REGULATING PROTEOSTASIS

FELIPE OSSA



A study of interactions between HSP70 proteins, HSP40 proteins, and the σ_1 receptor

Doctor in Structural Biology
University of Oxford

Supervisor: Professor Jason R. Schnell
Co-Supervisors: Professor Christina Redfield and Dr Sergi Padilla-Parra

October 2018

Felipe Ossa: *Regulating Proteostasis*, A study of interactions between HSP70 proteins, HSP40 proteins, and the σ_1 receptor, © October 2018

To my parents Gina and Eduardo Ossa without thier support I would not have gotten this far.

To Anna, whose companionship has helped me bear the trials and tribulations of the last four years.

ABSTRACT

Proteostasis encompasses a number of cellular processes that are essential for maintaining cellular protein machinery, ranging from synthesis and folding through to degradation, prevention of aggregation and even the mediation of constitutive function. This fundamental cellular process is linked to several clinically important phenomena, in particular neurodegeneration, cancer and ageing. Understanding the proteostatic machinery requires not just an individual examination of the key players in the system, namely the protein chaperones, but also an examination of the vast network of regulatory machinery that underpins their function. The work presented here examines two types of regulators: the HSP40 proteins and the σ_1 receptor, in the context of their impact on the activity of HSP70 proteins, one of the key hubs of the proteostatic network. The σ_1 receptor is a small (25 kDa) protein of significant clinical importance, and recent work has highlighted its links to the proteostatic machinery and its interaction with HSP70 proteins. With its unique pharmacology, σ_1 is now emerging as a novel ligand-operated regulator of proteostasis. Our work on the σ_1 receptor was hindered by the difficulties of expressing and isolating a functional integral human membrane protein; despite extensive optimisation of the construct and procedures for expression in E. coli and reconstitution into detergent we remained unable to demonstrate functionality. The data presented here on balance argues that even under our optimised conditions the receptor structure is significantly disrupted, and that activity is largely absent. More successful studies in cellulo show the response of the oligomeric state of σ_1 to its ligands using a novel number and brightness-based methodology. We next turned to HSP40 regulation of HSP70s. HSP40s are canonical regulators of HSP70s and act by stimulating ATPase activity using J-domains. We hypothesised that because of their different roles in the cell, the human cytosolic HSP70 and HSC70 may preferentially interact with or be stimulated by J-domains from different HSP40s. Using the J-domains of DNAJB2, DNAJB6, and DNAJC5, we show that the rank order of ATPase stimulation is preserved for HSP70 and HSC70, but that the stimulated activity of HSP70 is less than that of HSC70. Furthermore, we present data indicating that the HSP70 C-terminal domain plays a role in determining the response of HSC70 and HSP70 to regulation by HSP40 co-chaperone J-domains. Finally, we confirm the importance of the G/F rich region for the structure and function of a B class J-domain from humans, DNAJB6.

Many thanks go first and foremost to my supervisor Jason Schnell whose guidance, support and encouragement were essential to completing the work presented here.

Special thanks go to Jose Ortega-Roldan whose direct supervision at the lab bench helped shape my experimental skills and whose optimism and encouragement helped me endure working on σ_1.

Special thanks also go to Jolyon Claridge for his help with all things NMR and the many drinks shared in good company throughout my PhD.

I am grateful to Sergi Padilla-Parra for his help with the fluorescence microscopy of σ_1.

Thanks go to the students I supervised Katharina Warinner, Jack Granby, Jose-Carlos Gonzalez, Xenia Massana-Munoz and Charlie Whitehouse both for their contributions to the work presented here and to my development as a scientist.

To Rory Nolan for helpful discussions about N&B

My thanks as well to Luis Alvarez for the hints and tricks he showed me on the microscopes.

I would also like to extend my thanks to Anu Chandran and Faiz Kipli for making the lab a pleasure to work in.

Thanks go to Jacob Brady for warning me about the road ahead.

Thanks as well to Nader Amin who shared the many frustrations of working on σ_1 with Jose and I.

I would also like to thank Shahid Mehmood for his help with native mass spectrometry experiments on σ_1.

I gratefully acknowledge the provision of human σ_1 in a mammalian expression plasmid for microscopy experiments by Radu Aricescu.

I am indebted to Eduardo Ossa for proof reading this manuscript

CONTENTS

I OVERVIEW OF PROTEOSTASIS 1

1 INTRODUCTION 3
- 1.1 The σ_1 Receptor . . . 4
  - 1.1.1 Discovery of the σ_1 Receptor . . . 5
  - 1.1.2 Pharmacological Properties . . . 6
  - 1.1.3 The σ_1 receptor in disease . . . 6
  - 1.1.4 The erg2 family and the cellular role of σ_1 . . . 7
  - 1.1.5 Structure of σ_1 and Insights into Ligand Binding . . . 12
  - 1.1.6 Topology and Localisation . . . 16
  - 1.1.7 Link to Proteostasis . . . 18
  - 1.1.8 Studying σ_1 . . . 19
- 1.2 Regulation of HSP70 Proteins . . . 20
  - 1.2.1 Anatomy of a HSP70 Protein . . . 21
  - 1.2.2 The HSP70 ATPase Cycle . . . 22
  - 1.2.3 Regulators of the HSP70 ATPase Cycle . . . 23
  - 1.2.4 The HSP40 Interaction . . . 25
  - 1.2.5 Target Selection . . . 31
  - 1.2.6 Studying HSP40-HSP70 Interactions . . . 33

2 MATERIALS & METHODS 35
- 2.1 Protein Expression and Purification . . . 35
  - 2.1.1 Expression and Purification of σ_1 . . . 35
  - 2.1.2 Expression and Purification of HSP70 Constructs . . . 36
  - 2.1.3 Expression and Purification of HSP40 Constructs . . . 37
- 2.2 Biophysical Techniques . . . 38
  - 2.2.1 Chaperone Assays . . . 38
  - 2.2.2 Photo-Crosslinking . . . 38
  - 2.2.3 SEC-MALS . . . 38
  - 2.2.4 Microscale Thermophoresis . . . 39
  - 2.2.5 Circular Dichroism . . . 39
  - 2.2.6 Radio-Ligand Binding . . . 40
  - 2.2.7 ATPase Assays . . . 40
- 2.3 NMR . . . 41
  - 2.3.1 TRACT . . . 41
  - 2.3.2 σ_1 . . . 41
  - 2.3.3 σ_1-HSC70 Titrations . . . 42
  - 2.3.4 TALOS . . . 42
  - 2.3.5 HSC70 . . . 42
  - 2.3.6 DNAJC5 . . . 44
  - 2.3.7 DNAJB6 . . . 44
- 2.4 Cell Culture . . . 45
- 2.5 Constructs for Microscopy . . . 45

2.6 Microscopy Sample Preparation . . . 45
2.6.1 Imaging . . . 46
2.6.2 FRET-FLIM . . . 46
2.6.3 Number and Brightness . . . 47

II THE σ_1 RECEPTOR 49

3 BIOPHYSICAL STUDIES OF σ_1 51
3.1 Expressing and Purifying σ_1 . . . 51
3.1.1 Prior approaches to Expressing σ_1 . . . 51
3.1.2 Developing a Native Purification Strategy . . . 53
3.1.3 Testing the Functionality of Full Length σ_1 . . . 54
3.2 DPC . . . 56
3.2.1 Oligomeric State . . . 56
3.2.2 NMR Characterisation of Secondary Structure . . . 58
3.2.3 Functional Assays . . . 62
3.2.4 On the Functional State of σ_1 in DPC . . . 64
3.3 DDM . . . 64
3.3.1 Ligand Binding . . . 65
3.3.2 Monitoring Oligomerisation . . . 66
3.3.3 Chaperone Activity . . . 69
3.3.4 σ_1-Client Protein Interactions . . . 71
3.3.5 Functional state of σ_1 in DDM . . . 77
3.4 LMG Extraction and Periplasmic Targeting . . . 82
4 FLUORESCENCE MICROSCOPY OF σ_1 85
4.1 Overview . . . 85
4.2 Expression and Localisation of σ_1 . . . 86
4.3 Interactions by FRET-FLIM . . . 87
4.4 Oligomeric State of σ_1 by Number and Brightness . . . 90
4.4.1 Theoretical and Practical Considerations for N&B . . . 91
4.4.2 Applying Number and Brightness to σ_1 . . . 93
4.4.3 Interpreting the Results of Number and Brightness . . . 97
5 LESSONS FROM σ_1 99
5.1 Functional State of Isolated σ_1 . . . 99
5.2 Evaluating the *in cellulo* System . . . 101
5.3 Lessons Learned . . . 102

III REGULATION OF HSP70 PROTEINS BY HSP40 CO-CHAPERONES 103
6 EXPRESSION AND CHARACTERISATION OF HSP40 AND HSP70 TARGETS 105
6.1 HSC70 . . . 105
6.1.1 Oligomerisation of HSC70 . . . 106
6.1.2 NMR . . . 109
6.2 HSP40 Proteins . . . 114
6.2.1 Expression and Purification of DNAJB6 and DNAJB2 . . . 114
6.2.2 J-Domain Constructs . . . 120
6.2.3 NMR and Assignments of DNAJC5 and DNAJB6 . . . 121
7 ANALYSIS OF HSP40-HSP70 INTERACTIONS AND THEIR SPECIFICITY 127

7.1 Interaction by SEC . . . 127
7.2 Binding by NMR: Observing the J-Domain . . . 131
7.3 Binding by NMR: Observing HSC70 . . . 131
7.4 ATPase Assays . . . 137
7.4.1 Differences Between J-Domain Constructs: The G/F Rich Region . . . 137
7.4.2 Differences between HSC70 and HSP70 . . . 139
7.4.3 Role of the C-Terminal Domain . . . 140
8 TOWARDS A COMPLETE MAP OF J-DOMAIN-HSC70 ALLOSTERY 143
8.1 The Importance of the HSP70 C-Terminal Domain . . . 143
8.2 The role of the G/F Rich Region in DNAJB6 . . . 145
8.3 Delineating an Allosteric Network . . . 145

IV APPENDICES 147
A NMR ASSIGNMENT INFORMATION 149
A.1 HSC70 SBD Assignments . . . 149
A.1.1 Assigment Statistics . . . 149
A.1.2 Assigment Table . . . 149
A.2 DNAJB6b 1-130 . . . 156
A.2.1 Assigment Statistics . . . 156
A.2.2 Assigment Table . . . 156
B CONSTRUCT SEQUENCES 161
B.1 HSC70 (HSPA8) . . . 161
B.1.1 Full-Length . . . 161
B.1.2 NBDSBD . . . 162
B.1.3 NBD . . . 163
B.1.4 SBD . . . 164
B.1.5 SBDCTD . . . 164
B.2 HSP70 (HSPA1A/B) . . . 165
B.2.1 Full-Length . . . 165
B.3 HSP70-HSC70 Chimera . . . 166
B.3.1 1ANS8CT . . . 166
B.4 DNAJC5 . . . 167
B.4.1 DNAJC5 1-100 . . . 167
B.5 DNAJB6 . . . 168
B.5.1 DNAJB6 1-77 . . . 168
B.5.2 DNAJB6 1-100 . . . 168
B.5.3 DNAJB6 1-130 . . . 169
B.6 DNAJB2 . . . 169
B.6.1 DNAJB2 1-100 . . . 169
B.6.2 DNAJB2 1-130 . . . 170
B.7 BiP . . . 170
B.7.1 Full-Length . . . 170
B.8 σ_1 . . . 171
B.8.1 GFP-σ_1 . . . 171
B.8.2 σ_1-Δ35 . . . 172

BIBLIOGRAPHY 175

LIST OF FIGURES

Figure 1 Model of the Cellular Role of the σ_1 Receptor 9
Figure 2 The σ_1 Chaperone Domain in a Structural Context 11
Figure 3 Structure of the σ_1 Receptor 13
Figure 4 Prior and Current Topologies of σ_1 16
Figure 5 Anatomy of a HSP70 Protein 21
Figure 6 Schematic of the HSP70 ATPase Cycle 24
Figure 7 Structure of the J-domain 26
Figure 8 Schematic of Class A B and C HSP40 Proteins 28
Figure 9 Purification Scheme for σ_1 55
Figure 10 Oligomeric State in DPC 57
Figure 11 Comparison σ_1-Δ35 of and Full-Length σ_1 HSQC Spectra . 59
Figure 12 Assigments for Full-Length σ_1 in DPC 60
Figure 13 Secondary Structure Prediction and Actual Structure Comparison 63
Figure 14 Crosslinking and TRACT in DDM 67
Figure 15 SEC-MALS and Native MS in DDM 68
Figure 16 Chaperone Assays in DDM 70
Figure 17 σ_1-HSP70 Protein Interactions by MST 73
Figure 18 ^{1}H ^{15}N TROSY titration of HSC70 NBD into σ_1 75
Figure 19 ^{1}H ^{13}C HSQC titration of HSC70 NBD into σ_1 76
Figure 20 ^{1}H ^{15}N TROSY titration of σ_1 into HSC70 NBD 78
Figure 21 ^{1}H ^{13}C HSQC titration of σ_1 into HSC70 NBD 79
Figure 22 CD spectra in DDM and DPC 80
Figure 23 DDM and DPC ^{1}H ^{15}N TROSY Overlay 81
Figure 24 SEC-MALS of full-length σ_1 in LMNG 83
Figure 25 Localisation of σ_1 in Different Cell Lines 86
Figure 26 ER Localisation of σ_1 in COS7 Cells 87
Figure 27 RRQQ Localisation Signal Mutant 88
Figure 28 Interactions with BiP and HSC70 by FRET-FLIM 90
Figure 29 Explanation of N&B 92
Figure 30 Brightness Boxplots Before and After Addition of Different σ_1 ligands 95
Figure 31 Oligomeric Behaviour of Full-Length HSC70 106
Figure 32 Oligomeric Behaviour of HSC70 With and Without ATP . . 107
Figure 33 Oligomeric Behaviour of HSC70 Truncation Constructs . . . 108
Figure 34 Spectra of ^{15}N Labelled Full-Length HSC70 110
Figure 35 Backbone Spectrum of the NBDSBD of HSC70 112
Figure 36 Backbone Spectrum of Selectively Labelled Deuterated Full-Length HSC70 113
Figure 37 Overlay of Backbone Spectra of Truncations and Full-Length HSC70 115

Figure 38 Assignments of the HSC70 SBD 116
Figure 39 Native Purification of DNAJB6 and DNAJB2 117
Figure 40 Purification of Full Length DNAJB6 with 2 M Urea 119
Figure 41 Purification of J-Domain Constructs 122
Figure 42 Assignments of the J-domain of DNAJC5 (1-100) 124
Figure 43 Assignments of the J-domain of DNAJB6 (1-130) 125
Figure 44 NMR Comparison of DNAJB6 Truncation Constructs 126
Figure 45 The HSC70-DNAJC5 J-Domain Interaction by SEC 128
Figure 46 Titration of HSC70 into ^{15}N Labelled DNAJC5 (1-100) . . . 129
Figure 47 Titration of HSC70 into ^{15}N Labelled DNAJB6 (1-77) 130
Figure 48 ^{1}H ^{15}N TROSY Titration of DNAJC5 J-domain into HSC70 NBDSBD 132
Figure 49 ^{1}H ^{15}N TROSY Titration of DNAJC5 J-domain into Full-Length HSC70 134
Figure 50 ^{1}H ^{13}C SOFAST HMQC Titration of DNAJC5 J-domain into Full-Length HSC70 135
Figure 51 Structural Map of Shifts for DNAJC5 1-100 - HSC70 Interaction 136
Figure 52 Effects of the G/F Rich Region on J-Domain Function 138
Figure 53 Stimulation of HSC70 and HSP70 by Different J-domains . . 139
Figure 54 Effects of the CTD on the Response of HSC70 and HSP70 to different J-domains 141

LIST OF TABLES

Table 1 Brightnesses Before and After Addition of Different σ_1 Ligands 96
Table 2 Activity of HSC70 and HSP70 in the Presence of Different J-domains 140

ACRONYMS

ADP Adenosine DiPhosphate

AFM Atomic Force Microscopy

ALS Amylotrophic Lateral Sclerosis

APH AmphiPathic Helix

ATP Adenosine TriPhosphate

BiP Binding immunoglobulin Protein

CD Circular Dichroism

CNS Central Nervous System

CSP Cystine String Protein

CTD C-Terminal Domain

DDM DoDecyl Maltoside

DMT N,N-DiMethylTryptamine

DPC DodecylPhosphoCholine

EGFP Enhanced Green Fluorescent Protein

ER Endoplasmic Reticulum

ERAD ER-Associated protein Degradation

FCCS Fluorescence Cross Correlation Spectroscopy

FLIM Fluorescence Lifetime Imaging Microscopy

FRAP Fluoresence Recovery After Photobleaching

FRET Förester Resonance Energy Transfer

GPCR G-Protein Coupled Receptor

hERG human Ether-à-go-go-Related Gene

HSJ1 *Homo Sapiens* dnaJ 1

HPLC High-Pressure Liquid Chromatography

HSC70 Heat Shock Cognate 70 kDa

HSP70 Heat Shock Protein 70 kDa

HSQC Heteronuclear Single Quantum Coherence spectroscopy

IMAC Immobilised Metal Affinity Chromatography

LMNG Lauryl Maltose Neopentyl Glycol

IP_3R Inositol trisPhosphate Receptor

IST Iterative Soft Thresholding

MALS Multi-Angle Light Scattering

MAM Mitochondrial Associated er Membrane

MBP Maltose Binding Protein

MS Mass Spectrometry

MST MicroScale Theromophoresis

N&B Number and Brightness

NBD Nucleotide Binding Domain

NEF Nucleotide Exchange Factor

NMDAR N-Methyl-D-Aspartic acid Receptor

NMR Nuclear Magnetic Resonance

NUS Non-Uniform Sampling

PM Plasma Membrane

σ_1-Δ_{35} 35-Residue N-terminal Truncation ofσ_1

σ_1-CD σ_1 Chaperone Domain

SBD Substrate Binding Domain

SBDL I Substrate Binding Domain Like I

SBDL II Substrate Binding Domain Like II

SEC Size Exclusion Chromatography

SSRI Selective Seratonin Re-uptake Inhibitor

$STIM_1$ Stromal Interaction Molecule 1

TCEP Tris(2-CarboxyEthyl)Phosphine

TEV Tobacco Etch Virus protease

TM Trans-Membrane helix

TRACT TROSY for RotAtional Correlation Times

TROSY Transverse Relaxation Optimized SpectroscopY

UIM Ubiquitin Interacting Motif

Part I

OVERVIEW OF PROTEOSTASIS

1

INTRODUCTION

Proteins are for the large part the agents that facilitate the plethora of processes essential for life. Like all cellular systems they are subject to attrition as part of normal operation and to stress induced damage under unfavourable conditions. Ensuring that the constant turnover of proteins is well regulated and that any damage is duly mitigated is essential for maintaining a healthy cell.

At the heart of both protein synthesis and degradation, as well as the cells response to proteomic damage lie the cellular chaperones. These proteins are tasked with aiding folding during protein synthesis, preventing and dealing with the aggregation of cellular proteins and modulating certain protein-protein interactions, both under normal and stress induced conditions [1, 2]. Understanding the way these chaperones function and how they are regulated is key to developing a fully formed picture of proteostasis.

Not only is such an understanding crucial to basic biological science, but it also provides avenues for the treatment of numerous associated pathologies [3]. The effects of proteinopathies such as Alzheimer's or ALS can be mitigated or even avoided by robust chaperone function. Conversely, pathologies such as cancer and infectious disease rely on the chaperone machinery, either their own or their hosts', to proliferate and spread. In both cases chaperones and their regulators present potential therapeutic targets [4] .

This network of chaperone proteins and their regulators hinges on several key operational hubs. One such hub of great importance is provided by the HSP70 family of proteins [5]. These proteins present across all domains of life and in almost every cellular compartment, represent versatile ATP driven machines. Their specificity and function are controlled through a large host of interacting proteins that serve to channel their efforts to where they are needed in the cell. The work

detailed here aimed to expand our structural understanding of the interactions of two regulators of HSP70 proteins with their HSP70 partners. The first of these, the σ_1 receptor, represents a non-canonical ligand-regulated HSP70 partner protein implicated in the cellular stress response. The second, the HSP40 family, are archetypical regulators of HSP70 proteins. The three members studied are DNAJC5, and the closely related pair of proteins DNAJB6 and DNAJB2. They represent both a modulator of a constitutive cellular process (DNAJC5) [6] as well as proteins involved in response to stress mainly in the context of poly-Q protein aggregation (DNAJB2 and DNAJB6) [7, 8].

The weak, transient and often fuzzy complexes that underlie the regulation of the HSP70 family provide fruitful grounds for investigation by NMR [9]. Equally the prevalence of allostery both within the HSP70 machines and their complexes with regulatory partners is well complemented by the capacity of solution state NMR to track dynamic structural changes at atomic resolution. Understanding the regulation of the chaperone machinery therefore provides not only an interesting and relevant process to study, but one which we possess, at least in theory, the tools to dissect in detail.

1.1 THE σ_1 RECEPTOR

The σ_1 receptor represents one of the most pharmacologically diverse and poorly understood drug targets in humans. This small (25 kDa) single transmembrane receptor is found in high abundance in the Central Nervous System (CNS) and appears to have strong neuroprotective effects and associations with a number of pathologies, ranging from stroke and cancer through to depression and addiction [10]. Behind this large array of cellular properties lie its interactions with a myriad of other proteins and small molecules. It is these intriguing properties of σ_1 that make it such a compelling target for investigation.

Since its discovery in the late 1970s [11] research on σ_1 has expanded substantially. Despite the now enormous volume of papers that abound, structural inves-

tigations of the receptor have been few and far between. Most published work has been constrained to pharmacological and cell biological studies and despite substantial interest the structure of σ_1 remained unsolved until 2016 [12]. The structure itself resolved several questions in the field explaining how σ_1 can accommodate such a diverse array of ligands, and validating prior work developing pharmacological profiles of its binding pocket [13]. Nonetheless, the structure left some questions unanswered and created new ones. The single transmembrane homotrimer, defied earlier predictions of a two-transmembrane dimer [14–17] and opened up the question of receptor topology. Furthermore, the observed changes between agonist- and antagonist-bound structures are small, leaving the structural demarcation of this classification unclear [12, 18]. This seemingly unending nature of the σ_1 mystery is most likely what serves to maintain those who work on it enthralled, despite the many difficulties and setbacks that the field often throws up.

1.1.1 *Discovery of the σ_1 Receptor*

Initial identification of σ receptors stemmed from their drug binding properties [11]. Much of the early research in the field focused on fully characterising this behaviour. The reversed stereospecificity relative to the opioid receptors and the failure of σ_1 sites to bind several key opioid receptor ligands led to the recognition of σ sites as being novel pharmacological entities [19]. Further work on the drug binding profiles of these sites demonstrated the existence of two receptors [20], now termed σ_1 and σ_2, with similar but distinct drug binding profiles and physiological effects. The σ_1 receptor gene was successfully identified and cloned in the early 1990s [21] while σ_2 proved a tougher target with its gene remaining unidentified until 2017 [22]. Cloning of σ_1 substantially expanded the field with much more detailed studies and made structural investigations possible.

1.1.2 *Pharmacological Properties*

Since the initial pharmacological work identifying σ_1, its drug binding profile has expanded substantially with σ_1 being shown to interact with a large number of compounds from different chemical classes. These include therapeutic agents such as haloperidol [19] (a psycholeptic anti-psychotic) and fluvoxamine [23, 24] (an SSRI anti-depressant) as well as recreational drugs such as N,N-dimethyltryptamine [25] (DMT) and cocaine [26]. In many cases these clinically important ligands have been shown to exert at least part of their effect through σ_1. In addition to its synthetic drug binding profile several endogenous ligands have been proposed for σ_1 including steroid hormones [27] such as progesterone [28], long acyl chain compounds such as sphingosine [29] and myristic acid [30], and endogenously produced DMT [25].

In general, only the broadest classification of σ_1 ligands into agonists and antagonists has achieved wide acceptance. Archetypal antagonists of σ_1 include haloperidol and NE-100 while archetypal agonists include pentazocine and SKF-10047. In both cases the definition remains strictly empirical and based on the capacity of the compounds to elicit (agonist) or inhibit (antagonist) often different σ_1-mediated responses in an organism or specific cellular pathway; as a result, different ligands will be classified in different contexts with no systematic definition prevailing.

1.1.3 *The σ_1 receptor in disease*

The clinical importance of σ_1 is underpinned by its remarkable pharmacology, which provides much of the impetus for its study. σ_1 agonists for example appear to possess strong neuroprotective effects [31]. The most dramatic demonstration of this can be found in rat models of ischemic stroke, with σ_1 ligands reducing infarct volumes by 80% over the control group [32]. Continued research has strengthened this link and highlights the potential of σ_1 agonists as treatments for stroke. In a similar manner, the neuroprotective effects of σ_1 can be leveraged to

tackle Alzheimer's. Animal models have shown that symptoms associated with Alzheimer's can be alleviated by selective σ_1 agonists [33, 34] and further work in cells has demonstrated that σ_1 can protect against peptide toxicity [35].

The clinical potential of σ_1 goes further than its impressive neuroprotective properties. σ_1 is known to modulate a range of signalling pathways from GPCRs through to ion channels [10, 36]. Modulation of these pathways underlies a number of other potential therapeutic uses for its ligands. σ_1 antagonists for example potentiate opioid analgesia through stimulation of opioid receptor signalling and inhibition of NMDAR responses [37].

The above examples are by no means exhaustive. To give but a few other examples, the role of σ_1 in cell survival implicates it in cancer and its protective effects have been posited to aid in treatment of ALS [38]. σ_1 modulation of signalling pathways also provides avenues for treating depression and addiction — two further pathologies where σ_1 ligands have shown therapeutic potential [10].

1.1.4 *The erg2 family and the cellular role of σ_1*

σ_1 is a member of the erg2 family and shares ~30% identity and ~66% similarity to its closest homologues, the fungal C8-C7 sterol isomerases [21]. These enzymes are essential for ergosterol synthesis, the fungal homologue of cholesterol. Their catalytic activity hinges on the stabilisation of a positively charged carbocationic intermediate formed while shifting a double bond from the C8 to the C7 position.

Traces of this ancestry are evident in the ligand binding profile of σ_1. Studies have shown that potent inhibitors of sterol isomerases (mimics of the positively charged intermediate) also bind σ_1 with high affinity [39]. The profile of these molecules being a close fit for the σ_1 pharmacophore, which favours a positive ionisable feature flanked by two hydrophobic structural elements. In most high affinity σ_1 ligands this positive ionisable feature is a nitrogen atom [40]. The similarity in binding between the two sets of molecules extends further: hydrophobicity and homology analysis have identified two hydrophobic regions in σ_1 homologous

to the substrate binding regions of the erg2 family, termed Substrate Binding Domain Like (SBDL) I and II [41], and many steroid and sterol like molecules show significant affinity for σ_1 [27, 28, 42] .

Despite similarities in ligand binding, no enzymatic activity in σ_1 has been detected and the protein seems to act instead as a receptor. Naturally a receptor must form part of a signalling pathway and much of the work on σ_1 has addressed the question of what this pathway looks like. Current models posit that an inactive form of σ_1 is held at the Mitochondrial Associated ER Membrane (MAM) [14, 36]. Several lines of evidence indicate that an association with the ER resident HSP70 protein BiP, [14, 43] as well as the formation of large immobile oligomeric assemblies [17, 44, 45] are responsible for maintaining this inactive state of the receptor. Activation of the receptor either by agonists or through cellular stress results in the disassembly of these complexes and subsequent movement of the receptor away from the MAM as shown in Figure 1 [36, 46]. Antagonists conversely seem to function by blocking activation of σ_1 causing the protein to remain in inactive assemblies.

Although the pathways of σ_1 activation in response to stress have yet to be well delineated, several lines of evidence provide a strong link to intracellular calcium levels and related signalling pathways. Disruption of the σ_1-BiP interaction has been observed to occur both in response to agonists and to falling ER calcium concentrations, which are symptomatic of cell stress [14]. Active σ_1 has also been observed to modulate Inositol Trisphosphate Receptors (IP3Rs) at the MAM, potentiating calcium signalling [14] into mitochondria as well as participating in STIM1 (Stromal Interaction Molecule 1) mediated calcium signalling [47]. Relocation of the active receptor either purely within the ER or to the Plasma Membrane (PM) also allows for the multiple modulatory interactions between $\sigma_1$and its numerous putative client proteins to take place [36]. Work with fluorescently labelled σ_1 has shown that once activated, σ_1 can also move away from the MAM [46]. Further cell biological work has supported this notion; activation with cocaine, for example, has been reported to disassociate σ_1 from BiP allowing it to relocate to the

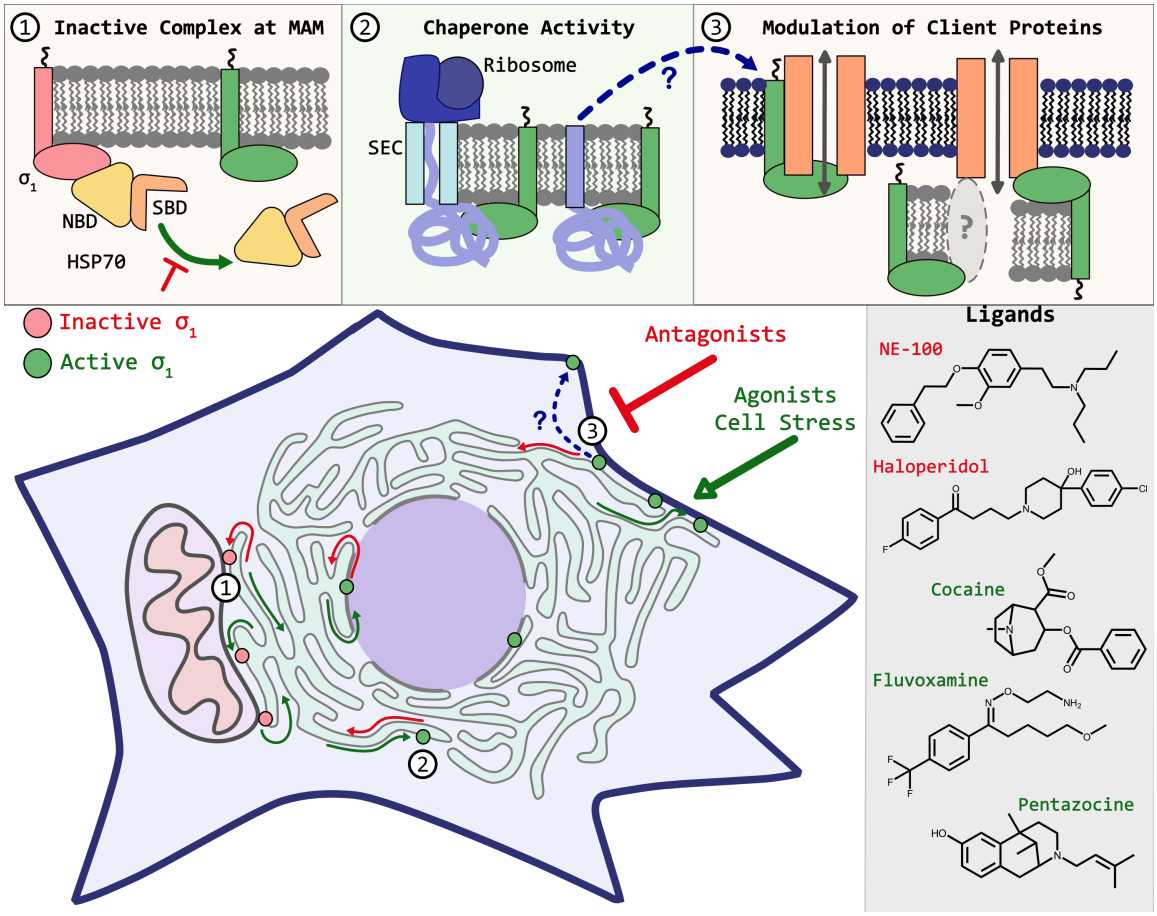


Figure 1: Main Panel: Schematic showing re-localisation of σ_1 by agonists or through cell stress. Inactive σ_1 (red circles) is present at the MAM and activation by agonists results in disassembly of these complexes causing σ_1 to spread throughout the ER. Translocation to the PM (blue dotted line) on activation may occur but is yet to be conclusively proved. This activation process is blocked by antagonists. Numbered regions are shown in more detail in the corresponding top panel. **Top Left Panel:** Inactive σ_1 at the MAM is present either as oligomeric assemblies or, as shown, in a complex with a HSP70 protein family partner (likely BiP). Dissociation of these complexes is stimulated by agonists and cell stress resulting in activation and re-localisation of σ_1. This process is blocked by antagonists. **Top Middle Panel:** Modulation of client protein folding by σ_1 through possible chaperone interactions with a partially folded client. A putative mechanism for translocation to the PM while bound to a nascent PM destined client is shown (blue dotted line). **Top Right Panel:** Modulation of client proteins by active σ_1, in most cases this occurs through direct protein-protein interactions either in the ER, Nuclear Envelope or PM. Whether these interactions with PM resident clients occur in trans or within the same membrane remains to be confirmed. Some interactions may also be indirect through a yet uncharacterised ternary complex. All three possibilities are shown. **Right Inset:** Example σ_1 ligands with antagonists labelled in red and agonists labelled in green. Apart from cocaine (micromolar affinity) all ligands shown are tight binders (nanomolar affinity). The ligands shown contain large hydrophobic moieties flanking a positive ionisable feature in the form of a nitrogen atom. These features are typical of most high affinity σ_1 ligands.

nuclear envelope and interact with emerin to affect transcription [43]. While this provides a useful framework with which to begin to understand σ_1, many aspects of this model are yet to be fully resolved.

1.1.4.1 *Protein-Protein Interactions*

The protein-protein interactions of the activated form of σ_1 appear to be the key to its cellular role. The list of σ_1 clients is extensive, including several signalling proteins and ion channels which seem to be directly modulated by σ_1. This list includes GPCRs [48], the HSP70 chaperone BiP [14], numerous ion channels from the NMDA receptor [37] through to voltage gated sodium [28, 49] and potassium channels [15, 50, 51], protein kinases [52] and transcription factors [43]. With the exception of its interaction with BiP, these interactions have generally been seen to be stimulated by agonists and inhibited by antagonists.

Since the initial observation of σ_1 directly modulating Kv 1.4 and Kv 1.5 ion channels [15], evidence for a direct interaction between σ_1 and its client proteins has been steadily building. Patch clamp studies have shown the ability of σ_1 to directly modulate numerous other ion channels [28, 50, 51]. Coimmunoprecipitation [15, 50] and FRET [53, 54] have also confirmed direct interactions.

Further evidence for a direct interaction has been provided by Atomic Force Microscopy (AFM). To date AFM has shown the direct association of σ_1 with hERG [53], Nav 1.5 [55], ASIC [56] and NMDAR [57] ion channels. In addition to this, AFM has yielded some insight in the structure of these complexes. In all cases a 1:1 stoichiometry of σ_1 molecules to channel subunits was observed, however in the case of the GluN1 subunits of the NMDA receptor both 2:1 and 1:1 stoichiometries were detected.

Though by no means a complete list, the above examples demonstrate the striking versatility of the small σ_1 protein to interact with various clients. How σ_1 identifies its clients and achieves specificity remains completely unanswered. Detailed structural characterisation of these interactions has also proved elusive; and,

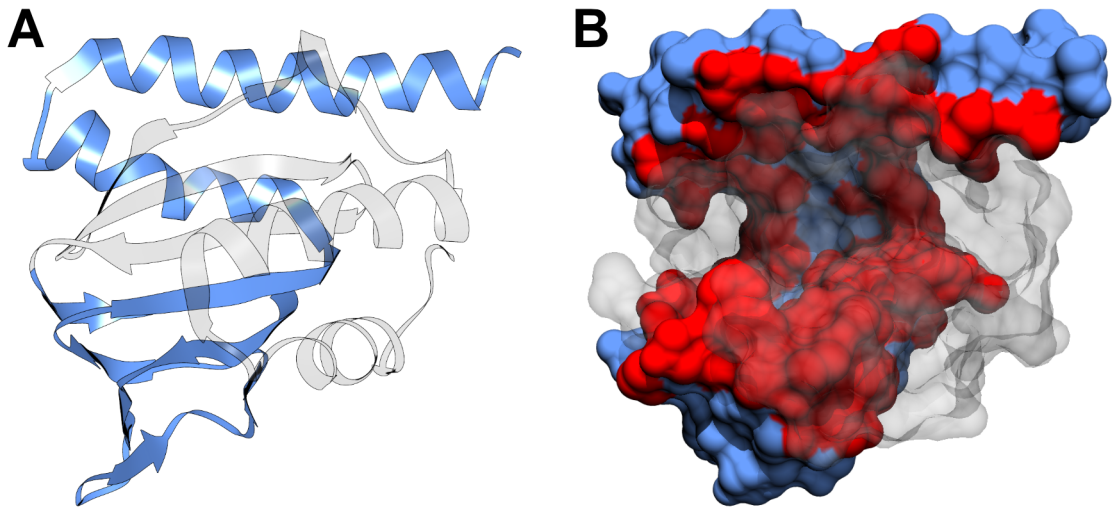


Figure 2: Structure of σ_1 (PDB: 5HK1) focusing on the β-barrel domain with the TM helix removed for clarity. **A:** Ribbon diagram with the chaperone domain (residues 112-223) shown in blue and the rest of the β-barrel domain shown in semi-transparent grey (residues 35-111). **B:** Surface plot produced using Chimera [58] in which the opaque portion of the plot in blue and red represents the chaperone domain, the semi-transparent grey portion corresponds to the rest of the β-barrel domain, and contact residues between the chaperone domain and the rest of the β-barrel domain are shown in red.

beyond some idea of the stoichiometry of some of these interactions offered by AFM, the structural details remain a black box in our understanding of σ_1.

1.1.4.2 *Chaperone Activity*

In addition to its ligand binding activity a now emerging aspect of σ_1 is its potential activity as a cellular chaperone. σ_1 has been shown to bind to and affect the trafficking and maturation properties of hERG channels [53]. Furthermore, a more targeted study with isolated σ_1 has provided evidence that σ_1 may possess chaperone activity. A construct comprising residues 116-223 of σ_1 showed the ability to supress the heat induced aggregation of citrate synthase [14]. This anti-aggregation activity was shown to be additive with Binding immunoglobulin Protein (BiP, HSPA5) an ER resident member of the HSP70 family of chaperones.

Although the possibility of σ_1 chaperone activity is compelling in light of the strong neuroprotective activity displayed by σ_1 agonists; the relationship between σ_1 and the cellular chaperone machinery remains poorly examined. On the struc-

tural front our own work on a construct of σ_1 comprising residues 112-223 (σ_1-CD) was able to recapitulate the interaction between σ_1 and BiP *in vitro* [59]. The putative chaperone activity of the protein however has not been reproduced in any subsequent biophysical studies and is complicated by the requirement to stabilise σ_1 or the σ_1-CD in detergent.

The recently determinined structure of the receptor further confuses matters. At the time of the publication of the structure, a region starting at about residue 116 and ending at the C-terminus of the receptor, was expected to form a self-contained domain within σ_1 based on protease protection assays [14]. This, however, cannot be the case, since a construct comprising these residues would cut the neatly folded β-barrel of σ_1 in half exposing a substantial amount of buried surface area (Figure 2). Although chaperone activity for peptide segments of known chaperones has been observed it seems difficult to imagine that this segment of the receptor would fold well enough to recapitulate native-like activity. This raises questions about the observation of co-operativity between BiP and σ_1-CD.

The potential for chaperone activity in σ_1 and its many apparent similarities to elements of the cells proteostasis machinery have not escaped notice with a recent review in the field highlighting a potential link between σ_1 and the small heat shock proteins [44]. Nonetheless the role of σ_1 in protein folding requires further characterisation, and with no clear demonstration of chaperone activity in isolated full-length protein, the question of whether σ_1 directly prevents aggregation and or stimulates folding remains.

1.1.5 *Structure of σ_1 and Insights into Ligand Binding*

The structure of σ_1, published part way through our work on σ_1 in 2016 [12], helped explain its extraordinary ligand binding profile. The receptor adopts a β-barrel fold, with the interior of the β-barrel serving as a large hydrophobic pocket capable of binding many diverse compounds (Figure 3). This structure fits well with the previously characterised pharmacophore for the receptor and mutagene-

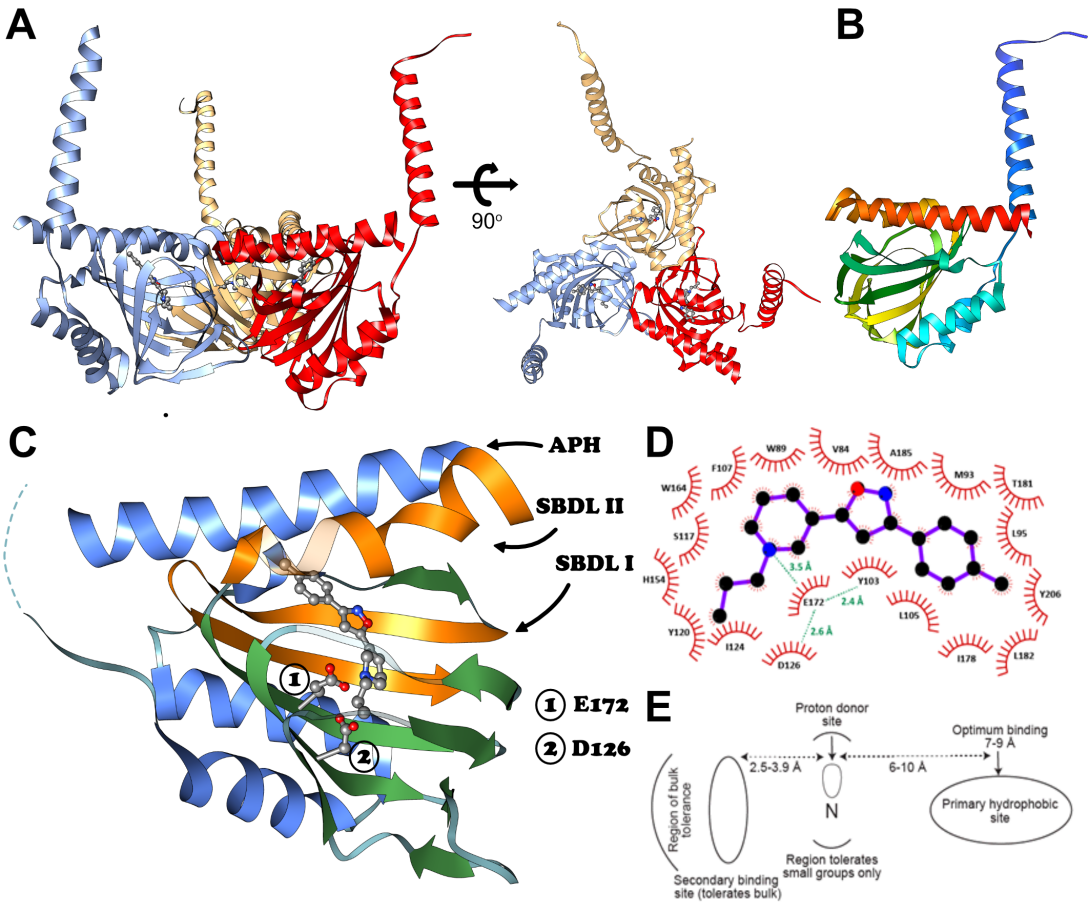


Figure 3: **A:** *Left*, homotrimeric structure of σ_1 (PDB: 5HK1) viewed in the plane of the membrane. *Right*, homotrimer viewed from above. **B:** σ_1 asymmetric unit showing an individual monomer. (PDB: 5HK1) **C:** Focused view of the ligand binding pocket. For clarity, the TM helix has been removed and loops obscuring the ligand have been made partly transparent. Residues critical for binding, E172 and D126 have been highlighted and regions of interest labelled. **D:** Modified LigPlot output [60] showing the ligand binding pocket with bound PD144418. Distances are shown from Cδ of E172 to N4 of PD144418, from an Oϵ of D126 to an Oϵ of E172, and from OH of Y103 to an Oϵ of E172. Contacts on the far left correspond to the bottom of the β-barrel, and contacts on the far right correspond to SBDL II (Helix 4) and the Amphipathic Helix (APH, Helix 5) which form the membrane attachment region. **E:** Proposed pharmacophore model for σ_1. [40] On the structure shown in Figure 3C the primary binding site can be seen to occur at the bottom of the σ_1 β-barrel oriented towards D126, while the larger region of bulk tolerance can be seen to occur towards the top of the β-barrel extending out to SBDLII and the APH. E172 a residue known to be critical for ligand binding can be seen to form a charge-charge interaction with the nitrogen atom of PD144418 corresponding to the proton donor site in the σ_1 pharmacophore. The structure shows that E172 is stabilised by hydrogen bonds both to Y103 and D126 the latter of which has also been shown to be critical for ligand binding.

sis data [13]. Glutamate 172, a residue essential for ligand binding [61], provides a means of interacting with the positively ionisable feature common to many σ_1 ligands (Figure 3C).

The large hydrophobic pocket at the centre of the barrel can flexibly accommodate the flanking hydrophobic regions of σ_1 ligands with the smaller region oriented towards the bottom of the barrel and away from the membrane. While the larger region is oriented towards the top of the barrel and the membrane anchoring face of σ_1 (Figure 3D). This mode of binding appears to completely occlude ligands from the surrounding solvent which may explain the unusually slow binding kinetics seen for σ_1 [18].

Nonetheless, not all the questions about the pharmacology of σ_1 have been resolved by the published structures. Exactly how ligands enter the occluded cavity of σ_1 remains unaddressed. Furthermore, and perhaps more crucially, the mechanistic distinction between agonists and antagonists has not been clarified. Structures of the trimeric receptor bound to either class of molecule show only small differences. The largest of these is a 1.8 Å pushing apart of helices 4 and 5 [18]. A larger version of this rearrangement could disrupt the oligomeric interface and may explain one difference between the two compounds (see below), but this remains untested.

1.1.5.1 *Oligomerisation*

Evidence for σ_1 oligomers existed for some time prior to the publication of its structure. A study using selective photoreactive-probes showed labelling of higher molecular weight species [62]. Additional evidence for oligomerisation was observed directly in protein recombinantly expressed in *E. coli*, with two distinct oligomeric species being observed and isolated, a smaller potential dimer and a larger possibly hexameric or octameric assembly [17]. Tests on different isolated oligomeric forms of the receptor have suggested that only oligomeric σ_1 is competent to bind ligands [17].

Furthermore, several drug binding assays suggested stoichiometries of 2:1 protein to drug [17, 63]. Work with 4-NCCP-12, a photoreactive ligand, showed that only 50% of receptor molecules were labelled both for isolated protein and protein in guinea pig liver microsomal extracts [64]. Again, this is all suggestive of dimers or possibly larger, even-numbered assemblies of σ_1 being functionally important.

Although the trimeric structure of σ_1 supports the importance of oligomerisation, as do the differences between agonist and antagonist bound structures discussed in subsection 1.1.5, they also present inconsistencies with prior observations. The crystal structures show each individual monomer bound to a ligand that is not at the subunit interface, which coupled to the trimeric structure seems to contradict a 2:1 stoichiometry for binding. Thus, the apparent requirement of oligomerisation for ligands to bind remains unexplained.

The importance of oligomerisation does not rely on structural and biochemical work alone. *In cellulo* work using YFP and GFP tagged σ_1 in COS-7 cells showed the presence of monomers, dimers and higher oligomeric assemblies of σ_1 [45]. This indicates that the various oligomeric forms of σ_1 are likely to be biologically relevant. Oligomerisation in this system was shown to be stimulated by antagonists.

Most of this data supports the model of active monomers and inactive oligomers of σ_1. There is however a patent gap in our understanding both structurally and biologically. What the structural implications of these diverse assemblies are remains unknown. More importantly the presence of a substantial quantity of monomers in live cells and the apparent 1:1 stoichiometry of many σ_1 client protein interactions highlight the importance of this so far uncharacterised conformation of the receptor. Understanding the structure and function of smaller σ_1 assemblies such as dimers or monomers thus seems the next logical step in clarifying the cell biology of σ_1 and understanding the mechanistic differences between agonists and antagonists.

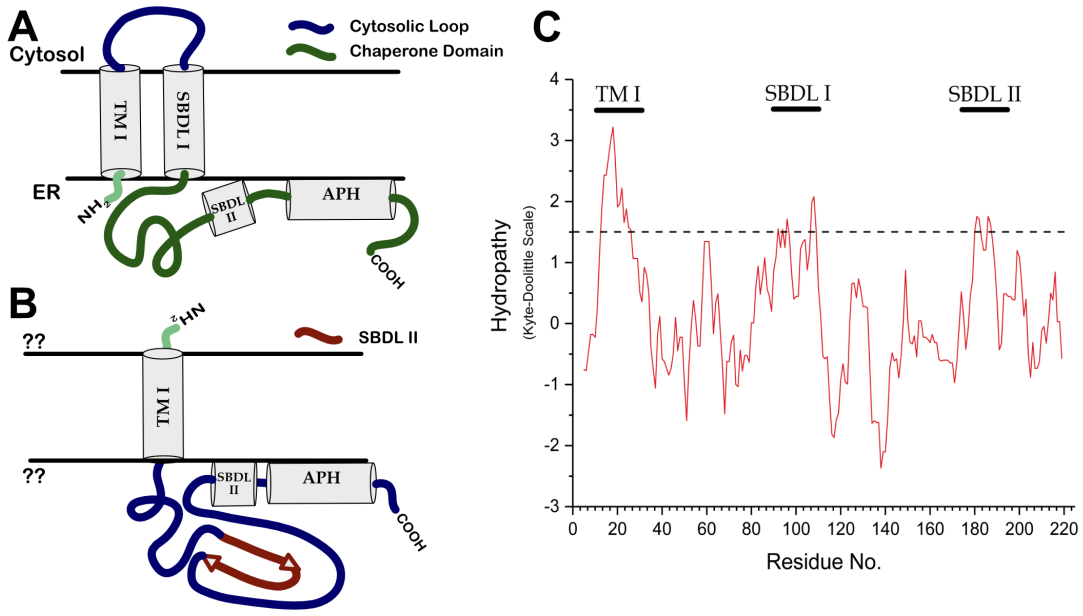


Figure 4: **A:** Previously proposed topology for σ_1 based on protease accessibility studies by Hayashi and Su [14], as well as secondary structure prediction by NMR in DPC detergent [16, 59]. SBDL I and II are highlighted and juxtaposed as suggested by photocrosslinking by Ruoho et. al. [65] **B:** Actual topology of σ_1. The SBDL I and II regions remain juxtaposed, though the protein contains only one TM and SBDL II comprises a pair of strands in the soluble β-barrel of the receptor. **C:** Hydropathy plot of σ_1 using the Kyte-Doolittle hydrophobicity scale and a window function of 9. The first TM and two SBDL regions are highlighted.

1.1.6 *Topology and Localisation*

A hydropathy plot of σ_1 (Figure 4A) reveals three distinct regions of high hydrophobicity; the strongest of these occurs proximal to the N terminus between residues 10 and 30. Unsurprisingly the earliest proposed topology for σ_1, based on sequence analysis, predicted a single transmembrane domain in this region. Subsequent work in the field however did much to advance a two-transmembrane model of the receptor (Figure 4A) which had become widely assumed by the time the structure was published [36, 44].

The two-transmembrane model hinged largely on two accessibility studies. The first in Xenopus oocytes used fluorescent antibodies to establish that both N and C termini of the protein were on the exterior of the cell. This was then further corroborated through a biotinylation assay showing that residues 30 to 80 were accessible on the exterior of the cell [15]. Further antibody accessibility work in

Chinese Hamster Ovary (CHO) Cells supported the notion of two TMs but in this case σ_1 was found to be localised exclusively to the ER and not to the PM [14].Despite the differing localisation in both studies the location of the N and C termini on the same side of the membrane appeared clear. Additionally, the two TMs were further corroborated through sequence based transmembrane predictions with 10 of 11 TM predictors tested showing a TM in or near SBDL I [16], a notion that was further supported by hydropathy plots of σ_1 (Figure 4C).

Studies into the localisation of σ_1 confirmed an ER localisation in a number of mammalian cell lines. Localisation of the receptor to the PM however remains unclear in part due to difficulties in differentiating the PM from regions of the ER closely associated with it. Several studies have been unable to detect σ_1 at the PM [14, 46] and are potentially contradicted by other studies that detect σ_1 at or near the PM [15, 52, 66]. In several cases evidence has been put forward that σ_1 localises to regions of the ER that are closely juxtaposed to the PM [67, 68].

1.1.6.1 *Orientation*

The single TM structure of the receptor may well have laid the question of the number of TMs in σ_1 to rest but it has thrown open the question of receptor orientation. The prior model placed a significant portion of the protein on either side of the membrane allowing for a straightforward rationalisation of how interactions could simultaneously be observed in the cytosol and ER lumen. This is no longer the case, since the short eight residue tail which remains on one side of the membrane is poorly conserved and seems very unlikely to be responsible for any protein-protein interactions. Equally the first TM does not demonstrate strong conservation and would not serve to explain interactions with soluble cytosolic regions of client proteins.

Recent work using a novel EM based labelling technique addressed questions related to the orientation of σ_1 and has put forward an ER- facing orientation for the bulk of σ_1 [69]. Although attractive, this hypothesis does not offer a complete explanation since this orientation would require either the re-localisation of σ_1 to

the PM to affect PM resident clients or the formation of some as yet unknown ternary complex at the ER-PM interface.

Additionally, at least two interactions in the literature seem irreconcilable with an ER orientation. The first is an interaction with the cytosolic C-terminus of the NR1 subunit of the NDMA receptor [37]. The second is with the nuclear envelope resident protein emirin [43]. In the latter case it would be possible to postulate a TM-TM interaction though for reasons discussed above this seems unlikely. In the absence of this it is hard to see how two proteins which reside largely on opposite sides of the membrane could interact.

A cytosolic orientation of σ_1 hence remains tenable. An interaction in trans between a cytosol facing σ_1 and the cytosolic domains of its clients at the PM sits well with observations of σ_1 moving to regions of the ER closely juxtaposed to the PM and not the PM itself [67, 68]. Furthermore, as the cytosol is topologically continuous with the nucleus a direct interaction with emirin would be well explained. The main contention with this orientation would be the observed interaction with ER-resident BiP. With both orientations posing unresolved clashes with current data the orientation of σ_1 remains an important area in need of clarification.

1.1.7 *Link to Proteostasis*

That σ_1 seems to be directly involved in the cellular stress response [14], that it can be shown to bind proteins during their maturation [53], and may possess chaperone activity [14], suggests a link between σ_1 and proteostasis. Its strong neuroprotective effects, its involvement in neurodegenerative diseases and its involvement in cell survival [10] are biological behaviours with close associations to cellular chaperones.

Perhaps the most direct link seems to be the interaction of σ_1 with members of the HSP70 family, the main candidate being the ER resident BiP [14]. As such σ_1 may act as another modulatory protein which feeds in a distinct subset of signals into the HSP70 machinery. Coupled with the possibility of σ_1 possessing

its own chaperone activity, the notion of σ_1 acting as a co-chaperone to the cellular protein folding machinery seems particularly pertinent. The prominence of σ_1 in electrically excitable cells, and its regulation of the ion channels that are responsible for excitability elicits the possibility that the native role of σ_1 is to provide a link between the chaperone machinery and electrical excitability.

1.1.8 *Studying* σ_1

Given the gaps present in our structural and functional understanding of σ_1 we sought to address several outstanding questions by developing a system for the study of σ_1 using solution state NMR as a central tool to study its dynamics and structure:

- What are the structural differences between agonist-bound and antagonist-bound σ_1?
- How do small molecules gain access to the interior cavity of σ_1?
- Does full length purified σ_1 exhibit chaperone activity?
- What is the structural basis of this activity? Can it be characterised by NMR?
- How does σ_1 interact with client proteins (both *in cellulo* and *in vitro*), in particular the HSP70 family, and what are the structural changes that enable this?
- Can we validate the interactions observed through structural work with *in cellulo* fluorescence microscopy experiments?
- Coupling *in vitro* and *in cellulo* work can we address the functional significance of different oligomeric states of σ_1?

1.2 REGULATION OF HSP70 PROTEINS

Proteins of the HSP70 family are highly expressed and ubiquitous. Originally identified as heat shock inducible proteins in drosophila [70], they have subsequently become a textbook staple due to their essential role in protein synthesis and degradation. Acting as a workhorse of protein folding, the HSP70 proteins form a central hub through which many proteins in the cell pass, either during protein synthesis, as a result of misfolding and aggregation or as part of constitutive regulatory interactions [71–73].

Entry and exit to this protein folding hub is mediated through interactions with numerous regulatory proteins including multiple co-chaperones, chaperonins and other HSP proteins such as HSP90 [72]. This regulation allows the cell to react to the changing demands made of its proteome as well as to any damage induced by cellular stress. The HSP70 proteins themselves are comprised of a well conserved two domain architecture with an ATP hydrolysing Nucleotide Binding Domain (NBD) and a protein or peptide binding Substrate Binding Domain (SBD) [74]. The interplay between these two domains links ATP hydrolysis to substrate binding and release, forming the core mechanism by which these proteins assist protein folding [75].

A result of this cycle is that HSP70 proteins function as an assembly of moving parts and adopt substantially flexible and heterogenous conformations during their catalytic cycle [9]. The allosteric pathways that co-ordinate the movements of these parts have proved a fruitful subject of investigation by a combination of X-ray crystallography and solution state NMR, which have done much to expand our understanding of how these molecular machines function [9]. HSP70 substrate capture and ATP hydrolysis are coordinated by interactions with the HSP40 co-chaperones [76] and understanding the allosteric pathways formed in those complexes represents a natural next step in developing our understanding of the HSP70 family and provided the central motivation for our study of this system.

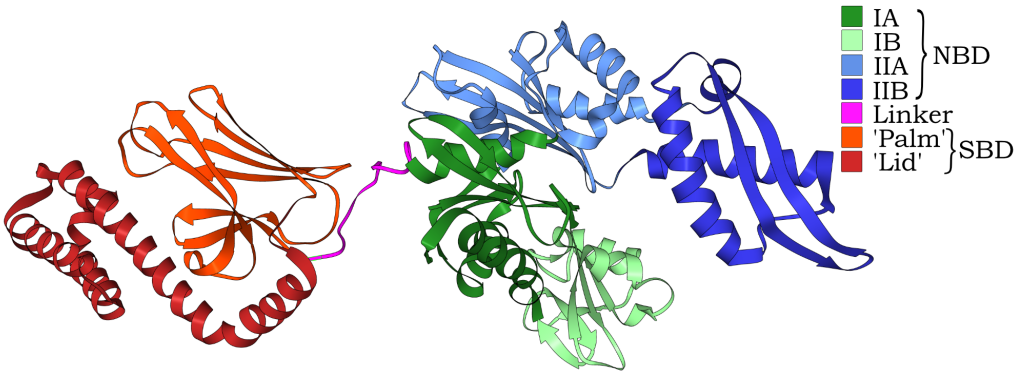


Figure 5: Structure of DNAK bound to ADP and Substrate (PDB: 2KHO) with the Nucleotide Binding (NBD) and Substrate Biding (SBD) and sub-domains highlighted. The inter-domain linker is shown in magenta. 598 residues of 638 are shown and almost all of the missing residues comprise a disordered C-terminal extension of ~33 residues.

1.2.1 *Anatomy of a HSP70 Protein*

At its most basic level the architecture of HSP70 proteins is dominated by two domains tethered to each other through a linker of around 15 residues (Figure 5). The NBD is the larger of the two domains comprising around 380 out of approximately 640 residues. It is composed of four distinct $\alpha\beta$ bundles (each one numbered either I or II) arranged into two lobes (subdomains A and B) [77]. The cleft between these two structures serves to bind ATP and provides the active site responsible for its hydrolysis [78]. The two lobes and their corresponding subdomains possess substantial flexibility allowing for internal motions to transmit changes in bound nucleotide [79].

The smaller SBD is comprised of two distinct structures a large β-sandwich that forms the substrate binding groove and an attached helical bundle that is referred to as the helical lid domain [80]. It is the disposition of these two subdomains which varies in response to the nucleotide bound state of the NBD. Aside from substrate binding, the SBD and its helical lid domain in particular, are implicated in the formation of oligomeric assemblies of HSP70 proteins [81–84]. The presence of these assemblies has been known for some time having been observed both in isolated proteins as well as in a cellular context [84–87].

The formation of oligomeric assemblies like almost everything else in HSP70 proteins appears to be linked to its nucleotide bound state. ADP binding has been observed to favour oligomers while ATP binding shifts the equilibrium towards monomers [86, 87]. The biological purpose of this oligomerisation appears to be regulatory in nature, with post translational modifications having been suggested to act as one mode of modulating this phenomenon [88]. Nonetheless the exact impact of these assemblies on the functional cycle of HSP70 is an area of active investigation.

1.2.2 *The HSP70 ATPase Cycle*

Substrate binding and release in HSP70 proteins is directly linked to their ATP hydrolysis cycle (Figure 6) [1]. This cycle depends on alternation between an ATP and an ADP bound state which possess distinct structural and mechanistic characteristics. When bound to ATP the protein adopts a docked conformation with its two domains, the NBD and the SBD, locked onto each other [89, 90]. This docked conformation results in low affinity (low micromolar range) and high rates of exchange for substrates.

Hydrolysis of ATP results in a switch to the ADP bound state where both the NBD and SBD dissociate, assuming a conformation where both domains act more independently of each other [91, 92]. This undocking releases the helical lid domain and allows it to associate directly with the substrate binding groove in the SBD, resulting in high affinity (low nanomolar range) and low rates of exchange for substrates.

Numerous observations of the behaviour of HSP70 proteins demonstrate the importance of allostery between the two domains. Not only does the nucleotide bound state alter substrate binding but substrate binding itself has been observed to increase hydrolysis of nucleotide by 2- to 10-fold [93, 94]. Recent research has begun to bring the pathways of communication between the two HSP domains into focus [9]. A central aspect of this is the role of the interdomain linker [95, 96].

This linker adopts a flexible conformation when ADP is bound allowing for the independent behaviour of the NBD and SBD. Binding of ATP however results in a rearrangement of the four subdomains of the NBD which opens a hydrophobic patch between sub domains IA and IIA. This patch is bound by the interdomain linker which becomes rigid and immobile as a result and aids in the process of docking the SBD onto the NBD [97, 98].

The basic model of HSP70 proteins alternating between an ATP-bound and ADP-bound conformation during their catalytic cycle has proved a powerful one and many observations have borne out the importance of these two states. Additional experiments however indicate that the two-state model is a simplification of a much more conformationally complex cycle. Even when bound to either nucleotide, NMR and FRET studies have shown that substantial conformational heterogeneity exists [99–101]. At best the ATP and ADP bound conformations seem to represent an average, or energy minimum, within a fairly large conformational ensemble. This inherent flexibility in the behaviour of these proteins makes them amenable to understanding by solution state NMR.

1.2.3 *Regulators of the HSP70 ATPase Cycle*

Although hydrolysis of ATP and hence the substrate binding cycle can occur with an isolated HSP70 protein, turnover is limited by both the slow rate of ATP hydrolysis and nucleotide exchange. Within the cell both these processes are subject to regulation through protein-protein interactions.

HSP40 co-chaperones are responsible for stimulating ATP hydrolysis [5, 76, 93], in some cases rendering the rate of nucleotide exchange limiting. Efficient operation is hence also dependent on interaction with a third protein, a Nucleotide Exchange Factor (NEF), which is responsible for stimulating exchange of ADP for ATP [102]. By alternating interactions with these two regulators HSP70 substantially speeds up its functional cycle [76], the complete canonical form of which is

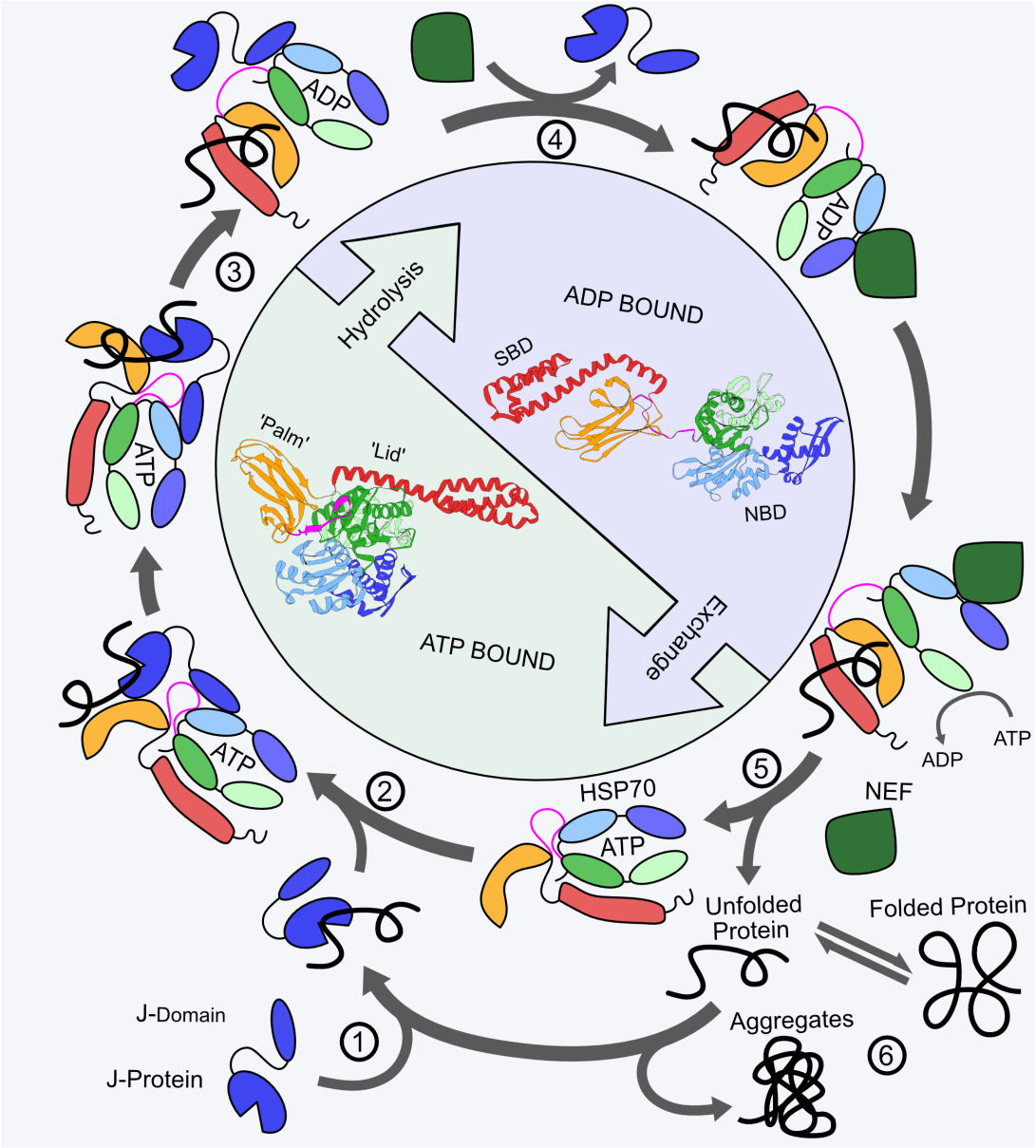


Figure 6: Schematic of the canonical ATP driven substrate binding and unbinding cycle of a HSP70 protein. **Center:** Structures shown are both for DNAK and represent the ADP bound (PDB: 2KHO) and ATP bound (PBD: 4B9Q) forms. Domains and sub domains are highlighted as in Figure 5. Hydrolysis of ATP causes a change in the conformation of the flexible linker (magenta) which docks onto subdomain IA of the NBD which in turn drives the docking of the two domains of HSP70. **1:** Substrate is picked up by a J-domain containing protein. **2:** J-domain protein binds its cognate HSP70 partner in the ATP bound form. Close juxtaposition of substrate with the low affinity high exchange rate form of the HSP70 SBD drives transfer of the substrate to HSP70. **3:** ATP hydrolysis is stimulated by the J-domain. NBD interactions couple substrate transfer to ATP hydrolysis and ensure efficient unidirectional transfer of substrate to HSP70. **4:** J-protein unbinds while the ADP and substrate bound HSP70 is targeted by an NEF (Nucleotide Exchange Factor). The NEF interacts with the NBD causing a rotation of the lobes of its subdomains which opens the nucleotide binding cavity stimulating exchange of ADP for ATP. **5:** Nucleotide exchange allows substrate and NEF to unbind. **6:** Substrate now has a second chance to attempt successful folding, and either folding or aggregation will remove the substrate from the cycle.

laid out in Figure 6. More importantly still, these interactions provide a means of regulating and targeting the activity of the HSP70 molecule [76].

The modularity of the HSP40-HSP70 interaction relies on a conserved interaction domain, called the J-domain [76, 103], and by co-opting this domain into a diverse range of proteins the cell is able to regulate the targeting of the HSP70 system in response to its requirements. This renders the HSP40 interaction the gate keeper to the HSP70 protein folding hub and makes it central to understanding the regulation of HSP70 proteins [76], a process which presents numerous opportunities for therapeutic intervention [4]. Its biological importance coupled to its suitability as a target for NMR studies drove our interest in this over the NEF interaction. Hence, while it is important in the context of HSP70 regulation, discussion of the NEF-HSP70 interaction here will be limited.

1.2.4 *The HSP40 Interaction*

The diverse HSP40 family is unified by its interactions with HSP70 proteins and their capacity to stimulate ATPase activity. Across the whole HSP40 family this is achieved through the J-domain, which adopts a small highly conserved four helical bundle fold (Figure 7A). The defining feature of this domain is the three-amino acid HPD motif found between helices II and III (Figure 7B) that is essential for the stimulatory capacity of all HSP40 proteins [76, 104, 105].

Understanding the ternary complex formed when a HSP70 protein binds to a HSP40 co-chaperone has to date proven difficult. This in part stems from the transient nature of the interaction [106–109]. Nonetheless several studies have converged on a number of key features of the interaction. These include a requirement for a coupled NBD and SBD [110], a localisation of J-domain binding to a region near the NBD-SBD interface and the formation of critical contacts between the HPD motif and an arginine residue in the NBD [104, 111].

There is however a general lack of consistency in proposed models for this interaction [112]. It is possible that some of this stems from an inherently dynamic

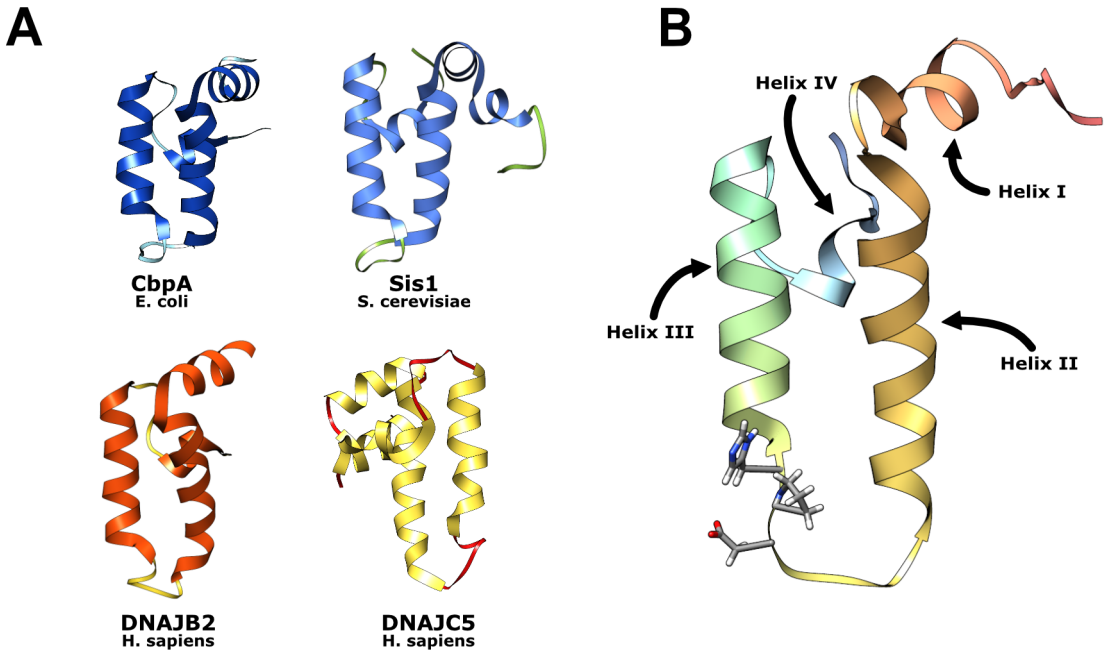


Figure 7: **A:** Structures of a range of J-domains from different proteins and organisms (PDB codes are as follows: CbpA: 2KQX, Sis1: 4RWU, DNAJB2: 2M6Y, and DNAJC5: 2N04) showing the strong conservation of J-domain architecture from bacteria through to humans. **B:** Annotated structure of the J-domain of *E. coli* protein DNAJ (PDB: 1BQZ), the founding member of the HSP40 family, showing the canonical 4 helix structure of the J-domain. The critical HPD motif which lies between helices II and III is highlighted with the side chains shown as stick models.

interaction. NMR work on the *E. coli* HSP70 DNAK was unable to resolve a static bound conformation suggesting instead that while DNAK is bound to ADP a dynamic complex is formed as the J-domain effectively scans the surface of DNAK [109]. There is also evidence that the interaction varies as a function of bound nucleotide [113].

The majority of NMR work has been done from the perspective of the J-domain and on simpler bacterial systems. Focusing on how a eukaryotic HSP70 responds to J-domain binding poses an attractive avenue of inquiry.

1.2.4.1 *The HSP70 Family*

The HSP70 family is one of the most highly conserved protein families: even between *E. coli* and humans the proteins have an almost identical length and the sequences are ~50% identical. Within humans several HSP70 variants exist and expression of 13 family members has been detected [114, 115]. Of the 13 expressed proteins, three seem to clearly derive from cellular compartmentalisation. BiP (HSPA5), GRP75 (HSPA9) and HSP13 (HSPA13) are resident in the ER [116], mitochondria [117], and microsomes [118], respectively. A further two cytosolic variants seem to be tissue specific: HS71L (HSPA1L)and HSP72 (HSPA2) are found in testis, HSP72 can also be found in brain tissue [114]. The more distantly related HS12A and B (HSPA12A and B) as well as HSP7E (HSPA14) proteins have a less clearly defined function and their sequences are the most diverged from the rest of the human HSP family suggesting a substantial alteration of function. HSP76 (HSPA6) and HSP77 (HSPA7) represent two seldom expressed variants with high homology to the stress induced HSP70 (HSPA1) proteins the latter corresponding to a pseudogene requiring bypass of a frame-shift mutation to express [119].

HSC70 (HSPA8) and the HSP70 (HSPA1) proteins are two of the most closely related and often co-expressed members of the human HSP70 family [120, 121]. HSC70 represents a constitutive and ubiquitous cytosolic HSP70 family member while the very closely related HSP70-1A and 1B (HSPA1A and B) proteins represent stress induced variants of HSC70. In the case of these proteins along with

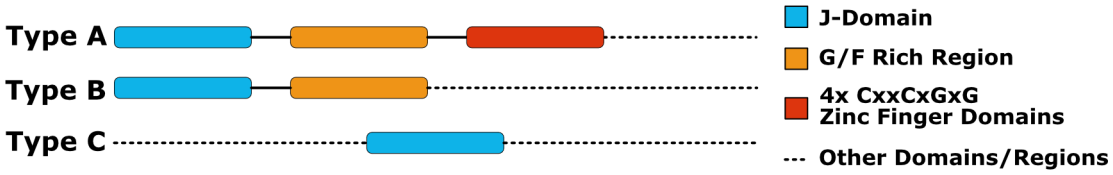


Figure 8: Schematic of the structures of the three different classes of HSP40 family proteins indicating characteristic domains. Dotted black lines show variable regions. The J-domain shown in blue comprises approximately 70 residues and is located on the N terminus of class A and B HSP40 proteins (~ residues 1-70) but can be located anywhere within the sequence of the C type HSP40 proteins. Both A and B class proteins contain a G/F rich region (which is lacking in C type proteins), shown in yellow, this can vary in length and generally starts immediately after the J-Domain. In A type proteins this is then followed by four CxxCxGxG repeats that comprise a zinc finger like domain. Across all three classes of HSP40 proteins the remaining regions show substantial variability in size, sequence and function with additional domains being inserted in line with the function of each particular HSP40 variant.

HS71L, HSP72, HSP76, and HSP77, expression will often occur at the same time and within the same cellular compartment [114]. The high identity and similarity between these proteins presents the question of why so many variants are needed.

1.2.4.2 *The HSP40 Family*

By contrast to the HSP70 Family the HSP40 family of proteins is enormous and highly divergent [114]. The family itself is divided into three sub families A, B and C according to the internal architecture of the proteins [76] (Figure 8). The A subfamily represents the most canonical branch of the HSP40s with members bearing a close resemblance to *E. coli* DNAJ. Their conserved architecture consists of an N-terminal J domain followed by a G/F rich region and a conserved cystine containing zinc finger region with a final variable C-terminal extension known to bind substrates. The B family follows a similar arrangement though it lacks the zinc finger domain and has greater variability in their C-termini. Finally, the C family is the most divergent and represents all HSP40 members that contain a J-domain but do not follow either the A or B architectures.

The requirement for many of these HSP40 family members is likely to be context specific, with many of them being expressed only under specific conditions. Many members have specific roles in diverse physiological processes ranging from

the cell cycle, through to development and the cellular stress response [76, 114]. Their interaction with HSP70 family members appears to act in most cases as a context specific substrate delivery system. The HSP70 family proteins themselves are able to bind a wide array of substrate peptides. Though they show a preference for largely hydrophobic peptides usually with a flanking charge [122, 123], they are nonetheless able to bind almost any presented peptide including polar peptides upon ATP hydrolysis [124]. As many HSP40 proteins possess specific substrate binding activity of their own, the stimulation of ATP hydrolysis by HSP40 J-domains is likely coupled to the juxtaposition of a particular type of substrate with the HSP70 SBD. By coupling spatial proximity of a particular substrate to ATP hydrolysis the HSP40s effectively ensure efficient substrate loading of the HSP70 machinery [125].

1.2.4.3 *HSP70-HSP40 Pairings*

Given the large number of HSP40 proteins and an array of HSP70 proteins which are co-expressed in the cytosol, questions regarding the specificity of these interactions arise. A number of experiments in which J-domains were swapped between different HSP40 proteins have shown that specificity can be localised to the J- domain [126]. In yeast, the HSP40 Ydj1 was shown to interact with the HSP70 proteins Ssa1p/2p but not with Ssb1/2p [127]. Interestingly the authors did not extend comparisons to Ssa3p/4p, the stress induced analogues of Ssa1p/2p.

Further work in yeast was also able to demonstrate a lack of interchangeability between compartment specific HSP70/40 pairings. In particular, HSP40 proteins from mitochondria or the cytosol were unable to regulate the ER-resident HSP70 BiP. Again, the J-domain was found to drive specificity with the interactions being recovered by substitution of three amino acids in the J domain that were found to be a common difference between cytosolic and ER resident HSP40s [128].

These studies in yeast demonstrate a clear divergence in pairings driven by compartmentalisation. However, the question of pairings in HSP70/40 proteins that share the same compartment is less well answered. The yeast studies simulta-

neously show the presence of J-domain driven specificity between the more distantly related Ssb and Ssa HSP70s and crosstalk at least between the closely related Ssa1p/2p proteins [127, 128]. Given biology's avoidance of needless redundancy and the presence of at least some sequence differences between even the most similar HSP70 proteins, the need for their existence and the biological roles of the differences between them is an intriguing question.

As in yeast, humans possess a number of often co-expressed and extremely closely related cytosolic HSP70 proteins [114, 115, 129]. Of particular interest are the distinctions between the closely related HSP70-1A and HS70-1B stress inducible proteins and the constitutively expressed HSC70. These proteins possess greater than 80% identity yet are expressed for distinct reasons with HSC70 acting as a house keeping chaperone primarily concerned with aiding folding during protein synthesis, while the HSP70-1A and HS70-1B proteins are concerned with protecting the cell from stress induced protein aggregation. HSP70-1A and HS70-1B are particularly closely related differing only in three positions.

The astounding similarity between HSC70 and HSP70-1A/1B (they are 85% identical and 93% similar) has led to them being commonly treated as interchangeable despite evidence that these proteins behave differently and favour different biological pathways [120]. For example, the two sets of proteins differ in basal hydrolysis rates as well as in their response to HSP40 mediated regulation. Furthermore, there is at least some evidence that the constitutively expressed HSC70 preferentially targets proteins for degradation while HSP70-1A/1B allows the same protein to refold [130]. While this observation has only been made in the specific context of hERG channel folding it highlights the capacity of subtle differences in the proteins to target distinct biological pathways. It is easy to rationalise why this may be the case, since under stressful circumstances it seems logical for the cell to prioritise the rescue of misfolding proteins. Whether it is a general property of the constitutive HSP machinery to dispose of proteins that mis-fold and of the stress induced HSP machinery to be more tolerant remains to be seen.

1.2.5 *Target Selection*

With the observations discussed above serving as a motivating factor we focused part of our study of the HSP40-HSP70 interaction on addressing questions of specificity and of what structural differences may underlie the interaction of distinct J-domains with the same and different HSP70 proteins. We chose the central cytosolic pair of HSP70 proteins HSC70 and HSP70-1A/1B for study because of their importance and poorly characterised specificity with regards to HSP40s despite being regulated in different ways.

Using somewhat similar logic we chose two closely related B family HSP40 proteins for comparison, DNAJB6 and DNAJB2. Both proteins represent stress induced variants of the HSP40 family which are abundantly expressed in the CNS particularly in the brain [114] and both possess an S/T and S rich stretch, respectively. In both cases the proteins have been associated with the management of poly(Q) protein aggregation [7, 8] though DNAJB6 displays a clearer pathological link [131]. DNAJB2 exists in two isoforms, one of which is nuclear and cytoplasmic while the other is membrane anchored and localises to the ER [132]. Both of these isoforms possess identical J-domains. Its exact cellular role remains to be completely delineated though recent work has highlighted the presence and role of its Ubiquitin Interacting Motifs (UIMs) in targeting misfolded proteins for degradation [7, 133] with the ER localised isoform being implicated in the neuronal ER-Associated Degradation (ERAD) pathway [133]. Like DNAJB2, DNAJB6 exists as two isoforms differing in cellular localisation with the 'a' isoform containing a nuclear localisation signal which confines it to the nucleus while the 'b' isoform lacks this and is also found in the cytoplasm [134]. In a similar fashion to DNAJB2, the exact cellular role of DNAJB6 has not been pinned down. Its importance in numerous pathological process, in particular those associated with ploy(Q) protein aggregation has focused studies on its role during pathological events [134–138], though its high levels of endogenous expression suggest it may also play a housekeeping role in neuronal cells [114]. Recent work has shown several interesting

properties of DNAJB6 including its propensity to form polydisperse oligomeric assemblies [137] and the importance of its S/T rich region in binding substrate and attenuating poly(Q) aggregation [131, 139, 140].

To provide an example with greater contrast we also chose DNAJC5 from the C subfamily of HSP40 proteins. Again, following our neurological focus carried over from work on σ_1, DNAJC5 represents a target which is closely tied to the CNS. It is a distinctive 20 kDa J-protein easily identifiable through the presence of a contiguous stretch of ~12 cystine residues, hence its cognomen Cystine String Protein (CSP) [141, 142]. Like DNAJB6 and DNAJB2 it has been found to possess neuroprotective properties and has been implicated in various neurodegenerative pathologies [6]. In contrast to both DNAJB6 and DNAJB2 however, DNAJC5 is much better characterised having been directly shown to interact with and stimulate HSC70 [143], and having a more clearly delineated role in partnership with HSC70 regulating synaptic vesicle fusion [106]. Its list of characterised client proteins includes a number of proteins involved in vesicular fusion such as VAMP and syntaxin [6], with recent work highlighting SNAP-25 as its primary target [144, 145]. Its client profile, cell biological role and its interaction with HSC70 suggest a primarily constitutive role for DNAJC5 in contrast to the cell-stress related roles of DNAJB2 and DNAJB6.

Adding further weight to their selection, all three targets also presented interesting proteins from a structural perspective with structures available only for the J-domains of DNAJC5 [141] and DNAJB2 [146] and no structural information available for DNAJB6. The selection of these three proteins hence not only provides three interesting biological and structural targets but provides three proteins which co-exist within the neuronal ecosystem and play distinct biological roles. The co-expression of all three in similar tissues and cells and their contrasting roles from housekeeping through to mediating responses to cellular stress makes the question of their specificity and targeting of HSC70 and HSP70 apposite.

1.2.6 *Studying HSP40-HSP70 Interactions*

The power of NMR in studying the interactions of co-chaperones combined with the emerging view of σ_1 as a co-chaperone/chaperone served both to direct our interest towards the field and as the final impetus for the change in direction of our work. With our selected target proteins, HSC70 (HSPA8), HSP70 (HSPA1A/B), DNAJB2, DNAJB6 and DNAJC5 we set out to recapitulate the regulatory interactions of the HSP70 proteins with their HSP40 partners in vitro. The lack of information on the HSP70-HSP40 complex from the perspective of the HSP70 protein prompted us to become interested in this aspect of their interaction. Due to our prior work with HSC70 and the recently published assignments for its NBD [147], we chose to focus on HSC70 with regards to structural studies by NMR. Throughout the work present here we hoped to address several key questions:

- What are the structural changes through which J-domain binding is communicated throughout the structure of HSC70?
- What are differences, if any, in J-domain specificity between HSC70 and HSP70?
- What are the structural correlates behind any observed differences between these two proteins?
- Can the functional properties of different HSP40 J-domains be related to their structural effects on HSC70?

2

MATERIALS & METHODS

2.1 PROTEIN EXPRESSION AND PURIFICATION

2.1.1 *Expression and Purification of* σ_1

An ACA free gene construct (GeneArt) was ordered and used to clone N-terminal TwinStrep-GFP-TEV(ENLYFQG)-σ_1 full-length R7QC94A in a pCold I (Takara) background and was transfected into C43(DE3) *E. coli* cells. A selection of colonies was picked and inoculated into either LB (unlabelled growths) or unlabelled M9 (labelled growths) media. Inoculated media was grown overnight at 37 °C and used to inoculate either TB or M9 media. Flasks were grown to an O.D at 600 nm of ~0.8 in M9 media, or ~1.2 in TB, at 37 °C prior to being cold shocked on Ice for 15 minutes. Unlabelled growths were subsequently induced with IPTG to a final concentration of 1 mM and allowed to express overnight at 15 °C.

Labelled growths were pelleted by centrifugation and washed with 1x M9 salts prior to 10x condensation and resuspension in labelled M9 media labelled with ISOGRO (Sigma-Aldrich) according to the manufacturer's instructions. Condensed cultures were allowed to grow a further hour at 15 °C prior to induction with 2.5 mM IPTG and overnight growth.

Biomass was harvested by centrifugation and cells were resuspended into lysis buffer (HEPES 50 mM pH 8, 300 mM NaCl, EDTA 1 mM, 0.03% Azide) and lysed using a cell disruptor. Cell membranes were then harvested by ultracentrifugation for 1 hour at 234,998 x g (45000 rpm) in a Type 45 Ti rotor (Beckman Coulter). Membranes were resuspended by homogenisation into lysis buffer, supplemented

with detergent, either 1% DPC or 1% DDM and 1% Triton X-100 and incubated overnight for extraction.

Extracted mixtures were pelleted as before by ultracentrifugation to remove debris and loaded onto a Strep-Tactin® (IBA) column. Columns were washed four times with 25 ml of buffer containing 0.5% detergent (either DPC or DDM without Triton X-100) and a further four times with 25 ml of buffer with 0.25% detergent. Protein was eluted into buffer with 0.05% detergent using 5 mM d-desthiobiotin. Eluted protein was concentrated and run down a Superdex S200 Increase 10/300 column (GE) in NMR buffer (20 mM HEPES pH 6.5 or 7.2, 20 mM NaCl, 0.03% Azide and 0.03% Detergent).

2.1.2 *Expression and Purification of HSP70 Constructs*

Template DNAs for HSC70 (pcDNA5/FRT/TO HSPA8) and HSP70 (pcDNA5/FRT/TO GFP HSPA1A) were obtained as a kind gift from professor Harm Kampinga [114] (Addgene plasmids #19483 and #19460), template DNA for BiP was graciously provided by Herwig Schüler [148]. All HSP70 family proteins and constructs follow the same purification scheme. pPSG-IBA 105 (IBA Lifesciences) plasmids containing the desired construct with a leading TEV sequence (ENLYFQG) were cloned and transformed into BL21 DE3 *E. coli* cells (New England Biolabs). A selection of colonies was then picked, inoculated into LB, grown overnight at 30 °C, and then used to inoculate either TB (unlabelled growths) or LB (labelled Growths). Cultures were grown to an O.D at 600 nm of ~0.8, LB or ~1.2 TB. Unlabelled growths were then induced with IPTG to a final concentration of 1 mM and left to express overnight at 37 °C.

Labelled growths were pelleted by centrifugation and washed twice with 1x M9 salts prior to 4x condensation and resuspension in labelled media. Condensed cultures were allowed to grow 15 minutes at 37 °C prior to induction with 2 mM IPTG and overnight expression.

Biomass was harvested by centrifugation and cells were resuspended into lysis buffer (HEPES 50 mM pH 8, 300 mM NaCl, 1 mM EDTA, 0.03% Azide, 5 mM β-Me) and lysed using a cell disruptor. Lysates were clarified by ultracentrifugation for 1 hour at 234,998 x g (45000 rpm) in a Type 45 Ti rotor (Beckman Coulter) and filtering through glass wool. Clarified lysates were then loaded onto Strep-Tactin® (IBA) columns and washed four times with 25 ml of lysis buffer prior to elution with lysis buffer supplemented with 5 mM d-desthiobiotin. Eluted protein was concentrated and run down a Superdex S200 Increase 10/300 column (GE) into NMR buffer (20 mM HEPES p.H 7.2, 50 mM NaCl, 2 mM $MgCl^{2+}$, 1 mM TCEP, 0.03% Azide).

2.1.3 *Expression and Purification of HSP40 Constructs*

Template DNAs for DNAJB6a (pcDNA5/FRT/TO GFP DNAJB6a), DNAJB6b (pcDNA5/FRT/TO DNAJB6b), DNAJB2a (pcDNA5/FRT/TO GFP DNAJB2a) and DNAJB2b (pcDNA5/FRT/TO GFP DNAJB2b) were obtained as a kind gift from professor Harm Kampinga [114] (Addgene plasmids #19502, #19474, #19496 and #19497). The template DNA for DNAJC5 was obtained as a synthetic gene (GeneArt). Expression and purification of HSP40 J-domains followed the same scheme as HSP70 constructs with an added cleavage step after the affinity column. Eluted protein was supplemented with 4 ml of 1 $mg\,ml^{-1}$ TEV per 3 litres of biomass and dialysed overnight at 4 °C in TEV cleavage buffer (20 mM HEPEs, pH 7.5, 300 mM Nacl, 5 mM DTT, 5% Glycerol). Cleaved protein was then passed through a Ni affinity column to remove His-tagged TEV. Eluted cleaved protein was then concentrated and run through SEC as described for the HSP70 constructs.

2.2 BIOPHYSICAL TECHNIQUES

2.2.1 *Chaperone Assays*

Chaperone assays were carried out by mixing Bovine Insulin Solution (Sigma-Aldrich) with either pH 6.5 NMR buffer (20 mM HEPES pH 7.2, 20 mM NaCl, 0.03% Azide and 0.03% Detergent) or pH 6.5 NMR buffer and σ_1 to give a solution with a final concentration of 100 µM insulin in a final volume of 100 µl. Aggregation was triggered through the addition of 10 mM TCEP and was monitored by following the O.D of the sample at 360 nm using a Cary 50 UV-Vis Spectrophotometer (Agilent) with an averaging time of 100 ms for 25 minutes.

2.2.2 *Photo-Crosslinking*

Photo-crosslinking experiments were performed using a final concentration of ~13 µM protein and 1.25 $mg\,ml^{-1}$ Sulfo-SANPAH (ThermoFIsher) in a 20 µl reaction volume. All samples were crosslinked in NMR buffer (20 mM HEPES, pH 6.5). Samples were allowed to react in the dark at 4 °C on ice for an hour. The lysine conjugation reaction was quenched through the addition of 1 µl of 1 M Tris. Samples were then kept on ice and given three rounds of UV exposure at 999.9 $mJ\,cm^{-2}$ per round on a Spectrolinker XL-1500 for protein crosslinking (Spectronics Corporation). Crosslinked samples were then run on an SDS-PAGE gel for analysis.

2.2.3 *SEC-MALS*

SEC-MALS experiments for σ_1 were carried out in NMR buffer(20 mM HEPES pH 6.5, 20 mM NaCl, 0.03% Detergent) supplement with either, DDM, DPC or LMNG. Peak fractions of σ_1 from preparative SEC were taken and pooled and loaded onto a Superdex S200 Increase 10/300 column. Light scattering was detected at 662 nm at eight different scattering angles ranging between 20.6° and 149.1° using a Heleos

8 Instrument (Wyatt Technology). Data was analysed using ASTRA 6.1 software (Wyatt Technology) using protein conjugate analysis in order to calculate total mass, detergent mass and protein mass. Published values of dn/dc of detergents were used. For HSC70 an identical proceedure was used though samples were run in NMR buffer at pH 7.2 without detergent (20 mM HEPES pH 7.2, 50 mM NaCl, 2 mM $MgCl^{2+}$, 1 mM TCEP, 0.03% Azide) and standard molar mass calculations from light scattering were used during data analysis.

2.2.4 *Microscale Thermophoresis*

Fluorescently labelled σ_1 was prepared using the Alexa Fluor^{TM} 594 Microscale Protein Labelling Kit (Invitrogen) according to the manufacturer's instructions. Micro Scale Thermophoresis (MST) experiments were carried out on a Monolith NT.115 Instrument (Nano Temper Technologies). Experiments were performed using 50 – 200 nM labelled protein in pH 7.2 NMR buffer (20 mM HEPES p.H 7.2, 20 mM NaCl) with protein concentration and LED laser power adjusted to achieve a signal of 500-1500 counts. In all cases temperature was set to 25 °C. Initial experiments were carried out using a series of MST powers, 20, 50 and 80% with the power for subsequent experiments being set to maximise signal. Analysis of data and fitting was performed using OriginPro 2017 (OriginLab Corporation). For the peptide binding assays NR1 peptides RHKDARRKQMQLAFAAVNVWR and WRKNLQDRKSGRAEPDPKKKATFRAI were commercially synthesised (thinkpeptides)

2.2.5 *Circular Dichroism*

All spectra were collected using a J-815 circular dichroism spectropolarimeter (JASCO). Scanning was performed from 190 to 250 nm with 10 accumulations per experiment. σ_1 samples were prepared either in buffer with DPC:DPPC to a q of 0.1 (σ_1-Δ35) or in DDM (Full-length σ_1). Spectra for full-length σ_1 were calculated us-

ing the DichroCalc [149, 150] web service based on the published crystal structure (PDB: 5HK1).

2.2.6 *Radio-Ligand Binding*

Radioligand assays were performed based on previously reported methods used for σ_1 [17, 63]. Briefly, glass microfiber filters grade GF/B (Whatman®) were soaked in 1% polyethylenimine (PEI) for at least 1 hour at room temperature prior to the experiment and washed with buffer immediately before use. Purified σ_1 receptor in NMR buffer at either pH 6.5 or 7.2 at a concentration of 2-10 µM in 100 µl was incubated with 0.2 – 100 nM ^{3}H-(+)-pentazocine with a specific activity of 32 $Ci\,mmol^{-1}$(Biotrend Germany) for 1 hour at 30 °C. Samples for calculating background were prepared as above but supplemented with 1 µM Haloperidol or 10 µM NE-100. Incubated samples were applied to washed filters using a vacuum manifold, the filters were then washed a further 3x with 400 µl of buffer and ^{3}H signal was measured by liquid scintillation counting.

2.2.7 *ATPase Assays*

ATPase assays were carried out using a commercially available High Throughput Colorimetric ATPase Assay Kit (Expedeon) according to the manufacturer's instructions with the following modification: quantities of all reagents used were halved (100 µl Reactions, 25 µl PiColorLock™ and 10 µl of stabiliser). All assays were performed in NMR Buffer (20 mM HEPES pH 7.2, 50 mM NaCl, 2 mM MgCl2+, 1 mM TCEP, 0.03% Azide) using 2 µM HSC/HSP70 with or without 25 µM of J-domain, background was measured in reactions without protein. Reactions were carried out in 96-well plates (Corning®) and initiated through the addition of ATP to a final concentration of 1 mM prior to incubation at 30 °C for 1 hour, samples were then treated according to the kit instructions. The amount

of inorganic phosphate released was measured spectrophotometrically at 600 nm using a SpectraMax M3 micro plate reader (Molecular Devices).

2.3 NMR

All NMR experiments were carried out on either a 500 MHz (^{1}H) Bruker Advance II, a 600 MHz (^{1}H) Bruker Advance III HD, or a 750 MHz (^{1}H) Bruker Advance III HD spectrometer equipped with TCI cryoprobes, or on a 950 MHz (^{1}H) Bruker Advance III HD spectrometer equipped with a room temperature probe as indicated. All data was processed using NMRPipe [151] and analysed using NMRFx Viewer [152] (previously NMRViewJ, One Moon Scientific) and OriginPro 2017 (OriginLab Corporation). Where required reconstruction of Non-Uniformly Sampled (NUS) spectra was performed using Iterative Soft Thresholding (IST) with software provided by the Gerhard Wagner Lab [153].

2.3.1 *TRACT*

Experiments were performed at 600 MHz (^{1}H) using a published ^{1}H ^{15}N-TRACT sequence [154] previously implemented in our laboratory. Data was collected at 41 °C on ^{2}H/^{15}N/^{13}C labelled full-length σ_1 in 20 mM potassium phosphate at pH 6.5 supplemented with either 50 mM DPC and 5 mM DPPC or 10 mM DDM. Data processing and fitting was performed using a combination of NMRPipe [151] and scripts developed in house by Dr. Jacob Brady.

2.3.2 σ_1

Backbone ^{1}H ^{15}N spectra of σ_1-Δ35 were collected in 20 mM potassium phosphate at pH 6.5 supplemented with 50 mM DPC and 5 mM DPPC at 600 MHz (^{1}H) using a BEST-TROSY pulse sequence recorded at 41 °C. Backbone ^{1}H ^{15}N spectra of ^{2}H/^{13}C/^{15}N-labelled full- length σ_1 were collected in 20 mM potassium phos-

phate at pH 6.5 supplemented with either 50 mM DPC and 5 mM DPPC or 10 mM DDM at 750 MHz (^{1}H) using a BEST-TROSY pulse sequence recorded at 41 °C.

2.3.3 σ_1-*HSC70 Titrations*

Both sets of titration spectra were collected in 20 mM HEPES pH 7.2, 50 mM NaCl in 0.03% DDM at 750 MHz (^{1}H) . Titration of unlabelled HSC 70-NBD (residues 1-385) into a 100 µM ^{15}N/^{13}C-labelled full-length σ_1 was performed at 37 °C. NBD was added to molar ratios of 0, 0.5, 1 and 2 each point was monitored using both a ^{1}H ^{15}N best-TROSY experiment to monitor back bone resonances and a sensitivity enhanced ^{1}H ^{13}C HSQC to follow side chain methyl resonances with a total acquisition time per titration point of 5 hours. Titration of unlabelled full-length σ_1 into a 200 µM ^{15}N/^{13}C-labelled HSC70 was performed as above with the following modifications: experiments were carried out at 25 °C at molar ratios of 0, 0.25 and 0.5 and the total acquisition time per titration point was reduced to 2 hours 15 minutes.

2.3.4 *TALOS*

Predictions of σ_1-Δ35 secondary structure based on secondary shifts were made using TALOS-N [155] based on backbone ^{1}HN ^{15}N ^{13}C $^{13}C_\alpha$ and sidechain $^{13}C_\beta$ chemical shifts. Resonance assignments for the prediction were made by Dr. Jose Ortega-Roldan.

2.3.5 *HSC70*

All HSC70 NMR experiments were carried out in 20 mM HEPES pH 7.2, 50 mM NaCl, 1 mM TCEP, 0.03% azide. Where stated, samples were supplemented with 2 mM $MgCl^{2+}$ and ADP. ^{1}H ^{15}N backbone spectra of full-length HSC70 were col-

lected using a BEST-TROSY sequence on either a ^{15}N-labelled sample at 600 MHz (^{1}H) and 35 °C or on a ^{2}H/^{15}N/ILV-labelled sample supplemented with ADP and $MgCl^{2+}$ at 950 MHz MHz (^{1}H) and 30 °C using 50% NUS. ^{1}H ^{15}N backbone spectra of the NBDSBD construct of HSC70 (residues 1-555) were collected using a BEST-TROSY pulse sequence at 750 MHz (^{1}H) and 30 °C. Test spectra looking at ^{1}H ^{15}N backbone resonances were collected at 600 MHz (^{1}H) for the SBDCTD (residues 386-646) and at 500 MHz (^{1}H) for the CTD (residues 556-646).

Backbone ^{1}H ^{15}N spectra of the NBD of HSC70 alone (residues 1-385) were collected on a ^{15}N/^{13}C-labelled sample with a BEST-TROSY sequence at 750 MHz (^{1}H) and 30 °C. Using this sample an additional set of three TROSY based triple resonance experiments collected with 10% NUS: HNCO HNCA and a HN(CO)CA to help with the transfer of assignments based on published data.

Backbone ^{1}H ^{15}N spectra of the core of the SBD of HSC70 (residues 386-555) were collected on a ^{15}N/^{13}C and a ^{15}N-labelled sample using a best-HSQC sequence. The ^{15}N labelled sample was further used to collect a standard ^{15}N edited TOCSY and two standard ^{15}N edited NOESY experiments with 75 and 100 ms mixing times for assignment, all three experiments were collected with 33% NUS. The ^{15}N ^{13}C sample was further used to collect a set of five TROSY based triple resonance experiments for assignment using 25% NUS: a HNCO, HNCA, HNCACB, HN(CO)CA and a HN(CO)CACB. All assignment experiments on the SBD were performed on at 750 MHz (^{1}H) and 25 °C. Assignments for ^{1}HN ^{15}N ^{13}C $^{13}C_{\alpha}$ backbone and side-chain $^{13}C_{\beta}$ were made using data from triple resonance experiments.

2.3.5.1 *HSC70 Titrations*

Titration of unlabelled DNAJC5 J-domain (Residues 1-100) was performed both into a 120 µM ^{15}N-labelled HSC70 NBDSBD sample and into a 210 µM ^{2}H/^{15}N/methyl-labelled sample of full-length HSC70. Both samples were supplemented with 2 mM $MgCl^{2+}$ and ADP. For the former sample, backbone ^{1}H ^{15}N resonances were monitored using BEST-TROSY experiments collected at molar ratios of 0, 0.5, 1

and 2 at 750 MHz (^{1}H) and 30 °C. Total acquisition time for each point was 1 hour 30 minutes. For the latter sample, backbone ^{1}H ^{15}N resonances were monitored using 50% NUS BEST-TROSY experiments and methyl resonances were followed using 50% NUS carbon SOFAST experiments at molar ratios of 0, 0.75, 1.5 and 2.5 at 950 MHz (^{1}H) and 30 °C. Total acquisition time for each point was 5 hours 30 minutes.

2.3.6 *DNAJC5*

Spectra for DNAJC5 were collected in 20 mM HEPES pH 7.2, 50 mM NaCl buffer. Backbone ^{1}H ^{15}N spectra of ^{15}N-labelled DNAJC5 J-domain (Residues 1-100) were collected using a BEST-HSQC sequence at 750 MHz (^{1}H) and 25 °C. Assignments were transferred from previously published data based on backbone ^{1}H ^{15}N resonances alone. Titration of unlabelled full-length HSC70 was performed into a 210 µM ^{15}N-labelled DNAJC5 sample supplemented with 2 mM $MgCl^{2+}$ and ADP. Backbone ^{1}H ^{15}N resonances were monitored using a BEST-TROSY pulse sequence at molar ratios of 0, 0.5, 1 and 2 at 750 MHz (^{1}H) and 25 °C. Total acquisition time for each point was 1 hour 30 minutes.

2.3.7 *DNAJB6*

All spectra for DNAJB6 were collected in 20 mM HEPES pH 7.2, 50 mM NaCl buffer. Backbone ^{1}H ^{15}N spectra were collected at 25 °C using a BEST- HSQC sequence for ^{15}N-labelled DNAJB6 1-77, 1-100 and 1-130. In the case of 1-130 spectra were also collected on a ^{15}N/^{13}C-labelled sample. ^{15}N edited NOESY (100 ms mixing time) and TOCSY (45 ms mixing time) pairs were collected on a ^{15}N labelled DNAJB6 1-130 sample at 750 MHz (^{1}H) a further 130 ms mixing time NOESY was also collected at 600 MHz (^{1}H), these experiments were recorded at 25 °C using 33% NUS. A standard suite of five TROSY based triple resonance spectra were collected on the ^{15}N/^{13}C-labelled 1-130 DNAJB6 sample at 600 MHz

(^{1}H) and 25 °C using 25% NUS: HNCO, HNCA, HNCACB, HN(CO)CA and a HN(CO)CACB. Triple resonance, NOESY and TOCSY experiments were used for assignment of DNAJB6B 1-130 ^{1}HN ^{15}N ^{13}C $^{13}C_\alpha$ backbone and side-chain $^{13}C_\beta$ resonances. Titration of unlabelled full-length HSC70 was performed into a 250 µM ^{15}N-labelled sample of DNAJB6 1-77 supplemented with 2 mM $MgCl^{2+}$ and ADP. Backbone ^{1}H ^{15}N resonances were monitored using a best-TROSY pulse sequence at molar ratios of 0, 0.5, 1 and 2 at 750 MHz (^{1}H) and 25 °C. Total acquisition time for each point was 1 hour 30 minutes.

2.4 CELL CULTURE

Cos-7 and TZM cells were grown in Dulbecco's Modified Eagle Medium (DMEM; Life Technologies), SH-SY5Y cells were grown in a 50:50 mixture of DMEM and Ham's F-12 Nutrient Mixture (Thermo Fisher). In both cases media was supplemented with 10% Fetal Bovine Calf Serum (FBS), 1% penicillin and streptomycin, and 1% L-Glutamine. Cells were maintained at 37 °C in the presence of 5% CO_2 and were passaged upon reaching ~80% confluence.

2.5 CONSTRUCTS FOR MICROSCOPY

mVenus tagged human σ_1 in a mammalian expression plasmid was a kind gift of Dr. Radu Aricescu and was used as a template for all microscopy constructs. KCNA5 and KCNA4 template DNAs were obtained as synthetic genes (GeneArt) and HSC70 and BiP DNAS were obtained as outlined in subsection 2.1.2. Fluorescent protein fusions were made using these constructs as required.

2.6 MICROSCOPY SAMPLE PREPARATION

Cells were plated two days prior to imaging onto 10 mm glass bottom Poly-d-lysine coated dishes (MatTek Corporation) at a total cell count of $\sim 2 \times 10^5$ cells

and ~24 hours after plating cells were transfected with the desired constructs using GeneJuice transfection reagent (Novagen) according to the manufacturer's instructions. Transfected cells were imaged the next day. Prior to imaging cells were washed with PBS (pH 7.4) and switched into DMEM media without phenol red and supplemented with 25 mM HEPES. Staining of cells with ER-Tracker Blue-White DPX (Life Technologies) was performed as per manufacturer's instructions.

2.6.1 *Imaging*

Images were produced either using a ZEISS ELYRA TIRF microscope (Zeiss) or on an SP8-X-SMD confocal microscope (Leica Microsystems). In both cases cells were imaged at 37 °C focused on a 512x512 region using either a 100x/1.46 (zoom/aperture) in the case of the former or a 63x/1.4 in the case of the latter oil immersion objective. Excitation of mClover was achieved at 488 nm using either a white light laser set to 488 nm on the SP8-X-SMD or a 488 nm 100 mW laser on the ZEISS ELYRA TIRF. Images of cells stained with ER- Tracker Blue-White (ThermoFisher) were acquired as above on a ZEISS ELYRA TIRF microscope with excitation using a 405 nm 50 mW diode laser. Cells for these experiments were simultaneously imaged in both green and blue channels to correlate σ_1-mClover staining and ER-Tracker Blue-White staining. Imaging on the SP8-X-SMD was performed using its internal hybrid detectors while imaging on the ZEISS ELYRA TIRF was performed using an iXon DU 897 EM-CCD camera (Andor).

2.6.2 *FRET-FLIM*

FRET-FLIM experiments were performed on an SP8-X-SMD confocal microscope (Leica Microsystems) using a 63x/1.4 oil immersion objective. Excitation of mClover was achieved using a white light laser set at 488 nm wavelength tuned to 80 MHz. Excitation of TFP was achieved using a 440 nm picosecond-pulsed diode laser (PDL 800-B, PicoQuant) tuned to 40 MHz. In both cases photon counting was

performed using a Time-Correlated Single Photon Counting (TCSPC) system operated by a PicoHarp 300 module (PicoQuant) with external hybrid detectors set to photon counting mode. 300-500 photons were acquired per pixel and data was analysed and fitted to a mono-exponential model using SymPhoTime (PicoQuant).

2.6.3 *Number and Brightness*

Number and brightness experiments were performed by collecting a series of 500 128 x 128-pixel images (NE-100 and Pentazocine experiments) or 200 x 200-pixel images (Cocaine Experiments) on an SP8-X- SMD confocal microscope (Leica Microsystems). Excitation of mClover was performed by a white light laser set at 488 nm wavelength tuned to 80 MHz, detection was performed at 500 – 550 nm using internal hybrid detectors set to single-photon counting mode. Acquisition settings were adjusted to produce a pixel dwell time of 9.75 µs (NE-100 and pentazocine experiments) or 4.88 µs (cocaine Experiments). Image series were collected for a selection of cells on each plate before and after the addition of ligands, 20 µM NE-100 or pentazocine and 200 µM cocaine. Imaging after ligand addition was performed after a 30-minute incubation. All experiments were performed on live COS-7 cells which remained at 37 °C throughout. Data analysis was performed using Image J and R Studio with the nandb R package produced by Rory Nolan [156].

Part II

THE σ_1 RECEPTOR

3

BIOPHYSICAL STUDIES OF σ_1

3.1 EXPRESSING AND PURIFYING σ_1

Expression of all constructs of σ_1 was constrained by the need to produce isotopically labelled proteins (^{15}N and ^{13}C) required for NMR studies. This inevitably favoured expression in *E. coli* over other systems (mammalian, insect or yeast cells) in which, although labelling is possible, it is both much harder to achieve and, in the case of low yields, likely to be prohibitively expensive. Initial construct design was guided by the two-TM topology of the receptor derived from previous studies (discussed in more detail in subsection 1.1.6). Based on this, earlier work on σ_1 took a divide and conquer approach seeking to produce viable truncation constructs which could be assigned and used as stepping stones on the road to characterising the full-length protein [16, 59].

3.1.1 *Prior approaches to Expressing* σ_1

Two truncation constructs served as the basis for the purification strategy initially employed with full-length protein. The first of these constructs was thought to comprise the entirety of the membrane associated portion of the receptor and is referred to as the Chaperone Domain (σ_1-CD). This encompasses residues 112-223 [59]. The second of these constructs was thought to contain one of the two putative transmembrane helices and lacked only the first 35 residues of the protein now know to contain its only transmembrane helix, this construct termed σ_1-Δ35 [16] contains residues 36-223.

Expression of σ_1 was achieved by relying on the pCold system which brings expression under the additional control of a cold shock activated promoter. Induction is hence achieved through a combination of both cold shock and addition of IPTG. Use of the pCold system proved particularly useful, not only as a tight regulator of protein expression, but also due to the high levels of culture condensation tolerated by this system (up to ten-fold). The use of highly condensed cultures proved key in generating usable yields of isotopically labelled protein despite poor expression.

Earlier work on σ_1-Δ35 carried out during my master's degree highlighted the importance of removing a cysteine at position 94. This mutation had been previously shown not to affect σ_1 function [65]. Exchange of this cysteine for alanine prevented substantial protein aggregation, improved refolding and prevented the formation of artificial cross-linked dimers. As such all further work on σ_1-Δ35 and longer constructs maintained this C94A background.

Once expressed, prior work with truncation constructs (both my own prior to beginning my PhD and work carried out by Dr Jose Ortega-Roldan) and initial studies with the full-length protein (carried out during my PhD) relied on the use of a denaturing protein preparation. Extraction from harvested membranes post-lysis was achieved using 6 M guanadine hydrochloride. His-tagged constructs were then pulled out of the unfolded membrane fraction using IMAC chromatography. Protein was then cleaned up with reverse phase HPLC with the final product of purification being pure lyophilised protein powder.

Refolding of truncation constructs was previously achieved by solubilisation of thin-filmed detergent protein mixtures into 8 M urea. This mixture was subsequently either dialysed or buffer exchanged using SEC into NMR buffer. Extensive screening at the refolding stage during purification of σ_1-CD showed effective solubilisation only in DPC [59]. Given its favourable properties for NMR and its ability to extract all tested σ_1 constructs, DPC was initially carried forward into purification of full length σ_1.

3.1.2 *Developing a Native Purification Strategy*

Purification of full-length σ_1 using a denaturing approach failed at the final refolding step. Several attempts were made to try and solubilise the lyophilised full-length σ_1. Neither on-column refolding nor refolding using dialysis proved successful. Alteration of the buffer conditions at this step, including variation of detergent mixtures and detergent concentrations was also to no avail. The protein remained unable to refold even in high concentrations of the relatively harsh detergent DPC, which was able to solubilise the σ_1-Δ35 construct. This drastic change in detergent preference between σ_1-Δ35 and full-length was somewhat surprising at the time but, in the context of the structure of σ_1, can now be rationalised in that the construct does not contain a transmembrane domain as was expected, whereas full-length does.

Unable to refold full length σ_1, our approach shifted to extracting the protein in a folded state directly from membranes. Efforts focused on two detergents, although others were also tried. The first of these, DPC, was chosen due to its success at solubilising prior σ_1-Δ35 constructs and its favourable properties for NMR. The second of these, DDM, was chosen due to published work demonstrating its ability to extract functional σ_1 expressed in *E. coli* [17, 29].

Extraction of the protein was achieved by incubation of isolated *E. coli* membranes overnight in either 1% DPC or 1% DDM and Triton X-100. Extraction in DPC tended to be more efficient than in DDM (data not shown). Post extraction, His-tagged proteins were isolated with IMAC followed by SEC to clean up the samples. The yields remained in the low microgram range (10 litre cultures) using either detergent.

The first 8 residues of σ_1 contain a twin-arginine motif [157] (residues 7 and 8) that may interfere with processing of the protein in *E. coli*; consequently, both mutation of arginine 7 to Q and truncation of the first eight residues (σ_1-Δ8) were tested for expression. Modest increases in yields were observed for both constructs.

The less disruptive R7Q construct was hence carried forward in further expression tests with the full-length protein.

The low final yields were compounded by poor performance at the IMAC step likely due to binding affinity and specificity being hindered by high detergent concentrations. Higher affinity tags were explored to circumvent this. Initially an increase in tag affinity was achieved by extension of the His-tag from 6 residues to 9. Further increases in performance at this stage as well as improvements in expression were achieved by switching to a Twin-Strep tagged construct. Finally, addition of a cleavable EGFP tag to the Twin-Strep tagged σ_1 construct further increased yields particularly when extracting with DDM.

In both purifications with DDM and DPC, EGFP fluorescence was preserved under extraction and subsequent purification conditions; and coupled with the successful insertion of the protein into membranes, was consistent with correct protein folding. The finalised purification scheme, shown in Figure 9, with EGFP fused receptor comprised four steps: an initial antibody affinity step followed by TEV cleavage, reverse affinity, and SEC. Use of this scheme allowed the expression deadlock to be broken, yielding sufficient protein for further biophysical assays and allowing the continuation of the NMR work started with prior truncation constructs of σ_1.

3.1.3 *Testing the Functionality of Full Length* σ_1

Having successfully established a methodology for producing both labelled and unlabelled σ_1 in *E. coli* further work on the protein hinged on demonstrating that the protein so produced was functional. To do so we aimed to re-capitulate *in vitro* previously established properties of σ_1. Based on the published literature four such properties were deemed to be readily testable:

1. High affinity binding to archetypal σ_1 ligands
2. Chaperone activity

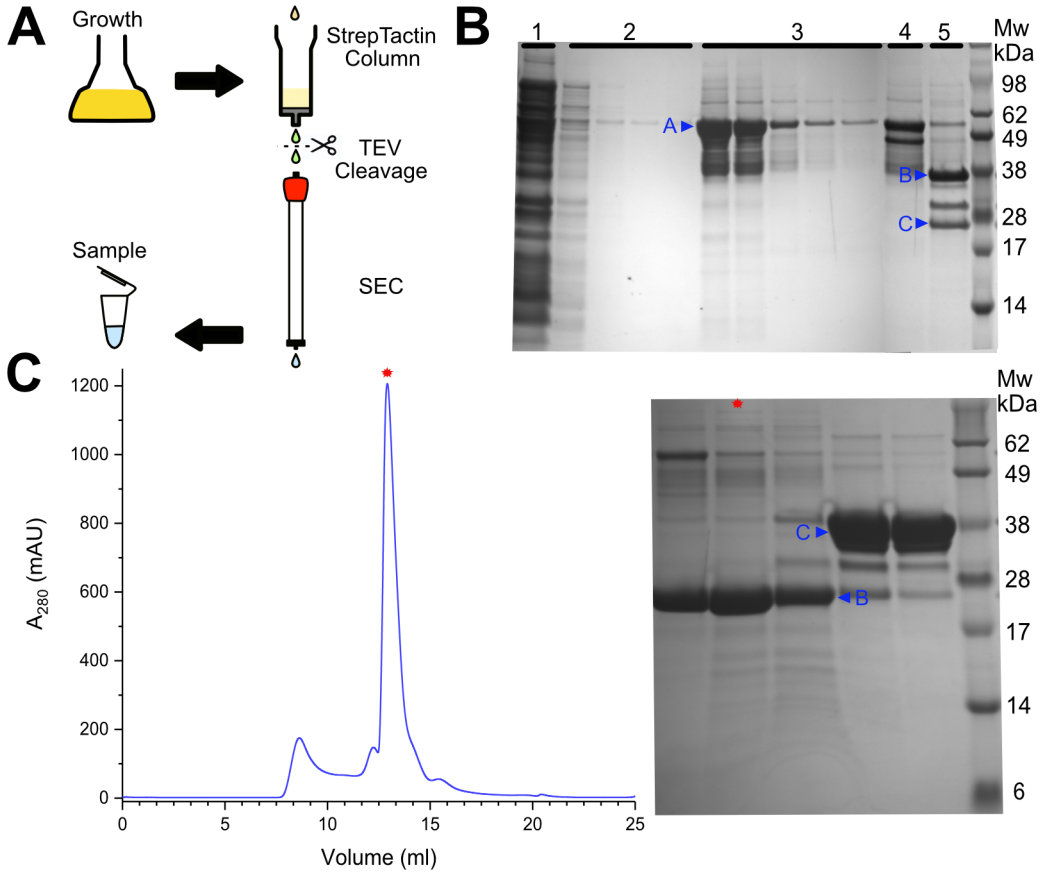


Figure 9: Summary of purification scheme for full length σ_1 using a Twin-Strep-GFP-R7QC94Aσ_1 construct. **A:** Schematic of steps in the protocol. After growth, cell lysates are run through a Strep-Tactin column, elutions are cleaved with TEV at 4 °C overnight, concentrated and run through SEC using a Superdex S200 Increase 10/300 column. Peak fractions are then used to make samples as required. **B:** Gels showing progress of purification through the affinity column and cleavage steps. Going from left to right, labelled wells show: *Lane 1:* Flow through from Strep-Tactin column, *Lane 2:* Wash steps, *Lane 3:* Elution of σ_1 with 5 mM d-desthiobiotin, *Lane 4:* pooled elutions prior to TEV cleavage, *Lane 5:* cleaved protein after incubation with TEV overnight at 4 °C. Labelled bands are as follows: *Band A:* GFP-σ_1, *Band B:* σ_1, *Band C:* GFP **C:** *Left:* SEC chromatogram of cleaved protein starred peak corresponds to σ_1 with a right hand tail/shoulder containing free GFP. *Right:* Gel of the main peak from the SEC chromatogram with wells proceeding from left to right across the peak. Bands are labelled as for B.

3. Protein oligomerisation

4. Any of the many protein-protein interactions previously reported

3.2 DPC

Due to its favourable NMR properties, namely its small dynamic micelle [158] and its successful use in NMR characterisation of σ_1 truncation constructs, initial work on the full-length protein continued using DPC. Use of DPC during the early stages of work was also favoured due to its more complete extraction of σ_1 from bacterial membranes, which helped alleviate problems with yields. A number of attempts were thus made to establish the functionality of natively purified full-length protein in DPC.

3.2.1 *Oligomeric State*

With an increasing number of papers reporting σ_1 oligomers [17, 62, 64] and the hypothesis that these oligomers are required for ligand binding [17], establishing the oligomeric state of our samples became a central concern. Earlier work on σ_1-Δ35 had shown that cystine crosslinked dimers eluted at ~12.7 ml while monomers eluted at ~14.4 ml. Elution of full length σ_1 occurs at ~13.3 ml which is larger than the smaller σ_1-Δ35 species but smaller than the cystine crosslinked species (Figure 10B). The ~1 ml shift in retention volume relative to σ_1-Δ35 can be explained by a combination of increased size, since full-length is 35 residues longer, and additional detergent associated with the added TM. Coupled with the reduced size relative to σ_1-Δ35 dimers this shift does not seem consistent with a dimeric receptor.

In addition to inferences made from the SEC profile of the protein we also employed NMR TRACT measurements to get some idea of the protein size in solution. For samples of full-length σ_1 a size estimate of ~50 kDa was obtained (Figure 10A).

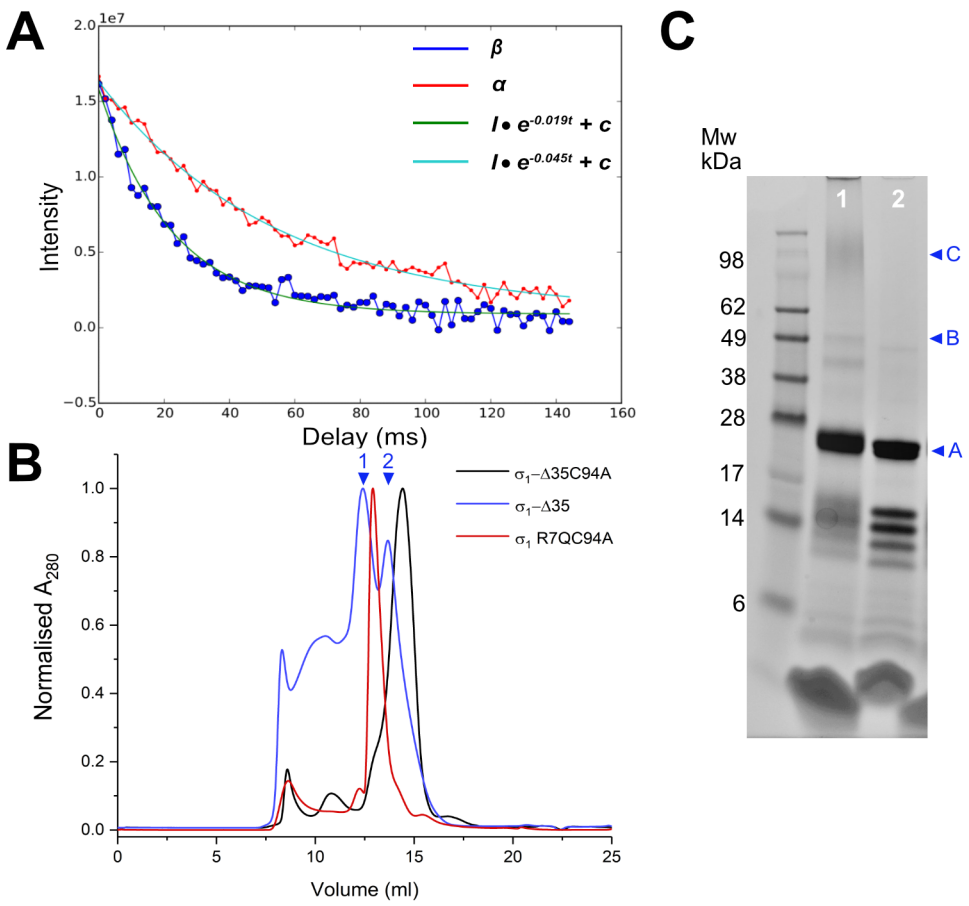


Figure 10: A summary of data on the oligomeric state of σ_1 in DPC. **A:** NMR TRACT data for full-length σ_1 in DPC collected at 600 MHz (^{1}H) and 41 °C. Fitting of the data gives a τ_c of 12.3 ns with a corresponding size estimate of 46 – 56 kDa. **B:** Comparison of SEC chromatograms for different σ_1 constructs. For σ_1-Δ35, peaks labelled 1 and 2 correspond to cystine crosslinked dimer and monomer, respectively. **C:** Photo-crosslinking gel for σ_1-Δ8 in DPC (at a q of 0.1 with DPPC). *Lane 1:* Crosslinked, and *Lane 2:* Un-crosslinked. *Band A:* Monomer, *Band B:* Dimer, and *Band C:* Higher order assemblies.

This estimate is indirect as TRACT measures rotational correlation times (τ_c), not mass, hence the mass values rely on modelling the relationship between τ_c and mass [154]. Uncertainties in formulating this model translate into uncertainties in mass. Additionally, TRACT experiments are unable to separate contributions to τ_c from the surrounding detergent micelle making accurate size estimates of a membrane proteins problematic. Nonetheless a rough estimate can be made using the measured size of a DPC micelle alone (~21 kDa [159]) this gives an approximate protein mass of 30 kDa which is more consistent with a monomer than a dimer.

Additional characterisation of the oligomeric state of the protein was attempted through crosslinking. Previously developed UV-crosslinking methodologies on σ_1-Δ35 were able to demonstrate that this truncation construct has some propensity to self-associate despite being predominantly monomeric. Application of this technique to σ_1-Δ8 samples showed that full length protein maintains a slight propensity to self-associate with crosslinking generating a faint dimer band as well as a smear above 98 kDa (Figure 10C). Compared to σ_1-Δ35, laddering of σ_1-Δ8 was almost completely abolished, and the majority of the protein remained in a monomer band.

All of the initial data therefore supported a monomeric state of σ_1 samples in DPC. In particular, the TRACT gives little room for anything larger than a monomer while crosslinking results in a much weaker self-association than that observed for σ_1-Δ35, a construct known to be monomeric.

3.2.2 *NMR Characterisation of Secondary Structure*

Previous work on truncation constructs indicated that these constructs in DPC were not completely unfolded. This came both from CD spectra with a strong helical signal as well as from TALOS predictions of secondary structure from NMR secondary chemical shifts. Thanks to work by Dr. Jose Ortega-Roldan, the longest of the truncation constructs σ_1-Δ35 was almost completely assigned [16, 59] which allowed secondary structure predictions for most of σ_1 to be made.

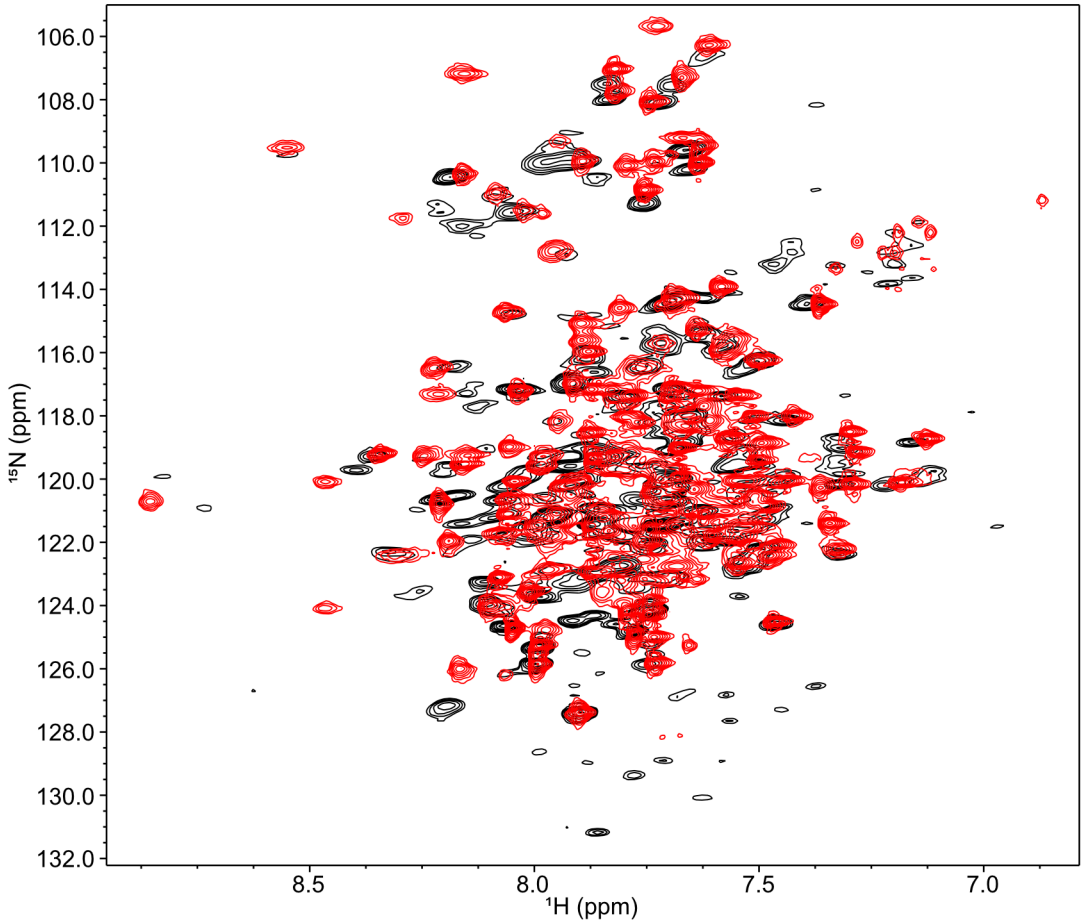


Figure 11: Overlay of ^{1}H ^{15}N TROSY for σ_1-Δ35 (Red) collected on a 600 MHz spectrometer and a ^{1}H ^{15}N BEST-TROSY of full-length σ_1 (Black) collected on a 750 MHz (^{1}H) spectrometer. Both spectra were collected at 41 °C in 50 mM DPC and 5 mM DPPC.

Initial HSQC spectra of the full-length protein were promising showing several new peaks as expected from a longer construct. With the majority of peaks from full-length σ_1 overlaying closely with spectra of σ_1-Δ35 further work by Dr. Jose Ortega-Roldan allowed for the transfer of assignments between the two constructs (Figure 11 and Figure 12). Hence, addition of the first TM of σ_1 does not seem to have a major impact on the structure of the rest of the receptor. Based on this and on secondary structure predictions for σ_1-Δ35, the evidence seemed to argue that in DPC the full-length receptor was folded (Figure 13).

Secondary structure predictions gave a mostly α-helical structure for the receptor, with the σ_1-Δ35 construct containing eight predicted α-helices. Based on this

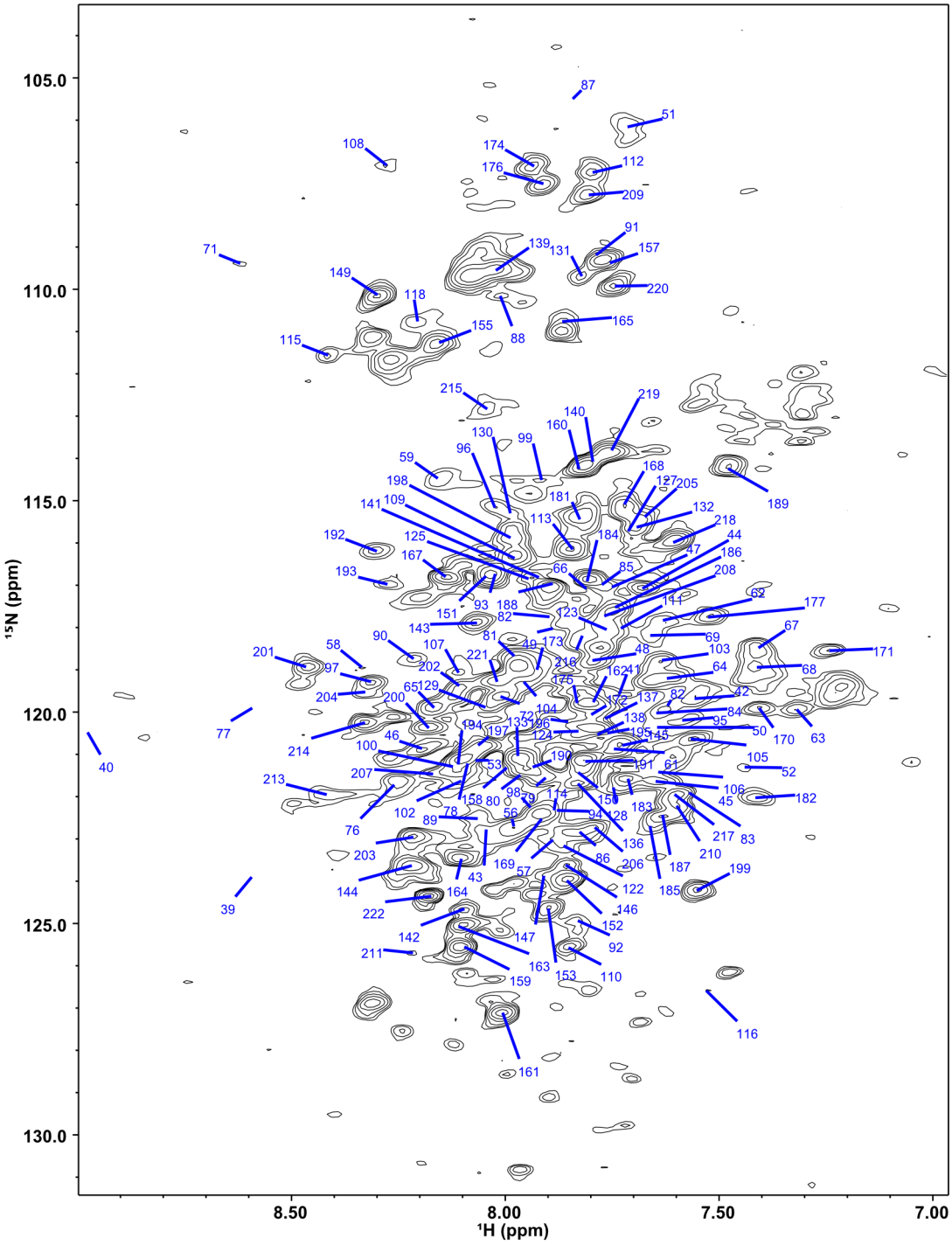


Figure 12: ^{1}H ^{15}N BEST-TROSY HSQC of full-length σ_1 at 41 °C in 50 mM DPC and 5 mM DPPC collected on a 750 MHz (^{1}H) spectrometer with backbone amide assignments shown. Assignments were made by Dr. Jose Ortega-Roldan.

data a helix predicted between residues 91 and 107 and covering most of the SBDL I region of the receptor was interpreted as the second transmembrane domain of σ_1 [16]. Publication of the crystallographic structure of σ_1 revealed a structure completely at odds with this result.

In addition, almost the entirety of one face of the σ_1 β-barrel was predicted to be α-helical based on chemical shifts in DPC. Strands surrounding this sheet were also poorly predicted showing up as loops or disordered instead (Figure 13). Correct secondary structure prediction occurred mainly in two segments of the receptor. The first of these are the two helices forming the hydrophobic membrane anchoring face of the β-barrel. The longer of these two helices was correctly delineated and identified by our earlier work as being amphipathic and probably attached to the membrane based on NOEs to detergent [59]. The second region of accurate prediction consists of the two side-facing helices that pack against the β-barrel.

Another point of congruence between structural predictions in DPC and crystallographic work lies at the bottom of the β-barrel. The latest structures of σ_1 suggest that this region is flexible and acts as a lid regulating access to the ligand binging cavity of the receptor [18]. This flexibility is born out in the NMR data both for σ_1 in DDM and DPC which shows this region to be dynamic, as well as in TALOS predictions in DPC which show a high degree of disorder for these residues.

The overall failure of secondary structure prediction is somewhat hard to interpret. Two separate factors are likely to be at play. Firstly, there could be a failure of the TALOS based prediction method for this particular fold. Secondly, it could be that the bacterially expressed receptor in DPC is aberrantly folded and the secondary structure adopted by the protein in our NMR samples is different from that in membranes or lipidic cubic phase.

Differentiating between these two possibilities is not trivial and cannot be conclusively answered with the data collected. The unusual fold of σ_1 may have contributed to inaccuracies in TALOS which relies on previously solved structures to make predictions. TALOS is not the only prediction method to have trouble with

σ_1, in particular the region incorrectly identified as TM2 was robustly predicted as a TM by numerous other TM helix prediction and secondary structure prediction software [16].

On the other hand, that σ_1 in DPC samples possess a disrupted structure is supported by the high degree of disorder seen in σ_1 samples in DPC: Hydrogen-deuterium (HD) exchange experiments carried out by Dr. Nader Amin failed to observe significant protection of backbone amide protons.

Detergents are known to be able to induce changes in protein secondary structure in certain cases. In the case of DPC, numerous examples have surfaced recently which demonstrate its tendency to disrupt both structure and function [158]. On balance, it seems likely that that the poverty of DPC as a membrane mimetic along with an inability of *E. coli* to properly fold the protein are the major contributing factors to a disrupted structure which is reflected in the secondary structure predicted in our samples.

3.2.3 *Functional Assays*

High affinity ligand binding is one of the best tests for protein functionality; as such, this was addressed from the outset of our work on full-length σ_1. Prior work looking at drug binding in σ_1-Δ35 in DPC had been unable to detect high affinity binding by either NMR, ITC or MST. In the case of the latter two techniques signals suggesting low affinity binding were probably due to interference from drug-detergent interactions.

With the difficulties experienced due to the presence of DPC in other work neither MST nor ITC were considered viable methods to test ligand binding. Seeking to generate more conclusive data we turned to radio-ligand binding experiments, which having been well established in the field, represent the gold standard for measuring drug binding to σ_1. In DPC samples of either σ_1-Δ35 or full length we were unable to detect any high affinity binding. As radio-ligand binding is limited

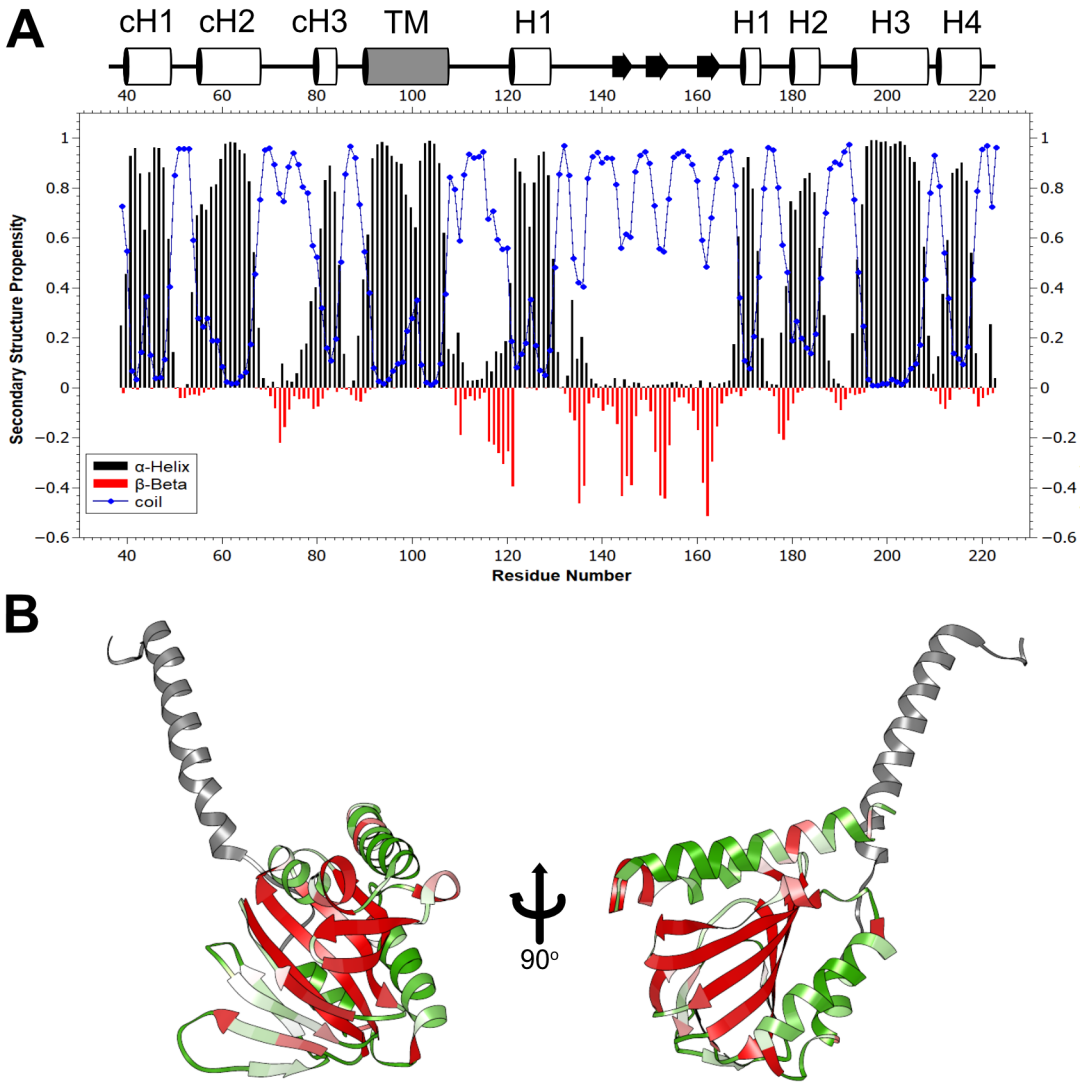


Figure 13: Comparison of predicted secondary structure from NMR on σ_1-Δ35 in DPC with structure solved by X-Ray crystallography. **A:** TALOS-N prediction [155]. Red bars below zero show β-sheet propensity, black bars above zero show α-helical propensity, and the blue line shows propensity for unstructured loop regions. *Top:* Schematic of predicted secondary structure. **B:** Heat map of prediction accuracy shown on σ_1 structure (PDB: 5HK1). Red shows regions where prediction is incorrect while green shows areas where accurate prediction has been made.

in the range of testable concentrations the experiments were unable to rule out a lower affinity interaction.

Additional attempts at measuring functionality were also made through chaperone assays aimed at measuring the attenuation of protein aggregation by σ_1. Such assays however proved impossible as even at low concentrations of DPC the strength of the detergent effectively prevented aggregation.

3.2.4 *On the Functional State of σ_1 in DPC*

The lack of high affinity ligand binding in DPC does not stand alone; the data on oligomerisation and the published link between oligomerisation and ligand binding as well as the failure of NMR secondary structure prediction provide strong support for a non-functional receptor. Though some observations suggest that the structure of σ_1 in DPC retains certain characteristics of its native fold it it was considered insufficient to justify further biophysical work.

3.3 DDM

With samples in DPC not providing positive functional corroboration, we had already started to shift towards preparing samples in an alternative detergent, even prior to publication of the structure. Literature reports of functional *E. coli* expressed σ_1 extracted in DDM [17, 29] served as the basis for selecting DDM as a DPC alternative. This was also supported by cursory detergent screening which showed little extraction in alternative detergents such as DM and OG.

With the transition to GFP fused receptor and the expression deadlock completely broken, we were able to produce substantial quantities of σ_1. This allowed for significant yields even with the poorer extraction observed in DDM. With these samples in hand corroborating the functionality of the protein, once again, became the primary goal.

3.3.1 *Ligand Binding*

In addition to NMR, two different biophysical techniques were explored: Isothermal Titration Calorimetry (ITC) and Microscale Thermophoresis (MST). ITC continued to be hampered by drug–detergent interactions and showed no clear high affinity binding. MST largely followed this trend as well, with observed unsaturable signal at high (millimolar) concentrations of ligand. This likely reflects changes in the bulk properties of the solution as ligand concentration enters the low millimolar range.

Doubts over the clarity of the negative results seen by ITC and MST were raised by the observation of Chemical Shift Perturbations (CSPs) during drug titrations in DDM, carried out in conjunction with Dr. Jose Ortega-Roldan. Shifts in carbonyl spectra of the protein were in slow exchange on the NMR time scale (~ms) which could indicate high affinity binding. The much larger apparent size of the protein was also encouraging, though this increase in size could be explained, at least in part, due to a larger less mobile micelle; it could also be the result of oligomerisation which was suspected to be necessary for high affinity binding.

There were reasons to be cautious: though encouraging, the larger size also resulted in poor coverage of the protein hampering a clear rationalisation of the observed CSPs. These difficulties were compounded by earlier experiences with σ_1-Δ35 showing that at concentrations starting in the mid micromolar range, hydrophobic molecules could cause spectral shifts in σ_1. The persistence of theses shifts, even when clearly known non-binders such as naloxone were used, argues against the observation reflecting native-like ligand binding.

Given the capacity of these hydrophobic molecules to interact with detergent one possible explanation for observed shifts could lie in the perturbation of the detergent micelle around σ_1 at higher drug concentrations. Equally the possibility of a non-native interaction driven by non-specific sticking to exposed hydrophobic regions on a poorly folded protein provided yet another possible explanation for

observed shifts. With difficulties in saturating observed shifts persisting even in DDM, the NMR observations were unable to stand on their own.

Given the uncertainty surrounding our attempts to measure ligand binding through alternative biophysical techniques it seemed pertinent to return to the gold standard established in the field, radio-ligand binding. Using tritiated pentazocine several experiments were carried out to try and establish binding with protein produced in DDM with sample in DPC serving as a non-functional control. The experiments showed only background radioactive counts, indicating that at least for the high affinity range testable by radioligand binding no interaction could be detected.

3.3.2 *Monitoring Oligomerisation*

As with samples in DPC, initial estimates of protein size were obtained using NMR TRACT experiments. In the case of samples in DDM a size of around 70 kDa was obtained (Figure 14B). The increase in size relative to protein in DPC agreed well with the decrease in spectral quality. Given this estimate however it proved difficult to ascertain the size of the protein itself since DDM micelles without protein are reported as being ~70 kDa in size [160]. Here, the approximation of the behaviour of the protein-detergent micelle as an additive mixture of the two contributions clearly breaks down.

Application of photo-crosslinking to samples in DDM yielded more promising results (Figure 14A). Unlike samples of σ_1-Δ8 in DPC, full-length σ_1 in DDM had an increased propensity to form oligomers when compared to σ_1-Δ35. The monomer band could be seen to almost completely disappear contrasting with prior crosslinking experiments where even though higher bands appeared, the monomer band remained dominant. Although photo crosslinking does reflect some tendency of a protein to self-associate it does not provide a direct readout of the predominant state of the protein in solution, as transient and strong interactions cannot be differentiated.

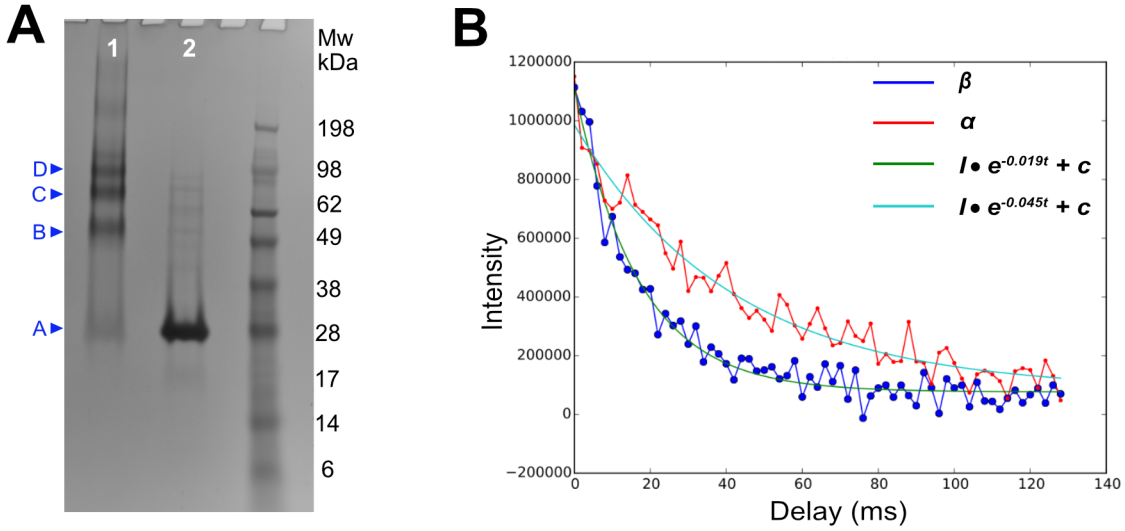


Figure 14: Crosslinking and NMR TRACT data for full-length σ_1 in DDM. **A:** Photo-crosslinking gel of full-length σ_1 in DDM. Bands corresponding to monomers, dimers, trimers, and tetramers are labelled A, B, C and D, respectively. *Lane 1:* Crosslinked, *Lane 2:* Un-crosslinked. **B:** NMR TRACT data for full-length σ_1 in DDM collected at 600 MHz (^{1}H) and 41 °C. Fitting of the data gives a τ_c of 16.8 ns with a corresponding size estimate of 66 – 78 kDa.

Accurately determining the oligomeric state of σ_1 in our samples hinged both on getting a direct readout of the size of the predominant species in solution as well as on being able to separate detergent and protein contributions to size. Two techniques provided the ability to achieve this, SEC coupled to Multi Angle Laser Scattering (SEC-MALS) and Native Mass Spectrometry (Native MS).

Both SEC-MALS and Native MS measurements required tweaking of detergent concentrations in samples. Initial measurements suffered from high detergent and low protein concentration. This prevented dissociation of the protein and detergent in Native MS experiments and caused scattering by the protein to be swamped by contributions from the adjacent free micelle peak in SEC-MALS.

Better protein expression as well as steps taken to suppress detergent concentration with a view to improving outcomes in chaperone assays (discussed below in subsection 3.3.3) alleviated these difficulties. Final mass estimates for the protein in DDM gave a size of ~22 kDa corresponding to a monomer (Figure 15A). This is further corroborated through Native MS measurements showing only monomeric protein with a mass of 25251.5 Da (Figure 15B).

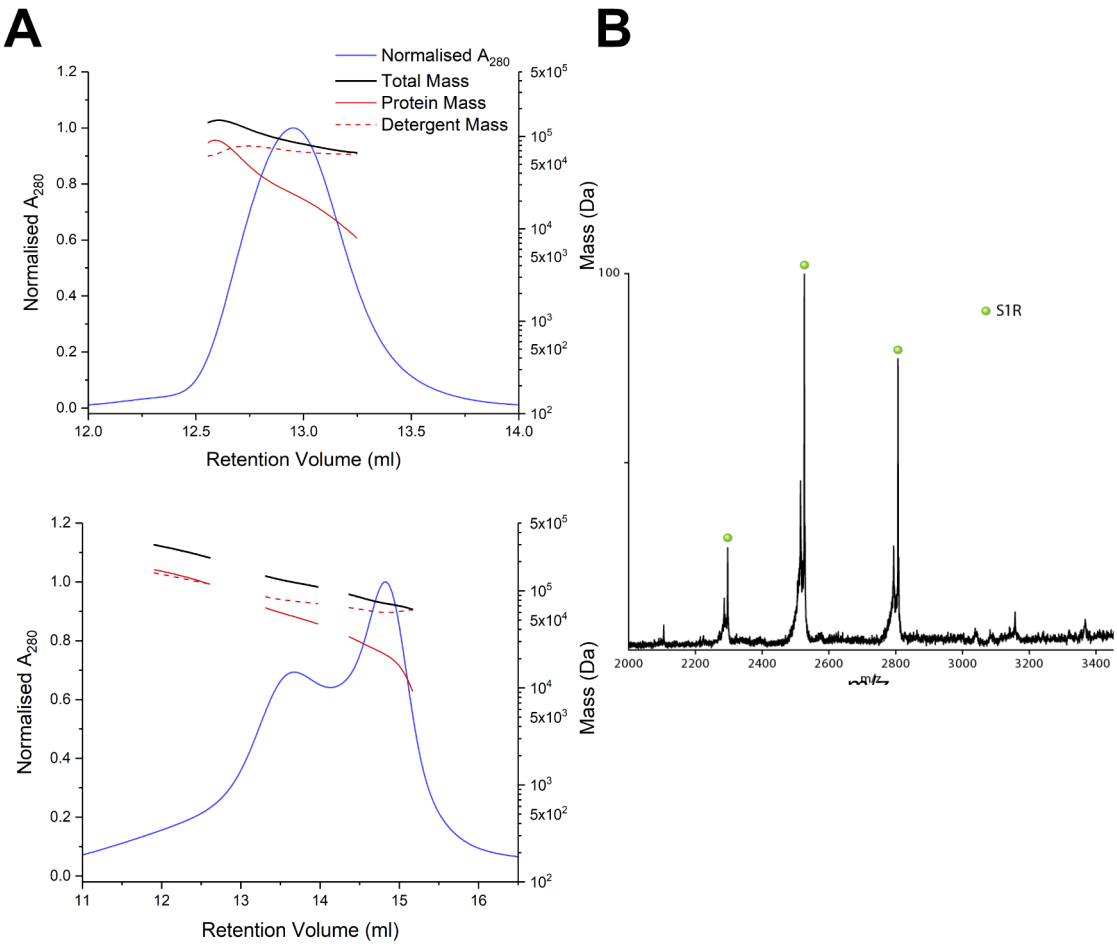


Figure 15: SEC-MALS and native mass spectrometry (MS) data for full-length σ_1 **A:** SEC-MALS protein conjugate analysis of full-length σ_1 wild type (top) and E172A (bottom) in DDM. In the case of E172A, mass calculations are shown for both peaks and the trailing shoulder. Both wild type and E172A show a monomer peak with a calculated protein mass of ~22 kDa, the second peak for E172A yields a mass of ~50 kDa corresponding to a dimer while the shoulder yields a mass >100kDa **B:** Native MS spectrum of full-length σ_1 in DDM. The fitted mass is 25251.5 Da and the expected mass is 25252.76 Da.

Though full-length σ_1 samples in DDM are unequivocally monomeric at least one observation of an oligomeric species was possible. The drug binding knockout mutant E172A [61] generated to help validate attempts to follow ligand binding showed distinct shouldering in its SEC profile. Further SEC-MALS measurements confirmed this shoulder as stemming from an ~50 kDa species most likely a dimer (Figure 15A).

3.3.3 *Chaperone Activity*

Previously reported chaperone activity for a σ_1 construct comprised of residues 116-223 [14] provided an additional means for testing the functionality of σ_1 in tandem with drug binding experiments. A straight forward assay hinging on monitoring the aggregation of insulin in the presence of a reducing agent [14, 161] was chosen to try and monitor chaperone activity in σ_1. Aggregation could be followed across a time course through measurement of the OD of the sample at 360 nm.

Despite earlier failures at reproducing this assay in DPC we were able to produce the required aggregation of insulin in the presence of DDM. A number of steps had to be taken to overcome inhibition of aggregation by DDM. DDM concentrations were reduced below 0.1%, the insulin concentration was raised to 100 µM and 10 mM of the very strong reducing agent TCEP was used.

Initial trials with σ_1 samples appeared to show a strong anti-aggregation effect even at very low molar ratios of σ_1. Given the anti-aggregation properties of detergent and in order to control for this we measured the concentration of detergent in our samples using NMR. NMR measurements showed that even subsequent to SEC in low detergent concentrations (0.03% DDM), the concentration of DDM in fractions containing σ_1 remained well above the target concentration contained in the buffer, going as high as ~0.3% (Figure 16B).

Two approaches were taken in tandem to try and address this issue, the first was to establish the degree of sensitivity of the insulin assay to DDM concentrations, the second was to try to reduce DDM concentrations during purification as extensively as possible. Reduction of the DDM concentration at the affinity step below 1% proved viable within certain limits.

High concentrations of DDM were maintained during column loading and initial washing in order to avoid precipitation of extracted proteins. Once σ_1 was isolated on the column subsequent washes could further reduce detergent concentrations to around 0.2%; below which, precipitation started to become an issue as a result of the increasing protein to DDM ratio when concentrating samples for

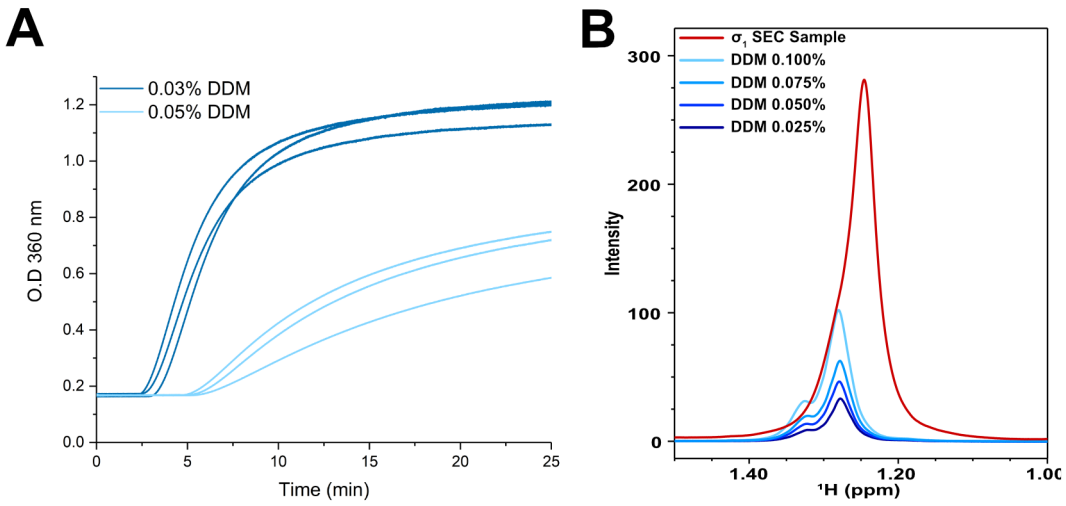


Figure 16: Effects of DDM on the insulin aggregation assay for chaperone activity. **A:** Increased attenuation of insulin aggregation on going from 0.03% to 0.05% DDM. Three technical repeats are shown at each concentration. Assays were carried out in HEPES pH 6.5 buffer with 20 mM NaCl. **B:** 1D ^{1}H NMR spectrum collected on samples in assay buffer at 600 MHz (^{1}H) and 25 °C. Blue curves show buffer standards at different concentrations of DDM while the red curve shows the DDM concentration in the peak fraction of σ_1 from SEC. The upfield shift observed is a result of protein-detergent interactions. A rough estimate based on intensity shows that the concentration is ~0.3%.

SEC. A scheme using 4 washes with 0.5% DDM buffer and 4 washes with 0.25% DDM buffer yielded lower peak fraction DDM concentrations. Despite this the concentration of DDM in peak SEC fractions persisted well above the target buffer concentration of 0.03%.

Tests of the sensitivity of the insulin aggregation assay to DDM showed a strong effect of even fairly small changes in DDM concentration. An increase from 0.03% (588 µM) to 0.05% (972 µM) DDM results in a substantial attenuation of aggregation. Hence changes on the order of 0.01% (~200 µM) DDM have a substantial effect on aggregation during the assay (Figure 16A).

Further difficulties were encountered when raising the molar ratio of insulin to σ_1. Ideally at least a ratio of 1 would be tested, corresponding to previously published data on σ_1 [14] and in line with the stoichiometric requirement expected of most chaperone proteins. With the high concentrations of insulin needed to overcome inhibition by detergent equally high concentrations of σ_1 were required. Concentrating protein samples to achieve this however resulted in concentration

of the detergent as well as protein. Mitigating this through dilution and removal of detergent resulted in precipitation of the protein sample.

The sensitivity to DDM concentration coupled with trouble controlling the DDM concentration in samples of σ_1, as well as reaching high concentrations of protein at low concentrations of DDM, compounded difficulties in getting an accurate measurement of chaperone activity. As a result, we were unable to conclusively test chaperone activity in our samples.

3.3.4 σ_1-*Client Protein Interactions*

One of the central remits of proposed work on σ_1 was to study its interactions with client proteins. With the majority of σ_1 clients being integral membrane proteins which would represent tough targets to work with we chose to focus on interactions with soluble partners. To that end two possible targets were identified as viable subjects of study.

The first interaction examined was that of σ_1 with the C-terminus of the NMDA receptor. Published peptide scanning experiments provided evidence localising this interaction to two particular sequences within the C-terminus of the NR1 subunit of NMDA receptors [37]. These sequences (of around 10 amino acids) were tested for binding to σ_1 using an MST based approach that proved successful for monitoring the interaction between σ_1 and the HSP70 family proteins BiP and HSC70. These experiments were unable to demonstrate binding.

3.3.4.1 σ_1-*HSP70 Family Interaction by MST*

Of the originally considered interactions, those with the HSP70 family appeared not only central to the function of σ_1 but also showed positive preliminary results and were considered amenable to work by NMR since previous work had shown an interaction between BiP and the σ_1-CD construct [59].

Titration of BiP into full length σ_1 fused to GFP led to no observable binding signal. Worries that this may be due to interference from the GFP tag were then

addressed by using cleaved σ_1 chemically labelled with a red Alexa Fluor NHS Ester dye (Alexa Fluor 594). Due to the low number of free primary amines present in σ_1 only two positions were available for derivatisation, either lysine 147 or the N terminus. These conditions provided well controlled labelling without the need for further mutagenesis of the protein.

Using chemically labelled protein binding signal was obtained for the σ_1-BiP interaction, giving a low micro molar range K_d (Figure 17A). Exact K_d values proved difficult to obtain as complete and consistent saturation of the binding signal required for fitting could not be achieved. Nonetheless an apparent low affinity interaction could be observed. This low affinity range is consistent with prior NMR work with σ_1-CD [59] and is not unexpected for a cochaperone interaction, some of which have been shown in the literature to be transient with fast on and off rates and high K_d values [106–109].

The publication of the single TM structure of σ_1 re-raised the question of the receptors topology and as a result also brought into question many of its reported interactions. With an initial proposal of a cytosolic orientation of the σ_1 β-barrel by Kruse *et al.* [12] being incompatible with a BiP interaction we decided to fully explore alternative explanations for both this data and our and previously published observation. With the very high degree of similarity between human HSP70 proteins we hypothesised that the BiP interaction may actually reflect a native interaction with a cytosolic HSP70 protein.

To that end we selected HSC70 (HSPA8) as an alternative with which to test σ_1-HSP70 interactions as it is the most abundant of the cytosolic HSP70s and plays a central constitutive role. Titration of HSC70 into fluorescently labelled σ_1 also yielded a binding curve with a similar affinity to that observed for BiP (Figure 17A). Like BiP, the interaction could be localised to the NBD. Binding, albeit at a lower affinity than for full length, could still be observed on titration of just the HSC70 NBD into σ_1 and was not observed for a construct comprising the SBD and C-terminus of HSC70 (Figure 17B) consistent with the interaction being regulatory rather than σ_1 binding as an HSC70 substrate.

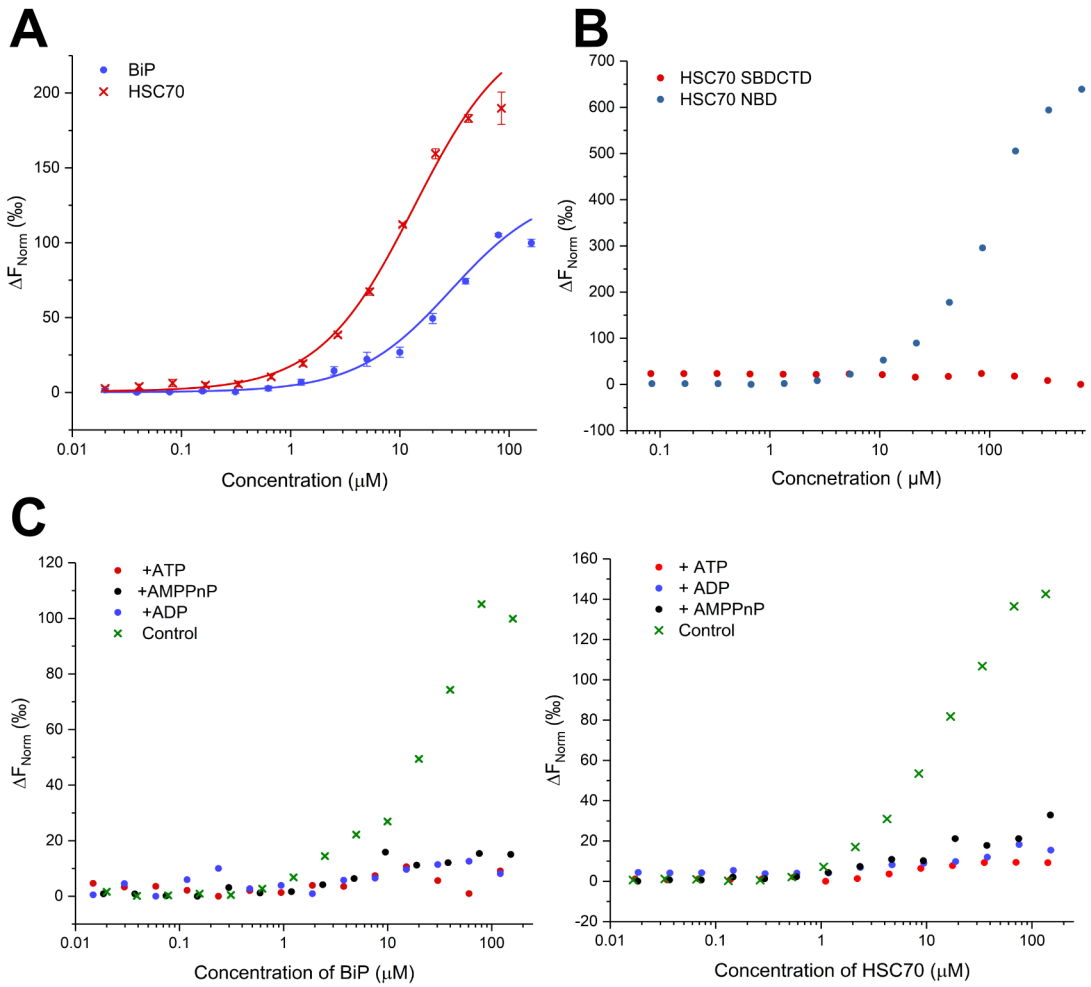


Figure 17: Summary of MST binding experiments on σ_1 and the HSP70 proteins BiP and HSC70 **A:** Binding curves for both BiP (Blue) and HSC70 (Red). In the absence of saturation fitting was unreliable but provided an estimate of the lower limit for affinity. Fitting yielded a K_d of ~13 µM for HSC70, and a K_d of ~30 µM for BiP. **B:** Comparison of binding against the NBD of HSC70 and the SBDCTD, a construct comprising the rest of the protein. Although a slight response occurs at very high concentrations there is no clear binding curve for the SBDCTD. **C:** Effects of nucleotide on BiP and HSC70 binding. In both cases, the presence of nucleotide, as well as the non-hydrolysable ATP analogue AMPPnP, appears to abrogate binding.

Given the importance of nucleotide in the behaviour of the HSP70 family further work was carried out to ascertain the effects of both ATP and ADP on this interaction. Addition of either nucleotide abrogated binding as did the non-hydrolysable analogue of ATP, AMPPnP (Figure 17C). That σ_1 possesses affinity for the apo form of the HSP70 presents an unresolved conundrum given that in a cellular context HSP70 proteins are unlikely to exist in this state.

Overall the data shows a fairly clear and likely regulatory association between HSP70 family proteins and σ_1. Association of σ_1 with both HSC70 and BiP shows that either a cytosolic or ER facing orientation of σ_1 are compatible with a σ_1-HSP70 family interaction, though no indication of which is correct is given by this work. In the absence of any clarifying cell biological work we chose to focus NMR based structural efforts on the HSC70 interaction given that of the numerous proposed σ_1 interactions a larger volume appeared compatible with a cytosolic orientation.

3.3.4.2 σ_1-*HSP70 Family Interactions by NMR*

Based on our MST data we carried forward the use of the apo form of HSC70 into our NMR interaction studies. An initial NMR titration was carried out looking at $^{15}N/^{13}C$ labelled σ_1. As with prior work on the σ_1-CD looking at the Bip interaction [59], sample stability was a problem, although we were able to alleviate this to some extent through increasing the sample pH from 6.5 to 7.2. The resulting data proved inconclusive. Slight shifts were observed in both backbone amide resonances and the methyl resonances of CHSQC spectra of the protein on addition of HSC70 NBD (Figure 18 and Figure 19) but these were small and not well clustered in the sequence. Furthermore, loss of signal resulting from sample instability and precipitation prevented the observation of chemical shift perturbations (CSPs) at high molar ratios of HSC70 NBD and σ_1.

A clear interpretation of the experiment was also obfuscated by poor coverage of the protein in DDM due to the large size of the protein-micelle complex. This resulted in only the more flexible regions of σ_1 being clearly observable. Recently

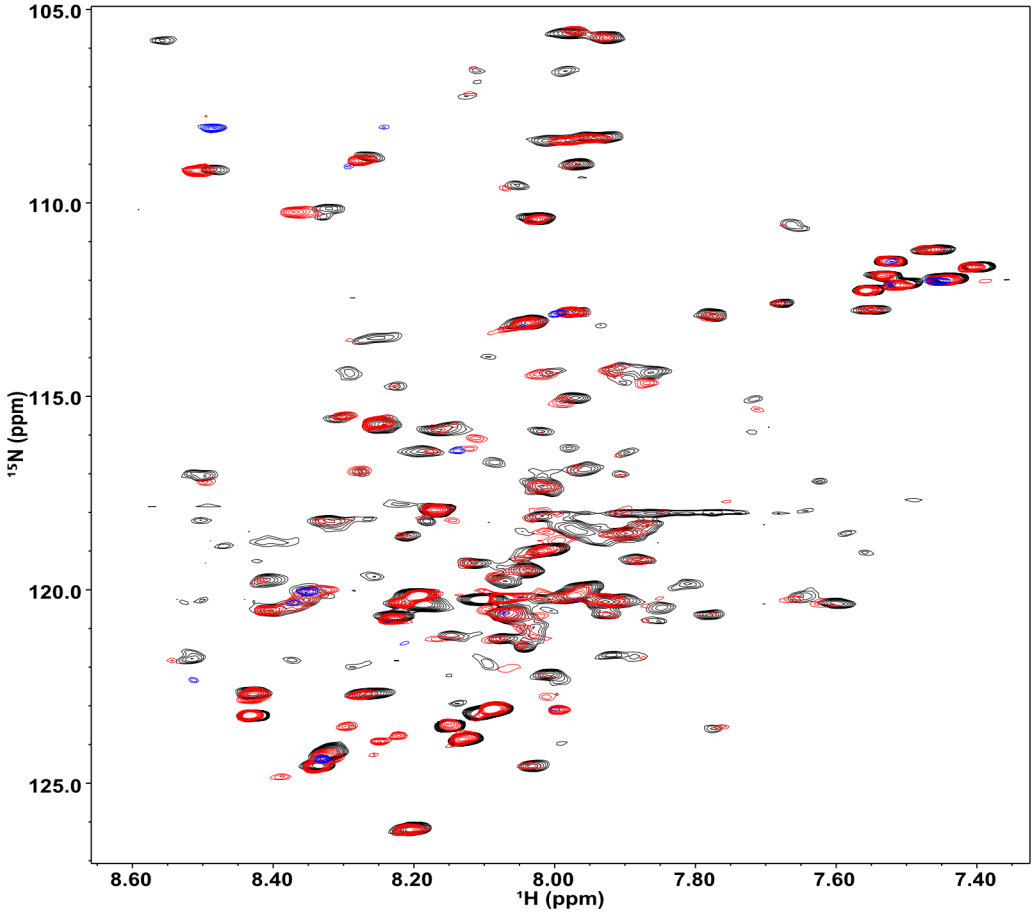


Figure 18: Titration of HSC70 NBD into a 100 µM $^{15}N/^{13}C$ labelled sample of full-length σ_1. Titration points were followed with ~1.5 hour ^{1}H ^{15}N BEST-TROSY experiments at 750 MHz (^{1}H) and 37 °C. Spectra for molar ratios of 0, 0.5 and 1 σ_1:HSC70 NBD are shown in black, red and blue, respectively. Extensive signal loss occurred throughout the titration and few resonances can be seen at a molar ratio of 1. Data for an molar ratio of 2 was collected but showed no signal.

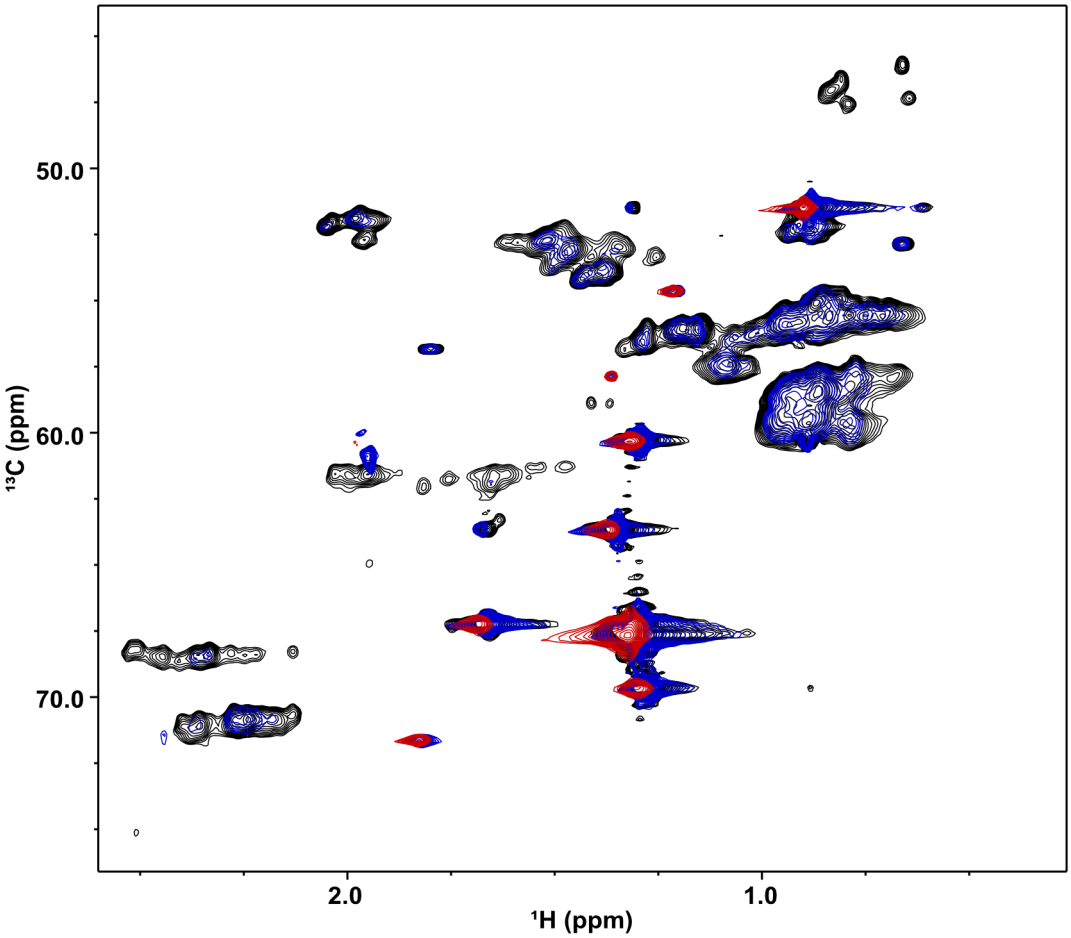


Figure 19: Titration of HSC70 NBD into a 100 µM $^{15}N/^{13}C$ labelled full-length σ_1. Titration points were followed with ~1.5 hour ^{1}H ^{13}C HSQC experiments collected on a 750 MHz (^{1}H) spectrometer at 37 °C. Spectra for molar ratios of 0, 0.5 and 1 σ_1:HSC70 NBD are shown in black, blue and red, respectively. Extensive signal loss occurred throughout the titration and few resonances can be seen at a molar ratio of 1. Data for a molar ratio of 2 was collected but showed no signal.

published assignments for the NBD of HSC70 [147] and the localisation of the σ_1-HSP70 interaction to the NBD provided a means of easily circumventing these difficulties. Using $^{15}N/^{13}C$ labelled HSC70 NBD, which yielded good spectra, we performed the titration in reverse (Figure 20 and Figure 21).

In the case of the reverse titration limitations occurred due to the yields of σ_1 which from a 10-litre growth were insufficient to fully saturate a 210 µM HSC70 NBD sample. Nonetheless a titration was performed up to a molar ratio of 0.5 which at these protein concentrations should be sufficient to start to show spectral changes resulting from even weak interactions between the two proteins. Further improvements on sample stability were observed as the temperature was dropped from 37 °C to 25 °C. No precipitation was observed up to an NBD:σ_1 molar ratio of 0.5 in the sample over the course of two days. Addition of σ_1 to this sample however resulted in no significant spectral changes in either the backbone amide resonances or the methyl region of HSC NBD carbon spectra.

3.3.5 *Functional state of σ_1 in DDM*

Establishing the functional state of σ_1 in DDM proved substantially more difficult than doing so in DPC. Some notion of the state of the protein could be obtained through CD where a largely α-helical signal is obtained comparable to signal obtained for σ_1-Δ35 in DPC (Figure 22). This signal seems in line with the fairly close overlap of spectra in DDM and in DPC (Figure 23) which may well suggest that excess helical structure is present. A notion that is further supported by comparing the CD spectra collected with a CD spectrum for σ_1 simulated on the basis of the X-ray crystal structure (Figure 22).

With the poor spectral quality obtained in DDM however, assignment of resonances and hence any attempt at accurately predicting secondary structure proved impossible. It is difficult to say therefore how closely the structure being observed approximates the known crystal structure.

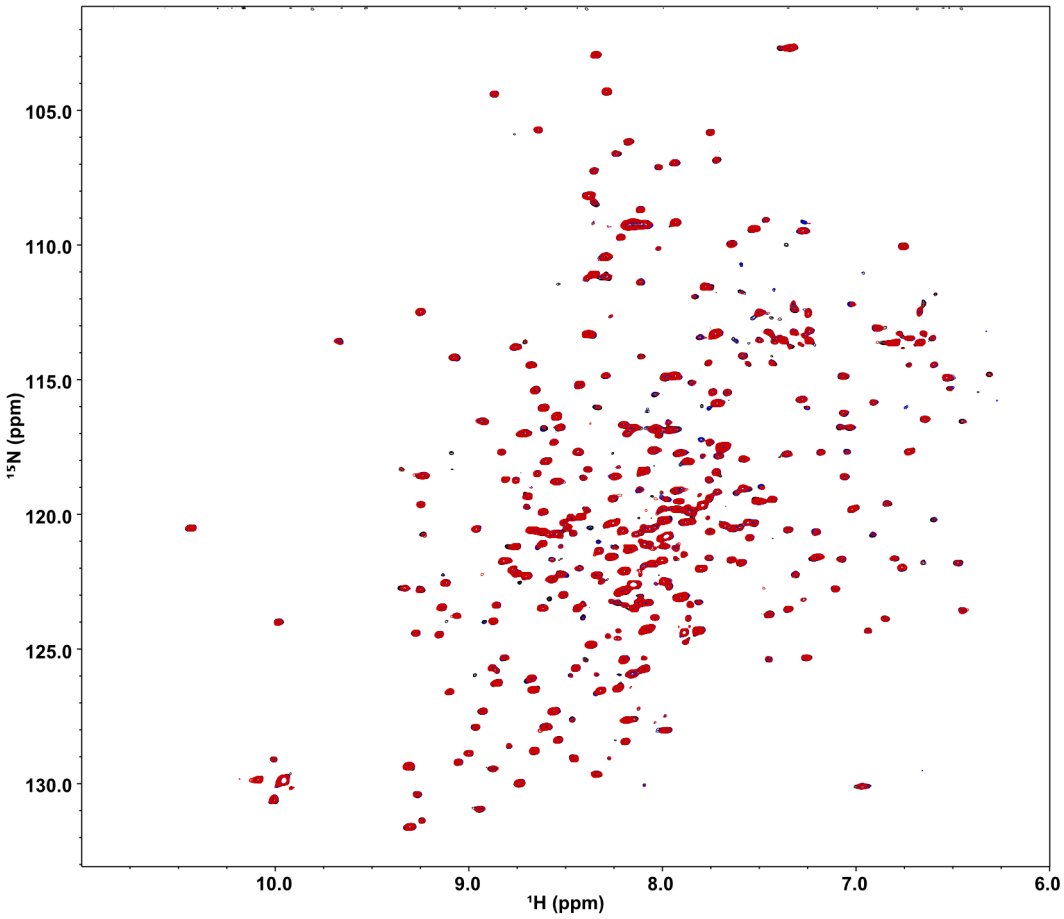


Figure 20: Titration of full-length σ_1 into a 210 µM $^{15}N/^{13}C$ labelled HSC70 NBD. Titration points were followed with ~1.5 hour ^{1}H ^{15}N Best TROSY experiments (750 MHz ^{1}H) at 25 °C. Spectra for molar ratios of 0, 0.25 and 0.5 HSC70 NBD:σ_1 are shown in black, blue and red, respectively. The maximum concentration of σ_1 was limited by yields from its purification.

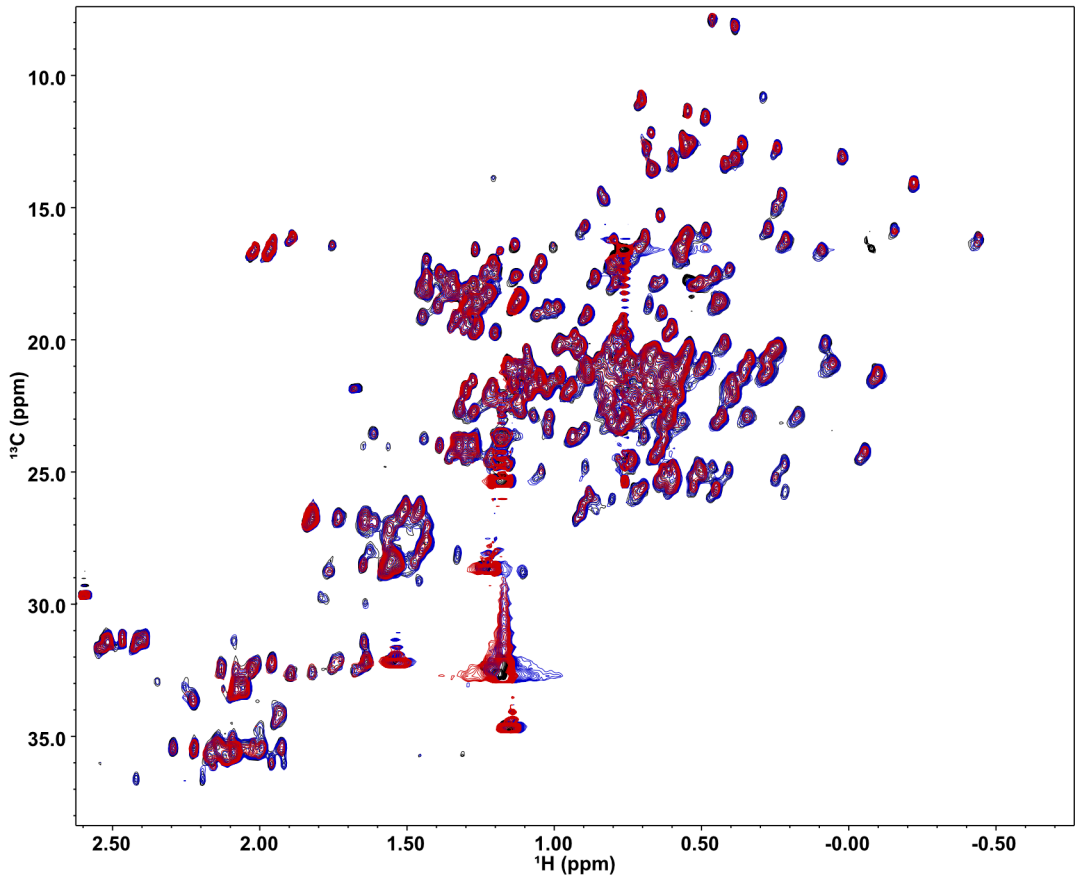


Figure 21: Titration of full-length σ_1 into a 210 µM $^{15}N/^{13}C$ labelled sample of HSC70 NBD. Titration points were followed with ~1.5 hour ^{1}H ^{13}C HSCQ experiments (750 MHz ^{1}H) at 25 °C. Spectra for molar ratios of 0, 0.25 and 0.5 HSC70 NBD:σ_1 are shown in black, blue and red, respectively. The maximum concentration of σ_1 was limited by yields from its purification.

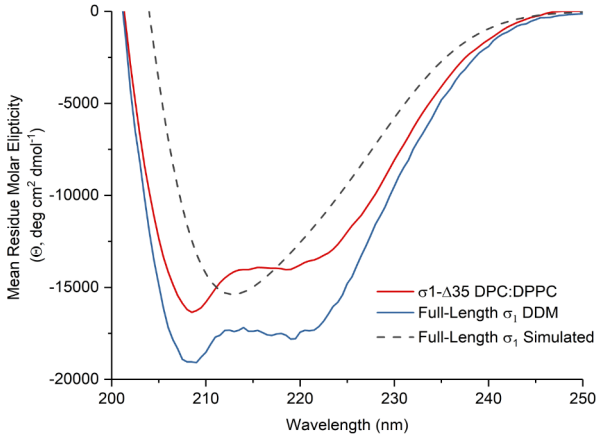


Figure 22: Far UV circular dichroism spectra for σ_1 full length in 0.05% (972 µM) DDM (Blue) and σ_1-Δ35 (Red) in 50 mM DPC supplemented with 5 mM DPPC. Both spectra are collected in 5 mM potassium phosphate buffer at pH 6.5 and plotted as mean molar ellipticity per residue. Simulated CD spectrum (produced using DichroCalc [149, 150]) based on the crystal structure of σ_1 (PDB 5HK1) is also shown (Black Dotted Line) for comparison.

The functional data available in DDM is also not completely clear, though the general tone remains negative. With no ligand binding observable by either radioligand assays or alternative biophysical methods the observed ligand dependent NMR shifts could not be corroborated. These shifts are also of dubious providence since previous work on σ_1-Δ35 was able to generate similar shifts that were later shown to be artefactual in nature. In general, the data do not strongly support high affinity ligand binding and suggest that if there is an interaction present it is likely occurring at a lower affinity than would be expected for a physiologically relevant interaction.

The oligomeric state of the protein is conclusively monomeric. Some propensity to oligomerise was observed in samples through crosslinking and the behaviour of the E172A mutant in SEC and SEC-MALS experiments, but there was no indication of a stable trimeric receptor as reported in the crystal structure of σ_1 [12]. Although cell biological work has demonstrated the relevance of monomeric σ_1 [45] and reports in the literature indicate that a dimeric form of σ_1 [17] may be sufficient to bind ligands with high affinity, our data shows both a lack of oligomerisation and high affinity ligand binding, making the relevance of the σ_1 monomers obtained here questionable.

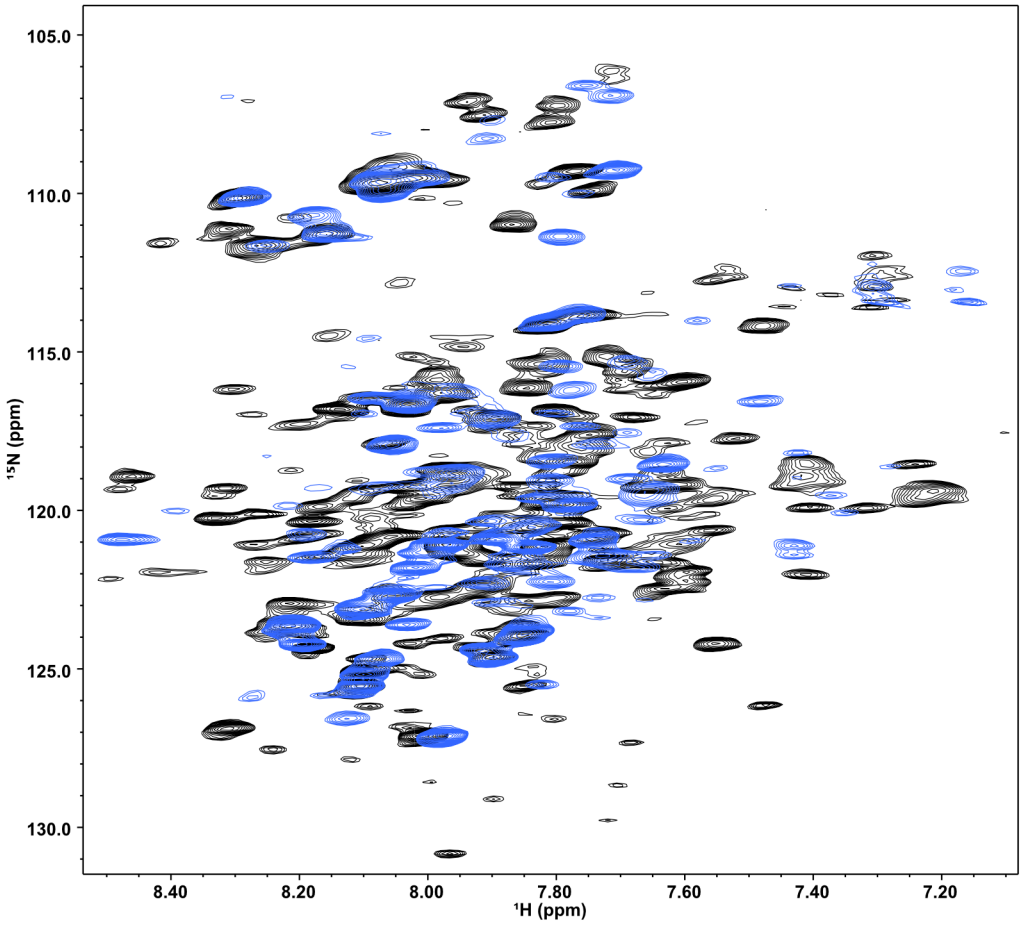


Figure 23: Overlay of a ^{1}H ^{15}N BEST-TROSY of full-length σ_1 in 50 mM DPC and 5 mM DPPC (Black) and a ^{1}H ^{15}N BEST-TROSY of σ_1 in 10 mM DDM both spectra were collected on a 750 MHz (^{1}H) spectrometer in 20 mM potassium phosphate buffer at pH 6.5 and 41 °C.

The observation of a protein-protein interaction between σ_1 and a HSP70 family protein provided some grounds for optimism. Issues however exist with the data since the interaction observed by MST was low affinity and could not be corroborated by NMR. In addition, the interaction appeared to require the biologically less relevant apo form of the HSP70 protein .

Overall it seems unlikely that the σ_1 we were able to produce and isolate from bacterial membranes in DDM was fully functional. The complete lack of any oligomeric behaviour and high affinity ligand binding not only contradicts the crystal structure but also published work with the MBP fused receptor in *E. coli* [17, 63–65]. In addition to this, CD and NMR provide some evidence that the conformation of the protein in DDM is not very different to that seen in DPC.

3.4 LMG EXTRACTION AND PERIPLASMIC TARGETING

At the time of the publication of the structure of σ_1 we possessed no functional samples of σ_1. Understanding why our protein was non-functional hinged on comparisons with previously published work showing active σ_1. The active protein used for structure determination was produced in insect cells rather than *E. coli* and extracted in LMNG rather than DDM. GFP fused σ_1 expressed into *E. coli* membranes was thus extracted into LMNG to test whether the difference in detergent was the major contributor to functionality. Although extraction was poorer than that observed for either DDM or DPC we were able to extract a reasonable amount of protein. The produced protein however was observed to be monomeric by SEC-MALS (Figure 24) and as such we were unable to reproduce the extraction of trimers as reported in the literature. This seems to suggest that the detergent was not the main issue with retaining function in σ_1 but rather the expression system.

That σ_1 would not functionally express in our hands in *E. coli* proved puzzling given the literature reports of functional DDM extracted receptor expressed in *E. coli* membranes [17, 63–65]. The major difference between our methodology and

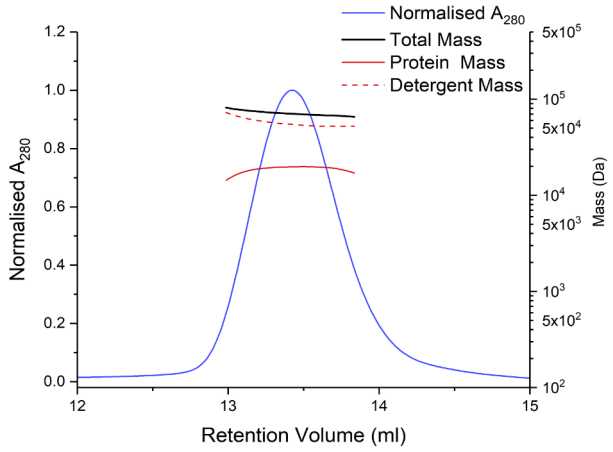


Figure 24: SEC-MALS protein conjugate analysis of full-length σ_1 extracted and purified in LMNG.Protein mass estimate is ~22 kDa in line with expectations for a monomeric form of σ_1, which has an expected mass of ~25 kDa

the reported methodology lies in the expression tag used. The literature reports the use of an MBP fusion as opposed to a GFP fusion. Extensive work in our laboratory had failed to reproduce this MBP fusion prep and as such we were reluctant to go over already covered ground.

An examination of the properties of the two fusion tags, MBP and GFP, shows them to be fairly similar in size and behaviour however one key difference is the targeting of the proteins. MBP is periplasmically targeted in *E. coli*. This raised the possibility that insertion of σ_1 into *E. coli* membranes was occurring in an orientation unfavourable for folding. To test this, we used a DsbA leader sequence inserted N terminal to our GFP which has been reported to target GFP to the bacterial periplasm [162]. Introduction of this modification however abolished expression of σ_1 and we were unable to proceed further with expression in *E. coli*.

4

FLUORESCENCE MICROSCOPY OF σ_1

4.1 OVERVIEW

The high degree of uncertainty surrounding the cell biology of σ_1 underlined the importance of corroborating and contextualising its structural biology and biophysics. Quantitative fluorescence microscopy offers a powerful and complementary approach with which to try and derive *in cellulo* correlates for our *in vitro* work. Two areas, in particular, presented compelling and amenable questions for *in cellulo* work. The first of these is the interactions of σ_1 with its client proteins. The application of Förster Resonance Energy Transfer (FRET) to measure protein-protein interactions is well established and can be extended to make measurements in live cells. Combining FRET with a Fluorescence Lifetime Imaging Microscopy (FLIM) approach allows many of the problems associated with ratiometric FRET measurements to be circumvented: briefly, the fluorescence lifetime, unlike intensity which is used for ratiometric measurements is independent of concentration and optical settings, in particular laser power, as such less stringent calibration is required for data acquisition and interpretation. This renders FLIM a powerful means with which to model protein-protein interactions in a cellular context [163–165]. The second area we hoped to address was that of the oligomeric state of σ_1. Determining the oligomeric properties of a protein directly within a cell presents a challenging but important problem in linking structural and cell biological studies. Number and Brightness (N&B) offers a way of monitoring self-association in live cells [166–169], providing the opportunity to investigate the relationship between oligomerisation and the biological role of σ_1.

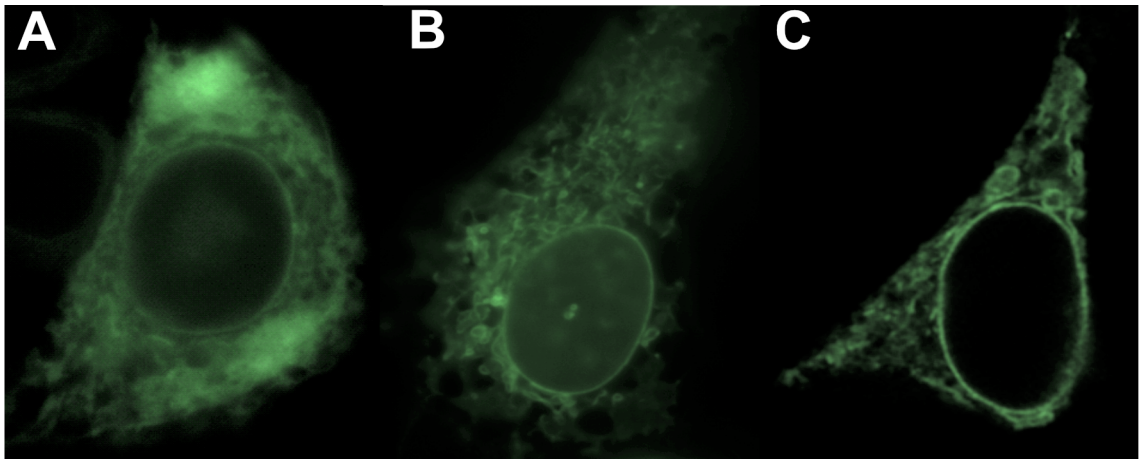


Figure 25: Expression of σ_1 in three different cell lines. In all cases images are of transiently transfected live cells, and are selected as representative of the general population of successfully transfected cells. **A:** Wide field fluorescence image of a TZM cell expressing σ_1-mVenus. **B:** Confocal image of a COS7 cell expressing σ_1-mClover **C:** Confocal image of an undifferentiated SH-SY5Y cell expressing σ_1-mClover.

4.2 EXPRESSION AND LOCALISATION OF σ_1

Live cell fluorescence imaging is dependent on both the appropriate tagging of the molecule of interest and on being able to induce expression in a desired cell line. Tagging of σ_1 was addressed by using either a C-terminal mVenus fusion or a C-terminal mClover fusion which have the advantage of forming strong FRET pairings with either mTFP (Turquoise Fluorescent Protein) [170] or mRuby, respectively [171].

Transient expression of σ_1 proved straightforward with σ_1 transfecting and expressing robustly in three different cell lines: TZMs, COS7s and undifferentiated SH-SY5Y cells (Figure 25). In all cases a diffuse thread like pattern of expression is observed with strong staining of the nuclear envelope. Comparison of this with ER-Tracker staining in COS7 cells (Figure 26) as well as the observed morphology of staining shows that under our conditions σ_1 populates the ER. This is in strong agreement with published imaging studies performed on σ_1 as well as cellular fractionation studies [14, 46, 66, 69, 172].

Subsequent studies were carried forward using COS7 cells which due to their flattened shape and large volume of ER, produce very clear images of the ER and

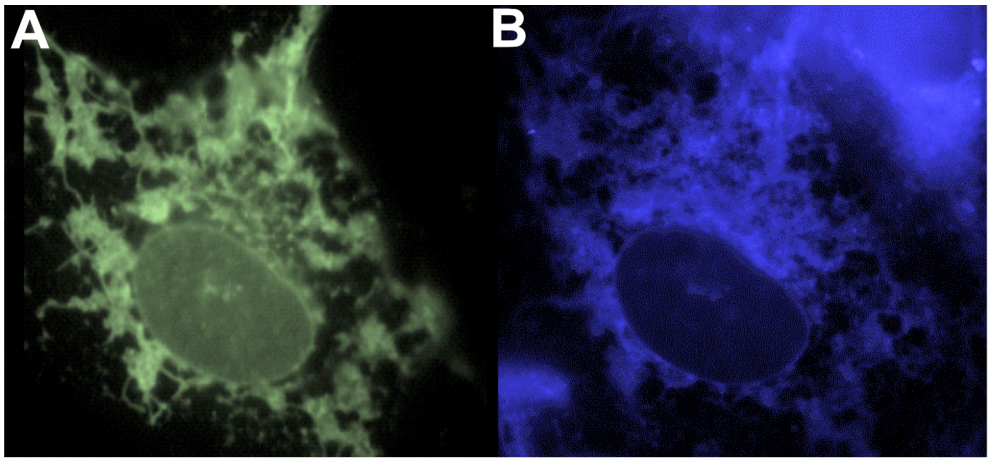


Figure 26: Localisation test of σ_1-mClover in COS7 cells. Both are live-cell widefield images taken simultaneously of the same cell. **A:** Green channel image showing the staining pattern of σ_1–mClover. **B:** Blue channel image showing the staining pattern of ER-Tracker blue-white. A high degree of congruence between both localisation patterns was observed.

consequently of fluorescently labelled σ_1. Further testing showed this ER staining pattern is maintained under a variety of conditions including incubation of cells with different σ_1 ligands, either agonists or antagonists, which produced no obvious changes in localisation. Notably, at no point was there a bulk staining of the PM, although low level re-localisation cannot be ruled out by these experiments. Modification of an N-terminal RR motif in σ_1 that is a potential ER retention signal [157] also failed to alter the staining pattern achieved (Figure 27).

The tendency for improperly folded membrane proteins to be retained in the ER [173] advocates for a certain degree of caution in interpreting the above data. Nonetheless, ER localisation is consistent with previous observations [14, 46, 66, 69, 172] and suggests that σ_1 behaves appropriately in our system. Ultimately, complete validation of the transiently expressed *in cellulo* system is reliant upon data showing that the receptor behaves in a functional manner under these conditions.

4.3 INTERACTIONS BY FRET-FLIM

One functional test of σ_1 is to monitor its interactions with partner proteins. To achieve this, we fluorescently labelled the target client with an appropriate FRET

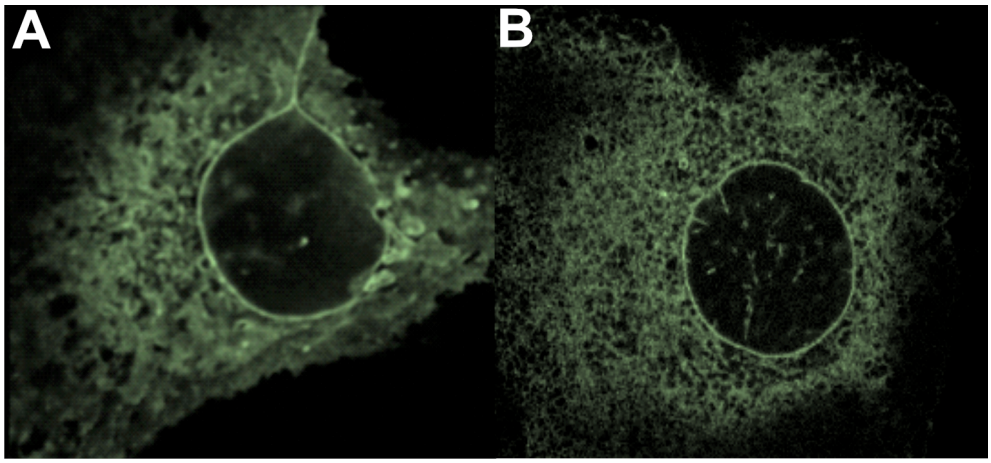


Figure 27: Comparison of the staining pattern observed for σ_1-mClover wild type and the putative localisation signal mutant RRQQ. Both images are representative confocal images of live COS7 cells. **A:** σ_1-mClover wild type, and **B:** σ_1-mClover RRQQ

.

acceptor prior to co-transfection of donor fused σ_1. As FRET provides an alternate de-excitation pathway for donor molecules, the fluorescence lifetime of these molecules decreases when in close proximity to an appropriate acceptor. Comparison of the donor fluorescence lifetime in cells transfected with σ_1 alone or with both σ_1 and its client hence allows for FRET, and consequently an interaction between the two proteins, to be detected.

Initial studies used a C-terminal mVenus tagged σ_1 construct paired with mTFP tagged Kv 1.4 or Kv 1.5. The Kv channels were chosen for two reasons, the first being that they represent one of the better characterised σ_1 clients with both electrophysiology experiments and crosslinking demonstrating a direct association between the two proteins [15, 50, 51]. The second is that the channels are homotetramers and as such require only transfection and expression of one construct, in contrast to many other σ_1 ion channel clients which form hetero-oligomeric assemblies. Unfortunately, effective measurement of an interaction was hampered by low levels of channel expression which precluded continued experiments with these two channels.

With the difficulties encountered with Kv 1.4 and 1.5 as well as the renewed ambiguity surrounding the question of topology resulting from the single TM struc-

ture of σ_1, work shifted focus to look at the interaction of σ_1 with the HSP70 family of proteins. The HSP70 family represents not only a well characterised interaction but also one where the identity of the interacting protein is dependent on the topology of σ_1.

Two candidates presented themselves as putative clients, the first being the ER resident protein BiP, which has been widely reported in the literature as a σ_1 client [14, 36, 43]. The second is the cytosolic cognate of BiP, HSC70, which we proposed as an alternative interaction partner for a cytosolic orientation of the receptor (see subsection 1.1.6 and subsubsection 1.1.4.1).

We set out to test both interactions in live cells using mRuby fused constructs of both BiP and HSC70. Cells could be robustly transfected with BiP to high levels although transfection with HSC70 proved more problematic, and expression levels remained significantly lower than those seen for BiP despite attempts at optimisation. Co-transfection with σ_1-mClover proved an additional obstacle with most cells showing either green or red fluorescence and almost none showing both. Staggering the transfection of the mRuby fused HSP70 partner and σ_1 improved matters by allowing more time for mRuby maturation but did not completely resolve the problem.

Experimental work was hampered by having to select the few cells that possessed both red and green fluorescence particularly in the case of HSC70 co-transfection. In all cases these cells had much stronger fluorescence in the green channel than in the red. Despite this, shorter lifetimes, 7.69 ns and 7.72 ns for cells co-transfected with HSC70 and BiP respectively, versus 7.77 ns for the σ_1-mClover alone control were detected (Figure 28). Unexpectedly, both BiP and HSC70 showed FRET with σ_1, with a stronger effect observed for BiP. Under native conditions, interaction with both cytosolic and ER lumen localised soluble proteins should not occur for a uniformly inserted membrane protein with a topology like that of σ_1. With the formation of some unknown and peculiar ternary complex representing a weak hypothesis, it seems likely that either an artefactual interaction is being generated by overexpression or that the localisation of one or both co-

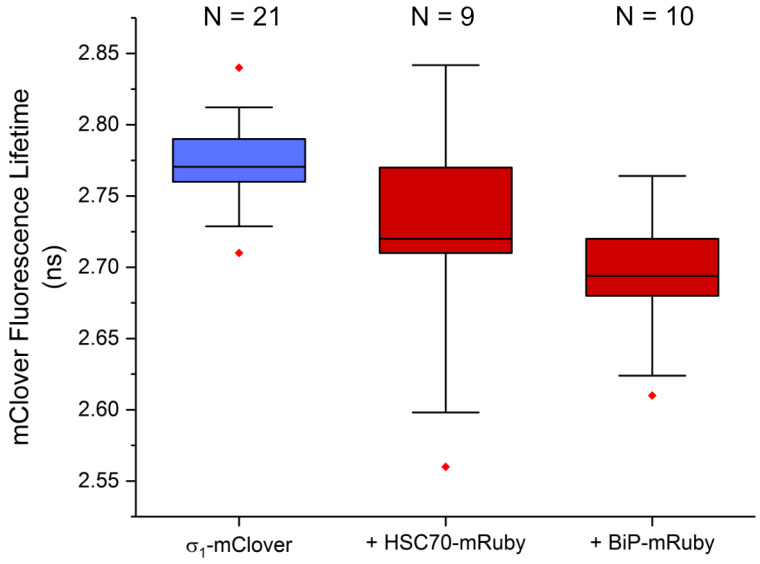


Figure 28: Box plot comparison of mClover fluorescence life time in COS7 cells expressing either σ_1-mClover alone (Blue) or co-transfected with both σ_1-mClover and either HSC70-mRuby or BiP-mRuby (Red). Box plot whiskers represent 1.5σ. Lifetimes were measured by fitting the decay curves obtained from FLIM experiments to a single exponential fit. In the case of co-transfected cells those chosen for imaging were selected on the basis of strong green and red channel fluorescence. In general BiP was expressed to higher levels than HSC70.

transfected proteins is not occurring correctly. In addition, the drops in lifetimes were slight: 2σ for the control cell population's lifetime was 0.06 ns, which places the mean lifetime for cells co-transfected with BiP just below the cut-off for significance and those co-transfected with HSC70 just above the cut-off for significance (2.71 ns). This difference in significance may be a result of the poorer expression of HSC70 resulting in less favourable donor:acceptor ratios in these cells.

4.4 OLIGOMERIC STATE OF σ_1 BY NUMBER AND BRIGHTNESS

N&B represents a method in the family of Fluorescence Correlation Spectroscopy (FCS) techniques that is dependent on the analysis of the fluctuations in intensity that occur within a time series of fluorescent images. The central point of this type of analysis is that the amplitude of these intensities is related to the brightness of the fluorescent particles which are diffusing in the volume being imaged. The molecular brightness expressed in photons per molecule per second is related to

the number of fluorescent molecules co-diffusing together which for a fluorescently tagged protein will directly correlate with its oligomeric state [166, 167] (Figure 29).

4.4.1 *Theoretical and Practical Considerations for N&B*

Using data from a photon counting detector N&B can be used to calculate the quantity B (apparent brightness) from a given set of images collected over a span of time t. B is defined by Equation 1 as the variance in intensity in one pixel (σ^2) divided by the average intensity of that pixel ($<I>$) and is related to the quantity ϵ, the molecular brightness of the particles being imaged. Defining N (apparent number) in an analogous manner as in Equation 2, gives a quantity which is proportional to the number of particles n being imaged.

$$B = \frac{\sigma^2}{<I>} = 1 + \epsilon \tag{1}$$

$$N = \frac{<I>^2}{\sigma^2} = \frac{\epsilon n}{1 + \epsilon} \tag{2}$$

From these definitions, a pair of equations for ϵ and n defined in terms of both average intensity and variance results (Equation 3 and Equation 4). These four equations and the relationships between them provide the central tenants of N&B analysis [167]. A result of this treatment is that any immobile particles will yield a brightness of one given that in this case $\sigma^2 = <I>$.

$$n = \frac{<I>^2}{\sigma^2 - <I>} \tag{3}$$

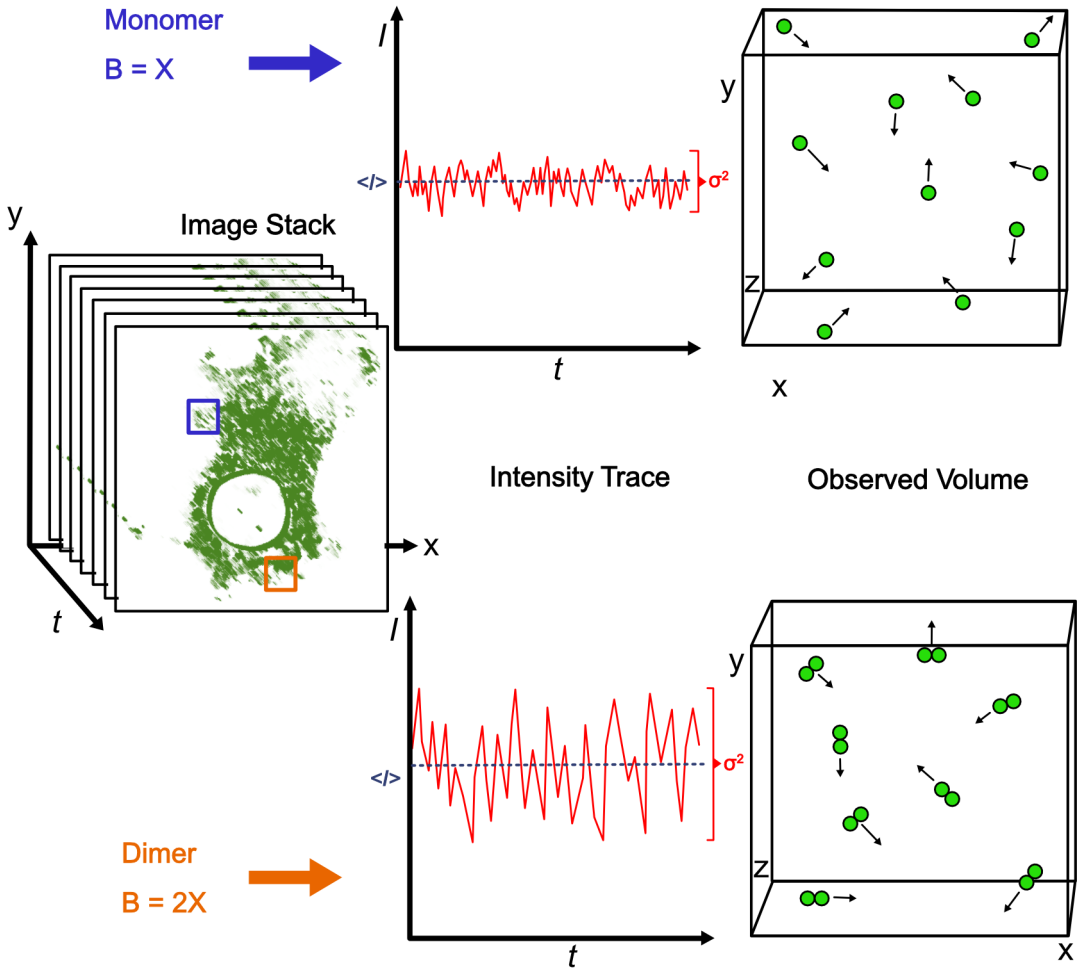


Figure 29: Schematic showing the principals of Number and Brightness. A series of images are collected over time t to form an image stack where the z-dimension corresponds to time. The gap between images, dt, is kept faster than the rate of diffusion but large enough for movement of particles in and out of the observation volume to be detected. At a given pixel the intensity over time t is analysed. Idealised intensity traces showing both the mean intensity $\langle I \rangle$ and variance σ^2 are shown; at the top for a pixel containing fluorescent particles diffusing as monomers, and at the bottom for a pixel containing fluorescent particles diffusing as dimers. The fluctuation of the intensity is larger for the observation volume containing dimers, corresponding to a larger calculated brightness. As a linear relationship is maintained, B in the case of the dimer, will be two times that for the monomer. Idealised observation volumes corresponding to each trace are also shown, the intensity fluctuations in each case arise from the movement of fluorescent particles in and out of the observation volume.

$$\epsilon = \frac{\sigma^2}{<I>} - 1 \tag{4}$$

It is important to consider however that immobility in N&B is a relative term. In order to apply N&B, sampling must occur rapidly enough for diffusion related fluctuations to be detectable [166]. The sampling speed (pixel dwell time in the case of rasterised images) needs to be fast enough that diffusion in and out of the acquisition volume during sampling does not average out the diffusive contribution to σ^2. This will depend on the rate of diffusion of the molecules being observed, being slower for example for membrane proteins relative to soluble proteins in the cytosol. Equally, sampling must be slow enough that movement of the molecules of interest during the acquisition time t is significant, lest they effectively appear as immobile.

In addition to the above theoretical considerations, several practical considerations are also required to yield useful N&B data. The majority of these considerations come from contributions to the variance (σ^2) which do not come from molecular diffusion. Fluctuations in intensity will also arise due to phenomena such as photobleaching and movement of the cell. For data to accurately reflect the aggregation state of a protein these contributions must be corrected for. In general, this can be achieved through the application of an effective high band-pass filter since most extraneous contributions to σ^2 will occur on a much longer time scale than molecular diffusion. There are several methodologies which can be applied to achieve this; in our case a previously developed method based on single exponential filtering was used [156].

4.4.2 *Applying Number and Brightness to* σ_1

The current cell biological model of σ_1 function postulates that inactive receptor is held in an immobile conformation likely associated with large oligomeric as-

semblies of σ_1 [44–46]. N&B on live cells offered us a direct way of testing this hypothesis. Since brightness is related to the oligomeric state of the protein the assembly and disassembly of σ_1 oligomers could be followed in live cells through measurements of brightness before and after the addition of σ_1 ligands. In this case a decrease in the measured brightness would correspond to the disassembly of σ_1 oligomers while an increase would indicate the formation of larger assemblies.

In total four different σ_1 ligands were assayed in this way: the strong and archetypal σ_1 antagonist NE-100 as well as three different agonists the high affinity ligand pentazocine and two lower affinity ligands cocaine and fluvoxamine. Data acquisition with fluvoxamine was hampered by its toxicity at the 20 µM concentration, previously used for NE-100 and pentazocine on live cells. The net result was that pre-incubation for 30 minutes with 20 µM fluvoxamine resulted in the majority of cells being dead or obviously unhealthy.

For both the antagonist NE-100 and the agonist pentazocine robust data could be collected (Figure 30) and a very clear difference was observed in data with the antagonist NE-100. For NE-100, a decrease in the average brightness from 1.09 to 1.02 could be observed, with all the cells monitored showing a decrease in brightness (Table 1). Addition of pentazocine resulted in a much subtler decrease in average brightness from 1.14 to 1.09, and in this case 12 out of 16 cells showed a decrease in brightness (Table 1).

Given that the data is paired with the brightness being collected for the same cell before and after drug addition, it is appropriate to use a paired statistical test to analyse the difference. Application of the Wilcoxon Signed rank test to the data for pentazocine addition (Table 1) gives a significant difference (two-tailed $p < 0.05$) for the brightness after the addition of pentazocine being lower than before. Application of the same test to the results with NE-100 also gives a significant difference (two-tailed $p < 0.01$).

The situation for cocaine however remained ambiguous with roughly half the cells showing an increase in brightness while the other half show a decrease (Figure 30 and Table 1). Though some degree of heterogeneity in response was ex-

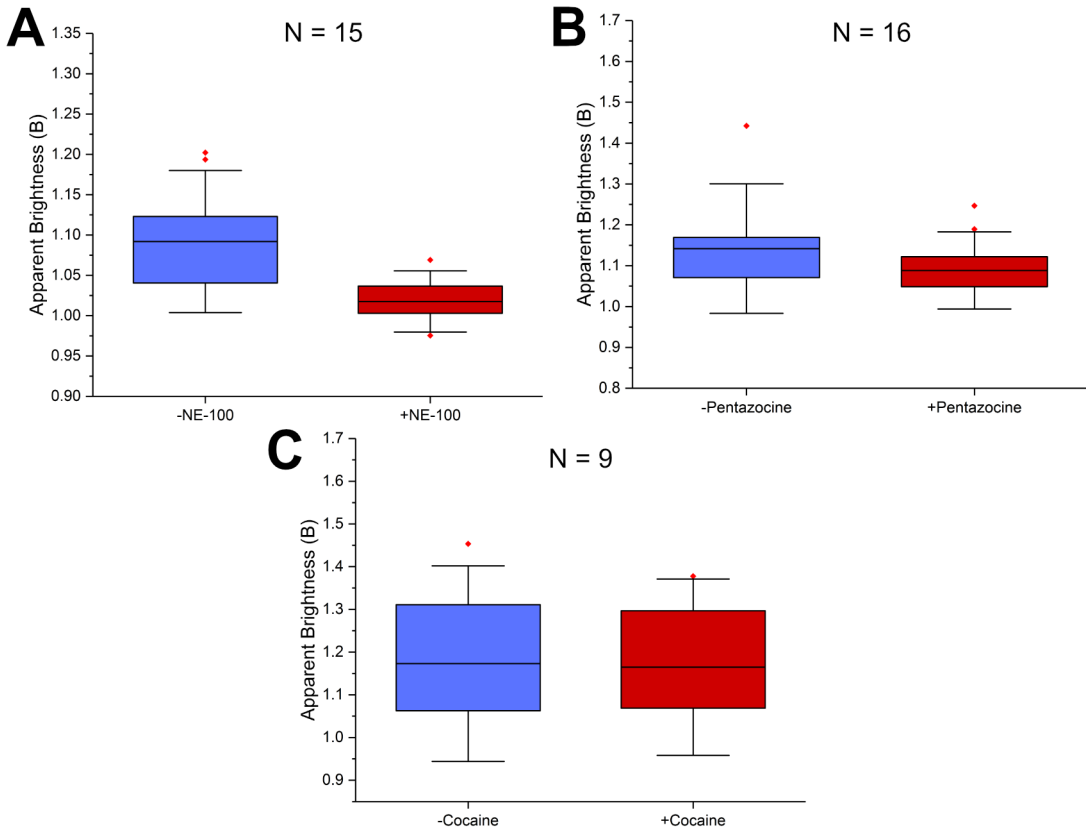


Figure 30: Boxplot comparison of average brightness calculated for cells before and after addition of various σ_1 ligands. Box plot whiskers correspond to 1.5σ, red dots correspond to data points outside the 1.5σ range. In all cases, 500-frame acquisitions were made on live COS7 cells using a confocal microscope with a pixel dwell time of 9.75 µs. Imaging after addition of ligand was performed after a 30 minute incubation at 37 °C. **A:** Brightness before (Blue) and after (Red) the addition of 20 µM NE-100. **B:** Brightness before (Blue) and after (Red) the addition of 20 µM pentazocine. **C:** Brightness before (Blue) and after (Red) the addition of 200 µM cocaine.

	Apparent Brightness					
	NE-100		Pentazocine		Cocaine	
Cell	−	+	−	+	−	+
1	1.011	1.010	1.055	1.084	1.325	1.378
2	1.015	0.993	1.104	1.125	1.063	1.108
3	1.041	1.015	1.005	1.020	1.311	1.354
4	1.046	1.014	1.058	1.065	1.015	1.046
5	1.097	1.053	1.251	1.247	1.050	1.074
6	1.098	1.049	1.052	1.047	1.079	1.069
7	1.123	1.069	1.151	1.133	1.147	1.117
8	1.057	1.002	1.148	1.118	1.115	1.040
9	1.040	0.975	1.091	1.057	1.454	1.297
10	1.098	1.008	1.104	1.066		
11	1.105	1.007	1.256	1.189		
12	1.102	0.991	1.084	1.017		
13	1.151	1.027	1.142	1.075		
14	1.202	1.037	1.187	1.098		
15	1.194	1.013	1.139	1.020		
16			1.442	1.051		

Table 1: Brightnesses before and after addition of different σ_1 ligands are shown to 4 s.f. Entries are ordered in ascending order of difference between brightness before and after ligand addition. Horizontal lines in the body of the table are inserted where the difference changes from positive to negative.

pected in live cells, and indeed was seen in the case of clearer data with pentazocine and NE-100, in the case of cocaine the heterogeneity prevented any effect from being detected. Application of the Wilcoxen signed rank test to the cocaine data shows no significant difference.

This highly heterogenous response to cocaine suggests that some confounding variable was at play in these experiments. One possibility may be related to the concentrations used and the affinity of the differnt ligands to σ_1. The affinity of NE-100 is sub nanomolar [174], and that of pentazocine is low nanomolar [20] while cocaine has reported affinities in the low to mid micromolar range [26]. As such, the 20 µM addition of NE-100 and pentazocine is likely far more than required to saturate σ_1, whereas cocaine may not be able to saturate σ_1 at 200 µM. This issue is highlighted further by any consideration of the bioavailability of the ligands

in our experimental conditions. Any deviation from ideal (100%) will reduce the quantity of ligand available to receptors in the cells. While the exact bioavailability of ligands in our conditions is unknown, both NE-100 and pentazocine are high enough above their expected K_d for all but the most dismal level of bioavailability to impede a response. For cocaine, where the experiments were carried out much closer to its K_d the situation is more ambiguous. The use of cocaine at higher concentrations becomes limited by both solubility and toxicity considerations.

4.4.3 *Interpreting the Results of Number and Brightness*

The data appears to indicate that application of either an agonist or an antagonist will result in disaggregation of σ_1. This effect also provides good evidence that the transiently expressed receptor is functional and binds ligands in live cells. That both agonists and antagonists disaggregate σ_1 contradicts previously published *in cellulo* work [45] and is not congruent with the prevalent model of σ_1 function [44]. There is however an alternate interpretation that must be considered. The decrease in brightness observed upon addition of the antagonist NE-100 is far sharper than that observed with pentazocine and brings the average brightness close to 1. As the immobile fraction yields a brightness of 1 [167], a possible alternate interpretation is that addition of antagonist is causing an immobilisation of σ_1 on the timescale of sampling used for these experiments. Published cell biological data would tend to support this immobilisation interpretation, and though this makes it highly attractive a clear conclusion would require further work and characterisation of this behaviour.

5

LESSONS FROM σ_1

5.1 FUNCTIONAL STATE OF ISOLATED σ_1

Initial indications suggested that the hurdles of ensuring the proper folding of full-length σ_1 had been surmounted by our novel GFP-tagged expression and purification approach: the protein readily inserted into *E. coli* membranes upon expression, the fused GFP fluoresced even after extraction with the mild detergent DDM, and the CD spectra of the protein appeared to show a folded α-helical protein. The poor dispersion seen in our NMR spectra fell in line with what was expected for an α-helical membrane protein. These observations were supported by the most prevalent structural model of the protein at the time, that of a two TM α-helical receptor. Furthermore, our earlier work with a truncation construct in DPC had validated at least one expected behaviour based on cell biological data within our system, namely the interaction with the ER chaperone BiP [59].

Nonetheless we were unable to unambiguously observe high affinity ligand binding. Often our attempts to do so were vexed by the age old scientific quandary of discerning when it is the method that is at fault or the system which is not behaving as expected. This issue was compounded by the difficulties associated with working on a poorly expressed membrane protein solubilised in detergent, which often meant that a conclusive negative observation could not be readily made. Throughout the extensive attempts to measure functional data in both DPC and DDM we were unable to, at any point observe ligand binding, chaperone activity or protein-protein interactions in a clear and reproducible manner that could be orthogonally corroborated.

Although the publication of the structure of σ_1 [12] was a frustrating and disappointing answer at the time, it did resolve this impasse. Despite many aspects of the published structure being unexpected, the structure explained well the body of mutagenesis data that had been published previously and precluded any serious doubts about its veracity [13]. The differences seen between the structure and our predictions made by NMR in DPC suggest that the receptor we were isolating was not in a natively folded state. This supports what has since become a well-documented aspect of DPC, namely its tendency to induce non-native structure and abrogate proper folding and function [158]. Though less conclusive due to the more limited NMR information available in the much larger micelle, it seems likely that the receptor extracted from *E. coli* membranes into DDM also suffers from some degree of compromised folding since there does not seem to be a significant difference in CD spectra of the protein in the two detergents and in both cases the receptor is clearly monomeric. Furthermore, the differences between the NMR spectra in DDM and DPC could well be accounted for by differences in the micelle size, dynamics, and different protein-detergent interactions. Crucially, the receptor isolated in DDM did not show any of the functional properties expected of it.

Despite these negatives there are hints of some native structure being retained throughout our work. Observed NMR shifts may indicate some residual binding to hydrophobic molecules and certain segments of the receptor even in DPC retain a native like fold. Furthermore, though weak and not fully corroborated, there is some evidence of native-like protein-protein interactions in our conditions and perhaps most intriguingly a drug binding knock-out mutant forms oligomers when extracted from *E. coli* in DDM. Ultimately however these hints of native-like function do not outweigh the strong negative evidence stacked against them. As such, whatever native-like structure may remain in the receptor would provide a very poor foundation for further studies.

A post-mortem of this work raises the question of why our protein is non-functional. Examining the differences in expression and purification between the

structure papers and our work highlights key differences in the detergent and expression system: Kruse *et al.* used insect cell expressed protein extracted in LMNG [12, 18] while we employed *E. coli* expressed protein extracted in DPC or DDM. Our cursory experiments with *E. coli* expressed receptor in LMNG did not detect the formation of oligomers as reported in the structure papers, or indeed by our collaborators working on σ_1 produced in mammalian cells and isolated in OG (Aricescu group, unpublished data). Were we to base our conclusions on this alone it would seem evident that the issues most likely stem from the expression system, *E. coli*. Work carried out prior to and during our studies of σ_1 by the Ruoho group however was able to produce an MBP-σ_1 fusion in *E. coli* [63] and demonstrate both ligand binding [17, 63] and oligomerisation [17, 64] when the protein was extracted into DDM. Prior attempts by our group to reproduce these results on MBP-σ_1 construct were marred by very low yields and problems with protein degradation. Why our receptor produced in *E. coli* does not behave as the Ruoho group reports is a hard question to address. Though they use guinea pig σ_1 the protein is 93% identical and 98% similar to the human receptor we used, and it is hard to countenance that the changes between the two would suffice to account for the radically different behaviours observed. Additionally, not all of the Ruoho group's observations tally up completely with the structure, in particular their radioligand binding data suggested a 2:1 receptor to ligand stoichiometry and they detected only even numbered assemblies of σ_1, never trimers.[17, 63]. The exact reasons behind these differences may well not be readily ascertainable and from a pragmatic standpoint the point may be moot as the most logical next step in any future work on σ_1 in our hands would be to change to a eukaryotic expression system.

5.2 EVALUATING THE *in cellulo* SYSTEM

The *in cellulo* work on the other hand stands on better foundations given the response of σ_1 in COS7 cells to ligands seen in N&B experiments. There are however

several points that would need to be addressed before solid conclusions could be drawn from these studies. The N&B data raises questions around the capacity of the technique to discern between disaggregation and immobilisation, and while this is not an insurmountable question it does require further experimental work; Fluorescence Recovery After Photobleaching (FRAP) and Fluoresence Cross Correlation Spectroscopy (FCCS) would provide alternative means of measuring protein mobility which could be well correlated with N&B data. The second question that remains open is why a FRET-FLIM interaction is observed both with HSC70 and BiP. The central problem for these experiments remains the inability to reliably produce cells co-expressing both σ_1 and either BiP or HSC70. Without reliable co-expression the number of observations is limited and subject to variations in cell morphology and apparent expression levels. The latter point in particular is problematic given the propensity of transient over expression to induce non-native protein-protein interactions.

5.3 LESSONS LEARNED

σ_1 proved an extremely frustrating system to study, although we no doubt have much company in this regard. Our earlier work highlights the perils of confirmation bias in science whereby tenuous observations that support accepted models are favoured despite them being better explained as either negative results or indicative of problems with the prevailing view of the system. In the case of our early work with σ_1, and of science in general, these observations are often unable to survive the paradigm shifts that move scientific work forward. In spite of these setbacks the σ_1 receptor remains intriguing both despite this and partly because of this, serving as an excellent example of the maxim that the most interesting questions are the most difficult ones.

Part III

REGULATION OF HSP70 PROTEINS BY HSP40 CO-CHAPERONES

6

EXPRESSION AND CHARACTERISATION OF HSP40 AND HSP70 TARGETS

Planned work on the HSP70-HSP40 system relied not only on the expression and purification of the proteins of interest but also on producing good quality, assignable spectra suitable for monitoring interactions within the system. Fortunately, recombinant expression in *E. coli* and subsequent purification of HSP70 family proteins proved straightforward. The three chosen HSP40 proteins on the other hand presented a more significant challenge at least with regards to obtaining the full-length proteins. From the outset of our work on the HSP40/HSP70 system we leveraged the modularity of the HSP70 and HSP40 proteins by using a divide and conquer approach both for the assignment of HSC70 and for the expression of tractable J-domain constructs of our selected HSP40 targets: DNAJB2, DNAJB6 and DNAJC5.

6.1 HSC70

In the case of HSC70–our primary target for NMR studies–we subdivided the protein into three domains: the NBD (residues 1-385), the core of the SBD domain (residues 386-555) referred to here as the SBD, and a C-terminal helical and disordered domain (residues 556-646) referred to here as the CTD. This further allowed for the generation of constructs consisting of combinations of these domains, namely the N-terminal NBDSBD (residues 1-555) and the C-terminal SBDCTD (residues 386-646) containing the complete SBD and a flexible C-terminal extension. Expression of the full-length protein as well as its individual domains

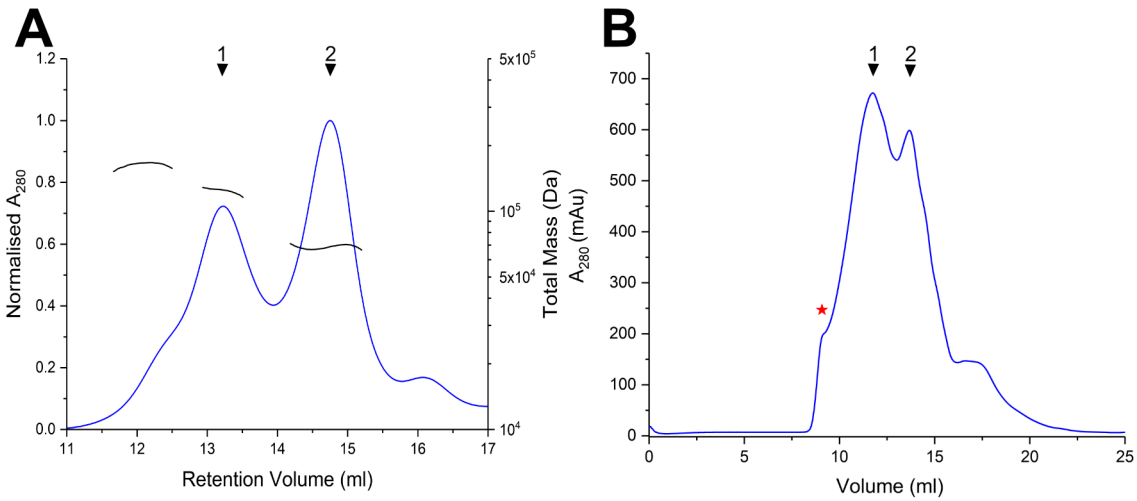


Figure 31: Oligomeric state of HSC70 by SEC and SEC-MALS. **A:** SEC-MALS trace for full-length HSC70. Mass calculations based on light scattering are shown for the two main peaks as well as the far left shoulder. Peaks labelled 1 and 2 correspond to dimeric and monomeric species with calculated masses of ~120 and ~69 kDa, respectively. Mass for the far left shoulder is ~165 kDa. **B:** Example SEC chromatogram obtained during purification of full-length HSC70 peaks corresponding to the two species characterised by SEC-MALS and labelled as in A. The shoulder corresponding to the void volume peak is marked with a red star.

and combinations thereof was robust in *E. coli* with favourable yields allowing for the straightforward production of NMR samples.

6.1.1 *Oligomerisation of HSC70*

From the outset, the oligomeric behaviour of full length HSC70 was evident, with SEC chromatograms of pure protein showing a characteristic combination of two peaks (Figure 31B). Additional asymmetry in SEC chromatogram peaks is also evident with shouldering being commonly observed. Peak asymmetry and multiplicity at least in the case of peaks with small elution volume variations (and hence only a minimal apparent size difference) seems likely to arise from conformational heterogeneity. The dominant and well-spaced peaks within the chromatogram however are best explained by the presence of different oligomeric forms of HSC70.

Further characterisation of HSC70 by SEC-MALS confirms oligomerisation of HSC70 (Figure 31A), showing a monomer with a molecular weight calculated from

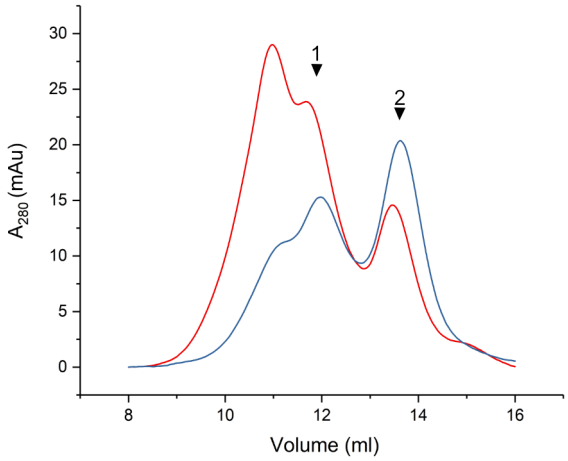


Figure 32: Comparison of the SEC profiles of HSC70 with (blue) and without (red) prior incubation (for ~30 minutes) with 2 mM ATP and 2 mM $MgCl^2$. Dimeric and monomeric species as identified by SEC-MALS (Figure 31) are labelled 2 and 1 respectivley. Experiments were performed using two identical samples of HSC70 in NMR (20 mM HEPES pH 7.2, 50 mM NaCl, 1 mM TCEP, 0.03% azide) buffer supplemented as indicated, on an Superdex S200 Increase 10/300 column.

light scattering of ~69 kDa (expected weight of 74 kDa) and a larger species with a weight of ~120 kDa (expected dimer weight of 150 kDa). Although the weight derived from light scattering for the larger species deviates significantly from the expected dimer weight it is likely that some degree of averaging between dimer and monomer is occurring as these two species do not separate completely through SEC. A further light scattering peak corresponding to a subtle shoulder on the larger of the two SEC peaks yields a mass estimate in excess of 160 kDa and seems likely to correspond to another larger oligomeric form. Extensive overlap with the adjacent dimer peak however precludes a clear assignment of its oligomeric state. This oligomeric behaviour was affected by the presence of nucleotide, in general agreement with published data [86, 87]: the addition of ATP favours the monomerisation of HSC70 and increases the height of the monomer peak while reducing the height of the larger molecular weight peaks in the chromatogram (Figure 32).

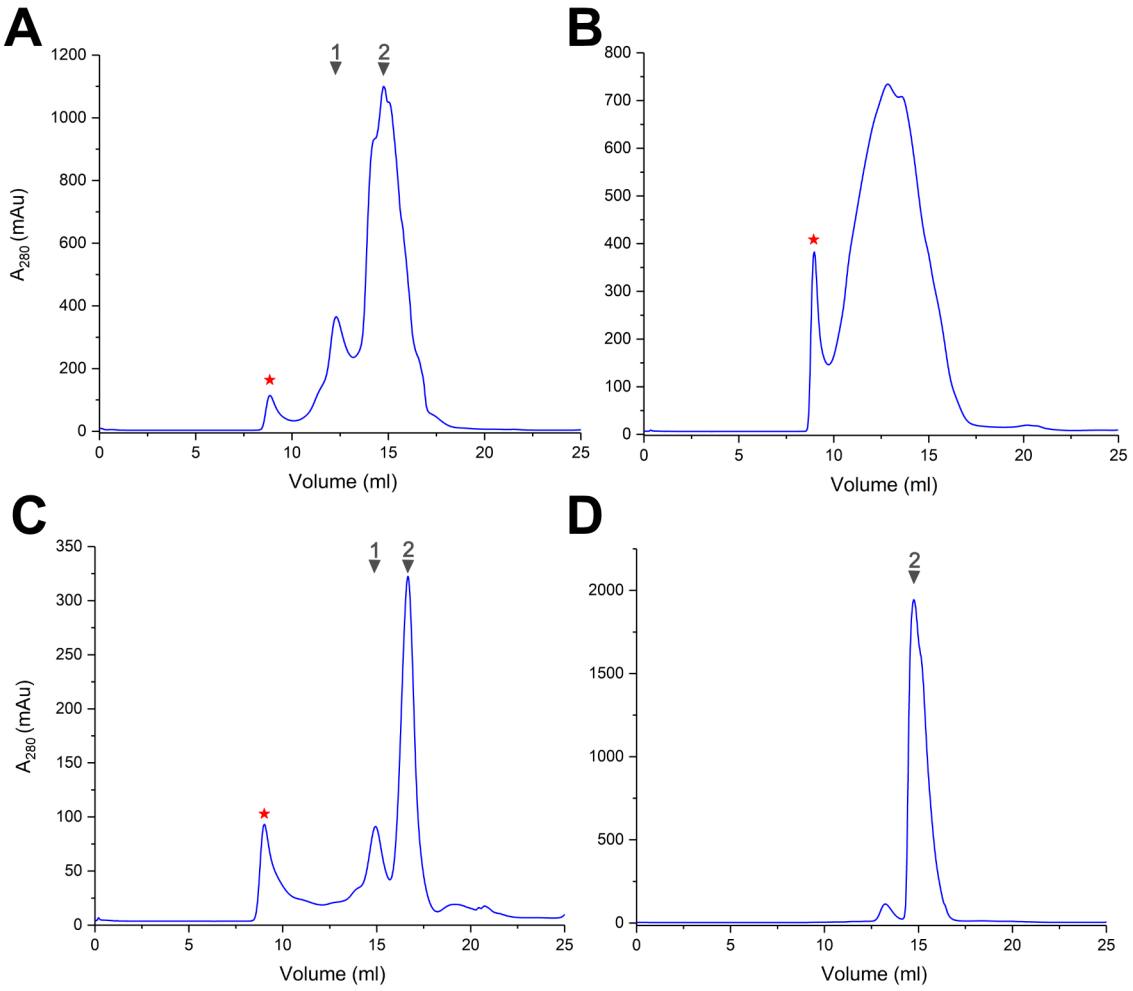


Figure 33: Comparison of the SEC profiles of several HSC70 truncation constructs. All experiments were performed in NMR buffer (20 mM HEPES pH 7.2, 50 mM NaCl, 1 mM TCEP, 0.03% azide) on an Superdex S200 Increase 10/300 column. **A:** NBDSBD (residues 1-555), **B:** SBDCTD (residues 556-646), **C:** SBD (residues 386-555), and **D:** NBD (residues 1-385). Possible dimeric and monomeric species are labelled 1 and 2 in dark grey, respectively. In the case of B, where one broad-shouldered peak is observed assigning individual species was not possible. Where applicable void volume peaks are marked with a red star.

6.1.1.1 *Truncation Constructs*

As with the full length almost all truncation constructs of HSC70 (both N and C terminal) display some degree of multiplicity in their SEC chromatograms likely stemming both from oligomerisation as well as from conformational heterogeneity (Figure 33). Of the array of HSC70 truncation constructs only the NBD displayed a clear single peak corresponding to a monomer (Figure 33D). The larger NBDSBD construct displayed a greatly reduced tendency to oligomerise relative to the full length with a similar though diminished pattern of larger species peaks being observed (Figure 33A). Expression of the core of the SBD on its own produces two distinct peaks (Figure 33C), the smaller of the two peaks almost certainly corresponds to monomeric SBD while the larger peak seems likely to correspond to a far less abundant dimeric species. The C-terminal end of the protein, the SBDCTD, produced a large broad peak indicative of polydisperse oligomerisation (Figure 33B).

In general, the behaviours observed for both the full-length protein and its truncations are in line with expectations based on published data. The C-terminus of HSP70 family proteins has generally been associated with the tendency of these proteins to oligomerise [81–84]. In our hands, removal of the C-terminus greatly diminishes oligomerisation while removal of the NBD appears to have a deregulating effect, stimulating the formation of multiple oligomeric assemblies.

6.1.2 *NMR*

Size constitutes one of the major barriers to NMR studies of proteins, with larger proteins displaying not only increased spectral complexity but also unfavourable relaxation properties which contribute to line broadening and poor signal intensity. The ~70 kDa full-length HSC70 hence presents a challenging target, with issues of its already large size being compounded by its tendency to form oligomeric assemblies. Initial tests with ^{15}N labelled HSC70 resulted in poor quality spectra not suitable for structural studies (Figure 34). Two possible approaches were explored

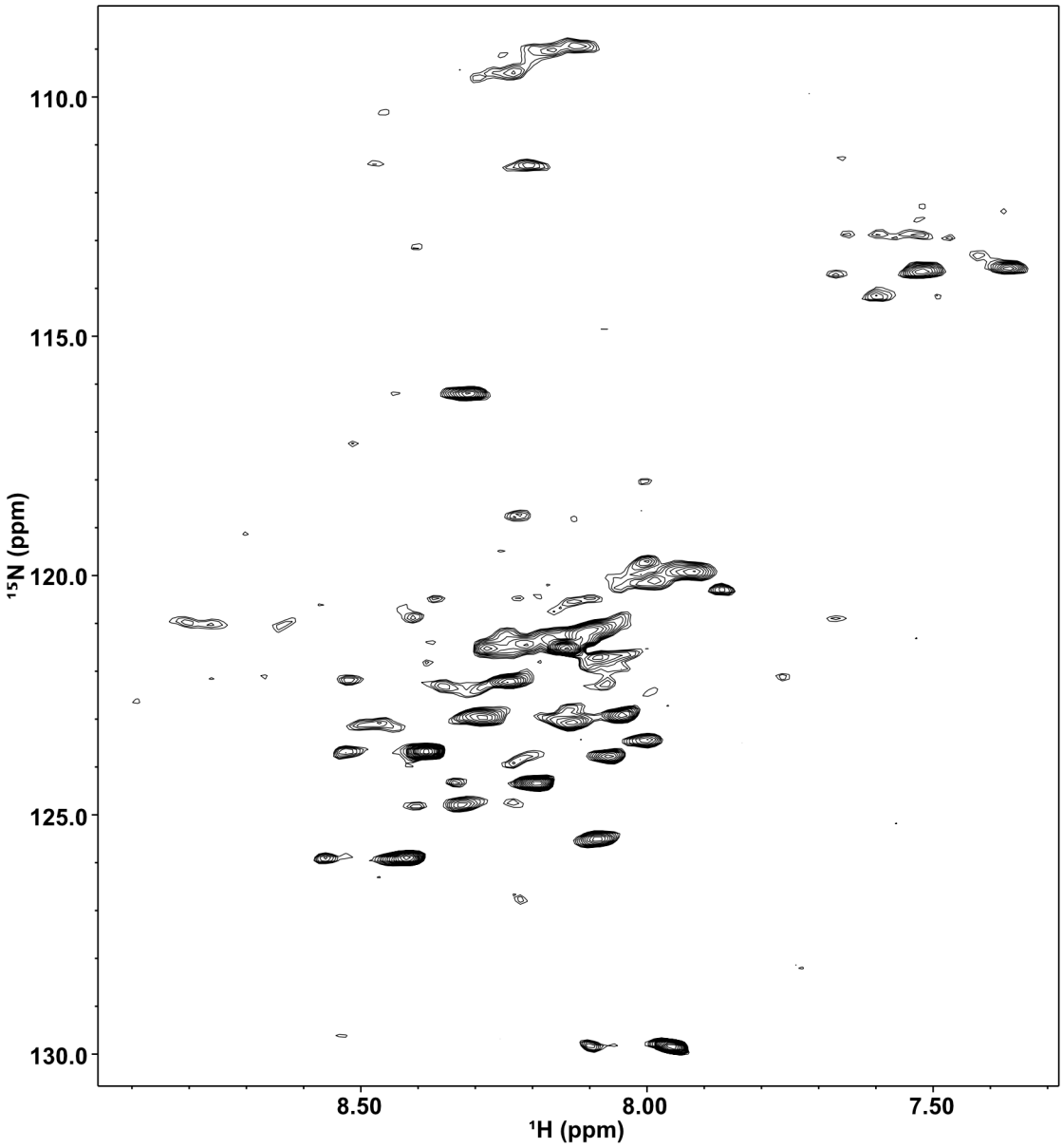


Figure 34: Initial test spectrum of uniformly ^{15}N labelled full-length HSC70 showing backbone ^{1}H ^{15}N resonances. The experiment was carried out on a 600 MHz (^{1}H) spectrometer at 35 °C using a BEST-TROSY sequence.

to circumvent this and improve spectral coverage of the protein. The first was to use the NBDSBD truncation construct since removal of the C-terminal end of the protein both reduces size down to ~65 kDa and more importantly, significantly reduces oligomerisation (see above, subsection 6.1.1). Consequently, the NBDSBD construct produces good quality NMR spectra covering most of the residues in the construct (Figure 35). The second approach was to employ a combination of deuteration, selective methyl labelling and TROSY experiments to improve the spectral quality for the full-length protein. Application of these techniques coupled with the use of a high magnetic field substantially improved spectral quality making the unaltered full-length protein amenable to study by NMR (Figure 36).

6.1.2.1 *Assigning HSC70*

Interpreting spectra of full-length HSC70 poses several challenges. The high spectral complexity and polydispersity of the sample, seen as characteristic peak streaking, complicate the clear assignment of resonances. The size of the protein also poses an obstacle for traditional triple-resonance based assignment techniques. The problem of assignment can be, at least for the ADP-bound form of HSC70, readily surmounted by exploiting the independent behaviour of its domains in this undocked conformation. Spectra both for the isolated NBD and the core of the SBD overlay well with spectra of the full-length ADP-bound protein (Figure 37). This allows for many assignments to be transferred from smaller constructs based on HSQC spectra alone. Assignment of individual domains hence provides the path of least resistance to assigning the full length. Assignments for the ADP bound NBD of HSC70 have been recently published [147], and while the SBDCTD and CTD on their own have oligomeric properties that hinder their use in NMR studies (data not shown) the core of the SBD alone (residues 386-555) generates high quality spectra (Figure 38). Use of a standard triple resonance suite of experiments could be used to assign ~90% of SBD backbone resonances. Coupled with published assignments for the NBD, theoretical coverage of ~80% of the protein can be achieved by overlays with the full-length spectrum and in practice should allow

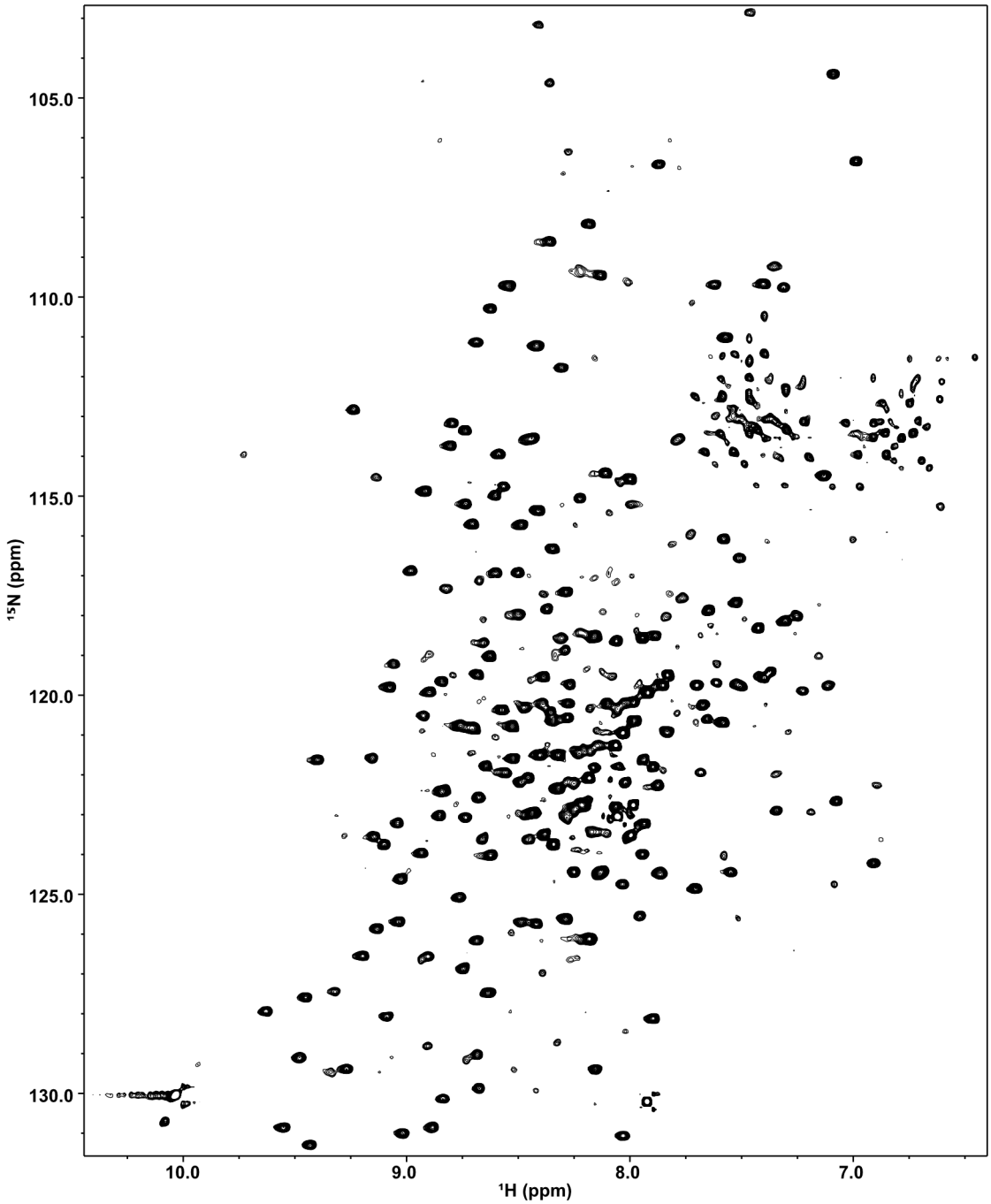


Figure 35: ^{1}H ^{15}N backbone BEST-TROSY spectrum of the NBDSBD construct (residues 1-555) collected on a 750 MHz (^{1}H) spectrometer at 30 °C.

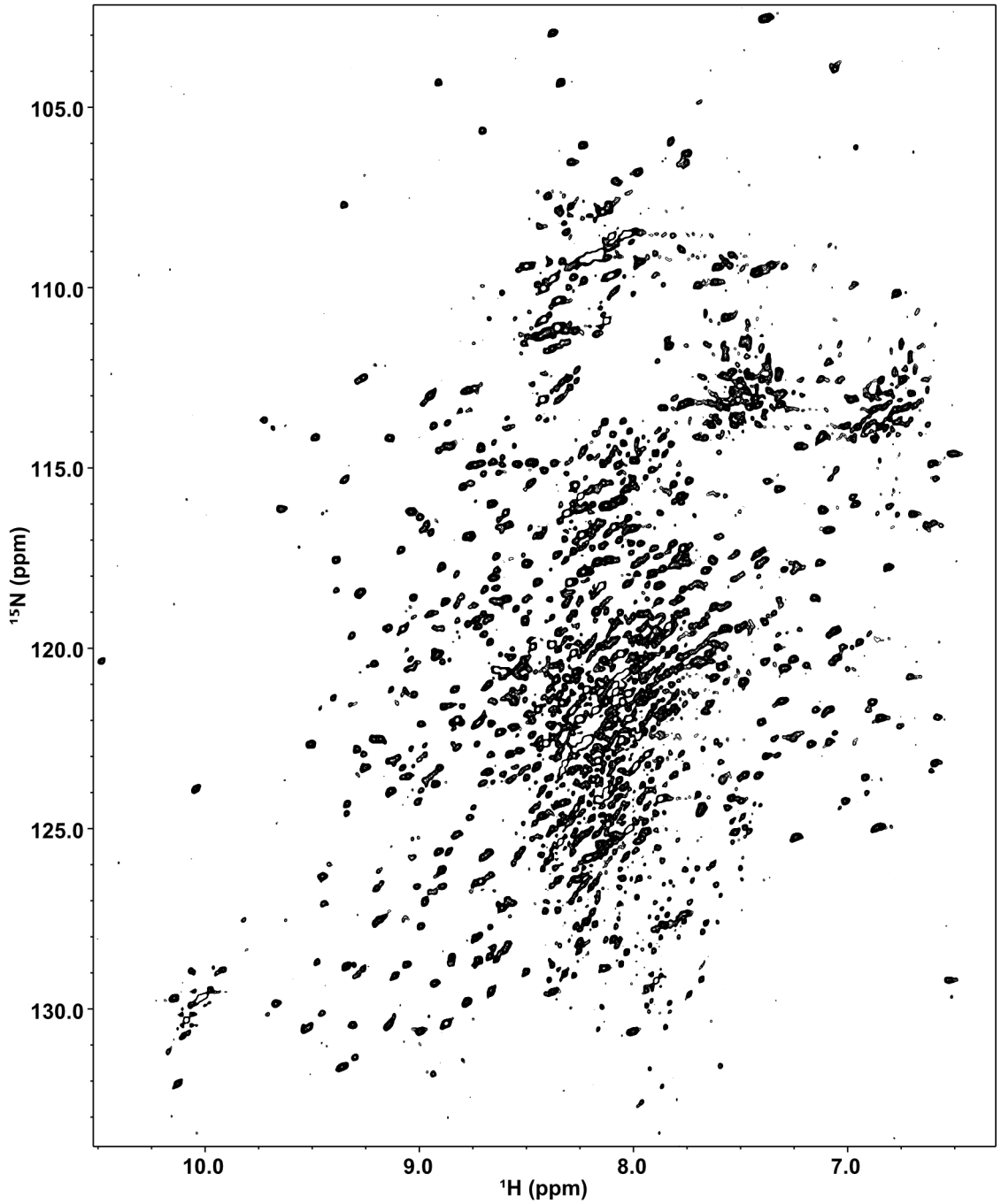


Figure 36: Backbone ^{1}H ^{15}N spectrum of deuterated, methyl labelled and ^{15}N labelled full-length HSC70. The experiment was carried out on a 950 MHz (^{1}H) spectrometer at 30 °C using a BEST-TROSY sequence.

for around 50% of residues in the full length to be assigned. While numerous avenues exist to extend the number of assigned resonances in full-length HSC70 the coverage given by a straight forward divide and conquer approach was sufficient to enable a first-pass structural analysis to be conducted.

6.2 HSP40 PROTEINS

Throughout the work carried out on the HSP70-HSP40 system the possibility of extending our model system from one involving just J-domain binding to one examining the binding of full-length HSP40 proteins was investigated. In the case of DNAJB6 this was further driven by both interest in its biology as well as the distinct lack of structural information available. The full-length versions of all three HSP40 proteins studied represent challenging targets for both expression and purification. Prior published structural work on DNAJC5 has shown the full-length protein to be highly aggregation prone and unstable even in the absence of its cysteine rich region prompting the authors to generate a minimal functional J-domain construct (residues 1-100) [141]. With DNAJC5 acting as both as a point of comparison and positive control during our studies (it being the most well characterised of the three target proteins) and with its 1-100 J-domain construct expressing and purifying in a straightforward manner in our hands we made no further attempts to acquire the full-length protein.

6.2.1 *Expression and Purification of DNAJB6 and DNAJB2*

For DNAJB2 and DNAJB6 there is a general paucity of published protein work on full-length constructs. In the case of the latter, a published protocol for the full-length protein exists which relies upon the protein being targeted to inclusion bodies when expressed in *E. coli* and a subsequent denaturing purification being employed [137]. Expression of Twin-Strep tagged constructs of full-length proteins did not result in the formation of inclusion bodies, hence a native protein prepara-

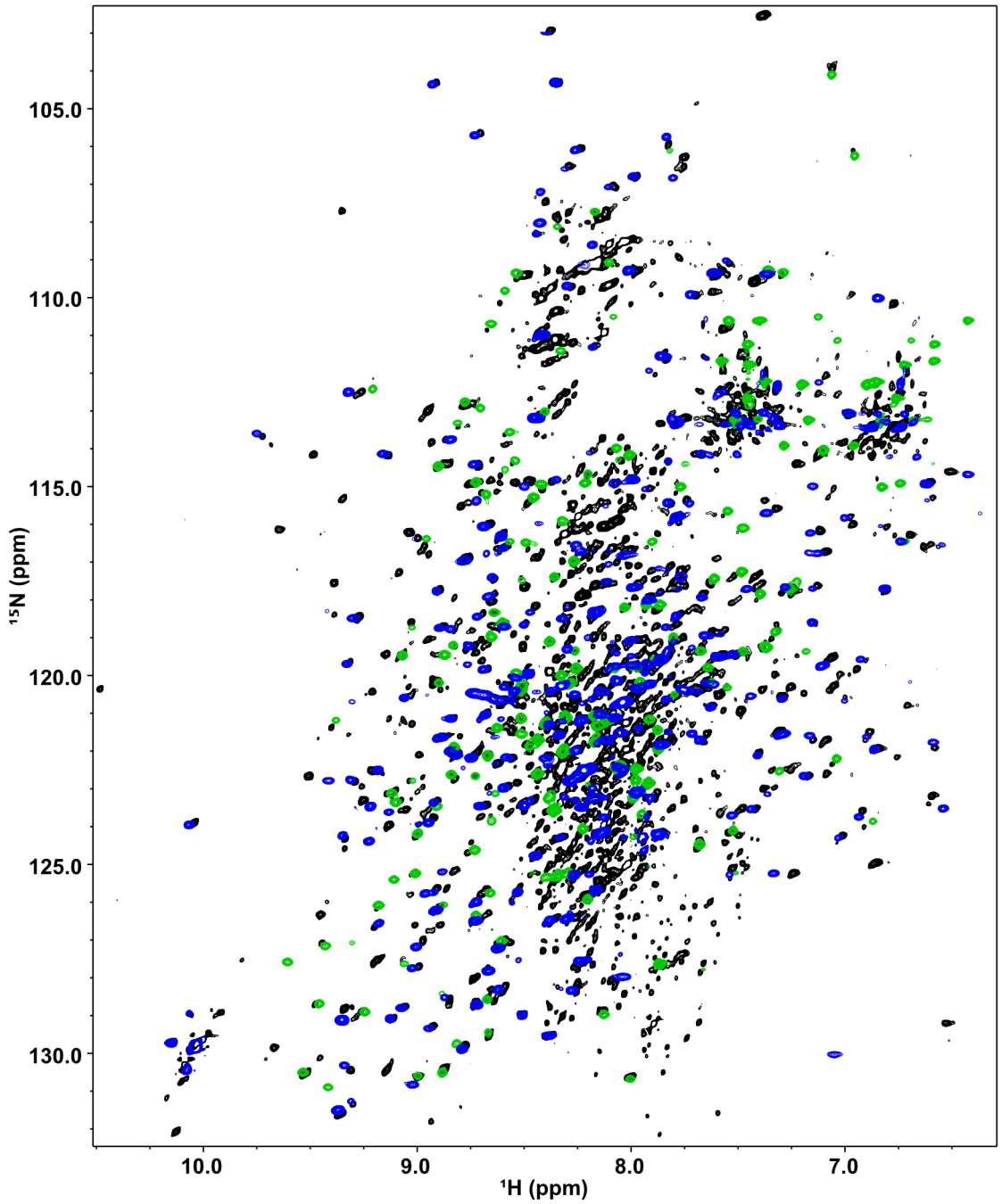


Figure 37: Overlay of spectra for the SBD (Green), NBD (Blue), and full-length HSC70 (Black). All spectra are BEST-TROSY experiments, NBD and full-length spectra were recorded at 30 °C, and the spectrum for the SBD was recorded at 25 °C. SBD and NBD spectra were collected on a 750 MHz (^{1}H) spectrometer and the full length spectrum was collected on a 950 MHz (^{1}H) spectrometer.

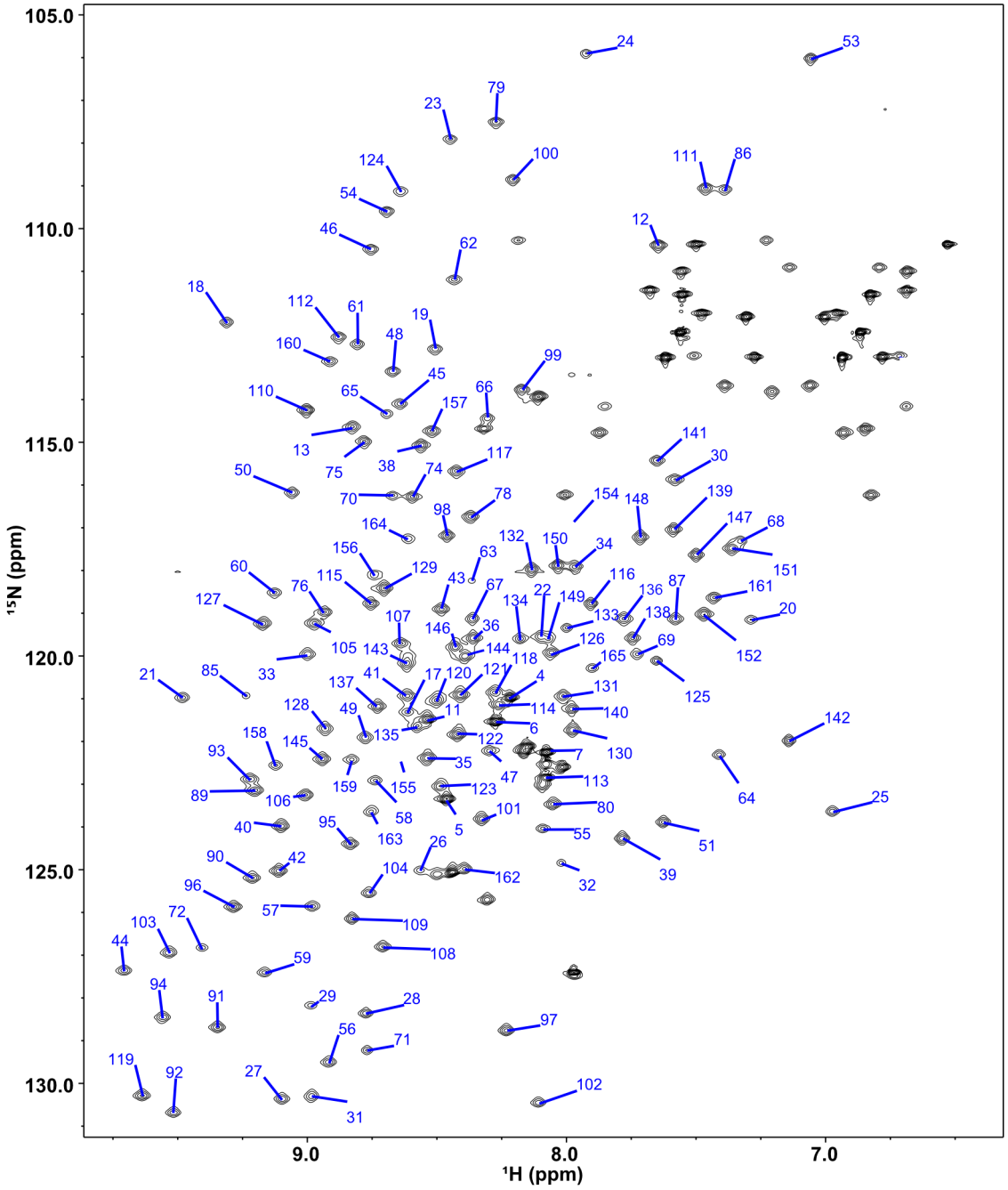


Figure 38: Backbone ^{1}H ^{15}N spectrum of the SBD (residues 386-555) shown with assignments. The spectrum was produced on a 750 MHz (^{1}H) spectrometer at 25 °C. A total of 148 out of 165 assignable backbone resonances (~90%) have been assigned using a standard suite of triple resonance based experiments.

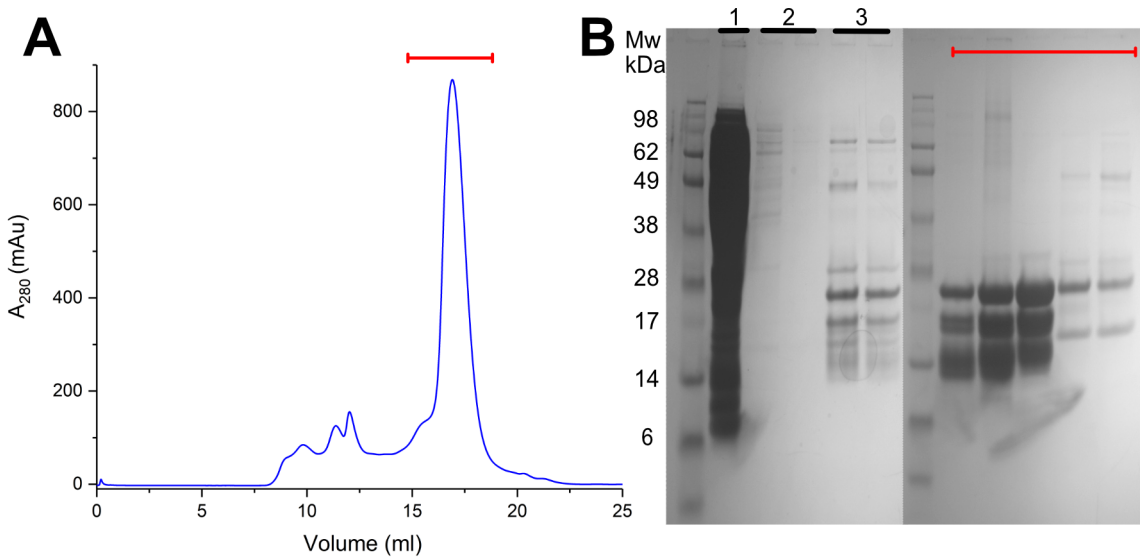


Figure 39: Example purification showing the behaviour of DNAJB6 1-200. Almost identical results were obtained for 1-156 and full-length DNAJB6 as well as full-length DNAJB2. **A:** SEC-chromatogram after purification on a Strep-Tactin column showing a dominant elution peak at ~17 ml. **B:** SDS-PAGE gels of the affinity step (*left*) and SEC step (*right*) during purification. Lanes labelled 1, 2 and 3 correspond to the affinity step flow through, washes and elutions, respectively. The red bar indicates samples taken from the region of the SEC chromatogram shown by a red bar in A. The banding pattern observed here is typical of that observed for 1-156 and full-length DNAJB6, as well as for full-length DNAJB2.

tion was initially used. Strep-Tactin column elutions for both proteins were marked by numerous impurity bands ranging in molecular weight from much higher than expected to lower, with a handful of bands being of approximately the right molecular weight for the target protein. Truncation of DNAJB6 to generate constructs 1-156, which contains the J-domain and all the G/F rich region, and 1-200 which contains the large disordered G/F rich and S/T rich regions and is missing a short C-terminal segment predicted to be structured, were tried as a means of alleviating these problems. Both these constructs however behaved in a manner identical to full-length DNAJB6 and DNAJB2 (Figure 39). This behaviour strongly suggests that DNAJB2 as well as DNAJB6 and its two truncation constructs 1-156 and 1-200 are prone to sticking to *E. coli* proteins.

Subsequent SEC of the mixture of proteins eluted from affinity columns did not produce clear results either. SEC chromatograms for the above constructs are dominated by an extremely large absorbance peak at ~17 ml retention volume. Analysis by SDS-PAGE shows that the peak is composed of a set of co-eluting bands rang-

ing in molecular weights from ~10–20 kDa. Although segregated into three major bands, the bands themselves show multiplicity (Figure 39B). Analysis of the remaining far smaller peaks in the chromatogram did not show any species that could be unambiguously assigned to full length target proteins. Attempts at mass spectrometry to further characterise the dominant peak failed as the polypeptides did not 'fly' under the conditions used.

With the extremely similar behaviour during purification of both DNAJB2 and DNAJB6 and our strong interest in full-length DNAJB6, subsequent attempts at obtaining full-length protein focused on DNAJB6. It was d eemed that if a successful approach could be developed for DNAJB6 it would likely be applicable to JB2. Two different approaches were explored to produce full-length DNAJB6. The first was to try using a GFP fusion construct, N- and C-terminal fusions were tried and resulted in no appreciable expression of the protein. Additional attempts at isolating the protein were made by performing the purification under at least partially denaturing conditions in the absence of GFP. Resuspension of cells in 2 M urea followed by a Strep-Tactin column step resulted in one clear protein band of ~32 kDa in mass by SDS-PAGE (the expected mass of DNAJB6b is 31 kDa) but several impurities remained. Dialysis of the Strep-Tactin column elutions to remove urea proved successful without any protein precipitation though substantial precipitation issues were encountered when concentrating the protein for SEC.

The observed SEC chromatogram (Figure 40A) showed signs of substantial protein aggregation with a large void volume peak; furthermore, full-length DNAJB6 showed no clear elution profile being found across the whole chromatogram from the void onwards (Figure 40B). Elution of DNAJB6 between ~9 ml (overlapping with the void volume peak) and ~14 ml occurred as a series of bumps most likely representing multiple different oligomeric species, a result congruent with published data on the isolated full-length protein where polydisperse oligomerisation resulting in at least one case in oligomers of up to 540 kDa has been observed [137, 139]. Such a size is at the limits of resolution for an S200 column and explains the elution of DNAJB6 overlapping with the void volume peak. As with previous

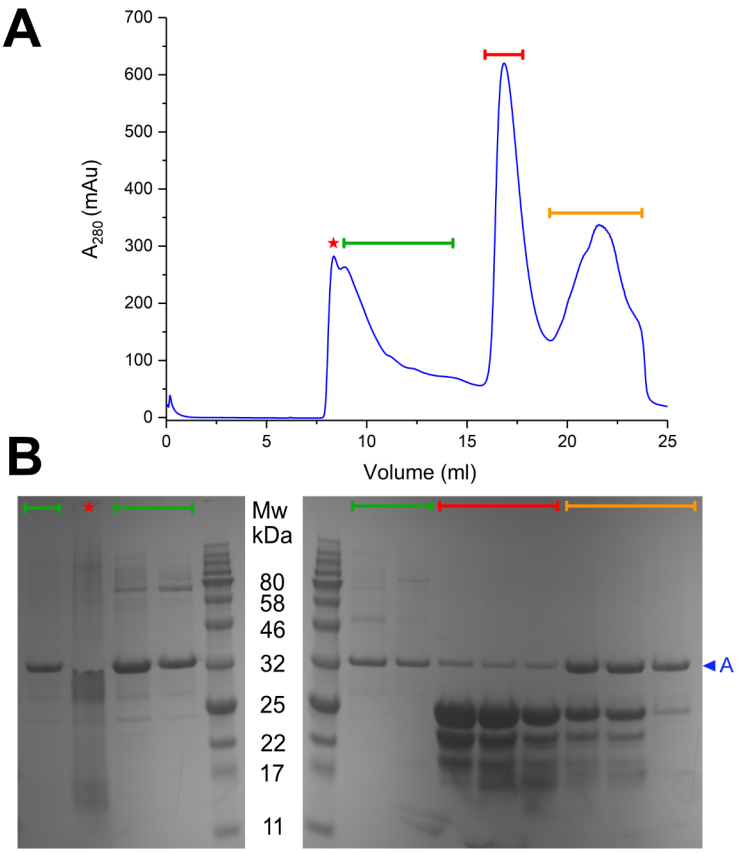


Figure 40: Chromatogram and SDS-PAGE results obtained from full-length DNAJB6 purified in the presence of 2 *M* urea. **A:** SEC chromatogram, in which the region indicated by the green bar appears to correspond to a series of polydisperse peaks of different oligomeric species of DNAJB6. The region indicated by the orange bar elutes after the salt peak and shows sticking of the protein to the column resin most likely as a result of aggregation. The void volume peak is indicated by a red star. **B:** SDS-PAGE analysis of the corresponding SEC chromatogram shown in A with the regions of the chromatogram analysed indicated by coloured bars. The red star shows the fraction from the void volume peak corresponding to aggregated protein, and the band corresponding to full-length DNAJB6 is labelled A. The peak indicated by the red bar in the chromatogram in A, corresponds to the dominant peak observed in previous native preparations of full-length DNAJB6. As with previous experiments the same banding pattern is observed for this peak, though in this case a clear band corresponding to full-length protein can be seen.

attempts the SEC profile remained dominated by one large peak at ~17 ml corresponding to the characteristic collection of three bands seen previously, though this time, co-elution with the larger full-length band could be seen. A further peak at around one column volume was also produced and contained a large amount of full-length DNAJB6. This very late elution of the protein seems to indicate sticking to the column, likely due to protein aggregation, and in most fractions containing full-length DNAJB6 precipitation was clearly visible.

6.2.2 *J-Domain Constructs*

With the difficulties encountered in isolating full length DNAJB6 and DNAJB2 and for the sake of expedience we turned again to the old maxim of divide and conquer seeking to produce constructs of the J-domains of our target proteins in isolation. In the case of DNAJB2, prior published work has shown that a minimal J-domain construct consisting of residues 1-99 will express in *E. coli* and can bind to HSP70 proteins [146]. Expression and purification of this construct proved straightforward and as for the DNAJC5 1-100 construct yields a clear monodisperse peak after SEC. For DNAJB6 our data had already shown that truncation down to 1-156 was insufficient to surmount problems at the purification stage (see above subsection 6.2.1). Initial testing was therefore performed using a minimal J-domain construct, 1-77, this behaved well and expressed at substantial levels producing a monodisperse and clear SEC peak. The result from the purification of 1-77 placed the longest easily obtainable construct of DNAJB6 somewhere between 1-77 and 1-156. A further run of purifications was hence performed on two more constructs, 1-100 and 1-130, the latter of these adds a significant portion of the G/F rich region of the protein. Both these constructs behaved much like the isolated J-domain though the 1-130 construct presented both lower yields and a tendency to degrade towards a mass consistent with 1-100, presumably due to the loss of C-terminal unstructured residues. Transfer of the latter construct boundary to DNAJB2 also proved successful with DNAJB2 1-130 expressing and behaving in a similar fashion to DNAJB6

1-130. With this protein work in hand we were able to produce a range of tractable constructs of all three target HSP40 proteins DNAJC5 1-100, DNAJB6 1-77, 100 and 130, and DNAJB2 1-100 and 130 (Figure 41).

The work outlined above (subsection 6.2.1 and subsection 6.2.2) on DNAJB6 provides an interesting functional inference. That the 1-156 construct sticks to *E. coli* proteins in a similar manner as 1-200 and full-length while the 1-130 construct does not suggest a protein binding role for the middle region of residues 131-156.

6.2.3 *NMR and Assignments of DNAJC5 and DNAJB6*

Having produced constructs for our target J-proteins, efforts turned to the characterisation of these constructs by NMR to lay the ground work for being able to monitor the HSP70-HSP40 interaction from either side. This work focused on both DNAJC5 and DNAJB6; the former representing a well characterised HSC70 interaction partner which could serve as a positive control for future work and the latter representing a more biologically interesting and structurally unknown target. ^{15}N labelled samples were produced for DNAJC5 and all three constructs of DNAJB6. The DNAJC5 1-100 spectrum obtained closely matches previously published data and assignment transfer was performed in a straight forward manner based on published backbone shifts [141] and our HSQC (Figure 42). For DNAJB6, assignments for the longest construct (1-130) were performed using a triple resonance-based strategy with 100 out of 124 assignable backbone resonances being assigned (Figure 43). Comparison of the three truncation constructs studied for DNAJB6 shows a striking difference between the spectra for 1-77 and 1-100 when compared to 1-130 (Figure 44), extensive peak shifts occur across the spectrum and prevented the straightforward transfer of assignments from 1-130 to the shorter DNAJB6 constructs. The addition of these 30 residues substantially alters the majority of backbone resonances throughout residues 1-100 of DNAJB6 indicative of a significant structural change in this region of the protein. This data, suggestive of an interaction between the G/F rich region and the J-domain is backed by prior work on a

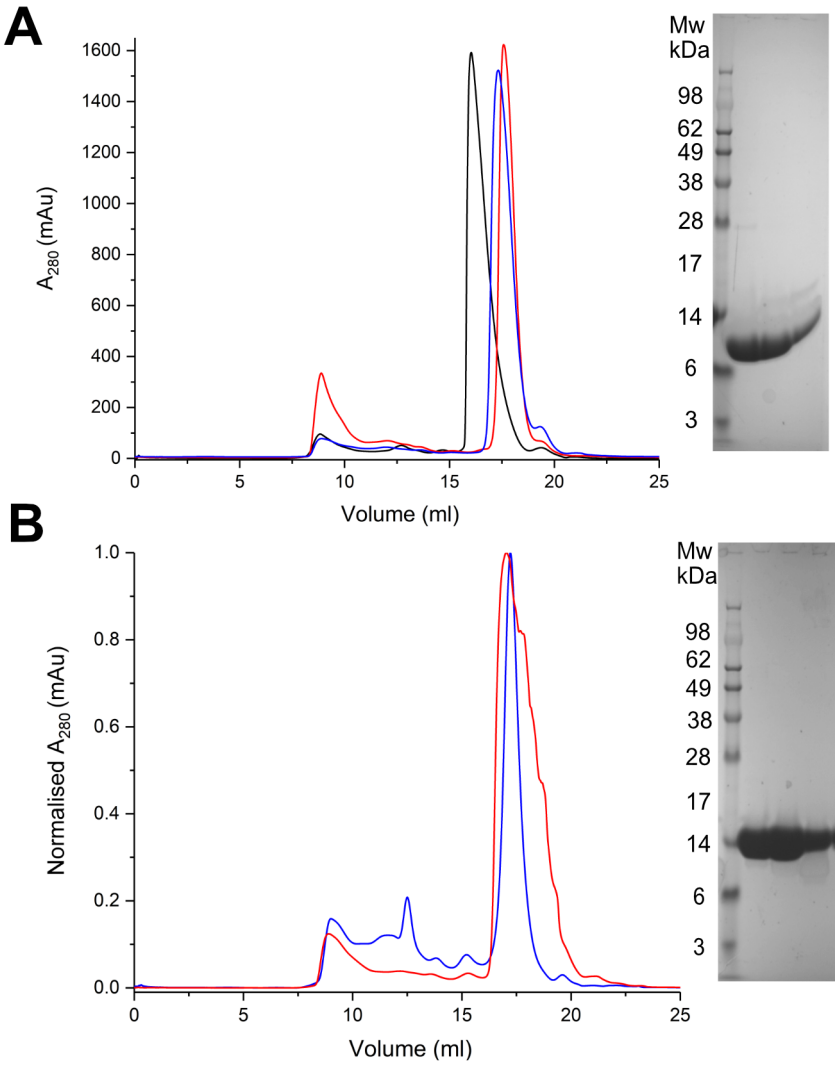


Figure 41: Example SEC chromatograms and SDS-PAGE results for the purification of isolated J-domain constructs corresponding to residues 1-100 and 1-130 of DNAJB2 and DNAJB6 and residues 1-100 of DNAJC5. **A:** Overlay of SEC chromatograms for 1-100 constructs of DNAJC5 (Black), DNAJB6 (Blue), and DNAJB2(Red). An example gel is shown for the peak of DNAJB2 1-100. **B:** Overlay of SEC chromatograms for 1-130 constructs of DNAJB6 (Blue) and DNAJB2 (Red). The absorbances have been normalised for clarity as the DNAJB2 construct is more over-expressed than the DNAJB6 construct. An example gel is shown for the DNAJB6 1-130 peak.

J-protein from a thermophilic organism which shows the G/F rich region packing in a structured fashion onto the J-domain [175]. Interestingly our data shows that many of the peaks gained on extending from 1-100 to 1-130 occur towards the centre of the spectrum and have poor dispersion suggestive of disorder. Comparison with the crystal structure further validates our conclusions based on the 1-130 truncation as it shows that aside from J-domain there are no stabilising intermolecular contacts with the G/F rich region, supporting both the notion of a largely disordered G/F rich region and that the disorder we observe is not a truncation artefact. Additionally, during purification C-terminal degradation was seen for 1-130 that was not present in shorter constructs further supporting the notion that these G/F rich region residues are dynamic and disordered. This suggests that a dynamic interaction between the J-domain and the G/F rich region may operate in J-domains which are not adapted to function at high temperatures.

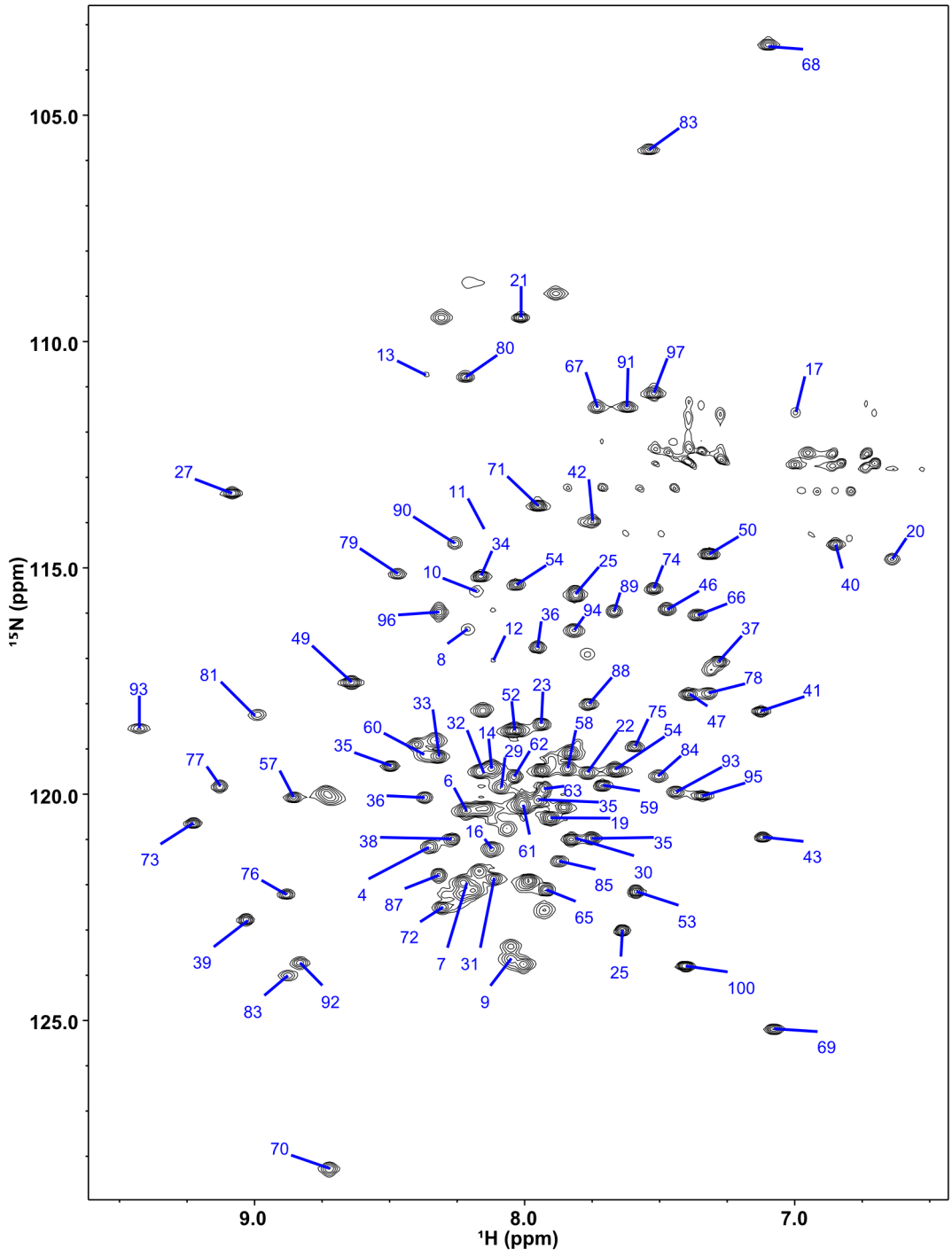


Figure 42: Backbone ^{1}H ^{15}N BEST-HSQC spectrum of the J-domain of DNAJC5 (residues 1-100) shown with assignments. The experiment was collected on a 750 MHz (^{1}H) spectrometer at 25 °C. Assignments were transferred to our conditions from work published by Patel et al. [141] based on backbone resonances.

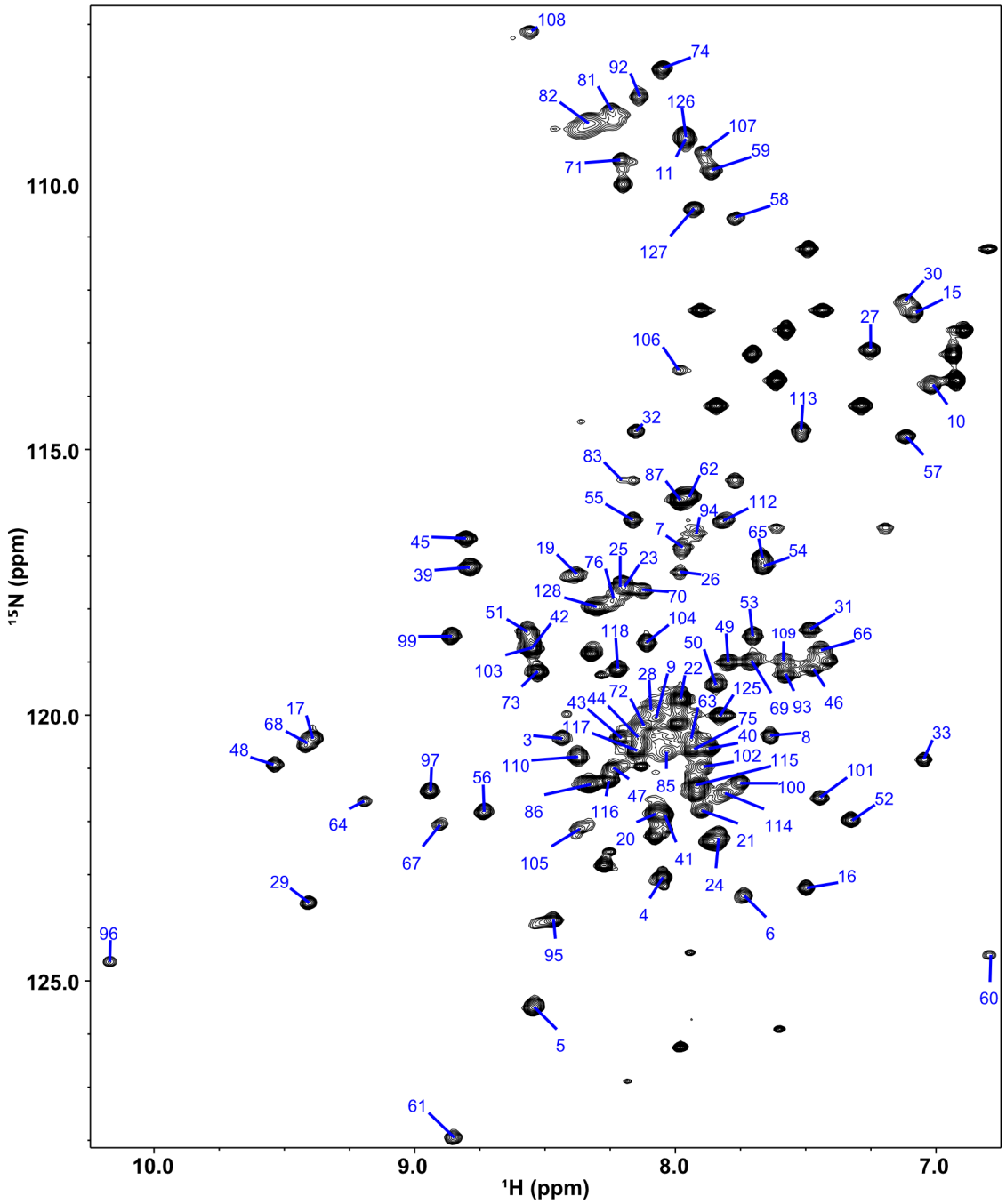


Figure 43: Backbone ^{1}H ^{15}N BEST-HSQC spectrum of the J-domain of DNAJB6 (residues 1-130) shown with assignments. The experiment was collected on a 750 MHz (^{1}H) spectrometer at 25 °C. A total of 100 out of 124 assignable backbone resonances (~80%) have been assigned using a standard suite of triple resonance based experiments.

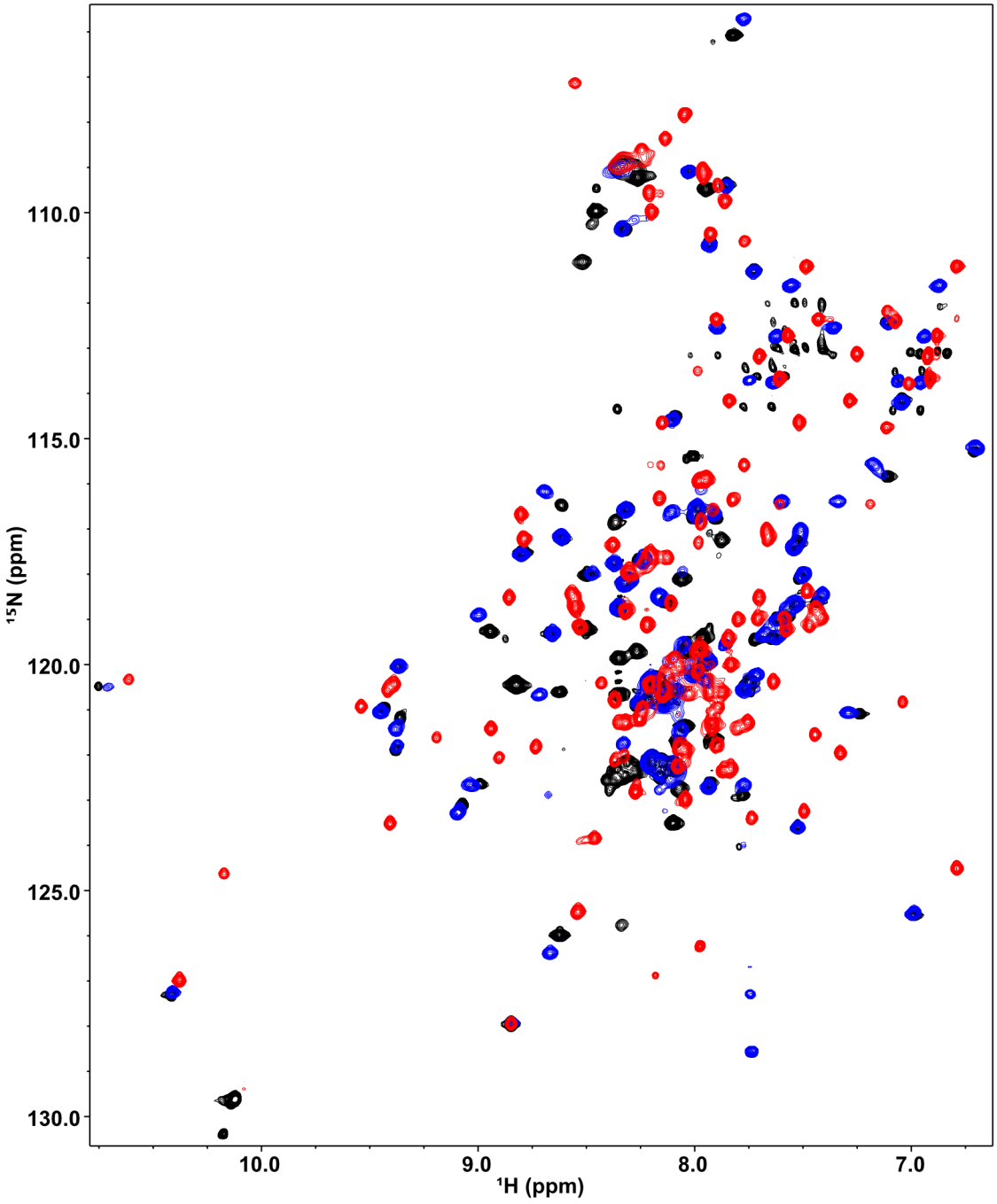


Figure 44: Comparison of ^{1}H ^{15}N resonances for three different DNAJB6 J-domain constructs: 1-77 (Black), 1-100 (Blue) and 1-130 (Red). Experiments are all collected on a 750 MHz(^{1}H) spectrometer at 25 °C using a BEST-HSQC pulse sequence.

7

ANALYSIS OF HSP40-HSP70 INTERACTIONS AND THEIR SPECIFICITY

Comprehending the regulation of the canonical HSP70 cycle presented earlier in Figure 6 requires not only the study of its components in isolation but also of their actions in concert. Our work focused on the interaction between HSP40 J-domains and HSP70 that stimulates HSP70 ATPase activity and is central to this proteostatic hub, but remains incompletely understood. While there are numerous experimental approaches which could be taken to demonstrate and analyse a protein-protein interaction, NMR spectroscopy is a particularly useful approach to study transient interactions of this ilk. NMR measurements were complemented by the use of ATPase assays, which provided a straightforward method for measuring the stimulatory efficacy of the interaction.

7.1 INTERACTION BY SEC

As both DNAJC5 J-domain and full-length HSC70 were readily obtainable, an early attempt was made at isolating a complex of the two proteins: both proteins were mixed in the presence of ATP and subjected to SEC. No stable co-complex of the proteins could be isolated by these means, and most of the J-domain eluted independently of HSC70 as a dominant peak around 15 ml retention volume (Figure 45A). Interestingly, analysis of HSC70 peak fractions by SDS page showed a small though detectable amount of DNAJC5 J-domain across these fractions (Figure 45B). Elution of DNAJC5 in this volume range does not normally occur, and analysis of elution fractions of an identical HSC70 sample alone rules out the presence of a contaminant with similar molecular weight to the J-Domain of DNAJC5

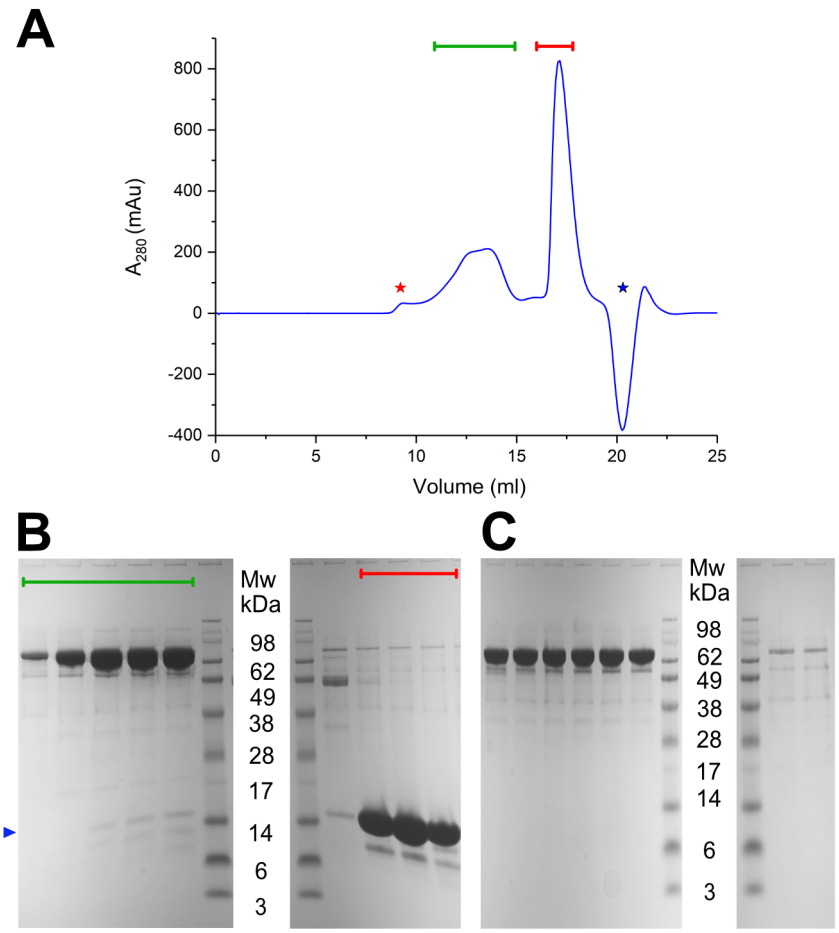


Figure 45: Analysis of the HSC70–DNAJC5 J-domain (Residues 1-100) interaction by SEC. **A:** SEC chromatogram for a high concentration mixture of HSC70 and DNAJC5 J-domain. SEC was performed in NMR buffer (HEPES p.H 7.2, 50 mM NaCl) supplemented with 2 mM ADP and 2 mM $MgCl^{2+}$. The green bar marks the peaks associated with HSC70 while the red bar marks the peak associated with DNAJC5 J-domain. The void peak is marked by a red star. The negative peak marked by a blue star corresponds to a peak that arises from a band of buffer containing less ADP than the rest. **B:** SDS-PAGE analysis of the SEC chromatogram shown in A with coloured bars corresponding to the regions marked in A. The band corresponding to DNAJC5 J-domain is marked by a blue arrow. **C:** Control gels from chromatograms of HSC70 (Left) and DNAJC5 J-domain (Right) on their own. The region corresponding to the green bar in A has been sampled for both.

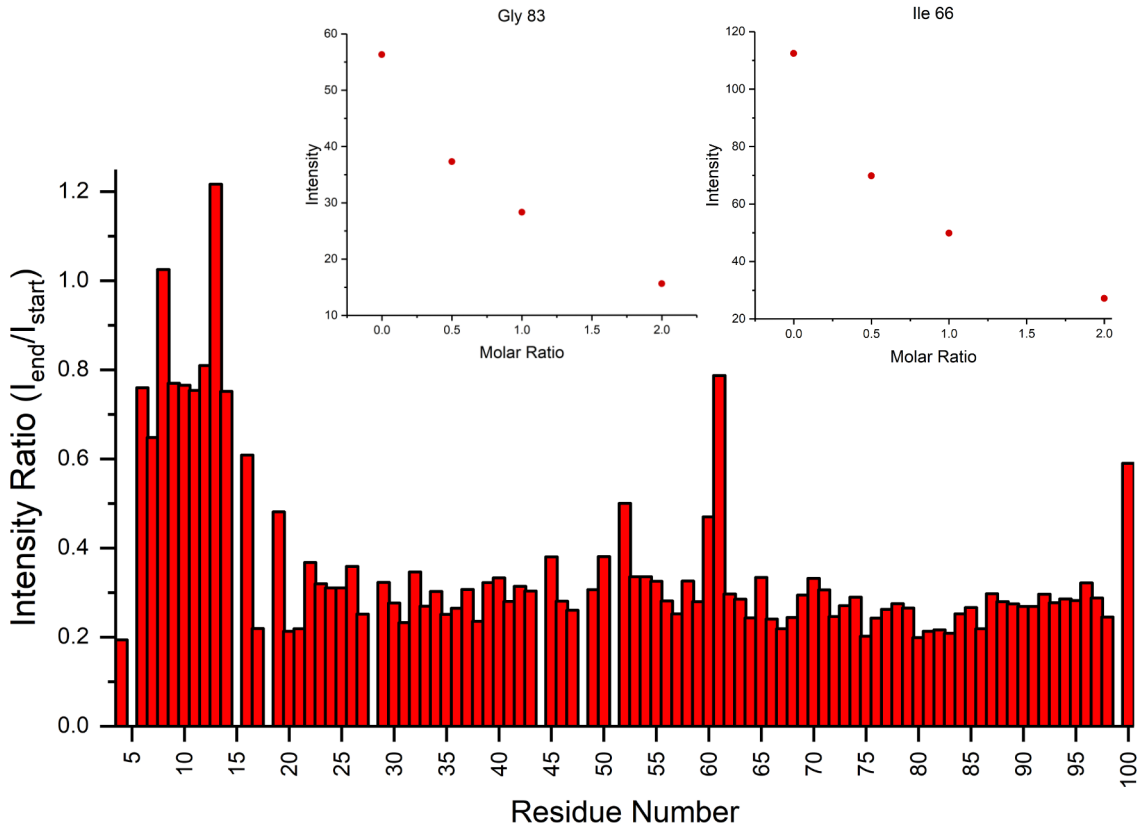


Figure 46: Analysis of the NMR titration of unlabelled full-length HSC70 into a sample of ^{15}N-labelled DNAJC5 J-domain (residues 1-100) in the presence of 2 mM ADP and 2 mM $MgCl^{2+}$. ^{1}H ^{15}N backbone resonances were followed using a BEST-TROSY pulse sequence on a 750 MHz (^{1}H) spectrometer at 25 °C . Data was collected at molar ratios of 0, 0.5, 1 and 2 DNAJC5 J-domain:HSC70. The bar chart shows the change in intensity for each residue, corrected for dilution, between the start of the titration and the end (molar ratio of 2). Two inset scatter graphs show the dilution corrected intensity for two sample residues across the entire titration. The intensity points indicate saturable behaviour.

(Figure 45C). The partial co-elution of DNAJC5 J-domain with HSC70 both hints at an interaction and suggests that any such interaction is likely low affinity and hence dilution as well as rapid exchange during SEC prevent the complex of the two proteins from being isolated. Though this behaviour poses a challenge for structural and biophysical characterisation it is not at all unexpected and falls in line with previous reports on HSP40-HSP70 interactions [106–109].

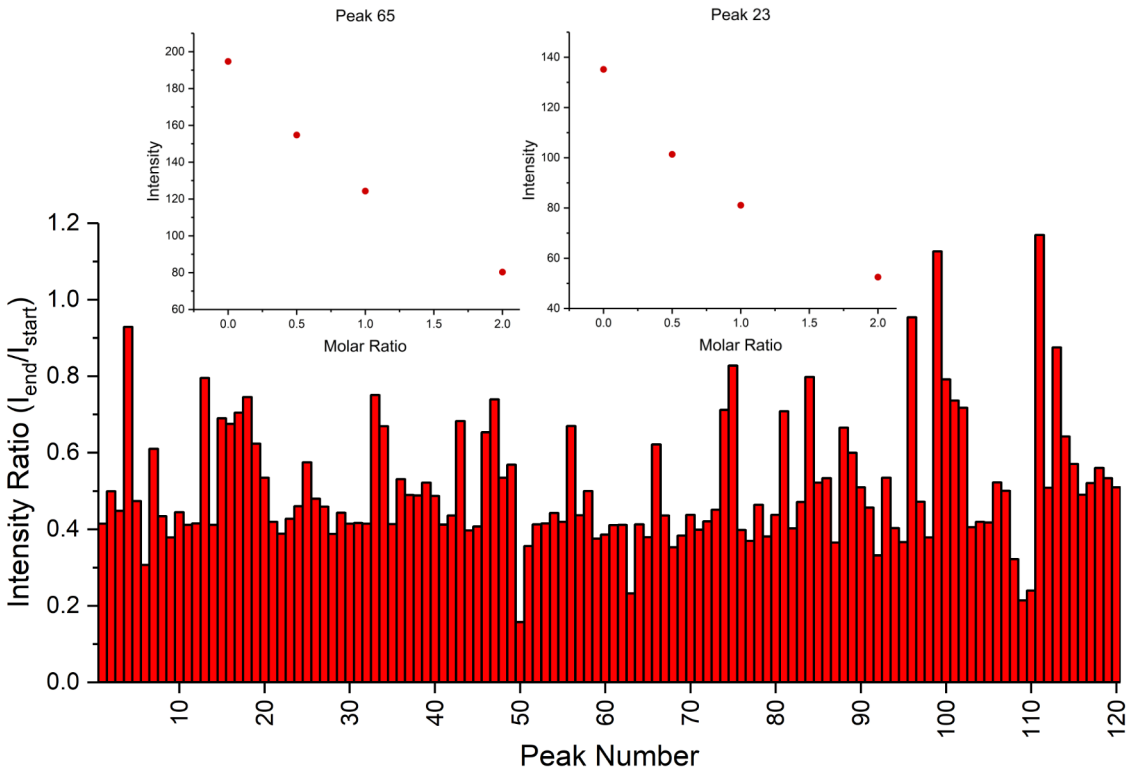


Figure 47: Analysis of the NMR titration of unlabelled full-length HSC70 into a sample of ^{15}N-labelled DNAJB6 J-domain (residues 1-77) in the presence of 2 mM ADP and 2 mM MgCl2+. ^{1}NH ^{15}N backbone resonances were followed using a BEST-TROSY pulse sequence on a 750 MHz (^{1}H) spectrometer at 25 °C . Data was collected at molar ratios of 0, 0.5, 1 and 2 DNAJB6 J-domain:HSC70. The bar chart shows the change in intensity for each residue corrected for dilution, between the start of the titration and the end (molar ratio of 2). Two inset scatter graphs show the dilution corrected intensity for two sample peaks across the entire titration. The intensity points indicate saturable behaviour.

7.2 BINDING BY NMR: OBSERVING THE J-DOMAIN

NMR binding experiments were first carried out monitoring ^{15}N labelled J-domain constructs of DNAJC5 1-100 and DNAJB6 1-77 upon the addition of unlabelled full-length HSC70. In neither case were any chemical shifts observed during these experiments though a sharp drop in the signal for backbone resonances of both J-domains did result from the addition of HSC70 (Figure 46 and Figure 47). In both cases the drop-in signal could not be completely accounted for by dilution from the titration procedure. For DNAJC5 1-100 a generalised drop can be seen across the entire J-domain although the N-terminus between residues 5 and 15 sees a fall in intensity that is on average half that experienced throughout the rest of the protein. For DNAJB6 1-77 the fall in intensity is fairly uniform across the peaks analysed with a handful of peaks showing a substantially smaller drop than the rest, though in this case a lack of assignments for this shorter DNAJB6 construct (discussed in subsection 6.2.3) prevents any structural analysis. The intensity drop for DNAJB6 1-77 is slightly less pronounced than for DNAJC5 1-100 with resonances falling to ~40% initial intensity for the former and ~30% initial intensity for the latter. This generalised drop of intensity could reflect either the increase in mass on binding to HSC70 or an unfavourable exchange regime for the complex or a combination of both. A similar behaviour has been reported previously for the interaction of bacterial HSP40 and DNAK [109].

7.3 BINDING BY NMR: OBSERVING HSC70

Our primary interest was in analysing the interaction from the perspective of the HSP70 protein. While the fuzzy nature of HSP40-HSP70 complexes has been previously examined [109], and models have been proposed for the interaction [9, 109], there is little data about the internal allostery of HSP70 in these complexes which transmits the effects of J-domain binding to the active site of the NBD.

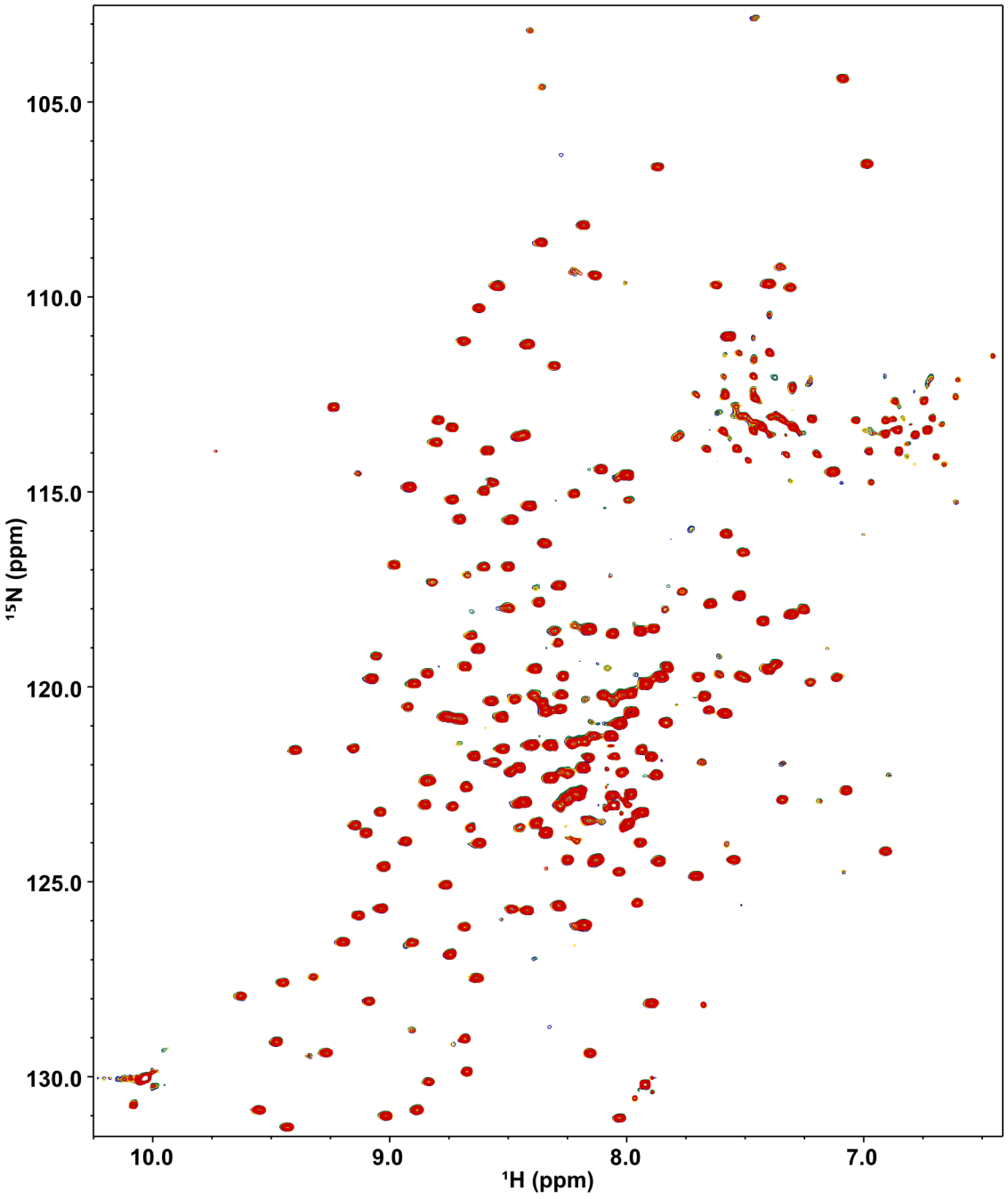


Figure 48: Titration of unlabelled DNAJC5 J-domain (residues 1-100) into ^{15}N labelled ADP bound HSC70 NBDSBD (residues 1-555). Backbone ^{1}H ^{15}N resonances were followed using a BEST-TROSY pulse sequence collected at 750 MHz (^{1}H) and 30 °C. Spectra are shown for molar ratios of 0, 0.5, 1 and 2 in blue, green, yellow and red, respectively.

Mindful of the potential issues that could be encountered when working with a large protein with a propensity for polydisperse oligomerisation we first selected the NBDSBD construct for NMR studies. With most evidence indicating that the lynch pin of the HSP40-HSP70 interaction lies in the interaction of the J-domain with the NBD, we speculated that this truncation construct with more favourable oligomeric properties subsubsection 6.1.1.1 and spectra subsection 6.1.2 may be sufficient for J-domain binding and should contain enough of the protein, including the crucial interdomain linker, to allow for a detailed analysis of the allostery involved during binding. Titration of DNAJC5 into HSC70-NBDSBD produced no spectral changes (Figure 48). Taken together with previous titration data indicating an interaction with full-length HSC70 under otherwise identical conditions this suggests that the NBDSBD construct alone contains regions that are necessary for binding, but are insufficient on their own for the interaction. The requirement of the C-terminus for a productive interaction between HSP40 proteins and HSP70 proteins has been reported before in the context of DNAJB2 where a similar set of experiments showed that DNAJB2 binding is diminished when HSP70 is truncated to a construct consisting of residues 1-565 [146]. Our NBDSBD construct corresponds to residues 1-555.

With confirmation that the full-length protein is necessary for an interaction between the J-domain and HSC70 we turned to a full-length sample produced with selective methyl protonation and deuteration at other non-exchangeable sites. Titration of unlabelled DNAJC5 into such a sample produced clear spectral changes that appeared to saturate at a molar ratio above 1.5 (Figure 49 and Figure 50) corresponding to a concentration of J-domain of 310 µM or higher. This is in line with expectations based on literature reports of K_ds for the HSP40-HSP70 interaction which are usually in the low to mid micromolar range [176]. A range of different exchange regimes (slow, intermediate, and fast) were observed depending on the magnitude of the chemical shift difference indicating that binding was occurring close to the NMR timescale. Several shifting peaks were easily identifiable using spectral overlays with previously characterised HSC70 truncation constructs

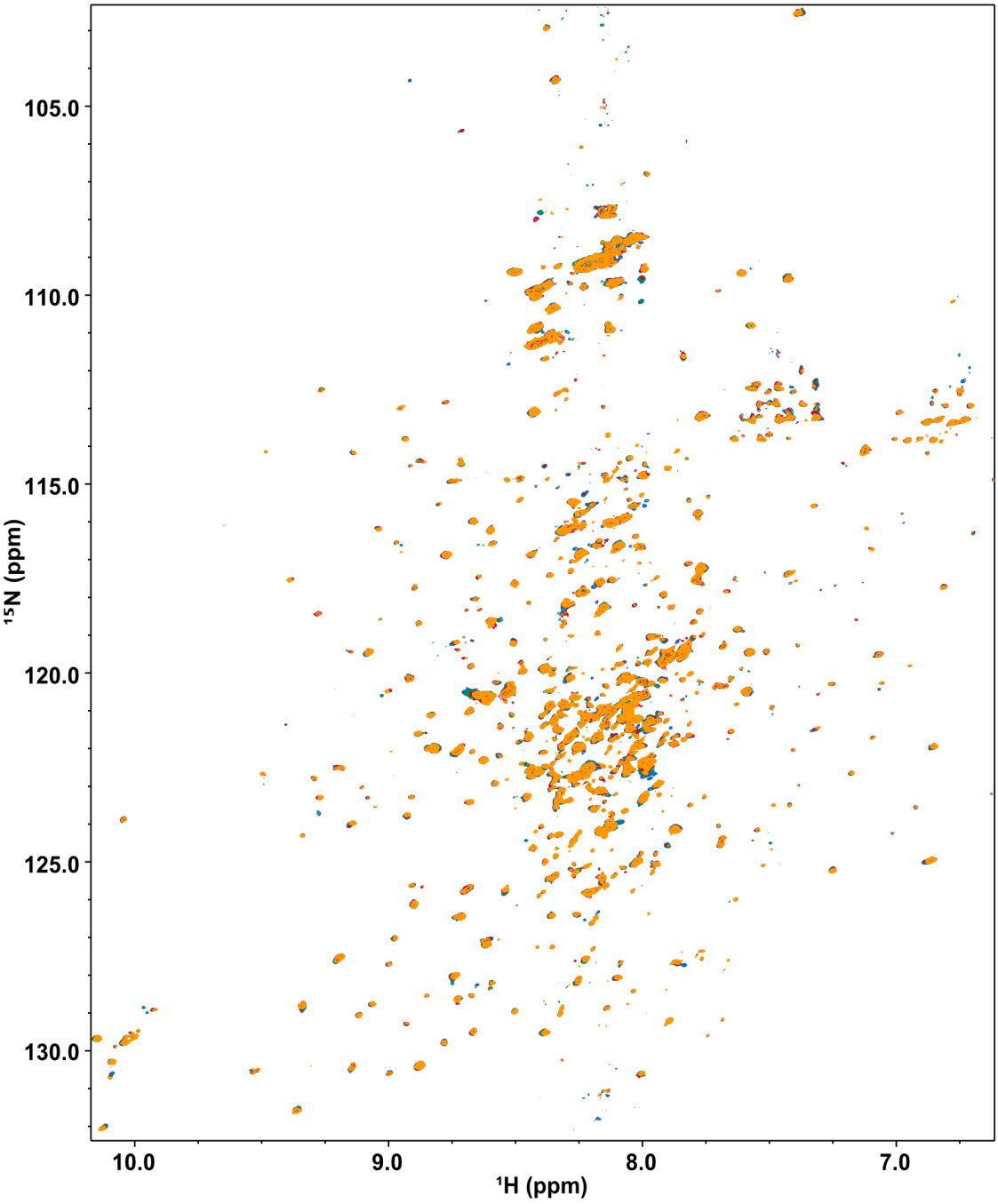


Figure 49: Titration of unlabelled DNAJC5 J-domain (residues 1-100) into deuterated, ^{15}N and selectively methyl labelled ADP bound HSC70. Backbone ^{1}H ^{15}N resonances were followed using a BEST-TROSY pulse sequence collected at 950 MHz (^{1}H) and 30 °C. Spectra are shown for molar ratios of 0, 0.75, 1.5 and 2.5 in blue, green, red and orange, respectively, corresponding to J-domain concentrations of 0 µM, 157.5 µM, 310 µM and 420 µM.

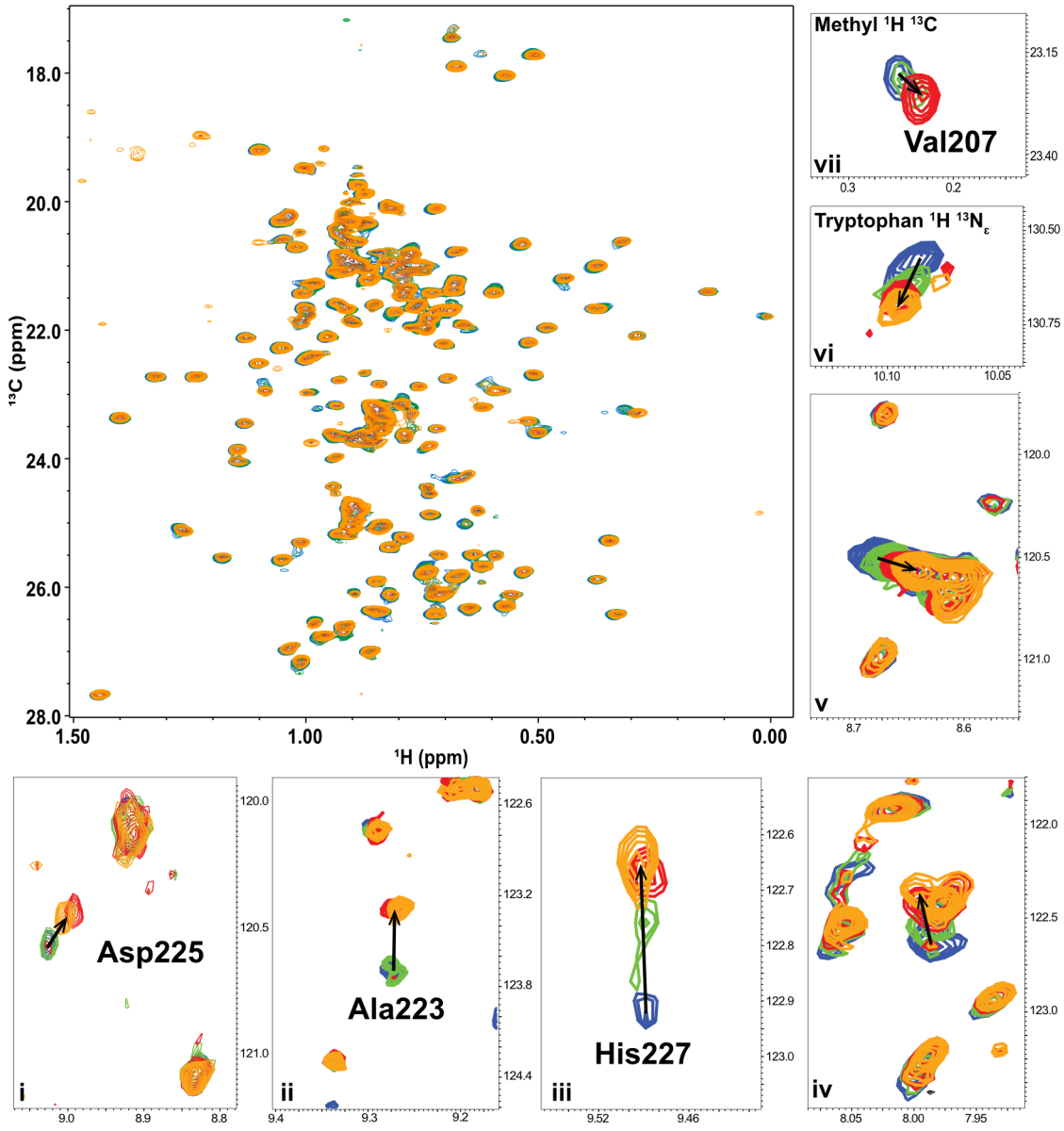


Figure 50: Methyl spectrum of full-length deuterated ^{15}N and selectively methyl labelled ADP bound HSC70. Spectra were collected using a ^{1}H ^{13}C SOFAST HMQC pulse sequence on a 950 MHz (^{1}H) spectrometer at 30 °C. Titration points shown in blue, green, red and orange, correspond to molar ratios of 0, 0.75, 1.5 and 2 respectively, corresponding to J-domain concentrations of 0 µM, 157.5 µM, 310 µM and 420 µM. The panels along the side show specific spectral changes in both the ^{1}H ^{15}N TROSY spectra backbone (panels i, ii, iii and iv) and tryptophan indole resonances (panel vi) as well as changes in the methyl spectra (panel vii). Panels i, iv, v and vi show peaks in fast exchange while panel iii shows intermediate exchange and panels ii and vii show slow exchange.

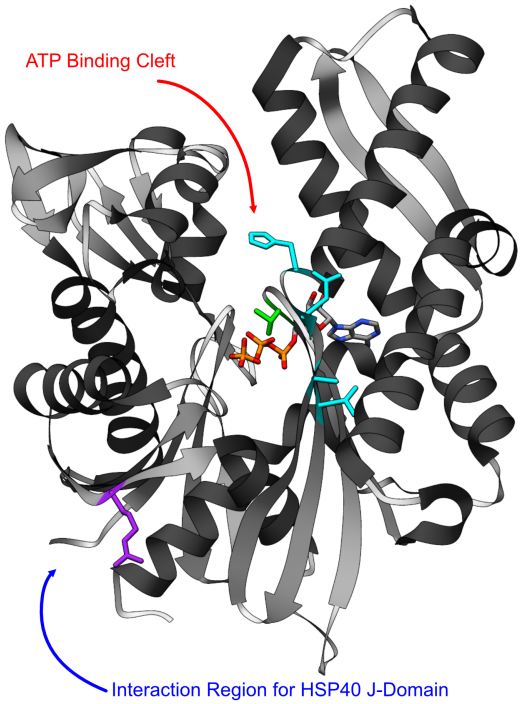


Figure 51: X-ray crystal structure of the Human HSC70 NBD bound to ADP (PDB: 4H5T) with bound ADP shown. Residues identified as shifting by titration experiments shown (Cyan). R171 (Purple) which has been demonstrated to interact directly with the HPD motif in J-domains [111] is also shown, indicating the J-domain binding site on the NBD. Additionally T204 (Green) which has been identified as catalytically important is highlighted.

(subsection 6.1.2). Residues with perturbations included V207, A223, D225 and H227 (Figure 51), the latter three residues agree with the binding site observed for HSP40 on DNAK by NMR [109]; interestingly this portion of the protein is largely occluded by the SBD in an ATP bound DNAK J-domain complex solved by X-ray crystallography [177]. Apart from the three residues proximal to the suspected J-domain binding site we also observed a clear shift in V207 which is located proximal to the catalytic site of the NBD (T204 is a key catalytic residue [178], Figure 51). Numerous observed spectral changes occur in regions which have not yet been assigned, including several changes in leucine and valine methyls as well as changes in tryptophan indole groups. Strikingly, a chain of allostery can be seen between the J-domain binding site and the catalytic site of the NBD.

7.4 ATPASE ASSAYS

A modified colorimetric method based on the detection of phosphomolybdate complexes and hence free inorganic phosphate by malachite green, an approach that has a long history of use in biochemistry [179], was taken in order to measure ATPase activity. As expected HSC70 on its own showed only a sluggish rate of ATP hydrolysis. Furthermore, in all cases, all J-domain constructs of the three HSP40 proteins of interest (DNAJB2, DNAJB6 and DNAJC5) stimulated the ATPase activity of HSC70. With the assays confirming the presence of a J-domain HSP70 interaction under the conditions we had been using for NMR, we were able to begin applying the assays to gather relevant functional data.

7.4.1 *Differences Between J-Domain Constructs: The G/F Rich Region*

Comparison of the three DNAJB6 truncations studied by NMR (subsection 6.2.3) seems to show a small but significant difference between the stimulatory activities of DNAJB6 1-77 and 1-100 when compared to 1-130, the latter possessing a stronger stimulatory rate than the other two (Figure 52A). These correlates seem to fall into place with NMR observations where the largest perturbation of backbone resonances is seen on extending out from residue 100 to residue 130 (Figure 52B). The salient interpretation of this is that structural changes induced by the G/F rich region have a direct impact on the way the J-domain interacts with HSC70. Nonetheless some caution is required as this data does not rule out the possibility that the disordered G/F rich-region extension is binding HSC70 as a substrate and hence contributing to the observed stimulation through an alternate pathway.

A similar though less distinct difference may exist upon extending DNAJB2 out from residue 100 to 130. No solid conclusion with regards to the behaviour of DNAJB2 can however be drawn as the size of the observed difference is not above the variation of these experiments, and unlike for DNAJB6 we do not have spectral data that correlates with this change.

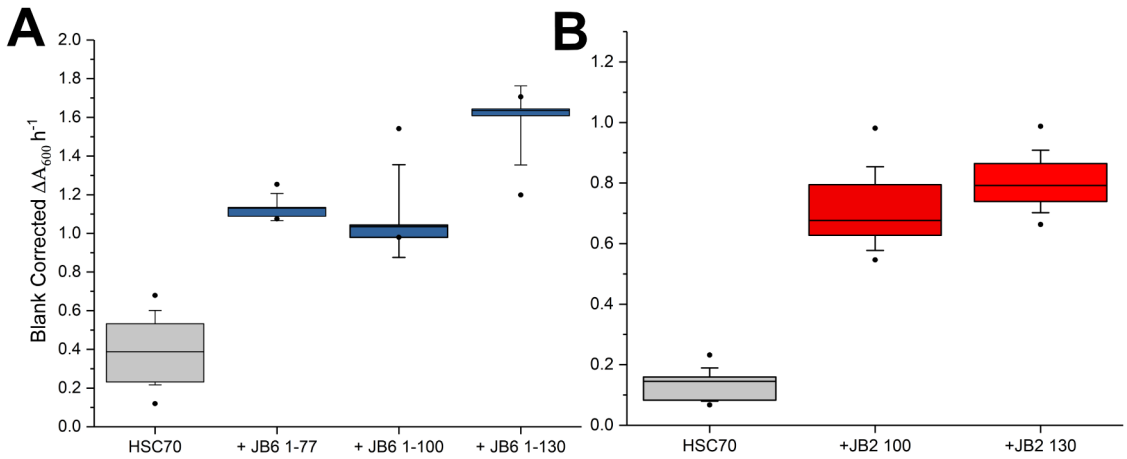


Figure 52: Box plots showing the effects of different length DNAJB6 and DNAJB2 J-domain constructs on the ATPase activity of HSC70. Box plot whiskers represent 1σ, the central line shows the median and the boxes extend from the 1^{st} to the 3^{rd} quartile of the data. Filled circles show the 1^{st} and 99^{th} percentiles of the data respectively. ATPase activity was measured through a colorimetric assay. Values for the box plots are blank adjusted by subtracting the average absorbance of five samples without any protein. A: Comparison of the stimulatory activity of three DNAJB6 constructs, 1-77 (J-domain alone), 1-100 (J-domain and linker), and 1-130 (J-domain, linker and a portion of the G/F rich region). The boxplot for HSC70 alone is based on 18 data points whereas all others are based on 5 data points. B: Comparison of the stimulatory activity of two DNAJB2 constructs, 1-100 (J-domain and linker) and 1-130 (J-domain, linker and a portion of the G/F rich region). All box plots are based off 5 data points.

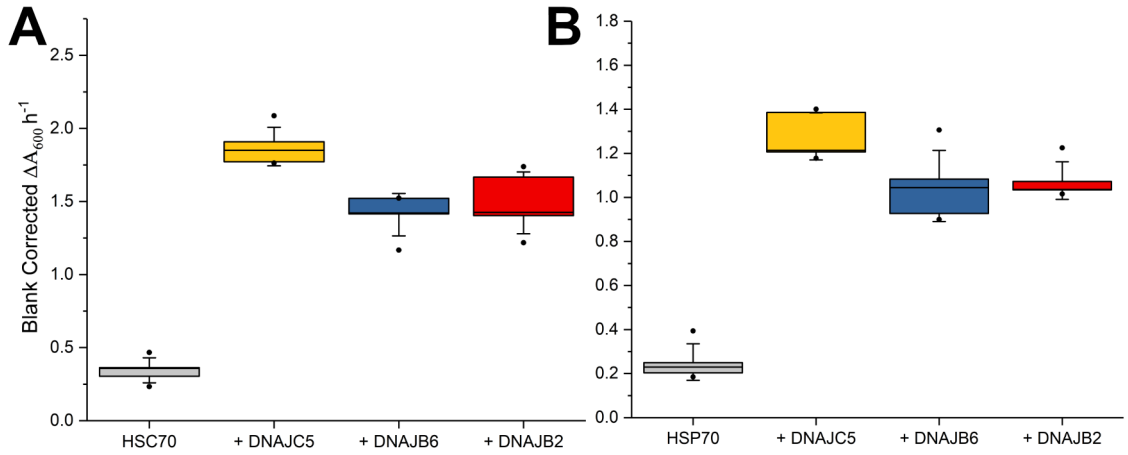


Figure 53: Box plots showing the effects of DNAJB2 (Red), DNAJB6 (Blue) and DNAJC5 (Yellow) J-domains (residues 1-100) on the ATPase activity of HSC70 shown in A and HSP70 shown in B, activity of the HSP70 protein on its own is shown in grey. Box plot whiskers represent 1σ, the central line shows the median, and the boxes extend from the 1^{st} to the 3^{rd} quartile of the data. Filled circles show the 1^{st} and 99^{th} percentiles of the data respectively. ATPase activity was measured through a colorimetric assay. Values for the box plots are blank adjusted by subtracting the average absorbance of five samples without any protein. All box plots are based off 5 data points.

7.4.2 *Differences between HSC70 and HSP70*

The straight forward nature of the activity assays employed offered a good starting point not only for comparing the activities of our different target J-domains but also for the comparison of HSC70 and HSP70. Though biologically distinct entities with distinct regulation and function as discussed previously (subsubsection 1.2.4.3), HSP70 and HSC70 share 85% identity and 95% similarity. With the differing roles of our selected target J-proteins a key point was to ascertain whether HSP70 and HSC70 differed in relation to their interaction with these proteins. Production of HSP70 using an identical method to that used for HSC70 proved viable and the protein behaved indistinguishably from HSC70 throughout its purification displaying the same characteristic polydisperse SEC chromatogram associated with oligomerisation observed for HSC70.

A comparison of the activity of 1-100 constructs of all J-domains against HSC70 shows that DNAJB6 and DNAJB2 possess similar stimulatory activity while DNAJC5 possesses significantly stronger stimulatory activity (Figure 53). This trend was

robustly reproducible across a range of different protein preps and repeat experiments. The observed behaviour correlates well with sequence similarity, with DNAJB2 and DNAJB6 possessing far more similar J-domains with respects to each other than to DNAJC5.

J-domain	HSC70	HSP70
DNAJB2	1.49	1.08
DNAJB6	1.41	1.05
DNAJC5	1.88	1.28

Table 2: Mean activity (ΔA_{600} h^{-1}) from 5 repeats is shown (to 3 s.f.) for the six different J-domain (residues 1-100) and HSP70 protein combinations tested. Activity was measured using a colourimetric assay and data shown is blank subtracted against samples containing no protein. Data corresponds to that shown in the box-plots in Figure 53.

Side by side comparison of the stimulatory activity of 1-100 constructs of all three HSP40 targets with HSP70, repeats the pattern observed for HSC70 with DNAJB2 and DNAJB6 possessing similar activity and being weaker than DNAJC5 (Figure 53). Comparison of the ATPase rates of HSP70 and HSC70 shows that for each J-domain the stimulated activity of HSP70 is lower relative to that observed with HSC70 (Table 2). This difference appears most evident for DNAJC5 with the gap in activity between DNAJC5 and DNAJB2/6 being substantially narrower for HSP70. This lack of difference in the relative preferences for different J-domains exhibited by either HSP70 or HSC70 tentatively supports the widely accepted though untested notion that these proteins are interchangeable. This degree of interchangeability would support the notion that cellular control over which protein is selected may be applied predominantly at the level of protein expression.

7.4.3 *Role of the C-Terminal Domain*

A salient point in any comparison of HSP70 and HSC70 is the similarity of the two proteins, and analysis of sequence alignments shows that vast swathes of both the NBD and SBD are virtually identical. Differences between the two tend to cluster in a small handful of regions with the largest systematic deviation between the two

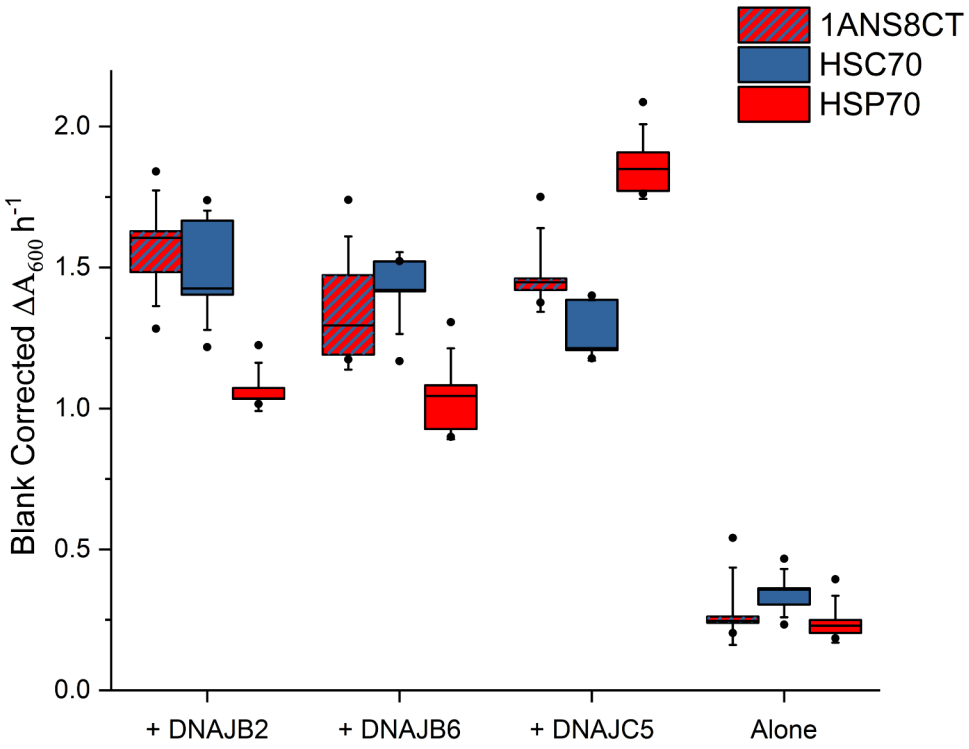


Figure 54: Boxplots showing the effects of the C-terminal domain on the behaviour of HSP70 when stimulated by DNAJB2, DNAJB6 and DNAJC5. The 1ANS8CT construct is a chimera with the NBDSBD of HSP70 (residues 1-555) and the CTD of HSC70 (residues 556-646). Boxplot whiskers represent 1σ, the central line shows the median, and the boxes extend from the 1^{st} to the 3^{rd} quartile of the data. Filled circles show the 1^{st} and 99^{th} percentiles of the data respectively. ATPase activity was measured through a colorimetric assay. Values for the box plots are blank adjusted by subtracting the average absorbance of five samples without any protein. All box plots are based off 5 data points.

sequences occurring at the CTD past residue 550. Given the modularity of these proteins, a quick avenue for testing the significance of these C-terminal differences existed in the creation of a chimeric construct possessing the NBDSBD (residues 1-555) of HSP70 and the CTD (residues 556-646) of HSC70. Unsurprisingly given the behaviour of the two proteins it is derived from, the chimeric construct behaved and purified identically to both HSP70 and HSC70 displaying the same characteristic polydispersity and oligomeric properties. Its rate of unstimulated ATP hydrolysis is also in line with that of its parent proteins (Figure 54). Its behaviour upon J-protein stimulation however is far more interesting. In the case of stimulation by DNAJB6 and DNAJB2 switching the CTD of HSP70 for that of HSC70 appears to produce stimulated rates identical to those observed with HSC70, i.e. higher than those seen for HSP70 (Figure 54). Stimulation by DNAJC5 is also raised by switching the CTD, though unlike with DNAJB6 and DNAJB2 the increase is not sufficient to match the difference observed between HSC70 and HSP70 (Figure 54).

This data shows that the CTD is a defining feature in determining the response of HSP70 and HSC70 to J-domains and that differences between the CTD of these two proteins account at least in part for their different behaviours. The importance of the CTD is further corroborated by NMR data which shows that this region is essential for the binding of DNAJC5 (our data) and DNAJB2 (previously published data by another group[146]) J-domains. Whether this effect is direct resulting from contacts between the J-domain and the CTD or indirect resulting from an allosteric change in the NBD remains to be tested. Another open question is why a complete change of behaviour is not seen for DNAJC5 while it is seen for DNAJB6 and DNAJB2, a phenomenon that may suggest a slightly different mode of interaction for the latter two proteins when compared to the former. Although it requires further corroboration one could hypothesise that such a behaviour may be a general difference between B and C class HSP40 proteins.

8

TOWARDS A COMPLETE MAP OF J-DOMAIN-HSC70 ALLOSTERY

Although we were unable to produce full length versions of any of the selected target HSP40 proteins, the extensive protein work carried out did allow us to isolate the required components for reconstituting a minimal HSC70-J-domain interaction. Furthermore, work with truncations of DNAJB6 enabled the acquisition of data examining the effects of the G/F rich region on the J-domain and its stimulatory activity. The data collected on the HSP70-HSP40 system confirms that we are able to reconstitute this interaction *in vitro*. SEC of DNAJC5 J-domain and full length HSC70 shows a weak interaction while NMR from both the perspective of the J-domain and HSC70 confirm binding of DNAJC5 and DNAJB6 J-domains to HSC70. In all cases (DNAJB2, DNAJB6 and DNAJC5), we were able to confirm interactions through the use of a colorimetric ATPase assay. Though many avenues remain open to further investigation the wealth of data produced so far allows for several key observations to be made.

8.1 THE IMPORTANCE OF THE HSP70 C-TERMINAL DOMAIN

One of the main points highlighted by the data is the importance of the C-terminal domain of HSC70 and HSP70 in mediating interactions with the J-domain. Truncation of HSC70 to remove this domain abrogated DNAJC5 binding, in agreement with previously published data on DNAJB2 [146]. Our analysis of the activities of different J-domains with respects to HSC70 and HSP70 shows that a similar pattern of activation is observed for both proteins, with the DNAJC5 J-domain being the strongest stimulator of activity while DNAJB2 and DNAJB6 J-domains show a

weaker effect. Although the same relative pattern is observed in both cases, stimulation of HSC70 is stronger for all the J-domains tested. Using a chimeric construct, we were able to demonstrate that attaching the C-terminal domain of HSC70 onto HSP70 results in a protein whose response to J-domain stimulation most closely resembles that of the C-terminal donor. This observation is supported by analysis of protein sequences which shows that many of the differences between HSP70 and HSC70 are in their C-terminal domains. Notably while the responses of this chimeric protein to DNAJB2 and DNAJB6 were identical to those of HSC70 the response to DNAJC5 was somewhere in between HSP and HSC70. This suggests that at least for this interaction further determinants of specificity likely lie outside the C-terminal domain.

How the C-terminal domain imparts these behaviours onto HSP70 merits further investigation. Whether the effects observed are direct, occurring through CTD-J-domain contacts or indirect, occurring due to alterations in the conformation and dynamics of the rest of the protein provides a particularly intriguing avenue for further study. An avenue that could be addressed by NMR comparisons between a chimera with the CTD of HSP70 and NBDSBD of HSC70 and wild type HSC70. That the same pattern of stimulatory activity is observed for both HSP70 and HSC70 suggests that some degree of interchangeability exists between these two proteins *in vitro* and may exist in a cellular context; this highlights the importance of the control of relative expression levels of the two in defining which of the them is most likely to be used in a given situation. Future work using a more expansive screen of J-domains would be the next key step in finding out whether the interchangeable responses seen here can be generalised. Ultimately a complete test of this interchangeability will require more *in vivo* or *in cellulo* work, an area where fluorescence microscopy studies may be particularly fruitful.

8.2 THE ROLE OF THE G/F RICH REGION IN DNAJB6

An unexpected result of our work on DNAJB6 truncations came in the form of observed spectral changes as the J-domain is extended out towards the G/F rich region. Changes in backbone resonances can be seen to occur on going out from residue 77 to 100 and again when the construct is extended out to residue 130. The backbone resonance changes are particularly stark between 1-100 and 1-130 constructs and precluded a straight forward transfer of assignments from the longer construct to the truncations. This data coincides with ATPase assay data which shows a small but significant increase in stimulatory capacity between these two constructs–a change that is not observed between 1-77 and 1-100. Taken together these observations suggest that structural alterations occurring within the J-domain as a result of the G/F rich region, have a functional impact on the protein. The location of peaks pertaining to the last thirty residues of the 1-130 construct as well as their propensity for degradation suggests that this region remains highly dynamic and contrasts with the crystal structure of the J-domain of a thermophilic organism where rigid packing of the G/F rich region onto the J-domain is observed [175]. The exact nature of the structural changes requires future work, and NMR-based dynamics experiments between the different constructs as well as chemical shift assignments for the truncations would serve to better elucidate the molecular details of the observed changes. Further protein work following up on the urea purification method preliminarily developed here may also allow these structural observations to be extended out further into full-length DNAJB6 which poses a highly interesting structural target.

8.3 DELINEATING AN ALLOSTERIC NETWORK

The data collected using selective labelling and deuteration on the interaction of DNAJC5 J-domain (residues 1-100) with HSC70 is particularly promising. Though the main hurdle, assignment of the backbone resonances of full length HSC70, is

yet to be completely surmounted we were able to begin to extract some atomistic data from the perspective of HSC70. In particular, we were able to map a series of shifts ranging from the proposed J-domain interaction site through to the catalytic site of the NBD. This data serves as a valuable proof of principle and hints at the more complete mapping of the allosteric changes induced in HSC70 upon J-domain binding, which could be achieved as more assignments are produced. A new avenue of research may be opened up by investigating any conformational differences when full-length DNAJB6 binds to HSC70 in the absence and presence of substrate. A concerted divide and conquer approach should allow for a larger amount of data to be mined from the system. With the SBD now assigned complete assignment of the NBD under our conditions provides the next logical step in this process.

Part IV

APPENDICES

A

NMR ASSIGNMENT INFORMATION

A.1 HSC70 SBD ASSIGNMENTS

A.1.1 *Assigment Statistics*

	HN	N	Cα	Cβ	C′
Total	164	171	171	160	171
Assigned	151	151	155	150	147
Percent Assigned	92.1	88.3	90.6	93.8	86.0

A.1.2 *Assigment Table*

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
ASN	3	-	-	50.557	36.196	172.611
VAL	4	8.221	121.004	60.075	29.983	173.630
GLN	5	8.463	123.379	53.539	26.793	173.092
ASP	6	8.271	121.544	51.958	38.467	173.586
LEU	7	8.077	122.282	52.813	39.761	-
LEU	10	-	-	52.179	-	174.026
ASP	11	8.537	121.514	51.655	38.514	172.509
VAL	12	7.645	110.396	54.974	34.110	174.525
THR	13	8.825	114.670	55.980	66.281	-

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
SER	16	-	-	57.667	60.051	169.064
LEU	17	8.613	121.334	50.154	45.254	174.738
GLY	18	9.309	112.215	44.006	N/A	168.443
ILE	19	8.504	112.812	54.830	39.839	119.148
GLU	20	7.285	119.160	52.997	29.867	174.614
THR	21	9.480	120.976	57.883	67.499	171.673
ALA	22	8.097	119.549	50.938	16.319	175.050
GLY	23	8.446	107.914	41.667	N/A	171.700
GLY	24	7.920	105.901	42.868	N/A	171.444
VAL	25	6.970	123.652	59.436	31.435	171.826
MET	26	8.563	125.026	52.056	31.123	171.508
THR	27	9.097	130.388	60.089	66.099	171.117
VAL	28	8.774	128.369	62.262	29.891	173.131
LEU	29	8.985	128.212	53.830	40.979	173.693
ILE	30	7.578	115.894	56.902	37.849	173.180
LYS	31	8.982	130.301	54.686	30.917	174.184
ARG	32	8.019	124.856	55.071	27.591	171.605
ASN	33	9.000	119.990	51.857	33.956	171.867
THR	34	7.966	117.915	61.267	66.781	171.639
THR	35	8.536	122.394	64.567	-	176.156
ILE	36	8.357	119.603	56.316	-	-
PRO	37	N/A	-	59.692	66.097	172.871
THR	38	8.558	115.078	59.137	66.069	167.155
LYS	39	7.782	124.267	52.565	32.980	172.786
GLN	40	9.099	123.986	52.153	31.085	171.244

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
THR	41	8.614	120.931	58.718	69.409	171.441
GLN	42	9.108	125.026	52.563	31.668	171.002
THR	43	8.482	118.894	60.652	66.655	170.805
PHE	44	9.707	127.379	54.314	40.831	170.365
THR	45	8.642	114.105	56.204	67.914	171.607
THR	46	8.754	110.483	57.643	66.500	172.003
TYR	47	8.297	122.225	56.368	38.785	172.065
SER	48	8.669	113.351	54.101	63.315	170.626
ASP	49	8.773	121.909	52.836	38.384	174.037
ASN	50	9.058	116.177	52.234	33.820	171.352
GLN	51	7.625	123.901	-	27.904	-
PRO	52	N/A	-	61.513	29.448	173.101
GLY	53	7.056	106.038	42.832	N/A	169.588
VAL	54	8.691	109.605	56.194	32.801	170.017
LEU	55	8.090	124.061	50.920	40.400	172.054
ILE	56	8.915	129.532	56.124	31.684	171.004
GLN	57	8.978	125.862	51.906	31.012	171.218
VAL	58	8.736	122.918	59.242	29.450	171.431
TYR	59	9.163	127.425	54.637	43.374	171.573
GLU	60	9.125	118.524	50.437	31.610	172.991
GLY	61	8.805	112.711	41.304	N/A	170.157
GLU	62	8.431	111.195	51.722	28.331	174.875
ARG	63	8.363	118.242	50.455	24.783	175.332
ALA	64	7.409	122.336	51.882	18.605	173.568
MET	65	8.691	114.340	49.239	27.796	175.361

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
THR	66	8.303	114.462	63.437	62.622	174.117
LYS	67	8.362	119.140	55.610	28.418	173.489
ASP	68	7.326	117.314	51.875	39.706	172.021
ASN	69	7.724	119.957	49.036	38.238	170.180
ASN	70	8.668	116.243	50.100	38.354	171.360
LEU	71	8.768	129.232	52.900	38.881	173.895
LEU	72	9.406	126.867	52.088	41.308	174.993
GLY	73	7.164	103.888	42.545	N/A	166.796
LYS	74	8.592	116.272	51.964	34.435	171.071
PHE	75	8.779	114.996	54.109	37.327	168.878
GLU	76	8.932	118.995	51.857	30.885	-
LEU	77	-	-	51.918	41.895	172.530
THR	78	8.367	116.761	57.518	68.120	171.279
GLY	79	8.272	107.503	43.425	N/A	172.988
ILE	80	8.051	123.468	56.970	35.272	-
PRO	84	N/A	-	60.089	29.396	172.299
ARG	85	9.235	120.939	53.104	26.647	173.661
GLY	86	7.388	109.124	43.934	N/A	172.440
VAL	87	7.578	119.155	63.245	27.310	-
PRO	88	N/A	-	61.865	29.329	173.352
GLN	89	9.200	123.150	51.895	28.203	170.436
ILE	90	9.210	125.197	55.842	36.497	173.170
GLU	91	9.346	128.697	53.558	29.010	174.045
VAL	92	9.517	130.694	58.905	31.095	171.574
THR	93	9.220	122.909	58.943	68.694	171.148

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
PHE	94	9.558	128.480	52.989	38.257	170.122
ASP	95	8.831	124.411	50.553	42.504	171.941
ILE	96	9.281	125.876	57.465	37.966	173.649
ASP	97	8.228	128.768	50.235	38.871	174.440
ALA	98	8.459	117.190	50.171	15.898	175.178
ASN	99	8.169	113.771	50.174	36.852	172.851
GLY	100	8.204	108.857	42.946	N/A	171.198
ILE	101	8.327	123.851	59.576	33.227	172.088
LEU	102	8.104	130.474	50.747	40.913	170.833
ASN	103	9.530	126.955	49.887	39.362	172.104
VAL	104	8.759	125.554	57.932	30.660	170.924
SER	105	8.973	119.244	53.207	64.179	170.139
ALA	106	9.007	123.258	47.778	20.502	171.949
VAL	107	8.639	119.723	56.670	33.923	172.351
ASP	108	8.707	126.822	50.541	39.274	174.706
LYS	109	8.827	126.154	55.399	28.860	175.790
SER	110	9.001	114.253	58.490	60.452	173.285
THR	111	7.461	109.056	59.352	68.623	174.073
GLY	112	8.877	112.548	42.746	N/A	171.408
LYS	113	8.079	122.849	55.172	30.252	171.711
GLU	114	8.260	121.127	51.311	32.463	173.572
ASN	115	8.754	118.771	48.901	38.645	169.118
LYS	116	7.903	118.775	52.285	32.927	171.354
ILE	117	8.423	115.693	57.686	40.138	170.063
THR	118	8.274	120.873	59.251	68.014	171.137

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
ILE	119	9.634	130.301	57.347	36.894	172.464
THR	120	8.501	121.065	59.300	36.879	170.932
ASN	121	8.408	120.916	49.927	66.583	171.946
ASP	122	8.421	121.816	51.703	36.422	174.272
LYS	123	8.486	123.043	54.910	38.637	174.817
GLY	124	8.637	109.141	42.800	N/A	171.906
ARG	125	7.651	120.113	54.785	27.869	173.550
LEU	126	8.060	119.976	51.180	41.618	173.880
SER	127	9.171	119.268	54.035	63.381	172.299
LYS	128	8.929	121.693	57.522	29.680	176.226
GLU	129	8.700	118.423	57.571	26.166	176.560
ASP	130	7.975	121.740	54.992	38.695	175.800
ILE	131	8.014	120.954	62.993	35.837	174.908
GLU	132	8.132	117.996	57.041	26.690	176.569
ARG	133	7.997	119.347	57.082	27.068	176.233
MET	134	8.178	119.633	57.078	30.467	177.084
VAL	135	8.575	121.669	60.973	-	170.467
GLN	136	8.727	121.184	55.524	25.525	177.192
GLU	137	7.777	119.141	57.071	27.441	175.330
ALA	138	7.743	119.584	52.649	15.982	177.435
GLU	139	7.583	117.038	56.524	26.687	176.513
LYS	140	7.977	121.244	56.544	29.793	175.936
TYR	141	7.650	115.442	55.219	35.449	172.505
LYS	142	7.140	121.994	58.042	30.350	171.867
ALA	143	8.615	120.172	51.857	-	178.389

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
GLU	144	6.715	113.016	56.820	26.740	176.891
ASP	145	8.942	122.415	54.569	37.509	175.810
GLU	146	8.427	119.797	56.966	26.704	175.944
LYS	147	7.497	117.632	56.395	29.747	176.865
GLN	148	7.713	117.231	55.811	26.103	175.943
ARG	149	8.070	119.610	56.677	27.572	175.778
ASP	150	8.031	117.915	53.240	38.129	174.499
LYS	151	7.360	117.483	54.843	30.277	174.259
VAL	152	7.469	119.047	60.002	-	-
SER	153	-	-	56.518	61.010	172.344
SER	154	7.971	116.862	55.257	61.240	171.968
LYS	155	8.644	122.470	53.816	30.557	173.688
ASN	156	8.741	118.124	50.843	35.488	171.365
SER	157	8.519	114.737	55.483	62.896	169.929
LEU	158	9.123	122.576	50.363	42.086	174.557
GLU	159	8.828	122.454	53.287	29.921	173.304
SER	160	8.912	113.109	55.817	61.915	170.702
TYR	161	7.427	118.637	53.539	37.654	171.318
ALA	162	8.392	124.995	49.512	16.869	174.247
PHE	163	8.753	123.658	57.493	36.947	173.592
ASN	164	8.612	117.272	50.461	36.124	172.287
MET	165	7.898	120.296	52.876	-	-

A.2 DNAJB6B 1-130

A.2.1 *Assigment Statistics*

	HN	N	Cα	Cβ	C′
Total	124	132	132	117	132
Assigned	104	102	110	109	101
Percent Assigned	83.9	77.3	83.3	93.2	76.5

A.2.2 *Assigment Table*

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
ALA	2	-	-	49.685	16.835	174.780
MET	3	8.434	120.441	52.729	30.444	173.034
VAL	4	8.048	123.051	59.111	30.284	171.694
ASP	5	8.542	125.490	49.667	38.899	174.171
TYR	6	7.736	123.394	56.590	35.253	174.524
TYR	7	7.971	116.840	59.819	33.194	175.470
GLU	8	7.636	120.383	56.184	26.617	177.684
VAL	9	8.071	120.051	63.644	28.874	174.890
LEU	10	7.012	113.781	51.787	39.886	172.995
GLY	11	7.960	109.154	43.047	N/A	-
ARG	14	-	-	55.640	26.386	173.745
HIS	15	7.077	112.423	51.909	26.255	172.241
ALA	16	7.495	123.250	50.205	16.820	175.045
SER	17	9.390	120.446	53.706	60.783	-
PRO	18	N/A	-	63.297	29.213	176.929

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
GLU	19	8.377	117.347	57.394	26.496	175.859
ASP	20	8.075	121.858	54.833	38.193	177.150
ILE	21	7.898	121.795	62.532	35.391	174.215
LYS	22	7.980	119.694	58.301	30.389	175.443
LYS	23	8.195	117.587	57.265	30.292	176.239
ALA	24	7.834	122.308	52.363	16.770	176.684
TYR	25	8.212	117.587	59.648	36.256	172.520
ARG	26	7.983	117.300	57.140	27.534	174.175
LYS	27	7.251	113.116	56.505	30.245	177.107
LEU	28	8.094	119.899	54.950	40.638	175.647
ALA	29	9.405	123.528	52.797	14.643	176.543
LEU	30	7.117	112.194	54.697	38.843	176.531
LYS	31	7.483	118.384	56.251	30.517	176.015
TRP	32	8.150	114.652	53.765	26.505	170.304
HIS	33	7.047	120.824	-	27.917	-
PRO	38	N/A	-	62.706	28.733	175.769
GLU	39	8.789	117.217	54.014	26.587	173.346
ASN	40	7.873	120.630	49.320	36.033	172.497
LYS	41	8.040	121.877	58.251	30.628	175.465
GLU	42	8.549	118.728	57.519	26.532	176.848
GLU	43	8.206	120.434	56.083	36.770	175.547
ALA	44	8.143	120.484	53.152	15.297	175.785
GLU	45	8.803	116.672	57.716	26.671	176.930
ARG	46	7.471	119.129	56.958	27.791	176.619
LYS	47	8.238	120.984	55.658	28.641	175.580

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
PHE	48	9.541	120.930	59.440	37.268	175.491
LYS	49	7.797	118.996	56.874	30.222	175.439
GLN	50	7.844	119.434	56.707	26.198	175.921
VAL	51	8.567	118.433	63.330	29.266	174.291
ALA	52	7.322	121.979	53.078	15.227	176.161
GLU	53	7.703	118.513	56.558	26.983	174.276
ALA	54	7.661	117.193	51.461	16.571	176.111
TYR	55	8.163	116.324	59.091	37.083	175.466
GLU	56	8.733	121.818	59.223	28.043	175.062
VAL	57	7.115	114.757	63.193	28.914	172.802
LEU	58	7.771	110.629	53.724	38.776	174.841
SER	59	7.859	109.738	56.336	61.054	170.775
ASP	60	6.790	124.510	49.490	40.662	172.070
ALA	61	8.848	127.951	53.107	16.013	177.161
LYS	62	7.947	115.897	56.376	29.174	177.041
LYS	63	7.939	120.445	57.533	30.058	175.356
ARG	64	9.194	121.615	57.725	27.264	174.863
ASP	65	7.665	117.046	55.088	38.921	175.921
ILE	66	7.442	118.778	61.990	35.628	175.389
TYR	67	8.903	122.061	59.257	36.610	175.820
ASP	68	9.417	120.554	54.161	37.604	176.037
LYS	69	7.708	118.975	56.345	30.495	175.408
TYR	70	8.128	117.637	56.453	36.480	174.498
GLY	71	8.211	109.550	43.107	N/A	171.741
LYS	72	8.120	120.203	54.942	30.200	174.913

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
GLU	73	8.528	119.173	54.920	27.133	174.951
GLY	74	8.048	107.821	42.987	N/A	171.757
LEU	75	7.928	120.643	52.913	39.361	174.628
ASN	76	8.236	117.807	50.718	36.256	-
GLY	80	-	-	42.968	N/A	172.150
GLY	81	8.245	108.618	42.743	N/A	172.174
GLY	82	8.332	108.864	42.545	N/A	171.523
SER	83	8.205	115.575	55.448	61.246	-
HIS	84	-	-	53.684	38.623	172.207
PHE	85	8.029	120.689	54.865	36.862	172.531
ASP	86	8.325	121.302	51.684	38.646	172.889
SER	87	7.981	115.952	57.083	61.215	-
PHE	91	-	-	55.468	36.556	173.513
GLY	92	8.136	108.360	42.645	N/A	171.243
PHE	93	7.577	119.223	53.643	37.090	172.286
THR	94	7.919	116.584	57.973	67.691	170.700
PHE	95	8.462	123.843	55.562	38.088	174.053
ARG	96	10.170	124.643	52.979	30.203	173.930
ASN	97	8.939	121.421	49.365	35.991	-
PRO	98	N/A	-	62.819	29.952	175.884
ASP	99	8.857	118.512	54.983	38.164	176.010
ASP	100	7.752	121.280	54.732	37.018	176.065
VAL	101	7.445	121.546	64.081	29.415	175.078
PHE	102	7.894	120.958	59.056	37.669	174.458
ARG	103	8.550	118.731	56.707	27.534	177.128

Residue	Number	^{1}HN ppm	^{15}N ppm	^{13}Cα ppm	^{13}Cβ ppm	^{13}C′ ppm
GLU	104	8.111	118.632	56.357	27.233	176.301
PHE	105	8.367	122.142	59.249	36.914	174.720
PHE	106	7.987	113.505	57.119	-	173.968
GLY	107	7.893	109.390	44.198	N/A	172.545
GLY	108	8.552	107.128	42.676	N/A	171.243
ARG	109	7.586	118.977	52.548	29.155	172.331
ASP	110	8.373	120.782	48.367	38.932	-
PRO	111	-	-	60.842	28.979	173.789
PHE	112	7.815	116.337	55.479	34.929	173.673
SER	113	7.518	114.635	55.811	61.345	171.607
PHE	114	7.809	121.466	55.050	36.849	173.105
ASP	115	7.920	121.313	53.968	28.023	173.145
PHE	116	8.255	121.233	51.418	38.676	173.654
PHE	117	8.146	120.666	56.058	36.540	173.289
GLU	118	8.221	119.122	55.534	-	-
PHE	124	-	-	51.388	38.396	173.013
PHE	125	7.829	120.016	55.484	36.868	173.707
GLY	126	7.959	109.124	43.978	N/A	172.924
ASN	127	7.927	110.463	55.584	32.068	172.123
ARG	128	8.302	117.962	51.855	27.024	-

B

CONSTRUCT SEQUENCES

B.1 HSC70 (HSPA8)

Construct sequences are shown including purification tags, where relevant TEV cleavage sites are shown, cleavage was performed on constructs as per chapter 2, Materials & Methods.

B.1.1 *Full-Length*

```
   −35  −30  −25  −20  −15  −10   −5     5   10   15
MASAWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGMRDAMSKGPAVGIDLGTTYSC
```

Purification Tag

```
   20   25   30   35   40   45   50   55   60   65   70
VGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVAMNPTNTVFDAKR

   75   80   85   90   95  100  105  110  115  120  125
LIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGETKSFYPEEVSSMVLTKM

  130  135  140  145  150  155  160  165  170  175  180
KEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAAIA

  185  190  195  200  205  210  215  220  225  230  235
YGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDTHLGGEDFDNRM

  240  245  250  255  260  265  270  275  280  285  290
VNHFIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGID

  295  300  305  310  315  320  325  330  335  340  345
FYTSITRARFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRIPKIQ

  350  355  360  365  370  375  380  385  390  395  400
KLLQDFFNGKELNKSINPDEAVAYGAAVQAAILSGDKSENVQDLLLLDVTPLSLG
```

```
  405       410       415       420       425       430       435       440       445       450       455
IETAGGVMTVLIKRNTTIPTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLG

  460       465       470       475       480       485       490       495       500       505       510
KFELTGIPPAPRGVPQIEVTFDIDANGILNVSAVDKSTGKENKITITNDKGRLSK

  515       520       525       530       535       540       545       550       555       560       565
EDIERMVQEAEKYKAEDEKQRDKVSSKNSLESYAFNMKATVEDEKLQGKINDEDK

  570       575       580       585       590       595       600       605       610       615       620
QKILDKCNEIINWLDKNQTAEKEEFEHQQKELEKVCNPIITKLYQSAGGMPGGMP

  625       630       635       640       645
GGFPGGGAPPSGGASSGPTIEEVD
```

B.1.2 *NBDSBD*

```
  −35       −30       −25       −20       −15       −10       −5          5         10        15
MASAWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGMRDAMSKGPAVGIDLGTTYSC
```

Purification Tag

```
  20        25        30        35        40        45        50        55        60        65        70
VGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVAMNPTNTVFDAKR

  75        80        85        90        95        100       105       110       115       120       125
LIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGETKSFYPEEVSSMVLTKM

  130       135       140       145       150       155       160       165       170       175       180
KEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAAIA

  185       190       195       200       205       210       215       220       225       230       235
YGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDTHLGGEDFDNRM

  240       245       250       255       260       265       270       275       280       285       290
VNHFIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGID

  295       300       305       310       315       320       325       330       335       340       345
FYTSITRARFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRIPKIQ

  350       355       360       365       370       375       380       385       390       395       400
KLLQDFFNGKELNKSINPDEAVAYGAAVQAAILSGDKSENVQDLLLLDVTPLSLG
```

```
   405       410       415       420       425       430       435       440       445       450       455
IETAGGVMTVLIKRNTTIPTKQTQTFTTYSDNQPGVLIQVYEGERAMTKDNNLLG

   460       465       470       475       480       485       490       495       500       505       510
KFELTGIPPAPRGVPQIEVTFDIDANGILNVSAVDKSTGKENKITITNDKGRLSK

   515       520       525       530       535       540       545       550       555       560       565
EDIERMVQEAEKYKAEDEKQRDKVSSKNSLESYAFNMKATVESEQUENCEGRLSK

   570       575       580       585       590       595       600       605
EDIERMVQEAEKYKAEDEKQRDKVSSKNSLESYAFNMKATVE
```

B.1.3 *NBD*

```
  −35       −30       −25       −20       −15       −10       −5                5         10        15
MASAWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGMRDAMSKGPAVGIDLGTTYSC
```

Purification Tag

```
   20        25        30        35        40        45        50        55        60        65        70
VGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVAMNPTNTVFDAKR

   75        80        85        90        95       100       105       110       115       120       125
LIGRRFDDAVVQSDMKHWPFMVVNDAGRPKVQVEYKGETKSFYPEEVSSMVLTKM

  130       135       140       145       150       155       160       165       170       175       180
KEIAEAYLGKTVTNAVVTVPAYFNDSQRQATKDAGTIAGLNVLRIINEPTAAAIA

  185       190       195       200       205       210       215       220       225       230       235
YGLDKKVGAERNVLIFDLGGGTFDVSILTIEDGIFEVKSTAGDTHLGGEDFDNRM

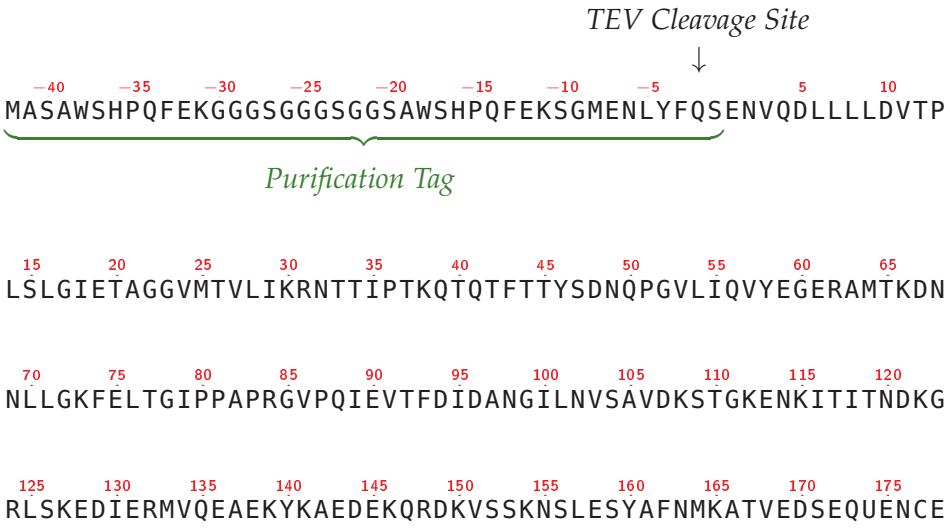
  240       245       250       255       260       265       270       275       280       285       290
VNHFIAEFKRKHKKDISENKRAVRRLRTACERAKRTLSSSTQASIEIDSLYEGID

  295       300       305       310       315       320       325       330       335       340       345
FYTSITRARFEELNADLFRGTLDPVEKALRDAKLDKSQIHDIVLVGGSTRIPKIQ

  350       355       360       365       370       375       380       385       390       395       400
KLLQDFFNGKELNKSINPDEAVAYGAAVQAAILSGDKSSEQUENCEGGSTRIPKI

  405       410       415       420       425       430       435       440
QKLLQDFFNGKELNKSINPDEAVAYGAAVQAAILSGDKS
```

B.1.4 *SBD*

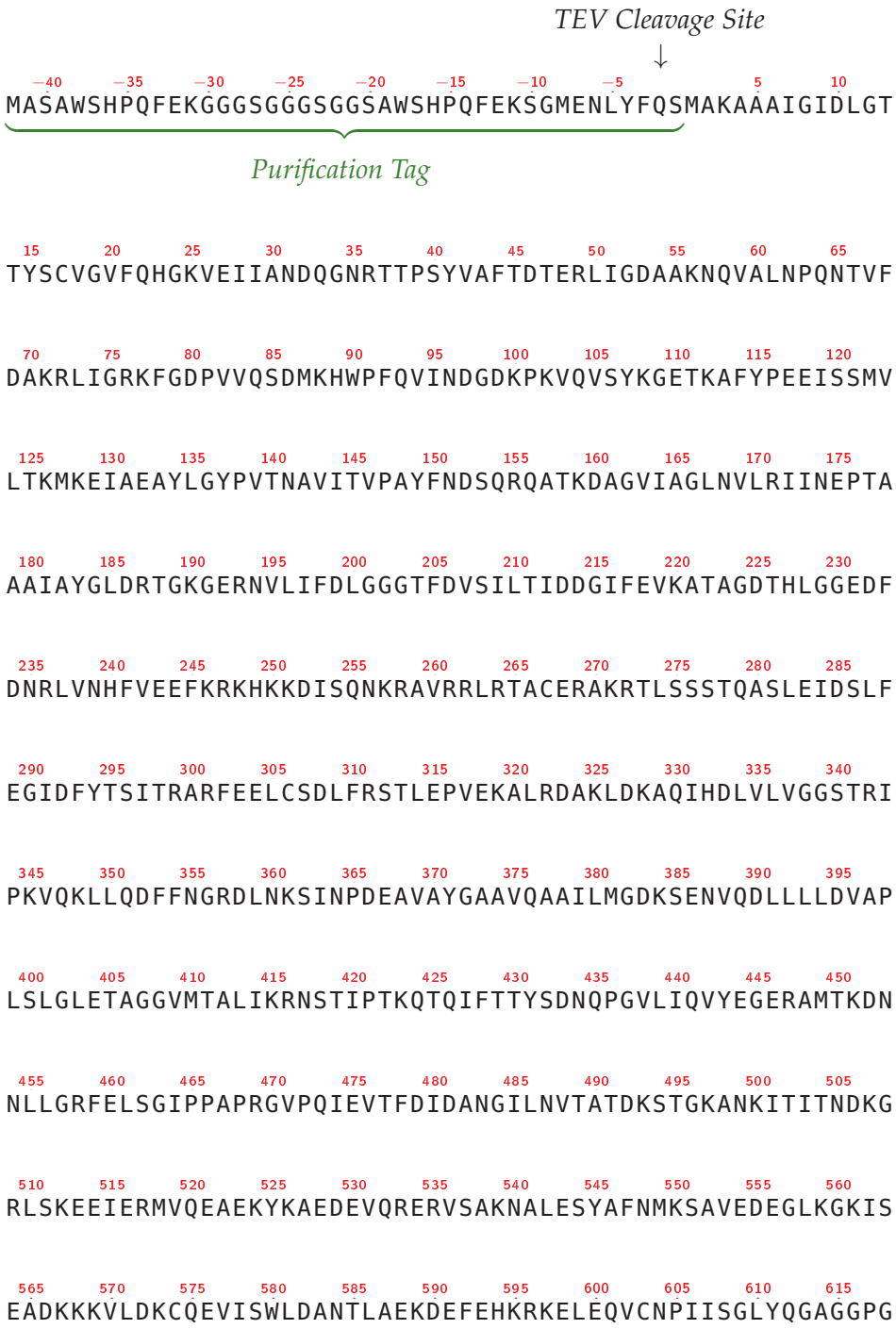


B.1.5 *SBDCTD*



B.2 HSP70 (HSPA1A/B)

B.2.1 *Full-Length*



```
620       625       630       635       640       645       650       655
PGGFGAQGPKGGSGSGPTIEEVDSEQUENCEGPTIEEVD
```

B.3 HSP70-HSC70 CHIMERA

B.3.1 *1ANS8CT*

TEV Cleavage Site ↓

```
 −40       −35       −30       −25       −20       −15       −10       −5        5         10
MASAWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGMENLYFQSMAKAAAIGIDLGT
```

Purification Tag

```
 15        20        25        30        35        40        45        50        55        60        65
TYSCVGVFQHGKVEIIANDQGNRTTPSYVAFTDTERLIGDAAKNQVALNPQNTVF

 70        75        80        85        90        95        100       105       110       115       120
DAKRLIGRKFGDPVVQSDMKHWPFQVINDGDKPKVQVSYKGETKAFYPEEISSMV

 125       130       135       140       145       150       155       160       165       170       175
LTKMKEIAEAYLGYPVTNAVITVPAYFNDSQRQATKDAGVIAGLNVLRIINEPTA

 180       185       190       195       200       205       210       215       220       225       230
AAIAYGLDRTGKGERNVLIFDLGGGTFDVSILTIDDGIFEVKATAGDTHLGGEDF

 235       240       245       250       255       260       265       270       275       280       285
DNRLVNHFVEEFKRKHKKDISQNKRAVRRLRTACERAKRTLSSSTQASLEIDSLF

 290       295       300       305       310       315       320       325       330       335       340
EGIDFYTSITRARFEELCSDLFRSTLEPVEKALRDAKLDKAQIHDLVLVGGSTRI

 345       350       355       360       365       370       375       380       385       390       395
PKVQKLLQDFFNGRDLNKSINPDEAVAYGAAVQAAILMGDKSENVQDLLLLDVAP

 400       405       410       415       420       425       430       435       440       445       450
LSLGLETAGGVMTALIKRNSTIPTKQTQIFTTYSDNQPGVLIQVYEGERAMTKDN

 455       460       465       470       475       480       485       490       495       500       505
NLLGRFELSGIPPAPRGVPQIEVTFDIDANGILNVTATDKSTGKANKITITNDKG

 510       515       520       525       530       535       540       545       550       555       560
RLSKEEIERMVQEAEKYKAEDEVQRERVSAKNALESYAFNMKSAVEDEKLQGKIN
```

```
565       570       575       580       585       590       595       600       605       610       615
DEDKQKILDKCNEIINWLDKNQTAEKEEFEHQQKELEKVCNPIITKLYQSAGGMP

620       625       630       635       640       645       650       655       660       665       670
GGMPGGFPGGGAPPSGGASSGPTIEEVDSEQUENCEELEKVCNPIITKLYQSAGG

675       680       685       690       695       700
MPGGMPGGFPGGGAPPSGGASSGPTIEEVD
```

B.4 DNAJC5

B.4.1 *DNAJC5 1-100*

TEV Cleavage Site ↓

```
−45       −40       −35       −30       −25       −20       −15       −10       −5          5         10
MASAWSHPQFEKGGGSGGGSGGSAWSHPQFEKSGMGENLYFQGAMADQRQRSLST
```

Purification Tag

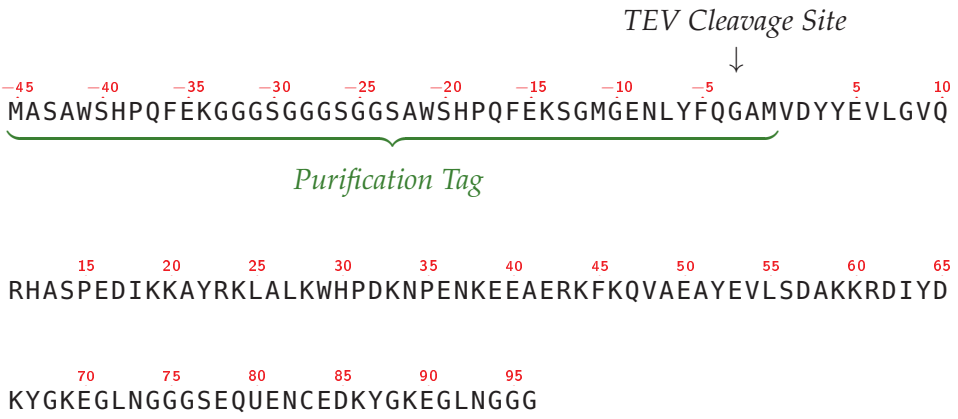
```
      15        20        25        30        35        40        45        50        55        60        65
SGESLYHVLGLDKNATSDDIKKSYRKLALKYHPDKNPDNPEAADKFKEINNAHAI

      70        75        80        85        90        95       100       105       110       115       120
LTDATKRNIYDKYGSLGLYVAEQFGEENVNTYFVSEQUENCENAHAILTDATKRN

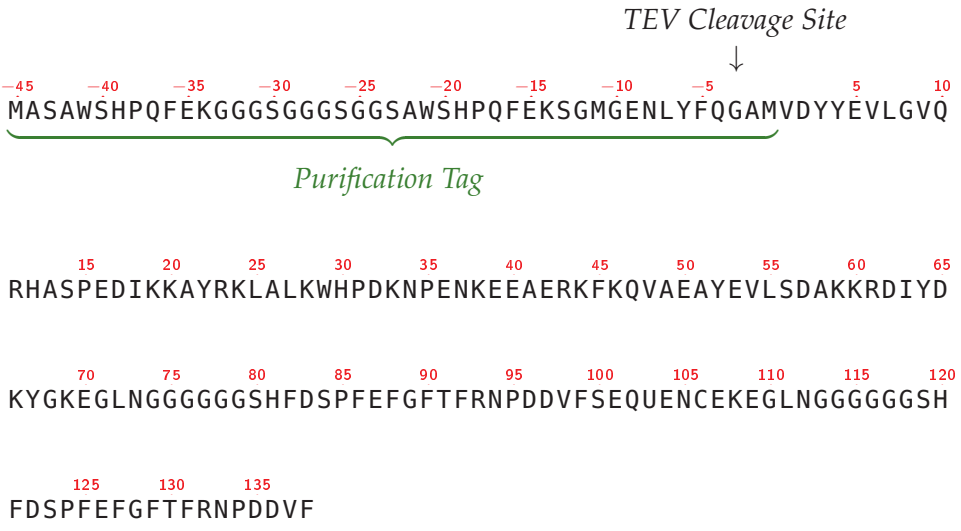
      125       130       135       140       145
IYDKYGSLGLYVAEQFGEENVNTYFV
```

B.5 DNAJB6

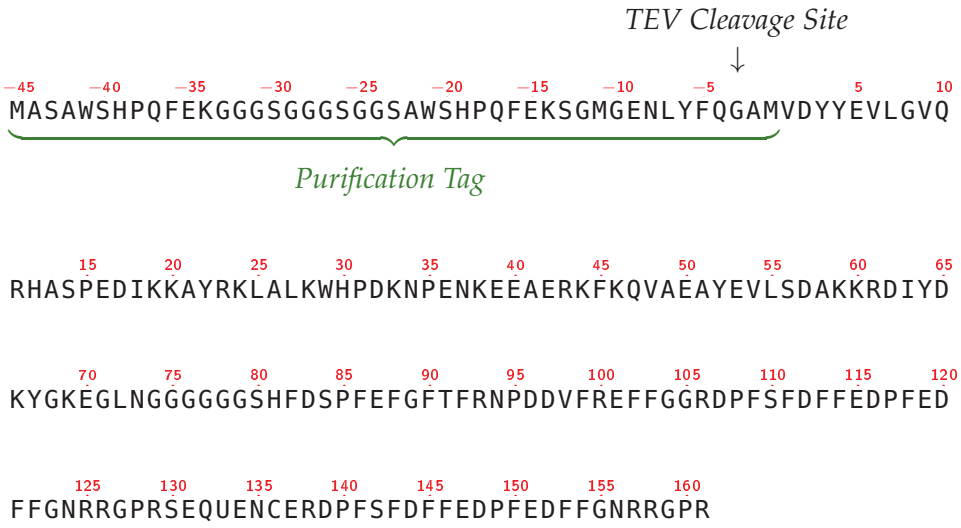
B.5.1 *DNAJB6 1-77*



B.5.2 *DNAJB6 1-100*

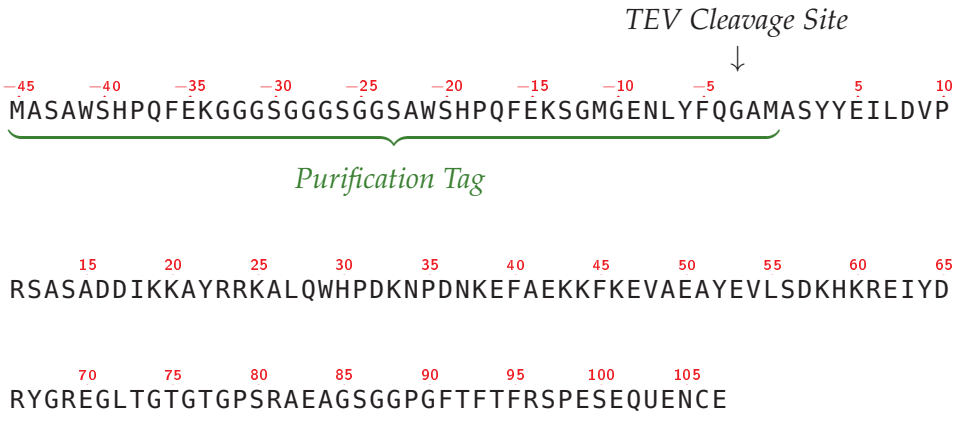


B.5.3 *DNAJB6 1-130*

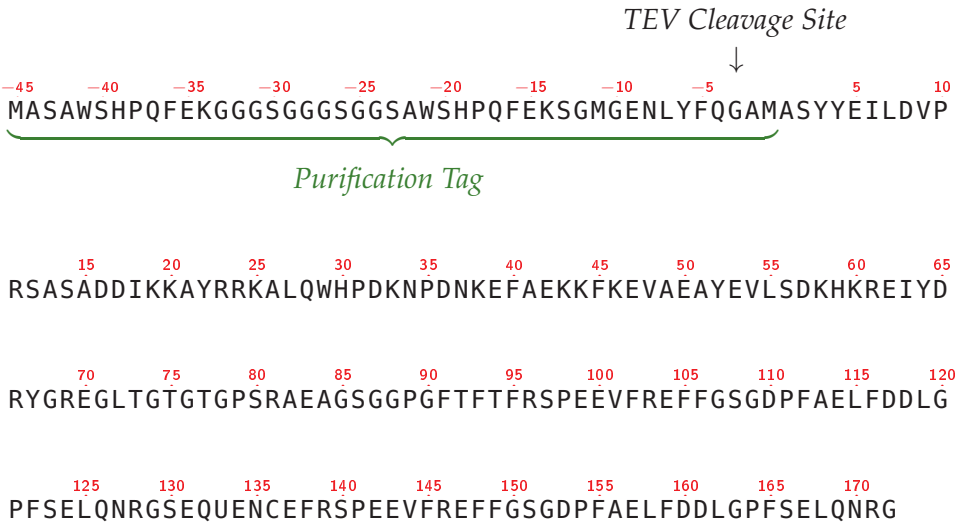


B.6 DNAJB2

B.6.1 *DNAJB2 1-100*



B.6.2 *DNAJB2 1-130*



B.7 BIP

B.7.1 *Full-Length*



```
295       300       305       310       315       320       325       330       335       340       345
FSETLTRAKFEELNMDLFRSTMKPVQKVLEDSDLKKSDIDEIVLVGGSTRIPKIQ

350       355       360       365       370       375       380       385       390       395       400
QLVKEFFNGKEPSRGINPDEAVAYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIE

405       410       415       420       425       430       435       440       445       450       455
TVGGVMTKLIPRNTVVPTKKSQIFSTASDNQPTVTIKVYEGERPLTKDNHLLGTF

460       465       470       475       480       485       490       495       500       505       510
DLTGIPPAPRGVPQIEVTFEIDVNGILRVTAEDKGTGNKNKITITNDQNRLTPEE

515       520       525       530       535       540       545       550       555       560       565
IERMVNDAEKFAEEDKKLKERIDTRNELESYAYSLKNQIGDKEKLGGKLSSEDKE

570       575       580       585       590       595       600       605       610       615       620
TMEKAVEEKIEWLESHQDADIEDFKAKKKELEEIVQPIISKLYGSAGPPPTGEED

625       630       635
TAEKDELSEQUENC
```

B.8 σ_1

B.8.1 *GFP*-σ_1

```
−300      −295      −290      −285      −280      −275      −270      −265      −260      −255      −250
MNHKVHHHHHHIEGRHSSAWSHPQFEKGGGSGGGSGGSAWSHPQFEKQRQALKAS
```

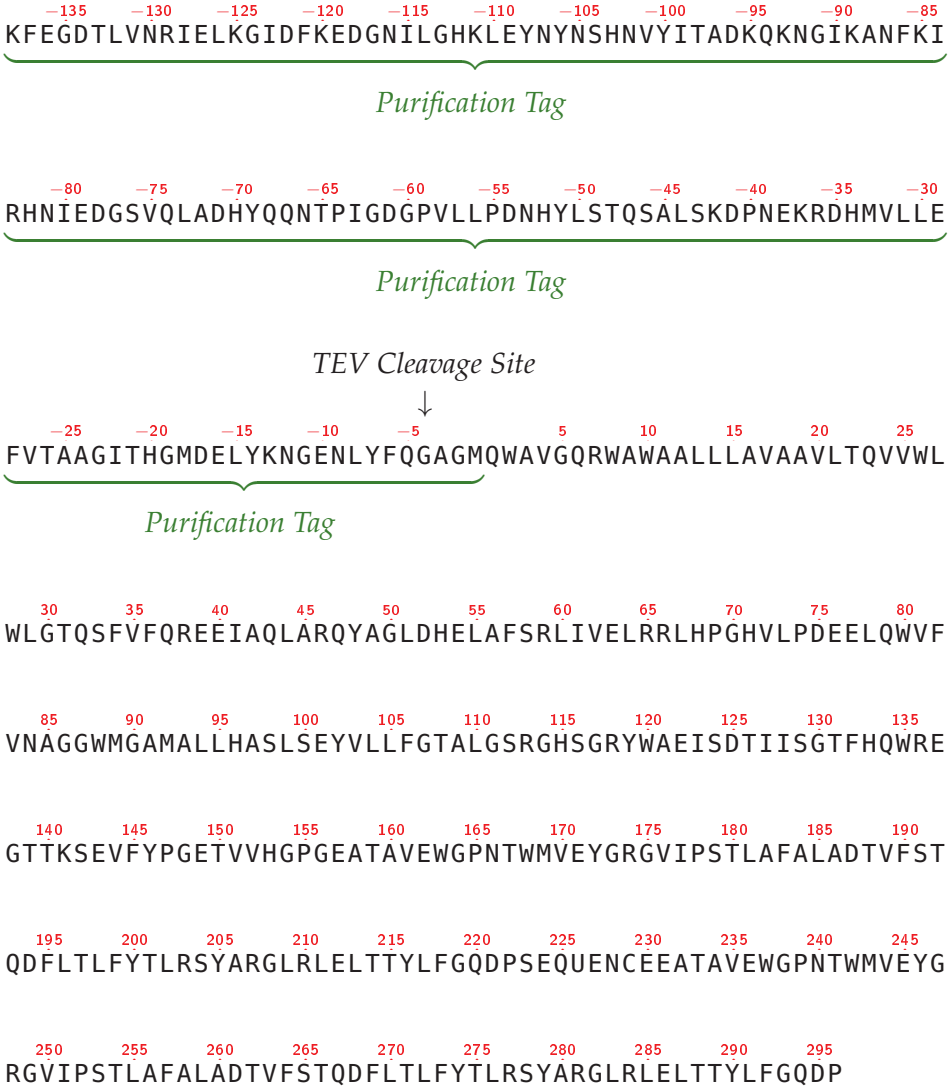
Purification Tag

```
−245      −240      −235      −230      −225      −220      −215      −210      −205      −200      −195
KGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPW
```

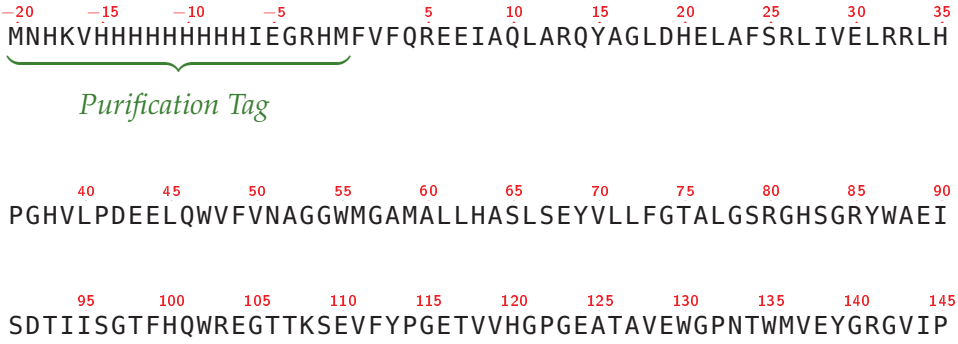
Purification Tag

```
−190      −185      −180      −175      −170      −165      −160      −155      −150      −145      −140
PTLVTTLTYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEV
```

Purification Tag



B.8.2 σ_1-Δ_{35}



```
      150       155       160       165       170       175       180       185
STLAFALADTVFSTQDFLTLFYTLRSYARGLRLELTTYLFGQDP
```

BIBLIOGRAPHY

1. Clerico, E. M., Tilitsky, J. M., Meng, W. & Gierasch, L. M. How Hsp70 Molecular Machines Interact with Their Substrates to Mediate Diverse Physiological Functions. en. *Journal of Molecular Biology* **427,** 1575–1588. ISSN: 00222836 (Apr. 2015).

2. Dobson, C. M. Protein folding and misfolding. eng. *Nature* **426,** 884–890. ISSN: 1476-4687 (Dec. 2003).

3. Liu, T., Daniels, C. K. & Cao, S. Comprehensive review on the HSC70 functions, interactions with related molecules and involvement in clinical diseases and therapeutic potential. *Pharmacology & Therapeutics* **136,** 354–374. ISSN: 01637258 (Dec. 2012).

4. Evans, C. G., Chang, L. & Gestwicki, J. E. Heat Shock Protein 70 (Hsp70) as an Emerging Drug Target. *Journal of Medicinal Chemistry* **53,** 4585–4602. ISSN: 0022-2623, 1520-4804 (June 2010).

5. Mayer, M. P. & Bukau, B. Hsp70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences* **62,** 670–684. ISSN: 1420-682X, 1420-9071 (Mar. 2005).

6. Burgoyne, R. D. & Morgan, A. Cysteine string protein (CSP) and its role in preventing neurodegeneration. *Seminars in Cell & Developmental Biology* **40,** 153–159. ISSN: 10849521 (Apr. 2015).

7. Gao, X.-C., Zhou, C.-J., Zhou, Z.-R., Zhang, Y.-H., Zheng, X.-M., Song, A.-X. & Hu, H.-Y. Co-Chaperone HSJ1a Dually Regulates the Proteasomal Degradation of Ataxin-3. *PLoS ONE* **6** (ed Chu, C. T.) e19763. ISSN: 1932-6203 (May 2011).

8. Gillis, J. *et al.* The DNAJB6 and DNAJB8 Protein Chaperones Prevent Intracellular Aggregation of Polyglutamine Peptides. *Journal of Biological Chemistry* **288,** 17225–17237. ISSN: 0021-9258, 1083-351X (June 2013).

9. Alderson, T., Kim, J. & Markley, J. Dynamical Structures of Hsp70 and Hsp70-Hsp40 Complexes. *Structure* **24,** 1014–1030. ISSN: 09692126 (July 2016).

10. Maurice, T. & Su, T.-P. The pharmacology of sigma-1 receptors. *Pharmacology & Therapeutics* **124,** 195–206. ISSN: 01637258 (Nov. 2009).

11. Martin, W. R., Eades, C. G., Thompson, J. A., Huppler, R. E. & Gilbert, P. E. The effects of morphine- and nalorphine- like drugs in the nondependent and morphine-dependent chronic spinal dog. *The Journal of Pharmacology and Experimental Therapeutics* **197,** 517–532. ISSN: 0022-3565 (June 1976).

12. Schmidt, H. R., Zheng, S., Gurpinar, E., Koehl, A., Manglik, A. & Kruse, A. C. Crystal structure of the human σ_1 receptor. *Nature* **532,** 527–530. ISSN: 0028-0836, 1476-4687 (Apr. 2016).

13. Ossa, F., Schnell, J. R. & Ortega-Roldan, J. L. in *Sigma Receptors: Their Role in Disease and as Therapeutic Targets* 15–29 (Springer International Publishing, Cham, 2017). ISBN: 978-3-319-50174-1.

14. Hayashi, T. & Su, T.-P. Sigma-1 Receptor Chaperones at the ER- Mitochondrion Interface Regulate Ca^{2+} Signaling and Cell Survival. *Cell* **131,** 596–610. ISSN: 00928674 (Nov. 2007).

15. Aydar, E., Palmer, C. P., Klyachko, V. A. & Jackson, M. B. The Sigma Receptor as a Ligand-Regulated Auxiliary Potassium Channel Subunit. *Neuron* **34,** 399–410. ISSN: 08966273 (Apr. 2002).

16. Ortega-Roldan, J. L., Ossa, F., Amin, N. T. & Schnell, J. R. Solution NMR studies reveal the location of the second transmembrane domain of the human sigma-1 receptor. *FEBS Letters* **589,** 659–665. ISSN: 1873-3468 (Feb. 2015).

17. Gromek, K. A., Suchy, F. P., Meddaugh, H. R., Wrobel, R. L., LaPointe, L. M., Chu, U. B., Primm, J. G., Ruoho, A. E., Senes, A. & Fox, B. G. The Oligomeric

States of the Purified Sigma-1 Receptor Are Stabilized by Ligands. *The Journal of Biological Chemistry* **289,** 20333–20344. ISSN: 0021-9258 (July 2014).

18. Schmidt, H. R., Betz, R. M., Dror, R. O. & Kruse, A. C. Structural basis for sigma-1 receptor ligand recognition (June 2018).

19. Su, T. P. Evidence for sigma opioid receptor: binding of [^{3}H]SKF-10047 to etorphine-inaccessible sites in guinea-pig brain. *The Journal of Pharmacology and Experimental Therapeutics* **223,** 284–290. ISSN: 0022-3565 (Nov. 1982).

20. Hellewell, S. B. & Bowen, W. D. A sigma-like binding site in rat pheochromocytoma (PC12) cells: decreased affinity for (+)-benzomorphans and lower molecular weight suggest a different sigma receptor form from that of guinea pig brain. *Brain Research* **527,** 244–253. ISSN: 0006-8993 (Sept. 1990).

21. Hanner, M., Moebius, F. F., Flandorfer, A., Knaus, H. G., Striessnig, J., Kempner, E. & Glossmann, H. Purification, molecular cloning, and expression of the mammalian sigma1-binding site. *Proceedings of the National Academy of Sciences* **93,** 8072–8077. ISSN: 0027-8424, 1091-6490 (July 1996).

22. Alon, A., Schmidt, H. R., Wood, M. D., Sahn, J. J., Martin, S. F. & Kruse, A. C. Identification of the gene that codes for the σ_2 receptor. *Proceedings of the National Academy of Sciences of the United States of America* **114,** 7160–7165. ISSN: 1091-6490 (July 2017).

23. Ishikawa, M. *et al.* High Occupancy of Sigma-1 Receptors in the Human Brain after Single Oral Administration of Fluvoxamine: A Positron Emission Tomography Study Using [11C]SA4503. *Biological Psychiatry* **62,** 878–883. ISSN: 00063223 (Oct. 2007).

24. Sugimoto, Y., Tagawa, N., Kobayashi, Y., Mitsui-Saito, K., Hotta, Y. & Yamada, J. Involvement of the sigma1 receptor in the antidepressant-like effects of fluvoxamine in the forced swimming test in comparison with the effects elicited by paroxetine. *European Journal of Pharmacology* **696,** 96–100. ISSN: 00142999 (Dec. 2012).

25. Fontanilla, D., Johannessen, M., Hajipour, A. R., Cozzi, N. V., Jackson, M. B. & Ruoho, A. E. The Hallucinogen N,N-Dimethyltryptamine (DMT) Is an Endogenous Sigma-1 Receptor Regulator. *Science* **323,** 934–937. ISSN: 0036-8075, 1095-9203 (Feb. 2009).

26. Sharkey, J., Glen, K. A., Wolfe, S. & Kuhar, M. J. Cocaine binding at sigma receptors. *European Journal of Pharmacology* **149,** 171–174. ISSN: 0014-2999 (Apr. 1988).

27. Su, T. P., London, E. D. & Jaffe, J. H. Steroid binding at sigma receptors suggests a link between endocrine, nervous, and immune systems. *Science* **240,** 219–221. ISSN: 0036-8075, 1095-9203 (Apr. 1988).

28. Johannessen, M., Fontanilla, D., Mavlyutov, T., Ruoho, A. E. & Jackson, M. B. Antagonist action of progesterone at σ-receptors in the modulation of voltage-gated sodium channels. *American Journal of Physiology-Cell Physiology* **300,** C328–C337. ISSN: 0363-6143, 1522-1563 (Feb. 2011).

29. Ramachandran, S., Chu, U. B., Mavlyutov, T. A., Pal, A., Pyne, S. & Ruoho, A. E. The sigma1 receptor interacts with N-alkyl amines and endogenous sphingolipids. *European Journal of Pharmacology* **609,** 19–26. ISSN: 00142999 (May 2009).

30. Tsai, S.-Y. A., Pokrass, M. J., Klauer, N. R., Nohara, H. & Su, T.-P. Sigma-1 receptor regulates Tau phosphorylation and axon extension by shaping p35 turnover via myristic acid. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 6742–6747. ISSN: 0027-8424 (May 2015).

31. Nguyen, L., Lucke-Wolds, B. P., Mookerjee, S., Kaushal, N. & Matsumoto, R. R. Sigma-1 Receptors and Neurodegenerative Diseases: Towards a Hypothesis of Sigma-1 Receptors as Amplifiers of Neurodegeneration and Neuroprotection. *Advances in experimental medicine and biology* **964,** 133–152. ISSN: 0065-2598 (2017).

32. Ajmo, C. T., Vernon, D. O. L., Collier, L., Pennypacker, K. R. & Cuevas, J. Sigma receptor activation reduces infarct size at 24 hours after permanent

middle cerebral artery occlusion in rats. *Current Neurovascular Research* **3,** 89–98. ISSN: 1567-2026 (May 2006).

33. Maurice, T., Su, T. P. & Privat, A. Sigma1 (sigma 1) receptor agonists and neurosteroids attenuate B25-35-amyloid peptide-induced amnesia in mice through a common mechanism. *Neuroscience* **83,** 413–428. ISSN: 0306-4522 (Mar. 1998).

34. Meunier, J., Ieni, J. & Maurice, T. The anti-amnesic and neuroprotective effects of donepezil against amyloid beta25-35 peptide-induced toxicity in mice involve an interaction with the sigma1 receptor. eng. *British Journal of Pharmacology* **149,** 998–1012. ISSN: 0007-1188 (Dec. 2006).

35. Marrazzo, A., Caraci, F., Salinaro, E. T., Su, T.-P., Copani, A. & Ronsisvalle, G. Neuroprotective effects of sigma-1 receptor agonists against beta-amyloid-induced toxicity. *Neuroreport* **16,** 1223–1226. ISSN: 0959-4965 (Aug. 2005).

36. Su, T.-P., Hayashi, T., Maurice, T., Buch, S. & Ruoho, A. E. The sigma-1 receptor chaperone as an inter-organelle signaling modulator. *Trends in Pharmacological Sciences* **31,** 557–566. ISSN: 01656147 (Dec. 2010).

37. Rodríguez-Muñoz, M., Sánchez-Blázquez, P., Herrero-Labrador, R., Martínez-Murillo, R., Merlos, M., Vela, J. M. & Garzón, J. The σ_1 Receptor Engages the Redox-Regulated HINT1 Protein to Bring Opioid Analgesia Under NMDA Receptor Negative Control. *Antioxidants & Redox Signaling* **22,** 799–818. ISSN: 1523-0864 (Jan. 2015).

38. Mavlyutov, T. A., Guo, L.-W., Epstein, M. L. & Ruoho, A. E. Role of the Sigma-1 receptor in Amyotrophic Lateral Sclerosis (ALS). eng. *Journal of Pharmacological Sciences* **127,** 10–16. ISSN: 1347-8648 (Jan. 2015).

39. Moebius, F. F., Reiter, R. J., Hanner, M. & Glossmann, H. High affinity of sigma1-binding sites for sterol isomerization inhibitors: evidence for a pharmacological relationship with the yeast sterol C8–C7 isomerase. *British Journal of Pharmacology* **121,** 1–6. ISSN: 0007-1188 (May 1997).

40. Glennon, R. A., Ablordeppey, S. Y., Ismaiel, A. M., El-Ashmawy, M. B., Fischer, J. B. & Howie, K. B. Structural Features Important for σ1 Receptor Binding. *Journal of Medicinal Chemistry* **37,** 1214–1219. ISSN: 0022-2623, 1520-4804 (Apr. 1994).

41. Chen, Y., Hajipour, A. R., Sievert, M. K., Arbabian, M. & Ruoho, A. E. Characterization of the Cocaine Binding Site on the Sigma-1 Receptor. *Biochemistry* **46,** 3532–3542. ISSN: 0006-2960, 1520-4995 (Mar. 2007).

42. Palmer, C. P., Mahen, R., Schnell, E., Djamgoz, M. B. & Aydar, E. Sigma-1 Receptors Bind Cholesterol and Remodel Lipid Rafts in Breast Cancer Cell Lines. *Cancer Research* **67,** 11166–11175. ISSN: 0008-5472, 1538-7445 (Dec. 2007).

43. Tsai, S.-Y. A., Chuang, J.-Y., Tsai, M.-S., Wang, X.-f., Xi, Z.-X., Hung, J.-J., Chang, W.-C., Bonci, A. & Su, T.-P. Sigma-1 receptor mediates cocaine-induced transcriptional regulation by recruiting chromatin-remodeling factors at the nuclear envelope. *Proceedings of the National Academy of Sciences* **112,** E6562–E6570. ISSN: 0027-8424, 1091-6490 (Nov. 2015).

44. Chu, U. B. & Ruoho, A. E. Biochemical Pharmacology of the Sigma-1 Receptor. *Molecular Pharmacology* **89,** 142–153. ISSN: 0026-895X, 1521-0111 (Jan. 2016).

45. Mishra, A. K., Mavlyutov, T., Singh, D. R., Biener, G., Yang, J., Oliver, J., Ruoho, A. & Raicu, V. The sigma-1 receptors are present in monomeric and oligomeric forms in living cells in the presence and absence of ligands. *Biochemical Journal* **466,** 263–271. ISSN: 0264-6021, 1470-8728 (Mar. 2015).

46. Hayashi, T. Intracellular Dynamics of -1 Receptors (1 Binding Sites) in NG108-15 Cells. *Journal of Pharmacology and Experimental Therapeutics* **306,** 726–733. ISSN: 0022-3565, 1521-0103 (Apr. 2003).

47. Srivats, S., Balasuriya, D., Pasche, M., Vistal, G., Edwardson, J. M., Taylor, C. W. & Murrell-Lagnado, R. D. Sigma1 receptors inhibit store-operated

Ca^{2+} entry by attenuating coupling of STIM1 to Orai1. *The Journal of Cell Biology* **213,** 65–79. ISSN: 0021-9525 (Apr. 2016).

48. Kim, F. J., Kovalyshyn, I., Burgman, M., Neilan, C., Chien, C.-C. & Pasternak, G. W. 1 Receptor Modulation of G-Protein-Coupled Receptor Signaling: Potentiation of Opioid Transduction Independent from Receptor Binding. *Molecular Pharmacology* **77,** 695–703. ISSN: 0026-895X (Apr. 2010).

49. Gao, X.-F., Yao, J.-J., He, Y.-L., Hu, C. & Mei, Y.-A. Sigma-1 Receptor Agonists Directly Inhibit NaV1.2/1.4 Channels. *PLOS ONE* **7,** e49384. ISSN: 1932-6203 (Nov. 2012).

50. Kourrich, S., Hayashi, T., Chuang, J.-Y., Tsai, S.-Y., Su, T.-P. & Bonci, A. Dynamic Interaction between Sigma-1 Receptor and Kv1.2 Shapes Neuronal and Behavioral Responses to Cocaine. *Cell* **152,** 236–247. ISSN: 00928674 (Jan. 2013).

51. Kinoshita, M., Matsuoka, Y., Suzuki, T., Mirrielees, J. & Yang, J. Sigma-1 receptor alters the kinetics of Kv1.3 voltage gated potassium channels but not the sensitivity to receptor ligands. *Brain Research* **1452,** 1–9. ISSN: 00068993 (May 2012).

52. Morin-Surun, M. P., Collin, T., Denavit-Saubie, M., Baulieu, E.-E. & Monnet, F. P. Intracellular σ_1 receptor modulates phospholipase C and protein kinase C activities in the brainstem. *Proceedings of the National Academy of Sciences* **96,** 8196–8199. ISSN: 0027-8424, 1091-6490 (July 1999).

53. Balasuriya, D., D'Sa, L., Talker, R., Dupuis, E., Maurin, F., Martin, P., Borgese, F., Soriani, O. & Edwardson, J. M. A Direct Interaction between the Sigma-1 Receptor and the hERG Voltage-gated K^+ Channel Revealed by Atomic Force Microscopy and Homogeneous Time-resolved Fluorescence (HTRF). *Journal of Biological Chemistry* **289,** 32353–32363. ISSN: 0021-9258, 1083-351X (Nov. 2014).

54. Navarro, G. *et al.* Cocaine Inhibits Dopamine D2 Receptor Signaling via Sigma-1-D2 Receptor Heteromers. *PLoS ONE* **8,** e61245. ISSN: 1932-6203 (Apr. 2013).

55. Balasuriya, D., Stewart, A. P., Crottès, D., Borgese, F., Soriani, O. & Edwardson, J. M. The Sigma-1 Receptor Binds to the Nav1.5 Voltage-gated Na^+ Channel with 4-Fold Symmetry. *Journal of Biological Chemistry* **287,** 37021–37029. ISSN: 0021-9258, 1083-351X (Oct. 2012).

56. Carnally, S. M., Johannessen, M., Henderson, R. M., Jackson, M. B. & Edwardson, J. M. Demonstration of a Direct Interaction between σ-1 Receptors and Acid-Sensing Ion Channels. *Biophysical Journal* **98,** 1182–1191. ISSN: 00063495 (Apr. 2010).

57. Balasuriya, D., Stewart, A. P. & Edwardson, J. M. The σ-1 Receptor Interacts Directly with GluN1 But Not GluN2A in the GluN1/GluN2A NMDA Receptor. *Journal of Neuroscience* **33,** 18219–18224 (Nov. 2013).

58. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. UCSF Chimera–A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25,** 1605–1612. ISSN: 0192-8651, 1096-987X (Oct. 2004).

59. Ortega-Roldan, J. L., Ossa, F. & Schnell, J. R. Characterization of the Human Sigma-1 Receptor Chaperone Domain Structure and Binding Immunoglobulin Protein (BiP) Interactions. *Journal of Biological Chemistry* **288,** 21448–21457. ISSN: 0021-9258, 1083-351X (July 2013).

60. Laskowski, R. A. & Swindells, M. B. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *Journal of Chemical Information and Modeling* **51,** 2778–2786. ISSN: 1549-9596, 1549-960X (Oct. 2011).

61. Seth, P., Ganapathy, M. E., Conway, S. J., Bridges, C. D., Smith, S. B., Casellas, P. & Ganapathy, V. Expression pattern of the type 1 sigma receptor in the brain and identity of critical anionic amino acid residues in the ligand-

binding domain of the receptor. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1540,** 59–67. ISSN: 01674889 (July 2001).

62. Pal, A., Hajipour, A. R., Fontanilla, D., Ramachandran, S., Chu, U. B., Mavlyutov, T. & Ruoho, A. E. Identification of Regions of the σ-1 Receptor Ligand Binding Site Using a Novel Photoprobe. *Molecular Pharmacology* **72,** 921–933. ISSN: 0026-895X (July 2007).

63. Ramachandran, S., Lu, H., Prabhu, U. & Ruoho, A. E. Purification and characterization of the guinea pig sigma-1 receptor functionally expressed in Escherichia coli. *Protein Expression and Purification* **51,** 283–292. ISSN: 1046-5928 (Feb. 2007).

64. Chu, U. B., Ramachandran, S., Hajipour, A. R. & Ruoho, A. E. Photoaffinity Labeling of the Sigma-1 Receptor with *N* -[3-(4-Nitrophenyl)propyl]- *N* -dodecylamine: Evidence of Receptor Dimers. *Biochemistry* **52,** 859–868. ISSN: 0006-2960, 1520-4995 (Feb. 2013).

65. Pal, A., Chu, U. B., Ramachandran, S., Grawoig, D., Guo, L.-W., Hajipour, A. R. & Ruoho, A. E. Juxtaposition of the Steroid Binding Domain-like I and II Regions Constitutes a Ligand Binding Site in the σ_1 Receptor. *The Journal of Biological Chemistry* **283,** 19646–19656. ISSN: 0021-9258 (July 2008).

66. Hayashi, T. & Su, T. P. Regulating ankyrin dynamics: Roles of sigma-1 receptors. *Proceedings of the National Academy of Sciences of the United States of America* **98,** 491–496. ISSN: 0027-8424 (Jan. 2001).

67. Mavlyutov, T. A., Duellman, T., Kim, H. T., Epstein, M. L., Leese, C., Davletov, B. A. & Yang, J. Sigma-1 receptor expression in the dorsal root ganglion: Reexamination using a highly specific antibody. *Neuroscience* **331,** 148–157. ISSN: 03064522 (Sept. 2016).

68. Wong, A. Y. C., Hristova, E., Ahlskog, N., Tasse, L.-A., Ngsee, J. K., Chudalayandi, P. & Bergeron, R. Aberrant Subcellular Dynamics of Sigma-1 Receptor Mutants Underlying Neuromuscular Diseases. *Molecular Pharmacology* **90,** 238–253. ISSN: 1521-0111 (Aug. 2016).

69. Mavylutov, T., Chen, X., Guo, L. & Yang, J. APEX2- tagging of Sigma 1-receptor indicates subcellular protein topology with cytosolic N-terminus and ER luminal C-terminus. *Protein & Cell* **9,** 733–737. ISSN: 1674-800X (Aug. 2018).

70. Ritossa, F. A new puffing pattern induced by temperature shock and DNP in drosophila. *Experientia* **18,** 571–573. ISSN: 0014-4754, 1420-9071 (Dec. 1962).

71. Blatch, G. L. & Edkins, A. L. *The Networking of Chaperones by Co-chaperones: Control of Cellular Protein Homeostasis* ISBN: 978-3-319-11731-7 (Springer, Dec. 2014).

72. Young, J. C., Agashe, V. R., Siegers, K. & Hartl, F. U. Pathways of chaperone-mediated protein folding in the cytosol. *Nature Reviews Molecular Cell Biology* **5,** 781–791 (Oct. 2004).

73. Finka, A., Sharma, S. K. & Goloubinoff, P. Multi-layered molecular mechanisms of polypeptide holding, unfolding and disaggregation by HSP70/HSP110 chaperones. *Frontiers in Molecular Biosciences* **2.** ISSN: 2296-889X. (2018) (June 2015).

74. Chappell, T. G., Konforti, B. B., Schmid, S. L. & Rothman, J. E. The ATPase core of a clathrin uncoating protein. *The Journal of Biological Chemistry* **262,** 746–751. ISSN: 0021-9258 (Jan. 1987).

75. Mayer, M. P. Hsp70 chaperone dynamics and molecular mechanism. *Trends in Biochemical Sciences* **38,** 507–514. ISSN: 09680004 (Oct. 2013).

76. Kampinga, H. H. & Craig, E. A. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology* **11,** 579–592. ISSN: 1471-0072, 1471-0080 (Aug. 2010).

77. Flaherty, K. M., DeLuca-Flaherty, C. & McKay, D. B. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* **346,** 623–628. ISSN: 0028-0836 (Aug. 1990).

78. Flaherty, K. M., Wilbanks, S. M., DeLuca-Flaherty, C. & McKay, D. B. Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. II. Structure of the active site with ADP or ATP bound to wild type and mutant ATPase fragment. *The Journal of Biological Chemistry* **269,** 12899–12907. ISSN: 0021-9258 (Apr. 1994).

79. Gołaś, E. I., Czaplewski, C., Scheraga, H. A. & Liwo, A. Common functionally important motions of the nucleotide-binding domain of Hsp70. *Proteins: Structure, Function, and Bioinformatics* **83,** 282–299. ISSN: 1097-0134 (Feb. 2015).

80. Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E. & Hendrickson, W. A. Structural Analysis of Substrate Binding by the Molecular Chaperone DnaK. en. *Science* **272,** 1606–1614. ISSN: 0036-8075, 1095-9203 (June 1996).

81. Aprile, F. A., Dhulesia, A., Stengel, F., Roodveldt, C., Benesch, J. L. P., Tortora, P., Robinson, C. V., Salvatella, X., Dobson, C. M. & Cremades, N. Hsp70 Oligomerization Is Mediated by an Interaction between the Interdomain Linker and the Substrate-Binding Domain. *PLOS ONE* **8,** e67961. ISSN: 1932-6203 (June 2013).

82. Chou, C.-C., Forouhar, F., Yeh, Y.-H., Shr, H.-L., Wang, C. & Hsiao, C.-D. Crystal Structure of the C-terminal 10-kDa Subdomain of Hsc70. *Journal of Biological Chemistry* **278,** 30311–30316. ISSN: 0021-9258, 1083-351X (Aug. 2003).

83. Nemoto, T. K., Fukuma, Y., Itoh, H., Takagi, T. & Ono, T. A Disulfide Bridge Mediated by Cysteine 574 Is Formed in the Dimer of the 70-kDa Heat Shock Protein. *The Journal of Biochemistry* **139,** 677–687. ISSN: 1756-2651, 0021-924X (Apr. 2006).

84. Benaroudj, N., Fouchaq, B. & Ladjimi, M. M. The COOH-terminal peptide binding domain is essential for self-association of the molecular chaperone HSC70. *The Journal of Biological Chemistry* **272,** 8744–8751. ISSN: 0021-9258 (Mar. 1997).

85. Angelidis, C. E., Lazaridis, I. & Pagoulatos, G. N. Aggregation of hsp70 and hsc70 *in vivo* is distinct and temperature-dependent and their chaperone function is directly related to non-aggregated forms. *European Journal of Biochemistry* **259,** 505–512. ISSN: 00142956, 14321033 (Jan. 1999).

86. Kim, D., Lee, Y. J. & Corry, P. M. Constitutive HSP70: Oligomerization and its dependence on ATP binding. *Journal of Cellular Physiology* **153,** 353–361. ISSN: 0021-9541, 1097-4652 (Nov. 1992).

87. Sarbeng, E. B., Liu, Q., Tian, X., Yang, J., Li, H., Wong, J. L., Zhou, L. & Liu, Q. A Functional DnaK Dimer Is Essential for the Efficient Interaction with Hsp40 Heat Shock Protein. *Journal of Biological Chemistry* **290,** 8849–8862. ISSN: 0021-9258, 1083-351X (Apr. 2015).

88. Morgner, N. *et al.* Hsp70 Forms Antiparallel Dimers Stabilized by Post-translational Modifications to Position Clients for Transfer to Hsp90. *Cell Reports* **11,** 759–769. ISSN: 22111247 (May 2015).

89. Sousa, R. A Dancer Caught Midstep: The Structure of ATP-Bound Hsp70. *Molecular Cell* **48,** 821–823. ISSN: 10972765 (Dec. 2012).

90. Kityk, R., Kopp, J., Sinning, I. & Mayer, M. Structure and Dynamics of the ATP-Bound Open Conformation of Hsp70 Chaperones. *Molecular Cell* **48,** 863–874. ISSN: 10972765 (Dec. 2012).

91. Bertelsen, E. B., Chang, L., Gestwicki, J. E. & Zuiderweg, E. R. P. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proceedings of the National Academy of Sciences* **106,** 8471–8476. ISSN: 0027-8424, 1091-6490 (May 2009).

92. Chang, Y.-W., Sun, Y.-J., Wang, C. & Hsiao, C.-D. Crystal Structures of the 70-kDa Heat Shock Proteins in Domain Disjoining Conformation. *Journal of Biological Chemistry* **283,** 15502–15511. ISSN: 0021-9258, 1083-351X (May 2008).

93. McCarty, J. S., Buchberger, A., Reinstein, J. & Bukau, B. The Role of ATP in the Functional Cycle of the DnaK Chaperone System. *Journal of Molecular Biology* **249,** 126–137. ISSN: 00222836 (May 1995).

94. Flynn, G. C., Pohl, J., Flocco, M. T. & Rothman, J. E. Peptide-binding specificity of the molecular chaperone BiP. *Nature* **353,** 726–730. ISSN: 0028-0836 (Oct. 1991).

95. Gruschus, J. M., Greene, L. E., Eisenberg, E. & Ferretti, J. A. Experimentally biased model structure of the Hsc70/auxilin complex: Substrate transfer and interdomain structural change. *Protein Science : A Publication of the Protein Society* **13,** 2029–2044. ISSN: 0961-8368 (Aug. 2004).

96. Swain, J. F., Dinler, G., Sivendran, R., Montgomery, D. L., Stotz, M. & Gierasch, L. M. Hsp70 Chaperone Ligands Control Domain Association via an Allosteric Mechanism Mediated by the Interdomain Linker. *Molecular Cell* **26,** 27–39. ISSN: 1097-2765 (Apr. 2007).

97. Vogel, M., Bukau, B. & Mayer, M. P. Allosteric Regulation of Hsp70 Chaperones by a Proline Switch. *Molecular Cell* **21,** 359–367. ISSN: 10972765 (Feb. 2006).

98. Vogel, M., Mayer, M. P. & Bukau, B. Allosteric Regulation of Hsp70 Chaperones Involves a Conserved Interdomain Linker. *Journal of Biological Chemistry* **281,** 38705–38711. ISSN: 0021-9258, 1083-351X (Dec. 2006).

99. Marcinowski, M., Höller, M., Feige, M. J., Baerend, D., Lamb, D. C. & Buchner, J. Substrate discrimination of the chaperone BiP by autonomous and cochaperone-regulated conformational transitions. *Nature Structural & Molecular Biology* **18,** 150–158. ISSN: 1545-9993, 1545-9985 (Feb. 2011).

100. Sikor, M., Mapa, K., von Voithenberg, L. V., Mokranjac, D. & Lamb, D. C. Real-time observation of the conformational dynamics of mitochondrial Hsp70 by spFRET. en. *The EMBO Journal* **32,** 1639–1649. ISSN: 0261-4189, 1460-2075 (Apr. 2013).

101. Wieteska, L., Shahidi, S. & Zhuravleva, A. Allosteric fine-tuning of the conformational equilibrium poises the chaperone BiP for post-translational regulation. *eLife* **6,** e29430. ISSN: 2050-084X (Oct. 2017).

102. Bracher, A. & Verghese, J. in *The Networking of Chaperones by Co-chaperones* 1–33 (Springer, Cham, 2015). ISBN: 978-3-319-11730-0 978-3-319-11731-7.

103. Wall, D., Zylicz, M. & Georgopoulos, C. The NH2-terminal 108 amino acids of the Escherichia coli DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. *The Journal of Biological Chemistry* **269,** 5446–5451. ISSN: 0021-9258 (Feb. 1994).

104. Greene, M. K., Maskos, K. & Landry, S. J. Role of the J-domain in the cooperation of Hsp40 with Hsp70. *Proceedings of the National Academy of Sciences* **95,** 6108–6113. ISSN: 0027-8424, 1091-6490 (May 1998).

105. Pellecchia, M., Szyperski, T., Wall, D., Georgopoulos, C. & Wüthrich, K. NMR Structure of the J-domain and the Gly/Phe-rich Region of theEscherichia coliDnaJ Chaperone. *Journal of Molecular Biology* **260,** 236–250. ISSN: 00222836 (July 1996).

106. Braun, J. E. A., Wilbanks, S. M. & Scheller, R. H. The Cysteine String Secretory Vesicle Protein Activates Hsc70 ATPase. *Journal of Biological Chemistry* **271,** 25989–25993. ISSN: 0021-9258, 1083-351X (Oct. 1996).

107. Liberek, K., Wall, D. & Georgopoulos, C. The DnaJ chaperone catalytically activates the DnaK chaperone to preferentially bind the sigma 32 heat shock transcriptional regulator. *Proceedings of the National Academy of Sciences* **92,** 6224–6228. ISSN: 0027-8424, 1091-6490 (July 1995).

108. Cyr, D. M., Lu, X. & Douglas, M. G. Regulation of Hsp70 function by a eukaryotic DnaJ homolog. *Journal of Biological Chemistry* **267,** 20927–20931. ISSN: 0021-9258, 1083-351X (Oct. 1992).

109. Ahmad, A., Bhattacharya, A., McDonald, R. A., Cordes, M., Ellington, B., Bertelsen, E. B. & Zuiderweg, E. R. P. Heat shock protein 70 kDa chaperone/DnaJ cochaperone complex employs an unusual dynamic interface. *Proceedings of the National Academy of Sciences* **108,** 18966–18971. ISSN: 0027-8424, 1091-6490 (Nov. 2011).

110. Gassler, C. S., Buchberger, A., Laufen, T., Mayer, M. P., Schroder, H., Valencia, A. & Bukau, B. Mutations in the DnaK chaperone affecting interaction with the DnaJ cochaperone. *Proceedings of the National Academy of Sciences* **95,** 15229–15234. ISSN: 0027-8424, 1091-6490 (Dec. 1998).

111. Suh, W.-C., Burkholder, W. F., Lu, C. Z., Zhao, X., Gottesman, M. E. & Gross, C. A. Interaction of the Hsp70 molecular chaperone, DnaK, with its cochaperone DnaJ. *Proceedings of the National Academy of Sciences* **95,** 15223–15228. ISSN: 0027-8424, 1091-6490 (Dec. 1998).

112. Sousa, R., Jiang, J., Lafer, E. M., Hinck, A. P., Wang, L., Taylor, A. B. & Maes, E. G. Evaluation of Competing J domain:Hsp70 Complex Models in Light of Existing Mutational and NMR Data. *Proceedings of the National Academy of Sciences* **109,** E734–E734. ISSN: 0027-8424, 1091-6490 (Mar. 2012).

113. Kim, J. H., Alderson, T. R., Frederick, R. O. & Markley, J. L. Nucleotide-Dependent Interactions within a Specialized Hsp70/Hsp40 Complex Involved in Fe–S Cluster Biogenesis. *Journal of the American Chemical Society* **136,** 11586–11589. ISSN: 0002-7863, 1520-5126 (Aug. 2014).

114. Hageman, J. & Kampinga, H. H. Computational analysis of the human HSPH/HSPA/DNAJ family and cloning of a human HSPH/HSPA/DNAJ expression library. *Cell Stress and Chaperones* **14,** 1–21. ISSN: 1355-8145, 1466-1268 (Jan. 2009).

115. Radons, J. The human HSP70 family of chaperones: where do we stand? *Cell Stress and Chaperones* **21,** 379–404. ISSN: 1355-8145, 1466-1268 (May 2016).

116. Bole, D. G., Dowin, R, Doriaux, M & Jamieson, J. D. Immunocytochemical localization of BiP to the rough endoplasmic reticulum: evidence for protein sorting by selective retention. en. *Journal of Histochemistry & Cytochemistry* **37,** 1817–1823. ISSN: 0022-1554, 1551-5044 (Dec. 1989).

117. Bhattacharyya, T., Karnezis, A. N., Murphy, S. P., Hoang, T., Freeman, B. C., Phillips, B. & Morimoto, R. I. Cloning and subcellular localization of human

mitochondrial hsp70. eng. *The Journal of Biological Chemistry* **270,** 1705–1710. ISSN: 0021-9258 (Jan. 1995).

118. Otterson, G. A., Flynn, G. C., Kratzke, R. A., Coxon, A., Johnston, P. G. & Kaye, F. J. Stch encodes the 'ATPase core' of a microsomal stress 70 protein. *The EMBO journal* **13,** 1216–1225. ISSN: 0261-4189 (Mar. 1994).

119. Brocchieri, L., Conway de Macario, E. & Macario, A. J. hsp70 genes in the human genome: Conservation and differentiation patterns predict a wide array of overlapping and specialized functions. *BMC Evolutionary Biology* **8,** 19. ISSN: 1471-2148 (2008).

120. Zuiderweg, E. R. P., Hightower, L. E. & Gestwicki, J. E. The remarkable multivalency of the Hsp70 chaperones. *Cell Stress & Chaperones* **22,** 173–189. ISSN: 1355-8145 (Mar. 2017).

121. Grennan, A. K. Genevestigator. Facilitating Web-Based Gene-Expression Analysis. en. *PLANT PHYSIOLOGY* **141,** 1164–1166. ISSN: 0032-0889, 1532-2548 (Aug. 2006).

122. Rudiger, S. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *The EMBO Journal* **16,** 1501–1507. ISSN: 14602075 (Apr. 1997).

123. Gragerov, A., Zeng, L., Zhao, X., Burkholder, W. & Gottesman, M. E. Specificity of DnaK-peptide binding. *Journal of Molecular Biology* **235,** 848–854. ISSN: 0022-2836 (Jan. 1994).

124. Misselwitz, B., Staeck, O. & Rapoport, T. A. J Proteins Catalytically Activate Hsp70 Molecules to Trap a Wide Range of Peptide Sequences. *Molecular Cell* **2,** 593–603. ISSN: 10972765 (Nov. 1998).

125. Kelley, W. L. Molecular chaperones: How J domains turn on Hsp70s. *Current biology: CB* **9,** R305–308. ISSN: 0960-9822 (Apr. 1999).

126. Hennessy, F., Nicoll, W. S., Zimmermann, R., Cheetham, M. E. & Blatch, G. L. Not all J domains are created equal: Implications for the specificity of Hsp40-

Hsp70 interactions. *Protein Science* **14,** 1697–1709. ISSN: 09618368, 1469896X (July 2005).

127. Cry, D. M. Cooperation of the molecular chaperone Ydj1 with specific Hsp70 homologs to suppress protein aggregation. *FEBS Letters* **359,** 129–132. ISSN: 1873-3468 (Feb. 1995).

128. Schlenstedt, G., Harris, S., Risse, B., Lill, R. & Silver, P. A. A yeast DnaJ homologue, Scj1p, can function in the endoplasmic reticulum with BiP/Kar2p via a conserved domain that specifies interactions with Hsp70s. *The Journal of Cell Biology* **129,** 979–988. ISSN: 0021-9525, 1540-8140 (May 1995).

129. Tavaria, M., Gabriele, T., Kola, I. & Anderson, R. L. A Hitchhiker's Guide to the Human Hsp70 Family. *Cell Stress & Chaperones* **1,** 23–28. ISSN: 1355-8145 (1996).

130. Goldfarb, S. B., Kashlan, O. B., Watkins, J. N., Suaud, L., Yan, W., Kleyman, T. R. & Rubenstein, R. C. Differential effects of Hsc70 and Hsp70 on the intracellular trafficking and functional expression of epithelial sodium channels. *Proceedings of the National Academy of Sciences* **103,** 5817–5822. ISSN: 0027-8424, 1091-6490 (Apr. 2006).

131. Kakkar, V. *et al.* The S/T-Rich Motif in the DNAJB6 Chaperone Delays Polyglutamine Aggregation and the Onset of Disease in a Mouse Model. *Molecular Cell* **62,** 272–283. ISSN: 10972765 (Apr. 2016).

132. Chapple, J. P. & Cheetham, M. E. The Chaperone Environment at the Cytoplasmic Face of the Endoplasmic Reticulum Can Modulate Rhodopsin Processing and Inclusion Formation. *Journal of Biological Chemistry* **278,** 19087–19094. ISSN: 0021-9258, 1083-351X (May 2003).

133. Westhoff, B., Chapple, J. P., van der Spuy, J., Höhfeld, J. & Cheetham, M. E. HSJ1 Is a Neuronal Shuttling Factor for the Sorting of Chaperone Clients to the Proteasome. *Current Biology* **15,** 1058–1064. ISSN: 09609822 (June 2005).

134. Meng, E., Shevde, L. A. & Samant, R. S. Emerging roles and underlying molecular mechanisms of DNAJB6 in cancer. *Oncotarget* **7.** ISSN: 1949-2553. (2018) (Aug. 2016).

135. Ruggieri, A. *et al.* Complete loss of the DNAJB6 G/F domain and novel missense mutations cause distal-onset DNAJB6 myopathy. *Acta Neuropathologica Communications* **3.** ISSN: 2051-5960 (Dec. 2015).

136. Durrenberger, P. F., Filiou, M. D., Moran, L. B., Michael, G. J., Novoselov, S., Cheetham, M. E., Clark, P., Pearce, R. K. B. & Graeber, M. B. DnaJB6 is present in the core of Lewy bodies and is highly up-regulated in parkinsonian astrocytes. *Journal of Neuroscience Research* **87,** 238–245. ISSN: 1097-4547 (Jan. 2009).

137. Månsson, C., Kakkar, V., Monsellier, E., Sourigues, Y., Härmark, J., Kampinga, H. H., Melki, R. & Emanuelsson, C. DNAJB6 is a peptide-binding chaperone which can suppress amyloid fibrillation of polyglutamine peptides at substoichiometric molar ratios. *Cell Stress and Chaperones* **19,** 227–239. ISSN: 1355-8145, 1466-1268 (Mar. 2014).

138. Kakkar, V., Kuiper, E. F. E., Pandey, A., Braakman, I. & Kampinga, H. H. Versatile members of the DNAJ family show Hsp70 dependent anti-aggregation activity on RING1 mutant parkin C289G. *Scientific Reports* **6.** ISSN: 2045-2322 (Dec. 2016).

139. Söderberg, C. A. G. *et al.* Structural modelling of the DNAJB6 oligomeric chaperone shows a peptide-binding cleft lined with conserved S/T-residues at the dimer interface. *Scientific Reports* **8.** ISSN: 2045-2322 (Mar. 2018).

140. Månsson, C. *et al.* Conserved S/T Residues of the Human Chaperone DNAJB6 Are Required for Effective Inhibition of Aβ42 Amyloid Fibril Formation. *Biochemistry* **57,** 4891–4902. ISSN: 0006-2960 (Aug. 2018).

141. Patel, P., Prescott, G., Burgoyne, R., Lian, L.-Y. & Morgan, A. Phosphorylation of Cysteine String Protein Triggers a Major Conformational Switch. en. *Structure* **24,** 1380–1386. ISSN: 09692126 (Aug. 2016).

142. Zinsmaier, K. E., Hofbauer, A., Heimbeck, G., Pflugfelder, G. O., Buchner, S. & Buchner, E. A cysteine-string protein is expressed in retina and brain of Drosophila. *Journal of Neurogenetics* **7,** 15–29. ISSN: 0167-7063 (Nov. 1990).

143. Stahl, B., Tobaben, S. & Sudhof, T. C. Two distinct domains in hsc70 are essential for the interaction with the synaptic vesicle cysteine string protein. *European Journal of Cell Biology* **78,** 375–381. ISSN: 01719335 (June 1999).

144. Sharma, M., Burré, J. & Südhof, T. C. CSPα promotes SNARE-complex assembly by chaperoning SNAP-25 during synaptic activity. *Nature Cell Biology* **13,** 30–39. ISSN: 1465-7392, 1476-4679 (Jan. 2011).

145. Sharma, M., Burré, J., Bronk, P., Zhang, Y., Xu, W. & Südhof, T. C. CSPα knockout causes neurodegeneration by impairing SNAP-25 function: Mechanism of CSPα KO-induced neurodegeneration. *The EMBO Journal* **31,** 829–841. ISSN: 02614189 (Feb. 2012).

146. Gao, X.-C., Zhou, C.-J., Zhou, Z.-R., Wu, M., Cao, C.-Y. & Hu, H.-Y. The C-terminal Helices of Heat Shock Protein 70 Are Essential for J-domain Binding and ATPase Activation. *Journal of Biological Chemistry* **287,** 6044–6052. ISSN: 0021-9258, 1083-351X (Feb. 2012).

147. Zuiderweg, E. R. P. & Gestwicki, J. E. Backbone and methyl resonance assignments of the 42 kDa human Hsc70 nucleotide binding domain in the ADP state. *Biomolecular NMR Assignments* **11,** 11–15. ISSN: 1874-2718, 1874-270X (Apr. 2017).

148. Wisniewska, M., Karlberg, T., Lehtiö, L., Johansson, I., Kotenyova, T., Moche, M. & Schüler, H. Crystal Structures of the ATPase Domains of Four Human Hsp70 Isoforms: HSPA1L/Hsp70-hom, HSPA2/Hsp70-2, HSPA6/Hsp70B', and HSPA5/BiP/GRP78. *PLoS ONE* **5** (ed Gay, N.) e8625. ISSN: 1932-6203 (Jan. 2010).

149. Bulheller, B. M. & Hirst, J. D. DichroCalc–circular and linear dichroism online. *Bioinformatics* **25,** 539–540. ISSN: 1367-4803, 1460-2059 (Feb. 2009).

150. Jasim, S. B., Li, Z., Guest, E. E. & Hirst, J. D. DichroCalc: Improvements in Computing Protein Circular Dichroism Spectroscopy in the Near-Ultraviolet. *Journal of Molecular Biology* **430,** 2196–2202. ISSN: 00222836 (July 2018).

151. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J. & Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR* **6.** ISSN: 0925-2738, 1573-5001. (2018) (Nov. 1995).

152. Norris, M., Fetler, B., Marchant, J. & Johnson, B. A. NMRFx Processor: a cross-platform NMR data processing program. *Journal of Biomolecular NMR* **65,** 205–216. ISSN: 0925-2738, 1573-5001 (Aug. 2016).

153. Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H. & Wagner, G. Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *Journal of Biomolecular NMR* **52,** 315–327. ISSN: 0925-2738, 1573-5001 (Apr. 2012).

154. Lee, D., Hilty, C., Wider, G. & Wüthrich, K. Effective rotational correlation times of proteins from NMR relaxation interference. *Journal of Magnetic Resonance* **178,** 72–76. ISSN: 10907807 (Jan. 2006).

155. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *Journal of Biomolecular NMR* **56,** 227–241. ISSN: 0925-2738, 1573-5001 (July 2013).

156. Nolan, R., Alvarez, L. A. J., Elegheert, J., Iliopoulou, M., Jakobsdottir, G. M., Rodriguez-Muñoz, M., Aricescu, A. R., Padilla-Parra, S. & Murphy, R. nandb—number and brightness in R with a novel automatic detrending algorithm. *Bioinformatics* **33,** 3508–3510. ISSN: 1367-4803 (Nov. 2017).

157. Schutze, M. P., Peterson, P. A. & Jackson, M. R. An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum. *The EMBO journal* **13,** 1696–1705. ISSN: 0261-4189 (Apr. 1994).

158. Chipot, C. *et al.* Perturbations of Native Membrane Protein Structure in Alkyl Phosphocholine Detergents: A Critical Assessment of NMR and Biophysical Studies. *Chemical Reviews* **118,** 3559–3607. ISSN: 1520-6890 (Apr. 2018).

159. Le Maire, M., Champeil, P. & Møller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochimica et Biophysica Acta (BBA) - Biomembranes. Detergents in Biomembrane Studies* **1508,** 86–111. ISSN: 0005-2736 (Nov. 2000).

160. Strop, P. & Brunger, A. T. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Science* **14,** 2207–2211. ISSN: 09618368, 1469896X (Aug. 2005).

161. Horwitz, J., Huang, Q.-L., Ding, L. & Bova, M. P. in *Methods in Enzymology* 365–383 (Academic Press, Jan. 1998). (2018).

162. Aronson, D. E., Costantini, L. M. & Snapp, E. L. Superfolder GFP is Fluorescent in Oxidizing Environments when Targeted via the Sec Translocon. *Traffic (Copenhagen, Denmark)* **12,** 543–548. ISSN: 1398-9219 (May 2011).

163. Borst, J. W. & Visser, A. J.W. G. Fluorescence lifetime imaging microscopy in life sciences. *Measurement Science and Technology* **21,** 102002. ISSN: 0957-0233, 1361-6501 (Oct. 2010).

164. Margineanu, A., Chan, J. J., Kelly, D. J., Warren, S. C., Flatters, D., Kumar, S., Katan, M., Dunsby, C. W. & French, P. M. W. Screening for protein-protein interactions using Förster resonance energy transfer (FRET) and fluorescence lifetime imaging microscopy (FLIM). *Scientific Reports* **6.** ISSN: 2045-2322. (2018) (Sept. 2016).

165. Sun, Y., Hays, N. M., Periasamy, A., Davidson, M. W. & Day, R. N. in *Methods in Enzymology* 371–391 (Elsevier, 2012). ISBN: 978-0-12-391857-4.

166. Nolan, R., Iliopoulou, M., Alvarez, L. & Padilla-Parra, S. Detecting protein aggregation and interaction in live cells: A guide to number and brightness. *Methods. Developments in Fluorescence Correlation Spectroscopy and related techniques* **140-141,** 172–177. ISSN: 1046-2023 (May 2018).

167. Digman, M. A., Dalal, R., Horwitz, A. F. & Gratton, E. Mapping the Number of Molecules and Brightness in the Laser Scanning Microscope. en. *Biophysical Journal* **94,** 2320–2332. ISSN: 00063495 (Mar. 2008).

168. Hellriegel, C., Caiolfa, V. R., Corti, V., Sidenius, N. & Zamai, M. Number and brightness image analysis reveals ATF-induced dimerization kinetics of uPAR in the cell membrane. *The FASEB Journal* **25,** 2883–2897. ISSN: 0892-6638, 1530-6860 (Sept. 2011).

169. Hilsch, M., Goldenbogen, B., Sieben, C., Höfer, C., Rabe, J., Klipp, E., Herrmann, A. & Chiantia, S. Influenza A Matrix Protein M1 Multimerizes upon Binding to Lipid Membranes. *Biophysical Journal* **107,** 912–923. ISSN: 00063495 (Aug. 2014).

170. Padilla-Parra, S., Audugé, N., Lalucque, H., Mevel, J.-C., Coppey-Moisan, M. & Tramier, M. Quantitative Comparison of Different Fluorescent Protein Couples for Fast FRET-FLIM Acquisition. *Biophysical Journal* **97,** 2368–2376. ISSN: 00063495 (Oct. 2009).

171. Lam, A. J. *et al.* Improving FRET dynamic range with bright green and red fluorescent proteins. *Nature Methods* **9,** 1005–1012. ISSN: 1548-7091, 1548-7105 (Oct. 2012).

172. Hayashi, T. σ_1-1 Receptors (1 Binding Sites) Form Raft-Like Microdomains and Target Lipid Droplets on the Endoplasmic Reticulum: Roles in Endoplasmic Reticulum Lipid Compartmentalization and Export. *Journal of Pharmacology and Experimental Therapeutics* **306,** 718–725. ISSN: 0022-3565, 1521-0103 (Apr. 2003).

173. Braakman, I. & Hebert, D. N. Protein Folding in the Endoplasmic Reticulum. *Cold Spring Harbor Perspectives in Biology* **5,** a013201–a013201. ISSN: 1943-0264 (May 2013).

174. Tanaka, M., Shirasaki, T., Kaku, S., Muramatsu, M. & Otomo, S. Characteristics of binding of [3H]NE-100, a novel sigma-receptor ligand, to guinea-pig

brain membranes. *Naunyn-Schmiedeberg's Archives of Pharmacology* **351,** 244–251. ISSN: 0028-1298 (Mar. 1995).

175. Barends, T. R. M. *et al.* Combining crystallography and EPR: crystal and solution structures of the multidomain cochaperone DnaJ. en. *Acta Crystallographica Section D Biological Crystallography* **69,** 1540–1552. ISSN: 0907-4449, 1399-0047 (Aug. 2013).

176. Wittung-Stafshede, P., Guidry, J., Horne, B. E. & Landry, S. J. The J-Domain of Hsp40 Couples ATP Hydrolysis to Substrate Capture in Hsp70. *Biochemistry* **42,** 4937–4944 (May 2003).

177. Kityk, R., Kopp, J. & Mayer, M. P. Molecular Mechanism of J-Domain-Triggered ATP Hydrolysis by Hsp70 Chaperones. *Molecular Cell* **69,** 227–237.e4. ISSN: 10972765 (Jan. 2018).

178. Barthel, T. K., Zhang, J. & Walker, G. C. ATPase-Defective Derivatives of Escherichia coli DnaK That Behave Differently with Respect to ATP-Induced Conformational Change and Peptide Release. *Journal of Bacteriology* **183,** 5482–5490. ISSN: 0021-9193 (Oct. 2001).

179. Geladopoulos, T. P., Sotiroudis, T. G. & Evangelopoulos, A. E. A malachite green colorimetric assay for protein phosphatase activity. *Analytical Biochemistry* **192,** 112–116. ISSN: 0003-2697 (Jan. 1991).