein HMGB1 by necrotic cells triggers inflammation.** *Nature* 2002, **418**:191–195.

69. Eberharter A, Becker PB: **Histone acetylation: A switch between repressive and permissive chromatin. Second in review on chromatin dynamics**. *EMBO Rep* 2002, **3**:224–229.

70. Kew RR, Penzo M, Habiel DM, Marcu KB: **The IKKα-dependent NF-κB p52/RelB noncanonical pathway is essential to sustain a CXCL12 autocrine loop in cells migrating in response to HMGB1.** *J Immunol* 2012, **188**:2380–2386.

71. Ottestad W, Rognes IN, Pischke SE, Mollnes TE, Andersson U, Eken T: **Biphasic Release of the Alarmin High Mobility Group Box 1 Protein Early After Trauma Predicts Poor Clinical Outcome**. *Crit Care Med* 2019, **47**:e614–e622.

72. Yang H, Wang H, Wang Y, Addorisio M, Li J, Postiglione MJ, Chavan SS, Al-Abed Y, Antoine DJ, Andersson U, et al.: **The haptoglobin beta subunit sequesters HMGB1 toxicity in sterile and infectious inflammation**. *J Intern Med* 2017, **282**:76–93.

73. Yang H, Wang H, Levine YA, Gunasekaran MK, Wang Y, Addorisio M, Zhu S, Li W, Li J, de Kleijn DP V, et al.: **Identification of CD163 as an antiinflammatory receptor for HMGB1-haptoglobin complexes.** *JCI insight* 2016, **1**:1–14.

74. Zandarashvili L, Sahu D, Lee K, Lee YS, Singh P, Rajarathnam K, Iwahara J: **Real-time kinetics of high-mobility group box 1 (HMGB1) oxidation in extracellular fluids studied by in situ protein NMR spectroscopy**. *J Biol Chem* 2013, **288**:11621–11627.

75. Andersson U, Wang H, Palmblad K, Aveberger A, Bloom O, Erlandsson-harris H, Janson A, Kokkola R, Zhang M, Yang H, et al.: **High Mobility Group 1 Protein (HMG-1) Stimulates Proinflammatory Cytokine Synthesis in Human Monocytes**. *J Exp Med* 2000, **192**:565–570.

76. Yang H, Wang H, Czura CJ, Tracey KJ: **The cytokine activity of HMGB1**. *J Leukoc Biol* 2005, **78**:1–8.

77. Antoine DJ, Williams DP, Jenkins AK, Regan SL, Sathish JG, Kitteringham NR, Park BK: **High-mobility group box-1 protein and keratin-18, circulating serum proteins informative of acetaminophen-induced necrosis and apoptosis in vivo**. *Toxicol Sci* 2009, **112**:521–531.

78. Evankovich J, Cho SW, Zhang R, Cardinal J, Dhupar R, Zhang L, Klune JR, Zlotnicki J, Billiar T, Tsung A: **High mobility group box 1 release from hepatocytes during ischemia and reperfusion injury is mediated by decreased histone deacetylase activity**. *J Biol Chem* 2010, **285**:39888–39897.

79. Sachdev U, Cui X, Hong G, Namkoong S, Karlsson JM, Baty CJ, Tzeng E: **High mobility group box 1 promotes endothelial cell angiogenic behavior in vitro and improves muscle perfusion in vivo in response to ischemic injury**. *J Vasc Surg* 2012, **55**:180–191.

80. Maugeri N, Campana L, Gavina M, Covino C, De Metrio M, Panciroli C, Maiuri L, Maseri A, D'Angelo A, Bianchi ME, et al.: **Activated platelets present high mobility group box 1 to neutrophils, inducing autophagy and promoting the extrusion of neutrophil extracellular traps**. *J Thromb Haemost* 2014, **12**:2074–2088.

81. Park BS, Song DH, Kim HM, Choi B-S, Lee H, Lee J-O: **The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex**. *Nature* 2009, **458**:1191–1195.

82. Xu D, Young JH, Krahn JM, Song D, Corbett KD, Chazin WJ, Pedersen LC, Esko JD: **Stable RAGE-heparan sulfate complexes are essential for signal transduction**. *ACS Chem Biol* 2013, **8**:1611–1620.

83. van Beijnum JR, Buurman WA, Griffioen AW: **Convergence and amplification of toll-like receptor (TLR) and receptor for advanced glycation end products (RAGE) signaling pathways via high mobility group B1 (HMGB1)**. *Angiogenesis* 2008, **11**:91–99.

84. Sakaguchi M, Murata H, Yamamoto K, Ono T, Sakaguchi Y, Motoyama A, Hibino T, Kataoka K, Huh N: **TIRAP, an adaptor protein for TLR2/4, transduces a signal from RAGE phosphorylated upon ligand binding.** *PLoS One* 2011, **6**:e23132.

85. Tang D, Kang R, Iii HJZ, Lotze MT: **High-Mobility Group Box 1, Oxidative Stress, and Disease**. *Antioxid Redox Signal* 2011, **14**.

86. Park JS, Gamboni-robertson F, He Q, Svetkauskaite D, Kim J, Strassheim D, Sohn J, Yamada S, Maruyama I, Banerjee A, et al.: **High mobility group box 1 protein interacts with multiple Toll-like receptors**. *Am J Physiol* 2006, **3**:917–924.

87. Yang H, Wang H, Ju Z, Ragab AA, Lundbäck P, Long W, Valdes-Ferrer SI, He M, Pribis JP, Li J, et al.: **MD-2 is required for disulfide HMGB1–dependent TLR4 signaling**. *J Exp Med* 2015, **212**:5–14.

88. He M, Bianchi ME, Coleman TR, Tracey KJ, Al-abed Y: **Exploring the biological functional mechanism of the HMGB1 / TLR4 / MD-2 complex by surface plasmon resonance**. *Mol Med* 2018, **24**:1–9.

89. Hreggvidsdóttir HS, Lundberg AM, Aveberger A-C, Klevenvall L, Andersson U, Harris HE: **High mobility group box protein 1 (HMGB1)-partner molecule complexes enhance cytokine production by signaling through the partner molecule receptor.** *Mol Med* 2012, **18**:224–30.

90. Das N, Dewan V, Grace PM, Gunn RJ, Tamura R, Tzarum N, Watkins LR, Wilson IA, Yin H: **HMGB1 Activates Proinflammatory Signaling via TLR5 Leading to Allodynia**. *Cell Rep* 2016, **17**:1128–1140.

91. Penzo M, Molteni R, Suda T, Samaniego S, Raucci A, Habiel DM, Miller F, Jiang H, Li J, Pardi R, et al.: **Inhibitor of NF-κB Kinases α and β Are Both Essential for High Mobility Group Box 1-Mediated Chemotaxis**. *J Immunol* 2010, **184**:4497–4509.

92. Campana L, Bosurgi L, Bianchi ME, Manfredi AA, Patrizia R-Q: **Requirement of {HMGB1} for stromal cell-derived {factor-1/CXCL12-dependent} migration of macrophages and dendritic cells.** *J Leukoc Biol* 2009, **86**:609–615.

93. Brinkmann V, Reichard U, Goosmann C, Fauler B, Uhlemann Y, Weiss DS, Weinrauch Y, Zychlinsky A: **Neutrophil extracellular traps kill bacteria.** *Science* 2004, **303**:1532–5.

94. Brill A, Fuchs TA, Savchenko AS, Thomas GM, Martinod K, De Meyer SF, Bhandari AA, Wagner DD: **Neutrophil extracellular traps promote deep vein thrombosis in mice**. *J Thromb Haemost* 2012, **10**:136–144.

95. Metz AK, Diaz JA, Obi AT, Wakefield TW, Myers DD, Henke PK: **Venous Thrombosis and Post-Thrombotic Syndrome: From Novel Biomarkers to Biology.** *Methodist Debakey Cardiovasc J* 2018, **14**:173–181.

96. Schiraldi M, Raucci A, Muñoz LM, Livoti E, Celona B, Venereau E, Apuzzo T, De Marchis F, Pedotti M, Bachi A, et al.: **{HMGB1} promotes recruitment of inflammatory cells to damaged tissues by forming a complex with {CXCL12} and signaling via {CXCR4.}**. *J Exp Med* 2012, **209**:551–563.

97. Venereau E, Schiraldi M, Uguccioni M, Bianchi ME: **HMGB1 and leukocyte migration during trauma and sterile inflammation**. *Mol Immunol* 2013, **55**:76–82.

98. Rodgers JT, King KY, Brett JO, Cromie MJ, Charville GW, Maguire KK, Brunson C, Mastey N, Liu L, Tsai C-R, et al.: **mTORC1 controls the adaptive transition of quiescent stem cells from G0 to G(Alert).** *Nature* 2014, **509**:393–6.

99. Joseph J, Dyson M: **The effect of abdominal wounding on the rate of tissue regeneration**. *Experientia* 1970, **26**:66–67.

100. Rodgers JT, Schroeder MD, Ma C, Rando TA: **HGFA is an injury-regulated systemic factor that induces the transition of stem cells into GAlert**. *Cell Rep* 2017, **19**:479–486.

101. Yellowley C: **CXCL12/CXCR4 signaling and other recruitment and homing pathways in fracture repair.** *Bonekey Rep* 2013, **2**:300.

102. Balabanian K, Lagane B, Infantino S, Chow KYC, Harriague J, Moepps B, Arenzana-Seisdedos F, Thelen M, Bachelerie F: **The chemokine SDF-1/CXCL12 binds to and signals through the orphan receptor RDC1 in T lymphocytes**. *J Biol Chem* 2005, **280**:35760–35766.

103. Yang OO, Swanberg SL, Lu Z, Dziejman M, McCoy J, Luster AD, Walker BD, Herrmann SH: **Enhanced inhibition of human immunodeficiency virus type 1 by Met-stromal-derived factor 1beta correlates with down-modulation of CXCR4**. *J Virol* 1999, **73**:4582–4589.

104. Doms RW, Moore JP: **HIV-1 membrane fusion: Targets of opportunity**. *J Cell Biol* 2000, **151**:9–13.

105. Bleul CC, Farzan M, Choe H, Parolin C, Clark-Lewis I, Sodroski J, Springer T a: **The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry.** *Nature* 1996, **382**:829–833.

106. Wang Z, Wang C, Mo B, Wei J, Ma L, Wang J: **The SDF-1 / CXCR4 / AKT signaling regulates bone marrow mesenchymal stem cell trafficking to promote airway remodeling in asthmatic rats**. *Int J Clin Exp Med* 2017, **10**:507–515.

107. Ratajczak MZ, Zuba-Surma E, Kucia M, Reca R, Wojakowski W, Ratajczak J: **The pleiotropic effects of the SDF-1-CXCR4 axis in organogenesis, regeneration and tumorigenesis.** *Leukemia* 2006, **20**:1915–1924.

108. Sierro F, Biben C, Martı L, Woehl B, Leung H, Groom J, Batten M, Harvey RP, Martı C, Mackay CR, et al.: **Disrupted cardiac development but normal hematopoiesis in mice deficient in the second CXCL12/SDF-1 receptor, CXCR7.** *Proc Natl Acad Sci U S A* 2007, **104**:14759–64.

109. Nagasawa T, Hirota S, Tachibana K, Takakura N, Nishikawa S, Kitamura Y, Yoshida N, Kikutani H, Kishimoto T: **Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1.** *Nature* 1996, **382**:635–638.

110. Tachibana K, Hirota S, Iizasa H, Yoshida H, Kawabata K, Kataoka Y, Kitamura Y, Matsushima K, Yoshida N, Nishikawa S, et al.: **The chemokine receptor CXCR4 is essential for vascularization of the gastrointestinal tract**. *Nature* 1998, **393**:591–594.

111. Crump MP, Gong JH, Loetscher P, Rajarathnam K, Amara A, Arenzana-Seisdedos F, Virelizier JL, Baggiolini M, Sykes BD, Clark-Lewis I: **Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1.** *EMBO J* 1997, **16**:6996–7007.

112. Veldkamp CT, Ziarek JJ, Su J, Basnet H, Lennertz R, Weiner JJ, Peterson FC, Baker JE, Volkman BF: **Monomeric structure of the cardioprotective chemokine SDF-1/CXCL12**. *Protein Sci* 2009, **18**:1359–1369.

113. Veldkamp CT, Seibert C, Peterson FC, De NB, Haugner JC, Basnet H, Sakmar TP, Volkman BF: **Structural Basis of CXCR4 Sulfotyrosine Recognition by the Chemokine SDF-1/CXCL12**. *Sci Signal* 2009, **1**:1–18.

114. Veldkamp CT, Peterson FC, Pelzek AJ, Volkman BF: **The monomer – dimer equilibrium of stromal cell-derived factor-1 (CXCL12) is altered by pH, phosphate, sulfate, and heparin**. *Protein Sci* 2005, **1**:1071–1081.

115. Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, et al.: **Structures of the CXCR4 Chemokine GPCR with Small-molecule and Cyclic Peptide Agonists - SOM**. *Science* 2010, **330**.

116. Qin L, Kufareva I, Holden LG, Wang C, Zheng Y, Zhao C, Fenalti G, Wu H, Han GW, Cherezov V, et al.: **Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine.** *Science* 2015, **347**:1117–22.

117. Ziarek JJ, Kleist AB, London N, Raveh B, Montpas N, Bonneterre J, St-onge G, Dicosmo-ponticello CJ, Koplinski CA, Roy I, et al.: **Structural basis for chemokine recognition by a G protein – coupled receptor and implications for receptor activation**. *Sci Signal* 2017, **5756**.

118. Skieterska K, Rondou P, Craenenbroeck K Van: **Regulation of G Protein-Coupled Receptors by Ubiquitination**. 2017, doi:10.20944/preprints201704.0135.v1.

119. Bhandari D, Robia SL, Marchese A: **The E3 Ubiquitin Ligase Atrophin Interacting Protein 4 Binds Directly To The Chemokine Receptor CXCR4 Via a Novel WW Domain-mediated Interaction**. *Mol Biol Cell* 2009, **20**:1324–1339.

120. Janssens R, Struyf S, Proost P: **The unique structural and functional features of CXCL12**. *Cell Mol Immunol* 2017, doi:10.1038/cmi.2017.107.

121. Oberlin E, Amara A, Bachelerie F, Bessia C, Virelizier JL, Arenzana-Seisdedos F, Schwartz O, Heard JM, Clark-Lewis I, Legler DF, et al.: **The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1.** *Nature* 1996, **382**:833–835.

122. Vega B, Muñoz LM, Holgado BL, Lucas P, Rodríguez-Frade JM, Calle A, Rodríguez-Fernández JL, Lechuga LM, Rodríguez JF, Gutiérrez-Gallego R, et al.: **Technical advance: Surface plasmon resonance-based analysis of CXCL12 binding using immobilized lentiviral particles.** *J Leukoc Biol* 2011, **90**:399–408.

123. Gustavsson M, Dyer DP, Zhao C, Handel TM: **Kinetics of CXCL12 binding to atypical chemokine receptor 3 reveal a role for the receptor N terminus in chemokine binding**. *Sci Signal* 2019, **12**.

124. Mirzabekov T, Bannert N, Farzan M, Hofmann W, Kolchinsky P, Wu L, Wyatt R, Sodroski J: **Enhanced expression, native purification, and characterization of CCR5, a principal HIV-1 coreceptor**. *J Biol Chem* 1999, **274**:28745–28750.

125. Carpenter B, Tate CG: **Engineering a minimal G protein to facilitate crystallisation of G protein-coupled receptors in their active conformation**. *Protein Eng Des Sel* 2016, doi:10.1093/protein/gzw049.

126. Kofuku Y, Yoshiura C, Ueda T, Terasawa H, Hirai T, Tominaga S, Hirose M, Maeda Y, Takahashi H, Terashima Y, et al.: **Structural basis of the interaction between chemokine stromal cell-derived factor-1/CXCL12 and its G-protein-coupled receptor CXCR4**. *J Biol Chem* 2009, **284**:35240–35250.

127. Janssens R, Mortier A, Boff D, Ruytinx P, Gouwy M, Vantilt B, Larsen O, Daugvilaite V, Rosenkilde MM, Parmentier M, et al.: **Truncation of CXCL12 by CD26 reduces its CXC chemokine receptor 4- and atypical chemokine receptor 3-dependent activity on endothelial cells and lymphocytes**. *Biochem Pharmacol* 2017, doi:10.1016/j.bcp.2017.03.009.

128. Richter R, Jochheim-Richter A, Ciuculescu F, Kollar K, Seifried E, Forssmann U, Verzijl D, Smit MJ, Blanchet X, von Hundelshausen P, et al.: **Identification and Characterization of Circulating Variants of CXCL12 from Human Plasma: Effects on Chemotaxis and Mobilization of Hematopoietic Stem and Progenitor Cells**. *Stem Cells Dev* 2014, **23**:1959–1974.

129. Chang S, Li Y, Yuan F, Qu M, Song Y, Zhang Z, Yang GY, Wang Y: **Monomeric CXCL12 outperforms its dimeric and wild type variants in the promotion of human endothelial progenitor cells' function**. *Biochem Biophys Res Commun* 2017, **488**:303–310.

130. Drury LJ, Ziarek JJ, Gravel S, Veldkamp CT, Takekoshi T, Hwang ST, Heveker N, Volkman BF, Dwinell MB: **Monomeric and dimeric CXCL12 inhibit metastasis through distinct CXCR4 interactions and signaling pathways**. *Proc Natl Acad Sci U S A* 2011, **108**:17655–17660.

131. Livoti E: **Experimentally validated computational docking to characterize protein-protein interactions**. [date unknown],

132. Mollica L, De Marchis F, Spitaleri A, Dallacosta C, Pennacchini D, Zamai M, Agresti A, Trisciuoglio L, Musco G, Bianchi ME: **Glycyrrhizin binds to high-mobility group box 1 protein and inhibits its cytokine activities.** *Chem Biol* 2007, **14**:431–41.

133. D'Agostino G, Artinger M, Locati M, Perez L, Legler DF, Bianchi ME, Rüegg C, Thelen M, Marchese A, Rocchi MBL, et al.: **β-Arrestin1 and β-Arrestin2 Are Required to Support the Activity of the CXCL12/HMGB1 Heterocomplex on CXCR4**. *Front Immunol* 2020, **11**:1–16.

134. Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, et al.: **Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists**. *Science* 2010, **330**:1066–1071.

135. Huttunen HJ, Fages C, Kuja-Panula J, Ridley AJ, Rauvala H: **Receptor for advanced glycation end products-binding COOH-terminal motif of amphoterin inhibits invasive migration and metastasis**. *Cancer Res* 2002, **62**:4805–4811.

136. Perrin CL: **Linear or Nonlinear Least-Squares Analysis of Kinetic Data?** *J Chem Educ* 2017, **94**:669–672.

137. Lee S, Lee DK: **Multiple Comparison Test and Its Imitations What is the proper way to apply the multiple comparison test? KJA**. *Korean J Anesth* 2018, doi:10.4097/kja.d.18.00242.

138. Williams DR, Rapley R: **Agarose Gel Electrophoresis of Nucleic Acids BT - The Nucleic Acid Protocols Handbook**. In *Nucleic Acid Protocols Handbook, The*. . 2000:67–70.

139. Garca-Descalzo L, Garca-Lpez E, Alczar A, Baquero F, Ci C: **Gel Electrophoresis of Proteins**. In *Protein Protocols Handbook, The*. . 2012:221–238.

140. Dyballa N, Metzger S: **Fast and sensitive colloidal Coomassie G-250 staining for proteins in polyacrylamide gels**. *J Vis Exp* 2009, doi:10.3791/1431.

141. Chalk R: **Mass Spectrometric Analysis of Proteins**. Humana Press, New York, NY; 2017:373–395.

142. Chalk R, Berridge G, Shrestha L, Strain-Damerell C, Mahajan P, Yue W, Gileadi O, Burgess-Brown N: **High-Throughput Mass Spectrometry Applied to Structural Genomics**. *Chromatography* 2014, **1**:159–175.

143. Testa L, Brocca S, Grandori R: **Charge-surface correlation in electrospray ionization of folded and unfolded proteins**. *Anal Chem* 2011, **83**:6459–6463.

144. Kaltashow IA, Mohimen A: **Electrospray ionization mass spectrometry can provide estimates of protein surface areas in solution**. *Anal Chem* 2005, **77**:5370–5379.

145. Estrada J, Bernadó P, Blackledge M, Sancho J: **ProtSA: A web application for calculating sequence specific protein solvent accessibilities in the unfolded ensemble**. *BMC Bioinformatics* 2009, **10**:1–8.

146. Savitsky P, Bray J, Cooper CDO, Marsden BD, Mahajan P, Burgess-Brown NA, Gileadi O: **High-throughput production of human proteins for crystallization: The SGC experience**. *J Struct Biol* 2010, **172**:3–13.

147. Temple G, Gerhard DS, Rasooly R, Feingold EA, Good PJ, Robinson C, Mandich A, Derge JG, Lewis J, Shoaf D, et al.: **The completion of the Mammalian Gene Collection**. *Genome Res* 2009, **19**:2324–2333.

148. Schwarz H, Schmittner M, Duschl A, Horejs-Hoeck J: **Residual endotoxin contaminations in recombinant proteins are sufficient to activate human CD1c+ dendritic cells.** *PLoS One* 2014, **9**:e113840.

149. Mamat U, Wilke K, Bramhill D, Schromm AB, Lindner B, Kohl TA, Corchero JL, Villaverde A, Schaffer L, Head SR, et al.: **Detoxifying Escherichia coli for endotoxin-free production of recombinant proteins**. *Microb Cell Fact* 2015, **14**:1–15.

150. Aida Y, Pabst MJ: **Removal of endotoxin from protein solutions by phase separation using triton X-114**. *J Immunol Methods* 1990, **132**:191–195.

151. Falagas ME, Kasiakou SK: **Toxicity of polymyxins: a systematic review of the evidence from old and recent studies.** *Crit Care* 2006, **10**:R27.

152. Li J, Wang H, Mason JM, Levine J, Yu M, Ulloa L, Czura CJ, Tracey KJ, Yang H: **Recombinant HMGB1 with cytokine-stimulating activity**. *J Immunol Methods* 2004, **289**:211–223.

153. Teodorowicz M, Perdijk O, Verhoek I, Govers C, Savelkoul HFJ, Tang Y, Wichers H, Broersen K: **Optimized Triton X-114 assisted lipopolysaccharide (LPS) removal method reveals the immunomodulatory effect of food proteins**. *PLoS One* 2017, **12**:e0173778.

154. Ackland GL, Gutierrez Del Arroyo A, Yao ST, Stephens RC, Dyson A, Klein NJ, Singer M, Gourine A V.: **Low-molecular-weight polyethylene glycol improves survival in experimental sepsis**. *Crit Care Med* 2010, **38**:629–636.

155. Abate W, Sattar AA, Liu J, Conway ME, Jackson SK: **Evaluation of recombinant factor C assay for the detection of divergent lipopolysaccharide structural species and comparison with Limulus amebocyte lysate-based assays and a human monocyte activity assay**. *J Med Microbiol* 2017, **66**:888–897.

156. Niesen FH, Berglund H, Vedadi M: **The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability**. *Nat Protoc* 2007, **2**:2212–2221.

157. Vivoli M, Novak HR, Littlechild JA, Harmer NJ: **Determination of protein-ligand interactions using differential scanning fluorimetry**. *J Vis Exp* 2014, doi:10.3791/51809.

158. Nady N, Min J, Kareta MS, Chédin F, Arrowsmith CH: **A SPOT on the chromatin landscape? Histone peptide arrays as a tool for epigenetic research**. *Trends Biochem Sci* 2008, **33**:305–313.

159. Hulme EC, Trevethick MA: **Ligand binding assays at equilibrium: Validation and interpretation**. *Br J Pharmacol* 2010, **161**:1219–1237.

160. Knapp S, Müller S, Digilio G, Bonaldi T, Bianchi ME, Musco G: **The long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes**. *Biochemistry* 2004, **43**:11992–11997.

161. Gaspar P, Oliveira JL, Frommlet J, Santos MAS, Moura G: **EuGene: Maximizing synthetic gene design for heterologous expression**. *Bioinformatics* 2012, **28**:2683–2684.

162. Liao JK, Seto M, Noma K: **ROCK inhibitors**. *J Cardiovasc Pharmacol* 2007, **50**:17–24.

163. Claassen DA, Desler MM, Rizzino A: **ROCK inhibition enhances the recovery and growth of cryopreserved human embryonic stem cells and human induced pluripotent stem cells**. *Mol Reprod Dev* 2009, **76**:722–732.

164. Kodaka M, Yang Z, Nakagawa K, Maruyama J, Xu X, Sarkar A, Ichimura A, Nasu Y, Ozawa T, Iwasa H, et al.: **A new cell-based assay to evaluate myogenesis in mouse myoblast C2C12 cells**. *Exp Cell Res* 2015, **336**:171–181.

165. Wang Z, Liang Q, Zhang Y, Yang J, Li M: **An optimized HMGB1 expressed by recombinant rabies virus enhances immunogenicity through activation of dendritic cells in mice**. 2017,

166. Merenmies J, Pihlaskari R, Laitinien J, Wartiovaarall J, Rauvala H: **30-kDa Heparin-binding Protein of Brain (Amphoterin) Involved in Neurite Outgrowth**. *J Biol Chem* 1991, **266**:16722–16729.

167. Hori O, Brett J, Slattery T, Cao R, Zhang J, Chen JX, Stern D, Schmidt AM: **The Receptor for Advanced Glycation End Products (RAGE) Is a Cellular Binding Site for Amphoterin**. *J Biol Chem* 1995, **270**:25752–25761.

168. Humphreys DP, Sehdev M, Chapman AP, Ganesh R, Smith BJ, King LM, Glover DJ, Reeks DG, Stephens PE: **High-level periplasmic expression in Escherichia coli using a eukaryotic signal peptide: importance of codon usage at the 5' end of the coding sequence**. *Protein Expr Purif* 2000, **20**:252–264.

169. Gustafsson C, Minshull J, Govindarajan S, Ness J, Villalobos A, Welch M, Drive OB, Suite A, Park M: **Engineering Genes for Predictable Protein Expression**. *Protein Expre* 2012, **83**:37–46.

170. Burgess-Brown NA, Sharma S, Sobott F, Loenarz C, Oppermann U, Gileadi O: **Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study**. *Protein Expr Purif* 2008, **59**:94–102.

171. Rapp C, Snow S, Laufer T, McClendon CL: **The role of tyrosine sulfation in the dimerization of the CXCR4:SDF-1 complex**. *Protein Sci* 2013, **22**:1025–1036.

172. Sivashanmugam A, Murray V, Cui C, Zhang Y, Wang J, Li Q: **Practical protocols for production of very high yields of recombinant proteins using Escherichia coli**. *Protein Sci* 2009, **18**:936–948.

173. Di Maggio S, Milano G, De Marchis F, D'Ambrosio A, Bertolotti M, Palacios BS, Badi I, Sommariva E, Pompilio G, Capogrossi MC, et al.: **Non-oxidizable HMGB1 induces cardiac fibroblasts migration via CXCR4 in a CXCL12-independent manner and worsens tissue remodeling after myocardial infarction**. *Biochim Biophys Acta - Mol Basis Dis* 2017, **1863**:2693–2704.

174. Yang H, Lundback P: **Redox Modification of Cysteine Residues Regulates the Cytokine Activity of High Mobility Group Box-1 (HMGB1) Huan**. *Mol Med* 2012, **18**:1.

175. Gorman JJ, Wallis TP, Pitt JJ: **Protein disulfide bond determination by mass spectrometry**. *Mass Spectrom Rev* 2002, **21**:183–216.

176. Hernandez C, Huebener P, Pradere J-P, Antoine DJ, Friedman RA, Schwabe RF: **HMGB1 links chronic liver injury to progenitor responses and hepatocarcinogenesis.** *J Clin Invest* 2018, **128**:2436–2451.

177. Lundbäck P, Stridh P, Klevenvall L, Jenkins RE, Fischer M, Sundberg E, Andersson U, Antoine DJ, Harris HE: **Characterization of the Inflammatory Properties of Actively Released HMGB1 in Juvenile Idiopathic Arthritis.** *Antioxid Redox Signal* 2014, **00**:1–15.

178. Scientific T: *Iodoacetamide, Single-Use Instructions 90034*. [date unknown].

179. Arumugam T, Ramachandran V, Gomez SB, Schmidt AM, Logsdon CD: **S100P-derived RAGE antagonistic peptide reduces tumor growth and metastasis**. *Clin Cancer Res* 2012, **18**:4356–4364.

180. Xu D, Young J, Song D, Esko JD: **Heparan sulfate is essential for high mobility group protein 1 (HMGB1) signaling by the receptor for advanced glycation end products (RAGE)**. *J Biol Chem* 2011, **286**:41736–41744.

181. Palumbo R, Sampaolesi M, De Marchis F, Tonlorenzi R, Colombetti S, Mondino A, Cossu G, Bianchi ME: **Extracellular HMGB1, a signal of tissue damage, induces mesoangioblast migration and proliferation**. *J Cell Biol* 2004, **164**:441–449.

182. Fassi EMA, Sgrignani J, D'Agostino G, Cecchinato V, Garofalo M, Grazioso G, Uguccioni M, Cavalli A: **Oxidation State Dependent Conformational Changes of HMGB1 Regulate the Formation of the CXCL12/HMGB1 Heterocomplex**. *Comput Struct Biotechnol J* 2019, **17**:886–894.

183. De Leo F, Quilici G, Tirone M, Mannella V, De Marchis F, Preti A, Gori A, Casalgrandi M, Mezzapelle R, Bianchi M, et al.: **Diflunisal targets the HMGB1/CXCL12 heterocomplex and blocks immune cell recruitment - SM**. *bioRxiv* 2019, doi:10.1101/563890.

184. Held PG: **Nucleic Acid Purity Assessment using A 260 / A 280 Ratios**. *BioTek-Application Note* 2001,

185. Glasel JA: **Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios.** *Biotechniques* 1995, **18**:62–3.

186. Armbrecht M: **Detection of contamination in DNA and protein samples by photometric measurements**. *Appl Note 279* 2013,

187. Faust G, Stand A, Weuster-Botz D: **IPTG can replace lactose in auto-induction media to enhance protein expression in batch-cultured *Escherichia coli***. *Eng Life Sci* 2015, **15**:824–829.

188. Wanner BL, Kodaira R, Neidhardt FC: **Regulation of lac operon expression: Reappraisal of the theory of catabolite repression**. *J Bacteriol* 1978, **136**:947–954.

189. Hirel PH, Schmitter MJ, Dessen P, Fayat G, Blanquet S: **Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid.** *Proc Natl Acad Sci U S A* 1989, **86**:8247–51.

190. Kapust RB, Copeland TD, Waugh DS: *The P1 0 specificity of tobacco etch virus protease*. [date unknown].

191. Harvey Lodish, Arnold Berk, S Lawrence Zipursky, Paul Matsudaira, David Baltimore JD: **Noncovalent Bonds**. In *Molecular Cell Biology. 4th edition.* . WH Freeman; 2000.

192. Benco L, Tunega D, Hafner J, Lischka H: **Upper limit of the O-H···O hydrogen bond. Ab initio study of the kaolinite structure**. *J Phys Chem B* 2001, **105**:10812–10817.

193. Ito T, Kawahara K, Okamoto K, Yamada S, Yasuda M, Imaizumi H, Nawa Y, Meng X, Shrestha B, Hashiguchi T, et al.: **Proteolytic cleavage of high mobility group box 1 protein by thrombin-thrombomodulin complexes.** *Arterioscler Thromb Vasc Biol* 2008, **28**:1825–30.

194. GE Healthcare Life Sciences: **Superdex High-performance Columns - Data File 18-1163-79 AD**. 2007,

195. Hagemans D, van Belzen IAEM, Luengo TM, Rüdiger SGD: **A script to highlight hydrophobicity and charge on protein surfaces**. *Front Mol Biosci* 2015, **2**:1–11.

196. Anggayasti WL, Ogino K, Yamamoto E, Helmerhorst E, Yasuoka K, Mancera RL: **The acidic tail of HMGB1 regulates its secondary structure and conformational flexibility: A circular dichroism and molecular dynamics simulation study**. *Comput Struct Biotechnol J* 2020, **18**:1160–1172.

197. Schnier PD, Gross DS, Williams ER: **On the Maximum Charge State and Proton Transfer Reactivity of Peptide and Protein Ions Formed By Electrospray Ionization**. 1995, **0305**.

198. Anggayasti WL, Mancera RL, Bottomley S, Helmerhorst E: **The effect of physicochemical factors on the self-association of HMGB1: A surface plasmon resonance study**. *Biochim Biophys Acta - Proteins Proteomics* 2016, **1864**:1620–1629.

199. Smith RD, Light-Wahl KJ, Winger BE, Loo JA: **Preservation of non-covalent associations in electrospray ionization mass spectrometry: Multiply charged polypeptide and protein dimers**. *Org Mass Spectrom* 1992, **27**:811–821.

200. Smith RD, Light-Wahl KJ: **The observation of non-covalent interactions in solution by electrospray ionization mass spectrometry: Promise, pitfalls and prognosis**. *Biol Mass Spectrom* 1993, **22**:493–501.

201. Kebarle P, Verkerk UH: **ELECTROSPRAY: FROM IONS IN SOLUTION TO IONS IN THE GAS PHASE, WHAT WE KNOW NOW**. *Mass Spec Rev* 2009, doi:10.1002/mas.

202. Erickson HP: **Size and shape of protein molecules at the nanometer level determined by sedimentation, gel filtration, and electron microscopy**. *Biol Proced Online* 2009, **11**:32–51.

203. Fassi EMA, Sgrignani J, D'Agostino G, Cecchinato V, Garfalo M, Grazioso G, Uguccioni M, Cavalli A: **Oxidation state dependent conformational changes of HMGB1 regulates the formation of the CXCL12/HMGB1 heterocomplex**. *bioRxiv Biochem* 2019, doi:10.1101/555946.

204. Payne SH, Bonissone S, Wu S, Brown RN, Ivankov DN, Frishman D, Paša-Tolić L, Smith RD, Pevzner PA: **Unexpected diversity of signal peptides in prokaryotes**. *MBio* 2012, **3**:1–6.

205. Wang W, Choi BK, Li W, Lao Z, Lee AYH, Souza SC, Yates NA, Kowalski T, Pocai A, Cohen LH: **Quantification of intact and truncated stromal cell-derived factor-1?? in circulation by immunoaffinity enrichment and tandem mass spectrometry**. *J Am Soc Mass Spectrom* 2014, **25**:614–625.

206. Wescott MP, Kufareva I, Paes C, Goodman JR, Thaker Y, Puffer BA, Berdougo E, Rucker JB, Handel TM, Doranz BJ: **Signal transmission through the CXC chemokine receptor 4 (CXCR4) transmembrane helices**. *Proc Natl Acad Sci* 2016, **113**:9928–9933.

207. Ziarek JJ, Getschman AE, Butler SJ, Taleski D, Stephens B, Kufareva I, Handel TM, Payne RJ, Volkman BF: **Sulfopeptide probes of the CXCR4/CXCL12 interface reveal oligomer-specific contacts and chemokine allostery**. *ACS Chem Biol* 2013, **8**:1955–1963.

208. Veldkamp CT, Peterson FC, Hayes PL, Mattmiller JE, Iii CH, Cruz N De, Volkman BF: **On-column refolding of recombinant chemokines for NMR studies and biological assays**. *Protein Expr Purif* 2007, **52**:202–209.

209. Yamamoto Y, Ritz D, Planson A, Jönsson TJ, Melinda J, Boyd D, Beckwith J, Poole LB: **Mutant AhpC peroxiredoxins suppress thiol-disulfide redox deficiencies and acquire deglutathionylating activity**. *Mol Cell* 2009, **29**:36–45.

210. Hatahet F, Nguyen VD, Salo KEH, Ruddock LW: **Disruption of reducing pathways is not essential for efficient disulfide bond formation in the cytoplasm of E. coli.** *Microb Cell Fact* 2010, **9**:67.

211. Matos CFRO, Robinson C, Alanen HI, Prus P, Uchida Y, Ruddock LW, Freedman RB, Keshavarz-Moore E: **Efficient export of prefolded, disulfide-bonded recombinant proteins to the periplasm by the Tat pathway in Escherichia coli CyDisCo strains**. *Biotechnol Prog* 2014, **30**:281–290.

212. Nguyen VD, Hatahet F, Salo KEH, Enlund E, Zhang C, Ruddock LW: **Pre-expression of a sulfhydryl oxidase significantly increases the yields of eukaryotic disulfide bond containing proteins expressed in the cytoplasm of E.coli**. *Microb Cell Fact* 2011, **10**:1–13.

213. Natale P, Brüser T, Driessen AJM: **Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-Distinct translocases and mechanisms**. *Biochim Biophys Acta - Biomembr* 2008, **1778**:1735–1756.

214. Waugh DS: **An overview of enzymatic reagents for the removal of affinity tags**. *Protein Expr Purif* 2011, **80**:283–293.

215. Turchetto J, Sequeira AF, Ramond L, Peysson F, Brás JLA, Saez NJ, Duhoo Y, Blémont M, Guerreiro CIPD, Quinton L, et al.: **High-throughput expression of animal venom toxins in Escherichia coli to generate a large library of oxidized disulphide-reticulated peptides for drug discovery**. *Microb Cell Fact* 2017, **16**:6.

216. Nozach H, Fruchart-Gaillard C, Fenaille F, Beau F, Ramos OHP, Douzi B, Saez NJ, Moutiez M, Servent D, Gondry M, et al.: **High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in E. coli.** *Microb Cell Fact* 2013, **12**:37.

217. Sequeira AF, Turchetto J, Saez NJ, Peysson F, Ramond L, Duhoo Y, Blémont M, Fernandes VO, Gama LT, Ferreira LMA, et al.: **Gene design, fusion technology and TEV cleavage conditions influence the purification of oxidized disulphide-rich venom peptides in Escherichia coli**. *Microb Cell Fact* 2017, **16**:1–16.

218. Crane JM, Randall LL: **The Sec System: Protein Export in Escherichia coli**. *EcoSal Plus* 2017, **7**:1–73.

219. Thein M, Sauer G, Paramasivam N, Grin I, Linke D: **Efficient subfractionation of gram-negative bacteria for proteomics studies**. *J Proteome Res* 2010, **9**:6135–6147.

220. Eisenberg D, Schwarz E, Komaromy M, Wall R: **Analysis of membrane and surface protein sequences with the hydrophobic moment plot**. *J Mol Biol* 1984, **179**:125–142.

221. Takekoshi T, Ziarek JJ, Volkman BF, Hwang ST: **A Locked, Dimeric CXCL12 Variant Effectively Inhibits Pulmonary Metastasis of CXCR4-Expressing Melanoma Cells Due to Enhanced Serum Stability**. *Mol Cancer Ther* 2012, **11**:2516–2525.

222. Gustavsson M, Wang L, Gils N Van, Stephens BS, Zhang P, Schall TJ, Yang S, Abagyan R, Chance MR, Kufareva I, et al.: **Structural basis of ligand interaction with atypical chemokine receptor 3**. 2017, doi:10.1038/ncomms14135.

223. Zhang S, Tao F, Qing R, Tang H, Skuhersky M, Corin K, Tegler L, Wassie A, Wassie B, Kwon Y, et al.: **QTY code enables design of detergent-free chemokine receptors that retain ligand-binding activities**. *Proc Natl Acad Sci* 2018, doi:10.1073/pnas.1811031115.

224. Zhang S, Tao F, Qing R, Tang H, Skuhersky M, Corin K, Tegler L, Wassie A, Wassie B, Kwon Y, et al.: **QTY code enables design of detergent-free chemokine receptors that retain ligand-binding activities - SM**. *Proc Natl Acad Sci* 2018, doi:10.1073/pnas.1811031115.

225. Dumitriu IE, Bianchi ME, Bacci M, Manfredi A a, Rovere-Querini P: **The secretion of HMGB1 is required for the migration of maturing dendritic cells.** *J Leukoc Biol* 2007, **81**:84–91.

226. Yun J, Jiang G, Wang Y, Xiao T, Zhao Y, Sun D, Kaplan HJ, Shao H: **The HMGB1-CXCL12 complex promotes inflammatory cell infiltration in uveitogenic T cell-induced chronic experimental autoimmune uveitis**. *Front Immunol* 2017, **8**:1–11.

227. Magna M, Pisetsky DS: **The role of HMGB1 in the pathogenesis of inflammatory and autoimmune diseases**. *Mol Med* 2014, **20**:138–146.

228. Tirone M, Tran NL, Ceriotti C, Gorzanelli A, Canepari M, Bottinelli R, Raucci A, Maggio S Di, Santiago C, Mellado M, et al.: **High mobility group box 1 orchestrates tissue regeneration via CXCR4**. *J Exp Med* 2018, doi:10.1084/jem.20160217.

229. Gurevich DB, Nguyen PD, Siegel AL, Ehrlich O V., Sonntag C, Phan JMN, Berger S, Ratnayake D, Hersey L, Berger J, et al.: **Asymmetric division of clonal muscle stem cells coordinates muscle regeneration in vivo**. *Science* 2016, **353**.

230. Haller S, Kapuria S, Riley RR, O'Leary MN, Schreiber KH, Andersen JK, Melov S, Que J, Rando TA, Rock J, et al.: **mTORC1 Activation during Repeated Regeneration Impairs Somatic Stem Cell Maintenance**. *Cell Stem Cell* 2017, **21**:806-818.e5.

231. Di Trani JM, De Cesco S, O'Leary R, Plescia J, Do Nascimento CJ, Moitessier N, Mittermaier AK: **Rapid measurement of inhibitor binding kinetics by isothermal titration calorimetry /631/45/607 /631/1647/2204 /631/92/613 /631/154 /9 article**. *Nat Commun* 2018, **9**.

232. De Leo F, Quilici G, Tirone M, Mannella V, De Marchis F, Preti A, Gori A, Casalgrandi M, Mezzapelle R, Bianchi M, et al.: **Diflunisal targets the HMGB1/CXCL12 heterocomplex and blocks immune cell recruitment**. *bioRxiv* 2019, doi:10.1101/563890.

233. Davenport AP, Russell FD: **Radioligand Binding Assays: Theory and Practice**. In *Current Directions in Radiopharmaceutical Research and Development*. Edited by Mather SJ. Springer Netherlands; 1996:169–179.

234. Navratilova I, Besnard J, Hopkins AL: **Screening for GPCR Ligands Using Surface Plasmon Resonance**. *ACS Med Chem Lett* 2011,

235. Cheng Y, Grigorieff N, Penczek PA, Walz T: **A primer to single-particle cryo-electron microscopy**. *Cell* 2015, **161**:439–449.

236. Henderson R: **The Potential and Limitations of Neutrons, Electrons and X-Rays for Atomic Resolution Microscopy of Unstained Biological Molecules**. *Q Rev Biophys* 1995, **28**:171–193.

237. Wu M, Herzik MA, Lander GC: **Pushing Single-Particle Cryo-EM To The Theoretical Size and Resolution Limits At 200 keV**. [date unknown], doi:10.1101/457861.

238. Herzik MA, Wu M, Lander GC: **High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM**. *Nat Commun* 2019, **10**:1–9.

239. Stark H: *GraFix: Stabilization of fragile macromolecular complexes for single particle Cryo-EM*. Elsevier Masson SAS; 2010.

240. Bianchi ME, Falciola L, Ferrari S, Lilley DM: **The DNA binding site of HMG1 protein is composed of two similar segments (HMG boxes), both of which have counterparts in other eukaryotic regulatory proteins**. *Embo J* 1992, **11**:1055–1063.

241. Wang J, Tochio N, Takeuchi A, Uewaki J ichi, Kobayashi N, Tate S ichi: **Redox-sensitive structural change in the A-domain of HMGB1 and its implication for the binding to cisplatin modified DNA**. *Biochem Biophys Res Commun* 2013, **441**:701–706.

242. Redfield C: **Assignment of Protein NMR Spectra using Heteronuclear NMR - A Tutorial**. In *Protein NMR: Modern Techniques and Biomedical Applications*. . 2015:1–185.

243. Krishnan V, Rupp B: **Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy**. *eLS* 2012, doi:10.1002/9780470015902.a0002716.pub2.

244. Puthenveetil R, Vinogradova O: **Solution NMR: A powerful tool for structural and functional studies of membrane proteins in reconstituted environments**. *J Biol Chem* 2019, **294**:15914–15931.

245. Balagurunathan K, Nakato H, Desai UR: **Using Isothermal Titration Calorimetry to Determine Thermodynamic Parameters of Protein–Glycosaminoglycan Interactions**. *Methods Mol Biol* 2015, **1229**:v.

246. Steigenberger B, van den Toorn HWP, Bijl E, Greisch JF, Räther O, Lubeck M, Pieters RJ, Heck AJR, Scheltema RA: **Benefits of Collisional Cross Section Assisted Precursor Selection (caps-PASEF) for Cross-linking Mass Spectrometry**. *Mol Cell Proteomics* 2020, **19**:1677–1687.

247. Ahrens I, Agrotis A, Topcic D, Bassler N, Chen YC, Bobik A, Bode C, Peter K: **HMGB1 binds to activated platelets via platelet-expressed receptor for advanced glycation end products (RAGE) and is highly expressed in platelet rich coronary artery thrombi**. *Thromb Haemost* 2015, **14**:994–1003.

248. Vaure C, Liu Y: **A comparative review of toll-like receptor 4 expression and functionality in different animal species.** *Front Immunol* 2014, **5**:316.

249. Lee K-M, Seong S-Y: **Partial role of TLR4 as a receptor responding to damage-associated molecular pattern**. *Immunol Lett* 2009, **125**:31–39.

250. Pisetsky DS, Erlandsson-Harris H, Andersson U: **High-mobility group box protein 1 (HMGB1): an alarmin mediating the pathogenesis of rheumatic disease.** *Arthritis Res Ther* 2008, **10**:209.

251. Jovanović Stojanov S, Martinović V, Bogojević D, Poznanović G, Petrović A, Ivanović-Matić S, Grigorov I: **Modulation of diabetes-related liver injury by the HMGB1/TLR4 inflammatory pathway**. *J Physiol Biochem* 2018, doi:10.1007/s13105-018-0626-0.

252. Li J, Kokkola R, Tabibzadeh S, Yang R, Ochani M, Qiang X, Harris HE, Czura CJ, Wang H, Ulloa L, et al.: **Structural basis for the proinflammatory cytokine activity of high mobility group box 1.** *Mol Med* 2003, **9**:37–45.

253. Ibrahim ZA, Armour CL, Phipps S, Sukkar MB: **RAGE and TLRs: Relatives, friends or neighbours?** *Mol Immunol* 2013, **56**:739–744.

254. Yu M, Wang H, Ding A, Golenbock DT, Latz E, Czura CJ, Fenton MJ, Tracey KJ, Yang H: **HMGB1 signals through toll-like receptor (TLR) 4 and TLR2**. *Shock* 2006, **26**:174–179.

255. Teng W, He X, Shan L, Fan C, Peng S, Wang X, Shan Z, Wang W, Liu X, Liu S, et al.: **Glycyrrhizin, a Direct HMGB1 Antagonist, Ameliorates Inflammatory Infiltration in a Model of Autoimmune Thyroiditis via Inhibition of TLR2-HMGB1 Signaling**. *Thyroid* 2017, **27**:722–731.

256. Chen X, Cheng F, Liu Y, Zhang L, Song L, Cai X, You T, Fan X, Wang D, Gong A, et al.: **Toll-like receptor 2 and Toll-like receptor 4 exhibit distinct regulation of cancer cell stemness mediated by cell death-induced high-mobility group box 1**. *EBioMedicine* 2019, doi:10.1016/J.EBIOM.2018.12.016.

257. Lee S, Piao C, Kim G, Kim JY, Choi E, Lee M: **Production and application of HMGB1 derived recombinant RAGE-antagonist peptide for anti-inflammatory therapy in acute lung injury**. *Eur J Pharm Sci* 2018, **114**:275–284.

258. Yang H, Liu H, Zeng Q, Imperato GH, Addorisio ME, Li J, He M, Cheng KF, Al-abed Y, Harris HE, et al.: **Inhibition of HMGB1 / RAGE-mediated endocytosis by HMGB1 antagonist box A , anti-HMGB1 antibodies , and cholinergic agonists suppresses inflammation**. *Mol Med* 2019,

259. Zhang CL, Shu MG, Qi HW, Li LW: **Inhibition of tumor angiogenesis by HMGB1 A box peptide**. *Med Hypotheses* 2008, **70**:343–345.

260. LeBlanc PM, Doggett TA, Choi J, Hancock MA, Durocher Y, Frank F, Nagar B, Ferguson TA, Saleh M: **An immunogenic peptide in the A-box of HMGB1 protein reverses apoptosis-induced tolerance through RAGE receptor**. *J Biol Chem* 2014, **289**:7777–7786.

261. Mitola S, Belleri M, Urbinati C, Coltrini D, Sparatore B, Pedrazzi M, Melloni E, Presta M: **Cutting edge: extracellular high mobility group box-1 protein is a proangiogenic cytokine**. *J Immunol* 2006, **176**:12–15.

262. Tang D, Billiar T: **A Janus Tale of Two Active High Mobility Group Box 1 (HMGB1) Redox States**. *Mol Med* 2012, **18**:1.

263. Rui Liu, SM, C, Wake H, Zhang J, Liu, KeyueIZushi Y, Takahashi HK, Peng B, Nishibori M: **Establishment of in Vitro Binding Assay of High Mobility Group Box-1 and S100A12 to Receptor for Advanced Glycation Endproducts: Heparin's Effect on Binding**. *Acta Med Okayama* 2009,

264. He M, Bianchi ME, Coleman TR, Tracey KJ, Al-abed Y, Unit CD, Biology C: **Exploring the biological functional mechanism of the HMGB1 / TLR4 / MD-2 complex by surface plasmon resonance**. [date unknown],

265. Bianchi ME, CASALGRANDI M, VENEREAU EJ, Brunelli S: **Hmgb1 variants and uses thereof**. 2014,

266. Gustafsson C, Minshull J, Govindarajan S, Ness J, Villalobos A, Welch M: **Engineering genes for predictable protein expression**. *Protein Expr Purif* 2012, **83**:37–46.

267. Faure G, Ogurtsov AY, Shabalina SA, Koonin E V.: **Role of mRNA structure in the control of protein folding**. *Nucleic Acids Res* 2016, doi:10.1093/nar/gkw671.

268. Gould N, Hendy O, Papamichail D: **Computational Tools and Algorithms for Designing Customized Synthetic Genes**. *Front Bioeng Biotechnol* 2014, **2**.

269. Welch M, Villalobos a., Gustafsson C, Minshull J: **You're one in a googol: optimizing genes for protein expression**. *J R Soc Interface* 2009, **6**:S467–S476.

270. Parret AH, Besir H, Meijers R: **Critical reflections on synthetic gene design for recombinant protein expression**. *Curr Opin Struct Biol* 2016, **38**:155–162.

271. Stark K, Philippi V, Stockhausen S, Busse J, Antonelli A, Miller M, Schubert I, Hoseinpour P, Chandraratne S, Von Bruhl ML, et al.: **Disulfide HMGB1 derived from platelets coordinates venous thrombosis in mice - SM**. *Blood* 2016, **128**:2435–2449.

272. Robert SM, Sjodin H, Fink MP, Aneja RK: **Preconditioning with high mobility group Box 1 (HMGB1) induces lipoteichoic acid (LTA) tolerance**. *J Immunother* 2010, doi:10.1097/CJI.0b013e3181dcd111.

273. Akira S: **Toll-like Receptors and Innate Immunity**. In Edited by Dixon FJBT-A in I. Academic Press; 2001:1–56.

274. *Animals (Scientific Procedures) Act, rev. April 2020*. Statute Law Database; 1986.

275. Miyake T, McDermott JC, Gramolini AO: **A method for the direct identification of differentiating muscle cells by a fluorescent mitochondrial dye**. *PLoS One* 2011, **6**.

276. Grune T, Ott C, Häseli S, Höhn A, Jung T: **The "MYOCYTER" – Convert cellular and cardiac contractions into numbers with ImageJ**. *Sci Rep* 2019, **9**:1–13.

277. Golforoush P, Schneider MD: **Intensive care for human hearts in pluripotent stem cell models**. *npj Regen Med* 2020, **5**:1–9.

278. Yang H, Hreggvidsdottir HS, Palmblad K, Wang H, Ochani M, Li J, Lu B, Chavan S, Rosas-Ballina M, Al-Abed Y, et al.: **A critical cysteine is required for HMGB1 binding to Toll-like receptor 4 and activation of macrophage cytokine release.** *Proc Natl Acad Sci U S A* 2010, **107**:11942–7.

279. Lian YJ, Gong H, Wu TY, Su WJ, Zhang Y, Yang YY, Peng W, Zhang T, Zhou JR, Jiang CL, et al.: **Ds-HMGB1 and fr-HMGB induce depressive behavior through neuroinflammation in contrast to nonoxid-HMGB1**. *Brain Behav Immun* 2017, **59**:322–332.

280. Couture LA, Piao W, Ru LW, Vogel SN, Toshchakov VY: **Targeting toll-like receptor (TLR) signaling by toll/interleukin-1 receptor (TIR) domain-containing adapter protein/MyD88 adapter-like (TIRAP/Mal)-derived decoy peptides**. *J Biol Chem* 2012, **287**:24641–24648.

281. Xu L, Li Y, Sun H, Li D, Hou T: **Structural basis of the interactions between CXCR4 and CXCL12/SDF-1 revealed by theoretical approaches.** *Mol Biosyst* 2013, **9**:2107–2117.

282. Blanca MJ, Alarcón R, Arnau J, Bono R, Bendayan R: **Datos no normales: ¿es el ANOVA una opción válida?** *Psicothema* 2017, **29**:552–557.

283. Skoog B: **Removal of polyethylene glycols from immunoglobulin samples by adsorption chromatography on polystyrene beads.** *Experientia* 1980, **36**:1157–8.

284. **[EN] ToxinSensor Chromogenic LAL Endotoxin Assay Kit 1 * 1 Each**. 2017,

285. Strickland S, Palmer G, Massey V: **Determination of dissociation constants and specific rate constants of enzyme substrate (or protein ligand) interactions from rapid reaction kinetic data**. *J Biol Chem* 1975, **250**:4048–4052.

286. Hara T, Tanegashima K: **CXCL14 antagonizes the CXCL12-CXCR4 signaling axis**. *Biomol Concepts* 2014, **5**:167–173.

287. Pozzobon T, Goldoni G, Viola A, Molon B: **CXCR4 signaling in health and disease**. *Immunol Lett* 2016, **177**:6–15.

288. Rueda P, Balabanian K, Lagane B, Staropoli I, Chow K, Levoye A, Laguri C, Sadir R, Delaunay T, Izquierdo E, et al.: **The CXCL12γ chemokine displays unprecedented structural and functional properties that make it a paradigm of chemoattractant proteins**. *PLoS One* 2008, **3**.

289. **2020 Allen Institute for Cell Science. Methods for Cells in the lab.** [date unknown],

290. Lian X, Zhang J, Azarin SM, Zhu K, Hazeltine LB, Bao X, Hsiao C, Kamp TJ, Palecek SP: **Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions**. *Nat Protoc* 2013, **8**:162–175.

291. Burridge PW, Matsa E, Shukla P, Lin ZC, Churko JM, Ebert AD, Lan F, Diecke S, Huber B, Mordwinkin NM, et al.: **Chemically defined generation of human cardiomyocytes**. *Nat Methods* 2014, **11**:855–860.

292. Rouhiainen A, Tumova S, Valmu L, Kalkkinen N, Rauvala H: **Analysis of proinflammatory activity of highly purified eukaryotic recombinant HMGB1 (amphoterin)**. *J Leukoc Biol* 2007, **81**:49–58.

293. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, Antunes R, Arganiska J, Bely B, Bingley M, et al.: **UniProt: A hub for protein information**. *Nucleic Acids Res* 2015, **43**:D204–D212.

294. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al.: **Clustal W and Clustal X version 2.0**. *Bioinformatics* 2007, **23**:2947–2948.

295. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ: **Jalview Version 2--a multiple sequence alignment editor and analysis workbench**. *Bioinformatics* 2009, **25**:1189–1191.

296. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A: **The Proteomics Protocols Handbook**. *Proteomics Protoc Handb* 2005, doi:10.1385/1592598900.

297. Bagos PG, Nikolaou EP, Liakopoulos TD, Tsirigos KD: **Combined prediction of Tat and Sec signal peptides with hidden Markov models**. *Bioinformatics* 2010, **26**:2811–2817.

298. Damerell DR, Strain-Damerell C, Garsot S, Joyce SP, Barrett P, Marsden BD: **SATurn: A modular bioinformatics framework for the design of robust maintainable web-based and standalone applications**. *Bioinformatics* 2019, **35**:349–351.

299. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H: **SignalP 5.0 improves signal peptide predictions using deep neural networks**. *Nat Biotechnol* 2019, **37**:420–423.

300. Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: discriminating signal peptides from transmembrane regions**. *Nat Methods* 2011, **8**:785–786.

301. Castiñeiras TS, Williams SG, Hitchcock AG, Smith DC: **E. coli strain engineering for the production of advanced biopharmaceutical products**. *FEMS Microbiol Lett* 2018, **365**:1–10.

Appendices (Chapter 8): The HMGB1-CXCL12 heterocomplex and its interaction with CXCR4 as a target for tissue repair



Álvaro Viñals Guitart
Saint Hilda's College

In partial fulfilment of the requirements for the title of DPhil in Molecular and Cellular Medicine

Supervised by Prof. Jagdeep Nanchahal, Dr. Nicola Burgess-Brown, Dr. Katharina Duerr

APPENDIX - TABLE OF CONTENTS

Appendix A1 Supplemental Methods ..... - 3 -

Chapter 2 Reagents and primers ..... - 3 -

Suppl. data A1-2-A Reagents and equipment ..... - 3 -

Suppl. data A1-2-B Primer and template combinations for constructs designed ..... - 7 -

Chapter 3 Designed constructs and sequences ..... - 9 -

Suppl. data A1-3-A Entry clone DNA (sequence verified) ..... - 9 -

Suppl. data A1-3-B HMGB1 constructs designed ..... - 12 -

Suppl. data A1-3-C DNA sequences of HMGB1 constructs ..... - 13 -

Suppl. data A1-3-D Protein sequences of HMGB1 constructs (full protein) ..... - 21 -

Suppl. data A1-3-E Designed CXCL12 constructs ..... - 24 -

Suppl. data A1-3-F DNA sequences of CXCL12 constructs ..... - 25 -

Suppl. data A1-3-G Protein sequences of CXCL12 constructs ..... - 26 -

Chapter 5 Sequences ..... - 27 -

Suppl. data A1-5-A HMGB1-C1 construct panel construct list ..... - 27 -

Suppl. data A1-5-B DNA sequences for HMGB1-C1 construct panel ..... - 29 -

Suppl. data A1-5-C Protein sequences of HMGB1-C1 constructs (tagged) ..... - 47 -

Appendix A2 Mass spectrometry data ..... - 51 -

Chapter 3 MS data ..... - 51 -

Suppl. data A2-3-A : Figure 3.3. ..... - 51 -

Suppl. data A2-3-B Figure 3.7: DNA contamination of HMGB1 after size exclusion chromatography (SEC)., pp. 83, tryptic digest MS/MS verification ..... - 52 -

Suppl. data A2-3-C Table 8.3, LC-ESI-TOF/MS verification of HMGB1A constructs. ..... - 53 -

Suppl. data A2-3-D Figure 3.18, pp. 98, all replicate spectra ..... - 55 -

Suppl. data A2-3-E Figure 3.19, pp. 100, all spectra ..... - 59 -

Suppl. data A2-3-F Figure 3.20, pp. 102; all spectra ..... - 61 -

Suppl. data A2-3-G Figure 3.25A, pp.116, tryptic digest MS/MS verification ..... - 65 -

Suppl. data A2-3-H Figure 3.25B, pp.116, tryptic digest MS/MS verification ..... - 65 -

Suppl. data A2-3-I Figure 3.26, pp. 117, tryptic digest MS/MS verification ..... - 66 -

Suppl. data A2-3-J Figure 3.33, pp. 129, all spectra ..... - 68 -

Chapter 4 MS data ..... - 69 -

Suppl. data A2-4-A Figure 4.3, pp. 141 (full range deconvoluted spectra) ..... - 69 -

Chapter 5 MS data ..... - 70 -

Suppl. data A2-5-A Table 8.5, pp. - 26 - (LC-ESI-TOF/MS verification of HMGB1-C1 constructs) ..... - 70 -

Suppl. data A2-5-B Figure 5.18, pp. 205, all spectra (HMGB1-c048). ..... - 76 -

Suppl. data A2-5-C Figure 5.19, pp 205, all spectra (HMGB1-c048). ..... - 77 -

Suppl. data A2-5-D Figure 5.21, pp. 211 (full range deconvolution) ..... - 78 -

Appendix A3 Supplemental data ..... - 81 -

Chapter 3 Figure supplements ..... - 81 -

Suppl. data A3-1-A Function test of Superdex S75 column ..... - 81 -

Suppl. data A3-1-B Figure 3.16, pp. 94 ..... - 82 -

Suppl. data A3-1-C Preparative SEC runs ..... - 84 -

Suppl. data A3-1-D Figure 3.23, pp108; DSF curves for HMGB1 c002 and c031 (rejected) ..... - 85 -

Chapter 4 Figure supplements ..... - 86 -

Suppl. data A3-1-E Figure 4.5, pp. 146, sensor coating interferogram ..... - 86 -

Suppl. data A3-1-F Figure 4.12, pp. 163, sensor coating interferogram and raw traces ..... - 88 -

Appendix A4 NMR tables ..... - 92 -

Chapter 8 Appendices

Chapter 4 NMR spectra (Figure 4.8/Figure 4.9) ..... - 92 -

Suppl. data A4-1-A HMGB1-c028 (94-162, biotinylated) titration with CXCL12A-c021 at 0, 0.42, 0.82 and 1.42 molar equivalents. ..... - 92 -

Suppl. data A4-1-B HMGB1A-c007 (3S, 1-184) with 1:2 molar ratio CXCL12 (1:1 molar ratio CXCL12 to HMG Box), in 10 mM HEPES pH 7.5 150 mM NaCl ..... - 103 -

Suppl. data A4-1-C HMGB1-c038 (89-174, biotinylated) titration with CXCL12A-c021 at 0, 0.42, 0.82 and 1.42 molar equivalents ..... - 113 -

Appendix A5 Statistics ..... - 124 -

Chapter 3 Data analysis ..... - 124 -

Suppl. data A5-3-A Figure 3.6, pp. 82 ..... - 124 -

Suppl. data A5-3-B Figure 3.9, pp. 86 ..... - 125 -

Suppl. data A5-3-C Figure 3.15, pp. 92 ..... - 125 -

Suppl. data A5-3-D Figure 8.1, pp. - 81 - ..... - 126 -

Suppl. data A5-3-E Figure 3.18B, pp. 98 ..... - 126 -

Suppl. data A5-3-F Figure 3.19B, pp. 100 ..... - 128 -

Suppl. data A5-3-G Figure 3.20, pp. 102 ..... - 129 -

Suppl. data A5-3-H Figure 3.23,pp. 110 ..... - 131 -

Suppl. data A5-3-I Figure 3.34, pp. 131 ..... - 131 -

Chapter 4 Data analysis ..... - 132 -

Suppl. data A5-4-A Figure 4.1, pp. 137 ..... - 132 -

Suppl. data A5-4-B Figure 4.2, pp. 139 ..... - 133 -

Suppl. data A5-4-C Figure 4.5, pp. 144 ..... - 134 -

Suppl. data A5-4-D Figure 4.12, pp.162 ..... - 134 -

Chapter 5 Data analysis ..... - 137 -

Suppl. data A5-5-A Figure 5.7, pp. 187 ..... - 137 -

Suppl. data A5-5-B Figure 5.8, pp. 187 ..... - 137 -

Suppl. data A5-5-C Figure 5.9, pp. 188 ..... - 138 -

Suppl. data A5-5-D Figure 5.10, pp. 189 ..... - 138 -

Suppl. data A5-5-E Figure 5.12, pp. 192 ..... - 140 -

Suppl. data A5-5-F Figure 5.13, pp. 196. ..... - 142 -

Suppl. data A5-5-G Figure 5.14, pp. 198 ..... - 146 -

Suppl. data A5-5-H Figure 5.15, pp. 199 ..... - 146 -

Suppl. data A5-5-I Figure 5.16, pp. 201 ..... - 147 -

Suppl. data A5-5-J Figure 5.18, pp. 205. ..... - 148 -

Suppl. data A5-5-K Figure 5.19, pH 2.2, pp. 206 ..... - 148 -

Suppl. data A5-5-L Figure 5.20 , pp. 208 ..... - 148 -

Appendix A1 Supplemental Methods

Chapter 2 Reagents and primers

Suppl. data A1-2-A Reagents and equipment

Unless containing detergents (TX-100, TX-114, DDM, Tween, Antifoam-204 solution), all stocks and buffer solutions were prepared in 18.2 mΩ water and vacuum-filtered through a 0.22µm membrane. Buffers for pump-coupled chromatography were degassed for one hour in the same setup. All chemicals used were anhydrous unless otherwise indicated.

PBS was ubiquitously used, prepared with the formula 2.88 g/L Na2HPO4, 160 g/L NaCl, 4 g/L KH2PO4, 4g KCl, pH 7.4 (dissolved in this order), filter-sterilised and autoclaved, except for LPS-free applications or tissue culture where a commercial sterile formulation was used (Gibco, 10010023) .

Powdered chemicals			
	Stock	**Supplier**	**Item**
$(NH_4)_2SO_4$	1 M (autoclaved)	Sigma	A4915
^{15}N NH_4Cl, U-99	25%	Cambridge Isotopes	NLM-467
Agar	No	Acros Organics	40040050
Agarose, low melting point	No	ThermoFisher	16520
Aluminium sulphate, 18-hydrate	No	Acros Organics	400590025
Ampicillin, sodium salt	50 mg/mL	Sigma	A9518
AIM-TB base	No	ForMedium	AIMTB0210
Barium chloride, powder (99%)	1.2 %	Sigma	202738
BSA, heat-shock fraction	No	Sigma	A9647
$CaCl_2$, $•2H_2O$	2 M	VWR	437055N
Carbenicillin, sodium salt	50 mg/mL	Melford	C46000
CBB-G250	No	MP Biomedicals	190343
Chloramphenicol	34 mg/mL, ethanol	Sigma	C0378
$CoCl_2•6H_2O$	0.2 M	MP Biomedicals	203149
$CuCl_2•6H_2O$	0.1 M	Sigma	194642
d-Biotin	No	VWR	0340
DDM	10%	Generon	D310LA
DTT	1 M	Fluorochem	H02712
EDTA, sodium salt	0.5 M	Fisher Scientific	BP120
$FeCl_3•6H_2O$	No	Sigma	31232
Glucose	25% (1.25 M)	Acros Organics	450740010
Guanidinium-HCl	6M	Apollo Scientific	BIMB2002
HEPES	1 M, pH 7.5	Apollo Scientific	BI8181
Imidazole	3 M, pH 7.5	Sigma	S202
IPTG	1 M	Apollo Scientific	BIMB1008
K_2HPO_4	No	Fisher Scientific	10101490
Kanamycin sulphate	50 mg/mL	Apollo Scientific	BIK0126
KCl	No	Sigma	P9333
KH_2PO_4	No	Merck	1.05101
$MgCl$, $• 6H_2O$	2 M	Sigma	M2670
$MgSO_4$	1 M (autoclaved)	Sigma	M2643
$MnCl_2•4H_2O$	1 M	Acros Organics	20589
Na_2HPO_4, $•2H_2O$	No	Sigma	71643
NaCl	5 M	Sigma	S3014
NaOH	10 M	Sigma	221465
NH_4Cl	25%	Sigma	213330
$NiCl_2•6H_2O$	1 M	Sigma	364304
$NiSO_4$	1 M, in 0.2 M Tris pH 8.0	Fisher	10055413
Peptone	No	Merck	107213
Streptomycin sulphate	50 mg/mL	Fisher Scientific	BP910
Sucrose	25%	Fisher Scientific	BP220
TCEP	0.5 M	Apollo Scientific	BIT0122
Tris base	2 M, pH 7.5	Fisher Scientific	BP152
Urea	9 M	Sigma	U5378
Yeast extract	No	Melford	420025
$ZnSO_4•7H_2O$	1 M	Sigma	74750

Commercial buffers and solutions			
	Concentration/%	**Supplier**	**Item**
1 kB plus DNA standard	1X	ThermoFisher	10787018
Acetic acid	33%	SLS	CHE1008
Acetonitrile (ACN)	LC-MS, 99.9%	Fisher Scientific	10001334
Ammonium acetate	7.5 M, pH 6.8	Sigma	A2706
Ammonium carbonate	1 M, pH 8.0	Boston Bioproducts	BB-2609
Antifoam-204	100%	Sigma	A6476
BSA, molecular biology grade	1X	NEB	B9000S
cOmplete protease inhibitor tablets, EDTA-free	Tablet	Roche	11836145001
DNA loading buffer, 2X		NEB	B7024S
DMSO	100%, for cell culture	Sigma	D2650
dNTPs	100 mM	Invitrogen	10927018
DPBS, sterile, no calcium, magnesium, or phenol red	10X/1X	Gibco	14200075/ 14190144
Ethanol	99.9%	Sigma	32221
Formic acid (FA)	LC-MS, 99.9%	Fisher Scientific	10325960
Glycerol	98%	Fisher Scientific	10336040
H_2SO_4	95-98%	Sigma	S9300
HCl	37%	Fisher Scientific	10596814
HEPES	1 M, pH (6.2/8.0)	Molecular Dimensions	MD2-011-6.2/8.0
Isopropanol	99.9%	Sigma	59300
LDS loading buffer	4X	ThermoFisher	NP0008
MES/SDS buffer	20X	ThermoFisher	NP0002
Methanol	99.99%	Acros Organics	326950025
Methanol	LC-MS	Fisher Scientific	A456-212
Paraformaldehyde	3.7%, in PBS	Santa Cruz Biotech.	sc-281692
PeqGreen DNA dye	15000X	VWR	732-3196
Phosphoric acid, 85% w/v	85% v/v	Fisher Scientific	10792212
Protease inhibitors, set III, EDTA-free	1000X in DMSO	Merck	**539134**
Protein MW standard, SDS-PAGE	1X	BioRad	1610363
SEC MW standards	1X	BioRad	1511901
SPG buffer	1 M, pH (4.0/10.0)	Molecular Dimensions	MD2-59-4.0/10.0
Tris-HCl	1 M, pH (7.0/9.0)	Molecular Dimensions	MD2-027-7.0/9.0
Triton X-110	100%	Sigma	T8787
Triton X-114	100%	Sigma	X114
Tween-20	100%	Sigma	P9416
Water	HPLC grade	Fisher Scientific	10777404

Eukaryotic cell culture		
Media components		
	Supplier	**Item**
DMEM, high glucose	Sigma	D5796
Trypsin Solution, 1X (TrypLE)	Thermo Fisher	12563011
Heat inactivated HI-FBS	Thermo Fisher	10082147
Pen/Strep 100X mixture	Thermo Fisher	15140122
Trypan Blue, 0.4% solution	Thermo Fisher	T10282
Disposables		
T75 flask, vented cap, cell culture treated	Thermo Fisher	156499
60 mm Petri dish, tissue culture treated	StemCell Technologies	100-0099

E. coli cell culture		
	Supplier	**Item**
AeroTop air porous seal	Thomson Scientific	899423 (250 mL) 899423 (2.5 L)
UltraYield plastic flasks		931141 (0.5 L) 931136-B (2.5 L)

Commercial kits		
	Supplier	**Item**
4-12% Midi 26-well acrylamide gradient gels	ThermoFisher	WG1403Box
PureLink PCR purification kit	Invitrogen	K310001
NucleoSpin Miniprep kit	Macherey-Nagel	740588

Disposable plasticware consumables		
	Supplier	**Item**
0.2 mL PCR strips	Appleton Woods	B5151
0.22 µM syringe Luer-Lock filters, 30 mm, sterile	Sartorius	16541Q
0.22 µM syringe Luer-Lock filters, 4 mm, sterile	Merck	SLGV004SL
1.2 µM syringe Luer-Lock filters, 30 mm, sterile	Sartorius	17593K
30/10/3 kDa MWCO concentrator, 4 mL	Merck	UFC 803096/801096/800396
30/10/3 kDa MWCO concentrator, 15 mL	Merck	UFC903096/904096/900396
384-format 0.1 mL filter pipette tips, sterile	Eppendorf	0030078411
5 µM syringe Luer-Lock filters, 30 mm, sterile	Sartorius	17594Q
5 kDa MWCO concentrator, 0.5 mL	GE Healthcare	28-9322-23
5 mL filter pipette tips, sterile	Eppendorf	0030077725
96-well 2 mL deep-well blocks	Greiner Bio-One	780270
96-well PCR plates, skirted	Appleton Woods	TA841
96-well V-bottom plate	Greiner Bio-One	651161
Dialysis membrane, 3.5 kDa MWCO	ThermoFisher	68035
DNAse, RNAse, Pyrogen-free filter for water source	Merck	CDUFBI001
Fluorescence optical seal for 96-well plates	BioRad	223944
Heat resistant PCR foil seal, strong	4Titude	4ti-0550FI
Low-protein binding centrifuge microtubes (0.5, 1.5, 2.0 and 5.0 mL)	Eppendorf	00301(08094/08116/08132/08302
Paraffin film	Appleton Woods	6C783
Polypropylene centrifuge 50 mL tubes (rated for 17000 *x g)*	Sigma	CLS430828
Sterile Luer-Lock syringes (5, 20, and 60 mL)	Fisher Scientific	15869152, 15889152, 15899152
Sterile Luer-Lock syringe (1 mL)	Medisave BD	SYR283
Sterile filter pipette tips, RPT (1 mL, 0.2 mL, 10/20 µL), sterile	TipOne	S1182-1370,S1180-8810, S1180-3710
Sterile plastic bottle, 1 L	ThermoFisher	455-1000
Vacuum filter, 0.22 µM, 0.5 L, sterile	ThermoFisher	291-3320
Wide-bore pipette tips, filter, 1 mL	TipOne	E1011-9618
Machinery and electronics		
18.2 mΩ water source	Elga	ELGA_21026
MultiScreen HTS Manifold	Merck	MSVMHTS00
Benchtop microcentrifuge, refrigerated (example)	Axygen	601-05-021
Benchtop centrifuge, refrigerated, for microplate, 15- and 50- mL conical tubes (example)	Eppendorf	5811000065
Floor-top high-speed centrifuge	Beckman Coulter	Avanti JXN-30 B38621
JLA 8.1000; *Rotor, 1 L x6*	Beckman Coulter	969329
JS 5.3; *Rotor, multipurpose*	Beckman Coulter	368690
Centrifuge tubes	Beckman Coulter	A98814 (1 L)
	Corning	431123 (500 mL)
Absorbance-capable plate reader, 96-/384- well plates	BMG Labtech	Fluostar OMEGA
Agarose gel tanks, casting mould and combs	BioRad	1704487
Electrophoresis power supply (300W/3A)	BioRad	1645052
Gel tanks (SDS)	ThermoFisher	WR0100
Laminar airflow tissue culture cabin		
New Brunswick Innova 44 incubation shaker	Eppendorf	M1282-0002
Tissue culture incubator, with temperature (37ºC) and CO_2 (5%) control. *These conditions were used for all eukaryotic cell cultures*		
Automated cell counter (Luna) and slides	Logos Biosystems	L10001 and L12001
FPLC system, AKTA Xpress (UNICORN 5.31 software)	GE Healthcare	28-9089-36
FPLC system, NGC Discover	Biorad	7880009
Surgical tools and consumables		
	Manufacturer	**Item**
0.27G needle, MicroLance	Medisave BD	ND635
0.29G needle and syringe, MicroFine+	Medisave BD	320806
Sterile swabs	Premier	1860
Sterile saline solution		
Surgical drapes	Bastos Viegas	488-101

Suppl. data A1-2-B Primer and template combinations for constructs designed

Primer ID	*Primer sequence (5`->3`)*
HMGB1WTLICF	**ATGGGCAAAGGCGATCCGAAAAAACCGC**
HMGB1C106S	**GCGATACTCAGAGCTGAAGAGGAAGAAGGC**
HMGB1C45S	**TTTTCTAAGAAGAGCTCAGAGAGGTGGAA**
HMGB1WTLICR	**TTATTCATCATCATCATCTTCTTCTTCA**
HMGB1C23S	**TTTTTGTGCAGACCAGCCGCGAAGAACA**
CXCL12A-f001	TTAAGAAGGAGATATACTATGAACGCCAAGGTCGTGGTCG
CXCL12A-f003	TTAAGAAGGAGATATACTATGAAGCCCATCAGCCTGAGCTAC
CXCL12A-f004	TTAAGAAGGAGATATACTATGAAGCCCGTCAGCCTGAGCTAC
CXCL12A-r001	GATTGGAAGTAGAGGTTCTCTGCCTTGTTTAAAGCTTTCTCCAGG
CXCL12A-r003	GATTGGAAGTAGAGGTTCTCTGCCTTTAGCTTCGGGTCAATGC
CXCL12RevLicDSB	TATCCACCTTTACTGTCATTACTTGTTTAAAGCTTTCTCCAGGTA
CXCL12RevLicOmpA	ACCACCACCACTACTGCCTTACTTGTTTAAAGCTTTCTCCAGGTA
HMGB1A-f001	TTAAGAAGGAGATATACTATGGGCAAAGGAGATCCTAAGAAGCCG
HMGB1A-f002	TACTTCCAATCCATGGGCAAAGGAGATCCTAAGAAGC
HMGB1A-f003	TTAAGAAGGAGATATACTATGAGAGGCAAAATGTCATCATATGCATT
HMGB1A-f004	TTAAGAAGGAGATATACTATGCCGAGAGGCAAAATGTCATCATATGC
HMGB1A-f005	TACTTCCAATCCATGCCGAGAGGCAAAATGTCA
HMGB1A-f006	TACTTCCAATCCATGCCGAGAGGCAAAATGTCATCATATG
HMGB1A-f007	TACTTCCAATCCATGCCCAAGAGGCCTCCTTCG
HMGB1A-f008	TACTTCCAATCCATGGGCAAAGGAGATCCTAAGAAGC
HMGB1A-f009	TTAAGAAGGAGATATACTATGGGCAAAGGAGATCCTAAGAAGCCG
HMGB1A-f010	TTAAGAAGGAGATATACTATGCCGAGAGGCAAAATGTCATCATATGC
HMGB1A-f011	TTAAGAAGGAGATATACTATGAAGGATCCCAATGCACCCAAGAG
HMGB1A-f012	TACTTCCAATCCATGAAGGATCCCAATGCACCC
HMGB1A-f013	TTAAGAAGGAGATATACTATGCCCAAGAGGCCTCCTTCG
HMGB1A-r001	GATTGGAAGTAGAGGTTCTCTGCTTCATCATCATCATCTTCTTCTTC
HMGB1A-r002	TATCCACCTTTACTGCTTTCATCATCATCATCTTCTTCTTCATCTTCATC
HMGB1A-r003	GATTGGAAGTAGAGGTTCTCTGCCTTCTTTTTCTTGCTTTTTTCAGC
HMGB1A-r004	TATCCACCTTTACTGCTTCGATATGCAGCAATATCCTTTTCG
HMGB1A-r005	GATTGGAAGTAGAGGTTCTCTGCTTTAGCTCGATATGCAGCA
HMGB1A-r006	TATCCACCTTTACTGCTCTTCTTTTTCTTGCTTTTTTCAGCCT
HMGB1A-r007	TATCCACCTTTACTGCTGATATAGGTTTTCATTTCTCTTTCATAACGGG
HMGB1A-r008	TATCCACCTTTACTGCTTTTAGCTCGATATGCAGCAATATCCT
HMGB1A-r009	TATCCACCTTTACTGCTAACTCCCTTTTTTGCTGCATCA
HMGB1A-r010	GATTGGAAGTAGAGGTTCTCTGCAACTCCCTTTTTTGCTGCA
HMGB1A-r011	GATTGGAAGTAGAGGTTCTCTGCTCGATATGCAGCAATATCCT
HMGB1A-r012	TATCCACCTTTACTGCTGAACTTCTTTTTTGTCTCCCCT
HMGB1A-r013	GATTGGAAGTAGAGGTTCTCTGCGAACTTCTTTTTTGTCTCCCC
HMGB1A-r014	GATTGGAAGTAGAGGTTCTCTGCGATATAGGTTTTCATTTCTCTTTCATAACG
SUMOFwdLicDSB_OMPA	TACTTCCAATCCATGTCTGACCAGGAGGCAAAACCTTCAACTG
Sumo_CXCL12c003	GCATCTGTAGCTCAGGCTGACGGGCTTACCTCCCGTTTGCTCCTGGTAAACTTC
Sumo_CXCL12c006	GCATCTGTAGCTCAGGCTGATGGGCTTACCTCCCGTTTGCTCCTGGTAAACTTC
CXCL12c003_SUMO	GAAGTTTACCAGGAGCAAACGGGAGGTAAGCCCGTCAGCCTGAGCTACAGATGC
CXCL12c006_SUMO	GAAGTTTACCAGGAGCAAACGGGAGGTAAGCCCATCAGCCTGAGCTACAGATGC

Table 8.1: List of primers used in generation of PCR products (alleles)

Allele ID	Forward primer ID	Reverse primer ID	Entry clone ID	Vector	Construct ID
HMGB1A-a001	HMGB1A-f001	HMGB1A-r001	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c001**
HMGB1A-a002	HMGB1A-f001	HMGB1A-r003	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c002**
HMGB1A-a003	HMGB1A-f003	HMGB1A-r003	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c003**
HMGB1A-a004	HMGB1A-f003	HMGB1A-r001	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c004**
HMGB1A-a005	HMGB1A-f003	HMGB1A-r005	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c005**
HMGB1A-a006	HMGB1A-f001	HMGB1A-r001	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c006**
HMGB1A-a007	HMGB1A-f001	HMGB1A-r003	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c007**
HMGB1A-a008	HMGB1A-f003	HMGB1A-r003	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c008**
HMGB1A-a009	HMGB1A-f003	HMGB1A-r001	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c009**
HMGB1A-a010	HMGB1A-f003	HMGB1A-r005	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c010**
HMGB1A-a018	HMGB1A-f002	HMGB1A-r002	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c011**
HMGB1A-a012	HMGB1A-f004	HMGB1A-r003	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c012**
HMGB1A-a013	HMGB1A-f004	HMGB1A-r001	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c013**
HMGB1A-a014	HMGB1A-f004	HMGB1A-r005	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c014**
HMGB1A-a015	HMGB1A-f004	HMGB1A-r003	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c015**
HMGB1A-a016	HMGB1A-f004	HMGB1A-r001	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c016**
HMGB1A-a017	HMGB1A-f004	HMGB1A-r005	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c017**
HMGB1A-a019	HMGB1A-f002	HMGB1A-r006	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c018**
HMGB1A-a020	HMGB1A-f006	HMGB1A-r006	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c019**
HMGB1A-a021	HMGB1A-f006	HMGB1A-r002	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c020**
HMGB1A-a022	HMGB1A-f006	HMGB1A-r008	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c021**
HMGB1A-a023	HMGB1A-f002	HMGB1A-r002	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c022**
HMGB1A-a024	HMGB1A-f002	HMGB1A-r006	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c023**
HMGB1A-a025	HMGB1A-f006	HMGB1A-r006	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c024**
HMGB1A-a026	HMGB1A-f006	HMGB1A-r002	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c025**
HMGB1A-a027	HMGB1A-f006	HMGB1A-r008	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c026**
HMGB1A-a028	HMGB1A-f005	HMGB1A-r007	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c027**
HMGB1A-a029	HMGB1A-f007	HMGB1A-r004	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c028**
HMGB1A-a030	HMGB1A-f005	HMGB1A-r007	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c029**
HMGB1A-a031	HMGB1A-f007	HMGB1A-r004	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c030**
HMGB1A-a032	HMGB1A-f001	HMGB1A-r005	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c031**
HMGB1A-a033	HMGB1A-f002	HMGB1A-r008	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c032**
HMGB1A-a034	HMGB1A-f009	HMGB1A-r013	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c033**
HMGB1A-a035	HMGB1A-f011	HMGB1A-r010	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c034**
HMGB1A-a036	HMGB1A-f010	HMGB1A-r014	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c035**
HMGB1A-a037	HMGB1A-f013	HMGB1A-r011	HMGB1A-s001	pNIC-CTHF	**HMGB1A-c036**
HMGB1A-a038	HMGB1A-f008	HMGB1A-r012	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c037**
HMGB1A-a039	HMGB1A-f012	HMGB1A-r009	HMGB1A-s001	pNIC-Bio3	**HMGB1A-c038**
HMGB1A-a040	HMGB1A-f009	HMGB1A-r013	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c039**
HMGB1A-a041	HMGB1A-f011	HMGB1A-r010	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c040**
HMGB1A-a042	HMGB1A-f010	HMGB1A-r014	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c041**
HMGB1A-a043	HMGB1A-f013	HMGB1A-r011	HMGB1A-s002	pNIC-CTHF	**HMGB1A-c042**
HMGB1A-a044	HMGB1A-f008	HMGB1A-r012	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c043**
HMGB1A-a045	HMGB1A-f012	HMGB1A-r009	HMGB1A-s002	pNIC-Bio3	**HMGB1A-c044**
CXCL12A-a001	CXCL12A-f001	CXCL12A-r001	CXCL12A-s001	pNIC-CTHF	**CXCL12A-c001**
CXCL12A-a002	CXCL12A-f004	CXCL12A-r003	CXCL12A-s001	pNIC-CTHF	**CXCL12A-c002**
CXCL12A-a003	CXCL12A-f004	CXCL12A-r001	CXCL12A-s001	pNIC-CTHF	**CXCL12A-c003**
CXCL12A-a004	CXCL12A-f001	CXCL12A-r001	CXCL12A-s002	pNIC-CTHF	**CXCL12A-c004**
CXCL12A-a005	CXCL12A-f003	CXCL12A-r003	CXCL12A-s002	pNIC-CTHF	**CXCL12A-c005**
CXCL12A-a006	CXCL12A-f003	CXCL12A-r001	CXCL12A-s002	pNIC-CTHF	**CXCL12A-c006**
CXCL12A-a008	CXCL12A-f004	CXCL12A-r003	CXCL12A-s001	pFB-CT10HF-LIC	**CXCL12A-c008**
CXCL12A-a009	CXCL12A-f004	CXCL12A-r001	CXCL12A-s001	pFB-CT10HF-LIC	**CXCL12A-c009**
CXCL12A-a011	CXCL12A-f003	CXCL12A-r003	CXCL12A-s002	pFB-CT10HF-LIC	**CXCL12A-c011**
CXCL12A-a012	CXCL12A-f003	CXCL12A-r001	CXCL12A-s002	pFB-CT10HF-LIC	**CXCL12A-c012**
CXCL12A-a013	SUMOFwdLicDSB_OMPA	CXCL12RevLicDSB	CXCL12A-s001	pDsbC-HT-CBio	**CXCL12A-c021**
CXCL12A-a014		CXCL12RevLicOmpA		pNIC-OmpA-NHT-cTStII	**CXCL12A-c022**
				pNIC-SufI-NHT-cTStII	**CXCL12A-c023**
				pNIC-TolB-NHT-cTStII	**CXCL12A-c024**
CXCL12A-a015		CXCL12RevLicDSB	CXCL12A-s002	pDsbC-HT-CBio	**CXCL12A-c025**
CXCL12A-a016		CXCL12RevLicOmpA		pNIC-OmpA-NHT-cTStII	**CXCL12A-c026**
				pNIC-SufI-NHT-cTStII	**CXCL12A-c027**
				pNIC-TolB-NHT-cTStII	**CXCL12A-c028**

Table 8.2: Primer and vector combinations for allele and construct generation

Chapter 3 Designed constructs and sequences

Suppl. data A1-3-A Entry clone DNA (sequence verified)

Entry clones comprise direct amplification of the complementary DNA stored in the MGC. Construct DNA sequences comprise the coding sequence of the construct, without LIC adaptions or additions from the vector. Both types of sequence have been colored by nucleotide. CXCR4A-c038 has been added despite not being an entry clone as I only did the design of that construct but did not express it myself.

Entry clone ID	DNA sequence
HMGB1A-s001	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG GCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCT GCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAA AGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAG GAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAATAA
HMGB1A-s002	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG GCCTTCTTCCTCTTCAGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCT GCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAA AGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAG GAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAATAA
CXCL12A-s001 (signal peptide underlined)	ATGAACGCCAAGGTCGTGGTCGTGCTGGTCCTCGTGCTGACCGCGCTCTGCCTCAGCGAC GGGAAGCCCGTCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAG GAGTACCTGGAGAAAGCTTTAAACAAGTAA
CXCL12A-s002 (signal peptide underlined)	ATGAACGCCAAGGTCGTGGTCGTGCTGGTCCTCGTGCTGACCGCGCTCTGCCTCAGCGAC GGGAAGCCCATCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAG GAGTACCTGGAGAAAGCTTTAAACAAGTAA

Entry clone ID	DNA sequence
CXCR4A-c038 (L125W, T240P, 1-325, Isoform I)	ATGGAGGGGATCAGTATATACACTTCAGATAACTACACCGAGGAAATGGGCTCAGGGGAC TATGACTCCATGAAGGAACCCTGTTTCCGTGAAGAAAATGCTAATTTCAATAAAATCTTC CTGCCCACCATCTACTCCATCATCTTCTTAACTGGCATTGTGGGCAATGGATTGGTCATC CTGGTCATGGGTTACCAGAAGAAACTGAGAAGCATGACGGACAAGTACAGGCTGCACCTG TCAGTGGCCGACCTCCTCTTTGTCATCACGCTTCCCTTCTGGGCAGTTGATGCCGTGGCA AACTGGTACTTTGGGAACTTCCTATGCAAGGCAGTCCATGTCATCTACACAGTCAACCTC TACAGCAGTGTCTGGATCCTGGCCTTCATCAGTCTGGACCGCTACCTGGCCATCGTCCAC GCCACCAACAGTCAGAGGCCAAGGAAGCTGTTGGCTGAAAAGGTGGTCTATGTTGGCGTC TGGATCCCTGCCCTCCTGCTGACTATTCCCGACTTCATCTTTGCCAACGTCAGTGAGGCA GATGACAGATATATCTGTGACCGCTTCTACCCCAATGACTTGTGGGTGGTTGTGTTCCAG TTTCAGCACATCATGGTTGGCCTTATCCTGCCTGGTATTGTCATCCTGTCCTGCTATTGC ATTATCATCTCCAAGCTGTCACACTCCAAGGGCCACCAGAAGCGCAAGGCCCTCAAGCCC ACAGTCATCCTCATCCTGGCTTTCTTCGCCTGTTGGCTGCCTTACTACATTGGGATCAGC ATCGACTCCTTCATCCTCCTGGAAATCATCAAGCAAGGGTGTGAGTTTGAGAACACTGTG CACAAGTGGATTTCCATCACCGAGGCCCTAGCTTTCTTCCACTGTTGTCTGAACCCCATC CTCTATGCTTTCCTTGGAGCCAAATTTAAAACCTCTGCCCAGCACGCACTCACCTCTGTG AGCAGAGGGTCCAGCGCAGAGAACCTCTACTTCCAATCGGGCAGCGCGGTGAGCAAGGGC GAGGAGCTGTTCACGGGTGTCGTACCGATTTTGGTAGAACTGGATGGGGACGTAAACGGA CACAAGTTCAGCGTATCCGGGGAAGGGGAGGGGGACGCCACATATGGAAAACTGACCTTG AAGTTCATCTGTACGACAGGTAAACTTCCAGTGCCATGGCCTACTCTCGTCACGACTTTG ACTTATGGAGTACAGTGTTTCTCGAGATATCCGGACCACATGAAACAGCACGATTTCTTC AAATCAGCGATGCCCGAAGGTTATGTACAAGAAAGGACCATCTTCTTTAAGGATGACGGT AACTACAAGACACGGGCTGAGGTGAAATTCGAGGGTGACACGCTGGTCAACCGCATTGAG CTTAAGGGGATCGACTTTAAGGAGGACGGGAACATCCTCGGACATAAACTCGAATACAAT TACAATTCGCACAATGTCTACATCATGGCGGATAAGCAAAAGAACGGTATCAAAGTAAAC TTCAAAATCAGACACAACATCGAGGACGGGTCCGTGCAATTGGCAGATCACTACCAGCAG AATACACCTATTGGTGACGGGCCTGTGCTGCTGCCCGATAACCACTACTTGTCGACGCAG TCGAAACTGTCAAAAGACCCCAACGAGAAGAGGGATCACATGGTCTTGCTCGAGTTCGTC ACTGCGGCAGGGATTACCCTTGGGATGGACGAACTCTACAAGTCAGGACTTCGCTCGTGG TCACATCCTCAGTTCGAGAAGGGTGGTGGTTCTGGTGGTGGTTCTGGTGGTTCTGCTTGG TCACATCCTCAGTTCGAGAAGCATCACCACCATCATCATCACCACCATCAC

Suppl. data A1-3-B HMGB1 constructs designed

Construct ID	mW (Da)	mW (Da), no tag	ε (M*cm)$^{-1}$	Vector	Start	End	Use	Notes
HMGB1A-c001	27664.7	25759.8	22920	pNIC-CTHF	Met (0)	Glu 214	I, X	WT, initial panel
HMGB1A-c002	23931.5	22026.6	22920	pNIC-CTHF	Met (0)	Lys 184	I, X	WT, initial panel
HMGB1A-c003	*23123.5*	*21218.6*	*22920*	*pNIC-CTHF*	*Arg 9*	*Lys 184*	*I, X*	*WT, initial panel*
HMGB1A-c004	26856.7	24951.8	22920	pNIC-CTHF	Arg 9	Glu 214	I, X	WT, initial panel
HMGB1A-c005	21015.9	19111	22920	pNIC-CTHF	Arg 9	Lys 164	I, X	WT, initial panel
HMGB1A-c006	27616.5	25711.6	22920	pNIC-CTHF	Met (0)	Glu 214	I, X	3S, initial panel
HMGB1A-c007	23883.3	21978.4	22920	pNIC-CTHF	Met (0)	Lys 184	I, X	3S, initial panel
HMGB1A-c008	23075.3	21170.4	22920	pNIC-CTHF	Arg 9	Lys 184	I, X	3S, initial panel
HMGB1A-c009	26808.5	24903.6	22920	pNIC-CTHF	Arg 9	Glu 214	I, X	3S, initial panel
HMGB1A-c010	20967.7	19062.9	22920	pNIC-CTHF	Arg 9	Lys 164	I, X	3S, initial panel
HMGB1A-c011	29837.2	27371.5	28420	pNIC-Bio3	Met (0)	Glu 214	I, X	WT, v.3 (interaction test)
HMGB1A-c012	23220.6	21315.7	22920	pNIC-CTHF	Pro 8	Lys 184	I, X	WT, v.2 (revision)
HMGB1A-c013	26953.8	25048.9	22920	pNIC-CTHF	Pro 8	Glu 214	I, X	WT, v.2 (revision)
HMGB1A-c014	21113.1	19208.2	22920	pNIC-CTHF	Pro 8	Lys 164	I, X	WT, v.2 (revision)
HMGB1A-c015	23172.4	21267.5	22920	pNIC-CTHF	Pro 8	Lys 184	I, X	3S, v.2 (revision)
HMGB1A-c016	26905.6	25000.7	22920	pNIC-CTHF	Pro 8	Glu 214	I, X	3S, v.2 (revision)
HMGB1A-c017	21064.9	19160	22920	pNIC-CTHF	Pro 8	Lys 164	I, X	3S, v.2 (revision)
HMGB1A-c018	26104	23638.3	28420	pNIC-Bio3	Met (0)	Lys 184	B, N, P	WT, v.3 (interaction test)
HMGB1A-c019	25393.1	22927.5	28420	pNIC-Bio3	Pro 8	Lys 184	B, N, P	WT, v.3 (interaction test)
HMGB1A-c020	29126.3	26660.7	28420	pNIC-Bio3	Pro 8	Glu 214	B, N, P	WT, v.3 (interaction test)
HMGB1A-c021	23285.6	20819.9	28420	pNIC-Bio3	Pro 8	Lys 164	B, N, P	WT, v.3 (interaction test)
HMGB1A-c022	29789	27323.3	28420	pNIC-Bio3	Met (0)	Glu 214	B, N, P	3S, v.3 (interaction test)
HMGB1A-c023	26055.8	23590.1	28420	pNIC-Bio3	Met (0)	Lys 184	B, N, P	3S, v.3 (interaction test)
HMGB1A-c024	25344.9	22879.3	28420	pNIC-Bio3	Pro 8	Lys 184	B, N, P	3S, v.3 (interaction test)
HMGB1A-c025	*29078.1*	*26612.5*	*28420*	*pNIC-Bio3*	*Pro 8*	*Glu 214*	*B, N, P*	*3S, v.3 (interaction test)*
HMGB1A-c026	23237.4	20771.7	28420	pNIC-Bio3	Pro 8	Lys 164	B, N, P	3S, v.3 (interaction test)
HMGB1A-c027	13536.3	11070.6	16960	pNIC-Bio3	Pro 8	Ile 78	B, N, P	WT, v.4 - NMR
HMGB1A-c028	12975.6	10510	18450	pNIC-Bio3	Pro 94	Arg 162	B, N, P	WT, v.4 - NMR
HMGB1A-c029	13504.1	11038.5	16960	pNIC-Bio3	Pro 8	Ile 78	B, N, P	3S, v.4 - NMR
HMGB1A-c030	12959.6	10493.9	18450	pNIC-Bio3	Pro 94	Arg 162	B, N, P	3S, v.4 - NMR
HMGB1A-c031	21823.9	19919	22920	pNIC-CTHF	Met (0)	Lys 164	I, N	WT, v.2 (revision)
HMGB1A-c032	23996.4	21530.7	28420	pNIC-Bio3	Met (0)	Lys 164	B, N, P	WT, v.3 (interaction test)
HMGB1A-c033	13216	11311.1	11460	pNIC-CTHF	Met (0)	Phe 88	I, N	WT, v.5
HMGB1A-c034	12480.1	10575.2	12950	pNIC-CTHF	Lys 89	Arg 174	I, N	WT, v.5
HMGB1A-c035	11363.8	9458.87	11460	pNIC-CTHF	Pro9	Ile 78	I, N	WT, v.5
HMGB1A-c036	10803.2	8898.26	12950	pNIC-CTHF	Lys 89	Val 162	I, N	WT, v.5
HMGB1A-c037	15388.5	12922.8	16960	pNIC-Bio3	Met (0)	Phe 88	B, N, P	WT, v.5
HMGB1A-c038	14652.6	12186.9	18450	pNIC-Bio3	Lys 89	Val 174	B, N, P	WT, v.5
HMGB1A-c039	13183.8	11279	11460	pNIC-CTHF	Met (0)	Phe 89	I, N	3S, v.5
HMGB1A-c040	12464	10559.1	12950	pNIC-CTHF	Lys 89	Val 174	I, N	3S, v.5
HMGB1A-c041	11331.6	9426.74	11460	pNIC-CTHF	Pro9	Ile 78	I, N	3S, v.5
HMGB1A-c042	10787.1	8882.19	12950	pNIC-CTHF	Pro94	Arg 162	I, N	3S, v.5
HMGB1A-c043	*15356.3*	*12890.7*	*16960*	*pNIC-Bio3*	*Met (0)*	*Phe 88*	*B, N, P*	*3S, v.5*
HMGB1A-c044	*14636.5*	*12170.9*	*18450*	*pNIC-Bio3*	*Lys 89*	*Val 174*	*B, N, P*	*3S, v.5*

Table 8.3: Designed HMGB1 constructs for characterisation of the HMGB1-CXCL12 interaction

Legend key: B=Biolayer interferometry, N=NMR, P=Peptide arrays, I=*In vitro* or *in vivo* assays, X=X-ray crystallography. Proteins in pNIC-Bio3 were purified from *E. coli* strains overexpressing BirA. Version numbers correspond to specific experiments which led to the development of that construct series. Version 2 was designed after the initial test expression and upscales of the first HMGB1 panel. Version 3 is a re-cloning of HMGB1A-c001, c002, c006, c007, and c012-c017 in biotinylation vectors. Version 4 consists of minimal Box A and Box B constructs for NMR. Version 5 is a revision based on the NMR and peptide array data. Constructs in italics were not completed and they were discarded as new data made them obsolete. Constructs were verified by LC-ESI/QTOF-MS (Suppl. data A1-3-B). Full DNA and protein sequences can be found in Suppl. data A2-1-B**.** and Suppl. data A2-1-C

Suppl. data A1-3-C DNA sequences of HMGB1 constructs

Construct ID	DNA sequence
HMGB1A-c001	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG GCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCT GCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAA AGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAG GAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A-c002	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG GCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCT GCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAA AGCAAGAAAAAGAAG
HMGB1A-c003	ATGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCATAAG AAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAGAGG TGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAG GCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAG TTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCT GATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG
HMGB1A-c004	ATGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCATAAG AAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAGAGG TGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAG GCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAG TTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCT GATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAG GAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAA GAAGAAGATGATGATGATGAA
HMGB1A-c005	ATGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCATAAG AAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAGAGG TGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAG GCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAG TTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA

Construct ID	DNA sequence
HMGB1A-c006	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG GCCTTCTTCCTCTTCAGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCT GCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAA AGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAG GAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A-c007	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG GCCTTCTTCCTCTTCAGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCT GCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAA AGCAAGAAAAAGAAG
HMGB1A-c008	ATGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAGGAGCATAAG AAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGCTCAGAGAGG TGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAG GCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAG TTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCT GATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG
HMGB1A-c009	ATGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAGGAGCATAAG AAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGCTCAGAGAGG TGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAG GCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAG TTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCT GATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAG GAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAA GAAGAAGATGATGATGATGAA
HMGB1A-c010	ATGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAGGAGCATAAG AAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGCTCAGAGAGG TGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAG GCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAG TTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA

Construct ID	DNA sequence
HMGB1A-c011	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTT TTTGTGCAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTC TCAGAGTTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTAT ATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCT CCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCT GGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCA GATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGAT ATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCT GAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAG GAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A-c012	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAG AAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCT GAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAG AAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAG GCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAG CCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG
HMGB1A-c013	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAG AAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCT GAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAG AAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAG GCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAG CCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAG GAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAAGATGAAGAAGATGAAGAT GAAGAAGAAGATGATGATGATGAA
HMGB1A-c014	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAG AAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCT GAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAG AAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAG GCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA
HMGB1A-c015	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAG AAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCT GAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAG AAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAG GCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAG CCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG

Construct ID	DNA sequence
HMGB1A-c016	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAG AAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCT GAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAG AAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAG GCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAG CCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAG GAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGAT GAAGAAGAAGATGATGATGATGAA
HMGB1A-c017	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAG AAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCT GAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAG AAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAG GCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA
HMGB1A-c018	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTT TTTGTGCAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTC TCAGAGTTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTAT ATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCT CCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCT GGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCA GATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGAT ATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCT GAAAAAAGCAAGAAAAAGAAG
HMGB1A-c019	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACA AAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTC TGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTT GCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAA AAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA GGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG
HMGB1A-c020	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACA AAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTC TGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTT GCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAA AAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA GGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG GAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGAT GAAGATGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A-c021	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACA AAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTC TGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTT GCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAA AAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA
HMGB1A-c022	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTT TTTGTGCAAACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTC TCAGAGTTTTCTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTAT ATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCT CCTTCGGCCTTCTTCCTCTTCAGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCT GGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCA GATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGAT ATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCT GAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAG GAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A-c023	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTT TTTGTGCAAACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTC TCAGAGTTTTCTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTAT ATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCT CCTTCGGCCTTCTTCCTCTTCAGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCT GGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCA GATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGAT ATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCT GAAAAAAGCAAGAAAAAGAAG
HMGB1A-c024	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACA AAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTC AGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTT GCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAA AAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA GGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG

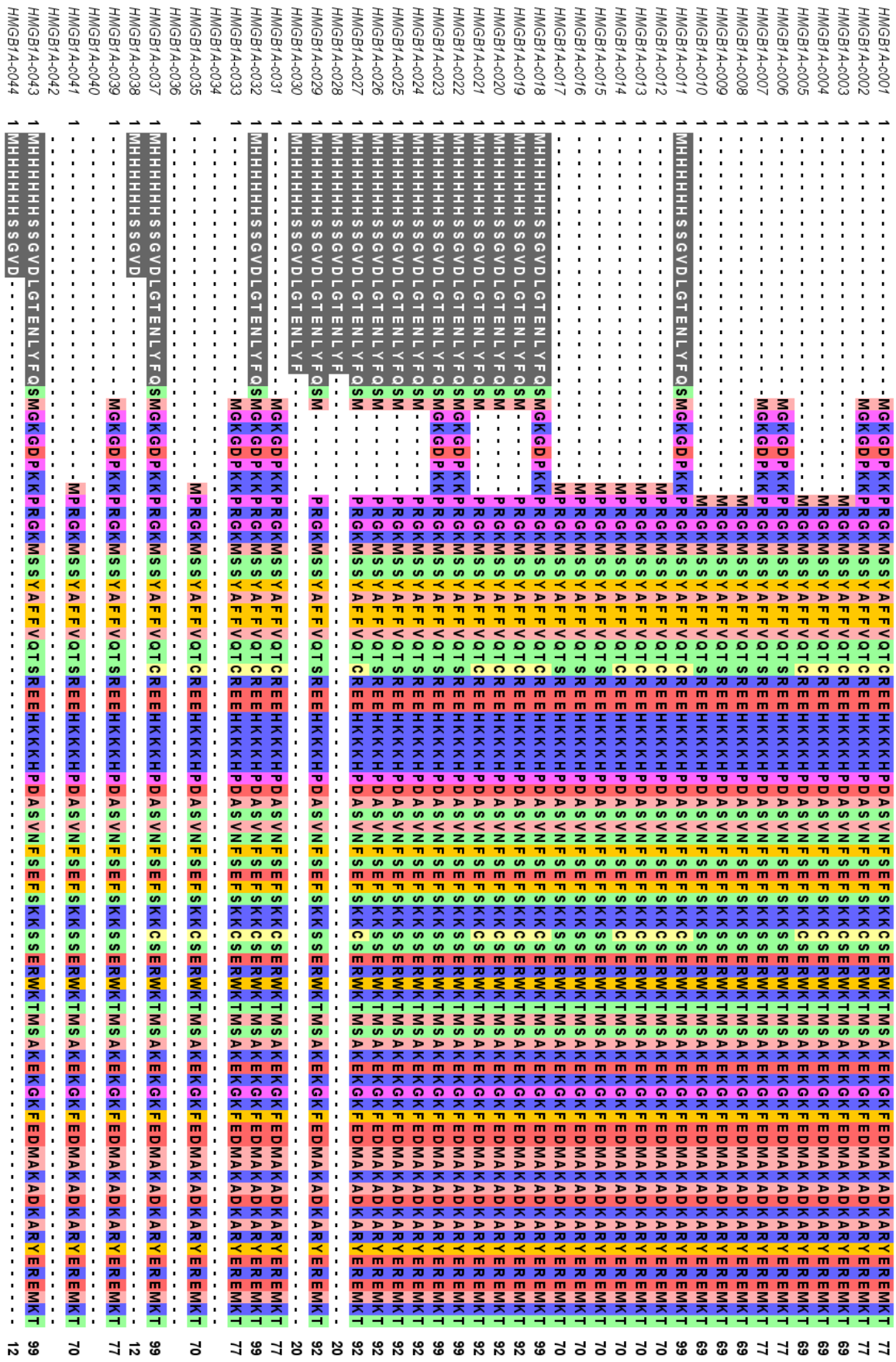
Construct ID	DNA sequence
HMGB1A-c025	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACA AAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTC AGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTT GCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAA AAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA GGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG GAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGAT GAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A-c026	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACA AAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTC AGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTT GCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAA AAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA
HMGB1A-c027	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATC
HMGB1A-c028	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCA AAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAG ATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTG AAGGAAAAATACGAAAAGGATATTGCTGCATATCGA
HMGB1A-c029	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAG GAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGC TCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATC
HMGB1A-c030	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGCCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCTGAGTATCGCCCA AAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAG ATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTG AAGGAAAAATACGAAAAGGATATTGCTGCATATCGA
HMGB1A-c031	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCG GCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGAC AAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCT GCATATCGAGCTAAA

Construct ID	DNA sequence
HMGB1A-c032	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTT TTTGTGCAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTC TCAGAGTTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTAT ATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCT CCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCT GGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCA GATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGAT ATTGCTGCATATCGAGCTAAA
HMGB1A-c033	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTC
HMGB1A-c034	ATGAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCT GATGCAGCAAAAAAGGGAGTT
HMGB1A-c035	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATC
HMGB1A-c036	ATGCCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATC AAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGG AATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAA AAATACGAAAAGGATATTGCTGCATATCGA
HMGB1A-c037	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTT TTTGTGCAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTC TCAGAGTTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTAT ATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTC
HMGB1A-c038	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGC TCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCG AAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAG AAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGA AAGCCTGATGCAGCAAAAAAGGGAGTT
HMGB1A-c039	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTG CAAACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAG TTTTCTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTT GAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCT CCCAAAGGGGAGACAAAAAAGAAGTTC
HMGB1A-c040	ATGAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCTGAG TATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCT GCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCT GATGCAGCAAAAAAGGGAGTT

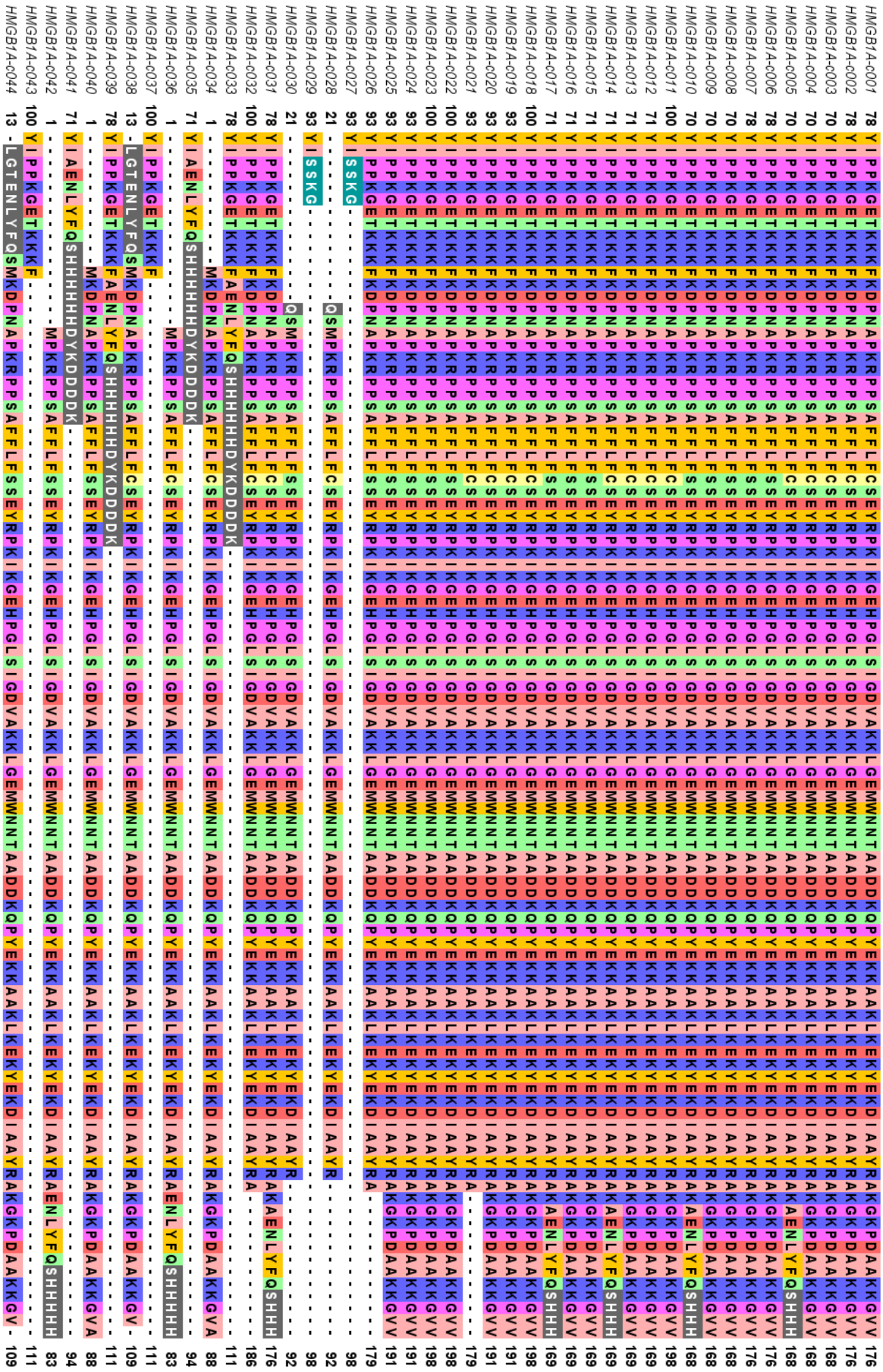
Construct ID	DNA sequence
HMGB1A-c041	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTAGTCGGGAGGAGCAT AAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGAGCTCAGAG AGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATC
HMGB1A-c042	ATGCCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGCTCTGAGTATCGCCCAAAAATC AAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGG AATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAA AAATACGAAAAGGATATTGCTGCATATCGA
HMGB1A-c043	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTT TTTGTGCAAACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTC TCAGAGTTTTCTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTAT ATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTC
HMGB1A-c044	ATGCACCATCATCATCATCATTCTTCTGGTGTAGATCTGGGTACCGAGAACCTGTACTTC CAATCCATGAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCAGC TCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCG AAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAG AAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGA AAGCCTGATGCAGCAAAAAAGGGAGTT

Suppl. data A1-3-D Protein sequences of HMGB1 constructs (full protein)

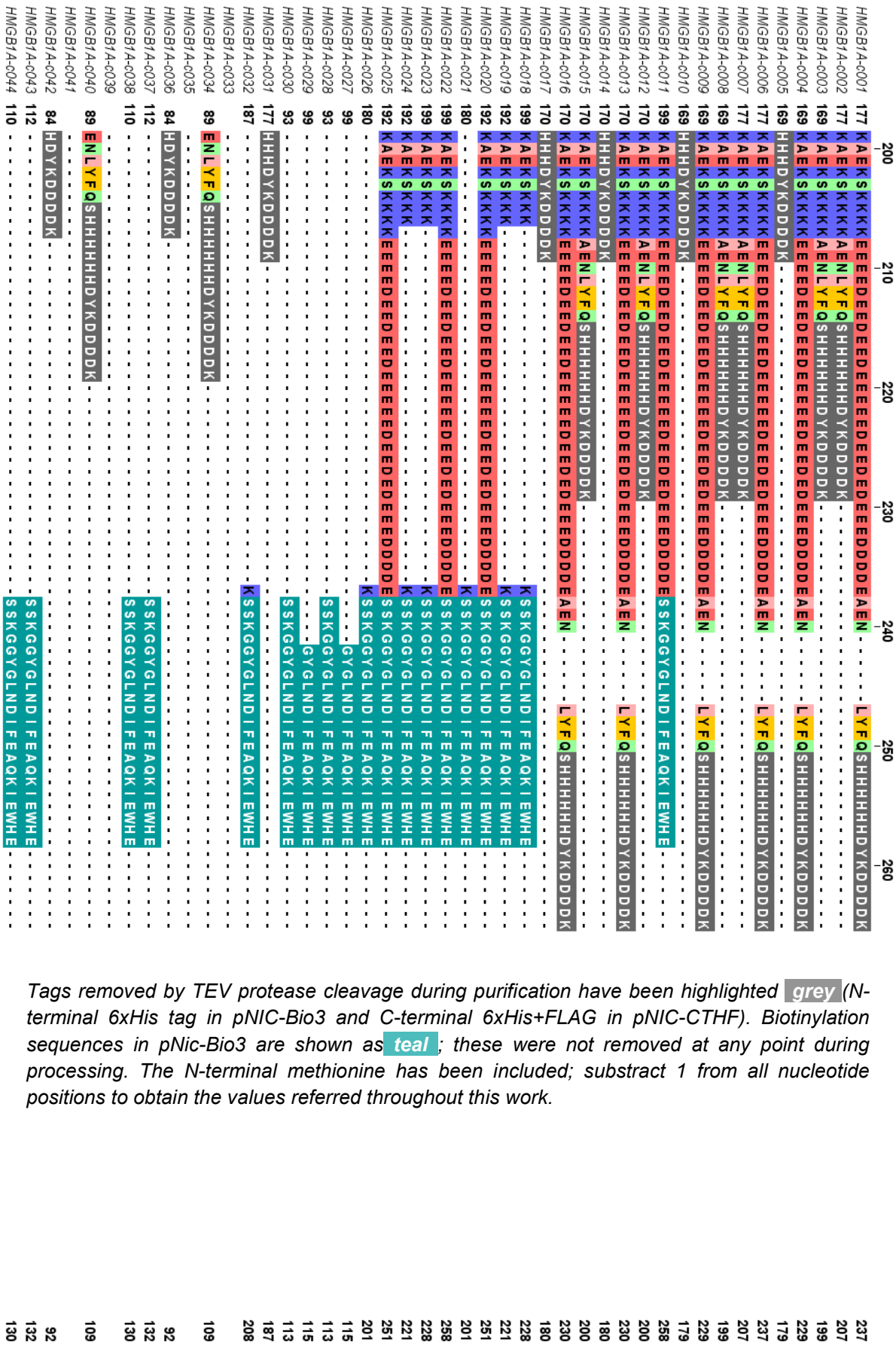
Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**



Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**



Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**



Tags removed by TEV protease cleavage during purification have been highlighted grey (N-terminal 6xHis tag in pNIC-Bio3 and C-terminal 6xHis+FLAG in pNIC-CTHF). Biotinylation sequences in pNic-Bio3 are shown as teal; these were not removed at any point during processing. The N-terminal methionine has been included; substract 1 from all nucleotide positions to obtain the values referred throughout this work.

Suppl. data A1-3-E Designed CXCL12 constructs

Construct ID	mW (Da)	mW (Da), no tag	ε $(M*cm)^{-1}$	Vector	Start	End	Notes
CXCL12A-c001	*12874*	*10969.1*	*11460*	*pNIC-CTHF*	*Met(-22)*	*Lys 68*	*Wt, initial panel*
CXCL12A-c002	*9348.73*	*7443.84*	*4470*	*pNIC-CTHF*	*Lys 1*	*Lys 56*	*Wt, initial panel*
CXCL12A-c003	10865.5	8960.6	11460	pNIC-CTHF	Lys 1	Lys 68	Wt, initial panel; for CyDisCO
CXCL12A-c004	*12888*	*10983.1*	*11460*	*pNIC-CTHF*	*Met(-22)*	*Lys 68*	*V3I, initial panel*
CXCL12A-c005	*9362.76*	*7457.87*	*4470*	*pNIC-CTHF*	*Lys 1*	*Lys 56*	*V3I, initial panel*
CXCL12A-c006	10879.5	8974.63	11460	pNIC-CTHF	Lys 1	Lys 68	V3I, initial panel; for CyDisCO
CXCL12A-c021	47465.5	7963.43	8740	pDsbC-HT-CBio	Lys 1	Lys 68	Wt, final expression construct
CXCL12A-c022	*21630.7*	*7963.43*	*8740*	*pNIC-OmpA-NHT-cTStII*	*Lys 1*	*Lys 68*	*Wt, periplasmic export*
CXCL12A-c023	21630.7	7963.43	8740	pNIC-SufI-NHT-cTStII	Lys 1	Lys 68	Wt, periplasmic export
CXCL12A-c024	*21630.7*	*7963.43*	*8740*	*pNIC-TolB-NHT-cTStII*	*Lys 1*	*Lys 68*	*Wt, periplasmic export*
CXCL12A-c025	47465.5	7977.46	8740	pDsbC-HT-CBio	Lys 1	Lys 68	V3I, final expression construct
CXCL12A-c026	*21644.7*	*7977.46*	*8740*	*pNIC-OmpA-NHT-cTStII*	*Lys 1*	*Lys 68*	*V3I, periplasmic export*
CXCL12A-c027	21644.7	7977.46	8740	pNIC-SufI-NHT-cTStII	Lys 1	Lys 68	V3I, periplasmic export
CXCL12A-c028	*21644.7*	*7977.46*	*8740*	*pNIC-TolB-NHT-cTStII*	*Lys 1*	*Lys 68*	*V3I, periplasmic export*

Table 8.4: Designed CXCL12 constructs for characterisation of the HMGB1-CXCL12 interaction.

Constructs in italics did not express and were not used beyond test expression.

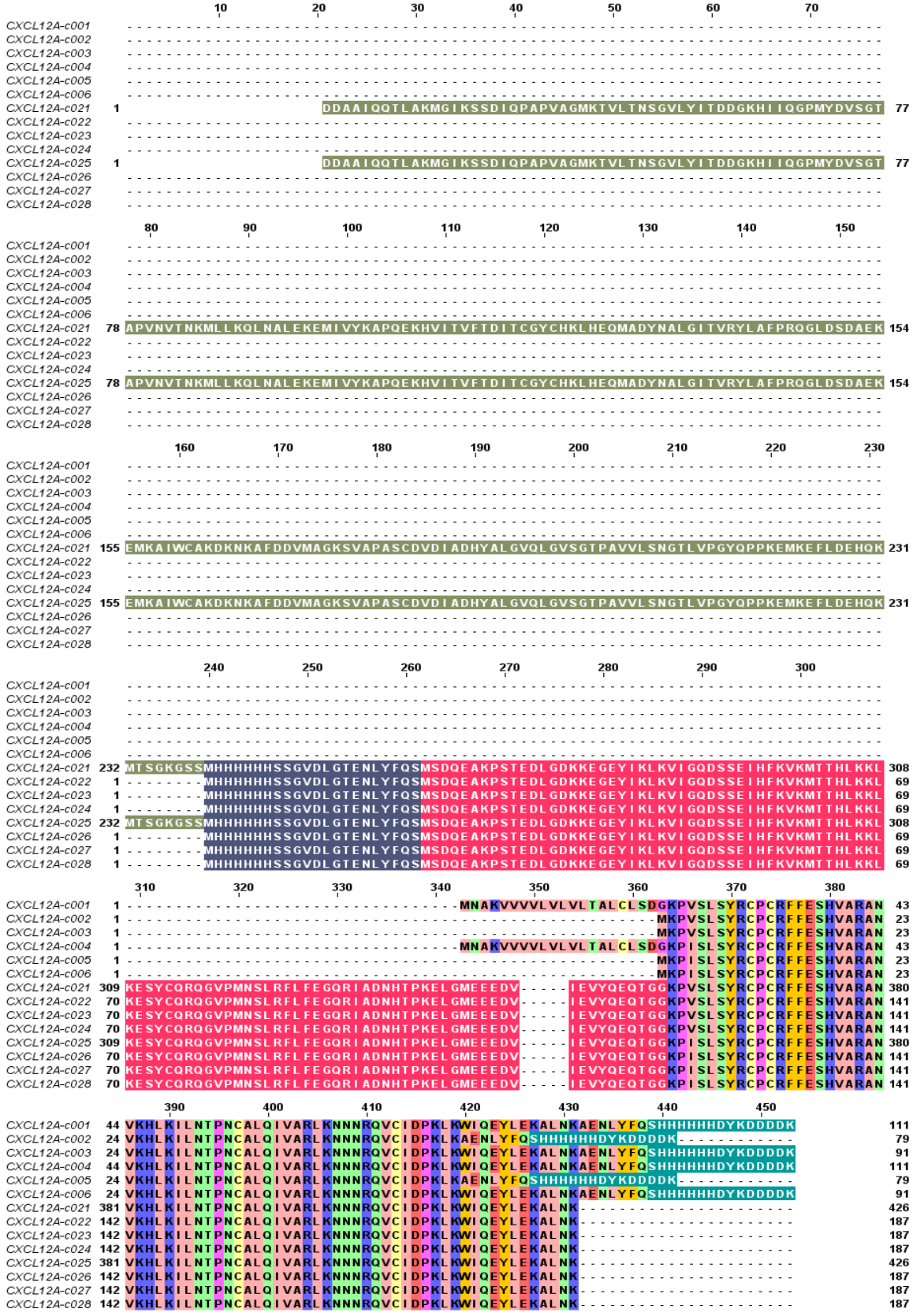
Suppl. data A1-3-F DNA sequences of CXCL12 constructs

Construct ID	DNA sequence
CXCL12A-c001 (signal peptide underlined)	ATGAACGCCAAGGTCGTGGTCGTGCTGGTCCTCGTGCTGACCGCGCTCTGCCTCAGCGAC GGGAAGCCCGTCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAG GAGTACCTGGAGAAAGCTTTAAACAAG
CXCL12A-c002	ATGAAGCCCGTCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAG
CXCL12A-c003	ATGAAGCCCGTCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAG GAGTACCTGGAGAAAGCTTTAAACAAG
CXCL12A-c004 (signal peptide underlined)	ATGAACGCCAAGGTCGTGGTCGTGCTGGTCCTCGTGCTGACCGCGCTCTGCCTCAGCGAC GGGAAGCCCATCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAG GAGTACCTGGAGAAAGCTTTAAACAAG
CXCL12A-c005	ATGAAGCCCATCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAG
CXCL12A-c006	ATGAAGCCCATCAGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCC AGAGCCAACGTCAAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTA GCCCGGCTGAAGAACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAG GAGTACCTGGAGAAAGCTTTAAACAAG
CXCL12A-c021 CXCL12A-c022 CXCL12A-c023 CXCL12A-c024 (SUMO fusion underlined)	ATGTCTGACCAGGAGGCAAAACCTTCAACTGAGGACTTGGGGGATAAGAAGGAAGGTGAA TATATTAAACTCAAAGTCATTGGACAGGATAGCAGTGAGATTCACTTCAAAGTGAAAATG ACAACACATCTCAAGAAACTCAAAGAATCATACTGTCAAAGACAGGGTGTTCCAATGAAT TCACTCAGGTTTCTCTTTGAGGGTCAGAGAATTGCTGATAATCATACTCCAAAAGAACTG GGAATGGAGGAAGAAGATGTGATTGAAGTTTACCAGGAGCAAACGGGAGGTAAGCCCGTC AGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCCAGAGCCAACGTC AAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTAGCCCGGCTGAAG AACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAGGAGTACCTGGAG AAAGCTTTAAACAAG
CXCL12A-c025 CXCL12A-c026 CXCL12A-c027 CXCL12A-c028 (SUMO fusion underlined)	ATGTCTGACCAGGAGGCAAAACCTTCAACTGAGGACTTGGGGGATAAGAAGGAAGGTGAA TATATTAAACTCAAAGTCATTGGACAGGATAGCAGTGAGATTCACTTCAAAGTGAAAATG ACAACACATCTCAAGAAACTCAAAGAATCATACTGTCAAAGACAGGGTGTTCCAATGAAT TCACTCAGGTTTCTCTTTGAGGGTCAGAGAATTGCTGATAATCATACTCCAAAAGAACTG GGAATGGAGGAAGAAGATGTGATTGAAGTTTACCAGGAGCAAACGGGAGGTAAGCCCATC AGCCTGAGCTACAGATGCCCATGCCGATTCTTCGAAAGCCATGTTGCCAGAGCCAACGTC AAGCATCTCAAAATTCTCAACACTCCAAACTGTGCCCTTCAGATTGTAGCCCGGCTGAAG AACAACAACAGACAAGTGTGCATTGACCCGAAGCTAAAGTGGATTCAGGAGTACCTGGAG AAAGCTTTAAACAAG

Suppl. data A1-3-G Protein sequences of CXCL12 constructs

Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**

DsbC tag coloured khaki; **6xHis tag and TEV site** in navy blue, **SUMO tag** in salmon, and **C-terminal 6xHIS-FLAG tag** (which was not removed) in teal.



Chapter 5 Sequences

Suppl. data A1-5-A HMGB1-C1 construct panel construct list

Construct family	ID	mW (tag)	ε $(M*cm)^{-1}$	mW (no tag)	Construct properties	Notes
Positive control constructs	HMGB1A_C1_c001	26669.3	22920	25759.4	FL-HMGB1,WT (=HMGB1A-c001)	
	HMGB1A_C1_c002	22936.2	22920	22026.3	1-184 HMGB1 (=HMGB1A-c002)	
	HMGB1A_C1_c003	20828.6	22920	19918.7	1-164 HMGB1 (=HMGB1A-c031)	
Caspase-activated RAGE binding domain mutations	HMGB1A_C1_c004	26625.3	22920	25715.4	Asp66Ala	*Flagged*
	HMGB1A_C1_c005	26683.4	22920	25773.5	Asp66Glu	Same charge, different length
	HMGB1A_C1_c006	26668.4	22920	25758.4	Asp66Asn	Same length, different charge
	HMGB1A_C1_c007	26683.4	22920	25773.5	Asp157Glu	Homologous Box A residue
	HMGB1A_C1_c008	26668.4	22920	25758.4	Asp157Asn	
	HMGB1A_C1_c009	26683.4	22920	25773.5	Asp168Glu	Homologous Box A residue
	HMGB1A_C1_c010	26668.4	22920	25758.4	Asp168Asn	
	HMGB1A_C1_c011	26711.4	22920	25801.5	Triple (66,157,168) Asp->Glu	
Serial linker truncations, C- to N-terminal direction	HMGB1A_C1_c012	25528	22920	24618.1	**PPKGETKKKF**K->K	Bold region corresponds to altered fragment. Peptide shown starts at position 79 of HMGB1 wt.
	HMGB1A_C1_c013	25803.3	22920	24893.4	**PPKGETKK**KFK->KFK	
	HMGB1A_C1_c014	26059.7	22920	25149.8	**PPKGET**KKKFK->KKKFK	
	HMGB1A_C1_c015	26289.9	22920	25380	**PPKG**ETKKKFK->ETKKKFK	
	HMGB1A_C1_c016	26475.1	22920	25565.2	**PP**KGETKKKFK->KGETKKKFK	
Serial linker truncations, N- to C-terminal direction	HMGB1A_C1_c017	26394	22920	25484.1	PPKGETKKKFK->PPKGETKKK	
	HMGB1A_C1_c018	26137.6	22920	25227.7	PPKGETK**KKFK**->PPKGETK	
	HMGB1A_C1_c019	25908.4	22920	24998.5	PPKGE**TKKKFK**->PPKGE	
	HMGB1A_C1_c020	25722.2	22920	24812.3	PPK**GETKKKFK**->PPK	
	HMGB1A_C1_c021	25496.9	22920	24587	P**PKGETKKKFK**->P	
	HMGB1A_C1_c022	25399.8	22920	24489.9	**PPKGETKKKFK**->-	
GSGSG linker substitution	HMGB1A_C1_c023	26506.1	22920	25596.2	**PPKGE**TKKKFK->**GSGSG**TKKKFK	
	HMGB1A_C1_c024	26381.9	22920	25471.9	PPKGE**TKKKF**K->PPKGE**GSGSG**K	
	HMGB1A_C1_c025	26489.1	22920	25579.2	PPKGETKKKFKDPNA->PPKGETKKKF**GSGSG**	
TLR-4 site mutations, controls	HMGB1A_C1_c026	26637.3	22920	25727.4	Cys105Ala	*Flagged*
	HMGB1A_C1_c027	26722.4	22920	25812.5	Cys105Arg	
	HMGB1A_C1_c028	26665.3	22920	25755.4	Cys105Val	
TLR-4 site mutations, Phe101/102	HMGB1A_C1_c029	26634.2	22920	25724.2	PPKGE**TKKK**FK-PPKGE**FSSE**FK	Decoy FSSE peptide
	HMGB1A_C1_c030	27119.8	22920	26209.9	MGKGDPKK-MG**FSSE**KGDPKK	
	HMGB1A_C1_c031	26651.3	22920	25741.4	Phe102Glu	
	HMGB1A_C1_c032	26685.3	24410	25775.4	Phe102Tyr	Aromatic group substitution for another
	HMGB1A_C1_c033	26708.4	28420	25798.5	Phe102Trp	
	HMGB1A_C1_c034	26651.3	22920	25741.4	Phe101Glu	
	HMGB1A_C1_c035	26685.3	24410	25775.4	Phe101Tyr	Aromatic group substitution
	HMGB1A_C1_c036	26708.4	28420	25798.5	Phe101Trp	
	HMGB1A_C1_c037	26701.3	25900	25791.4	Phe101/102->Tyr	
	HMGB1A_C1_c038	26747.4	33920	25837.5	Phe101/102->Trp	
	HMGB1A_C1_c039	26724.4	29910	25814.5	Phe101Tyr; Phe102Trp	
	HMGB1A_C1_c040	26724.4	29910	25814.5	Phe101Trp; Phe102Tyr	
	HMGB1A_C1_c041	26698.3	28420	25788.4	Phe101His; Phe102Trp	
	HMGB1A_C1_c042	26675.3	24410	25765.4	Phe101His; Phe102Tyr	
	HMGB1A_C1_c043	26659.3	22920	25749.4	Phe102His	Aromatic group to imidazole ring
	HMGB1A_C1_c044	26659.3	22920	25749.4	Phe101His	
	HMGB1A_C1_c045	26649.3	22920	25739.4	Phe101/102->His	
	HMGB1A_C1_c046	26675.3	24410	25765.4	Phe101Tyr; Phe103His	
	HMGB1A_C1_c047	26698.3	28420	25788.4	Phe101Trp; Phe102His	

Construct family	ID	mW (tag)	ε $(M*cm)^{-1}$	mW (no tag)	Construct properties	Notes
Double Box B constructs	HMGB1A_C1_c048	21044.9	24410	20135	Double Box B construct (90-175--90-175)	
	HMGB1A_C1_c049	21534.3	24410	20624.4	Double Box B construct (90-175-GSGSGSG-90-175)	
Main RAGE-binding site mutations (150-184)	HMGB1A_C1_c050	26684.4	22920	25774.4	Leu150Lys	Different charge and length
	HMGB1A_C1_c051	26643.3	22920	25733.3	Leu150Ser	
	HMGB1A_C1_c052	26671.3	22920	25761.4	Leu150Asp	Same length, different charge
	HMGB1A_C1_c053	26670.3	22920	25760.4	Leu150Asn	
	HMGB1A_C1_c054	26684.4	22920	25774.4	Ile158Lys	Different charge and length
	HMGB1A_C1_c055	26643.3	22920	25733.3	Ile158Ser	
	HMGB1A_C1_c056	26740.5	22920	25830.5	Gly165Lys	Different charge or length
	HMGB1A_C1_c057	26683.4	22920	25773.5	Gly165Ala	
	HMGB1A_C1_c058	26709.4	22920	25799.5	Gly165Pro	
	HMGB1A_C1_c059	26656.3	22920	25746.3	Lys172Asp	Lys->Arg/Gln conserved in other RAGE binding motif proteins
	HMGB1A_C1_c060	26670.3	22920	25760.4	Lys172Glu	
	HMGB1A_C1_c061	26656.3	22920	25746.3	Lys176Asp	
	HMGB1A_C1_c062	26670.3	22920	25760.4	Lys176Glu	
Peptide array alanine scan negative control mutations	HMGB1A_C1_c063	26625.3	22920	25715.4	Asp4Ala	Avoided Proline mutations
	HMGB1A_C1_c064	26612.2	22920	25702.3	Lys7Ala	
	HMGB1A_C1_c065	26612.2	22920	25702.3	Lys8Ala	
	HMGB1A_C1_c066	26584.2	22920	25674.3	Arg10Ala	
	HMGB1A_C1_c067	26683.4	22920	25773.5	Gly11Ala	
	HMGB1A_C1_c068	26625.3	22920	25715.4	Asp90Ala	
	HMGB1A_C1_c069	26626.3	22920	25716.4	Asn92Ala	
	HMGB1A_C1_c070	26754.4	22920	25844.5	Ala93Arg	
	HMGB1A_C1_c071	26612.2	22920	25702.3	Lys95ala	
	HMGB1A_C1_c072	26584.2	22920	25674.3	Arg96Ala	
Glycyrrhizin binding site mutations (potential TLR-2 signalling alteration)	HMGB1A_C1_c073	26584.2	22920	25674.3	Arg24Ala	Guanidinium group critical for glycyrrhizin binding
	HMGB1A_C1_c074	26650.3	22920	25740.4	Arg25His	
	HMGB1A_C1_c075	26641.3	22920	25731.4	Arg24Lys	
	HMGB1A_C1_c076	26584.2	22920	25674.3	Arg109Ala	
	HMGB1A_C1_c077	26650.3	22920	25740.4	Arg109His	
	HMGB1A_C1_c078	26641.3	22920	25731.4	Arg109Lys	
	HMGB1A_C1_c079	26672.4	22920	25762.5	Lys42Met	Amine in Lys accepts carbonyl from glycyrrhizin
	HMGB1A_C1_c080	26669.3	22920	25759.4	Lys42Gln	
	HMGB1A_C1_c081	26672.4	22920	25762.5	Lys127Met	
	HMGB1A_C1_c082	26669.3	22920	25759.4	Lys127Gln	
Cysteine-3S HMGB1	HMGB1A_C1_c083	26596.1	22920	25686.2	Cys22,44,105->Ser, Lys86Cys	Thiol labelling
Peptide array alanine scan N-terminal serial truncation	HMGB1A_C1_c084	26612.3	22920	25702.4	**G**KGDPKK - KGDPKK (HMGB1 2-214)	Bold region corresponds to deleted fragment. Peptide represents residues 1-7 of wt. HMGB1
	HMGB1A_C1_c085	26484.1	22920	25574.2	**GK**GDPKK - GDPKK (HMGB1 3-214)	
	HMGB1A_C1_c086	26427.1	22920	25517.2	**GKG**DPKK - DPKK (HMGB1 4-214)	
	HMGB1A_C1_c087	26312	22920	25402.1	**GKGD**PKK - PKK (HMGB1 5-214)	
	HMGB1A_C1_c088	26214.9	22920	25304.9	**GKGDP**KK - KK (HMGB1 6-214)	
	HMGB1A_C1_c089	26086.7	22920	25176.8	**GKGDPK**K - K (HMGB1 7-214)	
	HMGB1A_C1_c090	25958.5	22920	25048.6	GKGDPKK - (HMGB1 8-214)	

Table 8.5: HMGB1-C1 construct panel

Summary of expected MW (in Da) of the designed constructs with and without TEV cleavage. The tag did not contain any residues absorbing at 280nm and therefore extinction coefficients are valid for both tag and tag-less protein. *Flagged* = mutations which have been described in the literature but the data in Chapter 4 suggest could affect CXCL12 binding. A short description of the rationale for each mutation has been provided. Sequences are collected in

Suppl. data A1-5-B DNA sequences for HMGB1-C1 construct panel

All these sequences were produced by DNA synthesis and thus were considered verified.

Construct ID	DNA sequence
HMGB1A_C1_c001	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c002	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAG
HMGB1A_C1_c003	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAA
HMGB1A_C1_c004	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGCGAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c005	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGAGAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c006	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGAACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c007	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGAGATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c008	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGAACATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c009	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGAAGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c010	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTAACGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c011	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGAGAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGAGATTGCTGCATATCGAGCTAAAG GAAAGCCTGAAGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c012	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCAAGGATCCCAA TGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCA AAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAAT AACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATA CGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTG TCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGAT GAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c013	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCAAGTTCAAGGA TCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAA AAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATG TGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGA AAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGG GAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAA GAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGA A
HMGB1A_C1_c014	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCAAAAAGAAGTT CAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATC GCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGA GAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCT GAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAA AAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAA GATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGA TGATGAA
HMGB1A_C1_c015	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCGAGACAAAAAA GAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTG AGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGC GAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATG CAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGAT GAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGA TGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c016	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCAAAGGGGAGAC AAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCT GCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCG AAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAA GGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGC CTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAG GAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGA AGAAGATGATGATGATGAA
HMGB1A_C1_c017	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCT GCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCG AAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAA GGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGC CTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAG GAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGA AGAAGATGATGATGATGAA
HMGB1A_C1_c018	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTG AGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGC GAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATG CAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGAT GAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGA TGATGATGATGAA
HMGB1A_C1_c019	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATC GCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGA GAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCT GAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAA AAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAA GATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGA TGATGAA
HMGB1A_C1_c020	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGA TCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAA AAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATG TGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGA AAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGG GAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAA GAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGA A

Construct ID	DNA sequence
HMGB1A_C1_c021	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTGATCCCAA TGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCA AAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAAT AACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATA CGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTG TCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGAT GAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c022	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCGATCCCAATGC ACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAG GAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAAC ACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGA AAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCA AGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAG GAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c023	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCGGCTCTGGATC AGGTACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c024	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGGGATCTGGTTCAGGCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c025	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCGGATCCGGAAGTGGTCCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c026	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCGCCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c027	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCAGGTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c028	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCGTGTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c029	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGTTCAGCAGCGAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c030	ATGGGCTTCAGCAGCGAGAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGC ATTTTTTGTGCAAACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACT TCTCAGAGTTTTCTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGA AAATTTGAAGATATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATAT CCCTCCCAAAGGGGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTT CGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTG TCCATTGGTGATGTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAA GCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCAT ATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAG AAAAAGAAGGAAGAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGA AGAAGATGAAGATGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c031	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCGAGC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c032	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c033	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTGGC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c034	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCGAGTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c035	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTACTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c036	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTGGTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c037	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTACTACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c038	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTGGTGGC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c039	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTACTGGC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c040	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTGGTACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c041	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCCACTGGC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c042	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCCACTACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c043	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCCACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c044	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCCACTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c045	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCCACCACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c046	*ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTACCACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA*
HMGB1A_C1_c047	*ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTGGCACC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA*
HMGB1A_C1_c048 (Optimised sequence)	ATGAAGGACCCGAACGCGCCGAAGCGCCCACCGTCCGCTTTTTTTCTGTTCTGTAGCGAATA CCGTCCAAAAATTAAAGGTGAACACCCAGGCCTTTCTATCGGCGATGTTGCGAAAAAACTGG GTGAGATGTGGAACAATACTGCAGCAGATGATAAACAACCGTATGAAAAGAAAGCGGCGAAG CTGAAAGAAAAATACGAAAAAGATATCGCTGCGTATCGGGCCAAGGGGAAACCGGATGCGGC GAAAAAAGGCGTCAAGGATCCCAACGCGCCGAAACGTCCACCGAGCGCGTTCTTTCTGTTTT GTTCAGAATACCGCCCGAAAATTAAAGGCGAACACCCGGGCCTTTCGATTGGTGACGTTGCG AAAAAACTTGGCGAAATGTGGAATAACACGGCAGCGGACGATAAGCAGCCTTATGAAAAAAA AGCCGCCAAACTTAAAGAAAAATACGAGAAAGACATTGCAGCTTACCGCGCGAAGGGTAAAC CTGACGCTGCAAAAAAAGGTGTTTAA
HMGB1A_C1_c049 (Optimised sequence)	ATGAAGGACCCGAATGCACCAAAACGCCCGCCATCGGCTTTTTTTTTATTCTGTAGCGAATA TCGGCCGAAGATTAAGGGTGAACACCCTGGCTTATCTATCGGCGACGTCGCTAAGAAACTGG GCGAAATGTGGAACAACACAGCCGCGGATGATAAACAACCGTATGAAAAAAAAGCGGCGAAA CTTAAAGAAAAATATGAAAAGGACATCGCCGCATATCGCGCGAAGGGTAAACCGGACGCCGC GAAAAAAGGTGTGGGCTCGGGGTCCGGTTCAGGCAAAGACCCGAACGCGCCAAAGCGGCCGC CCTCTGCATTTTTCCTTTTTTGTTCCGAATACCGGCCAAAAATCAAAGGTGAACATCCGGGC TTATCGATTGGTGACGTCGCGAAAAAGTTGGGCGAAATGTGGAATAATACTGCTGCAGATGA TAAACAGCCGTATGAAAAAAAAGCCGCGAAATTAAAGGAGAAATATGAAAAAGACATCGCAG CCTATCGCGCAAAGGGCAAACCTGACGCGGCGAAAAAAGGTGTG
HMGB1A_C1_c050	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGAAGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c051	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGTCGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c052	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGGACAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c053	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGAACAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c054	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATAAGGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c055	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATAGTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c056	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAA AGAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c057	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG CAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c058	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAC CAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c059	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAGACGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c060	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAGAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c061	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCGACGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c062	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCGAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c063	ATGGGCAAAGGAGCTCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c064	ATGGGCAAAGGAGATCCTGCCAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c065	ATGGGCAAAGGAGATCCTAAGGCCCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c066	ATGGGCAAAGGAGATCCTAAGAAGCCGGCAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c067	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGCCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c068	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGCTCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c069	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCGCTGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c070	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATAGACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c071	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCGCGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c072	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGGCGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c073	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTGCGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c074	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCACGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c075	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTAAGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c076	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATGCGCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c077	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCACCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c078	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATAAGCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c079	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTATGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c080	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTCAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c081	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGATGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c082	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGCAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c083	ATGGGCAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCA AACTAGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTT CTAAGAAGAGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGAT ATGGCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGG GGAGACAAAATGTAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCC TCTTCAGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGAT GTTGCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGA AAAGAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAG GAAAGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAA GAGGAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGA TGAAGAAGAAGATGATGATGATGAA
HMGB1A_C1_c084	ATGAAAGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAAC TTGTCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTA AGAAGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATG GCAAAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGA GACAAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCT TCTGCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTT GCGAAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAA GAAGGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAA AGCCTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAG GAGGAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGA AGAAGAAGATGATGATGATGAA
HMGB1A_C1_c085	ATGGGAGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTG TCGGGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGA AGTGCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCA AAAGCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGAC AAAAAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCT GCTCTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCG AAGAAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAA GGCTGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGC CTGATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAG GAAGATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGA AGAAGATGATGATGATGAA

Construct ID	DNA sequence
HMGB1A_C1_c086	ATGGATCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCG GGAGGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGT GCTCAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAA GCGGACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAA AAAGAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCT CTGAGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAG AAACTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGC TGCGAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTG ATGCAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAA GATGAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGA AGATGATGATGATGAA
HMGB1A_C1_c087	ATGCCTAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGA GGAGCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCT CAGAGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCG GACAAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAA GAAGTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTG AGTATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAA CTGGGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGC GAAGCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATG CAGCAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGAT GAGGAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGA TGATGATGATGAA
HMGB1A_C1_c088	ATGAAGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGA GCATAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAG AGAGGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGAC AAGGCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAA GTTCAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGT ATCGCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTG GGAGAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAA GCTGAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAG CAAAAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAG GAAGATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGA TGATGATGAA
HMGB1A_C1_c089	ATGAAGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCA TAAGAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAGA GGTGGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAG GCCCGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAGTT CAAGGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATC GCCCAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGA GAGATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCT GAAGGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAA AAAAGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAA GATGAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGA TGATGAA
HMGB1A_C1_c090	ATGCCGAGAGGCAAAATGTCATCATATGCATTTTTTGTGCAAACTTGTCGGGAGGAGCATAA GAAGAAGCACCCAGATGCTTCAGTCAACTTCTCAGAGTTTTCTAAGAAGTGCTCAGAGAGGT GGAAGACCATGTCTGCTAAAGAGAAAGGAAAATTTGAAGATATGGCAAAAGCGGACAAGGCC CGTTATGAAAGAGAAATGAAAACCTATATCCCTCCCAAAGGGGAGACAAAAAAGAAGTTCAA GGATCCCAATGCACCCAAGAGGCCTCCTTCGGCCTTCTTCCTCTTCTGCTCTGAGTATCGCC CAAAAATCAAAGGAGAACATCCTGGCCTGTCCATTGGTGATGTTGCGAAGAAACTGGGAGAG ATGTGGAATAACACTGCTGCAGATGACAAGCAGCCTTATGAAAAGAAGGCTGCGAAGCTGAA GGAAAAATACGAAAAGGATATTGCTGCATATCGAGCTAAAGGAAAGCCTGATGCAGCAAAAA AGGGAGTTGTCAAGGCTGAAAAAAGCAAGAAAAAGAAGGAAGAGGAGGAAGATGAGGAAGAT GAAGAGGATGAGGAGGAGGAGGAAGATGAAGAAGATGAAGATGAAGAAGAAGATGATGATGA TGAA

Suppl. data A1-5-C Protein sequences of HMGB1-C1 constructs (tagged)

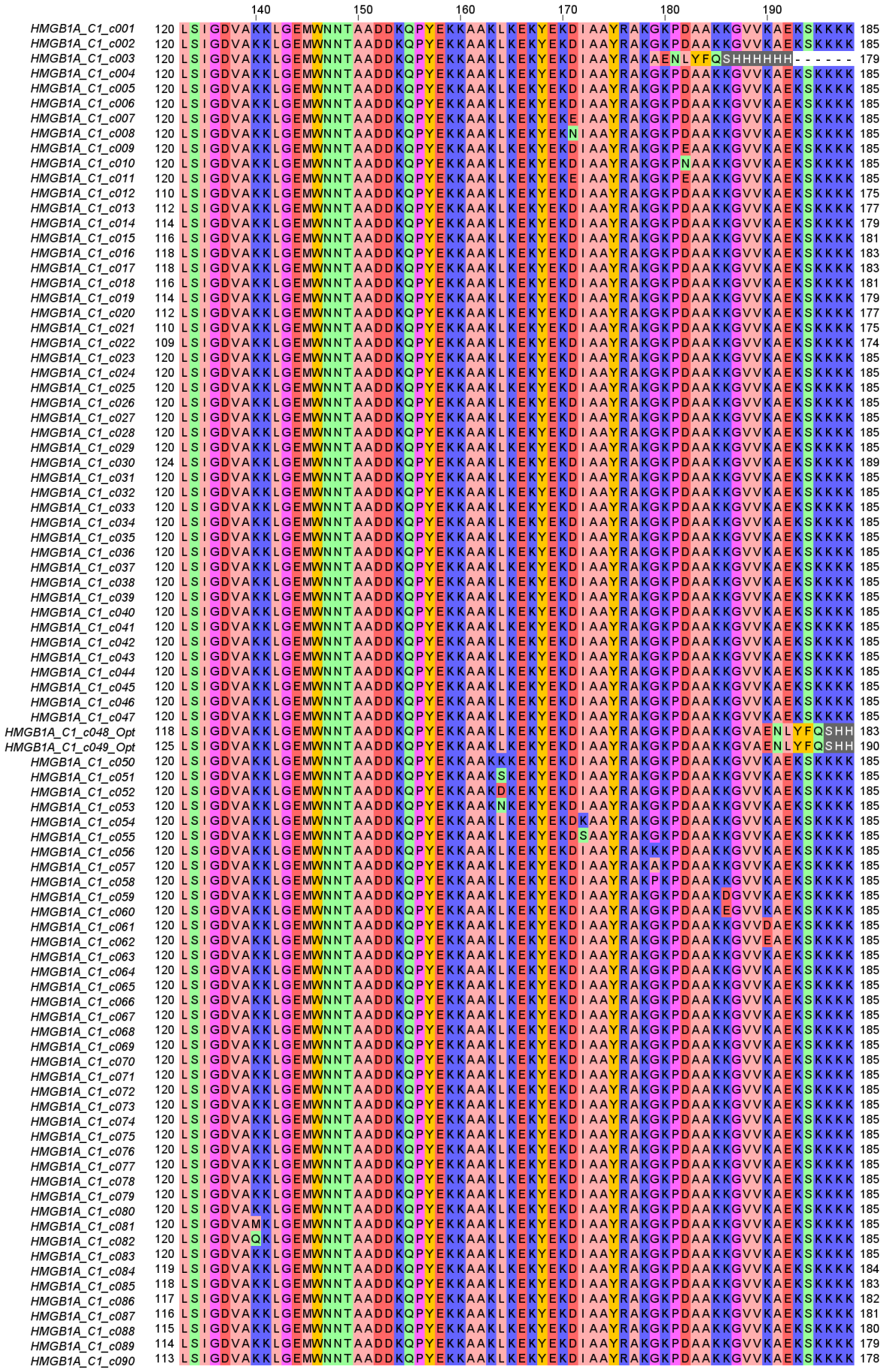
Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**



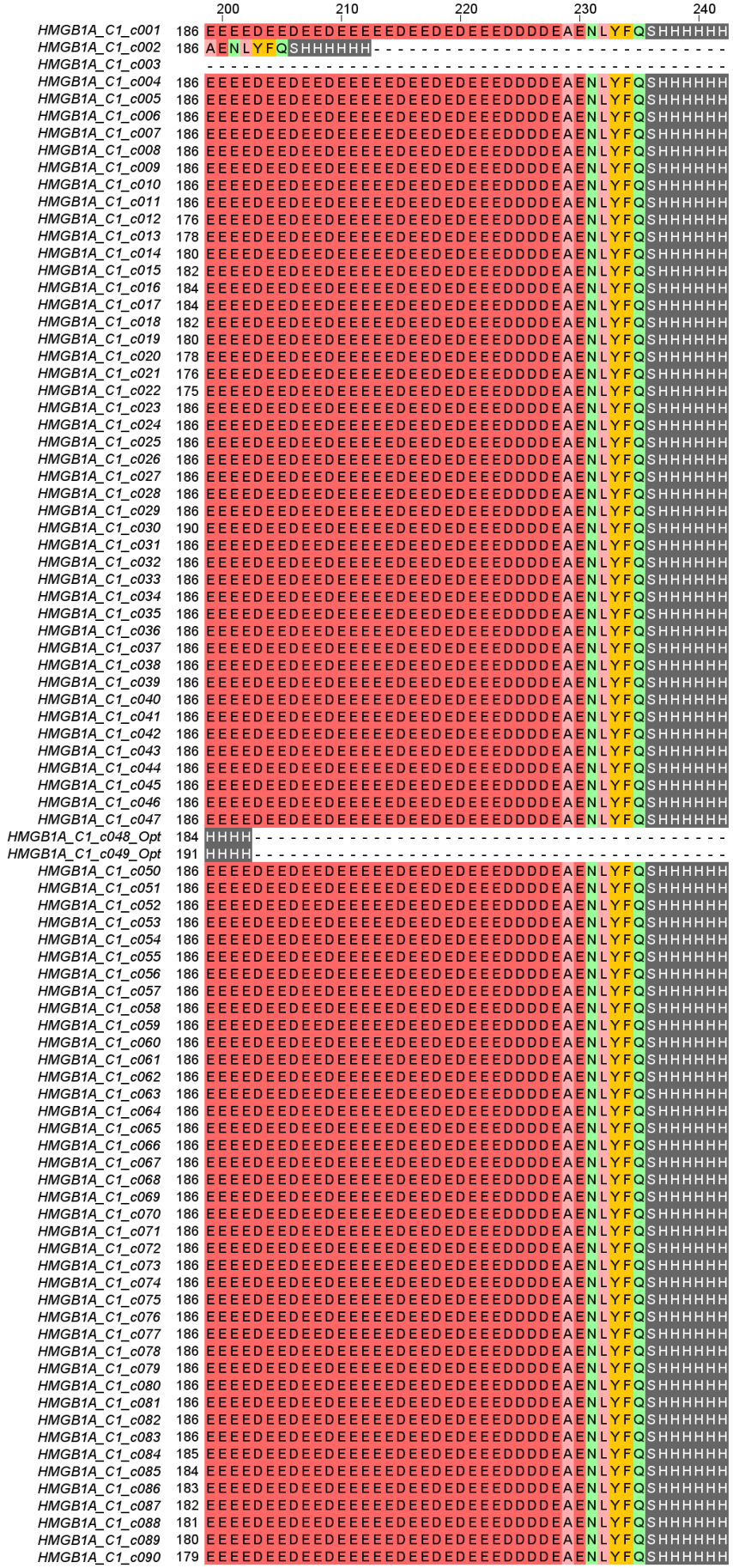
Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**



Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**



Colour key: **Aliphatic**, **Gly/Pro**, **Aromatic**, **Acidic**, **Basic**, **Polar**, **Cysteine**

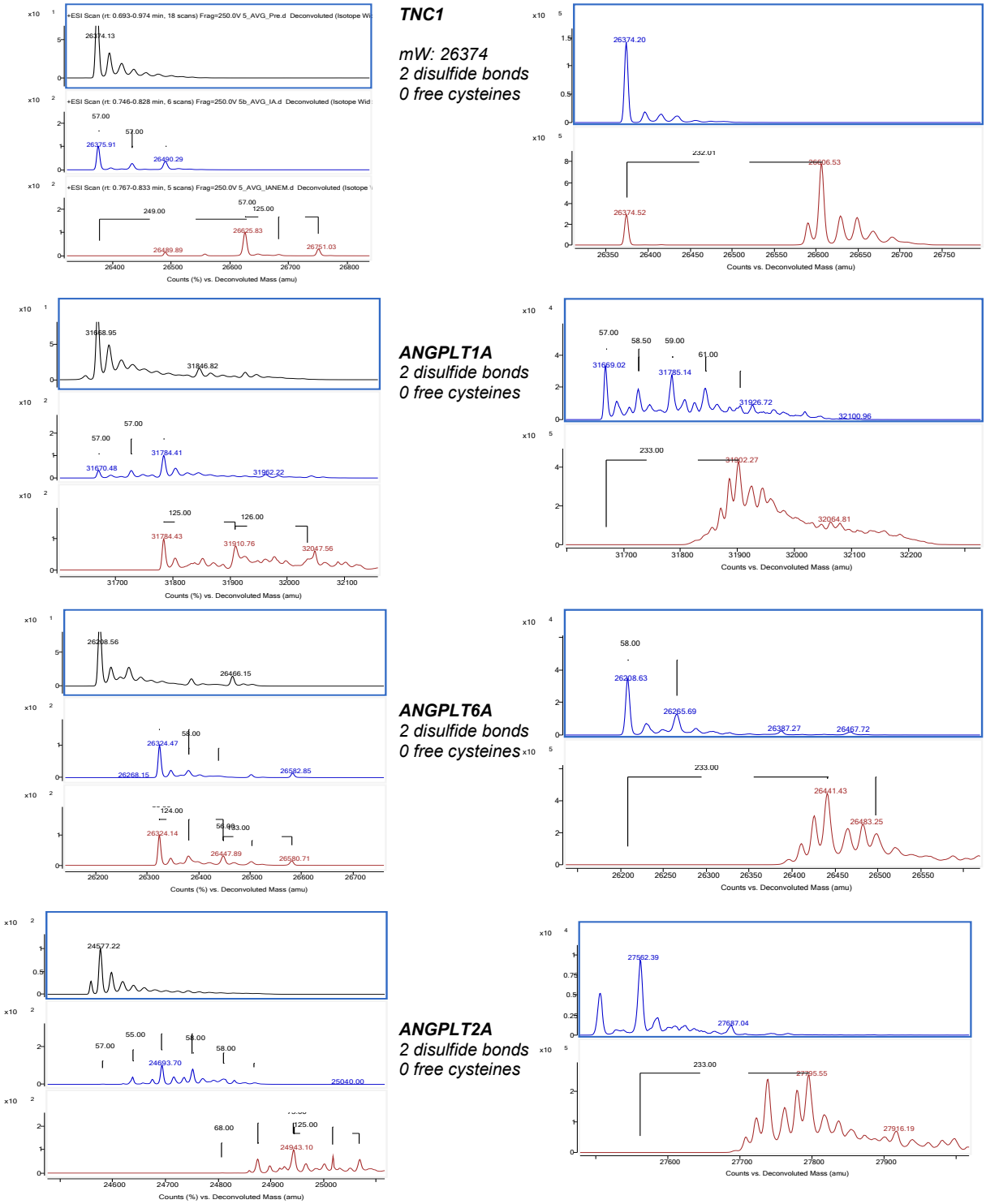


Tags removed by TEV protease cleavage during purification have been highlighted grey (C-terminal 6xHis in pNIC-CTH0). The N-terminal methionine has been included; substract 1 from all nucleotide positions to obtain the values referred throughout this work.

Appendix A2 Mass spectrometry data

Chapter 3 MS data

Suppl. data A2-3-A : Figure 3.3



Black: original samples. Left hand side: published method [24]. Blue: iodoacetamide only, Red: samples treated with IA after DTT dialysis and N-ethyl maleimide treatment. Right hand side: urea-denatured samples. Blue: Iodoacetamide only, Red: Iodoacetamide + TCEP. These protein samples were a gift from DPhil student Jesse A. Coker.

Suppl. data A2-3-B Figure 3.7: DNA contamination of HMGB1 after size exclusion chromatography (SEC)., pp. 81, tryptic digest MS/MS verification

Construct		MASCOT score	Match	Output
HMGB1A-c006, band 1	**SGC constructs**	481	HMGB1A-c006	Contaminating proteins predominate, target is present ☺
	Swissprot sequences (top 5)	571	50S ribosomal protein L2 OS=Escherichia coli O139:H28 (strai	
		550	Ribose-phosphate pyrophosphokinase OS=Escherichia coli (stra	
		407	50S ribosomal protein L2 OS=Photorhabdus luminescens subsp.	
		331	High mobility group protein B1 OS=Bos taurus GN=HMGB1 PE=1 S	
		323	High mobility group protein B1 OS=Gallus gallus GN=HMGB1 PE=	
HMGB1A-c006 Band 2	**SGC constructs**	277	HMGB1A-c003	Contaminating proteins predominate, target is present ☹
	Swissprot sequences (top 5)	442	30S ribosomal protein S4 OS=Citrobacter koseri (strain ATCC	
		435	30S ribosomal protein S4 OS=Salmonella agona (strain SL483)	
		354	30S ribosomal protein S3 OS=Serratia proteamaculans (strain	
		352	30S ribosomal protein S3 OS=Citrobacter koseri (strain ATCC	
		273	High mobility group protein B1 OS=Bos taurus GN=HMGB1 PE=1 S	
HMGB1A-c006 Band 3	**SGC constructs**	142	HMGB1A-c001	Contaminating proteins predominate, target is present ☹
	Swissprot sequences (top 5)	565	30S ribosomal protein S7 OS=Escherichia coli O139:H28 (strai	
		316	Ferric uptake regulation protein OS=Escherichia coli (strain	
		301	30S ribosomal protein S5 OS=Citrobacter koseri (strain ATCC	
		229	30S ribosomal protein S7 OS=Tolumonas auensis (strain DSM 91	
		216	50S ribosomal protein L13 OS=Citrobacter koseri (strain ATCC	

Band 1 MS/MS details (59% coverage):

```
  1  MGKGDPKKPR  GKMSSYAFFV  QTSREEHKKK  HPDASVNFSE  FSKKSSERWK
 51  TMSAKEKGKF  EDMAKADKAR  YEREMKTYIP  PKGETKKKFK  DPNAPKRPPS
101  AFFLFSSEYR  PKIKGEHPGL  SIGDVAKKLG  EMWNNTAADD  KQPYEKKAAK
151  LKEKYEKDIA  AYRAKGKPDA  AKKGVVKAEK  SKKKKEEEED  EEDEEDEEEE
201  EDEEDEDEEE  DDDDEAENLY  FQSHHHHHHD  YKDDDDK
```

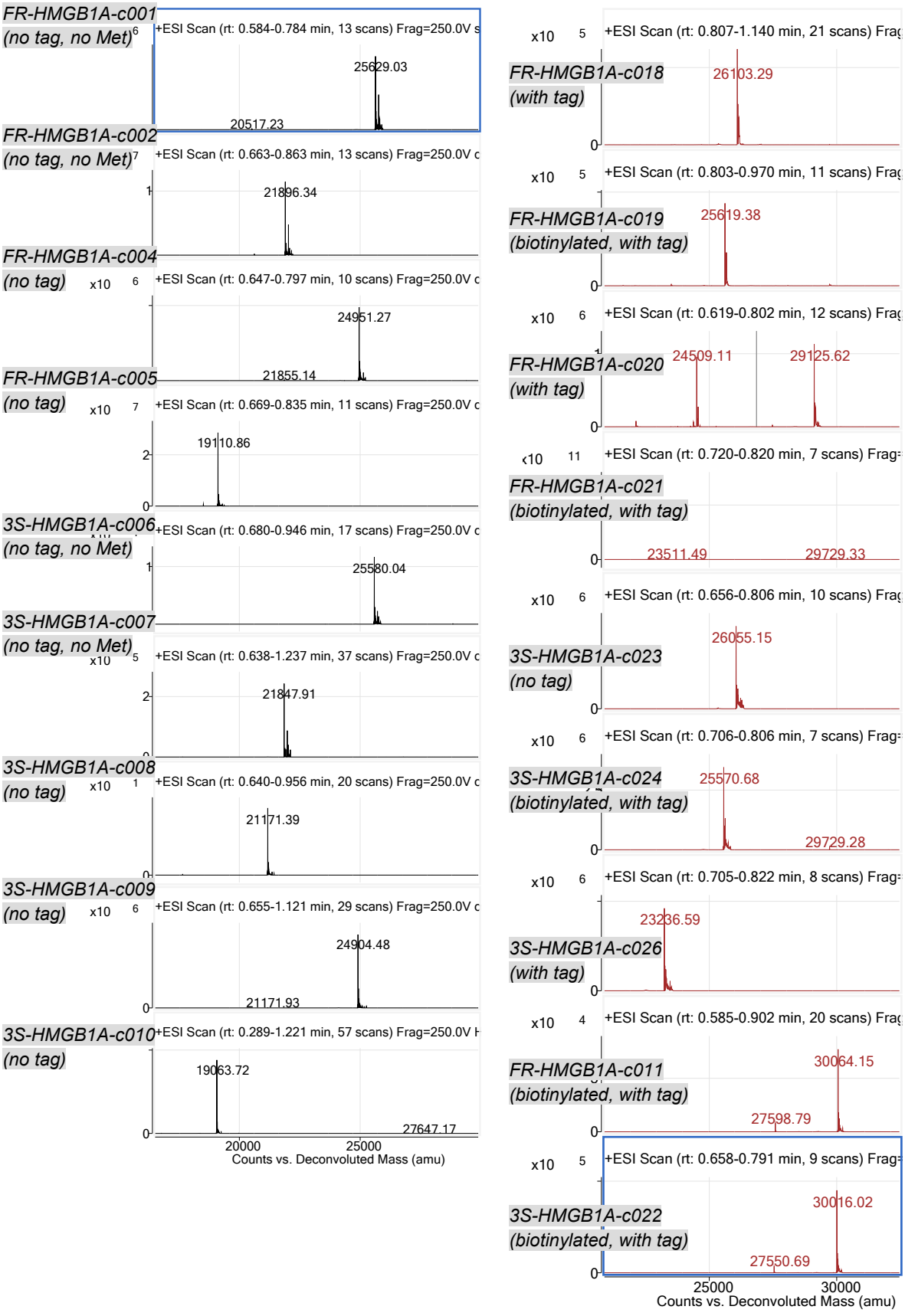
Band 2 MS/MS coverage (37% coverage)

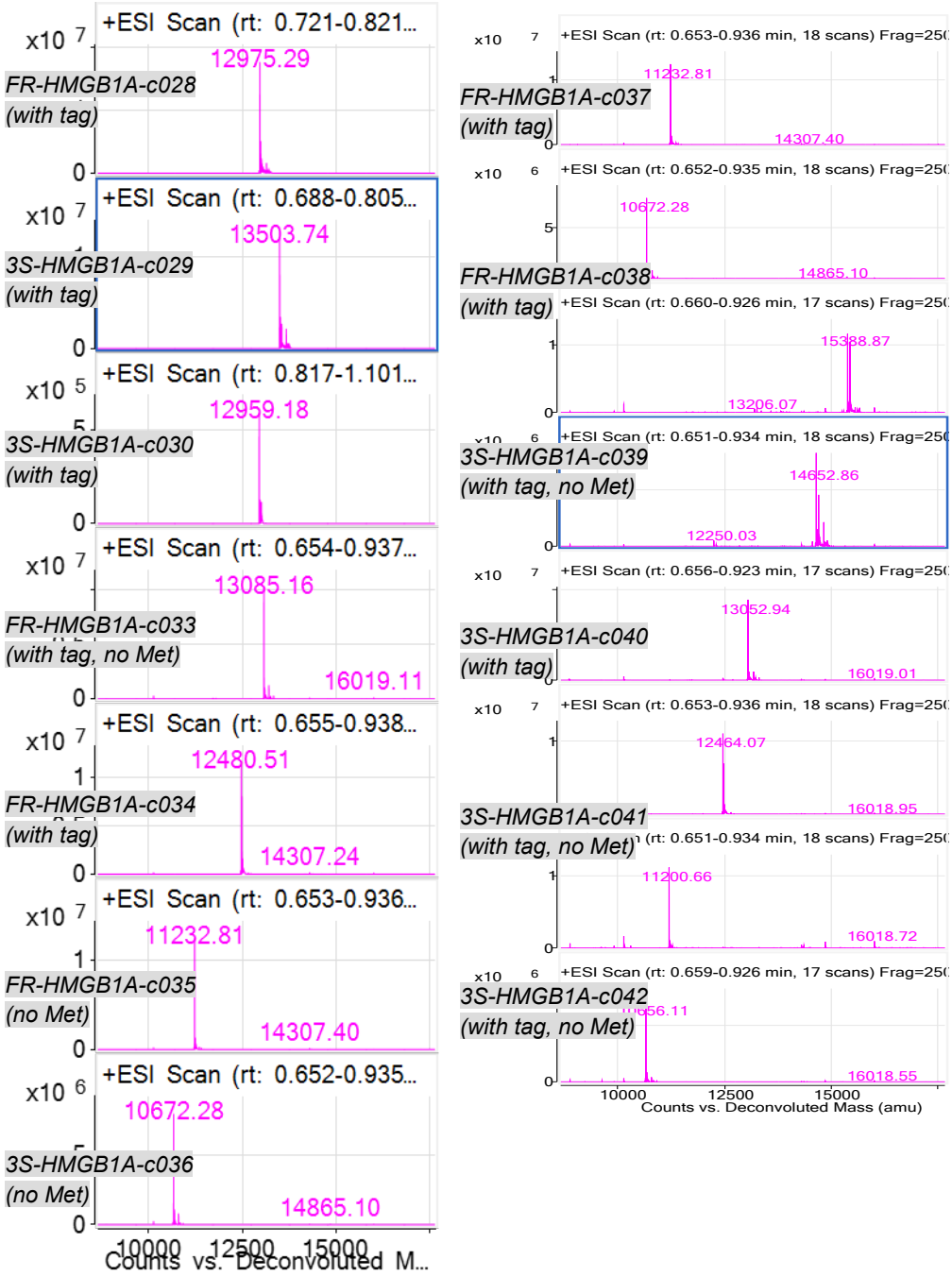
```
  1  MGKGDPKKPR  GKMSSYAFFV  QTSREEHKKK  HPDASVNFSE  FSKKSSERWK
 51  TMSAKEKGKF  EDMAKADKAR  YEREMKTYIP  PKGETKKKFK  DPNAPKRPPS
101  AFFLFSSEYR  PKIKGEHPGL  SIGDVAKKLG  EMWNNTAADD  KQPYEKKAAK
151  LKEKYEKDIA  AYRAKGKPDA  AKKGVVKAEK  SKKKKEEEED  EEDEEDEEEE
201  EDEEDEDEEE  DDDDEAENLY  FQSHHHHHHD  YKDDDDK
```

Band 3 MS/MS coverage (27% coverage)

```
  1  MGKGDPKKPR  GKMSSYAFFV  QTSREEHKKK  HPDASVNFSE  FSKKSSERWK
 51  TMSAKEKGKF  EDMAKADKAR  YEREMKTYIP  PKGETKKKFK  DPNAPKRPPS
101  AFFLFSSEYR  PKIKGEHPGL  SIGDVAKKLG  EMWNNTAADD  KQPYEKKAAK
151  LKEKYEKDIA  AYRAKGKPDA  AKKGVVKAEK  SKKKKEEEED  EEDEEDEEEE
201  EDEEDEDEEE  DDDDEAENLY  FQSHHHHHHD  YKDDDDK
```

Suppl. data A2-3-C Table 8.3, LC-ESI-TOF/MS verification of HMGB1A constructs.

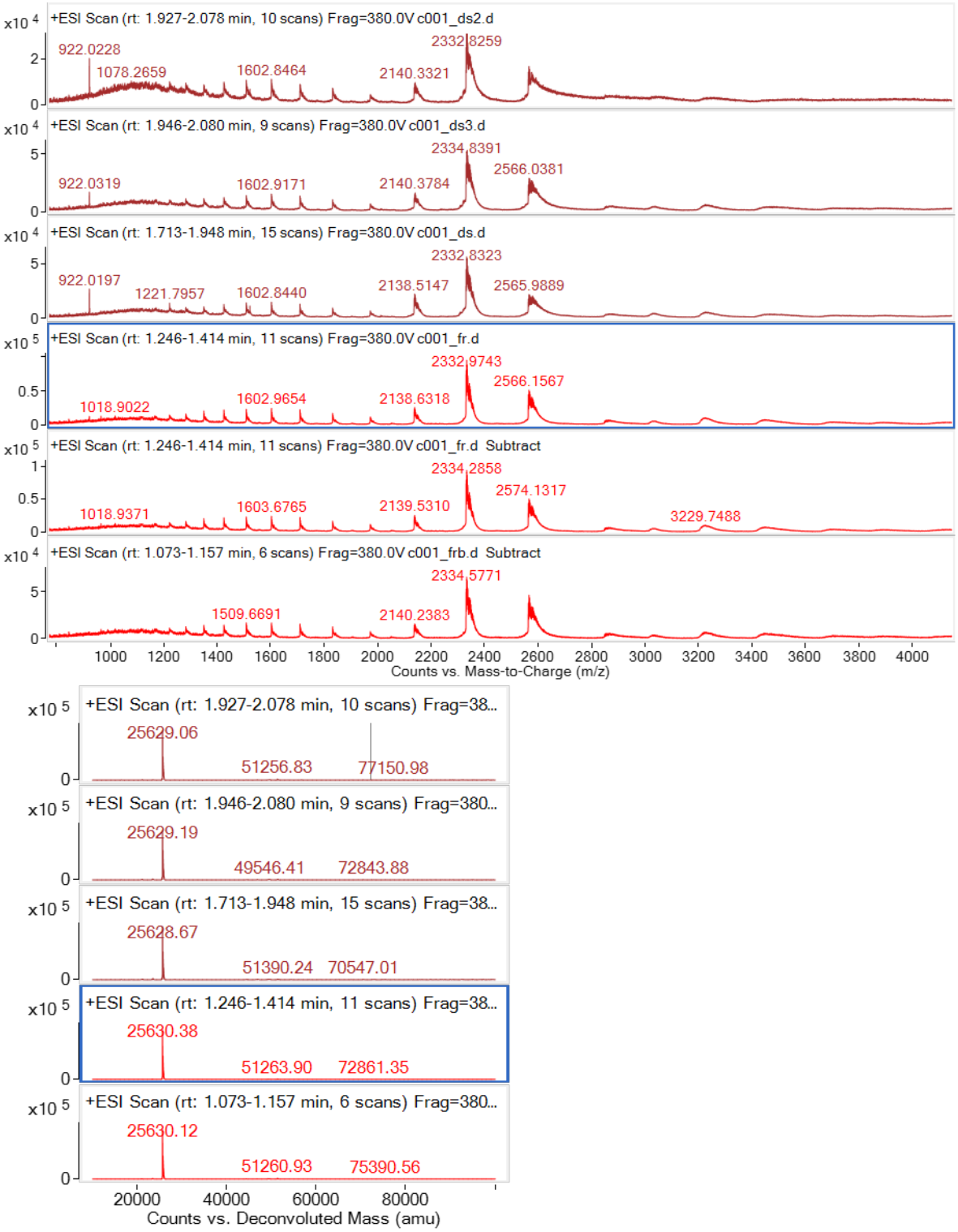



+ESI Scan (rt: 0.721-0.821...
12975.29
FR-HMGB1A-c028
(with tag)
+ESI Scan (rt: 0.688-0.805...
13503.74
3S-HMGB1A-c029
(with tag)
+ESI Scan (rt: 0.817-1.101...
12959.18
3S-HMGB1A-c030
(with tag)
+ESI Scan (rt: 0.654-0.937...
13085.16
16019.11
FR-HMGB1A-c033
(with tag, no Met)
+ESI Scan (rt: 0.655-0.938...
12480.51
14307.24
FR-HMGB1A-c034
(with tag)
+ESI Scan (rt: 0.653-0.936...
11232.81
14307.40
FR-HMGB1A-c035
(no Met)
+ESI Scan (rt: 0.652-0.935...
10672.28
14865.10
3S-HMGB1A-c036
(no Met)
10000
12500
15000
Counts vs. Deconvoluted M...
+ESI Scan (rt: 0.653-0.936 min, 18 scans) Frag=250
11232.81
14307.40
FR-HMGB1A-c037
(with tag)
+ESI Scan (rt: 0.652-0.935 min, 18 scans) Frag=250
10672.28
14865.10
FR-HMGB1A-c038
(with tag)
+ESI Scan (rt: 0.660-0.926 min, 17 scans) Frag=250
15388.87
13206.07
+ESI Scan (rt: 0.651-0.934 min, 18 scans) Frag=250
14652.86
12250.03
3S-HMGB1A-c039
(with tag, no Met)
+ESI Scan (rt: 0.656-0.923 min, 17 scans) Frag=250
13052.94
16019.01
3S-HMGB1A-c040
(with tag)
+ESI Scan (rt: 0.653-0.936 min, 18 scans) Frag=250
12464.07
16018.95
3S-HMGB1A-c041
(with tag, no Met)
+ESI Scan (rt: 0.651-0.934 min, 18 scans) Frag=250
11200.66
16018.72
+ESI Scan (rt: 0.659-0.926 min, 17 scans) Frag=250
16018.55
3S-HMGB1A-c042
(with tag, no Met)
10000
12500
15000
Counts vs. Deconvoluted Mass (amu)

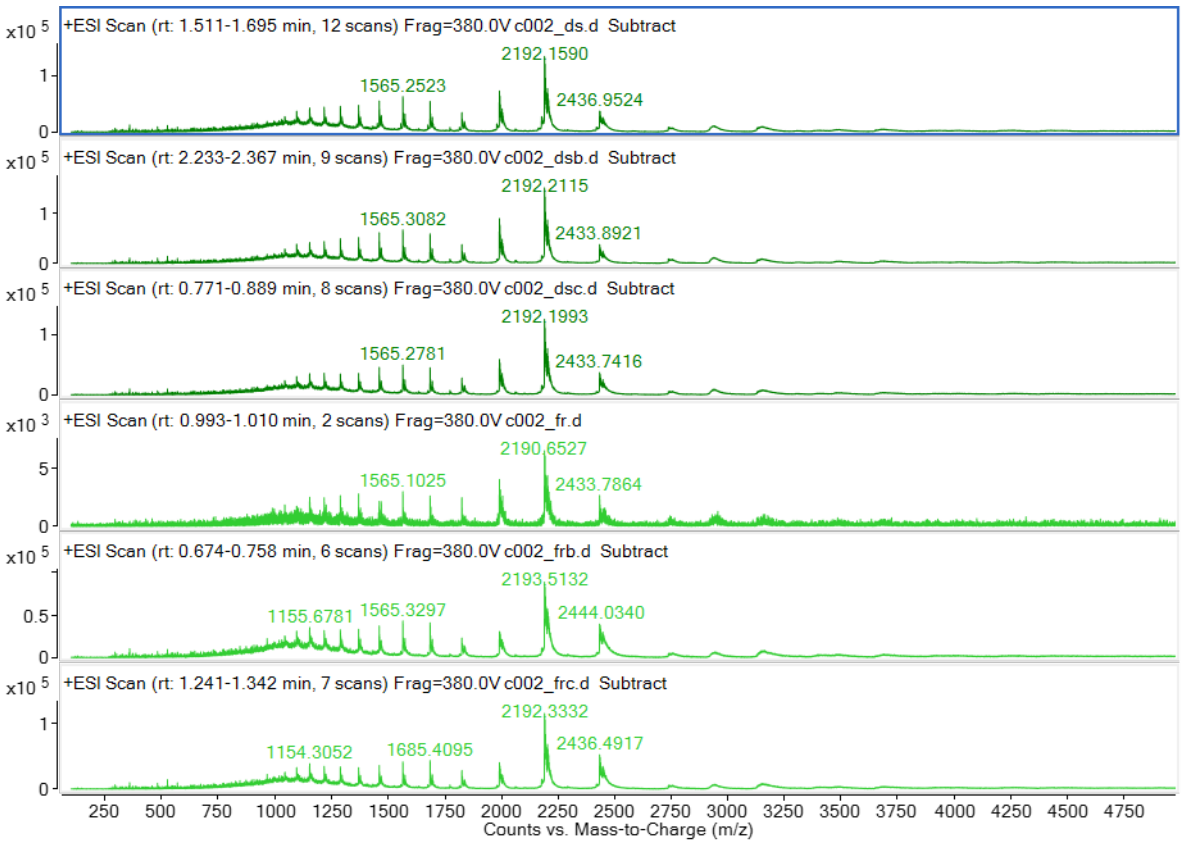
Suppl. data A2-3-D Figure 3.18, pp. 96, all replicate spectra

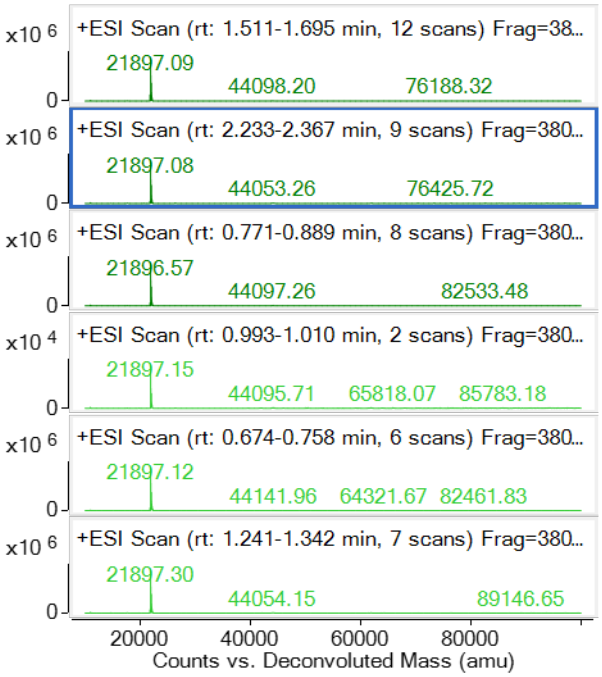
Top set of graphs; ESI-TOF spectra; bottom, deconvolution.

HMGB1A-c001 (DS/FR)

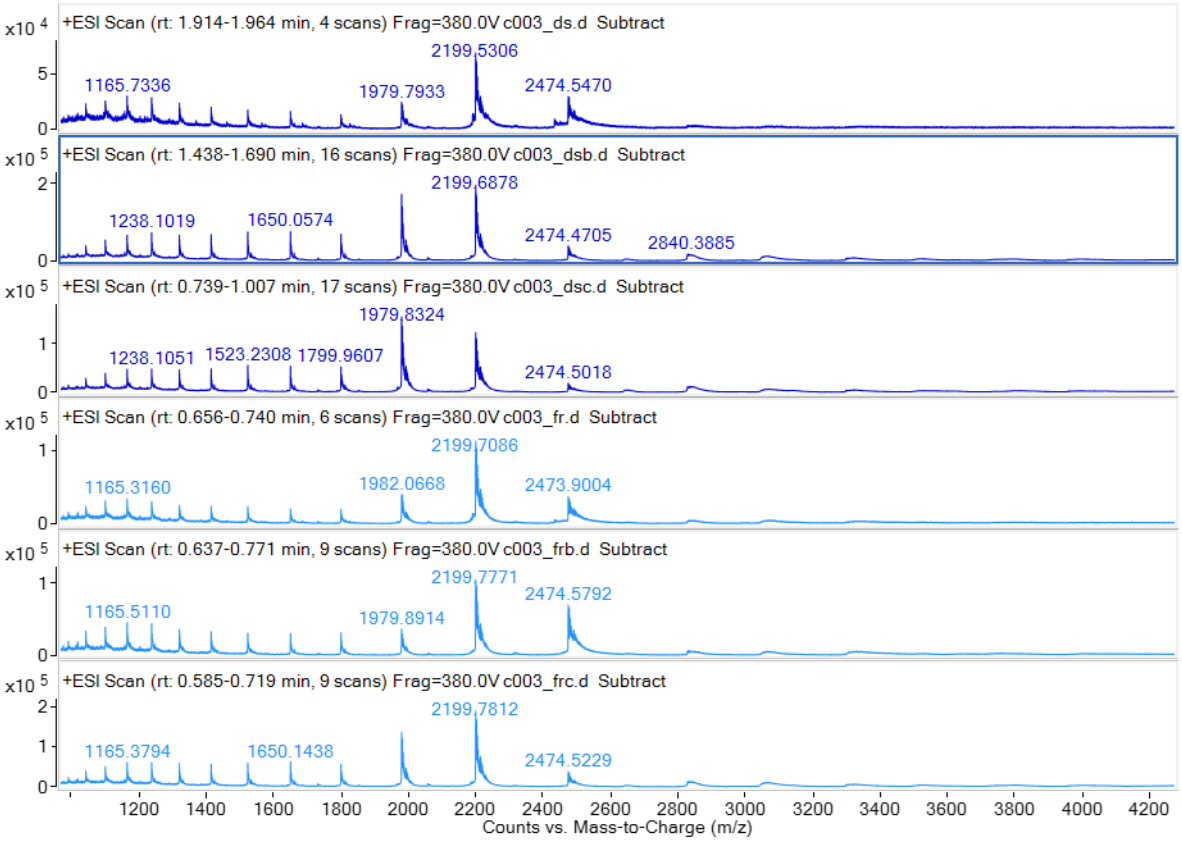


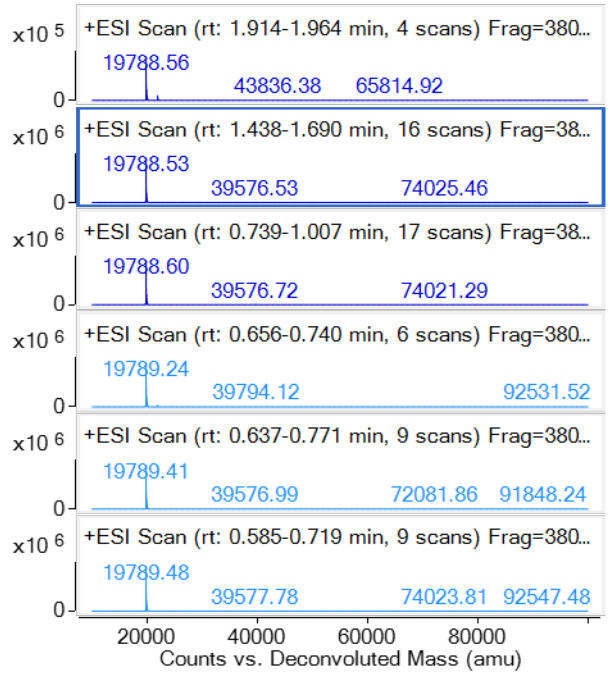
HMGB1A-c002 (DS/FR)



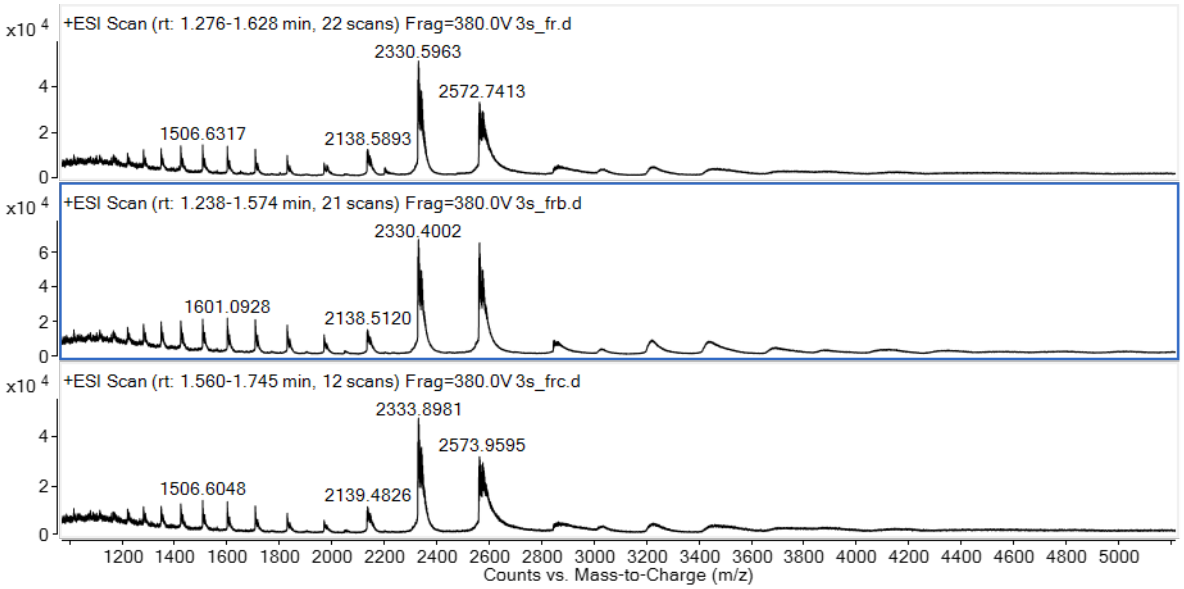


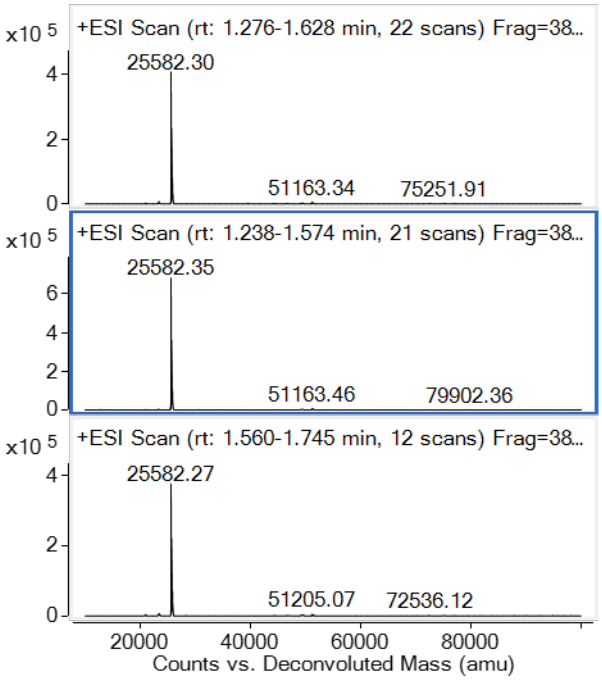
HMGB1A-c031 (DS/FR)





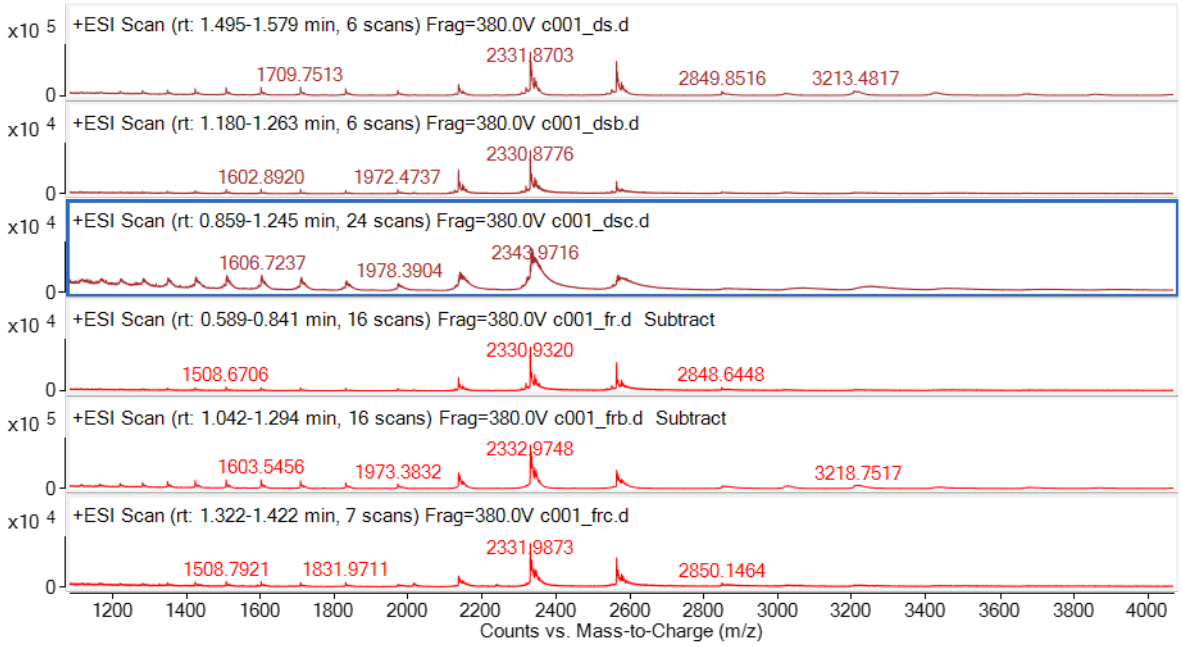
HMGB1A-c006 (3S)

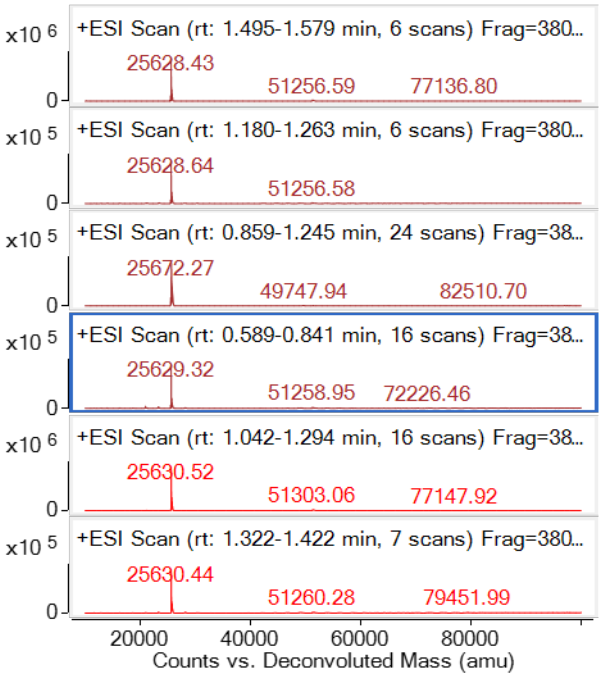




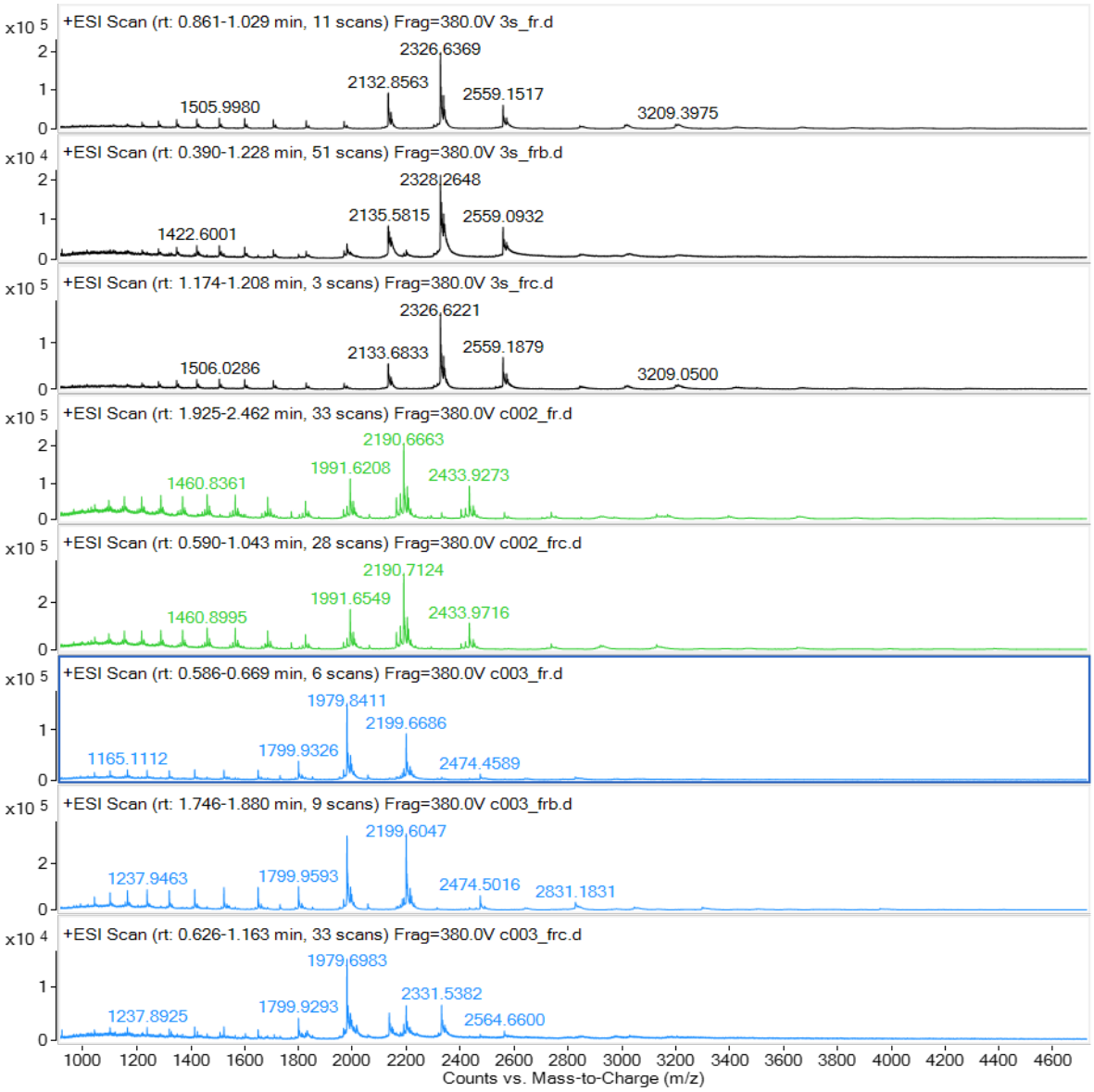
Suppl. data A2-3-E Figure 3.19, pp. 98, all spectra

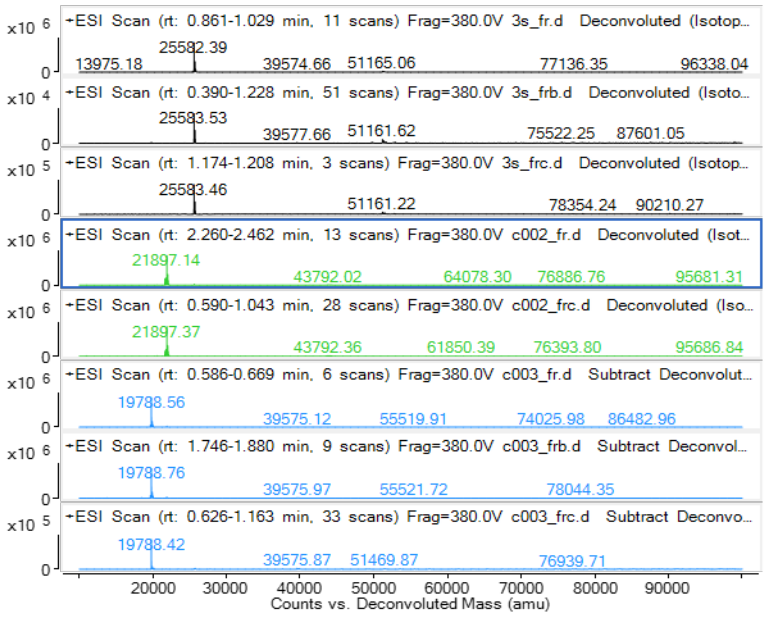
HMGB1A-c001 (DS/FR)





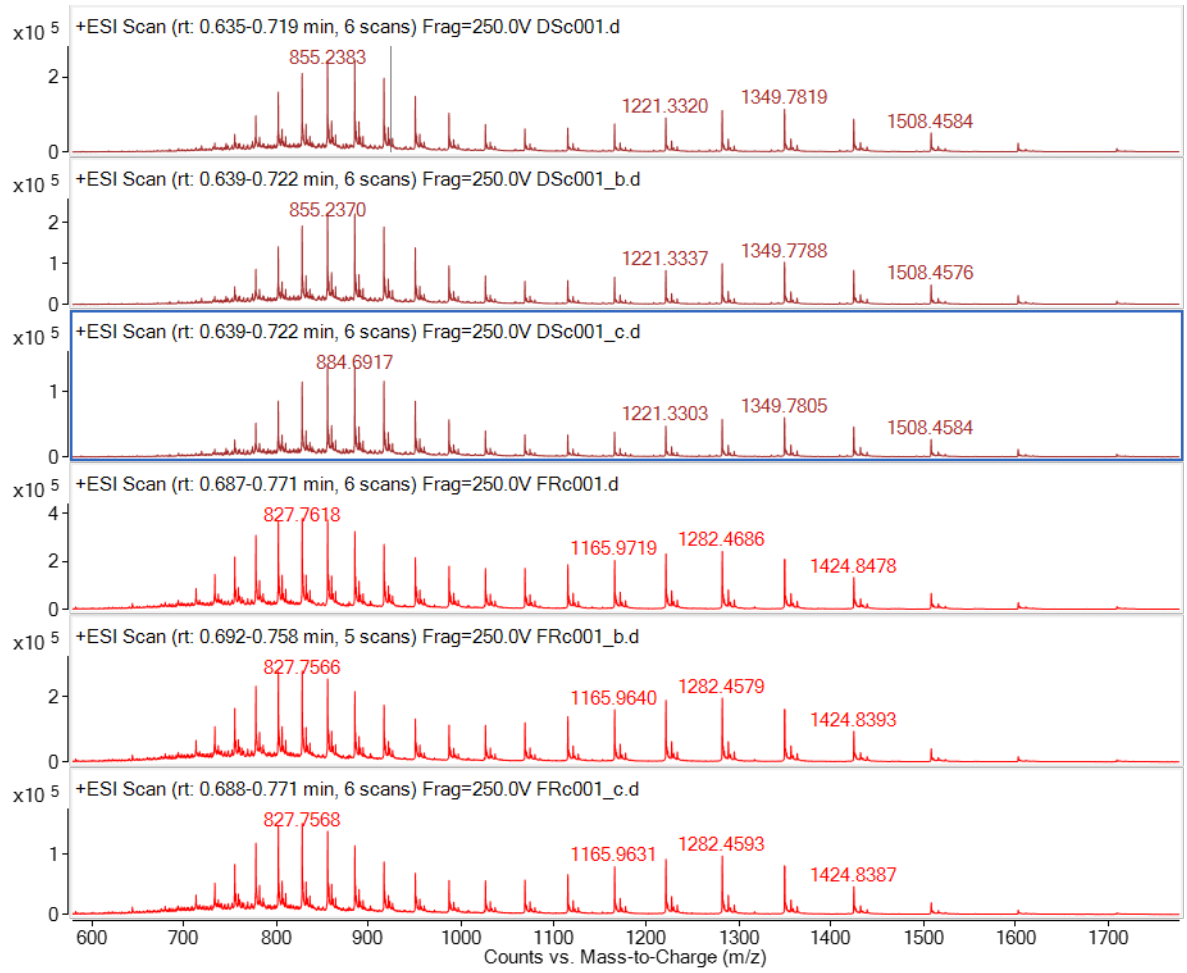
HMGB1A-c006 (3S), c002, c031 (FR)

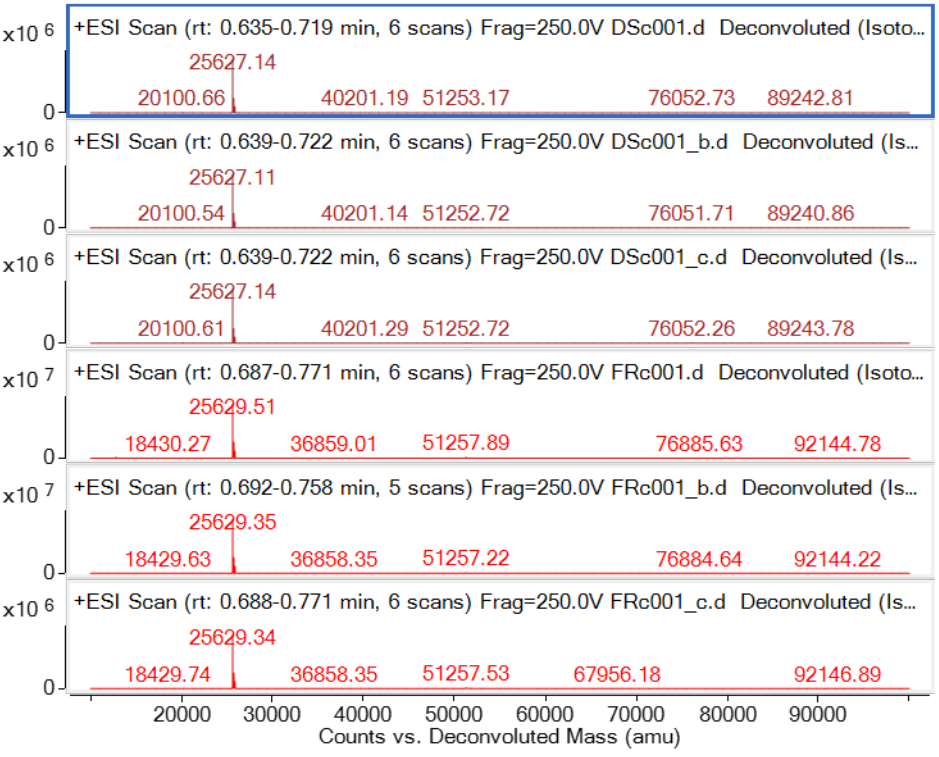




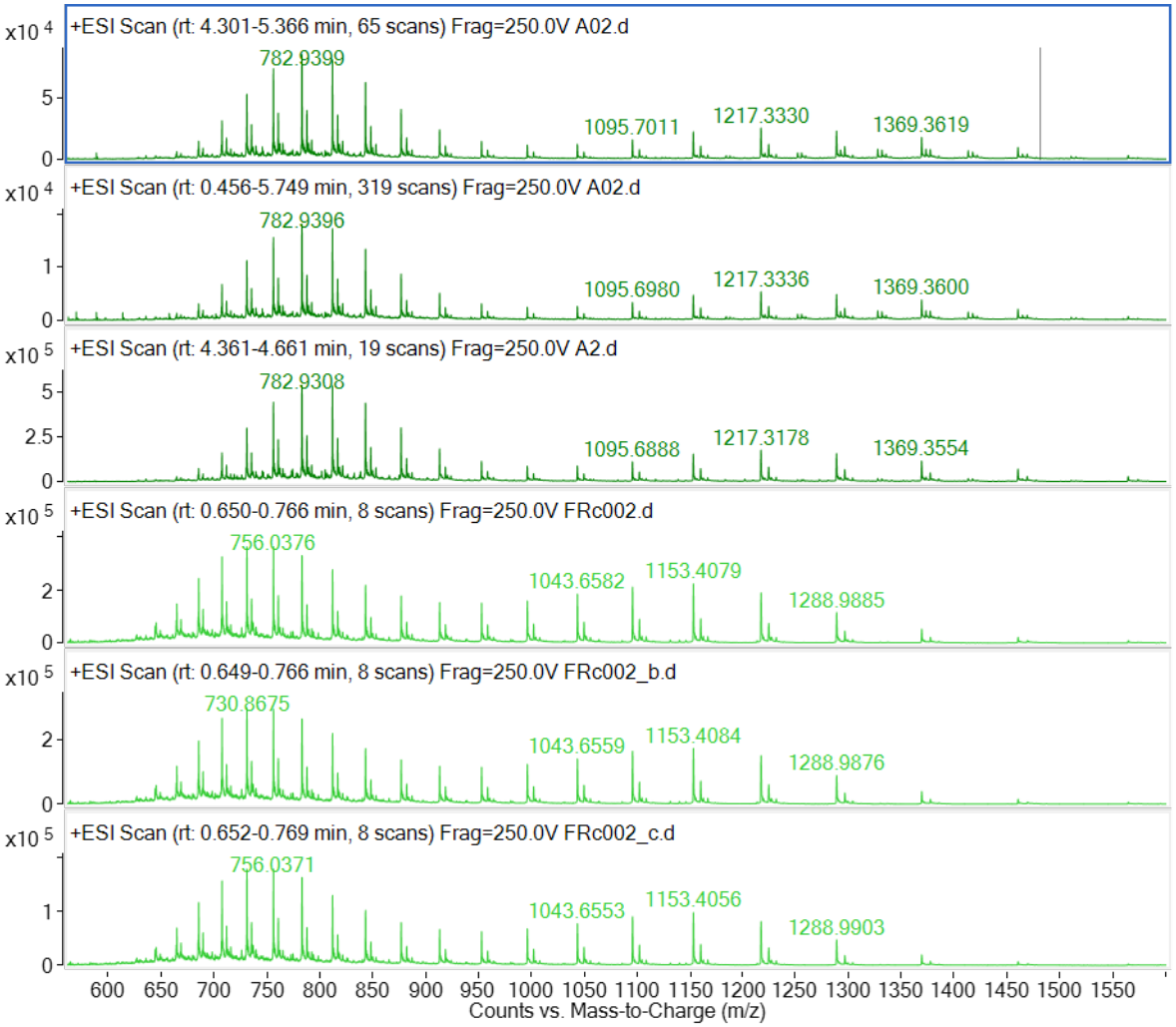
Suppl. data A2-3-F Figure 3.20, pp. 100; all spectra

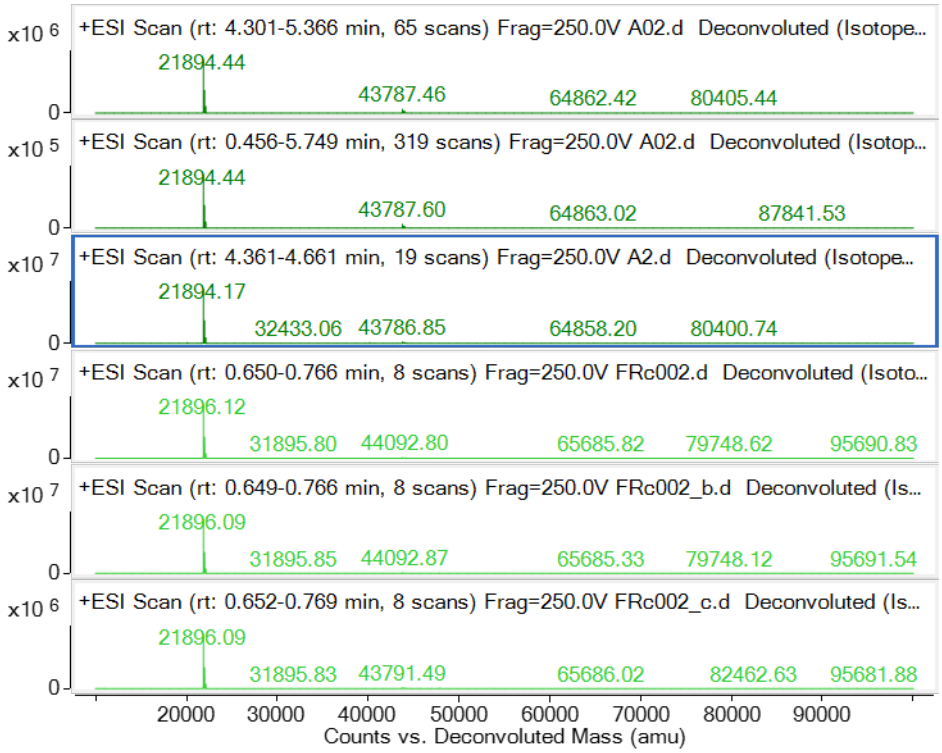
HMGB1A-c001 (DS/FR)



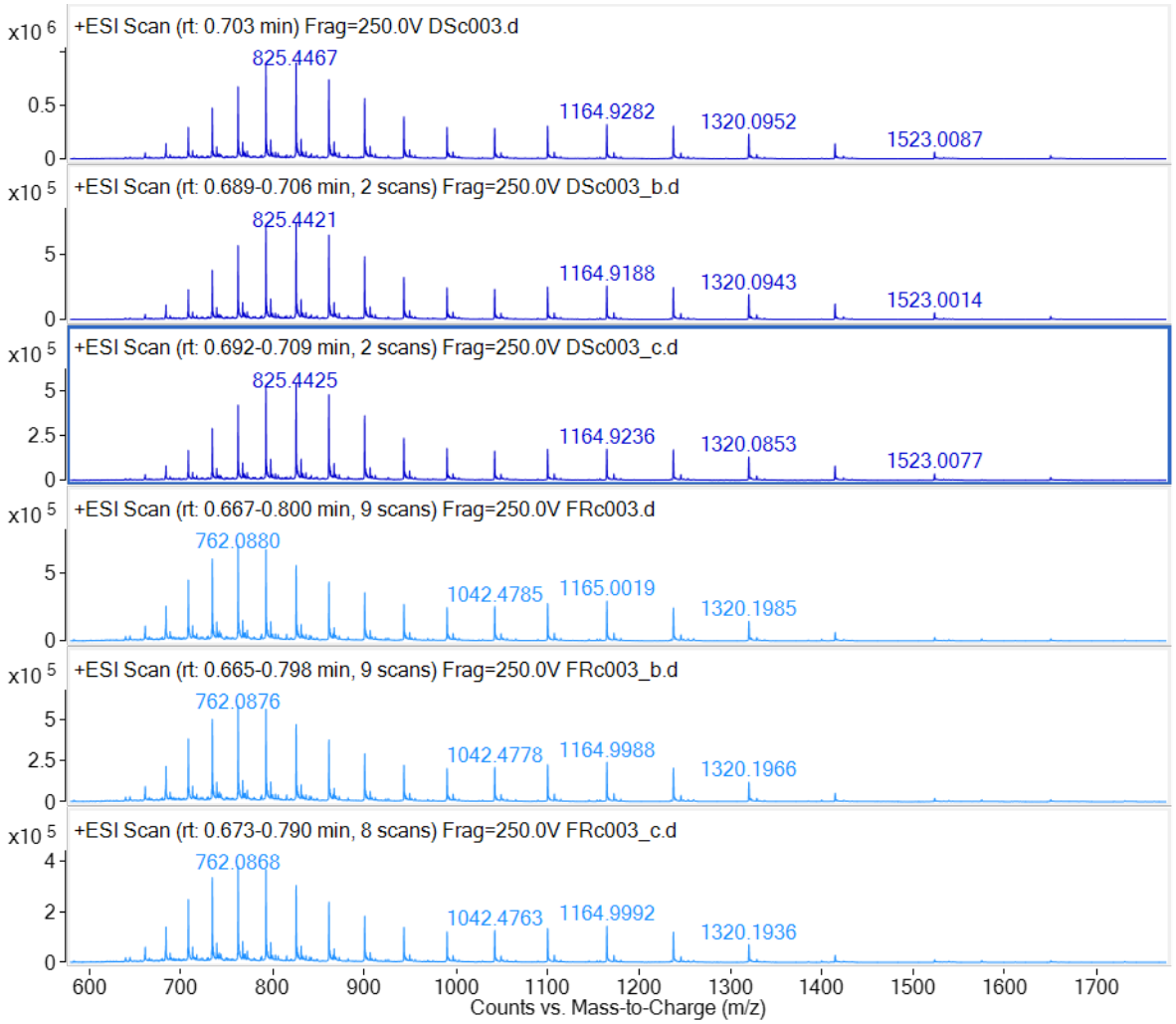


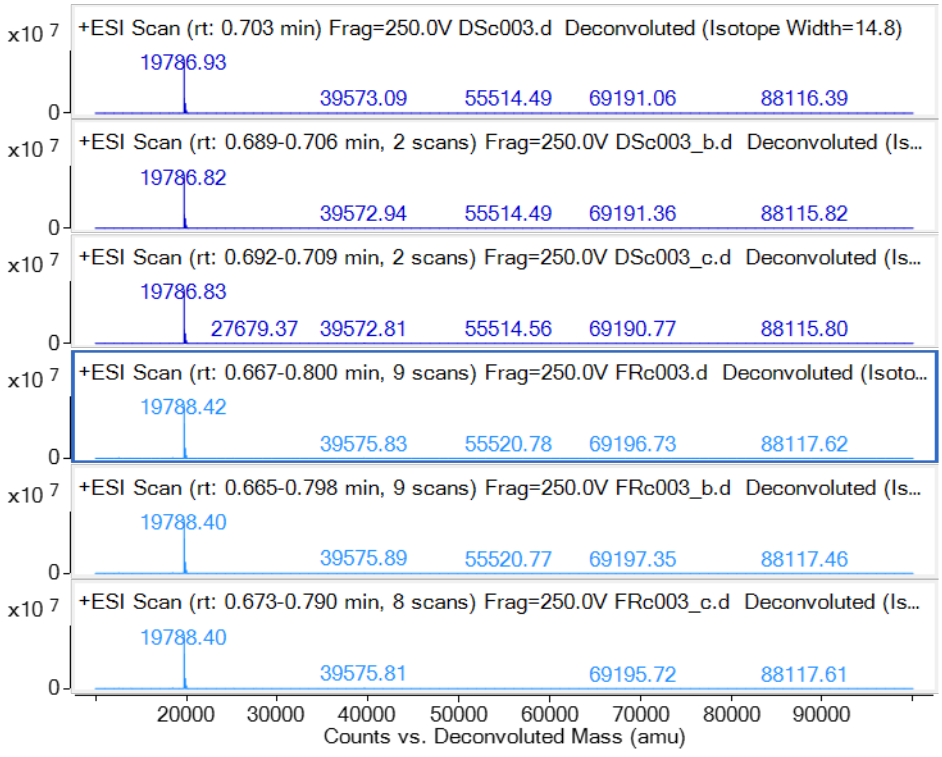
HMGB1A-c002 (DS/FR)



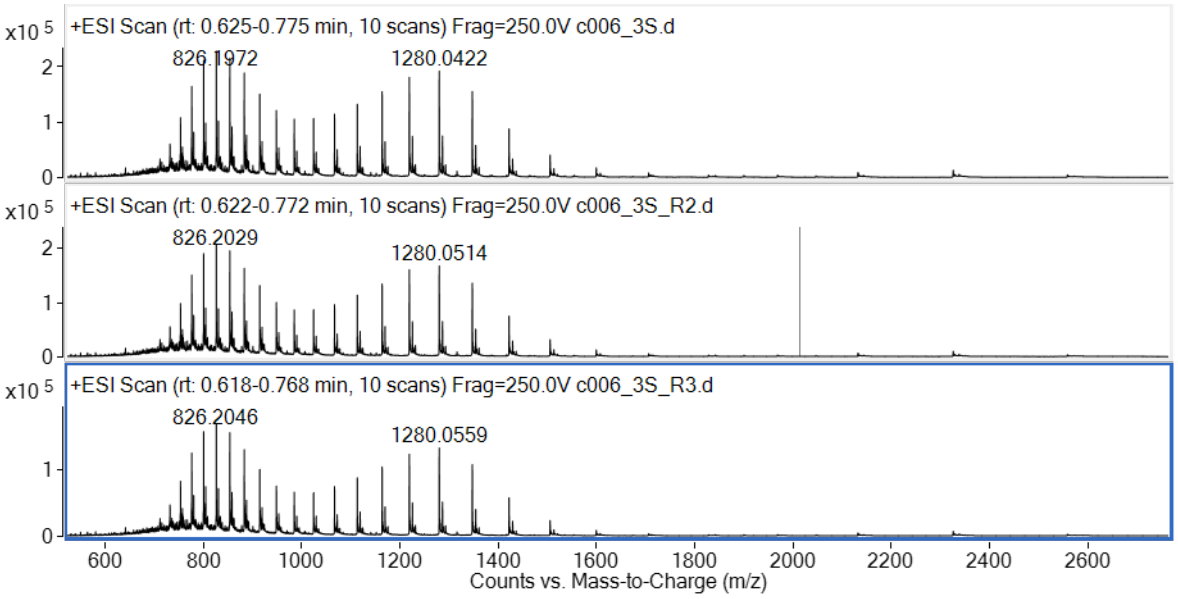


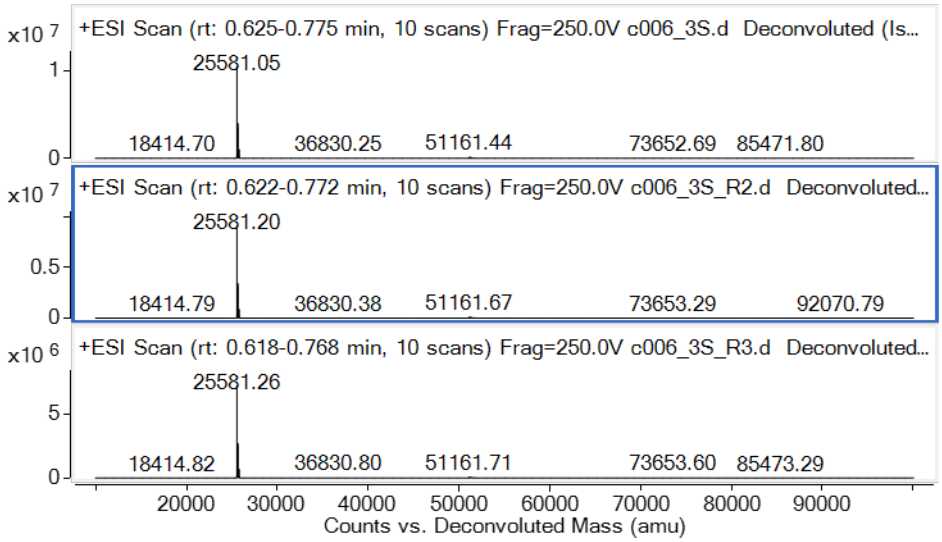
HMGB1A-c003 (DS/FR)





HMGB1A-c006 (3S)





Suppl. data A2-3-G Figure 3.25A, pp.114, tryptic digest MS/MS verification

Construct		MASCOT score	Match	Output
CXCL12A-C002	**SGC constructs**	135	SLFN5A-c001	No proteins detected ☹
	Swissprot sequences (top 5)	105	Schlafen family member 5 OS=Homo sapiens GN=SLFN5 PE=1 SV=1	
		76	50S ribosomal protein L7/L12 OS=Escherichia coli O139:H28 (s	
		71	50S ribosomal protein L28 OS=Escherichia coli O139:H28 (stra	
		71	50S ribosomal protein L28 OS=Ruthia magnifica subsp. Calypto	
		58	50S RIBOSOMAL PROTEIN L28 OS=RUTHIA MAGNIFICA SUBSP. CALYPT	
CXCL12A-C003	**SGC constructs**	77	SLFN5A-c001	No proteins detected ☹
	Swissprot sequences (top 5)	77	Schlafen family member 5 OS=Homo sapiens GN=SLFN5 PE=1 SV=1	
		61	SCHLAFEN FAMILY MEMBER 5 OS=HOMO SAPIENS GN=SLFN5 PE=1 SV=1	
		56	THREONYLCARBAMOYL-AMP SYNTHASE OS=VIBRIO FISCHERI (STRAIN A	
		54	METHIONINE AMINOPEPTIDASE OS=STAPHYLOCOCCUS AUREUS (STRAIN	
		53	AUTOPHAGY-RELATED PROTEIN 2 OS=NEOSARTORYA FISCHERI (STRAIN	
CXCL12A-C004	**SGC constructs**	207	SLFN5A-c001	No proteins detected ☹
	Swissprot sequences (top 5)	183	Schlafen family member 5 OS=Homo sapiens GN=SLFN5 PE=1 SV=1	
		50	Enoyl-[acyl-carrier-protein] reductase [NADH] OS=Idiomarina	
		50	ENOYL-[ACYL-CARRIER-PROTEIN] REDUCTASE [NADH] OS=IDIOMARINA	
		49	AUTOPHAGY-RELATED PROTEIN 2 OS=MAGNAPORTHE ORYZAE (STRAIN 7	
		49	BACULOVIRAL IAP REPEAT-CONTAINING PROTEIN 5 OS=BOS TAURUS G	
CXCL12A-C005	**SGC constructs**	109	SLFN5A-c001	Target protein only ☺
	Swissprot sequences (top 5)	109	Schlafen family member 5 OS=Homo sapiens GN=SLFN5 PE=1 SV=1	
		68	Stromal cell-derived factor 1 OS=Felis catus GN=CXCL12 PE=3	
		51	STROMAL CELL-DERIVED FACTOR 1 OS=FELIS CATUS GN=CXCL12 PE=3	
		49	50S RIBOSOMAL PROTEIN L17 OS=PARVIBACULUM LAVAMENTIVORANS (	
		49	GANGLIOSIDE-INDUCED DIFFERENTIATION-ASSOCIATED PROTEIN 2 OS	
CXCL12A-C006	**SGC constructs**	144	CXCL12A-c006	Target protein only ☺
	Swissprot sequences (top 5)	94	Stromal cell-derived factor 1 OS=Felis catus GN=CXCL12 PE=3	
		64	Schlafen family member 5 OS=Homo sapiens GN=SLFN5 PE=1 SV=1	
		55	SCHLAFEN FAMILY MEMBER 5 OS=HOMO SAPIENS GN=SLFN5 PE=1 SV=1	
		50	M1 family aminopeptidase OS=Plasmodium falciparum (isolate F	
		49	M1 FAMILY AMINOPEPTIDASE OS=PLASMODIUM FALCIPARUM (ISOLATE	

Suppl. data A2-3-H Figure 3.25B, pp.114, tryptic digest MS/MS verification

Construct	MASCOT score		Match	Output
AV1 CXCL12A-c003 Main	**SGC constructs**	247	CXCL12A-c003	Target protein only ☺
	Swissprot sequences (top 5)	207	Stromal cell-derived factor 1 OS=Felis catus GN=CXCL12 PE=3	
		99	Trypsin OS=Sus scrofa PE=1 SV=1	
		66	TRYPSIN OS=SUS SCROFA PE=1 SV=1	
		66	RNA polymerase sigma factor SigF OS=Mycobacterium tuberculos	
		66	Putative pre-16S rRNA nuclease OS=Cupriavidus metallidurans	
AV2 CXCL12A-c003 Cont	**SGC constructs**	89	CDPK5PFA-c001	No proteins detected ☹
	Swissprot sequences (top 5)	95	Trypsin OS=Sus scrofa PE=1 SV=1	
		62	Trypsin-1 OS=Homo sapiens GN=PRSS1 PE=1 SV=1	
		52	Probable ribosomal RNA small subunit methyltransferase A OS=	
		52	PROBABLE RIBOSOMAL RNA SMALL SUBUNIT METHYLTRANSFERASE A OS	
		52	GLYCINE DEHYDROGENASE (DECARBOXYLATING) OS=GRAMELLA FORSETI	

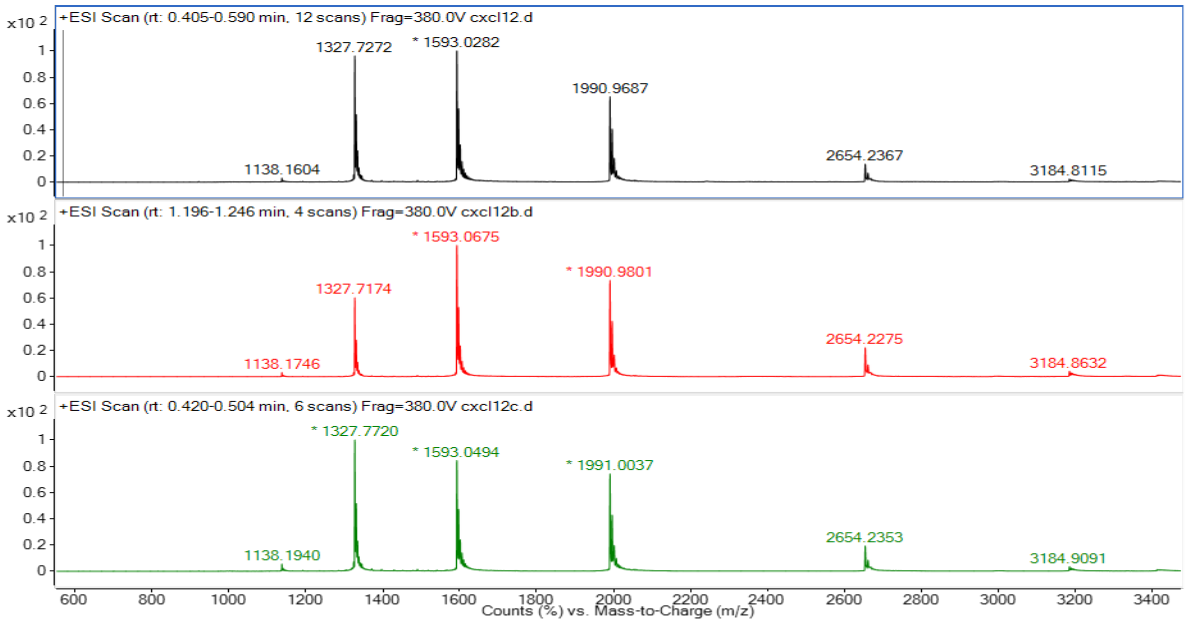
Suppl. data A2-3-I Figure 3.26, pp. 115, tryptic digest MS/MS verification

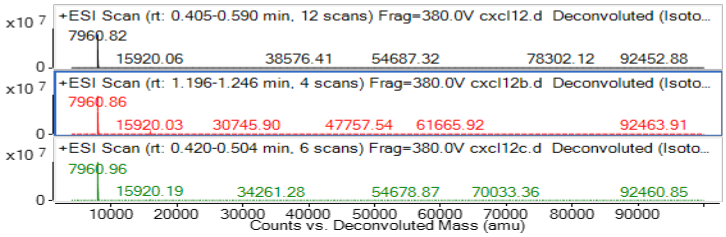
Construct		MASCOT score	Match	Output
AV1 CXCl12A-c021 ULPDSBc Top band	**SGC constructs**	606	GRNA-c011	Contaminating proteins predominate, target is present ☺
	Swissprot sequences (top 5)	556	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		194	Small ubiquitin-related modifier 1 OS=Bos taurus GN=SUMO1 PE	
		125	Stromal cell-derived factor 1 OS=Felis catus GN=CXCL12 PE=3	
		92	Trypsin OS=Sus scrofa PE=1 SV=1	
		68	Thioredoxin H-type OS=Ricinus communis PE=3 SV=1	
AV2 CXCL12a-c025 ULPDSBc Top band	**SGC constructs**	636	FCN1A-c001	Contaminating proteins predominate, target is present ☺
	Swissprot sequences (top 5)	514	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		136	Stromal cell-derived factor 1 OS=Felis catus GN=CXCL12 PE=3	
		83	Small ubiquitin-related modifier 1-A OS=Xenopus laevis GN=su	
		63	Thioredoxin H-type OS=Ricinus communis PE=3 SV=1	
		43	Teichoic acids export ATP-binding protein TagH OS=Staphyloco	
AV1 CXCl12A-c021 ULPDSBc Bottom band 1	**SGC constructs**	387	GRNA-c127	Contaminating proteins only, target is not detected ☹
	Swissprot sequences (top 5)	363	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		292	Elongation factor Tu OS=Salmonella arizonae (strain ATCC BAA	
		286	Elongation factor Tu 2 OS=Escherichia coli O139:H28 (strain	
		280	Elongation factor Tu 1 OS=Shigella flexneri serotype 5b (str	
		208	Elongation factor Tu OS=Sodalis glossinidius (strain morsita	
AV2 CXCL12a-c025 ULPDSBc Bottom band 1	**SGC constructs**	460	FCN1A-c001	Contaminating proteins only, target is not detected ☹
	Swissprot sequences (top 5)	450	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		130	Small ubiquitin-related modifier 1-B OS=Xenopus laevis GN=su	
		79	Trypsin OS=Sus scrofa PE=1 SV=1	
		74	Elongation factor Tu OS=Salmonella arizonae (strain ATCC BAA	
		68	Thioredoxin H-type OS=Ricinus communis PE=3 SV=1	
AV5 CXCl12A-c003 40,ULPDSBc Bottom band 2	**SGC constructs**	159	FCN1A-c001	Contaminating proteins only, target is not detected ☹
	Swissprot sequences (top 5)	156	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		101	Trypsin OS=Sus scrofa PE=1 SV=1	
		62	Thioredoxin H-type OS=Ricinus communis PE=3 SV=1	
		33	Accession: PGK_DEIRA description retrieval failed	
		32	Accession: FOLD_BORBR description retrieval failed	
AV6 CXCL12a-c006 40,ULPDSBc Bottom band 2	**SGC constructs**	168	RETC-c017	Contaminating proteins only, target is not detected ☹
	Swissprot sequences (top 5)	155	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		62	Trypsin OS=Sus scrofa PE=1 SV=1	
		55	Thioredoxin H-type OS=Ricinus communis PE=3 SV=1	
		47	Accession: G3P1_ECO57 description retrieval failed	
		34	Accession: FOLD_BORBR description retrieval failed	

Construct		MASCOT score	Match	Output
AV5 CXCl12A-c003 40,ULPDSBc Bottom band 3	**SGC constructs**	347	EPHA2A-c006	Contaminating proteins only, target is not detected ☹
	Swissprot sequences (top 5)	291	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		63	Thioredoxin H-type OS=Ricinus communis PE=3 SV=1	
		57	Trypsin OS=Sus scrofa PE=1 SV=1	
		38	Accession: FOLD_BORBR description retrieval failed	
		34	Fusaric acid resistance protein FusC OS=Burkholderia cepacia	
AV6 CXCL12a-c006 40,ULPDSBc Bottom band 3	**SGC constructs**	76	RETC-c017	Contaminating proteins only, target is not detected ☹
	Swissprot sequences (top 5)	72	Trypsin OS=Sus scrofa PE=1 SV=1	
		67	Thiol:disulfide interchange protein DsbC OS=Escherichia coli	
		46	Thioredoxin H-type OS=Ricinus communis PE=3 SV=1	
		36	Mediator of RNA polymerase II transcription subunit 4 OS=Can	
		30	ATP-dependent RNA helicase DBP10 OS=Lodderomyces elongisporu	
AV9 CXCL12a-c027 Band 1 (Red)	**SGC constructs**	157	CXCL12A-c001	Target protein with contaminating proteins ☺
	Swissprot sequences (top 5)	147	30S ribosomal protein S3 (Fragment) OS=Buchnera aphidicola s	
		120	Small ubiquitin-related modifier 1 OS=Bos taurus GN=SUMO1 PE	
		77	Trypsin OS=Sus scrofa PE=1 SV=1	
		65	30S ribosomal protein S2 OS=Escherichia coli O127:H6 (strain	
		55	Stromal cell-derived factor 1 OS=Felis catus GN=CXCL12 PE=3	

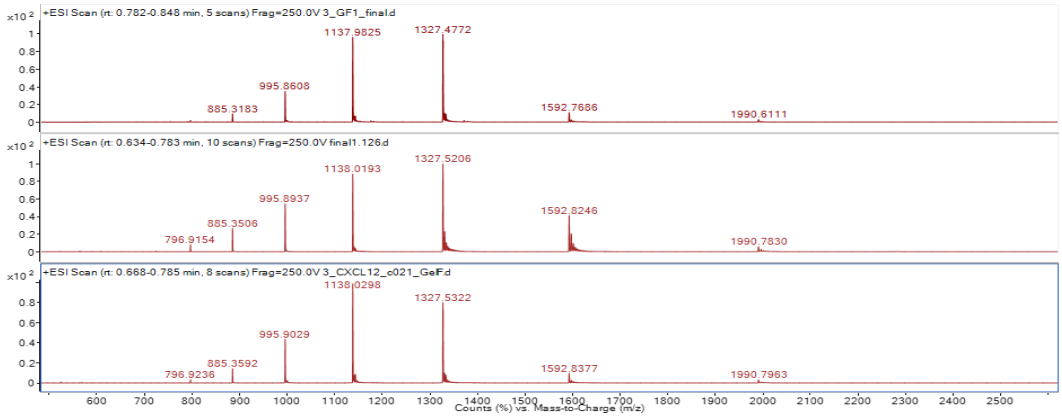
Suppl. data A2-3-J Figure 3.33, pp. 127, all spectra

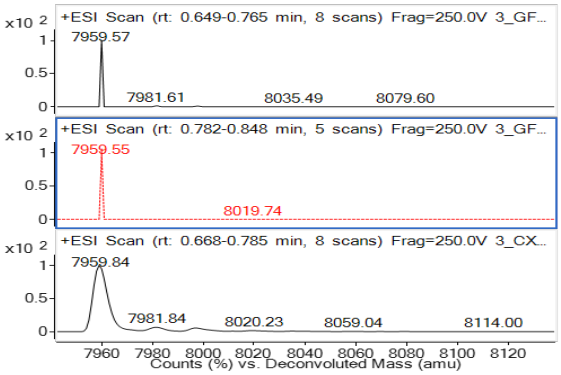
Native ESI/MS





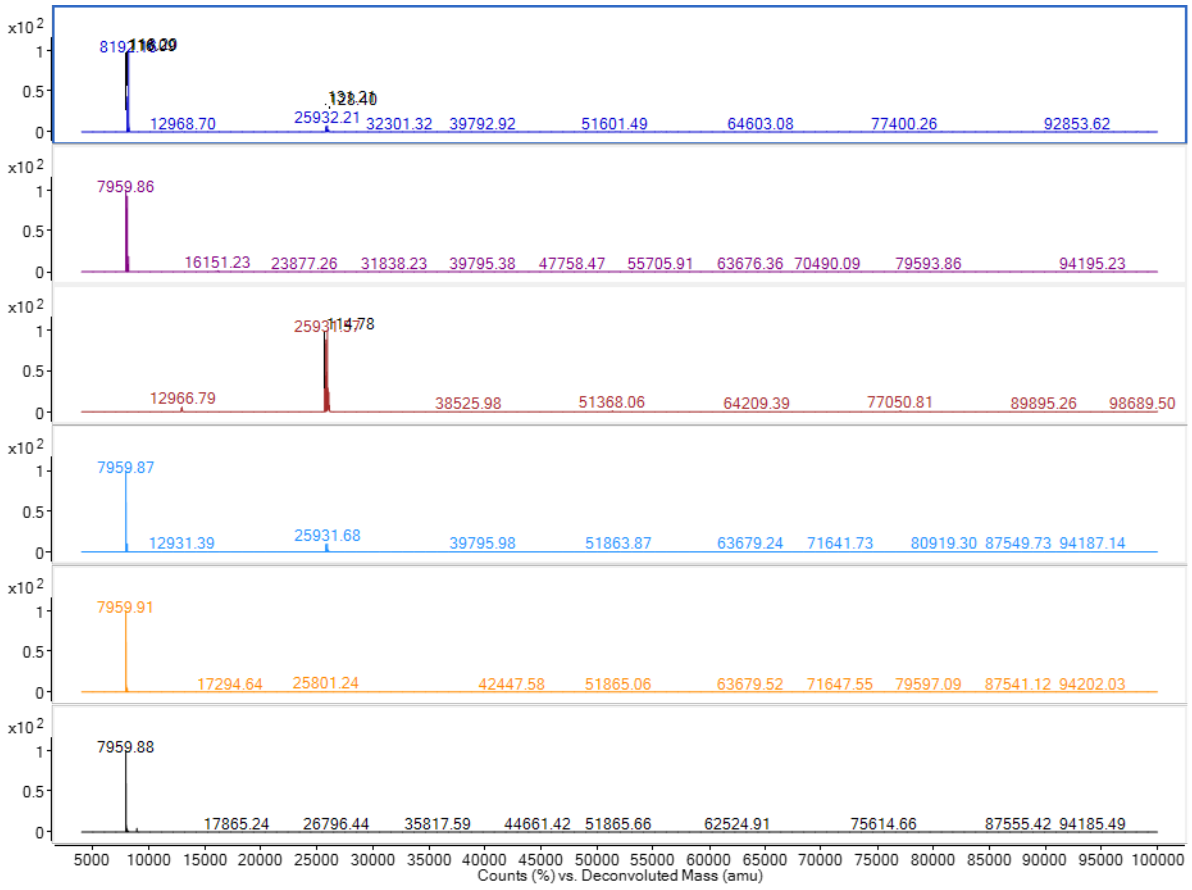
Denaturing ESI/MS





Chapter 4 MS data

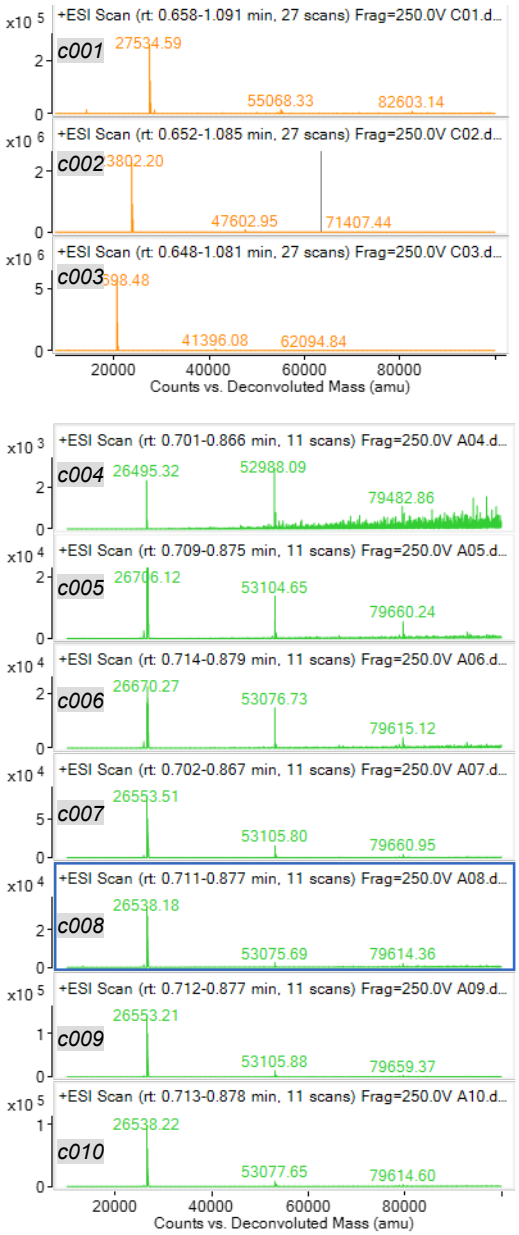
Suppl. data A2-4-A Figure 4.3, pp. 139 (full range deconvoluted spectra)

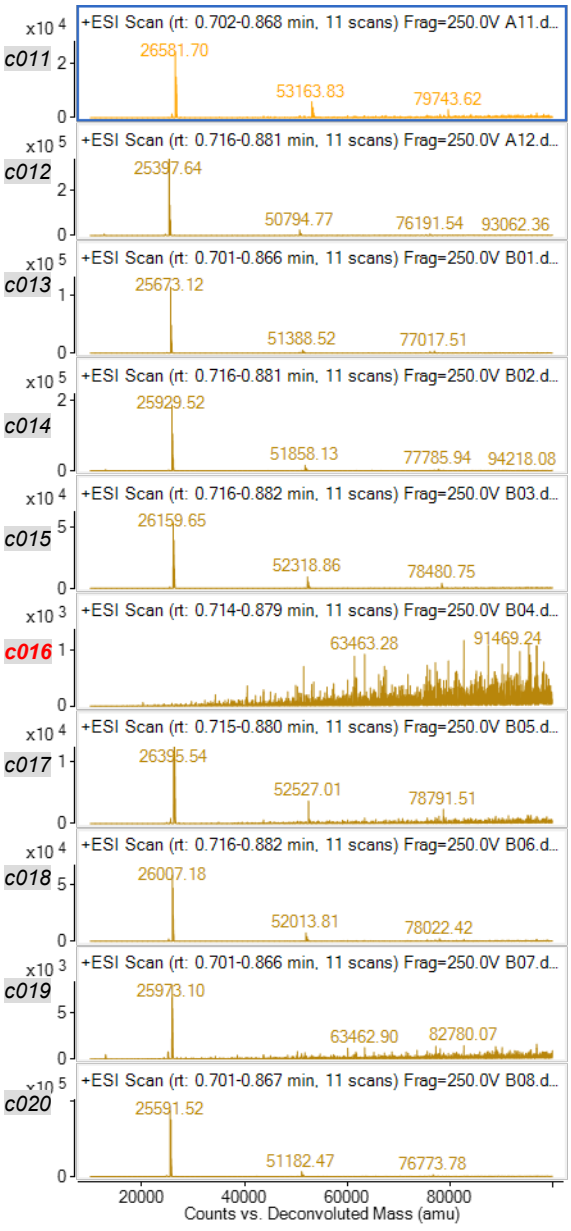


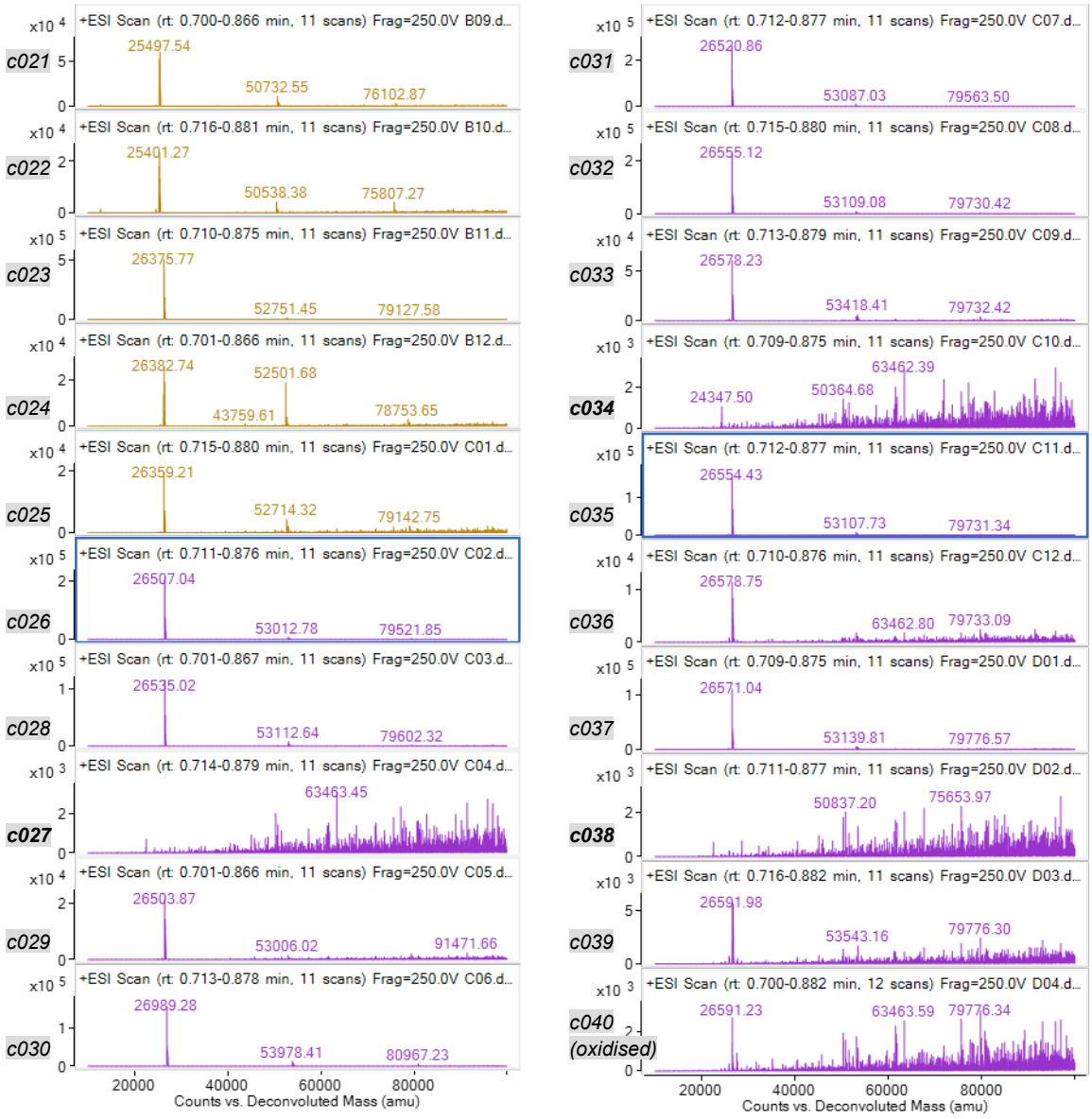
Chapter 5 MS data

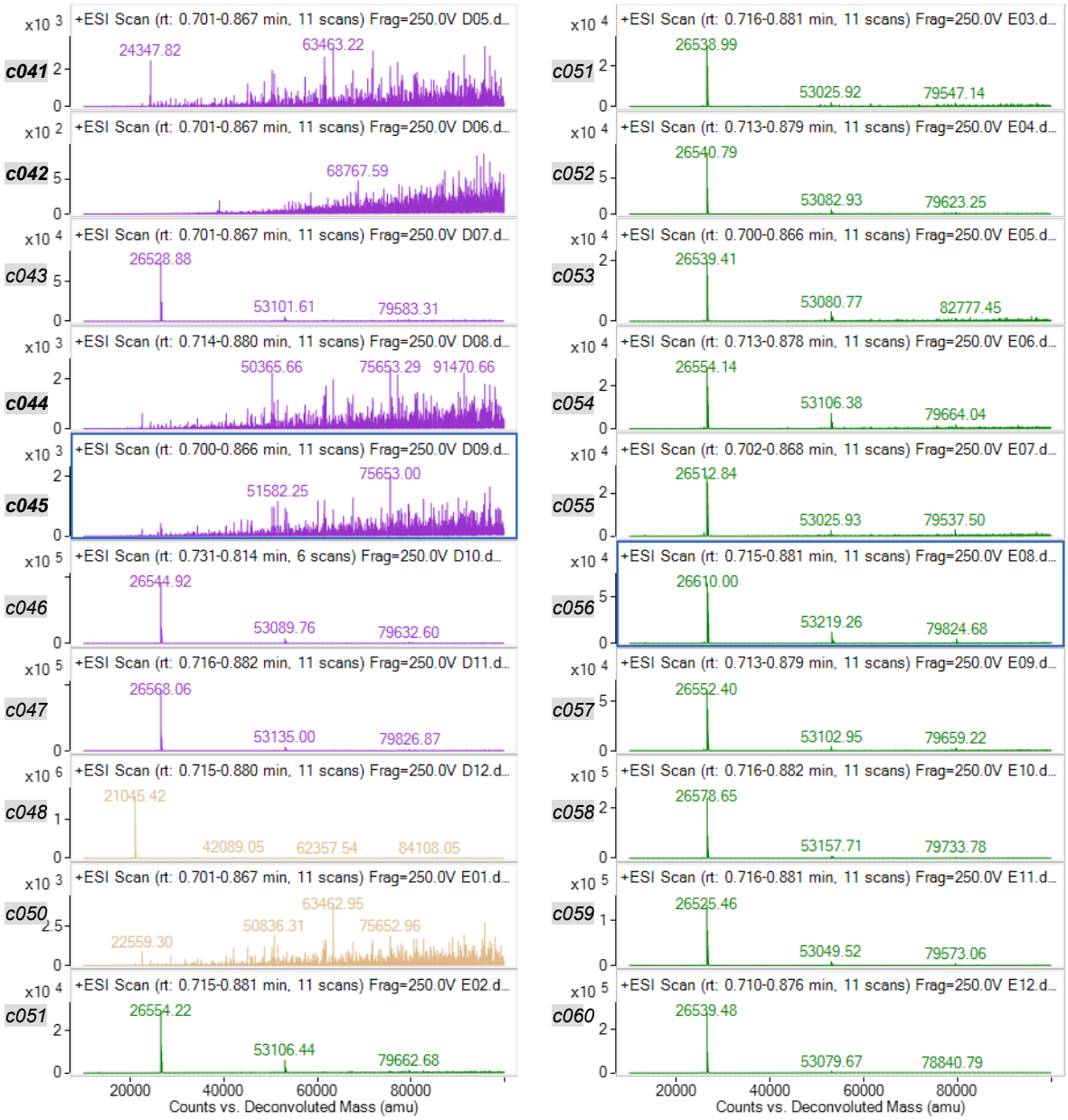
Suppl. data A2-5-A Table 8.5, pp. - 28 - (LC-ESI-TOF/MS verification of HMGB1-C1 constructs)

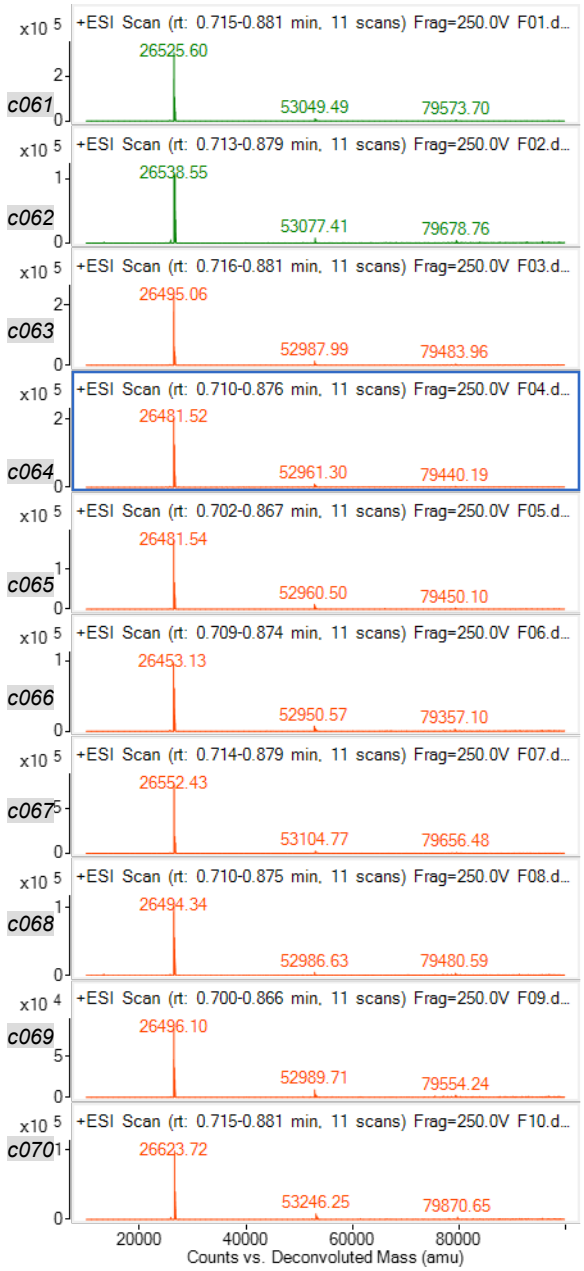
*Most constructs experience N-terminal Met loss (-131 Da); all have the His6 tag present. Identifiers in **bold red** designate no expression; identifiers in **bold** only were from failed sample runs (due to low protein yield) which were repeated in a separate experiment with larger scale.*

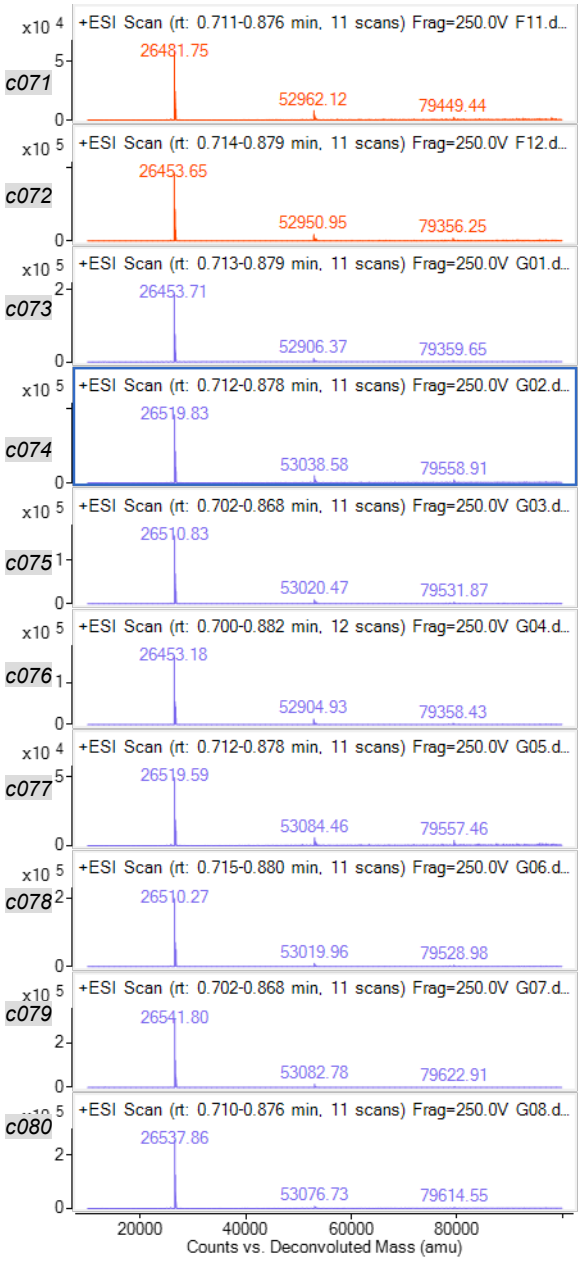


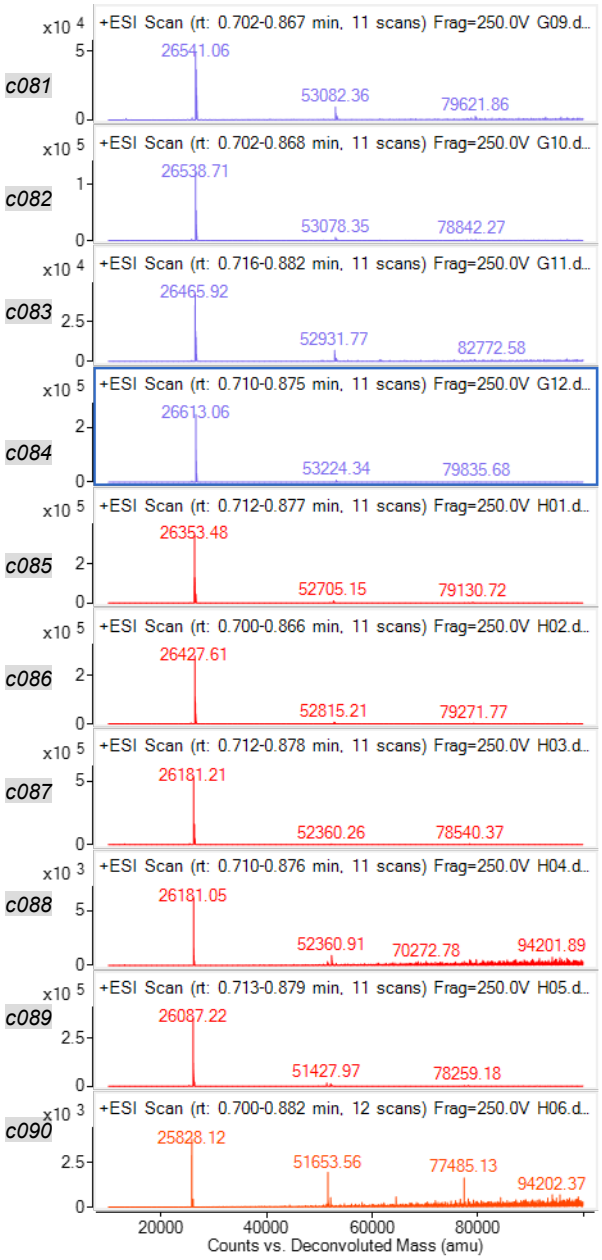



c021
+ESI Scan (rt: 0.700-0.866 min, 11 scans) Frag=250.0V B09.d...
25497.54
50732.55
76102.87
c022
+ESI Scan (rt: 0.716-0.881 min, 11 scans) Frag=250.0V B10.d...
25401.27
50538.38
75807.27
c023
+ESI Scan (rt: 0.710-0.875 min, 11 scans) Frag=250.0V B11.d...
26375.77
52751.45
79127.58
c024
+ESI Scan (rt: 0.701-0.866 min, 11 scans) Frag=250.0V B12.d...
26382.74
52501.68
43759.61
78753.65
c025
+ESI Scan (rt: 0.715-0.880 min, 11 scans) Frag=250.0V C01.d...
26359.21
52714.32
79142.75
c026
+ESI Scan (rt: 0.711-0.876 min, 11 scans) Frag=250.0V C02.d...
26507.04
53012.78
79521.85
c028
+ESI Scan (rt: 0.701-0.867 min, 11 scans) Frag=250.0V C03.d...
26535.02
53112.64
79602.32
c027
+ESI Scan (rt: 0.714-0.879 min, 11 scans) Frag=250.0V C04.d...
63463.45
c029
+ESI Scan (rt: 0.701-0.866 min, 11 scans) Frag=250.0V C05.d...
26503.87
53006.02
91471.66
c030
+ESI Scan (rt: 0.713-0.878 min, 11 scans) Frag=250.0V C06.d...
26989.28
53978.41
80967.23
20000
40000
60000
80000
Counts vs. Deconvoluted Mass (amu)
c031
+ESI Scan (rt: 0.712-0.877 min, 11 scans) Frag=250.0V C07.d...
26520.86
53087.03
79563.50
c032
+ESI Scan (rt: 0.715-0.880 min, 11 scans) Frag=250.0V C08.d...
26555.12
53109.08
79730.42
c033
+ESI Scan (rt: 0.713-0.879 min, 11 scans) Frag=250.0V C09.d...
26578.23
53418.41
79732.42
c034
+ESI Scan (rt: 0.709-0.875 min, 11 scans) Frag=250.0V C10.d...
63462.39
24347.50
50364.68
c035
+ESI Scan (rt: 0.712-0.877 min, 11 scans) Frag=250.0V C11.d...
26554.43
53107.73
79731.34
c036
+ESI Scan (rt: 0.710-0.876 min, 11 scans) Frag=250.0V C12.d...
26578.75
63462.80
79733.09
c037
+ESI Scan (rt: 0.709-0.875 min, 11 scans) Frag=250.0V D01.d...
26571.04
53139.81
79776.57
c038
+ESI Scan (rt: 0.711-0.877 min, 11 scans) Frag=250.0V D02.d...
50837.20
75653.97
c039
+ESI Scan (rt: 0.716-0.882 min, 11 scans) Frag=250.0V D03.d...
26591.98
53543.16
79776.30
c040
(oxidised)
+ESI Scan (rt: 0.700-0.882 min, 12 scans) Frag=250.0V D04.d...
26591.23
63463.59
79776.34
20000
40000
60000
80000
Counts vs. Deconvoluted Mass (amu)

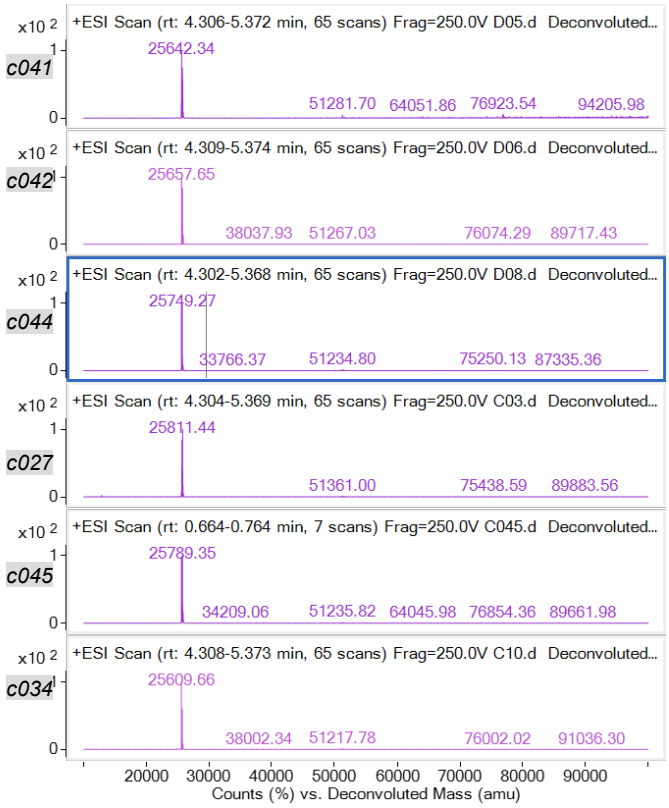

c041
+ESI Scan (rt: 0.701-0.867 min, 11 scans) Frag=250.0V D05.d...
24347.82
63463.22
c042
+ESI Scan (rt: 0.701-0.867 min, 11 scans) Frag=250.0V D06.d...
68767.59
c043
+ESI Scan (rt: 0.701-0.867 min, 11 scans) Frag=250.0V D07.d...
26528.88
53101.61
79583.31
c044
+ESI Scan (rt: 0.714-0.880 min, 11 scans) Frag=250.0V D08.d...
50365.66
75653.29
91470.66
c045
+ESI Scan (rt: 0.700-0.866 min, 11 scans) Frag=250.0V D09.d...
75653.00
51582.25
c046
+ESI Scan (rt: 0.731-0.814 min, 6 scans) Frag=250.0V D10.d...
26544.92
53089.76
79632.60
c047
+ESI Scan (rt: 0.716-0.882 min, 11 scans) Frag=250.0V D11.d...
26568.06
53135.00
79826.87
c048
+ESI Scan (rt: 0.715-0.880 min, 11 scans) Frag=250.0V D12.d...
21045.42
42089.05
62357.54
84108.05
c050
+ESI Scan (rt: 0.701-0.867 min, 11 scans) Frag=250.0V E01.d...
63462.95
50836.31
75652.96
22559.30
c051
+ESI Scan (rt: 0.715-0.881 min, 11 scans) Frag=250.0V E02.d...
26554.22
53106.44
79662.68
20000
40000
60000
80000
Counts vs. Deconvoluted Mass (amu)
c051
+ESI Scan (rt: 0.716-0.881 min, 11 scans) Frag=250.0V E03.d...
26538.99
53025.92
79547.14
c052
+ESI Scan (rt: 0.713-0.879 min, 11 scans) Frag=250.0V E04.d...
26540.79
53082.93
79623.25
c053
+ESI Scan (rt: 0.700-0.866 min, 11 scans) Frag=250.0V E05.d...
26539.41
53080.77
82777.45
c054
+ESI Scan (rt: 0.713-0.878 min, 11 scans) Frag=250.0V E06.d...
26554.14
53106.38
79664.04
c055
+ESI Scan (rt: 0.702-0.868 min, 11 scans) Frag=250.0V E07.d...
26512.84
53025.93
79537.50
c056
+ESI Scan (rt: 0.715-0.881 min, 11 scans) Frag=250.0V E08.d...
26610.00
53219.26
79824.68
c057
+ESI Scan (rt: 0.713-0.879 min, 11 scans) Frag=250.0V E09.d...
26552.40
53102.95
79659.22
c058
+ESI Scan (rt: 0.716-0.882 min, 11 scans) Frag=250.0V E10.d...
26578.65
53157.71
79733.78
c059
+ESI Scan (rt: 0.716-0.881 min, 11 scans) Frag=250.0V E11.d...
26525.46
53049.52
79573.06
c060
+ESI Scan (rt: 0.710-0.876 min, 11 scans) Frag=250.0V E12.d...
26539.48
53079.67
78840.79
20000
40000
60000
80000
Counts vs. Deconvoluted Mass (amu)


c061
+ESI Scan (rt: 0.715-0.881 min, 11 scans) Frag=250.0V F01.d...
26525.60
53049.49
79573.70
c062
+ESI Scan (rt: 0.713-0.879 min, 11 scans) Frag=250.0V F02.d...
26538.55
53077.41
79678.76
c063
+ESI Scan (rt: 0.716-0.881 min, 11 scans) Frag=250.0V F03.d...
26495.06
52987.99
79483.96
c064
+ESI Scan (rt: 0.710-0.876 min, 11 scans) Frag=250.0V F04.d...
26481.52
52961.30
79440.19
c065
+ESI Scan (rt: 0.702-0.867 min, 11 scans) Frag=250.0V F05.d...
26481.54
52960.50
79450.10
c066
+ESI Scan (rt: 0.709-0.874 min, 11 scans) Frag=250.0V F06.d...
26453.13
52950.57
79357.10
c067
+ESI Scan (rt: 0.714-0.879 min, 11 scans) Frag=250.0V F07.d...
26552.43
53104.77
79656.48
c068
+ESI Scan (rt: 0.710-0.875 min, 11 scans) Frag=250.0V F08.d...
26494.34
52986.63
79480.59
c069
+ESI Scan (rt: 0.700-0.866 min, 11 scans) Frag=250.0V F09.d...
26496.10
52989.71
79554.24
c070
+ESI Scan (rt: 0.715-0.881 min, 11 scans) Frag=250.0V F10.d...
26623.72
53246.25
79870.65
20000
40000
60000
80000
Counts vs. Deconvoluted Mass (amu)


c071
+ESI Scan (rt: 0.711-0.876 min, 11 scans) Frag=250.0V F11.d...
26481.75
52962.12
79449.44
c072
+ESI Scan (rt: 0.714-0.879 min, 11 scans) Frag=250.0V F12.d...
26453.65
52950.95
79356.25
c073
+ESI Scan (rt: 0.713-0.879 min, 11 scans) Frag=250.0V G01.d...
26453.71
52906.37
79359.65
c074
+ESI Scan (rt: 0.712-0.878 min, 11 scans) Frag=250.0V G02.d...
26519.83
53038.58
79558.91
c075
+ESI Scan (rt: 0.702-0.868 min, 11 scans) Frag=250.0V G03.d...
26510.83
53020.47
79531.87
c076
+ESI Scan (rt: 0.700-0.882 min, 12 scans) Frag=250.0V G04.d...
26453.18
52904.93
79358.43
c077
+ESI Scan (rt: 0.712-0.878 min, 11 scans) Frag=250.0V G05.d...
26519.59
53084.46
79557.46
c078
+ESI Scan (rt: 0.715-0.880 min, 11 scans) Frag=250.0V G06.d...
26510.27
53019.96
79528.98
c079
+ESI Scan (rt: 0.702-0.868 min, 11 scans) Frag=250.0V G07.d...
26541.80
53082.78
79622.91
c080
+ESI Scan (rt: 0.710-0.876 min, 11 scans) Frag=250.0V G08.d...
26537.86
53076.73
79614.55
20000
40000
60000
80000
Counts vs. Deconvoluted Mass (amu)


c081
+ESI Scan (rt: 0.702-0.867 min, 11 scans) Frag=250.0V G09.d...
26541.06
53082.36
79621.86
c082
+ESI Scan (rt: 0.702-0.868 min, 11 scans) Frag=250.0V G10.d...
26538.71
53078.35
78842.27
c083
+ESI Scan (rt: 0.716-0.882 min, 11 scans) Frag=250.0V G11.d...
26465.92
52931.77
82772.58
c084
+ESI Scan (rt: 0.710-0.875 min, 11 scans) Frag=250.0V G12.d...
26613.06
53224.34
79835.68
c085
+ESI Scan (rt: 0.712-0.877 min, 11 scans) Frag=250.0V H01.d...
26353.48
52705.15
79130.72
c086
+ESI Scan (rt: 0.700-0.866 min, 11 scans) Frag=250.0V H02.d...
26427.61
52815.21
79271.77
c087
+ESI Scan (rt: 0.712-0.878 min, 11 scans) Frag=250.0V H03.d...
26181.21
52360.26
78540.37
c088
+ESI Scan (rt: 0.710-0.876 min, 11 scans) Frag=250.0V H04.d...
26181.05
52360.91
70272.78
94201.89
c089
+ESI Scan (rt: 0.713-0.879 min, 11 scans) Frag=250.0V H05.d...
26087.22
51427.97
78259.18
c090
+ESI Scan (rt: 0.700-0.882 min, 12 scans) Frag=250.0V H06.d...
25828.12
51653.56
77485.13
94202.37
20000
40000
60000
80000
Counts vs. Deconvoluted Mass (amu)

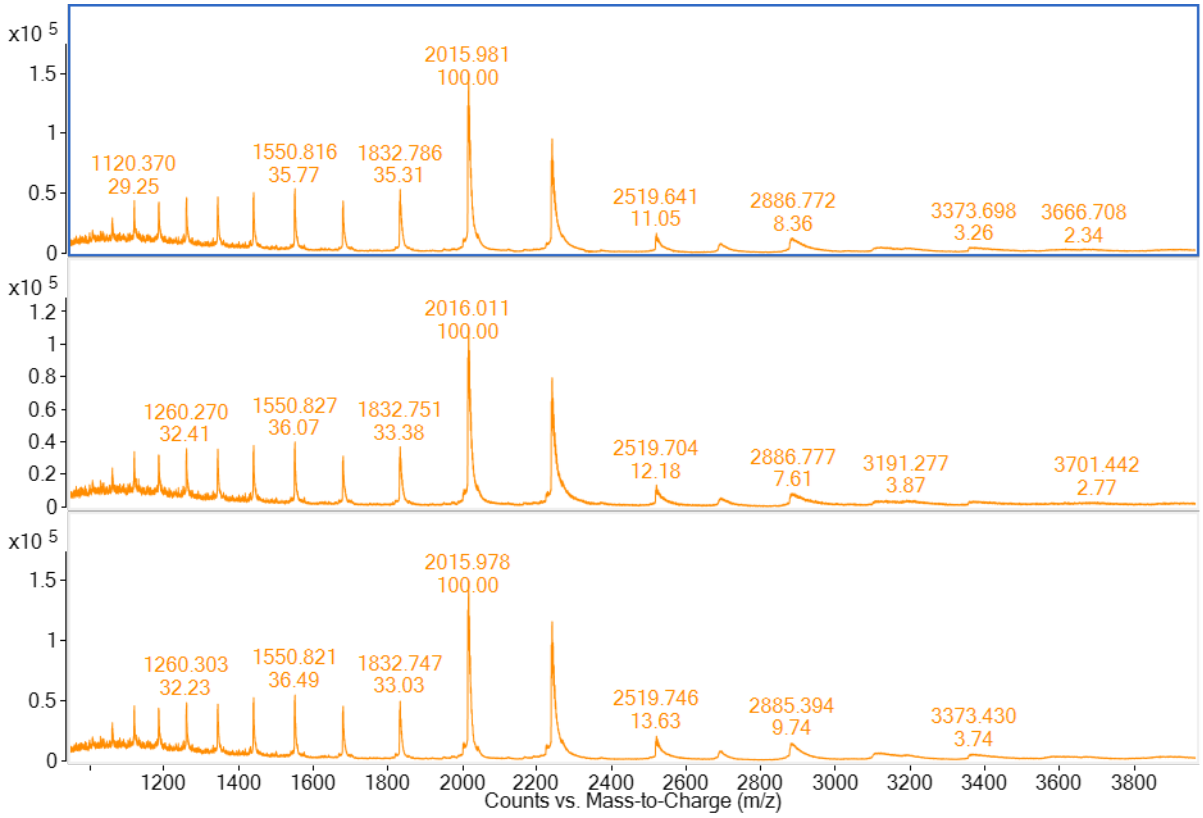
Repeat purifications in large scale of constructs for which no clear ESI-TOF/MS identification could be obtained from test expression (no tag, no N-terminal Met). Construct HMGB1-C1-c038 (D2) was lost upon depyrogenation with SM-2 biobeads and could not be detected in ESI-TOF/MS so identity is suspect, although significant protein could be observed.



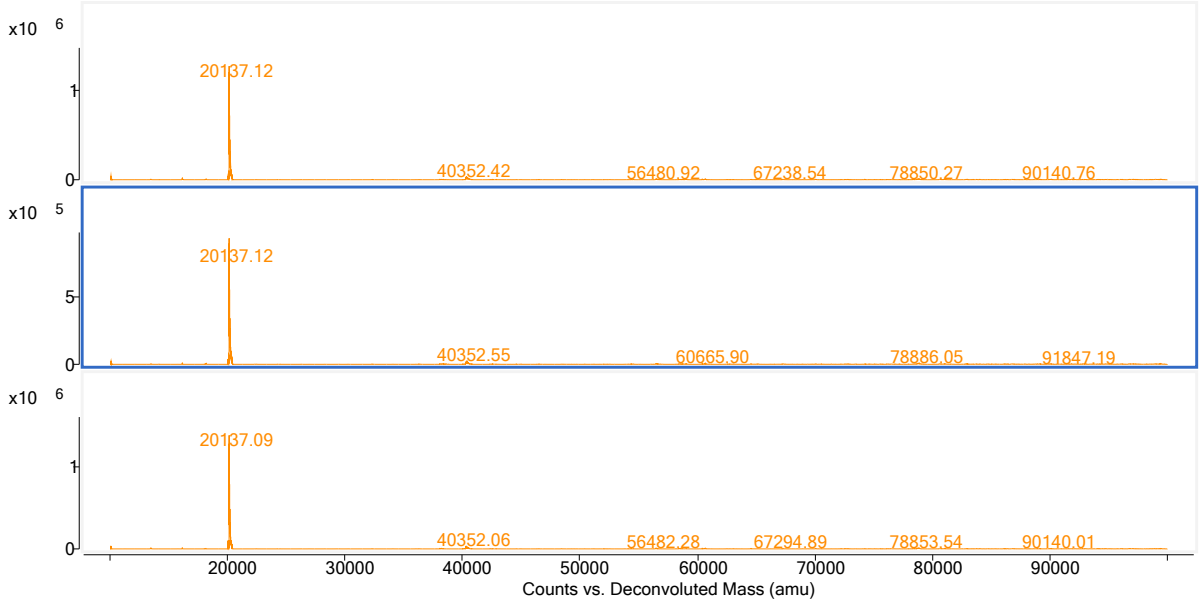
Suppl. data A2-5-B Figure 5.18, pp. 203, all spectra (HMGB1-c048).

Spectra for FR-HMGB1 c001, c002 and c031 can be found in Suppl. data A2-3-D

M/Z spectra



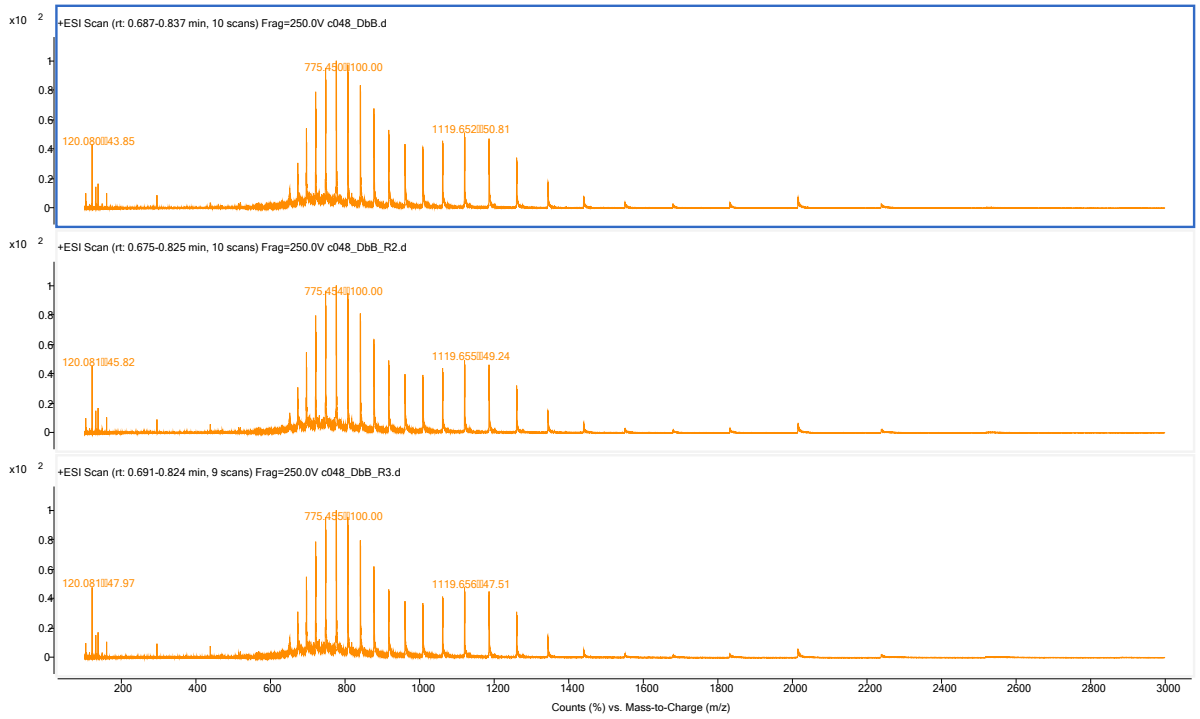
Deconvoluted spectra



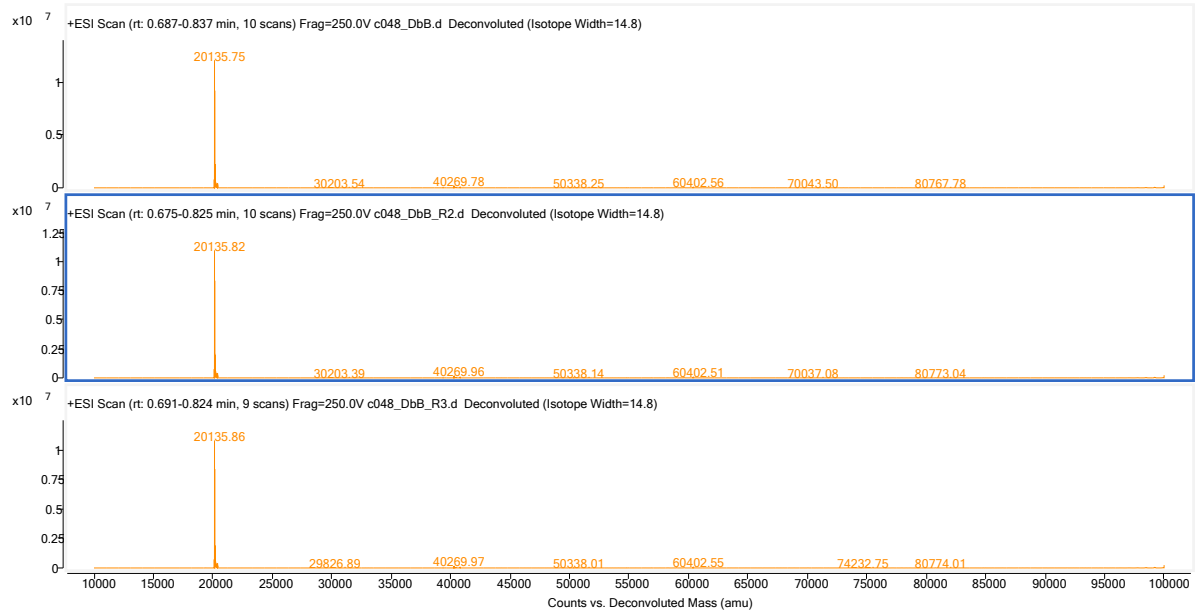
Suppl. data A2-5-C Figure 5.19, pp 203, all spectra (HMGB1-c048).

Spectra for FR-HMGB1 c001, c002 and c031 can be found in Suppl. data A2-3-F.

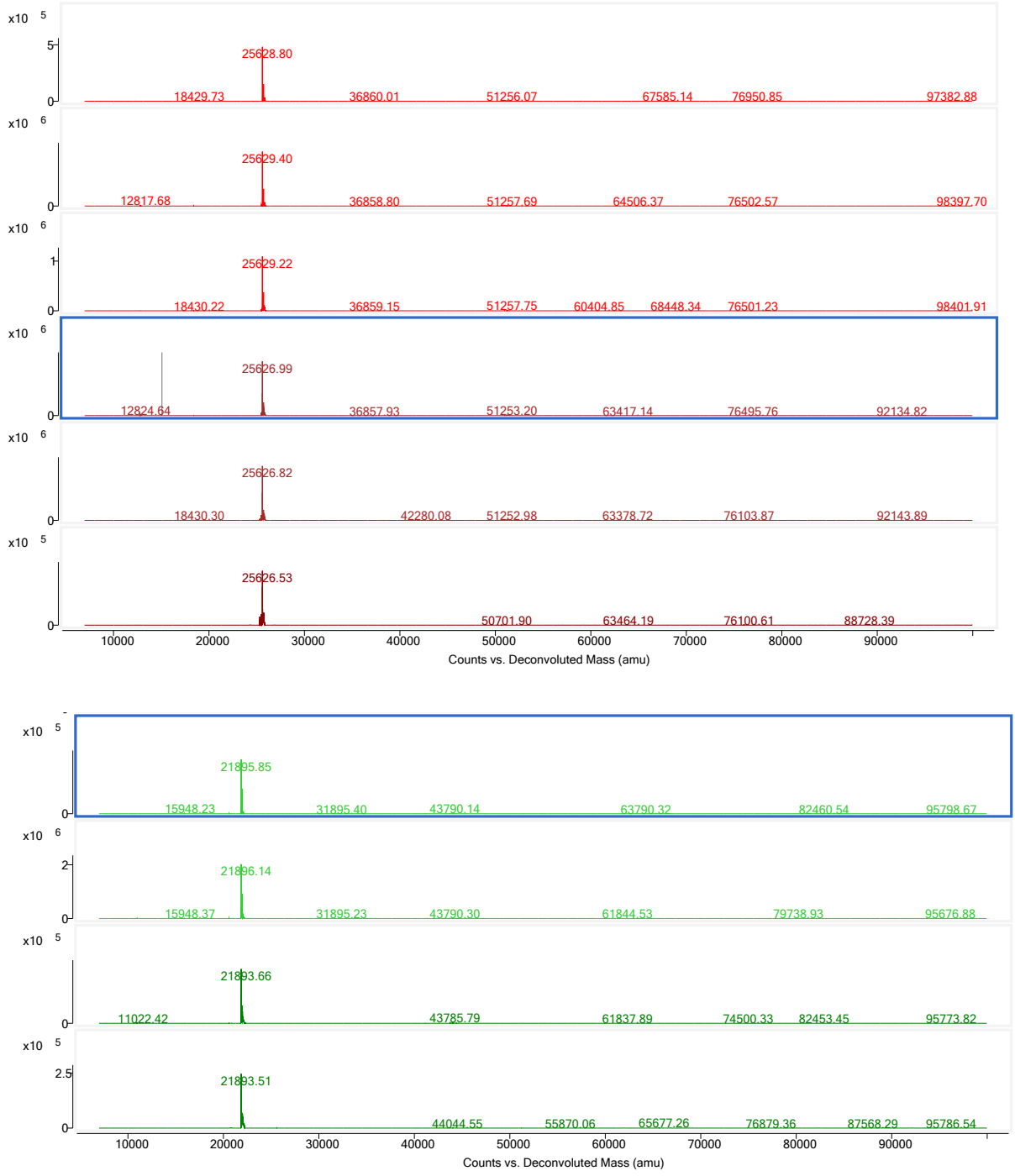
M/Z spectra

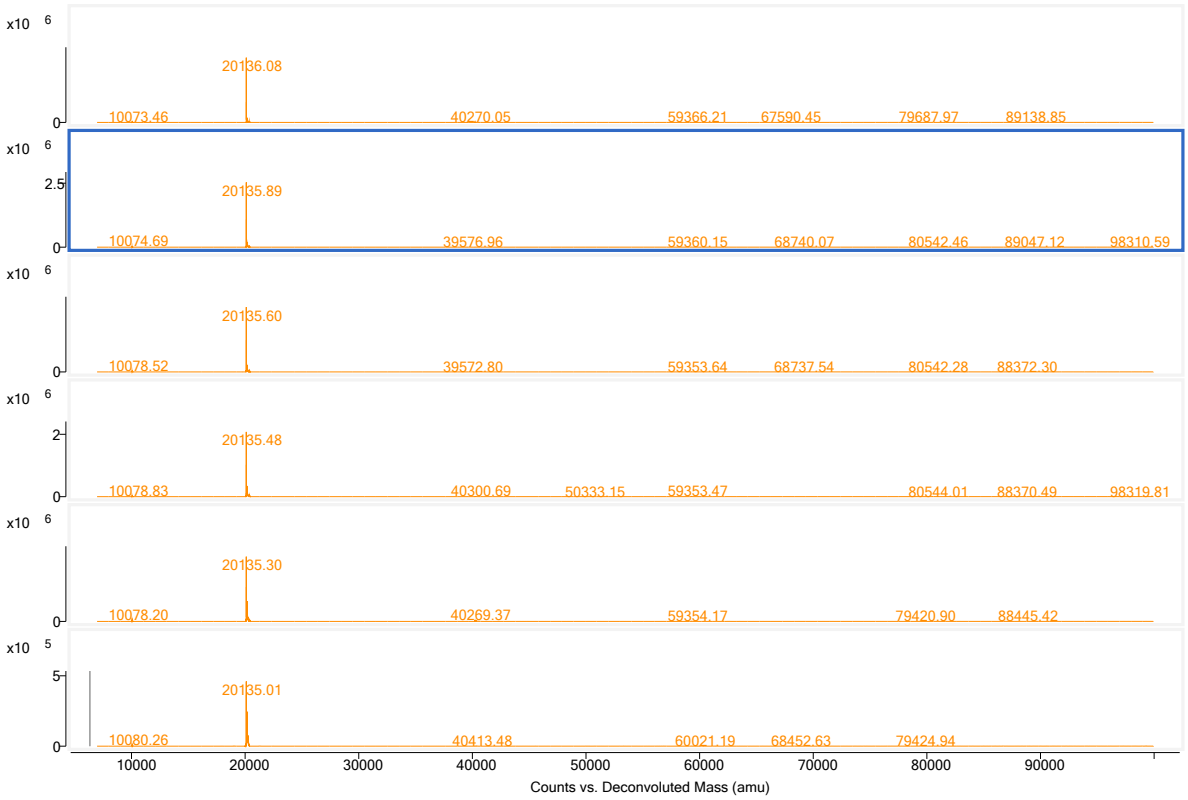


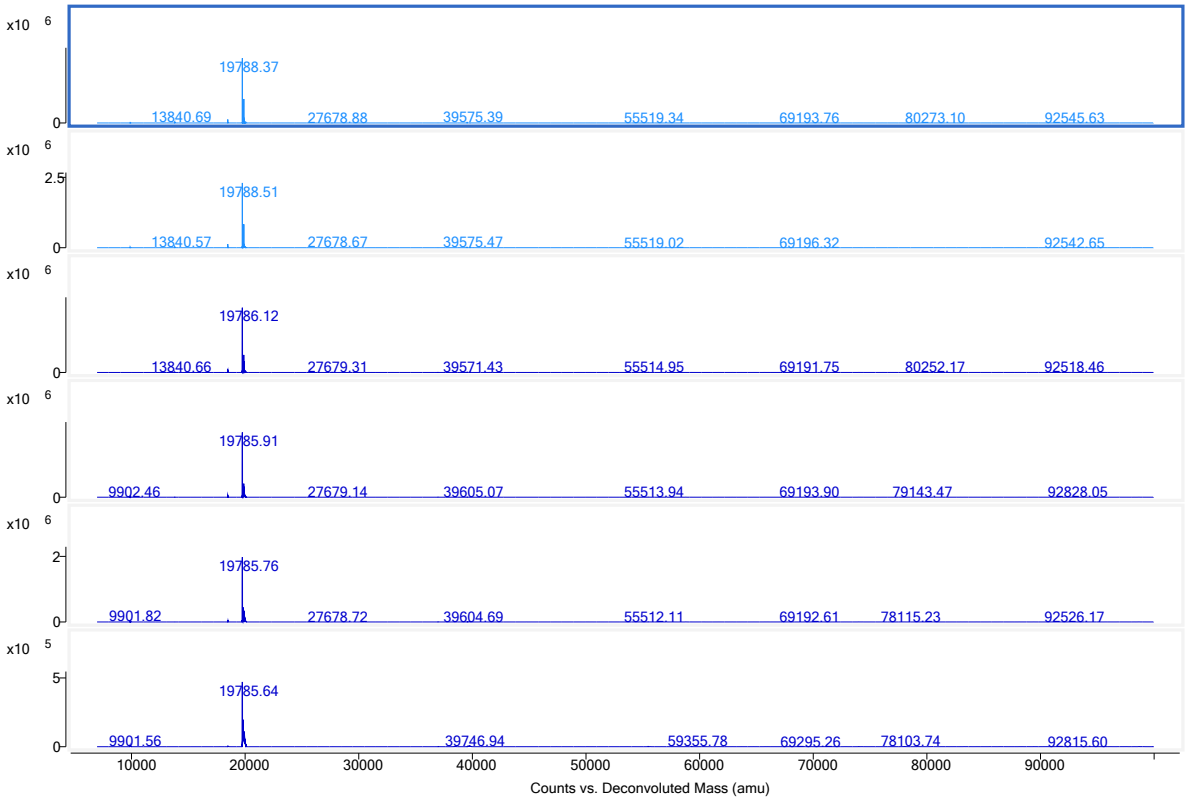
Deconvoluted spectra



Suppl. data A2-5-D Figure 5.21, pp. 209 (full range deconvolution)




x10 6
20136.08
10073.46
40270.05
59366.21
67590.45
79687.97
89138.85
x10 6
2.5
20135.89
10074.69
39576.96
59360.15
68740.07
80542.46
89047.12
98310.59
x10 6
20135.60
10078.52
39572.80
59353.64
68737.54
80542.28
88372.30
x10 6
2
20135.48
10078.83
40300.69
50333.15
59353.47
80544.01
88370.49
98319.81
x10 6
20135.30
10078.20
40269.37
59354.17
79420.90
88445.42
x10 5
5
20135.01
10080.26
40413.48
60021.19
68452.63
79424.94
10000
20000
30000
40000
50000
60000
70000
80000
90000
Counts vs. Deconvoluted Mass (amu)

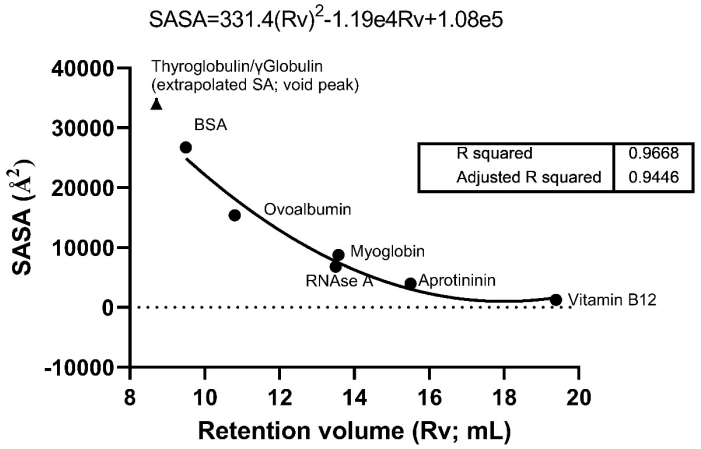

x10 6
19788.37
13840.69
27678.88
39575.39
55519.34
69193.76
80273.10
92545.63
x10 6
2.5
19788.51
13840.57
27678.67
39575.47
55519.02
69196.32
92542.65
x10 6
19786.12
13840.66
27679.31
39571.43
55514.95
69191.75
80252.17
92518.46
x10 6
19785.91
9902.46
27679.14
39605.07
55513.94
69193.90
79143.47
92828.05
x10 6
2
19785.76
9901.82
27678.72
39604.69
55512.11
69192.61
78115.23
92526.17
x10 5
5
19785.64
9901.56
39746.94
59355.78
69295.26
78103.74
92815.60
10000
20000
30000
40000
50000
60000
70000
80000
90000
Counts vs. Deconvoluted Mass (amu)

Appendix A3 Supplemental data

Chapter 3 Figure supplements

Suppl. data A3-1-A Function test of Superdex S75 column

A.



B.

PDB structure	BSA, 3V03	Ovalbumin, 1JTI	*Myoglobin, 2V1I	RNAseA, 1A5P	Aprotinin, 1NAG	**Vitamin B12, 3BUL
Rv (mL)	9.5	10.8	13.5	15.5	13.575	19.39
SASA (Å²)	26727	15384	8761	6820	3963	1290
mW (Da)	63095	43116	16894	15882	6316	1330

Figure 8.1: Function test of Superdex 75 Increase, 10/300

A: Standard curve interpolation of molecular weight standards eluted from the SEC column to solvent accessible surface area of each protein. **B**. SASA calculations for protein monomers from PDB structures. * The SASA calculation for myoglobin was performed including the heme group as the sample was visibly red. ** The structure of vitamin B12 was extracted from the referenced PDB file and calculated separately to exclude interference from the ligand. Void peak eluted at 8.71 mL; theoretical SASA of this peak would be 34000 Å^2 (exclusion limit)

Suppl. data A3-1-B Figure 3.16, pp. 92

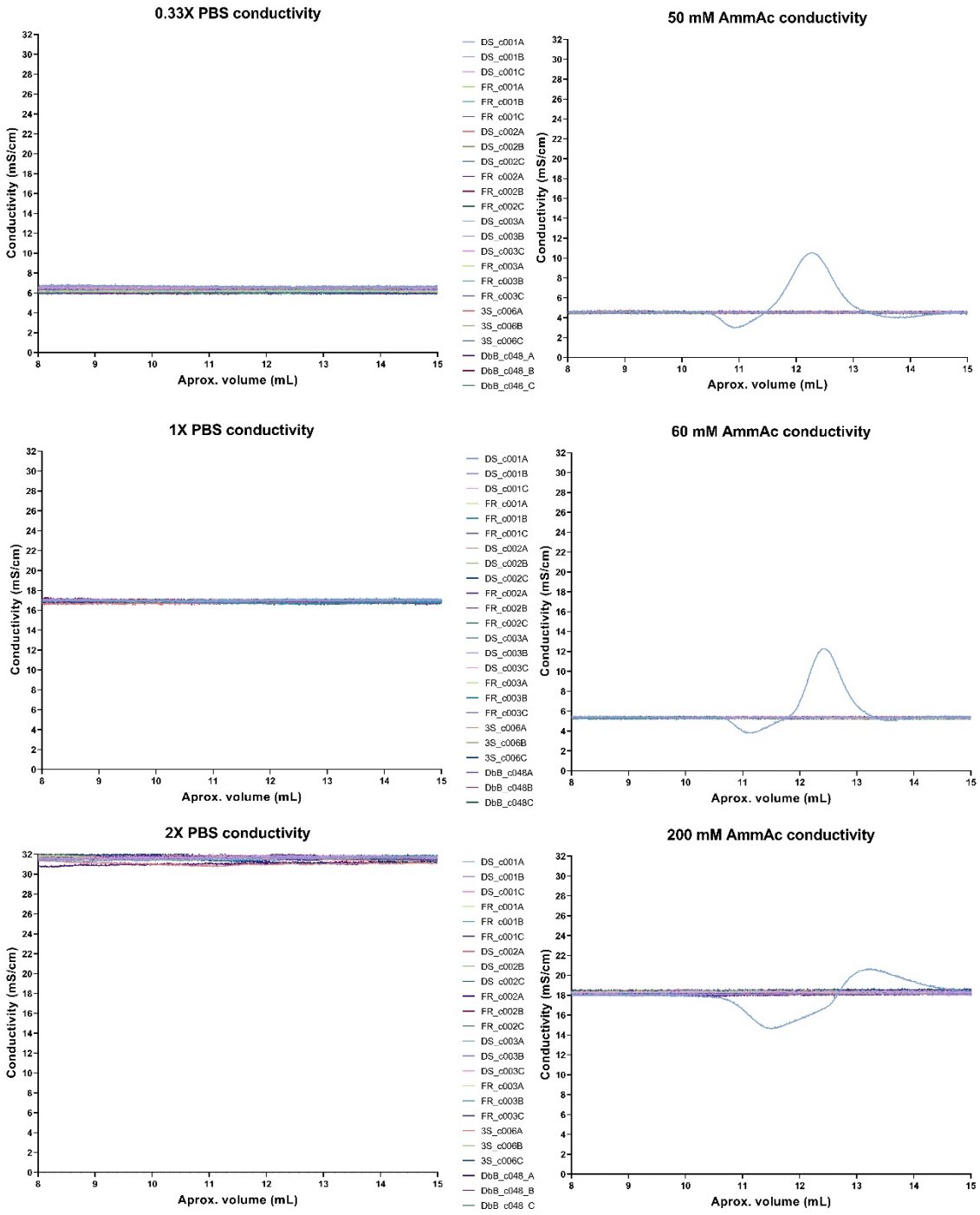


Figure 8.2: Conductivity traces and sample exclusion criteria

For the ammonium acetate runs, the first chromatogram for DS-HMGB1-c001 (1-214, light blue) had to be excluded from analysis due to unacceptable variations in conductivity during the run, which overlapped with the peak for HMGB1

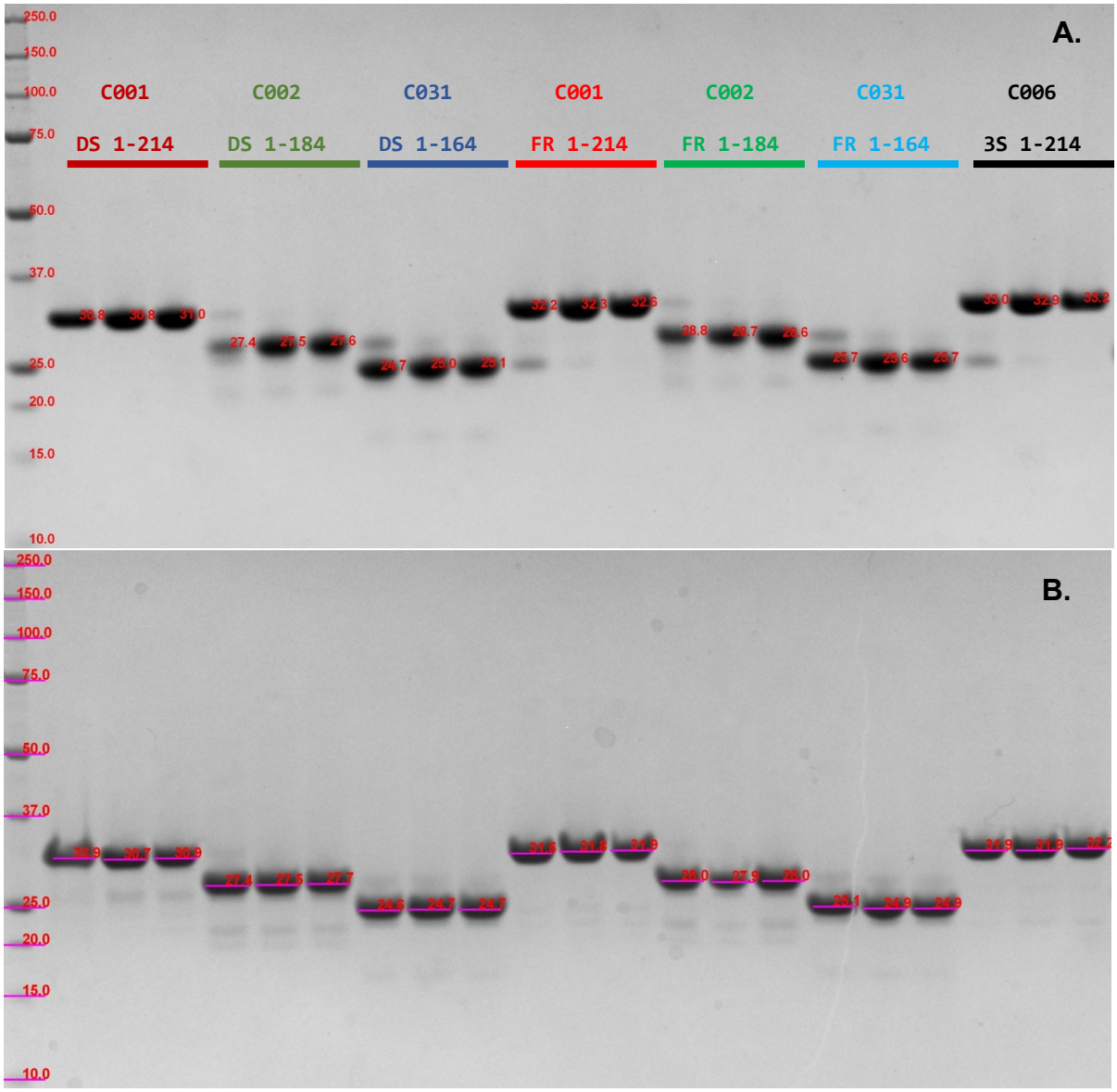


Figure 8.3: Nonreducing denaturing SDS-PAGE of HMGB1 constructs in pH 6.8 ammonium acetate (Figure 3.19)

Non-reducing, denaturing SDS-PAGE of peak fractions of the SEC runs described in Figure 3.19. HMGB1A-c001 (FL, red), c002 (1-184, green), c031 (1-164, blue), both reduced and oxidised (darker hue), and HMGB1-c006 (3S, FL, black). Fractions are loaded in the sequence in which chromatography was performed; this was a requirement due to the different redox states to minimise cross-contamination via sample buffer. MW of each band has been extrapolated from the standard; disulfide-HMGB1 bands have slightly faster migration owing to the more compact shape of linearised peptides in SDS as the disulfide bonds are not broken. **A.** 50 mM ammonium acetate. There is minor carryover from samples due to nonspecific binding of HMGB1 to the SEC column, particularly for tail-less constructs (c002/c031); this was also observable due to a back-pressure increase after these and also observed in 0.33X PBS. **B.** 200 mM ammonium acetate.

Suppl. data A3-1-C Preparative SEC runs

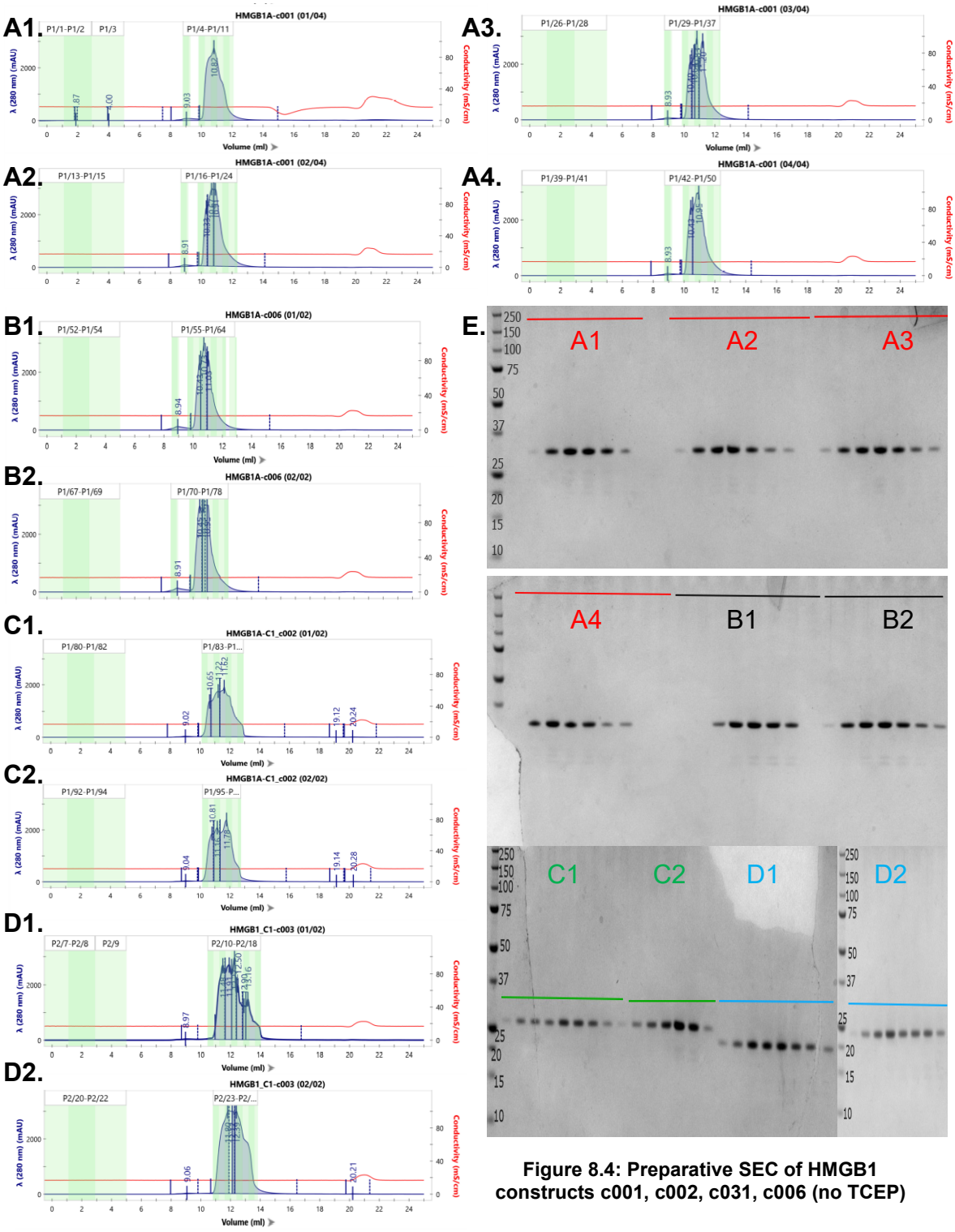


Figure 8.4: Preparative SEC of HMGB1 constructs c001, c002, c031, c006 (no TCEP)

SEC chromatograms for 0.3 mL injected into a Superdex 75 10/300 column of HMGB1A-c001 2 mM (**A1-A4**), HMGB1A-c006 2.2 mM (**B1-B2**), HMGB1A-c002 2.5 mM (**C1-C2**) and HMGB1A-c031 3 mM (**D1-D2**). **E.** SDS-PAGE gels of the peak fractions (0.5 µg protein at peak load, stained with colloidal Coomassie (1-4 ng/band sensitivity). Each peak fraction set has been labelled with the parent chromatogram letter and number and the colour coding used in Figure 3.19.

Suppl. data A3-1-D Figure 3.23, pp108; DSF curves for HMGB1 c002 and c031 (rejected)

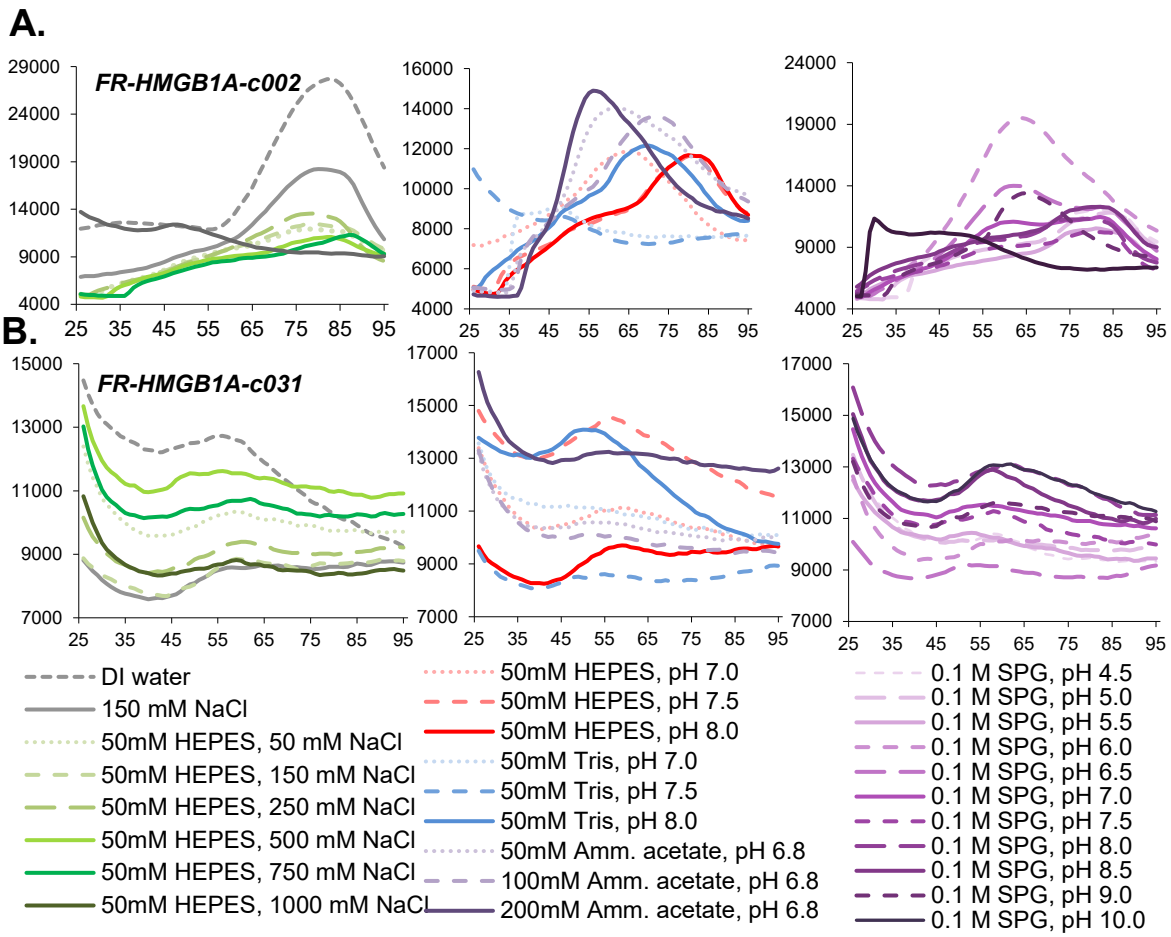


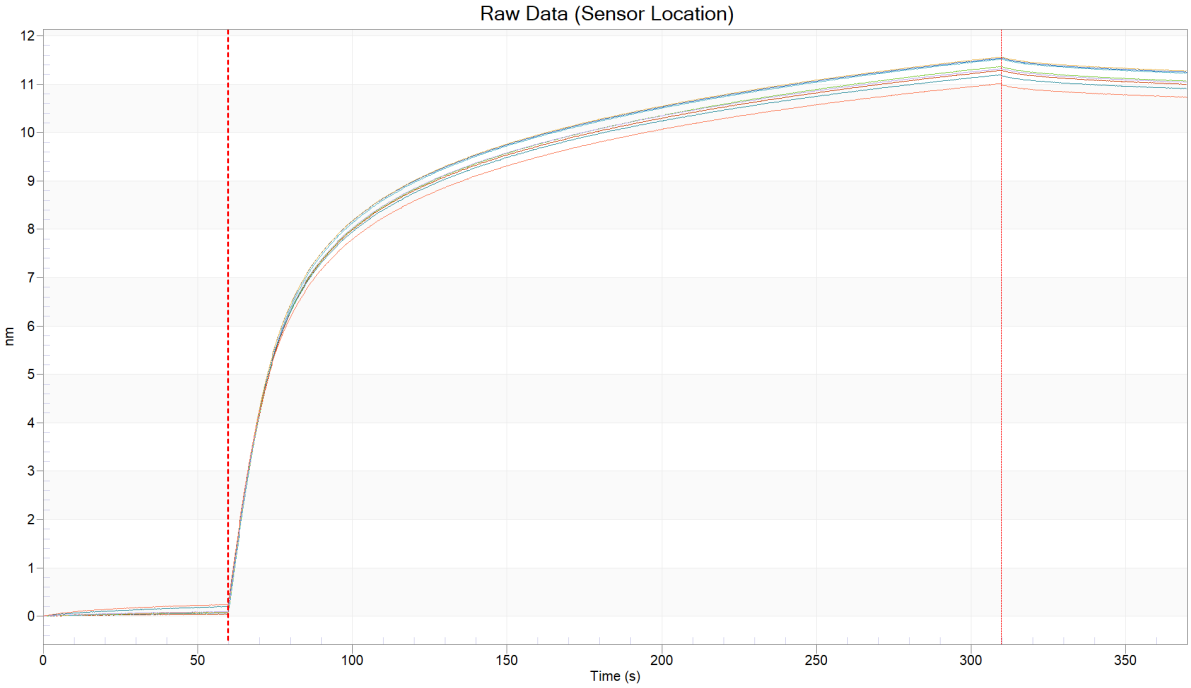
Figure 8.5: DSF melting curve profiles and Tm_{50} analysis of HMGB1A constructs

A: Thermal denaturation profiles for FR-HMGB1 c002 in my custom screen; conditions have been coloured in the same code as shown in Table 3.4. Each row of graphs corresponds to the listed construct; all experiments performed with 8 µM protein (desalted in 10 mM HEPES pH 7.5) and 1X SyPro Orange. **B:** Thermal denaturation profiles for FR-HMGB1 c031

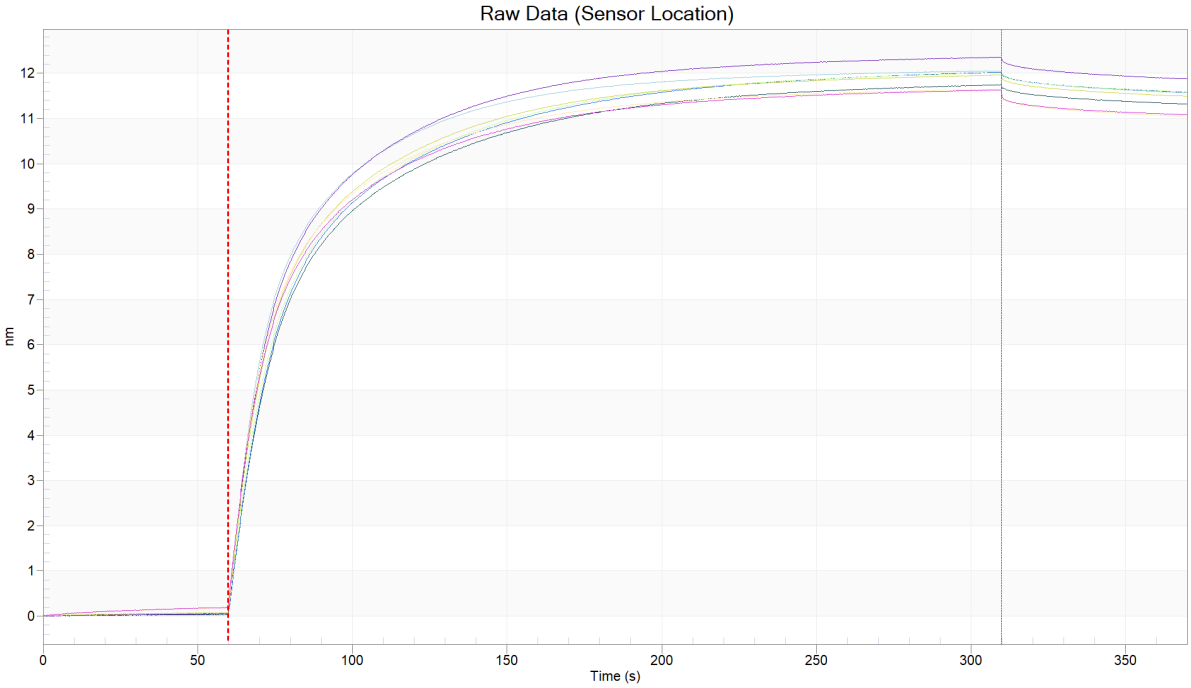
Chapter 4 Figure supplements

Suppl. data A3-1-E Figure 4.5, pp. 144, sensor coating interferogram

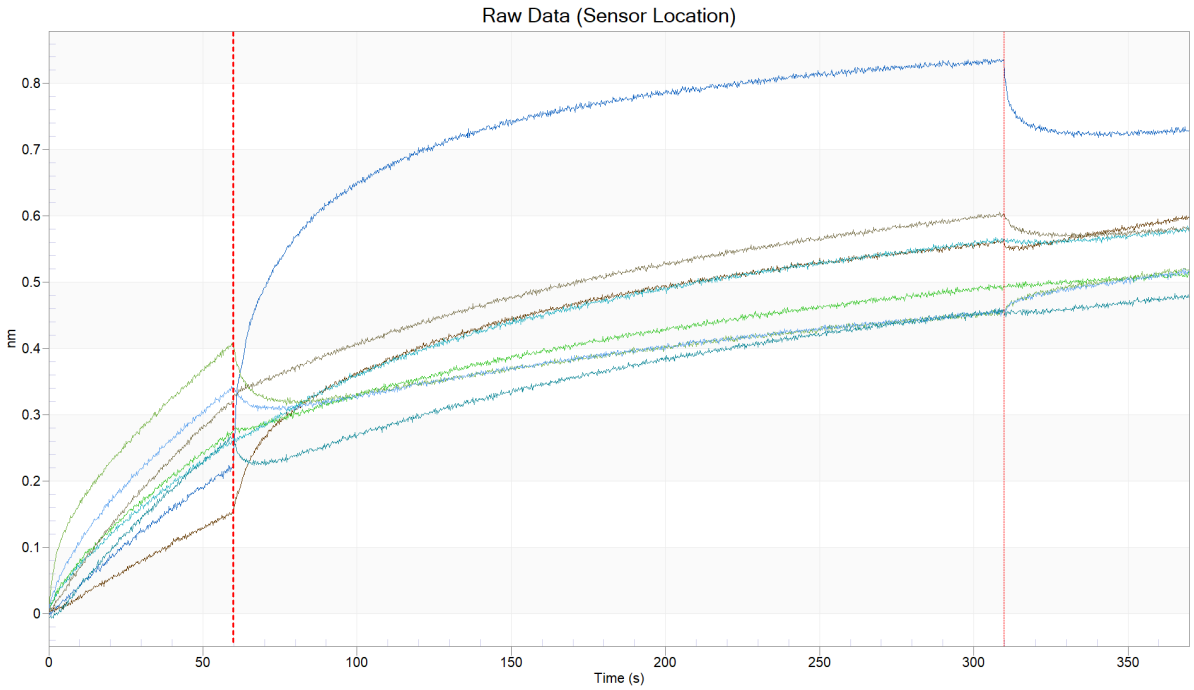
HMGB1A-c011 (FR, full-length, biotinylated)



HMGB1A-c022 (3S, full-length, biotinylated)

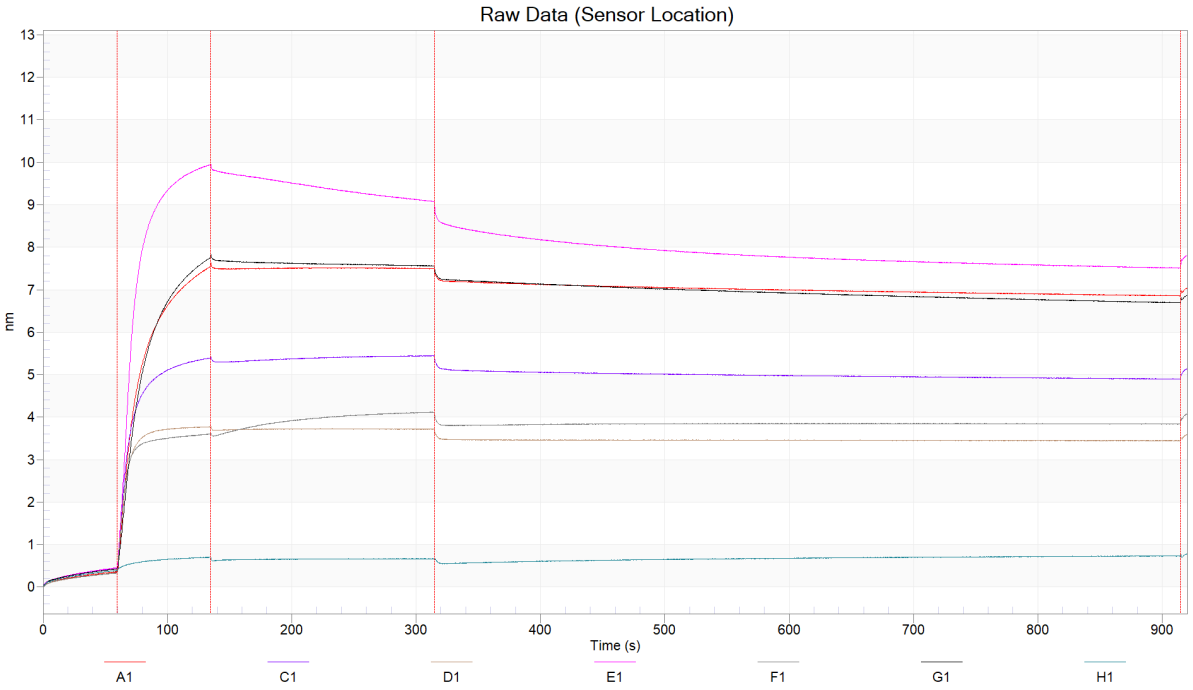


Reference row

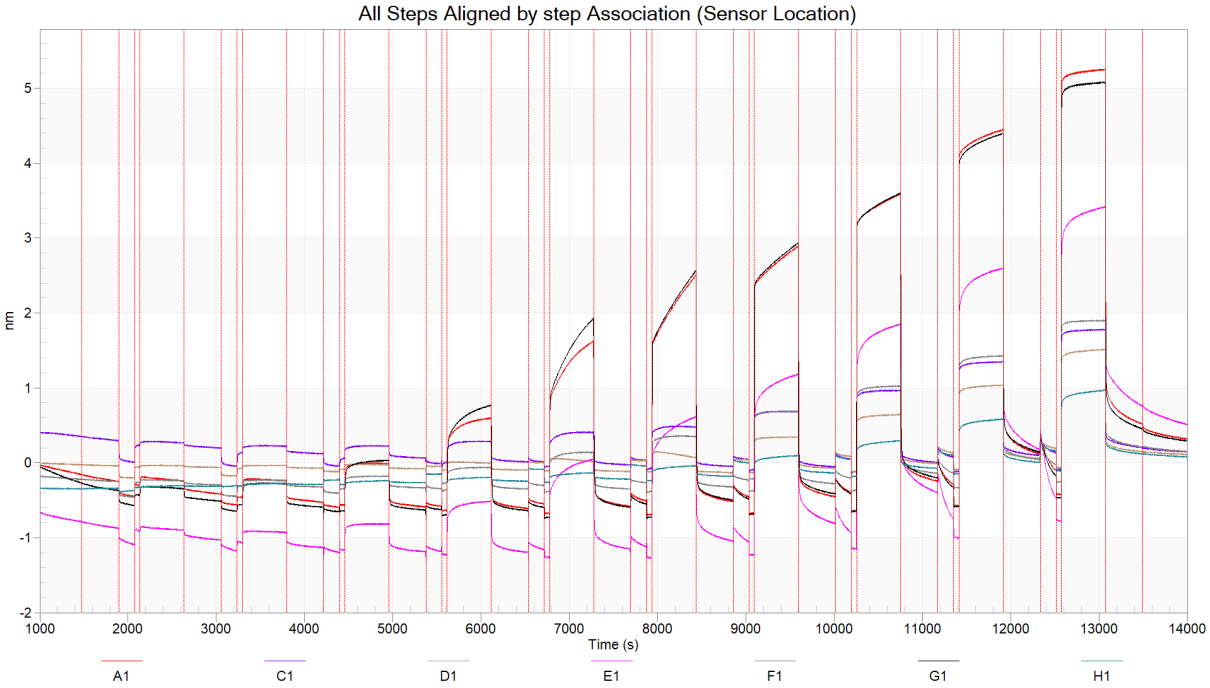


Suppl. data A3-1-F Figure 4.11, pp. 159, sensor coating interferogram and raw traces

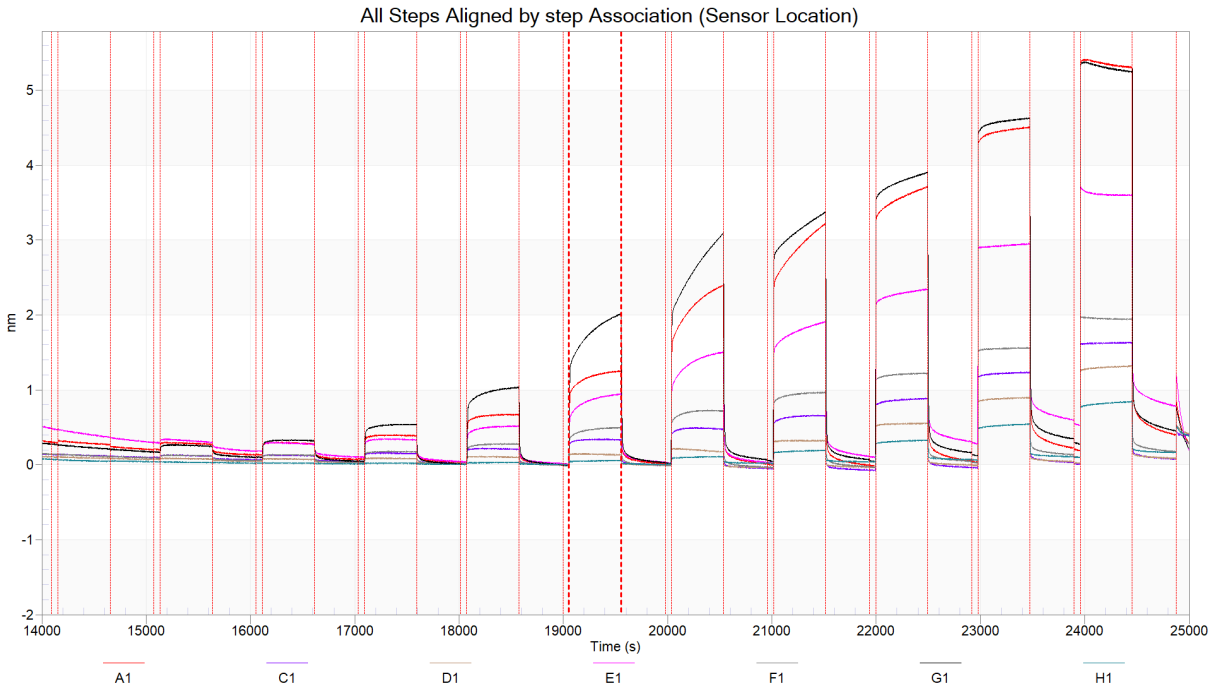
Coating interferogram; note that HMGB1A-c027/c037 gives a much higher coating response than HMGB1A-c028/c038, despite all constructs being in equimolar concentrations in the solution (4 µM) and all 4 constructs being of nearly equal mW. Blank sensor is shown in green.



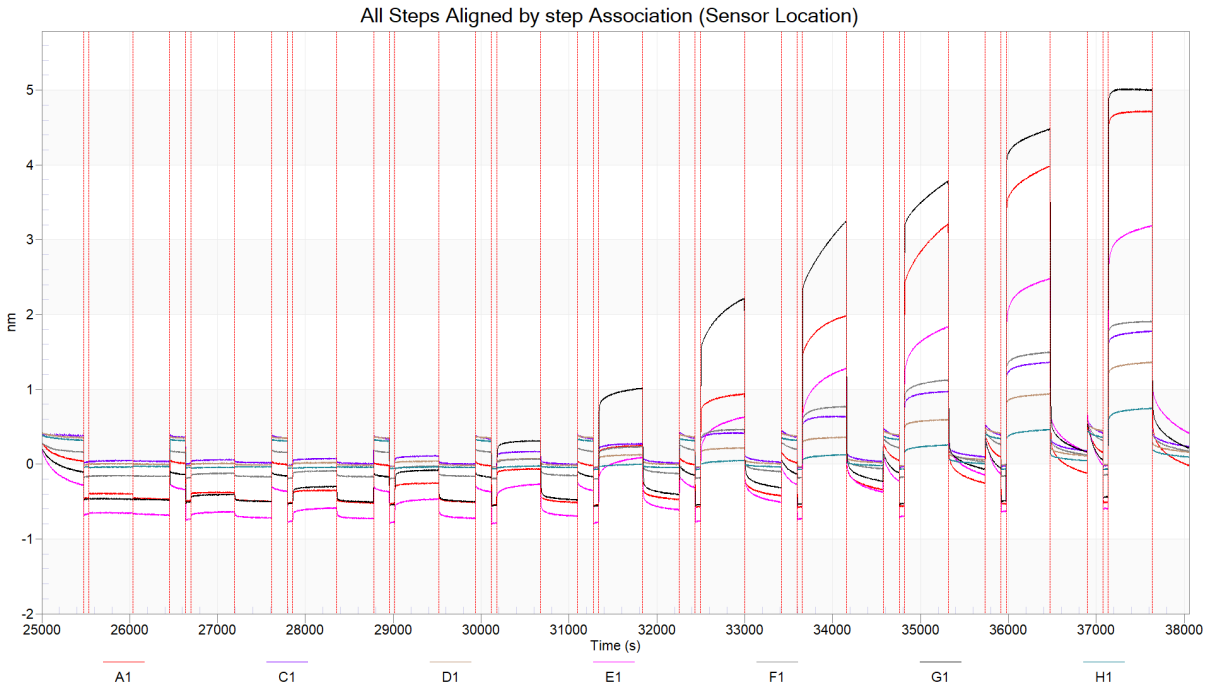
Cycle 1 interferogram



Cycle 2 interferogram



Cycle 3 interferogram



Appendix A4 NMR tables

Chapter 4 NMR spectra (Figure 4.8/Figure 4.9)

Suppl. data A4-1-A HMGB1-c028 (94-162, biotinylated) titration with CXCL12A-c021 at 0, 0.42, 0.82 and 1.42 molar equivalents.

Due to protein amount limitations, titration was performed by sequential addition of CXCL12 to HMGB1 samples, resulting in sample dilution. As the calculations in the chemical shift tracking module in CCPNMR are independent of peak, this does not alter the results; the median change has been indicated in the volume comparisons.

	BRUKER AVIII HD 500, 5 MM CPTCI 1H-13C/15N/D Z-GRD PROBE
^{1}H frequency (Hz)	500.01233645
^{1}H sweep width (ppm)	11.9044681870088
^{1}H sweep width (Hz)	5952.38095238095
^{1}HN sweep width (ppm)	10.9286921061064
^{15}N frequency (Hz)	50.6715211382702
^{15}N sweep width (ppm)	32.0372585410812
^{15}N sweep width (Hz)	1623.37662337662
Water suppression gradient pulse width (Hz)	2336.45
Buffer	10 mM HEPES, pH 7.5, 150 mM NaCl

Baseline 1 (Day 1)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.45 mM, HMGB1-c028 (Box B 94-162), C-terminal Avi tag
CXCL12 construct and concentration	0 mM
Temperature	25 ºC
Volume	300 µL
Sample pH (after adjusting)	7.7
Number of scans	8

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.95.LYS.N	8.47816	122.87351	16.67468	12.22962	108864312
BoxBFL.96.ARG.N	8.30965	123.9316	14.96291	11.62421	144139104
BoxBFL.99.SER.N	7.95885	114.22726	22.5074	15.03176	16802856
BoxBFL.100.ALA.N	9.0643	122.83852	12.15917	12.63149	138071536
BoxBFL.101.PHE.N	8.44324	116.14361	13.00283	12.61897	92877416
BoxBFL.102.PHE.N	8.26304	122.13933	14.85477	12.7452	88662424
BoxBFL.103.LEU.N	8.27748	121.1555	13.77712	12.44091	112430712
BoxBFL.104.PHE.N	8.15422	122.51526	17.30965	12.77095	208687344
BoxBFL.105.CYS.N	8.49978	118.38542	12.79441	13.0002	114087696
BoxBFL.106.SER.N	8.12921	115.05247	11.81293	12.01762	185743392
BoxBFL.107.GLU.N	7.15444	119.18761	15.14408	11.95417	174642688
BoxBFL.108.TYR.N	7.944	114.61662	17.02596	12.46204	118306368
BoxBFL.109.ARG.N	9.23012	123.69626	13.5621	12.34021	130691488
BoxBFL.111.LYS.N	6.80053	117.70067	15.27665	12.39016	143073536
BoxBFL.112.ILE.N	8.1617	119.74645	14.25943	12.83013	126874280
BoxBFL.113.LYS.N	8.24755	118.47981	17.01044	13.8802	71528184
BoxBFL.114.GLY.N	7.6363	103.71903	15.16014	12.02905	120973232
BoxBFL.115.GLU.N	7.50091	119.58688	15.18257	12.11395	146458640
BoxBFL.116.HIS.N	7.8198	114.78302	14.39534	11.92497	117505296
BoxBFL.119.LEU.N	7.36285	121.19842	12.72385	11.91713	222764784
BoxBFL.120.SER.N	9.31614	121.05943	13.74707	12.14133	126779632
BoxBFL.121.ILE.N	8.7484	120.66613	14.38154	12.69317	61496764
BoxBFL.122.GLY.N	8.70938	109.44017	17.19009	12.98604	15080434
BoxBFL.123.ASP.N	7.96042	124.28299	14.86623	12.46038	123154200
BoxBFL.124.VAL.N	8.63541	124.07082	13.35656	11.92144	172075232
BoxBFL.125.ALA.N	7.92766	121.34983	12.19988	11.82723	225860864
BoxBFL.126.LYS.N	8.01853	119.68738	12.11349	12.34419	232838720
BoxBFL.127.LYS.N	8.03865	121.30632	12.54305	12.01855	257706416
BoxBFL.128.LEU.N	8.6442	120.16728	12.9965	12.28793	158055184
BoxBFL.129.GLY.N	8.26246	106.03693	14.92945	11.88241	144422592
BoxBFL.130.GLU.N	8.07143	123.2476	13.07349	12.04715	202424016
BoxBFL.131.MET.N	8.77019	118.9603	12.65128	12.17389	190716976
BoxBFL.132.TRP.N	8.7155	122.68159	13.66438	12.0197	125464952
BoxBFL.133.ASN.N	8.17909	117.36655	11.90513	11.83855	191488080
BoxBFL.134.ASN.N	7.58752	115.82865	19.62284	12.10783	166835600
BoxBFL.135.THR.N	7.30723	119.95451	15.53231	12.28047	175465168
BoxBFL.136.ALA.N	9.24925	131.44171	12.05279	11.8376	183664944
BoxBFL.138.ASP.N	8.99204	115.70293	13.82959	13.53115	66696728
BoxBFL.139.ASP.N	7.33843	117.87656	17.35284	12.00825	172632800
BoxBFL.140.LYS.N	7.8113	119.11469	15.11934	12.32708	144059264
BoxBFL.141.GLN.N	7.42681	118.19987	12.23114	12.40879	203757632
BoxBFL.143.TYR.N	7.21691	115.66054	15.9391	12.2	135669520
BoxBFL.144.GLU.N	8.18987	120.68771	12.95833	12.4239	148173872
BoxBFL.145.LYS.N	9.18301	121.23245	13.19779	12.25029	171717696
BoxBFL.146.LYS.N, BoxBFL.156.LYS.N	7.69692	120.57372	12.71031	12.47383	196874368
BoxBFL.147.ALA.N	8.48249	121.09351	12.35114	12.39023	191332288
BoxBFL.148.ALA.N	8.36965	122.00069	12.19025	12.15233	220547328

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.149.LYS.N	7.97224	120.72971	13.19222	12.42346	213554816
BoxBFL.150.LEU.N	8.28652	120.07973	14.56355	13.14455	145004448
BoxBFL.151.LYS.N	8.42411	123.51065	12.54219	12.56125	167598160
BoxBFL.152.GLU.N	8.08663	119.9904	12.53193	12.7395	223413808
BoxBFL.153.LYS.N	7.75259	119.74852	13.195	12.42426	187833136
BoxBFL.154.TYR.N	8.13041	120.47623	13.12379	12.69255	159451584
BoxBFL.155.GLU.N	8.52484	117.50521	13.60921	12.71924	154408976
BoxBFL.157.ASP.N	8.85459	123.10409	12.45843	12.23066	174923680
BoxBFL.158.ILE.N	9.0667	122.28071	13.73538	12.53844	157092272
BoxBFL.159.ALA.N	7.45046	124.74939	12.51372	12.04599	200120928
BoxBFL.160.ALA.N	7.74386	120.563	14.38809	12.29212	198915568
BoxBFL.161.TYR.N	8.20147	120.05331	14.39746	13.18223	167227632
BoxBFL.162.ARG.N	8.26383	118.79358	13.15573	12.41703	173894048

3D experiments (Day 1)

These spectra are not shown due to their 3D nature; data available if requested. Sample is that of Baseline 1.

Experiment type	**3D, ^{15}N EDITED ^{1}H-^{1}H TOCSY-HSQC, THROUGH-SPACE INTERACTION**
Temperature	25 ºC
Volume	300 µL
Number of scans	12

Experiment type	**3D, ^{15}N EDITED ^{1}H-^{1}H NOESY-HSQC, THROUGH-BOND INTERACTION**
Temperature	25 ºC
Volume	300 µL
Number of scans	16

0.42 molar ratio CXCL12/HMG Box (Day 2)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.38 mM, U-15N HMGB1A-c028 (Box B 94-162), C-terminal biotinylation tag (300 µL of 0.45 mM stock)
CXCL12 construct and concentration	0.162 mM, CXCL12A-c021, wt 1-67, no tag (50 µL of 1.14 mM stock)
Volume	350 µL
Sample pH (after adjusting)	7.71
Temperature	25 ºC
Number of scans	8

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.95.LYS.N	8.47806	122.8889	15.74549	12.59893	1.04E+08
BoxBFL.96.ARG.N	8.31955	123.9958	15.32754	11.79906	1.23E+08
BoxBFL.99.SER.N	7.95975	114.213	18.91249	16.00748	19314842
BoxBFL.100.ALA.N	9.06482	122.8403	12.12244	12.51287	1.23E+08
BoxBFL.101.PHE.N	8.44328	116.1449	14.02058	12.94259	83270120
BoxBFL.102.PHE.N	8.2626	122.1026	14.57979	12.61002	83484184
BoxBFL.103.LEU.N	8.27876	121.1543	13.21545	12.41884	98850448
BoxBFL.104.PHE.N	8.1723	122.6037	17.85543	16.70344	1.63E+08
BoxBFL.105.CYS.N	8.49894	118.3888	12.67018	13.00159	1.03E+08
BoxBFL.106.SER.N	8.12904	115.0558	11.922	12.17556	1.66E+08
BoxBFL.107.GLU.N	7.15246	119.1841	15.05754	12.02976	1.54E+08
BoxBFL.108.TYR.N	7.94272	114.6176	17.11915	12.55614	1.04E+08
BoxBFL.109.ARG.N	9.23036	123.7013	13.06226	12.53714	1.15E+08
BoxBFL.111.LYS.N	6.79969	117.703	15.74346	12.62776	1.25E+08
BoxBFL.112.ILE.N	8.1614	119.7481	15.19044	12.89975	1.06E+08
BoxBFL.113.LYS.N	8.25273	118.4894	16.03035	15.20607	71142304
BoxBFL.114.GLY.N	7.63597	103.7241	14.57231	12.02882	1.1E+08
BoxBFL.115.GLU.N	7.49968	119.5819	15.16195	12.5197	1.35E+08
BoxBFL.116.HIS.N	7.82235	114.787	15.19416	11.98251	1.02E+08
BoxBFL.119.LEU.N	7.3624	121.2015	12.53706	12.04396	2.03E+08
BoxBFL.120.SER.N	9.31525	121.061	13.81227	12.2652	1.13E+08
BoxBFL.121.ILE.N	8.74925	120.6769	12.5753	12.7098	61919540
BoxBFL.122.GLY.N	8.71207	109.4584	25.89949	13.91125	16954762
BoxBFL.123.ASP.N	7.95784	124.2831	14.47967	12.59113	1.13E+08
BoxBFL.124.VAL.N	8.63489	124.0688	13.34502	12.07557	1.51E+08
BoxBFL.125.ALA.N	7.92846	121.3424	11.96993	12.06949	1.99E+08
BoxBFL.126.LYS.N	8.01973	119.6945	12.24395	12.62986	2.08E+08
BoxBFL.127.LYS.N	8.03911	121.3059	12.17418	12.02556	2.24E+08
BoxBFL.128.LEU.N	8.64346	120.1657	12.73252	12.33711	1.36E+08
BoxBFL.129.GLY.N	8.26155	106.0419	14.87585	12.01202	1.26E+08
BoxBFL.130.GLU.N	8.07094	123.248	12.8194	12.1305	1.81E+08
BoxBFL.131.MET.N	8.77033	118.9628	12.47681	12.18055	1.65E+08
BoxBFL.132.TRP.N	8.71521	122.6821	13.43252	12.15419	1.09E+08
BoxBFL.133.ASN.N	8.17879	117.3654	11.81466	11.94108	1.72E+08
BoxBFL.134.ASN.N	7.58738	115.8377	20.34615	11.95454	1.47E+08
BoxBFL.135.THR.N	7.30659	119.9511	15.61869	12.3645	1.57E+08
BoxBFL.136.ALA.N	9.2484	131.4382	11.79903	12.0221	1.7E+08
BoxBFL.138.ASP.N	8.99068	115.7088	13.61787	13.15666	66921064
BoxBFL.139.ASP.N	7.33797	117.8781	18.54773	12.14095	1.52E+08
BoxBFL.140.LYS.N	7.81083	119.1164	14.67427	12.45762	1.24E+08
BoxBFL.141.GLN.N	7.42677	118.2031	12.06407	12.71254	1.77E+08
BoxBFL.143.TYR.N	7.21631	115.6619	15.45741	12.35014	1.19E+08
BoxBFL.144.GLU.N	8.18784	120.6799	13.13889	12.30444	1.27E+08
BoxBFL.145.LYS.N	9.183	121.238	13.17983	12.43273	1.47E+08

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.146.LYS.N, BoxBFL.156.LYS.N	7.69709	120.5754	12.94948	12.65531	1.73E+08
BoxBFL.147.ALA.N	8.48271	121.0951	12.06422	12.53668	1.63E+08
BoxBFL.148.ALA.N	8.37014	122.0096	12.01385	12.60817	1.9E+08
BoxBFL.149.LYS.N	7.97339	120.7383	13.42094	12.64648	1.82E+08
BoxBFL.150.LEU.N	8.28676	120.0845	14.4103	13.05275	1.22E+08
BoxBFL.151.LYS.N	8.42457	123.5222	12.66033	12.96606	1.42E+08
BoxBFL.152.GLU.N	8.08704	119.9998	12.5224	13.21315	1.98E+08
BoxBFL.153.LYS.N	7.75272	119.7489	13.27981	12.63584	1.63E+08
BoxBFL.154.TYR.N	8.12892	120.4783	12.99281	12.73019	1.34E+08
BoxBFL.155.GLU.N	8.52395	117.5111	13.25587	12.99665	1.34E+08
BoxBFL.157.ASP.N	8.85469	123.1047	12.50397	12.25568	1.51E+08
BoxBFL.158.ILE.N	9.06751	122.2825	13.25933	12.71521	1.38E+08
BoxBFL.159.ALA.N	7.44862	124.7591	12.4575	12.38733	1.75E+08
BoxBFL.160.ALA.N	7.7427	120.5687	14.46225	12.64065	1.71E+08
BoxBFL.161.TYR.N	8.19946	120.0547	15.35591	13.74534	1.39E+08
BoxBFL.162.ARG.N	8.2623	118.7997	12.52352	12.69996	1.54E+08

0.84 molar ratio CXCL12/HMG Box (Day 2)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.337 mM, U-15N HMGB1A-c028 (Box B 94-162), C-terminal biotinylation tag (300 µL of 0.45 mM stock)
CXCL12 construct and concentration	0.285 mM, CXCL12A-c021, wt 1-67, no tag (100 µL of 1.14 mM stock)
Volume	350 µL
Sample pH (after adjusting)	7.73
Temperature	25 ºC
Number of scans	8

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.95.LYS.N	8.47806	122.8889	15.74549	12.59893	1.04E+08
BoxBFL.96.ARG.N	8.31955	123.9958	15.32754	11.79906	1.23E+08
BoxBFL.99.SER.N	7.95975	114.213	18.91249	16.00748	19314842
BoxBFL.100.ALA.N	9.06482	122.8403	12.12244	12.51287	1.23E+08
BoxBFL.101.PHE.N	8.44328	116.1449	14.02058	12.94259	83270120
BoxBFL.102.PHE.N	8.2626	122.1026	14.57979	12.61002	83484184
BoxBFL.103.LEU.N	8.27876	121.1543	13.21545	12.41884	98850448
BoxBFL.104.PHE.N	8.1723	122.6037	17.85543	16.70344	1.63E+08
BoxBFL.105.CYS.N	8.49894	118.3888	12.67018	13.00159	1.03E+08
BoxBFL.106.SER.N	8.12904	115.0558	11.922	12.17556	1.66E+08
BoxBFL.107.GLU.N	7.15246	119.1841	15.05754	12.02976	1.54E+08
BoxBFL.108.TYR.N	7.94272	114.6176	17.11915	12.55614	1.04E+08
BoxBFL.109.ARG.N	9.23036	123.7013	13.06226	12.53714	1.15E+08
BoxBFL.111.LYS.N	6.79969	117.703	15.74346	12.62776	1.25E+08
BoxBFL.112.ILE.N	8.1614	119.7481	15.19044	12.89975	1.06E+08
BoxBFL.113.LYS.N	8.25273	118.4894	16.03035	15.20607	71142304
BoxBFL.114.GLY.N	7.63597	103.7241	14.57231	12.02882	1.1E+08
BoxBFL.115.GLU.N	7.49968	119.5819	15.16195	12.5197	1.35E+08
BoxBFL.116.HIS.N	7.82235	114.787	15.19416	11.98251	1.02E+08
BoxBFL.119.LEU.N	7.3624	121.2015	12.53706	12.04396	2.03E+08
BoxBFL.120.SER.N	9.31525	121.061	13.81227	12.2652	1.13E+08
BoxBFL.121.ILE.N	8.74925	120.6769	12.5753	12.7098	61919540
BoxBFL.122.GLY.N	8.71207	109.4584	25.89949	13.91125	16954762
BoxBFL.123.ASP.N	7.95784	124.2831	14.47967	12.59113	1.13E+08
BoxBFL.124.VAL.N	8.63489	124.0688	13.34502	12.07557	1.51E+08
BoxBFL.125.ALA.N	7.92846	121.3424	11.96993	12.06949	1.99E+08
BoxBFL.126.LYS.N	8.01973	119.6945	12.24395	12.62986	2.08E+08
BoxBFL.127.LYS.N	8.03911	121.3059	12.17418	12.02556	2.24E+08
BoxBFL.128.LEU.N	8.64346	120.1657	12.73252	12.33711	1.36E+08
BoxBFL.129.GLY.N	8.26155	106.0419	14.87585	12.01202	1.26E+08
BoxBFL.130.GLU.N	8.07094	123.248	12.8194	12.1305	1.81E+08
BoxBFL.131.MET.N	8.77033	118.9628	12.47681	12.18055	1.65E+08
BoxBFL.132.TRP.N	8.71521	122.6821	13.43252	12.15419	1.09E+08
BoxBFL.133.ASN.N	8.17879	117.3654	11.81466	11.94108	1.72E+08
BoxBFL.134.ASN.N	7.58738	115.8377	20.34615	11.95454	1.47E+08
BoxBFL.135.THR.N	7.30659	119.9511	15.61869	12.3645	1.57E+08
BoxBFL.136.ALA.N	9.2484	131.4382	11.79903	12.0221	1.7E+08
BoxBFL.138.ASP.N	8.99068	115.7088	13.61787	13.15666	66921064
BoxBFL.139.ASP.N	7.33797	117.8781	18.54773	12.14095	1.52E+08
BoxBFL.140.LYS.N	7.81083	119.1164	14.67427	12.45762	1.24E+08
BoxBFL.141.GLN.N	7.42677	118.2031	12.06407	12.71254	1.77E+08
BoxBFL.143.TYR.N	7.21631	115.6619	15.45741	12.35014	1.19E+08

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.144.GLU.N	8.18784	120.6799	13.13889	12.30444	1.27E+08
BoxBFL.145.LYS.N	9.183	121.238	13.17983	12.43273	1.47E+08
BoxBFL.146.LYS.N, BoxBFL.156.LYS.N	7.69709	120.5754	12.94948	12.65531	1.73E+08
BoxBFL.147.ALA.N	8.48271	121.0951	12.06422	12.53668	1.63E+08
BoxBFL.148.ALA.N	8.37014	122.0096	12.01385	12.60817	1.9E+08
BoxBFL.149.LYS.N	7.97339	120.7383	13.42094	12.64648	1.82E+08
BoxBFL.150.LEU.N	8.28676	120.0845	14.4103	13.05275	1.22E+08
BoxBFL.151.LYS.N	8.42457	123.5222	12.66033	12.96606	1.42E+08
BoxBFL.152.GLU.N	8.08704	119.9998	12.5224	13.21315	1.98E+08
BoxBFL.153.LYS.N	7.75272	119.7489	13.27981	12.63584	1.63E+08
BoxBFL.154.TYR.N	8.12892	120.4783	12.99281	12.73019	1.34E+08
BoxBFL.155.GLU.N	8.52395	117.5111	13.25587	12.99665	1.34E+08
BoxBFL.157.ASP.N	8.85469	123.1047	12.50397	12.25568	1.51E+08
BoxBFL.158.ILE.N	9.06751	122.2825	13.25933	12.71521	1.38E+08
BoxBFL.159.ALA.N	7.44862	124.7591	12.4575	12.38733	1.75E+08
BoxBFL.160.ALA.N	7.7427	120.5687	14.46225	12.64065	1.71E+08
BoxBFL.161.TYR.N	8.19946	120.0547	15.35591	13.74534	1.39E+08
BoxBFL.162.ARG.N	8.2623	118.7997	12.52352	12.69996	1.54E+08

1.43 molar ratio CXCL12/HMG Box (Day 2)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.28 mM, U-15N HMGB1A-c028 (Box B 94-162), C-terminal biotinylation tag, 300 µL of 0.45 mM stock
CXCL12 construct and concentration	0.412 mM, CXCL12A-c021, wt 1-67, no tag, 170 µL of 1.14 mM stock
Temperature	25 ºC
Volume	350 µL
Sample pH (after adjusting)	7.69
Number of scans	8

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.95.LYS.N	8.47778	122.9103	15.37565	12.39804	80050272
BoxBFL.96.ARG.N	8.33007	124.0602	14.92746	12.00118	82355536
BoxBFL.99.SER.N	7.96392	114.2034	N/A	N/A	14006653
BoxBFL.100.ALA.N	9.06578	122.8423	12.95809	12.6583	87116544
BoxBFL.101.PHE.N	8.44289	116.1454	13.47969	12.85253	61494512
BoxBFL.102.PHE.N	8.26155	122.0515	14.49526	12.56936	63868476
BoxBFL.103.LEU.N	8.28051	121.1503	14.84782	12.62126	65927464
BoxBFL.104.PHE.N	8.18982	122.6828	20.78995	15.31712	1E+08
BoxBFL.105.CYS.N	8.49843	118.3925	13.1053	13.13926	70645200
BoxBFL.106.SER.N	8.1287	115.0622	12.0922	12.5294	1.18E+08
BoxBFL.107.GLU.N	7.15039	119.185	14.99568	11.952	1.12E+08
BoxBFL.108.TYR.N	7.9412	114.6199	17.85822	12.78123	73545448
BoxBFL.109.ARG.N	9.23053	123.7071	13.66025	13.12308	78615272
BoxBFL.111.LYS.N	6.79981	117.7058	16.96135	12.39598	86118672
BoxBFL.112.ILE.N	8.16151	119.7459	14.73583	13.27134	71542432
BoxBFL.113.LYS.N	8.25631	118.4871	16.07722	14.62748	55390008
BoxBFL.114.GLY.N	7.63682	103.7319	15.93962	12.34204	83576064
BoxBFL.115.GLU.N	7.49983	119.5699	15.3896	12.37429	96911488
BoxBFL.116.HIS.N	7.82619	114.7906	14.24248	12.35403	74137880
BoxBFL.119.LEU.N	7.36266	121.2025	12.44261	12.21601	1.49E+08
BoxBFL.120.SER.N	9.31408	121.062	12.93199	12.44319	84698824
BoxBFL.121.ILE.N	8.75038	120.6813	14.48357	12.89635	45051028
BoxBFL.122.GLY.N	8.71118	109.4804	16.4655	14.14866	16957850
BoxBFL.123.ASP.N	7.95741	124.2905	15.0023	12.39813	80795464
BoxBFL.124.VAL.N	8.63479	124.0656	13.47634	12.10198	1.07E+08
BoxBFL.125.ALA.N	7.92779	121.3329	13.01599	12.5379	1.33E+08
BoxBFL.126.LYS.N	8.02119	119.7027	12.30924	12.8741	1.48E+08
BoxBFL.127.LYS.N	8.03986	121.3062	12.4847	12.16844	1.55E+08
BoxBFL.128.LEU.N	8.64351	120.1664	12.98093	12.34936	95774424
BoxBFL.129.GLY.N	8.26062	106.0395	14.57273	12.29455	88870768
BoxBFL.130.GLU.N	8.07098	123.2514	12.85049	12.30451	1.28E+08
BoxBFL.131.MET.N	8.77103	118.9657	12.95376	12.1418	1.15E+08
BoxBFL.132.TRP.N	8.71439	122.6801	14.10342	12.19979	76116600
BoxBFL.133.ASN.N	8.17827	117.365	12.53999	12.19573	1.2E+08
BoxBFL.134.ASN.N	7.58849	115.8472	19.64464	11.72657	1.07E+08
BoxBFL.135.THR.N	7.30642	119.9458	15.21751	12.24418	1.12E+08
BoxBFL.136.ALA.N	9.24753	131.4348	12.65711	12.04662	1.19E+08
BoxBFL.138.ASP.N	8.99115	115.7039	12.95325	13.45469	55523416
BoxBFL.139.ASP.N	7.33717	117.8755	17.56433	12.18072	1.09E+08
BoxBFL.140.LYS.N	7.80982	119.1179	15.67278	12.39083	82706752
BoxBFL.141.GLN.N	7.42724	118.2084	12.20378	12.60333	1.22E+08
BoxBFL.143.TYR.N	7.21658	115.6631	17.78092	12.5108	80885808
BoxBFL.144.GLU.N	8.18558	120.6672	12.65054	12.60763	91853544
BoxBFL.145.LYS.N	9.1835	121.244	13.52547	12.49804	1.01E+08

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.146.LYS.N, BoxBFL.156.LYS.N	7.69757	120.5767	13.43106	13.08878	1.22E+08
BoxBFL.147.ALA.N	8.48332	121.097	12.54408	12.70391	1.12E+08
BoxBFL.148.ALA.N	8.37158	122.0199	12.14356	12.99704	1.29E+08
BoxBFL.149.LYS.N	7.97496	120.7468	12.94812	12.72338	1.28E+08
BoxBFL.150.LEU.N	8.28799	120.0826	13.26184	13.21101	83975808
BoxBFL.151.LYS.N	8.42544	123.5327	13.10847	12.87192	96388256
BoxBFL.152.GLU.N	8.08805	120.0097	13.34088	13.96487	1.37E+08
BoxBFL.153.LYS.N	7.75361	119.7537	13.4631	12.86615	1.13E+08
BoxBFL.154.TYR.N	8.12773	120.4857	13.65716	12.83688	92183904
BoxBFL.155.GLU.N	8.522	117.5173	12.9949	13.07435	92028096
BoxBFL.157.ASP.N	8.85612	123.1054	13.64785	12.54087	99610256
BoxBFL.158.ILE.N	9.06839	122.2825	13.74015	12.91936	94902504
BoxBFL.159.ALA.N	7.44725	124.7733	12.59718	12.9055	1.24E+08
BoxBFL.160.ALA.N	7.74231	120.5797	15.01396	13.34523	1.18E+08
BoxBFL.161.TYR.N	8.19948	120.0505	15.71875	13.97336	94945976
BoxBFL.162.ARG.N	8.26146	118.804	12.47976	12.94011	1.12E+08

Baseline 2 (Day 2)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.45 mM, HMGB1-c028 (Box B 94-162), C-terminal biotinylation tag
CXCL12 construct and concentration	0 mM
Temperature	25 ºC
Volume	300 µL
Sample pH (after adjusting)	7.7
Number of scans	8

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.95.LYS.N	8.47836	122.8704	15.45555	12.31895	1.15E+08
BoxBFL.96.ARG.N	8.30905	123.928	14.69757	11.67602	1.52E+08
BoxBFL.99.SER.N	7.95323	114.2133	15.92668	15.37737	19244536
BoxBFL.100.ALA.N	9.06443	122.8375	12.04384	12.58362	1.45E+08
BoxBFL.101.PHE.N	8.44327	116.1467	13.37094	12.57032	99456376
BoxBFL.102.PHE.N	8.26291	122.1401	13.91643	12.69516	93582568
BoxBFL.103.LEU.N	8.27739	121.1559	13.54853	12.32854	1.2E+08
BoxBFL.104.PHE.N	8.15312	122.5067	16.98582	12.61768	2.23E+08
BoxBFL.105.CYS.N	8.49958	118.3879	12.45479	13.04816	1.23E+08
BoxBFL.106.SER.N	8.12962	115.053	11.71118	11.95992	2.01E+08
BoxBFL.107.GLU.N	7.15443	119.1857	15.15976	11.875	1.85E+08
BoxBFL.108.TYR.N	7.94312	114.6181	16.58447	12.32968	1.25E+08
BoxBFL.109.ARG.N	9.2303	123.6965	13.1266	12.39272	1.36E+08
BoxBFL.111.LYS.N	6.80044	117.7026	15.51854	12.47282	1.51E+08
BoxBFL.112.ILE.N	8.16214	119.7445	14.38896	12.72125	1.33E+08
BoxBFL.113.LYS.N	8.24788	118.4867	17.04567	13.89265	75341960
BoxBFL.114.GLY.N	7.63545	103.7167	16.12903	11.9801	1.23E+08
BoxBFL.115.GLU.N	7.50084	119.5887	15.33825	12.18563	1.53E+08
BoxBFL.116.HIS.N	7.81992	114.7797	14.07903	12.00191	1.27E+08
BoxBFL.119.LEU.N	7.36245	121.1993	12.49408	11.92292	2.35E+08
BoxBFL.120.SER.N	9.31637	121.0591	13.96408	12.31166	1.32E+08
BoxBFL.121.ILE.N	8.74889	120.6692	14.04081	13.00014	64885424
BoxBFL.122.GLY.N	8.71247	109.452	14.17681	13.60233	17286296
BoxBFL.123.ASP.N	7.95965	124.2811	13.89649	12.49605	1.34E+08
BoxBFL.124.VAL.N	8.63524	124.0725	13.34482	12.04966	1.82E+08
BoxBFL.125.ALA.N	7.92792	121.3503	11.86774	11.79008	2.42E+08
BoxBFL.126.LYS.N	8.01841	119.6864	11.86371	12.33853	2.45E+08
BoxBFL.127.LYS.N	8.03849	121.3063	12.1092	12.01079	2.71E+08
BoxBFL.128.LEU.N	8.64384	120.1687	12.82931	12.23453	1.65E+08
BoxBFL.129.GLY.N	8.26299	106.0359	14.80106	11.90484	1.55E+08
BoxBFL.130.GLU.N	8.07168	123.2465	12.61599	12.00749	2.15E+08
BoxBFL.131.MET.N	8.76967	118.9613	12.58221	12.16772	2E+08
BoxBFL.132.TRP.N	8.71566	122.6842	13.26303	12.06045	1.35E+08
BoxBFL.133.ASN.N	8.17935	117.3676	11.78966	11.94741	2.04E+08
BoxBFL.134.ASN.N	7.58744	115.8288	20.03676	12.02406	1.76E+08
BoxBFL.135.THR.N	7.30721	119.9545	14.90053	12.28766	1.9E+08
BoxBFL.136.ALA.N	9.2491	131.4409	11.88622	11.94243	1.93E+08
BoxBFL.138.ASP.N	8.99179	115.703	14.27261	13.33337	68493488
BoxBFL.139.ASP.N	7.33887	117.8767	18.17146	11.95364	1.82E+08
BoxBFL.140.LYS.N	7.81148	119.1162	15.20448	12.14857	1.5E+08
BoxBFL.141.GLN.N	7.4268	118.1975	11.86842	12.43174	2.19E+08
BoxBFL.143.TYR.N	7.21681	115.6633	16.06949	12.06739	1.45E+08
BoxBFL.144.GLU.N	8.19039	120.6913	12.43979	12.40533	1.6E+08
BoxBFL.145.LYS.N	9.18288	121.2343	13.02849	12.2664	1.81E+08
BoxBFL.146.LYS.N, BoxBFL.156.LYS.N	7.69649	120.5726	12.4096	12.47784	2.08E+08

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.147.ALA.N	8.48223	121.0936	11.71767	12.54903	2.05E+08
BoxBFL.148.ALA.N	8.36946	122.0011	12.08468	12.21854	2.33E+08
BoxBFL.149.LYS.N	7.9718	120.7293	12.63602	12.4563	2.26E+08
BoxBFL.150.LEU.N	8.28613	120.0804	13.87273	13.26832	1.52E+08
BoxBFL.151.LYS.N	8.42463	123.5087	12.67228	12.49152	1.76E+08
BoxBFL.152.GLU.N	8.08647	119.9901	12.2303	12.58471	2.36E+08
BoxBFL.153.LYS.N	7.75247	119.748	13.27241	12.43002	1.97E+08
BoxBFL.154.TYR.N	8.13017	120.4741	12.9694	12.56059	1.68E+08
BoxBFL.155.GLU.N	8.52488	117.5048	13.26926	12.59952	1.65E+08
BoxBFL.157.ASP.N	8.85488	123.1048	12.53023	12.19821	1.85E+08
BoxBFL.158.ILE.N	9.06773	122.2845	12.97983	12.55507	1.7E+08
BoxBFL.159.ALA.N	7.45038	124.7516	12.10878	12.01767	2.17E+08
BoxBFL.160.ALA.N	7.74351	120.5631	14.03576	12.41792	2.12E+08
BoxBFL.161.TYR.N	8.20134	120.056	14.08572	13.06065	1.75E+08
BoxBFL.162.ARG.N	8.26373	118.7947	12.78627	12.3573	1.81E+08

Suppl. data A4-1-B HMGB1A-c007 (3S, 1-184) with 1:2 molar ratio CXCL12 (1:1 molar ratio CXCL12 to HMG Box), in 10 mM HEPES pH 7.5 150 mM NaCl

	BRUKER BIOSPIN 750 MHZ BRUKER AVANCE SPECTROMETER WITH 5MM TCI CRYOPROBE
^{1}H frequency (Hz)	749.91352
^{1}H sweep width (ppm)	12.122609
^{1}H sweep width (Hz)	9090.9090
^{1}HN sweep width (ppm)	12.122609
^{15}N frequency (Hz)	75.99661
^{15}N sweep width (ppm)	32.01577
^{15}N sweep width (Hz)	2433.09002
Water suppression gradient pulse width (Hz)	3527.31

Baseline 1 (Day 1)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.30 mM, HMGB1-c007 (3S 1-184), TEV-cleaved (200 µL 0.45 mM stock + 185 µL buffer)
CXCL12 construct and concentration	0 mM
Temperature	25 ºC
Volume	350 µL
Sample pH (after adjusting)	7.66
Number of scans	16
Buffer	10 mM HEPES, pH 7.5

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.3.GLY.H, TL.10.GLY.H	8.6061	111.0604	24.6728	23.8470	4.24E+06
TL.6.LYS.H	8.3808	120.7275	19.3276	19.0515	1.26E+07
TL.7.LYS.H	7.9672	123.9738	18.9452	18.4814	1.22E+07
TL.9.ARG.H	8.9422	124.5857	30.5185	22.5922	2.68E+06
TL.11.LYS.H	7.6533	119.5407	18.7001	20.1359	3.98E+06
TL.13.SER.H, TL.26.HIS.H	8.1990	119.2086	26.4193	23.6946	6.96E+06
TL.14.SER.H	9.3037	116.5728	17.7398	20.7154	1.52E+06
TL.15.TYR.H	7.9749	122.3508	21.5872	20.0836	9.77E+06
TL.16.ALA.H	7.8898	122.0710	18.6785	21.9854	1.16E+07
TL.17.PHE.H	8.3224	118.1346	25.4892	18.2495	5.66E+06
TL.18.PHE.H	8.1246	125.6475	19.4733	26.9249	6.38E+06
TL.21.THR.H	8.1820	116.6683	21.3497	19.6855	7.47E+06
TL.23.ARG.H	8.9337	123.5827	19.6637	19.5765	1.34E+07
TL.25.GLU.H	8.3297	119.6119	17.9559	20.8464	4.41E+06
TL.27.LYS.H	7.9199	118.2035	17.4027	24.5903	1.32E+07
TL.28.LYS.H	7.4656	117.6101	21.4259	22.0467	2.19E+06
TL.29.LYS.H	7.5008	116.9197	26.2391	24.4263	3.84E+06
TL.30.HIS.H	7.9045	116.8771	21.2366	19.0908	3.22E+07
TL.32.ASP.H	8.5014	116.2147	21.6349	20.6257	1.15E+07
TL.33.ALA.H	7.6405	123.1947	27.1908	18.4877	6.05E+06
TL.36.ASN.H	8.5631	124.5616	21.3601	19.8298	1.05E+07
TL.38.SER.H	8.5266	116.5707	20.6963	19.7437	1.75E+07
TL.39.GLU.H	7.8017	121.4763	24.3844	20.7391	5.60E+06
TL.41.SER.H	8.4328	114.6754	21.9986	19.2519	7.61E+06
TL.43.LYS.H	7.9047	120.0881	24.5054	21.5027	8.26E+06
TL.48.TRP.H	8.6119	121.3719	27.7321	19.6527	6.67E+06
TL.49.LYS.H	7.5392	114.9177	22.4466	26.2994	6.95E+06
TL.50.THR.H	7.4184	106.6545	19.9202	26.9454	3.55E+06
TL.51.MET.H	7.1286	124.1028	34.8660	19.9547	3.59E+06
TL.52.SER.H	9.0214	120.2824	25.7658	22.2592	3.14E+06
TL.54.LYS.H	8.2612	118.6673	21.4006	20.8276	1.72E+07
TL.55.GLU.H, TL.146.LYS.H, TL.156.LYS.H	7.6743	120.5743	39.0944	20.4649	4.08E+07
TL.57.GLY.H	7.8481	107.2718	24.8060	21.1106	4.92E+06
TL.58.LYS.H	7.9240	118.8658	23.1257	30.2328	5.35E+06
TL.60.GLU.H	8.4408	121.0589	18.5170	21.8991	2.43E+07
TL.61.ASP.H	8.5793	121.9550	25.9461	21.0215	2.41E+06
TL.63.ALA.H	8.0272	123.2985	18.8092	20.8840	3.01E+07
TL.64.LYS.H	8.5925	122.4879	24.6164	26.1814	3.30E+06
TL.65.ALA.H	7.9334	123.0276	19.4945	20.6248	1.77E+07
TL.66.ASP.H	8.3419	120.0786	22.2268	23.0110	5.95E+06
TL.67.LYS.H	8.2124	120.8922	19.2067	21.7276	1.54E+07
TL.68.ALA.H	7.4653	120.9539	26.2159	21.1243	4.51E+06

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.71.GLU.H	8.4294	117.1838	20.4156	25.0309	7.16E+06
TL.72.ARG.H	8.0622	120.1074	18.6953	20.2934	2.46E+07
TL.73.GLU.H	8.4749	120.0201	22.5415	19.5028	9.11E+06
TL.74.MET.H	8.4409	117.6522	#VALUE!	#VALUE!	6.40E+06
TL.75.LYS.H	7.5100	119.0899	22.5073	22.2834	3.08E+06
TL.77.TYR.H	7.5024	123.9352	30.9003	17.6211	3.53E+06
TL.78.ILE.H	7.8034	129.2718	20.0866	19.3005	1.72E+07
TL.88.PHE.H	8.3596	122.1156	17.9211	19.8933	2.66E+07
TL.89.LYS.H	8.2181	123.3632	21.8312	19.6249	1.19E+07
TL.90.ASP.H	8.0110	119.1287	20.5283	23.7061	9.96E+06
TL.93.ALA.H	7.3090	123.8485	18.6517	18.5426	2.41E+07
TL.95.LYS.H	8.5094	123.7507	19.5433	18.6325	1.67E+07
TL.99.SER.H	7.8684	114.9049	20.2451	19.8589	2.20E+07
TL.100.ALA.H	9.0706	122.9058	17.9269	19.6686	1.43E+07
TL.102.PHE.H	8.2289	122.4276	19.9106	19.5888	1.37E+07
TL.103.LEU.H	8.2529	120.0969	21.6806	25.4949	1.99E+07
TL.104.PHE.H	8.1450	122.5447	19.7785	17.8570	3.42E+06
TL.106.SER.H	7.9866	116.8141	16.6297	18.4050	2.77E+07
TL.107.GLU.H	7.1139	119.4856	18.6763	18.2751	2.77E+07
TL.111.LYS.H	6.8070	117.8973	20.9968	19.4882	1.97E+07
TL.112.ILE.H	8.1471	119.7094	19.6956	20.7896	2.21E+07
TL.113.LYS.H	8.2277	118.6620	18.9931	20.0020	2.62E+07
TL.114.GLY.H	7.6276	103.8053	19.2245	18.7965	2.17E+07
TL.115.GLU.H	7.4519	119.5050	19.0651	18.6530	2.50E+07
TL.116.HIS.H	7.8557	114.5584	21.3730	26.3315	1.95E+07
TL.119.LEU.H	7.3441	121.1282	16.9517	18.0939	4.11E+07
TL.120.SER.H	9.2783	120.9530	19.8514	19.2891	1.57E+07
TL.121.ILE.H	8.7093	120.6277	19.2620	19.2230	1.13E+07
TL.122.GLY.H	8.6846	109.4587	26.8288	22.3716	4.51E+06
TL.123.ASP.H	7.9382	124.3219	18.3663	18.7325	2.13E+07
TL.124.VAL.H	8.6040	124.0412	18.2886	18.7654	2.25E+07
TL.125.ALA.H	7.8785	121.3328	16.5530	19.1095	3.63E+07
TL.126.LYS.H	7.9989	119.7206	19.3006	19.3155	7.93E+07
TL.127.LYS.H	8.0145	121.3551	17.2034	17.8224	7.22E+07
TL.128.LEU.H	8.6505	120.2524	18.6803	18.7237	2.09E+07
TL.129.GLY.H	8.2032	106.2602	19.1920	18.8485	2.46E+07
TL.131.MET.H	8.7231	118.8760	18.0856	18.4060	2.38E+07
TL.132.TRP.H	8.7095	122.5759	19.0186	19.0721	1.59E+07
TL.133.ASN.H	8.1328	117.2422	19.0104	18.8528	3.15E+07
TL.134.ASN.H	7.5431	115.7695	19.1824	18.8449	2.98E+07
TL.135.THR.H	7.2603	119.9046	19.3742	19.5603	2.60E+07
TL.136.ALA.H	9.2045	131.3561	17.2667	18.5735	2.23E+07
TL.137.ALA.H	8.8342	124.9935	33.3526	21.2936	1.19E+06
TL.138.ASP.H	8.9528	115.6245	18.8620	19.6655	1.34E+07
TL.139.ASP.H	7.2986	117.8616	19.8829	18.6225	2.53E+07
TL.140.ALA.H	7.7883	119.0980	20.0357	19.4836	1.79E+07
TL.141.GLN.H	7.4069	118.1124	18.1893	19.2718	2.30E+07
TL.143.TYR.H	7.1940	115.5652	20.7215	19.7599	1.68E+07
TL.145.LYS.H	9.1712	121.2666	18.7285	20.6529	1.58E+07
TL.147.ALA.H	8.3287	120.5992	22.2428	21.8046	1.24E+07
TL.149.LYS.H	7.9709	120.8132	19.5777	23.5142	2.79E+07
TL.150.LEU.H	8.2579	119.8355	19.6870	24.4644	1.95E+07
TL.151.LYS.H	8.3943	123.7275	19.8464	22.6956	2.13E+07
TL.152.GLU.H	8.0005	120.0908	18.5048	18.9795	4.56E+07
TL.153.LYS.H	7.7378	119.7316	20.1171	19.9428	2.47E+07
TL.154.TYR.H	8.1211	120.2560	27.4207	25.0214	4.56E+07
TL.155.GLU.H	8.4779	117.6292	20.1839	20.3442	1.81E+07
TL.157.ASP.H	8.8596	123.1691	18.1489	19.7980	1.95E+07
TL.158.ILE.H	9.1467	122.4891	19.2484	19.6793	1.60E+07

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.159.ALA.H	7.3324	124.8211	18.1448	19.5515	2.61E+07
TL.160.ALA.H	7.7523	120.3814	18.8194	20.3354	2.87E+07
TL.162.ARG.H	8.2660	119.1276	20.2383	21.2173	2.28E+07
TL.163.ALA.H	7.5998	121.5022	18.4569	19.0784	1.90E+07
TL.164.LYS.H	7.6444	118.6458	18.5535	19.1662	1.93E+07
TL.166.LYS.H	8.0107	121.8105	21.0399	22.8760	1.20E+07
TL.168.ASP.H	8.4217	120.2067	20.8680	20.3203	1.18E+07
TL.169.ALA.H	8.2713	124.9694	23.2824	19.9815	1.63E+07
TL.170.ALA.H	8.1977	121.6039	19.5714	19.2612	1.48E+07
TL.174.VAL.H	7.9243	119.5174	23.1927	24.7751	6.62E+06
TL.175.VAL.H	8.3105	125.5123	16.0756	16.4540	6.65E+07
TL.176.LYS.H	8.4655	126.6128	28.7022	20.5965	3.74E+06
TL.177.ALA.H	8.3773	125.8179	31.4688	20.8340	3.36E+06
TL.178.GLU.H	8.4450	120.6799	#VALUE!	#VALUE!	9.04E+06
TL.179.LYS.H	8.4425	122.3029	19.9586	24.6041	1.34E+07
TL.185.GLU.H	8.2782	122.8421	19.8975	21.0209	1.81E+07

3D experiments (Day 1)

These spectra are not shown due to their 3D nature; data available if requested. Sample is that of Baseline 1.

Experiment type	**3D, ^{15}N EDITED ^{1}H-^{1}H TOCSY-HSQC, THROUGH-SPACE INTERACTION (1 MAR 19)**
Temperature	25 ºC
Volume	300 µL
Number of scans	8
Experiment type	**3D, ^{15}N edited ^{1}H-^{1}H NOESY-HSQC, through-bond interaction (28 FEB 19)**
Temperature	25 ºC
Volume	300 µL
Number of scans	16

Baseline 2 (Day 3)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.30 mM, HMGB1-c007 (3S 1-184), TEV-cleaved (200 µL 0.45 mM stock + 185 µL buffer)
CXCL12 construct and concentration	0 mM
Temperature	25 ºC
Sample pH (after adjusting)	7.66
Volume	350 µL
Number of scans	16
Buffer	10 mM HEPES, pH 7.5

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.3.GLY.H, TL.10.GLY.H	8.6065	111.0499	23.1613	23.4948	4.43E+06
TL.6.LYS.H	8.3807	120.7320	19.6417	18.6975	1.24E+07
TL.7.LYS.H	7.9677	123.9750	18.5170	18.1121	1.19E+07
TL.9.ARG.H	8.9453	124.5715	24.2572	23.7356	2.61E+06
TL.11.LYS.H	7.6528	119.5498	17.4576	20.3878	3.89E+06
TL.13.SER.H, TL.26.HIS.H	8.1991	119.1995	26.6500	22.3922	6.54E+06
TL.14.SER.H	9.2956	116.5826	16.8521	20.2197	1.33E+06
TL.15.TYR.H	7.9757	122.3477	21.7938	19.9370	9.68E+06
TL.16.ALA.H	7.8896	122.0725	19.8069	21.8309	1.12E+07
TL.17.PHE.H	8.3224	118.1373	25.3998	18.6893	5.67E+06
TL.18.PHE.H	8.1241	125.6449	20.5389	27.4613	6.00E+06
TL.21.THR.H	8.1825	116.6667	21.1937	19.7947	7.42E+06
TL.23.ARG.H	8.9329	123.5826	19.1345	19.5535	1.33E+07
TL.25.GLU.H	8.3294	119.6141	20.8289	20.5729	4.29E+06
TL.27.LYS.H	7.9196	118.2075	17.7875	24.4847	1.31E+07
TL.28.LYS.H	7.4656	117.5781	22.7406	23.1308	2.32E+06
TL.29.LYS.H	7.4989	116.9374	23.0209	25.0045	4.05E+06
TL.30.HIS.H	7.9042	116.8739	21.0887	18.9026	3.21E+07
TL.32.ASP.H	8.5004	116.2125	20.0730	20.5524	1.17E+07
TL.33.ALA.H	7.6424	123.1937	25.8311	18.9977	6.03E+06
TL.36.ASN.H	8.5629	124.5621	20.2634	20.1827	1.01E+07
TL.38.SER.H	8.5266	116.5762	20.3927	20.2102	1.75E+07
TL.39.GLU.H	7.8015	121.4813	25.4307	21.5986	5.49E+06
TL.41.SER.H	8.4315	114.6769	24.0654	18.8439	7.20E+06
TL.43.LYS.H	7.9025	120.0835	25.2682	22.0179	8.40E+06
TL.48.TRP.H	8.6129	121.3660	22.6050	19.2575	6.76E+06
TL.49.LYS.H	7.5396	114.9173	20.9942	25.1553	6.76E+06
TL.50.THR.H	7.4186	106.6388	24.1325	25.7242	3.44E+06
TL.51.MET.H	7.1310	124.1048	22.9068	19.6438	3.75E+06
TL.52.SER.H	9.0209	120.2880	28.9274	22.9000	2.95E+06
TL.54.LYS.H	8.2621	118.6676	20.9063	21.0484	1.74E+07
TL.55.GLU.H, TL.146.LYS.H, TL.156.LYS.H	7.6741	120.5737	35.8117	20.3103	4.07E+07
TL.57.GLY.H	7.8487	107.2598	28.3292	20.2673	4.70E+06
TL.58.LYS.H	7.9229	118.8497	24.2289	27.7675	5.26E+06
TL.60.GLU.H	8.4408	121.0595	18.5310	21.9295	2.41E+07
TL.61.ASP.H	8.5746	121.9512	28.9797	22.9634	2.66E+06
TL.63.ALA.H	8.0269	123.2962	18.7324	21.1710	2.95E+07
TL.64.LYS.H	8.5909	122.4962	24.9341	24.0694	2.94E+06
TL.65.ALA.H	7.9332	123.0304	20.3519	20.5074	1.74E+07
TL.66.ASP.H	8.3409	120.0857	22.2493	23.7054	5.80E+06
TL.67.LYS.H	8.2130	120.8955	19.6156	22.3669	1.54E+07
TL.68.ALA.H	7.4616	120.9561	37.9570	19.8115	4.30E+06
TL.71.GLU.H	8.4286	117.1801	21.4128	25.4454	6.76E+06
TL.72.ARG.H	8.0621	120.1083	18.7908	20.0752	2.45E+07

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.73.GLU.H	8.4738	120.0142	23.7008	19.2906	8.71E+06
TL.74.MET.H	8.4409	117.6522	#VALUE!	#VALUE!	6.44E+06
TL.75.LYS.H	7.5068	119.0858	26.7860	22.3687	3.23E+06
TL.77.TYR.H	7.5032	123.9308	38.9311	17.2302	3.49E+06
TL.78.ILE.H	7.8036	129.2703	19.9513	19.4233	1.70E+07
TL.88.PHE.H	8.3595	122.1143	17.6506	19.7390	2.67E+07
TL.89.LYS.H	8.2182	123.3647	19.6893	20.1167	1.23E+07
TL.90.ASP.H	8.0126	119.1425	20.7630	25.2387	9.55E+06
TL.93.ALA.H	7.3090	123.8482	18.8016	18.5673	2.40E+07
TL.95.LYS.H	8.5090	123.7493	18.6872	18.3947	1.70E+07
TL.99.SER.H	7.8687	114.9053	20.2014	20.0612	2.21E+07
TL.100.ALA.H	9.0705	122.9049	17.6281	19.4111	1.43E+07
TL.102.PHE.H	8.2287	122.4261	20.3823	19.7821	1.38E+07
TL.103.LEU.H	8.2533	120.0963	22.3290	25.7323	1.99E+07
TL.104.PHE.H	7.9865	116.8146	16.6032	18.4916	2.71E+07
TL.106.SER.H	7.1139	119.4847	18.9087	18.2576	2.71E+07
TL.107.GLU.H	6.8074	117.8959	20.3505	19.3071	1.99E+07
TL.111.LYS.H	8.1473	119.7047	19.6728	20.5823	2.19E+07
TL.112.ILE.H	8.2277	118.6631	19.2041	19.9527	2.58E+07
TL.113.LYS.H	7.6278	103.8060	20.2430	18.4257	2.16E+07
TL.114.GLY.H	7.4522	119.5041	18.8464	18.7187	2.51E+07
TL.115.GLU.H	7.8558	114.5590	21.3428	26.4292	1.94E+07
TL.116.HIS.H	7.3444	121.1280	17.0541	18.0097	4.07E+07
TL.119.LEU.H	9.2779	120.9556	19.1260	18.9896	1.56E+07
TL.120.SER.H	8.7096	120.6282	19.3699	19.3593	1.16E+07
TL.121.ILE.H	8.6834	109.4592	23.3770	21.1555	4.56E+06
TL.122.GLY.H	7.9382	124.3228	18.0611	18.8783	2.14E+07
TL.123.ASP.H	8.6038	124.0408	18.3007	18.9123	2.25E+07
TL.124.VAL.H	7.8787	121.3329	16.2690	19.0869	3.68E+07
TL.125.ALA.H	7.9987	119.7210	19.0408	19.1946	7.83E+07
TL.126.LYS.H	8.0145	121.3562	17.2124	17.7746	7.14E+07
TL.127.LYS.H	8.6507	120.2533	18.8465	18.5636	2.06E+07
TL.128.LEU.H	8.2034	106.2583	19.1542	18.5945	2.46E+07
TL.129.GLY.H	7.9939	122.7099	31.0113	24.8633	2.83E+06
TL.131.MET.H	8.7234	118.8753	17.8826	18.2976	2.41E+07
TL.132.TRP.H	8.7097	122.5757	19.4286	19.1503	1.60E+07
TL.133.ASN.H	8.1327	117.2411	19.1545	18.7949	3.10E+07
TL.134.ASN.H	7.5430	115.7691	19.2434	18.7462	2.97E+07
TL.135.THR.H	7.2604	119.9055	18.9668	19.2131	2.62E+07
TL.136.ALA.H	9.2048	131.3556	16.6770	18.6275	2.27E+07
TL.137.ALA.H	8.8215	124.9797	28.9631	18.8998	9.53E+05
TL.138.ASP.H	8.9534	115.6244	19.3079	19.6390	1.34E+07
TL.139.ASP.H	7.2990	117.8613	19.5011	18.4882	2.54E+07
TL.140.ALA.H	7.7881	119.0997	19.6218	19.4625	1.80E+07
TL.141.GLN.H	7.4070	118.1139	18.2290	19.3873	2.28E+07
TL.143.TYR.H	7.1941	115.5640	20.6936	19.5627	1.75E+07
TL.145.LYS.H	9.1714	121.2694	18.9207	20.5145	1.55E+07
TL.147.ALA.H	8.3294	120.5969	21.2897	22.3662	1.25E+07
TL.149.LYS.H	7.9711	120.8080	19.0205	22.9038	2.77E+07
TL.150.LEU.H	8.2586	119.8352	19.6218	24.8315	1.93E+07
TL.151.LYS.H	8.3946	123.7260	20.5952	23.0334	2.12E+07
TL.152.GLU.H	8.0004	120.0915	18.6445	18.7793	4.53E+07
TL.153.LYS.H	7.7381	119.7331	20.2180	19.8851	2.46E+07
TL.154.TYR.H	8.1209	120.2560	28.0283	25.0515	4.45E+07
TL.155.GLU.H	8.4777	117.6292	19.6687	20.1632	1.81E+07
TL.157.ASP.H	8.8597	123.1714	18.3048	19.6916	1.96E+07
TL.158.ILE.H	9.1469	122.4919	19.3081	19.5812	1.61E+07
TL.159.ALA.H	7.3325	124.8205	17.9084	19.2404	2.59E+07
TL.160.ALA.H	7.7526	120.3794	18.6070	20.3336	2.91E+07

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.162.ARG.H	8.2667	119.1253	20.5620	21.2604	2.28E+07
TL.163.ALA.H	7.5996	121.4996	18.4419	19.1701	1.89E+07
TL.164.LYS.H	7.6446	118.6434	18.3150	18.9990	1.94E+07
TL.166.LYS.H	8.0103	121.8118	20.8605	22.0326	1.21E+07
TL.168.ASP.H	8.4199	120.2046	21.4934	19.9990	1.18E+07
TL.169.ALA.H	8.2714	124.9725	23.9158	19.9230	1.64E+07
TL.170.ALA.H	8.1980	121.6044	19.8101	19.2270	1.50E+07
TL.174.VAL.H	7.9239	119.5328	23.8954	29.0599	6.20E+06
TL.175.VAL.H	8.3108	125.5127	16.0825	16.4752	6.74E+07
TL.176.LYS.H	8.4651	126.6236	22.4539	20.5401	4.14E+06
TL.177.ALA.H	8.3781	125.8126	25.5412	22.8124	3.71E+06
TL.178.GLU.H	8.4450	120.6799	#VALUE!	#VALUE!	9.25E+06
TL.179.LYS.H	8.4424	122.3058	19.6823	24.3392	1.27E+07
TL.185.GLU.H	8.2779	122.8405	19.1833	22.8993	1.91E+07

1:1 molar ratio CXCL12/HMG Box (1:2 ratio HMGB1A-c007/CXCL12A-c021, Day 4)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.30 mM, HMGB1-c007 (3S 1-184), TEV-cleaved (200 µL 0.45 mM stock)
CXCL12 construct and concentration	0.6 mM, CXCL12A-c021, wt 1-67, no tag (185 µL 1.14 mM stock)
Temperature	25 ºC
Volume	350 µL
Sample pH (after adjusting)	7.65
Number of scans	16
Buffer	10 mM HEPES, pH 7.5

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.3.GLY.H, TL.10.GLY.H	8.6044	111.0338	27.0394	23.1335	5.94E+06
TL.6.LYS.H	8.3809	120.6948	22.5007	18.0287	1.34E+07
TL.7.LYS.H	7.9643	123.9632	20.3742	18.4134	1.54E+07
TL.9.ARG.H	8.9503	124.5797	31.7140	24.0875	3.87E+06
TL.11.LYS.H	7.6557	119.5615	16.0398	21.6463	5.42E+06
TL.13.SER.H, TL.26.HIS.H	8.2004	119.2105	23.4564	24.3179	1.04E+07
TL.14.SER.H	9.3238	116.5556	14.7611	15.5346	1.67E+06
TL.15.TYR.H	7.9778	122.3535	21.1121	19.3913	1.51E+07
TL.16.ALA.H	7.8925	122.0870	19.3520	22.0446	1.84E+07
TL.17.PHE.H	8.3240	118.1495	27.5926	18.0447	9.03E+06
TL.18.PHE.H	8.1278	125.6553	19.8250	26.5265	1.13E+07
TL.21.THR.H	8.1835	116.6724	21.5629	19.5787	9.24E+06
TL.23.ARG.H	8.9348	123.5796	18.8791	19.0832	2.40E+07
TL.25.GLU.H	8.3321	119.6265	22.0156	18.9298	6.69E+06
TL.27.LYS.H	7.9208	118.2122	17.2389	24.5277	2.00E+07
TL.28.LYS.H	7.4693	117.6188	20.5149	21.7352	3.16E+06
TL.29.LYS.H	7.5019	116.9291	20.0353	22.0649	5.99E+06
TL.30.HIS.H	7.9083	116.8927	20.9767	19.1705	5.23E+07
TL.32.ASP.H	8.5041	116.2183	21.3681	20.8293	1.80E+07
TL.33.ALA.H	7.6486	123.1983	33.3099	19.1418	7.80E+06
TL.36.ASN.H	8.5655	124.5629	18.7902	19.8770	1.75E+07
TL.38.SER.H	8.5296	116.5760	19.5825	19.3443	3.10E+07
TL.39.GLU.H	7.8000	121.4911	27.7099	20.3638	8.06E+06
TL.41.SER.H	8.4349	114.6887	23.1352	19.2946	1.18E+07
TL.43.LYS.H	7.9049	120.0830	25.4006	22.5369	1.19E+07
TL.48.TRP.H	8.6160	121.3765	22.4001	19.4978	1.10E+07
TL.49.LYS.H	7.5429	114.9206	21.7660	26.6768	1.04E+07
TL.50.THR.H	7.4225	106.6598	25.9188	26.8533	5.09E+06
TL.51.MET.H	7.1282	124.0983	25.1134	21.1386	5.30E+06
TL.52.SER.H	9.0270	120.2729	23.7729	19.7149	4.97E+06
TL.54.LYS.H	8.2637	118.6728	21.3251	20.6048	2.61E+07
TL.55.GLU.H, TL.146.LYS.H, TL.156.LYS.H	7.6767	120.5832	30.2398	20.3845	6.71E+07
TL.57.GLY.H	7.8522	107.2682	26.6428	20.8923	8.03E+06
TL.58.LYS.H	7.9244	118.8701	22.7163	29.6656	7.94E+06
TL.60.GLU.H	8.4437	121.0610	18.2434	20.3448	4.09E+07
TL.61.ASP.H	8.5815	121.9597	33.1388	21.6707	3.39E+06
TL.63.ALA.H	8.0294	123.2984	18.2389	19.6294	5.05E+07
TL.64.LYS.H	8.5980	122.5146	26.0412	25.2214	4.71E+06
TL.65.ALA.H	7.9368	123.0339	18.6025	19.9725	2.82E+07
TL.66.ASP.H	8.3457	120.0879	22.5058	23.2892	7.23E+06
TL.67.LYS.H	8.2154	120.8999	18.9826	22.1852	2.52E+07

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.68.ALA.H	7.4697	120.9568	23.2933	19.0137	6.76E+06
TL.71.GLU.H	8.4315	117.1750	21.7295	26.3586	1.06E+07
TL.72.ARG.H	8.0645	120.1128	18.0569	19.4262	4.23E+07
TL.73.GLU.H	8.4787	120.0270	23.1364	19.3270	1.51E+07
TL.74.MET.H	8.4409	117.6522	#VALUE!	#VALUE!	9.83E+06
TL.75.LYS.H	7.5049	119.1098	31.7802	21.8213	4.59E+06
TL.77.TYR.H	7.5105	123.9396	38.1218	16.9483	4.14E+06
TL.78.ILE.H	7.8084	129.2854	19.5578	19.9301	2.60E+07
TL.88.PHE.H	8.3628	122.1204	16.9053	19.1178	4.49E+07
TL.89.LYS.H	8.2210	123.3706	19.7502	19.2731	2.10E+07
TL.90.ASP.H	8.0147	119.1518	21.2737	24.1312	1.43E+07
TL.93.ALA.H	7.3126	123.8531	19.3345	18.4537	3.57E+07
TL.95.LYS.H	8.5127	123.7578	18.2433	18.4107	2.55E+07
TL.99.SER.H	7.8719	114.9159	19.1639	18.9030	3.80E+07
TL.100.ALA.H	9.0737	122.9134	17.9203	18.9328	2.47E+07
TL.102.PHE.H	8.2316	122.4356	22.4797	20.9918	1.28E+07
TL.103.LEU.H	8.2553	120.1006	21.3513	23.9579	3.34E+07
TL.104.PHE.H	8.1434	122.5532	26.6494	18.4699	3.85E+06
TL.106.SER.H	7.9897	116.8195	16.1819	18.1182	4.66E+07
TL.107.GLU.H	7.1171	119.4906	17.8505	18.0469	4.73E+07
TL.111.LYS.H	6.8097	117.9049	19.8483	19.1076	3.40E+07
TL.112.ILE.H	8.1508	119.7121	19.4877	20.4021	3.72E+07
TL.113.LYS.H	8.2307	118.6666	18.3996	19.2564	4.37E+07
TL.114.GLY.H	7.6300	103.8047	18.7943	18.4306	3.60E+07
TL.115.GLU.H	7.4552	119.5222	19.2213	18.9640	4.04E+07
TL.116.HIS.H	7.8582	114.5580	20.7604	23.4804	3.16E+07
TL.119.LEU.H	7.3458	121.1340	16.1812	17.8123	6.86E+07
TL.120.SER.H	9.2812	120.9606	18.7787	19.3743	2.64E+07
TL.121.ILE.H	8.7114	120.6306	20.5564	19.1568	1.71E+07
TL.122.GLY.H	8.6882	109.4736	24.3584	20.5405	5.40E+06
TL.123.ASP.H	7.9414	124.3316	18.9583	18.7555	3.43E+07
TL.124.VAL.H	8.6075	124.0518	17.5597	18.0450	4.01E+07
TL.125.ALA.H	7.8803	121.3377	16.3135	18.8634	6.13E+07
TL.126.LYS.H	8.0024	119.7348	20.5705	18.3671	1.04E+08
TL.127.LYS.H	8.0190	121.3711	17.2894	17.8534	1.01E+08
TL.128.LEU.H	8.6532	120.2570	18.0062	18.2445	3.66E+07
TL.129.GLY.H	8.2053	106.2627	18.0361	18.4504	4.24E+07
TL.131.MET.H	8.7256	118.8819	16.6866	17.9953	4.34E+07
TL.132.TRP.H	8.7119	122.5814	18.4799	18.7626	2.80E+07
TL.133.ASN.H	8.1348	117.2495	17.4409	18.4657	5.40E+07
TL.134.ASN.H	7.5461	115.7741	19.5366	18.8391	4.91E+07
TL.135.THR.H	7.2627	119.9122	18.4851	18.5540	4.44E+07
TL.136.ALA.H	9.2071	131.3622	17.0411	18.2227	3.66E+07
TL.137.ALA.H	8.8274	124.9859	14.4547	22.1480	1.45E+06
TL.138.ASP.H	8.9571	115.6309	18.9345	19.4487	1.95E+07
TL.139.ASP.H	7.3010	117.8635	19.8679	18.2453	4.22E+07
TL.140.ALA.H	7.7905	119.1034	19.3181	18.7588	3.13E+07
TL.141.GLN.H	7.4097	118.1173	17.1880	18.8418	4.01E+07
TL.143.TYR.H	7.1967	115.5677	19.8319	19.1363	2.97E+07
TL.145.LYS.H	9.1741	121.2738	17.8887	20.0482	2.77E+07
TL.147.ALA.H	8.3318	120.6021	21.9785	20.7642	1.98E+07
TL.149.LYS.H	7.9737	120.8159	17.7823	22.3931	4.89E+07
TL.150.LEU.H	8.2612	119.8328	19.4856	23.9293	2.86E+07
TL.151.LYS.H	8.3980	123.7315	19.0907	22.3943	3.45E+07
TL.152.GLU.H	8.0036	120.1009	18.5759	18.4348	5.46E+07
TL.153.LYS.H	7.7406	119.7378	19.3471	19.4416	4.18E+07
TL.154.TYR.H	8.1247	120.2565	24.0128	22.4221	7.15E+07
TL.155.GLU.H	8.4809	117.6318	18.3609	19.8254	3.19E+07
TL.157.ASP.H	8.8614	123.1717	17.8349	19.3205	3.35E+07

F1 assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
TL.158.ILE.H	9.1480	122.4894	19.2904	19.2156	2.73E+07
TL.159.ALA.H	7.3355	124.8233	17.3550	19.1822	4.38E+07
TL.160.ALA.H	7.7549	120.3869	17.8564	19.8791	4.70E+07
TL.162.ARG.H	8.2685	119.1365	19.6143	21.0415	3.36E+07
TL.163.ALA.H	7.6030	121.5109	18.4123	19.0334	2.70E+07
TL.164.LYS.H	7.6483	118.6508	18.5044	19.7095	2.57E+07
TL.166.LYS.H	8.0157	121.7960	20.9936	22.5167	1.29E+07
TL.168.ASP.H	8.4254	120.2086	22.2976	21.9565	1.05E+07
TL.169.ALA.H	8.2750	124.9759	19.4096	19.6958	2.71E+07
TL.170.ALA.H	8.2009	121.6082	19.1281	18.5205	2.58E+07
TL.174.VAL.H	7.9258	119.5189	26.7477	26.1396	8.34E+06
TL.175.VAL.H	8.3137	125.5231	16.9953	16.9413	7.98E+07
TL.176.LYS.H	8.4697	126.6299	26.1100	20.5983	2.90E+06
TL.177.ALA.H	8.3806	125.7978	28.0424	18.8813	2.71E+06
TL.178.GLU.H	8.4483	120.7051	#VALUE!	#VALUE!	9.69E+06
TL.179.LYS.H	8.4452	122.3144	19.8139	22.4131	2.10E+07
TL.185.GLU.H	8.2803	122.8473	18.9947	20.4627	2.54E+07

Suppl. data A4-1-C HMGB1-c038 (89-174, biotinylated) titration with CXCL12A-c021 at 0, 0.42, 0.82 and 1.42 molar equivalents

Due to protein amount limitations, titration was performed by sequential addition of CXCL12 to HMGB1 samples, resulting in sample dilution. As the calculations in the chemical shift tracking module in CCPNMR are independent of peak, this does not alter the results; the median change has been indicated in the volume comparisons.

	BRUKER AVIII HD 500, 5 MM CPTCI 1H-13C/15N/D Z-GRD PROBE
^{1}H frequency (Hz)	500.01233645
^{1}H sweep width (ppm)	11.9044681870088
^{1}H sweep width (Hz)	5952.38095238095
^{1}HN sweep width (ppm)	10.9286921061064
^{15}N frequency (Hz)	50.6715211382702
^{15}N sweep width (ppm)	32.0372585410812
^{15}N sweep width (Hz)	1623.37662337662
Water suppression gradient pulse width (Hz)	2336.45
Buffer	10 mM HEPES, pH 7.5, 150 mM NaCl

Baseline 1 (Day 1)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.81 mM, HMGB1-c038 (Box B 89-174), C-terminal biotinylation tag
CXCL12 construct and concentration	0 mM
Temperature	25 ºC
Volume	300 µL
Sample pH (after adjusting)	7.8
Number of scans	8

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.89.LYS.H	8.22877	123.15	13.43706	11.76465	2.84E+08
BoxBFL.90.ASP.H	8.06507	119.1836	17.47528	13.0415	71477080
BoxBFL.92.ASN.H	8.38837	116.1932	26.72516	13.63039	40775376
BoxBFL.93.ALA.H	7.31759	123.8361	12.03289	11.75968	2.88E+08
BoxBFL.95.LYS.H	8.54394	123.6499	15.56655	12.66447	1.57E+08
BoxBFL.96.ARG.H	8.31244	123.9103	13.30976	11.8029	3.19E+08
BoxBFL.99.SER.H	7.89322	114.8385	12.94573	13.41732	83696192
BoxBFL.100.ALA.H	9.06159	122.8259	12.5281	12.3534	1.5E+08
BoxBFL.101.PHE.H	8.44257	116.1455	14.2976	12.88207	1.18E+08
BoxBFL.102.PHE.H	8.26477	122.1713	14.09453	12.29629	2.95E+08
BoxBFL.103.LEU.H	8.27759	121.1528	15.09128	13.06243	1.22E+08
BoxBFL.104.PHE.H	8.13387	122.4785	13.88173	12.22486	2.22E+08
BoxBFL.105.CYS.H	8.50131	118.3952	14.09029	13.6882	1.26E+08
BoxBFL.106.SER.H	8.1324	115.0572	11.86819	12.2455	2.18E+08
BoxBFL.107.GLU.H	7.15821	119.2011	15.07209	12.23544	1.97E+08
BoxBFL.108.TYR.H	7.94721	114.6268	17.12946	12.56173	1.29E+08
BoxBFL.109.ARG.H	9.22527	123.6917	14.35796	12.40158	1.33E+08
BoxBFL.111.LYS.H	6.80367	117.7105	15.83919	12.77867	1.48E+08
BoxBFL.112.ILE.H	8.16585	119.7589	15.93766	14.10622	1.3E+08
BoxBFL.113.LYS.H	8.24667	118.3979	14.18155	12.282	1.95E+08
BoxBFL.114.GLY.H	7.6419	103.7578	15.9097	12.86625	1.49E+08
BoxBFL.115.GLU.H	7.50312	119.5592	15.3641	12.25047	1.76E+08
BoxBFL.116.HIS.H	7.82141	114.7108	15.60333	12.79576	1.45E+08
BoxBFL.119.LEU.H	7.36645	121.1961	12.25315	11.84566	2.79E+08
BoxBFL.120.SER.H	9.31869	121.0561	13.67984	12.45399	1.5E+08
BoxBFL.121.ILE.H	8.7507	120.6719	13.37699	12.68743	93999208
BoxBFL.122.GLY.H	8.71122	109.4492	18.2158	13.7384	40704368
BoxBFL.123.ASP.H	7.96237	124.2817	14.31821	13.03276	1.64E+08
BoxBFL.124.VAL.H	8.6379	124.0546	13.5712	12.12567	1.86E+08
BoxBFL.125.ALA.H	7.93334	121.344	12.65998	12.7035	2.76E+08
BoxBFL.126.LYS.H	8.02273	119.6775	12.14567	12.20318	2.67E+08
BoxBFL.127.LYS.H	8.04282	121.3008	13.53438	16.44289	2.72E+08
BoxBFL.128.LEU.H	8.64819	120.1742	13.66523	12.49721	1.57E+08
BoxBFL.129.GLY.H	8.26351	106.0328	15.35471	12.18689	1.69E+08
BoxBFL.130.GLU.H	8.07435	123.2474	13.07314	12.31509	2.26E+08
BoxBFL.131.MET.H	8.77512	118.9617	13.39615	12.51007	2.05E+08
BoxBFL.132.TRP.H	8.71654	122.6792	14.56383	12.46287	1.3E+08
BoxBFL.133.ASN.H	8.17988	117.3613	12.44756	12.16788	2.1E+08
BoxBFL.134.ASN.H	7.59333	115.8407	17.61059	12.39088	1.9E+08
BoxBFL.135.THR.H	7.31004	119.9467	15.60722	12.12025	2.02E+08
BoxBFL.136.ALA.H	9.25013	131.437	12.17023	12.15813	2.27E+08
BoxBFL.137.ALA.H	8.85817	124.999	15.52392	16.42924	11136171
BoxBFL.138.ASP.H	8.99358	115.7103	13.01277	13.43146	1.17E+08
BoxBFL.139.ASP.H	7.34269	117.8836	17.65566	12.28682	1.93E+08
BoxBFL.140.LYS.H	7.81346	119.1053	15.13034	12.58181	1.53E+08

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.141.GLN.H	7.43127	118.216	12.22585	12.9068	2.03E+08
BoxBFL.143.TYR.H	7.21936	115.67	15.42101	12.81604	1.43E+08
BoxBFL.144.GLU.H	8.1673	120.3573	19.36893	15.09454	2.23E+08
BoxBFL.145.LYS.H	9.19147	121.2642	13.95073	13.10497	1.73E+08
BoxBFL.146.LYS.H, BoxBFL.156.LYS.H	7.70091	120.5702	14.84964	12.91125	3.78E+08
BoxBFL.147.ALA.H	8.47542	120.5044	16.38659	13.35601	61059660
BoxBFL.148.ALA.H	8.38049	122.0622	13.43968	12.61324	2.27E+08
BoxBFL.149.LYS.H	7.98075	120.7812	13.45082	13.37624	2.21E+08
BoxBFL.150.LEU.H	8.29642	120.0874	14.65693	13.57267	1.39E+08
BoxBFL.151.LYS.H	8.4373	123.66	13.67398	13.22858	1.71E+08
BoxBFL.152.GLU.H	8.0907	120.0978	14.82538	13.76241	2.37E+08
BoxBFL.153.LYS.H	7.76448	119.7662	14.67797	13.05371	1.95E+08
BoxBFL.154.TYR.H	8.14526	120.4576	#VALUE!	#VALUE!	1.83E+08
BoxBFL.155.GLU.H	8.52465	117.6406	14.85626	13.30374	1.65E+08
BoxBFL.157.ASP.H	8.90183	123.2173	13.33878	13.30286	1.79E+08
BoxBFL.158.ILE.H	9.20351	122.5498	13.98115	12.71351	1.81E+08
BoxBFL.159.ALA.H	7.36561	124.8259	12.81959	12.53986	2.31E+08
BoxBFL.160.ALA.H	7.78333	120.4019	13.47788	13.29924	2.25E+08
BoxBFL.161.TYR.H	8.19889	120.0636	16.81491	18.13778	2.19E+08
BoxBFL.162.ARG.H	8.26173	118.8178	12.98114	12.65685	1.88E+08
BoxBFL.163.ALA.H	7.63926	121.3704	13.43373	12.39034	1.48E+08
BoxBFL.164.LYS.H	7.65612	118.3003	14.35415	12.33745	1.22E+08
BoxBFL.165.GLY.H	8.05505	108.6412	27.92431	13.8698	20239174
BoxBFL.166.LYS.H	8.04195	121.5844	15.77089	13.22022	4.24E+08
BoxBFL.168.ASP.H	8.4885	121.098	12.61027	13.02291	1.96E+08
BoxBFL.169.ALA.H	8.27056	124.8752	14.6408	13.13492	53273208
BoxBFL.170.ALA.H	8.13312	121.3951	14.00858	12.18198	1.5E+08
BoxBFL.172.LYS.H	8.4277	124.6737	7.48588	9.20329	3518457
BoxBFL.173.GLY.H	8.57505	111.1121	15.89317	13.54035	25440228

3D experiments (Day 1-5)

These spectra are not shown due to their 3D nature; data available if requested. Sample is that of Baseline 1.

Experiment type	**3D, ^{15}N EDITED ^{1}H-^{1}H TOCSY-HSQC, THROUGH-SPACE INTERACTION**
Temperature	25 ºC
Volume	300 µL
Number of scans	32

Experiment type	**3D, ^{15}N EDITED ^{1}H-^{1}H NOESY-HSQC, THROUGH-BOND INTERACTION**
Temperature	25 ºC
Volume	300 µL
Number of scans	32

0.42 molar ratio CXCL12/HMG Box (Day 5)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.7 mM, HMGB1-c038 (Box B 94-162), C-terminal biotinylation tag (285 µL of 0.8 mM stock)
CXCL12 construct and concentration	0.29 mM, CXCL12A-c021, wt 1-67, no tag (41 µL of 1.8 mM stock)
Volume	25 ºC
Sample pH (after adjusting)	320 µL
Temperature	7.8
Number of scans	32

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.89.LYS.H	8.23317	123.0882	15.99121	13.36991	8.54E+08
BoxBFL.90.ASP.H	8.07425	119.195	17.03548	12.94404	3.32E+08
BoxBFL.92.ASN.H	8.39252	116.22	20.52278	13.76841	1.57E+08
BoxBFL.93.ALA.H	7.32309	123.8461	13.23603	12.04781	8.1E+08
BoxBFL.95.LYS.H	8.55152	123.6861	15.235	12.47899	5.7E+08
BoxBFL.96.ARG.H	8.33348	124.0077	14.64892	11.53036	6.66E+08
BoxBFL.99.SER.H	7.89458	114.8324	17.06332	13.29929	3.9E+08
BoxBFL.100.ALA.H	9.07237	122.8303	12.45091	12.65528	6.1E+08
BoxBFL.101.PHE.H	8.45267	116.1672	12.95977	12.54238	5.08E+08
BoxBFL.102.PHE.H	8.2712	122.1831	13.80017	13.34009	8.34E+08
BoxBFL.103.LEU.H	8.28642	121.127	15.05638	12.72235	4.44E+08
BoxBFL.104.PHE.H	8.154	122.5034	15.15558	12.55092	4.93E+08
BoxBFL.105.CYS.H	8.50821	118.3965	13.60565	14.11493	5.43E+08
BoxBFL.106.SER.H	8.14042	115.0895	12.31066	12.82245	7.92E+08
BoxBFL.107.GLU.H	7.1606	119.2044	15.04631	12.06946	7.79E+08
BoxBFL.108.TYR.H	7.95517	114.6331	17.1485	12.51247	5.1E+08
BoxBFL.109.ARG.H	9.24196	123.7109	14.77907	12.44116	5.14E+08
BoxBFL.111.LYS.H	6.80383	117.7113	16.35257	12.50023	5.75E+08
BoxBFL.112.ILE.H	8.17228	119.7666	15.12234	13.68265	5.3E+08
BoxBFL.113.LYS.H	8.26058	118.4342	14.68114	12.6851	5.11E+08
BoxBFL.114.GLY.H	7.64965	103.7836	14.15269	12.03673	6.32E+08
BoxBFL.115.GLU.H	7.50893	119.5449	15.10884	12.14579	7E+08
BoxBFL.116.HIS.H	7.83669	114.7278	#VALUE!	#VALUE!	3.31E+08
BoxBFL.119.LEU.H	7.37631	121.2095	11.82599	11.95517	1.11E+09
BoxBFL.120.SER.H	9.32759	121.0738	13.32814	12.27617	6.4E+08
BoxBFL.121.ILE.H	8.75541	120.6669	13.45308	12.74538	4.21E+08
BoxBFL.122.GLY.H	8.72192	109.4783	18.02792	13.29589	1.9E+08
BoxBFL.123.ASP.H	7.97614	124.3289	15.78848	12.93298	5.54E+08
BoxBFL.124.VAL.H	8.6467	124.0535	13.39795	11.91108	7.39E+08
BoxBFL.125.ALA.H	7.93311	121.3373	20.94036	12.50751	7.5E+08
BoxBFL.126.LYS.H	8.03459	119.7401	12.11644	12.969	9.82E+08
BoxBFL.127.LYS.H	8.05395	121.3131	13.08531	12.0213	1.04E+09
BoxBFL.128.LEU.H	8.65647	120.1666	12.77726	12.56328	6.48E+08
BoxBFL.129.GLY.H	8.26041	106.0026	16.56723	12.47707	5.83E+08
BoxBFL.130.GLU.H	8.0809	123.2669	12.42406	12.29425	8.83E+08
BoxBFL.131.MET.H	8.78921	118.9705	13.20277	12.33316	7.89E+08
BoxBFL.132.TRP.H	8.71893	122.6622	13.85963	12.75907	4.98E+08
BoxBFL.133.ASN.H	8.18222	117.3634	13.90003	12.0061	7.53E+08
BoxBFL.134.ASN.H	7.6005	115.8484	17.9579	11.96318	7.22E+08
BoxBFL.135.THR.H	7.31604	119.9459	15.68761	11.89589	7.84E+08
BoxBFL.136.ALA.H	9.25887	131.4427	12.04897	11.90835	9.4E+08
BoxBFL.137.ALA.H	8.86544	124.9937	18.19465	14.33638	71698952
BoxBFL.138.ASP.H	9.00067	115.712	12.28957	13.24047	5.28E+08

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.139.ASP.H	7.34953	117.8877	17.87632	12.18195	7.56E+08
BoxBFL.140.LYS.H	7.82254	119.1251	15.16611	12.34065	6.15E+08
BoxBFL.141.GLN.H	7.44009	118.2284	11.84562	12.39199	8.33E+08
BoxBFL.143.TYR.H	7.22876	115.6684	16.28895	12.57946	5.48E+08
BoxBFL.144.GLU.H	8.18291	120.313	24.30189	16.31598	7.26E+08
BoxBFL.145.LYS.H	9.20235	121.2804	13.87277	13.31669	6.47E+08
BoxBFL.146.LYS.H, BoxBFL.156.LYS.H	7.70895	120.5874	14.10502	13.01551	1.38E+09
BoxBFL.147.ALA.H	8.48426	120.552	16.0703	14.31819	2.91E+08
BoxBFL.148.ALA.H	8.39134	122.0914	13.96231	13.19139	7.81E+08
BoxBFL.149.LYS.H	7.99377	120.7978	13.72404	13.22412	7.88E+08
BoxBFL.150.LEU.H	8.30338	120.106	14.84471	13.15063	5.18E+08
BoxBFL.151.LYS.H	8.44582	123.6944	13.55023	13.11479	6.01E+08
BoxBFL.152.GLU.H	8.10037	120.1197	14.86852	13.77938	8.15E+08
BoxBFL.153.LYS.H	7.77409	119.7673	13.08363	13.46475	7.36E+08
BoxBFL.154.TYR.H	8.15602	120.4585	24.22866	16.12356	6.94E+08
BoxBFL.155.GLU.H	8.53056	117.6515	13.90664	13.6262	5.87E+08
BoxBFL.157.ASP.H	8.91261	123.2303	14.37862	13.03578	5.44E+08
BoxBFL.158.ILE.H	9.21552	122.5667	15.94888	12.94384	4.83E+08
BoxBFL.159.ALA.H	7.36871	124.8542	14.00455	12.99954	6.93E+08
BoxBFL.160.ALA.H	7.79079	120.4098	13.48341	13.45411	7.53E+08
BoxBFL.161.TYR.H	8.20906	120.0582	17.84468	15.5777	8.17E+08
BoxBFL.162.ARG.H	8.26922	118.8094	13.23997	12.86921	7.66E+08
BoxBFL.163.ALA.H	7.64365	121.401	13.62381	13.15062	4.69E+08
BoxBFL.164.LYS.H	7.66395	118.3578	14.08222	14.60296	3.61E+08
BoxBFL.165.GLY.H	8.06949	108.6489	23.00117	14.26944	84792952
BoxBFL.166.LYS.H	8.06804	121.6875	16.22206	12.67618	1.02E+09
BoxBFL.168.ASP.H	8.49698	121.1145	12.93091	12.80312	7.46E+08
BoxBFL.169.ALA.H	8.27748	124.8788	14.08935	13.77469	2.91E+08
BoxBFL.170.ALA.H	8.13889	121.4079	15.50116	12.51434	5.16E+08
BoxBFL.172.LYS.H	8.44213	124.8148	14.07011	12.03688	79426568
BoxBFL.173.GLY.H	8.58133	111.1267	16.71041	13.95112	1.23E+08

0.84 molar ratio CXCL12/HMG Box (Day 5)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.58 mM, HMGB1-c038 (Box B 94-162), C-terminal biotinylation tag (285 µL of 0.8 mM stock)
CXCL12 construct and concentration	0.482 mM, CXCL12A-c021, wt 1-67, no tag (104 µL of 1.8 mM stock)
Volume	25 ºC
Sample pH (after adjusting)	320 µL
Temperature	7.8
Number of scans	32

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.89.LYS.H	8.22679	122.98562	19.92847	15.30362	6.04E+08
BoxBFL.90.ASP.H	8.07636	119.19941	17.45582	13.22061	2.24E+08
BoxBFL.92.ASN.H	8.39239	116.22614	23.07716	14.06705	1.06E+08
BoxBFL.93.ALA.H	7.32259	123.8468	13.46498	12.02831	6.25E+08
BoxBFL.95.LYS.H	8.55168	123.69223	15.31781	12.58495	4.41E+08
BoxBFL.96.ARG.H	8.34299	124.0494	14.88799	11.9875	4.06E+08
BoxBFL.99.SER.H	7.89423	114.84247	#VALUE!	#VALUE!	3.07E+08
BoxBFL.100.ALA.H	9.07369	122.83003	11.94043	12.61474	5.02E+08
BoxBFL.101.PHE.H	8.45401	116.1693	13.37439	12.53442	3.98E+08
BoxBFL.102.PHE.H	8.27231	122.15451	13.98629	14.5081	5.11E+08
BoxBFL.103.LEU.H	8.28889	121.12078	14.37061	12.72043	3.73E+08
BoxBFL.104.PHE.H	8.16438	122.51969	15.78414	13.39462	2.93E+08
BoxBFL.105.CYS.H	8.50866	118.40024	13.19461	13.85193	4.4E+08
BoxBFL.106.SER.H	8.1423	115.09842	11.9209	12.43735	6.58E+08
BoxBFL.107.GLU.H	7.16005	119.20473	14.95609	12.03505	6.36E+08
BoxBFL.108.TYR.H	7.95609	114.63347	16.8564	12.6025	4.15E+08
BoxBFL.109.ARG.H	9.24419	123.71548	14.94347	12.70613	4.15E+08
BoxBFL.111.LYS.H	6.80375	117.70835	16.02119	12.42776	4.77E+08
BoxBFL.112.ILE.H	8.17303	119.76733	15.02702	13.64763	4.37E+08
BoxBFL.113.LYS.H	8.26665	118.45194	14.62975	13.06685	3.35E+08
BoxBFL.114.GLY.H	7.64978	103.7787	14.53777	12.13962	5.15E+08
BoxBFL.115.GLU.H	7.50983	119.55375	15.22445	11.99963	5.72E+08
BoxBFL.116.HIS.H	7.84327	114.73761	#VALUE!	#VALUE!	2.63E+08
BoxBFL.119.LEU.H	7.37738	121.21246	11.7834	12.0448	9.01E+08
BoxBFL.120.SER.H	9.32827	121.07659	13.31064	12.27752	5.09E+08
BoxBFL.121.ILE.H	8.7561	120.66594	13.60291	12.71327	3.22E+08
BoxBFL.122.GLY.H	8.72507	109.47718	17.30426	13.44831	1.43E+08
BoxBFL.123.ASP.H	7.97787	124.33699	16.31857	12.97771	4.33E+08
BoxBFL.124.VAL.H	8.6484	124.05721	12.86612	11.89578	6.24E+08
BoxBFL.125.ALA.H	7.92439	121.32801	19.18773	11.92173	6.11E+08
BoxBFL.126.LYS.H	8.0362	119.75212	12.45829	12.85125	7.81E+08
BoxBFL.127.LYS.H	8.05575	121.31674	12.43276	12.04567	8.57E+08
BoxBFL.128.LEU.H	8.65767	120.16374	13.11087	12.67279	5.26E+08
BoxBFL.129.GLY.H	8.25962	105.99904	16.21275	12.32688	4.86E+08
BoxBFL.130.GLU.H	8.08214	123.27002	12.67708	12.37817	7.09E+08
BoxBFL.131.MET.H	8.79146	118.97292	13.56391	12.24356	6.42E+08
BoxBFL.132.TRP.H	8.71828	122.65646	13.66393	12.75142	4.11E+08
BoxBFL.133.ASN.H	8.1825	117.36324	13.94004	11.96683	6.19E+08
BoxBFL.134.ASN.H	7.60102	115.8515	19.86751	11.87762	5.95E+08
BoxBFL.135.THR.H	7.31691	119.94345	15.84212	11.78759	6.45E+08
BoxBFL.136.ALA.H	9.25983	131.44194	11.68396	11.88563	7.59E+08
BoxBFL.137.ALA.H	8.86664	124.98894	17.53135	14.89353	51062092
BoxBFL.138.ASP.H	9.0022	115.71168	12.51951	13.18965	4.13E+08

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.139.ASP.H	7.3505	117.8869	19.06678	12.10936	6.11E+08
BoxBFL.140.LYS.H	7.82332	119.12697	14.98048	12.40774	4.99E+08
BoxBFL.141.GLN.H	7.44172	118.2299	11.81727	12.36063	6.81E+08
BoxBFL.143.TYR.H	7.23052	115.66761	15.72802	12.52565	4.53E+08
BoxBFL.144.GLU.H	8.17729	120.35761	20.14083	15.18319	5.71E+08
BoxBFL.145.LYS.H	9.20417	121.2843	13.91087	13.13152	5.12E+08
BoxBFL.146.LYS.H, BoxBFL.156.LYS.H	7.70982	120.59018	14.12825	13.12581	1.11E+09
BoxBFL.147.ALA.H	8.48613	120.5644	16.41386	14.34424	2.09E+08
BoxBFL.148.ALA.H	8.39403	122.09686	13.11396	13.17476	6.52E+08
BoxBFL.149.LYS.H	7.99631	120.80367	13.52788	13.03509	6.35E+08
BoxBFL.150.LEU.H	8.30472	120.10823	15.05319	13.15084	4.31E+08
BoxBFL.151.LYS.H	8.44699	123.70209	13.71114	13.30569	4.8E+08
BoxBFL.152.GLU.H	8.10211	120.12205	13.45462	13.68813	6.67E+08
BoxBFL.153.LYS.H	7.77558	119.76647	13.7091	13.64753	5.96E+08
BoxBFL.154.TYR.H	8.15632	120.46388	21.61287	16.30266	5.62E+08
BoxBFL.155.GLU.H	8.53123	117.6528	13.66215	13.47564	4.65E+08
BoxBFL.157.ASP.H	8.91586	123.23537	14.3729	12.90289	4.42E+08
BoxBFL.158.ILE.H	9.21936	122.58191	16.30721	13.2986	3.82E+08
BoxBFL.159.ALA.H	7.36846	124.86417	13.94085	12.75957	5.54E+08
BoxBFL.160.ALA.H	7.79128	120.4108	13.753	13.44963	5.87E+08
BoxBFL.161.TYR.H	8.21265	120.05605	16.82215	14.29685	6.46E+08
BoxBFL.162.ARG.H	8.27145	118.80874	12.38965	12.75367	6.47E+08
BoxBFL.163.ALA.H	7.64337	121.40884	14.1601	13.35172	3.57E+08
BoxBFL.164.LYS.H	7.66408	118.38396	14.21213	14.54701	2.69E+08
BoxBFL.165.GLY.H	8.07305	108.67299	20.01365	13.50363	59059740
BoxBFL.166.LYS.H	8.07769	121.72575	15.87162	13.20124	6.24E+08
BoxBFL.168.ASP.H	8.4982	121.11929	12.80238	12.73258	6.09E+08
BoxBFL.169.ALA.H	8.28032	124.89477	15.56254	15.33651	1.81E+08
BoxBFL.170.ALA.H	8.14155	121.40211	15.22818	12.63552	3.56E+08
BoxBFL.172.LYS.H	8.44239	124.80508	14.17767	12.23471	67949384
BoxBFL.173.GLY.H	8.58353	111.11154	15.74456	14.03704	74405800

1.43 molar ratio CXCL12/HMG Box (Day 5)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.48 mM, HMGB1-c038 (Box B 94-162), C-terminal biotinylation tag (285 µL of 0.8 mM stock)
CXCL12 construct and concentration	0.69 mM, CXCL12A-c021, wt 1-67, no tag (171 µL of 1.8 mM stock)
Temperature	25 ºC
Volume	320 µL
Sample pH (after adjusting)	7.8
Number of scans	32

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.89.LYS.H	8.21895	122.9077	18.92121	17.02404	5.39E+08
BoxBFL.90.ASP.H	8.07743	119.2044	17.38461	13.70778	1.57E+08
BoxBFL.92.ASN.H	8.39578	116.2228	22.65251	14.06046	78748568
BoxBFL.93.ALA.H	7.3206	123.844	13.85399	11.97782	5.01E+08
BoxBFL.95.LYS.H	8.55161	123.6949	15.60312	12.56883	3.5E+08
BoxBFL.96.ARG.H	8.34894	124.0718	15.43549	12.28544	2.6E+08
BoxBFL.99.SER.H	7.89399	114.8337	#VALUE!	#VALUE!	2.52E+08
BoxBFL.100.ALA.H	9.07401	122.83	12.0362	12.69173	4.12E+08
BoxBFL.101.PHE.H	8.45444	116.1709	13.37824	12.49155	3.25E+08
BoxBFL.102.PHE.H	8.27168	122.1244	14.24496	16.08647	3.25E+08
BoxBFL.103.LEU.H	8.28943	121.1158	14.34292	12.71141	3.12E+08
BoxBFL.104.PHE.H	8.17112	122.5365	17.78199	16.29353	1.94E+08
BoxBFL.105.CYS.H	8.50875	118.4054	12.72396	13.22091	3.64E+08
BoxBFL.106.SER.H	8.14279	115.1046	11.79716	12.39792	5.45E+08
BoxBFL.107.GLU.H	7.15923	119.2044	15.5058	12.08644	5.2E+08
BoxBFL.108.TYR.H	7.95617	114.6327	17.84136	12.54391	3.41E+08
BoxBFL.109.ARG.H	9.24629	123.7188	14.08981	12.70662	3.49E+08
BoxBFL.111.LYS.H	6.8034	117.7047	16.29311	12.28933	4E+08
BoxBFL.112.ILE.H	8.17337	119.7674	14.76928	13.50054	3.6E+08
BoxBFL.113.LYS.H	8.27134	118.4462	#VALUE!	#VALUE!	2.17E+08
BoxBFL.114.GLY.H	7.64931	103.7749	14.43105	12.23031	4.15E+08
BoxBFL.115.GLU.H	7.50986	119.5617	14.96744	11.95792	4.69E+08
BoxBFL.116.HIS.H	7.87127	114.7661	19.92977	12.94964	2.92E+08
BoxBFL.119.LEU.H	7.37742	121.2158	11.72553	12.11307	7.45E+08
BoxBFL.120.SER.H	9.32858	121.0789	13.18459	12.45593	4.18E+08
BoxBFL.121.ILE.H	8.7553	120.6649	13.54667	12.77952	2.56E+08
BoxBFL.122.GLY.H	8.72559	109.4867	18.08209	13.61802	1.03E+08
BoxBFL.123.ASP.H	7.98069	124.3565	15.26493	12.68456	3.59E+08
BoxBFL.124.VAL.H	8.6488	124.0599	12.70186	11.8932	5.19E+08
BoxBFL.125.ALA.H	7.92234	121.3274	19.67901	12.01544	5.27E+08
BoxBFL.126.LYS.H	8.03667	119.7712	12.36815	13.1321	6.59E+08
BoxBFL.127.LYS.H	8.05631	121.3212	12.09821	12.05342	7.14E+08
BoxBFL.128.LEU.H	8.65776	120.1602	13.10106	12.67615	4.37E+08
BoxBFL.129.GLY.H	8.25895	105.9973	16.02096	12.32461	4.1E+08
BoxBFL.130.GLU.H	8.08217	123.2734	12.61882	12.35737	5.86E+08
BoxBFL.131.MET.H	8.79223	118.9771	12.92803	12.14025	5.33E+08
BoxBFL.132.TRP.H	8.71806	122.6543	13.59301	12.72152	3.45E+08
BoxBFL.133.ASN.H	8.18174	117.3624	14.1966	12.05599	5.03E+08
BoxBFL.134.ASN.H	7.60142	115.8532	19.14308	11.90938	4.93E+08
BoxBFL.135.THR.H	7.31694	119.9423	15.52944	11.84047	5.33E+08
BoxBFL.136.ALA.H	9.25989	131.4404	11.56485	11.90864	6.26E+08
BoxBFL.137.ALA.H	8.86475	125.0108	16.06168	15.67273	35378044
BoxBFL.138.ASP.H	9.00211	115.711	12.86091	13.34747	3.22E+08

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.139.ASP.H	7.35096	117.8866	19.0693	12.15302	5.06E+08
BoxBFL.140.LYS.H	7.82374	119.1295	14.6706	12.41802	4.18E+08
BoxBFL.141.GLN.H	7.44219	118.232	11.85067	12.27391	5.63E+08
BoxBFL.143.TYR.H	7.23072	115.6698	15.74818	12.51927	3.75E+08
BoxBFL.144.GLU.H	8.17663	120.3647	19.90498	14.8137	4.57E+08
BoxBFL.145.LYS.H	9.20511	121.2878	14.30353	13.0189	4.26E+08
BoxBFL.146.LYS.H, BoxBFL.156.LYS.H	7.71006	120.5924	14.14949	13.23139	9.14E+08
BoxBFL.147.ALA.H	8.48712	120.5785	16.9957	14.34865	1.49E+08
BoxBFL.148.ALA.H	8.39484	122.1015	12.93976	12.98248	5.41E+08
BoxBFL.149.LYS.H	7.99731	120.8071	13.2779	13.16866	5.32E+08
BoxBFL.150.LEU.H	8.3049	120.1088	14.94034	13.2186	3.55E+08
BoxBFL.151.LYS.H	8.4477	123.7082	13.98336	13.38566	3.93E+08
BoxBFL.152.GLU.H	8.10253	120.1254	13.30973	13.69574	5.51E+08
BoxBFL.153.LYS.H	7.77618	119.767	13.49419	13.56752	4.88E+08
BoxBFL.154.TYR.H	8.15722	120.4673	22.68082	16.71144	4.54E+08
BoxBFL.155.GLU.H	8.53108	117.6557	13.40247	13.49932	3.86E+08
BoxBFL.157.ASP.H	8.91702	123.2376	13.94689	12.93051	3.63E+08
BoxBFL.158.ILE.H	9.22334	122.5976	15.59077	13.35733	3.16E+08
BoxBFL.159.ALA.H	7.36744	124.8739	13.98725	12.61412	4.56E+08
BoxBFL.160.ALA.H	7.79157	120.411	13.64762	13.47195	4.83E+08
BoxBFL.161.TYR.H	8.21447	120.0592	16.95219	14.11511	5.1E+08
BoxBFL.162.ARG.H	8.27207	118.8079	12.06017	12.68513	5.39E+08
BoxBFL.163.ALA.H	7.64246	121.4127	13.9688	13.29345	2.85E+08
BoxBFL.164.LYS.H	7.66442	118.4079	14.86267	15.95431	2.05E+08
BoxBFL.165.GLY.H	8.07387	108.6809	24.49469	14.82814	37366064
BoxBFL.166.LYS.H	8.08076	121.7464	18.19369	13.05761	3.83E+08
BoxBFL.168.ASP.H	8.49868	121.1201	12.66135	12.73158	5.01E+08
BoxBFL.169.ALA.H	8.28128	124.9183	14.6585	15.5819	1.24E+08
BoxBFL.170.ALA.H	8.14173	121.3982	16.83083	12.7598	2.45E+08
BoxBFL.172.LYS.H	8.44324	124.8133	14.79008	12.58927	50445744
BoxBFL.173.GLY.H	8.57949	111.0939	20.28256	15.57476	46002192

Baseline 2 (Day 5)

Experiment type	**2D ^{1}H-^{15}N HSQC**
HMGB1 construct and concentration	0.8 mM, HMGB1-c038 (Box B 94-162), C-terminal biotinylation tag
CXCL12 construct and concentration	0 mM
Temperature	25 ºC
Volume	285 µL
Sample pH (after adjusting)	7.8
Number of scans	32

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.89.LYS.H	8.23102	123.1415	14.76968	12.01671	1.13E+09
BoxBFL.90.ASP.H	8.06917	119.1829	17.02515	12.22975	3.96E+08
BoxBFL.92.ASN.H	8.38929	116.2163	20.63226	13.77677	2.02E+08
BoxBFL.93.ALA.H	7.32084	123.8384	14.33885	11.63902	8.9E+08
BoxBFL.95.LYS.H	8.5478	123.6697	15.96042	12.62552	6.07E+08
BoxBFL.96.ARG.H	8.3189	123.9301	14.50868	11.3829	1.03E+09
BoxBFL.99.SER.H	7.89207	114.8246	15.31161	12.80624	3.66E+08
BoxBFL.100.ALA.H	9.06739	122.8207	14.07546	12.19279	5.84E+08
BoxBFL.101.PHE.H	8.44676	116.1525	14.74864	12.91541	4.61E+08
BoxBFL.102.PHE.H	8.26698	122.1883	14.22121	12.44991	1.11E+09
BoxBFL.103.LEU.H	8.28133	121.1339	16.10868	12.48103	4.42E+08
BoxBFL.104.PHE.H	8.13782	122.4713	14.37141	12.44314	7.5E+08
BoxBFL.105.CYS.H	8.50419	118.3867	15.14844	14.2131	4.87E+08
BoxBFL.106.SER.H	8.13524	115.0683	13.16398	12.31969	7.42E+08
BoxBFL.107.GLU.H	7.15834	119.1995	15.9825	11.95677	7.25E+08
BoxBFL.108.TYR.H	7.9497	114.6278	17.47581	12.33864	4.82E+08
BoxBFL.109.ARG.H	9.23234	123.6965	16.59866	12.29412	4.81E+08
BoxBFL.111.LYS.H	6.80327	117.711	16.90439	12.46518	5.6E+08
BoxBFL.112.ILE.H	8.16856	119.7594	16.15624	13.91137	4.88E+08
BoxBFL.113.LYS.H	8.24926	118.3993	14.3358	12.1708	6.95E+08
BoxBFL.114.GLY.H	7.64664	103.781	15.76333	12.14132	5.87E+08
BoxBFL.115.GLU.H	7.50528	119.5289	15.88641	12.53165	6.74E+08
BoxBFL.116.HIS.H	7.82554	114.7129	17.52591	12.91339	4.1E+08
BoxBFL.119.LEU.H	7.37163	121.1972	13.24197	11.65462	1.03E+09
BoxBFL.120.SER.H	9.32283	121.0625	15.14154	12.10113	5.93E+08
BoxBFL.121.ILE.H	8.7522	120.6657	14.87619	12.56095	4.04E+08
BoxBFL.122.GLY.H	8.71514	109.4591	18.21003	13.47553	2E+08
BoxBFL.123.ASP.H	7.96996	124.3035	16.15828	12.66772	5.78E+08
BoxBFL.124.VAL.H	8.64104	124.0413	14.34107	12.30741	7.01E+08
BoxBFL.125.ALA.H	7.93386	121.3372	15.70871	12.46088	8.51E+08
BoxBFL.126.LYS.H	8.02815	119.7021	13.46969	12.7483	9.18E+08
BoxBFL.127.LYS.H	8.0478	121.2943	14.12423	15.42969	9.72E+08
BoxBFL.128.LEU.H	8.65186	120.1671	14.50135	12.59408	5.79E+08
BoxBFL.129.GLY.H	8.26077	106.0107	16.79447	12.54485	5.75E+08
BoxBFL.130.GLU.H	8.07653	123.2546	14.00815	12.1069	8.25E+08
BoxBFL.131.MET.H	8.7821	118.9601	14.7908	12.21398	7.26E+08
BoxBFL.132.TRP.H	8.71726	122.6675	15.33509	12.4674	4.67E+08
BoxBFL.133.ASN.H	8.17966	117.3578	14.21937	12.09116	7.35E+08
BoxBFL.134.ASN.H	7.59622	115.839	17.11067	12.24948	7.11E+08
BoxBFL.135.THR.H	7.31205	119.9435	16.01333	11.84476	7.59E+08
BoxBFL.136.ALA.H	9.25411	131.4374	13.61094	12.00251	8.78E+08
BoxBFL.137.ALA.H	8.86096	124.9828	16.96746	13.9175	82783192
BoxBFL.138.ASP.H	8.99667	115.7095	14.2258	12.81374	5.22E+08
BoxBFL.139.ASP.H	7.34551	117.8838	17.77517	12.06753	7.26E+08
BoxBFL.140.LYS.H	7.81795	119.1111	16.26883	12.4296	5.8E+08

Assign	F1 position	F2 position	LW F1 (Hz)	LW F2 (Hz)	Height
BoxBFL.141.GLN.H	7.4356	118.2188	13.44592	12.84284	7.56E+08
BoxBFL.143.TYR.H	7.22342	115.6648	17.19563	12.6802	5.23E+08
BoxBFL.144.GLU.H	8.17886	120.3107	24.6535	15.3662	8.3E+08
BoxBFL.145.LYS.H	9.19676	121.2667	15.33294	12.93118	6.33E+08
BoxBFL.146.LYS.H, BoxBFL.156.LYS.H	7.70415	120.5771	15.33781	12.75619	1.39E+09
BoxBFL.147.ALA.H	8.47838	120.5221	16.72412	13.05465	3.2E+08
BoxBFL.148.ALA.H	8.38608	122.0761	15.38089	12.69376	7.72E+08
BoxBFL.149.LYS.H	7.98664	120.7873	15.37662	13.186	7.8E+08
BoxBFL.150.LEU.H	8.29876	120.0933	15.78188	13.31339	5.08E+08
BoxBFL.151.LYS.H	8.44117	123.678	14.86959	12.979	6.28E+08
BoxBFL.152.GLU.H	8.09634	120.1123	17.64454	13.74299	8.61E+08
BoxBFL.153.LYS.H	7.76844	119.7644	14.87574	12.75453	7.39E+08
BoxBFL.154.TYR.H	8.14945	120.4238	#VALUE!	#VALUE!	7.23E+08
BoxBFL.155.GLU.H	8.52687	117.6413	15.0031	13.04017	6.17E+08
BoxBFL.157.ASP.H	8.906	123.2206	15.30218	13.15	6.04E+08
BoxBFL.158.ILE.H	9.20733	122.5499	17.72024	12.87546	5.48E+08
BoxBFL.159.ALA.H	7.36719	124.835	14.8669	12.75735	7.55E+08
BoxBFL.160.ALA.H	7.78634	120.3982	14.86213	13.04298	7.84E+08
BoxBFL.161.TYR.H	8.20162	120.0455	17.64211	17.63803	8.72E+08
BoxBFL.162.ARG.H	8.26363	118.8125	15.04827	12.83358	7.06E+08
BoxBFL.163.ALA.H	7.64059	121.3835	14.5465	12.45633	5.34E+08
BoxBFL.164.LYS.H	7.65996	118.3228	15.81194	12.97592	4.19E+08
BoxBFL.165.GLY.H	8.0589	108.6345	19.5729	13.76712	1.13E+08
BoxBFL.166.LYS.H	8.04566	121.5924	15.74549	13.70187	1.31E+09
BoxBFL.168.ASP.H	8.49244	121.104	14.18502	12.88319	7.2E+08
BoxBFL.169.ALA.H	8.2712	124.8749	14.52477	12.28691	3.58E+08
BoxBFL.170.ALA.H	8.13452	121.4021	14.64616	11.71461	6.15E+08
BoxBFL.172.LYS.H	8.42733	124.7554	12.80903	10.72519	46624344
BoxBFL.173.GLY.H	8.57669	111.124	19.67674	13.22967	1.55E+08

Appendix A5 Statistics

Chapter 3 Data analysis

Suppl. data A5-3-A Figure 3.6, pp. 80

A. Residual distribution was Gaussian, and all groups had equal variance (linear QQ plot, rectangular homoscedasticity plot). A nested one-way ANOVA was applied with treatment (HMGB1 construct; n=number of animals, 5) as the main column variable and individual animals as subcolumns ($n_{subcolumn}$=number of cells, ≈ 8-10 per animal). 1 family.

Table Analysed	**Data 7**	
Nested one-way ANOVA		
P value	<0.0001	
P value summary	****	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	10.08, 5, 24	
Random effects	SD	Variance
Variation within subcolumns	335.9	112811
Variation among subcolumn means	131.0	17165
Do the subcolumns differ (within each column)?		
Chi-square, df	9.615, 1	
P value	0.0019	
P value summary	**	

Holm-Sidak's multiple comparisons test	**Mean Diff.**	**Adjusted P Value**	**N1**	**N2**	**Mean 1**	**Mean 2**
Control vs. c001	-595.2	0.0002	42	44	1012	1607
Control vs. c002	-445.3	0.0053	42	43	1012	1457
Control vs. c004	-109.7	0.7990	42	43	1012	1122
Control vs. c005	-81.41	0.8494	42	43	1012	1093
Control vs. c006	-458.6	0.0043	42	43	1012	1471
c001 vs. c002	149.8	0.7077	44	43	1607	1457
c001 vs. c004	485.5	0.0024	44	43	1607	1122
c001 vs. c005	513.8	0.0013	44	43	1607	1093
c001 vs. c006	136.6	0.7218	44	43	1607	1471
c002 vs. c004	335.6	0.0383	43	43	1457	1122
c002 vs. c005	363.9	0.0265	43	43	1457	1093
c002 vs. c006	-13.21	0.9598	43	43	1457	1471
c004 vs. c005	28.29	0.9598	43	43	1122	1093
c004 vs. c006	-348.9	0.0329	43	43	1122	1471
c005 vs. c006	-377.2	0.0220	43	43	1093	1471

B. Residual distribution was Gaussian with unequal variance (linear QQ plot with skew at the edges, triangular homoscedasticity plot, N insufficient to compute variance tests). Data were analysed via Kruskal-Wallis test, with n= number of fields. α= 0.05.

Kruskal-Wallis test	
P value	0.0005
Exact or approximate P value?	Approximate
P value summary	***
Do the medians vary signif. (P < 0.05)?	Yes
Number of groups	8
Kruskal-Wallis statistic	25.79

Dunn´s Multiple comparison Test	**Mean rank 1**	**Mean rank 2**	**Adj.P value**	**Mean rank diff.**	**n1**	**n2**	**Z**
0% HI-FBS vs. 20% HI-FBS	3.750	40.17	0.0001	-36.42	6	6	4.338
0% HI-FBS vs. HMGB1-c006 (FL 3S)	3.750	31.83	0.0008	-28.08	6	12	3.863
0% HI-FBS vs. HMGB1-c007 (1-184 3S)	3.750	30.83	0.0088	-27.08	6	6	3.226
0% HI-FBS vs. HMGB1-c008 (9-184 3S)	3.750	18.50	0.6571	-14.75	6	5	1.675
0% HI-FBS vs. HMGB1-c009 (9-215 3S)	3.750	23.42	0.1340	-19.67	6	6	2.343
0% HI-FBS vs. HMGB1-c010 (9-164 3S)	3.750	14.50	>0.9999	-10.75	6	3	1.046
0% HI-FBS vs. CXCL12	3.750	28.00	0.0271	-24.25	6	6	2.889

Suppl. data A5-3-B Figure 3.9, pp. 84

Residual distribution was non-Gaussian with unequal variance between groups (non-uniform heteroskedasticity plot, sigmoidal QQ plot). Data analysed by Kruskal-Wallis test, 1 family.

A.

Kruskal-Wallis test	
P value	0.0178
Exact or approximate P value?	Approximate
P value summary	*
Do the medians vary signif. (P < 0.05)?	Yes
Number of groups	3
Kruskal-Wallis statistic	8.062

Dunn's multiple comparisons test	Mean rank diff.	Mean rank 1	Mean rank 2	n1	n2	Adjusted P Value
Original protocol vs. High salt wash	19.99	25.24	5.250	17	4	0.0124
Original protocol vs. Split-wash	0.8603	25.24	24.38	17	24	>0.9999

B.

Kruskal-Wallis test	
P value	<0.0001
Exact or approximate P value?	Approximate
P value summary	****
Do the medians vary signif. (P < 0.05)?	Yes
Number of groups	3
Kruskal-Wallis statistic	30.53

Dunn's multiple comparisons test

Test details	**Mean rank 1**	**Mean rank 2**	**Mean rank diff.**	**n1**	**n2**	**Adjusted P Value**
Original protocol vs. High salt wash	9.471	23.88	-14.40	17	4	0.0963
Original protocol vs. Split-wash	9.471	32.44	-22.97	17	24	<0.0001

Suppl. data A5-3-C Figure 3.15, pp. 90

Residual distribution was Gaussian, and all groups had equal variance (linear QQ plot, rectangular homoscedasticity plot). A nested one-way ANOVA was applied with treatment (HMGB1 construct; n=number of animals, 5) as the main column variable and individual animals as subcolumns ($n_{subcolumn}$=number of cells, ≈ 40 per animal), 1 family.

Nested one-way ANOVA		
P value	<0.0001	
P value summary	****	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	20.39, 5, 24	
Random effects	SD	Variance
Variation within subcolumns	311.2	96839
Variation among subcolumn means	96.25	9264
Do the subcolumns differ (within each column)?		
Chi-square, df	87.47, 1	
P value	<0.0001	
P value summary	****	

Holm-Sidak's multiple comparisons test	Mean Diff.	Adjusted P Value	Mean 1	Mean 2	N1	N2
Control (original experiment) vs. c001 (original experiment)	-596.1	<0.0001	1011	1607	42	44
Control (original experiment) vs. c002 (original experiment)	-446.0	0.0003	1011	1457	42	43
Control (original experiment) vs. c031	-237.4	0.0181	1011	1249	42	482
c001 (original experiment) vs. c002 (original experiment)	150.1	0.2075	1607	1457	44	43
c001 (original experiment) vs. c031	358.7	0.0006	1607	1249	44	482
c001 (original experiment) vs. c001 (parallel to c031)	559.6	<0.0001	1607	1048	44	609
c001 (parallel to c031) vs. c031	-200.9	0.0175	1048	1249	609	482
Control (parallel to c031) vs. c001 (parallel to c031)	-70.90	0.2753	977.0	1048	618	609

Suppl. data A5-3-D Figure 8.1, pp. - 81 -

Retention volumes (average of 3 runs) were interpolated in GraphPad via the Interpolate Standard Curve function, against surface area of each standard protein marker calculated from PDB files by nonlinear least squares estimation, to a third order polynomial function. Adjusted R^2=0.9615, runs test p= 0.9. Second order polynomial fitting resulted in lower MW proteins having too high theoretical RV compared to the experimental data.

Suppl. data A5-3-E Figure 3.18B, pp. 96

This analysis includes the data referenced in 0 (HMGB1-C1-c048).

For all groups: α= 0.05, 3 replicates (injection events) each collecting the stream for at least 10 counts after the ion stream became stable.

B. Ion count comparisons Residual distribution was Gaussian and variances between groups were equal according to both statistical tests and residual/Q-Q plots. Data were analysed by 2-way ANOVA with the conformation (2 groups, extended and compact monomer) and the constructs (8 groups) as the variables. 1 family per column (construct).

	% of total variation	P value	Significant?	
Interaction	31.05	<0.0001	Yes	
Construct	1.994	0.4657	No	
Conformation	53.99	<0.0001	Yes	
ANOVA table	SS (Type III)	DF	MS	F (DFn, DFd)
Interaction	0.04723	7	0.006747	F (7, 30) = 15.21
Construct	0.003034	7	0.0004334	F (7, 30) = 0.9772
Conformation	0.08214	1	0.08214	F (1, 30) = 185.2
Residual	0.01330	30	0.0004435	
Difference between column means				
Predicted (LS) mean of Compact monomer			0.4364	
Predicted (LS) mean of Partially unfolded/extended monomer			0.5217	
Difference between predicted means			-0.08528	
SE of difference			0.006266	
95% confidence interval of difference			-0.09808 to -0.07248	
Spearman's test for heteroscedasticity				
Rs of predicted Y vs. |residual|			0.02614	
P value (one tailed)			0.4316	
Passed (P > 0.05)?			Yes	
Normality of Residuals	Statistics	P value	P value summary	
Anderson-Darling (A2*)	0.6344	0.0922	ns	
D'Agostino-Pearson omnibus (K2)	3.577	0.1672	ns	
Shapiro-Wilk (W)	0.9670	0.2127	ns	
Kolmogorov-Smirnov (distance)	0.1008	0.1000	ns	

Holm-Sidak's multiple comparisons test	Predicted (LS) mean 1	Predicted (LS) mean 2	Predicted (LS) mean diff.	Adjusted p value	N1	N2	t	DF
Compact monomer								
FR, 1-214 vs. DS, 1-214	0.4692	0.4584	0.01080	0.9987	3	2	0.5618	30
FR, 1-214 vs. FR, 1-184	0.4692	0.3887	0.08047	0.0016	3	3	4.680	30
FR, 1-214 vs. DS, 1-184	0.4692	0.3999	0.06927	0.0095	3	3	4.028	30
FR, 1-214 vs. FR, 1-164	0.4692	0.4407	0.02853	0.8378	3	3	1.659	30
FR, 1-214 vs. DS, 1-164	0.4692	0.4441	0.02510	0.9050	3	3	1.460	30
FR, 1-214 vs. 3S, 1-214	0.4692	0.4432	0.02597	0.8985	3	3	1.510	30
FR, 1-214 vs. DbB, c048	0.4692	0.4471	0.02207	0.9527	3	3	1.283	30
DS, 1-214 vs. FR, 1-184	0.4584	0.3887	0.06967	0.0272	2	3	3.624	30
DS, 1-214 vs. DS, 1-184	0.4584	0.3999	0.05847	0.1016	2	3	3.041	30
DS, 1-214 vs. FR, 1-164	0.4584	0.4407	0.01773	0.9956	2	3	0.9225	30
DS, 1-214 vs. DS, 1-164	0.4584	0.4441	0.01430	0.9982	2	3	0.7439	30
DS, 1-214 vs. 3S, 1-214	0.4584	0.4432	0.01517	0.9982	2	3	0.7889	30
DS, 1-214 vs. DbB, c048	0.4584	0.4471	0.01127	0.9987	2	3	0.5861	30
FR, 1-184 vs. DS, 1-184	0.3887	0.3999	-0.01120	0.9986	3	3	0.6514	30
FR, 1-184 vs. FR, 1-164	0.3887	0.4407	-0.05193	0.1022	3	3	3.020	30
FR, 1-184 vs. DS, 1-164	0.3887	0.4441	-0.05537	0.0713	3	3	3.220	30
FR, 1-184 vs. 3S, 1-214	0.3887	0.4432	-0.05450	0.0775	3	3	3.170	30
FR, 1-184 vs. DbB, c048	0.3887	0.4471	-0.05840	0.0474	3	3	3.396	30
DS, 1-184 vs. FR, 1-164	0.3999	0.4407	-0.04073	0.3437	3	3	2.369	30
DS, 1-184 vs. DS, 1-164	0.3999	0.4441	-0.04417	0.2558	3	3	2.569	30
DS, 1-184 vs. 3S, 1-214	0.3999	0.4432	-0.04330	0.2704	3	3	2.518	30
DS, 1-184 vs. DbB, c048	0.3999	0.4471	-0.04720	0.1841	3	3	2.745	30
FR, 1-164 vs. DS, 1-164	0.4407	0.4441	-0.003433	0.9998	3	3	0.1997	30
FR, 1-164 vs. 3S, 1-214	0.4407	0.4432	-0.002567	0.9998	3	3	0.1493	30
FR, 1-164 vs. DbB, c048	0.4407	0.4471	-0.006467	0.9994	3	3	0.3761	30
DS, 1-164 vs. 3S, 1-214	0.4441	0.4432	0.0008667	0.9998	3	3	0.05040	30
DS, 1-164 vs. DbB, c048	0.4441	0.4471	-0.003033	0.9998	3	3	0.1764	30
3S, 1-214 vs. DbB, c048	0.4432	0.4471	-0.003900	0.9998	3	3	0.2268	30
Partially unfolded/extended monomer								
FR, 1-214 vs. DS, 1-214	0.4606	0.4868	-0.02617	0.7583	3	2	1.361	30
FR, 1-214 vs. FR, 1-184	0.4606	0.5817	-0.1211	<0.0001	3	3	7.043	30
FR, 1-214 vs. DS, 1-184	0.4606	0.5723	-0.1116	<0.0001	3	3	6.492	30
FR, 1-214 vs. FR, 1-164	0.4606	0.5360	-0.07533	0.0030	3	3	4.381	30
FR, 1-214 vs. DS, 1-164	0.4606	0.5227	-0.06210	0.0217	3	3	3.612	30
FR, 1-214 vs. 3S, 1-214	0.4606	0.4971	-0.03650	0.4035	3	3	2.123	30
FR, 1-214 vs. DbB, c048	0.4606	0.5164	-0.05573	0.0499	3	3	3.241	30
DS, 1-214 vs. FR, 1-184	0.4868	0.5817	-0.09493	0.0007	2	3	4.938	30
DS, 1-214 vs. DS, 1-184	0.4868	0.5723	-0.08547	0.0027	2	3	4.446	30
DS, 1-214 vs. FR, 1-164	0.4868	0.5360	-0.04917	0.2003	2	3	2.558	30
DS, 1-214 vs. DS, 1-164	0.4868	0.5227	-0.03593	0.5232	2	3	1.869	30
DS, 1-214 vs. 3S, 1-214	0.4868	0.4971	-0.01033	0.9290	2	3	0.5375	30
DS, 1-214 vs. DbB, c048	0.4868	0.5164	-0.02957	0.7276	2	3	1.538	30
FR, 1-184 vs. DS, 1-184	0.5817	0.5723	0.009467	0.9290	3	3	0.5506	30
FR, 1-184 vs. FR, 1-164	0.5817	0.5360	0.04577	0.1703	3	3	2.662	30
FR, 1-184 vs. DS, 1-164	0.5817	0.5227	0.05900	0.0331	3	3	3.431	30
FR, 1-184 vs. 3S, 1-214	0.5817	0.4971	0.08460	0.0007	3	3	4.920	30
FR, 1-184 vs. DbB, c048	0.5817	0.5164	0.06537	0.0137	3	3	3.802	30
DS, 1-184 vs. FR, 1-164	0.5723	0.5360	0.03630	0.4035	3	3	2.111	30
DS, 1-184 vs. DS, 1-164	0.5723	0.5227	0.04953	0.1100	3	3	2.881	30
DS, 1-184 vs. 3S, 1-214	0.5723	0.4971	0.07513	0.0030	3	3	4.370	30
DS, 1-184 vs. DbB, c048	0.5723	0.5164	0.05590	0.0499	3	3	3.251	30
FR, 1-164 vs. DS, 1-164	0.5360	0.5227	0.01323	0.9068	3	3	0.7696	30
FR, 1-164 vs. 3S, 1-214	0.5360	0.4971	0.03883	0.3390	3	3	2.258	30
FR, 1-164 vs. DbB, c048	0.5360	0.5164	0.01960	0.8402	3	3	1.140	30
DS, 1-164 vs. 3S, 1-214	0.5227	0.4971	0.02560	0.7276	3	3	1.489	30
DS, 1-164 vs. DbB, c048	0.5227	0.5164	0.006367	0.9290	3	3	0.3703	30
3S, 1-214 vs. DbB, c048	0.4971	0.5164	-0.01923	0.8402	3	3	1.119	30

Suppl. data A5-3-F Figure 3.19B, pp. 98

For all groups: α= 0.05, 3 replicates (injection events) each collecting the stream for at least 10 counts after the ion stream became stable. Residual distribution was Gaussian and variances between groups were equal according to both statistical tests and residual/Q-Q plots. Data were analysed by 2-way ANOVA with the conformation (2 groups, extended and compact monomer) and the constructs (5 groups) as the variables. 1 family per column (construct).

Source of Variation	**% of total variation**	**P value**	**P value summary**	**Significant?**
Interaction	32.64	0.0006	***	Yes
Construct	0.1618	0.9965	ns	No
Conformation	52.89	<0.0001	****	Yes
ANOVA table	SS (Type III)	DF	MS	F (DFn, DFd)
Interaction	0.1315	4	0.03288	F (4, 18) = 8.235
Construct	0.0006521	4	0.0001630	F (4, 18) = 0.04082
Conformation	0.2131	1	0.2131	F (1, 18) = 53.37
Residual	0.07188	18	0.003993	

Difference between column means	
Predicted (LS) mean of Compact monomer	0.5686
Predicted (LS) mean of Partially unfolded/extended monomer	0.3918
Difference between predicted means	0.1768
SE of difference	0.02420
95% confidence interval of difference	0.1260 to 0.2276
Spearman's test for heteroscedasticity	
Rs of predicted Y vs. |residual|	0.03740
P value (one tailed)	0.4251
Passed (P > 0.05)?	Yes

Normality of Residuals	Statistics	P value	P value summary
Anderson-Darling (A2*)	0.5810	0.1181	ns
D'Agostino-Pearson omnibus (K2)	1.408	0.4947	ns
Shapiro-Wilk (W)	0.9574	0.3025	ns
Kolmogorov-Smirnov (distance)	0.1578	0.0721	ns

Holm-Sidak's multiple comparisons test	**Predicted (LS) mean 1**	**Predicted (LS) mean 2**	**Predicted (LS) mean diff.**	**Adjusted P Value**	**N1**	**N2**	**t**	**DF**
Compact monomer								
FR, 1-214 vs. DS, 1-214	0.6250	0.6241	0.0008667	0.9882	3	2	0.01502	18.00
FR, 1-214 vs. FR, 1-184	0.6250	0.4451	0.1799	0.0260	3	3	3.487	18.00
FR, 1-214 vs. FR, 1-164	0.6250	0.5662	0.05873	0.8485	3	3	1.138	18.00
FR, 1-214 vs. 3S, 1-214	0.6250	0.5825	0.04250	0.8875	3	3	0.8237	18.00
DS, 1-214 vs. FR, 1-184	0.6241	0.4451	0.1790	0.0538	2	3	3.104	18.00
DS, 1-214 vs. FR, 1-164	0.6241	0.5662	0.05787	0.8641	2	3	1.003	18.00
DS, 1-214 vs. 3S, 1-214	0.6241	0.5825	0.04163	0.8875	2	3	0.7217	18.00
FR, 1-184 vs. FR, 1-164	0.4451	0.5662	-0.1212	0.1948	3	3	2.348	18.00
FR, 1-184 vs. 3S, 1-214	0.4451	0.5825	-0.1374	0.1200	3	3	2.663	18.00
FR, 1-164 vs. 3S, 1-214	0.5662	0.5825	-0.01623	0.9408	3	3	0.3146	18.00

Partially unfolded/extended monomer								
FR, 1-214 vs. DS, 1-214	0.3351	0.3323	0.002883	0.9607	3	2	0.04998	18.00
FR, 1-214 vs. FR, 1-184	0.3351	0.5218	-0.1867	0.0195	3	3	3.618	18.00
FR, 1-214 vs. FR, 1-164	0.3351	0.4067	-0.07157	0.7012	3	3	1.387	18.00
FR, 1-214 vs. 3S, 1-214	0.3351	0.3629	-0.02780	0.9344	3	3	0.5388	18.00
DS, 1-214 vs. FR, 1-184	0.3323	0.5218	-0.1896	0.0364	2	3	3.286	18.00
DS, 1-214 vs. FR, 1-164	0.3323	0.4067	-0.07445	0.7012	2	3	1.291	18.00
DS, 1-214 vs. 3S, 1-214	0.3323	0.3629	-0.03068	0.9344	2	3	0.5319	18.00
FR, 1-184 vs. FR, 1-164	0.5218	0.4067	0.1151	0.2411	3	3	2.231	18.00
FR, 1-184 vs. 3S, 1-214	0.5218	0.3629	0.1589	0.0506	3	3	3.079	18.00
FR, 1-164 vs. 3S, 1-214	0.4067	0.3629	0.04377	0.8767	3	3	0.8483	18.00

Suppl. data A5-3-G Figure 3.20, pp. 100

This analysis includes the data referenced in Suppl. data A5-5-K (HMGB1-C1-c048).

For all groups: α= 0.05, 3 replicates (injection events) each containing data from ≈20 counts (0.3 min) corresponding to the largest peak in the TIC, at 0.6 min. Residual distribution was Gaussian and variances between groups were equal according to both statistical tests and residual/Q-Q plots. Data were analysed by 2-way ANOVA with the conformation (2 groups, extended and compact monomer) and the constructs (8 groups) as the variables. 1 family per column (construct). Data for the fully folded monomer were analysed separately as otherwise ANOVA assumptions would have been violated (C).

Two-way ANOVA	**Ordinary**			
Alpha	0.05			
Source of Variation	% of total variation	P value	P value summary	Significant?
Interaction	6.037	<0.0001	****	Yes
Construct	0.01048	0.9506	ns	No
Conformation	93.79	<0.0001	****	Yes
ANOVA table	SS	DF	MS	F (DFn, DFd)
Interaction	0.1716	7	0.02451	F (7, 32) = 170.3
Construct	0.0002979	7	4.256e-005	F (7, 32) = 0.2957
Conformation	2.665	1	2.665	F (1, 32) = 18518
Residual	0.004606	32	0.0001439	
Difference between column means				
Mean of Fully unfolded				0.7291
Mean of Extended				0.2578
Difference between means				0.4713
SE of difference				0.003463
95% confidence interval of difference				0.4642 to 0.4783
Spearman's test for heteroscedasticity				
Rs of predicted Y vs. |residual|				0.1550
P value (one tailed)				0.1464
Passed (P > 0.05)?				Yes
Normality of Residuals		Statistics	P value	P value summary
Anderson-Darling (A2*)		0.4834	0.2191	ns
D'Agostino-Pearson omnibus (K2)		2.351	0.3086	ns
Shapiro-Wilk (W)		0.9734	0.3418	ns
Kolmogorov-Smirnov (distance)		0.09824	0.1000	ns

Holm-Sidak's multiple comparisons test	**Mean 1**	**Mean 2**	**Mean Diff.**	**Adjusted P value.**	**N1**	**N2**	**t**	**DF**
Fully unfolded								
FR, 1-214 vs. DS, 1-214	0.6660	0.7207	-0.05467	<0.0001	3	3	5.581	32.00

FR, 1-214 vs. FR, 1-184	0.6660	0.7130	-0.04700	0.0003	3	3	4.798	32.00
FR, 1-214 vs. DS, 1-184*	0.6660	0.7703	-0.1043	<0.0001	3	3	10.65	32.00
FR, 1-214 vs. FR, 1-164	0.6660	0.7747	-0.1087	<0.0001	3	3	11.09	32.00
FR, 1-214 vs. DS, 1-164	0.6660	0.7927	-0.1267	<0.0001	3	3	12.93	32.00
FR, 1-214 vs. 3S, 1-214	0.6660	0.6180	0.04800	0.0002	3	3	4.900	32.00
FR, 1-214 vs. DbB, c048	0.6660	0.7777	-0.1117	<0.0001	3	3	11.40	32.00
DS, 1-214 vs. FR, 1-184	0.7207	0.7130	0.007667	0.9014	3	3	0.7827	32.00
DS, 1-214 vs. DS, 1-184*	0.7207	0.7703	-0.04967	0.0002	3	3	5.070	32.00
DS, 1-214 vs. FR, 1-164	0.7207	0.7747	-0.05400	<0.0001	3	3	5.513	32.00
DS, 1-214 vs. DS, 1-164	0.7207	0.7927	-0.07200	<0.0001	3	3	7.350	32.00
DS, 1-214 vs. 3S, 1-214	0.7207	0.6180	0.1027	<0.0001	3	3	10.48	32.00
DS, 1-214 vs. DbB, c048	0.7207	0.7777	-0.05700	<0.0001	3	3	5.819	32.00
FR, 1-184 vs. DS, 1-184*	0.7130	0.7703	-0.05733	<0.0001	3	3	5.853	32.00
FR, 1-184 vs. FR, 1-164	0.7130	0.7747	-0.06167	<0.0001	3	3	6.295	32.00
FR, 1-184 vs. DS, 1-164	0.7130	0.7927	-0.07967	<0.0001	3	3	8.133	32.00
FR, 1-184 vs. 3S, 1-214	0.7130	0.6180	0.09500	<0.0001	3	3	9.698	32.00
FR, 1-184 vs. DbB, c048	0.7130	0.7777	-0.06467	<0.0001	3	3	6.602	32.00
DS, 1-184* vs. FR, 1-164	0.7703	0.7747	-0.004333	0.9014	3	3	0.4424	32.00
DS, 1-184* vs. DS, 1-164	0.7703	0.7927	-0.02233	0.1886	3	3	2.280	32.00
DS, 1-184* vs. 3S, 1-214	0.7703	0.6180	0.1523	<0.0001	3	3	15.55	32.00
DS, 1-184* vs. DbB, c048	0.7703	0.7777	-0.007333	0.9014	3	3	0.7486	32.00
FR, 1-164 vs. DS, 1-164	0.7747	0.7927	-0.01800	0.3754	3	3	1.838	32.00
FR, 1-164 vs. 3S, 1-214	0.7747	0.6180	0.1567	<0.0001	3	3	15.99	32.00
FR, 1-164 vs. DbB, c048	0.7747	0.7777	-0.003000	0.9014	3	3	0.3063	32.00
DS, 1-164 vs. 3S, 1-214	0.7927	0.6180	0.1747	<0.0001	3	3	17.83	32.00
DS, 1-164 vs. DbB, c048	0.7927	0.7777	0.01500	0.5172	3	3	1.531	32.00
3S, 1-214 vs. DbB, c048	0.6180	0.7777	-0.1597	<0.0001	3	3	16.30	32.00
Extended								
FR, 1-214 vs. DS, 1-214	0.3267	0.2709	0.05583		3	3	5.700	32.00
FR, 1-214 vs. FR, 1-184	0.3267	0.2780	0.04873	<0.0001	3	3	4.975	32.00
FR, 1-214 vs. DS, 1-184*	0.3267	0.2082	0.1185	0.0002	3	3	12.09	32.00
FR, 1-214 vs. FR, 1-164	0.3267	0.2084	0.1183	<0.0001	3	3	12.08	32.00
FR, 1-214 vs. DS, 1-164	0.3267	0.1887	0.1380	<0.0001	3	3	14.09	32.00
FR, 1-214 vs. 3S, 1-214	0.3267	0.3723	-0.04563	<0.0001	3	3	4.659	32.00
FR, 1-214 vs. DbB, c048	0.3267	0.2095	0.1172	0.0004	3	3	11.96	32.00
DS, 1-214 vs. FR, 1-184	0.2709	0.2780	-0.007100	<0.0001	3	3	0.7248	32.00
DS, 1-214 vs. DS, 1-184*	0.2709	0.2082	0.06263	0.9234	3	3	6.394	32.00
DS, 1-214 vs. FR, 1-164	0.2709	0.2084	0.06247	<0.0001	3	3	6.377	32.00
DS, 1-214 vs. DS, 1-164	0.2709	0.1887	0.08220	<0.0001	3	3	8.391	32.00
DS, 1-214 vs. 3S, 1-214	0.2709	0.3723	-0.1015	<0.0001	3	3	10.36	32.00
DS, 1-214 vs. DbB, c048	0.2709	0.2095	0.06137	<0.0001	3	3	6.265	32.00
FR, 1-184 vs. DS, 1-184*	0.2780	0.2082	0.06973	<0.0001	3	3	7.119	32.00
FR, 1-184 vs. FR, 1-164	0.2780	0.2084	0.06957	<0.0001	3	3	7.102	32.00
FR, 1-184 vs. DS, 1-164	0.2780	0.1887	0.08930	<0.0001	3	3	9.116	32.00
FR, 1-184 vs. 3S, 1-214	0.2780	0.3723	-0.09437	<0.0001	3	3	9.633	32.00
FR, 1-184 vs. DbB, c048	0.2780	0.2095	0.06847	<0.0001	3	3	6.989	32.00
DS, 1-184* vs. FR, 1-164	0.2082	0.2084	-0.0001667	<0.0001	3	3	0.01701	32.00
DS, 1-184* vs. DS, 1-164	0.2082	0.1887	0.01957	0.9989	3	3	1.997	32.00
DS, 1-184* vs. 3S, 1-214	0.2082	0.3723	-0.1641	0.2761	3	3	16.75	32.00
DS, 1-184* vs. DbB, c048	0.2082	0.2095	-0.001267	<0.0001	3	3	0.1293	32.00
FR, 1-164 vs. DS, 1-164	0.2084	0.1887	0.01973	0.9989	3	3	2.014	32.00
FR, 1-164 vs. 3S, 1-214	0.2084	0.3723	-0.1639	0.2761	3	3	16.74	32.00
FR, 1-164 vs. DbB, c048	0.2084	0.2095	-0.001100	<0.0001	3	3	0.1123	32.00
DS, 1-164 vs. 3S, 1-214	0.1887	0.3723	-0.1837	0.9989	3	3	18.75	32.00
DS, 1-164 vs. DbB, c048	0.1887	0.2095	-0.02083	<0.0001	3	3	2.127	32.00
3S, 1-214 vs. DbB, c048	0.3723	0.2095	0.1628	0.2553	3	3	16.62	32.00

Suppl. data A5-3-H Figure 3.23,pp. 108

Data extracted from each replicate as described in the Methods section (real-time fluorescence graph, ΔF, against temperature) were independently fitted to a sigmoidal Boltzmann equation. Individual calculated Tm_{50} values were then averaged by nonlinear least squares fit to a horizontal line to propagate the SEM of the values. N= 3 replicates (independent fluorescence experiments) per buffer condition and construct. R^2 values have not been listed due to the large number of runs performed (> 1500); fits with R^2 below 0.8 were excluded from further analysis.

Suppl. data A5-3-I Figure 3.34, pp. 129

Tm_{50} values were fitted against CXCL12 concentration in a modified Langmuir binding isotherm [157], accounting for concentration of CXCR4. 2 replicates per CXCL12 concentration were performed, with and without CXCR4. Residual distribution from the fit was Gaussian and variance remained constant between concentration points (no heteroskedasticity). Adjusted R^2=0.96.

Chapter 4 Data analysis

Suppl. data A5-4-A Figure 4.1, pp. 135

In both cases, residual distribution was Gaussian with equal variances between groups (linear QQ plot, rectangular homoscedasticity plot). A nested one-way ANOVA was applied with either dose or time (n=number of animals, 5) as the main column variable and individual animals as subcolumns ($n_{subcolumn}$=number of cells, ≈ 20 per animal). α= 0.05, 1 family. *Note that this approach is technically incorrect, but the slope of the curves did not allow for reliable least squares curve fitting which would be the optimal approach.*

A. Optimal HMGB1 dose		
P value	<0.0001	
P value summary	****	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	5.7, 8, 36	
Random effects	SD	Variance
Variation within subcolumns	358	128338
Variation among subcolumn means	197	38969
Do the subcolumns differ (within each column)?		
Chi-square, df	128, 1	
P value	<0.0001	
P value summary	****	
Is there significant difference between subcolumns (P < 0.05)?	Yes	

Holm-Sidak's multiple comparisons test	**Mean 1**	**Mean 2**	**Mean Diff.**	**Adjusted P Value**	**N1**	**N2**	**t**	**DF**
0 vs. 0.075	942	972	-30	0.9997	40	120	0.21	36
0 vs. 0.225	942	954	-12	0.9997	40	213	0.086	36
0 vs. 0.45	942	1193	-251	0.5341	40	207	1.8	36
0 vs. 0.75	942	1473	-531	0.0160	40	40	3.6	36
0 vs. 1.5	942	1424	-482	0.0348	40	40	3.2	36
0 vs. 3	942	1424	-482	0.0348	40	40	3.2	36
0 vs. 4.5	942	1475	-533	0.0114	40	85	3.7	36
0 vs. 6	942	1396	-454	0.0348	40	114	3.2	36
0.075 vs. 0.225	972	954	18	0.9997	120	213	0.14	36
0.075 vs. 0.45	972	1193	-221	0.5758	120	207	1.7	36
0.075 vs. 0.75	972	1473	-501	0.0162	120	40	3.6	36
0.75 vs. 1.5	1473	1424	49	0.9997	40	40	0.33	36
0.75 vs. 3	1473	1424	49	0.9997	40	40	0.33	36
0.75 vs. 4.5	1473	1475	-2.3	0.9997	40	85	0.016	36
0.75 vs. 6	1473	1396	77	0.9981	40	114	0.54	36
0.225 vs. 0.45	954	1193	-239	0.5341	213	207	1.8	36
0.225 vs. 0.75	954	1473	-519	0.0114	213	40	3.7	36
0.45 vs. 0.75	1193	1473	-280	0.4449	207	40	2.0	36

B. Maximum administration delay post-injury		
P value	0.0001	
P value summary	***	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	5.310, 9, 37	
Random effects	SD	Variance
Variation within subcolumns	345.0	119054
Variation among subcolumn means	208.9	43635
Do the subcolumns differ (within each column)?		
Chi-square, df	640.7, 1	
P value	<0.0001	
P value summary	****	
Is there significant difference between subcolumns (P < 0.05)?	Yes	

Holm-Sidak's multiple comparisons test	**Mean 1**	**Mean 2**	**Mean Diff.**	**Adjusted P Value**	**N1**	**N2**	**t**	**DF**
Baseline vs. 0	942.2	1473	-530.9	0.0093	40	40	3.470	37
Baseline vs. 1	942.2	1609	-666.3	0.0004	40	650	4.639	37
Baseline vs. 2	942.2	1429	-487.2	0.0093	40	457	3.386	37
Baseline vs. 3	942.2	1545	-602.4	0.0014	40	307	4.162	37
Baseline vs. 5	942.2	1461	-518.8	0.0093	40	300	3.417	37
Baseline vs. 6	942.2	1081	-138.3	0.5488	40	223	0.9068	37
Baseline vs. 8	942.2	1168	-225.4	0.4688	40	258	1.484	37
Baseline vs. 16	942.2	1119	-176.4	0.5488	40	314	1.213	37
Baseline vs. 24	942.2	1114	-172.0	0.5488	40	488	1.196	37

Suppl. data A5-4-B Figure 4.2, pp. 137

Residual distribution was Gaussian with equal variance (linear QQ plot, rectangular homoscedasticity plot), although a small number of outliers were present in both curves. A nested one-way ANOVA was applied with treatment (HMGB1 construct; n=number of animals, 5) as the main column variable and individual animals as subcolumns ($n_{subcolumn}$=number of cells, ≈ 40 per animal). α= 0.05, 1 family.

P value	**0.0001**	
P value summary	***	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	7.2, 6, 27	
Random effects	SD	Variance
Variation within subcolumns	232	53872
Variation among subcolumn means	78	6012
Do the subcolumns differ (within each column)?		
Chi-square, df	92, 1	
P value	<0.0001	
P value summary	****	

Holm-Sidak's multiple comparisons test	**Mean 1**	**Mean 2**	**Mean Diff.**	**Adjusted P Value**	**N1**	**N2**	**t**	**DF**
FR 1.5 mg/kg vs. Control	1153	778	375	0.0006	40	297	4.8	27
FR 1.5 mg/kg, WT CXCL12 vs. Control	1007	778	229	0.0228	234	297	3.3	27
FR 1.5 mg/kg, V3I CXCL12 vs. Control	796	778	19	0.9134	312	297	0.26	27
FR 0.45 mg/kg vs. Control	949	778	171	0.1322	204	297	2.5	27
FR 0.45 mg/kg, WT CXCL12 vs. Control	807	778	29	0.9134	322	297	0.42	27
FR 0.45 mg/kg, V3I CXCL12 vs. Control	818	778	41	0.9134	466	297	0.59	27
FR 1.5 mg/kg, WT CXCL12 vs. FR 1.5 mg/kg	1007	1153	-146	0.3058	234	40	1.9	27
FR 1.5 mg/kg, V3I CXCL12 vs. FR 1.5 mg/kg	796	1153	-356	0.0017	312	40	4.4	27
FR 0.45 mg/kg, WT CXCL12 vs. FR 0.45 mg/kg	807	949	-142	0.2628	322	204	2.1	27
FR 0.45 mg/kg, V3I CXCL12 vs. FR 0.45 mg/kg	818	949	-130	0.3058	466	204	1.9	27
FR 1.5 mg/kg, V3I CXCL12 vs. FR 1.5 mg/kg, WT CXCL12	796	1007	-211	0.0594	312	234	2.9	27

Suppl. data A5-4-C Figure 4.5, pp. 144

Each concentration point per construct was performed with a separate sensor, referenced against sensors uncoated with HMGB1. N= 8 sensors per construct (0+ 7 concentration points). Fit was done to a saturation curve : $Interference = \frac{B_{max} * CXCL12}{K_d + CXCL12}$

Comparison of Fits (is Kd equal?)	Kd	BMax
Simpler model	Kd same for all data sets	Bmax same for all data sets
Probability it is correct	62.36%	<0.01%
Alternative model	Kd different for each data set	Bmax different for each data set
Probability it is correct	37.64%	>99.99%
Ratio of probabilities	1.657	
Preferred model	Kd same for all data sets	Bmax different for each data set
Difference in AICc	-1.009	18.98

	FR	3S
Degrees of Freedom	5	5
R squared	0.8900	0.9102

Suppl. data A5-4-D Figure 4.12, pp.161

Each sensor coated with HMGB1 constructs was measured in triplicate over a concentration range of CXCL12 (0+10 concentrations). Signal of an uncoated sensor was subtracted in all cases. Concentrations corresponding to each replicate were independently fitted to a saturation fit (see previous section); likewise, the apparent first order rate constants for each step were fitted to a linear equation where $k_{Obs} = k_{Off} + k_{On} * CXCL12$ to estimate kinetics rates. Measurements with error higher than the measurement itself were excluded in the first order plot (concentration points 1 to 4, which were concentrations several orders of magnitude under $K_{d)}$. R^2 values for these fits are listed below; fits in red were excluded:

	DF	c011			c022			c027			c028			c037			c038		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
Sat. fit	8	0.99	0.99	0.99	0.99	0.98	0.99	0.96	0.99	0.98	0.98	0.98	0.96	0.99	0.99	0.98	0.94	0.95	0.94
1st order fit	4	0.88	0.94	0.9	0.91	0.91	0.94	0.84	0.81	0.71	0.9	0.56	-0.13	0.92	0.92	0.9	0.98	0.94	0.94

Kinetics rates were also estimated by nonlinear least squares fit to a horizontal line of the directly measured k_{Off} and k_{On} parameters, respectively, from the instrument in real-time; this was done instead of a geometric average in order to factor the standard errors of each measurement. Broad comparisons (test for differences) were done via AICc comparison of all curve fits for the saturation fit (K_d , B_{Max}), the first order fit (k_{Off}, k_{On}) and the horizontal averages of the directly measured k_{Off}, and k_{On}:

Comparison of Fits	kOff	kOn	kOff	kOn	Kd	BMax
	1st order fit		Horizontal average		Saturation fit	
Simpler model	kOff same for all data sets	kOn same for all data sets	One curve for all data sets	One curve for all data sets	Kd same for all data sets	Bmax same for all data sets
Probability it is correct	<0.01%	30.36%	6.56%	>99.99%	<0.01%	<0.01%
Alternative model	kOff different for each data set	kOn different for each data set	Different curve for each data set	Different curve for each data set	Kd different for each data set	Bmax different for each data set
Probability it is correct	>99.99%	69.64%	93.44%	<0.01%	>99.99%	>99.99%
Ratio of probabilities		2.294	14.24			
Preferred model	kOff different for each data set	kOn different for each data set	Different curve for each data set	One curve for all data sets	Kd different for each data set	Bmax different for each data set
Difference in AICc	19.07	1.660	5.313	-884.1	19.16	142.5

As in all cases differences were found, data were analysed via 2-way ANOVA (1 family), where each calculation for the $k_{Off,}$ or k_{On} (first order fit and horizontal average) or K_d (saturation fit and k_{Off}/ k_{On} ratio from each pair) was treated as one grouping variable, with construct as the second grouping variable. This was done to account for the variations from each calculation method; comparisons were performed by grand row means (construct vs. construct).

Comparison of kOn values				
Source of Variation	**% of total variation**	**P value**	**P value summary**	**Significant?**
Interaction	6.393	0.0058	**	Yes
Construct	10.19	<0.0001	****	Yes
Calculation	14.96	<0.0001	****	Yes
ANOVA table	**SS (Type III)**	**DF**	**MS**	**F (DFn, DFd)**
Interaction	0.0009445	5	0.0001889	F (5, 182) = 3.402
Construct	0.001505	1	0.001505	F (1, 182) = 27.10
Calculation	0.002210	5	0.0004421	F (5, 182) = 7.961
Residual	0.01011	182	5.553e-005	

Holm-Sidak's multiple comparisons test	Predicted (LS) mean 1	Predicted (LS) mean 2	Predicted (LS) mean diff.	Adjusted P Value	N1	N2	t	DF
c011 vs. c022	0.008802	0.01194	-0.003135	0.6260	33	32	1.685	182.0
c011 vs. c027	0.008802	0.01149	-0.002686	0.7696	33	32	1.444	182.0
c011 vs. c028	0.008802	0.01048	-0.001677	0.9622	33	31	0.8918	182.0
c011 vs. c037	0.008802	0.01075	-0.001945	0.9372	33	33	1.056	182.0
c011 vs. c038	0.008802	0.01931	-0.01051	<0.0001	33	33	5.706	182.0
c022 vs. c027	0.01194	0.01149	0.0004488	0.9754	32	32	0.2390	182.0
c022 vs. c028	0.01194	0.01048	0.001458	0.9703	32	31	0.7680	182.0
c022 vs. c037	0.01194	0.01075	0.001189	0.9754	32	33	0.6395	182.0
c022 vs. c038	0.01194	0.01931	-0.007377	0.0012	32	33	3.966	182.0
c027 vs. c028	0.01149	0.01048	0.001009	0.9754	32	31	0.5315	182.0
c027 vs. c037	0.01149	0.01075	0.0007407	0.9754	32	33	0.3982	182.0
c027 vs. c038	0.01149	0.01931	-0.007826	0.0005	32	33	4.208	182.0
c028 vs. c037	0.01048	0.01075	-0.0002681	0.9754	31	33	0.1426	182.0
c028 vs. c038	0.01048	0.01931	-0.008835	<0.0001	31	33	4.698	182.0
c037 vs. c038	0.01075	0.01931	-0.008567	<0.0001	33	33	4.650	182.0

Comparison of kOff values				
Source of Variation	**% of total variation**	**P value**	**P value summary**	**Significant?**
Interaction	2.995	0.0118	*	Yes
Construct	5.572	<0.0001	****	Yes
Calculation	53.35	<0.0001	****	Yes
ANOVA table	**SS (Type III)**	**DF**	**MS**	**F (DFn, DFd)**
Interaction	0.5716	5	0.1143	F (5, 203) = 3.022
Construct	1.063	1	1.063	F (1, 203) = 28.11
Calculation	10.18	5	2.037	F (5, 203) = 53.83
Residual	7.680	203	0.03783	

Holm-Sidak's multiple comparisons test	Predicted (LS) mean 1	Predicted (LS) mean 2	Predicted (LS) mean diff.	Adjusted P Value	N1	N2	t	DF
c011 vs. c022	0.1796	0.1659	0.01371	0.7593	38	38	0.3069	203.0
c011 vs. c027	0.1796	0.4344	-0.2548	<0.0001	38	38	5.702	203.0
c011 vs. c028	0.1796	0.7775	-0.5979	<0.0001	38	33	12.88	203.0
c011 vs. c037	0.1796	0.1315	0.04815	0.6825	38	30	1.001	203.0
c011 vs. c038	0.1796	0.2535	-0.07385	0.3438	38	38	1.653	203.0
c022 vs. c027	0.1659	0.4344	-0.2685	<0.0001	38	38	6.009	203.0
c022 vs. c028	0.1659	0.7775	-0.6116	<0.0001	38	33	13.18	203.0
c022 vs. c037	0.1659	0.1315	0.03444	0.7240	38	30	0.7162	203.0
c022 vs. c038	0.1659	0.2535	-0.08756	0.2320	38	38	1.960	203.0
c027 vs. c028	0.4344	0.7775	-0.3431	<0.0001	38	33	7.392	203.0
c027 vs. c037	0.4344	0.1315	0.3029	<0.0001	38	30	6.300	203.0
c027 vs. c038	0.4344	0.2535	0.1809	0.0005	38	38	4.049	203.0
c028 vs. c037	0.7775	0.1315	0.6460	<0.0001	33	30	13.00	203.0
c028 vs. c038	0.7775	0.2535	0.5240	<0.0001	33	38	11.29	203.0
c037 vs. c038	0.1315	0.2535	-0.1220	0.0694	30	38	2.537	203.0

Comparison of Kd values				
Source of Variation	**% of total variation**	**P value**	**P value summary**	**Significant?**
Interaction	10.32	<0.0001	****	Yes
Construct	17.94	<0.0001	****	Yes
Calculation	6.844	<0.0001	****	Yes
ANOVA table	**SS (Type III)**	**DF**	**MS**	**F (DFn, DFd)**
Interaction	193755	10	19376	F (10, 357) = 5.451
Construct	336906	5	67381	F (5, 357) = 18.96
Calculation	128550	2	64275	F (2, 357) = 18.08
Residual	1268863	357	3554	

Test details	Predicted (LS) mean 1	Predicted (LS) mean 2	Predicted (LS) mean diff.	Adjusted P Value	N1	N2	t	DF
c011 vs. c022	19.56	12.40	7.162	0.9864	63	63	0.6461	357.0
c011 vs. c027	19.56	41.85	-22.30	0.2757	63	63	2.011	357.0
c011 vs. c028	19.56	101.1	-81.58	<0.0001	63	60	7.232	357.0
c011 vs. c037	19.56	13.59	5.971	0.9864	63	63	0.5386	357.0
c011 vs. c038	19.56	12.27	7.287	0.9864	63	63	0.6574	357.0
c022 vs. c027	12.40	41.85	-29.46	0.0768	63	63	2.658	357.0
c022 vs. c028	12.40	101.1	-88.74	<0.0001	63	60	7.867	357.0
c022 vs. c037	12.40	13.59	-1.191	0.9992	63	63	0.1075	357.0
c022 vs. c038	12.40	12.27	0.1256	0.9992	63	63	0.01133	357.0
c027 vs. c028	41.85	101.1	-59.29	<0.0001	63	60	5.256	357.0
c027 vs. c037	41.85	13.59	28.27	0.0861	63	63	2.550	357.0
c027 vs. c038	41.85	12.27	29.58	0.0768	63	63	2.669	357.0
c028 vs. c037	101.1	13.59	87.55	<0.0001	60	63	7.761	357.0
c028 vs. c038	101.1	12.27	88.87	<0.0001	60	63	7.878	357.0
c037 vs. c038	13.59	12.27	1.317	0.9992	63	63	0.1188	357.0

Chapter 5 Data analysis

Suppl. data A5-5-A Figure 5.7, pp. 185

3 replicates over 8 concentrations (0 + 7 concentration points) were used for fitting a Michaelis-Menten curve in the assay performed in 96-well plates; 4 replicates over the same concentration ranges were used in the 384-well plate assay. In both cases a parallel number of uncoated (blank) wells were exposed to the same RAGE concentration and subtracted as references. As each point was independent, all values were introduced in the same curve fit. R^2 for 96-well plate fit; 0.98, for 384-well plate fit: 0.97; DF (degrees of freedom)=30. To facilitate comparisons, signal was normalised assigning to DS-HMGB1 FL signal an arbitrary value of 1 at maximum intensity.

Suppl. data A5-5-B Figure 5.8, pp. 185

Data were processed as described in Suppl. data A5-5-B (384-well assay). Fits for HMGB1A c037 and c038 were of very low quality as the binding did not reach saturation (poor binders); this is reflected in the adjusted R^2 and several fitted parameters having no computable confidence intervals. B_{max} and K_d for HMGB1A-c001 1-214 and c006 1-214 were compared by Brown-Forsythe/Welch ANOVA. DF= 30.

	DS-HMGB1A-C1-c001 (1-214)	DS-HMGB1A-c037 (1-89)	"DS"-HMGB1A-c038 (89-174)	"DS" 3S-HMGB1A-c006 (1-214)	FR-HMGB1A-C1-c001 (1-214)	FR-HMGB1A-c037 (1-89)	"FR"-HMGB1A-c038 (89-174)	"FR" 3S-HMGB1A-c006 (1-214)
Adjusted R^2	0.96	0.43	-0.014	0.97	0.99	0.61	0.40	0.99

Kd comparison	Brown-Forsythe ANOVA	Welch's ANOVA
F*/W (DFn, DFd)	78.64 (3.000, 36.28)	30.30 (3.000, 64.22)
P value	<0.0001	<0.0001
P value summary	****	****
Significant diff. among means (P < 0.05)?	Yes	Yes

Dunnett's T3 multiple comparisons test	Mean 1	Mean 2	Mean Diff.	SE of diff.	n1	n2	t	DF
DS HMGB1A-C1-c001 (1-214) vs. "DS" 3S-HMGB1A-c006 (1-214)	4.383	7.216	-2.833	1.152	32	32	2.459	53.21
DS HMGB1A-C1-c001 (1-214) vs. FR HMGB1A-C1-c001 (1-214)	4.383	45.70	-41.31	4.363	32	32	9.468	32.31
DS HMGB1A-C1-c001 (1-214) vs. "FR" 3S-HMGB1A-c006 (1-214)	4.383	6.119	-1.736	0.7894	32	32	2.200	57.94
"DS" 3S-HMGB1A-c006 (1-214) vs. FR HMGB1A-C1-c001 (1-214)	7.216	45.70	-38.48	4.425	32	32	8.696	34.10
"DS" 3S-HMGB1A-c006 (1-214) vs. "FR" 3S-HMGB1A-c006 (1-214)	7.216	6.119	1.097	1.078	32	32	1.017	45.35
FR HMGB1A-C1-c001 (1-214) vs. "FR" 3S-HMGB1A-c006 (1-214)	45.70	6.119	39.58	4.344	32	32	9.110	31.76

BMax comparison	Brown-Forsythe ANOVA	Welch's ANOVA
F*/W (DFn, DFd)	2.853 (3.000, 109.7)	2.301 (3.000, 66.34)
P value	0.0406	0.0852
P value summary	*	ns
Significant diff. among means (P < 0.05)?	Yes	No

All post-hoc comparisons were non-significant for B_{max}

Suppl. data A5-5-C Figure 5.9, pp. 186

Data were processed as described in Suppl. data A5-5-B (384-well assay). All curves were identical except c003 (DS) vs c001 (DS), no post-hoc comparisons were run save for AICc between those two curves; these are shown in the figure itself.

	c001 (plate 1)	c013	c017	c018	c020	c021	c023	c024	c025	c003	c001 (plate 2)	c056	c057	c058	c059	c060	c061	c062
DF	30	14	14	14	14	14	14	14	14	30	30	30	30	30	30	30	30	30
Adj. R^2	0.99	1.0	1.0	1.0	1.0	1.0	0.99	0.98	0.99	0.99	0.98	0.97	0.97	0.95	0.97	0.98	0.98	0.97

Suppl. data A5-5-D Figure 5.10, pp. 187

Data were processed as described in Suppl. data A5-5-B (384-well assay). Both B_{Max} and K_d were different across constructs; the best-fit parameters were extracted and analysed via one-way Brown-Forsythe/Welch ANOVA (unequal SDs). “FR” HMGB1-C1-c048 was excluded from this analysis as the binding was insufficient to achieve reliable fits, indicative of low RAGE binding capacity (R^2 too low).

	DS HMGB1-C1-c001 (1-214)	DS HMGB1-C1-c002 (1-184)	DS HMGB1-C1-c003 (1-164)	"DS" HMGB1-C1-c048 (DbB)	FR HMGB1-C1-c001 (1-214)	FR HMGB1-C1-c002 (1-184)	FR HMGB1-C1-c003 (1-164)	"FR" HMGB1-C1-c048 (DbB)
Adj. R^2	0.9870	0.9804	0.8735	0.5955	0.9897	0.9848	0.6187	**0.3389**

AICc comparisons	**BMax**	**Kd**
Simpler model	Bmax same for all data sets	Kd same for all data sets
Probability it is correct	<0.01%	<0.01%
Alternative model	Bmax different for each data set	Kd different for each data set
Probability it is correct	>99.99%	>99.99%
Ratio of probabilities		
Preferred model	Bmax different for each data set	Kd different for each data set
Difference in AICc	386.0	480.7

BMax comparison	Brown-Forsythe ANOVA	Welch's ANOVA
F*/W (DFn, DFd)	799.8 (6.000, 120.7)	948.2 (6.000, 94.79)
P value	<0.0001	<0.0001
P value summary	****	****
Significant diff. among means (P < 0.05)?	Yes	Yes

Dunnett's T3 multiple comparisons test	Mean 1	Mean 2	Mean Diff.	Adjusted P Value	n1	n2	t	DF
DS HMGB1-C1-c001 (1-214) vs. DS HMGB1-C1-c002 (1-184)	3.133	1.836	1.297	<0.0001	32	32	19.22	62.00
DS HMGB1-C1-c001 (1-214) vs. DS HMGB1-C1-c003 (1-164)	3.133	0.6405	2.492	<0.0001	32	32	41.60	58.12
DS HMGB1-C1-c001 (1-214) vs. "DS" HMGB1-C1-c048 (DbB)	3.133	0.1939	2.939	<0.0001	32	32	56.36	43.13
DS HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c001 (1-214)	3.133	3.185	-0.0529	>0.9999	32	32	0.5758	50.98
DS HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c002 (1-184)	3.133	1.562	1.570	<0.0001	32	32	26.22	58.10
DS HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c003 (1-164)	3.133	0.1783	2.954	<0.0001	32	32	56.75	42.91
DS HMGB1-C1-c002 (1-184) vs. DS HMGB1-C1-c003 (1-164)	1.836	0.6405	1.195	<0.0001	32	32	19.85	57.92
DS HMGB1-C1-c002 (1-184) vs. "DS" HMGB1-C1-c048 (DbB)	1.836	0.1939	1.642	<0.0001	32	32	31.29	42.96
DS HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c001 (1-214)	1.836	3.185	-1.350	<0.0001	32	32	14.66	51.21
DS HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c002 (1-184)	1.836	1.562	0.2736	0.0006	32	32	4.546	57.89
DS HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c003 (1-164)	1.836	0.1783	1.657	<0.0001	32	32	31.64	42.74
DS HMGB1-C1-c003 (1-164) vs. "DS" HMGB1-C1-c048 (DbB)	0.6405	0.1939	0.4466	<0.0001	32	32	10.55	50.14
DS HMGB1-C1-c003 (1-164) vs. FR HMGB1-C1-c001 (1-214)	0.6405	3.185	-2.545	<0.0001	32	32	29.36	43.76
DS HMGB1-C1-c003 (1-164) vs. FR HMGB1-C1-c002 (1-184)	0.6405	1.562	-0.9216	<0.0001	32	32	17.87	62.00
DS HMGB1-C1-c003 (1-164) vs. FR HMGB1-C1-c003 (1-164)	0.6405	0.1783	0.4622	<0.0001	32	32	10.95	49.85
"DS" HMGB1-C1-c048 (DbB) vs. FR HMGB1-C1-c001 (1-214)	0.1939	3.185	-2.991	<0.0001	32	32	36.71	35.59
"DS" HMGB1-C1-c048 (DbB) vs. FR HMGB1-C1-c002 (1-184)	0.1939	1.562	-1.368	<0.0001	32	32	32.35	50.17
"DS" HMGB1-C1-c048 (DbB) vs. FR HMGB1-C1-c003 (1-164)	0.1939	0.1783	0.01561	>0.9999	32	32	0.5171	61.99
FR HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c002 (1-184)	3.185	1.562	1.623	<0.0001	32	32	18.73	43.73
FR HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c003 (1-164)	3.185	0.1783	3.007	<0.0001	32	32	36.92	35.50
FR HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c003 (1-164)	1.562	0.1783	1.384	<0.0001	32	32	32.80	49.88

Kd comparison	Brown-Forsythe ANOVA	Welch's ANOVA
F*/W (DFn, DFd)	44.19 (6.000, 137.0)	47.01 (6.000, 83.80)
P value	<0.0001	<0.0001
P value summary	****	****
Significant diff. among means (P < 0.05)?	Yes	Yes

Dunnett's T3 multiple comparisons test	Mean 1	Mean 2	Mean Diff.	Adjusted P Value	n1	n2	t	DF
DS HMGB1-C1-c001 (1-214) vs. DS HMGB1-C1-c002 (1-184)	4.351	33.87	-29.52	<0.0001	32	32	8.288	31.61
DS HMGB1-C1-c001 (1-214) vs. DS HMGB1-C1-c003 (1-164)	4.351	16.56	-12.20	0.1327	32	32	2.866	31.42
DS HMGB1-C1-c001 (1-214) vs. "DS" HMGB1-C1-c048 (DbB)	4.351	7.978	-3.626	>0.9999	32	32	0.8272	31.40
DS HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c001 (1-214)	4.351	86.11	-81.76	<0.0001	32	32	11.57	31.15
DS HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c002 (1-184)	4.351	34.37	-30.01	<0.0001	32	32	9.299	31.74
DS HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c003 (1-164)	4.351	7.711	-3.360	>0.9999	32	32	0.7334	31.36
DS HMGB1-C1-c002 (1-184) vs. DS HMGB1-C1-c003 (1-164)	33.87	16.56	17.32	0.0535	32	32	3.132	60.10
DS HMGB1-C1-c002 (1-184) vs. "DS" HMGB1-C1-c048 (DbB)	33.87	7.978	25.89	0.0005	32	32	4.602	59.47
DS HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c001 (1-214)	33.87	86.11	-52.24	<0.0001	32	32	6.617	45.71
DS HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c002 (1-184)	33.87	34.37	-0.4937	>0.9999	32	32	0.1033	61.39
DS HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c003 (1-164)	33.87	7.711	26.16	0.0006	32	32	4.525	58.40
DS HMGB1-C1-c003 (1-164) vs. "DS" HMGB1-C1-c048 (DbB)	16.56	7.978	8.578	0.9663	32	32	1.408	61.95
DS HMGB1-C1-c003 (1-164) vs. FR HMGB1-C1-c001 (1-214)	16.56	86.11	-69.56	<0.0001	32	32	8.449	50.83
DS HMGB1-C1-c003 (1-164) vs. FR HMGB1-C1-c002 (1-184)	16.56	34.37	-17.81	0.0291	32	32	3.348	57.72
DS HMGB1-C1-c003 (1-164) vs. FR HMGB1-C1-c003 (1-164)	16.56	7.711	8.844	0.9639	32	32	1.419	61.67
"DS" HMGB1-C1-c048 (DbB) vs. FR HMGB1-C1-c001 (1-214)	7.978	86.11	-78.14	<0.0001	32	32	9.415	51.73
"DS" HMGB1-C1-c048 (DbB) vs. FR HMGB1-C1-c002 (1-184)	7.978	34.37	-26.39	0.0002	32	32	4.867	56.90
"DS" HMGB1-C1-c048 (DbB) vs. FR HMGB1-C1-c003 (1-164)	7.978	7.711	0.2666	>0.9999	32	32	0.0421	61.88
FR HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c002 (1-184)	86.11	34.37	51.75	<0.0001	32	32	6.677	43.30
FR HMGB1-C1-c001 (1-214) vs. FR HMGB1-C1-c003 (1-164)	86.11	7.711	78.40	<0.0001	32	32	9.328	53.10
FR HMGB1-C1-c002 (1-184) vs. FR HMGB1-C1-c003 (1-164)	34.37	7.711	26.65	0.0003	32	32	4.775	55.60

Suppl. data A5-5-E Figure 5.12, pp. 190

Responses at equilibrium were fitted to saturation curves ; kinetics rates were estimated by nonlinear least squares fit to a horizontal line of the directly measured k_{Off} and k_{On} parameters, both as described in Suppl. data A5-4-D. RAGE-Fc coated sensors were dipped sequentially in a rising concentration of HMGB1, in triplicate; each replicate was independently fitted.

All curves were used in further calculations except those highlighted in red. Broad comparisons (test for differences) were done via AICc comparison of all curve fits for the saturation fit (K_d), and the horizontal averages of the directly measured k_{Off}, and k_{On}. An apparent first order fit was not performed due to poor quality of the measured k_{Obs} parameter in the instrument. Differences were observed between constructs for K_d, k_{Off}, and k_{On}.

		DS-C1-c001			FR-C1-c001			C1-c048			3S-c006			DS-C1-c002			DS-C1-c003			DS-C1-c003		
9 HMGB1	DF	2	6	6	5	6	6	6	6	6	2	5	5	5	6	6	6	6	6	4	6	6
concentrations	R^2	0.94	1.0	1.0	0.94	0.99	1.0	0.93	1.0	1.0	0.99	1.0	1.0	0.97	0.97	0.97	**-1.2**	0.98	0.99	0.99	0.99	0.99
7 HMGB1	DF	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	4	4	5	5	5
concentrations	R^2	0.99	1.0	1.0	0.93	0.98	0.99	0.98	0.99	0.99	0.94	1.0	1.0	0.97	0.95	0.95	1.0	1.0	1.0	0.99	0.98	0.98

AICc comparisons	**Kd**	**kOff**	**kOn**
Probability it is correct	<0.01%	One curve for all data sets	One curve for all data sets
Alternative model	Kd different for each data set	Different curve for each data set	Different curve for each data set
Probability it is correct	>99.99%	<0.0001	<0.0001
Ratio of probabilities		Reject null hypothesis	Reject null hypothesis
Preferred model	Kd different for each data set	Different curve for each data set	Different curve for each data set
Difference in AICc	96.08	17.49 (6, 252)	10.13 (6, 141)

Since two K_d parameters were estimated for each construct, they were compared via 2-way ANOVA with calculation and construct as grouping variables. All comparisons were done as a single family.

Kd comparisons	**% of total variation**	**P value**	**P value summary**	**Significant?**
Interaction	5.419	<0.0001	****	Yes
Row Factor	13.88	<0.0001	****	Yes
Column Factor	13.73	<0.0001	****	Yes
ANOVA table	**SS (Type III)**	**DF**	**MS**	**F (DFn, DFd)**
Interaction	219.2	5	43.84	F (5, 612) = 9.886
Row Factor	561.5	5	112.3	F (5, 612) = 25.32
Column Factor	555.4	1	555.4	F (1, 612) = 125.2
Residual	2714	612	4.435	

Holm-Sidak's multiple comparisons test	**Pred.(LS) mean 1**	**Pred. (LS) mean 2**	**Pred. (LS) mean diff.**	**Adjusted P Value**	**N1**	**N2**	**t**	**DF**
DS-HMGB1 vs. FR-HMGB1	0.7583	1.542	-0.7834	0.0392	123	94	2.662	612.0
DS-HMGB1 vs. dBB12L-HMGB1	0.7583	3.403	-2.645	<0.0001	123	88	8.835	612.0
DS-HMGB1 vs. 3S-HMGB1	0.7583	1.937	-1.179	0.0004	123	95	4.016	612.0
DS-HMGB1 vs. DS 1-184	0.7583	0.3396	0.4187	0.4544	123	113	1.476	612.0
DS-HMGB1 vs. DS 1-164	0.7583	1.957	-1.199	0.0002	123	111	4.214	612.0
FR-HMGB1 vs. dBB12L-HMGB1	1.542	3.403	-1.862	<0.0001	94	88	5.954	612.0
FR-HMGB1 vs. 3S-HMGB1	1.542	1.937	-0.3954	0.4544	94	95	1.288	612.0
FR-HMGB1 vs. DS 1-184	1.542	0.3396	1.202	0.0004	94	113	4.038	612.0
FR-HMGB1 vs. DS 1-164	1.542	1.957	-0.4158	0.4544	94	111	1.393	612.0
dBB12L-HMGB1 vs. 3S-HMGB1	3.403	1.937	1.466	<0.0001	88	95	4.700	612.0
dBB12L-HMGB1 vs. DS 1-184	3.403	0.3396	3.064	<0.0001	88	113	10.12	612.0
dBB12L-HMGB1 vs. DS 1-164	3.403	1.957	1.446	<0.0001	88	111	4.763	612.0
3S-HMGB1 vs. DS 1-184	1.937	0.3396	1.597	<0.0001	95	113	5.379	612.0
3S-HMGB1 vs. DS 1-164	1.937	1.957	-0.02036	0.9455	95	111	0.0683	612.0
DS 1-184 vs. DS 1-164	0.3396	1.957	-1.618	<0.0001	113	111	5.615	612.0

As only one k_{Off} and k_{On} method of calculation was applied, best-fit parameters were compared by Brown-Forsythe/Welch one-way ANOVA (unequal SDs). P values in red were considered significant with an α=0.1. All comparisons were done as a single family.

kOn comparison	Brown-Forsythe ANOVA test	Welch's ANOVA test
F*/W (DFn, DFd)	11.00 (5.000, 74.51)	19.67 (5.000, 46.94)
P value	<0.0001	<0.0001
P value summary	****	****
Significant diff. among means (P < 0.05)?	Yes	Yes

Dunnett's T3 multiple comparisons test	Mean 1	Mean 2	Mean Diff.	Adj. P value	n1	n2	t	DF
DS-HMGB1 vs. FR-HMGB1	0.2146	0.05148	0.1631	0.0048	25	19	4.091	28.14
DS-HMGB1 vs. dBB12L-HMGB1	0.2146	0.1635	0.05112	0.9944	25	14	1.005	35.88
DS-HMGB1 vs. 3S-HMGB1	0.2146	0.01749	0.1971	0.0004	25	20	5.153	24.08
DS-HMGB1 vs. DS 1-184	0.2146	0.3493	-0.1347	0.5861	25	27	1.912	43.94
DS-HMGB1 vs. DS 1-164	0.2146	0.1783	0.03631	>0.9999	25	27	0.6860	49.66
FR-HMGB1 vs. dBB12L-HMGB1	0.05148	0.1635	-0.1120	**0.0767**	19	14	3.160	16.00
FR-HMGB1 vs. 3S-HMGB1	0.05148	0.01749	0.03399	0.1012	19	20	2.963	18.64
FR-HMGB1 vs. DS 1-184	0.05148	0.3493	-0.2978	0.0005	19	27	4.944	27.90
FR-HMGB1 vs. DS 1-164	0.05148	0.1783	-0.1268	0.0338	19	27	3.307	30.84
dBB12L-HMGB1 vs. 3S-HMGB1	0.1635	0.01749	0.1460	0.0106	14	20	4.344	13.05
dBB12L-HMGB1 vs. DS 1-184	0.1635	0.3493	-0.1858	0.1266	14	27	2.731	37.64
dBB12L-HMGB1 vs. DS 1-164	0.1635	0.1783	-0.01481	>0.9999	14	27	0.2980	36.50
3S-HMGB1 vs. DS 1-184	0.01749	0.3493	-0.3318	0.0001	20	27	5.607	26.03
3S-HMGB1 vs. DS 1-164	0.01749	0.1783	-0.1608	0.0025	20	27	4.387	26.09
DS 1-184 vs. DS 1-164	0.3493	0.1783	0.1710	0.2266	27	27	2.457	43.37

kOn comparison	Brown-Forsythe ANOVA test	Welch's ANOVA test
F*/W (DFn, DFd)	11.52 (5.000, 46.66)	25.07 (5.000, 48.94)
P value	<0.0001	<0.0001
P value summary	****	****
Significant diff. among means (P < 0.05)?	Yes	Yes

Dunnett's T3 multiple comparisons test	Mean 1	Mean 2	Mean Diff.	Adj. P value	n1	n2	t	DF
DS-HMGB1 vs. FR-HMGB1	0.04251	0.04093	0.001579	>0.9999	25	19	0.1448	21.69
DS-HMGB1 vs. dBB12L-HMGB1	0.04251	0.2062	-0.1637	0.0040	25	14	4.890	13.26
DS-HMGB1 vs. 3S-HMGB1	0.04251	0.01499	0.02751	<0.0001	25	20	7.642	31.91
DS-HMGB1 vs. DS 1-184	0.04251	0.06731	-0.02480	0.2447	25	27	2.443	32.10
DS-HMGB1 vs. DS 1-164	0.04251	0.1518	-0.1093	**0.0653**	25	27	3.083	26.46
FR-HMGB1 vs. dBB12L-HMGB1	0.04093	0.2062	-0.1653	0.0032	19	14	4.737	15.55
FR-HMGB1 vs. 3S-HMGB1	0.04093	0.01499	0.02593	0.2572	19	20	2.473	18.64
FR-HMGB1 vs. DS 1-184	0.04093	0.06731	-0.02638	0.6230	19	27	1.865	41.09
FR-HMGB1 vs. DS 1-164	0.04093	0.1518	-0.1109	**0.0712**	19	27	3.013	30.37
dBB12L-HMGB1 vs. 3S-HMGB1	0.2062	0.01499	0.1912	0.0010	14	20	5.735	13.05
dBB12L-HMGB1 vs. DS 1-184	0.2062	0.06731	0.1389	0.0155	14	27	4.007	15.19
dBB12L-HMGB1 vs. DS 1-164	0.2062	0.1518	0.05438	0.9848	14	27	1.120	35.94
3S-HMGB1 vs. DS 1-184	0.01499	0.06731	-0.05231	0.0002	20	27	5.396	27.09
3S-HMGB1 vs. DS 1-164	0.01499	0.1518	-0.1368	0.0094	20	27	3.873	26.08
DS 1-184 vs. DS 1-164	0.06731	0.1518	-0.08451	0.3176	27	27	2.310	29.82

Suppl. data A5-5-F Figure 5.13, pp. 194

A, B: HEK-Dual NF-κβ Reporter assays

Residual distribution was not Gaussian with unequal variance (heavy sigmoidal QQ plot, triangular homoscedasticity plot, due to the difference of the positive controls with treatment). Data were analyzed via Kruskal-Wallis test, with n= number of wells (8 per condition). α= 0.05.

HEK-Dual hTLR-2 (B)	
P value	<0.0001
Exact or approximate P value?	Approximate
P value summary	****
Do the medians vary signif. (P < 0.05)?	Yes
Number of groups	5
Kruskal-Wallis statistic	39.73
Data summary	
Number of treatments (columns)	5
Number of values (total)	42

Dunn's multiple comparisons test	Mean rank 1	Mean rank 2	Mean rank diff.	Adjusted P Value	n1	n2	Z
Control vs. FR-HMGB1-C1-c001	5.000	23.00	-18.00	0.0176	9	9	3.128
Control vs. HMGB1-C1-c048	5.000	14.00	-9.000	>0.9999	9	9	1.564
Control vs. DS-HMGB1	5.000	31.50	-26.50	<0.0001	9	8	4.467
Control vs. LTA	5.000	39.00	-34.00	<0.0001	9	7	5.527
FR-HMGB1-C1-c001 vs. HMGB1-C1-c048	23.00	14.00	9.000	>0.9999	9	9	1.564
FR-HMGB1-C1-c001 vs. DS-HMGB1	23.00	31.50	-8.500	>0.9999	9	8	1.433
FR-HMGB1-C1-c001 vs. LTA	23.00	39.00	-16.00	0.0930	9	7	2.601
HMGB1-C1-c048 vs. DS-HMGB1	14.00	31.50	-17.50	0.0318	9	8	2.950
HMGB1-C1-c048 vs. LTA	14.00	39.00	-25.00	0.0005	9	7	4.064
DS-HMGB1 vs. LTA	31.50	39.00	-7.500	>0.9999	8	7	1.187

HEK-Dual mTLR-4 (C)	
P value	<0.0001
Exact or approximate P value?	Approximate
P value summary	****
Do the medians vary signif. (P < 0.05)?	Yes
Number of groups	5
Kruskal-Wallis statistic	42.02
Data summary	
Number of treatments (columns)	5
Number of values (total)	45

Dunn's multiple comparisons test	Mean rank 1	Mean rank 2	Mean rank diff.	Adjusted P Value	n1	n2	Z
Control vs. FR-HMGB1-C1-c001	5.000	22.33	-17.33	0.0494	9	9	2.811
Control vs. HMGB1-C1-c048	5.000	14.67	-9.667	>0.9999	9	9	1.568
Control vs. DS-HMGB1	5.000	32.00	-27.00	0.0001	9	9	4.378
Control vs. LPS	5.000	41.00	-36.00	<0.0001	9	9	5.838
FR-HMGB1-C1-c001 vs. HMGB1-C1-c048	22.33	14.67	7.667	>0.9999	9	9	1.243
FR-HMGB1-C1-c001 vs. DS-HMGB1	22.33	32.00	-9.667	>0.9999	9	9	1.568
FR-HMGB1-C1-c001 vs. LPS	22.33	41.00	-18.67	0.0247	9	9	3.027
HMGB1-C1-c048 vs. DS-HMGB1	14.67	32.00	-17.33	0.0494	9	9	2.811
HMGB1-C1-c048 vs. LPS	14.67	41.00	-26.33	0.0005	9	9	4.270
DS-HMGB1 vs. LPS	32.00	41.00	-9.000	>0.9999	9	9	1.459

C. Monocyte TNF secretion assays

Due to violation of heteroskedasticity, data pertaining to TLR-2 (LTA conditions) and TLR-4 (LPS conditions) were analysed separately with the control (HMGB1 constructs alone, negative control, negative control with proteinase K and recombinant HMGB1. Under these conditions residual distribution was Gaussian, and all groups had equal variance (linear QQ plot, rectangular homoscedasticity plot), although in both cases these assumptions were borderline and likely more donors will be required. A nested one-way ANOVA was applied with treatment (n=number of donors, 3) as the main column variable and individual animals as subcolumns ($n_{subcolumn}$=number of wells, ≈ 3 per condition). All post hoc comparisons within each test were done as part of a single family. No differences were detected between individual donors.

TLR-2 conditions	Nested table	
Data sets analysed	A, C-M	
Nested one-way ANOVA		
P value	<0.0001	
P value summary	****	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	261.3, 9, 20	
Random effects	SD	Variance
Variation within subcolumns	1.445	2.088
Variation among subcolumn means	0.6430	0.4135
Do the subcolumns differ (within each column)?		
Chi-square, df	1.752, 1	
P value	0.1856	
P value summary	ns	
Is there significant difference between subcolumns (P < 0.05)?	No	

Test details	Mean 1	Mean 2	Mean Diff.	Adjusted P Value	N1	N2	t	DF
Ctrl + Proteinase K vs. Ctrl	0.8490	1.000	-0.1510	>0.9999	9	9	0.1756	20
LTA10 vs. Ctrl	1.511	1.000	0.5106	>0.9999	9	9	0.5937	20
LTA50 vs. Ctrl	3.472	1.000	2.472	0.2098	9	9	2.874	20
FR vs. Ctrl	0.6903	1.000	-0.3097	>0.9999	9	9	0.3601	20
DS vs. Ctrl	21.12	1.000	20.12	<0.0001	9	9	23.39	20
dBB vs. Ctrl	0.6184	1.000	-0.3816	>0.9999	9	9	0.4437	20
FR+LTA10 vs. Ctrl	1.261	1.000	0.2608	>0.9999	9	9	0.3032	20
DS+LTA10 vs. Ctrl	27.44	1.000	26.44	<0.0001	9	9	30.75	20
dBB+LTA10 vs. Ctrl	1.145	1.000	0.1455	>0.9999	9	9	0.1691	20
LTA10 vs. Ctrl + Proteinase K	1.511	0.8490	0.6616	>0.9999	9	9	0.7693	20
LTA50 vs. Ctrl + Proteinase K	3.472	0.8490	2.623	0.1520	9	9	3.050	20
FR vs. Ctrl + Proteinase K	0.6903	0.8490	-0.1587	>0.9999	9	9	0.1846	20
DS vs. Ctrl + Proteinase K	21.12	0.8490	20.27	<0.0001	9	9	23.57	20
dBB vs. Ctrl + Proteinase K	0.6184	0.8490	-0.2306	>0.9999	9	9	0.2681	20
FR+LTA10 vs. Ctrl + Proteinase K	1.261	0.8490	0.4118	>0.9999	9	9	0.4788	20
DS+LTA10 vs. Ctrl + Proteinase K	27.44	0.8490	26.60	<0.0001	9	9	30.92	20
dBB+LTA10 vs. Ctrl + Proteinase K	1.145	0.8490	0.2965	>0.9999	9	9	0.3447	20
LTA50 vs. LTA10	3.472	1.511	1.961	0.5293	9	9	2.281	20
FR vs. LTA10	0.6903	1.511	-0.8203	0.9998	9	9	0.9538	20
DS vs. LTA10	21.12	1.511	19.61	<0.0001	9	9	22.80	20
dBB vs. LTA10	0.6184	1.511	-0.8922	0.9996	9	9	1.037	20
FR+LTA10 vs. LTA10	1.261	1.511	-0.2498	>0.9999	9	9	0.2905	20
DS+LTA10 vs. LTA10	27.44	1.511	25.93	<0.0001	9	9	30.15	20
dBB+LTA10 vs. LTA10	1.145	1.511	-0.3651	>0.9999	9	9	0.4246	20
FR vs. LTA50	0.6903	3.472	-2.782	0.1064	9	9	3.234	20
DS vs. LTA50	21.12	3.472	17.65	<0.0001	9	9	20.52	20
dBB vs. LTA50	0.6184	3.472	-2.854	0.0918	9	9	3.318	20
FR+LTA10 vs. LTA50	1.261	3.472	-2.211	0.3450	9	9	2.571	20
DS+LTA10 vs. LTA50	27.44	3.472	23.97	<0.0001	9	9	27.87	20
dBB+LTA10 vs. LTA50	1.145	3.472	-2.327	0.2805	9	9	2.705	20
DS vs. FR	21.12	0.6903	20.43	<0.0001	9	9	23.75	20
dBB vs. FR	0.6184	0.6903	-0.07183	>0.9999	9	9	0.08353	20
FR+LTA10 vs. FR	1.261	0.6903	0.5705	>0.9999	9	9	0.6633	20
DS+LTA10 vs. FR	27.44	0.6903	26.75	<0.0001	9	9	31.11	20
dBB+LTA10 vs. FR	1.145	0.6903	0.4552	>0.9999	9	9	0.5293	20
dBB vs. DS	0.6184	21.12	-20.50	<0.0001	9	9	23.84	20
FR+LTA10 vs. DS	1.261	21.12	-19.86	<0.0001	9	9	23.09	20
DS+LTA10 vs. DS	27.44	21.12	6.325	<0.0001	9	9	7.355	20
dBB+LTA10 vs. DS	1.145	21.12	-19.97	<0.0001	9	9	23.22	20
FR+LTA10 vs. dBB	1.261	0.6184	0.6423	>0.9999	9	9	0.7469	20
DS+LTA10 vs. dBB	27.44	0.6184	26.83	<0.0001	9	9	31.19	20
dBB+LTA10 vs. dBB	1.145	0.6184	0.5270	>0.9999	9	9	0.6128	20
DS+LTA10 vs. FR+LTA10	27.44	1.261	26.18	<0.0001	9	9	30.45	20
dBB+LTA10 vs. FR+LTA10	1.145	1.261	-0.1153	>0.9999	9	9	0.1341	20
dBB+LTA10 vs. DS+LTA10	1.145	27.44	-26.30	<0.0001	9	9	30.58	20

TLR-4 conditions	Nested table	
Data sets analysed	A, C-P	
Nested one-way ANOVA		
P value	<0.0001	
P value summary	****	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	197.5, 8, 18	
Random effects	SD	Variance
Variation within subcolumns	2.459	6.049
Variation among subcolumn means	1.114	1.242
Do the subcolumns differ (within each column)?		
Chi-square, df	1.674, 1	
P value	0.1957	
P value summary	ns	
Is there significant difference between subcolumns (P < 0.05)?	No	

Test details	Mean 1	Mean 2	Mean Diff.	Adjusted P Value	N1	N2	t	DF
Ctrl + Proteinase K vs. Ctrl	0.8490	1.000	-0.1510	>0.9999	9	9	0.1024	18
LPS50 vs. Ctrl	37.43	1.000	36.43	<0.0001	9	9	24.72	18
FR vs. Ctrl	0.6903	1.000	-0.3097	>0.9999	9	9	0.2101	18
DS vs. Ctrl	21.12	1.000	20.12	<0.0001	9	9	13.65	18
dBB vs. Ctrl	0.6184	1.000	-0.3816	>0.9999	9	9	0.2589	18
FR+LPS50 vs. Ctrl	1.146	1.000	0.1456	>0.9999	9	9	0.09880	18
DS+LPS50 vs. Ctrl	28.70	1.000	27.70	<0.0001	9	9	18.79	18
dBB+LPS50 vs. Ctrl	1.443	1.000	0.4431	>0.9999	9	9	0.3006	18
LPS50 vs. Ctrl + Proteinase K	37.43	0.8490	36.58	<0.0001	9	9	24.82	18
FR vs. Ctrl + Proteinase K	0.6903	0.8490	-0.1587	>0.9999	9	9	0.1077	18
DS vs. Ctrl + Proteinase K	21.12	0.8490	20.27	<0.0001	9	9	13.75	18
dBB vs. Ctrl + Proteinase K	0.6184	0.8490	-0.2306	>0.9999	9	9	0.1564	18
FR+LPS50 vs. Ctrl + Proteinase K	1.146	0.8490	0.2966	>0.9999	9	9	0.2012	18
DS+LPS50 vs. Ctrl + Proteinase K	28.70	0.8490	27.85	<0.0001	9	9	18.89	18
dBB+LPS50 vs. Ctrl + Proteinase K	1.443	0.8490	0.5941	>0.9999	9	9	0.4031	18
FR vs. LPS50	0.6903	37.43	-36.74	<0.0001	9	9	24.93	18
DS vs. LPS50	21.12	37.43	-16.31	<0.0001	9	9	11.07	18
dBB vs. LPS50	0.6184	37.43	-36.81	<0.0001	9	9	24.98	18
FR+LPS50 vs. LPS50	1.146	37.43	-36.28	<0.0001	9	9	24.62	18
DS+LPS50 vs. LPS50	28.70	37.43	-8.733	0.0002	9	9	5.925	18
dBB+LPS50 vs. LPS50	1.443	37.43	-35.99	<0.0001	9	9	24.42	18
DS vs. FR	21.12	0.6903	20.43	<0.0001	9	9	13.86	18
dBB vs. FR	0.6184	0.6903	-0.07183	>0.9999	9	9	0.04874	18
FR+LPS50 vs. FR	1.146	0.6903	0.4553	>0.9999	9	9	0.3089	18
DS+LPS50 vs. FR	28.70	0.6903	28.01	<0.0001	9	9	19.00	18
dBB+LPS50 vs. FR	1.443	0.6903	0.7528	>0.9999	9	9	0.5108	18
dBB vs. DS	0.6184	21.12	-20.50	<0.0001	9	9	13.91	18
FR+LPS50 vs. DS	1.146	21.12	-19.97	<0.0001	9	9	13.55	18
DS+LPS50 vs. DS	28.70	21.12	7.577	0.0011	9	9	5.141	18
dBB+LPS50 vs. DS	1.443	21.12	-19.68	<0.0001	9	9	13.35	18
FR+LPS50 vs. dBB	1.146	0.6184	0.5272	>0.9999	9	9	0.3577	18
DS+LPS50 vs. dBB	28.70	0.6184	28.08	<0.0001	9	9	19.05	18
dBB+LPS50 vs. dBB	1.443	0.6184	0.8246	>0.9999	9	9	0.5595	18
DS+LPS50 vs. FR+LPS50	28.70	1.146	27.55	<0.0001	9	9	18.69	18
dBB+LPS50 vs. FR+LPS50	1.443	1.146	0.2974	>0.9999	9	9	0.2018	18
dBB+LPS50 vs. DS+LPS50	1.443	28.70	-27.25	<0.0001	9	9	18.49	18

Suppl. data A5-5-G Figure 5.14, pp. 196

Residual distribution was not Gaussian and variances between groups were equal according to both statistical tests and residual/Q-Q plots. Data were analysed by 2-way ANOVA with each combination of construct and time point as an individual. No differences were detected across constructs; post-hoc comparisons have been omitted due to this.

Source of Variation	% of total variation	P value	P value summary	
Interaction	0.9082	0.9965	ns	
Time	71.96	<0.0001	****	
Construct	0.9382	0.5104	ns	
ANOVA table	SS	DF	MS	F (DFn, DFd)
Interaction	223.3	15	14.88	F (15, 120) = 0.2774
Time	17689	3	5896	F (3, 120) = 109.9
Construct	230.6	5	46.13	F (5, 120) = 0.8597
Residual	6438	120	53.65	
Spearman's test for heteroscedasticity				
Rs of predicted Y vs. |residual|	-0.1318			
P value (one tailed)	0.0576			
Passed (P > 0.05)?	Yes			
Normality of Residuals	Statistics	P value	Passed normality test?	
Anderson-Darling (A2*)	5.075	<0.0001	No	
D'Agostino-Pearson omnibus (K2)	335.7	<0.0001	No	
Shapiro-Wilk (W)	0.9167	<0.0001	No	
Kolmogorov-Smirnov (distance)	0.1713	<0.0001	No	

Suppl. data A5-5-H Figure 5.15, pp. 197

Residual distribution was Gaussian, and all groups had equal variance; a one-way ANOVA was applied with all comparisons in one family.

ANOVA summary					
F				56.43	
P value				<0.0001	
P value summary				****	
Significant diff. among means (P < 0.05)?				Yes	
R squared				0.9261	
ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Treatment (between columns)	9.600	2	4.800	F (2, 9) = 56.43	P<0.0001
Residual (within columns)	0.7655	9	0.08506		
Total	10.37	11			

Holm-Sidak's multiple comparisons test	**Mean 1**	**Mean 2**	**Mean Diff.**	**Adjusted P Value**	**n1**	**n2**	**t**	**DF**
Control vs. FR 1-214, C1-c001	1.318	3.271	-1.953	<0.0001	4	4	9.469	9
Control vs. DBB, C1-c048	1.318	3.155	-1.837	<0.0001	4	4	8.906	9
FR 1-214, C1-c001 vs. DBB, C1-c048	3.271	3.155	0.1163	0.5867	4	4	0.5637	9

Suppl. data A5-5-I Figure 5.16, pp. 199

Each time point was treated as an individual dataset. Within each timepoint (day 7, 14 and 28), Residual distributions were Gaussian with equal variance (linear QQ plot with skew at the higher Q/Q end in the Day 7 dataset, rectangular homoscedasticity plot). A nested one-way ANOVA was applied to each timepoint with treatment (HMGB1 construct; n=number of animals, 4/5) as the main column variable and individual animals as subcolumns ($n_{subcolumn}$=number of cells, ≈ 10-40 per animal). Only post-hoc comparisons for day 14 are shown (non-significant differences in day 7 and 28)

Day 7		
P value	0.1166	
P value summary	ns	
Significantly different (P < 0.05)?	No	
F, DFn, Dfd	2.755, 2, 9	
Random effects	SD	Variance
Variation within subcolumns	179.3	32150
Variation among subcolumn means	109.7	12030
Do the subcolumns differ (within each column)?		
Chi-square, df	38.49, 1	
P value	<0.0001	
P value summary	****	
Goodness of fit		
Degrees of freedom		383
REML criterion		2563

Day 14		
P value	0.0172	
P value summary	*	
Significantly different (P < 0.05)?	Yes	
F, DFn, Dfd	6.271, 2, 10	
Random effects	SD	Variance
Variation within subcolumns	589.0	346934
Variation among subcolumn means	337.7	114041
Do the subcolumns differ (within each column)?		
Chi-square, df	49.27, 1	
P value	<0.0001	
P value summary	****	
Is there significant difference between subcolumns (P < 0.05)?	Yes	
Goodness of fit		
Degrees of freedom	310	
REML criterion	2451	

Holm-Sidak's multiple comparisons test	**Mean 1**	**Mean 2**	**Mean Diff.**	**Adjusted P Value**	**N1**	**N2**	**t**	**DF**
FR-HMGB1, c001 vs. PBS	1995	1380	614.9	0.0308	94	107	2.512	10
DBB-HMGB1, c048 vs. PBS	2204	1380	824.4	0.0142	114	107	3.371	10

Day 28		
P value	0.4587	
P value summary	ns	
Significantly different (P < 0.05)?	No	
F, DFn, Dfd	0.8508, 2, 9	
Random effects	SD	Variance
Variation within subcolumns	753.7	568076
Variation among subcolumn means	756.6	572398
Do the subcolumns differ (within each column)?		
Chi-square, df	102.4, 1	
P value	<0.0001	
P value summary	****	
Goodness of fit		
Degrees of freedom	521	
REML criterion	4232	

Suppl. data A5-5-J Figure 5.18, pp. 203.

See Suppl. data A5-3-E.

Suppl. data A5-5-K Figure 5.19: Denaturing ESI-QTOF/MS SASA plots of HMGB1 constructs and DbB-HMGB1 in 0.1% formic acid, pH 2.2, pp. 204

See Suppl. data A5-3-G.

Suppl. data A5-5-L Figure 5.20 , pp. 206

Data analysis methodology identical to that described in Suppl. data A5-3-H.