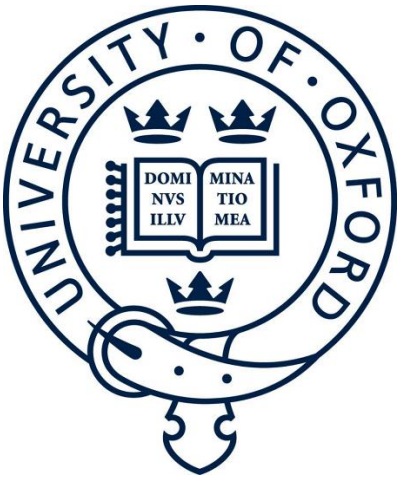
Investigating the role of the Abruptex region within the Notch receptor



Thomas J W Rowntree

Department of Biochemistry and Magdalen College

University of Oxford

A Thesis submitted for the degree of

Doctor of Philosophy (DPhil)

Trinity Term 2018

Abstract

Investigating the role of the Abruptex region within the Notch receptor

A thesis submitted for the degree of Doctor of Philosophy (DPhil) by Thomas J W Rowntree, Magdalen College, Trinity Term 2018

The aim of this project was to investigate the structure and function of the Abruptex (Ax) region within the Notch receptor. Previous data from *D. melanogaster* shows that missense mutations in this region (EGF24-29) lead to a phenotype indicative of increased signalling. A proposed mechanism for this involves an interaction between the Abruptex region and the ligand binding region of the receptor to negatively regulate signalling. In this project, hNotch1 protein constructs were expressed from prokaryotic and eukaryotic systems to investigate the structure of the region, and interactions with the ligand binding region. After establishing a pipeline for characterizing Ax EGF-like domain constructs by small angle X-ray scattering, structural analysis revealed an extended structure for human Notch1 and Notch2 EGF23-27, with a kink at the EGF25-26 interface inferred from bioinformatics analysis. This suggests that additional non-canonical cysteines in human Notch1 EGF25 and EGF27 do not have a structural role. Binding assays suggested an interaction between EGF12 and EGF24, with binding to EGF12 dissimilar to ligand binding and requiring both O-Fuc and O-Gluc modifications on opposite sides of the domain. A decrease in binding was observed for the EGF24 Abruptex mutant D909V suggesting involvement of the N-terminal of this domain. In addition, a cell-based transactivation assay utilizing hJagged1 as the ligand was established to investigate the impact of Abruptex variants on activation of the mammalian Notch1 receptor. This showed a decrease in activation associated with a variant containing Abruptex missense mutations in EGF24 and EGF25, and a lack of activation in an EGF24-25 deletion variant. These results suggest that in the case of Jagged1, the Abruptex region has a role in optimizing transactivation. The enhanced signalling phenotype indicative of Abruptex mutants may be mediated by other means such as modulating cis-inhibition.

Acknowledgements

I would like to thank my supervisor Penny Handford for her advice throughout the project, as well as the BBSRC for funding this research.

My thanks go out to past and present members of the Handford lab, in particular Pat Whiteman, Sacha Jensen, Boguslawa Korona, Ian Robertson, and Yao Meng for their help and advice. Furthermore, I would like to thank Rich Suckling, Matthias Jore, Steve Johnson and their group leader Susan Lea for their advice throughout the project as well as Christina Redfield for help and advice with NMR, and Ed Lowe for advice relating to SAXS.

Finally, I would like to thank my friends and family for their support, in particular my parents and Andrea.

Contents

List of Figures xi

List of Abbreviations xvi

1. Introduction 1

1.1. General Introduction 1

1.2. Notch signalling 3

1.2.1. Discovery and role of Notch signalling 3

1.2.2. Examples of Notch signalling 5

1.2.3. General scheme of Notch signalling 7

1.3. EGF domains 8

1.3.1. Structural Features 10

1.3.2. Calcium binding 11

1.3.2.1. Water as a ligand for Ca^{2+}. 12

1.3.3. Classes of cbEGF pairs 12

1.3.4. Post-translational modification 14

1.3.4.1. N-linked glycosylation 15

1.3.4.2. O-linked fucosylation 16

1.3.4.3. O-linked glucosylation 17

1.3.4.4. O-linked β-N-acetylglucosaminylation 19

1.3.4.5. B-hydroxylation 19

1.3.5. EGF-domain interfaces 20

1.3.5.1. Shapes of tandem EGF-domain constructs. 20

1.4. Notch receptors. 21

1.4.1. Introduction to Notch receptors. 21

1.4.2. Ligand-binding region. 24

1.4.3. Abruptex region. 26

1.4.3.1. The Abruptex phenotype. 26

1.4.3.2. Proposed interaction of AX with LBR and models. 30
1.4.4. Negative regulatory region. 31
1.4.5. Intracellular region. 33
1.5. Canonical Notch ligands. 34
1.5.1. Structure of ligands. 35
1.5.2. Delta/Serrate/Lag-2 domain. 35
1.5.3. C2 domain. 36
1.5.4. Mechanism of canonical Notch signalling. 37
1.6. Non-canonical modulators of Notch signalling. 39
1.7. Medical relevance of Notch signalling. 39
1.7.1. Alagille syndrome. 40
1.7.2. CADASIL syndrome. 41
1.7.3. Notch in cancer. 41
1.7.4. Other diseases. 42
1.8. Aims. 43

2. Materials and Methods **45**
2.1. DNA methods. 45
2.1.1. DNA preparation. 45
2.1.2. DNA gel electrophoresis. 46
2.1.3. DNA extraction from agarose TBE gel. 46
2.1.4. DNA concentration determination. 46
2.1.5. DNA sequencing. 47
2.1.6. Polymerase chain reactions. 47
2.1.7. Site directed mutagenesis. 48
2.1.7.1. Substitutions by overlap extension mutagenesis. 48
2.1.7.2. Substitutions by QuikChange XL site directed mutagenesis. 49
2.1.7.3. Substitutions using Q5 site directed mutagenesis. 49
2.1.7.4. Deletions. 50

2.1.8. Restriction digestion. 50
2.1.9. DNA ligation. 50
2.1.10. Primer sequences. 51
2.1.10.1. S2 hN1 11-13 directional mutants. 51
2.1.10.2. PR hN1 23-24 directional mutants. 52
2.1.10.3. PR hN1 23-27CSCS. 52
2.1.10.4. mN1 primers. 53
2.2. RNA methods. 54
2.2.1. RNA extraction. 54
2.2.2. RT-PCR. 55
2.2.2.1. RT-PCR primers. 55
2.3. Cell culture. 56
2.3.1. Insect cells. 56
2.3.1.1. S2 cells growth conditions and transfection protocol. . . . 56
2.3.2. Mammalian cells. 56
2.3.2.1. 293T cells growth conditions. 56
2.3.2.2. HEK-RBP cells growth conditions and transfection protocol.56
2.3.3. *E.coli* growth and transformations. 57
2.4. Protein methods. 57
2.4.1. General procedures. 57
2.4.1.1. Protein concentration determination. 57
2.4.1.2. SDS-PAGE. 58
2.4.1.3. Western blotting. 58
2.4.1.4. ESI-MS. 59
2.4.1.5. DTNB assay. 59
2.4.1.6. Biotinylation. 59
2.4.1.7. Factor Xa cleavage to remove His_6 tag. 60
2.4.1.8. EDTA treatment. 60
2.4.2. NMR. 61

2.4.2.1. Ca^{2+} titrations. 61
2.4.2.2. Protein titrations. 62
2.4.2.3. Checking S2 fold. 62
2.4.3. S2 expression and purification. 63
2.4.4. Prokaryotic refold expression and purification. 63
2.5. Small angle X-ray scattering. 64
2.5.1. Data collection. 64
2.5.2. Model determination. 64
2.6. Binding assays. 65
2.6.1. Plate binding assay. 65
2.6.2. Flow cytometry assay. 66
2.7. Activation assays. 67
2.7.1. Ligand-independent activation assay. 67
2.7.2. Ligand- dependent activation assay. 67

3. Protein Purification and Characterisation 69
3.1. Bioinformatics analysis of hN1 and other Notch receptors. 70
3.2. Expression of Notch constructs from the S2 system. 75
3.2.1. Overview of the S2 system. 75
3.2.2. Designing constructs for production in the S2 system. 75
3.2.3. Purification of S2 hN1 11-13 variants. 77
3.2.4. Purification of S2 hN1 Abruptex constructs. 81
3.2.4.1. S2 hN1 23-24 variants. 81
3.2.4.2. S2 hN1 23-29. 83
3.2.4.3. S2 hN1 23-27. 84
3.2.4.4. Purified S2 Ax constructs. 86
3.3. Expression of constructs from the prokaryotic/refold (PR) system. 86
3.3.1. Designing and expressing constructs from the prokaryotic/refold system. 87

3.3.2. Purification of PR hN1 23-24 variants. 88
3.3.3. Purification of PR hN1 23-27CSCS. 91
3.3.4. Purification of PR hN2 23-27cl. 93
3.3.5. Purified 23-27 constructs. 94
3.3.6. Electron spray ionization mass spectrometry of PR constructs. . . 95
3.4. Biotinylation and cleavage reactions for preparation of protein for use in assays. 96
3.4.1. Factor Xa cleavage reactions. 96
3.4.2. Biotinylation of LBR and Ax constructs for binding assays. 97
3.5. Investigating Ca^{2+} affinity and aromatic stacking in hN1 EGF24. 101
3.5.1. Characterisation of Ca^{2+} affinity of EGF24 by NMR. 101
3.5.2. The unusual intradomain aromatic in EGF24. 102
3.6. Biochemical studies of the region. 105
3.6.1. Investigating the non-canonical cysteines in the hN1 Abruptex region. 105

4. Low Resolution Structural Studies of the Abruptex region 109
4.1. Understanding receptor function in terms of shape. 109
4.2. Small angle X-ray scattering. 111
4.2.1. Theory. 111
4.2.2. SEC-SAXS. 112
4.2.3. Development of a pipeline for processing data. 114
4.2.3.1. Data quality screen. 114
4.2.3.2. Building reasonable EOM models. 116
4.2.3.3. P(r) differentiation. 117
4.2.3.4. OLIGOMER. 118
4.2.3.5. Generating models. 119
4.2.4. Testing the pipeline with hN1 11-13. 120
4.2.5. Structure of hN1 23-27CSCS. 128

4.2.5.1. Modelling of hN1 23-27CSCScl. 133
4.2.6. Structure of hN2 23-27cl. 133
4.2.7. Comparison of EGF23-27 in hN1 and hN2. 137

5. Introduction of the Abruptex region with the ligand-binding region 141
5.1. Background. 141
5.2. Investigating the Ax – LBR interaction in a plate binding assay. 144
5.2.1. The plate binding assay. 144
5.2.1.1. S2 hN1 11-13 interacts with PR hN1 23-24 and S2 hJag1. . 148
5.2.1.2. PR hN1 23-24 binds to S2 hN1 11-13 but not PR hN1 11-13. 149
5.2.1.3. Steric hindrance and orientation in the plate assay. 152
5.2.1.4. Interaction of hN1 11-13 with hN1 23-24 is dependent on orientation of hN1 23-24. 153
5.2.1.5. Interactions with PR hN1 23-24. 155
5.2.1.5.1. Interaction of hN1 11-13 variants. 155
5.2.1.5.2. Independent confirmation of S2 hN1 11-13 Q462R sample integrity. 160
5.2.1.5.3. Investigation into sugar-dependence of the interaction. 161
5.3. Investigating the interaction in the flow cytometry assay. 165
5.3.1. Establishing the flow cytometry assay for the investigating the interaction of LBR with Ax. 167
5.3.2. Interaction with other Abruptex constructs. 169
5.3.3. Identification of the interaction face of hN1 23-24. 170
5.4. Conclusions and discussion from binding assays. 173

6. Investigating the role of the Abruptex region in Notch activation 176
6.1. Introduction. 176

6.2. Advantages and disadvantages of the activation assay. 178
6.3. Establishing and validating the activation assay. 179
6.3.1. Using pcDNA5/FRT mN1 wt and L468A to validate the activation assay. 182
6.3.1.1. Assessment of L468A. 184
6.3.2. Optimising conditions for the hJag1 activation assay. 188
6.4. Designing Ax variant receptors to be investigated in the activation assay. 191
6.4.1. Abruptex EGF24 D909V variant. 192
6.4.2. Double Abruptex mutant. 196
6.4.3. Abruptex deletion mutant. 200
6.5. Discussion of Notch activation assays. 203

7. General discussion and future work **205**
7.1. Aims of this thesis. 205
7.2. Understanding the shape of the Notch receptor. 206
7.3. Understanding the function of the Abruptex region. 206
7.4. Future Work. 213
7.5. Summary. 216

Bibliography **217**

Appendix **230**

List of Figures

1.1. Schematic of Notch signalling. 2
1.2. The domain organisation of Notch receptors. 4
1.3. The roles of Notch signalling in development. 7
1.4. The structures of EGF and cbEGF domains. 10
1.5. The cbEGF Ca^{2+}-binding consensus. 13
1.6. The class II cbEGF pair. 14
1.7. EGF O-linked glycosylations. 15
1.8. Notch glycans can be sequentially modified. 18
1.9. cbEGF-EGF interfaces are variablie. 22
1.10. The structurally characterized EGF regions of the hN1 receptor. . . . 24
1.11. The ligand binding region of Notch1. 25
1.12. The extended interaction face of LBR and Jagged1. 26
1.13. Abruptex alleles identified in *D. melanogaster*. 27
1.14. The Abruptex phenotype. 29
1.15. hN1 negative regulatory region. 32
1.16. The structural organisation of the extracellular region of Notch ligands. 34

3.1. Summary, advantages and disadvantages of prokaryotic/refold and S2 protein production procedures. 70
3.2. Sequence analysis of the hN1, hN2, and dN receptors. 73
3.3. Constructs expressed in the S2 system. 77
3.4. General Purification procedure for S2 hN1 11-13 variants. 79
3.5. 16 % SDS-PAGE analysis showing purified products of hN1 11-13 variants. 80
3.6. Purification of S2 hN1 23-24. 82
3.7. Purification of S2 hN1 23-29. 83

3.8. Purification of S2 hN1 23-27. 85
3.9. Purified S2 hN1 Ax constructs. 87
3.10. Constructs expressed in prokaryotic/refold system. 88
3.11. Elution traces from multiple stages throughout the purification of PR hN1 23-24. 89
3.12. Final reverse-phase HPLC traces of PR hN1 23-24 variants. 90
3.13. 16% SDS-PAGE of purified PR hN1 23-24 variants. 91
3.14. Elution traces from multiple stages throughout the purification of PR hN1 23-27CSCS. 92
3.15. Purified PR hN1 Abruptex constructs. 93
3.16. Assessment of purity of PR hN2 23-27cl. 94
3.17. Results of ESI-MS used to accurately determine mass of prokaryotic/refold constructs. 95
3.18. Reverse-phase HPLC trace following Factor Xa cleavage of PR hN1 23-27CSCS to produce PR hN1 23-27CSCScl. 96
3.19. Reverse-phase HPLC trace following Factor Xa cleavage of PR hN1 23-24 to produce PR hN1 23-24cl. 97
3.20. Example biotinylation of Notch constructs for use in assays. 99
3.21. Western blot with streptavidin conjugate following 16 % SDS-PAGE shows equivalence of biotinylation across all S2 hN1 11-13 variants. 100
3.22. Quantification of the Ca^{2+} binding site of hN1 EGF24 at pH7.5 in the presence of 150 mM NaCl. 103
3.23. Comparison of typical and atypical hN1 EGF24 Ca^{2+}-consensus. . . . 104
3.24. Modelling the additional non-canonical cysteines in the hN1 onto a known structure. 106
3.25. DTNB assay to investigate the reduction-state of the thiols in the hN1 Ax region. 107

4.1. Schematic of SEC-SAXS data collection. 113

4.2. Pipeline for data assessment and processing for elongated multiple tandem-domain chain constructs. 115
4.3. The relationship between structure and P(r) 118
4.4. Comparison of independent SAXS experiments performed on S2 hN1 11-13cl. 122
4.5. OLIGOMER model fitting to S2 hN1 11-13cl. 125
4.6. Analyzing the S2 hN1 11-13cl SAXS data collected in 15 mM Ca^{2+}. . . 126
4.7. Investigating the Kratky plot. 127
4.8. Determination of hN1 11-13cl structure using SAXS. 128
4.9. SAXS analysis of hN1 23-27CSCScl (15 mM Ca^{2+}). 130
4.10. OLIGOMER analysis of hN1 23-27CSCScl in the presence of Ca^{2+}. . . . 132
4.11. The structure of hN1 23-27CSCScl at 15 mM Ca^{2+}. 134
4.12. Investigating the structure of hN2 23-27cl. 135
4.13. Analyzing SAXS data from hN2 23-27cl. 136
4.14. The structure of hN2 23-27cl. 137
4.15. The structures of hN1 and hN2 EGF23-27. 140

5.1. D909V in EGF24 disrupts the interaction with 11-13. 142
5.2. hN1 23-24 was used in interaction studies. 145
5.3. Overview of the plate binding assay. 146
5.4. Validating the plate binding assay with ligand. 147
5.5. Comparison of the interaction of PR hN1 23-24 in the plate binding assay. 149
5.6. Comparison of the interaction of PR and S2 proteins in the plate binding assay. 150
5.7. Locations of sugars on hN1 11-13 and 23-24. 152
5.8. Investigating different orientations of hN1 23-24 constructs in the plate assay. 154
5.9. The design of directional variants in EGF11 and EGF12. 156

5.10. Analysis hN1 11-13 EGF12 variants in the plate assay. 157
5.11. Analysis hN1 11-13 EGF11 variants in the plate assay. 158
5.12. Independent analysis of hN1 11-13 EGF12 variants in the plate assay. 159
5.13. Comparison of hJag1 and PR hN1 23-24 binding to hN1 11-13 wt and Q462R at 400 nM S2 hN1 11-13. 160
5.14. The two faces of EGF12 implicated in binding hN1 23-24. 161
5.15. O-glucosylated PR hN1 11-13 interacting in the plate assay. 162
5.16. The binding site of Ax in EGF12. 164
5.17. Schematic of the flow cytometry assay. 166
5.18. PR hN1 23-24 binds to cells, S2 hN1 11-13 does not bind to cells. . . .168
5.19. Antibody abrogation suggests a specific interaction of PR hN1 23-24 with EGF11-12 on the surface of the HEK293T cells. 169
5.20. hN1 23-24 expressed in both the prokaryotic/refold and S2 systems bind to cells in the flow cytometry assay. 170
5.21. Interactions of PR hN1 23-24 directional variants in the flow cytometry assay. 171
5.22. Elimination of the His-tag leads to a loss of binding. 173
5.23. Schematic models for the interaction of Ax with the LBR. 175

6.1. Schematic of the activation assay. 177
6.2. Preparation of stable clones for the activation assay. 180
6.3. Transfection of HEK-RBP cells with GFP-expressing plasmid. 180
6.4. Trial transfection to demonstrate FRT-integration capacity of cells. . 181
6.5. Assessment of accuracy of the activation assay. 184
6.6. RT-PCR to investigate expression and LBR sequence of mN1 variant. . 185
6.7. Sequence of mN1 wt L468A confirmed by RT-PCR and restriction digestion 187
6.8. Ligand independent activation assay with EGTA suggests equivalent

expression of mN1 wt and L468A. 188
6.9. Investigating optimal reading time and ligand concentration for activation assay. 190
6.10. Low signal for L468A suggests activation observed is ligand dependent. 191
6.11. RT-PCR to investigate expression and Ax sequence of mN1 variant. . 193
6.12. Sequence of mN1 D909V was confirmed by RT-PCR and restriction digestion. 194
6.13. Ligand dependent activation assay suggests mN1 D909V signals similarly mN1 wt. 195
6.14. Designing a double Ax mutant. 196
6.15. Expression of mN1 DVNI was confirmed by RT-PCR. 198
6.16. Ligand dependent activation assay suggests mN1 DVNI signals less than mN1 wt. 199
6.17. Ax deletion mutant engineered for investigation in the activation assay. 200
6.18. Expression of mN1 Δ24-25 was confirmed by RT-PCR. 201
6.19. Ligand dependent activation assay suggests mN1 Δ24-25 does not undergo activation on exposure to hJag1. 202

7.1. The proposed structural organisation of hN1 23-27 in hN1 207
7.2. Proposed organisation of the Notch receptor at the cell surface 208
7.3. Proposed models for Ax function in Jagged1-dependent ligand transactivation. 212
7.4. Proposed future investigations to further characterize the role of the Abruptex region within the Notch receptor. 215

List of Abbreviations

A	Adenine
Å	Ångstrom
ADAM	a disintegrin and metalloprotease-containing protein
ALGS	Alagille syndrome
ANK	Ankyrin repeat
ATP	Adenosine triphosphate
Ax	Abruptex
A_{xnm}	Absorbance at wavelength x nm
bp	base pairs
BSA	bovine serum albumin
C	Cytosine
CADASIL	cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy
cbEGF	Ca^{2+}-binding epidermal growth factor-like
CMV	Cytomegalovirus
CSL	CBF-1/Suppressor of Hairless/LAG-2 transcription factor
CV	Column volumes
d	*D. melanogaster*
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
Dll1	Delta like 1 ligand
Dll3	Delta like 3 ligand
Dll4	Delta like 4 ligand
dN	*D. melanogaster* Notch receptor
DNA	Deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates

DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
D_{max}	maximum dimension of the protein
DSL	Delta/Serrate/Lag-2
DTNB	5,5-dithio-bis-2-nitrobenzoic acid)
DTT	Dithiothreitol
DV	Dorsal/Ventral
ECM	extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor-like
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
EHBA	Extrahepatic biliary atresia
EMT	Epithelial-mesenchymal transitions
EOM	Ensemble optimization modelling
ER	Endoplasmic reticulum
ESI-MS	Electrospray ionization mass spectrometry
Fbn1	Fibrillin1
FCS	Fetal calf serum
FPLC	Fast protein liquid chromatography
Fuc	Fucose
G	Guanine
GFP	Green fluorescent protein
GlcNAc	N-acetylglucosamine
Gluc	Glucose
h	human
HPLC	High pressure liquid chromatography
HBSS	Hanks' balanced saline solution
hr	hour

HD	heterodimerization
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidase
HSQC	Heteronuclear single quantum coherence
Hz	hertz
I	ionic strength
Jag1	Jagged1
Jag2	Jagged2
K_d	Dissociation constant
L	Litre
LBR	Ligand binding region
LNR	Lin12/Notch repeat
LTBPs	Latent TGF-β binding proteins
Luc	luciferase
m	mouse
M	moles / dm^3
mAb	Monoclonal antibody
MAML	mastermind-like protein
min	minute
MNNL	Module at the N-terminus of Notch ligands domain
MW	Molecular weight
N	Newton
N	Notch haploinsufficiency phenotype
N1	Notch1 receptor
N2	Human Notch2 receptor
N3	Human Notch3 receptor
N4	Human Notch4 receptor
NICD	Notch intracellular region

NMR	Nuclear Magnetic resonance
NRR	Negative regulatory region
PCR	Polymerase chain reaction
PDB ID	Protein Data Bank Identification code
PEST	Proline, glutamic acid, serine, threonine-rich sequence
ppm	parts per million
PR	Prokaryotic/refold
$P_{(r)}$	Pair-distance distribution function
RAM	RBPJk-associated module
RMSD	Root-mean-square deviation
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Reverse transcription
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
secs	Seconds
SOP	Sensory organ precursor
SV40	Simian virus 40
S2	Schneider 2
T	Thymine
T-ALL	T-cell acute lymphoblastic leukaemia
TBE	Tris/borate/EDTA
TFA	Trifluoroacetic acid
T_m	Melting temperature
TM	Transmembrane
Tris	Tris (hydroxymethyl) aminomethane
UV	Ultraviolet

V	Volts
wt	Wild type
1D	One-dimensional
2D	Two-dimensional

List of amino acid abbreviations used

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Metric prefixes

p	10^{-12}
n	10^{-9}
µ	10^{-6}
m	10^{-3}
d	10^{-1}
k	10^{3}
M	10^{6}

1.

Introduction

1.1 General introduction

The Notch signalling pathway is a juxtacrine signalling system conserved throughout multicellular organisms. This pathway is involved in most aspects of cell biology including regulating cell differentiation, cell fate determination, and is implicated in a variety of acquired and genetic diseases. Canonical signalling takes place between a transmembrane Notch receptor and transmembrane ligand, each on different cells. The ligand can be of either the Delta or Serrate/Jagged families. The signal-sending cell expresses ligand at the cell surface, this subsequently binds to a receptor on the signal-receiving cell. This ligand-receptor complex then undergoes two proteolytic cleavages (S2, S3) near the transmembrane region of the receptor, and this releases the intracellular region of the receptor within the cell. This intracellular fragment then moves to the

nucleus to regulate transcription. In addition to transactivation complexes, ligand/receptor expressed on the same cell can lead to the formation of cis-inhibitory complexes which down-regulate the Notch signal (Figure 1.1).

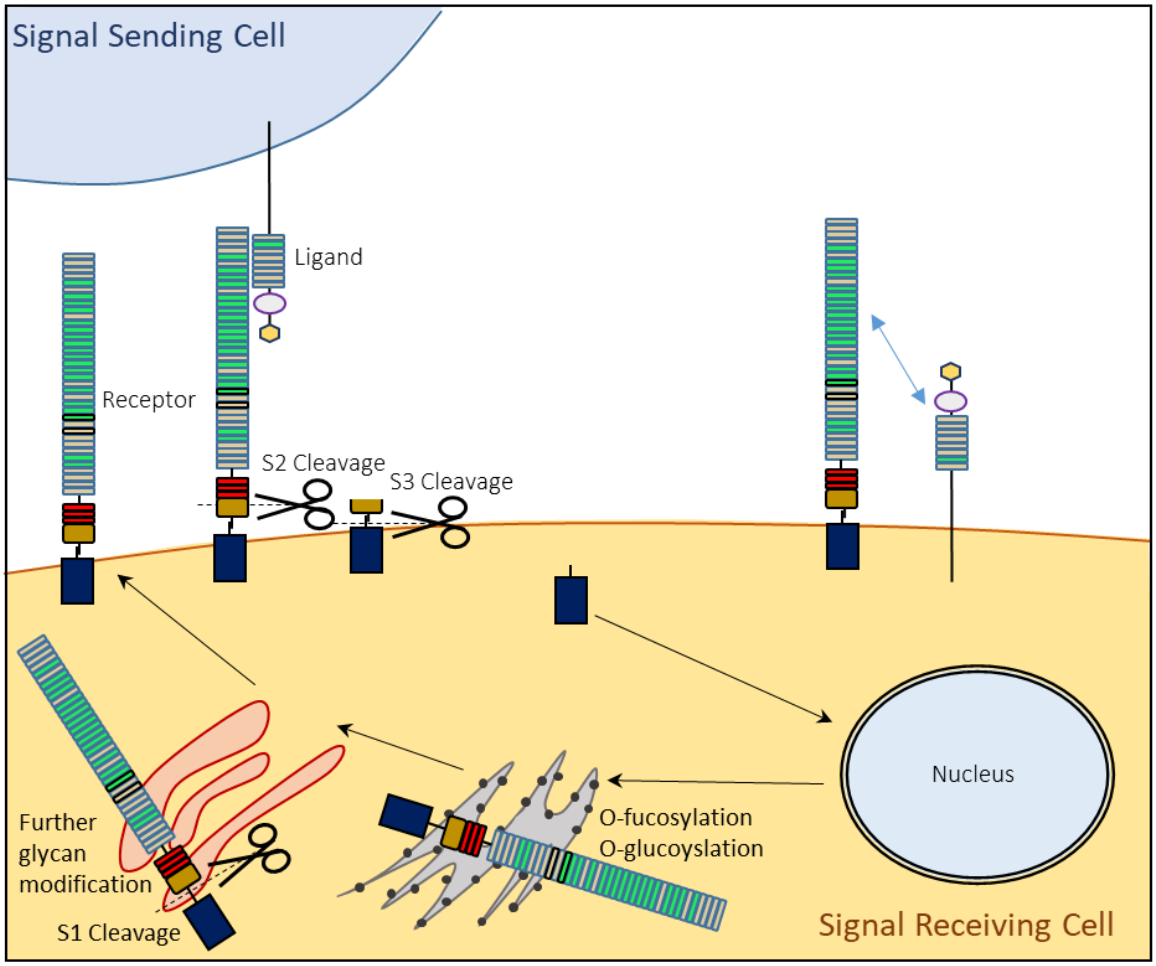


Figure 1.1 Schematic of Notch signalling: The Notch receptor is present on the cell surface after undergoing post-translational modification and processing in the endoplasmic reticulum (grey) and Golgi apparatus (red). On ligand induced activation, S2 and S3 proteolytic cleavage reactions occur at the cell surface, which lead to the release of the NICD (navy) to the nucleus. Receptor and ligand on the surface of the same cell can interact in cis (blue arrows) to decrease receptor activation.

Most of the extracellular region of the Notch receptor comprises of multiple tandem epidermal growth factor-like (EGF) domains and Ca^{2+}-binding EGF (cbEGF) domains, which is the location of the core ligand binding region EGF11-12 (LBR). The human genome encodes four paralogues of this receptor, which are varied in length. Receptors

and ligands are heavily post-translationally modified with N- and O-linked glycans, which provides structural stability, as well as mediating binding with other components of the pathway.

1.2 Notch signalling

1.2.1 Discovery and role of Notch signalling

The phenotypic consequences of defects in the Notch signalling pathway were first recorded over a century ago as sex-linked serration of wings in *D. melanogaster* mutants, known as the *Notch (N)* phenotype (Morgan, 1917) (Mohr, 1919). *D. melanogaster* remains the best characterised model organism for Notch function. The *Drosophila* Notch (dN) receptor is similar in structure and organisation to the human Notch1 (hN1) and Notch2 (hN2) receptors, unlike the *C. elegans* receptors (Lin-12 and Glp-1) which are shorter (Figure 1.2). These receptors are type 1 transmembrane proteins consisting of a large extracellular region of mainly tandem EGF domains, the NRR (negative regulatory region) and the NICD (Notch intracellular region).

Notch signalling is essential for development as indicated by embryonic lethality of Notch null mutants in male flies (Li, 1927) which lack an additional X-chromosome to complement the Notch null chromosome. Gradually, the broader role of Notch in many aspects of *Drosophila* development was elucidated, including the development of the wing veins, maintaining the Dorsal/Ventral (DV) boundary, and eye development (Artavanis-Tsakonas et al, 1999, Singh and Choi, 2003). The role of Notch in human

development is more complex considering that humans have four Notch genes, located on chromosomes 9, 1, 19 and 6, encoding receptors hN1 to hN4 (human Notch4) respectively. However, the fundamental role in organismal development is conserved. Specific examples of Notch signalling in mammals include development of the hair follicle cells in the inner ear (Yamamoto et al, 2003) in which Notch regulates the distribution of hair follicle cells, control of skin cell proliferation (Lowell et al, 2000), and fibre cell differentiation in the optical lens (Rowan et al, 2008).

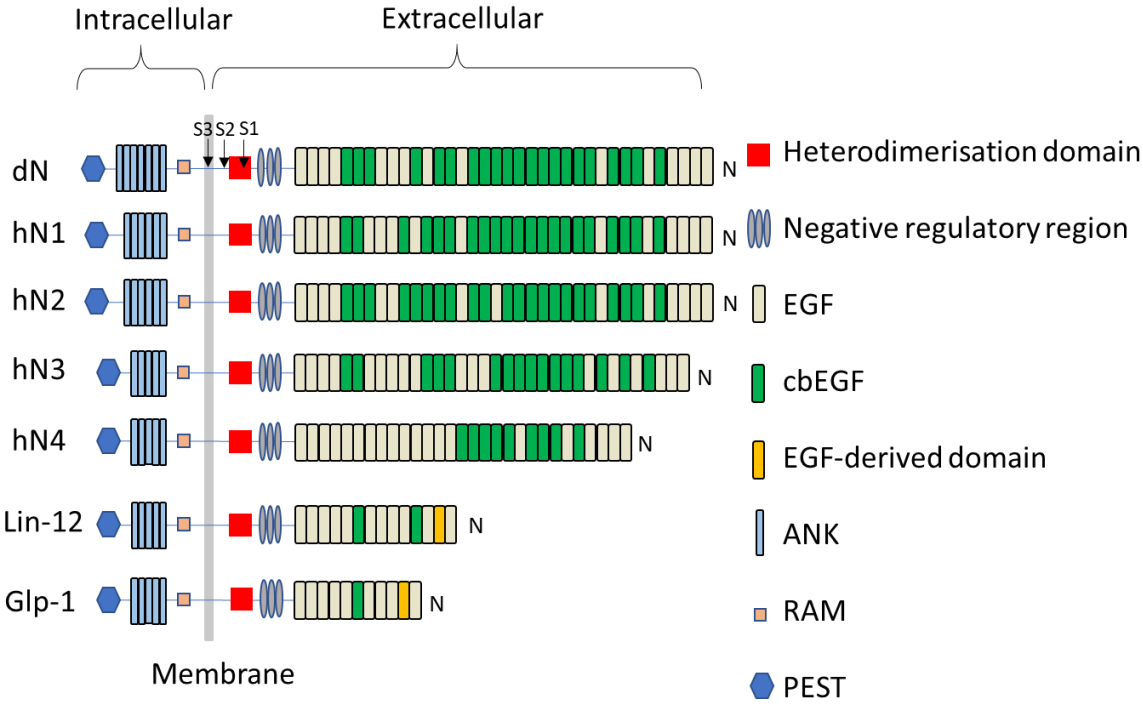


Figure 1.2 The domain organisation of Notch receptors: Notch receptors from *D. melanogaster* (dN), human (hN1, hN2, hN3, hN4), and *C. elegans* (Lin-12, Glp-1). ANK = Ankyrin repeat, RAM = RBPJK-associated module, PEST = Proline, glutamic acid, serine, threonine-rich sequence. The locations of the S1, S2, and S3 cleavage sites are indicated. The mammalian Notch receptor is cleaved at the S1 site during secretion and normally presents as a heterodimer at the cell surface (Gordon et al, 2009). dN, hN1, and hN2 consist of 36 EGF/cbEGF domains in the extracellular region.

1.2.2 Examples of Notch signalling

Despite the extensive involvement of Notch signalling in development, the mechanism of the pathway in any specific context can be grouped into one of three distinct modes: Boundary formation, binary cell fate decisions, and lateral inhibition. These are all brought about by the short-range cell-to-cell signalling between Notch ligand and receptor which is then propagated to produce a large developmental effect. It is this simplicity and generality that enables Notch signalling to be so widely applied throughout development.

Boundary formation is an important process during cell segregation in any complex organism. One of the best characterised examples is the role of Notch in defining the DV boundary in the wing of *D. melanogaster* (Kim et al, 1995, Becam et al, 2008). The GlcNAc transferase enzyme Fringe is involved in post-translational modification of Notch and expressed in dorsal cells. Notch modified by Fringe is able receive signal from the ligand Delta, but not the ligand Serrate. At a point far enough from the dorsal side, cells expressing Notch are not able to receive a signal from Delta because the Notch at their surface is not modified due to lack of Fringe. At this point, a boundary forms in which cells competent at Delta signalling are adjacent to cells competent at Serrate signalling, each ligand signal enhancing the effect of the other, leading to a spike in Notch signalling and instructing the position of the boundary (Figure 1.3).

Binary cell fate decisions can be facilitated by Notch through the asymmetric localisation of Notch-regulatory proteins between the two daughter cells following a cell division. The differences in Notch regulation between the two daughter cells leads to each adopting a different fate. This occurs in the development of the Sensory Organ Precursor (SOP) in *D. melanogaster* (Hutterer and Knoblich, 2005). The protein Numb acts to supress Notch signalling in the cell where it is expressed. Numb is segregated to one of the daughter cells (pIIb) but not the other (pIIa), and so only pIIa is able to receive a Notch signal from ligand. Meanwhile ligand is expressed equally on both cells, but the different fates of each daughter cell and their progeny form distinct parts of the SOP (Figure 1.3).

Lateral inhibition enables the formation of evenly dispersed cell lineages in a mosaic pattern. This occurs during sensory hair cell generation in the inner ear (Li et al, 2015). Stochastic fluctuations in expression of the ligand Jagged2 lead to surrounding cells receiving more Notch signal, this in turn leads to inhibition of ligand expression in these surrounding cells. This phenomenon occurs over the whole of the hair cell precursor layer and leads to uniformly separated cells expressing ligand, with cells around them receiving signal. The cells expressing ligand differentiate to become the hair cells and the surrounding cells become support cells (Lanford et al, 1999) (Figure 1.3). Recent evidence suggests that the area of cell contact is proportional to the strength of signal produced from that contact, from which it can be inferred that small cells are likely to become signal sending cells because they receive a weaker signal than a large cell in a similar environment (Shaya et al, 2017).

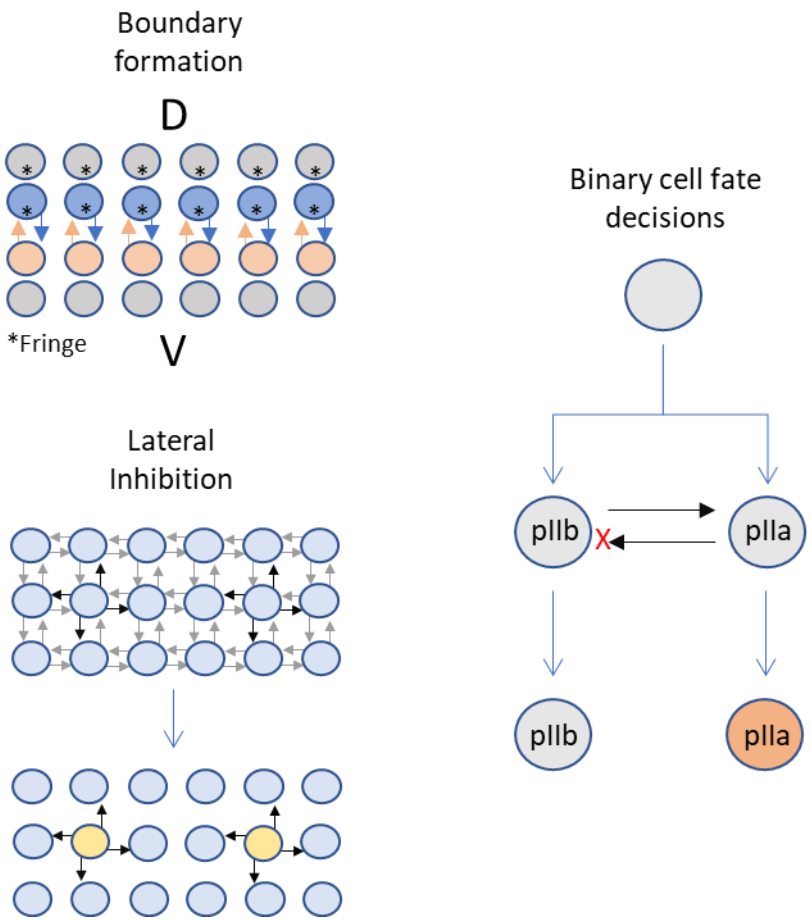


Figure 1.3 The roles of Notch signalling in development: Boundary formation) Dorsal (D) and Ventral (V) sides with Serrate signalling-competent cells coloured in blue and Delta signalling-competent cells coloured in orange. Fringe is present in the dorsal cells (blue). Orange arrows indicate signalling through Delta and blue arrows indicate signalling through Serrate. Lateral inhibition) Intensity of the arrows between cells representing signal strength. Yellow cells will differentiate to become hair cells. Binary cell fate decisions) Black arrows between cells representing Notch signalling and the red cross representing unsuccessful signalling. Cells are labelled and coloured according to their lineage with orange representing a pIIa cell.

1.2.3 General scheme of Notch signalling

Notch signalling between adjacent cells involves an interaction between a transmembrane receptor and ligand. The interaction varies with each ligand and occurs along an extended interface (Luca et al, 2017) despite being centred around the C2 and

DSL (Delta/Serrate/Lag-2) domain of ligand and EGF11-12 (LBR) of the receptor. The receptor-ligand complex then experiences a mechanical force, most likely mediated by the endocytosis of the transactivation complex, which leads to the exposure of a cryptic cleavage site (S2 cleavage site) within the negative-regulatory region (NRR) of the receptor, located adjacent to the transmembrane region on the extracellular side (Figure 1.2). Application of force to the receptor is sufficient to reveal this S2 cleavage site, and Ca^{2+} can act to stabilize this region (Stephenson and Avis, 2012). Cleavage at the S2 site by an ADAM metalloprotease leads allows the intracellular S3 site to be accessed by γ-secretase for cleavage (Van Tetering 2009). The Notch intracellular fragment (NICD) can then move to the nucleus to affect transcription by recruiting MAML (Mastermind-like protein) to form a complex with the transcription factor CSL (CBF-1/Suppressor of Hairless/Lag-2), leading to the transcription of HES/HEY genes, which themselves are transcription factors (Katoh and Katoh, 2007). There are no proliferative enzymatic steps in this signalling pathway, therefore the number of activated receptors corresponds directly to the number of the activated CSL complexes. This is supported by evidence showing the importance of dose (number of copies of the Notch gene being transcribed) (Welshons, 1965) and size of the contact area between signalling cells (Shaya et al, 2017).

1.3 EGF domains

The majority of the extracellular region of the Notch receptor is made up of EGF or cbEGF domains (Figure 1.2). The EGF domain module is present in many proteins of diverse functions, including selectins involved in the immune system, blood clotting factors, and components of the extracellular matrix such as fibrillin and the Latent TGF-β binding

proteins (LTBPs) (Wouters et al, 2005). The structure of the EGF domain is highly conserved but shows large sequence variability. The core central β-sheet structure of the EGF domain is maintained by several conserved residues including six cysteines in a 1-3, 2-4, 5-6 disulphide bonding pattern (Figure 1.4). Furthermore cbEGFs, contain additional residues which provide the electron lone pairs to coordinate Ca^{2+} at the N-terminal end of the domain. These distinct modules pack against each other to create a variety of angles and flexibilities (Weisshuhn et al, 2016, Downing et al, 1996).

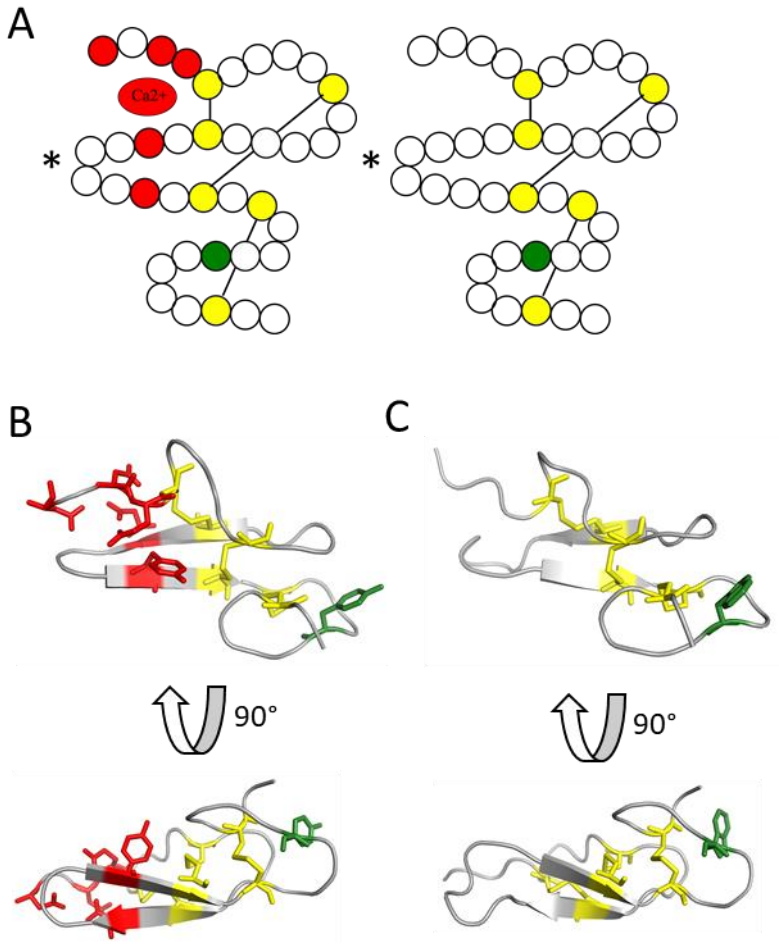


Figure 1.4 The structures of EGF and cbEGF domains: (A) Two-dimensional schematics of a cbEGF (with Ca^{2+} present) and an EGF are shown with consensus Ca^{2+}-binding residues and bound Ca^{2+} indicated (red), consensus cysteines (yellow) and connecting disulphide bonds (black lines), central β-sheet of the domain (*), and consensus packing aromatic (green) which can be important in interactions with the next EGF domain in a tandem repeat. (B) Three-dimensional model of a cbEGF, red circle represents the Ca^{2+} ion. (C) Three-dimensional model of an EGF, with colours correspond to the residues in panel A.

1.3.1 Structural Features

The core fold of the EGF domain consists of approximately 40 amino acid residues arranged in an N-terminal loop, followed by a major β-hairpin and then a minor β-hairpin held in position by three disulphide bonds securing the β-hairpin and the N-terminal loop

into a folded bundle (Figure 1.4). This is supported by structural evidence from NMR (Cooke et al, 1987) and X-ray crystallography (Rao *et al*, 1995). The six conserved cysteines are interspersed through the sequence forming the three canonical disulphide bonds in the endoplasmic reticulum during passage through the secretory pathway (Handford et al, 1990) which can be recapitulated in recombinant EGF constructs which are reduced and refolded *in vitro* (Whiteman et al, 2014).

1.3.2 Calcium binding

The Ca^{2+} binding ability of cbEGFs arises due to the presence of conserved carboxylate/carboxamide residues in the cbEGF domain. This consensus sequence was initially identified in human Factor IX (D/N-X-D/N-E/Q-X_a-D*/N*-X_b-Y/F) where a and b are variable and * denotes β-hydroxylation (Figure 1.4) (Handford et al, 1991, Rees et al, 1988, Baron et al, 1992). An essential feature are the lone pair-donor residues at the extreme N-terminus which act as ligands to correctly coordinate a Ca^{2+} ion usually in a pentagonal bipyramidal geometry. Up to seven O-ligands are provided by the sidechains of the residues in the consensus positions, as well as the backbone carbonyl oxygens of the residues adjacent to the metal ion (Lee et al, 2004), water may also act as a ligand as observed in cbEGF10 in Fibrillin-1 (Jensen et al, 2009) (Figure 1.5). In addition, two other consensus residues (second and fifth) stabilise the calcium-binding site. The second consensus residue forms an 'ASX' turn, and the fifth residue (an aromatic) packs against other hydrophobic residues between Cys1 and Cys2 (Rao et al, 1995).

1.3.2.1 Water as a ligand for Ca^{2+}

Although in all seven Ca^{2+}-ligands can be provided by peptide, it has been observed in X-ray crystallography, that water can act as one or two of the Ca^{2+}-ligands in a cbEGF (Cordle et al, 2008 Jensen et al, 2009). In the crystal structure of hN1 11-13, cbEGF11 is observed to involve one water molecule coordinating the consensus Ca^{2+}, as EGF12 is observed to involve two water molecules (Cordle et al, 2008). hN1 11-13 is suggested to be rigid-rod like using X-ray crystallography and NMR (Cordle et al, 2008 , Weisshuhn et al, 2016), which suggests that variations in the number of Ca^{2+}-ligands provided by water does not impact flexibility of an EGF interface. To further support this, the rigid crystal structures of tandem EGFs from Notch contain Ca^{2+} ions coordinated by peptide oxygens and one or two waters, and is therefore unlikely to be a key driver for flexibility.

1.3.3 Classes of cbEGF pairs

There are two classes of cbEGF pairs which differ by an extra residue in the connecting loop between domains (Downing et al, 1996), yet both form rod-shaped structures. Class I pairs contain a single-residue linker compared with a two-residue linker in class II pairs. In general, a multi cbEGF/EGF domain proteins contain only a single class of EGF pairs, and Notch proteins contain only class II EGF pairs. The classes have a difference in interdomain twist, 119° in class II pairs relative to 159° in class I. Both have a similar overall extended rod-like structure mediated by hydrophobic interactions between adjacent domains, although the consensus packing residues vary slightly. This interdomain interaction takes place between a conserved aliphatic residue in the β-

hairpin which interacts with the conserved aromatic at the end of the previous domain (Figure 1.6). Because of these inter-domain interactions, the shape of an inter-EGF domain hinge is determined by the context of the domain.

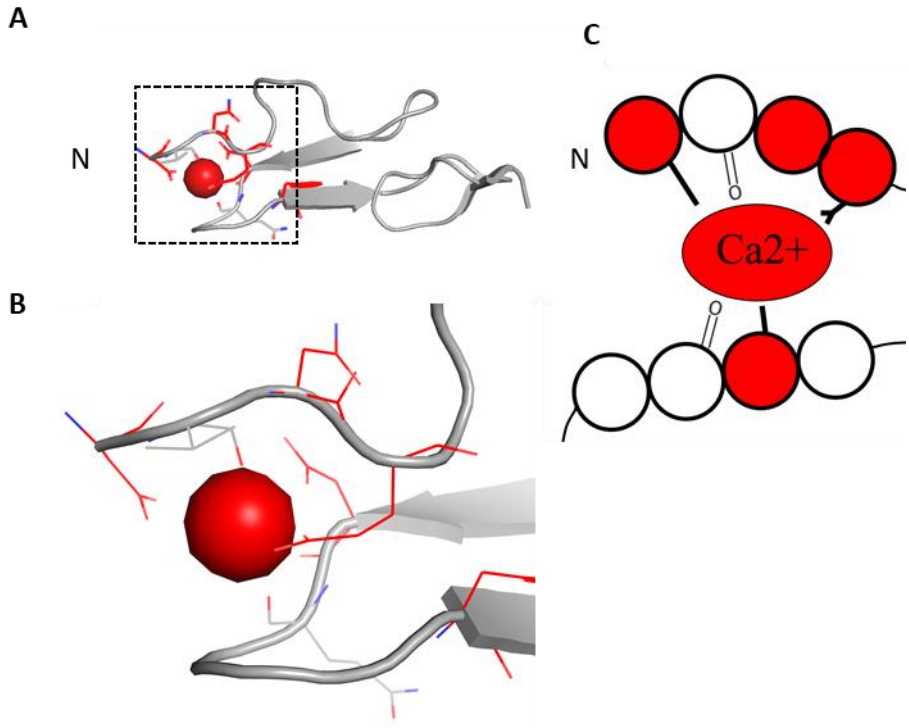


Figure 1.5 The cbEGF Ca^{2+}-binding consensus: (A) a cbEGF (grey) with N- and C- termini labelled and Ca^{2+} consensus indicated (red sticks). Three sidechain oxygens of the domain provide most of the ligands for the ion (red sphere), and two other residues provide a ligand through their backbone carbonyls. (B) Region enclosed by the dashed box enlarged to show Ca^{2+} coordination. (C) Schematic showing Ca^{2+} ligands from the protein.

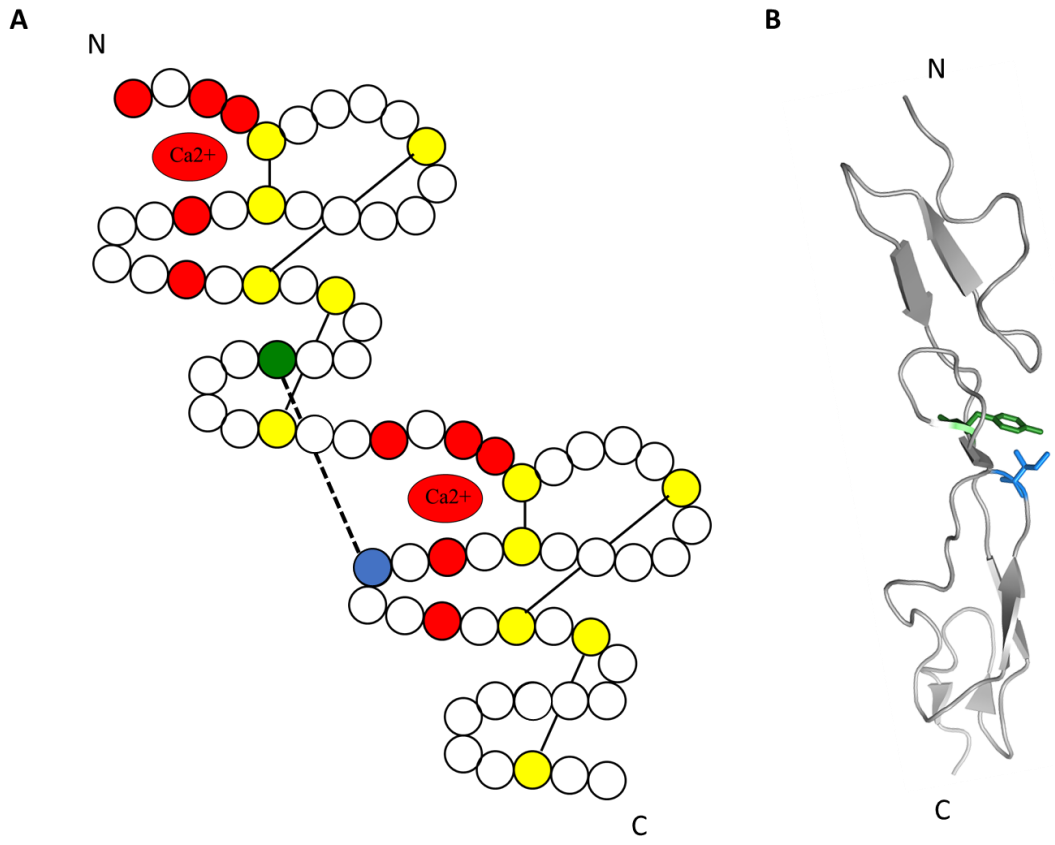


Figure 1.6 The class II cbEGF pair: A class II cbEGF pair typical of Notch is shown as a two-dimensional schematic (A) and as a three-dimensional structure (B). Conserved cysteines (yellow) are shown with connecting disulphides (black lines), with calcium binding-consensus residues (red). The conserved packing aromatic (green) in A and B, interacts with the conserved aliphatic in the following domain (blue). (B) Shows the proximity of these residues in the folded protein. (PDB ID: 2VJ3).

1.3.4 Post-translational modification

EGF domains are located outside the cell as part of extracellular or membrane proteins and therefore undergo oxidation in the secretory pathway, allowing disulphide bonds to form. These domains can also undergo a large variety of post-translational modifications during this process if the consensus sequences for recognition by modifying proteins are present (Figure 1.7). The impact of these modifications can be both structural and functional. The predicted EGF fucose and glucose modifications in hN1 are described in the Appendix (Table A.1).

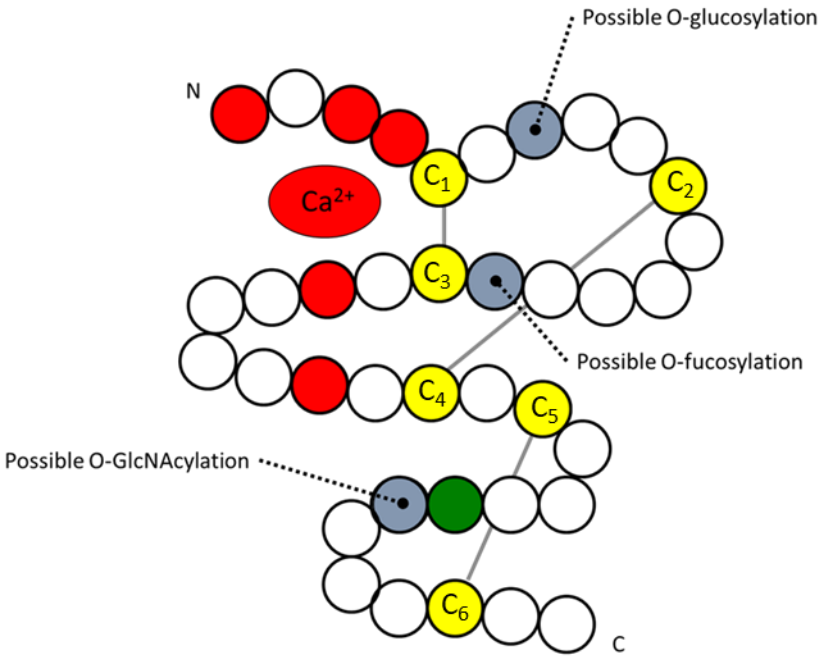


Figure 1.7 EGF O-linked glycosylations: A two-dimensional schematic of a cbEGF domain is displayed showing calcium binding consensus residues (red), conserved cysteines (yellow) and disulphides (grey lines), and conserved packing aromatic (green). Possible sites of O-linked glycosylation are indicated (blue).

1.3.4.1 N-linked glycosylation

Notch receptor EGF domains contain multiple N-linked glycan glycosylation consensus sequences (N-X-S/T, where X ≠ P) although it is widely agreed that these have no specific role in signalling. This is supported by evidence from CHO cells suggesting that alterations to the N-glycan pattern does not affect signalling (Takeuchi and Haltiwanger, 2010). Further support for this is the observation that mice unable to form complex N-linked glycans are not defective in Notch signalling during embryonic development (Stanley, 2007).

1.3.4.2 O-linked fucosylation

O-linked fucosylation is carried out by POFUT1 at the consensus sequence C_2-X-X-X-X-(S/T)-C_3 (Harvey et al, 2016). The expression of this enzyme is essential for viable embryonic development and knockout mice show multiple defects development comparable to knocking out components of the Notch signalling pathway, suggesting this effect is mediated through Notch (Shi and Stanley, 2003). O-fucose can be further sequentially modified by Fringe enzymes to produce glycans up to Siaα2-6Galβ1-4GlcNAcβ1-3Fuc (Figure 1.8). Mass spectrometric analysis has shown that all Notch1 O-fucose sites are modified in HEK-293T cells with fucose at high stoichiometry, except EGFs 23, 24, 26, 32 (Kakuda and Haltiwanger, 2017).

On a molecular level, O-fucosylation is important for Notch receptor production and function, and for trafficking of Notch to the cell membrane in HEK 293T cells (Takeuchi et al, 2017). O-linked fucoses are also involved in direct contacts in interactions with ligand for example T466-O-fuc in EGF12 is part of the binding interface for binding Dll4 (Delta-like4) and Jag1 (Jagged1) (Taylor et al, 2014 , Luca et al, 2015 , Luca et al, 2017). There are also ligand-specific effects of these modifications elsewhere in the receptor, which have been observed in Notch1 when substituting the consensus modified Thr for Val in EGF9 and EGF27. These substitutions reduced activation by ligand Jag1 but not Delta like1 (Dll1). On the other hand, the same substitution in EGF26 reduced activation by Dll1 but not Jag1 (Kakuda and Haltiwanger, 2017).

Further modification of these fucoses with GlcNAc by Fringe enzymes acts as an important regulator of Notch signalling specificity. *In vitro* binding assays show that modification of EGF12-fucose with GlcNAc leads to an enhancement of binding to cells expressing ligands such as Dll4 (Benedito et al, 2009), Jag1 and Dll1 (Taylor et al, 2014).

In the context of development, Fringe is essential for wing formation in *D. melanogaster*. Mammals have three Fringe enzymes Manic, Radical, and Lunatic, each responsible for different aspects of development and modifying different EGFs across the Notch receptors. An example of this is Manic and Lunatic acting in T-cell development by regulating modification of hN1 on T-cell progenitors for interaction with Dll4 on thymic cells (Stanley and Guidos, 2009). The extension of O-Fuc-GlcNAc with galactose has a marginal effect on development as seen in investigations in mice with the primary Notch-modifying galactosidase βGalT-1 knocked out (Chen et al, 2006).

1.3.4.3 O-linked glucosylation

Glucose can be added to EGF domains at the consensus region C1-X-S-X-P/A-C2 by protein O-glucosyltransferases (POGLUT1, POGLUT2, POGLUT3). This glucose can then be extended by the addition of two xyloses (Figure 1.8). Glucosylations have been observed to stabilize the fold of the modified EGF domain. For example, the addition of a glucose to EGF28 in mN1 leads to an increase in the resistance of the domain to unfolding in the presence of Dithiothreitol (DTT) (Takeuchi et al, 2017). This is supported by recent structural evidence of an EGF domain from human coagulation Factor IX which shows a trisaccharide O-Gluc glycan interacting with the surface of the domain (Takeuchi et al,

2017). This interaction involves conserved residues in the protein, suggesting that this stabilizing effect of the glycan may be widespread. Interestingly, the addition of the first xylose to glucose seems to reverse the stability-enhancing effect of the glucose as observed in unfolding stability experiments. This destabilisation can again be stabilised by the addition of the second xylose (Takeuchi et al, 2017), a phenomenon also observed in *D. melanogaster* development (Lee et al, 2013). In mammalian receptors, most O-Gluc sites are modified with two xyloses (Lee et al, 2013). However, this extension is restricted in dN to EGF14-20 only, with only EGF16 and EGF18 being modified with two xyloses, despite glucosylation occurring over the entire length of the receptor. A recent study investigating the O-Gluc modification in EGF11 suggests that the combination of O-linked glycans around the LBR can determine efficiency of signalling by ligand (Takeuchi et al, 2018).

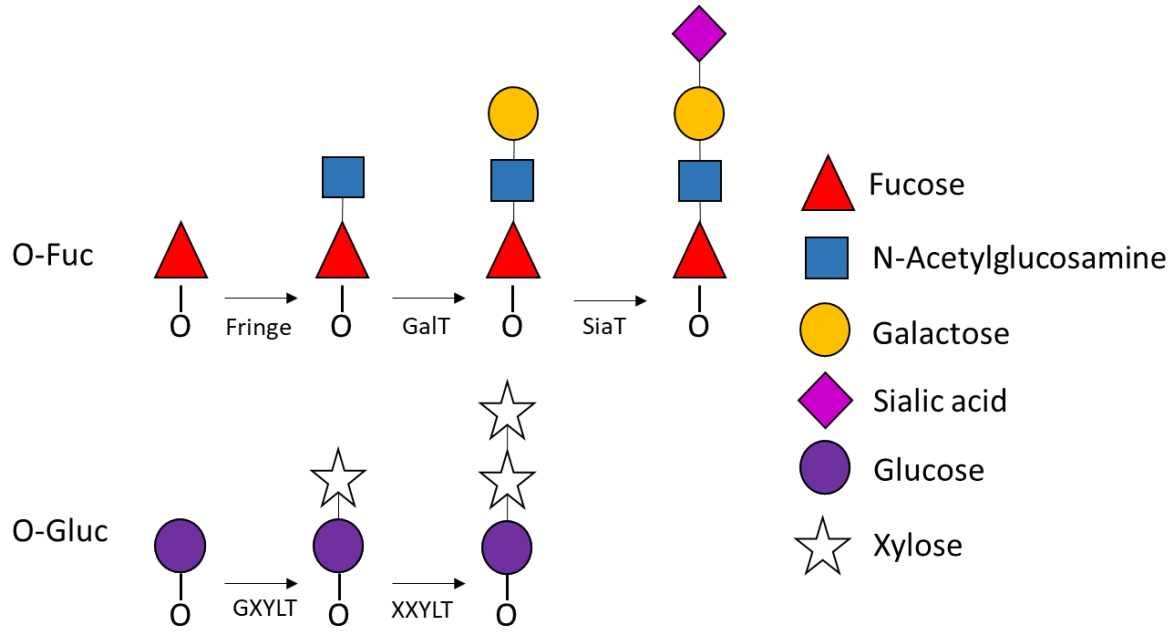


Figure 1.8 Notch glycans can be sequentially modified: O-Fucose modifications can be extended with GlcNAc by Fringe, and then with galactose by GalT and sialic acid by SiaT in mammals. O-Glucose glycans can be extended with xylose by GXYLT and XXYLT. Sugars are labelled according to Consortium for Functional Glycomics guidelines.

1.3.4.4 O-linked β-N-acetylglucosaminylation

A number of EGF domains of Notch are modified on Ser or Thr in the consensus sequence (C_5-X-X-G-(Y/F/L)-(T/S)-G-X_{2-3}-C_6) with N-acetylglucosamine (GlcNAc) by the enzyme glycosyltransferase (EOGT) (Matsuura et al, 2008) (Figure 1.7). Mutations affecting the gene encoding this enzyme are associated with Adams Oliver syndrome (Ogawa et al, 2015) although it is unclear if these are effects are due to an impact on Notch signalling. EOGT gene deletions in mice show phenotypes indicative of decreased ligand binding. Furthermore, cell co-culture assays show that EOGT-deficient cells expressing receptor show decreased binding to Delta-family ligands Dll1 and Dll4, while binding to Jag1 is unaffected, suggesting a role for these modifications in development (Sawaguchi et al, 2017).

1.3.4.5 β-hydroxylation

EGF domains can be modified by β -hydroxylation of asparagine and aspartate residues at the consensus sequence C_3-X-(D/N)-X-X-X-X-(Y/F)-X-C_4 by the enzyme aspartyl/asparaginyl B-hydroxylase (Harris and Spellman, 1993). A mutation in the gene expressing this enzyme in mice leads to a phenotype similar to that associated with mutations in Jagged2 (Dinchuk et al, 2002).

1.3.5 EGF-domain interfaces

Tandem EGF domain class II pairs (cbEGF-cbEGF, or EGF-cbEGF, see Section 1.3.3) present in the Notch receptor have interdomain contacts such as the interaction shown in Figure 1.9. The packing between domains in the cbEGF-EGF pairs can vary and lead to a range of shapes and flexibilities at inter-domain boundaries, influencing the shape of the receptor and any intramolecular interactions that might occur.

1.3.5.1 Shapes of tandem EGF-domain constructs

Rigid rod-like interfaces are well characterised in tandem cbEGF domain constructs. These rigid interfaces arise due to Ca^{2+} binding at the N-terminus of the domain rigidifying this part of the EGF for the hydrophobic packing interaction with the previous domain as shown in Figure 1.6 (Downing et al, 1996, Hambleton et al 2004). This elongated rod-like structure is observed in EGF11-13 of hN1 (all of which are cbEGFs) as observed in X-ray crystallography (Cordle et al, 2008), and supported by the observation that the interaction face with ligand extends across multiple EGFs (EGF11-12) which would only be compatible with an elongated conformation (Luca et al, 2017).

Interfaces between cbEGFs and EGFs are much more variable and can result in a variety of different shapes and properties. An example of this is the interface between hN1 EGF5-EGF6 (cbEGF-EGF) which produces a bent conformation and tilt angle of 80-90° between the two domains (Weisshuhn et al, 2016). This is brought about by two pairs of residues interacting at the interface, Y219 and A208 in EGF5 interacting with P221 and V239

respectively in EGF6 (Figure 1.9) (Weisshuhn et al, 2016). However, cbEGF-EGF interfaces can also be rod-like as in the case of EGF21-22 when EGF21 is Ca^{2+}-bound. This is due to an interaction between P280 near the C-terminus of EGF21 with Y850 on the tip of the β-hairpin in EGF22 (Suckling et al, unpublished data) (Figure 1.9). Interfaces can also be flexible, for example the EGF9-10 (cbEGF-EGF) interface in hN1, which lacks the packing aromatic shown in Figure 1.6 (Weisshuhn et al, 2016).

1.4 Notch receptors

1.4.1 Introduction to Notch receptors

Large parts of the extracellular tandem EGF region of hN1 have been structurally characterised (Figure 1.10) During synthesis in the endoplasmic reticulum, a furin-dependent S1 cleavage in the NRR must occur for successful trafficking to the cell surface in mammalian systems (Logeat et al 1998), but is not required for successful secretion in *D. melanogaster* (Kidd and Lieber, 2002). Two further proteolytic cleavages (S2 and S3) take place on interaction with ligand, and lead to successful signalling (Section 1.2.3).

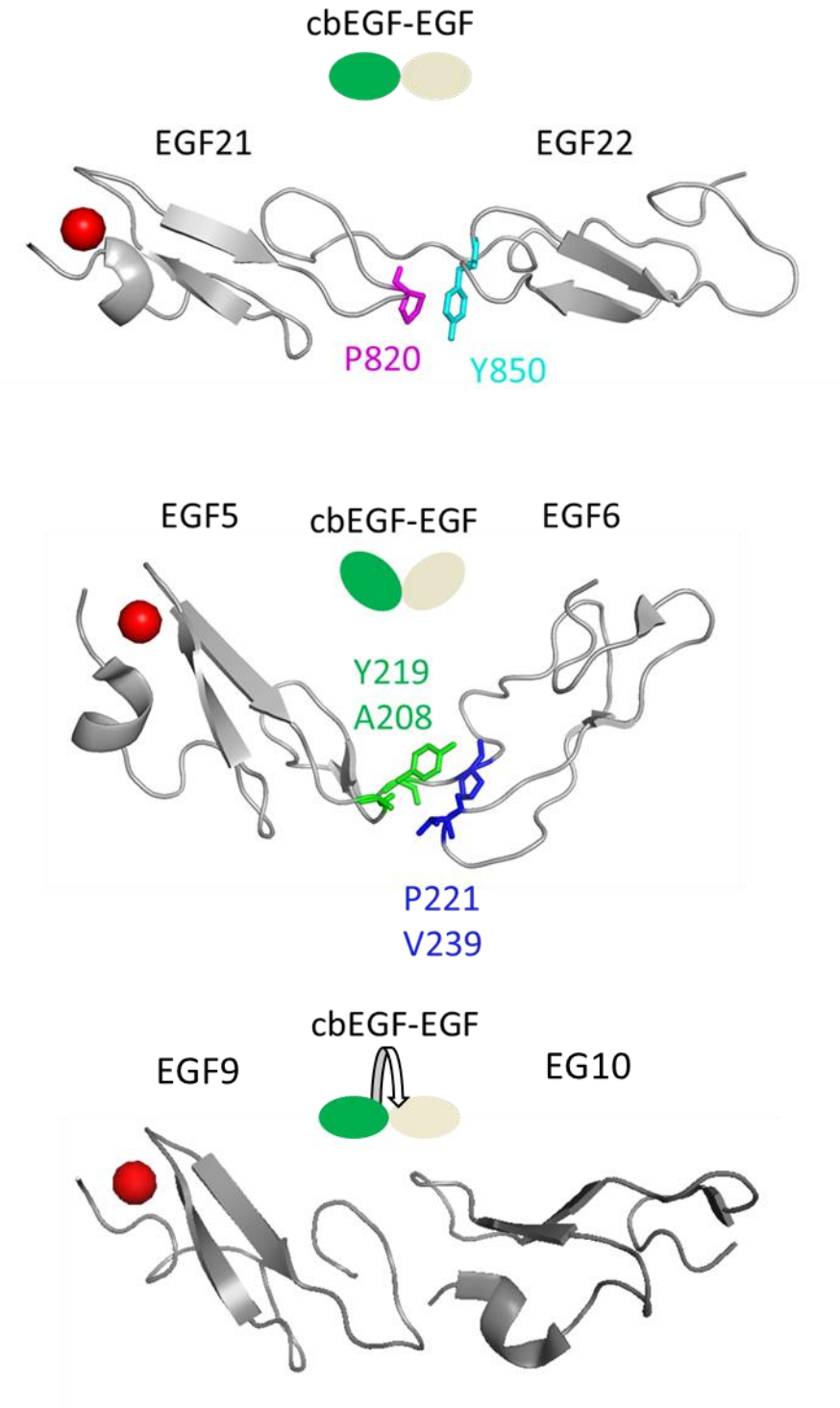


Figure 1.9 cbEGF-EGF interfaces are variable: hN1 EGF21-22 is shown as determined by X-ray crystallography in the context of EGF20-24 (Suckling et al, unpublished data). P820 (magenta) and Y850 (cyan) stack against one another, stabilising the rod-like interface. hN1 EGF5-6 is shown as determined by X-ray crystallography in the context of EGF4-7 (Weisshuhn et al, 2016, PDB ID: 5FMA). EGF9-10 has been determined to be flexible (indicated by white arrow) by NMR (Weisshuhn et al, 2016, structure for region PDB ID: 5UK5) with Ca^{2+} ions shown (red spheres). Interdomain packing residues are indicated.

Cryo-electron microscopy has been previously attempted to investigate the shape of the complete Notch receptor, suggesting the receptor is present as a dimer (Kelly et al, 2010), although the purity of the analysed protein could not be verified due to the affinity-grid selection approach used. Kelly et al. attempted to purify constructs containing the extracellular domain of Notch with a C-terminal His-tag, from growth medium of cells expressing very low amounts of the construct (<0.05 mg/L cell culture). This purification was carried out by incubating a Ni^{2+}-coated EM grid directly with the cell culture medium, without any additional purification steps, and without confirmation of purity of the material used in the EM. 3D reconstructions of the particles observed produces three structural models proposed to three distinct conformations of the receptor, with the mass of these models roughly corresponding to dimers of the receptor. Antibodies were used to determine the location of refence points of the receptor, however the location of the added mass is difficult to conclusively define at the low resolution of these models. Considering the lack of data supporting the purity of samples used in the EM experiment, and the poor expression of the cells used produce the construct, we are cautious about drawing conclusions from this study.

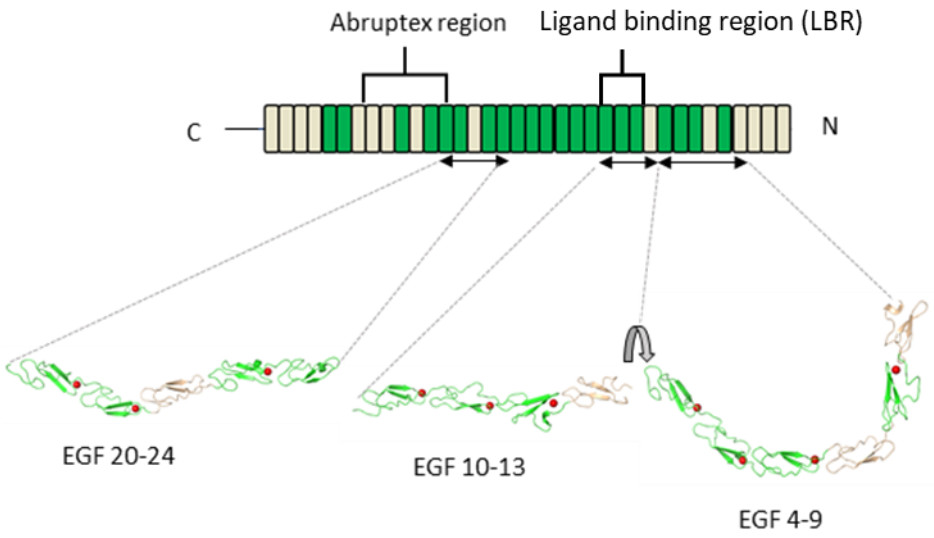


Figure 1.10 The structurally characterised EGF regions of the hN1 receptor: EGF (wheat) and cbEGF (green) domains are shown schematically and in three-dimensions for the structurally characterised sections EGF4-9 (PDB ID EGF4-7: 5FMA) and EGF10-13 (PDB ID: EGF11-13: 2VJ3) (Weisshuhn et al), EGF20-24 (Suckling et al, unpublished data) with Ca^{2+} ions shown (red spheres). The transmembrane region is located at the C-terminus of the receptor.

1.4.2 Ligand-binding region

The ligand binding region of the Notch receptor has been well characterised and consists of key sites of interaction in EGF11 and EGF12 of the receptor (Rebay et al, 1991, Whiteman et al, 2013, Luca et al, 2015), with some contribution from EGF8-10 (Luca et al, 2017). The primary sites in EGF11-12 can engage in ligand binding through an interface of conserved residues and O-linked sugars, centred around L468 and T466 in EGF12, and R448 in EGF11 (Figure 1.11). The construct EGF11-13 containing the ligand binding region of hN1 has been structurally characterised using X-ray crystallography, and NMR indicating a rod-like conformation of these cbEGFs. This structure is conserved in hN2 11-13, superimposing on hN1 with an RMSD of 1.55Å despite having 67% sequence identity,

(Sucking et al, 2017). The importance of sugars in this interaction has been demonstrated using flow cytometry to show interactions between purified ligand binding region with ligand on the surface of cells is enhanced when using O-linked glycosylated material. In these experiments, the addition of the disaccharide (Fringe modified) had the biggest effect on binding (Taylor et al, 2014).

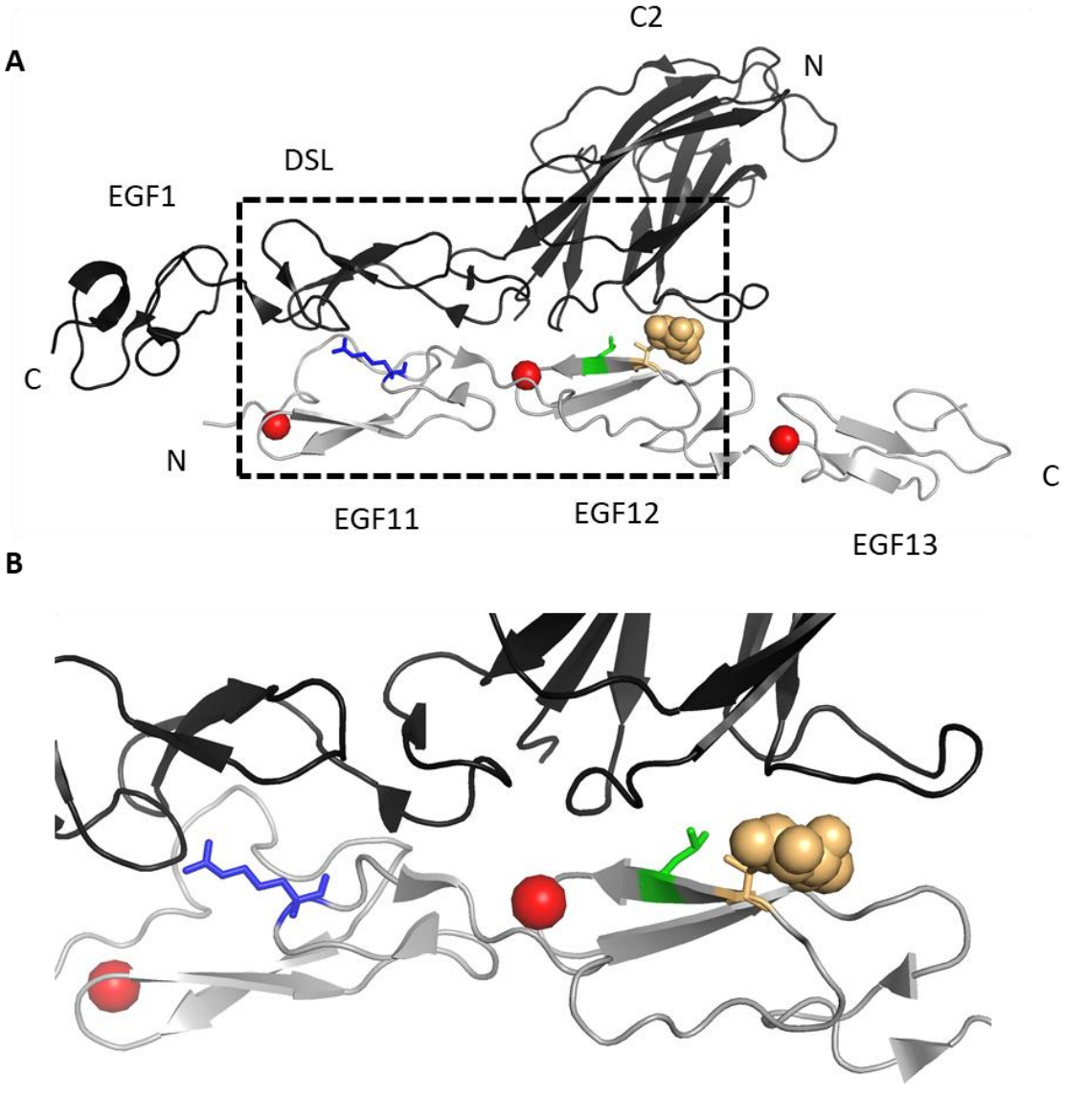


Figure 1.11 The ligand-binding region of Notch1: (A) The ligand Dll4 (black) interacts with the LBR of Notch1 (grey). Interaction sites on the receptor are indicated in EGF11 centred around R448 (blue, stick structure) and in EGF12 around L468 (green, stick structure) and T466-Fuc (orange, ball structure). Ca^{2+} ions are shown in red. PDB ID = 4XL1. (B) shows an enlargement of the region enclosed by the dashed box.

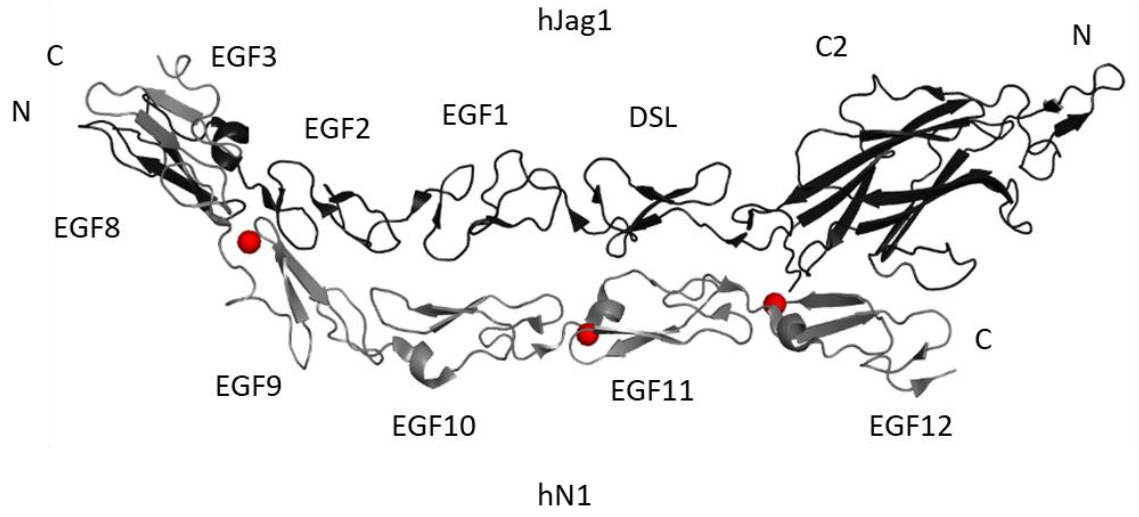


Figure 1.12 The extended interaction face of LBR and Jagged1: The ligand Jagged1 (hJag1, black) interacts with the LBR of Notch1 (hN1, grey) (PDB ID: 5UK5). Ca^{2+} ions are shown (red) and domains labelled.

1.4.3 Abruptex region

The Abruptex (Ax) region of the receptor was initially defined in *D. melanogaster* as a region consisting of EGF24-29. Within this region, missense mutations occur that lead to enhancement of signalling (de Celis and Garcia-Bellido, 1994) and more complex phenotypes when expressed over a *N* phenotype background (See Section 1.2.1). Sequence alignment suggests that the organisation of EGF domains is conserved amongst receptors of similar length.

1.4.3.1 The Abruptex phenotype

The Abruptex phenotype was first recognised as a phenotype in the wing (Nasarenko, 1930, Portin, 1975) and later observed to result from single nucleotide polymorphisms

located in extracellular EGFs of dN (Hartley et al, 1987). More recently, two additional mutants in EGF16 produced a signalling-enhancing phenotype (Yamamoto et al, 2012) although the classical Ax region is still defined as EGF24-29. These mutants are outlined in Figure 1.13.

Allele	EGF	Amino acid change	Effect on *Notch* (N) haplo-insufficiency phenotype*
egf16-AxGR	16	G668R	Not investigated
egf16-AxGD	16	G671D	Not investigated
NAx-9	24	D948V	*N* suppressor
NAx-59b	24	C972G	Lethal
NAx-59d	24	C972S	Lethal
NAx-1	25	N986I	*N* suppressor
NAx-M1	25	C999Y	Lethal
NAx-71d	27	S1088I	*N* enhancer
NAx-16	29	G1174A	*N* enhancer
NAx-E2	29	H1167Y	*N* enhancer

*Effect on *Notch* haplo-insufficiency phenotype (loss of the wing margin) when co-expressed on *Notch* haplo-insufficiency background.

Figure 1.13: Abruptex alleles identified in *D. melanogaster*: Mutations described in de Celis and Garcia-Bellido, 1994 , Yamamoto et al, 2012.

The Abruptex phenotype is generally accepted to result from too much signalling (De Celis and Garcia-Bellido, 1994 , De Celis and Bray, 2000), although there are theories explaining the Ax phenotype without implying increased signalling (Brennan et al, 1999). The impact of this allele in isolation is a decrease in the number of macro and microchatae on the

scutum of the fly, and a decrease in the length of the wing veins, the opposite to the *N* mutations (mutations associated with decreased Notch function due to deletions or null alleles (De Celis et al, 1991) (see Figure 1.14). There is some suggestion that Ax mutations produce their effect by having reduced susceptibility to the inhibitory impact of ligand. Investigating the effect of Ax mutants at the DV boundary in the wing shows that Ax alleles have the same impact as deficiencies in the ligands Delta and Serrate under conditions when these ligands are exerting inhibitory effects, which led to the suggestion that Ax is involved in mediating this interaction with ligand (De Celis and Bray, 2000). Furthermore, Ax mutant receptors bind *in vitro* with a lower affinity to Delta than wildtype receptor (Lieber et al, 1992). However, no interaction between Ax and ligand has been observed. Altered interplay between Ax alleles and the GlcNAc transferase enzyme Fringe (See Section 1.3.4.2) have also been suggested after observing that Ax alleles can restore normal function in altered Fringe phenotypes (Irvine and Wieschaus, 1994). In this case, each Ax allele has a characteristic effect, possibly suggesting that the conformation of the region is of great importance in either binding, or positioning the receptor to be supressed. It remains to be determined if similar molecular phenotypes associated with mutations in this region may occur in the mammalian Notch receptors.

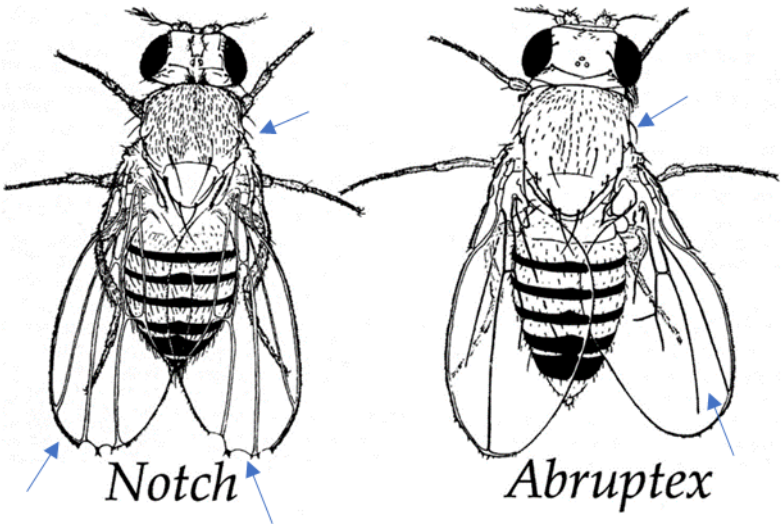


Figure 1.14 The Abruptex phenotype: The typical characteristics of the *N* phenotype (*Notch*) and Ax phenotype (*Abruptex*). The *N* mutant has enlarged wing veins and more bristles, as well as notches along the edge of the wing (blue arrows). The Abruptex mutant has thinner and fewer wing veins and fewer bristles. The wild type fly would fit between these descriptors, without notches in the wing and with moderate wing vein thickness and moderate frequency of bristles (adapted from Fryxell et al, 2001).

An important feature of these Abruptex mutants is the complex phenotypes when they are expressed in addition to other mutant NOTCH alleles (de Celis and Garcia-Bellido, 1994). As the NOTCH is present on the Drosophila X-chromosome, it is possible to easily see the impact of the mutant allele in isolation in hemizygous males. These combination effects enable categorisation into three classes of Abruptex allele; Lethal, *N*-enhancer, and *N*-suppressor, each with their own severity (see Figure 1.13).

Lethal alleles are not viable when homozygous, or in combination with the *N* allele or another Ax allele (Foster 1975, Portin 1975). These are all located in EGF24 and 25 and cause a substitution in a conserved cysteine residue involved in the canonical disulphide bonding in each cbEGF and likely to cause large structural disruption possibly leading to misfolding and faulty expression. *N*-enhancer and *N*-suppressor alleles act to respectively

enhance and supress the *N* phenotype when expressed on top of an *N* allele. An alteration in the phenotype suggests that signalling must involves interplay between multiple receptors from each allele interacting together to output a signal. *N*-enhancer and *N*-suppressor alleles negatively complement together and lead to lethality (de Celis and Garcia-Bellido, 1994), suggesting that there might be two distinct functions carried out by the Ax region at the N and C termini of the region.

1.4.3.2 Proposed interaction of Ax with LBR and models

Considering that Abruptex mutants act to increase signalling of the receptor, this suggests that the normal function of the region is to negatively regulate signalling. Sharma et al, 2013 and Pei and Baker, 2008 have both observed an interaction between constructs containing the LBR and the Ax region in different receptors and by different methods, providing a possible mechanism by which this negative regulation takes place. This proposed mechanism involves Ax competing with ligand for binding to the LBR.

Pei and Baker, 2008 observed this interaction between immobilised dN EGF11-14 and dN EGF21-30 in a plate-based assay. They further observed no interaction between EGF21-30 and the ligand Delta. They also observed that the Delta can compete with EGF21-30 for binding to EGF11-14, supporting the hypothesis that ligand and Ax compete for binding with the LBR. However little information on purity of these fragments was provided raising potential questions on the true nature of the interaction observed.

Sharma et al, 2013 observed a similar interaction using Surface Plasmon Resonance (SPR) between immobilized hN1 11-15 and hN1 21-30 and determined the K_d of this interaction to be 0.22 μM (K_d hJag1 to LBR = 2.93 μM) and observed no binding of the Abruptex region to ligand.

1.4.4 Negative regulatory region

The Negative Regulatory region is the site of the masked cleavage site (S2 cleavage site) which is revealed on binding to ligand and resulting mechanical pulling. This is cleaved by ADAM metalloprotease and leads to the release of the NICD triggering downstream events (Sanchez-Izirarry, 2004). X-ray crystallography shows the region consists of three Lin12/Notch repeats with a conserved hydrophobic plug which occludes access to the S2 site (Figure 1.15) (Gordon et al, 2009). Each LNR binds to Ca^{2+} and these ions are suggested to stabilize the structure and ensure shielding of the S2 cleavage site. During signalling, the application of force on ligand binding reveals this occluded site, which is then cleaved by an ADAM metalloprotease, and is subsequently cleaved by γ-secretase at the S3 site in the transmembrane region (not shown in Figure 1.15) to release the NICD from the membrane (De Strooper et al, 1999).

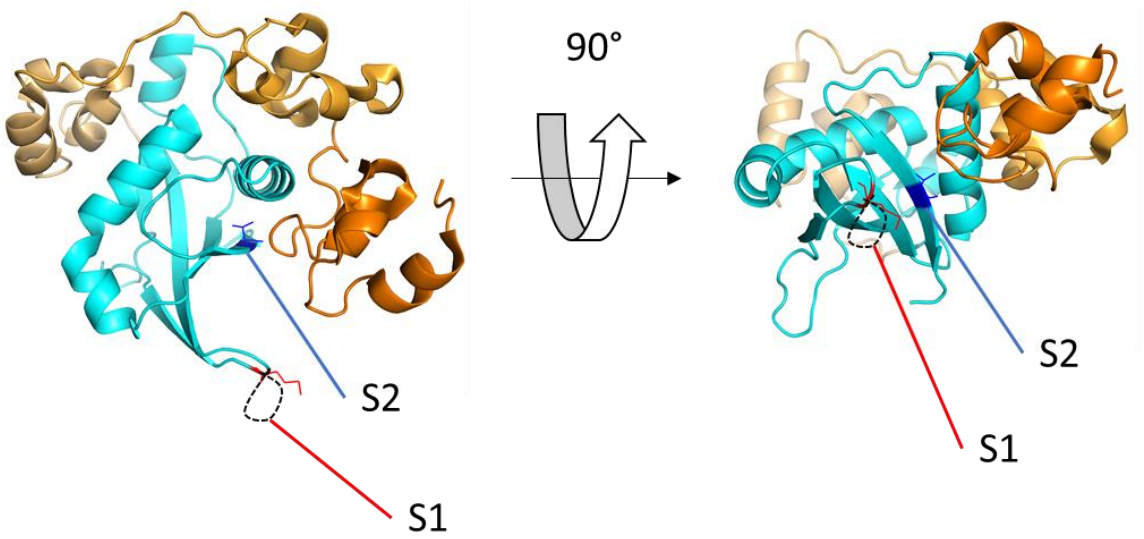


Figure 1.15 hN1 negative regulatory region: The NRR consists of three LNR domains (orange) each contain a Ca^{2+} ion (not shown) and the heterodimerisation domain (cyan). The sites of the cleavages are in the S1 Loop (red), and S2 (blue). PDB ID = 3ETO.

The requirement for force to reveal the S2 site for cleavage has recently been shown to be 3-4 pN, comparable to that of the myosin power stoke during muscle contraction (Gordon et al, 2015). The force required is proposed to be provided during endocytosis of the receptor-ligand complex based on the measured force for revealing the S2 site to be similar to that of Dll1 and EGF receptor endocytosis, and less than the force to break receptor-ligand interaction (Meloty-Kapella et al., 2012; Shergill et al., 2012, Gordon et al, 2015). This magnitude of force for activation is supported by evidence from Luca et al 2017, and Chowdhury et al, 2016, which show a force greater than 4 pN but less than 12 pN is required to initiate signalling with both Jag1 and Dll1.

1.4.5 Intracellular region

The NICD consists of multiple ANK repeats which are important for the interaction with the transcriptional co-effector RBPJK/CSL and with other proteins which are able to regulate Notch signalling. After cleavage by γ-secretase at the S3 site in the receptor following signalling, the NICD moves to the nucleus to alter the expression of genes, exerting the effect of the Notch signalling pathway on development and cell maintenance. The transcriptional co-effectors to which NICD binds are RBPJK/CSL along with co-activator MAML proteins. The resting state of RBPJK/CSL is bound to transcriptional repressors and inhibits expression from genes associated with Notch signalling (Borggrefe and Oswald, 2009). The RAM domain of NICD binds to the Beta-trefoil domain of RBPJK followed by the binding of the ANK domain of NICD to the transcriptional effector (Kopan and Ilagan, 2009). A large number of other transcriptional co-effectors can bind to and influence the complex, for example to modify chromatin or to potentiate transcription. An example of the genes regulated by Notch signalling are the Hes/Hey family of transcription factors which are involved in suppressing the expression of tissue-specific transcription factors in stem cell maintenance. Recent studies using real time analysis suggest that the presence of the co-effector complex increases chromatin accessibility and therefore enables rapid regulation of target gene transcription (Gomez-Lamarca et al, 2018).

1.5 Canonical Notch ligands

Notch ligands are type 1 transmembrane proteins which act to bind the Notch receptor primarily through a DSL domain (Figure 1.16). Exceptions to this are the soluble ligands in *C. elegans*, and the non-canonical Delta like 3 (Dll3) ligand. There are two families of ligands, the Delta and Serrate families classified according to similarity to the two *D. melanogaster* ligands. Mammals have five ligands; Dll1, Dll3, Dll4 (Delta family), Jag1, Jag2 (Jagged/Serrate family) (Kopan and Ilagan, 2009). Dll3 lacks a canonical DSL domain and is unable to activate Notch and is found mainly in the Golgi apparatus (Ladi et al, 2005; Geffers et al, 2007; Serth et al, 2015). The other four ligands are defined as canonical ligands, which interact with the LBR of the receptor on an adjacent cell (in trans) to initiate signalling (transactivation). Ligands can also interact with receptor on the same cell (in cis) to suppress signalling ability of those receptors (del Alamo et al, 2011).

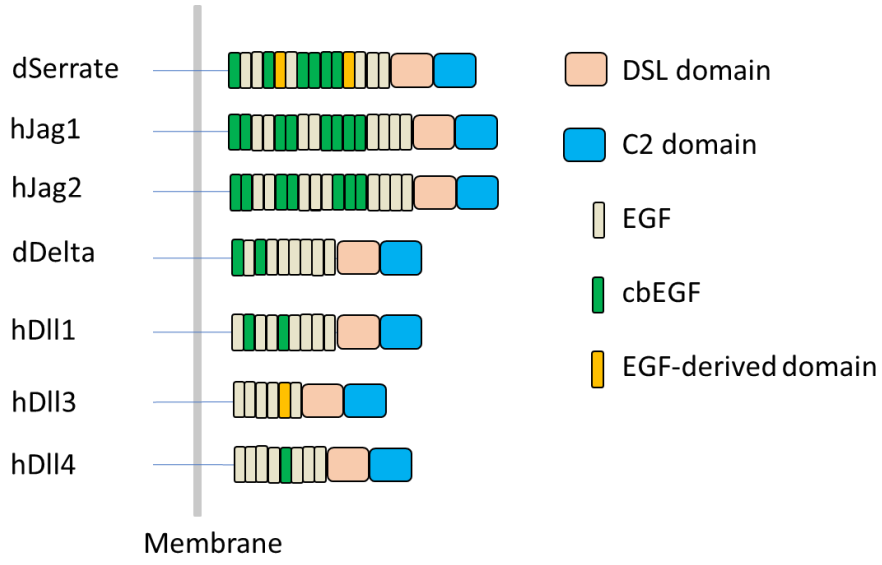


Figure 1.16 The structural organisation of the extracellular region of Notch ligands: Notch ligands from *D. melanogaster* (prefix d), human (prefix h) are shown. Domains constituting each receptor are described, DSL = Delta-Serrate-Lag2 domain. Jagged family ligands are longer than Delta

1.5.1 Structure of ligands

Notch ligands have a modular structure similar to the Notch receptor. The minimal binding module required for receptor binding is the DSL domain (Shimizu et al, 1999). Mutant studies also suggest that EGF1-2 (see Figure 1.16) are important in the interaction with Notch, also known as the conserved minimal EGF contribution (DOS), and this is supported by recent structural evidence detailing the involvement of these EGFs in binding to the receptor (Luca et al, 2017). The role of the DOS region can be substituted for non-canonical DOS-containing ligands when signalling occurs through DSL ligands lacking the DOS. (Komatsu et al, 2008). At the extreme N-terminal of each mammalian ligand is the C2 domain, which has been shown to bind lipids (Chillakuri et al, 2013, Suckling et al, 2017), as well as Notch (Luca et al, 2015, Luca et al, 2017). C2 domains are also present at the N-terminus of the *D. melanogaster* ligands (Suckling et al, unpublished data)

1.5.2 Delta/Serrate/Lag-2 domain

The DSL domain is essential but not sufficient for activation of ligand with the receptor. The crystal structure of this domain revealed a similar fold to an EGF domain but distinct, with conserved residues making up a Notch binding site located on one side of the domain supported by functional assays (hJag1 F199, R201, R203, F207) (Cordle et al, 2008). Along with EGF11-12 in the receptor, the DSL domain makes up the minimum unit for both cis and trans interactions (Becam et al, 2010). The nature of the differences in signalling arising from the different ligand families (Jagged and Serrate) suggest that the

interactions with the receptor are distinct. Until recently, isolation of a receptor ligand complex had been precluded by the relatively weak interaction affinity (130 μM) (Cordle et al, 2008). Recent co-crystal structures of receptor bound with affinity enhanced Dll4, and Jag1 give details of these interactions (Luca et al, 2015, Luca et al 2017) suggesting an extended interface centred around the DSL, with a greater dependence of DSL for binding to Dll4 than Jag1. (Luca et al, 2017).

1.5.3 C2 domain

The so-called 'Module at the N-terminus of Notch ligands' (MNNL) domain of Notch ligands was initially suggested to be important in binding because of its conserved position, always N-terminal to the DSL domain. Furthermore, deletion of the MNNL domain leads to loss of function in the *C. elegans* ligand LAG-2 (Henderson et al, 1997). The crystal structure of this domain revealed a C2-domain fold, a domain previously observed to be involved in Ca^{2+}-dependent phospholipid binding (Chillakuri et al, 2013 , Cho and Stahelin, 2006).

Ca^{2+}-binding loops of the C2 domain have been observed for the Jagged family of ligands, and these loops of Jag1 are proposed to be involved in phospholipid binding (strands 1-2, 3-4, 5-6) Sequence variation in loop sequences from different ligands has been suggested to optimize the mammalian Jagged ligands for binding to the specific phospholipid composition in the species in question. This specificity was recently investigated (Suckling et al, 2017) suggesting that each ligand has a different preference for lipid binding, thus providing another method of tissue-specific Notch regulation. This

is supported by the observation that some Jagged1 variants resulting in the disease extrahepatic biliary atresia (EHBA) have been shown to only impact lipid binding, and not Notch binding.

1.5.4 Mechanism of canonical Notch signalling

Initial studies in *Drosophila* S2 cells showed that ligand-expressing cells and receptor-expressing cells clump together with a dependence on Ca^{2+}, indicating that signalling involves the interaction of the extracellular regions of the transmembrane ligand and receptor (Fehon et al, 1990). Further studies showed that EGF11-12 in the Notch receptor were essential for binding to ligand (Rebay et al, 1991). To characterise the details of the interface between ligand and receptor, flow cytometry assays were used with hN1 constructs containing the LBR (EGF11-13) expressed in bacteria and refolded *in vitro*, interacting with cells expressing ligand. Alanine substitutions of conserved residues in the ligand binding region were used to probe which residues are essential for binding. These experiments identified L468 and I477 in EGF12 of hN1 as residues essential in the interaction with Jag1 (Whiteman et al, 2013). Although these experiments were crucial in identifying the details of the LBR, these constructs were expressed in bacteria, and not glycosylated and therefore residues important in stabilising the glycosylations could not be identified. For example, the conserved M479 which has a role in supporting the O-Fuc modification from T466, was not be observed to decrease binding in these initial assays. The role of modification in the interaction was elucidated by conducting in vitro glycosylation of these bacterially expressed hN1 constructs to modify potential glycosylation sites sequentially (Taylor et al, 2014).

There are two co-crystal structures of ligands interacting with the LBR of a Notch receptor, one with Dll4 and the other with Jag1 (Luca et al, 2015 , Luca et al 2017). These structures support the role of the previously identified residues in EGF11 and EGF12 (the secondary and primary ligand binding sites respectively) in the interface between ligand and receptor (Figures 1.11, 1.12). The longer ligand construct in the Jag1 co-crystal structure shows an extended antiparallel interaction, extending towards the N-terminus of the receptor involving EGF8-12 with the K_d calculated from SPR as 0.81 μM for EGF8-12 and 5.4 μM for EGF11-12 (Luca et al, 2017) demonstrating the importance of EGF8-10 in the interface. The structure suggests that an O-Fuc modification in EGF8 directly contacts side chains in Jag1 EGF3. This rationalises previously observed mutations, for example the interface of Val324 in EGF8 (equivalent to the site of the *jigsaw* missense mutation V361M in *D. melanogaste*r with Serrate specific effects (Yamamoto, 2012)) with Jag1 EGF3. This extended interface has variable roles for alternative ligands as data from SPR suggest that EGF8-10 does not contribute to binding to Dll4 (Luca et al, 2017). The recent structure of Notch/Jag1 complex proposes a catch-bond mechanism, in which force is required to trigger the formation of a ligand-active state through a motion at the C2-DSL hinge. Luca et al, 2017 go on to propose that the affinity enhanced Jag1 variant used to capture the complex is locked in this flexed 'active' state. Recent evidence from fluorescence co-culture assays suggests different ligands can each activate specific downstream genes by modulating the amplitude and pulse-frequency of signalling (Nandagopal et al, 2018).

1.6 Non-canonical modulators of Notch signalling

Non-canonical regulators of Notch signalling include non-DSL ligands which bind Notch to influence signalling efficiency. Examples of these include Dlk-1 and DNER which are membrane proteins similar to canonical ligands, but lack a DSL domain (Eiraku et al, 2005, Komotsu et al, 2008, D'Souza et al, 2010). These proteins can act to compete with ligand for binding to receptor or complement the effect of a ligand by providing non-DSL features of ligand to enhance signalling. Membrane tethered non-canonical ligands such as Contactin1 and Contactin6 bind receptor in regions other than the LBR and initiate signalling through co-effector proteins such as Deltex, an intracellular protein in the signal-receiving cell (Hu et al, 2003, Matsuno et al, 1998). There are also extracellular proteins which can activate Notch signalling, for example the extracellular matrix constituent Thrombospondin2 which has been shown to enhance the interaction between Notch3 and Jag1 (Meng et al, 2009).

1.7 Medical relevance of Notch signalling

The fundamental and universal role of Notch in metazoan development means that defects in this pathway manifest in many diseases and developmental defects. Although the characteristics of the Notch signalling pathway had been identified over a century ago, the role of the pathway in disease in humans was only recently first recognised in 1991 relating mutations in *NOTCH1* to T-ALL (T-cell acute lymphoblastic leukaemia) (Ellisen et al, 1991). This is likely because of the diverse and superficially unrelated phenotypes resulting from changes to the ways Notch impacts development. Since gene

sequencing technologies have become more available, relationships between mutations in the Notch receptor genes or ligand genes and disease have become clear, with some of these giving useful insight into the functional regions of these proteins and which ligands interact with which receptors in particular contexts.

1.7.1 Alagille syndrome

Alagille syndrome (ALGS) is a developmental disorder associated with the presentation of a variety of phenotypes in the heart, liver, vertebrae, eyes, face, and bile ducts (Alagille et al, 1987). This disease is primarily caused by mutations in the gene encoding ligand Jag1 and distributed throughout the entire gene, which can be missense, non-sense, splice site alterations, or deletions (Crosnier et al, 1999) thought to cause disease through a haploinsufficiency mechanism (Morrissette et al, 2001, Crosnier et al, 1999). More recently, mutations in *NOTCH2* have also been found to be associated with the disease, again across the length of the gene and of a variety of types. (Kamath et al, 2012). These findings suggest that the interaction between Jag1 and Notch2 is important for the correct development in the tissues in which ALGS manifests. The related disease EHBA, caused by mutations in the *JAG1* gene impacts only the extrahepatic bile duct, and is caused by mutations that prevent formation of the ternary complex of Notch-ligand-membrane bilayer (Suckling et al, 2017).

1.7.2 CADASIL syndrome

CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy) is a disease associated anomalies in the brain and increased chance of stroke and dementia caused by mutations in *NOTCH3* (Joutel et al, 1996). The mutations leading to this disease are largely structural changes affecting the extracellular domain including conserved Cys substitutions leading to misfolding and aggregation of the receptor (Joutel et al, 1997). The accepted hypothesis for the disease mechanism is that the accumulation of Notch3 leads to sequestering and accumulation of the extracellular matrix protease regulators TIMP3 and VTN. The accumulation of these regulators leads to a fragile blood vessel wall and therefore an increase chance of haemorrhage and stroke indicative of CADASIL.

1.7.3 Notch in cancer

The role of the Notch pathway as a regulator of growth suggests that cancer can be caused by defects in the pathway, and alterations in this pathway can be pre-requisites for the transformation of a healthy cell. Many different cancers are associated with defects in the signalling pathway acting in both a proto-oncogenic and tumour-suppressor context (Ranganathan et al, 2011). An important example is the involvement of Notch directing epithelial-mesenchymal transitions (EMT), which allows cancer cells to lose polarity and become mobile (Timmerman et al, 2004). Increasing expression of Jag1 has been shown to lead to EMT-like changes in kidney epithelia supporting the importance of Notch in this transition (Nyhan et al, 2010). Notch1 has been shown to

have an oncogenic role in T-ALL by enabling the cell to survive hypoxic conditions and increase invasiveness by promoting the expression of matrix-proteases to enable invasion of surrounding tissues. Knockdown of *NOTCH1* has been observed to inhibit this effect (Zou et al, 2013). There are multiple pharmaceutical approaches which aim to inhibit Notch signalling in cancer by targeting γ-secretase, decreasing expression of *NOTCH1* using siRNA, and antibodies targeting ligands to prevent interaction with the Notch receptor (Takebe et al, 2014).

Tumour-suppressor roles of Notch include the action of the pathway in supressing the effects of Sirtuin-1, a deacetylase which can act to deactivate p53 in Ewing sarcoma cells (Ban et al, 2014). Further supporting the tumour suppressor activity of Notch is the observation that expression of dominant negative Notch transcription co-effector MAML in the epidermis of a murine system leads to spontaneous formation of squamous cell carcinomas (Proweller et al, 2006).

1.7.4 Other diseases

A number of other diseases are associated with defects in the Notch signalling pathway, one of these is Spondylocostal Dystosis. This disease is associated with abnormal vertebrae segmentation and is associated with mutations in the genes encoding ligand Dll3 and Lunatic Fringe, and Notch target genes, which code for transcription factors (Dunwoodie, 2009, Turnpenny et al, 2007). Hajdu Cheney syndrome manifests in the kidneys, heart, and bone defects, all associated with mutations in the last exon of *NOTCH2* (Gray et al, 2012, Simpson et al, 2011). Adams Oliver Syndrome is a rare genetic

disorder characterised by scalp and limb defects can be associated with substitutions in *NOTCH1*, as well as other genes involved the Notch signalling pathway (Stittrich et al, 2014).

1.8 Aims

Reviewing the current knowledge of Notch signalling, despite the fundamental and essential role of this signalling pathway in development and tissue maintenance, it stands out how little we understand about the structure of the Notch receptor. The overall shape of the receptor, and in particular the shape and role of the Ax region remain elusive, despite Ax mutations suggesting an important role for this region in regulating Notch signalling. A likely contributor to this lack of understanding, is the challenge in requiring both structural and cellular techniques to investigate the system. Marrying-up crystallographic or NMR-derived structures with the impact of these structures in receptor signalling is not trivial. For example, the Abruptex region has been characterised from a phenotypic standpoint, and the impacts of Abruptex mutations on development has led to the proposal of the role for Abruptex in negatively regulating signalling. Despite this, we do not understand the mechanism for this de-regulation, or how the Abruptex mutations might cause this proposed alteration in signalling. In this study, investigations will be undertaken to bridge this gap, by investigating the structure of the Abruptex region and the role of this region in cellular signalling for the hN1 receptor. This insight into the Ax region will provide information into the overall shape of the receptor, the origin of the Ax mutants observed in *D. melanogaster*, and the conservation of these roles in the mammalian receptor hN1.

The research strategy undertaken will be three-fold. Firstly, structural analysis on purified protein fragments will be undertaken to understand the shape of the Ax region. Secondly interaction between the Ax and LBR will be probed with plate-based assays and flow cytometry assays, using purified protein fragments with the aim of identifying the interaction faces involved. Finally, the impact of homologous *D. melanogaster* Ax mutations on activation of the mammalian Notch receptor will be investigated in a cell activation assay involving full length N1 receptor.

In summary, the aims of this work are to:

- Investigate the structure of the Ax region to understand the cell surface organisation of the receptor
- Confirm and characterise the interaction between the ligand binding region and the Ax region of the Notch receptor
- Investigate the impact of Ax mutations identified in *D. melanogaster* on activation by hN1 and further investigate the role of the Ax in signalling

2.

Materials and Methods

2.1. DNA methods

2.1.1. DNA preparation

DNA preparations were carried out using Qiagen Plasmid Spin Miniprep, or Midiprep kits following the procedures found at https://www.qiagen.com/gb/products. All optional washes included in these protocols were carried out. Miniprep DNA was eluted in 30-50 µL buffer EB Elution Buffer (Qiagen). Midiprep DNA was dissolved in 250 µL Elution Buffer (Qiagen).

2.1.2. DNA gel electrophoresis

TBE/Agarose gels were prepared with 0.7-1.5 % agarose (depending on length of DNA analyzed) in TBE buffer (45 mM Tris-borate, 1 mM EDTA) with 1:10,000 Ethidium Bromide for DNA visualization. Samples were diluted 6x Purple Gel Loading Dye (NEB) and run alongside a 1 kbp or 100 bp DNA ladder (NEB) for calibration, using the most suitable marker range for the DNA construct being visualized. Gels were run at 75 V for 1 hr and then visualized on a UV-light box-camera system.

2.1.3. DNA extraction from agarose TBE gel

To elute DNA, the band of interest was cut out with a razor blade over a UV source and DNA eluted using a Qiagen QIAquick Gel Extraction kit following the protocol found at https://www.qiagen.com/gb/resources/resourcedetail?id=f4ba2d24-8218-452c-ad6f-1b6f43194425&lang=en. All optional washes included before DNA elution in 30 µL buffer Elution buffer (Qiagen).

2.1.4. DNA concentration determination

DNA concentration was determined by detection A_{260nm} using a Thermo Scientific NanoDrop Lite Spectrophotometer. Beer's law was used to calculate the DNA concentration,

A_{260nm}/A_{280nm} was also collected to assay the purity of DNA samples with a ratio of above 1.80 an indication of good DNA purity.

2.1.5. DNA Sequencing

DNA Sequencing was carried out by Source Biosciences PLC, Oxford. Samples were submitted as indicated on the Source Bioscience PLC website, and results interpreted using the software BioEdit.

2.1.6. Polymerase chain reactions

PCR reactions were carried out using a KOD Hot-Start Polymerase kit (Merck Millipore) following the manufacturer's suggested protocol, adjusting elongation time for length of DNA to be replicated. All PCR reactions were carried out in 50 µL volumes. The reactions contained 5 µL 10X Buffer for KOD Hot Start DNA Polymerase, 3 µL 25 mM $MgSO_4$, 5 µL dNTPs (stock 2 mM each), 1.5 µL of 10 µM forward primer, 1.5 µL of 10 µM reverse primer, 10 ng template DNA, 1 µL KOD Hot Start DNA Polymerase (stock 1U/µL), PCR grade water added to make reaction up to 50 µL.

The thermocycling procedure involved an activation step at 95 °C for 2 mins, followed by 30 repeats of: denaturation at 95 °C for 20 sec, annealing at 55 °C for 10 sec, extension at 70 °C for an appropriate time proportional to the length of DNA to be amplified according to the

product Protocol found at http://www.merckmillipore.com/GB/en/product/KOD-Hot-Start-DNA-Polymerase,EMD_BIO-71086#documentation. Reactions were then immediately used or stored at 4 °C for a short period before use.

2.1.7. Site directed mutagenesis

Site directed mutagenesis was carried out using either overlap extension mutagenesis, Agilent QuikChange XL kit mutagenesis, or Q5 kit mutagenesis. Primers used are outlined in Section 2.1.10.

2.1.7.1. Substitutions by overlap extension mutagenesis

Overlap extension mutagenesis was used to introduce hN1 11-13 substitutions. This method involved two stages of PCR. The first consisted of two reactions 1A and 1B. 1A) utilized an upstream primer and reverse mutagenesis primer with a mismatch corresponding to the substitution desired, and 1B) involved a downstream primer and forward mutagenesis primer. The products of these reactions were visualized on an agarose gel before undergoing gel extraction and then were subsequently mixed for the stage two PCR reaction with only the downstream and upstream primers. These downstream and upstream primers contained conveniently positioned and unique restriction digest consensus sites so that this mutagenized product could be ligated into a plasmid. The ligated construct was then sequenced from before the 5`-ligation site through the entire mutagenized fragment until

after the 3` ligation site to ensure mutagenesis had taken place as expected.

2.1.7.2. Substitutions by QuikChange XL site directed mutagenesis

The QuikChange XL Site-Directed Mutagenesis Kit (Agilent) was used to introduce substitutions following the manufacturer's instructions, each in 50 µL volumes. The reaction consisted of 5 µL 10X reaction buffer (Agilent), 20 ng template DNA, 1 µL of 10 µM forward primer, 1 µL of 10 µM reverse primer, 1 µL dNTP mix (Agilent), 3 µL Quiksolution (Agilent), PCR grade water added to make reaction up to 50 µL, 1 µL pfu Ultra HF DNA polymerase (Turbo).

The thermocycling procedure consisted of 1 min at 95 °C, followed by 18 repeats of: denaturation at 95 °C for 50 sec, annealing at 60 °C for 50 sec, extension at 68 °C for 4 min. Reactions were either immediately used, or stored at 4 °C for a short period before use. DNA sequencing was carried out to ensure mutagenesis had taken place as expected.

2.1.7.3. Substitutions using Q5 site directed mutagenesis

Substitutions were carried out using a Q5 site-directed mutagenesis kit (NEB) following the manufacturers instructions. DNA sequencing was carried out to ensure mutagenesis had taken place as expected.

2.1.7.4. Deletions

Deletions were carried out using a variation of the overlap extension mutagenesis method, introducing a 40 bp complimentary region and no mismatch in the mutagenesis primer to ensure products of the first PCR reaction steps interact in the second PCR step. DNA sequencing was carried out to ensure mutagenesis had taken place as expected.

2.1.8. Restriction digestion

Restriction digests were carried out using enzymes supplied by NEB. Reactions were carried out in 50 µL volumes under conditions recommended by NEB. The reactions contained 5 µL CutSmart buffer (NEB), PCR grade water added to make reaction 50 µL, 500 ng – 1 µg of DNA, 1 µL of each restriction enzyme (up to a maximum of two for a double digestion). Reactions were then incubated at 37 °C for 1 hr, followed by 65 °C for 20 mins. The entire reactions were then run on agarose TBE gels for visualization and DNA extraction if required.

2.1.9. DNA ligation

Ligation was carried out on complementary restriction digested products using the Rapid DNA Ligation Kit (Roche Diagnostics) following the manufacturer's recommended protocol.

2.1.10. Primer sequences

2.1.10.1. S2 hN1 11-13 directional mutants

hN1 11-13 A426R Substitution
Forward: 5` GCCAATCCATGCGAGCATCGCGGCAAGTGCATCAACACC 3`
Reverse: 5` GGTGTTGATGCACTTGCCGCGATGCTCGCATGGATTGGC 3`

hN1 11-13 K428E Substitution
Forward: 5` CCATGCGAGCATGCCGGCGAGTGCATCAACACCCTGGGC 3`
Reverse: 5` GCCCAGGGTGTTGATGCACTCGCCGGCATGCTCGCATGG 3`

hN1 11-13 E437R Substitution
Forward: 5` AACACCCTGGGCAGCTTCCGCTGCCAGTGCCTGCAGGGA 3`
Reverse: 5` TCCCTGCAGGCACTGGCAGCGGAAGCTGCCCAGGGTGTT 3`

hN1 11-13 R448E Substitution
Forward: 5` CAGGGATATACCGGCCCAGAGTGCGAGATCGATGTGAAC 3`
Reverse: 5` GTTCACATCGATCTCGCACTCTGGGCCGGTATATCCCTG 3`

hN1 11-13 N459R Substitution
Forward: 5` GTGAACGAGTGCGTGTCCCGCCCCTGCCAGAACGATGCC 3`
Reverse: 5` GGCATCGTTCTGGCAGGGGCGGGACACGCACTCGTTCAC 3`

hN1 11-13 Q462R Substitution
Forward: 5` TGCGTGTCCAACCCCTGCCGCAACGATGCCACCTGCCTG 3`
Reverse: 5` CAGGCAGGTGGCATCGTTGCGGCAGGGGTTGGACACGCA 3`

hN1 11-13 L468R Substitution
Forward: 5` CAGAACGATGCCACCTGCCGCGATCAGATCGGCGAGTTC 3`
Reverse: 5` GAACTCGCCGATCTGATCGCGGCAGGTGGCATCGTTCTG 3`

hN1 11-13 V485R Substitution
Forward: 5` ATGCCCGGCTATGAGGGCCGCCACTGCGAAGTGAACACC 3`
Reverse: 5` GGTGTTCACTTCGCAGTGGCGGCCCTCATAGCCGGGCAT 3`

hN1 11-13 Upstream primer
5` CGTCTACGGAGCGACAATTCAAT 3`

hN1 11-13 Downstream primer
5` GGAGTGTGTAAATGGACAATGTTG 3`

2.1.10.2. PR hN1 23-24 directional mutants

hN1 23-24 H917W Substitution
Forward: 5` TGCCGGCCCAACCCGTGTTGGAACGGGGGCTCCTGCACA 3`
Reverse: 5` TGTGCAGGAGCCCCCGTTCCAACACGGGTTGGGCCGGCA 3`

hN1 23-24 D932R Substitution
Forward: 5` ATCAACACGGCCTTCTGCCGCTGCCTGCCCGGCTTCCGG 3`
Reverse: 5` CCGGAAGCCGGGCAGGCAGCGGCAGAAGGCCGTGTTGAT 3`

hN1 23-24 L934R Substitution
Forward: 5` ACGGCCTTCTGCGACTGCCGGCCCGGCTTCCGGGGCACT 3`
Reverse: 5` AGTGCCCCGGAAGCCGGGCCGGCAGTCGCAGAAGGCCGT 3`

hN1 23-24 T940R Substitution
Forward: 5` CTGCCCGGCTTCCGGGGCCGGTTCTGTGAGCTCGACCTG 3`
Reverse: 5` CAGGTCGAGCTCACAGAACCGGCCCCGGAAGCCGGGCAG 3`

pQE-30 Sequencing Primers
Forward: 5` CCCGAAAAGTGCCACCTG 3`
Reverse: 5` GTTCTGAGGTCATTACTGG 3`

2.1.10.3. PR hN1 23-27CSCS

Primers for region amplification

Forward:
5` TAGTAGAGAGATCTATAGAAGGACGATCAGCAGTCGACATCAACGAGTGCG 3`

Reverse:
5` CACTGGCCCCAACTGCCAGCTCGAGTAGTAG 3`

EGF25 CS Substitution

Forward: 5` CACGGACAGCGTGGACAGC 3`
Reverse: 5` CAGTTGGCCCCGTTGCGG 3`

EGF27 CS Substitution

```
Forward: 5` GACGGCAGCGGCTCCTAC 3`
Reverse: 5` CTGACAGGTGCCGCCATGC 3`
```

2.1.10.4. mN1 primers

mN1 Sequencing primers

```
mN1_SeqRev_1 5` ACCCCCATTCTTGCAGTTGTTTCC 3`
mN1_SeqFor_1 5` GGAAACAACTGCAAGAATGGGGGT 3`
mN1_SeqFor_2 5` TTTCCAGGGTGCCACTTGCCACGA 3`
mN1_SeqFor_3 5` GAAATCAACACGGATGAGTGCGCC 3`
mN1_SeqFor_4 5` AGATCAACCTGGATGACTGCGCCA 3`
mN1_SeqFor_5 5` GGCCCTAATTGCCAGACCAACATC 3`
mN1_SeqFor_6 5` GGAGGACATCAATGAATGTGCCAG 3`
mN1_SeqFor_7 5` GCACAGAAGCGAGGCATTGACGTC 3`
mN1_SeqFor_8 5` TGACCAACTCCTACAAGTGTTCCT 3`
```

mN1 mutagenesis primers

mN1 L468A

```
mouseN1_L468A_F 5′ CAGAATGATGCCACTTGCGCCGACCAGATTGGGGAGTTC 3′
mouseN1_L468A_R 5′ GAACTCCCCAATCTGGTCGGCGCAAGTGGCATCATTCTG 3′
```

mN1 D909V

```
mouseN1_D909V_F 5′ AACTGTGAGAGTGACATCGTGGACTGCCGCCCCAACCCG 3′
mouseN1_D909V_R 5′ CGGGTTGGGGCGGCAGTCCACGATGTCACTCTCACAGTT 3′
```

mN1 N947I

```
mouseN1_N947I_F 5′ TTCTGTGAGGAGGACATCATTGAATGTGCCAGCAATCCC 3′
mouseN1_N947I_R 5′ GGGATTGCTGGCACATTCAATGATGTCCTCCTCACAGAA 3′
```

mN1 Δ24-25 deletion

D2425_For:
5′
AACACACCTGACTGTACTGAGAGCTCCTGCTTCAATGGTGGTACCTGTGTGGATGGTATC
3`

D2425_Rev:
5′
CACCATTGAAGCAGGAGCTCTCAGTACAGTCAGGTGTGTTACTCTCACAGTTGCGACCTG
3`

mN1 Δ24-29 deletion

D2429_For:
5′
GAGGTGGACGAGTGCTCACCTAACCCCTGCCAGAATGGAGCTACCTGCACTGACTATCTC
3`
D2429_Rev:
5′
CTCCATTCTGGCAGGGGTTAGGTGAGCACTCGTCCACCTCACTCTCACAGTTGCGACCTG
3`

mN1 end primers

For_end_Sbf1: 5′ ACACCTGCAGGCAAGATGTTAATGAGTGCAGCCAGAACCC 3`
Rev_end_BamH1: 5′ GAAGGATCCCAGGCATAGGCAGGTGGGACTACGTGGG 3`

2.2. RNA methods

2.2.1. RNA extraction

RNA was extracted from mammalian cells using a Qiagen QIAshredder column and RNeasy mini kit following the kit protocol found at https://www.qiagen.com/gb/shop//sample-

technologies/rna/rneasy-mini-kit/#resources including an on-column, DNAse1 digestion step. Extractions were carried out with 3 million cells and eluted in 50 µL RNAse free water.

2.2.2. RT-PCR

RT-PCRs were carried out immediately following RNA extraction using a Qiagen Omniscript RT kit following the manufacturer's instructions with total reaction volumes of 20 µL. Reactions were prepared on ice and consisted of: 2 µL 10X buffer RT, RNAse free water added to make reaction 20 µL, 2 µL dNTP mix, 1 µL 10 µM complementary reverse primer, 1 µL RNAse inhibitor (stock 10U/µL), 1 µL RTase, 2 µg template RNA. The reaction was then incubated at 37 °C for 1 hr, and then 2.5 µL of this reaction was immediately used for PCR (total volume 25 µL). Control PCRs were concurrently carried out without the RT-PCR step to show that no contaminating DNA was present. The products from these reactions were visualized on an agarose TBE gel.

2.2.2.1. RT-PCR primers

Primers to amplify the mN1 LBR

```
Forward: 5` CATTGATGACTGTGCCAGTGC 3`
Reverse: 5` CCGTTCTTGCATGGTGTG 3`
```

Primers to amplify mN1 Ax

```
Forward: 5` GTGCCCCTGGGTGGAGTGGAAC 3`
Reverse: 5` ATCCACACAGGTACCACCATTGAAGCAG 3`
```

2.3. Cell culture

2.3.1. Insect cells

2.3.1.1. S2 cells growth conditions and transfection protocol

S2 cells (Expres2ion Biotechnologies, Denmark) were grown and transfected as previously described in Suckling et al, 2017.

2.3.2. Mammalian cells

2.3.2.1. 293T cells growth conditions

293T cells were grown as previously described in Whiteman et al, 2013.

2.3.2.2. HEK-RBP cells growth conditions and transfection protocol

HEK-RBP cells grown under the same conditions as 293T cells, including 0.2 µg/ml puromycin. The cells were transfected using Mirus TransIT-LT1 Transfection Reagent following the recommended protocol found at https://www.mirusbio.com/products/transfection/transit-lt1-transfection-reagent. Transfections were carried out using 75 % confluent wells in a 6-well plate, using a total of 100 ng of DNA per transfection (pOG44 : pcDNA5/FRT mN1 at a

ratio of 9:1). 24 hours following transfection, cells were transferred to 15 cm dishes and again grown for 24 hrs. Subsequently, medium containing selection agents was used for growth (Dulbecco's modified Eagle's medium (Gibco), supplemented with 10% bovine calf serum (Gibco) + 0.2 μg/ml puromycin + 100 μg/ml hygromycin (after transfection)). Following establishment of resistant colonies, these were transferred to a 6-well plate and expanded for further investigation and grown in the presence of both 0.2 μg/ml puromycin + 100 μg/ml hygromycin. Site-directed mutagenesis to produce mN1 variants were carried out in a pUC19 plasmid.

2.3.3. *E.coli* growth and transformations

E.coli were grown and underwent transformations as described in Whiteman et al, 2014.

2.4. Protein methods

2.4.1. General procedures

2.4.1.1. Protein concentration determination

Protein concentration determination by absorbance at 280 nm was carried out using a Shimadzu Mini 1240 UV-VIS spectrophotometer and theoretical extinction coefficients

calculated using the peptide sequence in question. Theoretical extinction coefficients were calculated using ExPASy ProParam tool (found at https://web.expasy.org/protparam/).

2.4.1.2. SDS-PAGE

SDS-PAGE was carried out as described in Whiteman et al, 2014. When running under reducing conditions, sample buffer included 5 % β-mercaptoethanol and was heated to 100°C for 10 mins as described. For non-reducing samples, sample buffer did not include reducing agent. Samples were run at 180V for 60 mins, and either used for western blotting, or stained with Coomassie blue R-250.

2.4.1.3. Western blotting

Western blotting was carried out following SDS-PAGE, transferring to nitrocellulose membrane using the Bio-Rad Mini Trans-Blot cell system following the manufacturer's suggested protocol. After transfer, membrane was blocked in 5 % BSA in TBST (50 mM Tris pH 7.5, 150 mM NaCl, 0.1% Tween 20) for 1 hr at 4°C, and then washed three times in TBST. The membrane was then exposed to 1:20,000 Streptavidin-POD conjugate (Sigma Aldrich) in TBST for 1 hr. After washing three times in TBST, excess buffer was removed, and ECL Plus Western blotting detection reagent (Amersham Biosciences) was added to the nitrocellulose, before exposing to chemiluminescent film in a dark room. The film was developed using a compact X4 automated developer (Xenograph Imaging).

2.4.1.4. ESI-MS

Mass of purified protein constructs underwent ESI-MS which was carried out under the ESI-MS molecular biophysics suite service by Dr David Staunton, Department of Biochemistry, University of Oxford. A Micromass system was used to measure mass of a lyophilized sample dissolved (approx. 10 µg) in 50% acetonitrile, 0,1% formic acid. Data was processed using MassLynx software.

2.4.1.5. DTNB assay

DTNB (Ellman's reagent) assay was carried out to determine the number of free reduced thiols in a purified protein construct. Reactions with a range of concentrations of purified S2 hN1 Ax constructs, calibration solutions of cysteine were added to 2mM DTNB in 50 mM Tris pH8.0, 6M Guanidinium Hydrochloride. Products of the reaction were detected at an absorbance of 412 nm using an extinction coefficient of 14,150 $M^{-1}cm^{-1}$ (Shimadzu Mini 1240 UV-VIS spectrophotometer) using calibration cysteine solutions to enable determination of concentration of free thiols.

2.4.1.6. Biotinylation

Biotinylation was carried out using an Avidity BirA-500 kit following the manufacturer's instructions, incubating at room temperature for 16 hrs. The reaction was then acidified

(pH2) and purified by reverse-phase chromatography by HPLC after washing through in hydrophilic conditions for 45 mins (Varian Dynamax Reverse Phase Microsorb 300Å, C8 , 250 × 10 mm column), the collected peaks were then freeze dried. The freeze-dried peaks were then run on 16 % reducing SDS-PAGE and western blotted with Streptavidin-POD conjugate (Sigma Aldrich) to detect biotinylation. The highest intensity peak (representing fully biotinylated protein) was then used for future experiments.

2.4.1.7. Factor Xa cleavage to remove His_6 tag

Protein to be cleaved was diluted to ~3 mg/ml in a buffer containing 50 mM Tris–HCl pH 7.5, 0.1 M NaCl, and 1 mM $CaCl_2$. Factor Xa (Novagen) was added at a concentration of 1 U enzyme/1 mg protein and then incubated at 37 °C for 16 h. The reaction was then purified by reverse-phase chromatography by HPLC after washing through in hydrophilic conditions for 45 mins (Varian Dynamax Reverse Phase Microsorb 300Å, C8 , 250 × 10 mm column), the collected peak was then freeze dried.

2.4.1.8. EDTA treatment

In order to remove residual Ca^{2+} from a protein sample, EDTA was added to 15 mM, and then sample purified using reverse phase chromatography to separate the EDTA from the protein (Varian Dynamax Reverse Phase Microsorb 300Å, C8, 250 × 10 mm column). The peaks corresponding to protein were collected and then freeze dried.

2.4.2. NMR

2.4.2.1. Ca^{2+} titrations

Experiments were carried out in a Shigemi tube in matrix buffer (5 mM Tris, 150 mM NaCl, 5 % D_2O, pH7.5) at 25°C on a 500 MHz Bruker spectrometer at the Oxford University, Department of Biochemistry solution-state NMR facility under the supervision of Prof C Redfield, University of Oxford. Ca^{2+}-free samples were prepared by EDTA treatment (Section 2.4.1.8.) then an initial ^{1}H-^{15}N Heteronuclear Single Quantum Coherence experiment (HSQC) recorded before stepwise addition of Ca^{2+} at interval concentrations using a $CaCl_2$ stock solution diluted from a 1 M commercial stock solution (Fluka Analytical). After each addition, the pH was adjusted to 7.5±0.05 before another HSQC recorded.

An accurate estimation of the k_d (dissociation constant) was calculated by fitting the equation:

$$(\Delta/\Delta_0) = ([Ca]_f)/(k_d + [Ca]_f)$$

Where:

Δ = peak shift from [Ca] = 0 to a defined [Ca]

Δ_0 = peak shift from [Ca] = 0 to [Ca] = ∞

$[Ca]_f$ = [Ca] free in solution

Plotting:

$x = [Ca]_f$

$y = (\Delta/\Delta_0)$

Fitting to determine kd was performed using GraphPad Prism for Windows, GraphPad Software, San Diego California USA, www.graphpad.com

2.4.2.2. Protein titrations

An identical protocol was followed as for Ca^{2+} titrations described in section 2.4.2.1. with 1.5 mM Ca^{2+} included in matrix buffer, and S2 hN1 11-13 titrated into the sample of PR hN1 23-24.

2.4.2.3. Checking S2 protein fold

1D Experiments were carried out in matrix buffer (5 mM Tris, 150 mM NaCl, 5 % D_2O, pH7.5) at 25 °C on a 950 MHz Bruker spectrometer at the Oxford University, Department of Biochemistry solution-state NMR facility under the supervision of C Redfield. Expected shift in aromatic peaks were monitored on addition of 15 mM Ca^{2+} to a Ca^{2+}-free sample.

2.4.3. S2 expression and purification

Stable S2 cell lines were established as previously described (de Jongh et al, 2013) and transfected with Notch constructs in pEXS2-2 plasmids. 1 L of culture centrifuged at 1000 xg, 30 mins and loaded onto a Roche 1mL cOmplete His tag purification column (Roche) equilibrated in 50 mM Tris pH9.0, 200 mM NaCl, 2 mM $CaCl_2$. The column was washed with 2 x 30 CV (column volumes) of the same buffer before elution with 10 CV of the same buffer + 250 mM imidazole. The 10 CV was then filtered through a 0.45 µm syringe filter loaded onto an GE HiLoad 26/600 Superdex 75 or 200 pg gel filtration column (GE Life Sciences) equilibrated in 20 mM Tris pH7.5, 200 mM NaCl, 1 mM $CaCl_2$ before isocratic elution with this buffer. Fractions were then run on a 16% SDSPAGE gel and pure fractions pooled and concentrated to ~1 mL on a pre-washed Sartorius Vivaspin 20 MWCO 10 kDa spin concentrator. Aliquots of 50 µL were then frozen in liquid N_2 and stored at -80 °C until use.

2.4.4. Prokaryotic refold expression and purification

Proteins were expressed from a pQE-30 plasmid and purified following the protocol in Whiteman et al, 2014. ^{15}N-enriched proteins for 2D NMR spectrometry were expressed and purified as described in Weisshuhn et al, 2016.

2.5. Small angle X-ray scattering

2.5.1. Data collection

Data Collection was carried out at the Diamond Light Source with assistance from Dr E Lowe, University of Oxford, at Beamline B21 using a SEC-SAXS method. 100 µL samples were prepared (100 µL, 2.5 mg/mL, 5mM Tris pH7.5, 15 mM $CaCl_2$), including dialysis in 2 L of buffer overnight at 4°C. Samples were subsequently frozen at -80 °C for transport to the beamline and thawed immediately prior to use. SAXS samples (50 µL) were autoloaded onto a Shodex 402.5 analytical SEC column. The eluent from this column then passed through the X-ray beam and scatter pattern recorded. Data were collected up to a scattering vector (s) of 0.37 $Å^{-1}$.

Scatter resulting from buffer was subtracted from the peak corresponding to protein scatter using the program SCATTER 3.0 (Rambo et al) and the ATSAS software package was used for further analysis (Franke et al, 2017).

2.5.2. Model determination

Guinier analysis and ab initio modelling and structure fitting was carried out using the programs DAMMIF, DAMAVER and CRYSOL in the ATSAS software package (Franke et al,

2017). Modelling was performed at default settings using 10 repetitions followed by DAMAVER. CRYSOL was subsequently used to confirm the data fits the structure model.

2.6. Binding assays

2.6.1. Plate binding assay

His-tagged protein at 5 µg/ml in HBS (20 mM HEPES pH7.5, 200 mM NaCl) with 1 mM $CaCl_2$ was added to a clear bottom Ni^{2+} 96-well plate (Pierce) (100 µL/well) and left a 4°C overnight for surface adsorption. Plates were washed twice with HBS + 1 mM $CaCl_2$ and blocked for 1 hour in 4 % BSA in HBS + 1 mM $CaCl_2$. hN1 11-13 pre-clusters were prepared by addition of neutravidin horseradish peroxidase (HRP) conjugate (Pierce) to a biotinylated 11-13 variant at a ratio of 1:5 (eg. 20 nM neutravidin + 100 nM 11-13) to ensure saturation of the quadrivalent-neutravidin HRP in HBS + 5 mM $CaCl_2$ and incubated for one hour by mixing at room temperature. The plate was washed once in HBS with 5 mM $CaCl_2$ and hN1 11-13 preclusters added (50 µL/well) and incubated 1.5 hours at RT. The wells were then washed three times with HBS + 5 mM $CaCl_2$ + 0.05% Tween-20. 100 µL substrate solution (2 mg ABTS + 500 µL H_2O + 5.5 mL 0.1 M citric acid pH4, 5 µL H_2O_2 added just before use) was added to each well and incubated at room temperature for typically three minutes before quenching with 100 µL 1 % SDS. The A_{415nM} was recorded using Pherastar plate reader (provided by Prof L Cox, University of Oxford). Control experiments with hN1 EGF20-23 (provided by Mr J Hill, University of Oxford) were carried out with every experiment with Ax (EGF23-24) down on

the plate so that non-specific binding could be subtracted from the total binding. When combining data, data were scaled to wildtype to take into account day-to-day differences in temperature and other factors affecting the rate of HRP reaction. hJag1 N-EGF3 purified from the S2 system (S2 hJag1 N-EGF3) used as a control in these assays was produced by Dr B Korona, University of Oxford as used in Suckling et al, 2017.

2.6.2. Flow cytometry assay

Flow cytometry assays were carried out as described in Whiteman et al, 2013. 293T cells were grown to 90 % confluence in in 75 cm^2 flasks. Cells were then washed twice in 2 ml cold Hank's Balanced Saline Solution (HBSS) + 10 % fetal calf serum (FCS) before re-suspension in 1 ml, cold HBSS/FCS and incubated in the dark on ice for 60 mins.

Biotinylated hN1 23-24 or control construct were attached onto fluorescent purple avidin-coated (0.7-0.9 µm) beads. Beads were then washed twice in HBSS + 1 % BSA and following re-suspension in 50 µL per reaction in HBSS + BSA, 1 µg biotinylated protein was added. Beads were then incubate for 16 h in the dark at 4 °C before washing at mixing with 50 µL hN1-expressing HEK-293T cells. Probing interaction with antibodies was carried out with anti-LBR antibody (Ab3) added at a 1:300 dilution to cells before mixing with beads, or anti-Ax antibody (N1-162B) added at a 1:60 dilution to bead clusters before mixing with cells. Assays were carried out in pre-chilled FACs tubes (BD) after diluting with 0.5 mL cold HBSS. A FACSCalibur machine was use (Beckton Dickinson) with a flow rate < 2,500 cells/min.

Fluorescence measured with no gate in the FL1-H channel at 530 nm and FL3-H channel at >670 nm, with excitation at 488 nm.

2.7. Activation assays

2.7.1. Ligand-independent activation assay

HEK-RBP cells were provided from the Kopan lab, University of Cincinnati College of Medicine. Cells to be assayed were split 1:3 into DMEM + 10% FCS. After two days, 20,000 cells to be assayed added to each well of a 96-well Costar black clear bottom plate (DMEM +10 % FCS). 48 hrs later, wells were washed with HBS, and 100 µL Hanks Buffered Saline Solution + 150 µg/ml D-luciferin + 10 mM EGTA was added to each well pipetting to re-suspend cells, control wells had same solution without EGTA. Luciferase activity was monitored using a plate reader at 40 mins after addition of EGTA (Berthold Technologies TriStar LB 941, provided by Prof F Barr, University of Oxford) for monitoring luminescence for 30s per well.

2.7.2. Ligand-dependent activation assay

Cells to be assayed were split 1:3 into DMEM + 10% FCS. The following day, a 96-well Costar black clear bottom were coated with 50 µL of 10 µg/ml hJag1 (provided by Dr B Korona and Ms Y Meng, University of Oxford) in HBS (20 mM Hepes pH7.4, 150 mM NaCl, 0.9 mM Ca^{2+})

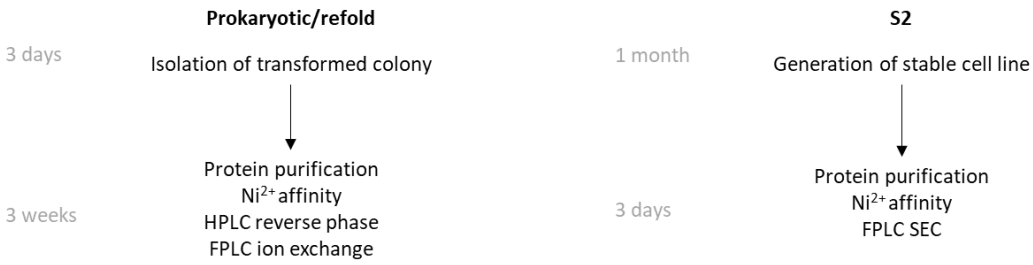
and incubated overnight at 4 °C, control wells were incubated without hJag1. The following day, wells were washed with HBS, and 20,000 cells to be assayed were added to each well (DMEM +10 % FCS). The following day, wells were washed with HBS, and 100 µL substrate solution added by pipetting to re-suspend cells (150 µg/ml D-luciferin in HBS + 1 mg/ml Mg^{2+}, 0.1 % glucose). Luciferase activity was monitored using a plate reader (Berthold Technologies TriStar LB 941, provided by Prof F Barr, University of Oxford) for monitoring luminescence for 30s per well.

3.

Protein Purification and Characterisation

Protein expression of a series of Ax and LBR constructs was carried out to prepare purified protein for molecular and cellular studies. In order to understand the nature of Ca^{2+}-binding at the N-terminal of the Ax region and for binding assays, a short Ax construct (EGF23-24) was expressed enriched with ^{15}N for 2D NMR. The aim of this was to investigate the affinity of the predicted Ca^{2+}-binding site in EGF24 and give insight into the shape of this region. Longer Ax and LBR constructs were purified for investigation by small angle X-ray scattering. Expression was carried out in a prokaryotic system followed by *in vitro* refolding (prokaryotic/refold, PR) or in a Drosophila Schneider 2 (S2) eukaryotic system. A summary comparing the systems as well as the advantages and disadvantages of each method are described in Figure 3.1.

A



B

	Advantages	Disadvantages
Prokaryotic/refold	Homogenous product Can routinely ^{15}N label for 2D NMR	No-glycosylation Disulphides need to be reformed Can only purify up to 5 EGF domains
S2	N-/O-linked glycosylation Can purify very long constructs	Heterogenous glycosylation prevents use of ESI-MS No assessment of fold directly Cannot easily ^{15}N label for 2D NMR

Figure 3.1 Summary, advantages and disadvantages of prokaryotic/refold and S2 protein production procedures: (A) Comparison of the expression and purification procedures. The prokaryotic/refold system involves a short expression period, followed by a long purification involving affinity chromatography, reverse phase high pressure liquid chromatography (HPLC), and ion exchange fast protein liquid chromatography (FPLC). The S2 system involves a long period to generate stable cell lines, followed by a purification involving affinity chromatography and size exclusion fast protein liquid chromatography (FPLC). (B) Comparison of the advantages and disadvantages of the systems.

3.1. Bioinformatics analysis of hN1 and other Notch receptors

Bioinformatics analysis of the 36 extracellular EGF/cbEGF domains was carried out with hN1, hN2, and dN to better inform design of constructs for structural and binding assays (Figure 3.2). Schematics of each receptor are shown in Chapter 1 – Introduction (Figure 1.2).

A) Human Notch 1

EGF1	RGPR	C	SQPGET	C	LNGGK	C	EAANGTEA	C	V	C	GGAFVGPR	C	QD
EGF2	PNP	C	LSTP	C	KNAGT	C	HVVDRRGVADYA	C	S	C	ALGFSGPL	C	LT
EGF3	PLDNA	C	LTNP	C	RNGGT	C	DLLTLTEYK	C	R	C	PPGWSGKS	C	QQ
EGF4	ADP	C	ASNP	C	ANGGQ	C	LPFEASYI	C	H	C	PPSFHGPT	C	RQ
EGF5	DVNE	C	GQKPGL	C	RHGGT	C	HNEVGSYR	C	V	C	RATHTGPN	C	ER
EGF6	PYVP	C	SPSP	C	QNGGT	C	RPTGDVTHE	C	A	C	LPGFTGQN	C	EE
EGF7	NIDD	C	PGNN	C	KNGGA	C	VDGVNTYN	C	R	C	PPEWTGQY	C	TE
EGF8	DVDE	C	QLMPNA	C	QNGGT	C	HNTHGGYN	C	V	C	VNGWTGED	C	SE
EGF9	NIDD	C	ASAA	C	FHGAT	C	HDRVASFY	C	E	C	PHGRTGLL	C	HL
EGF10	NDA	C	ISNP	C	NEGSN	C	DTNPVNGKAI	C	T	C	PSGYTGPA	C	SQ
EGF11	DVDE	C	SLGANP	C	EHAGK	C	INTLGSFE	C	Q	C	LQGYTGPR	C	EI
EGF12	DVNE	C	VSNP	C	QNDAT	C	LDQIGEFQ	C	I	C	MPGYEGVH	C	EV
EGF13	NTDE	C	ASSP	C	LHNGR	C	LDKINEFQ	C	E	C	PTGFTGHL	C	QY
EGF14	DVDE	C	ASTP	C	KNGAK	C	LDGPNTYT	C	V	C	TEGYTGTH	C	EV
EGF15	DIDE	C	DPDP	C	HYGS	C	KDGVATFT	C	L	C	RPGYTGHH	C	ET
EGF16	NINE	C	SSQP	C	RHGGT	C	QDRDNAYL	C	F	C	LKGTTGPN	C	EI
EGF17	NLDD	C	ASSP	C	DSGT	C	LDKIDGYE	C	A	C	EPGYTGSM	C	NI
EGF18	NIDE	C	AGNP	C	HNGGT	C	EDGINGFT	C	R	C	PEGYHDPT	C	LS
EGF19	EVNE	C	NSNP	C	VHGA	C	RDSLNGYK	C	D	C	DPGWSGTN	C	DI
EGF20	NNNE	C	ESNP	C	VNGGT	C	KDMTSGYV	C	T	C	REGFSGPN	C	QT
EGF21	NINE	C	ASNP	C	LNQGT	C	IDDVAGYK	C	N	C	LLPYTGAT	C	EV
EGF22	VLAP	C	APSP	C	RNGGE	C	RQSEDYESFS	C	V	C	PTGWQGQT	C	EV
EGF23	DINE	C	VLSP	C	RHGAS	C	QNTHGGYR	C	H	C	QAGYSGRN	C	ET
EGF24	DIDD	C	RPNP	C	HNGGS	C	TDGINTAF	C	D	C	LPGFRGTF	C	EE
EGF25	DINE	C	ASDP	C	RNGAN	C	TDCVDSYT	C	T	C	PAGFSGIH	C	EN
EGF26	NTPD	C	TESS	C	FNGGT	C	VDGINSFT	C	L	C	PPGFTGSY	C	QH
EGF27	DVNE	C	DSQP	C	LHGGT	C	QDGCGSYR	C	T	C	PQGYTGPN	C	QN
EGF28	LVHW	C	DSSP	C	KNGGK	C	WQTHTQYR	C	E	C	PSGWTGLY	C	DV
EGF29	PSVS	C	EVAAQR QGVDVA RL	C	QHGGL	C	VDAGNTHH	C	R	C	QAGYTGSY	C	ED
EGF30	LVDE	C	SPSP	C	QNGAT	C	TDYLGGYS	C	K	C	VAGYHGVN	C	SE
EGF31	EIDE	C	LSHP	C	QNGGT	C	LDLPNTYK	C	S	C	PRGTQGVH	C	EI
EGF32	NVDD	C	NPPVDP VSRSPK	C	FNNGT	C	VDQVGGYS	C	T	C	PPGFVGER	C	EG
EGF33	DVNE	C	LSNP	C	DARGTQN	C	VQRVNDFH	C	E	C	RAGHTGRR	C	ES
EGF34	VING	C	KGKP	C	KNGGT	C	AVASNTARGFI	C	K	C	PAGFEGAT	C	EN
EGF35	DART	C	GSLR	C	LNGGT	C	ISGPRSPT	C	L	C	LGPFTGPE	C	QF
EGF36	PASSP	C	LGGNP	C	YNQGT	C	EPTSESPFYR	C	L	C	PAKFNGLL	C	HI

Figure 3.2 (continued on next page)

B) Human Notch 2

EGF1	HALQ	C	RDGYEP	C	VNEGM	C	VTYHNGTGY	C	K	C	PEGFLGEY	C	QH
EGF2	RDP	C	EKNR	C	QNGGT	C	VAQAMLGKAT	C	R	C	ASGFTGED	C	QY
EGF3	STSHP	C	FVSRP	C	LNGGT	C	HMLSRDTYE	C	T	C	QVGFTGKE	C	QW
EGF4	TDA	C	LSHP	C	ANGST	C	TTVANQFS	C	K	C	LTGFTGQK	C	ET
EGF5	DVNE	C	DIPGH	C	QHGGT	C	LNLPGSYQ	C	Q	C	PQGFTGQY	C	DS
EGF6	LYVP	C	APSP	C	VNGGT	C	RQTGDFTFE	C	N	C	LPGFEGST	C	ER
EGF7	NIDD	C	PNHR	C	QNGGV	C	VDGVNTYN	C	R	C	PPQWTGQF	C	TE
EGF8	DVDE	C	LLQPNA	C	QNGGT	C	ANRNGGYG	C	V	C	VNGWSGDD	C	SE
EGF9	NIDD	C	AFAS	C	TPGST	C	IDRVASFS	C	M	C	PEGKAGLL	C	HL
EGF10	DDA	C	ISNP	C	HKGAL	C	DTNPLNGQYI	C	T	C	PQGYKGAD	C	TE
EGF11	DVDE	C	AMANSNP	C	EHAGK	C	VNTDGAFH	C	E	C	LKGYAGPR	C	EM
EGF12	DINE	C	HSDP	C	QNDAT	C	LDKIGGFT	C	L	C	MPGFKGVH	C	EL
EGF13	EINE	C	QSNP	C	VNNGQ	C	VDKVNRFQ	C	L	C	PPGFTGPV	C	QI
EGF14	DIDD	C	SSTP	C	LNGAK	C	IDHPNGYE	C	Q	C	ATGFTGVL	C	EE
EGF15	NIDN	C	DPDP	C	HHGQ	C	QDGIDSYT	C	I	C	NPGYMGAI	C	SD
EGF16	QIDE	C	YSSP	C	LNDGR	C	IDLVNGYQ	C	N	C	QPGTSGVN	C	EI
EGF17	NFDD	C	ASNP	C	IHGI	C	MDGINRYS	C	V	C	SPGFTGQR	C	NI
EGF18	DIDE	C	ASNP	C	RKGAT	C	INGVNGFR	C	I	C	PEGPHHPS	C	YS
EGF19	QVNE	C	LSNP	C	IHGN	C	TGGLSGYK	C	L	C	DAGWVGIN	C	EV
EGF20	DKNE	C	LSNP	C	QNGGT	C	DNLVNGYR	C	T	C	KKGFKGYN	C	QV
EGF21	NIDE	C	ASNP	C	LNQGT	C	FDDISGYT	C	H	C	VLPYTGKN	C	QT
EGF22	VLAP	C	SPNP	C	ENAAV	C	KESPNFESYT	C	L	C	APGWQGQR	C	TI
EGF23	DIDE	C	ISKP	C	MNHGL	C	HNTQGSYM	C	E	C	PPGFSGMD	C	EE
EGF24	DIDD	C	LANP	C	QNGGS	C	MDGVNTFS	C	L	C	LPGFTGDK	C	QT
EGF25	DMNE	C	LSEP	C	KNGGT	C	SDYVNSYT	C	K	C	QAGFDGVH	C	EN
EGF26	NINE	C	TESS	C	FNGGT	C	VDGINSFS	C	L	C	PVGFTGSF	C	LH
EGF27	EINE	C	SSHP	C	LNEGT	C	VDGLGTYR	C	S	C	PLGYTGKN	C	QT
EGF28	LVNL	C	SRSP	C	KNKGT	C	VQKKAESQ	C	L	C	PSGWAGAY	C	DV
EGF29	PNVS	C	DIAASRRG VLVEHL	C	QHSGV	C	INAGNTHY	C	Q	C	PLGYTGSY	C	EE
EGF30	QLDE	C	ASNP	C	QHGAT	C	SDFIGGYR	C	E	C	VPGYQGVN	C	EY
EGF31	EVDE	C	QNQP	C	QNGGT	C	IDLVNHFK	C	S	C	PPGTRGLL	C	EE
EGF32	NIDD	C	ARGPH	C	LNGGQ	C	MDRIGGYS	C	R	C	LPGFAGER	C	EG
EGF33	DINE	C	LSNP	C	SSEGSL D	C	IQLTNDYL	C	V	C	RSAFTGRH	C	ET
EGF34	FVDV	C	PQMP	C	LNGGT	C	AVASNMPDGFI	C	R	C	PPGFSGAR	C	QS
EGF35	S	C	GQVK	C	RKGEQ	C	VHTASGPR	C	F	C	PSPRD	C	ES
EGF36	G	C	ASSP	C	QHGGS	C	HPQRQPPYYS	C	Q	C	APPFSGSR	C	EL

Figure 3.2 (continued on next page)

C) Drosophila Notch

EGF1	VAAS	C	TSVG	C	QNGGT	C	VTQLNGKTY	C	A	C	DSHYVGDY	C	EH
EGF2	RNP	C	NSMR	C	QNGGT	C	QVTFRNGRPGIS	C	K	C	PLGFDESL	C	EI
EGF3	AVPNA	C	DHVT	C	LNGGT	C	QLKTLEEYT	C	A	C	ANGYTGER	C	ET
EGF4	KNL	C	ASSP	C	RNGAT	C	TALAGSSSFT	C	S	C	PPGFTGDT	C	SY
EGF5	DIEE	C	QSNP	C	KYGGT	C	VNTHGSYQ	C	M	C	PTGYTGKD	C	DT
EGF6	KYKP	C	SPSP	C	QNGGI	C	RSNGLSYE	C	K	C	PKGFEGKN	C	EQ
EGF7	NYDD	C	LGHL	C	QNGGT	C	IDGISDYT	C	R	C	PPNFTGRF	C	QD
EGF8	DVDE	C	AQRDHPV	C	QNGAT	C	TNTHGSYS	C	I	C	VNGWAGLD	C	SN
EGF9	NTDD	C	KQAA	C	FYGAT	C	IDGVGSFY	C	Q	C	TKGKTGLL	C	HL
EGF10	DDA	C	TSNP	C	HADAI	C	DTSPINGSYA	C	S	C	ATGYKGVD	C	SE
EGF11	DIDE	C	DQGSP	C	EHNGI	C	VNTPGSYR	C	N	C	SQGFTGPR	C	ET
EGF12	NINE	C	ESHP	C	QNEGS	C	LDDPGTFR	C	V	C	MPGFTGTQ	C	EI
EGF13	DIDE	C	QSNP	C	LNDGT	C	HDKINGFK	C	S	C	ALGFTGAR	C	QI
EGF14	NIDD	C	QSQP	C	RNRGI	C	HDSIAGYS	C	E	C	PPGYTGTS	C	EI
EGF15	NIND	C	DSNP	C	HRGK	C	IDDVNSFK	C	L	C	DPGYTGYI	C	QK
EGF16	QINE	C	ESNP	C	QFDGH	C	QDRVGSYY	C	Q	C	QAGTSGKN	C	EV
EGF17	NVNE	C	HSNP	C	NNGAT	C	IDGINSYK	C	Q	C	VPGFTGQH	C	EK
EGF18	NVDE	C	ISSP	C	ANNGV	C	IDQVNGYK	C	E	C	PRGFYDAH	C	LS
EGF19	DVDE	C	ASNP	C	VNEGR	C	EDGINEFI	C	H	C	PPGYTGKR	C	EL
EGF20	DIDE	C	SSNP	C	QHGGT	C	YDKLNAFS	C	Q	C	MPGYTGQK	C	ET
EGF21	NIDD	C	VTNP	C	GNGGT	C	IDKVNGYK	C	V	C	KVPFTGRD	C	ES
EGF22	KMDP	C	ASNR	C	KNEAK	C	TPSSNFLDFS	C	T	C	KLGYTGRY	C	DE
EGF23	DIDE	C	SLSSP	C	RNGAS	C	LNVPGSYR	C	L	C	TKGYEGRD	C	AI
EGF24	NTDD	C	ASFP	C	QNGGT	C	LDGIGDYS	C	L	C	VDGFDGKH	C	ET
EGF25	DINE	C	LSQP	C	QNGAT	C	SQYVNSYT	C	T	C	PLGFSGIN	C	QT
EGF26	NDED	C	TESS	C	LNGGS	C	IDGINGYN	C	S	C	LAGYSGAN	C	QY
EGF27	KLNK	C	DSNP	C	LNGAT	C	HEQNNEYT	C	H	C	PSGFTGKQ	C	SE
EGF28	YVDW	C	GQSP	C	ENGAT	C	SQMKHQFS	C	K	C	SAGWTGKL	C	DV
EGF29	QTIS	C	QDAADRKGLSLRQL	C	NNGT	C	KDYGNSHV	C	Y	C	SQGYAGSY	C	QK
EGF30	EIDE	C	QSQP	C	QNGGT	C	RDLIGAYE	C	Q	C	RQGFQGQN	C	EL
EGF31	NIDD	C	APNP	C	QNGGT	C	HDRVMNFS	C	S	C	PPGTMGII	C	EI
EGF32	NKDD	C	KPGA	C	HNNGS	C	IDRVGGFE	C	V	C	QPGFVGAR	C	EG
EGF33	DINE	C	LSNP	C	SNAGTLD	C	VQLVNNYH	C	N	C	RPGHMGRH	C	EH
EGF34	KVDF	C	AQSP	C	QNGGN	C	NIRQSGHH	C	I	C	NNGFYGKN	C	EL
EGF35	SGQD	C	DSNP	C	RVGN	C	VVADEGFGYR	C	E	C	PRGTLGEH	C	EI
EGF36	DTLDE	C	SPNP	C	AQGAA	C	EDLLGDYE	C	L	C	PSKWKGKR	C	DI

Figure 3.2 Sequence analysis of the hN1, hN2, and dN receptors: The amino acid sequences of hN1 (A), hN2 (B), and dN (C) are shown as one EGF domain per row. These include the Abruptex region EGF24-29 (blue), cysteines (yellow), with conserved cysteine aligned in columns. For the Abruptex region, the core ligand binding region (EGF11-12) and the region in between (EGF13-29), Ca^{2+} consensus residues D/N-X-D/N-E/Q-X_a-D*/N*-X_b-Y/F, where a and b are variable and * denotes β- hydroxylation are highlighted in red if the complete consensus is predicted or shown to bind Ca^{2+}. Consensus hydrophobic packing residues are shown (green), with substitutions of this conserved residue indicated (cyan).

The three canonical disulphide bonds are a core feature of EGF domains (See Chapter 1 – Introduction, Section 1.3). A surprising feature of the hN1 Abruptex region EGF24-29 is the presence of non-canonical additional cysteine residues in EGF25 and EGF27. The additional cysteine in EGF25 is also observed in mN1 but not dN. The presence of non-canonical cysteines is unusual because these highly reactive residues are normally involved in intramolecular disulphide bonding, and so the addition of these residues may suggest a structural role. However, the lack of these residues in any other Notch receptors and the presence of only a single extra cysteine in mN1 suggest that the extra cysteines in hN1 may not have a significant structural role.

Another interesting observation from the bioinformatics analysis is the presence of a threonine in the position of the conserved aromatic packing residue in EGF16. This substitution is also present in mN1, suggesting that it might lead to a conserved divergence in this region away from a canonical rigid interface (See Chapter 1 – Introduction, Figure 1.6). Possible consequences are an increase in flexibility at the EGF16-17 interface due to the potential lack of this conserved packing interaction between cbEGFs. From this analysis, it is possible to predict a rod-like structure for EGF11-15 and EGF17-21.

3.2. Expression of Notch constructs from the S2 system

3.2.1. Overview of the S2 system

The S2 system is a cell-based expression system derived from *D. melanogaster*, is able to produce milligram per litre amounts of glycosylated protein, and secrete this into the growth medium for purification (de Jongh et al, 2013). This system can produce long multiple tandem EGF-domain constructs up to 15 EGF domains (Suckling et al, unpublished data). Although glycosylation and folding are not directly measurable, crystal structures of tandem EGF-domain constructs from this system have been determined, for example S2 hN2 11-13 (Suckling et al, 2017), S2 hN1 20-24 (Suckling et al, unpublished data), suggesting that the system performs well in producing large amounts of pure folded protein, modified with at least one sugar at consensus modification sites (Suckling et al, unpublished data). To further support this, ligand constructs produced in the S2 system are natively folded and can be purified to homogeneity and crystallised (Suckling et al, 2017). The S2 system was used to produce protein for binding studies and structural studies.

3.2.2. Designing constructs for production in the S2 system

Following analysis of the sequence of the hN1 Abruptex region described in Table 3.2, the presence of the two non-canonical cysteines in EGF25 and EGF27 suggested that any constructs in this region should contain both of these cysteines, or neither, in anticipation

that these might form a disulphide bond between them. This would also minimise the chance of artificial disulphide bonding patterns brought about in this high-level expression system if only one of these cysteines were present. Taking this into account, constructs were designed to include either up to and including EGF24, or to include both EGF25 and EGF27. Furthermore, as the products of S2 cells are secreted, these constructs are more likely to form the correct disulphide bonding pattern, as opposed to the prokaryotic/refold system which are refolded *in vitro* in which the presence of additional non-canonical cysteines has been shown to decrease yield (Whiteman et al, unpublished data). The lack of canonical Ca^{2+}-binding consensus sequences in EGF28-29 also increases the chances of flexibility in this region so a construct was produced up to and including EGF27.

Constructs from the S2 system were expressed with a C-terminal BirA-His tag which aids identification of cell lines expressing complete constructs (See Appendix). These tags could then be cleaved with 3C protease (cleavage reactions were conducted by Dr R Suckling, University of Oxford). The constructs produced in the S2 system for investigation are described in Figure 3.3. and further rationale behind the design of these constructs is described in subsequent results chapters. Expression and purification of these constructs were carried out as described in Chapter 2 - Materials and Methods, Section 2.4.3.

Construct	MW
S2 hN1 11-13	17134
S2 hN1 11-13 A426R	17219
S2 hN1 11-13 K428E	17135
S2 hN1 11-13 E437R	17161
S2 hN1 11-13 R448E	17107
S2 hN1 11-13 N459R	17176
S2 hN1 11-13 Q462R	17162
S2 hN1 11-13 L468R	17177
S2 hN1 11-13 V485R	17192
S2 hN1 11-13cl*	13801
S2 hN1 23-24	12283
S2 hN1 23-24$_{CHis}$*	12327
S2 hN1 23-29	33618
S2 hN1 23-27	24251

*Purified by R Suckling

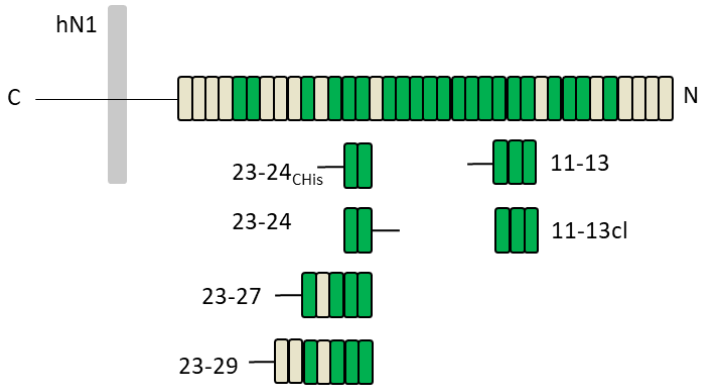


Figure 3.3 Constructs expressed in the S2 system: Table of expressed constructs, MW indicates mass of non-glycosylated material, including any tags present (for -cl constructs with tags cleaved, MW was calculated without tag). Schematic of constructs in the table, shown with in the context of full length hN1 (N and C termini indicated with membrane bilayer (grey)) EGFs (beige) and cbEGFS (green) and BirA-His tag (black line).

3.2.3. Purification of S2 hN1 11-13 variants

S2 hN1 11-13 variants were designed for investigation by small angle X-ray scattering and binding assays. Codon optimized cDNA was used to produce 11-13 wt pEXS2-2 plasmid in collaboration with Dr R Suckling, University of Oxford. This plasmid was used as template to produce S2 hN1 11-13 variant constructs by overlap extension mutagenesis (See Chapter 2 - Materials and Methods, Section 2.1.8.1). These variants were then expressed and purified following a protocol utilized previously for ligand constructs (Suckling et al, 2017), which is

described in Chapter 2 - Material and Methods. Plasmids were sequenced to ensure successful mutagenesis. 1 L of S2 culture expressing the protein was loaded on to a Ni^{2+}-affinity column, followed by washing and then elution with imidazole, the eluted fraction was then purified by size-exclusion chromatography on an S75 column, before pooling of fractions containing the protein of interest (see Figure 3.4). Figure 3.4 outlines the purification of the variant S2 hN1 11-13 N459R and reflects the standard protocol carried out for each variant. Details of purifications of the other S2 hN1 11-13 constructs are outlined in Figure 3.5. S2 hN1 11-13cl (BirA-His tag cleaved) was purified by Dr R Suckling, University of Oxford. All S2 hN1 11-13 variants purified to a near homogenous (> 95 %) product, with the exception of S2 hN1 11-13cl which shows a minor contaminating band, likely corresponding to uncleaved protein (Figure 3.5B). Non-reducing SDS-PAGE analysis and size exclusion chromatography profiles show that all S2 hN1 11-13 constructs run as a single species in a similar way to wt, suggesting these products are correctly folded (Figure 3.5). In addition, S2 hN1 11-13 has shown to be folded in crystal structures (Suckling et al, unpublished data). 1D NMR analysis (See Chapter 2 – Materials and Methods, Section 2.4.2.1.2) was used to observe a Ca^{2+}-dependent peak shift in S2 hN1 11-13 variants including substitutions at positions L468, R448, further supporting that these single residue substitution constructs are correctly folded.

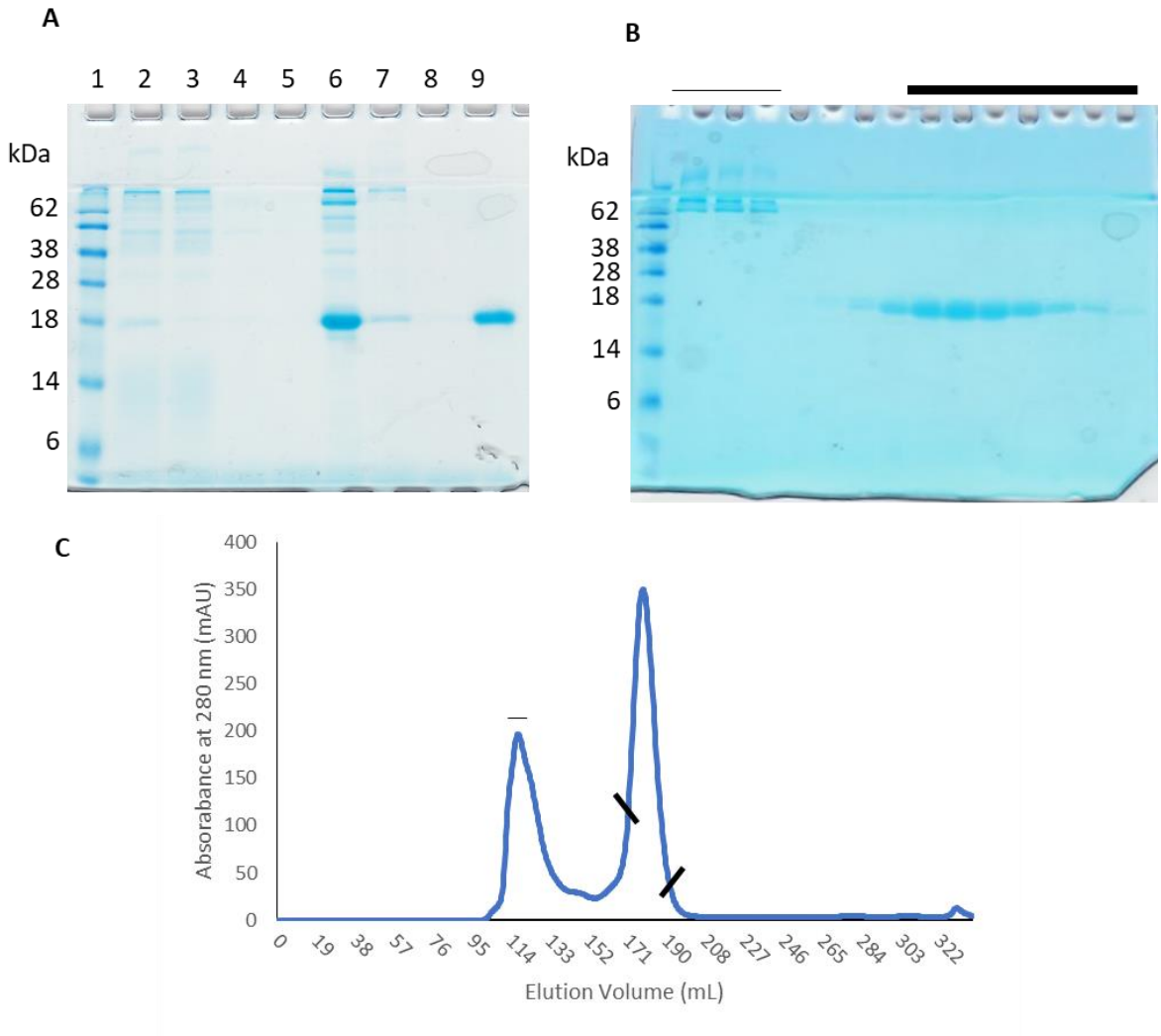


Figure 3.4 General purification procedure for S2 hN1 11-13 variants: Purification of the variant S2 hN1 11-13 N459R is shown, involving Ni^{2+} affinity chromatography followed by SEC, which is typical of all 11-13 variant purifications. (A) The purification procedure is sampled at multiple stages by reducing 16 % SDS-PAGE stained with Coomassie blue: 1) Ladder (NEB), 2) S2 culture supernatant, 3) Ni^{2+} flowthrough, 4) Ni^{2+} wash 1 (30ml), 5) Ni^{2+} wash 2 (30ml), 6) Ni^{2+} elution 1 (10ml), 7) Ni^{2+} elution 2 (10ml), 8) Ni^{2+} elution 3 (30ml), 9) Pooled SEC fractions. (C) Size exclusion chromatography (SEC) was carried out and absorbance at 280 nm recorded over the course of the isocratic elution. (B) Peaks from this elution were analyzed by reducing SDS-PAGE with void peak sampled (thin solid line), and fractions pooled for further experiments (thick solid line). MW from comparison with ladder matches with expected MW of 17176.

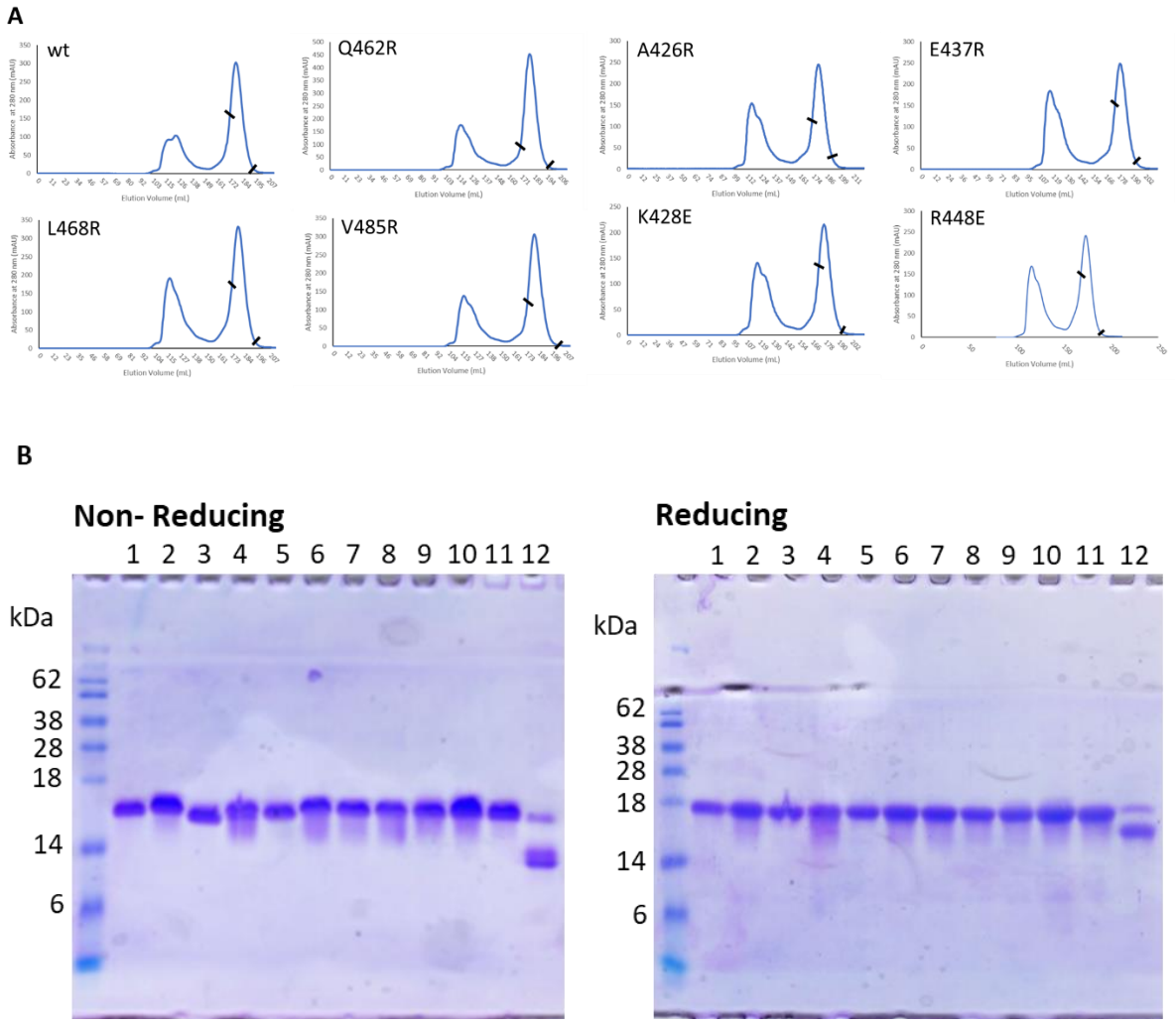


Figure 3.5 16 % SDS-PAGE analysis showing final purified products of hN1 11-13 variants: (A) SEC traces show elution profile with fractions pooled for final product indicated (black bars). First (void) peak was checked for eluted MW using SDS-PAGE as previously (data not shown). Q462R and L468R elution profiles shown are representative of both independent preps carried out for each (prep 1 and prep 2). (B) Reducing and non-reducing 16% SDS-PAGE gels stained with Coomassie blue showing final purified products. MW indicated by comparison with ladder markers matches expected MW of ~17200 for 11-13 variants, with band shift seen after cleavage for hN1 11-13cl. 1) S2 hN1 11-13, 2) S2 hN1 11-13 A426R, 3) S2 hN1 11-13 K428E, 4) S2 hN1 11-13 E437R, 5) S2 hN1 11-13 R448E, 6) S2 hN1 11-13 N459R, 7) S2 hN1 11-13 Q462R prep 1, 8) hN1 11-13 Q462R prep 2, 9) S2 hN1 11-13 L468R prep 1, 10) S2 hN1 11-13 L468R prep 2, 11) S2 hN1 11-13 V485R, 12) S2 hN1 11-13cl. The minor contaminant band in S2 hN1 11-13cl likely corresponds to uncleaved protein.

3.2.4. Purification of S2 hN1 Abruptex constructs

S2 hN1 Abruptex constructs were purified for use in binding assays as described in Chapter 2 – Materials and Methods. 1 L of S2 culture expressing the protein was loaded on to a Ni^{2+}-affinity column, followed by washing and then elution with imidazole, the eluted fraction was then purified by size-exclusion chromatography (SEC) on an S200 column (S75 for 23-24 constructs), before pooling of fractions containing the protein of interest.

3.2.4.1. S2 hN1 23-24 variants

S2 hN1 23-24 was expressed according to the scheme described for S2 purification, with the SEC purification conducted with an S75 column at pH9 (Figure 3.6). This pH was used for purification as a greater yield was observed at pH9 than pH7.5 when purifying a similar construct in the prokaryotic/refold system (Whiteman et al, data not shown). The similar construct with BirA-His tag at the C-terminal of the construct (S2 hN1 23-24C_{His}) was expressed and purified by Dr R Suckling, University of Oxford.

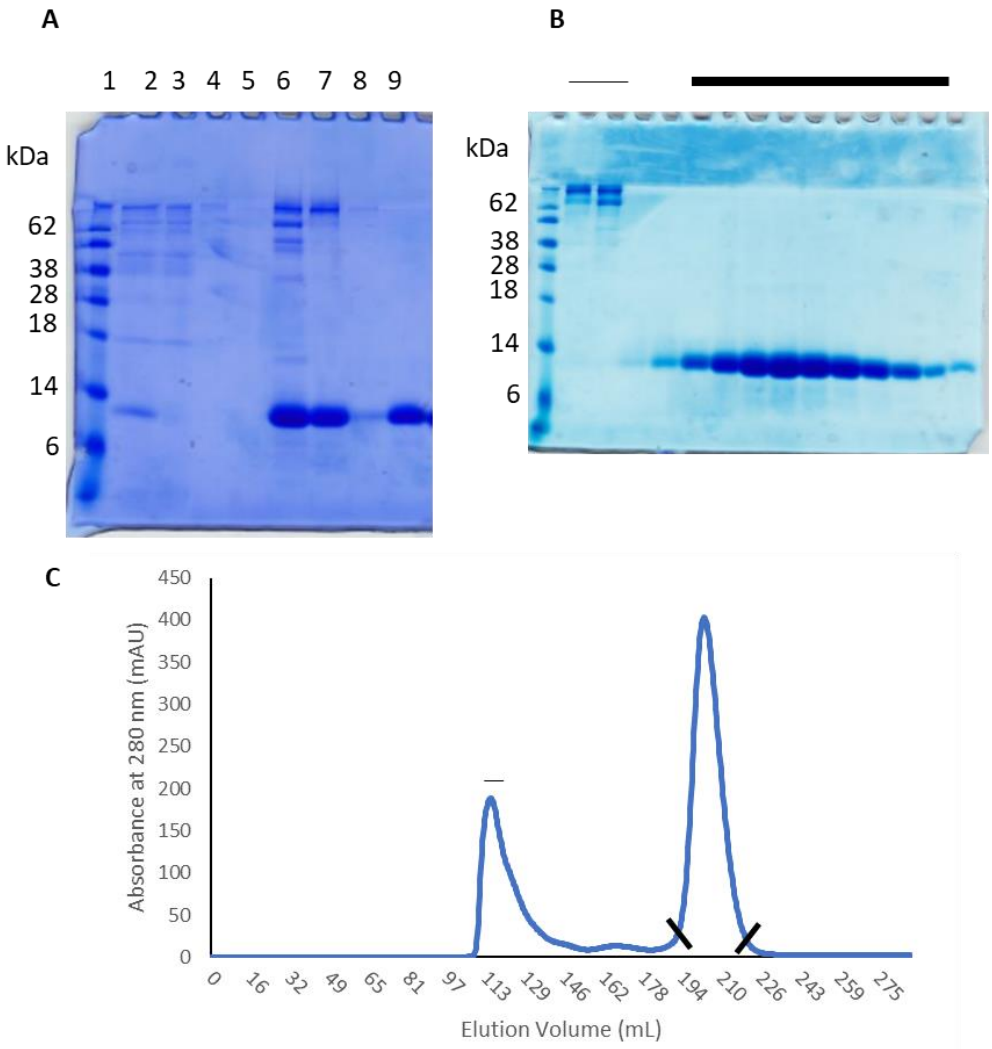


Figure 3.6 Purification of S2 hN1 23-24: (A) The purification procedure was sampled at multiple stages by reducing 16% SDS-PAGE stained with Coomassie blue. 1) Ladder (NEB), 2) S2 culture supernatant, 3) Ni^{2+}- flowthrough, 4) Ni^{2+} wash 1 (30ml), 5) Ni^{2+} wash 2 (30ml), 6) Ni^{2+} elution 1 (10ml), 7) Ni^{2+} elution 2 (10ml), 8) Ni^{2+} elution 3 (30ml), 9) Pooled SEC fractions. MW from gel ladder matches with expected MW of ~1200. (C) Size exclusion chromatography (SEC) was carried out using an S75 column and absorbance at 280 nm recorded over the course of the isocratic elution. (B) Peaks from this elution were analyzed by reducing SDS-PAGE with void peak sampled (thin solid line), and fractions pooled for further experiments (thick solid line).

3.2.4.2. S2 hN1 23-29

S2 hN1 23-29 was expressed according to the scheme described for S2 purification, with the SEC purification conducted with at pH9 using an S200 column (Figure 3.7). This pH was used for purification as a greater yield was observed at pH9 than pH7.5 when purifying a shorter Ax construct in the prokaryotic/refold system (Whiteman et al, data not shown).

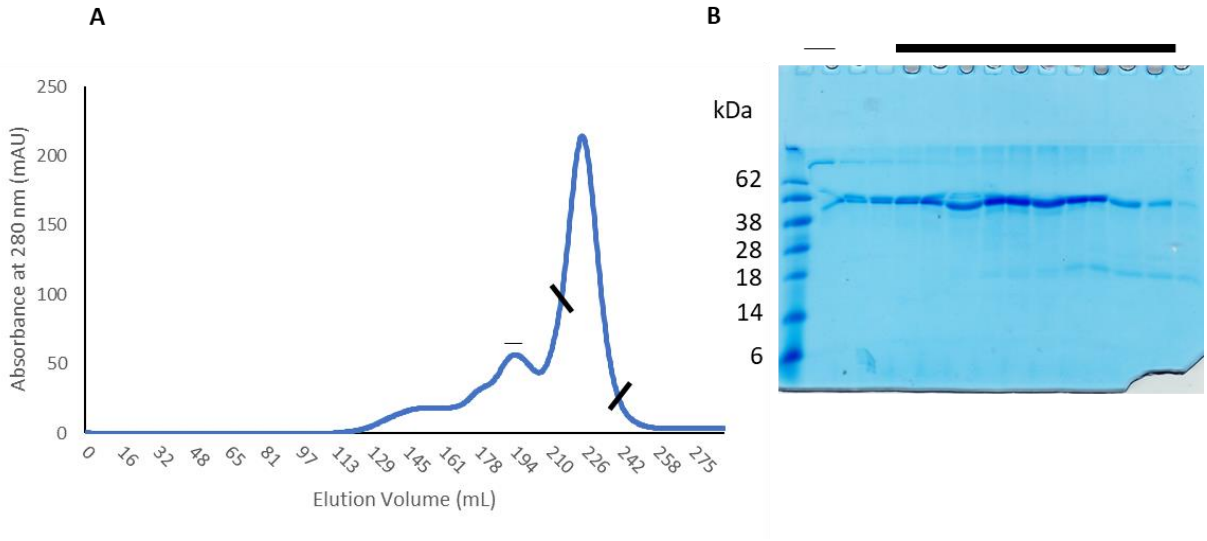


Figure 3.7 Purification of S2 hN1 23-29: (A) Following Ni^{2+} affinity chromatography, size exclusion chromatography (SEC) was carried out and absorbance at 280 nm recorded over the course of the isocratic elution. (B) Peaks from this elution were analyzed by reducing 16% SDS-PAGE stained with Coomassie blue with void peak sampled (thin solid line), and fractions pooled for further experiments (thick solid line). MW from gel ladder suggests a greater mass than the expected MW of 33618, however this is likely due to the presence of sugar modifications. Furthermore some distortion when running the gel appears to have pushed sample lanes to a higher position on the gel (See Figure 3.9 for pooled sample run on undistorted gel).

3.2.4.3. S2 hN1 23-27

S2 hN1 23-27 was expressed according to the scheme described for S2 purification, with the SEC purification conducted at pH9 using an S200 column (Figure 3.8). This purification shows a larger void peak, and broader (relative to height) peak of interest profile compared to previous purifications. In addition, the presence of a band similar in mass to the band of interest (between 28-38 kDa) in the void peak is apparent in the reducing SDS-PAGE (Figure 3.8B). This suggests the construct may interact to form complexes with itself or other cellular proteins, possibly mediated by the additional non-canonical cysteines in EGF25 and EGF27.

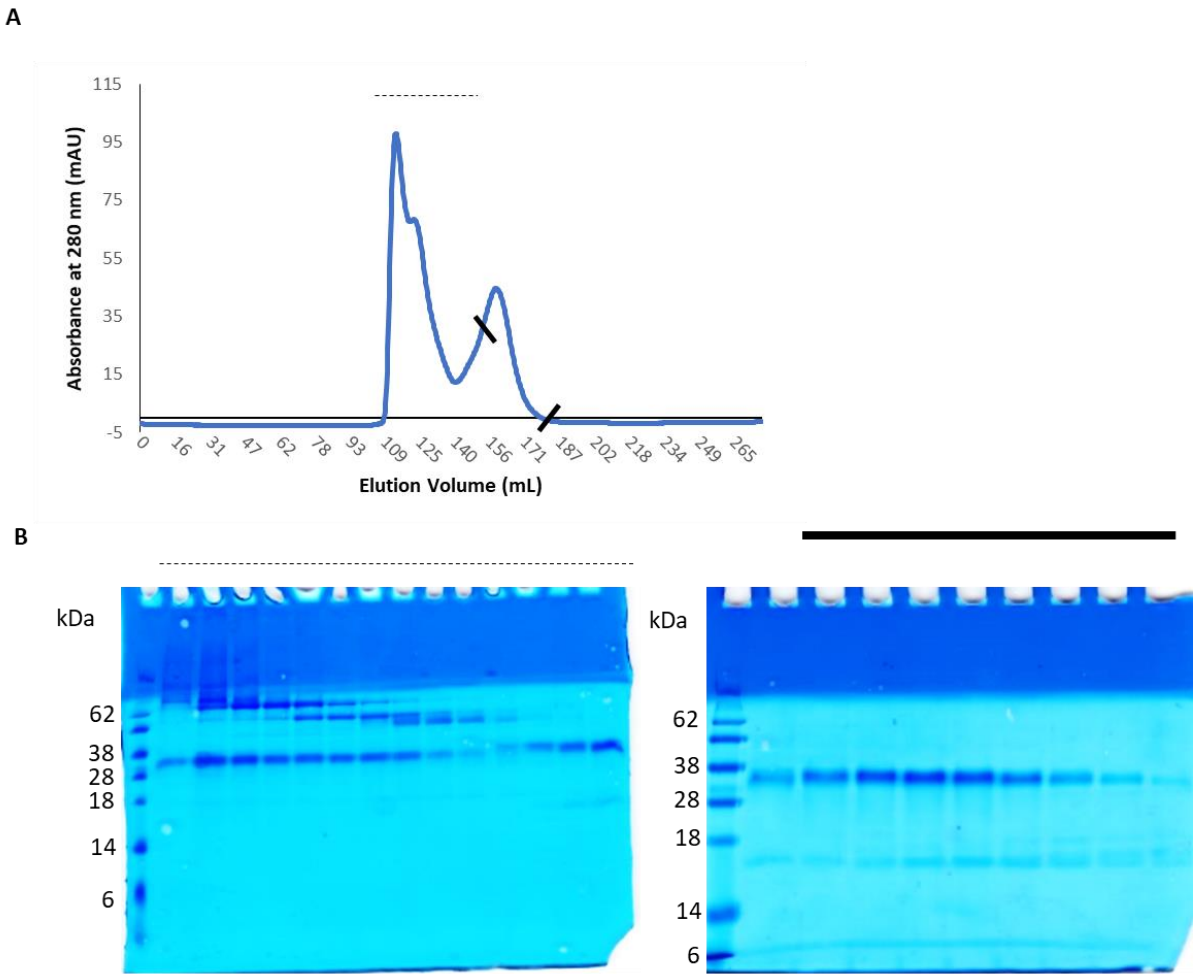


Figure 3.8 Purification of S2 hN1 23-27: (A) Following Ni^{2+} affinity chromatography, size exclusion chromatography (SEC) was carried out and absorbance at 280 nm recorded over the course of the isocratic elution. (B) Peaks from this elution were analyzed by reducing 16 % SDS-PAGE stained with Coomassie blue with void peak sampled (dashed line), and fractions pooled for further experiments (thick solid line). This elution shows that more than half of the material eluted from the Ni^{2+} column is discarded during SEC. MW inferred from ladder suggests a greater mass than the expected MW of 24251, however this is likely due to the presence of glycan modifications. The additional band at ~17 kDa which elutes with the protein of interest is likely a bound contaminant.

3.2.4.4. Purified S2 Ax constructs

The final products of each S2 hN1 Ax purification are shown in analysis by 16% SDS-PAGE (Figure 3.9). The SDS-PAGE gels indicate samples are of high purity and run with >90% of protein in a single band under reducing conditions, except for S2 hN1 23-27. The multiple bands observed under non-reducing conditions for S2 hN1 23-27 further supports that this construct has a variety of disulphide bond patterns leading to a range of conformations in non-reducing conditions. Furthermore, hN1 23-29 appears to run as multiple bands under non-reducing conditions, however appears to form fewer species than S2 hN1 23-27 suggesting the presence of EGF28-29 make the region more likely to fold into a single conformation. Due the heterogenous nature of the sample under non-reducing conditions, these constructs were not pursued for analysis in Small angle X-ray scattering.

3.3. Expression of constructs from the prokaryotic/refold (PR) system

The expression of multiple tandem EGF constructs from *E. coli* cells followed by reduction and refold *in vitro* is an established system for purifying milligram amounts of EGF-domain constructs (Whiteman et al, 2014, Cordle et al, 2008, Lee et al, 2004) (See Appendix). The lack of post-translational modifications allows characterisation of protein by NMR and mass spectrometry. This system is limited to a maximum length of ~5 EGF domains (~30 cysteines) due to the complexity of the refolding step, with yield significantly decreasing as number of domains and therefore the number of cysteines increases beyond this limit. The optimal tag-

organization for this construct is a Factor Xa-cleavable N-terminal His-tag, and a C-terminal BirA-tag as demonstrated previously to greatly increase stability and yield during the refold and purification (Cordle et al, 2008). Constructs were purified in this system for structural analysis in small angle X-ray scattering, and binding assays.

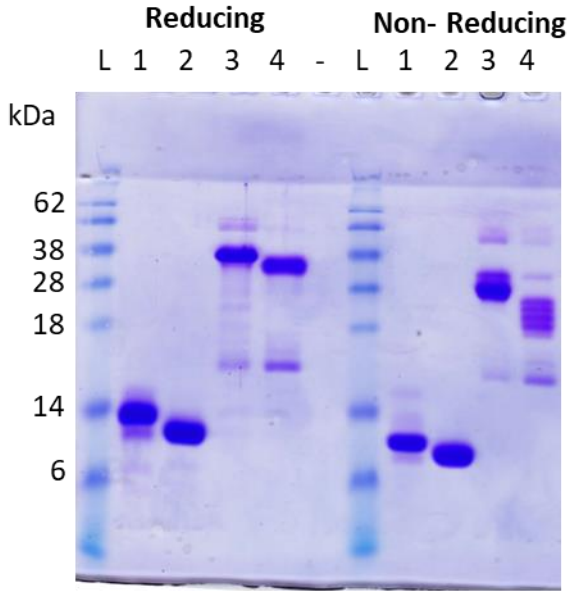


Figure 3.9 Purified S2 hN1 Ax constructs: Samples were separated on a 16% SDS-PAGE under reducing and non-reducing conditions and stained with Coomassie blue. MWs inferred from the pre-stained ladder match expected MWs when considering likely increase in apparent mass due to the presence of sugar modifications. L) Pre-stained ladder (NEB), 1) S2 hN1 23-24_{CHis}, 2) S2 hN1 23-24, 3) S2 hN1 23-29, 4) S2 hN1 23-27.

3.3.1. Designing and expressing constructs from the prokaryotic/refold system

Due to the nature of the refolding step, constructs containing the non-canonical additional cysteine residues in EGF25 and EGF27 (See Figure 3.2) were substituted for serine, a similarly

sized residue without the potential to form disulphide bonds. The constructs produced for this study in the prokaryotic/refold system are described in Figure 3.10.

Construct	MW
PR hN1 23-24^	12444
PR hN1 23-24 T940R	12338
PR hN1 23-24 H917W	12332
PR hN1 23-24 D932R	12324
PR hN1 23-24 L934R	12326
PR hN1 23-27CSCS	24145
PR hN2 23-27*	24300

^^{15}N enriched, * Purified by P Whiteman

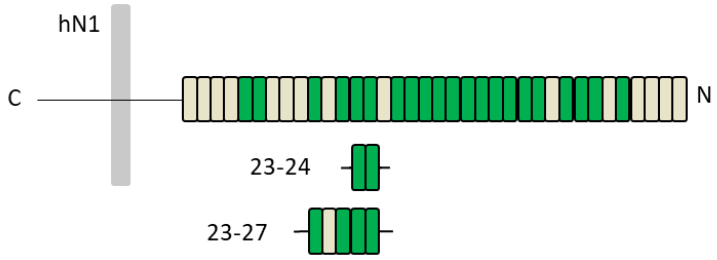


Figure 3.10 Constructs expressed in the prokaryotic/refold system: Table of expressed constructs with MW indicating mass including any tags present. Schematic of constructs in the table, shown within the context of full length hN1 (N and C termini indicated with membrane bilayer (grey)) EGFs (beige) and cbEGFS (green) and N-terminal His tag and C-terminal BirA tag (black lines). See Appendix for sequences of constructs produced.

3.3.2. Purification of PR hN1 23-24 variants

Purification of PR hN1 23-24 variants was carried out as described in Chapter 2 - Materials and Methods. The purification involved Ni^{2+} affinity chromatography, followed by reduction with dithiothreitol before numerous reverse phase and ion exchange purification steps. The protein is visualized between each stage during the reverse-phase HPLC and Figure 3.11 shows the traces recorded throughout the course of the purification of PR hN1 23-24. This purification reflects the procedure carried out for all PR hN1 23-24 variants. Details of purifications of the other variants are outlined in Figure 3.12. This purification method

enabled efficient production of mg amounts of highly pure natively folded proteins as shown in Figure 3.13, further supported by previous NMR and crystallographic studies (Whiteman et al, unpublished data).

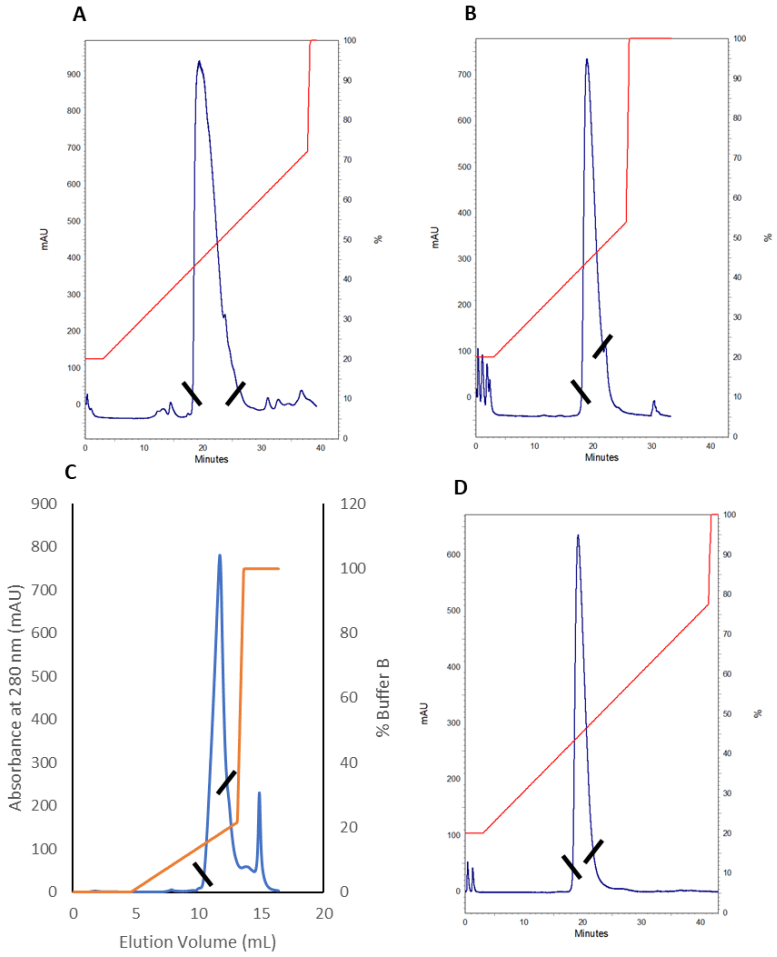


Figure 3.11 Elution traces from multiple stages throughout the purification of PR hN1 23-24: Peak traces throughout the course of the purification are shown with a sharp, symmetrical peak suggesting a pure protein sample. (A) Reverse-phase HPLC following reduction, (B) Reverse-phase HPLC after in-vitro refolding, (C) Ion exchange FPLC, (D) Final reverse-phase HPLC. A_{280nm} absorbance trace (blue) is shown with collected fractions corresponding to the area between the black bars, the red trace corresponds to elution gradient with hydrophobic solvent in A, B, D, or increasing ionic strength in C (See Chapter 2 – Materials and methods).

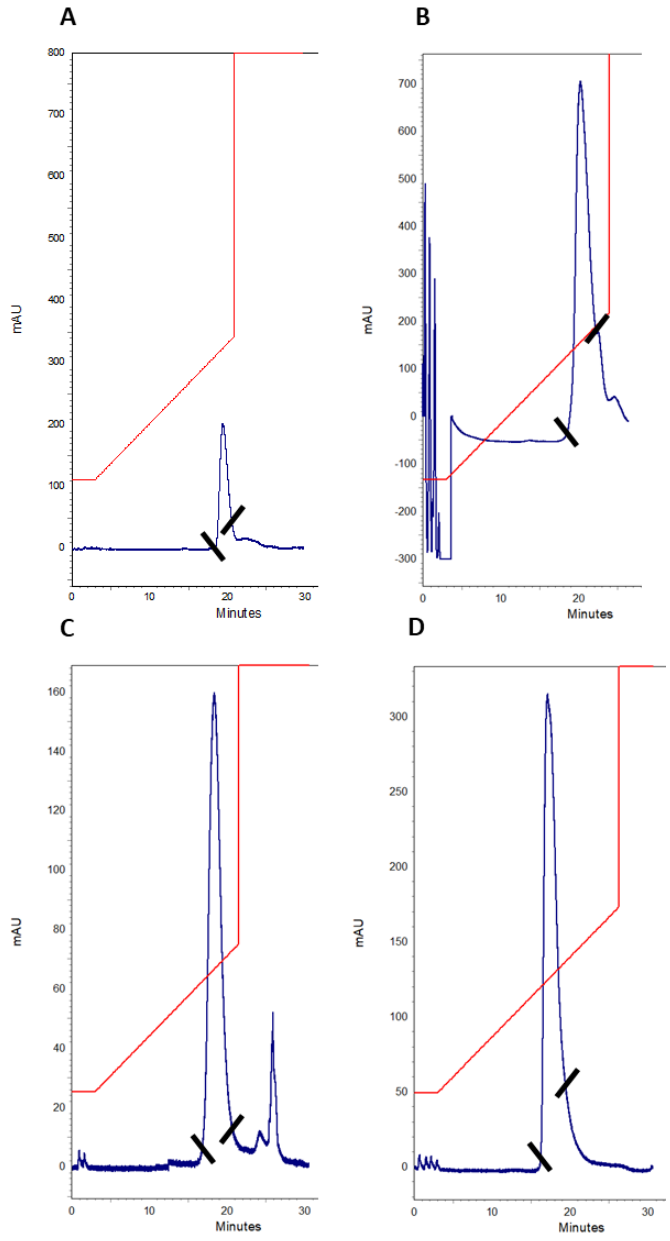


Figure 3.12 Final reverse-phase HPLC traces of PR hN1 23-24 variants: Final reverse-phase HPLC traces are shown for (A) PR hN1 23-24 T940R, (B) PR hN1 23-24 H917W, C) PR hN1 23-24 D932R, (D) PR hN1 23-24 L934R. A_{280nm} absorbance trace (blue) is shown with collected fraction corresponding to the area between the black bars, red trace corresponds to elution gradient with hydrophobic solvent. A sharp, symmetrical peak suggests a homogenous protein sample.

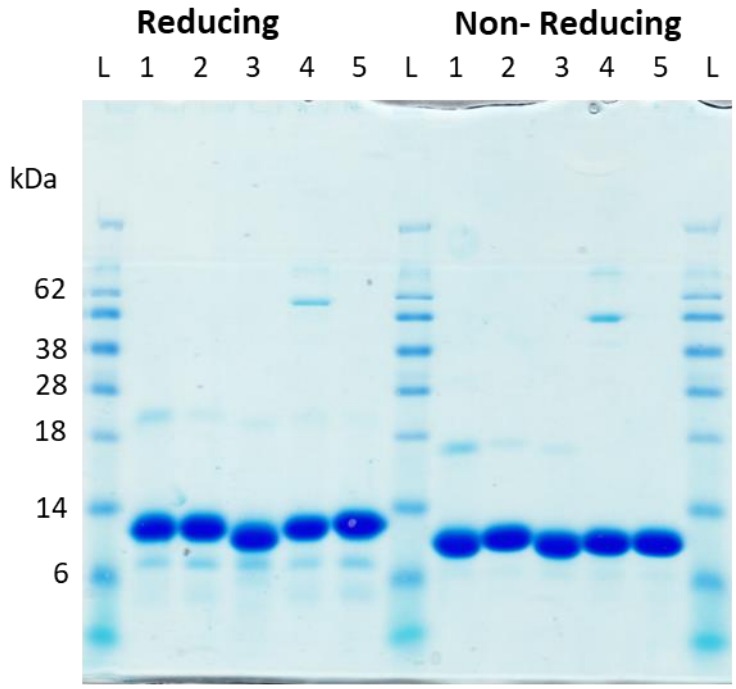


Figure 3.13 16% SDS-PAGE of purified PR hN1 23-24 variants: Following various chromatographic purifications by HPLC and FPLC. Samples were analyzed on a 16% SDS-PAGE under reducing and non-reducing conditions and stained with Coomassie blue. The theoretical MW of ~12500 matches the MW of the bands observed corresponding to purified product inferred from the ladder. L) Pre-stained ladder (NEB), 1) PR hN1 23-24, 2) PR hN1 23-24 T940R, 3) PR hN1 23-24 H917W, 4) PR hN1 23-24 D932R, 5) PR hN1 23-24 L934R.

3.3.3. Purification of PR hN1 23-27CSCS

PR hN1 23-27CSCS was designed to overcome potential difficulties of *in vitro* refolding with potentially unpaired cysteines step by substituting the non-canonical cysteines in EGF25 and EGF27 for serines. Following Ni^{2+} affinity chromatography and reduction, the steps of the purification are shown in Figure 3.14. Figure 3.15 shows the final product analyzed on 16% SDS-PAGE stained with Coomassie blue. The N-terminal His-tag of hN1 23-27CSCS was

subsequently cleaved with Factor Xa to produce hN1 23-27CSCScl, with the cleaved product shown in Figure 3.15.

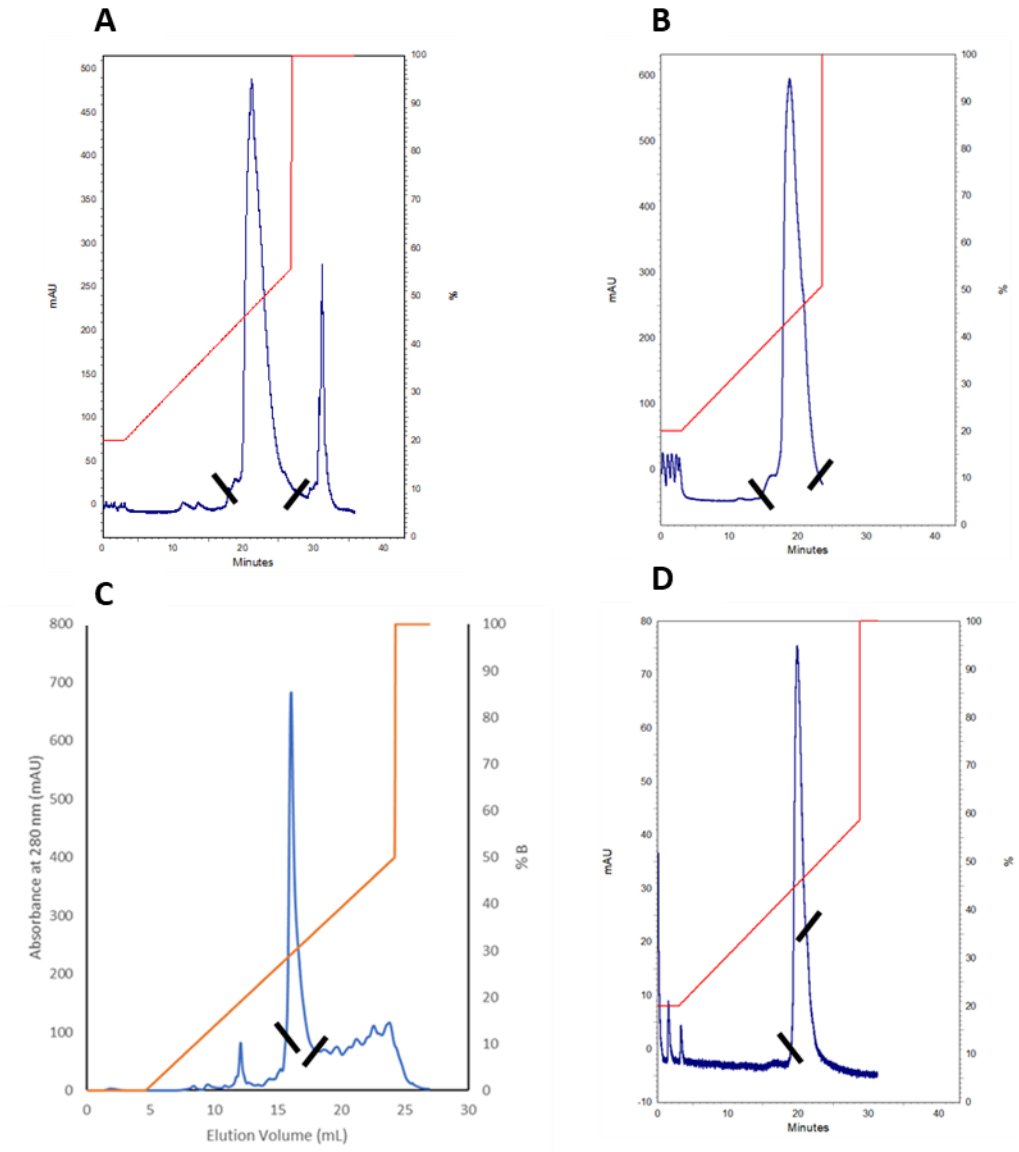


Figure 3.14 Elution traces from multiple stages throughout the purification of PR hN1 23-27CSCS: Visualisation throughout the course of the purification are shown, with a sharp symmetrical peak suggesting a pure protein sample. (A) Reverse-phase HPLC following reduction, (B) Reverse-phase HPLC after in-vitro refolding, (C) Ion exchange FPLC, (D) Final reverse-phase HPLC. A_{280nm} absorbance trace (blue) is shown with collected fraction corresponding to the area between the black bars, red trace corresponds to elution gradient with hydrophobic solvent in A, B, D, or increasing ionic strength in C.

3.3.4. Purification of PR hN2 23-27cl

PR hN2 23-27 was expressed and purified following the procedure outlined in Chapter 2 - Materials and Methods in collaboration with Dr P Whiteman, University of Oxford. This procedure was followed by a Factor Xa cleavage reaction and subsequent purification by reverse-phase HPLC outlined in the Chapter 2 - Materials and Methods (final reverse-phase HPLC shown in Figure 3.16), with the final product shown in Figure 3.15.

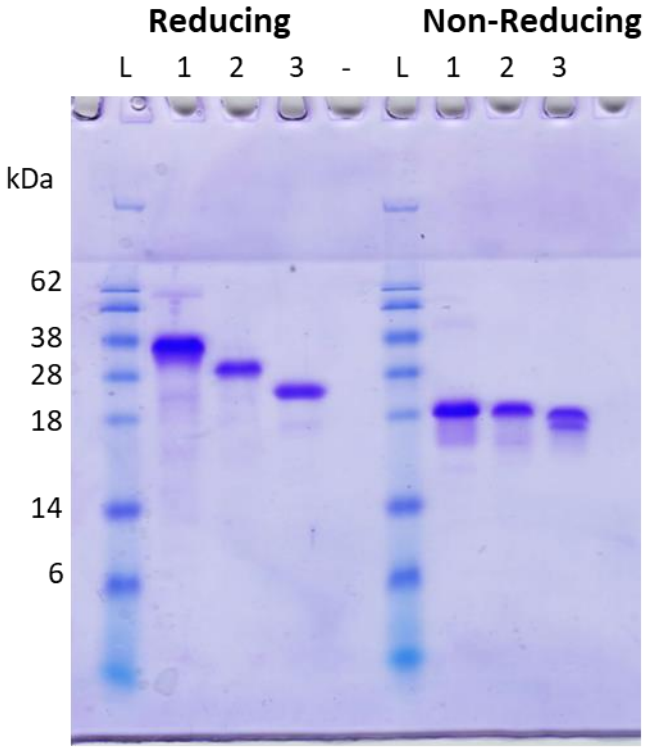


Figure 3.15 Purified PR hN1 Abruptex constructs: Following purification, samples were separated on a 16% SDS-PAGE under reducing and non-reducing conditions stained with Coomassie blue. L) Pre-stained ladder (NEB), 1) PR hN1 23-27CSCS, 2) PR hN1 23-27CSCScl, PR hN2 23-27cl. The bands run at a slightly greater MW than expected when inferring MWs using ladders, this is most likely due to the presence of a tags at the N and C-terminus.

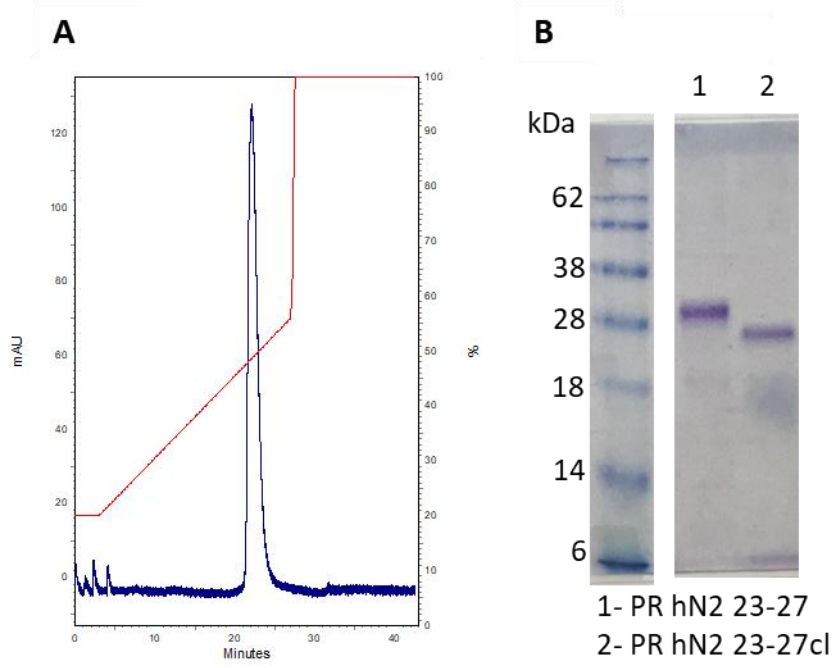


Figure 3.16 Assessment of purity of PR hN2 23-27cl: (A) Final reverse-phase HPLC carried out after Factor Xa cleavage. A_{280nm} absorbance is shown (blue trace), with elution gradient of hydrophobic solvent (red trace). (B) PR hN2 23-27 (before Factor Xa cleavage), and PR hN2 23-27cl (after cleavage) are shown on a 16% SDS-PAGE under reducing conditions and stained with Coomassie blue.

3.3.5. Purified 23-27 constructs

Final products of hN1 23-27CSCS and hN2 23-27 purifications are shown in analysis by 16% SDS-PAGE (Figure 3.15). The SDS-PAGE gels indicate samples are of high purity and run with >95% of protein in a single band under reducing conditions. The presence of a minor additional band in PR hN2 23-27cl when run on SDS-PAGE under non-reducing conditions possibly indicates a small fraction of the sample has a different disulphide bonding pattern.

3.3.6. Electron spray ionization mass spectrometry of PR constructs

Electron spray ionization mass spectrometry (ESI-MS) was carried out by Dr D Staunton, University of Oxford. Proteins purified from the prokaryotic/refold system can be assessed by this method due to their homogeneity and lack of glycosylation. Figure 3.17 shows the expected mass for all PR-expressed proteins matches well with the mass detected in ESI-MS, considering that the uncertainty in these measurements is around 0.01 % (Sharma et al, 1997). This confirms that the PR 23-27 constructs running at a higher mass than expected in 16 % SDS-PAGE (Figure 3.15) is not representative of MW.

Construct	**$MW_{Theoretical}$**	**$MW_{Experimental}$**
hN1 23-24^	12444	12443
hN1 23-24 T940R	12338	12340
hN1 23-24 H917W	12332	12334
hN1 23-24 D932R	12324	12326
hN1 23-24 L934R	12326	12328
hN1 23-27CSCS	24145	24144
hN2 23-27	24300	24298

^^{15}N enriched

Figure 3.17 Results of ESI-MS used to accurately determine mass of prokaryotic/refold constructs: ESI-MS was carried out to determine the true MWs of the purified PR Ax constructs. $MW_{calculated}$ represents theoretical MWs, MW_{ESI-MS} represents experimentally determined MWs. These determined MWs match expected MWs suggesting all constructs were expressed as expected.

3.4. Biotinylation and cleavage reactions for preparation of protein for use in assays

3.4.1. Factor Xa cleavage reactions

Factor Xa cleavage was carried out on PR hN1 23-27CSCS as described in Chapter 2 - Materials and Methods in order to cleave the N-terminal His-tag from this construct. Figure 3.18 shows the reverse-phase HPLC following the cleavage reaction.

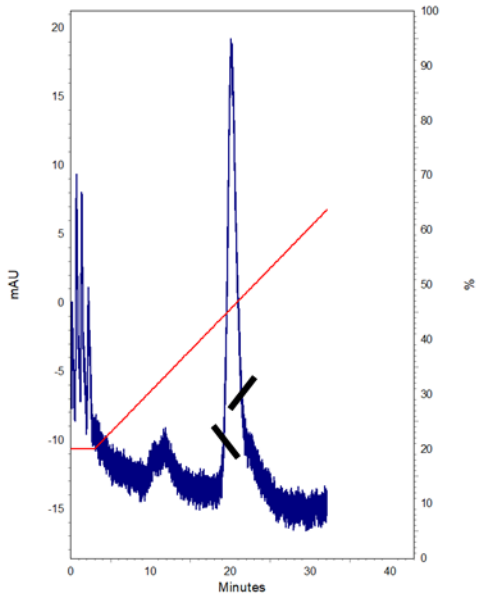


Figure 3.18 Reverse-phase HPLC trace following Factor Xa cleavage of PR hN1 23-27CSCS to produce PR hN1 23-27CSCScl: A_{280nm} absorbance is shown (blue trace) with elution gradient with hydrophobic solvent (red trace). Black bars correspond to fraction collected for further experiments.

Factor Xa cleavage was carried out on PR hN1 23-24 as described in Chapter 2 - Materials and Methods in order to cleave the N-terminal His-tag from this construct. Figure 3.19 shows the reverse-phase HPLC following the cleavage reaction. A band-shift demonstrating that

cleavage has occurred is shown in Chapter 5 - Interaction of the Abruptex region with the ligand-binding region, in the western blot presented with the flow cytometry data corresponding to this construct.

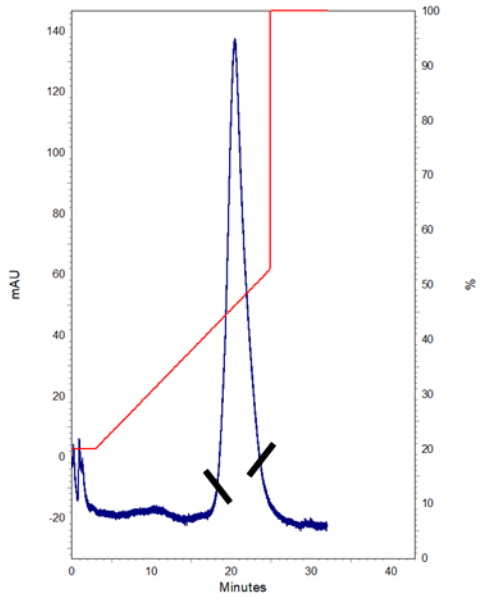


Figure 3.19: Reverse-phase HPLC trace following Factor Xa cleavage of PR hN1 23-24 to produce PR hN1 23-24cl: A_{280nm} absorbance is shown (blue trace) with elution gradient with hydrophobic solvent (red trace). Black bars correspond to fraction collected for further experiments.

3.4.2. Biotinylation of LBR and Ax constructs for binding assays

Biotinylation was possible because of the inclusion of a C-terminal BirA tag in both PR and S2-expressed constructs. This addition of biotin by biotin ligase enabled the clustering these constructs, either onto an avidin-coated fluorescent bead in the case of the flow cytometry assay, or onto neutravidin-HRP in the case of the plate binding assay. In each case, binding and equivalent loading for comparison between constructs depends on successful

biotinylation. An excess of biotinylated component was used in every clustering experiment, meaning that equal concentrations of streptavidin-coated bead component ensured parity between samples. Biotinylation was carried out following the procedure described in Chapter 2 - Materials and Methods. The reverse phase HPLC step following the biotinylation reaction gives an indication of the relative amounts of biotinylated vs unbiotinylated product (see Figure 3.20). Biotinylated fractions were determined by western blot, probing with a streptavidin-peroxidase conjugate (See Chapter 2 – Materials and Methods). The biotinylated fraction was then taken forward for use in binding assays. Biotinylation was undertaken for all S2 hN1 11-13 variants for use in the plate assay, and all hN1 23-24 Ax constructs for use in the flow cytometry assay. Figure 3.20 describes the biotinylation purification for S2 hN1 11-13 Q462R prep 2, which was typical of all biotinylation reactions carried out.

Biotinylation of S2 hN1 11-13 variants for plate assays was undertaken following the procedure outlined in Chapter 2 – Materials and Methods. The final purified products were blotted with a streptavidin-conjugate and shown in Figure 3.21. Near equivalent biotinylation is observed across all variants and preps of S2 hN1 11-13 variants, validating that these biotinylated constructs can be fairly compared for investigating binding to Ax in the plate assay. An additional minor contaminating band is seen (Figure 3.21), which can be biotinylated and possibly corresponding to a an 11-13 dimer. However, this contaminant may also be another protein which shows up in streptavidin blots because it is preferentially biotinylated during in vitro biotinylation.

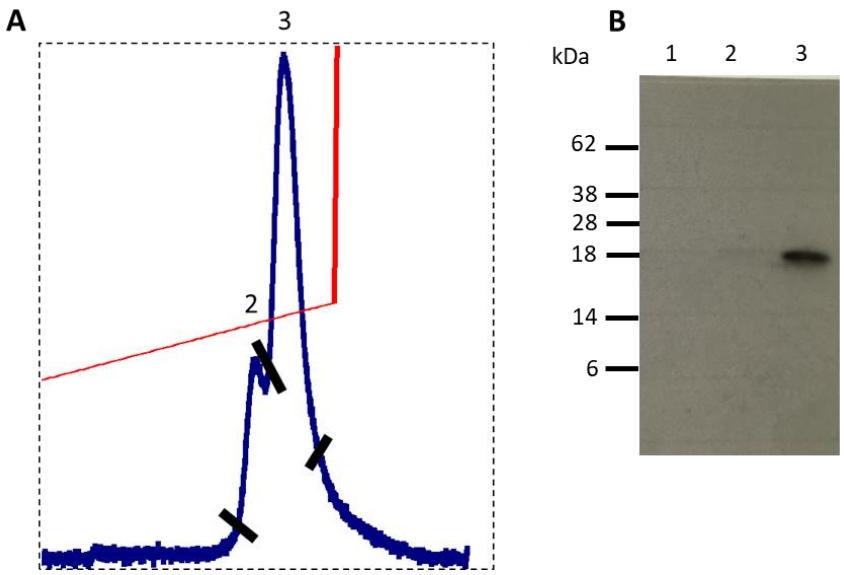


Figure 3.20 Example of biotinylation of Notch constructs for use in assays: The biotinylation of S2 hN1 11-13 Q462R prep 2 is shown and is representative of a typical biotinylation purification, (A) the A_{280nm} absorbance trace (blue) is shown from reverse phase HPLC following the biotinylation reaction. Elution with hydrophobic gradient (red) results in two peaks being eluted. Black bars correspond to fraction collected for western blot. (B) Western blot following 16 % SDS-PAGE of pre-biotinylation material, and the peaks eluted in A, 1) Purified S2 hN1 11-13 Q462R prep 2 (pre-biotinylation), 2) Peak 2 in trace, 3) Peak 3 in trace. Western blot was probed with a streptavidin-conjugate and suggests that peak 3 corresponds to biotinylated product, with no binding seen for pre-biotinylation sample and peak 2. MW matches MW expected from SDS-PAGE analysis of purified proteins (~17000).

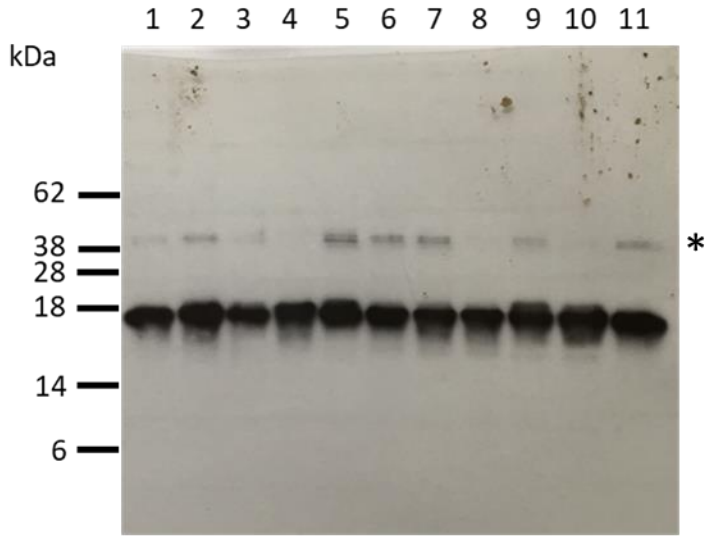


Figure 3.21 Western blot with streptavidin conjugate following 16 % SDS-PAGE shows equivalence of biotinylation across all S2 hN1 11-13 variants: Following biotinylation reactions and purification of product by reverse phase HPLC, a western blot was carried out following a reducing 16 % SDS-PAGE and probed with a streptavidin conjugate to show biotinylation when adding an ECL substrate before exposure to a chemiluminescent film. 1) S2 hN1 11-13, 2) S2 hN1 11-13 A426R, 3) S2 hN1 11-13 K428E, 4) S2 hN1 11-13 E437R, 5) S2 hN1 11-13 R448E, 6) S2 hN1 11-13 N459R, 7) S2 hN1 11-13 Q462R prep 1, 8) S2 hN1 11-13 Q426R prep 2, 9) S2 hN1 11-13 L468R prep 1, 10) S2 hN1 11-13 L468R prep 2, 11) S2 hN1 11-13 V485R. MW of major band matches expected MW of ~17000. The additional minor band at ~38 kDa (*) is an unknown minor contaminant which may be a dimer of hN1 11-13.

Biotinylation of Abruptex constructs was carried out for flow cytometry assays as described in Chapter 2 – Materials and Methods. Fluorescent bead-biotinylated protein complexes were run on a reducing 16% SDS-PAGE gel and blotted with streptavidin to show equivalent bead loading. These results are shown alongside the results of the experiment they refer to in Chapter 5 - Interaction of the Abruptex region with the ligand-binding region.

3.5. Investigating Ca^{2+} affinity and aromatic stacking in hN1 EGF24

3.5.1. Characterization of Ca^{2+} affinity of EGF24 by NMR

An NMR investigation with PR hN1 23-24 was undertaken to determine the affinity of the proposed hN1 EGF24 Ca^{2+} binding site, inferred from bioinformatics analysis (Figure 3.2). The N-terminal domain in this construct, EGF23 should have a weak affinity for Ca^{2+}, but EGF24 in the native-like context should have a higher affinity (See Chapter 1 – Introduction, Section 1.3.3). The PR expression system was chosen as this construct could be enriched with ^{15}N to facilitate 2D NMR experiments for monitoring this Ca^{2+} titration in ^{1}H-^{15}N heteronuclear single quantum coherence spectroscopy (HSQC). Previous studies to determine K_d of EGF23-24 binding to Ca^{2+} had been carried out at pH9, I = 0.15 (Barber and Redfield, unpublished data). This non-physiological pH was previously used due to the propensity for the hN1 23-24 construct to precipitate at near-neutral pHs. Here, the experiment was conducted at the more physiological pH7.5, I = 0.15. NMR Ca^{2+} titrations were performed with a ^{1}H-^{15}N HSQC recorded after each [Ca^{2+}] addition and enabled monitoring of peak shifts and plotting of a Ca^{2+} binding curve to saturation, (Figure 3.22). The NMR spectrum indicated a folded protein. This preparation of PR hN1 23-24, and a lowering of protein concentration to ~100 µM avoided protein precipitation. Following preliminary trials, Ca^{2+} was added in stepwise aliquots to a protein sample of ~100 µM with HSQCs recorded for $[Ca^{2+}]_{total}$ up to 10 mM. Data analysis suggested a K_d estimated to be 6.3 ± 0.9 µM (Figure 3.22). However, the first addition of Ca^{2+} saturated around 80% of hN1 23-24, suggesting more precise experiments

could be performed using the chromophoric chelator 5,5'-Br_2BAPTA (Jensen et al, 2005). Attempting to make smaller additions of Ca^{2+} in a similar NMR titration would be challenging considering the likely presence of residual Ca^{2+} on the surface of NMR tubes, and NMR buffers used for the experiment, which would partially saturate such a high-affinity site before any Ca^{2+} additions. The equations used to calculate K_d are described in Chapter 2 - Materials and methods, section 2.4.2.1, and are outlined in Jensen et al, 2005. This result supports previous evidence from NMR obtained at pH9 of a high affinity Ca^{2+} binding site in EGF24 and is in line with bioinformatics predictions. Despite any uncertainties associated with the lack of partially saturated points in the titration curve, a site with this affinity would be saturated at the cell surface where extracellular $[Ca^{2+}]_{free}$ is ~1.4 mM assuming pH ~7.4, (Weisshuhn et al, 2016) and suggest EGF23-24 adopts a rod-shape conformation typical of cbEGF pairs when saturated with Ca^{2+}.

3.5.2. The unusual intradomain aromatic in EGF24

The Ca^{2+} titration experiments demonstrate that EGF24 binds to Ca^{2+} with moderately high affinity. The Ca^{2+} consensus packing aromatic in this cbEGF is atypical due to a positional shift before Cys_4 by one position (see Chapter 1 - Introduction, and Figure 3.2 for details of the Ca^{2+} consensus aromatic). Every other Ca^{2+}-binding consensus region throughout hN1, hN2 and dN is typical (see Figure 3.2), however mN1 contains a similar atypical consensus in EGF24. Comparing the position of Ca^{2+} ligands in this atypical consensus domain to a typical consensus shows that the Ca^{2+} ligands remain in a typical arrangement. This is partly due to

the presence of an additional aromatic towards the domain C-terminus which packs into the space left by the shifted packing aromatic (Figure 3.23). This additional aromatic is also conserved in mN1 EGF24 and appears to normalize the position of the Ca^{2+} ligands, likely explaining the moderately high affinity of this site. This further supports that the EGF23-24 interface is a rigid-rod in the context of the receptor under physiological conditions in which this site will be saturated.

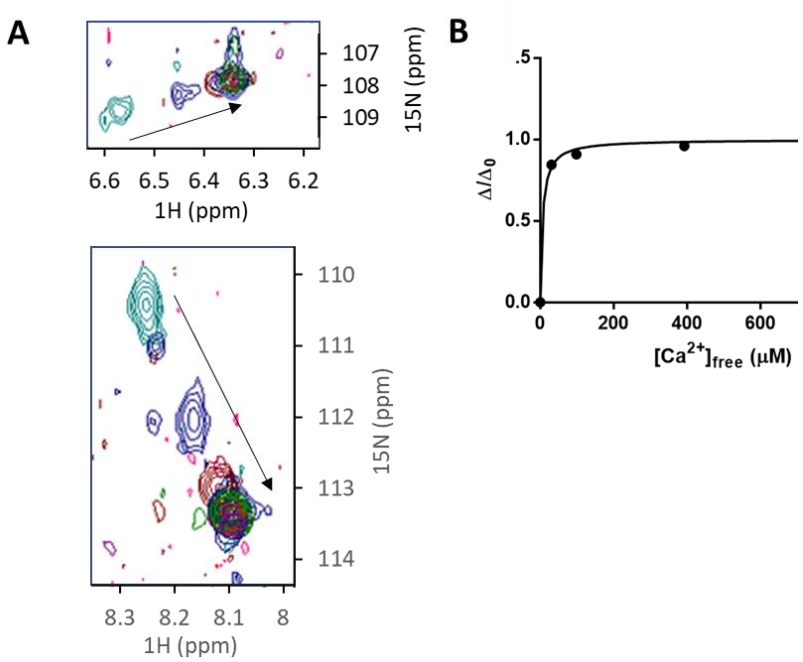
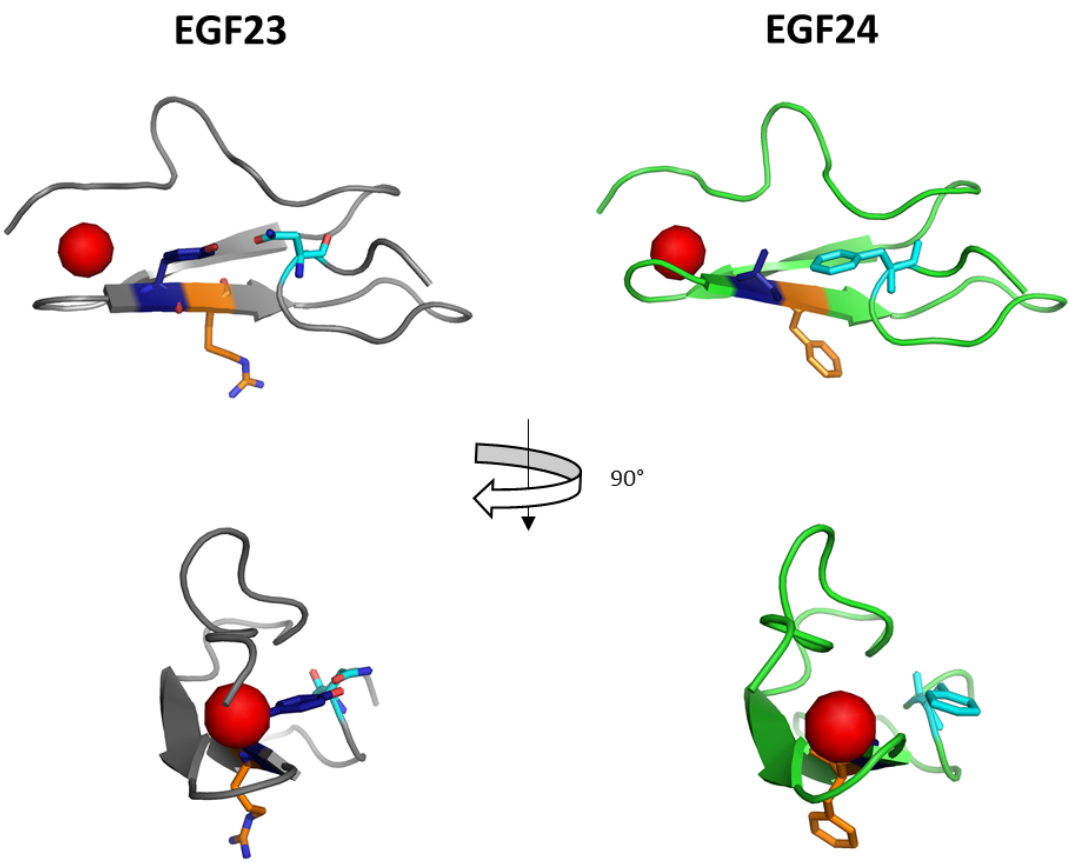


Figure 3.22: Quantification of the Ca^{2+} binding site of hN1 EGF24 at pH7.5 in the presence of 150 mM NaCl. Six peak shifts were followed on ^{1}H-^{15}N HSQCs, two of these peak shifts are shown in (A), arrows corresponding to direction of shifts (peaks are shown as cyan, blue, maroon, green, purple, and navy as $[Ca^{2+}]$ increases). B) K_d is calculated to be 6.3 ± 0.9 µM. Δ = peak shift from [Ca] = 0 to a defined [Ca], Δ_0 = maximal peak shift (See Chapter 2 - Materials and Methods). Regressions were carried out using GraphPad Prism 7 (GraphPad, San Diego, CA, USA).




```
EGF23 DINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCET
                                 **          |
EGF24 DIDDCRPNPCHNGGSCTDGINTAFCDCLPGFRGTFCEE
```

Figure 3.23 Comparison of typical and atypical hN1 EGF24 Ca^{2+}-consensus: A typical cbEGF is shown (hN1 EGF23 grey) and the atypical hN1 EGF24 is shown (green) with a bound Ca^{2+} ion (red sphere). Positions before Cys_4 with the shifted aromatic are shown in navy and orange and highlighted by * in the sequence. The position of the canonical aromatic in EGF24 is a small residue (Alanine, navy), which fills a similar space to the outwardly-pointing residue in EGF23 (navy). Interestingly, in order to fill the space created by the absent aromatic, an additional aromatic in EGF24 (cyan) points back into the core of the domain. This is fulfilled by a non-canonical aromatic in EGF24 (cyan, designated by | in the sequence).

3.6. Biochemical studies of the region

3.6.1. Investigating the non-canonical cysteines in the hN1 Abruptex region

The additional cysteines in hN1 EGF25 (Cys963) and EGF27 (Cys1040) (Figure 3.24) are an interesting feature because they are not conserved in hN2 or dN. Modelling these cysteines onto a similar region of the receptor which has been structurally determined shows that these two residues appear to be on the same face of the molecule. This indicates that an intramolecular disulphide bond involving these cysteines would be possible if there was flexibility or bending in the inter-domain hinges between EGF25 and EGF27 (Figure 3.24). However, if the region was a rigid rod as seen in the structure of the homology model an intra-molecular disulphide would not be possible. hN1 EGF21-23 was chosen for the model to predict the positions of the additional cysteines, as it has a similar domain organization (cbEGF-EGF-cbEGF) to EGF25-27 and a crystal structure exists for this fragment.

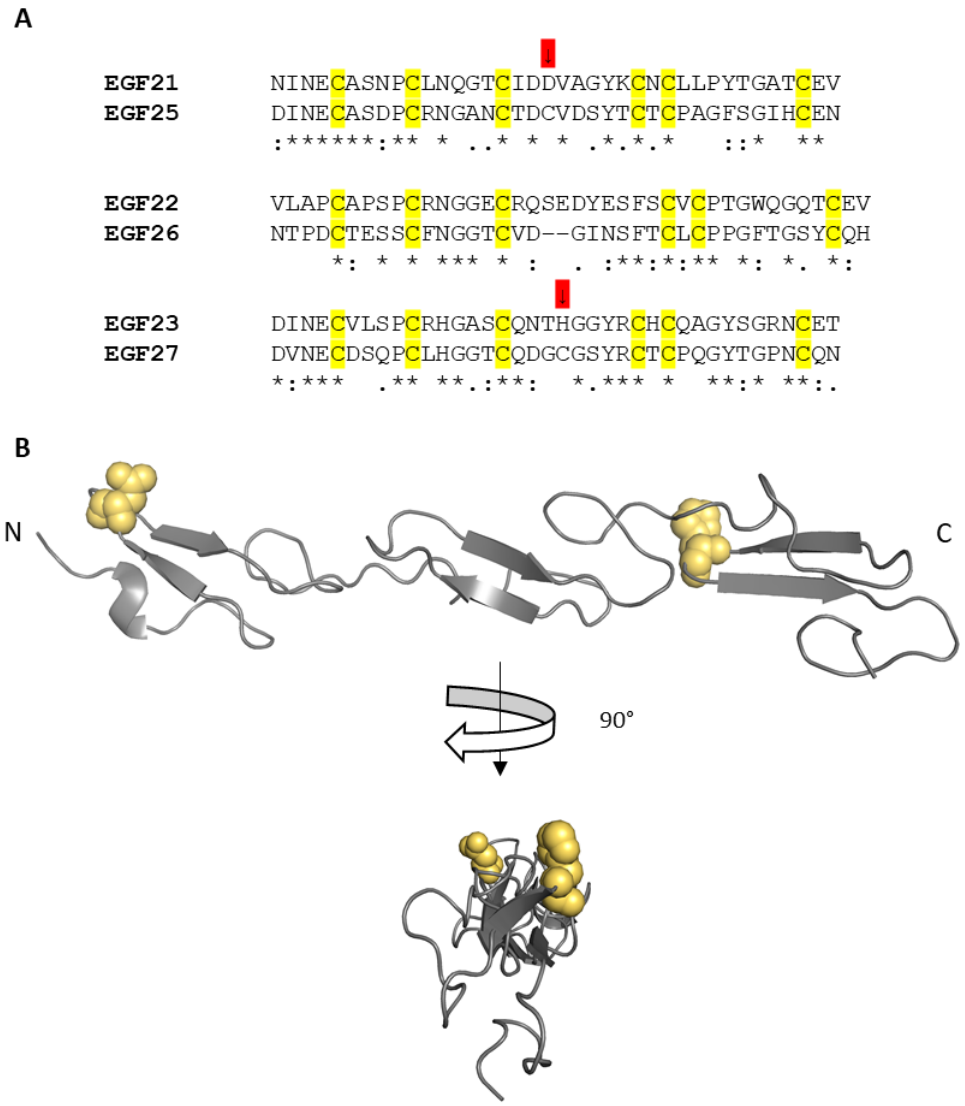


Figure 3.24 Modelling the additional non-canonical cysteines in hN1 onto a known structure: (A) The sequence of the region hN1 25-27 was aligned with a similar region (hN1 21-23) using sequence alignment software Clustal Omega 1.2.4 (Sievers et al, 2011, Li et al, 2015, McWilliam et al, 2013). The canonical cysteines are highlighted in yellow, the additional non-canonical cysteines are indicated by red arrows. Positions which are exactly conserved are indicated by '*', well conserved ':', and partially conserved '.'. (B) The residues in EGF21-23 corresponding to the positions of the additional cysteines in EGF25 and 27 are shown as yellow spheres. The structure of hN1 21-23 is a component of hN1 20-24 as determined by Suckling et al, unpublished data. N- and C-termini of the structure are indicated.

In order to further investigate the oxidation state and exposure of these additional cysteines in hN1 EGF25 and EGF27, a DTNB (5,5'-dithio-bis-(2-nitrobenzoic acid)) assay was carried out using Ellman's reagent on S2 hN1 23-29. This assay takes advantage of the ability of DTNB to bind to free thiol groups on exposed cysteines, and the subsequent absorbance of the product at 412 nm to quantify surface exposed reduced thiols. The results of this assay are displayed in Figure 3.25.

	A_{412} Experimental	**A_{412}** Theoretical	**Calculated concentration of free thiols (µM)**
10 µM Cysteine (-SH)	0.163	0.142	11.5
20 µM Cysteine (-SH)	0.276	0.283	23
10 µM S2 hN1 23-29	0.047	-	3.3

Theoretical readings were calculated using ε_{412nm} = 14,150 M^{-1} cm^{-1}

Figure 3.25 DTNB assay to investigate the reduction-state of the thiols in the hN1 Ax region: The assay was carried out in 6M GuHCl at pH8.0. A Panel of substrates for the DTNB assay, including Cys solutions at defined concentrations to confirm absorbance measurements, 10 µM and 20 µM Cys were used to confirm absorbance by this method in the assay as 10 µM hN1 23-29 construct is expected to have a maximum of 20 µM free thiols. Experimental and theoretical absorbances are not identical but similar enough to justify the use of this assay to give an insight into the free thiols in a denatured protein sample. The protein sample was prepared by treatment with 100 mM DTT in the presence of 6M GuHCl, followed by purification of the protein from the reducing buffer by reverse phase HPLC in 0.1% TFA.

The DTNB assay results in Figure 3.25 for S2 hN1 23-29 indicate 3.3 µM reduced cysteines per 10 µM protein, suggesting that there are on average, one free thiol for every three protein molecules. This suggests that the majority of thiols, including the non-canonical cysteines in EGF25 and EGF27 are likely shielded from interacting with the DTNB, either as

they are in a disulphide bond, or shielded from interaction with DTNB. This small number of free thiols may be due to a small proportion of the protein being misfolded with multiple free thiols in solution, possibly corresponding to the multiple bands observed for this construct on a non-reducing SDS-PAGE (Figure 3.9). It is unlikely that these thiols are modified with any other molecules considering that these cysteines are not conserved in other receptors.

4.

Low Resolution Structural Studies of the Abruptex region

4.1. Understanding receptor function in terms of shape

The Notch receptor is normally represented as a rod projecting away from the plasma membrane, but the true macromolecular structure of the receptor at the cell surface is unknown. A molecular dissection approach has been previously undertaken to investigate the structure of small sections of the receptor using a variety of structural techniques including NMR spectroscopy and X-ray crystallography (Weisshuhn et al, 2016). This method involves a combination of determining high-resolution structures by X-ray crystallography and investigating any flexibility by NMR. X-ray crystallographic approaches have so far proved unable to determine the structure of the Abruptex region. Furthermore, the narrow dimensions of EGF chains produce very low resolution structures by cryo-electron

microscopy (Kelly et al, 2010) (See Chapter 1 – Introduction, Section 1.4.1). In addition to obtaining evidence for the flexibility directly using a combination of these methods, the rod-like structure of an EGF interface can be inferred from bioinformatics analysis by the presence of conserved features such as the interdomain packing residues or calcium-binding consensus residues (See Chapter 1 – Introduction, Section 1.3). Ca^{2+} titration experiments to determine the affinity of the ion binding to the consensus sequence in a particular cbEGF can suggest if the domain will be Ca^{2+}-bound or unbound under physiological conditions. Considering previous evidence that Ca^{2+} binding leads to the rigidification of an interface (Weisshuhn et al, 2016) these affinity studies suggest if an interface is likely to be rigid or flexible under physiological conditions. NMR spectroscopy can therefore be used to determine Ca^{2+} affinity at the interfaces and to assess intrinsic flexibility. This method has produced a model for hN1 EGF4-13, which shows a non-rod like conformation with a likely flexible hinge at the EGF9-10 interface (See Chapter 1 – Introduction, Section 1.4.1). Furthermore, previous crystal structures of various multiple tandem EGF-domain constructs have revealed the possible conformations of adjacent cbEGF/EGF domains (Suckling et al, unpublished data, Weisshuhn et al, 2016, See Chapter 1 – Introduction, Figure 1.9).

Following the unsuccessful attempts to determine the structure of the Abruptex region (EGF24-29) of hN1 by X-ray crystallography and unsuccessful studies on similar regions of the receptor using electron microscopy, small angle X-ray scattering was used to determine a structural model of the region in solution. The details of the unsuccessful electron microscopy experiment are outlined in the Appendix.

4.2. Small angle X-ray scattering

Small angle X-ray scattering (SAXS) is a technique which gives low-moderate resolution (~10 Å) structural insights into a macromolecule in solution. Although this method yields a lower resolution structure than crystallography, the sample does not have to be grown in a crystal, which is a significant advantage for samples which have proven intractable to crystallography or electron microscopy, in order to obtain important moderate resolution structural data. Electron density is not directly observed, but rather an average shape of all objects in the solution, so it is essential that the sample is pure when data are collected.

4.2.1. Theory

SAXS involves exposing a sample in solution to an X-ray beam and collecting a scattering pattern. From this pattern, information relating to molecular dimensions, proportions and shape of the objects causing the scattering can be determined. In the case of multiple tandem EGF domains, the structure of each individual EGF is highly conserved with predictable dimensions, SAXS is able to provide information on how these domains link together and therefore the overall shape of the chain. Due to the random orientation of the proteins in solution, the scattering pattern is radially symmetric, and is described as a radial distribution curve of intensity known as a scatter curve (Figure 4.1). The gradient and form of various regions of the scattering curve give insight into the dimension, size and volume of the object which is scattering the X-ray beam. The scattering curve can be transformed into real space

to give a pair-distance distribution function ($P_{(r)}$) from which a model of the shape of the protein in solution can be determined. Another parameter which is derived from the $P_{(r)}$ is the maximum distance between any two points in the molecule, the maximum dimension of the protein (D_{max}). Multiple models can be produced, averaged, and then used to back-calculate a theoretical scattering curve to fit to the original data to assess the validity of the final model.

4.2.2. SEC-SAXS

Protein SAXS is routinely carried out with inline gel-filtration immediately preceding X-ray scattering (Figure 4.1), this ensures that the sample passing through the X-ray beam at any one time is highly likely to be homogenous, having been separated by size on the column. The alternative to SEC-SAXS is to directly subtract a buffer sample from a protein sample known as BioSAXS. BioSAXS relies on the protein sample being homogenous without any contamination. Although these samples are highly purified (See Chapter 3 – Protein Purification and Characterisation), to ensure homogeneity of the sample we conducted SEC-SAXS with a high-performance analytical gel filtration column suitable for resolving proteins over the suitable molecular weight range of our sample, monitoring the elution profile for scattering and absorbance at 280 nm (A_{280}) to identify the protein elution peak.

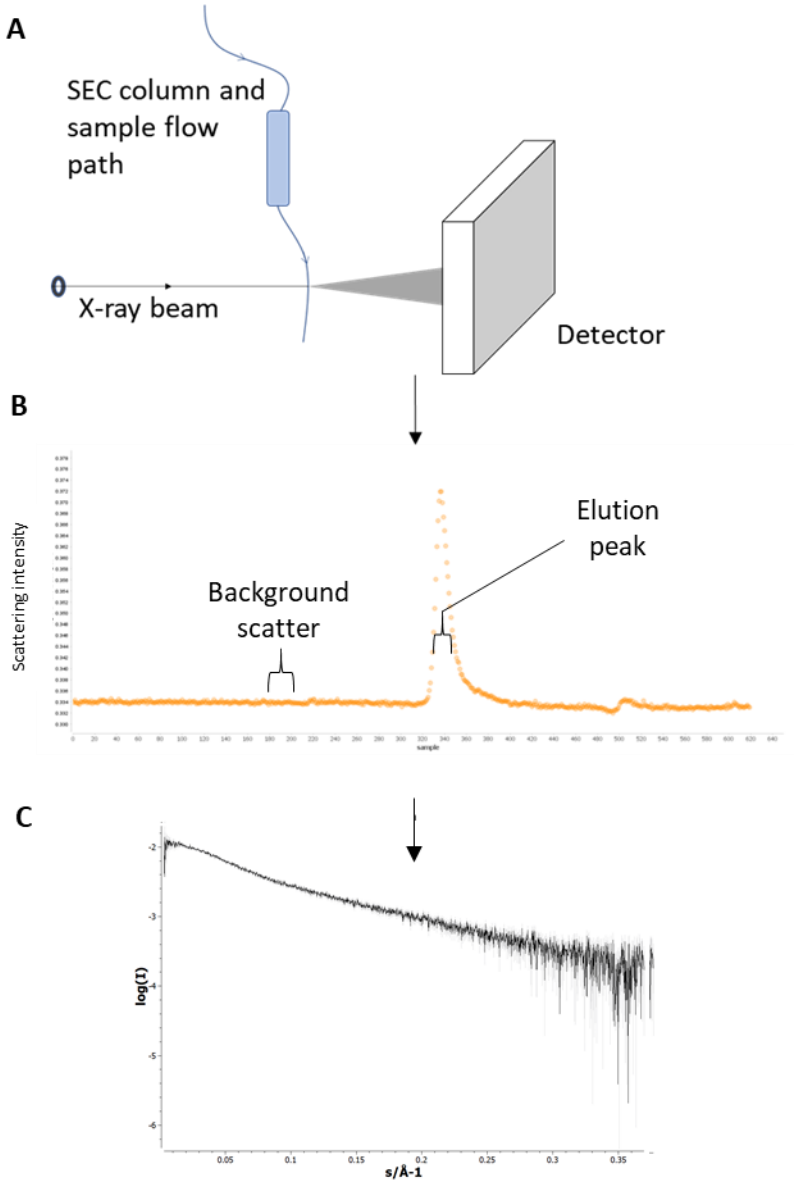


Figure 4.1 Schematic of SEC-SAXS data collection: (A) Experimental set up involves separation of sample on a Shodex 402.5 analytical size exclusion chromatography (SEC) column. Sample is then flowed through the X-ray beam, with scattering pattern collected over the duration of the SEC run. (B) Scattering pattern is collected over the elution from SEC column. (C) Subtraction of background scattering from elution peak scattering yields a scatter curve, which can be used to derive information on molecular dimensions.

4.2.3. Development of a pipeline for processing data

The different types of EGF domain interfaces (See Chapter 1 – Introduction, Section 1.3.5.1) mean that a large variety of structures are possible within a contiguous set of domains, for example linear or bent structures. The aim of this SAXS analysis is to determine the shape of the protein under analysis in solution. From this, novel information about the shape of the Abruptex region can be derived. The data processing pipeline is outlined in Figure 4.2.

4.2.3.1. Data quality screen

Initial SEC-SAXS data collection involves monitoring the A_{280} for the sample running through the column and eluting into the X-ray beamline, and concomitant collection of the total scattering intensity throughout the run. A successful sample run will show a single defined peak in the A_{280} absorption profile and scattering spectra. This allows an accurate buffer subtraction to be carried out, crucially, choosing the middle of the elution peak, avoiding the edges of the peak to minimize heterogeneity in the scatter frames averaged to determine the final scatter curve. All the samples run through SEC-SAXS described in this chapter eluted as a single sharp peak from the column, attesting to the purity of the proteins used for analysis, also inferred from SDS-PAGE analysis (See Chapter 3 - Protein Purification and Characterization). The scatter curve is then inspected to determine if the small angle intensities are too noisy to fit a precise curve.

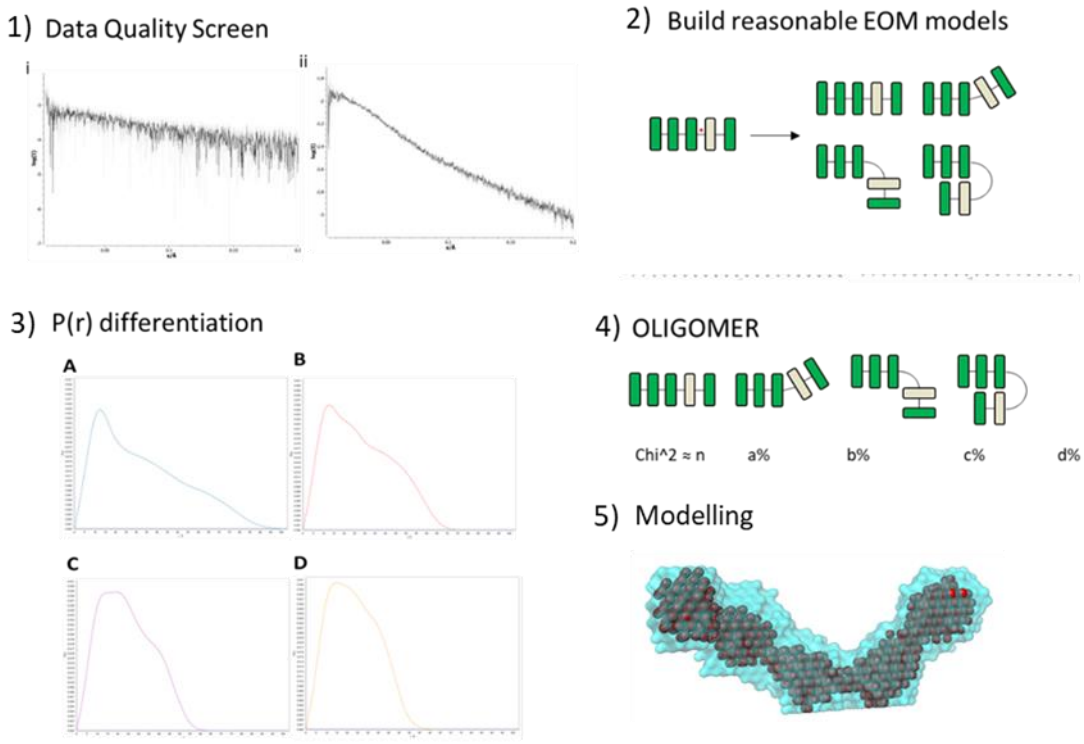


Figure 4.2 Pipeline for data assessment and processing for elongated multiple tandem-domain chain constructs: 1) Data Quality Screen: a clear single peak for absorbance at 280nm and scatter should be seen during column elution into X-ray beam (axes x = s/Å, y = Log(I)). The scatter curve should also have low noise at low angles (ii) as opposed to high noise (i). 2) The Ensemble Optimisation Modelling program (EOM) (Tria et al, 2015) is used to generate a number of feasible models based on predicted or observed sites of flexibility and produces a range of possible models. Four models are selected to include a range of flexibility. 3) The P(r) distributions of these models are assessed to confirm that the structures can be differentiated based on dimension factors observed from SAXS (axes x = r/Å, y = P(r)). 4) The program OLIGOMER (Svergun et al) is then used to look at the %-mix of the models which fit the data, and how well, ensuring that structure of the protein is well defined (a+b+c+d = 100, n ≈ 1). 5) Finally, a P(r) distribution for the data is generated and 10 models produced from the data using DAMMIF (Franke and Svergun, 2009). DAMAVER (Volkov and Svergun, 2003) is then used to produce an average structure from these. Following this, the best fit model is fitted to the original data (Chi^2 determined) using the program CRYSOL (Svergun et al, 1995) to ensure the data fit the model well.

4.2.3.2. Building reasonable EOM models

Using data from bioinformatics and Ca^{2+} titration assays, as well as any high-resolution structures within the region of interest, an ensemble of feasible models are produced. These models have constraints introduced, for example there may be previous studies that show a particular inter-domain interface is not rod-like, and so flexibility would be allowed here when generating a range of models, however there is little information from previous studies on the Abruptex region. These models are built using the Ensemble Optimization Modelling program within the ATSAS suite (See Chapter 2 – Materials and Methods). When generating models in EOM, in the case of hN1 11-13, the structure of this region was used to generate models. As no high resolution structures of the Abruptex region have been determined, the region hN1 20-24 was used to generate EOM models for EGF23-27 . The rationale for this was that EGF-domains are likely to appear similarly, independent of sequence, at the resolution of SAXS and that a range of structures representative of possible bends can be produced from this similar construct of five EGF-domain length. Future analyses could use homology models of the EGF23-27, which would presumably provide a more detailed range for possible structures of this region. Four models are selected from a pool generated using this program, representing a range of angles at a particular interface. These models are then fitted in linear combinations using the program OLIGOMER to the scatter curve to determine which weighted combination best fits the curve. The four models selected encompass a large range of structures to ensure that OLIGOMER suggests a similar result to the independently calculated DAMMIF model. A greater number of models at smaller increments of bentness

could be used in future studies to further support that the DAMMIF model was supported by the OLIGOMER prediction.

4.2.3.3. P(r) differentiation

The P(r) distribution is a frequency distribution plot of the distances between all pairs of electrons within the molecule. This leads to characteristic curves for different macromolecular shapes (see Figure 4.3). The models generated using EOM may only a have a marginally different maximum dimension (D_{max}), despite a different overall shape. Comparing P(r) distributions is more informative than comparing D_{max} alone, although looking at single dimensions alone (such as D_{max}) is commonly done (Baldock et al, 2006). The P(r) can be determined from the scatter curve, and this P(r) should be compared with the P(r) distributions of the EOM models to check that the data match the shapes suggested from the output of OLIGOMER. This also gives a qualitative insight into the shape of the molecule, for example extended or more compact (Figure 4.3).

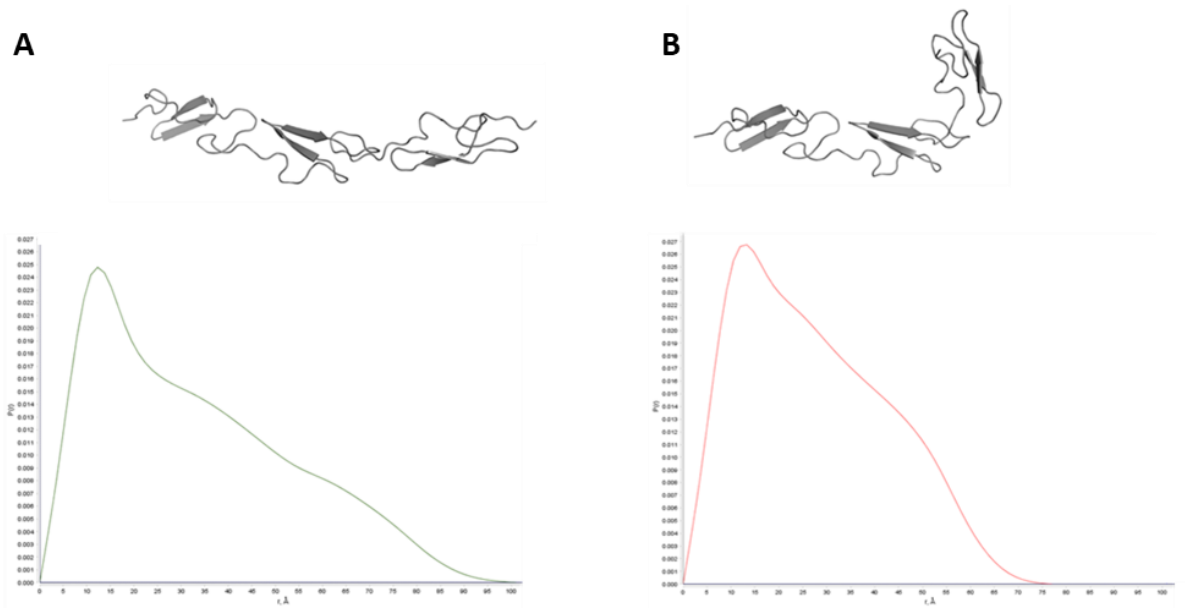


Figure 4.3 The relationship between structure and P(r): Two EGF structures are shown here (A) and (B), each of three domains. A) shows a straight model (derived from hN1 11-13 PDB ID: 2VJ3) an extended rod-like structure, for which the P(r) distribution is a sharp peak at short distances and then a smooth decrease in frequency to the largest intradomain distance (D_{max}). B) shows a more compact, bent model, generated with EOM with a smaller D_{max}. The P(r) distribution contains more area at short distances and a sharper drop off to smaller D_{max}. Axis labels: y-axis (P(r)), x-axis (r, Å)

4.2.3.4. OLIGOMER

OLIGOMER is a program for fitting an ensemble of structures by linear combination to form a best-fit average curve. This allows determination of whether the protein is present as a range of conformations in solution and is flexible. This can also determine if a protein is likely to be relatively rigid in solution and validate the use of a single model for the molecule in solution. OLIGOMER calculates the best fit combination of models and represents the weighted average as percentages of the different models representing what would be present in solution. The observation that the scatter curve can be explained by a single or

two similar models rather than a broad range of conformations suggests that the protein has a relatively rigid structure in solution.

4.2.3.5. Generating models

After concluding that the protein scatter can be explained by a relatively homogenous conformation, or a relatively fixed structure in which structural changes cannot be detected at the resolution of SAXS, a model for the structure in solution can be produced from the P(r) distribution using the program DAMMIF. 10 models are produced and then averaged to produce a final combined model which is then compared to the initial data using the program CRYSOL. This program produces a theoretical scatter curve from the final structure and then compares this to the experimental scatter curve and produces a Chi^2 value corresponding to how well the data fits the structure. A Chi^2 value close to 1 reflects a good fit, values greater than 1 indicate a less-good fit, and values less than 1 indicate that the data may be over-fitted. A good fit is defined to be between 0.8 – 1.2, considering the large Chi^2 value seen for a fit of the incorrect bent model in Figure 4.4, and Chi^2s fitted for published findings of similarly sized proteins (Robertson et al, 2017). A flexible protein would produce a scattering pattern from the combination of all possible structures present in solution and no single structure would be able to explain this scattering pattern (Kikhney and Svergun, 2015).

It is important to note that the generation of the models by DAMMIF is dependent on the experimental data only, any assumptions relating to the location of potentially flexible

interdomain hinges used during the generation of EOM models for OLIGOMER analysis are not included in calculating models in DAMMIF.

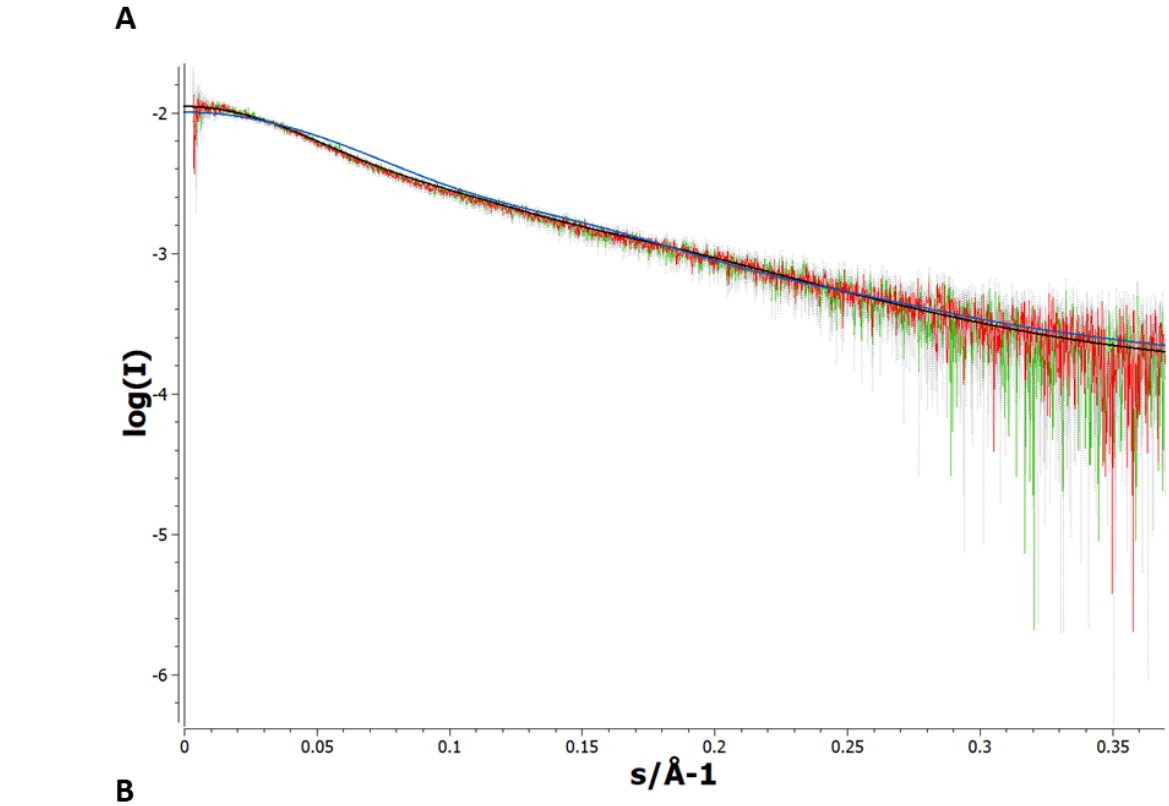
4.2.4. Testing the pipeline with hN1 11-13

hN1 11-13 was expressed and purified from the S2 system as a His-tag fusion protein, and subsequently cleaved (S2 hN1 11-13cl) using Factor 3C protease (see Chapter 3 - Protein Purification and Characterisation, Section 3.2.2). This construct includes domains 11 and 12 of the LBR of the receptor and has been very well characterized structurally (Cordle et al, 2008). The construct has been previously expressed using a prokaryotic/refold system as well as the S2 eukaryotic system (Taylor et al, 2014 , Suckling et al, 2017). S2 expressed hN2 11-13 has been crystallized and contains multiple O-glycosylations (Suckling et al, 2017) and the consensus glycosylation sequences are conserved in hN1 11-13 (EGF12: O-Fuc and O-Gluc, EGF13: O-Fuc). Although glycosylations are flexible, similar SAXS investigations on tandem EGF domains rich in N-linked glycans show that these glycans do not greatly increase the volume of SAXS models expected from protein alone (Troilo et al, 2016). Therefore, the significantly smaller O-glycans present in the constructs investigated here are unlikely to impact SAXS volumes. The Ca^{2+}-binding properties of the prokaryotic expressed and refolded version of this construct have also been well characterized by NMR, but not for the S2 expressed product.

Ten structures of this region exist in the Protein Data Bank (PDB) expressed with a variety of terminal tags, including the corresponding region from hN2 expressed from the identical S2

system, all of which overlay and can be represented by the structure in Figure 4.3A, suggesting that the structure is a rigid rod in solution. This is supported by NMR data from residual dipolar coupling experiments and heteronuclear NOE data suggesting rigidity in solution. This rigidity is Ca^{2+}-dependent, with the Ca^{2+} affinities of EGF11, 12, and 13 in the unmodified construct being 310 µM, 55 µM, and 30 µM respectively (Weisshuhn et al, 2016). Note that the affinity of EGF11 is sub-6 µM in constructs including EGF10, exemplifying the importance of the preceding domain in stabilizing the next EGF domain. Previous NMR studies show that the domains in this construct become more flexible at low Ca^{2+} concentrations when the Ca^{2+}-binding sites are not saturated (Weisshuhn et al, 2016).

Data were collected from two separate SAXS experiments for S2 hN1 11-13 in SAXS buffer (see Chapter 2 - Materials and methods, Section 2.5.1) which overlay very well, and each is able to clearly differentiate between straight and bent models, suggesting that SAXS is able to differentiate between straight and bent multi-domain EGF-fragments (Figure 4.4). Importantly, this demonstrated the reproducibility of the method for determining a scatter curve from the same protein in independent experiments. Although previous studies suggest that all the Ca^{2+} sites are likely to be saturated in the SAXS buffer condition (15 mM Ca^{2+}), EOM was used subsequently to generate a variety of feasible models to re-affirm that our SAXS pipeline could effectively differentiate straight and bent structures and correctly match these in the OLIGOMER step of the pipeline (Figure 4.5).



B

Straight structure



Bent structure



Data set	Straight structure (Chi^2)	Bent Structure (Chi^2)
11-13cl Sep17	0.97	2.38
11-13cl Dec17	0.97	2.07

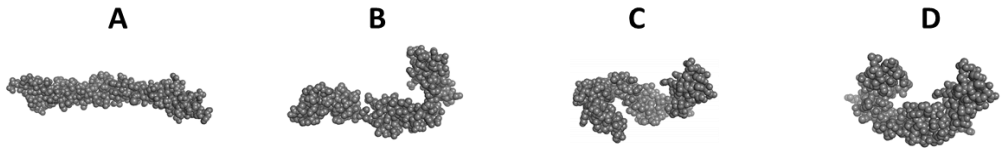
Figure 4.4 Comparison of independent SAXS experiments performed on S2 hN1 11-13cl: (A) scattering curves collected for S2 hN1 11-13cl in September 2017 (red) and December 2017 (green), error bars are shown (grey). The fits of the structure-derived theoretical scatter curves determined with CRYSOL are shown for the straight model (black, derived from PDB ID: 2VJ3) which fits better than the bent model (blue). (B) The Chi^2 values corresponding to these fits using the software CRYSOL. A Chi^2 value of 1 indicates a good fit.

For the two independent SAXS experiments, analysis showed that S2 hN1 11-13 was present as an extended structure; 80-94% linear as indicated by OLIGOMER (Figure 4.5). This agrees with previous structural data that this construct is an extended rod in solution. A greater number of models to represent smaller increments of bentness between models A and B could be used in future experiments to give a more precise insight into the accuracy of OLIGOMER. In this initial study, the four models were able to validate that OLIGOMER suggested a rod-like structure in solution. The variation in the proportions of model A and B in Figure 4.5 predicted in solution gives a useful insight into the resolution of the OLIGOMER fitting, considering that both scatter curves overlay well (Figure 4.4A). OLIGOMER gives a fit to multiple structures slightly less well ($Chi^2 = 1.06$) than the CRYSOL fit to the single structure in Figure 4.4 ($Chi^2 = 0.97$). This small difference in Chi^2 is negligible in terms of model validity as both programs suggest the rod-like model represents the real structure in solution. The cause of this small difference in fit is likely due to fundamental differences in the calculation steps in each program, CRYSOL used in Figure 4.4 takes into account hydration shell to calculate theoretical scattering pattern, compared with OLIGOMER which determines the optimum combination of radii of gyrations of input models. For further analysis, data collected in December 2017 was used. The P(r) distributions are also distinguishable for the range of EOM models produced for OLIGOMER, and the P(r) distribution produced from the experimental data (Figure 4.6C) supports a rod-like model as expected. Using the P(r) distribution, ten models were generated using DAMMIF and then averaged using DAMAVER producing a single combined model. This agrees with the experimental data well ($Chi^2 = 1.15$), suggesting that the single best-fit model (Figure 4.8A) describes the shape of S2 hN1 11-13cl

in solution, an elongated rigid rod (Figure 4.8). Importantly, the bent structure does not fit the data well when comparing to the scatter curve using CRYSOL (Figure 4.8B), further supporting that SAXS is able to clearly differentiate straight and bent EGF domain structures within a triple domain construct.

The Kratky plot (Figure 4.6D) is commonly quoted as a measure of unfoldedness of a protein (Putnam et al, 2007). A folded protein is widely understood to have a Kratky plot showing a peak (a fall after the initial rise in $I*s^2$ in the plot). All of the elongated constructs investigated in this chapter have an atypical Kratky plot which shows a plateau following the initial rise in $I*s^2$ (Figure 4.6D). This is commonly quoted as representative of an unfolded or disordered protein, however the S2 hN1 11-13 construct was shown to be natively-folded as supported by multiple techniques including proton NMR (Weisshuhn et al, 2016). This suggests that near linear elongated and non-spherical proteins produce a plot atypical of globular proteins and so care is required when using the Kratky plot as a marker of the foldedness of near-linear, elongated proteins. In fact, compressing an elongated model to produce a compact model shows that a peak in the Kratky is a measure of compactness/sphericity rather than unfoldedness/disorder (Figure 4.7).

i



		% of each model			
Data set	Chi² of combination fit	A	B	C	D
11-13cl Sep17	1.10	89	11	0	0
	1.10	89	11	0	
	1.10	89	11		
	1.11	100			
11-13cl Dec17	1.06	80	20	0	0
	1.06	80	20	0	
	1.06	80	20		
	1.09	100			

ii

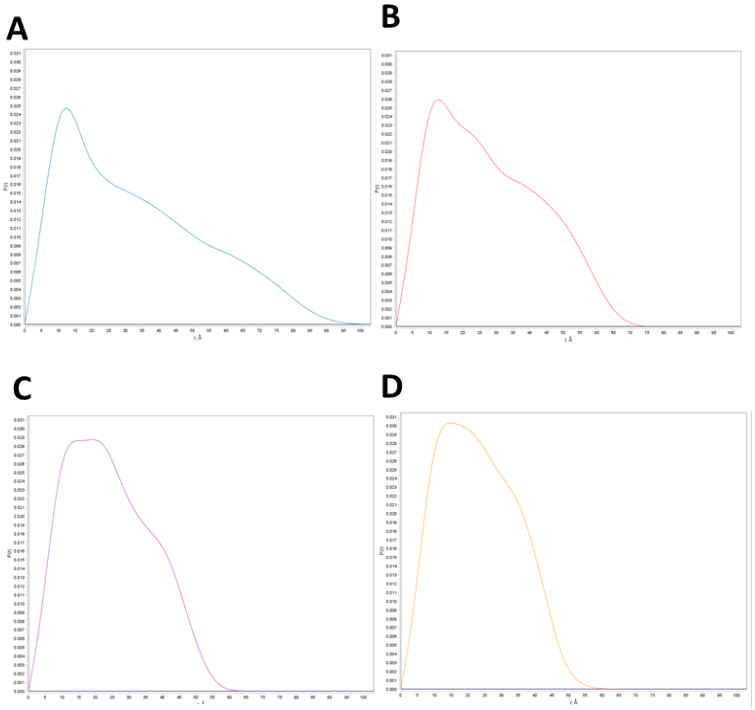


Figure 4.5 OLIGOMER model fitting to S2 hN1 11-13cl: (i) the selected range of models generated with EOM and the optimum combination (%) of models which fits the experimental scatter curves for the two S2 hN1 11-13cl data collections (calculated by OLIGOMER). Various combinations of models were processed through OLIGOMER. (ii) P(r) distributions of the four EOM-generated models with A, B, C, D corresponding to the P(r) for the respective model (axes x = r/Å, y = P(r)). A $Chi^2 \approx 1$ indicates a good fit. Note that model A which represents the largest structural contribution in the OLIGOMER fitting gives a P(r) distribution similar to the elongated model derived from the crystal structure of hN1 11-13 (Figure 4.3A).

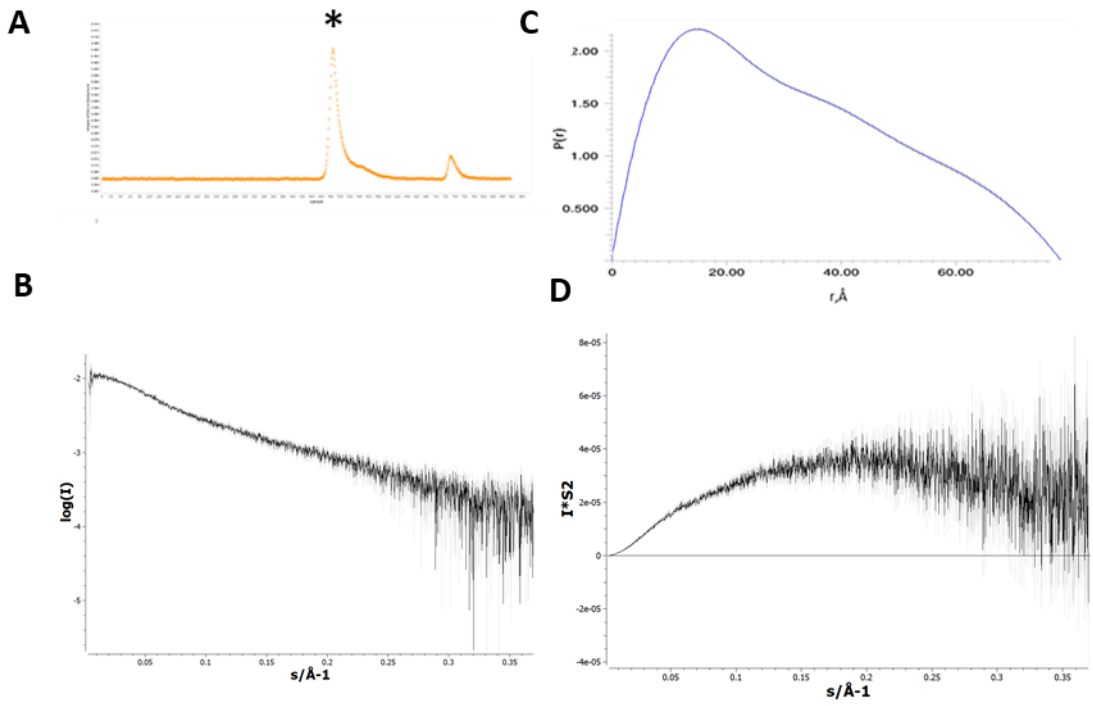


Figure 4.6 Analyzing the S2 hN1 11-13cl SAXS data collected in 15 mM Ca^{2+}: (A) The scatter profile from the SEC elution shows a clear initial peak (*) corresponding to protein (the smaller peak eluting later does not absorb at 280 nm) (axes x = volume/ml, y = scattering intensity). (B) The scatter curve from the eluted protein peak after subtracting background scatter. (C) P(r) distribution derived from the scatter curve, note that it is similar to the most elongated P(r)s in Figure 4.5. (D) Kratky plot for these data.

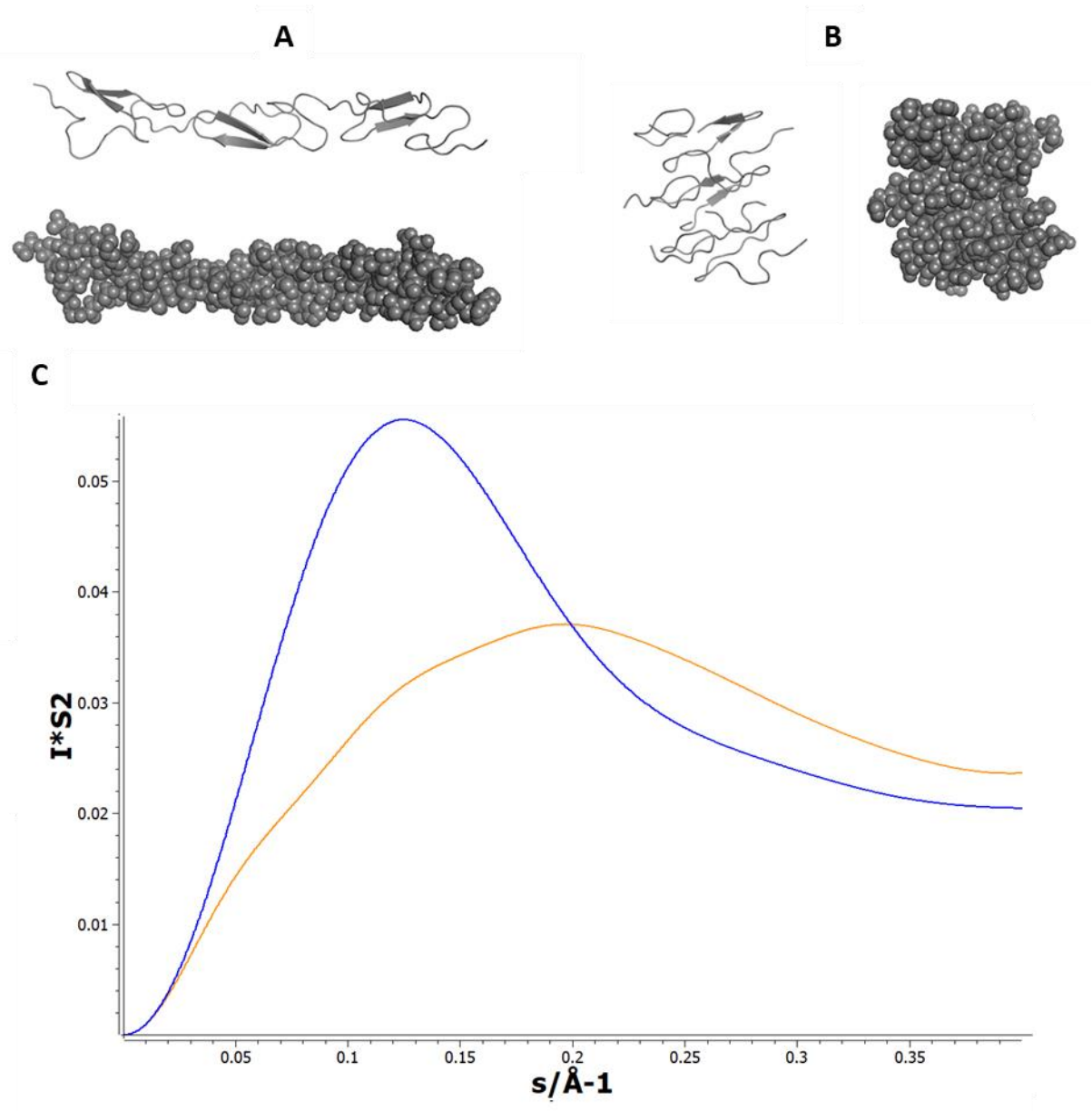


Figure 4.7 Investigating the Kratky plot: (A) Straight model of a triple domain construct derived from 2VJ3 (see Figure 4.3) shown as cartoon and sphere representation. (B) Model of the same three EGF domains as in model A, but stacked vertically, shown as cartoon and sphere representation. This is to produce a more compact shape with the same volume and constituent atoms as model A. (C) The theoretical Kratky plots for model A (orange) and model B (blue) produced using the ATSAS suite. Compacting the linear structure (A) into a less elongated structure (B), leads to an increase in the peak pronouncement in the Kratky plot. The Model B plot would commonly be regarded to show a folded protein and Model A plot to show an unfolded protein. It is clear from the input models that the Kratky plot gives information on compactness/sphericity, rather than unfoldedness/disorder, considering that the input models are of identical mass and volume.

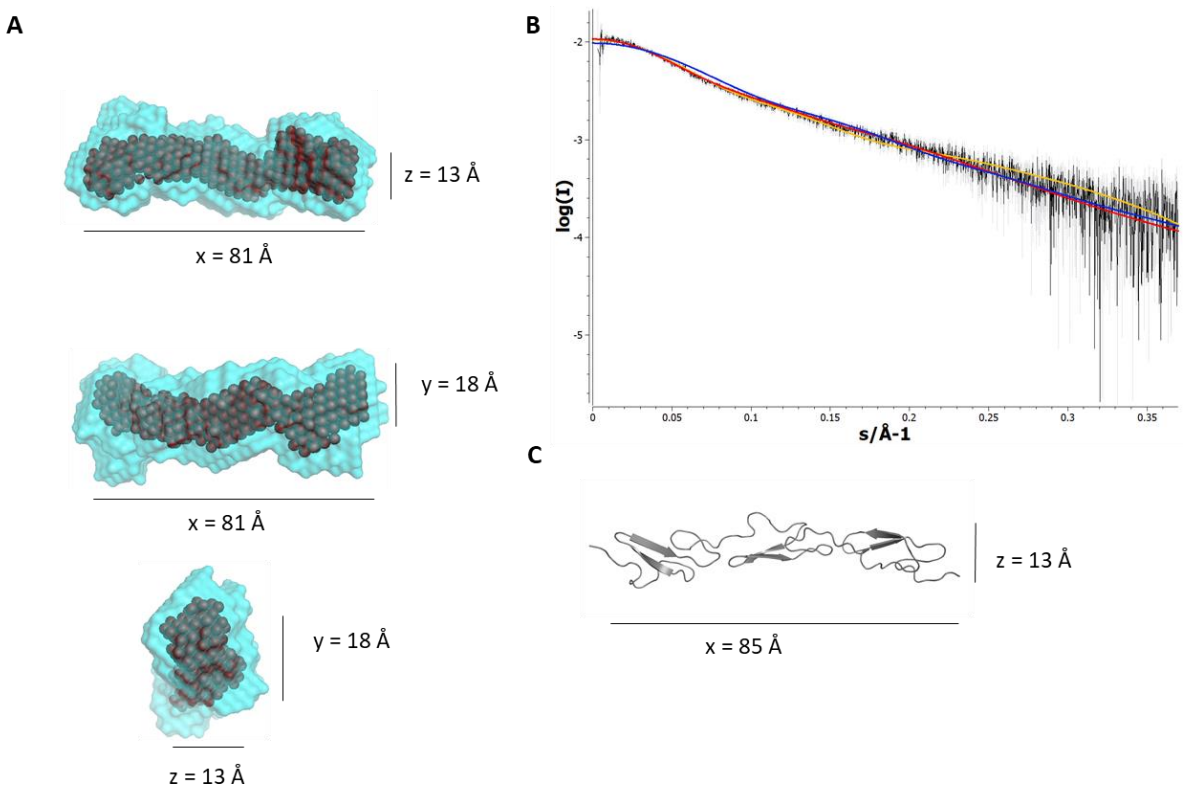


Figure 4.8 Determination of hN1 11-13cl structure using SAXS: Following determination of P(r) distribution, DAMMIF and DAMAVER were used to determine the averaged best fit model (A) shown from three perspectives, the maroon centre is the best fit model and the cyan volume is the space filled by the 10 models used for averaging. (B) shows the theoretical (CRYSOL) fit of this best fit model to the original experimental scatter data (Chi^2 = 1.15, orange line). The fit for the straight model derived from 2VJ3 is good (Chi^2 = 0.97, red line), and the bent model does not fit well (Chi^2 = 2.07, blue line) both described in Figure 4.4. (C) shows the dimensions of the rigid rod crystal structure 2VJ3 which are in good agreement with the generated model.

4.2.5. Structure of hN1 23-27CSCS

Having validated a pipeline for analyzing SAXS data collected from EGF domain fragments, this method was used to determine the structures of novel regions of the receptor. hN1 23-27CSCScl was designed and expressed in the prokaryotic/refold system yielding a final

product of at least 95% purity (see Chapter 3 – Protein Purification and Characterisation, Figure 3.16). EGF23 was included to ensure the native state of Ca^{2+}-bound EGF24. The non-canonical cysteines in EGF25 and EGF27 were excluded to ensure easier refolding. The exclusion of EGF28 and 29 from this construct was primarily because the yield from expressing this seven-domain construct in bacteria followed by in vitro oxidation would likely be significantly lower than the five-domain construct expressed here. The moderately high affinity cbEGF EGF24 (See Chapter 3 – Protein purification and characterisation, Section 3.5) provides some stability and rigidity to the construct compared to poorly characterized EGF28 and 29 which are predicted from bioinformatics analysis to be non-Ca^{2+} binding. SAXS was carried out on this construct using the pipeline developed with hN1 11-13cl (See Section 4.2.3). Data were collected in high Ca^{2+} conditions (15 mM Ca^{2+}) after eluting as a single peak from the in-line SEC column (Figure 4.9).

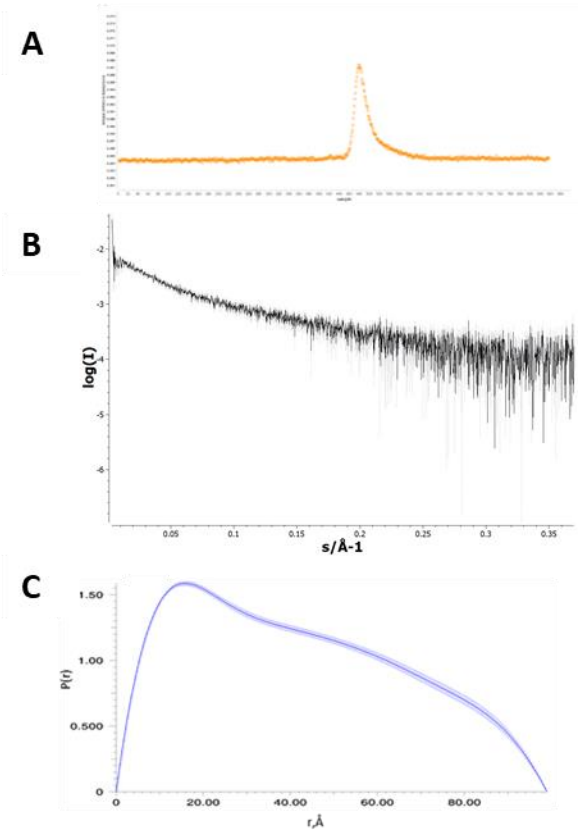
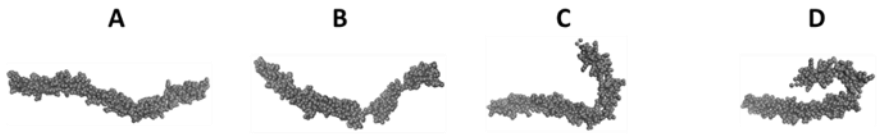


Figure 4.9 SAXS analysis of hN1 23-27CSCScl (15 mM Ca^{2+}): (A) The scatter profile from the SEC elution shows a clear elution peak corresponding to protein (axes x = volume/ml, y = scattering intensity). (B) the scatter curve from the eluted protein peak after subtracting background scatter. (C) P(r) distribution derived from scatter curve.

Generating rational models from EOM for this construct involved combining all the information known for this region of the Notch receptor and predicting where flexible inter-domain regions might occur. Although EGF domains 23-25 and 27 are shown or predicted to be Ca^{2+}-binding, the presence of non-Ca^{2+}-binding EGF26 could introduce a bend or flexibility. The Ca^{2+} titration data for hN1 EGF24 described in (see Chapter 3 - Protein Purification and Characterization, Section 3.5) shows that the atypical Ca^{2+} consensus binds Ca^{2+} with moderately high affinity (~6 µM). This suggests that the domain interfaces between EGF23-24 and 24-25 are rigid and likely lead to a rod-like structure for this region. EGF 26 is predicted

to not bind Ca^{2+} due to the lack of a cbEGF consensus (See Chapter 3 - Protein Purification and Characterisation, Figure 3.2). Therefore, flexibility or bending could be possible at the EGF25-26 interface (See Chapter 1 – Figure 1.9). Furthermore, when generating models in EOM, flexibility was allowed at the EGF26-27 interface (EGF-cbEGF). The EOM model generation procedure and four models were selected to represent a range of flexibilities. The range of models chosen is shown in Figure 4.10. Fitting combinations of these EOM models to the scatter curve using OLIGOMER suggests that the structure of this construct in solution is not fully elongated but slightly kinked in solution (Figure 4.10). Comparison of the P(r) distributions from the range of EOM models selected to represent a range of conformations shows that these models can be distinguished using this parameter and that the P(r) from the experimental data fits best with a slightly compacted P(r) in terms of D_{max} but is smoother, suggesting that the structure is slightly kinked in solution. Future studies could use an additional range of models between EOM structures A and B to further characterize the extent of this bend, however, in this initial study, a validation that the OLIGOMER result agreed with the DAMMIF result was achieved using these four models.

i



Data set	Chi² of combination fit	% of each model A	B	C	D
23-27CSCS 15mM Ca^{2+}	0.91	46	47	6	1
	0.91	42	49	1	
	0.91	36	64		
	0.92		100		

ii

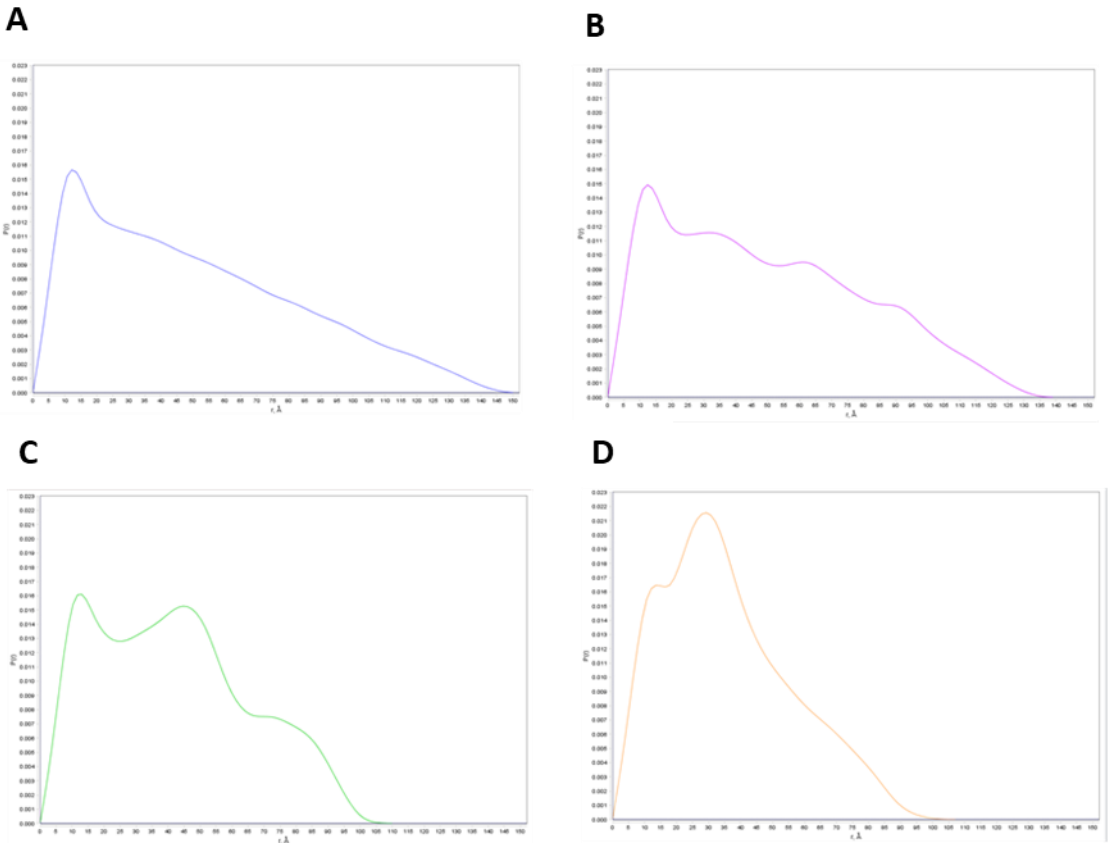


Figure 4.10 OLIGOMER analysis of hN1 23-27CSCScl in the presence of Ca^{2+}: (i) The selected range of models generated with EOM and the optimum combination (%) of models which fits the experimental scatter curve for the hN1 23-27CSCScl data (calculated by OLIGOMER). (ii) P(r) distributions of the four EOM-generated models with A, B, C, D corresponding to the P(r) for the respective model (axes x = r/Å, y = P(r)). A $Chi^2 \approx 1$ indicates a good fit.

4.2.5.1. Modelling of hN1 23-27CSCScl

The P(r) distribution (Figure 4.9) was used as the basis for modelling by DAMMIF to generate 10 models, which were then averaged using DAMAVER to produce a best fit-structure (see Figure 4.11). The structure of hN1 23-27CSCScl in solution appears to be relatively elongated with some bending. This structure fits the experimental data well suggesting the model is representative of the structure in solution (Chi^2 = 1.05). The cross-sectional distances (14 Å and 20 Å) also agree well with those for hN1 11-13 as would be expected for multiple tandem EGF constructs.

4.2.6. Structure of hN2 23-27cl

In order to further investigate the structure of the Abruptex region, the structure of hN2 23-27cl was investigated using SAXS. The rationale for investigating this construct was that the structure of the Abruptex region is likely to be conserved between hN1 and hN2 considering that these receptors both have 36 EGF domains in the extracellular region (See Chapter 1 – Introduction, Figure 1.2). However, the homologue hN2 lacks the additional non-canonical cysteines in the Abruptex region (See Chapter 3 – Protein purification and characterisation, Figure 3.2). The hN2 23-27 material was produced in the prokaryotic/refold system and purified in a similar manner to hN1 23-27CSCScl (see Chapter 3 – Protein Purification and Characterisation, Section 3.3.5).

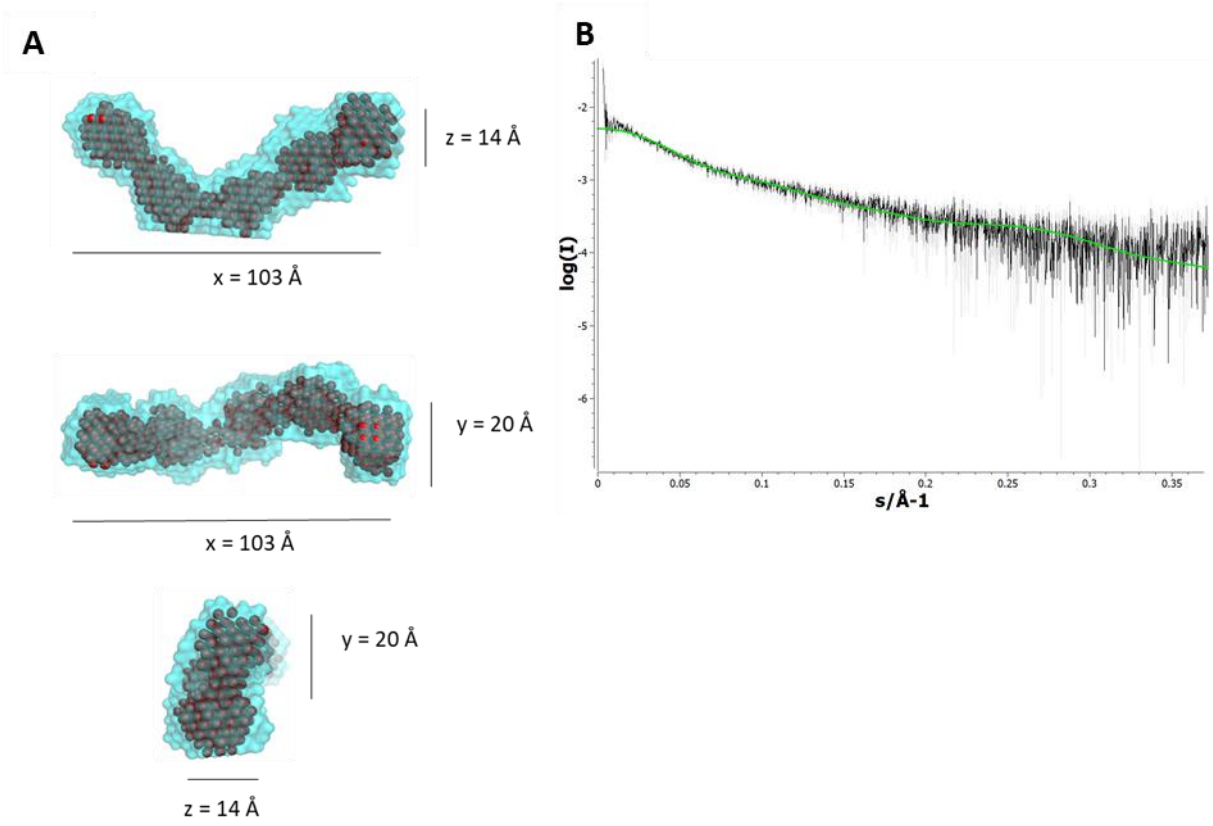
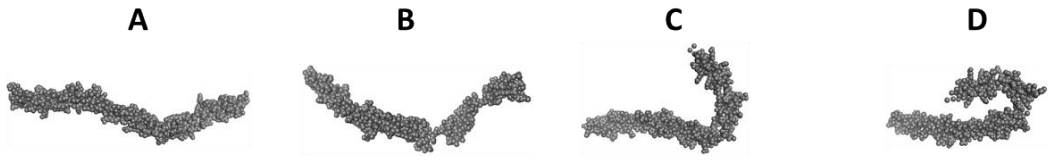


Figure 4.11 The structure of hN1 23-27CSCScl at 15 mM Ca^{2+}: Following determination of the P(r) distribution, DAMMIF and DAMAVER were used to determine the averaged best fit model (A, shown from three perspectives), the maroon centre is the best fit model and the cyan volume is the space filled by the 10 models used for averaging. (B) shows the theoretical (CRYSOL) fit of this best fit model to the original experimental scatter data (Chi^2 = 1.05, green line).

A similar SAXS experiment was carried out on hN2 23-27cl, also in 15 mM Ca^{2+} buffer conditions. Identical EOM models were used during OLIGOMER analysis for hN2 as for hN1. OLIGOMER analysis shows the optimum combination of EOM models required to fit the scatter curve well, suggests that the structure is not fully elongated and slightly bent, like hN1 23-27CSCScl (Figure 4.12). Although the data in Figure 4.12 suggest that the hN2 construct might be more compact due to the dominant contribution of model B, the data exclude a folded-back model (Figure 4.12, model D). Future studies could use an additional

range of models between EOM structures A and B to further characterize the extent of the bend, however, in this initial study, a validation that the OLIGOMER result agreed with the DAMMIF result was achieved using only four models.



		% of each model			
Data set	**Chi² of combination fit**	**A**	**B**	**C**	**D**
hN2 23-27 15mM Ca^{2+}	0.73	23	72	2	3
	0.73	25	68	7	
	0.73	20	80		
	0.74		100		

Figure 4.12 Investigating the structure of hN2 23-27cl: The selected range of models generated with EOM and the optimum combination (%) of models which fits the experimental scatter curve for the hN2 23-27cl data (calculated by OLIGOMER). These models are identical to those used in Figure 4.10. A $Chi^2 \approx 1$ indicates a good fit, and is low in this fitting, suggesting data overfitting.

The hN2 23-27cl data (Figure 4.13) were used to determine the structure of the region in solution in a similar manner as the previous experiments. The eluant scatter profile (Figure 4.13A) shows a single peak, suggesting that the additional minor band seen when analyzing by non-reducing 16% SDS-PAGE (Chapter 3 – Protein Purification and Characterisation, Figure 3.16), is a very small share of total protein in the sample, but possibly the cause of the slight asymmetry of the SEC elution peak (Figure 4.13). The best fit model was subsequently determined following the same procedure as for hN1 23-27CSCScl (Figure 4.14).

The dimensions fit well with what is expected for a multiple EGF-domain construct (16 Å and 20 Å). Overall, these data suggest the structure of hN2 23-27cl is an elongated bent rod (Figure 4.14). This model fits the starting data well with a Chi^2 of 0.82.

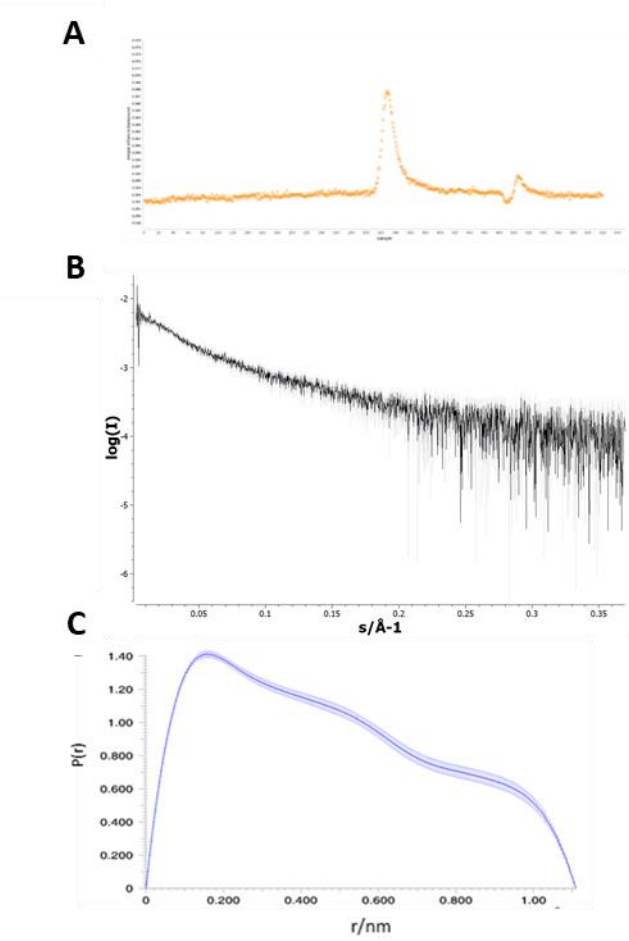


Figure 4.13 Analyzing SAXS data from hN2 23-27cl: (A) The scatter profile from the SEC elution and shows a clear peak corresponding to protein (axes x = volume/ml, y = scattering intensity). (B) the scatter curve from the eluted protein peak after subtracting background scatter. (C) P(r) distribution derived from the scatter curve.

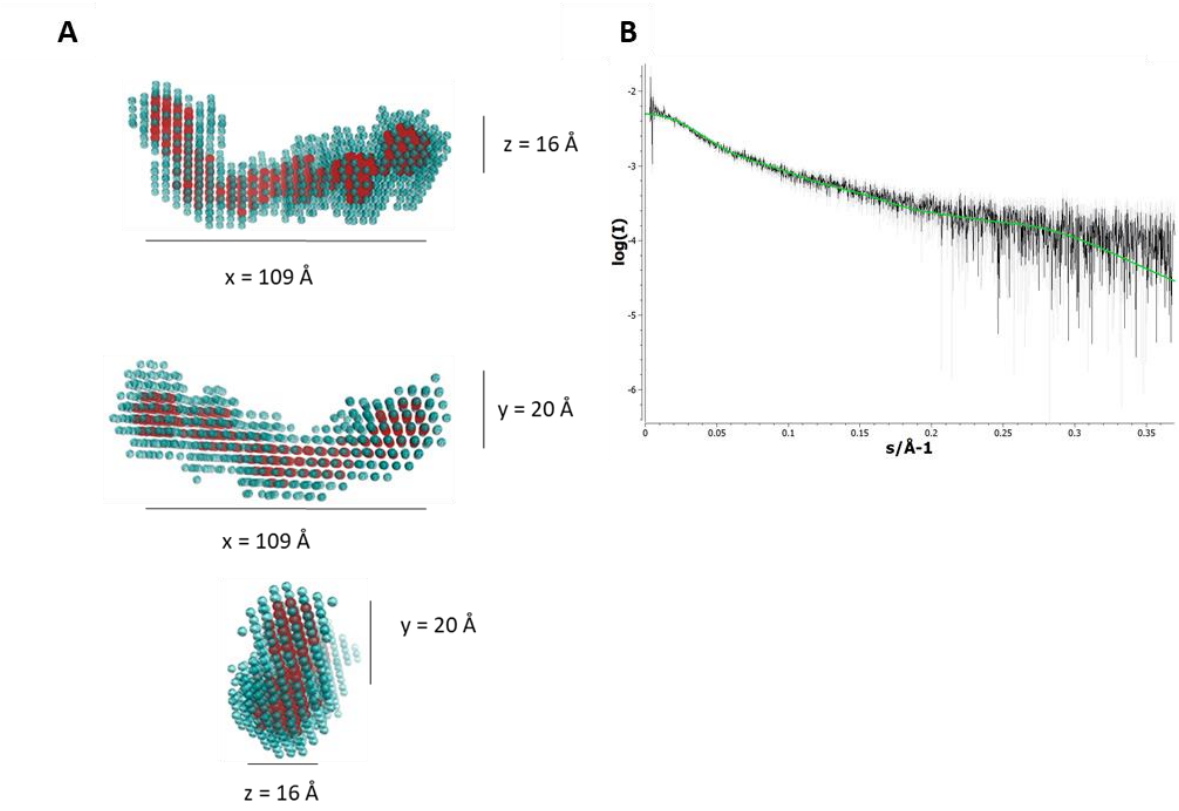


Figure 4.14 The structure of hN2 23-27cl: Following determination of the P(r) distribution, DAMMIF and DAMAVER were used to determine the averaged best fit model (A), shown from three perspectives), the maroon centre is the best fit model and the cyan volume is the space filled by the 10 models used for averaging. (B) The theoretical (CRYSOL) fit of this best fit model to the original experimental scatter data (Chi^2 = 0.82, green line).

4.2.7. Comparison of EGF23-27 in hN1 and hN2

The structures generated for hN1 23-27CSCScl and hN2 23-27cl are similar in terms of dimensions. Both structures are relatively elongated and rod-like with some bending. Both have a more pronounced bend in a similar position along the construct which corresponds with either the 24-25, or the 25-26 boundary. Considering what is known about Ca^{2+}

properties of this region in hN1, this bend is most likely at the EGF25-26 boundary because EGF26 is not predicted to bind Ca^{2+} in hN1 (See Chapter 3 – Protein purification and characterisation, Figure 3.2). Models were built taking this into account (Figure 4.15). EGF26 is predicted to be Ca^{2+}-binding in hN2 and dN, however these proteins, as well as hN1 all have a substitution of a conserved proline at the position immediately before Cys_2, part of a characteristic C_1-T-E-S-S-C_2 (see Figure 4.15). This feature means that EGF26 is not modified with O-Gluc in these proteins. The fully extended trisaccharide at the O-glucose modification has been recently shown to sit in the groove of the EGF stabilizing the β-hairpin to facilitate positioning of the Ca^{2+}-ligands for ion binding (Takeuchi et al, 2017). Furthermore, addition of O-glucose glycans has been previously shown to stabilize an isolated cbEGF from unfolding in the presence of 5 mM Ca^{2+} compared to without Ca^{2+} present (Chapter 1 – Introduction, Section 1.3.4.3, Takeuchi et al, 2017).

This suggests that Ca^{2+}-binding is possibly influenced by the presence or lack of O-Gluc modifications, but this could be indirect by influencing β-sheet stability. However, there is currently no evidence of O-Gluc modifications influencing EGF-interfaces as yet. The lack of this conserved proline and the conserved lack of an O-Gluc modification site may aid the formation of the kink between EGF25-26 in hN2 (and dN) proposed in the SAXS structures shown here. A recent structure of mN1 EGF26 (non-Ca^{2+} binding) in complex with POFUT1 has been determined (Li et al, 2017), however the structure is only visible from Cys1, the characteristic C_1-T-E-S-S-C_2 is at the extreme N-terminal of the construct and so the position

of the loop in this EGF-enzyme complex structure is unlikely to reflect its position in the context of a series of EGFs.

The fact that the structures of EGF23-27 in hN1 and hN2 are similar from SAXS analysis suggests that the non-canonical cysteines in hN1 do not have a role in holding the region in the correct 'conserved' structure, and that this structure is not bent back on itself to accommodate an interdomain disulphide bond. The normal fold of the Abruptex, as observed for the hN2 construct is attainable by hN1 even when the non-canonical Cys are substituted for Ser.

The observation that this region is extended complements the fact that missense mutations affecting the Abruptex in *D. melanogaster* are spread all over the region from EGF24-29 (See Chapter 1 – Introduction, Section 1.4.3.1). This suggests that any important interactions of this region with another tandem protein may be extended over several EGF domains in a similar manner to the interaction of the LBR with ligand (See Chapter 1 – Introduction, Section 1.4.2). These results suggest that SAXS is a valid technique for investigating the shape of very elongated molecules, such as multiple tandem EGF domains with one dimension $1/10^{th}$ smaller than other dimensions. As observed for the 11-13 construct, good quality data allows differentiation between straight and bent models, and can be used to investigate the shape of longer constructs which are not amenable for crystallography or electron microscopy. In addition, an atypical Kratky plot is observed for these elongated proteins, despite being rigid in solution, likely due to their lack of globular character.

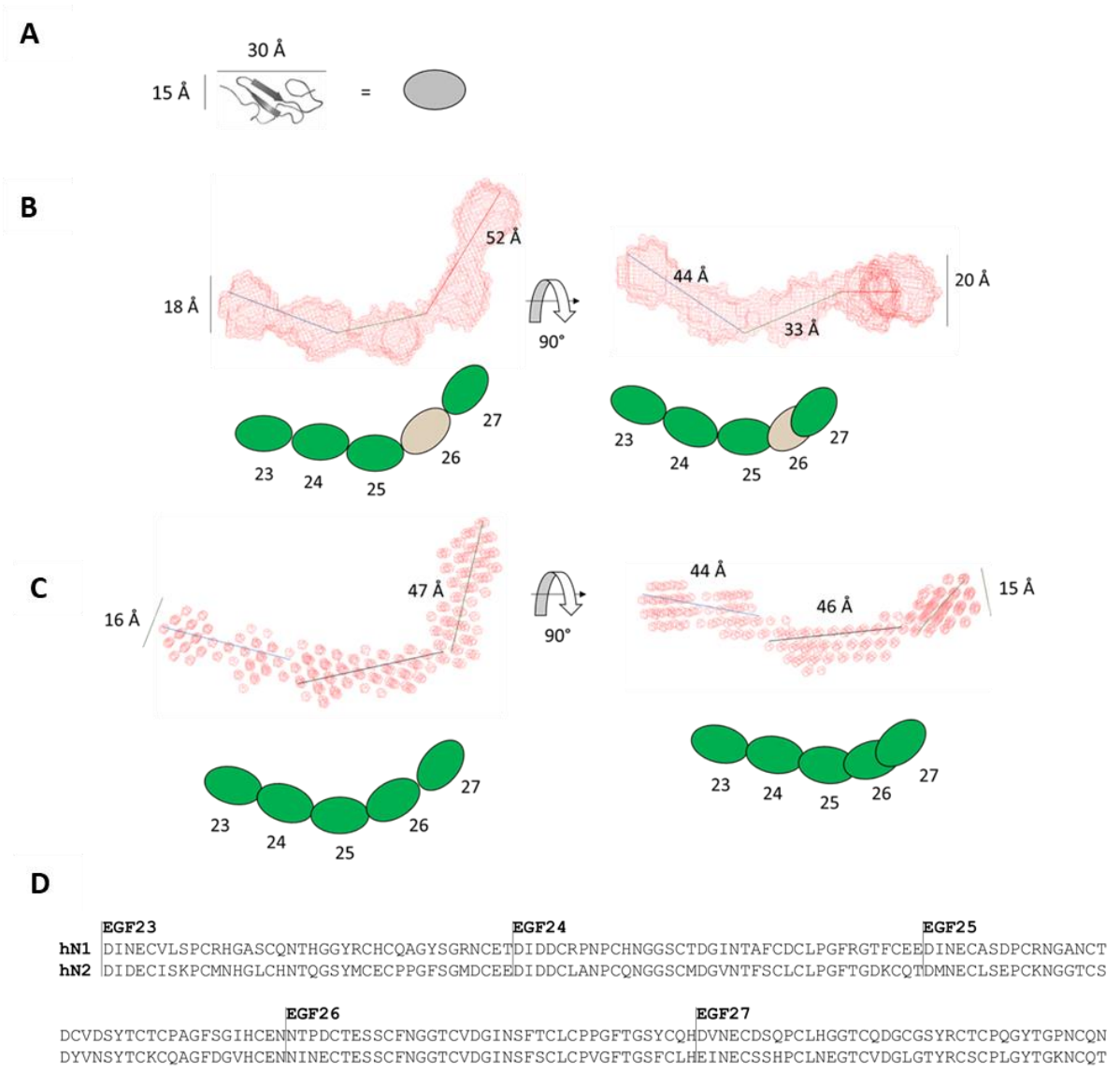


Figure 4.15 The structures of hN1 and hN2 EGF23-27: (A) The dimensions of a typical cbEGF/EGF domain are shown. (B) the detailed dimensions of hN1 23-27CSCScl and how a series of cbEGFs (green) and EGFs (wheat) would fit into this. EGF domains are numbered assuming the biggest kink is at the cbEGF-EGF interface between EGF25-26. (C) A schematic of hN2 23-27cl. In this case, the order of domains in the model is based on similarity with hN1. (D) The sequences of EGF23-27 from hN1 and hN2 are shown emphasizing their similarity.

5.

Interaction of the Abruptex region with the ligand binding region

5.1 Background

Previously, hN1 23-24 was shown to interact with hN1 11-13 on the surface of HEK293T cells in a flow cytometry assay (Whiteman et al, unpublished data). A calcium binding variant D909V in hN1 EGF24, homologous to the Abruptex substitution D948V in dN (See Chapter 1 – Introduction, Section 1.4.3.1) was shown to diminish this interaction (Figure 5.1, Whiteman et al, unpublished data). These data suggested a specific interaction between the N-terminal of EGF24 and the LBR.

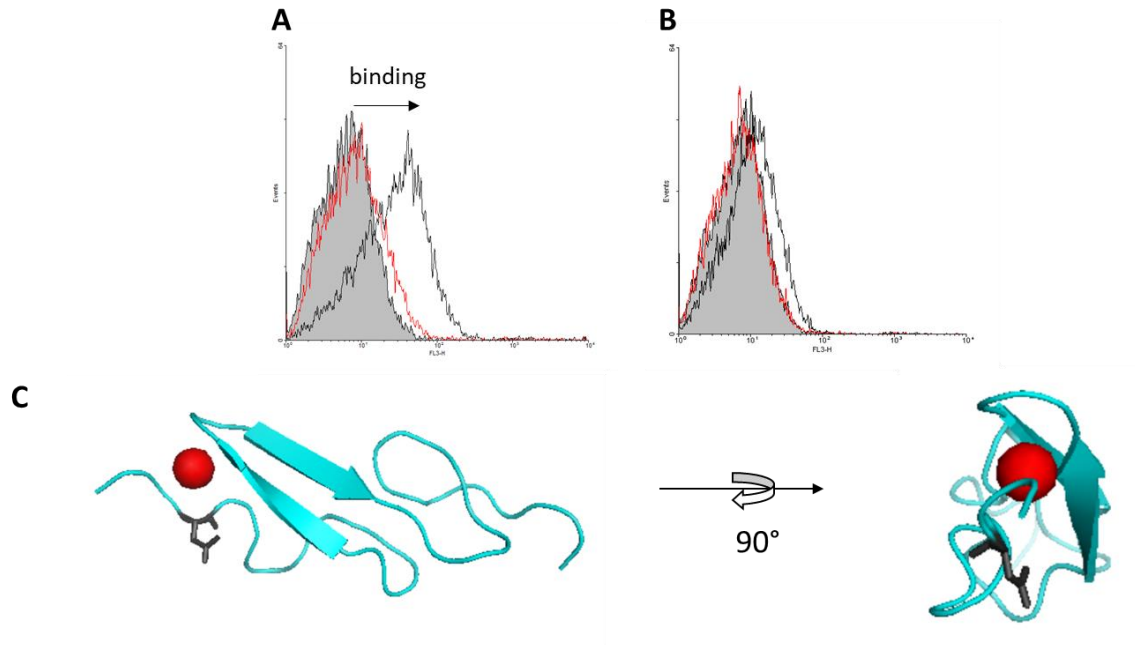


Figure 5.1 D909V in EGF24 disrupts the interaction with 11-13: A) PR hN1 23-24 wt binds (black) relative to background binding (grey) and binding can be disrupted by treating with an antibody against EGF24 (red). B) PR hN1 23-24 D909V (black) binds less than wt and can similarly be disrupted by treating with an anti-EGF24 mAb, equivalent biotinylation was shown by western blot. These experiments were carried out by Dr P Whiteman, University of Oxford. C) The position of the D909V substitution (black) is shown in on an EGF24 schematic (cyan, Ca^{2+} = red sphere).

The impact of *Abruptex* mutations in *D. melanogaster* is suggestive of too much Notch signalling (de Celis and Garcia-Bellido, 1994), which implies that the normal role of the Abruptex region is to negatively regulate signalling (see Chapter 1 – Introduction, Section 1.4.3). A proposed mechanism for this inhibitory effect is for the Abruptex region (EGF24-29) to compete with ligand for binding to the LBR, and this is supported by observations that the Abruptex region interacts with the LBR dN and hN1 (Sharma et al, 2013, Pei and Baker, 2008). Here, this interaction was investigated further, with the aim of characterising the binding interface using purified constructs of the Ax region and LBR. To do this, a plate-based assay and a flow-cytometry assay were employed, for which a

variety of Ax constructs were designed and expressed in the prokaryotic/refold and S2 systems (see Chapter 3 - Protein Purification and Characterisation).

Results from Chapter 4 – Low Resolution Structural Studies of the Abruptex region suggest EGF23-27 has an elongated shape (See Section 4.2.7). This may produce an extended binding surface to allow an interaction with the LBR in a similar manner to the LBR- ligand interaction. Considering that known Ax mutations occur throughout the length of the region (See Chapter 1 – Introduction, Section 1.4.3.1), this extended interface seems a plausible suggestion, however the spatial distributions of the different types of Ax mutations (*N*-suppressor, *N*-enhancer, lethal) suggest that the Ax region may have a number of distinct functional areas, an N-terminal site (encompassing EGF24-25), and a C-terminal interaction site (encompassing EGF27-29) (See Figure 5.2).

The N-terminal of the Abruptex region (EGF24) has been structurally characterised by X-ray crystallography (Suckling et al, unpublished data) and is known to bind Ca^{2+} with < 10µM affinity (Chapter 3 - Protein Purification and Characterisation Section 3.5). The remainder of the Abruptex region (EGF25-29) has not been well characterised until now, although EGF25 and EGF27 are predicted to be Ca^{2+}-binding in hN1 and hN2 from sequence analysis (see Chapter 3 - Protein Purification and Characterisation, Figure 3.2). Furthermore, the additional non-canonical Cys residues in EGF25 and EGF27 in hN1 add to the potential complexity of purifying this region, although these seem to not be present as exposed reduced sulfhydryl groups in purified S2 material (see Chapter 3 - Protein Purification and Characterisation, Section 3.6). Considering this, an S2 N-terminal construct of Ax (S2 hN1 23-24) and prokaryotic refold N-terminal construct of Ax (PR hN1

23-24) were used in the first instance to investigate the interaction of the Ax region with the LBR (See Figure 5.2), with the inclusion of the preceding domain (EGF23) to stabilise the calcium-binding fold of EGF24 (see Chapter 3 - Protein Purification and Characterisation). This eliminated any complications of purifying in the presence of the additional cysteines in EGF25 and EGF27.

5.2 Investigating the Ax - LBR interaction in a plate binding assay

5.2.1 The plate binding assay

The plate binding assay was established by Dr P Whiteman, University of Oxford and preliminary studies to investigate the interaction of Ax with the LBR were carried out in collaboration with Dr Whiteman. The plate binding assay involves the immobilisation of an Ax construct to the Ni^{2+}-coated plate via a His-tag. Soluble LBR constructs which are clustered on neutravidin via a biotin tag (see Chapter 3 – Protein Purification and Characterisation, Section 3.4.2) are then incubated in these wells and binding of the Ax region to the LBR is quantified by the amount of clustered LBR remaining after multiple washes. The neutravidin at the centre of the LBR cluster is conjugated to a horse radish peroxidase to provide a means of quantitative readout of the amount of LBR bound to the Ax construct. This method also allows a range of concentrations of soluble clusters to be tested to plot a saturation curve to confirm that the full binding-capability of the well is achieved (Figure 5.3). The conditions of this assay were pH7.5, I = 0.2, 1 mM Ca^{2+} (See Chapter 2 – Materials and Methods, Section 2.6.1).

The plate assay was initially validated using the interaction between immobilised ligand (S2 hJag1, provided by Dr B Korona, University of Oxford) and clustered S2 hN1 11-13. In addition, two LBR variants engineered to disrupt ligand binding were tested (Figure 5.4). S2 hN1 11-13 L468R disrupts the EGF12 ligand binding site (Site 1), and R448E disrupts the EGF11 ligand binding site (Site 2) (See Chapter 1 – Introduction, Section 1.4.2). wt LBR was found to interact with ligand in a concentration dependent manner. The LBR variants were shown to reduce binding to ligand, with L468R at site 1 being more disruptive than R448E at site 2. This preliminary study confirms that the disruption of the interaction between receptor and ligand caused by L468R and R448E can be observed in this assay.

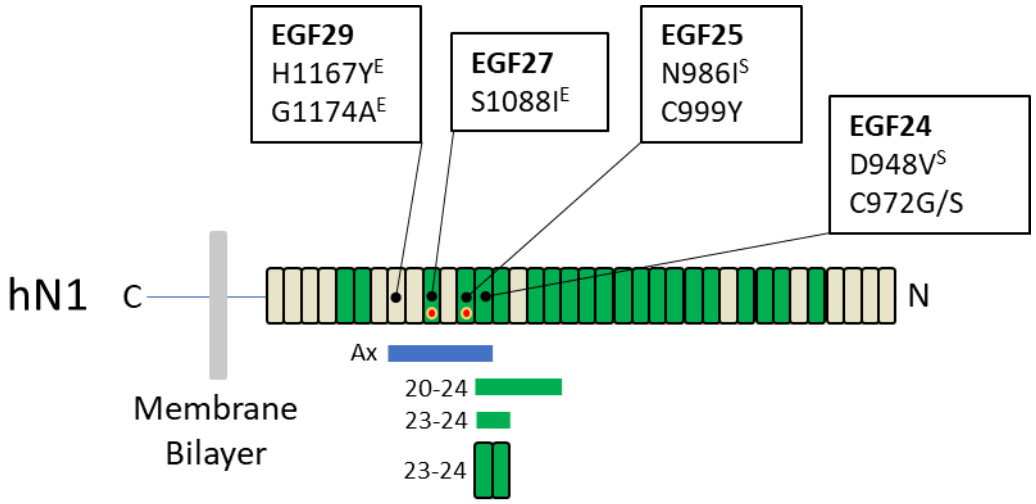


Figure 5.2 hN1 23-24 was used in interaction studies: The human Notch1 receptor hN1 is made up of cbEGFs (green rectangles) and EGFs (beige rectangles). The details of the Abruptex mutations in the Abruptex region (blue bar) are labelled for each domain with residue positions given as in dN, ^{S}*N*-suppressor mutations, ^{E}*N*-enhancer mutations). Constructs previously characterised by X-ray crystallography are shown (green strips, 20-24 and 23-24 (Suckling et al, unpublished data)). The Ax constructs designed for these interaction studies encompassed EGF23-24 (green rectangles) and expressed in both the S2 and prokaryotic refold systems.

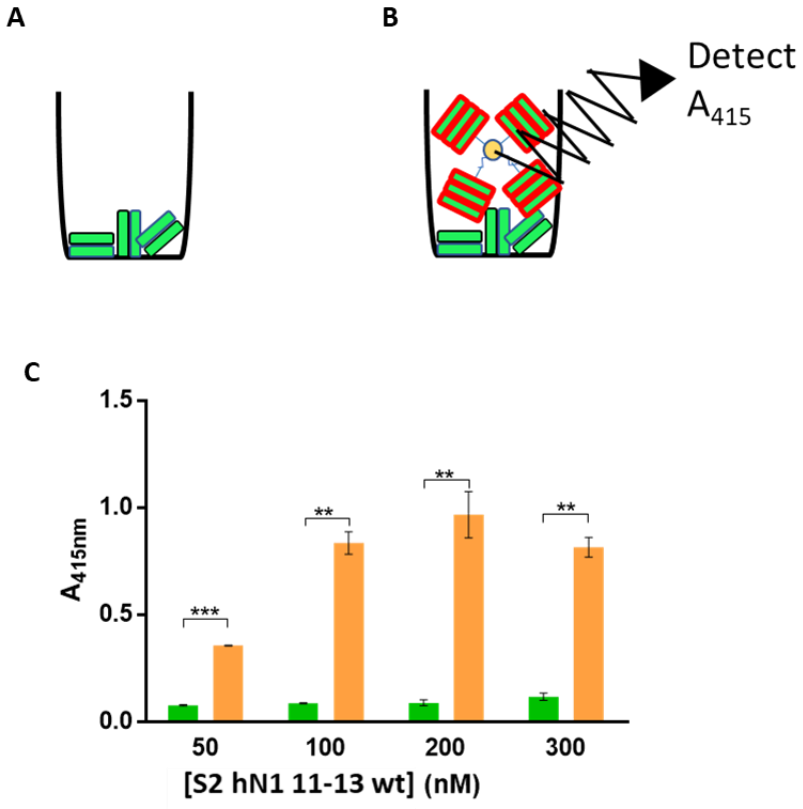


Figure 5.3 Overview of the plate binding assay: A) The Ni^{2+}-coated wells are coated with a His-tag containing protein (green rectangles, blue outline). B) Biotinylated S2 hN1 11-13 (green rectangles, red outline) clustered on neutravidin-HRP (yellow circle) is exposed to the protein in the well. Near equivalent biotinylation is shown in Chapter 3 – Figure 3.21. HRP signal is (A_{415nm}) quantified to determine the binding of S2 hN1 11-13 to the protein in the well. When comparing binding of variants to wt, relative binding is calculated by comparing variant binding to wt. The quantified HRP signal is collected over a range of concentrations of S2 hN1 11-13 incubated in wells (C) A typical graph is shown here, with higher signal indicating a greater binding affinity. Comparisons between groups are performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

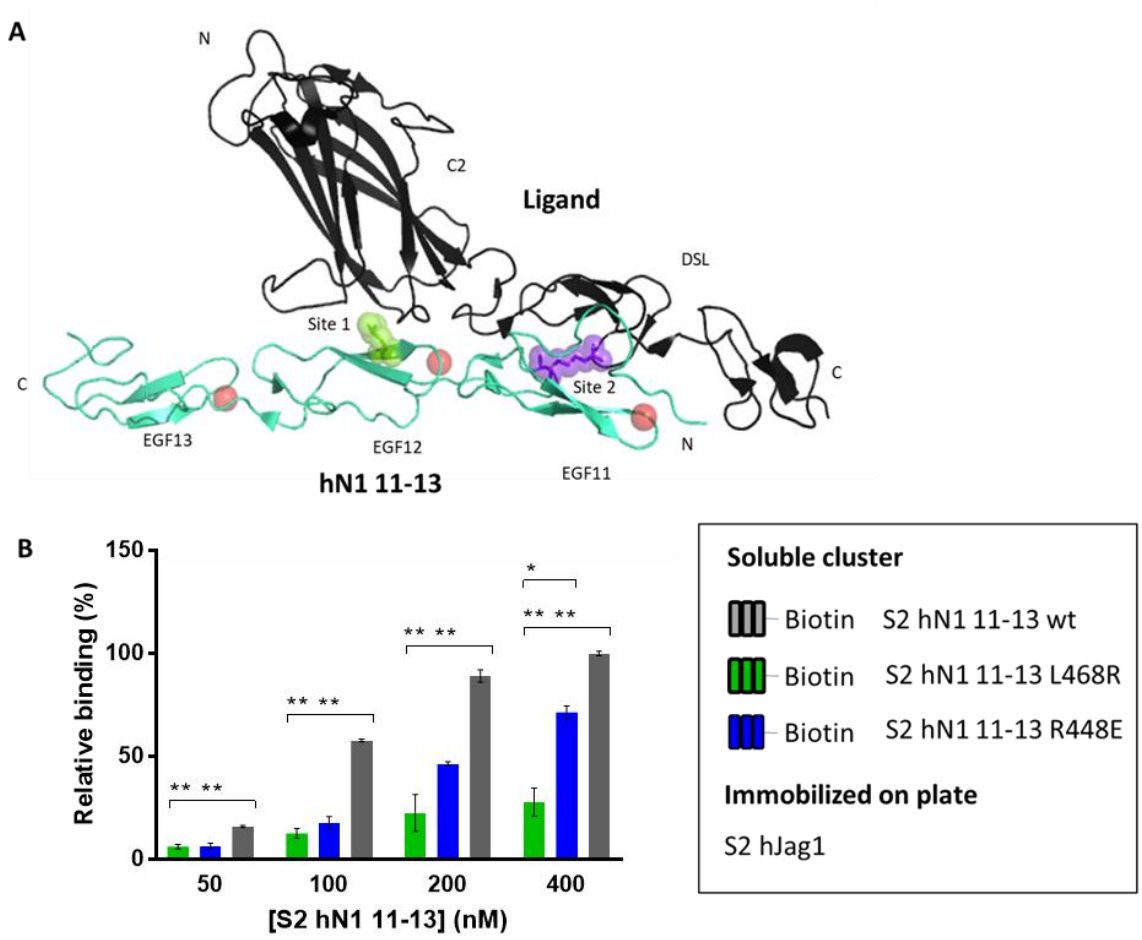


Figure 5.4 Validating the plate binding assay with ligand: A) The co-crystal structure of ligand (Dll4, black) and N1 (cyan) is shown (PDBID: 4XL1, Luca et al, 2015) with bound Ca^{2+} ions (red spheres). The two hN1 11-13 variants engineered for this assay, are L468R (L468 indicated by green spheres) to disrupt site 1 and R448E to disrupt site 2 (R448 indicated by blue spheres). These residues are at the centre of the LBR-ligand interface. B) Plate binding assay investigating binding. Binding is shown relative to binding of 400nM S2 hN1 11-13 wt. S2 hN1 11-13 wt (grey) binds better than both R448E (blue) and L468R (green). R448E binds better than L468R at the 400 nM concentration. Data show results from two independent plates, each of two repeat wells. Comparisons between variants was performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. **P<0.01; *P<0.05. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

5.2.1.1 S2 hN1 11-13 interacts with PR hN1 23-24 and S2 hJag1

hN1 23-24 was expressed in the prokaryotic/refold system having been labelled with ^{15}N and purified to at least 95% purity (Chapter 3 – Protein Purification and Characterisation, Section 3.3.2). Initial experiments were carried out to test whether PR hN1 23-24 binds to S2 hN1 11-13 specifically in the plate assay. PR hN1 20-23, provided by J Hill, University of Oxford which had been shown to be folded by NMR (data not shown) was used as a negative control alongside the protein bovine serum albumin (BSA, Sigma Aldrich). Both PR hN1 23-24 and hJag1 gave a reproducibly higher signal than controls in independent assays. However, it was noted that the signal developed in the case of PR hN1 23-24 slower than hJag1 suggesting less binding to 23-24 compared to ligand in this assay. Considering that PR hN1 20-23 does not bind, this suggests that EGF24 is responsible for the interaction observed (Figure 5.5).

Following the identification of this interaction in the plate assay, NMR was used to attempt to characterise the interaction and identify the binding face for LBR on hN1 23-24. In a similar manner to the Ca^{2+}-titration described in Chapter 3 - Protein Purification and Characterisation Section 3.5, peak shifts in a ^{1}H-^{15}N-HSQC can map an interaction by changes the shielding of nuclei in an HSQC spectrum. Subsequent assignment of the peaks in the enables identification of the residues involved in binding. Titration of S2 hN1 11-13 into a solution of ^{15}N-labelled PR hN1 23-24 did not result in any clear peak shifts (data not shown). Considering that in the plate assay, the S2 hN1 11-13 must be complexed on neutravidin to see an interaction (Whiteman et al, unpublished data). The observation that the interaction between hN1 23-24 and hN1 11-13 cannot be seen using NMR

suggests it is low affinity (> 100 µM) (Shortridge et al, 2008), and the avidity effect arising from the clustering in the plate assay is essential to observe this interaction.

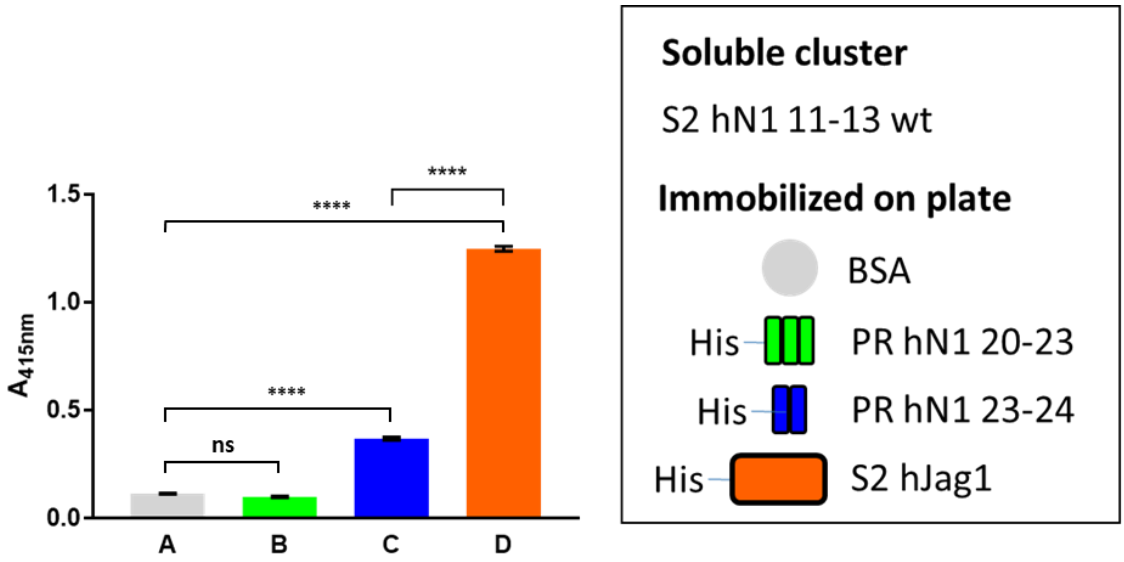


Figure 5.5 Comparison of the interaction of PR hN1 23-24 in the plate binding assay: BSA (A), PR hN1 20-23 (B), PR hN1 23-24 (C), hJag1 (D) were attached to a Ni^{2+} plate via a His-tag (except BSA). Clustered biotinylated S2 hN1 11-13 at 100 nM was then added to the wells in an identical manner and incubated with HRP for an identical length of time. Data shows results from two independent plates, each of three repeat wells. Statistical differences among various groups were assessed with ordinary one-way ANOVA. Values are presented together with the mean ± SD and compared with the BSA control (A) and each other if appropriate. ns = not significant; ****P < 0.0001. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

5.2.1.2 PR hN1 23-24 binds to S2 hN1 11-13 but not PR hN1 11-13

Further investigation was carried out to characterise the dependence on O-linked sugars for binding of LBR to Ax. The assay was carried using both PR and S2-expressed hN1 23-24 and LBR construct (hN1 11-13) (Figure 5.6). Both S2 and PR hN1 23-24 Ax constructs interact with S2 hN1 11-13, but neither interact with PR hN1 11-13 (purified by Dr P Whiteman, University of Oxford). These results suggest that the glycosylation state of 23-24 does not affect the interaction, as both S2-expressed and prokaryotic/refold 23-24

bind similarly in the plate assay, although it is necessary for 11-13 to be glycosylated for the interaction to occur. PR hN1 11-13 has been extensively characterized by NMR to show correct folding, and interacted with ligands in similar plate assay settings, confirming biotinylation (Whiteman et al, unpublished data).

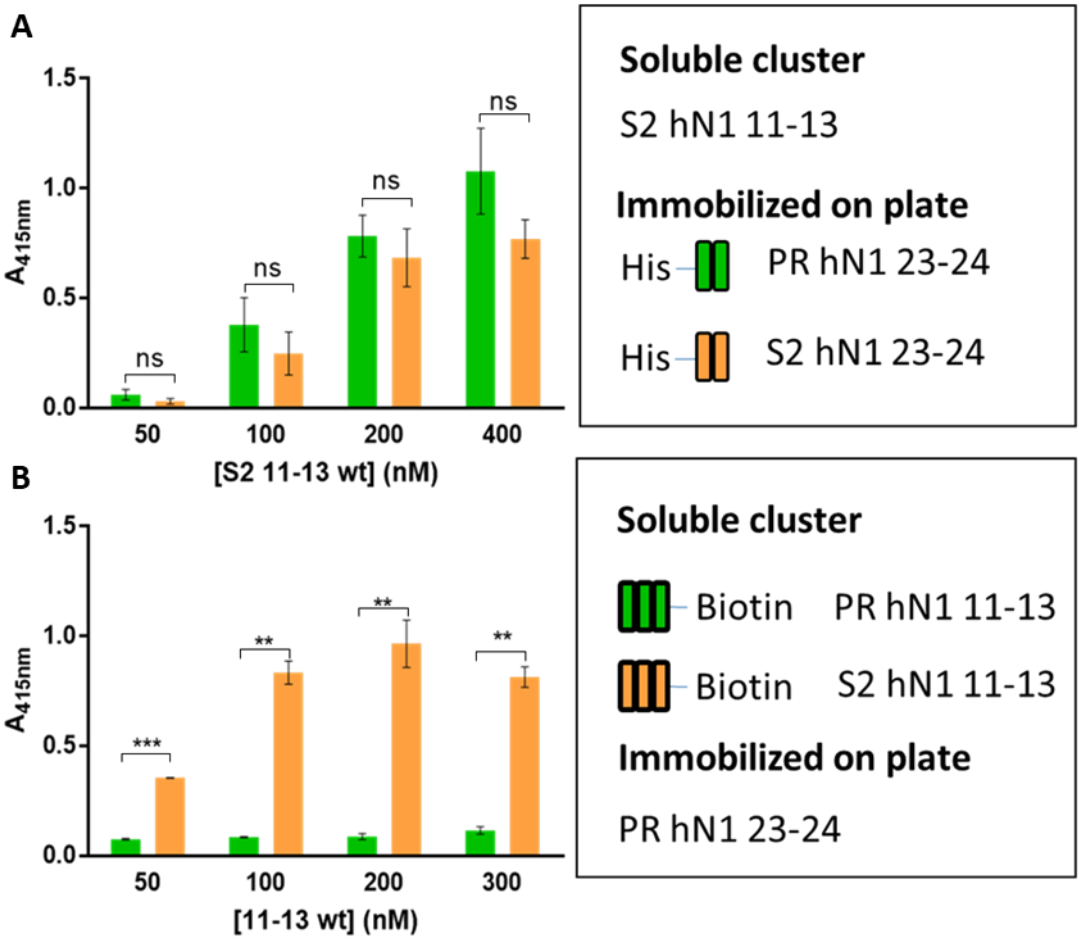


Figure 5.6 Comparison of the interaction of PR and S2 proteins in the plate binding assay: A) PR hN1 23-24 (green) or S2 hN1 23-24 (orange) were attached to a Ni^{2+} plate via am N-terminal His-tag. Biotinylated and clustered S2 11-13 wt was then added at a range of concentrations (50-400 nM). Data shown combines data from two independent plates, each of three repeat wells. B) PR hN1 23-24 was attached to a Ni^{2+} plate via a His-tag. S2 hN1 11-13 wt (orange) or PR 11-13 wt (green) were then added at a range of concentrations (50-300 nM). Data shown combines data from at two independent plates, each of at least three repeat wells. Comparisons between S2 and PR proteins were performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. ns = not significant; ***P < 0.001; **P<0.01. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

Considering the data in Figure 5.6, the location of possible O-linked sugar modification sites could potentially be used to infer the location of the interaction surface on hN1 23-24 and hN1 11-13 (no N-linked modifications are present in these regions). The structure of hN1 23-24 has been determined for the PR construct, and for the S2 construct in the context of hN1 20-24 (Suckling et al, unpublished data). Furthermore, the structure of hN1 11-13 has been structurally characterised for the PR construct (Cordle et al, 2008), and for the S2 construct of the identical hN2 region (Suckling et al, 2017) (Figure 5.7). Considering that the sugars are evenly dispersed over the surface of EGF11-13, from this analysis alone it is unclear exactly where the interaction face is located. This analysis also does not exclude the Ax-interaction face being the same as ligand considering that the glycan T466-O-Fuc is involved in direct ligand contacts (see Chapter 1 – Introduction, Section 1.4.2). Looking at hN1 23-24 (Figure 5.7B), considering that glycosylation is not required in EGF23-24 for binding, mapping the site of the interaction in this region is not possible from these data alone as large parts of this construct are distant from glycans.

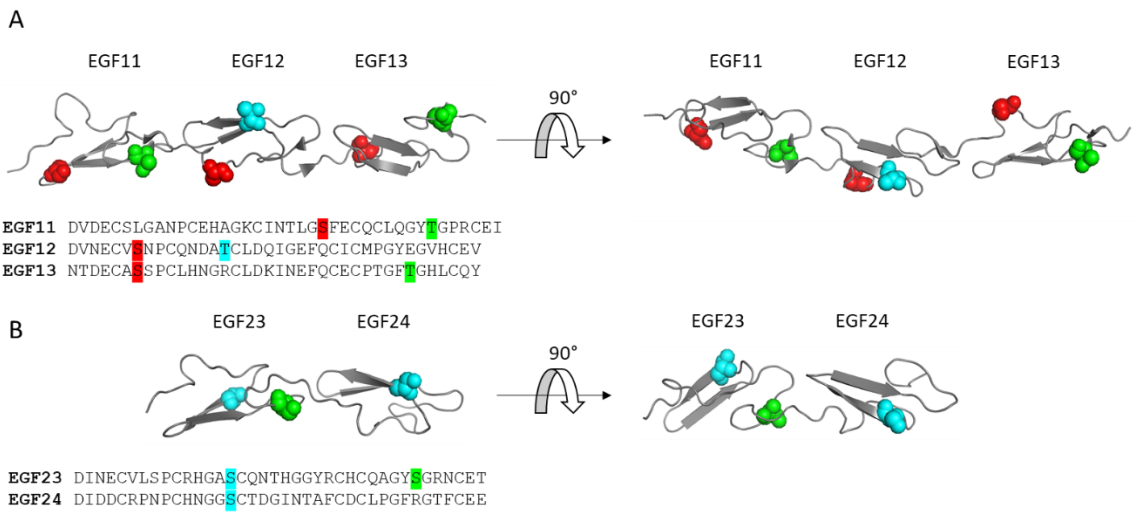


Figure 5.7 Locations of sugars on hN1 11-13 and 23-24: hN1 11-13 (A) and hN1 23-24 (B) are shown (grey) with glycosylated residues (predicted or observed) shown as coloured spheres with colours corresponding to O-Gluc (red), O-Fuc (cyan), O-GlcNAc (green). The interaction face of hN1 11-13 is sugar dependent and so is likely to involve or be close to a glycan. On the contrary, the glycan independent interaction of hN1 23-24 suggests the involvement of a region which does not have glycans.

5.2.1.3 Steric hindrance and orientation in the plate assay

Having investigated the dependence on sugars for binding of Ax to LBR, the plate assay was used to explore the dependence of binding on orientation of the immobilised construct.

The Ax region is made up entirely of EGF domains and suggested to be relatively elongated from SAXS analysis (see Chapter 4 – Low Resolution Structural studies of the Abruptex region, Section 4.2.7). The directionality of the interaction between LBR and Ax is unclear (parallel or antiparallel) and no co-crystal structure has been determined. In the case of the hN1 23-24 construct, the binding activity of the LBR appears to be centred around EGF24. The plate assay was used to investigate the impact of changing the

orientation of the hN1 23-24 construct immobilised on the plate in order to further optimize the assay.

5.2.1.4 Interaction of hN1 11-13 with hN1 23-24 is dependent on orientation of hN1 23-24

Considering that the His-tag on the Ax construct mediates the interaction with Ni^{2+}-coated plate, the position of the His-tag (N- or C-terminus) determines the orientation of hN1 23-24 on the surface of the plate. In addition to the S2 hN1 23-24 construct (N-terminal His-tag) shown in Figure 5.6, a construct with a C-terminal His-tag (hN1 23-24C_{His}) was designed and expressed to investigate the impact of orientation on binding (Figure 5.7) (See Chapter 3 – Protein Purification and Characterisation, Section 3.2.4.1). The differences in the His-tag lengths are unlikely to have an impact on the assay (N-terminal tag contains His_6, C-terminal tag contains His_8).

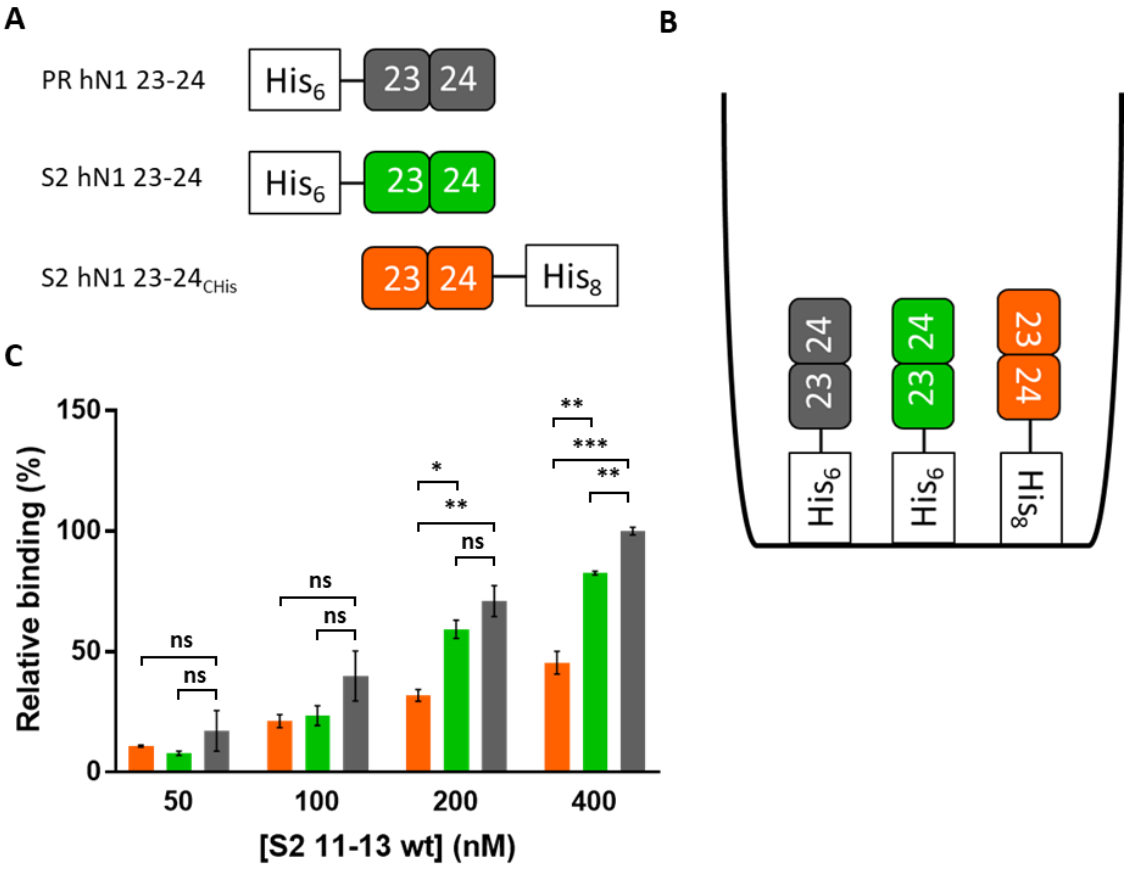


Figure 5.8 Investigating different orientations of hN1 23-24 constructs in the plate assay: A) The three hN1 23-24 constructs investigated in the plate assay were expressed from the prokaryotic/refold (PR) or S2 systems. These include His-tags in the positions depicted in the schematic. B) The proteins are likely to be presented on the Ni^{2+}-plate surface in different orientations due to differences in positions of the His-tag. C) PR hN1 23-24 (grey), S2 hN1 23-24 (green), S2 hN1 23-24$_{CHis}$ (orange) were attached to a Ni^{2+} plate via their His-tag. S2 11-13 wt was then added to the wells at a range of concentrations. Data shown combines results from three independent plates, each of four repeat wells. Binding is shown relative to PR hN1 23-24. Comparisons between groups were performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. ns = not significant; ***P < 0.001; **P<0.01, *P<0.05. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

The results (Figure 5.8) show a large increase in signal in S2 hN1 23-24 relative to S2 hN1 23-24 C_{His}. This suggests that either the interaction observed has a dependence on the orientation of EGF23-24, or difference in the relative exposure of domains in the two orientations leads to a decrease in accessibility to the binding site for the LBR in this assay. Considering that PR hN1 20-23 does not bind to the LBR (EGF23 is distal to the plate in

20-23), this suggests that EGF24 is the site of the interaction with hN1 11-13. In the S2 hN1 23-24 C_{His} construct, EGF24 is closest to the plate so it is feasible that EGF23 is occluding the LBR from binding to EGF24. Alternatively, the orientation of EGF24 could be important. Longer Ax constructs (C-terminal His-tag) were also investigated in the plate assay, in which EGF24 was not the most distal domain to the plate, these showed less binding than the binding of S2 hN1 23-24 C_{His} (data not shown), suggesting that occlusion by more distal domains in the Ax construct is likely a contributing factor to the reduced binding seen in these cases. Due to the higher signal achieved with PR hN1 23-24 compared and N-terminal His-tagged S2 material, these constructs were used for future studies to probe the interaction of the hN1 23-24 with the LBR. Furthermore, the lack of binding to PR hN1 20-23 confirms that the identical tags present are not responsible for the binding observed (Figure 5.5).

5.2.1.5 Interactions with PR hN1 23-24

5.2.1.5.1 Interaction of hN1 11-13 variants

In order to investigate the location of the interaction in the LBR, a number of S2 hN1 11-13 variants were generated. These variants each resulted in the disruption of a different face of EGF11 and EGF12 and henceforth referred to as directional variants (see Figure 5.9). EGF13 was not mutated because of previous evidence that an antibody which disrupted binding of the LBR and Ax region epitope mapped to the N-terminus of EGF12 at the EGF11-EGF12 interface (Whiteman et al, unpublished data). These directional

variants were expressed and purified in the S2 system (see Chapter 3 - Protein Purification and Characterisation, Section 3.2.3).

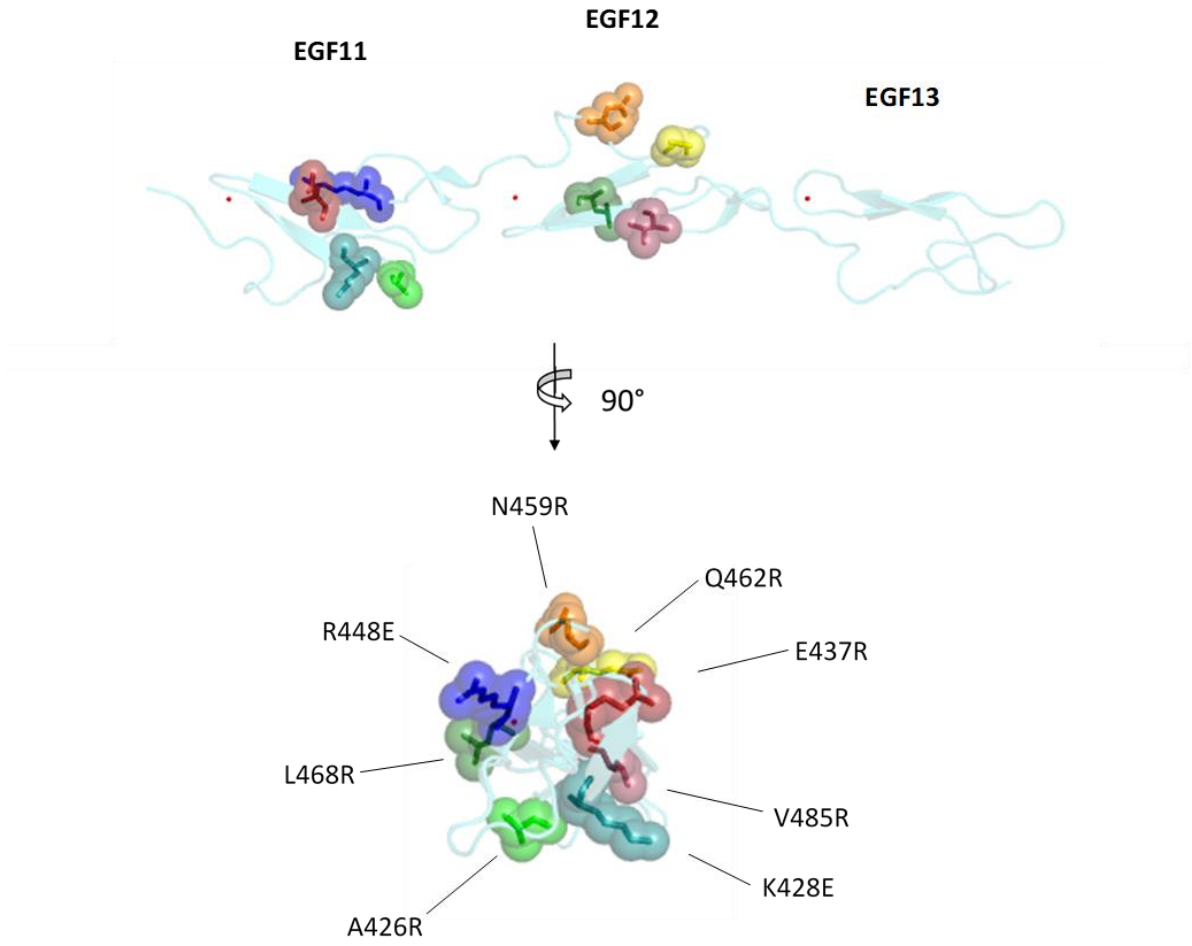


Figure 5.9 The design of directional variants in EGF11 and EGF12: hN1 11-13 (PDB ID: 2VJD) (cyan) is shown with bound Ca^{2+} ions (red dots). Residues of interest are coloured and shown as spheres and sticks. EGF11 = A426 (light green), K428 (teal), E437 (dark red), R448 (dark blue). EGF12 = N459 (orange), Q462 (yellow), L468 (dark green), V485 (pink). An end-on view is shown to demonstrate that each surface of each domain is disrupted.

Engineering disruptive surface residues eg Arg/Glu (long residues that introduce a reverse charge) into each of the four faces of the EGF domain enabled more precise localisation of binding site. The introduction of these surface residues was designed to disrupt the interaction with Ax if it occurred on this face. Plate assays were conducted to investigate

the impact of these substitutions in each face of 11-13 in EGF12 (Figure 5.10), and EGF11 (Figure 5.11).

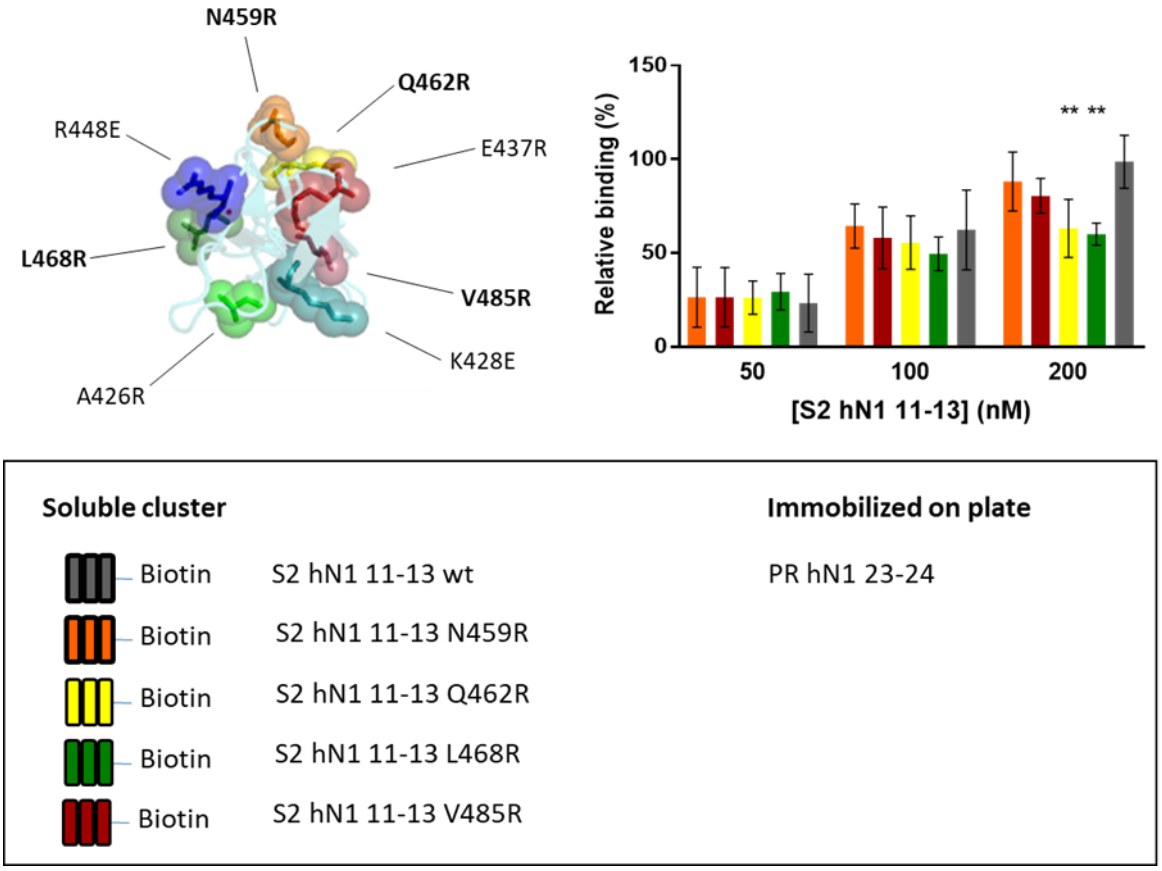


Figure 5.10 Analysis of hN1 11-13 EGF12 variants in the plate assay: Investigating the interaction with PR hN1 23-24. The structure of hN1 11-13 (PDBID = 2VJD) is shown with EGF12 substitutions labelled in bold. The graph shows binding of these EGF12 mutants in the plate assay. Data for wt (grey), N459R (orange), V485R (maroon), Q462R (yellow) and L468R (green) are shown combined from three independent plates, each of two wells. Near equivalent biotinylation of these S2 hN1 11-13 variants is shown in Chapter 3 – Figure 3.21. Comparisons between each variant and wt were performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. **P<0.01. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

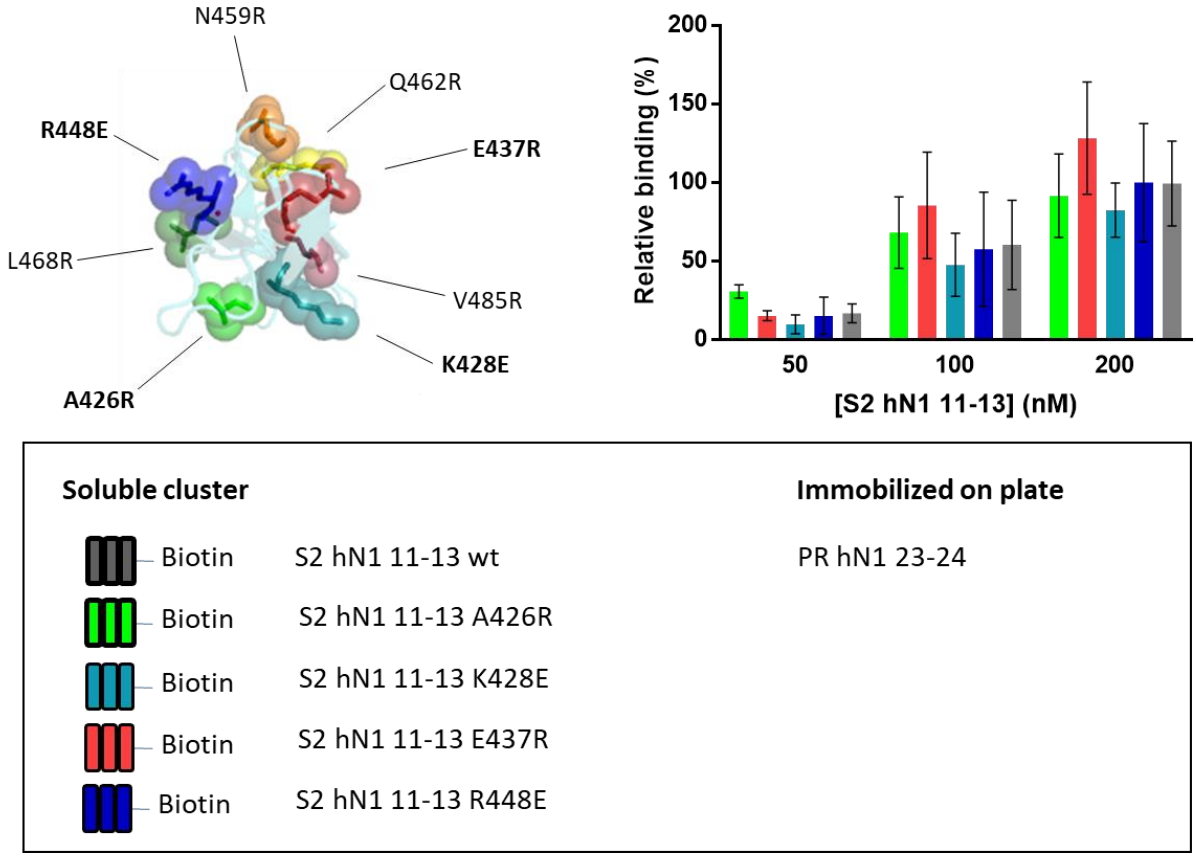


Figure 5.11 Analysis of hN1 11-13 EGF11 variants in the plate assay: Investigating the interaction with PR hN1 23-24. The structure of hN1 11-13 (PDBID = 2VJD) is shown with EGF11 substitutions labelled in bold. The graph shows binding of these EGF11 mutants in the plate assay. Data for wt (grey), A426R (green), E437R (red), K428E (cyan) and R448E (blue) are shown combined from three independent plates, each of two wells. Comparisons between each variant and wt were performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

After seeing some indication that two EGF12 variants (L468R and Q462R) disrupted the interaction with PR hN1 23-24, the plate assays were repeated to investigate the impact of EGF11 and EGF12 variants on the interaction with S2 hN1 23-24 (Figure 5.12). This second set of assays used independent preps (prep 2) of S2 hN1 11-13 L468R and S2 hN1 11-13 Q462R (see Chapter 3 – Protein purification and characterisation, Section 3.2.3).

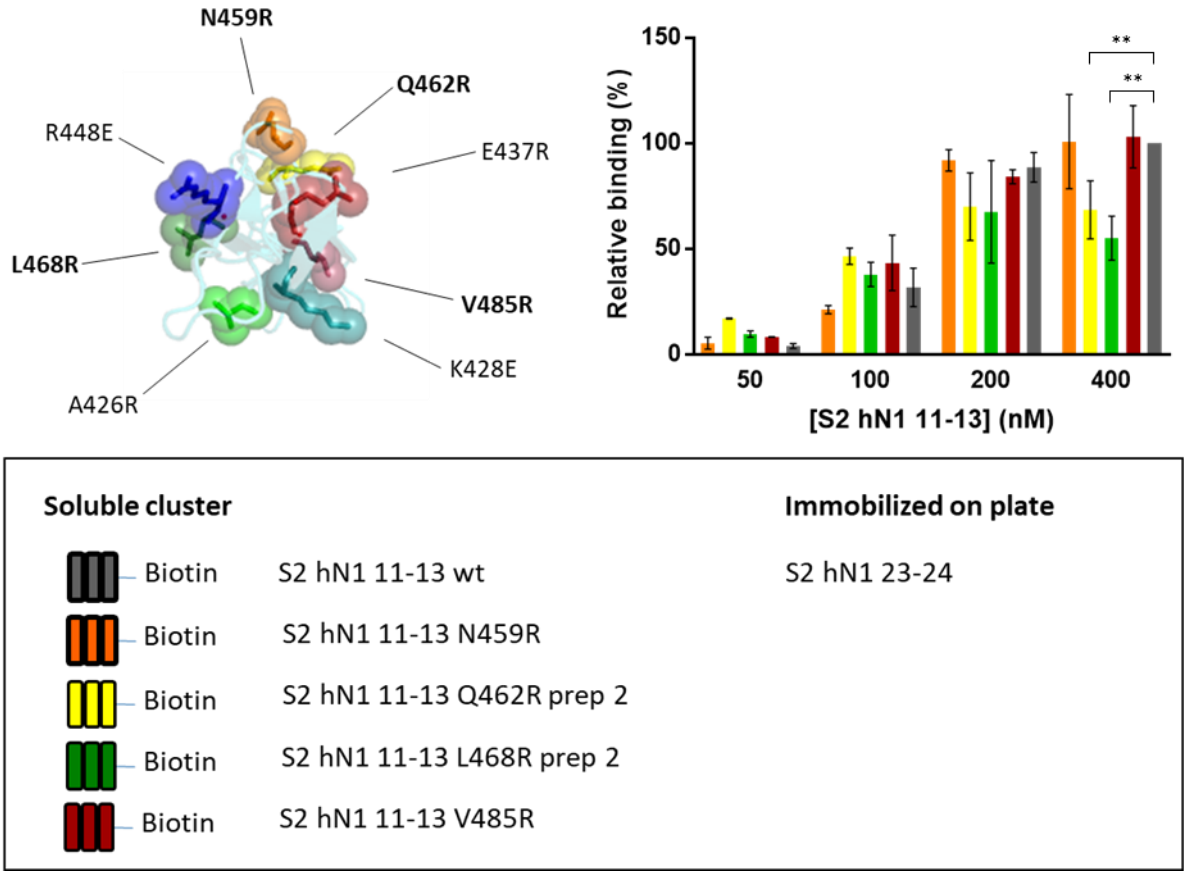


Figure 5.12 Independent analysis of hN1 11-13 EGF12 variants in the plate assay: Investigating the interaction with S2 hN1 23-24. The structure of hN1 11-13 (PDBID = 2VJD) is shown with EGF12 substitutions labelled in bold. The graph shows binding of these EGF12 mutants in the plate assay. Data for wt (grey), N459R (orange), V485R (maroon), Q462R prep 2 (yellow) and L468R prep 2 (green) are shown combined from two independent plates, each of one repeat well. Comparisons between variants and wt were performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. **P<0.01. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

The results of the plate assay suggest that EGF11 is unlikely to be involved in the interaction face since EGF11 variants do not lead to a significant decrease in LBR binding to the Ax region compared to wt 11-13. The results from EGF12 variants show a decrease in binding in both the prokaryotic/refold and S2 hN1 11-13 for L468R, which is an essential ligand binding residue on the face to which ligand is known to bind, and Q462R, which is located on the opposite face of EGF12 to L468R.

5.2.1.5.2 Independent confirmation of S2 hN1 11-13 Q462R sample integrity

Having already shown that S2 hN1 11-13 L468R binds ligand significantly weaker than S2 hN1 11-13 wt (Figure 5.4), the binding of S2 hN1 11-13 Q462R prep 1 to ligand was investigated (Figure 5.13). No impact of Q462R is expected on ligand binding as the substitution is on the opposite face to the ligand binding face. The results of this experiment demonstrated that hN1 11-13 Q462R does not disrupt the interaction with ligand hJag1, but does with hN1 23-24, suggesting binding of ligand and Ax to S2 hN1 11-13 is not identical.

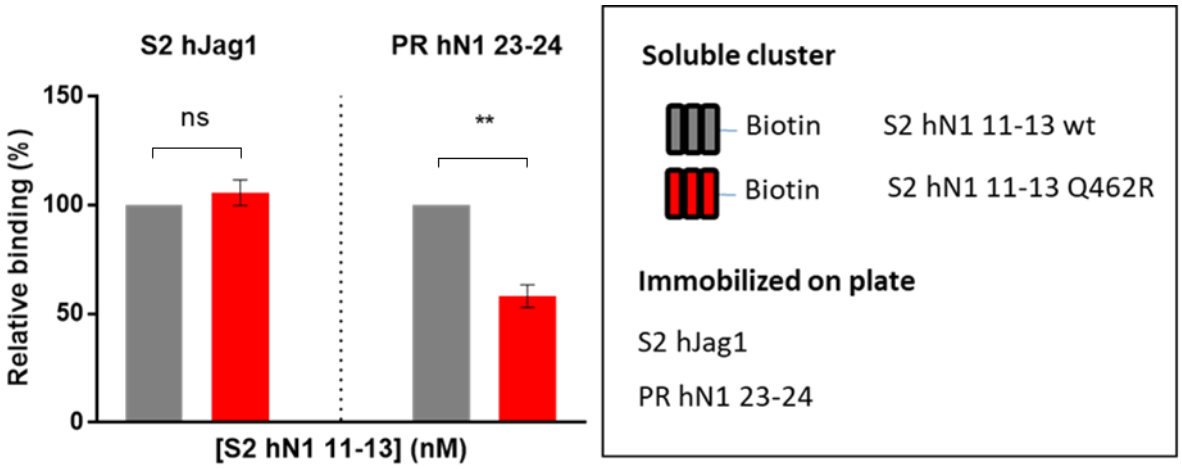


Figure 5.13 Comparison of hJag1 and PR hN1 23-24 binding to hN1 11-13 wt and Q462R at 400 nM S2 hN1 11-13: The graph shows binding of these EGF12 variants in the plate assay. Data for wt (grey), Q462R (red) are shown combined from two independent plates, each of one repeat well. Statistical significance was calculated compared to S2 hN1 11-13 wt. Comparisons between variants and wt were performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. **P<0.01. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

These results suggest that either the Ax region interacts across two faces of the LBR, or that one of these variants leads to an indirect effect on binding by causing a conformational change leading to the disruption of the binding face for Ax.

5.2.1.5.3 Investigation into sugar-dependence of the interaction

The two faces of EGF12 implicated in the plate assay to be involved in binding to the Ax region contain either an O-Fuc (L468 face), or an O-Gluc (Q462 face) modification (Figure 5.14).

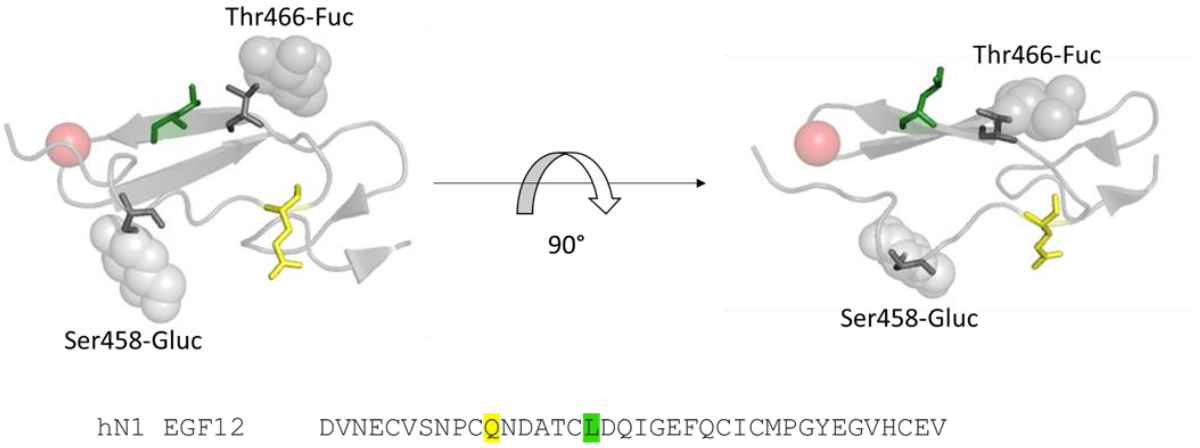


Figure 5.14: The two faces of EGF12 implicated in binding hN1 23-24: EGF12 is shown (grey) including a bound Ca^{2+} (red) the location of the substituted residues implicated to impact the interaction from the plate binding assay (Q462 (yellow), L468 (green)). These two residues are on opposite faces of EGF12, and each face includes a site of glycan modification (glycan presence has been observed in the crystal structure of S2 hN2 11-13 (Suckling et al, 2017). The residues and glycans on the two faces are labelled. The position of the substituted residues in the peptide sequence of hN1 EGF12 is shown.

Since S2-expressed (glycosylated) hN1 11-13 binds to Ax in the plate assay and PR-expressed protein does not, it can be inferred that the presence of glycans enhances the binding of 23-24 to 11-13. To explore the role of glycans further, in collaboration with Prof R Haltiwanger (Stony Brook University) specific glycosylation modifications were introduced to purified PR hN1 11-13 protein to investigate if the addition of any one modification would restore binding activity to levels similar to the fully modified S2 construct (Figure 5.16).

Investigation of the role of O-Fuc added to T466 in EGF12 was carried out by Dr P Whiteman. No improvement to binding PR hN1 23-24 was observed on introducing the O-Fuc modification at T466 (Whiteman, unpublished data), suggesting that the O-Fuc modification in isolation is not sufficient to restore binding. This is in contrast to ligand binding which showed improvement on addition of this modification (Taylor et al, 2014)

O-Gluc sites in EGF11 and EGF12 were enzymatically extended with single, double, and triple modifications corresponding to the length of the glycan (See Chapter 1 – Introduction, Section 1.3.4.3). The mono- di-, and tri- saccharide modified peptides were then investigated for binding to PR hN1 23-24 in the plate assay (Figure 5.15).

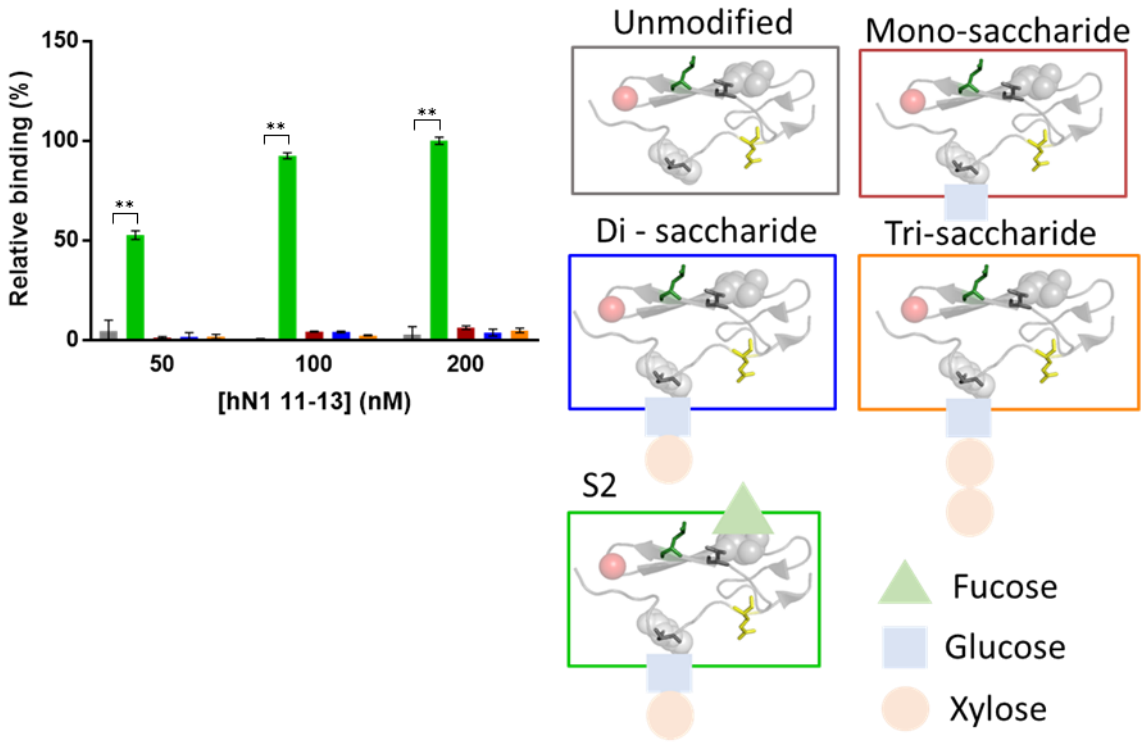


Figure 5.15 O-glucosylated PR hN1 11-13 interacting in the plate assay: Column colours in the graph correspond to the species shown within the box of that colour below the graph (EGF12 shown). Results shown from two independent repeats and shown relative to binding of 200nM S2 hN1 11-13 wt. Comparisons between groups were performed with a two-tailed unpaired t-test. Values are presented together with the mean ± SD. **P<0.01. Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

These results suggest that the O-Fuc or O-Gluc modifications separately cannot restore binding activity, and that both are required for binding. This supports the observation that both L468 and Q462 faces are essential in binding to the Ax region, either because of an extended interface which covers both faces, or that one of these faces (and glycan) provides an indirect structural effect to stabilise the conformation of the binding site to allow Ax binding. Considering EGF12 has been previously demonstrated to be involved in ligand binding (Luca et al, 2015), the residues inferred from the plate assay for binding to Ax were mapped onto EGF12 and compared to the previously characterised ligand-binding face (Figure 5.16).

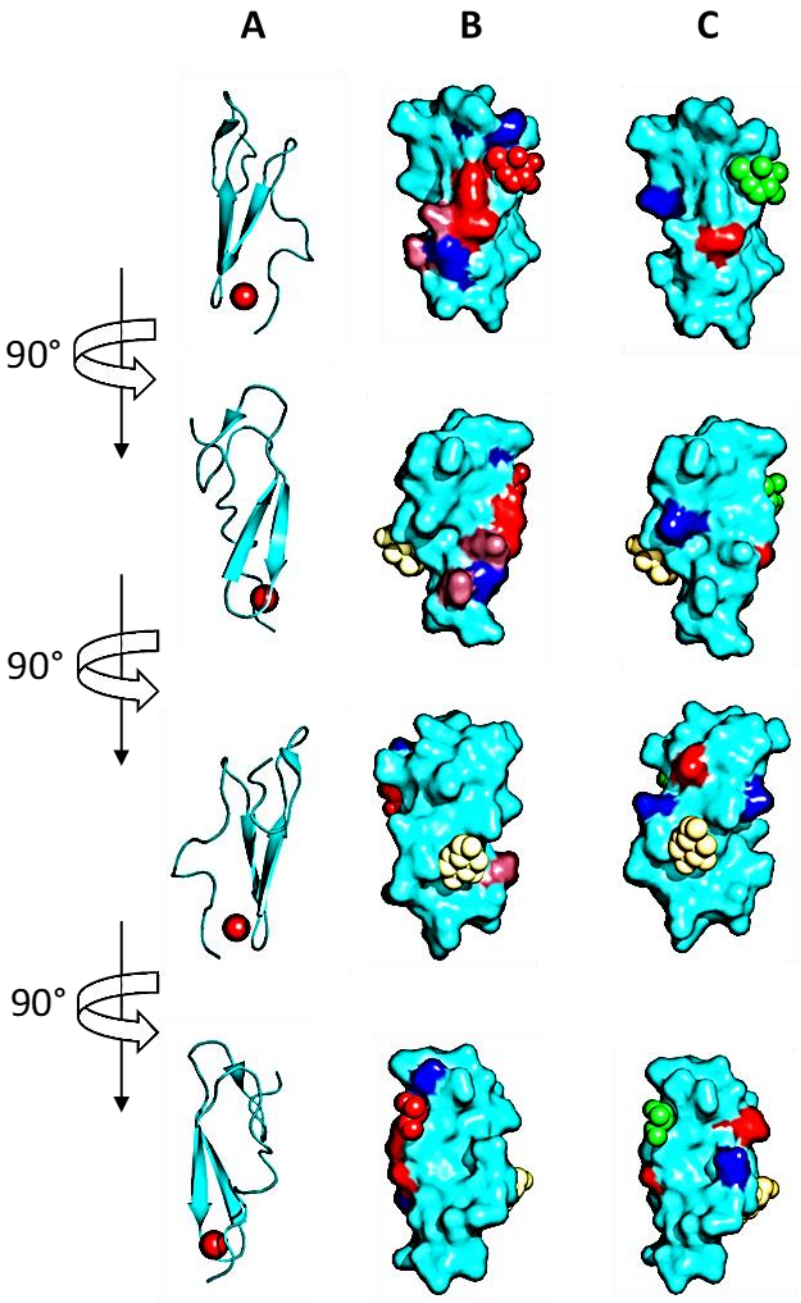


Figure 5.16 The binding site of Ax in EGF12: A) The four faces of EGF12 (cyan) are shown as a cartoon, along with the Ca^{2+} ion (red sphere). B) The surfaces of the four faces of EGF12 (cyan) are shown. Residues that have been shown to be essential for ligand binding are shown in red (L468, I477), residues that contribute to ligand binding are shown in purple (E473, Q475), and residues that are not involved in ligand binding are shown in blue (Q470, M479) (Whiteman et al, 2013). Fucose attached to Thr466 is shown as red spheres and glucose attached to Ser458 is shown as yellow spheres. C) The surfaces of the four faces of EGF12 (cyan) are shown. Residues possibly contributing to Ax binding are shown in red (L468, Q462), residues not involved in Ax binding are shown in blue (N459, V485). Fucose attached to Thr466 is shown as green spheres and glucose attached to Ser458 is shown as yellow spheres. L468 is involved in a hydrophobic interaction across the β-sheet, stabilizing the hairpin.

A limitation of the plate assay is the inability to guarantee that equivalent amounts of Ax are bound on the surface of the plate when comparing binding of different constructs to hN1 11-13, therefore different Ax variants cannot be accurately quantitatively compared. In order to address this, a flow cytometry assay was used to investigate the impact of substitutions in hN1 23-24 on binding to the LBR.

5.3 Investigating the interaction in the flow cytometry assay

The flow-cytometry assay is a method for investigating the interaction of multidomain constructs with cells expressing the Notch receptor. This enables the probing of interactions in the context of the transmembrane receptor rather than the interaction between small purified fragments.

The flow cytometry assay has been previously established for investigating interactions between ligand with Notch expressing cells (Cordle et al, 2008, Whiteman et al, 2013, Taylor et al, 2014). The cells used in this assay are HEK293T cells, which constitutively express hN1. The assay involves forming a complex of purified biotinylated protein by binding to a fluorescent bead coated in streptavidin. These coated beads are then mixed with cells and then cells are sorted after washing. The fluorescent signal associated with the cells corresponds to the number of fluorescent beads bound and so depends on the interaction of the protein coating the beads with the cell surface. Due to the absolute reading depending on many factors including the amount of Notch expressed on the cell surface, binding shifts on a day-by-day basis are variable and so results are always compared to a PR hN1 23-24 wt sample collected at the same time (Figure 5.17). To

confirm the precision of this assay, it is important to ensure that equivalent amounts of biotinylated protein is bound on each bead, and that the interaction with the cells is specific for the Notch receptor on the cell surface. To ensure that fluorescent beads are loaded equivalently, an excess of biotinylated protein is added to beads. In parallel, western blots for biotin on the washed beads were conducted for every experiment. This ensures that any differences in binding are not due to variable saturation of the beads.

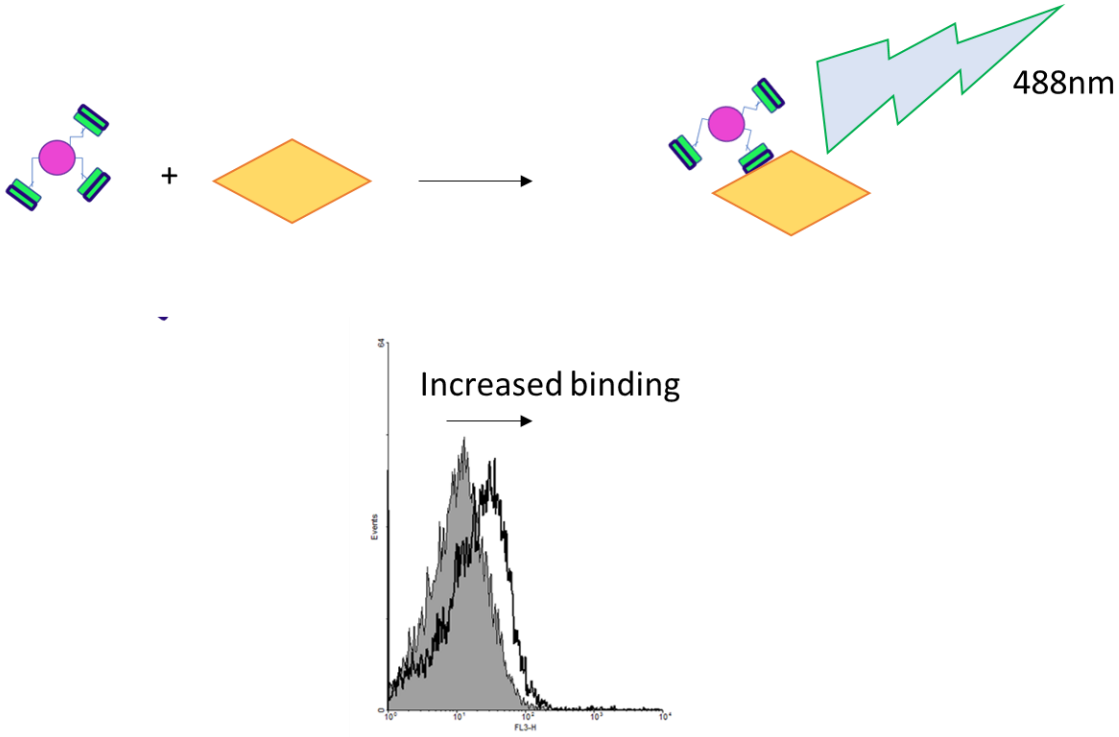


Figure 5.17 Schematic of the flow cytometry assay: The biotinylated Ax construct (green blocks) is attached to a fluorescent bead (purple circle) via streptavidin on the bead. These are then exposed to HEK293T cells (yellow rhombus) expressing hN1 on their cell surface. The cells are then measured in a flow cytometer for fluorescence corresponding to the bead. An increase in cell fluorescence corresponds to an interaction between the Ax construct and the cell. This increase in fluorescence corresponds to a shift to the right in the fluorescence histogram (grey area = background fluorescence of cells + beads, black line = cells + Ax-bound beads). 10,000 cells are sorted per experiment and repeating the non-specific binding reading produces a curve which overlays almost exactly with the first repeat indicating that even a small shift in fluorescence is significant.

5.3.1 Establishing the flow cytometry assay for investigating the interaction of LBR with Ax

The conformation or arrangement of the Notch receptor on the cell surface is yet to be determined. Initial experiments showed binding of PR hN1 23-24 to HEK293T cells but binding of S2 hN1 11-13 did not occur (Figure 5.18). This suggests either the Ax region of hN1 on the HEK293T cells is in a conformation not able to bind 11-13 on the bead, or that the complex including the fluorescent bead is too large to enable the S2 hN1 11-13 to access the Ax region which may be more proximal to the cell than 11-13 on the receptor. The control construct Fbn1 EGF32-34 (provided by Dr S Jensen, University of Oxford) was used to ensure that binding was due to the presence of the Abruptex domain rather than tags. This construct was similar to PR hN1 23-24 in terms of tag organisation (N-terminal His tag, C-terminal Biotin tag) and size. The lack of binding of this control construct suggested that the tags did not bind to the cells and the interaction with HEK293T cells requires EGF23-24 to be present in the construct (Figure 5.18). Beads treated with BSA were shown not to bind and were used for the negative control background binding in future assays. When blotting to show equivalent binding (Figure 5.18C), additional bands likely corresponding to dimers are seen in all three constructs. This dimerization is unlikely to contribute to binding of hN1 23-24, as the Fbn1 32-34 construct appears to show similar dimerization as hN1 23-24 but does not bind to cells. Dimers might possibly form mediated by the His-tag or through an interaction between biotinylated protein, although further investigation would be required to prove this.

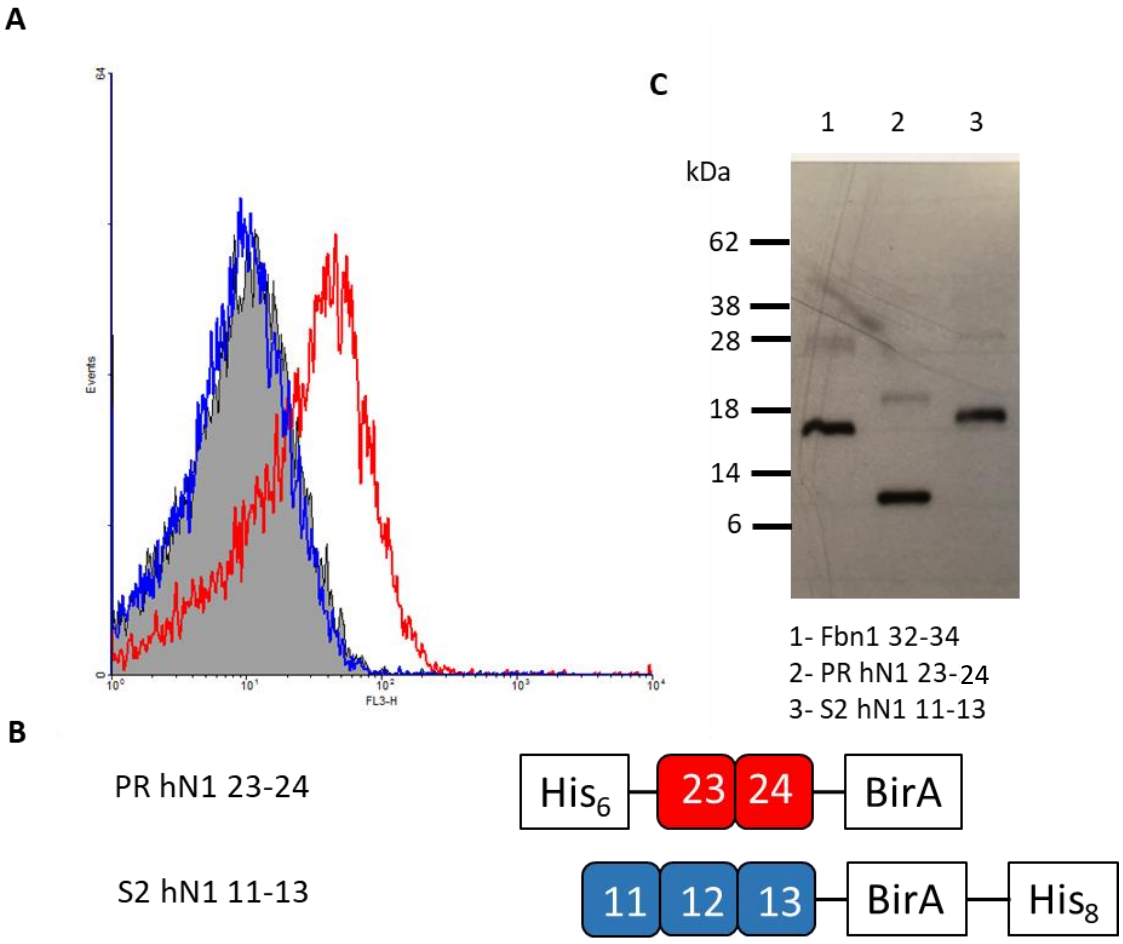


Figure 5.18 PR hN1 23-24 binds to cells, S2 hN1 11-13 does not bind to cells: A) The binding histograms show a shift indicative for binding for hN1 23-24 (red) but not S2 hN1 11-13 (blue) relative to background binding of control construct (Fbn1 32-34, grey) representative of two independent experiments. B) Schematic diagrams of PR hN1 23-24 (red) and S2 hN1 11-13 (blue) with colours corresponding to curves in panel (A). Beads from the flow cytometry show equal loading of biotinylated construct (C). Fbn1 32-34 has been extensively characterised to show it is correctly folded (Jensen et al, unpublished data).

The interaction observed with PR hN1 23-24 was probed using monoclonal antibodies against the N-terminus of EGF24, and the N-terminus of EGF12 (Figure 5.19). This enabled more precise definition of the site of interaction on the cell surface, which suggest that a specific interaction is taking place between Ax on the bead, and hN1 11-12 on the cell (Figure 5.19).

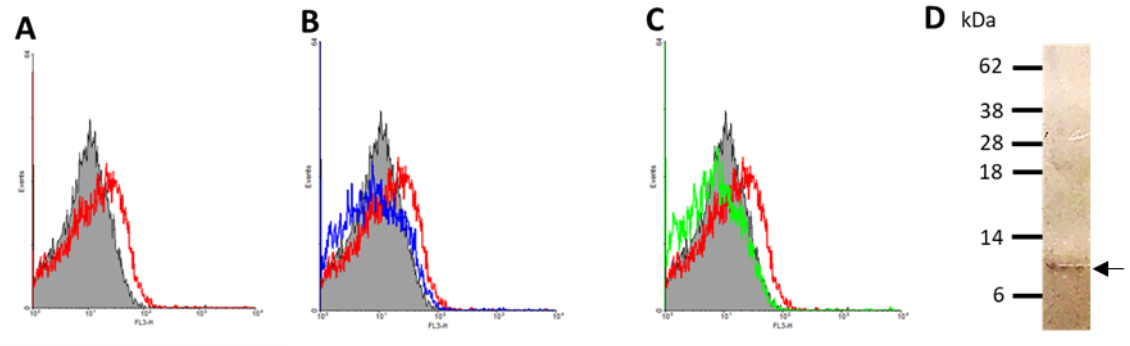


Figure 5.19 Antibody abrogation suggests a specific interaction of PR hN1 23-24 with EGF11-12 on the surface of the HEK293T cells: A) binding is observed for PR hN1 23-24 (red) relative to background binding (grey). The shift seen is small due to day-to-day variation but was observed in multiple independent repeats and with larger shifts. This binding can be prevented by treating with mAbs against B) hN1 EGF12 (blue), and C) hN1 EGF24 (green). D) Western blot probing with a streptavidin conjugate shows biotinylation of the PR hN1 23-24 construct used in this assay (indicated with black arrow), matching expected MW of ~12000. This knockdown in binding was demonstrated in multiple independent assays.

5.3.2 Interactions with other Abruptex constructs

The interactions in the flow cytometry assay mirror the interactions seen in the plate assay well, including binding of an N-terminal His-tag construct from the S2 system (S2 hN1 23-24) (Figure 5.20), but no binding was observed for an S2 hN1 23-24 C_{His} (Boyce et al, data not shown), suggesting either orientation of EGF24, or distance from the bead is important in enabling this interaction.

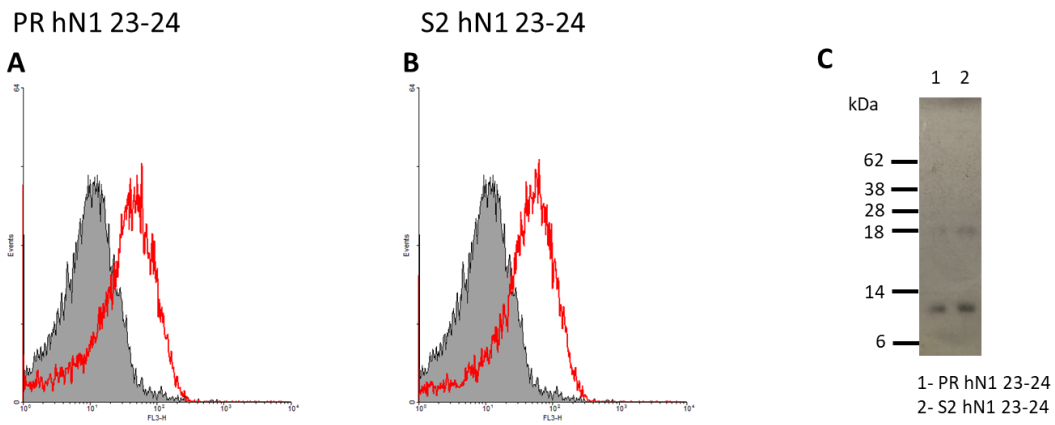


Figure 5.20 hN1 23-24 expressed in both the prokaryotic/refold and S2 systems bind to cells in the flow cytometry assay: A) Binding is observed for PR hN1 23-24 (red) relative to background binding (grey). B) Binding is observed for S2 hN1 23-24 (red) with background binding shown in grey. These fluorescence shifts are representative of independent assays. C) Western blot probing with a streptavidin conjugate shows biotinylation and equivalent loading of the constructs in this assay, matching expected MW of ~12000.

5.3.3 Identification of the interaction face of hN1 23-24

A decrease in binding has been previously observed for the D909V variant (Whiteman et al, unpublished data) (Figure 5.1). In order to further characterise the interaction face of EGF24 in more detail, a number of variants of hN1 23-24 were produced. EGF24 was targeted for substitutions as it is indicated in the plate assay to be essential for this interaction (see Section 5.1.1.1). The variants engineered have substitutions disrupting each face of EGF24. However, none of these variants showed a decrease in binding to cells compared to wt (Figure 5.21). As previously seen in Figure 5.18, bands likely corresponding to dimerization of the hN1 23-24 variant constructs are seen. This dimerization is unlikely to contribute to binding considering the control construct Fbn1 32-34 shows similar dimerization but does not bind to cells (Figure 5.18).

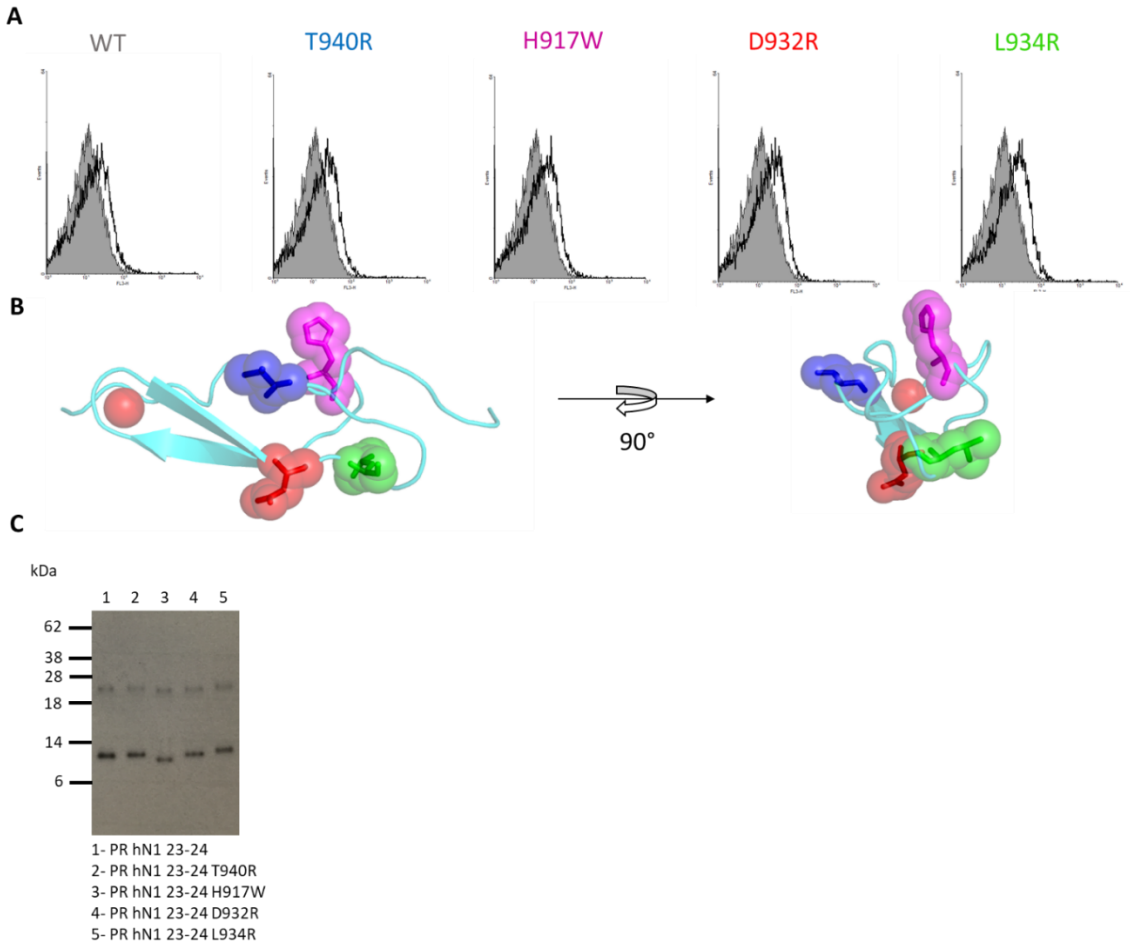


Figure 5.21 Interactions of PR hN1 23-24 directional variants in the flow cytometry assay. A) Binding of each EGF24 directional mutant is shown for each respective construct (black line) compared to background binding (grey). B) The EGF24 directional mutants are shown on EGF24 (cyan, Ca^{2+} = red sphere) with coloured residues corresponding to the titles of the graph in panel A: T940 (blue), H917 (purple), D932 (red), L934 (green). C) Western blot probing with a streptavidin conjugate shows biotinylation and equivalent loading of the constructs in this assay. Expected MWs match what is expected for these constructs.

Considering the position of these substituted residues on EGF24, these data indicate that the primary site of interaction is towards the N-terminus of EGF24, or dependent on the moderately-high affinity Ca^{2+} site being saturated. The D909V mutation corresponds to the Ax mutant D948V observed in EGF24 in *D. melanogaster* (See Chapter 1 – Introduction, Section 1.4.3.1) and has previously been shown to decrease Ca^{2+} affinity of EGF24 (Redfield et al, unpublished data) and so is likely to impact the rigidity of the

interface between EGF23 and 24. Therefore it is unclear if this mutation impacts binding because it is involved in direct contacts with the LBR, or if this impact is due to a change in flexibility at the EGF23-24 interface or both.

In addition to showing that EGF24 was essential for the interaction, the construct PR hN1 23-24cl was tested with the His tag cleaved off (PR hN1 23-24cl) (Figure 5.22). The cleavage of the His-tag eliminated binding of the construct to cells, suggesting that the His-tag must have a role in the interaction in this assay, possibly positioning or localising the bead complex for the interaction. However, considering the binding observed in the plate assay centred around EGF24 and the observation that the control construct EGF20-23 (with His-tag) does not bind in the plate assay (Figure 5.5) suggests that the role of the His-tag in binding in the plate assay is negligible. Furthermore, in the flow cytometry assay, EGF24 is required for the interaction, and the His-tag alone is not sufficient for binding (Fbn1 32-34 with His-tag does not bind to cells, see Figure 5.18). The D909V mutant disrupting binding furthermore supports that the EGF24 interaction with LBR observed here is the interaction observed in Sharma et al, 2013, and Pei and Baker, 2008 (See Chapter 1 – Introduction, Section 1.4.3.2).

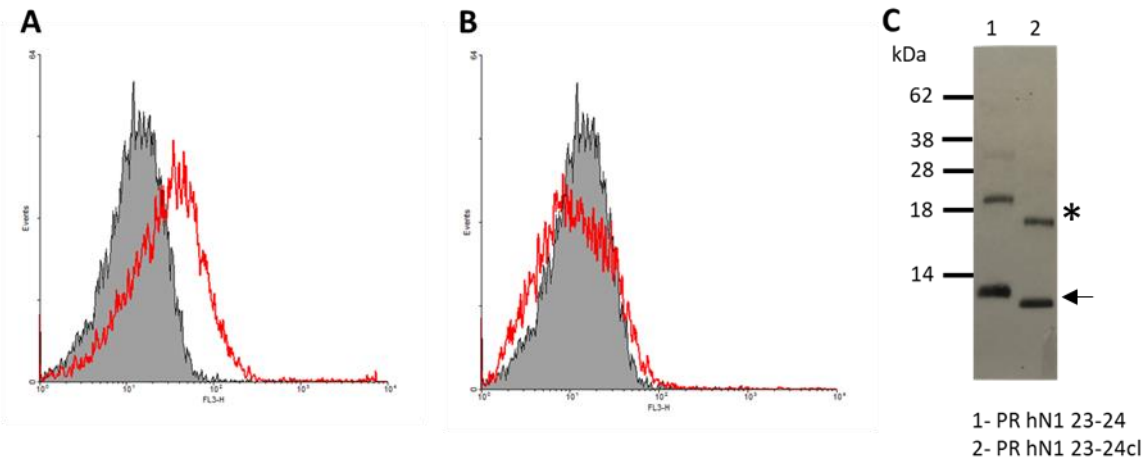


Figure 5.22 Elimination of the His-tag leads to a loss of binding: A) As seen previously, PR hN1 23-24 wt binds to cells (red) relative to background bead binding (grey). B) Cleavage of the His-tag from the PR hN1 23-24 construct to produce PR hN1 23-24cl leads to a loss of binding (red), the cleaved construct binds no better than background bead binding (grey). These fluorescent shifts are representative of multiple independent repeats. C) Western blot probing with a streptavidin conjugate shows biotinylation and equivalent loading of the constructs in this assay (indicated by black arrow). A minor contaminant is present in both samples which appears to shift with mass of hN1 23-24 (*). The main band matches expected MW for these constructs.

5.4 Conclusions and discussion from binding assays

The plate binding assays suggest that EGF12 is the centre of the interaction with Ax, and this interaction can be disrupted by substituting residues on two opposite faces of EGF12, one of which is the same face as ligand binding (L468 face). The Q462R substitution on the opposite face to L468 does not impact ligand binding, suggesting that binding of LBR to Ax is different to ligand binding. The glycan modifications on these two faces are also essential for Ax binding and binding only occurs when both faces are glycosylated. The plate assay results also suggest that the interaction is centred around EGF24 in the Ax.

Looking at the results from the flow cytometry assay, the equivalent loading of hN1 23-24 and Fbn1 32-34 on the fluorescent beads in the flow cytometry assay suggests that oligomerisation of hN1 23-24 does not occur and so does not impact binding. Furthermore, no evidence of oligomerisation of this construct is observed in the case of the S2 construct during gel filtration chromatography. The observation that other His-tag-containing constructs do not interact with the cells, and that antibodies specific to hN1 EGF12 or EGF24 abrogate the interaction, indicates that the interaction between Ax and LBR can be observed in this flow cytometry assay. The impact of the D909V mutant on binding supports that, even though a His-tag is required to observe binding, the interaction is centred around EGF24. This substitution also suggests the interaction involves the N-terminal region of EGF24, either directly contacting this area, or dependent on the Ca^{2+}-binding state of this area. The observation that the His-tag on the Ax is required for binding suggests that this tag has a role in mediating the interaction observed in this assay in some way, possibly localising the construct to the cell surface for long enough to allow binding to the LBR on the Notch receptor.

Other details of this interaction can be suggested, firstly the fact that the Ax mutations are spread along the surface of Ax and that the SAXS-derived structure of hN1 and hN2 23-27 is elongated (Chapter 4 – Low Resolution Structural studies of the Abruptex region, Section 4.2.7) it is likely that the interaction is spread over the length of Ax region in the receptor. This is supported by the observation that the interaction between EGF12 and EGF24 is weak compared to ligand in the plate assay, suggesting an extended face would be required to enhance affinity. Taking these observations into account, putative models for the interaction were generated corresponding to parallel and anti-parallel

orientations (Figure 5.23). In the anti-parallel model, the extended Jagged/Serrate family interaction face of the LBR (EGF8-9, See Chapter 1 – Section 1.4.2) would interact with the Ax (EGF27-28).

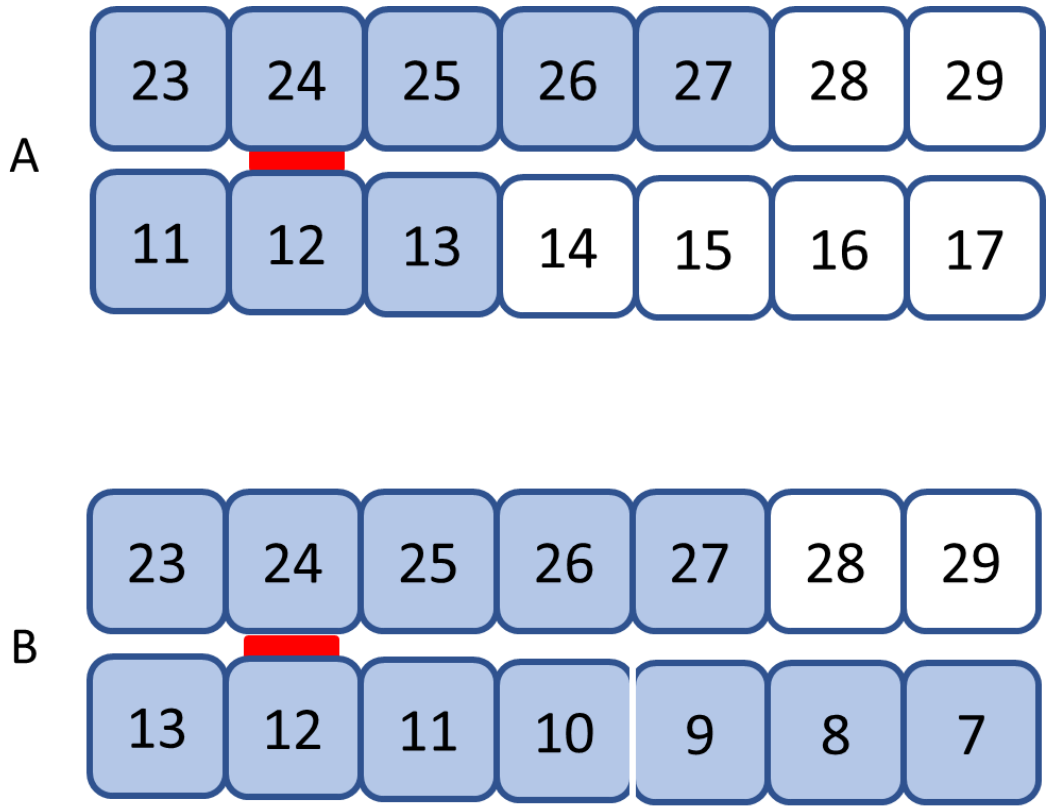


Figure 5.23 Schematic models for the interaction of Ax with the LBR: Regions containing the LBR and Ax are shown. Parts of these regions have been structurally characterised (blue shading) and some have unknown structures (white shading). The EGF9-10 interface is suggested to be flexible and is coloured white. The interaction between EGF12 and EGF24 observed in the binding assay is shown in red. Possible models are interaction which are parallel (A) or anti-parallel (B).

6.

Investigating the role of the Abruptex region in Notch activation

6.1. Introduction

In order to investigate the impact of Abruptex mutations on Notch activation at the cell surface, a cellular activation assay was used. This assay allows a quantitative readout of Notch activation using a luciferase complementation-based method (Ilagan et al, 2011).

The assay involves activation of a hybrid mouse Notch1 (mN1) receptor which includes half of luciferase at the C-terminal of Notch (on the NICD). On successful binding and cleavage of the receptor (see Chapter 1 – Introduction, Section 1.2.3), the NICD moves to the nucleus and interacts with the Notch transcriptional co-effector RBPJK which includes the other half of luciferase (RBPJK-luc). On interaction of the NICD and RBPJK that follows successful

activation, active luciferase is formed, and luminescence is produced on addition of substrate in proportion to the amount of Notch activation (Figure 6.1).

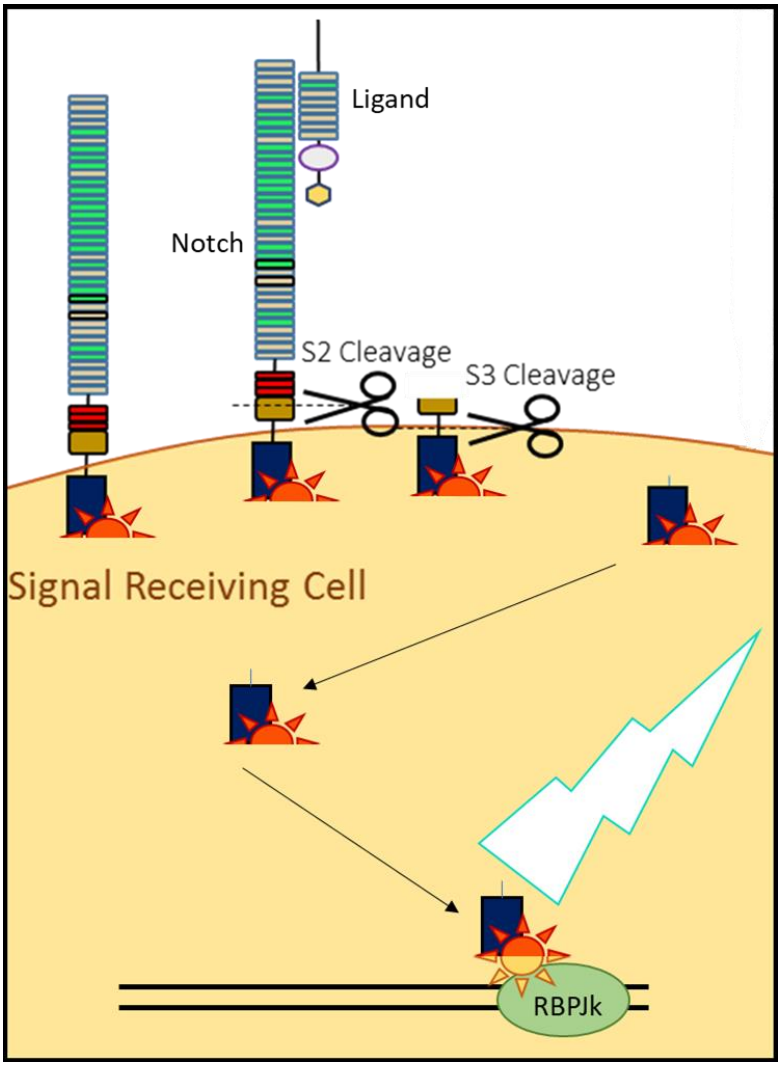


Figure 6.1 Schematic of the activation assay: The Notch receptor and NICD (dark blue) with N-terminal region of luciferase (orange). On ligand induced activation, S2 and S3 proteolytic cleavage reactions occur at the cell surface, which lead to the release of the NICD (navy) to the nucleus. The NICD interacts with RBPJK (green sphere with C-terminal region of luciferase (yellow) which is localized on the cellular DNA (black lines). This interaction of NICD with RBPJK leads to the interaction of the two halves of luciferase and therefore ability to produce a luminescent signal (white lightning) in the presence of substrate luciferin and ATP.

6.2. Advantages and disadvantages of the activation assay

This activation assay is useful because specific receptor variants can be engineered and then assayed for signal. This enables the investigation of the activation capability of designed receptor variants to probe the Abruptex region.

Although the mN1 variant receptor is expressed in large amounts in this system from a cytomegalovirus promoter, in addition to the endogenous constitutive hN1 expression, activation is only recorded through the variant receptor transfected into the cell because of the presence of luciferase. Furthermore, it cannot be ruled out that the presence of the luciferase alters normal cellular process to enhance or decrease Notch activation in these cell lines. A control cell line which is only transfected with RBPJK-luciferase is always assayed alongside cell lines containing the Notch variants to indicate background activation under the specific conditions of a particular assay on a particular day.

Although the system assays activation through mN1 rather than hN1, these mammalian Notch 1 receptors are identical in terms of length. One difference between these receptors is that the hN1 receptor contains two additional non-canonical Cys residues in the Ax region, one in EGF25 and one in EGF27 (See Chapter 3 – Protein purification and characterisation, Figure 3.2). mN1 contains the same additional Cys residue in EGF25 but does not include the additional Cys in EGF27. Considering the results of the Chapter 4 – Low Resolution Structural Studies of the Abruptex region, which suggest that these additional cysteines do not greatly

impact structure, this difference between hN1 and mN1 is not likely to greatly impact function of this region in activation of the Notch receptor and so the results from this activation assay are applicable to understanding the ability of the Ax region to affect Notch function.

6.3. Establishing and validating the activation assay

In order to validate the activation assay, initial experiments were carried out to produce cells expressing mN1 wt and mN1 L468A. This substitution Leu468Ala substitutes a residue previously shown to be essential for ligand binding (See Chapter 5 – Interaction of the Abruptex region with the ligand-binding region, Section 5.1.1) and provides a useful control to show that Notch ligand-mediated activation is responsible for the luminescent signal observed. The procedure for producing the Notch variant cell lines is outlined in Figure 6.2.

HEK293T cells with RBPJK-luc integrated into their genome (HEK-RBP cells) were grown in the presence of 200 ng/ml puromycin to ensure that the RBPJK-luc remains integrated. In order to investigate the transfection efficiency of these cells, a GFP-containing plasmid provided by Dr S Jensen, University of Oxford was stably transfected using the method described in the Chapter 2 – Materials and methods, Section 2.3.2.2. Subsequently, GFP expression was observed in transfected cells (green) but not in non-transfected cells showing the transfection protocol was effective at delivering plasmid into cells (Figure 6.3).

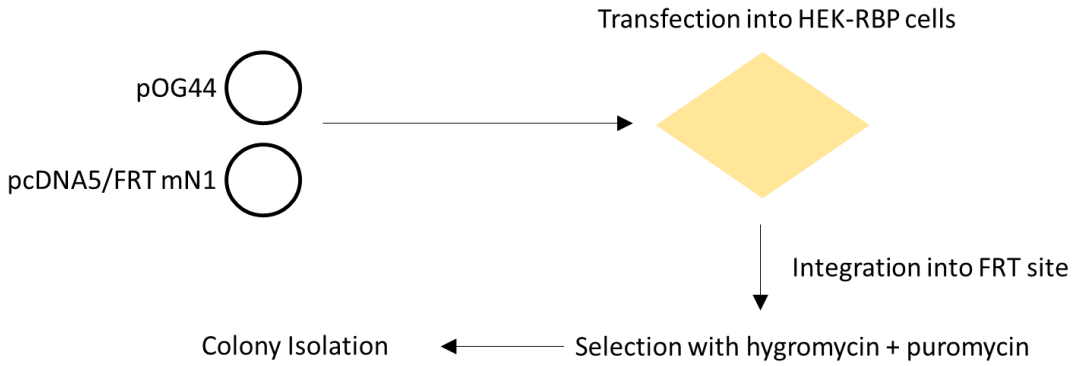


Figure 6.2 Preparation of stable clones for the activation assay: HEK-RBP cells (a Flp-in cell line containing an FRT site, expressing RBPJK-luc under puromycin selection) are transfected using the transfection reagent Mirus TransIT-LT1 with the plasmids pOG44 (coding for Flp recombinase) and pcDNA5/FRT mN1 containing an FRT site, mN1, and hygromycin resistance gene without a promoter. The Flp recombinase catalyzes the homologous recombination between the FRT sites, and so the mN1 and hygromycin resistance gene on the pcDNA5 plasmid are integrated into the host cell genome at the FRT site. Successful integration leads to the positioning of the hygromycin resistance gene downstream of an SV40 promoter, imbuing successful transfects with hygromycin resistance, as well as the mN1 cDNA sequence downstream of a CMV promoter. Following selection with hygromycin and puromycin, colonies are isolated for further study.

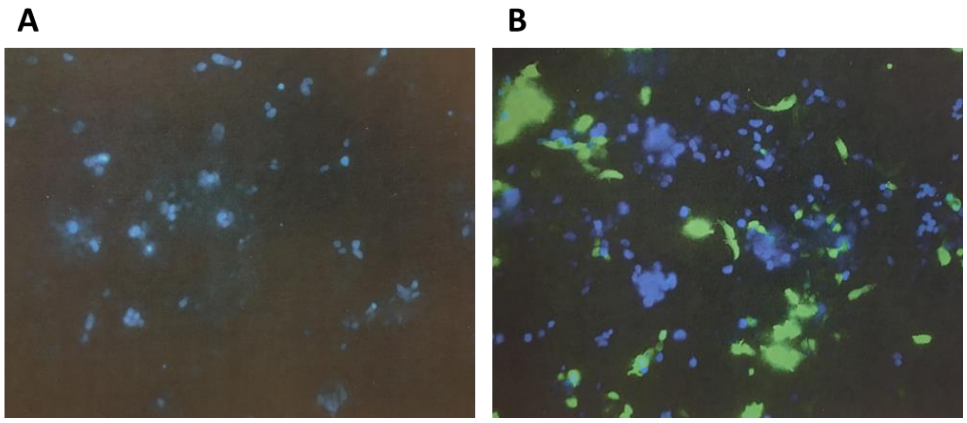


Figure 6.3 Transfection of HEK-RBP cells with GFP-expressing plasmid: Cells were transfected with TransIT-LT1 before visualization with a fluorescence microscope to observe DAPI (blue) and GFP (green). A) no transfection, B) HEK-RBP cells transfected with a GFP-containing plasmid provided by Dr S Jensen, University of Oxford.

In order to ensure that the intracellular apparatus to mediate the FRT integration of the Notch receptor were intact, a GFP-FRT plasmid, with an inducible GFP coded in the FRT site was tested for integration (provided by Dr A Gerondopoulos, Barr lab, University of Oxford). This FRT plasmid was transfected alongside the pOG44 plasmid coding for Flp recombinase for FRT integration at a ratio of 1:9 (pcDNA5/FRT:pOG44). The pcDNA5/FRT GFP plasmid allowed confirmation of FRT integration-ability which, unlike the mN1 plasmids would not have limited transfection efficiency due to Notch cDNA's large size (12 kbp). The transfection of multiple concentrations of this smaller GFP-containing plasmid showed that the ratio of 1:9 for pcDNA5 : pOG44 plasmids yielded successful clones with the transfection reagent Mirus TransIT-LT1 system (Figure 6.4). This suggests that the FRT integration site of the HEK-RBP cells was intact. Furthermore, these results suggest that as number of colonies did not greatly vary with 50-150 ng DNA, optimal transfection of this system should be carried out with 100 ng mass of pcDNA5/FRT GFP and a pcDNA5:pOG44 ratio of 1:9.

Transfected DNA (ng, pcDNA5/FRT GFP)	**pcDNA5/FRT GFP : pOG44**	**Number of colonies per 9 cm^2 confluent cells**
50	1:9	24
100	1:9	25
150	1:9	25

Figure 6.4 Trial transfection to demonstrate FRT-integration capacity of cells: HEK-RBP cells were transfected with pOG44 and pcDNA5/FRT GFP using Mirus TransIT-LT1 reagent. A range of DNA concentrations were tested to determine optimal concentration for this system.

The pcDNA5/FRT mN1 plasmid coding for the entire mN1 receptor is large (12.3 kbp). Therefore, the transfection efficiency of this large plasmid is likely to be significantly lower than the smaller GFP-containing plasmid used for test transfections, considering that the size of transfected plasmid is inversely proportional to the transfection efficiency (see www.thermofisher.com/uk/en/home/references/gibco-cell-culture-basics/transfection-basics/guidelines-for-plasmid-dna-transfection.html). Having demonstrated that the Flp-in system is intact in the host cells, and the optimum plasmid concentration determined, transfection with pcDNA5/FRT mN1, containing full length mN1 was undertaken.

6.3.1. Using pcDNA5/FRT mN1 wt and L468A to validate the activation assay

In order to validate that the signal observed in the activation assay is due to ligand-dependent activation, an mN1 variant defective in ligand binding (L468A) was created for use as a negative control. The importance of L468 in binding ligand has been previously observed (Whiteman et al, 2013, Luca et al 2017, see Chapter 5 - Interaction of the Abruptex region with the ligand-binding region). Site directed mutagenesis was carried out to engineer this L468A variant of mN1 in the pcDNA/FRT plasmid using overlap extension PCR and mN1 wt (provided by Dr B Korona, University of Oxford) as a template. This mutagenesis was carried out in the region of interest of mN1 in a smaller pUC19 plasmid (6.3 kbp), to increase ease of handling and processing. DNA sequencing was used to confirm that this substitution had taken place, and that the remainder of the sequence synthesized in the PCR was unchanged. This fragment was digested from the plasmid and then ligated back into pcDNA5/FRT mN1

treated with similar restriction enzymes, and the ligation sites sequenced to create the plasmid for transfection. wt and L468A plasmids were then transfected with pOG44 into HEK-RBP cells in the optimal conditions described in Section 6.2. in collaboration with Dr B Korona, University of Oxford. Subsequent selection with hygromycin enabled isolation of transfected colonies. Multiple checks were carried out at different stages over the course of generating these cell lines to ensure the wt and L468A plasmids had been transfected and to assess the amount of the variant receptors which were expressed on the surface of the cells (Figure 6.5). These checks include an RT-PCR to confirm the sequence of expressed mN1, and an EGTA-induced activation assay to investigate the amount of mutant receptor on the cell surface. The EGTA-induced ligand-independent activation assay induces activation by stripping the Ca^{2+} ions from the NRR of the receptor (See Chapter 1 – Introduction, Section 1.4.4). The loss of these ions leads to conformational changes that expose the S2 and S3 cleavage sites to proteolytic cleavage, which are subsequently cleaved and lead to NICD release. Therefore, EGTA treatment induces activation of all activation-capable receptor at the cell surface and provides a quantitative readout as to the total potential signal of receptors at the surfaces. This ensures that cell-surface expression of variants is similar between variants and so differences in ligand-induced activation can be compared.

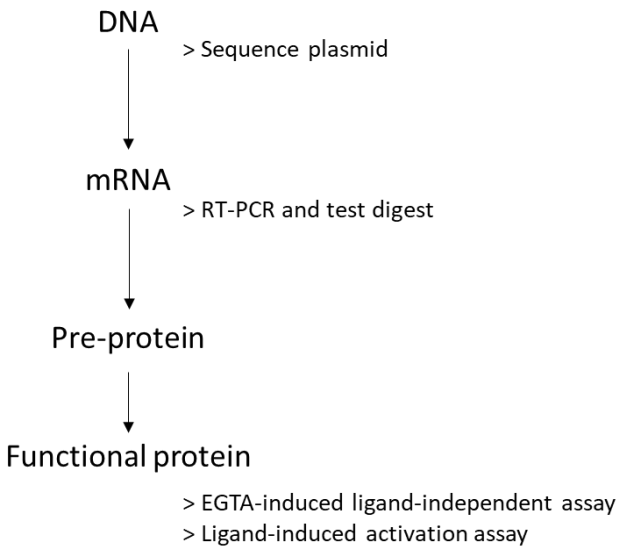


Figure 6.5 Assessment of accuracy of the activation assay: DNA sequencing was carried out before transfection to ensure correct sequence of the mN1 receptor to be transfected. Following transfection and selection, RNA was purified, and RT-PCR performed, and test digests carried out to confirm RNA sequence corresponds with expected variant. This also gives an insight into the relative level of expression of mN1 compared with endogenous hN1 expressed from the genome of the HEK-RBP cells. Finally, to ensure all mutants were expressed similarly, total mN1 at the cell surface was assessed by a ligand-independent activation method using EGTA. The EGTA-induced signal from different cell lines ensures expression of receptor at the cell surface and therefore valid comparison of signal between different variants.

6.3.1.1. Assessment of L468A

RT-PCR was carried out to confirm the variant protein was being expressed and provide an insight into the relative levels of mRNA expression between mN1 variants, and ensure the variant expressed contains expected substitution L468A in the LBR (Figure 6.6). The semi-quantitative nature of comparing RT-PCR product gel bands and the observation that mRNA amounts do not always correlate with protein amounts (Liu et al, 2016) suggest the EGTA-

induced ligand independent activation is the most accurate comparison of relative expression of mN1 variants.

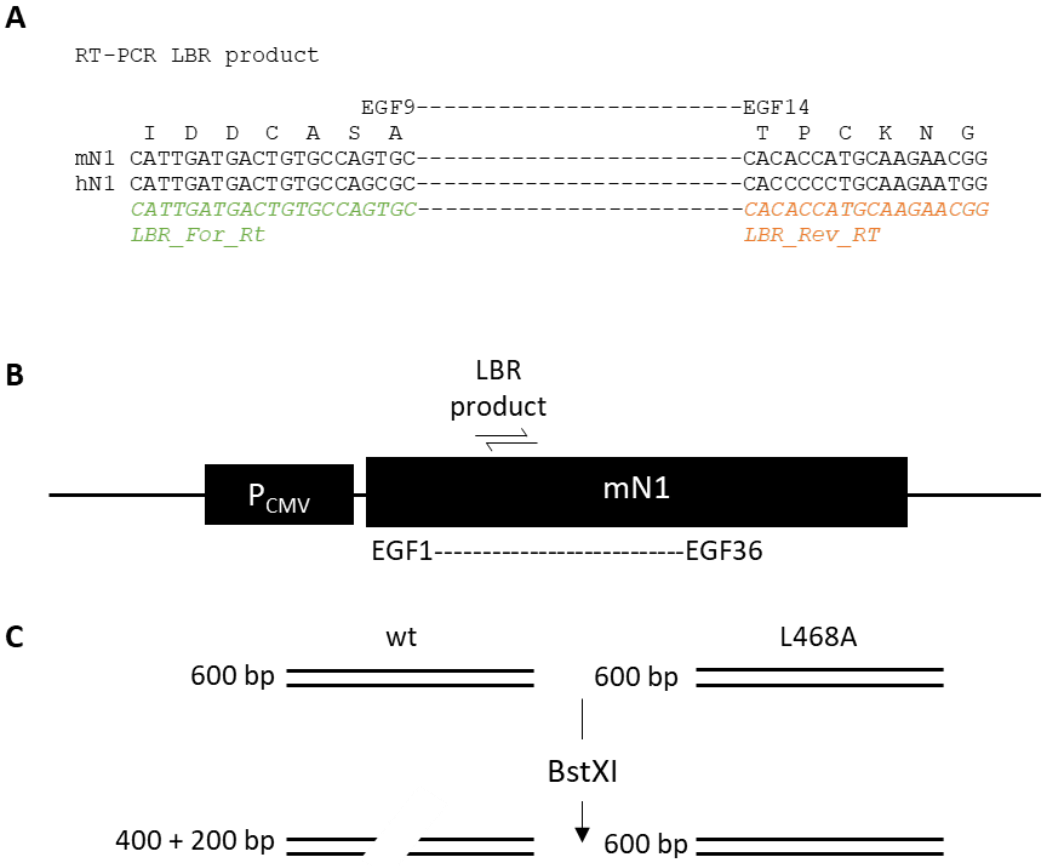


Figure 6.6 RT-PCR to investigate expression and LBR sequence of mN1 variant: A) Primers (green and orange) used for RT-PCR amplify the region EGF9-14 (LBR product) in mN1, downstream of the cytomegalovirus promoter (P_{CMV}) (B). C) Products of PCR following RT-PCR are ~600bp. The presence of the L468A substitution (CTG -> GCC) disrupts the consensus sequence of the restriction enzyme BstXI so cannot be cleaved.

Even though successful transfection and integration in selected colonies is implied by resistance to the selection agent hygromycin (see Chapter 2 -Materials and Methods), we further characterized the expression from the integrated mN1 construct to ensure that the correct variants were being expressed. Furthermore, although the levels of Notch expressed

on the surface should be similar considering the integrated constructs are all under the control of the same promoter, assays were carried out to check the relative levels of expression and these were then taken into account when comparing the activation levels. Expression of wt or L468A variants were confirmed using RT-PCR to investigate levels of mRNA expression, observing a band of interest at a length of ~600bp in the LBR, and then a test restriction digest with the restriction enzyme BstXI was used to ensure the presence of the variant L468A substitution in the LBR RT-PCR by observing lack of cleavage of the 600 bp product in the L468A mutant which disrupts the BstXI consensus sequence (Figure 6.7).

After confirming that the L468A variant was expressed in these cells, ligand-independent activation with EGTA was carried out to investigate the amount of activation-capable receptor present in each cell line. This involves treating the cells with 10 mM EGTA, which leads to the removal of the Ca^{2+} ions in the negative-regulatory region of the receptor, triggering the cleavage reactions and release of the NICD into the cell. The NICD then leads to activation following the method outlined in Figure 6.1. Quantifying the relative amounts of activation-capable receptor suggests comparable surface expression of variant constructs (Figure 6.8).

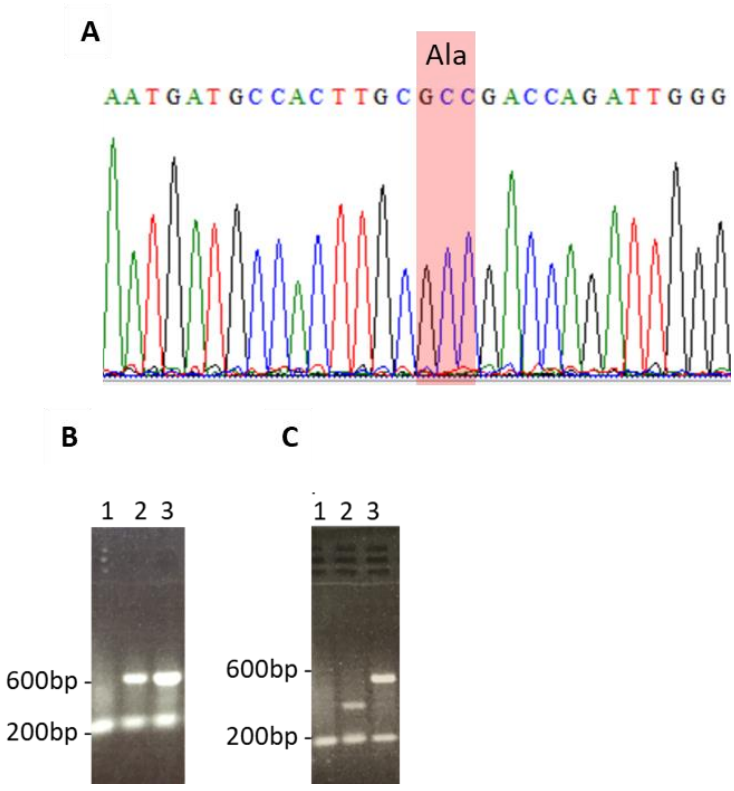


Figure 6.7 Sequence of mN1 wt L468A confirmed by RT-PCR and restriction digestion. A) DNA sequence of pcDNA5/FRT mN1 L468A confirms that mutations took place as expected, B) RT-PCR products were run on a 1.5 % agarose TBE gel stained with ethidium bromide. 1) HEK-RBP, 2) mN1 wt, 3) mN1 L468A. expected band of ~600bp is present in mN1 variants. Additional band at ~200 bp is a non-specific product which is amplified in HEK-RBP cells which have not been transfected with mN1. A PCR without reverse transcription was carried out concurrently to show the sample did not contain contaminating DNA. C) Restriction digest of RT-PCR with BstXI run on a 1.5 % agarose TBE gel. Band at 600 bp is not cleaved in mN1 L468A, but is cleaved in mN1 wt.

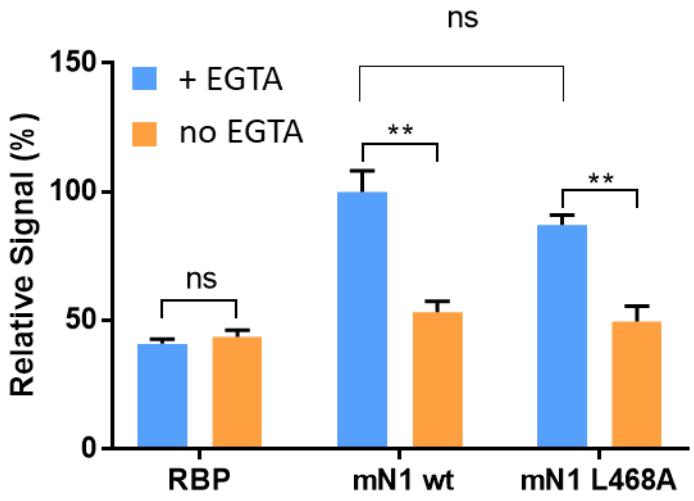


Figure 6.8: Ligand independent activation assay with EGTA suggests equivalent expression of mN1 wt and L468A: Ligand independent activation assays were carried out with +/- 10 mM EGTA with the graph showing the combined results of two independent plates, each of three repeat wells. Signal is shown relative to mN1 wt + EGTA. Statistical comparisons were performed with a two-way ANOVA test within groups and t-test between groups, Values are presented together with the mean ± SD. ns = not significant; **P < 0.0001; Data were analyzed with Prism 7 (GraphPad, San Diego, CA, USA).

6.3.2. Optimising conditions for the hJag1 activation assay

After showing comparable surface expression of mN1 variants using the EGTA assay, ligand-dependent activation using the ligand human Jagged1 (hJag1) was undertaken. These assays involved the binding of purified ligand to a plate, followed by washing and then seeding of cells to establish a steady state of activation. Following this, the luciferase substrate luciferin was added, and then luminescent signal recorded. Preliminary assays were carried out to determine the optimum concentration of ligand to coat the plate (Figure 6.9). In addition, these assays investigated the optimal incubation time for luminescent recording following

addition of substrate, corresponding to time taken for the substrate to permeate the plasma membrane and reach the active luciferase molecules in the nucleus (Figure 6.9). The results of these preliminary experiments showed that that the plate should be read within five minutes following addition of luciferin to the cells to achieve maximum signal. In addition, a similar signal was seen for the concentrations 5 μg/mL and 10 μg/mL of ligand when coating the plate. 10 μg/mL was selected for use in future assays considering 10 μg/mL was the optimal concentration for activation observed by Ilagan et al, 2011.

Following the confirmation that the mN1 L468A was expressed in these cells, and that the amount of receptor on the cell surface of the mN1 wt and L468A were equivalent in the two cell lines, and could be fairly compared, the ligand dependent activation assay was carried out (Figure 6.10). Considering the result of the experiments to optimize the activation reaction, the activation assay was conducted at 10 μg/ml ligand, reading the plate immediately following addition of luciferin substrate.

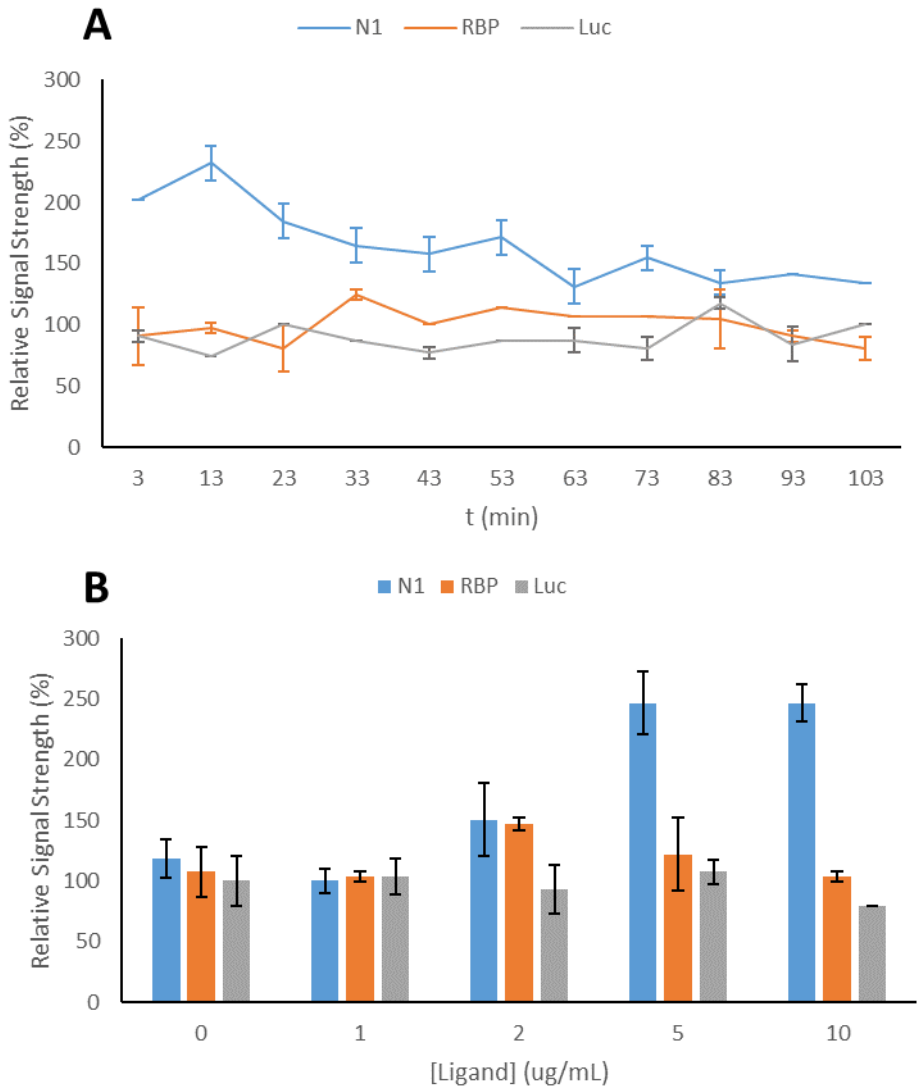


Figure 6.9 Investigating optimal reading time and ligand concentration for activation assay: hJag1 ligand was immobilized in the well and then cells added before addition of luciferin substrate. Luminescent signal was then recorded every 10 mins for ~100 mins, Signal is shown relative to HEK-RBP without ligand. A) Time course result for mN1 wt (N1), background signal from HEK-RBP (RBP), and empty wells (Luc), all at 10 µg/mL hJag1, B) Signal from a range of ligand concentrations at 13 mins into time course. Results shown are from a single plate, containing three repeat wells. Error bars show SD.

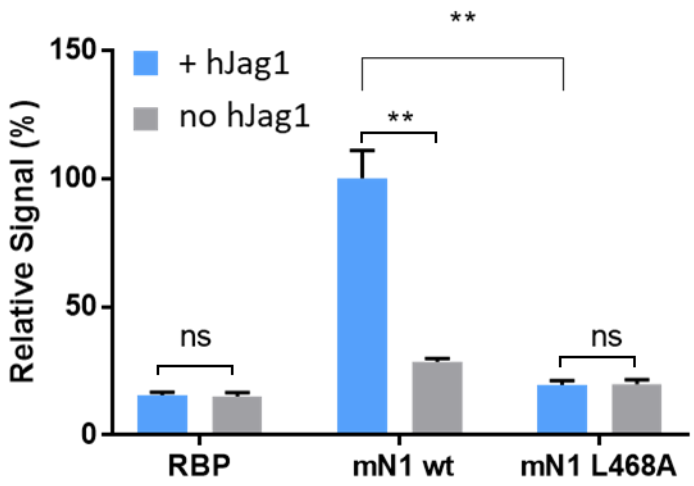


Figure 6.10 Low signal for L468A suggests activation observed is ligand dependent: Ligand dependent activation assays were carried out with +/- hJag1 with the graph showing the combined results of two independent plates, each of three repeat wells. Signal is shown relative to mN1 wt + hJag1. Statistical comparisons were performed with a two-way ANOVA test within groups and t-test between groups, Values are presented together with the mean ± SD. ns = not significant; **P < 0.0001; Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

These results suggest that the L468A substitution, which has been previously shown to abrogate ligand binding (Whiteman et al, 2013, Suckling et al, 2017), also abrogates the activation of the receptor. This finding confirms that the signal observed in these assays is ligand-dependent Notch activation.

6.4. Designing Ax variant receptors to be investigated in the activation assay

Having validated the activation assay using mN1 and mN1 L468A, several Abruptex variants were engineered in the mN1 system to investigate the impact of these mutations on Notch1 activation.

6.4.1. Abruptex EGF24 D909V variant

Considering that all of the Abruptex mutants observed in *D. melanogaster* are substitutions (see Chapter 1 – Introduction, Section 1.4.3.1), these substitutions served as a basis for the initial variants tested in this assay to investigate the impact of Abruptex mutations on activation.

The observations in Chapter 5 - Interaction of the Abruptex region with the ligand-binding region suggest that EGF24 is the site of an interaction between the Ax region and the LBR. Initial experiments were therefore carried out to investigate the impact of known Abruptex variants on activation. Adding to this is the observation that D909V (analogous to the D948V Abruptex mutation in *D. melanogaster*) led to a disruption of binding in the flow cytometry assay (see Chapter 5 - Interaction of the Abruptex region with the ligand-binding region, Section 5.2.3), this substitution was the first target for investigating the impact of Abruptex mutations on activation.

A mutant D909V construct was produced and transfected following the protocol described in the Chapter 2 – Materials and Methods, Section 2.3.2.2. In order to confirm this substitution had taken place, RT-PCR was carried out to amplify the Ax region of the mN1 RNA (Figure 6.11)

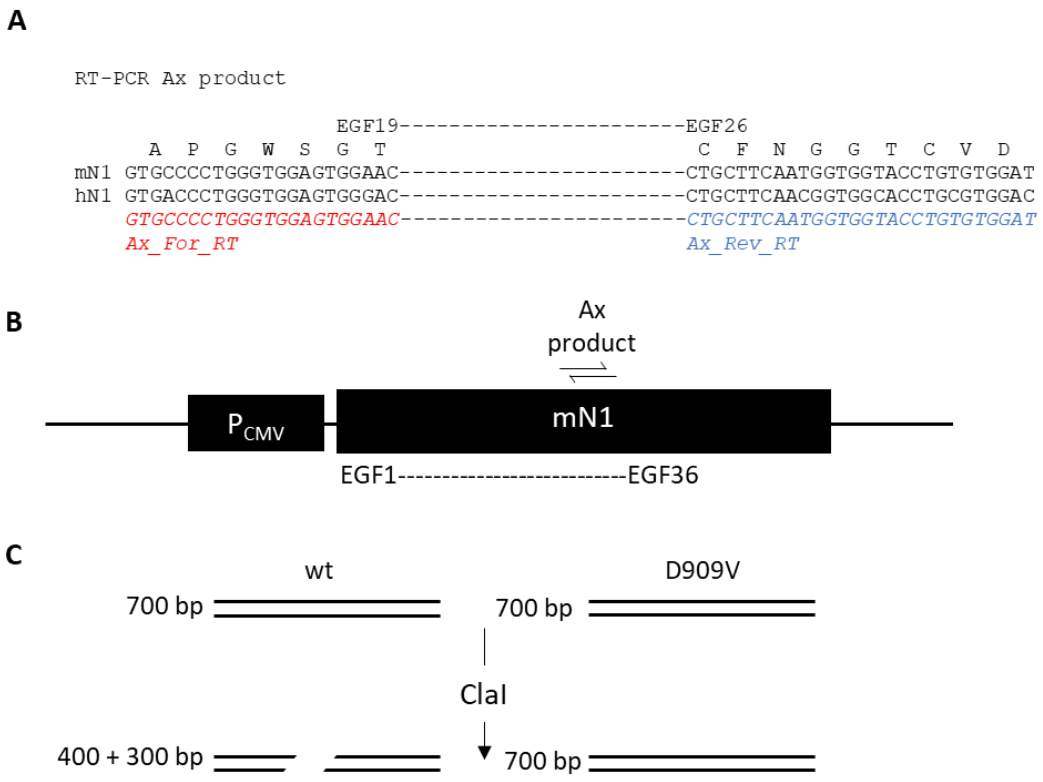


Figure 6.11 RT-PCR to investigate expression and Ax sequence of mN1 variant: A) Primers (green and orange) used for RT-PCR to amplify the region EGF19-26 (Ax product) of mN1, downstream from the cytomegalovirus promoter (P_{CMV}) (B). C) Products of PCR following RT-PCR are ~700bp. The presence of the D909V substitution (GAT -> GTG) disrupts the consensus sequence of the restriction enzyme ClaI so cannot be cleaved.

As described in Figure 6.11, RT-PCR and restriction digestion were used to demonstrate the presence of the mutation in mRNA (Figure 6.12). The lack of cleavage of the ~700bp product from the Ax RT-PCR product indicates the D909V mutation is present in the RNA expressed from the mN1 D909V cell line.

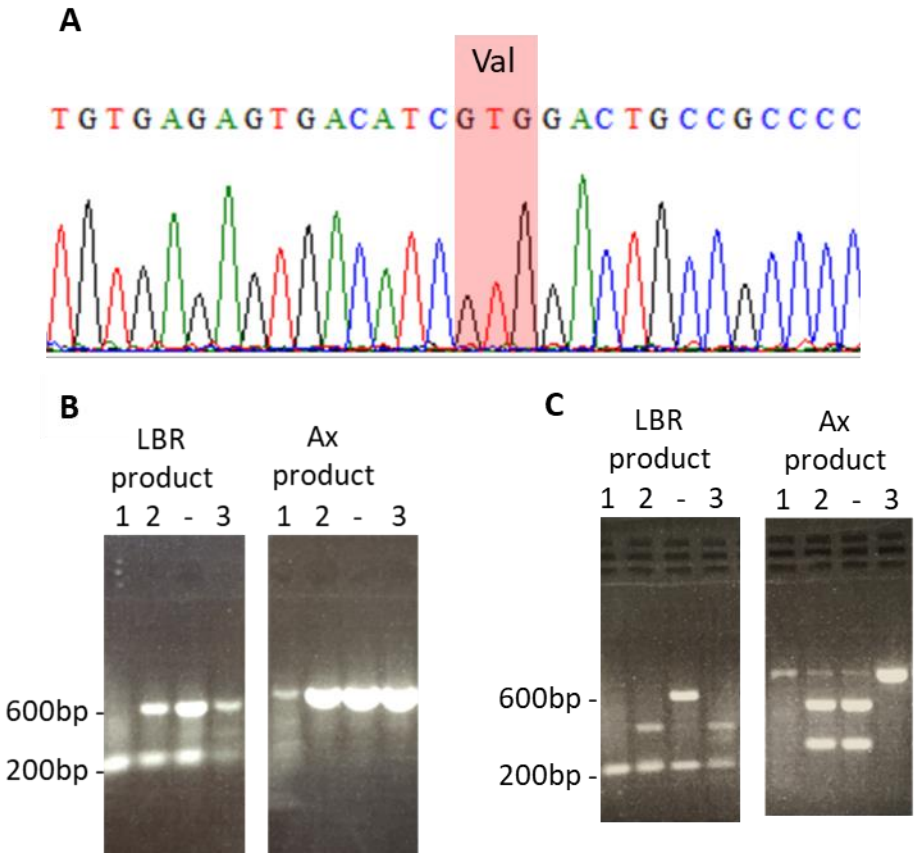


Figure 6.12 Sequence of mN1 D909V was confirmed by RT-PCR and restriction digestion: A) DNA sequence of pcDNA5/FRT mN1 D909V confirms that substitution took place as expected, B) RT-PCR products were run on a 1.5 % agarose TBE gel stained with ethidium bromide. 1) HEK-RBP, 2) mN1 wt, 3) mN1 D909V. expected band corresponding to Ax product (~700bp) is present in mN1 D909V. Furthermore, expected LBR product (~600bp) is also present with additional band at ~200 bp is a non-specific product which is amplified in HEK-RBP cells which have not been transfected with mN1. PCR reactions without reverse transcription was carried out concurrently to show the samples did not contain contaminating DNA. C) Restriction digestions of RT-PCR products with ClaI (Ax product), and BstXI (LBR product) run on a 1.5 % agarose TBE gel. mN1 D909V Ax product at 700 bp is not cleaved by ClaI, suggesting D909V substitution is present. Furthermore, cleavage of 600 bp LBR product confirms LBR region is wt.

Following confirmation that the mN1 D909V variant was expressed, ligand–independent activation with EGTA was carried out to investigate the relative amount of mN1 D909V receptor expressed on the cell surface (Figure 6.13A). The results of this ligand-independent

activation assay suggested that mN1 D909V was expressed on the cell surface in similar amounts to mN1 wt. Investigation into ligand-induced activation was subsequently carried out to investigate the impact of the D909V substitution on activation by ligand (Figure 6.13B).

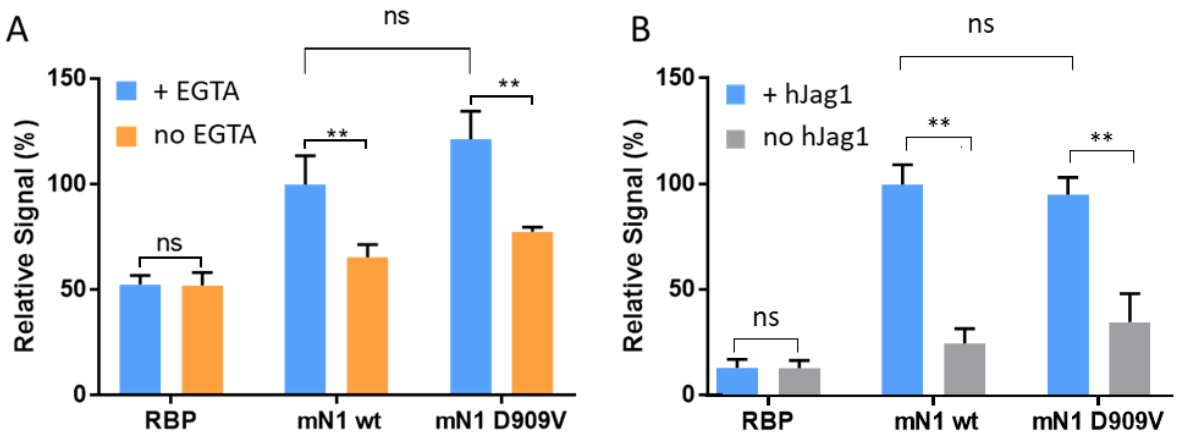


Figure 6.13 Ligand dependent activation assay suggests mN1 D909V signals similarly mN1 wt: A) Ligand independent activation assays were carried out +/- 10 mM EGTA with the graph showing the combined results of at least two independent repeats. Signal is shown relative to mN1 wt + EGTA. B) Ligand dependent activation assays were carried out +/- hJag1 with the graph showing the combined results of three independent plates, each of at least one repeat well. Signal is shown relative to mN1 wt + hJag1. Statistical comparisons were performed with a two-way ANOVA test within groups and t-test between groups, values are presented together with the mean ± SD. ns = not significant; **P < 0.0001; Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

Even though the L468A variant shows a significant decrease in activation compared to wt, the D909V variant shows no significance difference in signal compared with wt. This could be because there is no impact of this substitution on Notch activation in this context of overexpressed receptor on the HEK-RBP cell surface. Another possible explanation may be because the changes compared to wt are too subtle to be observed in this assay, however would impact development when this substitution occurs in an organism.

6.4.2. Double Abruptex mutant

In order to observe the impact of the Abruptex mutations on Notch activation, a double mutant receptor containing D909V (EGF24) and N947I was designed. This N947I substitution corresponds to the Ax mutant observed in EGF25 of the *D. melanogaster* receptor N986I (see Chapter 1 – Introduction, Section 1.4.3.1). These mutants were selected because their phenotypes are very similar, they are both *N*-suppressor mutations (see Chapter 1 – Introduction, Figure 1.13) and so the impact of these two substitutions on activation is likely to be similar and have an additive effect. Therefore, both were engineered into the receptor to produce a double Abruptex (DVNI) mutant (Figure 6.14).

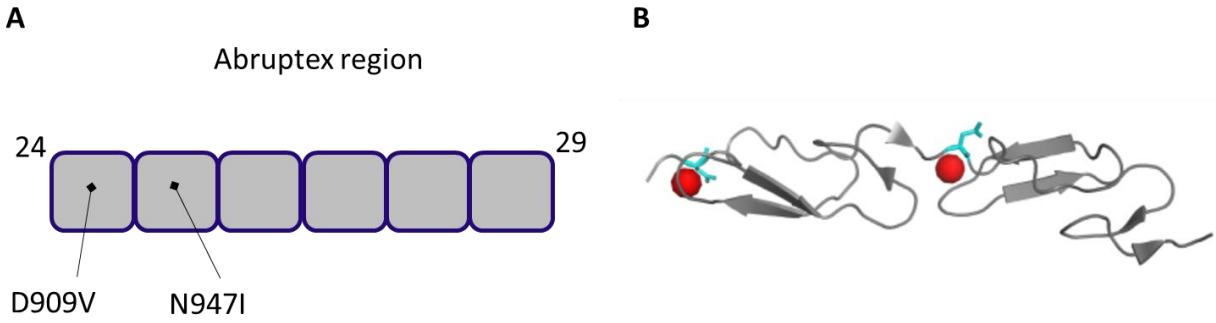


Figure 6.14 Designing a double Ax mutant: A) The location of the positions in mammalian Notch corresponding to dN suppressor mutations in the Ax region. B) Location of the mutations (cyan) shown in a cbEGF domain pair model (PDB ID = 2VJ3, grey) with Ca^{2+} shown (red).

The DVNI mutant was characterized using RT-PCR and test restriction digest following the protocol described in the Chapter 2 - Materials and Methods (Figure 6.15) to ensure the variant was expressed as expected.

The DVNI double Ax mutant was investigated using ligand-independent EGTA-induced activation to ensure expression of the mN1 receptor at the cell surface (Figure 6.16A). This EGTA assay confirmed that the surface expression of wt and DVNI were comparable. Activation with hJag1 was then assayed in the ligand-dependent activation assay (Figure 6.16B). The results of this assay suggest that activation by DVNI was significantly higher than background activation, and interestingly was also lower than wt ($P < 0.0001$). This is in contrast to the D909V variant, which appears to signal in a similar way to wt (Figure 6.13). These results suggest the addition of N947I to the D909V substitution leads to a decrease activation ability in the context of this assay. Although it is possible that the N947I in isolation would lead to a decrease in activation ability, it is likely that the D909V and N947I substitutions have the same impact as they are both Ax *N*-suppressor mutants (See Chapter 1 – Introduction, Figure 1.13).

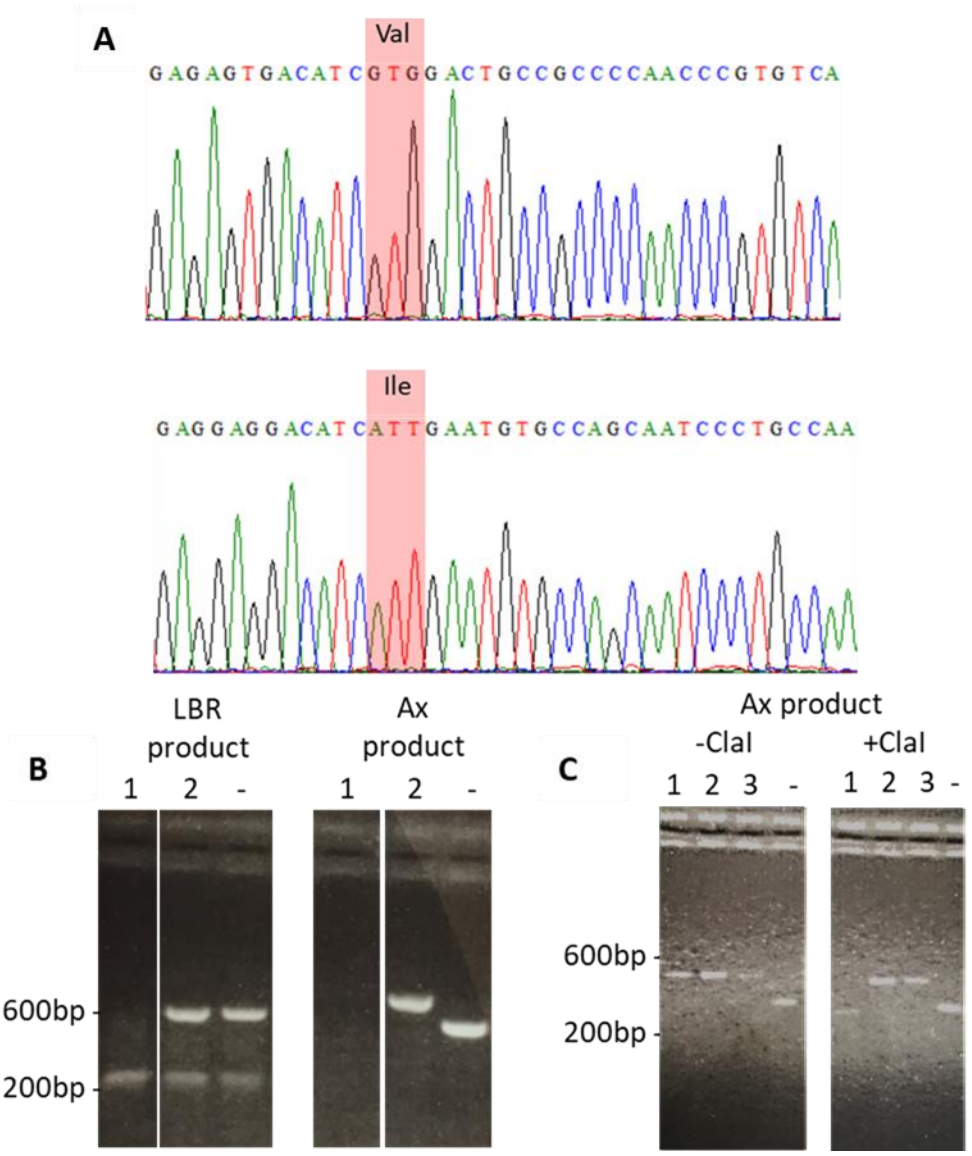


Figure 6.15 Expression of mN1 DVNI was confirmed by RT-PCR: A) DNA sequence of pcDNA5/FRT mN1 DVNI confirms that substitutions took place as expected, B) RT-PCR products were run on a 1.5 % agarose TBE gel stained with ethidium bromide. 1) HEK-RBP, 2) mN1 DVNI. Expected band corresponding to Ax product (~700bp) is present in mN1 DVNI. Furthermore, expected LBR product (~600bp) is also present with additional band at ~200 bp is a non-specific product which is amplified in HEK-RBP cells which have not been transfected with mN1. C) RT-PCR products underwent test digests with ClaI enzyme and run on a 1.5 % agarose TBE gel. 1) mN1 wt, 2) mN1 D909V, 3) mN1 DVNI. DVNI product at 700 bp is not cleaved confirming D909V substitution is present. PCR reactions without reverse transcription was carried out concurrently to show the samples did not contain contaminating DNA.

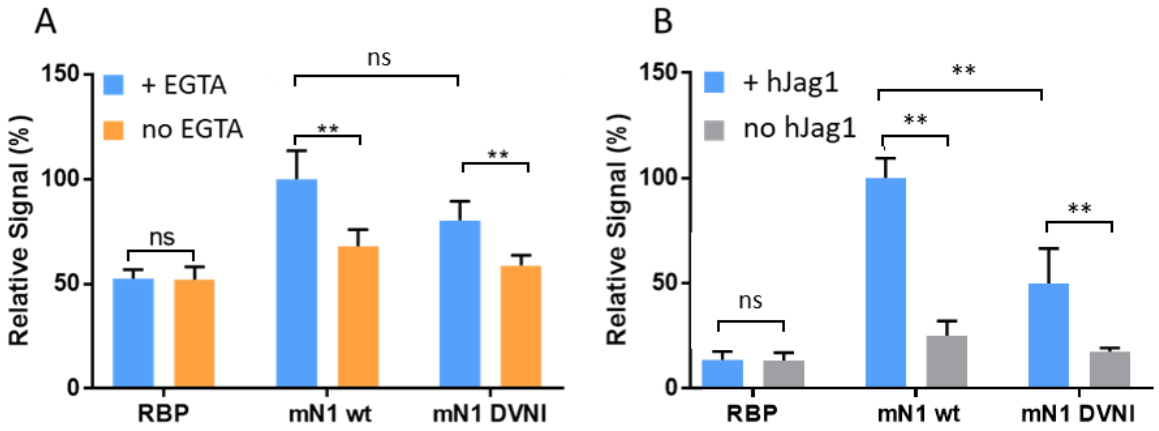


Figure 6.16 Ligand dependent activation assay suggests mN1 DVNI signals less than mN1 wt: Ligand independent activation assays were carried out +/- 10 mM EGTA with the graph showing the combined results of two independent plates, each of three repeat wells. Signal is shown relative to mN1 wt + EGTA. B) Ligand dependent activation assays were carried out +/- hJag1 with the graph showing the combined results of three independent plates, each of at least two repeat wells. Signal is shown relative to mN1 wt + hJag1. Statistical comparisons were performed with a two-way ANOVA test within groups and t-test between groups, values are presented together with the mean ± SD. ns = not significant; **P < 0.0001; Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

It might have been expected that D909V and DVNI would enhance activation as they act as *N*-suppressor substitutions, however, in this system in which mN1 is overexpressed, this double mutation leads to a decrease in activation. A further possibility is that the role of the Ax region is different in mammalian Notch receptors compared with dN, and it is possible that another mammalian Notch orthologue (N2, N3, N4) has adopted the Ax role of dN that is responsible for the Ax phenotypes. Furthermore, this assay is in trans, the impact of Ax mutations on regulation of the receptor in cis remains to be determined.

6.4.3. Abruptex deletion mutant

To role of the Abruptex region in activation was further explored by investigating activation in receptor variants with a section of the Ax region deleted. One deletion mutant was designed, corresponding to a deletion of EGF24-25 (Δ24-25) (Figure 6.17).

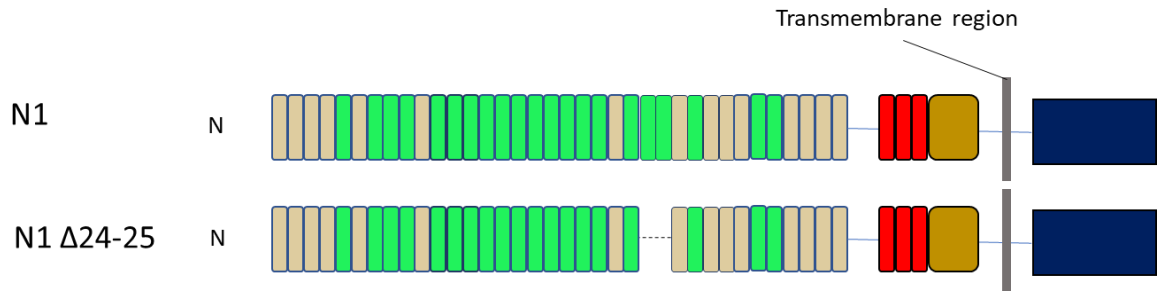


Figure 6.17 Ax deletion mutant engineered for investigation in the activation assay: The Notch receptor is shown as a schematic with EGF domains (beige) and cbEGF domains (green) making up the N-terminal of the receptor including the Ax region in which domains are deleted.

Deletion of EGF24-25 was designed as this deletes the domains containing *N*-suppressor Ax mutants (See Chapter 1 – Introduction, Figure 1.13). mN1 Δ24-25 was transfected following the protocol described in Chapter 2 – Materials and Methods and underwent RT-PCR to confirm a transcript of expected length was expressed (Figure 6.18).

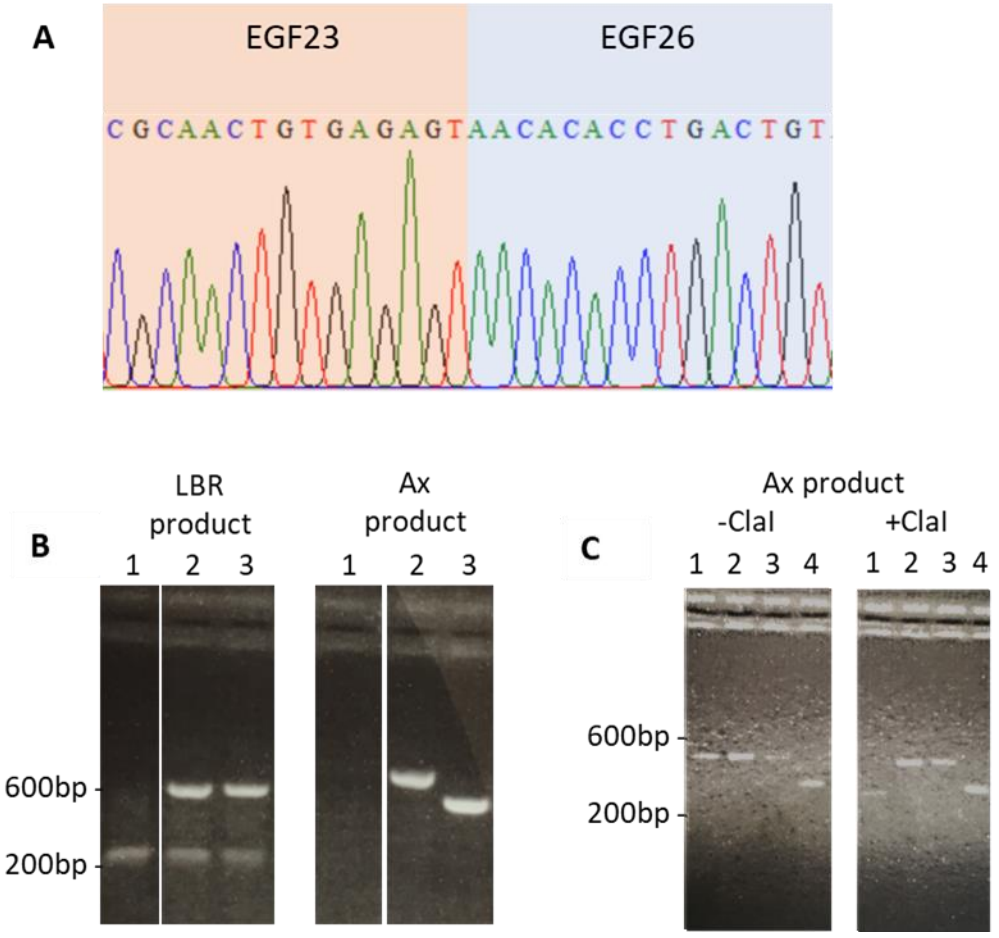


Figure 6.18 Expression of mN1 Δ24-25 was confirmed by RT-PCR: A) DNA sequence of pcDNA5/FRT mN1 Δ24-25 confirms that deletion took place as expected, B) RT-PCR products were run on a 1.5 % agarose TBE gel stained with ethidium bromide. 1) HEK-RBP, 2) mN1 DVNI 3) mN1 Δ24-25. Expected band corresponding to Ax product (~480bp) is present in mN1 Δ24-25. The band shift seen in the Ax product compared to mN1 DVNI confirms that EGF24-25 have been deleted in mN1 Δ24-25 (~220 bp deletion). Furthermore, expected LBR product (~600bp) is also present with additional band at ~200 bp is a non-specific product which is amplified in HEK-RBP cells which have not been transfected with mN1. C) RT-PCR products underwent test digests with ClaI enzyme and run on a 1.5 % agarose TBE gel. 1) mN1 wt, 2) mN1 D909V, 3) mN1 DVNI, 4) mN1 Δ24-25. Δ24-25 does not undergo cleavage on treatment with ClaI supporting that EGF24-25 is deleted in this construct. PCR reactions without reverse transcription were carried out concurrently to show the samples did not contain contaminating DNA.

In order to investigate the impact of deleting EGF24-25 on the ability of the receptor to signal, mN1 Δ24-25 was investigated in the activation assay, initially in the ligand independent EGTA

assay to confirm that this deletion construct was expressed in similar amounts as wt (Figure 6.19A). This suggests that the deletion of these domains does not greatly impact folding and secretion. Despite this similar expression, the activation assay with hJag1 suggests that the Δ24-25 construct is unable to signal on exposure to hJag1 (Figure 6.19B).

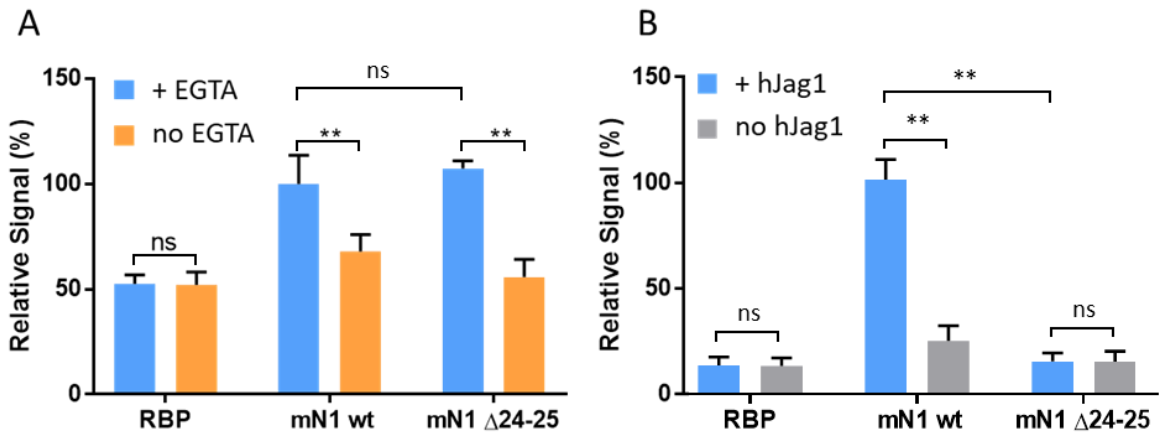


Figure 6.19 Ligand dependent activation assay suggests mN1 Δ24-25 does not undergo activation on exposure to hJag1: Ligand independent activation assays were carried out +/- 10 mM EGTA with the graph showing the combined results of two independent plates, each of three repeat wells. Signal is shown relative to mN1 wt + EGTA. B) Ligand dependent activation assays were carried out +/- hJag1 with the graph showing the combined results of three independent plates, each of at least two repeat wells. Signal is shown relative to mN1 wt + hJag1. Statistical comparisons were performed with a two-way ANOVA test within groups and t-test between groups, values are presented together with the mean ± SD. ns = not significant; **P < 0.0001; Data were analysed with Prism 7 (GraphPad, San Diego, CA, USA).

These results suggest that Δ24-25 is essential for activation in this context. The lack of activation of this variant could be due to differences in the length of this receptor compared with wt changing accessibility of the LBR to ligand. This deletion could also act to disrupt any conformational transmission through the Ax region from the LBR to the S2 and S3 cryptic cleavage sites near the transmembrane region.

6.5. Discussion of Notch activation assays

The Notch activation assays suggest that the L468A substitution in EGF12 in the LBR leads to a decrease in activation by ligand hJag1. Considering the previous observations that the L468A and substitutions of L468 leads to a decrease in binding to ligand (Whiteman et al, 2013, Chapter 5 – Interaction of the Abruptex region with the ligand binding region, Section 5.1.1), this confirms the activity observed in this assay corresponds to ligand dependent signalling.

These assays provide an insight into the impact of Ax substitutions on activation in a cellular assay. Furthermore, these substitutions have been observed to impact activation in mammalian Notch receptors. Investigating the impact of dN Ax mutations in mammalian N1 shows that the D909V (EGF24 Ax mutation) shows no decrease in activation compared to wt in this assay. However, combining this EGF24 Ax mutation with the N947I (EGF25 Ax mutation) in a DVNI double Ax variant, shows a decreased activation compared to mN1 wt. This suggests that the N947I is responsible for this decrease in activation, or that the impact on activation only manifests in the presence of both Ax substitutions. The mechanism for this decrease in activation is unclear and could be due to changes in structure or flexibility in the region EGF23-25 considering that the DV and NI substitutions are in the Ca^{2+} consensus region in these cbEGFs. Any structural changes relative to wt do not seem to impact secretion and surface expression as indicated by the ligand-independent EGTA assay.

The deletion of EGF24-25 leads to a decrease in activation to background levels suggesting that this deletion prevents signaling by hJag1 ligand. Considering that this deletion variant does not impact on secretion or surface expression, this suggests the variant receptors are not grossly misfolded. The LBR remains present in this deletion variant, therefore there are a number of possible mechanisms by which this deletion could prevent activation by ligand. These mechanisms include altering binding of ligand to the LBR so that ligand binding cannot occur, or that ligand binding occurs but this binding does not lead to cleavage at the S2 cleavage site.

7.

General discussion and future work

7.1. Aims of this thesis

The aim of this thesis was to investigate the shape and role of the previously uncharacterised Ax region within the Notch receptor. A range of Ax and LBR constructs were expressed using eukaryotic and prokaryotic systems and subsequently used for low resolution structural studies and binding assays. Structural studies involved the creation of a pipeline for processing data through SAXS and yielded a structural model of EGF23-27. Binding assays utilized site-directed mutagenesis of surface residues to produce a range of variants enabling identification of the location of the interaction between EGF24 in the Ax and EGF12 in the LBR. In addition, an activation assay was established to study the role of Ax substitutions on activation by Jag1. The findings support an interaction between Ax and the LBR, and a role for the Ax in facilitating successful activation following interaction with ligand Jag1.

7.2. Understanding the shape of the Notch receptor

The shape of the receptor at the cell surface is yet to be determined, and the mechanism by which ligand binding leads to successful signalling remains elusive. Structural analysis was conducted to investigate the shape of the Ax region and revealed an extended structure for EGF23-27. In light of bioinformatics analysis, these results suggest a kink at the EGF25-26 interface. Other than this kink, the extended structure of the region suggests that any binding interfaces are likely to extend across multiple EGF domains in a similar manner to the interaction between ligand and LBR (Luca et al, 2017). SAXS analysis suggests that EGF23-27 from hN1 and hN2 have a similar structure suggesting that the non-canonical cysteines in EGF25 and EGF27 do not form an interdomain disulphide bond, appear to have no major role in the structure of this region and do not cause a sharp bend in the receptor. The structure of EGF23-27 can be added to the emerging organisation of hN1 to further piece together the structure of the receptor at the cell surface (Figure 7.1).

7.3. Understanding the function of the Abruptex region

The interaction between Ax and the LBR was observed in plate binding and flow cytometry assays. The results reveal an interaction centered around EGF24 and EGF12, which is consistent with the observations of Pei and Baker, 2008, and Sharma et al, 2013, and is likely to be a contributor to the interactions observed in these studies. This is consistent with the extended structure of EGF23-27 which suggests that EGF25-29 may contribute to binding *in*

vivo. Furthermore, these investigations suggest that the location of the interaction face within EGF12 is dependent on both O-Fuc and O-Gluc modifications on two opposite sides of the domain. The O-Fuc/L468 side is the same side as ligand binding, however the observation that disruption of the opposite O-Gluc/Q462 side does not impact ligand binding suggests ligand binding and Ax binding to the LBR are fundamentally different.

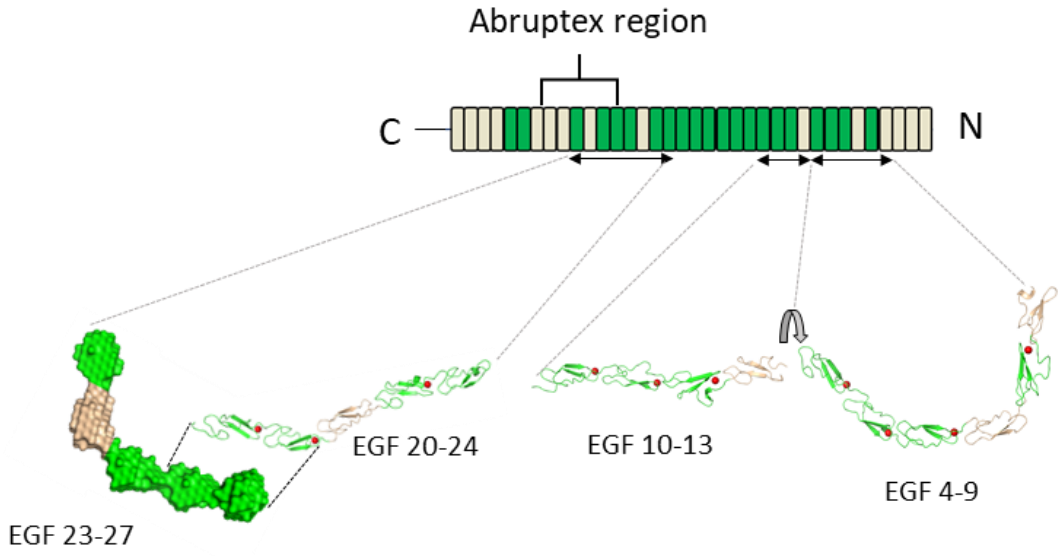


Figure 7.1 The proposed structural organisation of hN1 23-27 in hN1: EGF (wheat) and cbEGF (green) domains are shown schematically and in three-dimensions for the structurally characterised sections EGF 23-27 (See Chapter 4 - Structural Studies of the Abruptex region), EGF4-9 (PDB ID EGF4-7: 5FMA) and EGF10-13 (PDB ID EGF11-13: 2VJ3) (Weisshuhn et al, 2016), EGF20-24 (Suckling et al, unpublished data) with Ca^{2+} ions shown (red spheres). The transmembrane region is located at the C-terminal of the receptor.

This interaction between Ax and LBR could occur intra- or inter-molecularly, and future studies could be undertaken to further investigate the structure of the region between LBR and Ax (EGF 13-23), in particular EGF16, which lacks a conserved packing aromatic suggesting it may be the site of a conserved bend in the receptor (See Chapter 3). The presence of this

interaction raises the possibility of Notch dimerization at the cell surface (Figure 7.2). This dimerization is supported by the observation that *N*-enhancer and *N*-suppressor phenotypes only manifest when Ax variants are expressed alongside *N* mutant receptors, indicating interactions between receptors may occur (See Chapter 1).

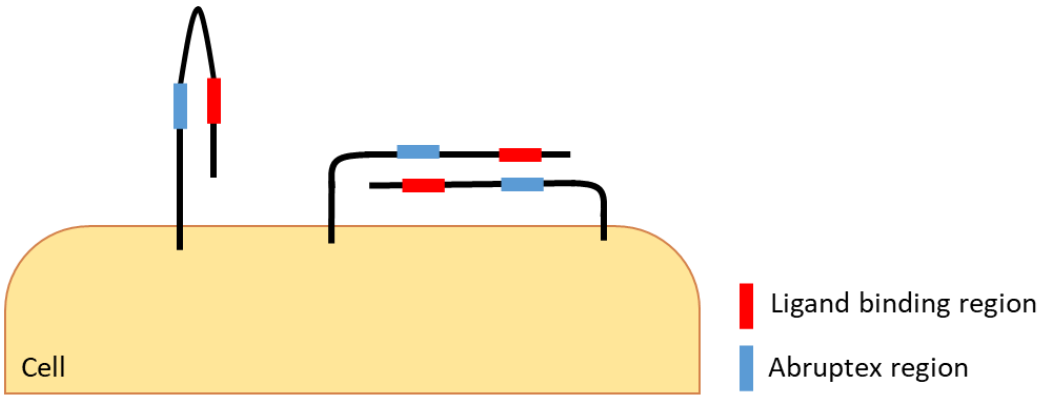


Figure 7.2 Proposed organisation of the Notch receptor at the cell surface: Interactions observed between the Ax region (blue) and LBR (red) suggest that the receptor may be present at the cell surface as a monomer as a result of intra-molecular interactions, or a dimer bought about by inter-molecular interactions.

Investigations into the function of the Ax region were carried out in a cell activation assay. The use of mN1 in this assay to infer function of hN1 is valid as the impact of the additional non-canonical cysteine residue in EGF27 in hN1 compared with mN1 does not appear to have a role in structuring the region considering the similarity of hN1 and hN2 23-27 structures as determined by SAXS (Chapter 4). It is an important observation that all variants tested in the activation assay were expressed on the surface of the cell in similar amounts to wildtype N1 (Chapter 6). This supports that the impacts of the DV and NI Abruptex mutations observed *in D. melanogaster* are due to changes in signalling ability rather than quantitative changes at the cell surface.

The observation that D909V does not alter activation relative to wt could be because this substitution does not lead to a change in activation relative to wt with hJag1, or that the change is so small that it cannot be observed in this assay system in which the receptor is overexpressed at the cell surface. The fact that the DVNI double mutant shows a decrease in activation relative to wt supports the latter, considering that the DV and NI substitutions are both *N*-suppressor mutants in *D. melanogaster* and are therefore likely to have a similar impact on signalling and have an additive effect in the DVNI variant. It is therefore likely that the inclusion of the NI substitution enables the impact of these Ax *N*-suppressor substitutions to be visible in the activation assay. An increase in activation might have been expected for Ax substitutions considering the Ax phenotype is indicative of too much signalling (Chapter 1), but a decrease in activation was observed. This may be because the ligand-dependent activation in this assay has some distinct differences compared to signalling in an organism. These differences include overexpression of receptor at the cell surface, and therefore any cis-mediated modulation would vary from an *in vivo* system. Although the potential impacts of these differences cannot be ruled out, the decrease in binding observed for the DVNI variant suggests that the role of Ax region is to enhance signalling during trans-activation by ligand. These results suggest the Ax phenotype, which is indicative of increased signalling, is caused by an alternative mechanism, possibly by changing cis-mediated modulation as suggested by de Celis, and Bray, 2000. A global decrease in signalling has also been proposed to lead to the Ax phenotype (Brennan et al, 1999) which would also be supported by these data.

The suggestion that Ax has a role in facilitating transactivation is consistent with the observation that the Δ24-25 deletion mutant does not activate on exposure to ligand. EGF24-25 is suggested to be rod-shaped from SAXS analysis (Chapter 4) supported by observations that EGF24 binds Ca^{2+} with moderately high affinity, and bioinformatics analysis (Chapter 3). The shortening of the receptor by ~60Å resulting from the deletion of EGF24-25 is in itself unlikely to alter ability to bind ligand. This is inferred from the observation that other active Notch receptors which successfully signal are of various lengths, with hN3 consisting of 34 cbEGF/EGF domains and hN4 consisting of 29 cbEGF/EGFs compared with the 36 cbEGF/EGFs in hN1 (See Figure 1.1). The DVNI substitutions which showed decreased activation in Chapter 6 are in Ca^{2+}-binding consensus sequences and therefore likely to change the rod-like nature of the domain interfaces in EGF24 (D909V) and EGF25 (N947I). Furthermore, deletion of EGF24-25, which has a rod-like structure (Chapter 4), also changes the N-terminal end of the Ax region and leads to an inability of the receptor to activate on exposure to ligand hJag1 (Chapter 6). This suggests that the structure of Ax, in particular the rod-like structure of N-terminus of Ax (EGF24-25) is important in facilitating activation. This function of the Ax could involve directly facilitating ligand binding, for example by orienting the LBR through the interaction with EGF24 (Figure 7.3A), or may enable successful receptor cleavage on interaction with ligand (Figure 7.3B).

Although it has been observed that Delta family ligands show similar preferences to the Jagged family in binding Notch (Luca et al, 2015, Luca et al, 2017) it remains to be determined if these similarities extend to the role of the Abruptex region in mediating activation.

Considering the primary ligand binding site is similar for Jagged and Delta family ligands (EGF11-12), and the structure of these ligands are similar (Figure 1.15), the role for Ax facilitating Delta-family transactivation is likely. This is supported by the observation that cells expressing Ax variant receptors bind less well than wt to cells expressing the ligand Delta (Lieber et al, 1992).

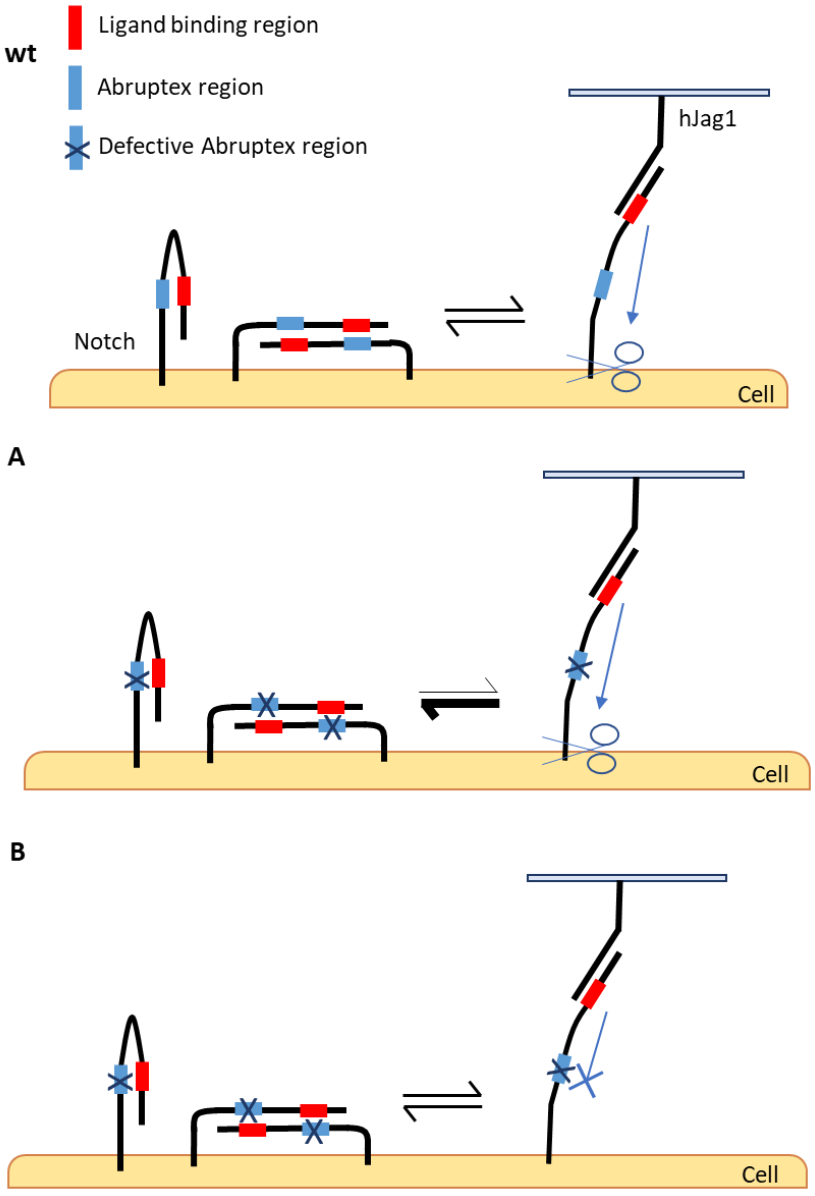


Figure 7.3: Proposed models for Ax function in Jagged1-dependent ligand transactivation: Notch receptor (Notch) is shown as a monomer or dimer at the cell surface. (wt) Normal Notch transactivation function is shown in which ligand (hJag1) interacts with receptor at the LBR (red). This interaction leads to cleavage of the receptor at the NRR (blue arrow and scissors). (A) Proposed model Ax (blue) function. Receptor with defective Ax region (blue with cross) is unable to effectively facilitate ligand binding. Therefore, activation is significantly decreased or absent depending on magnitude of the defect in Ax. Alternative model for Ax function (B) Receptor with defective Ax region (blue with cross) is able to bind ligand at the LBR (red), however the defective Ax region is less able to transduce the effect of ligand binding to the cleavage at the NRR (blue crossed arrow). Therefore, activation is decreased or absent depending on magnitude of the defect in Ax.

7.4. Future Work

The development of a SAXS pipeline for processing and evaluating data for multiple tandem EGF repeats developed in Chapter 4 - Structural Studies of the Abruptex region could be used for future studies of other regions of the receptor not amenable to crystallization, and systems where a low resolution model is sufficient for analysis of shape. This SAXS approach combined with NMR and X-ray crystallography could be used to investigate EGF16 which lacks the canonical interdomain packing aromatic (See Chapter 3). This region may be the location of a significant bend in the receptor which may facilitate an intramolecular interaction between Ax and LBR (See Figure 7.2). Furthermore, the Ax-like substitutions observed in dN EGF16 (Yamamoto et al, 2012) suggests that this domain may have an important structural role in regulating signalling.

Further investigation into the LBR binding site in EGF24 could also be carried out. This could involve observing the impact of substituting other Ca^{2+}-binding consensus residues in EGF24 using established assays as well as microscale thermophoresis to enable more precise determination of the interaction site. In addition, future experiments could focus on mapping other components of the interaction, by interacting longer multiple tandem EGF chains for example EGF7-13 with EGF23-29 and 11-17 with 23-29 to better understand the directionality of the observed interaction between Ax and LBR. This could ultimately involve affinity maturation of the interaction in an attempt to co-crystallize the Ax and LBR.

To understand more about the role of Ax in signalling, future experiments could focus on exploring the impact of Q462R in the activation assay. This residue is implicated to be involved in the interaction between LBR and the Ax (Chapter 5). Investigation into EGF28-29 at the extreme C-terminal end of Ax could be undertaken, which may not be rod-like as neither of these domains are cbEGFs. This is the location of the *N*-enhancer Ax substitutions in dN which, unlike the *N*-suppressor substitutions do not substitute residues in known conserved structural features.

A system to study the impact of Ax substitutions on cis-interactions could be developed, possibly involving co-transfection of a labelled receptor and ligand and looking for co-localisation, for example by Förster resonance energy transfer using fluorescent labels. Furthermore, development of a CRISPR/Cas9 cell-based system could be undertaken for replacing endogenous *NOTCH* genes with variants which are not overexpressed could also be implemented, which would more accurately model cis-interactions. This could extend to an organism system such as engineering mutations into *D. melanogaster* with CRISPR/Cas9 to enable screening for a phenotype indicative of Abruptex which could then be tested in cell-based or *in vitro* binding assays to infer molecular mechanism. These proposed future studies are described in Figure 7.4.

Investigation	Appropriate constructs	Proposed methodology
Investigating structure of the region between Ax and LBR	S2 hN1 EGF11-19 S2 hN1 EGF14-19	SAXS pipeline, X-ray crystallography, NMR
Investigating impact of EGF24 Ca^{2+}-binding variants on interaction with LBR	S2/PR hN1 EGF24 Ca^{2+}-binding variants	Flow cytometry, Microscale thermophoresis
Investigating impact of *N*-enhancer mutations on receptor activation	mN1 *N*-enhancer variants	Cell activation assay
Investigating impact of Q462R on receptor activation	mN1 Q462R variant	Cell activation assay
Investigating impact of Ax substitutions on cis-modulations	Labelled receptor variants and labelled ligand	Co-expression and observation (Förster resonance energy transfer)
Investigating impact of Ax substitutions on activation	Ax variant receptors	CRISPR/Cas9 to replace endogenous N1 with Ax variants, followed by luciferase complementation technique
Investigating the impact of proposed Ax substitutions on development	Variant dN	CRISPR/Cas9 to replace endogenous dN in *D. melanogaster*

Figure 7.4 Proposed future investigations to further characterize the role of the Abruptex region within the Notch receptor

7.5. Summary

This work suggests that EGF23-27 of the Ax region has an extended structure with a kink at the EGF25-26 interface inferred from bioinformatics analysis. The interaction between LBR and Ax involves EGF12 and EGF24, consisting of either the L468/O-Fuc or Q462/O-Gluc face of EGF12, but is dependent on both faces such that both modifications must be present for the interaction to occur. In addition, the interaction involves the N-terminal of EGF24. Although no impact of the EGF24 Ax substitution D909V was observed in the activation assay, the double Ax mutant DVNI showed a decrease in binding relative to wt and an EGF24-25 deletion mutant did not activate in this assay. These results suggest that in the case of Jagged1, the Abruptex region has a role in optimizing transactivation. Therefore, the enhanced signalling phenotype indicative of Abruptex mutants may be mediated by other means such as modulating cis-inhibition.

Bibliography

Alagille, D., Estrada, A., Hadchouel, M., Gautler, M., Odievre, M., & Dommergues, J. P. (1987). Syndromic paucity of interlobular bile ducts (Alagille syndrome or arteriohepatic dysplasia): review of 80 cases. *The Journal of pediatrics*, *110*(2), 195-200.

Artavanis-Tsakonas, S., Rand, M. D., & Lake, R. J. (1999). Notch signaling: cell fate control and signal integration in development. *Science*, *284*(5415), 770-776.

Baldock, C., Siegler, V., Bax, D. V., Cain, S. A., Mellody, K. T., Marson, A., Haston, J.L., Berry, R., Wang, M.C., Grossmann, J.G. & Roessle, M. (2006). Nanostructure of fibrillin-1 reveals compact conformation of EGF arrays and mechanism for extensibility. *Proceedings of the National Academy of Sciences*, *103*(32), 11922-11927.

Ban, J., Aryee, D. N., Fourtouna, A., Van Der Ent, W., Kauer, M., Niedan, S., Machado, I., Rodriguez-Galindo, C., Tirado, O.M., Schwentner, R. & Picci, P. (2014). Suppression of deacetylase SIRT1 mediates tumor suppressive NOTCH response and offers a novel treatment option in metastatic Ewing sarcoma. *Cancer research*, canres-1736.

Baron, M., Norman, D. G., Harvey, T. S., Campbell, I. D., Handford, P. A., Mayhew, M., G. G. Brownlee, & Tse, A. G. D. (1992). The three-dimensional structure of the first EGF-like module of human factor IX: Comparison with EGF and TGF-α. *Protein Science*, *1*(1), 81-90.

Becam, I., Fiuza, U. M., Arias, A. M., & Milán, M. (2010). A role of receptor Notch in ligand cis-inhibition in Drosophila. *Current Biology*, *20*(6), 554-560.

Becam, I., & Milán, M. (2008). A permissive role of Notch in maintaining the DV affinity boundary of the Drosophila wing. *Developmental biology*, *322*(1), 190-198.

Benedito, R., Roca, C., Sörensen, I., Adams, S., Gossler, A., Fruttiger, M., & Adams, R. H. (2009). The notch ligands Dll4 and Jagged1 have opposing effects on angiogenesis. *Cell*, *137*(6), 1124-1135.

Borggrefe, T., & Oswald, F. (2009). The Notch signaling pathway: transcriptional regulation at Notch target genes. *Cellular and molecular life sciences*, *66*(10), 1631-1646.

Brennan, K., Tateson, R., Lieber, T., Couso, J. P., Zecchini, V., & Arias, A. M. (1999). The abruptex mutations of notch disrupt the establishment of proneural clusters in Drosophila. *Developmental biology*, *216*(1), 230-242.

Chen, J., Lu, L., Shi, S., & Stanley, P. (2006). Expression of Notch signaling pathway genes in mouse embryos lacking β4galactosyltransferase-1. *Gene expression patterns*, *6*(4), 376-382.

Chillakuri, C. R., Sheppard, D., Ilagan, M. X. G., Holt, L. R., Abbott, F., Liang, S., Kopan, R., Handford, P.A. & Lea, S. M. (2013). Structural analysis uncovers lipid-binding properties of Notch ligands. *Cell reports*, *5*(4), 861-867.

Cho, W., & Stahelin, R. V. (2006). Membrane binding and subcellular targeting of C2 domains. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, *1761*(8), 838-849.

Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H. & Sheard, B. (1987). The solution structure of human epidermal growth factor. *Nature*, *327*(6120), 339.

Cordle, J., Johnson, S., Tay, J. Z. Y., Roversi, P., Wilkin, M. B., De Madrid, B. H., Shimizu, H., Jensen, S., Whiteman, P., Jin, B. & Redfield, C. (2008). A conserved face of the Jagged/Serrate DSL domain is involved in Notch trans-activation and cis-inhibition. *Nature Structural and Molecular Biology*, *15*(8), 849.

Crosnier, C., Driancourt, C., Raynaud, N., Dhorne–Pollet, S., Pollet, N., Bernard, O., Hadchouel, M. & Meunier–Rotival, M. (1999). Mutations in JAGGED1 gene are predominantly sporadic in Alagille syndrome. *Gastroenterology*, *116*(5), 1141-1148.

de Celis, J. F., & Bray, S. J. (2000). The Abruptex domain of Notch regulates negative interactions between Notch, its ligands and Fringe. *Development*, *127*(6), 1291-1302.

de Celis, J., & Garcia-Bellido, A. (1994). Roles of the Notch gene in Drosophila wing morphogenesis. *Mechanisms of development*, *46*(2), 109-122.

de Celis, J. F., Marí-Beffa, M., & García-Bellido, A. (1991). Cell-autonomous role of Notch, an epidermal growth factor homologue, in sensory organ differentiation in Drosophila. *Proceedings of the National Academy of Sciences*, *88*(2), 632-636.

de Jongh, W. A., dos SM Resende, M., Leisted, C., Strøbæk, A., Berisha, B., Nielsen, M. A., Salanti, A., Hjerrild, K., Draper, S. & Dyring, C. (2013, December). Development of a Drosophila S2 insect-cell based placental malaria vaccine production process. In *BMC proceedings* (Vol. 7, No. 6, p. P20). BioMed Central.

del Álamo, D., Rouault, H., & Schweisguth, F. (2011). Mechanism and significance of cis-inhibition in Notch signalling. *Current Biology*, *21*(1), R40-R47.

De Strooper, B., Annaert, W., Cupers, P., Saftig, P., Craessaerts, K., Mumm, J. S., Schroeter, E.H., Schrijvers, V., Wolfe, M.S., Ray, W.J. & Goate, A. (1999). A presenilin-1-dependent γ-secretase-like protease mediates release of Notch intracellular domain. *Nature*, *398*(6727), 518.

Dinchuk, J. E., Focht, R. J., Kelley, J. A., Henderson, N. L., Zolotarjova, N. I., Wynn, R., Neff, N.T., Link, J., Huber, R.M., Burn, T.C. & Rupar, M. J. (2002). Absence of post-translational aspartyl β-hydroxylation of epidermal growth factor domains in mice leads to developmental defects and an increased incidence of intestinal neoplasia. *Journal of Biological Chemistry*, *277*(15), 12970-12977.

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D., & Handford, P. A. (1996). Solution structure of a pair of calcium-binding epidermal growth factor-like domains: implications for the Marfan syndrome and other genetic disorders. *Cell*, *85*(4), 597-605.

D'Souza, B., Meloty-Kapella, L., & Weinmaster, G. (2010). Canonical and non-canonical Notch ligands. In *Current topics in developmental biology* (Vol. 92, pp. 73-129). Academic Press.

Dunwoodie, S. L. (2009). The role of Notch in patterning the human vertebral column. *Current opinion in genetics & development*, *19*(4), 329-337.

Eiraku, M., Tohgo, A., Ono, K., Kaneko, M., Fujishima, K., Hirano, T., & Kengaku, M. (2005). DNER acts as a neuron-specific Notch ligand during Bergmann glial development. *Nature neuroscience*, *8*(7), 873.

Ellisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D., & Sklar, J. (1991). TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell*, *66*(4), 649-661.

Fehon, R. G., Kooh, P. J., Rebay, I., Regan, C. L., Xu, T., Muskavitch, M. A., & Artavanis-Tsakonas, S. (1990). Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in Drosophila. *Cell*, *61*(3), 523-534.

Foster, G. G. (1975). Negative complementation at the Notch locus of Drosophila melanogaster. *Genetics*, *81*(1), 99-120.

Franke, D. and Svergun, D.I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. J. Appl. Cryst., 42, 342-346.

Fryxell, K. J., Soderlund, M., & Jordan, T. V. (2001). An animal model for the molecular genetics of CADASIL. *Stroke*, *32*(1), 6-11.

Geffers, I., Serth, K., Chapman, G., Jaekel, R., Schuster-Gossler, K., Cordes, R., Geffers, I., Serth, K., Chapman, G., Jaekel, R., Schuster-Gossler, K., Cordes, R., Sparrow, D.B., Kremmer, E., Dunwoodie, S.L., Klein, T. & Gossler, A. (2007). Divergent functions and distinct localization of the Notch ligands DLL1 and DLL3 in vivo. *The Journal of cell biology*, *178*(3), 465-476.

Gomez-Lamarca, M. J., Falo-Sanjuan, J., Stojnic, R., Rehman, S. A., Muresan, L., Jones, M. L., Pillidge, Z., Cerda-Moya, G., Yuan, Z., Baloul, S. & Valenti, P. (2018). Activation of the Notch

Signaling Pathway In Vivo Elicits Changes in CSL Nuclear Dynamics. *Developmental cell*, *44*(5), 611-623.

Gordon, W. R., Roy, M., Vardar-Ulu, D., Garfinkel, M., Mansour, M. R., Aster, J. C., & Blacklow, S. C. (2009). Structure of the Notch1-negative regulatory region: implications for normal activation and pathogenic signaling in T-ALL. *Blood*, *113*(18), 4381-4390.

Gordon, W. R., Zimmerman, B., He, L., Miles, L. J., Huang, J., Tiyanont, K., McArthur, D.G., Aster, J.C., Perrimon, N., Loparo, J.J. & Blacklow, S. C. (2015). Mechanical allostery: evidence for a force requirement in the proteolytic activation of Notch. *Developmental cell*, *33*(6), 729-736.

Gray, M. J., Kim, C. A., Bertola, D. R., Arantes, P. R., Stewart, H., Simpson, M. A., Irving, M.D. & Robertson, S. P. (2012). Serpentine fibula polycystic kidney syndrome is part of the phenotypic spectrum of Hajdu–Cheney syndrome. *European Journal of Human Genetics*, *20*(1), 122.

Hambleton, S., Valeyev, N. V., Muranyi, A., Knott, V., Werner, J. M., McMichael, A. J., Handford, P.A. & Downing, A. K. (2004). Structural and functional properties of the human notch-1 ligand binding region. *Structure*, *12*(12), 2173-2183.

Handford, P. A., Baron, M., Mayhew, M., Willis, A., Beesley, T., Brownlee, G. G., & Campbell, I. D. (1990). The first EGF-like domain from human factor IX contains a high-affinity calcium binding site. *The EMBO Journal*, *9*(2), 475-480.

Handford, P. A., Mayhew, M., Baron, M., Winship, P. R., Campbell, I. D., & Brownlee, G. G. (1991). Key residues involved in calcium-binding motifs in EGF-like domains. *Nature*, *351*(6322), 164.

Harris, R. J., & Spellman, M. W. (1993). O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology*, *3*(3), 219-224.

Hartley, D. A., Xu, T. A., & Artavanis-Tsakonas, S. (1987). The embryonic expression of the Notch locus of Drosophila melanogaster and the implications of point mutations in the extracellular EGF-like domain of the predicted protein. *The EMBO Journal*, *6*(11), 3407-3417.

Henderson, S. T., Gao, D., Christensen, S., & Kimble, J. (1997). Functional domains of LAG-2, a putative signaling ligand for LIN-12 and GLP-1 receptors in Caenorhabditis elegans. *Molecular biology of the cell*, *8*(9), 1751-1762.

Hu, Q. D., Ang, B. T., Karsak, M., Hu, W. P., Cui, X. Y., Duka, T., Takeda, Y., Chia, W., Sankar, N., Ng, Y.K. & Ling, E. A. (2003). F3/contactin acts as a functional ligand for Notch during oligodendrocyte maturation. *Cell*, *115*(2), 163-175.

Hutchinson, S., Aplin, R. T., Webb, H., Kettle, S., Timmermans, J., Boers, G. H., & Handford, P. A. (2005). Molecular effects of homocysteine on cbEGF domain structure: insights into the pathogenesis of homocystinuria. *Journal of molecular biology*, *346*(3), 833-844.

Hutterer, A., & Knoblich, J. A. (2005). Numb and α-Adaptin regulate Sanpodo endocytosis to specify cell fate in Drosophila external sensory organs. *EMBO reports*, *6*(9), 836-842.

Ilagan, M. X. G., Lim, S., Fulbright, M., Piwnica-Worms, D., & Kopan, R. (2011). Real-time imaging of notch activation with a luciferase complementation-based reporter. *Sci. Signal.*, *4*(181), rs7-rs7.

Irvine, K. D., & Wieschaus, E. (1994). fringe, a Boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during Drosophila wing development. *Cell*, *79*(4), 595-606.

Jensen S.A., Corbett A.R., Knott V., Redfield C. (2005). Handford P.A. Ca2+-dependent interface formation in fibrillin-1. *J. Biol. Chem,* 280, 14076–14084.

Jensen, S. A., Iqbal, S., Lowe, E. D., Redfield, C., & Handford, P. A. (2009). Structure and interdomain interactions of a hybrid domain: a disulphide-rich module of the fibrillin/LTBP superfamily of matrix proteins. *Structure*, *17*(5), 759-768.

Joutel, A., Corpechot, C., Ducros, A., Vahedi, K., Chabriat, H., Mouton, P., Alamowitch, S., Domenga, V., Cécillion, M., Maréchal, E. & Maciazek, J. (1996). Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature*, *383*(6602), 707.

Joutel, A., Vahedi, K., Corpechot, C., Troesch, A., Chabriat, H., Vayssière, C., Cruaud, C., Maciazek, J., Weissenbach, J., Bousser, M.G. & Bach, J. F. (1997). Strong clustering and stereotyped nature of Notch3 mutations in CADASIL patients. *The Lancet*, *350*(9090), 1511-1515.

Kakuda, S., & Haltiwanger, R. S. (2017). Deciphering the fringe-mediated notch code: identification of activating and inhibiting sites allowing discrimination between ligands. *Developmental cell*, *40*(2), 193-201.

Kamath, B. M., Podkameni, G., Hutchinson, A. L., Leonard, L. D., Gerfen, J., Krantz, I. D., Piccoli, D.A., Spinner, N.B., Loomes, K.M. & Meyers, K. (2012). Renal anomalies in Alagille syndrome: A disease-defining feature. *American journal of medical genetics Part A*, *158*(1), 85-89.

Katoh, M., & Katoh, M. (2007). Integrative genomic analyses on HES/HEY family: Notch-independent HES1, HES3 transcription in undifferentiated ES cells, and Notch-dependent HES1, HES5, HEY1, HEY2, HEYL transcription in fetal tissues, adult tissues, or cancer. International journal of oncology, 31(2), 461-466.

Kelly, D. F., Lake, R. J., Middelkoop, T. C., Fan, H. Y., Artavanis-Tsakonas, S., & Walz, T. (2010). Molecular structure and dimeric organization of the Notch extracellular domain as revealed by electron microscopy. *PloS one*, *5*(5), e10532.

Kidd, S., & Lieber, T. (2002). Furin cleavage is not a requirement for Drosophila Notch function. *Mechanisms of development*, *115*(1), 41-51.

Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. FEBS letters, 589(19), 2570-2577.

Kim, J., Irvine, K. D., & Carroll, S. B. (1995). Cell recognition, signal induction, and symmetrical gene activation at the dorsal-ventral boundary of the developing Drosophila wing. *Cell*, *82*(5), 795-802.

Komatsu, H., Chao, M. Y., Larkins-Ford, J., Corkins, M. E., Somers, G. A., Tucey, T., Dionne, H.M., White, J.Q., Wani, K., Boxem, M. & Hart, A. C. (2008). OSM-11 facilitates LIN-12 Notch signaling during Caenorhabditis elegans vulval development. *PLoS biology*, *6*(8), e196.

Kopan, R., & Ilagan, M. X. G. (2009). The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell*, *137*(2), 216-233.

Ladi, E., Nichols, J. T., Ge, W., Miyamoto, A., Yao, C., Yang, L. T., Boulter, J., Sun, Y.E., Kintner, C. & Weinmaster, G. (2005). The divergent DSL ligand Dll3 does not activate Notch signaling but cell autonomously attenuates signaling induced by other DSL ligands. *The Journal of cell biology*, *170*(6), 983-992.

Lanford, P. J., Lan, Y., Jiang, R., Lindsell, C., Weinmaster, G., Gridley, T., & Kelley, M. W. (1999). Notch signalling pathway mediates hair cell development in mammalian cochlea. *Nature genetics*, *21*(3), 289.

Lee, S. S., Knott, V., Jovanović, J., Harlos, K., Grimes, J. M., Choulier, L., Mardon, H.J., Stuart, D.I. & Handford, P. A. (2004). Structure of the integrin binding fragment from fibrillin-1 gives new insights into microfibril organization. *Structure*, *12*(4), 717-729.

Lee, T. V., Sethi, M. K., Leonardi, J., Rana, N. A., Buettner, F. F., Haltiwanger, R. S., Bakker, H. & Jafar-Nejad, H. (2013). Negative regulation of notch signaling by xylose. *PLoS genetics*, *9*(6), e1003547.

Li, J. C. (1927). Development in Drosophila melanogaster. *Genetics*, *12*(1), 1.

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N. & Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic acids research*, *43*(W1), W580-W584

Li, W., Wu, J., Yang, J., Sun, S., Chai, R., Chen, Z. Y., & Li, H. (2015). Notch inhibition induces mitotically generated hair cells in mammalian cochleae via activating the Wnt pathway. *Proceedings of the National Academy of Sciences*, *112*(1), 166-171.

Li, Z., Han, K., Pak, J. E., Satkunarajah, M., Zhou, D., & Rini, J. M. (2017). Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nature chemical biology*, *13*(7), 757.

Lieber, T., Wesley, C. S., Alcamo, E., Hassel, B., Krane, J. F., Campos-Ortega, J. A., & Young, M. W. (1992). Single amino acid substitutions in EGF-like elements of Notch and Delta modify Drosophila development and affect cell adhesion in vitro. *Neuron*, *9*(5), 847-859.

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the dependency of cellular protein levels on mRNA abundance. *Cell*, *165*(3), 535-550.

Logeat, F., Bessia, C., Brou, C., LeBail, O., Jarriault, S., Seidah, N. G., & Israël, A. (1998). The Notch1 receptor is cleaved constitutively by a furin-like convertase. *Proceedings of the National Academy of Sciences*, *95*(14), 8108-8112.

Lowell, S., Jones, P., Le Roux, I., Dunne, J., & Watt, F. M. (2000). Stimulation of human epidermal differentiation by delta–notch signalling at the boundaries of stem-cell clusters. *Current Biology*, *10*(9), 491-500.

Luca, V. C., Jude, K. M., Pierce, N. W., Nachury, M. V., Fischer, S., & Garcia, K. C. (2015). Structural basis for Notch1 engagement of Delta-like 4. *Science*, *347*(6224), 847-853.

Luca, V. C., Kim, B. C., Ge, C., Kakuda, S., Wu, D., Roein-Peikar, M., Haltiwanger, R.S., Zhu, C., Ha, T. & Garcia, K. C. (2017). Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. *Science*, *355*(6331), 1320-1324.

Matsuno, K., Eastman, D., Mitsiades, T., Quinn, A. M., Carcanciu, M. L., Ordentlich, P., Kadesch, T. & Artavanis-Tsakonas, S. (1998). Human deltex is a conserved regulator of Notch signalling. *Nature genetics*, *19*(1), 74.

Matsuura, A., Ito, M., Sakaidani, Y., Kondo, T., Murakami, K., Furukawa, K., K., Nadano, D., Matsuda, T. & Okajima, T. (2008). O-linked N-acetylglucosamine is present on the extracellular domain of notch receptors. *Journal of Biological Chemistry*, *283*(51), 35486-35495.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A.P. & Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. *Nucleic acids research*, *41*(W1), W597-W600..

Meloty-Kapella, L., Shergill, B., Kuon, J., Botvinick, E., & Weinmaster, G. (2012). Notch ligand endocytosis generates mechanical pulling force dependent on dynamin, epsins, and actin. *Developmental cell*, *22*(6), 1299-1312.

Meng, H., Zhang, X., Hankenson, K. D., & Wang, M. M. (2009). Thrombospondin 2 potentiates notch3/jagged1 signaling. *Journal of Biological Chemistry*, *284*(12), 7866-7874.

Mohr, O. L. (1919). Character changes caused by mutation of an entire region of a chromosome in Drosophila. *Genetics*, *4*(3), 275.

Morgan, T. H. (1917). The theory of the gene. *The American Naturalist*, *51*(609), 513-544.

Morrissette, J. J., Colliton, R. P., & Spinner, N. B. (2001). Defective intracellular transport and processing of JAG1 missense mutations in Alagille syndrome. *Human molecular genetics*, *10*(4), 405-413.

Nandagopal, N., Santat, L. A., LeBon, L., Sprinzak, D., Bronner, M. E., & Elowitz, M. B. (2018). Dynamic ligand discrimination in the Notch signaling pathway. *Cell*, *172*(4), 869-880.

Nasarenko, I. I. (1930). Ein Fall wahrscheinlicher Verdoppelung eines Chromosomsttlckes bei Drosophila melanogaster. *Biol. Zentr*, *50*, 385-392.

Nyhan, K. C., Faherty, N., Murray, G., Cooey, L. B., Godson, C., Crean, J. K., & Brazil, D. P. (2010). Jagged/Notch signalling is required for a subset of TGFβ1 responses in human kidney epithelial cells. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1803*(12), 1386-1395.

Ogawa, M., Sawaguchi, S., Kawai, T., Nadano, D., Matsuda, T., Yagi, H., Kato, K., Furukawa, K. & Okajima, T. (2015). Impaired O-linked N-acetylglucosaminylation in the endoplasmic reticulum by mutated epidermal growth factor (EGF) domain-specific O-linked N-acetylglucosamine transferase found in Adams-Oliver syndrome. *Journal of Biological Chemistry*, *290*(4), 2137-2149.

Pei, Z., & Baker, N. E. (2008). Competition between Delta and the Abruptex domain of Notch. *BMC developmental biology*, *8*(1), 4.

Portin, P. (1975). Allelic negative complementation at the Abruptex locus of Drosophila melanogaster. *Genetics*, *81*(1), 121-133.

Proweller, A., Tu, L., Lepore, J. J., Cheng, L., Lu, M. M., Seykora, J., Millar, S.E., Pear, W.S. & Parmacek, M. S. (2006). Impaired notch signaling promotes de novo squamous cell carcinoma formation. *Cancer research*, *66*(15), 7438-7444.

Putnam, C. D., Hammel, M., Hura, G. L., & Tainer, J. A. (2007). X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly reviews of biophysics*, *40*(3), 191-285.

Ranganathan, P., Weaver, K. L., & Ranganathan, P., Weaver, K. L., & Capobianco, A. J. (2011). Notch signalling in solid tumours: a little bit of everything but not all the time. *Nature Reviews Cancer*, *11*(5), 338.

Rao, Z., Handford, P., Mayhew, M., Knott, V., Brownlee, G. G., & StuartZ, D. (1995) "The structure of a Ca2+-binding epidermal growth factor-like domain: its role in protein-protein interactions." *Cell* 82.1, 131-141.

Rebay, I., Fleming, R. J., Fehon, R. G., Cherbas, L., Cherbas, P., & Artavanis-Tsakonas, S. (1991). Specific EGF repeats of Notch mediate interactions with Delta and Serrate: implications for Notch as a multifunctional receptor. *Cell*, *67*(4), 687-699.

Rees, D. J., Jones, I. M., Handford, P. A., Walter, S. J., Esnouf, M. P., Smith, K. J., & Brownlee, G. G. (1988). The role of beta-hydroxyaspartate and adjacent carboxylate residues in the first EGF domain of human factor IX. *The EMBO Journal*, *7*(7), 2053-2061.

Robertson, I. B., Dias, H. F., Osuch, I. H., Lowe, E. D., Jensen, S. A., Redfield, C., & Handford, P. A. (2017). The N-Terminal Region of Fibrillin-1 Mediates a Bipartite Interaction with LTBP1. *Structure*, *25*(8), 1208-1221.

Rowan, S., Conley, K. W., Le, T. T., Donner, A. L., Maas, R. L., & Brown, N. L. (2008). Notch signaling regulates growth and differentiation in the mammalian lens. *Developmental biology*, *321*(1), 111-122.

Sanchez-Irizarry, C., Carpenter, A. C., Weng, A. P., Pear, W. S., Aster, J. C., & Blacklow, S. C. (2004). Notch subunit heterodimerization and prevention of ligand-independent proteolytic activation depend, respectively, on a novel domain and the LNR repeats. *Molecular and cellular biology*, *24*(21), 9265-9273.

Sawaguchi, S., Varshney, S., Ogawa, M., Sakaidani, Y., Yagi, H., Takeshita, K., Murohara, T., Kato, K., Sundaram, S., Stanley, P. & Okajima, T. (2017). O-GlcNAc on NOTCH1 EGF repeats regulates ligand-induced Notch signaling and vascular development in mammals. *Elife*, *6*, e24419.

Serth, K., Schuster-Gossler, K., Kremmer, E., Hansen, B., Marohn-Köhn, B., & Gossler, A. (2015). O-fucosylation of DLL3 is required for its function during somitogenesis. *PloS one*, *10*(4), e0123776.

Sharma, A., Rangarajan, A., & Dighe, R. R. (2013). Antibodies against the extracellular domain of human Notch1 receptor reveal the critical role of epidermal-growth-factor-like repeats 25–26in ligand binding and receptor activation. *Biochemical Journal*, *449*(2), 519-530.

Sharma, J., Panico, M., Barber, J., & Morris, H. R. (1997). Purification and determination of intact molecular mass by electrospray ionization mass spectrometry of the photosystem II reaction center subunits. *Journal of Biological Chemistry*, *272*(52), 33153-33157.

Shaya, O., Binshtok, U., Hersch, M., Rivkin, D., Weinreb, S., Amir-Zilberstein, L., Khamaisi, B., Oppenheim, O., Desai, R.A., Goodyear, R.J. and Richardson, G.P. (2017). Cell-cell contact area affects Notch signaling and Notch-dependent patterning. *Developmental cell*, *40*(5), 505-511.

Shergill, B., Meloty-Kapella, L., Musse, A. A., Weinmaster, G., & Botvinick, E. (2012). Optical tweezers studies on Notch: single-molecule interaction strength is independent of ligand endocytosis. *Developmental cell*, *22*(6), 1313-1320.

Shi, S., & Stanley, P. (2003). Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways. *Proceedings of the National Academy of Sciences*, *100*(9), 5234-5239.

Shimizu, K., Chiba, S., Kumano, K., Hosoya, N., Takahashi, T., Kanda, Y., Hamada, Y., Yazaki, Y. & Hirai, H. (1999). Mouse jagged1 physically interacts with notch2 and other notch receptors assessment by quantitative methods. *Journal of Biological Chemistry*, *274*(46), 32961-329

Shortridge, M. D., Hage, D. S., Harbison, G. S., & Powers, R. (2008). Estimating protein-ligand binding affinity using high-throughput screening by NMR. Journal of combinatorial chemistry, 10(6), 948-58.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J. & Thompson, J. D. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, *7*(1), 539.

Simpson, M. A., Irving, M. D., Asilmaz, E., Gray, M. J., Dafou, D., Elmslie, F. V., Mansour, S., Holder, S.E., Brain, C.E., Burton, B.K. & Kim, K. H. (2011). Mutations in NOTCH2 cause Hajdu-Cheney syndrome, a disorder of severe and progressive bone loss. *Nature genetics*, *43*(4), 303.

Singh, A., & Choi, K. W. (2003). Initial state of the Drosophila eye before dorsoventral specification is equivalent to ventral. *Development*, *130*(25), 6351-6360.

Stanley, P. (2007). Regulation of Notch signaling by glycosylation. *Current opinion in structural biology*, *17*(5), 530-535.

Stanley, P., & Guidos, C. J. (2009). Regulation of Notch signaling during T-and B-cell development by O-fucose glycans. *Immunological reviews*, *230*(1), 201-215.

Stephenson, N. L., & Avis, J. M. (2012). Direct observation of proteolytic cleavage at the S2 site upon forced unfolding of the Notch negative regulatory region. *Proceedings of the National Academy of Sciences*.

Stittrich, A. B., Lehman, A., Bodian, D. L., Ashworth, J., Zong, Z., Li, H., Lam, P., Khromykh, A., Iyer, R.K., Vockley, J.G. & Baveja, R. (2014). Mutations in NOTCH1 cause Adams-Oliver syndrome. *The American Journal of Human Genetics*, *95*(3), 275-284.

Suckling, R. J., Korona, B., Whiteman, P., Chillakuri, C., Holt, L., Handford, P. A., & Lea, S. M. (2017). Structural and functional dissection of the interplay between lipid and Notch binding by human Notch ligands. *The EMBO journal*, *36*(15), 2204-2215.

Svergun D.I., Barberato C. and Koch M.H.J. (1995) CRYSOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates J. Appl. Cryst. , 28, 768-773.

Takebe, N., Nguyen, D., & Yang, S. X. (2014). Targeting notch signaling pathway in cancer: clinical development advances and challenges. *Pharmacology & therapeutics*, *141*(2), 140-149.

Takeuchi, H., & Haltiwanger, R. S. (2010, August). Role of glycosylation of Notch in development. In *Seminars in cell & developmental biology* (Vol. 21, No. 6, pp. 638-645). Academic Press.

Takeuchi, H., Schneider, M., Williamson, D. B., Ito, A., Takeuchi, M., Handford, P. A., & Haltiwanger, R. S. (2018). Two novel protein O-glucosyltransferases that modify sites distinct from POGLUT1 and affect Notch trafficking and signaling. *Proceedings of the National Academy of Sciences*, 201804005.

Takeuchi, H., Yu, H., Hao, H., Takeuchi, M., Ito, A., Li, H., & Haltiwanger, R. S. (2017). O-Glycosylation modulates the stability of epidermal growth factor-like repeats and thereby regulates Notch trafficking. *Journal of Biological Chemistry*, jbc-M117.

Taylor, P., Takeuchi, H., Sheppard, D., Chillakuri, C., Lea, S. M., Haltiwanger, R. S., & Handford, P. A. (2014). Fringe-mediated extension of O-linked fucose in the ligand-binding region of Notch1 increases binding to mammalian Notch ligands. *Proceedings of the National Academy of Sciences*, 201319683.

Timmerman, L. A., Grego-Bessa, J., Raya, A., Bertrán, E., Pérez-Pomares, J. M., Díez, J., Aranda, S., Palomo, S., McCormick, F., Izpisúa-Belmonte, J.C. & de la Pompa, J. L. (2004).

Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes & development*, *18*(1), 99-115.

Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. IUCrJ 2, 207-217

Troilo, H., Steer, R., Collins, R. F., Kielty, C. M., & Baldock, C. (2016). Independent multimerization of Latent TGFβ Binding Protein-1 stabilized by cross-linking and enhanced by heparan sulfate. *Scientific Reports*, *6*, 34347.

Turnpenny, P. D., Alman, B., Cornier, A. S., Giampietro, P. F., Offiah, A., Tassy, O., Pourquié, O., Kusumi, K. & Dunwoodie, S. (2007). Abnormal vertebral segmentation and the notch signaling pathway in man. *Developmental dynamics: an official publication of the American Association of Anatomists*, *236*(6), 1456-1474.

van Tetering, G., van Diest, P., Verlaan, I., van der Wall, E., Kopan, R., & Vooijs, M. (2009). The metalloprotease ADAM10 is required for notch1 S2 cleavage. *Journal of Biological Chemistry*, jbc-M109.

Volkov, V. V. and Svergun, D. I. (2003) Uniqueness of ab-initio shape determination in small-angle scattering. J. Appl. Cryst., 36, 860-864.

Weisshuhn, P. C., Sheppard, D., Taylor, P., Whiteman, P., Lea, S. M., Handford, P. A., & Redfield, C. (2016). Non-linear and flexible regions of the human Notch1 extracellular domain revealed by high-resolution structural studies. *Structure*, *24*(4), 555-566.

Welshons, W. J. (1965). Analysis of a gene in Drosophila. *Science*, *150*(3700), 1122-1129.

Whiteman, P., de Madrid, B. H., Taylor, P., Li, D., Heslop, R., Viticheep, N., Tan, J.Z., Shimizu, H., Callaghan, J., Masiero, M. & Li, J. L. (2013). Molecular basis for Jagged-1/Serrate ligand recognition by the Notch receptor. *Journal of Biological Chemistry*, *288*(10), 7305-7312.

Whiteman, P., Redfield, C., & Handford, P. A. (2014). Bacterial expression and in vitro refolding of limited fragments of the Notch receptor and its ligands. In *Notch Signaling* (pp. 193-208). Humana Press, New York, NY.

Wouters, M. A., Rigoutsos, I., Chu, C. K., Feng, L. L., Sparrow, D. B., & Dunwoodie, S. L. (2005). Evolution of distinct EGF domains with specific functions. *Protein Science*, *14*(4), 1091-1103.

Yamamoto, S., Charng, W. L., Rana, N. A., Kakuda, S., Jaiswal, M., Bayat, V., ... & Wang, H. (2012). A mutation in EGF repeat-8 of Notch discriminates between Serrate/Jagged and Delta family ligands. *Science*, *338*(6111), 1229-1232.

Yamamoto, N., Tanigaki, K., Han, H., Hiai, H., & Honjo, T. (2003). Notch/RBP-J signaling regulates epidermis/hair fate determination of hair follicular stem cells. *Current biology*, *13*(4), 333-338.

Zou, J., Li, P., Lu, F., Liu, N., Dai, J., Ye, J., Qu, X., Sun, X., Ma, D., Park, J. & Ji, C. (2013). Notch1 is required for hypoxia-induced proliferation, invasion and chemoresistance of T-cell acute lymphoblastic leukemia cells. *Journal of hematology & oncology*, *6*(1), 3.

Appendix

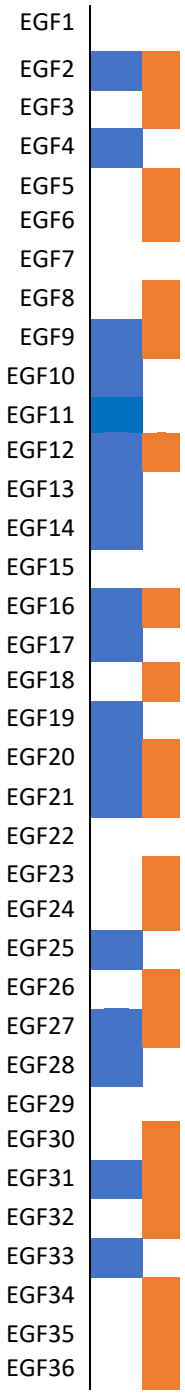


Table A.1 Glucose and fucose modification in hN1: Predicted and experimentally shown Glucose (blue) and Fucose (orange) modifications are shown for all EGFs in hN1.

Peptide sequences

S2 hN1 11-13

QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCICMPG
YEGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMV
WNHRGHHHHHHHH

S2 hN1 11-13 A426R

QDVDECSLGANPCEHRGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCICMPGY
EGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMVW
NHRGHHHHHHHH

S2 hN1 11-13 K428E

QDVDECSLGANPCEHAGECINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCICMPGY
EGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMVW
NHRGHHHHHHHH

S2 hN1 11-13 E437R

QDVDECSLGANPCEHAGKCINTLGSFRCQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCICMPG
YEGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMV
WNHRGHHHHHHHH

S2 hN1 11-13 R448E

QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPECEIDVNECVSNPCQNDATCLDQIGEFQCICMPGY
EGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMVW
NHRGHHHHHHHH

S2 hN1 11-13 N459R

QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSRPCQNDATCLDQIGEFQCICMPGY
EGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMVW
NHRGHHHHHHHH

S2 hN1 11-13 Q462R

QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCRNDATCLDQIGEFQCICMPGY
EGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMVW
NHRGHHHHHHHH

S2 hN1 11-13 L468R
QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCRDQIGEFQCICMPG
YEGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMV
WNHRGHHHHHHH

S2 hN1 11-13 V485R

QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCICMPG
YEGRHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQGPGSLHHILDAQKMV
WNHRGHHHHHHH

S2 hN1 11-13cl

QDVDECSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCICMPG
YEGVHCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLCQYDGSGLEVLFQ

S2 hN1 23-24

MRGSHHHHHHGSIEGRSAVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHN
GGSCTDGINTAFCDCLPGFRGTFCELDLHHILDAQKMVWNHR

S2 hN1 23-24CHis

DINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHNGGSCTDGINTAFCDCLPGFRG
TFCEGSGLEVLFQGPGSLHHILDAQKMVWNHRGHHHHHHHH

S2 hN1 23-29

DINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHNGGSCTDGINTAFCDCLPGFRG
TFCEEDINECASDPCRNGANCTDCVDSYTCTCPAGFSGIHCENNTPDCTESSCFNGGTCVDGINSFTCLCPPG
FTGSYCQHDVNECDSQPCLHGGTCQDGCGSYRCTCPQGYTGPNCQNLVHWCDSSPCKNGGKCWQTHTQ
YRCECPSGWTGLYCDVPSVSCEVAAQRQGVDVARLCQHGGLCVDAGNTHHCRCQAGYTGSYCEGSGLEVL
FQGPGSLHHILDAQKMVWNHRGHHHHHHHH

S2 hN1 23-27

DINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHNGGSCTDGINTAFCDCLPGFRG
TFCEEDINECASDPCRNGANCTDCVDSYTCTCPAGFSGIHCENNTPDCTESSCFNGGTCVDGINSFTCLCPPG
FTGSYCQHDVNECDSQPCLHGGTCQDGCGSYRCTCPQGYTGPNCQGSGLEVLFQGPGSLHHILDAQKMV
WNHRGHHHHHHHH

PR hN1 23-27CSCS

MRGSHHHHHHGSIEGRSAVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHN
GGSCTDGINTAFCDCLPGFRGTFCEEDINECASDPCRNGANCTDSVDSYTCTCPAGFSGIHCENNTPDCTESS
CFNGGTCVDGINSFTCLCPPGFTGSYCQHDVNECDSQPCLHGGTCQDGSGSYRCTCPQGYTGPNCQLDLH
HILDAQKMVWNHR

PR hN1 23-24

MRGSHHHHHHGSIEGRSAVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHN
GGSCTDGINTAFCDCLPGFRGTFCELDLHHILDAQKMVWNHR

PR hN1 23-24 T940R

MRGSHHHHHHGSIEGRSAVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHN
GGSCTDGINTAFCDCLPGFRGTRCELDLHHILDAQKMVWNHR

PR hN1 23-24 H917W

MRGSHHHHHHGSIEGRSAVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCWN
GGSCTDGINTAFCDCLPGFRGTFCELDLHHILDAQKMVWNHR

PR hN1 23-24 D932R

MRGSHHHHHHGSIEGRSAVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHN
GGSCTDGINTAFCRCLPGFRGTFCELDLHHILDAQKMVWNHR

PR hN1 23-24 L934R

MRGSHHHHHHGSIEGRSAVDINECVLSPCRHGASCQNTHGGYRCHCQAGYSGRNCETDIDDCRPNPCHN
GGSCTDGINTAFCDCRPGFRGTFCELDLHHILDAQKMVWNHR

Attempted Electron Microscopy experiment

A long EGF-domain constructed encompassing hN1 EGF10-24 was expressed in the S2 system as described in Chapter 3. Although this construct only included EGF24 of the Abruptex region, the purified protein was submitted to Prof Susan Lea, and Dr Rich Suckling, University of Oxford for cryo-EM analysis. The S2 purification is outlined in Chapter 3 and involved a Ni^{2+} affinity chromatography step followed by gel filtration purification on an S200 SEC column. The gel filtration chromatogram and SDS-PAGE suggest a heterogenous range of shapes and/or N-terminal truncations of the protein (Figure A.2). These N-terminal truncations would have an intact His-tag, enabling selection during the Ni^{2+} affinity chromatography step. The subsequent EM experiment was carried out by Lea et al. as a negative-stain EM experiment on a Tecnai T12 electron microscope. This experiment did not yield any useable data for structural analysis, and images were not collected for this negative result. Therefore, considering the lack of evidence that this technique would prove viable for structural analysis of the Abruptex region, cryo-EM was not further pursued in this project.

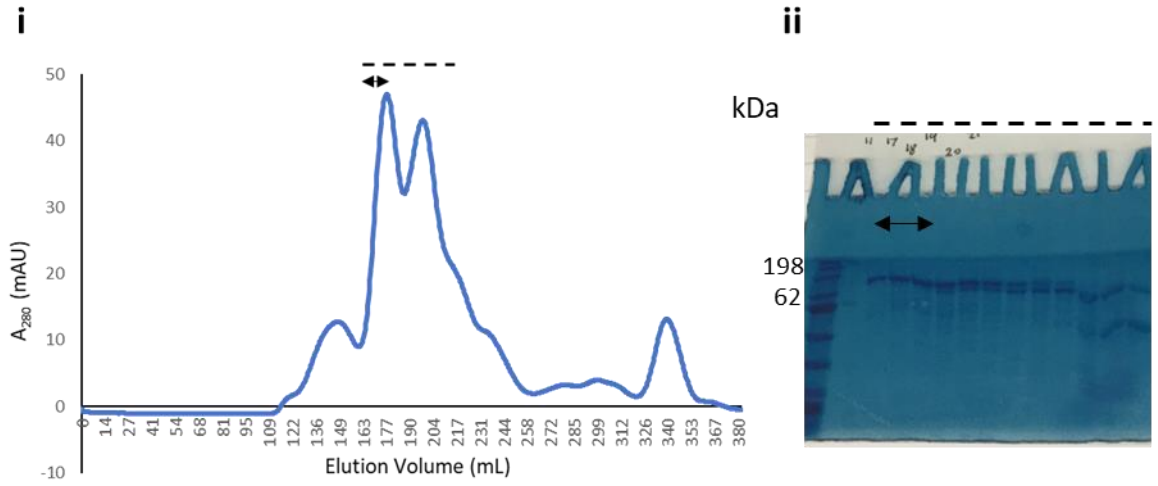


Figure A.2 Purification of S2 hN1 10-24: (i) Following Ni^{2+} affinity chromatography, size exclusion chromatography (SEC) was carried out and absorbance at 280 nm recorded over the course of the isocratic elution. (ii) Peaks from this elution (dashed line) were analyzed by reducing 16 % SDS-PAGE stained with Coomassie blue, and fractions pooled for further experiments (double headed arrow). This elution shows that more than half of the material eluted from the Ni^{2+} column is discarded during SEC. MW inferred from ladder indicates a mass greater than the expected MW of 67214, however glycosylations on the purified protein would add to the perceived mass in gel filtration. The complexity of the elution profile and multiple bands on the gel filtration column likely correspond to N-terminal truncations, which were selected for during Ni^{2+}-affinity chromatography.