Characterisation of the Notch receptor/ligand interaction using *C. elegans* as a model



Laurie Holt

Department of Biochemistry and Exeter College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy (DPhil)

Hilary Term 2016

Abstract

Characterisation of the Notch receptor/ligand interaction using *C. elegans* as a model

A thesis submitted for the degree of Doctor of Philosophy (DPhil) by Laurie Holt, Exeter College, Hilary Term 2016

The core aim of this project was to use the nematode worm *C. elegans* to probe the function of the N-terminus of Notch ligands using complementary molecular and *in vivo* approaches. Sequence analysis and structural predictions suggested that the N-termini of *C. elegans* Notch ligands LAG-2 and APX-1 may comprise a C2 domain, although it was concluded structural information would be necessary to confirm this. Eukaryotic expression systems were used to produce recombinant protein fragments, and ultimately yielded a large quantity of the fragment APX-1_{NE1FcHis} that was purified and used for crystallography studies, although no structural information has yet been obtained from this. In parallel, assays were established to assess rescue of the mutant phenotypes associated with the *lag-2(q420)* allele. Extrachromosomal expression of LAG-2 and APX-1 partially rescued the Lag phenotype, although overexpression of ligand appeared to cause an increase in embryonic lethality. An alternative assay was established, to assess rescue of the anchor cell duplication phenotype. Using this assay, it was demonstrated that deletion of the N-terminal MNNL or DSL domains of APX-1 abolished rescue of the phenotype observed for WT APX-1 confirming the importance of these domains in Notch signalling. Rescue was still achieved after deletion of the EGF-like 1 domain of APX-1, and substitutions targeting loop regions in the DSL and MNNL (DSL(ETSD), I80-K86/SS, respectively), but the extent of this rescue appeared to be reduced in all three constructs, suggesting these regions are required for optimised Notch signalling.

Acknowledgements

I would like to thank my supervisor Penny Handford for her contribution throughout this work. I am also extremely grateful to our collaborator Alison Woollard, for her valuable input. I would also like to thank the Medical Research Council, who have funded this research.

I'm very thankful for the support I've received from Handford lab members past and present, particularly to Mouli Chillakuri for plasmids and methods training. Many thanks to Richard Suckling, Matthias Jore and their group leader Susan Lea for their collaboration using the S2 expression system, along with Pat Whiteman and Bogusia Korona for their advice. I am also very grateful to Sacha Jensen for all the helpful discussions, proofreading, moral support and sweeties he has provided over the years.

I would also like to thank the members of the Woollard group both past and present, particularly Sophie Gilbert for microinjection and microscopy training, as well as Pete Appleford and Sam Hughes for their plasmids and mosSCI advice. I wish to extend my thanks to the worm community as a whole - who have provided a very pleasant and supportive environment in which to pursue research.

A particularly big 'thank you' to my extremely supportive family; Mum, Dad, Allie, Matt and last (but by no means least) Jake, who have been absolute stars. Thank you for the advice, the sympathy, and the attempts to understand my long explanations about green and red worms! I love you all very much, and I'm so grateful that you've all been by my side throughout this project.

Contents

List of Figures viii

List of Tables xiii

List of Abbreviations xv

1 Introduction 1

1.1 Overview . . . 1
1.2 Physiological roles of Notch signalling . . . 2
1.3 The Canonical Notch Pathway . . . 5
1.3.1 *Trans*-activation and *Cis*-inhibition . . . 5
1.3.2 Components of the canonical pathway and their conservation between species . . . 7
1.4 Notch signalling and disease . . . 9
1.4.1 Alagille syndrome . . . 12
1.4.2 CADASIL . . . 13
1.4.3 Cancer . . . 15
1.5 Structural and molecular studies of Notch receptors . . . 17
1.5.1 The EGF-like region . . . 17
1.5.2 Glycosylation of EGF-like repeats of Notch receptor . . . 20
1.5.3 The negative regulatory region (NRR) . . . 23
1.5.4 The Notch intracellular domain (NICD) . . . 24
1.6 Structural and molecular studies of Notch ligands . . . 28
1.6.1 The DSL domain . . . 28
1.6.2 The module at the N-terminus of Notch ligands (MNNL) . . . 33
1.6.3 The Notch receptor/ligand interaction: new insights by co-crystallisation . . . 37
1.7 Aims of the thesis . . . 42

2 Materials and Methods **43**
2.1 DNA methods . . . 43
2.1.1 General DNA methods . . . 43
2.1.1.1 Polymerase chain reaction (PCR) . . . 43
2.1.1.2 DNA gel electrophoresis . . . 43
2.1.1.3 Gel extraction . . . 44
2.1.1.4 T/A cloning with TOPO-XL . . . 44
2.1.1.5 Restriction digestion and ligation . . . 44
2.1.1.6 Quikchange mutagenesis . . . 45
2.1.1.7 Transformation of chemically competent cells . . . 45
2.1.1.8 Transformation of electro-competent cells . . . 46
2.1.1.9 Colony screening . . . 46
2.1.2 Generating constructs for expression in HEK293T cells . . . 46
2.1.3 Generating constructs for expression in *Drosophila* S2 cells . . . 48
2.1.4 Generating constructs for *C. elegans* microinjection . . . 48
2.1.4.1 Preparation of fosmid DNA . . . 48
2.1.4.2 Isolation of target sequences for use in rescue constructs . . . 49
2.1.4.3 Assembly of rescue constructs for microinjection and extrachromosomal array formation . . . 49
2.1.4.4 Assembly of rescue constructs for *mos-1* mediated single copy insertion (mosSCI) . . . 56
2.2 RNA methods . . . 59
2.2.1 Semi-quantitative gene expression analysis . . . 59
2.3 Protein methods . . . 60
2.3.1 Expression and purification of recombinant proteins . . . 60
2.3.1.1 HEK293T transfection for protein expression . . . 60
2.3.1.2 S2 transfection for protein expression . . . 60
2.3.1.3 HEK293T nickel purification . . . 60
2.3.1.4 S2 nickel purification . . . 61
2.3.1.5 Gel filtration . . . 61
2.3.2 Analysis of recombinant proteins . . . 62
2.3.2.1 SDS-Polyacrylamide gel electrophoresis (PAGE) . . . 62
2.3.2.2 Western blotting . . . 62
2.3.2.3 Limited proteolysis . . . 63

2.4 *C.elegans* methods . . . 64
2.4.1 Generation and maintenance of strains . . . 64
2.4.1.1 Growth and maintenance . . . 64
2.4.1.2 Freezing and thawing . . . 64
2.4.1.3 Crosses . . . 65
2.4.1.4 Genotyping by single worm lysis . . . 65
2.4.2 Generation of transgenic lines . . . 67
2.4.2.1 Generation of semi-stable extrachromosomal array transgenes . . . 67
2.4.2.2 Generation of mos-1 mediated single copy insertion transgenes (mosSCI) . . . 71
2.4.3 *C. elegans* imaging and assays . . . 72
2.4.3.1 Imaging . . . 72
2.4.3.2 Protruding vulva assay . . . 73
2.4.3.3 DAPI staining . . . 74
2.4.3.4 Transgene transmission efficiency assay . . . 74
2.4.3.5 Larval viability assay . . . 75
2.4.3.6 Brood fate assay . . . 75
2.4.3.7 Anchor cell duplication assay . . . 76

3 Expression and characterisation of *C. elegans* Notch proteins 77
3.1 Sequence analysis of *C. elegans* Notch proteins . . . 77
3.1.1 Analysis of *C. elegans* Notch receptors . . . 77
3.1.2 Analysis of the DSL domain of *C. elegans* Notch ligands . . . 81
3.1.3 Analysis of the MNNL region of *C. elegans* Notch ligands . . . 84
3.2 Expression of Notch protein fragments using eukaryotic cells . . . 89
3.2.1 Overview of target constructs . . . 89
3.2.2 Low levels of recombinant *C. elegans* Notch proteins are secreted by HEK293T cells . . . 91
3.2.3 Larger scale purifications of APX-1_{NE3FcHis} yield insufficient material for further characterisation . . . 94
3.2.4 APX-$_{\mathrm{NE3His}}$ can be expressed from S2 cells . . . 96
3.2.5 APX-1_{NE3His} can be expressed by S2 cells . . . 97
3.2.6 APX-1_{NE1His} can be purified from S2 cells with sufficiently high yield and quality for structural studies . . . 98

3.2.7 Limited proteolysis 98
3.3 Discussion 101

4 Developing an *in vivo* assay to study *C. elegans* Notch ligand function 103
4.1 Introduction 103
4.1.1 *C. elegans* as a model to study Notch signalling 103
4.1.1.1 The anatomy and life cycle of *C. elegans* 103
4.1.2 Identification and characterisation of Notch receptors and ligands in *C. elegans* 106
4.1.2.1 The *lag-2(q420)* allele is a temperature sensitive splice site mutation 108
4.1.3 Post embryonic developmental defects in *lag-2(q420)* mutants highlight the roles of Notch signalling in *C. elegans* 109
4.1.3.1 Anchor cell specification 110
4.1.3.2 Vulval precursor cell patterning 110
4.1.3.3 Germline proliferation 112
4.1.3.4 Summary of post embryonic developmental defects associated with *lag-2(q420)* mutants 114
4.2 Results 115
4.2.1 Quantification of *lag-2(q420)* mutant phenotypes 115
4.2.2 Rescue of the Lag phenotype of *lag-2(q420)* mutants 117
4.2.2.1 Generation of WT ligand::gfp constructs for rescue of the Lag phenotype of *lag-2(q420)* mutants 119
4.2.2.2 WT *lag-2::gfp* and *apx-1::gfp* can partially rescue the Lag phenotype in *lag-2(q420)* mutants 121
4.2.3 Extrachromosomal *lag-2::gfp* expression can cause gonad and vulval defects 124
4.2.4 Presence of the C-terminal GFP tag does not appear to alter the extent of rescue LAG-2 126
4.2.5 Dilution of *lag-2::gfp* is not sufficient to enhance rescue of the Lag phenotpye in *lag-2(q420)* mutants 127
4.2.6 *lag-2::gfp* expression is associated with increased embryonic lethality in *lag-2(q420)* mutants 130
4.3 Discussion 135

5 Single copy insertion and promoter considerations 140
5.1 *Mos-1* mediated single copy insertion as an alternative to extrachromosomal arrays . . . 140
5.1.1 Introduction . . . 140
5.1.2 Results . . . 141
5.2 The *lag-2* promoter - analysis and outcomes . . . 145
5.3 Discussion . . . 150

6 Rescue of the AC duplication phenotype to probe *C. elegans* Notch ligand function 153
6.1 Anchor cell duplication in *lag-2(q420)* mutants can be used to assay rescue constructs . . . 153
6.2 The effects of domain deletions in APX-1 . . . 156
6.3 Targeted mutagenesis of APX-1 . . . 161
6.4 Discussion . . . 165

7 General discussion and future directions 169

8 Appendices 178
8.1 Appendix 1 - Codon usage graphs . . . 178
8.2 Appendix 2 - FFAS primary sequence output . . . 179
8.3 Appendix 3 - Transgene transmission efficiency data . . . 181
8.4 Appendix 4 - Raw larval viability data and its normalisation . . . 184
8.5 Appendix 5 - Raw brood fate assay data and conversion to percentages 194
8.6 Appendix 6 - *lag-2* promoter considerations . . . 194
8.7 Appendix 7 - Raw AC duplication assay data and conversion to percentage reduction in phenotype . . . 204

Bibliography 208

List of Figures

1.1 The three modes of Notch signalling action 4
1.2 Domain organisation of Notch receptors and ligands 6
1.3 *Trans*-activation and *cis*-inhibition of the canonical Notch pathway . 8
1.4 The architecture of Notch receptors and their ligands 10
1.5 Structure of the Notch receptor major ligand-binding region 18
1.6 O-linked glycosylation modifications at amino acid consensus sequences in *Drosophila* 21
1.7 Structure of the negative regulatory region of Notch 25
1.8 Structure of *C. elegans* NICD in complex with CSL and LAG-3 and DNA 27
1.9 Model of ternary complex assembly 28
1.10 Structure of Jagged-1 DSL domain 31
1.11 Structure of the C2 domain at the N-terminus of human Jagged-1 . . 35
1.12 X-ray structure of rDLL-$4_{\text{SLP NE1}}$/$rN1_{11\text{-}13}$ complex 40

2.1 Scheme for generation of final constructs for microinjection 52
2.2 Scheme for generation of mosSCI constructs for microinjection 57
2.3 Example of digestion products for identification of the *lag-2(q420)* allele 66

3.1 Primary sequence alignment of LIN-12 and GLP-1 EGF-like regions . 78
3.2 Conservation of residues in the major-ligand-binding region of hNotch-1 and dNotch 79
3.3 Primary sequence alignment of LIN-12 and GLP-1 EGF-like regions with human and *Drosophila* Notch major ligand-binding domains . . 81
3.4 Primary sequence alignment of human and *Drosophila* Notch ligand DSL domains 82
3.5 Primary sequence alignment of *C. elegans* Notch ligand DSL domains 83

3.6 Primary sequence alignment of human and *C. elegans* Notch ligand C2/MNNL domains . . . 85
3.7 Superimposition of $\text{Serrate}_{\text{NE2}}$ and J1_{NE3} (4CC1) crystal structures . 86
3.8 *C. elegans* Notch ligand models generated using alignments to J1_{NE3} (4CBZ) . . . 87
3.9 Architecture of constructs designed for HEK293T expression . . . 91
3.10 Expression of *C. elegans* Notch proteins in HEK293T cells . . . 93
3.11 Analysis of purified APX-1 and Jag-1 recombinant proteins . . . 95
3.12 Comparison of levels of $\text{APX-1}_{\text{NE4FcHis}}$ before and after codon optimisation . . . 96
3.13 Analysis of purified APX-1NE3His from S2 cells . . . 98
3.14 Analysis of purified APX-1NE1His from S2 cells . . . 99
3.15 Tryspin and Endoproteinase Glu-C digestion time courses of $\text{APX-1}_{\text{NE1His}}$ . . . 100

4.1 Anatomy of an adult hermaphrodite . . . 104
4.2 Life cycle of *C. elegans* at 22°C . . . 105
4.3 The site of the *q420* allele in *lag-2* . . . 109
4.4 Specification of the AC/VU cells requires Notch signalling . . . 111
4.5 Notch provides the lateral signal for secondary cell fate specification during VPC patterning . . . 113
4.6 Anatomy of the adult hermaphrodite gonad . . . 114
4.7 The Lag phenotype is highly penetrant in *lag-2(q420)* mutants . . . 116
4.8 Post embryonic developmental defects associated with *lag-2(q420)* mutants . . . 118
4.9 Schematic of constructs generated to rescue *lag-2(q420)* mutants . . . 120
4.10 Rescue of the Lag phenotype in *lag-2(q420)* mutants by expression of WT Notch ligands . . . 122
4.11 Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by expression of WT Notch ligands . . . 123
4.12 Extrachromosomal *lag-2::gfp* expression can cause gonad migration defects . . . 125
4.13 Extrachromosomal *lag-2::gfp* expression can lead to formation of pseudovulvae . . . 126

4.14 Absence of the GFP tag does not affect the extent of rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 128
4.15 Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 with and without the GFP tag 129
4.16 Dilution of *lag-2::gfp* does not increase the extent of rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 131
4.17 Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 with and without the GFP tag 132
4.18 *lag-2::gfp* extrachromosomal array expression is associated with increased embryonic lethality in *lag-2(q420)* mutants 133
4.19 The effects of *lag-2::gfp* extrachromosomal array expression in the absence of the *lag-2(q420)* allele 134

5.1 Schematic of the generation of single copy insertions using mosSCI . . 142
5.2 Analysis of the *ttTi5605 lag-2::gfp* insertion 144
5.3 Analysis of the *ttTi5605 unc-62::gfp* insertion 146
5.4 Schematic of the *lag-2* promoter region 148
5.5 A reduced fragment of the *lag-2* promoter region cannot rescue the Lag phenotype 149
5.6 Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by a reduced *lag-2* promoter fragment 150

6.1 Expression of *lag-2::gfp*(AC) and *apx-1::gfp*(AC) can rescue the AC duplication phenotype in *lag-2(q420)* mutants 155
6.2 Statistical analysis suggests expression of *lag-2::gfp*(AC) and *apx-1::gfp*(AC) can rescue AC duplication phenotype to a similar extent . 156
6.3 Schematic of APX-1 whole domain deletion constructs and their expression 158
6.4 Deletion of the MNNL or DSL in APX-1 is associated with loss of AC duplication rescue in *lag-2(q420)* mutants 159
6.5 Statistical analysis of AC duplication rescue by domain deletion mutants 160
6.6 Generation of targeted loop region substitutions in the MNNL and DSL domain of APX-1 163
6.7 Specific mutations affecting the MNNL and DSL domain do allow some rescue of the AC duplication phenotype in *lag-2(q420)* mutants 164
6.8 Statistical analysis of AC duplication rescue by loop region mutants . 165

8.1 Frequency of optimal codons in APX-1_{NE4} sequence for human expression before and after optimisation by GenScript 178
8.2 Frequency of optimal codons in APX-1_{NE3His} sequence for *Drosophila* expression before and after optimisation by Geneart 179
8.3 Full LAG-2/4CBZ FFAS alignment 180
8.4 Full DSL-1/4CBZ FFAS alignment 180
8.5 Full APX-1/4CBZ FFAS alignment 181
8.6 Transgene transmission efficiency data 182
8.7 Transgene transmission efficiency data 183
8.8 Transgene transmission efficiency data 184
8.9 Maximum transgenic broods for each transgenic line calculated using transmission efficiency data 185
8.10 *unc-119* only rescued lines were not associated with 100% progeny survival after transgene transmission efficiency is accounted for 186
8.11 Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part I 188
8.12 Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part II 189
8.13 Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part III 190
8.14 Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part IV 191
8.15 Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part V 192
8.16 Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part VI 193
8.17 Raw data for brood fate assays 195
8.18 Brood fates for strains and transgenic lines expressed as percentages of the total brood 196
8.19 Sequence alignment of *Caenorhabditis lag-2* promoter regions part I . 197
8.20 Sequence alignment of *Caenorhabditis lag-2* promoter regions part II 198
8.21 Sequence alignment of *Caenorhabditis lag-2* promoter regions part II 199
8.22 Sequence alignment of *Caenorhabditis lag-2* promoter regions part IV 200
8.23 Sequence alignment of *Caenorhabditis lag-2* promoter regions part V 201
8.24 Sequence alignment of *Caenorhabditis lag-2* promoter regions part VI 202

8.25 Sequence alignment of *Caenorhabditis lag-2* promoter regions part VII 203
8.26 AC duplication assay data summary for control lines and WT ligand constructs 205
8.27 AC duplication assay data summary for whole domain deletion constructs 206
8.28 AC duplication assay data summary for loop mutant constructs . . . 207

List of Tables

1.1 Summary of the features associated with AGS 12
1.2 Summary of results reported by Cordle et al (2008a) for DSL mutant constructs 32

2.1 List of oligonucleotides used for amplification from cDNA library for cloning into modified pHLsec. 47
2.2 List of oligonucleotides used to generate final lag-2 promoter/42 bp synthetic intron region 50
2.3 List of oligonucleotides used to generate full length ligand cDNA sequences for rescue constructs. 51
2.4 List of oligonucleotides used to generate fusion constructs for insertion into *lag-2* promoter/TOPO-XL backbone. 53
2.5 List of oligonucleotides used to generate fusion constructs for insertion into *lag-2* promoter/TOPO-XL backbone. 54
2.6 List of oligonucleotides used to generate fusion constructs for insertion into *lag-2* promoter/TOPO-XL backbone. 55
2.7 List of oligonucleotides used to generate full length ligand cDNA sequences for rescue constructs. 58
2.8 New strains generated by genetic crossing 65
2.9 Transgenic lines generated using the unc-119$^+$ transformation marker, and their abbreviations 68
2.10 Transgenic lines generated using the unc-119$^+$ transformation marker, and their abbreviations 69
2.11 Transgenic lines generated with the *sur-5::gfp* transformation marker, and their abbreviations 70
2.12 List of oligonucleotides used to verify insertions at the *ttTi5605 mos-1* site. 72

3.1 Constructs for HEK293T expression properties 90

List of Abbreviations

A - adenine
AGS - Alagille syndrome
Å - *Å*ngstrom
AC - anchor cell
ADAM - A disintegrin and metalloprotease
ANK - ankyrin domain
APS - ammonium persulfate
APX - anterior pharynx in excess
bHLH - basic helix-loop-helix
bp - base pairs
BSA - bovine serum albumin
BTD - β trefoil domain
C - cytosine
cdh - cadherin
CADASIL - cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy
cb - calcium binding
C. elegans - *Caenorhabditis elegans*
CGC - Caenorhabditis Genetics Center
CIAP - calf intestinal alkaline phosphatase
CSL - CBF1/Suppressor of Hairless/LAG-1
CTD - C-terminal domain
d - *Drosophila*
ddH_20 - double-distilled water
dNTPs - deoxynucleotide 5′ triphsophates
Da - Dalton
DAPI - 4', 6-diamino-2-phenylindole
DIC - differential interference contrast
DLL - delta -like ligand
DMEM - Dulbecco's modified Eagle's medium
DMSO - dimethyl sulfoxide
DNA - deoxyribonucleic acid
DSL - Delta/Serrate/LAG-2
DTC - distal tip cell
DG - distal gonad
DTT- dithiothreitol
EDTA - 1-Ethyl-3(3-dimethylaminopropyl)-carbodiimmide hydrochloride
EGF - epidermal growth factor
ELISA - enzyme-linked immunosorbent assay
ER - endoplasmic reticulum
FBS - foetal bovine serum
FFAS - Fold and Function Alignment System
FITC - Fluorescein Isothiocyanate
FPLC - fast protein liquid chromatography

G - guanine
gf - gain of function
GFP - green fluorescent protein
GlcNAc - N-acetylglucosamine
GLP- germline proliferation defective
GSI - γ secretase inhibitor
h - human
HD - heterodimerisation
HEK - human embryonic kidney
HES - Hairy enhancer of split
him - high incidence of males
HRP - horseradish peroxidase
Kb - kilobase
kDa - kilodalton
LAG - lin-12 and glp-1
lf - loss-of-function
LIN - lineage defective
LNR - LIN-12/Notch repeats
mg - milligram
ml - millilitre
ms - millisecond
μ**g** - microgram
μ**l** - microlitre
$\mathbf{M_W}$ - molecular weight
MAML - Mastermind-like
MNNL - module at the N-terminus of Notch ligands
mosSCI - *mos-1* mediated single copy insertion
NEXT- Notch extracellular truncation
NGM - nematode growth media
NICD - Notch intracellular domain
NLS - nuclear localisation sequence
NMD - nonsense mediated decay
NRM - non red mover
NRR - negative regulatory region
NTD - N-terminal domain
OST - oligosaccharyl transferase
PAGE - polyacrylamide gel electrophoresis
PCR - polymerase chain reaction
PDB - protein data bank
PEI - polyethylenimine
PEST - proline, glutamate, serine, threonine rich
PLC - phospholipase C
PP - 1 phenoxy-2-propanol
PTM - post translational modification
PVul - protruding vulva
r - rat
RAM - $\mathrm{RBP_{j\kappa}}$ associated molecule
RNA - ribonucleic acid
rpm - revolutions per minute
SDS - sodium docecyl sulfate
SEM - standard error of the mean
SOP - sensory organ precursor
SPR - surface plasmon resonance
Su(H) - suppressor of hairless
T - thymine
T-ALL - T-cell acute lymphoblastic leukaemia
TAE - Tris/acetate/EDTA
TBS - Tris buffered saline
TE - Tris/EDTA (methods)
TE - transmission efficiency (*in vivo* data)
TEMED - N, N, N', N'- tetramethylethylenediamine
TM - transmembrane
Tris - tris(hydroxymethl) aminomethane

unc - uncoordinated
UTR - untranslated region
UV -ultra violet
V - volts
VPC - vulval precursor cell
VU - ventral uterine
WT - wild type
YT - yeast tryptone medium

1

Introduction

1.1 Overview

The Notch pathway is a fundamental intercellular signalling system that plays a critical role in metazoan development and adult tissue homeostasis. Notch signalling allows cells to adopt specific fates in a context-dependent manner, and its importance is illustrated by the severe human diseases that occur when it is improperly regulated. Activation of the canonical pathway occurs when a Notch receptor on the surface of a signal-receiving cell binds its corresponding ligand on an adjacent signal-sending cell, triggering a proteolytic cascade that ultimately leads to changes in the expression pattern of target genes in the signal-receiving cell.

Model organisms have been at the forefront of Notch signalling research and have provided us with essential insights into this well-conserved system. Observations of the notched-wing phenotype in *Drosophila* mutants were first recorded in the early 20th century (Dexter, 1914) (Morgan and Bridges, 1916) and the phenotypes associated with Notch mutations were subsequently characterised (Poulson, 1939). An advance that marked the beginning of a rapid expanse of our molecular knowledge of Notch was the cloning and sequence analysis of Notch in *Drosophila* (Wharton et al., 1985) complemented by studies of the homologue LIN-12 in *C.*

elegans (Greenwald, 1985). From these data, it was suggested that the Notch locus encodes a potential transmembrane protein with a large extracellular region, mostly consisting of EGF-like repeats. This led to a wealth of studies in both flies, nematodes, and subsequently mammalian models (Weinmaster et al., 1991) (de la Pompa et al., 1997) that contributed to the dissection of the pathway using both genetic and molecular techniques. The aim of the research in this thesis is to re-visit the model organism *C. elegans* as a system for studying Notch signalling, applying our current molecular understanding of the Notch receptor/ligand interaction to address questions surrounding the physiological importance of the N-terminal region of Notch ligands.

1.2 Physiological roles of Notch signalling

Notch signalling plays a key role in the development of multiple tissues across invertebrates and vertebrates alike. Notch signalling can be broadly divided into three modes of action: lateral inhibition, lineage decisions and boundary formation. There are many examples of cell fate decisions that are determined by each of these Notch signalling mechanisms from various species, however, only one example will be discussed for each mode below.

Lateral inhibition is one of the best characterised modes of Notch action, and occurs when signalling between cells of an initially equivalent population leads to the adoption of different and distinct cell fates within that population (Bray, 2006). One of the simplest examples of this is demonstrated by the AC/VU cell fate decision in *C. elegans*, in which Notch signalling between two initially equivalent cells leads to the subsequent specification of two distinct AC and VU cells (Figure 1.1A). A stochastic change in one of the cells (the pre-AC) triggers an increase in Notch signal-sending by this cell, and Notch signal-receiving in the other cell (the pre-

VU), and this is re-inforced by a positive Notch signalling feedback loop (Wilkinson et al., 1994). This example of lateral inhibition will be discussed further in Chapter 4.

Notch signalling is also required in certain lineage decisions, in which two non-equivalent daughter cells inherit Notch regulatory factors asymmetrically, leading to inductive signalling from one of the two cells to the other. A well-studied example of this is the segregation of the regulatory factor Numb to a single daughter cell after the division of the sensory organ precursor cell (SOP) PI during bristle formation in *Drosophila* (Hutterer and Knoblich, 2005). Numb subsequently down-regulates Notch signalling in the pIIb daughter cell, defining it as the signal-sending cell and the inductive Notch signal sent from this cell to the other daughter cell pIIa (in which Numb is absent) ensures that two distinct fates are adopted by the two cells (Figure 1.1B).

Signalling between two different cell populations can allow their distinct fates and the boundary between them to be maintained. *Drosophila* imaginal wing disc development provides an elegant example of boundary formation by Notch signalling (de Celis and Bray, 1997). Cells belonging to the dorsal (D) and ventral (V) compartments of the developing wing maintain a D/V boundary via Notch signalling between a single row of cells from each population (Figure 1.1C). Notch signalling is restricted to this boundary by expression of different Notch ligands and their regulatory factors. Expression of the ligand Serrate and the glycosyltransferase Fringe is restricted to the D compartment, whilst expression of the ligand Delta is restricted to V compartment. Fringe regulates Notch signalling in a ligand dependant manner, inhibiting Serrate-mediated Notch activation, whilst promoting Delta-mediated Notch activation (Panin et al., 1997). Therefore, Notch signalling is only potentiated at the boundary of these D and V cell populations, where Fringe-activated Notch

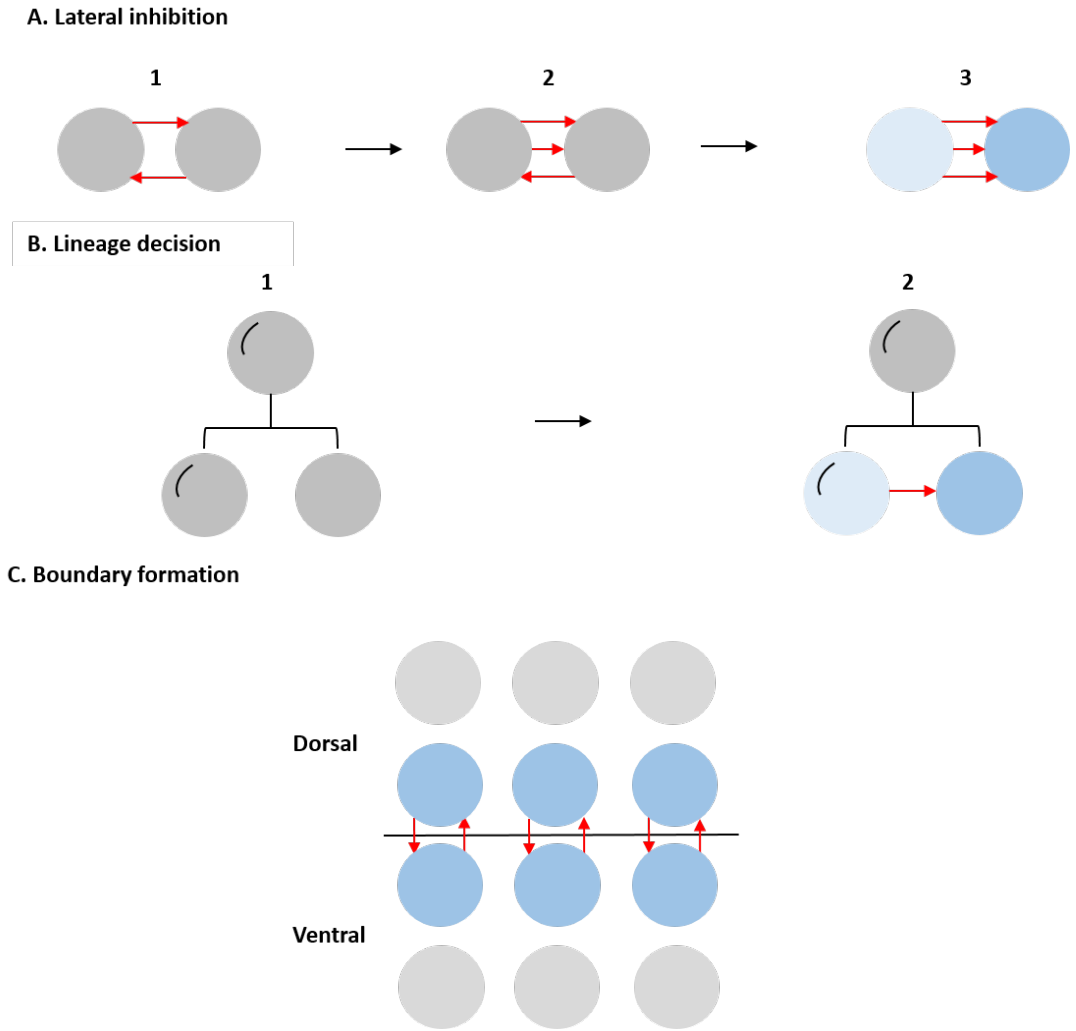


Figure 1.1: The three modes of Notch signalling action - (**A**) Lateral inhibition occurs between two initially equivalent cells (1) when a stochastic change leads to an increase in Notch signalling from one cell to the adjacent cell (2). This difference is amplified by a postive feedback loop of Notch signalling, ultimately resulting in the two cells adopting distinct fates (3), and a well documented example of this mechanism is the specification of the AC and VU cells in *C. elegans*. (**B**) Division of a cell such as PI in *Drosophila* can lead to asymmetric distribution of Notch regulators such as Numb (crescent) in daughter cells (1) The segregation of Numb to pIIb defines it as a signal sending cell, whilst its sister cell pIIa acts as the signal receiving cell (2). (**C**) The D/V boundary established in the *Drosophila* developing wing is maintained by Notch signals sent and received between cells lining the boundary, mediated by different ligands and the regulatory protein Fringe. Notch signals are represented as red arrows directed towards the signal receiving cell.

from D cells can interact with Delta from V cells, and non-Fringe-modified Notch from V cells can be activated by Serrate from D cells.

1.3 The Canonical Notch Pathway

1.3.1 *Trans*-activation and *Cis*-inhibition

Notch receptors and their ligands are single pass, type I transmembrane proteins (Figure 1.2). Newly translated Notch receptor polypeptide chains are trafficked to the surface of cells via the Golgi apparatus, and are cleaved at site S1 during this time by furin (Logeat et al., 1998), resulting in their presentation at the cell surface as a mature heterodimer. At the surface, Notch receptors can interact with their cognate membrane bound ligands, either on adjacent cells (*trans*-activation) or the same cell (*cis*-inhibition).

During *trans*-activation (Figure 1.3), binding of a Notch receptor via its extracellular EGF-like repeats to the Delta/Serrate/LAG-2 (DSL) domain of a ligand (Shimizu et al., 2000) causes a change in the conformation of the negative regulatory region (NRR) of the Notch receptor (Gordon et al., 2007). The mechanism of this conformational change has not been fully elucidated, however, studies of *Drosophila* and human Notch components have indicated that ligand endocytosis may contribute the physical force required for this process (Parks et al., 2000) (Stephenson and Avis, 2012) (Gordon et al., 2015). The consequence of this change is the exposure the S2 cleavage site, allowing the ADAM10/17 metalloproteases to cut the receptor, generating a Notch extracellular truncation (NEXT) fragment (Brou et al., 2000) (van Tetering et al., 2009). This fragment itself is a substrate for membrane-bound γ-secretase, whose activity processes the NEXT fragment by cutting within the transmembrane region (S3 cleavage site). This culminates in the

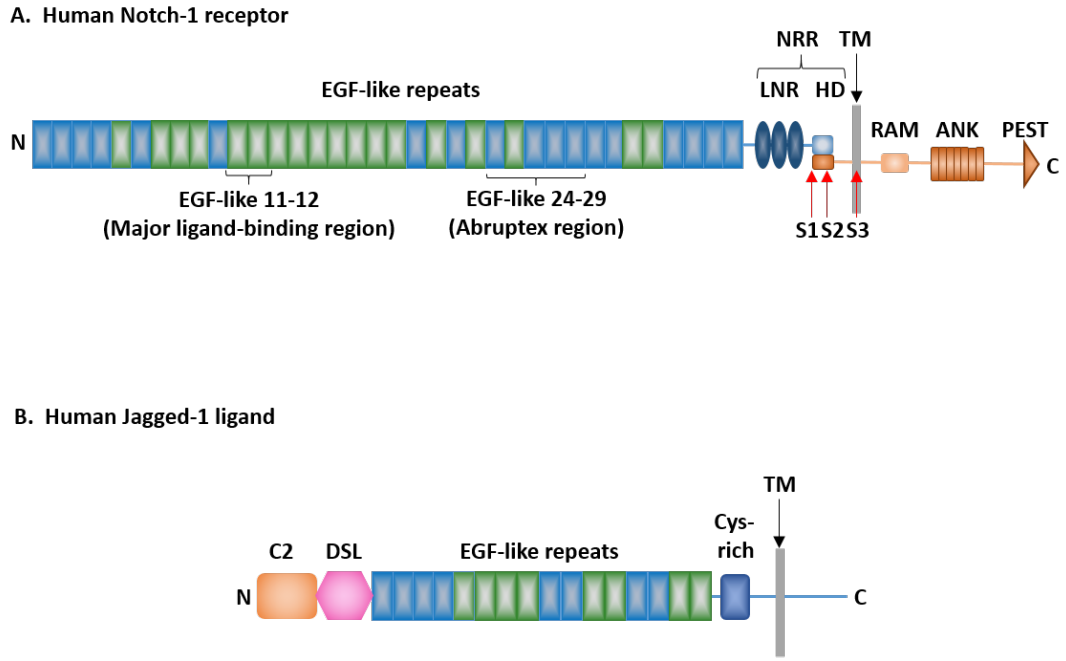


Figure 1.2: Domain organisation of Notch receptors and ligands - (**A**) Domain organisation of Notch receptors, using human Notch-1 as an example. The extracellular region comprises a series of EGF-like repeats (blue), a subset of which bind calcium (green). The negative regulatory region (NRR) includes three LIN-12/Notch repeats (LNR) followed by the heterodimerisation domain (HD) in which S1 cleavage by furin occurs. The HD also harbours the S2 cleavage site, which is processed during *trans*-activation of the receptor, whilst the S3 cleavage site occurs in the transmembrane region of the protein. The Notch intracellular domain region (NICD) (red) consists of a RBP$_{\text{j}}\kappa$ associated molecule (RAM) domain, followed by multiple ankyrin repeats (ANK) and a proline/glutamic acid/serine/threonine rich motif (PEST). (**B**) Domain organisation of Notch ligands using human Jagged-1 as an example. The module at the N-terminal of Notch ligands (MNNL) comprises a C2 domain (orange). This is followed by the Delta/Serrate/LAG-2 (DSL) domain (pink) and a region of EGF-like repeats. Ligands of the Jagged/Serrate class also contain a cysteine-rich (Cys-rich) motif prior to the transmembrane region.

release of a Notch intracellular domain (NICD) fragment (De Strooper et al., 1999) (Chandu et al., 2006) and its translocation to the nucleus (Struhl and Adachi, 1998). Here, the NICD interacts with the transcriptional regulator CBF1/Suppressor of Hairless/LAG-1 (CSL), switching it from a repressive to active state (Tamura et al., 1995). The NICD/CSL complex can then recruit co-activators including human Mastermind-like and p300 (Oswald et al., 2001) that promote the expression of target genes, largely basic-helix-loop-helix (bHLH) transcription factors such as those from the Hairy enhancer of split (HES) family (Jarriault et al., 1995).

By contrast, *cis*-inhibition occurs when Notch receptors and ligands that are expressed on the same cell interact. *Cis*-inhibitory interactions were identified in the model system of *Drosophila* wing development, whereby high levels of expression of ligand were seen to inhibit the expression of receptor within the same cell (de Celis and Bray, 1997). The precise mechanism underlying *cis*-inhibition is still not fully clear, however evidence suggests the ligand and receptor domains involved in *trans*-activation also play a role in *cis*-inhibition (Cordle et al., 2008a) (Becam et al., 2010). Additional regions, such as EGF-like 4-6 of Notch have also been implicated in the interaction (Fleming et al., 2013).

1.3.2 Components of the canonical pathway and their conservation between species

The general architecture of Notch receptors and their ligands are well conserved among metazoans (Figure 1.4). *Drosophila* possess a single Notch receptor (dNotch), whilst 4 homologues have been identified in mammals (Notch-1-4) and two in *C. elegans* (LIN-12 and GLP-1). Mammalian and *Drosophila* Notch contain 29-36 EGF-like repeats at the N-terminus of their extracellular region, whilst LIN-12 and GLP-1 have 13 and 10, respectively. This is followed by a negative regulatory

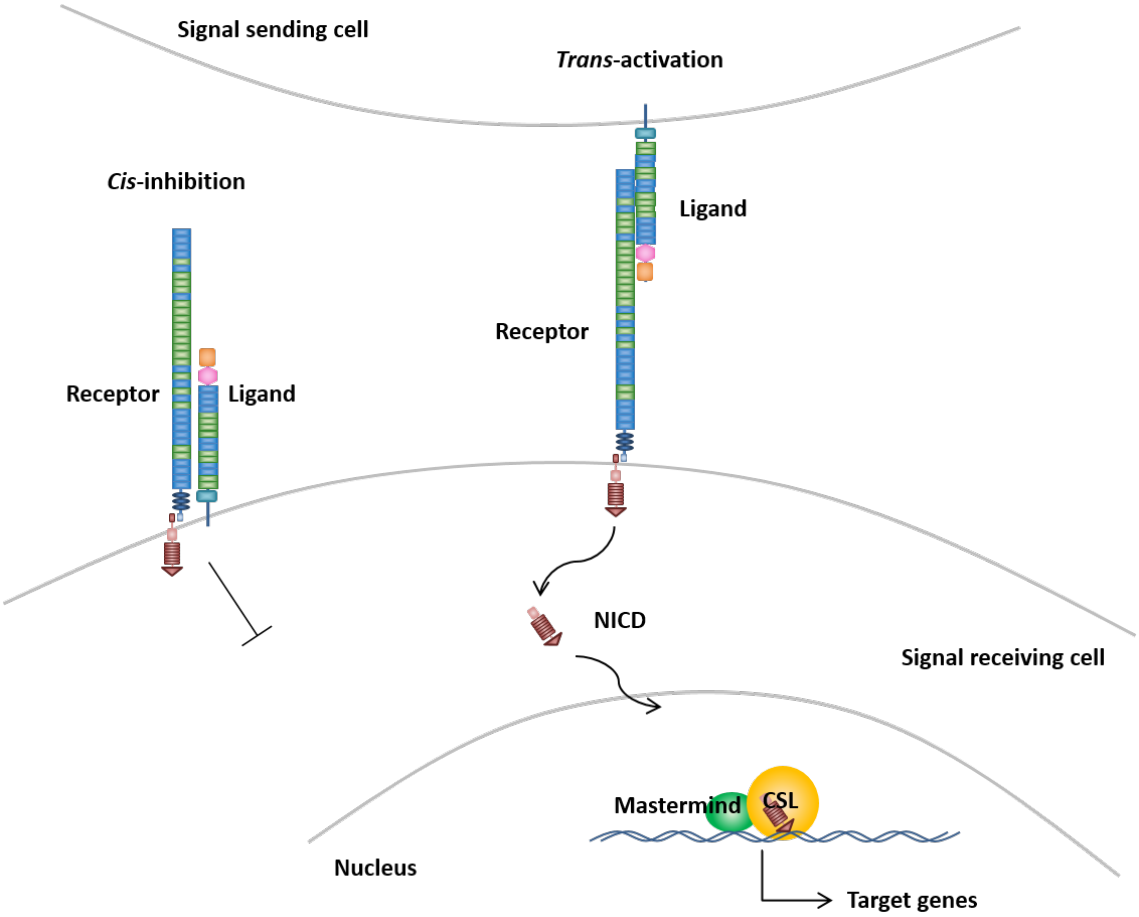


Figure 1.3: *Trans*-activation and *cis*-inhibition of the canonical Notch pathway - During transactivation, a ligand on the surface of a signal-sending cell binds to the Notch receptor on the surface of a signal-receiving cell, leading to two successive proteolytic cleavages of the Notch receptor to release the Notch intracellular domain (NICD). The NICD translocates to the nucleus and binds CBF1/Supressor of Hairless/LAG-1 (CSL), displaces repressors and recruiting co-activators such as Mastermind and its homologues in various species, to target genes of the Hairy Enhancer of Split (HES) family, resulting in their expression. Interaction of Notch receptors and ligands expressed on the surface of the same cell does not result in activation of the pathway, hence is termed *cis*-inhibition.

region which harbours the S2 cleavage site, a transmembrane region in which S3 cleavage occurs, and an intracellular domain whose components are well conserved between species. All canonical ligands are defined by the presence of their DSL domain, the minimal unit required for Notch receptor binding (Shimizu et al., 2000). Canonical ligands also contain varying number of EGF-like repeats and regions related to these, and in addition can be divided into two classes due to the presence (Jagged/Serrate class) or absence (Delta/Delta-like class) of a cysteine-rich (Cys-rich) region near the transmembrane domain. Whilst there are 5 human and 2 *Drosophila* ligands, *C. elegans* are predicted to have 10 ligands, a subset of which are soluble (Chen and Greenwald, 2004). An additional and important region of canonical ligands is the module at the N-terminus of Notch ligands (MNNL), which has been confirmed in structural studies to be a C2 domain in mammalian ligands of both the Jagged/Serrate and Delta/Delta-like classes, as well as *Drosophila* Serrate (Chillakuri et al., 2013) (Luca et al., 2015) (Kershaw et al., 2015) (R. Suckling, personal communication), and appears to be necessary for productive signalling.

1.4 Notch signalling and disease

The importance of the Notch system in both human development and adult tissue homeostasis is illustrated by several rare genetic disorders associated with aberrant signalling, which are generally caused by intragenic mutations or deletions within the receptor and ligand components of the canonical Notch pathway. Mutations in the *NOTCH1* locus associated with loss-of-function have been identified in aortic valve disease patients (McKellar et al., 2007), whilst mutations causing constitutive activation of Notch-1 receptor lead to T-cell acute lymphoblastic leukaemia (T-ALL) (Ellisen et al., 1991) (discussed further in section 1.4.3). Disease states brought about by either too much or too little activation of the pathway highlight

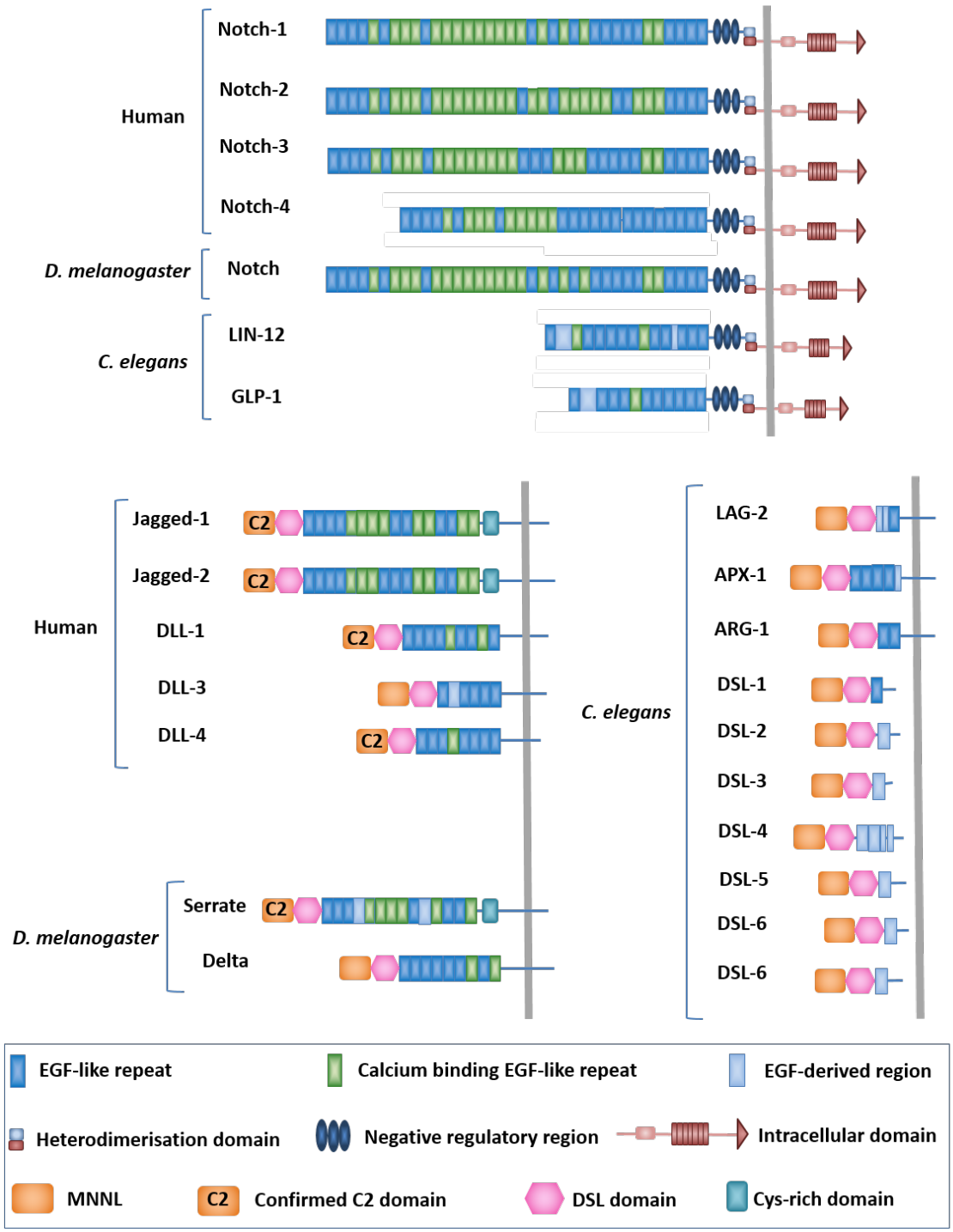


Figure 1.4: The architecture of Notch receptors and their ligands - Above, Notch receptors are single pass transmembrane proteins containing multiple extracellular EGF-like repeats, which are also present in Notch ligands (below), along with a DSL and MNNL domain. *C. elegans* also have soluble DSL ligands, not present in humans or *Drosophila*.

the importance of Notch signalling and the need for its strict regulation.

Other diseases associated with mutation in the Notch receptor genes include Hadju-Cheney disease, a severe osteoporosis disorder in which mutations in *NOTCH2* prevent degradation of the Notch-2 receptor, and leading to prolonged activation of the signalling pathway (Simpson et al., 2011). Mutations in the *NOTCH3* locus are associated with the stroke disorder CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy), and appear to act via a more complex mechanism than directly increasing or decreasing levels of Notch signalling (Joutel et al., 1996).

Mutations in the loci of Notch ligand genes have also been implicated in multiple diseases. Over 70% of mutations identified in patients with the severe skeletal disorder spondylocostal dysostosis occur in the *DLL-3* gene, and mutations at this locus are generally predicted to result in truncations or disruption of the domain structure of the ligand (Bulman et al., 2000) (Turnpenny et al., 2003). Jagged-1 mutations believed to cause loss of function have been identified in patients with isolated heart defects such as Tetralogy of Fallot (Eldadah et al., 2001), as well as other defects as part of the multisystemic disorder Alagille syndrome (Li et al., 1997). More recently, *DLL-4* mutations have been identified in a subset of patients suffering from the disease Adams-Oliver syndrome (affecting scalp, limb and often cardiac development), which has previously been associated with *NOTCH1* mutations (Meester et al., 2015). Below, the pathogenic mechanisms of Alagille syndrome and CADASIL are discussed in further detail as examples of developmental and adult-onset Notch signalling disorders, respectively.

1.4.1 Alagille syndrome

Mutations in the human *JAGGED1* gene have been identified in patients with both isolated and multi-systemic developmental defects. These include chronic structural heart defects (Krantz et al., 1999) (Bauer et al., 2010) as well as extra hepatic biliary atresia, brought about by malformation of the liver bile ducts (Kohsaka et al., 2002). In addition, these two features can contribute to Alagille syndrome (AGS) (Alagille et al., 1975) in which skeletal abnormalities, ocular defects and characteristic facies have also been identified (Table 1.1).

Table 1.1: Summary of the features associated with AGS -Patients with symptoms in three or more of the above categories are diagnosed with AGS.

Feature	Characteristics
Biliary atresia	Narrow malformed bile ducts, causing build up of bile in the liver. Associated with jaundice due to cholesterol deposits under the skin
Cardiac abnormalities	Failure of heart chambers to form correctly. Can give rise to pulmonic stenosis, ventricular septal defect, Tetralogy of fallot
Ocular defects	Such as posterior embryotoxon, characterised by an enlarged anterior limiting ring
Skeletal abnormalities	Inability for vertebrae to develop correctly, resulting in 'butterfly vertebrae', identifiable by X-ray
Characteristic facies	Broad prominent forehead with deepset eyes and small pointed chin

The underlying pathogenic mechanism of this autosomal dominant syndrome appears to be haploinsufficiency of the ligand Jagged-1. Approximately 70% of intragenic mutations (insertions, deletions, splice site mutations) were predicted to introduce a premature stop codon into the mRNA transcript encoded by *JAGGED1*

(Crosnier* et al., 1999), which are generally subjected to nonsense mediated decay (NMD) (Frischmeyer and Dietz, 1999). Molecular studies indicated that several of the point mutations result in amino acid substitutions likely to cause misfolding subsequent retention of Jagged-1 (Morrissette et al., 2001) (Lu et al., 2003). AGS mutations have been identified in intragenic regions corresponding to most domains of Jagged-1, although a particularly high frequency of substitutions are clustered in the C2 domain at the N-terminus of the ligand. Several AGS mutations have been mapped to the hydrophobic core of this domain, and it was demonstrated that several of these mutations lead to the retention of protein within the cell, most likely due to misfolding (Chillakuri et al., 2013).

A small number of AGS-causing mutations were identified in the *NOTCH2* gene (McDaniell et al., 2006) (Kamath et al., 2011), implying that the Notch-2 receptor is a major interaction partner of Jagged-1 in the tissues implicated in AGS. This discovery reconciled the observations from attempts to generate an AGS mouse model. It was previously shown that mice that were heterozygous for a null *JAGGED1* allele displayed only ocular defects (Xue et al., 1999). However, mice doubly heterozygous for the null *JAGGED1* allele and a null *NOTCH2* allele displayed tissue-specific defects consistent with those observed in human AGS (McCright et al., 2002), suggesting that *NOTCH2* acts as a genetic modifier of *JAGGED1*. This suggested that the integrity of multiple Notch pathway components may be important when influencing the phenotypic outcome associated with disease.

1.4.2 CADASIL

CADASIL is a heritable adult-onset disorder in which mutations in *NOTCH3* cause the build up of granular osmiophilic matter in the extracellular matrix of vascular smooth muscle cells, destroying the smooth muscle tissue. This leads to death of

white matter in the brain, resulting in ischemic stroke and dementia as a long term consequence (Joutel et al., 1996).

The majority of mutations identified in CADASIL patients result in an unpaired cysteine residue occurring in the extracellular region of the Notch-3 receptor, generally EGF-like repeats 2-5. Similar mutations identified in the EGF domains of other proteins, such as Fibrillin-1, have been shown to have more than one effect. Cysteine substitutions were shown to cause misfolding of the protein, which can lead to both the retention of the protein within the cell, or the secretion of defective protein which can produce a dominant-negative effect (Schrijver et al., 1999). The pathogenic mechanism of CADASIL is dominant negative, involving the deposition of extracellular fragments of the mutant Notch-3 receptors into the matrix, where they aggregate with additional matrix proteins to contributing to the deposition of granular matter in the vasculature (Joutel and Tournier-Lasserve, 2002).

Several CADASIL mutations occurring in EGF-like 2-5 of Notch-3 have been shown to bind ligand to a similar extent to wild type Notch-3 (Joutel et al., 2004), supporting the proposal that it is the deposition of mutant Notch-3 fragments into the matrix rather than a change in the activation of the canonical Notch pathway which leads to the disease state. Interestingly, extremely rare mutations in the major ligand-binding region of Notch-3 (EGF-like 10-11) are associated with a loss of activation of the signalling pathway. A study of CADASIL phenotypes in patients with EGF-like 10-11 mutations led to the proposal that the associated reduction of Notch-3 signalling led to a delayed onset of the more severe phenotypes of CADASIL (Monet-Leprêtre et al., 2009), highlighting the complex nature of this disease.

1.4.3 Cancer

Aberrant Notch signalling has been implicated in multiple forms of cancer, and the study of chromosomal rearrangements associated with T-ALL led to the initial identification of the *NOTCH1* gene in humans, (originally referred to as *TAN1*) (Ellisen et al., 1991). Subsequently, it was discovered that of the mutations associated with T-ALL, between 50-60% occurred in *NOTCH1*, with the majority of these localised to the region encoding the heterodimerisation domain (HD) of Notch-1 (Weng et al., 2004), and a subset of mutations that result in substitutions in the PEST domain of the receptor.

Several mutations associated with the HD region of *NOTCH1* result in single amino acid substitutions that alter the conformation the surrounding NRR region of the Notch-1 receptor, allowing S2 site cleavage even in the absence of ligand binding. This results in a greater number of NEXT fragments that can be subsequently recognised by γ-secretase, leading to increased release of NICD and a hyper activation phenotype. On the other hand, substitutions in the PEST domain were shown to act by impairing targeting of the NICD fragment for degradation (Bhanushali et al., 2010). This was supported by the consequences of T-ALL mutations localised to the gene encoding the E3 ubiquitin ligase FBXW7, which is required to target NICD for degradation (O'Neil et al., 2007). Disruption of the interaction between the two proteins results in a decreased turnover of NICD, propagating the Notch signal and resulting in over-activation. In cases where mutations affecting both HD and PEST regions in *NOTCH1* were present, they appeared to act synergystically to result in greater overall levels of Notch-1 activation (Weng et al., 2004).

Interestingly, *NOTCH1* and *NOTCH2* mutations associated with loss-of-function of Notch receptors have also been implicated in cancer. Analysis of patients with cutaneous squamous cell carcinomas (cSCCs) revealed a spectrum of mutations, rang-

ing from frameshifts likely to result in NMD, to substitutions spanning the length of the receptor, excepting the NRR and PEST regions (Wang, 2011). Dysregulated Notch signalling has also been linked to additional solid tumours, including but not limited to subtypes of breast, lung and colorectal cancer (Ntziachristos et al., 2014). Therefore, the Notch pathway is a high priority target for the generation of therapeutics. One therapeutic approach that was explored was the use of γ-secretase inhibitors (GSIs), to alleviate the over-activation of the pathway. However, a key limitation in their therapeutic use is their potential toxicity, demonstrated in mice (Wong et al., 2004) (van Es et al., 2005). Antibody-based therapies offered some improvement to specificity of targeting, and the generation of antibodies capable of discriminating between the Notch receptor paralogues, whilst recognising both human and rodent receptors demonstrated that this could provide a promising and powerful approach (Wu et al., 2010).

The search for suitable therapeutic reagents and the challenges presented by the need for highly specific delivery and targeting of these reagents highlights the importance of continued understanding and study of the core pathway of Notch itself. The greater our knowledge of the structure of the components, their regulation and the nuances of the system as a whole, the more likely it is that we will be able to find ways of successfully intercepting aberrant Notch signalling to ameliorate its detrimental effects.

1.5 Structural and molecular studies of Notch receptors

1.5.1 The EGF-like region

Structural and molecular studies have provided key insights into the role of the EGF-like repeats of Notch receptors, the sequences of which are generally closely related to consensus EGF domain sequence (Gray et al., 1983). Artavanis-Tsakonas and colleagues initially demonstrated that the extracellular region of *Drosophila* Notch (dNotch) mediated the receptor/ligand interaction in a calcium-dependent manner (Fehon et al., 1990) and subsequent deletion analysis and cell aggregation assays allowed the major ligand-binding region of dNotch to be localised to EGF-like 11-12 (Rebay et al., 1991). The conservation of this site between species was also suggested, as substitution of *Xenopus* Notch (Xotch) EGF-like 10-13 into a dNotch construct was still capable of promoting cell aggregation with the *Drosophila* ligands Delta and Serrate.

Alignment of the primary sequences of dNotch and human Notch-1 (hNotch-1) indicated a sequence identity of over 50% in the EGF-like region of the two proteins (Ellisen et al., 1991). Therefore, the suspected major ligand-binding region of hNotch-1 could be targeted for structural and molecular analysis (Hambleton et al., 2004). A fragment of human Notch-1 comprising EGF-like 11-13 ($N1_{11\text{-}13}$) was bacterially expressed, refolded, and its functionality confirmed by its ability to bind ligand-expressing cells in a calcium-dependent manner, using FACS. The NMR structure of this fragment was solved, revealing EGF-like 11-12 to adopt a fold consistent with a class II cbEGF-cbEGF domain pair (Downing et al., 1996). It was shown that each EGF-like domain bound a single calcium ion under saturating conditions, and complementary X-ray crystallography studies (Cordle et al., 2008a)

were consistent with these observations (Figure 1.5). It was further speculated that based on the structure of this specific pair, and the presence of multiple calcium binding EGF-like domains throughout the ectodomain of hNotch-1, that a linear conformation of the receptor at the cell surface was likely to be adopted (Hambleton et al., 2004). The equivalent ligand-binding region in *C. elegans* Notch ligands LIN-12 and GLP-1 has not been identified on the basis of sequence similarity, and is likely to differ to some extent in its structure, as it is predicted that there are no adjacent calcium-binding EGF-like repeats in either receptor. However, since LIN-12 and GLP-1 share approximately 50% amino acid identity, and can substitute for each other in cell fate decisions (Fitzgerald et al., 1993), it is likely they share a common mechanism of action.

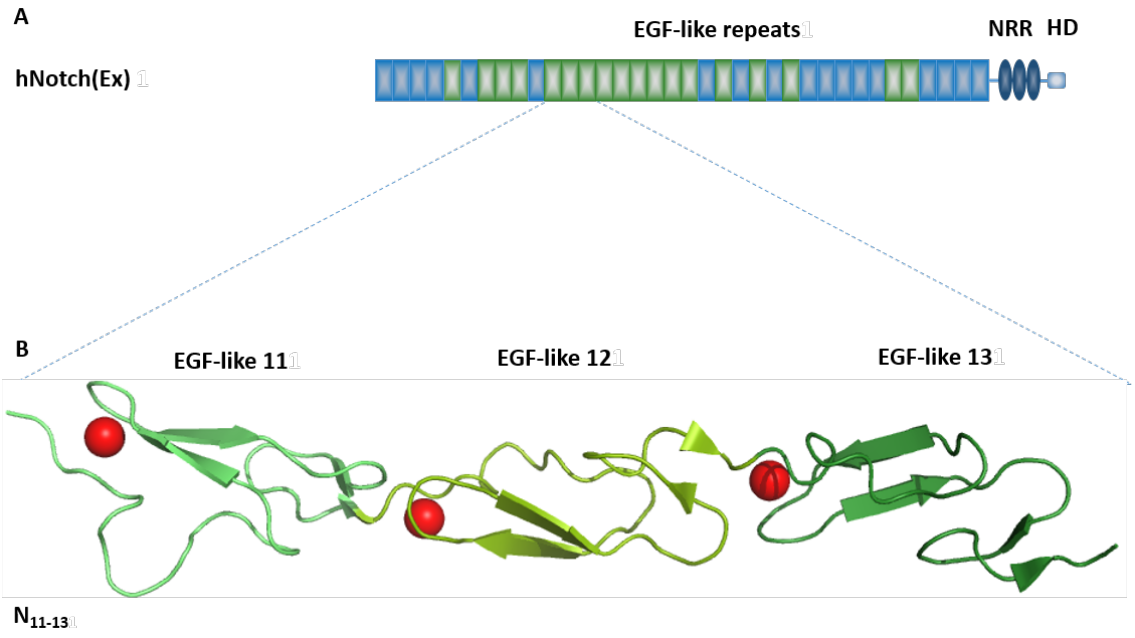


Figure 1.5: Structure of the Notch receptor major ligand-binding region - (**A**) the architecture of the extracellular region of human Notch 1 (hNotch1(Ex)), including 36 EGF repeats, followed by the NRR. (**B**) X-ray structure of the $N_{11\text{-}13}$ fragment at 2.5$\AA$ in the presence of calcium (red), as reported by Cordle et al. (2008a). PDB accession code 2VJ3.

In addition to structural information about the hNotch-1 ligand binding region,

molecular studies involving mutagenesis of calcium-binding sites and limited proteolysis suggested that the ligand binding is centred on EGF-like 12 (Cordle et al., 2008b). Insights into the specific residues of this domain involved have been afforded by studies of two substitutions in Notch-1 (L468A, I477A) and their counterparts in dNotch which result in abrogation of binding to Jagged-1 and Serrate, respectively (Whiteman et al., 2013). With our understanding of the mechanism of Notch receptor binding becoming increasingly detailed, the potential to target this site with highly specific antibodies is being further explored (Whiteman et al, personal communication).

An additional region of interest that was identified in *Drosophila* Notch is the Abruptex region, which spans EGF-like 24-29 (the equivalent region in human Notch-1 is highlighted in Figure 1.2). Dominant mutations occurring at the *dNOTCH* locus encoding this site are associated with distinct phenotypes, including a change in bristle pattern and a lack of wing veins in adult flies (Foster, 1975). The phenotypes associated with Abruptex mutants are generally the opposite of those associated with deletions in the *dNOTCH* locus, and it appears that the consequence of Abruptex mutations is an increased activation of dNotch by Delta, due to reduction in inhibition of the receptor by other components of the pathway (de Celis and Bray, 2000). It has been suggested that the Abruptex region of both dNotch and the equivalent region in hNotch-1 may interact with or compete with EGF-like 11-12 to serve as a second ligand-binding site (Pei and Baker, 2008) (Sharma et al., 2013). Whether these effects are functionally relevant *in vivo* will require further study. Our understanding of the role of the Abruptex region is not sufficient to exclude the possibility that a related region is present in LIN-12/GLP-1, although to date it has not been defined.

1.5.2 Glycosylation of EGF-like repeats of Notch receptor

EGF-like domains can be subjected to post-translational modifications including N- and O-linked glycosylation and also β-hydroxylation. The most functionally significant modification involved in the regulation of Notch receptor activity is O-glycosylation. Many of the EGF-like repeats present in Notch receptors contain consensus sequences for glycosylation, although not all of these sites are necessarily glycosylated to the same extent *in vivo* (Takeuchi and Haltiwanger, 2014). The two main forms of glycosylation that play an important role in the functionality of Notch receptors are O-linked glucosylation and O-linked fucosylation, which were simultaneously identified (Moloney et al., 2000), although our understanding of O-linked fucosylation is more advanced than O-linked glucosylation.

O-glucosylation links a glucose molecule to the serine of a consensus motif $C_1X\underline{S}X(P/A)C_2$ present in EGF-like domains (Rana et al., 2011) (Figure 1.6A). The *RUMI* gene was identified as a component of the Notch pathway in *Drosophila* when temperature sensitive mutations at this locus were found to abolish Notch signalling (Acar et al., 2008). dRumi is a glycosyltransferase that is capable of catalysing the O-linkage of glucose and the subsequent addition of xylose (Takeuchi et al., 2011). Mammalian Poglut1, the closest homologue for dRumi, has been shown to be necessary for embryonic development in mice (Fernandez-Valdivia et al., 2011), further indicating that O-glucosylation plays a key role in Notch signalling. Mass spectral analysis of mouse Notch-1 indicated that all O-glycosylation sites conserved among species in the region spanning EGF-like 10-20 were glycosylated *in vivo*, however, mutation of single glycosylation sites within this region did not reduce the receptor's activity. Interestingly, mutation of the non-conserved O-glycosylation site in EGF-like 28 of mouse Notch-1 was associated with decreased DLL-1 binding in a cell-binding assay (Rana et al., 2011), indicating the need for further studies to in-

vestigate the mechanism of action of O-glyosylation on Notch signalling regulation.

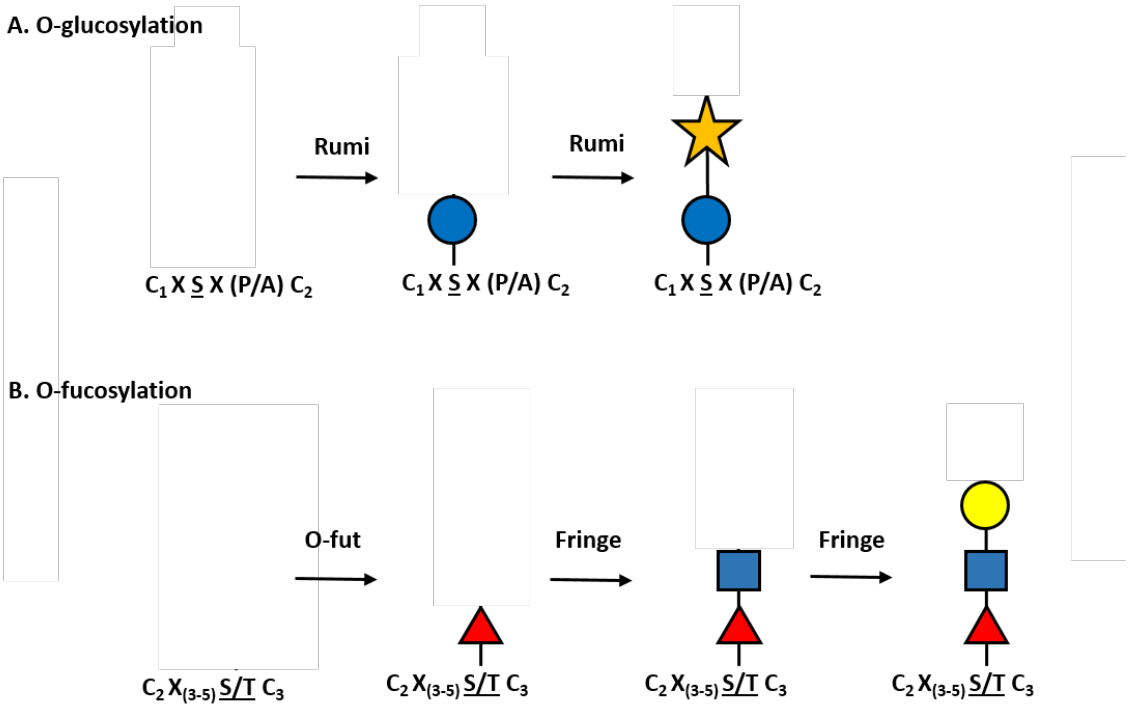


Figure 1.6: O-linked glycosylation modifications at amino acid consensus sequences in *Drosophila* - **A**) Rumi catalyses the O-linked addition of a single glucose (blue circle), which can be further linked to Xylose (yellow star) at the consensus sequence shown. **B**) O-fut catalyses the addition of a single fucose (red triangle) which can be further extended to a di-, tri-, or tetra-saccharide with GlcNAc (blue square) and galactose (yellow circle) respectively. In mammals, this can be further extended to a tetrasaccharide with the addition of sialic acid by sialyltransferases. Consensus sequences show sites of glycosylation within EGF-like repeats in relation to conserved cysteines (C_1, C_2, C_3). X= other amino acid.

O-fucosylation is the addition of a single fucose molecule to a conserved serine/threonine of the consensus sequence in EGF-like domains (Harris and Spellman, 1993) (Figure 1.6B). This reaction is catalysed by the enzyme Ofut-1, which acts as an O-fucosyltransferase (Wang et al., 2001) and a potential chaperone in the endoplasmic reticulum (Okajima et al., 2005). Knockout of the mouse homologue of Ofut-1 is associated with an embryonic lethal phenotype, indicating the importance of this enzyme. O-fucose modified EGF-like domains can act as a substrate for Fringe, an N-acetylglucosamine (GlcNAc)transferase (Moloney et al., 2000). Mam-

mals have three Fringe homologues (Manic, Lunatic and Radical (Johnston et al., 1997)), which appear to perform the same enzymatic function, although their substrate specificities appear to vary, with Lunatic fringe generally considered more active than Manic or Radical (Rampal et al., 2005). A Fringe homologue has yet to be identified in *C. elegans*.

The activity exerted by Fringe on Notch receptors has been a topic of particular interest due to the apparently opposing effects it may have on ligand mediated signalling. Studies of *Drosophila* wing development revealed that the ligands Delta and Serrate positively regulate each others expression at the dorsal/ventral (D/V) boundary at the third instar imaginal discs, and that Fringe participates in this process by positively modulating the Delta-mediated signals, whilst inhibiting Serrate-mediated signals (section 1.2), suggesting that the mechanism of Fringe is not trivial (Panin et al., 1997).

A more recent study that has given some insights to the elusive mechanism of Fringe activity examined the effects of Fringe modifications at a conserved threonine residue (T466) in EGF-like 12 of human Notch-1 and dNotch. The Notch-1 fragment $N1_{11\text{-}13}$ was modified *in vitro* with O-fucose monosaccharide, and subsequently extended with to a di-, tri-, or tetra-saccharide consistent with the activity of Fringe-mediated modification (Taylor et al., 2014). The effects of these modifications on the ligand-binding properties of $N1_{11\text{-}13}$ were assessed using flow cytometry and surface plasmon resonance (SPR), which both demonstrated enhanced binding to Jagged-1 and DLL-1 by the disaccharide modified fragment, compared to the monosaccharide modified fragment. X-ray structures reported in the same study indicate that modifications at T466 do not alter the backbone structure of the EGF-like domain. Taken as a whole, the data suggest that Fringe exerts its modulatory activity by altering the Notch receptor/ligand interaction. Since there are other EGF-like repeats aside

from the ligand-binding region at which Fringe could potentially act (including the Abruptex region), this may explain the differential outcomes of Fringe activity on Notch signalling.

1.5.3 The negative regulatory region (NRR)

The NRR is the site of S2 cleavage, and is essential for ligand-dependent activation of Notch receptors. Notch constructs lacking the extracellular domain, including the NRR have been shown to be active in the absence of ligand (Rebay et al., 1993) (Coffman et al., 1993), whilst truncations of the EGF-like repeats only (leaving an intact NRR) are not (Kopan et al., 1996). The NRR comprises three LIN-12/Notch repeats (LNRs), which are cysteine-rich modules that appear to be exclusively present in Notch receptors. These are followed by the heterodimerisation domain. An NMR structure of LNR-A from human Notch-1 indicated that these modules have irregular secondary structure, however calcium co-ordination allowed a well-defined tertiary structure to be formed (Vardar et al., 2003). The heterodimeristion domain is the region at which N- and C- terminal fragments of the receptor generated during the first receptor cleavage event (prior to receptor presentation at the cell surface) interact in a non-covalent manner (Blaumueller et al., 1997) (Sanchez-Irizarry et al., 2004). This cleavage event has not been fully confirmed in *C.elegans*, although it seems highly likely that this aspect of Notch receptors is conserved (Crittenden et al., 1994).

The importance of the NRR has been highlighted by the consequences of mutations in this region (section 1.3.3). Additional evidence for the importance of the NRR maintaining the integrity of ligand-dependent Notch signalling includes the phenotype associated with a *C. elegans lin-12* alleles containing mutations which map to the C-terminal portion of the heterodimerisation domain. Consequently,

these encode dominantly activated forms of LIN-12 that signal in the absence of ligand (Greenwald and Seydoux, 1990). In addition, the loss of a conserved serine residue between the third LNR repeat and transmembrane domain of GLP-1 results in constitutive activation of the receptor in the absence of its ligand, causing excessive proliferation of germ cells (Berry et al., 1997).

Before structural information was available for the NRR, popular models of receptor activation were based on mechanotransduction caused by binding of the ligand to the receptor (Parks et al., 2000). How this pull allowed for S2 cleavage, and how this was prevented in the absence of ligand was unknown. Insights into the mechanism came from a crystal structure of the NRR of human Notch-2 (excluding a poorly conserved and unstructured loop region) (Gordon et al., 2007) (Figure 1.7), and subsequently the NRR hNotch-1 (Gordon et al., 2009). These indicated that the S2 site lay in the C-terminal portion of the heterodimerisation domain, and that the LNR modules wrap around the site in the absence of ligand to protect the site from cleavage. These features led to the proposal of a lift and cut model of S2 cleavage (Gordon et al., 2007), whereby receptor binding leads to a conformational change within the NRR, resulting in the movement of a stabilizing helix in the C-terminal HD region. This in turn, exposes the S2 region sufficiently for recognition and cleavage by ADAM10/17. Whilst this is an attractive model, questions still remain as to how the binding of ligands to EGF-like 11-12, an architecturally distant region of the Notch receptor is transduced to this site.

1.5.4 The Notch intracellular domain (NICD)

The NICD that is released after S3 cleavage at the transmembrane region is smaller than its extracellular counterpart but critical for signalling. The conserved components of the NICD from different species include an RBPjk association molecule

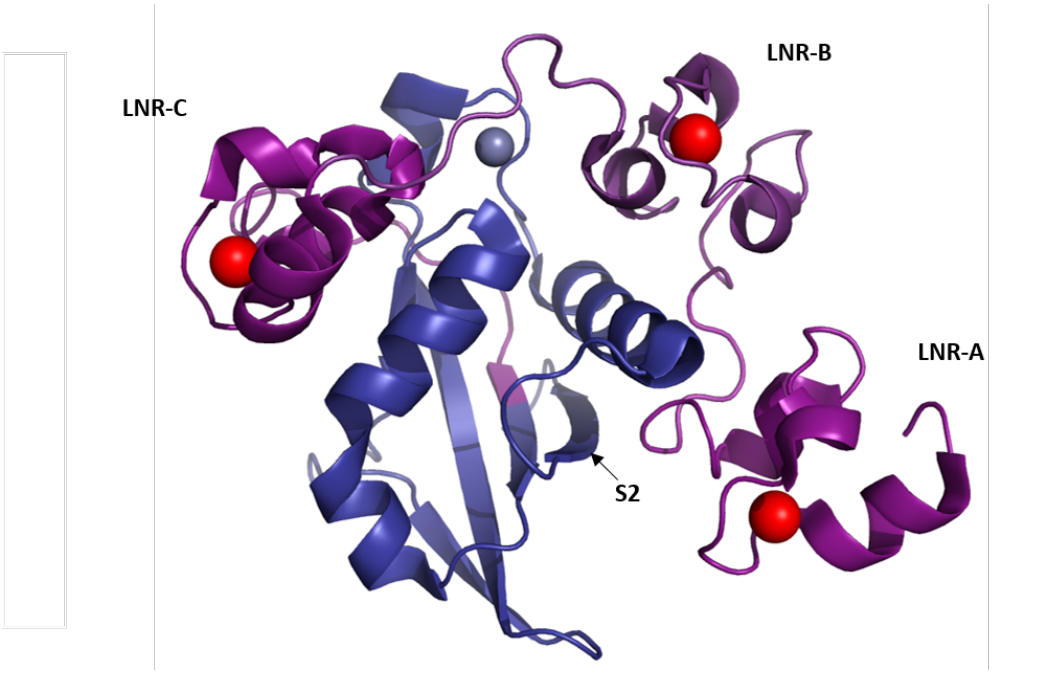


Figure 1.7: Structure of the negative regulatory region of Notch - The X-ray structure of human Notch-2 NRR revealed a protected S2 site, surrounded my LNR modules (purple) in the absence of ligand, that occludes metalloprotease enzymes from the cleavage site. Calcium ions shown in red, zinc in grey. PDB accession code 2OO4.

(RAM) domain followed by a linker region containing nuclear localisation signals, the ANK domain which is made up of multiple ankyrin (ANK) repeats, and a C-terminal proline/glutamic acid/serine/threonine-rich motif (PEST) domain.

Initial experiments conducted in *Drosophila* suggested an interaction between the transcriptional regulator CSL (Su(H) in *Drosophila*) and the dNICD were mediated by the ANK repeats (Fortini and Artavanis-Tsakonas, 1994). However this was subsequently disputed, and instead the RAM domain was implicated in the interaction (Tamura et al., 1995).

Additional studies of the interaction between *C. elegans* LAG-1 and the NICD of GLP-1 led to the proposal that direct binding to LAG-1 occurs via the RAM domain, and that the ANK repeats may be indirectly involved in binding (Roehl et al., 1996). It was subsequently discovered that the human homologue of *Drosophila* Master-

mind, MAML-1, binds to the NICD via the ANK domains (Wu et al., 2000), and that NICD binding to RBPjk appears to allow the co-operative recruitment of transcriptional co-activators such as MAML-1 to form an active transcriptional factor complex (Vardar et al., 2003). Insights into this model of assembly were provided by detailed structural studies of a co-crystallised complex of *C. elegans* Notch components CSL (LAG-1), NICD (NICD of LIN-12), and Mastermind (LAG-3) in the presence of DNA (Wilson and Kovall, 2006) (Figure 1.8). These studies indicated that the NICD and CSL initially interact via RAM domain binding to the β-trefoil domain (BTD) of CSL, leading to a conformational change that is likely to result in the initial displacement of co-repressors. The NICD/CSL complex then acts as a platform to which Mastermind can bind, via the N- and C-terminal domains of CSL (NTD and CTD, respectively), as well as the ANK domain of the NICD, an observation supported by structural studies of the ANK repeats of human Notch-1 in complex with CSL and Mastermind (MAML-1) (Nam et al., 2006). It has been proposed that the ternary complexes of the NICD/CSL/MAML can form functionally active dimers (Figure 1.9), which may be facilitated by additional unknown transcription factors (Nam et al., 2007) (Arnett et al., 2010).

The NICD also includes a C-terminal PEST domain, which contains WD-40 motifs, common in many proteins that require rapid turnover. These motifs are recognised by E3 ubiquitin ligases, which transfer ubiquitin to their substrate, and consequently targeting it for degradation by the proteasome (Ciechanover 1994). The interaction of the *C. elegans* protein SEL-10 with Notch receptors and its ability to regulate Notch activity indicated the importance of ubiquitination in the regulation of Notch signalling (Hubbard et al., 1997). Studies of the mammalian Sel-10 homologue (referred to as FBXW7 in humans) demonstrated that mSel-10 physically interacts with the NICD, and as a consequence, ubiquinates the NICD in

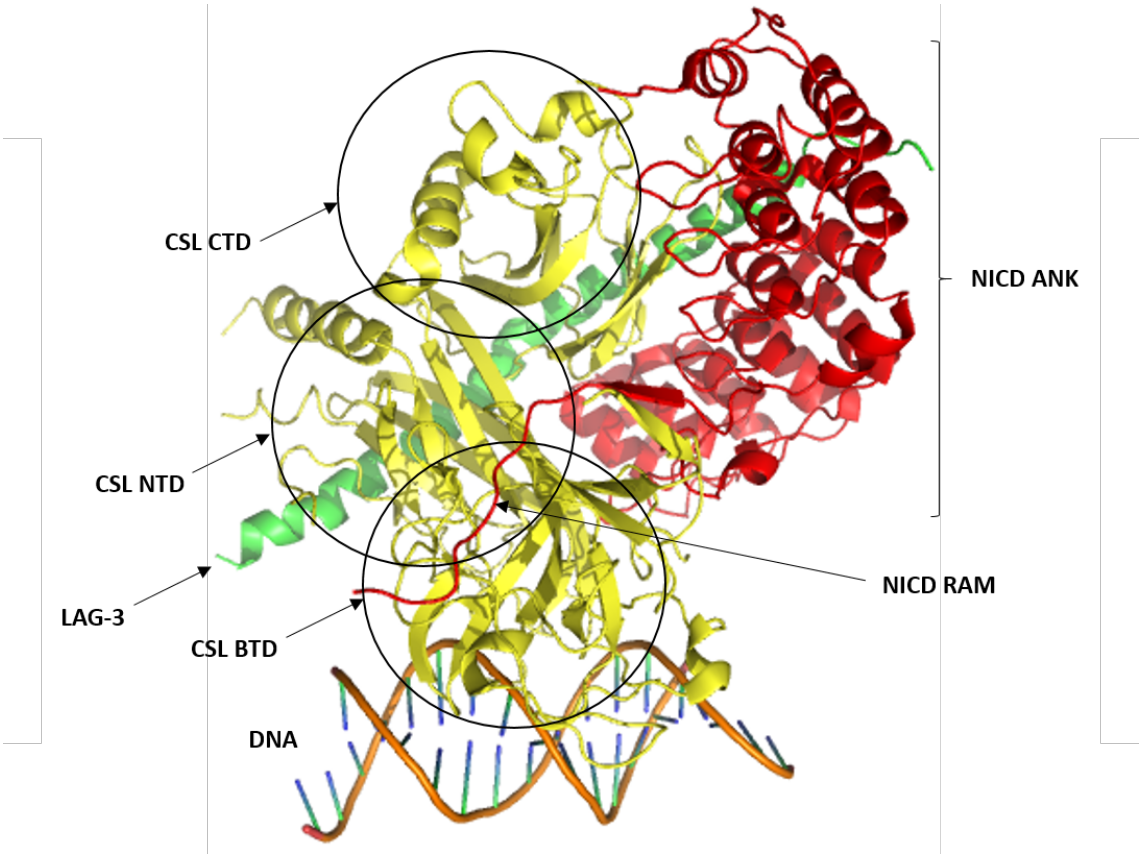


Figure 1.8: Structure of *C. elegans* NICD in complex with CSL and LAG-3 and DNA - ANK and RAM domains of NICD are shown in red, with the RAM interacting with the BTD domain of CSL (yellow). The Mastermind homologue LAG-3 (green) interacts with CSL via its NTD and CTD. PDB accession code F2O1.

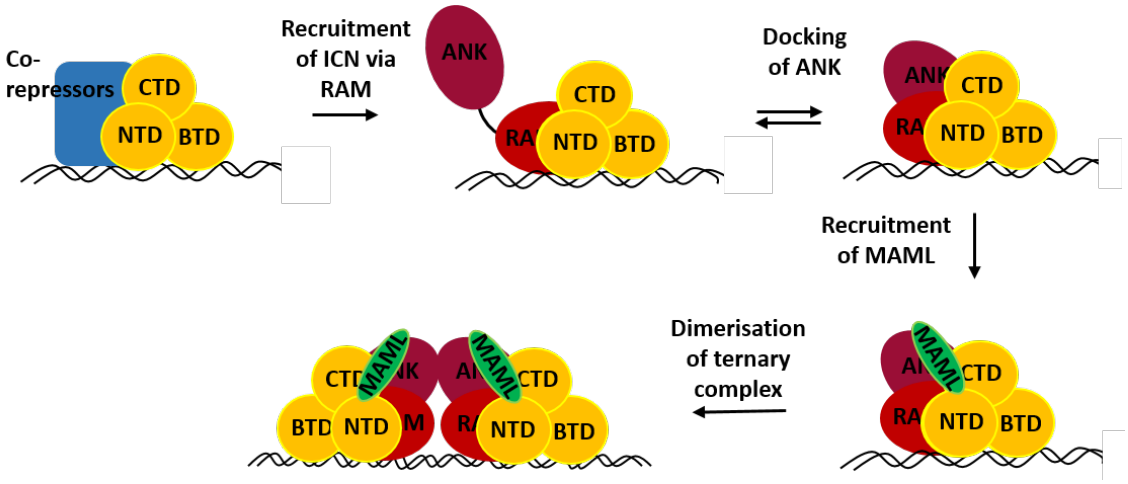


Figure 1.9: Model of ternary complex assembly - The binding of the NICD via the RAM domain to the BTD of CSL allows the ANK repeats to subsequently dock to the NTD and CTD of CSL. This complex then recruits MAML via contacts from both the NICD and CSL. An active complex may then be subsequently formed by dimerization of the ternary complexes. Adapted from Nam et al 2007.

its C-terminal region (Oberg et al., 2001). As previously described in section 1.3, mutations associated with T-ALL have been identified in both the PEST domain and the ubiquitin ligase FBXW7 that recognises it, illustrating the importance of the interaction of these two proteins for the attenuation of Notch signals.

1.6 Structural and molecular studies of Notch ligands

1.6.1 The DSL domain

Canonical Notch ligands are defined by the presence of the DSL in their extracellular region. This domain was identified in Delta, Serrate and LAG-2 (Thomas et al., 1991) (Tax et al., 1994) on the basis of sequence similarity, and contains six cysteine residues, but with longer linker regions in between these residues than is found in consensus EGF sequences. The importance of this region was indicated by rescue

studies in *C. elegans*, where transgenic constructs lacking the DSL domain were unable to rescue mutant phenotypes (Henderson et al., 1997).

Mammalian ligands were subsequently identified (Bettenhausen et al., 1995), and mouse Jagged-1 deletion constructs were used to demonstrate that the DSL domain is the minimal binding unit required for interaction with Notch receptors in cell-binding and solid-phase binding assays (Shimizu et al., 1999). This study was further expanded to show that human DLL-1, Jagged-1 and Jagged-2 interacted with Notch-1 and Notch-3 in cell binding and plate based assays, and that in addition to the presence of the DSL domain, the adjacent EGF-like 1 and 2 domains are required in these ligands to achieve a higher affinity interaction with Notch receptors (Shimizu et al., 2000). The extent to which neighbouring EGF-like repeats of ligands from various species participate in the Notch interaction is unclear, although it has been demonstrated that the deletion of the EGF-like regions present in a LAG-2 transgenic construct does not appear to hinder rescue of mutant phenotypes (Henderson et al., 1997). In addition, it has been shown that whilst EGF-like domains 4-6 of Serrate are not required for the *trans*-activation of Notch, they do play a role in *cis*-inhibition (Fleming et al., 2013). Therefore, there is still much to learn about the extent to which EGF-like domains are required by various ligands for interactions with receptors.

A greater understanding of the DSL domain and its participation in the Notch ligand/ receptor interaction has emerged from structure/function studies using human and *Drosophila* Notch ligands (Cordle et al., 2008a). The X-ray structure of a human Jagged-1 fragment containing the DSL domain followed by EGF-like 1-3 (J-1_{DE3}) at a resolution of 2.5$\mathring{A}$ (Figure 1.10) revealed a distinct DSL domain fold, with similarities to the EGF fold, and it was proposed that the DSL domain arose from a duplication event involving genes encoding EGF-like domains, and subsequently

lost some of the residues mediating disulfide bond formation. Features of the DSL domain include a double stranded antiparallel sheet and three essential disulfide bonds in the pattern C_1-C_2, C_3-C_4, C_5-C_6, as opposed to the C_1-C_3, C_2-C_4, C_5-C_6 pattern observed in EGF domains.

Sequence alignment of Jagged/Serrate and Delta/Delta-like ligands from various species allowed the identification of a series of highly conserved residues in the DSL domain (Cordle et al., 2008a). Residues corresponding to Y191 and Y192 in the human Jagged-1 fragment X-ray structure were predicted to perform structural roles through packing interactions. Residues corresponding to F199, R201, R203, D205 and F207 (Figure 1.10) of Jagged-1 were observed to be exposed on the same face of the DSL domain and were predicted to form the putative Notch binding site.

A well established *in vivo* assay (Panin et al., 1997) was used to investigate the importance of these residues in *trans*-activation and *cis*-inhibition of Notch, using the development of the *Drosophila* wing as a model. Full length WT and mutant Serrate proteins were expressed at anterior/posterior boundary of the third instar *Drosophila* imaginal wing disc of transgenic flies in which wingless was used as a reporter gene. Cell aggregation experiments using Serrate WT/mutant and Notch expressing cells, along with pull down experiments between $N1_{11\text{-}13}$ and WT/mutant $J1_{DE3}$ were used to observe ligand binding irrespective of whether this could lead to signalling, the outcomes of which are summarised in Table 1.2.

WT Serrate was shown to induce both *trans*-activation and *cis*-inhibition, as well as cell aggregation, whilst WT $J1_{DE3}$ was successfully pulled down by $N1_{11\text{-}13}$, demonstrating its binding ability. However, the F257A Serrate mutant (corresponding to the F207A Jagged-1 mutant) was unable to cause either *trans*-activation or *cis*-inhibition, and the F207A substitution in $J1_{DE3}$ appeared to abrogate $N1_{11\text{-}13}$ binding entirely. In addition, F249A(Serrate)/ F199A(Jagged-1) was defective in

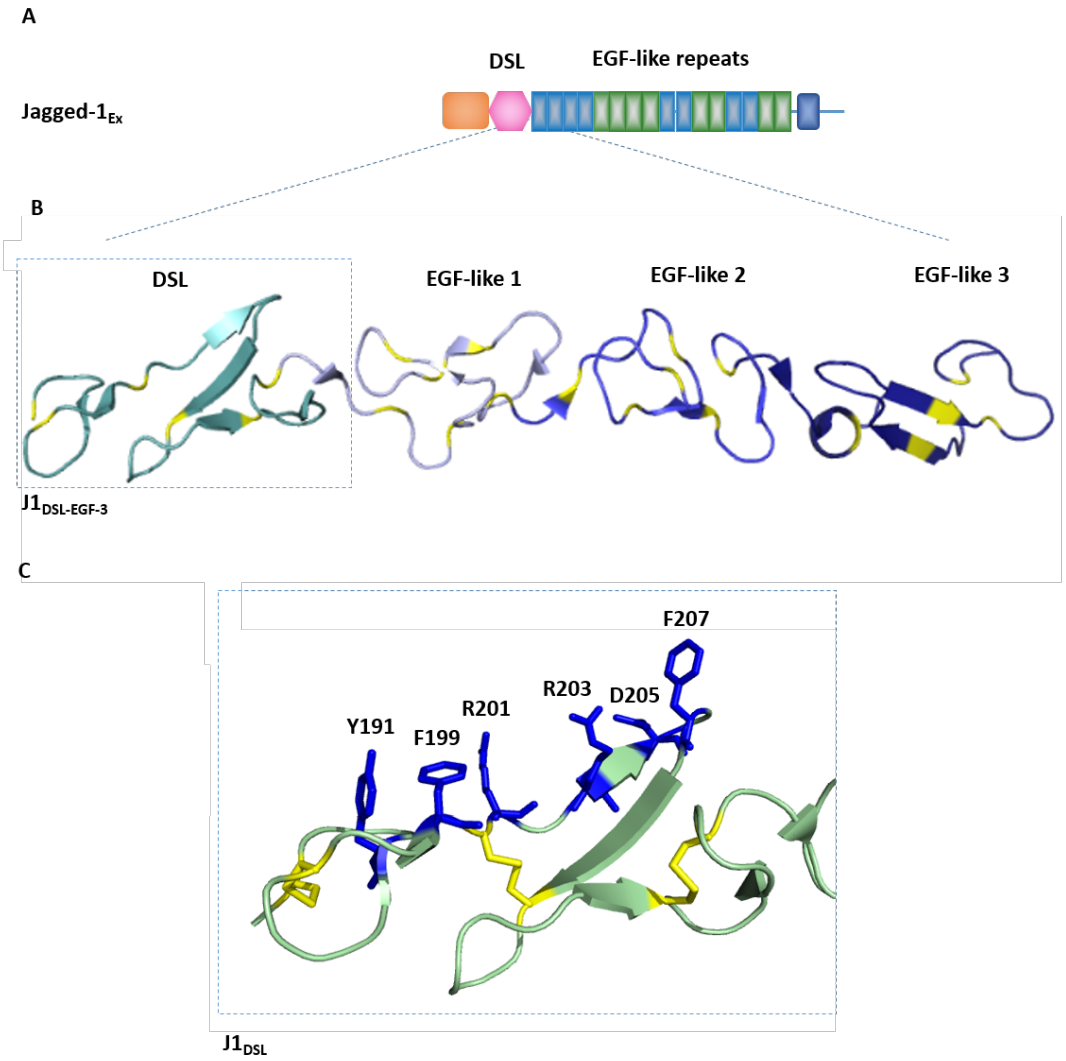


Figure 1.10: Structure of Jagged-1 DSL domain - (**A**) Cartoon represetation of the extracellular region of Jagged-1. **B**) High resolution structure of the $J1_{DE3}$ fragment (PDB accession code 2VJ2) of Jagged-1 reveals a rod-like module. (**C**) Residues F199A, R201A, R203A, D205A, F207A (blue) are predicted to form the putative Notch-1 binding site, and are conserved in *Drosophila* Serrate. In addition, conserved residues such as Y191A perform a structural role, to maintain the DSL domain fold. Cysteine residues contributing to disulfide bonds are shown in yellow.

Table 1.2: Summary of results reported by Cordle et al (2008a) for DSL mutant constructs -The table briefly summarises data collected from three assays; the *in vivo* assay based on *Drosophila* wing development using Serrate constructs, the Notch/Serrate cell aggregation assay, and the $J1_{DSL\text{-}EGF3}/N1_{11\text{-}13}$ pull-down assay.

Construct: hJagged1 (Serrate)	**Ability to be pulled down by $N1_{11\text{-}13}$**	**Ability to promote cell aggregation**	**Ability to *trans*-activate *in vivo***	**Ability to *cis*-inhibit *in vivo***
Wild type	Yes	Yes	Yes	Yes
F199A (F249A)	Yes	Yes, reduced	No	Only at 29°C
R201A (R251A)	Yes	Yes, reduced	Yes, reduced	Yes, reduced
R203A (R253A)	N/A	N/A	No	Only at 29°C
D205A (D255A)	N/A	N/A	Yes, reduced	Yes, reduced
F207A (F257A)	No	No	No	No

trans-activation, but still able to bind WT Notch-1 in cell aggregation experiments. Taken together, the structural and functional data provided in this paper allowed the identification of several key residues, conserved in both *Drosophila* and mammals that are involved in binding to Notch receptors. In particular, these data highlighted the importance of a conserved aromatic residue within the DSL (F207/F257).

1.6.2 The module at the N-terminus of Notch ligands (MNNL)

All canonical Notch ligands contain a region of over 100 amino acids at their N-termini, preceding the DSL domain, known as the MNNL. MNNL sequences show variability both between ligands of the same and differing species, although when analysed as two separate domains on the basis of cysteine content, fly and vertebrate MNNLs do show some relatively well conserved region of sequence similarity (Parks et al., 2006). It has been a region of particular interest due to the high frequency of AGS-associated mutations in this region (section 1.3.1), although computational predictions alone were insufficient to define the domain fold adopted by this region (C. Chillakuri, personal communication).

Early evidence indicating this region plays an important role in Notch signalling came from rescue studies, in which transgenic constructs encoding LAG-2 were modified to contain deletions of specific domains, and their ability to rescue mutant phenotypes assessed (Henderson et al., 1997). Deletion of the MNNL region of LAG-2 abrogated its ability to rescue phenotypes, suggesting a disruption in signalling. The requirement of the MNNL was supported by cell aggregation data indicating that the entire N-terminal region up to and including the DSL domain of Delta is necessary for dNotch binding (Parks et al., 2006).

The requirement for the MNNL domain for productive signalling was also demonstrated in assays designed to compare Notch-1 affinity for human DLL-1 and DLL-4

(Andrawes et al., 2013). Truncated fragments of both ligands spanning the MNNL-EGF-like 3 were found to be necessary and sufficient to achieve levels of activation associated with full length constructs in a luciferase reporter assay, whilst the MNNL-EGF-like 5 fragments were used with full length and Notch-1EGF-like 6-15 receptors to probe the differences of affinity between Notch-1/DLL-1 and Notch-1/DLL-4 interactions. It was concluded from these data that Notch-1 has a ten-fold greater affinity for DLL-4, compared to DLL-1 and speculated that the key to Notch receptor selectivity of various ligands may lie within the MNNL/DSL region, on the basis that these regions share a lower sequence identity than the EGF-like regions.

Structure/function analysis of the MNNL of human Jagged-1 has led to a significant increase in our understanding of this region (Chillarkuri et al., 2013). The X-ray structures of a fragment of human Jagged-1 including the MNNL, DSL and EGF-like 1-3 ($J1_{NE3}$) in the presence and absence of calcium revealed a C2 domain, which is likely to bind calcium in the matrix environment. Calcium dependence was confirmed using a limited proteolysis assay, and additional experiments revealed phospholipid binding properties also associated with this domain. This is consistent with the general properties of many C2 domains, which are often found in intracellular proteins with membrane-related functions (Lemmon, 2008).

The C2 domain of Jagged-1 adopts the common C2 domain fold of an eight-stranded anti-parallel β sandwich, which harbours a cationic β groove (Figure 1.11). This region is thought to play a role in making the initial contacts between the C2 domain and the cell membrane, allowing other regions to interact more selectively. AGS mutations tend to map to the core of the β sandwich, consistent with evidence that these mutations cause misfolding of Jagged-1 (Chillakuri et al., 2013). Since selectivity of C2 domains tends to occur through contacts made in their highly variable loop regions, site-directed mutagenesis was performed at specific residues in and

around the calcium-binding site. Aspartate residues were targeted for mutagenesis, based on their role on conferring integrity of the calcium-dependent phospholipid-binding site in a structurally homologous C2 domain of the human protein, Munc-13.

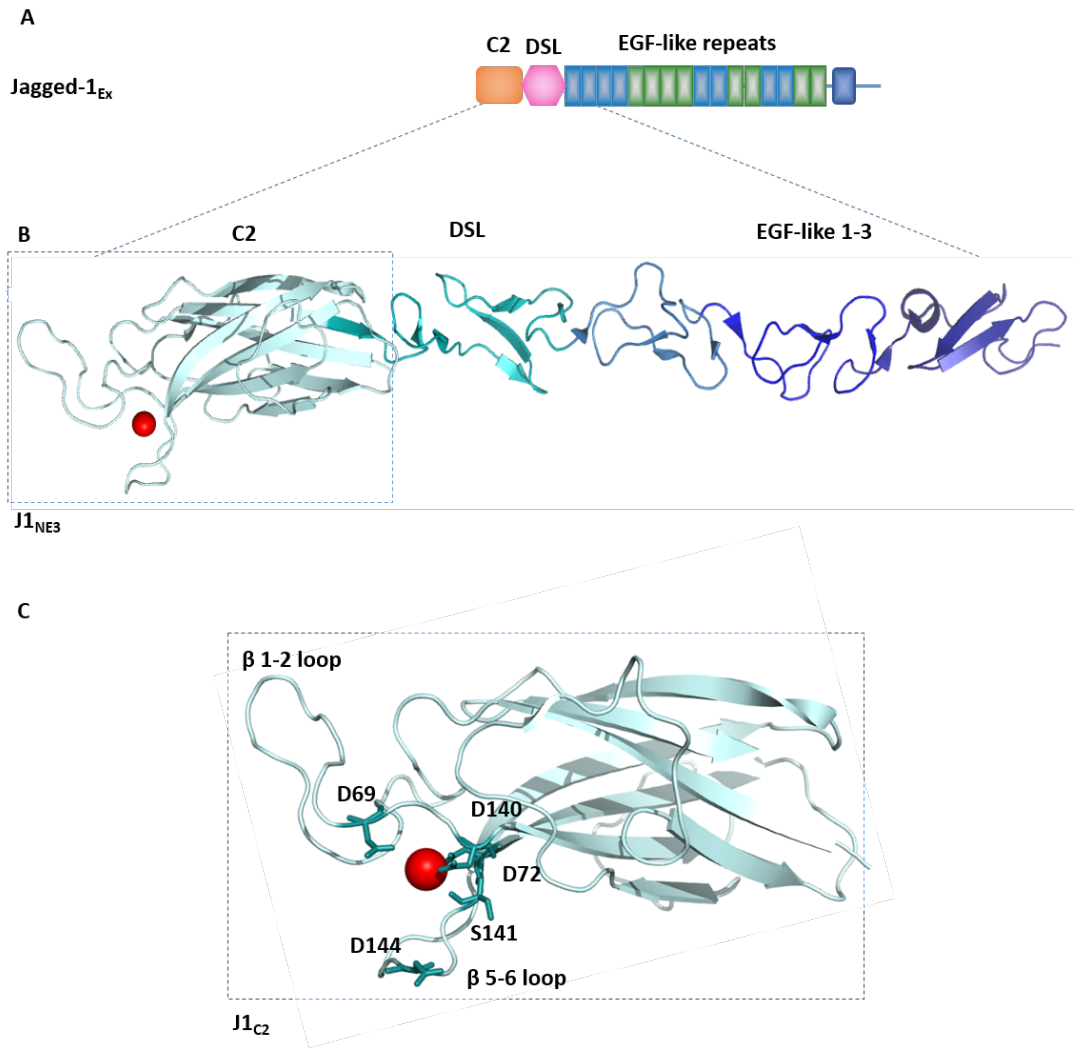


Figure 1.11: Structure of the C2 domain at the N-terminus of human Jagged-1 - (**A**) Cartoon representation of the extracellualar region of Jagged-1. (**B**) Full structure of the $J1_{NE3}$ fragment resolved to 2.5$\AA$. PDB accession code 4CC1. (**C**) Close view of the calcium-binding site of the C2 domain, highlighting residues that were selected for mutagenesis studies.

Several assays were used to demonstrate the calcium-dependent phospholipid-binding properties of the WT and mutant variant C2 domains of Jagged-1, to investigate the role of the C2 domain. ELISA data demonstrated that binding of the

$J1_{NE3}$ to a fragment of Notch-1 was dependent on the presence of calcium. Binding in this assay appeared unaffected in single and double D140/144 mutants, indicating that these residues are unlikely to be interacting with Notch receptors directly. However, in a cell signalling assay utilising a split-luciferase reporter system (Ilagan et al., 2011), recombinant D140/D144 $J1_{NE3}$ impaired activation of full length Notch-1-NLuc expressed on the surface of HelaTetON cells, suggesting that they play an important role to bring about productive activation of the receptor. Interestingly, a phospholipid-binding assay showed that the D140/D144 double mutant had substantially reduced phospholipid-binding abilities in comparison to the WT construct, although the *in vivo* relevance of this is unclear. The conclusions drawn from this initial structural study suggested that the C2 domain at the N-terminus of human Jagged-1 is likely to be a common feature of canonical Notch ligands, despite variations in sequence at loop regions between β strands, and this has been recently confirmed. Since the publication of this structure, improved protein expression technologies including use of the Expres^{2}ion *Drosophila* S2 expression system allowed the crystallisation of equivalent N-terminal fragments of human Jagged-2, DLL-4 and *Drosophila* Serrate (R. Suckling, personal communication). Consistent with predictions, both ligands contain a C2 domain at the MNNL, with a fold that is highly conserved with the human Jagged-1 C2, although a calcium-bound structure has not been observed for DLL-4.

Conservation of the C2 domain in Notch ligands was further supported by the structure of human DLL-1 (Kershaw et al., 2015). The full ectodomain of DLL-1 was expressed as an Fc tagged construct from HEK293F cells, and the X-ray structure solved to a resolution of 3.2$\mathring{A}$, revealing a C2 domain at the MNNL sharing significant structural homology with the Jagged-1. Despite the absence of some residues from the crystal structure, information from the visible residues along with super-

position of the higher resolution Jagged-1 structure discussed above and sequence alignment information allowed the authors to identify conserved regions between the two ligands. Interestingly, the DLL-1 was crystallised in the presence of 5 mM calcium, but failed to bind calcium in the crystal structure. Closer inspection of the conservation of calcium coordinating aspartate residues of Jagged-1 in primary sequence alignments revealed that these were replaced by histidine and arginine in the DLL-1 structure, and this absence of potential calcium binding residues is predicted to be a common feature of the Delta-like class of ligands. In addition to the structure of DLL-1, the study attempted to investigate the binding affinity of DLL-1 to Notch-1, using dimeric constructs in ELISA and monomeric, soluble constructs in analytical ultracentrifugation. However, the results from these efforts were inconsistent, highlighting the challenges faced when attempting to measure physiologically relevant affinities of interactions between Notch receptors and their ligands.

1.6.3 The Notch receptor/ligand interaction: new insights by co-crystallisation

A landmark event for our understanding of the Notch receptor/ligand interaction was the successful co-crystallization of N-terminal fragments of rat DLL-4 (rDLL-4) in complex with the major ligand-binding region of rat Notch-1 (rNotch-1) (Luca et al., 2015). The relatively low affinity of the Notch receptor/ligand interactions and the difficulties simulating the native environment of these proteins (including post-translational modification and calcium binding) have limited the ability to capture receptor and ligand fragments in complex. However, steps were taken in this study to enhance the affinity of the interaction. Mammalian Notch receptors are highly conserved, therefore rDLL-4 and rNotch-1 were selected for the study based on previous indications that the affinity of this interaction is higher than alternative

receptor/ligand pairings (Andrawes et al., 2013). A fragment of rDLL-4 spanning the MNNL-EGF-like 5 was expressed on the cell surface of yeast, and affinity maturation carried out. After multiple rounds of selection, several mutations were identified that enhanced binding to rNotch-1 binding. The relative binding affinities of these mutants were assessed comprehensively using SPR, and functionality confirmed in a Notch-1 activation assay based on the split luciferase reporter system previously discussed. This allowed identification of a SLP triple mutant (G28S, F107L, L206P), with enhanced affinity for rNotch-1. Two fragments of this mutant were expressed in SF9 insect cells, one spanning MNNL-EGF-like 2 (rDLL-$4_{\text{SLP NE2}}$) and a shorter one from MNNL-EGF-like 1(rDLL-$4_{\text{SLP NE1}}$). Both were crystallised in complex with $rN1_{11\text{-}13}$, with the rDLL-$4_{\text{SLP NE2}}/N1_{11\text{-}13}$ structure resolved at 3.4$\mathring{A}$ and rDLL-$4_{\text{SLP NE1}}/N1_{11\text{-}13}$ at 2.3$\mathring{A}$, giving atomic resolution insights into the contacts formed between the receptor and ligand fragments (Figure 1.12). The structural information supported several observations from previous studies (Cordle et al., 2008a) (Chillakuri et al., 2013) (Whiteman et al., 2013) and also expanded our understanding of the interaction interface, by identifying a second interaction site between the C2 domain of the ligand, and the major ligand-binding region of the receptor, in addition to the known DSL domain mediated interaction site (Figure 1.12A).

Surprisingly, the DSL domain residues including R191 and F195 of rDLL-4 (equivalent to Jagged-1 R201 AND F207, section 1.6.1) were shown to interact with residues located EGF-like 11 of rNotch-1, contradictory to the previous proposal that the DSL mediated receptor/ligand interaction centred on EGF-like 12 (Cordle et al., 2008a). Amongst the interactions mapped at the interface, R191 of rDLL-4 was shown to interact with the N421 of rNotch-1, whilst F195 interacts directly with R448 of the receptor, with an additional contribution from Y216 interacting with E450 (Figure 1.12B). A further interesting discovery from this co-crystal structure

was the presence of an O-linked glycan at S435 of rNotch-1, which is predicted to interact with D218 of DLL-4, however the precise sugar moiety present at this site under physiological conditions is unknown.

Interestingly, the additional receptor/ligand binding site identified in this structure, was shown to occur between the C2 domain of DLL-4 and Notch-1 EGF-like 12. A conserved loop between β strands 2 and 3 that is distant from the predicted Jagged-1 calcium/phospholipid-binding region presents a binding platform for Notch-1 EGF-like 12. The loop is supported by a disulphide bond between C61 and C74, presenting residues H64 and Y65 to make direct contacts with Notch-1 at I477, and accommodating the fucose moiety on Thr466 of Notch-1, with L468 and D469, previously shown to be important for Jagged-1 binding (Whiteman et al., 2013) occupying a hydrophobic cleft of the DLL-4 C2. H64A and Y65A substitutions were subsequently generated, and these mutants were associated with a significant reduction in Notch-1 binding.

Both sequence and structural modelling indicates this loop is a conserved feature of both Jagged and Delta-class mammalian ligands, with an aromatic residue (Y/F) is present at the Y65 site. In Jagged-1 and Jagged-2, the equivalent residue of H64 is substituted with a glutamate, which could perhaps contribute to differences in binding affinities between ligands, by modulating their interaction with fucosylated Notch. The authors computationally modelled the extension of the fucose monosaccharide present in the structure to a disaccharide, which revealed additional DLL-4 contacts, consistent with the enhanced ligand-binding affinity detected for disaccharide modified Notch-1 (Taylor et al., 2014).

Whilst these structural insights have enhanced our understanding of the initial binding event occurring between Notch receptors and their ligands, the unusual phospholipid-binding properties of the N-terminal region of Notch ligands have yet

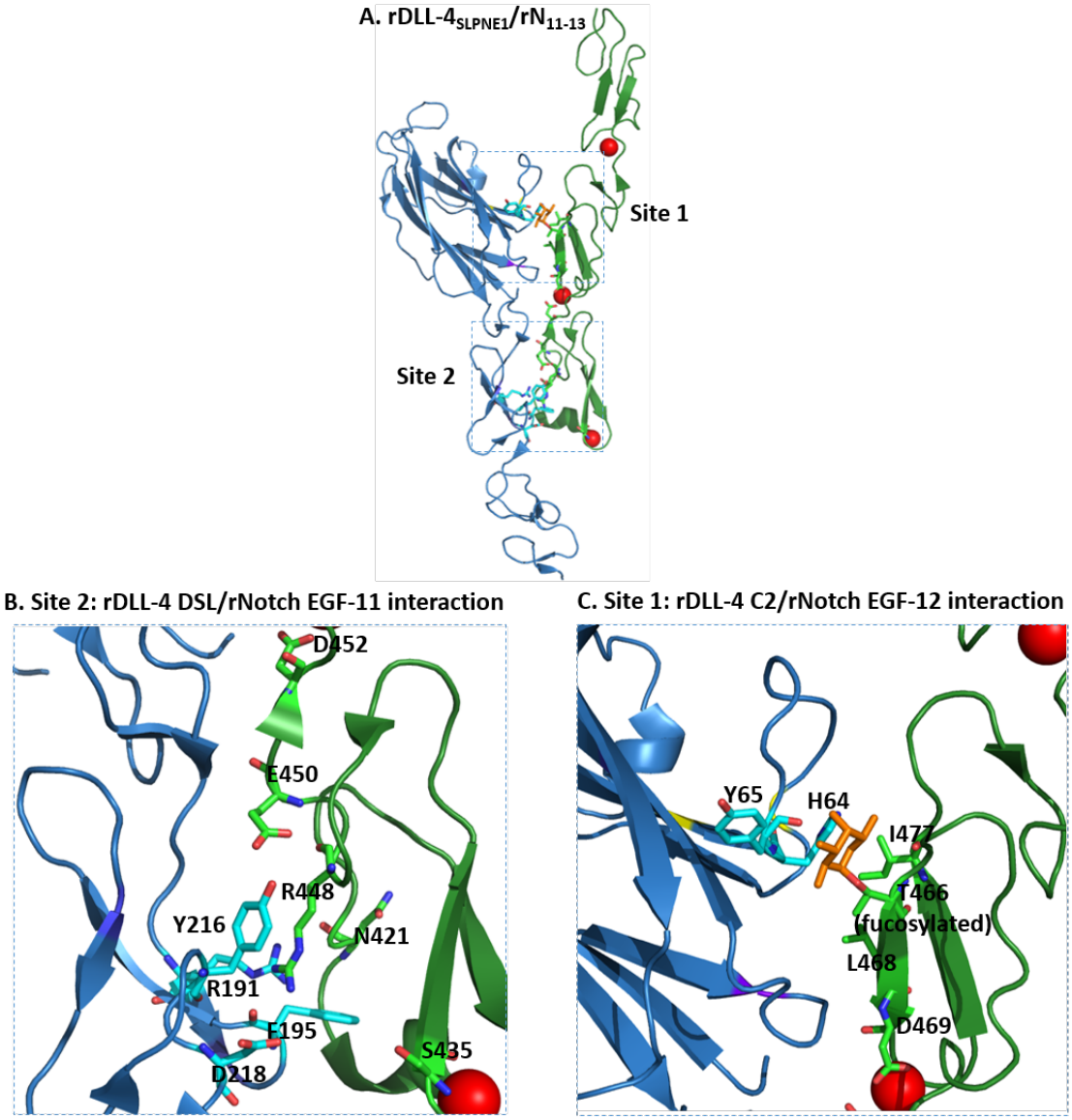


Figure 1.12: X-ray structure of rDLL-4$_{\text{SLP NE1}}$/rN1$_{11\text{-}13}$ complex - (**A**) Full structure of the complex resolved to 2.5$\AA$ (PDB accession code 4XL1). (**B**) Close view of rDLL-4 DSL/rNotch-1 EGF-like 12, highlighting key residues involved in the interaction. R448 of rNotch-1 interacts forms an H-bond with F195 and a salt bridge with D218, with additional contacts between N421 and R191, and E450 and Y216. (**C**) Close view of rDLL-4 C2/rNotch-1 EGF-like 11 interaction, highlighting key residues involved in the interaction. H64 directly interacts with I477 of rNotch EGF-like 12, whilst aromatic Y65 accommodates the fucose of T466. C63 and C74 (yellow) form a disufide bridge to stabilise loop.

to be fully understood. It has been suggested that the C2 domain of Notch ligands facilitates the step between the initial binding event between the ligand and receptor, and *trans*-activation of the Notch pathway. Despite the great advances to our understanding that this study contributed, there remain several unanswered questions as to the full role of this C2 domain.

1.7 Aims of the thesis

The aims of this work were:

To compare *C. elegans* Notch receptors and ligands to their *Drosophila* and mammalian counterparts, to assess the extent to which they will provide a suitable model for Notch signalling

To express recombinant fragments of *C. elegans* Notch ligands, and purify these constructs for biochemical and structural characterisation, allowing functionally significant regions in the N-terminus to be identified for mutagenesis

To establish an *in vivo* assay in which the effects of these mutations in *C. elegans* Notch ligands can be explored at the physiological level

To bring the complementary biochemical/structural and *in vivo* data together to gain novel insights into the functionality of the N-terminal region of Notch ligands

2

Materials and Methods

2.1 DNA methods

2.1.1 General DNA methods

2.1.1.1 Polymerase chain reaction (PCR)

PCR amplification of DNA was achieved using KOD Hot Start polymerase (Novagen). Reactions were carried out using 0.2 ml thin-walled strip tubes contain a final reaction volume of 50 μl, and prepared as advised by instructions accompanying the product.

2.1.1.2 DNA gel electrophoresis

Gels were prepared at a concentration of 0.7-2% w/v agarose in 1X TAE buffer (40 mM Tris pH 8.8, 20 mM acetic acid, and 1 mM EDTA). Ethidium bromide was added to a final concentration of 0.25%. KOD PCR products were mixed with an appropriate volume of 6 x loading DNA loading dye (30% (w/v) glycerol, 300 mM EDTA pH 8.0, 0.25% (w/v) bromophenol blue), loaded into wells and run at 80V for 1 hour. DNA bands were subsequently visualised using a UV transilluminator (UVItech).

2.1.1.3 Gel extraction

DNA bands were excised from gels during UV visualisation using a blade, and DNA was subsequently extracted from the gel slices using the QIAquick Gel extraction kit reagents and accompanying instructions (Qiagen). DNA products were eluted in 25 μl ddH_2O.

2.1.1.4 T/A cloning with TOPO-XL

PCR products generated for insertion into the TOPO-XL vector were eluted in 10 μl ddH_2O after gel extraction. To this, 1 μl dATP (10 mM stock) 2 μl Taq Thermopol buffer (New England Biolabs (NEB)) and 0.5 μl Taq polymerase (NEB) were added, and the reaction mixture was heated at 72°C for 10 minutes in a thermocycler to generate adenosine overhangs at the end of the PCR product. 4.5 μl of the final mixture was them used as the substrate for the TOPO-XL cloning reaction, performed using the TOPO-XL PCR Cloning Kit (Life technologies) according to instructions provided.

2.1.1.5 Restriction digestion and ligation

Restriction digestion was performed using enzymes and buffers provided by NEB. Buffer choices and volumes of enzyme were used as directed by the manufacturer. Digestions were performed using 0.5-1.5 μg of substrate DNA in a 50-100 μl final reaction volume, and allowed to proceed at 37°C for 60-90 minutes. To prevent self-ligation, digested vectors were treated with 0.5 μl of calf intestinal alkaline phosphatase (CIAP, 10 U/μl, NEB) after 60 minutes, and reaction continued for a further 30 minutes. Digested DNA products were then run on an appropriate percentage agarose gel and excised/extracted to ensure purity. Complementary overhangs were annealed using the Rapid ligation kit (Roche), as per the instructions

provided. The concentration of the vector and insert DNA was estimated from an agarose gel, and the two were added with an approximate ratio of 1:3 (vector:insert) molar ends in the 20 μl ligation reaction volume, with ~100 ng of vector used per reaction. A negative control reaction was also set up containing digested vector in the absence of insert. The ligation reaction was allowed to proceed at room temperature for no more than 10 minutes.

2.1.1.6 Quikchange mutagenesis

Single base changes were introduced into target DNA using the QuikChange II Site-Directed Mutagenesis Kit (Agilent). Mutagenic primers designed using guidelines suggested by the accompanying protocol were used for amplification by PCR using reagents supplied. Subsequently, parental DNA was removed from the reaction upon addition of 1 μl *Dpn*I (10 U/μl) supplied with the kit, and incubation for 1 hour at 37°C.

2.1.1.7 Transformation of chemically competent cells

High efficiency NEB 5-α competent *E. coli* cells were routinely used for the transformation of DNA. Either 2 μl T/A cloning product, 5 μl of ligated cloning product or 2 μl of *Dpn*I treated Quikchange reaction mixture was added to a single 50 μl aliquot of cells and incubated on ice for 30 seconds. This was followed be heat shock at 42°C for 30 seconds and a further 2 minutes on ice. 950 μl of room-temperature SOC outgrowth medium (NEB) was then added, and the cells incubated at 37°C for 1 hour at 300 rpm. 1:1, 1:10 and 1:100 dilutions of the recovered cells were made, and 100 μl of each spread on agar plates containing the appropriate antibiotic. Agar plates were prepared by autoclaving 2xTY broth with 1.5% w/v agar. The media was then cooled below 50°C, and antibiotics were added as necessary. Final concentrations of

antibiotics used were 100 μg/ml ampicillin or 50 μg/ml kanamycin.

2.1.1.8 Transformation of electro-competent cells

A 50 μl aliquot of OneShot TOP10 electro-competent cells (Life technologies) was thawed on ice, and 1 μl of reaction mixture added, mixed, and transferred to a 2 mm electroporation cuvette. Electroporation of cells was conducted using a single pulse of $\sim$5 ms on program EC2 of a Bio-Rad micropulser. Cells were then immediately recovered with 250 μl of room-temperature SOC outgrowth medium at 37°C, 300 rpm for 1 hour. The entire reaction volume of recovered cells were then spread on agar plates containing appropriate antibiotic.

2.1.1.9 Colony screening

Single colonies were picked with a sterile pipette tip and re-suspended in 10 μl ddH_2O. To 5 μl of this re-suspension, 50 μl 2xTY broth containing appropriate antibiotic was added, and the mixture incubated at 37°C for 6 hours to create a seed culture. During this incubation period, the other 5 μl of each mixture was used as a template for screening using PCR. Colony PCRs were performed using Thermopol Taq polymerase and buffer (NEB), and were performed using the reagents and conditions recommended by the manufacturer.

2.1.2 Generating constructs for expression in HEK293T cells

Constructs were designed to be compatible with a previously outlined expression protocol (Aricescu et al., 2006). DNA sequences were amplified by KOD Hot Start polymerase PCR from a yeast 2 hybrid *C. elegans* cDNA library using appropriate primers (Table 2.1). A native *Eco*RI site from the *apx-1* cDNA sequence was removed by silent mutagenesis using the QuikChange II kit, with primers also listed.

Table 2.1: List of oligonucleotides used for amplification from cDNA library for cloning into modified pHLsec. Primers were designed with a GC content of 40-60% and 65°C>Tm>60°C. Added restriction sites are shown in blue. Sites of point mutagenesis in red.

Target	Primer	Sequence (5'→3')
lin-12$_{1\text{-}13}$	*Mfe*I *lin-12*cDNA F	GC CAATTGATGCGGATCCCTACGATTTG
	*Kpn*I *lin-12*cDNA R	GGGTACCTTCTTCACACCTGGCTCCACTAAATC
lag-2$_{NE2}$	*Mfe*I *lag-2*cDNA F	GCCAATTGATGATCGCTTACTTCCTCTTACTCC
	*Kpn*I *lag-2*cDNA R	GCGGTACCGGTGGTGAGTGAGATTTCACAG
dsl-1$_{Full}$	*Eco*RI *dsl-1*cDNA F	GGAATTCATGCTCAAATATCTTATATTCCTTGCTATATT
	*Kpn*I *dsl-1*cDNA R	GCGGTACCCAATCGAGGAAGCGTGGCCAA
apx-1$_{NE4}$	*Eco*RI *apx-1*cDNA F	GGAATTCATGACTAACTTCTCTTCCCTTCTCACC
	*Kpn*I apx-1cDNA R	GCGGCATTGTGATTTCCAATTTCACAGAATTTTCCCG
	A819G F	CATTGCCCATTC**GA**G**TTC**GTCGGCTCTCAATG
	A819G R	CATTGAGAGCCGAC**GAA**C**TC**GAATGGGCAATG
apx-1$_{DE4}$	*Eco*RI *apx-1*DE4cDNA F	GGAATTCAATTTGTGCAGCTCAAACTACCACGG
	*Kpn*I *apx-1*cDNA R	GCGGCATTGTGATTTCCAATTTCACAGAATTTTCCCG

PCR products were purified and digested, and inserted into the pHLsec vector (previously modified by C. Chillakuri to contain an Igk leader sequence and C-terminal Fc-His tag sequence) using *Eco*RI and *Kpn*I restriction sites. Human Jagged-1 expression constructs were cloned and kindly provided for comparison by C. Chillakuri. The sequence encoding APX-1$_{NE4}$ cDNA was codon optimised for human cell line expression and synthesised using services provided by Genscript (www.genscript.com) (Appendix 1). The optimised DNA was delivered in pUC17, digested with *Eco*RI and *Kpn*I and inserted into the modified pHLsec vector as before.

2.1.3 Generating constructs for expression in *Drosophila* S2 cells

The sequence encoding APX-1_{NE3} was codon optimised and synthesised for *Drosophila* expression using Geneart services by Life technologies (www.lifetechnologies.com) (Appendix 1). The sequence submitted for synthesis included an N-terminal *Eco*RI site, followed by a Kozak sequence (GCCACC) the *Drosophila* BiP signal peptide sequence in place of the native APX-1 signal peptide sequence, and predicted cleavage confirmed using Signal.P 4.1 (Petersen et al., 2011). A C-terminal hexa-his tag was also included, between *Xba*I and *Not*I sites. The synthesised construct encoding APX-1_{NE3His} was delivered in a pMA-T vector backbone, and subsequently cloned using *Eco*RI and *Not*I restriction sites into the pExpres2-2 vector to give a final expression plasmid for transfection into S2 cells using Expres^{2}ion technology.

2.1.4 Generating constructs for *C. elegans* microinjection

2.1.4.1 Preparation of fosmid DNA

The ∼30kb fosmid; WRM0620cE03; spanning the *lag-2* gene and its flanking regions was obtained from a fosmid library (kindly provided by J. Hodgkin). Cells containg the fosmid were streaked out on an agar plate containing 12.5 μg/ml chloramphenicol (Chl) and grown overnight at 37°C. Single colonies were then picked into 5 ml of 2xYT media containing 12.5μg/ml Chl and grown overnight at 37°C at 200 rpm. 5 μl of 10% w/v arabinose solution was used to induce 500 μl of the overnight culture in a fresh 5 ml of 2xYT with 12.5 μg/ml Chl. This was shaken at 200 rpm at 37°C for 5 hours in 50 ml tubes to maximise aeration. The fosmid DNA was then mini-prepped the sample analysed by agarose gel electrophoresis to verify purity.

2.1.4.2 Isolation of target sequences for use in rescue constructs

The *lag-2* promoter region was amplified from the plasmid pJK520 (a gift from J. Kimble), to give the region spanning -3165 bp to -12 bp upstream of the *lag-2* gene. The reverse primer included the addition of a 42 bp synthetic intron as 3' tag to the promoter sequence (sequence from A. Fire), and the final product termed '*lag-2p*'. Sequencing of the PCR product indicated 3 deviations from the wild type sequence reported on Wormbase (www.wormbase.org), including two point mutations and a 16 bp deletion. The two point mutations were corrected to the WT sequence using the Quikchange II mutagenesis kit and primers listed (Table 2.2). The deletion was corrected by fusing two overlapping oligos (Table 2.2) and inserting the product at native *Xba*I sites flanking the deleted region. A subsequent shorter fragment of the *lag-2* promoter (-1597 bp to -12 bp with 42 bp synthetic intron) was amplified using an alternative forward primer and termed '*rlag-2p*'.

Full length *lag-2*, *dsl-1* and *apx-1* cDNA sequences were amplified by PCR from a yeast 2 hybrid cDNA library using primers shown in Table 2.3. These PCR products were then cloned into TOPO XL using T/A sites, and the final clone used as a template for subsequent PCR reactions to create rescue constructs.

2.1.4.3 Assembly of rescue constructs for microinjection and extrachromosomal array formation

An initial scheme to generate these constructs using Multisite Gateway Pro Four-fragment cloning kit (Life technologies) was pursued. Protocols were followed according to the accompanying manual, and all required entry clones were successfully made and confirmed with sequencing. However, recombination of the entry clones into the destination vector pcDNA6.2/V5-Pl-DEST (Life technologies) was unsuccessful, despite extensive trouble shooting, most likely caused by an extremely low

Table 2.2: List of oligonucleotides used to generate final lag-2 promoter/42 bp synthetic intron region. 42 bp synthetic region shown in red, sites of point mutation shown in red and *Xba*I restriction sites shown in blue.

Target	Primer	Sequence (5'→3')
lag-2p	-3165 *lag-2p* F	GGCAATTTGCCAAAACTTTGACCGGA
	42bp *lag-2p* R	CTGAAAATGTTCTATGTTATGTTAGTATCATTCGAAA CATACGGCAAATTTGAAAAGTGTTGTTGGCTTCTAG
rlag-2p	-1597 *lag-2p* F	CCATCAGCCATTTCCCCTGTCCAG
	42bp *lag-2p* R	CTGAAAATGTTCTATGTTATGTTAGTATCATTCGAAA CATACGGCAAATTTGAAAAGTGTTGTTGGCTTCTAG
lag-2p -601 C>T	-601 C>T F	GACAGACAATTACTTCTGCTTGAAATG
	-601 C>T R	CATTTCAAGCAGAAGTAATTGTCTGTC
lag-2p -1971 C>T	-1971 C>T F	GGTTTTTTTTTAATTCATAATTTAAAAAAATCTATAA AAATTG
	-1971 C>T R	CAATTTTTATAGATTTTTTTAAATTATGAATTAAA AAAAAACC
lag-2p 16bp insertion	16bp insert F	GTCTAGATTTGCAGAACATCTGTGTTAGG TCAAGCAGAGCGAGAATTTGAGCATTTGAACTTTAG
	-16bp insert R	GTCTAGACACCCCCCCCCCCCCCCCCCCTC CCCCCCCCAGGACCCTAAAGTTCAAATGCTCAAA

Table 2.3: List of oligonucleotides used to generate full length ligand cDNA sequences for rescue constructs. All ligand sequences amplified were full length.

Target	Primer	Sequence (5'→3')
lag-2 cDNA	*lag-2* cDNA F	ATGATCGCTTACTTCCTCTTACTCCTCAC
	lag-2 cDNA R	CTAGACATAGTGACAGGCTGGAATAGACG
dsl-1 cDNA	*dsl-1* cDNA F	ATGCTCAAATATCTTATATTCCTTGCTATATTGATCTCT
	dsl-1 cDNA R	TCACAATCGAGGAAGCGTGGC
apx-1 cDNA	*apx-1* cDNA F	ATGACTAACTTCTCTTCCCTTCTCACCA
	apx- 1 cDNA R	TTAGACTCGAAACGACGACTCGTC

efficiency of recombination of the 4 individual fragments.

Rescue constructs were designed to encode full length ligand cDNA::GFP fusions flanked at the 5' end by *lag-2p/rlag-2p*, and at the 3' end, the *unc-54 3' utr* region was present for regulation of expression (Figure 2.1). *ligand::gfp::unc-54 3' utr* fusion constructs were generated by overlapping PCR. Ligand cDNAs were amplified from their templates in TOPO-XL, whilst the *gfp::unc-54 3' utr* sequence was amplified from the plasmid pPD95.75 using primers shown (Table 2.4).These fragments were fused in a subsequent PCR reaction using outer primers, to give the final *ligand::gfp::unc-54 3' utr* fusion product. This was then digested using *Not*I and ligated into TOPO-XL that had *lag-2p* or *rlag-2p* inserted at the T/A site. The final *lag-2p::ligand::gfp::unc-54 3' utr* was midi-prepped and confirmed by sequencing prior to microinjection. Mutant constructs were also generated using overlapping PCR. In the first step, the ligand fragment was mutagenized using primers shown (Tables 2.5 and 2.6). Subsequently, these ligand fragments were fused to *gfp::unc-54 3' utr* and inserted into the TOPO-XL/ *lag-2* promoter backbone.

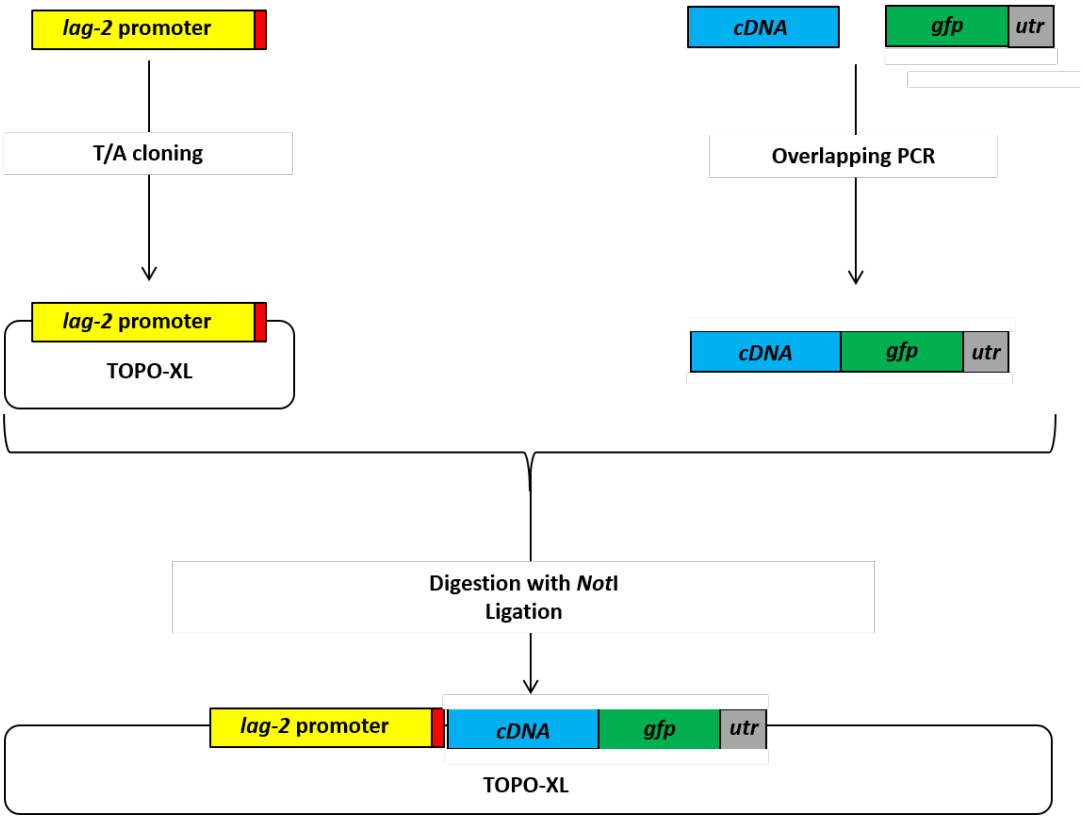


Figure 2.1: Scheme for generation of final constructs for microinjection - A fragment of the *lag-2* promoter including a 42 bp synthetic intron (red) was amplified and inserted into the TOPO-XL backbone. *ligand::gfp::unc-54 3' utr* fusion constructs were generated by overlapping PCR, with *Not*I restriction sites, allowing their insertion into the lag-2 promoter modified vector backbone. This general method was subsequently used for the creation of other variant constructs (including *lag-2* cDNA only, *rlag-2p* and *apx-1::gfp* mutants).

Table 2.4: List of oligonucleotides used to generate fusion constructs for insertion into *lag-2* promoter/TOPO-XL backbone. Complementary overlapping regions are underlined. *Not*I restriction sites shown in blue.

Target	Primer	Sequence (5'→3')
lag-2 ::gfp::unc-54 3'utr	*lag-2* Outer F	GAGCGGCCGCATGATCGCTTACTTCCTCTTACTC
	lag-2 Inner R	GTTCTTCTCCTTTACTGACATAGTGACAGGCTGGAATAG
	gfp Inner F	GCCTGTCACTATGTCAGTAAAGGAGAAGAACTTTTCAC
	gfp Outer R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATTG
dsl-1:: gfp::unc-54 3' utr	*dsl-1* Outer F	GAGCGGCCGCATGCTCAAATATCTTATATTCCTTGCTATAT
	dsl-1 Inner R	GTTCTTCTCCTTTACTCAATCGAGGAAGCGTGGCCAAC
	gfp Inner F	CACGCTTCCTCGATTGAGTAAAGGAGAAGAACTTTTCACTG
	gfp Outer R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATTG
apx-1:: gfp::unc-54 3' utr	*apx-1* Outer F	GAGCGGCCGCATGACTAACTTCTCTTCCCTTCTC
	apx-1 Inner R	GTTCTTCTCCTTTACTGACTCGAAACGACGACTCGTCATC
	gfp Inner F	GTCGTCGTTTCGAGTCAGTAAAGGAGAAGAACTTTTCACTG
	gfp Outer R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATTG
lag-2:: unc-54 3' utr (no gfp)	*lag-2* Out F	GAGCGGCCGCATGATCGCTTACTTCCTCTTACTC
	lag-2 In R	GAATTCTACGAATGCTAGACATAGTGACAGGCTGGAATA
	unc-54 In F	CAGCCTGTCACTATGTCTAGCATTCGTAGAATTCCAACTGAG
	unc-54 Out R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATTG

Table 2.5: List of oligonucleotides used to generate whole domain deletion constructs for insertion into *lag-2* promoter/TOPO-XL backbone. Complementary overlapping region are underlined. *Not*I restriction sites shown in blue.

Target	Primer	Sequence (5'→3')
Δ*MNNL apx-1::gfp ::unc-54 3' utr*	Δ*MNNL* Outer F	GAGCGGCCGCATGATCGCTTACTTCCTCTTACTC
	Δ*MNNL* Inner R	GTTCTTCTCCTTTACTGACATAGTGACAGGCTGGAATAG
	gfp Inner F	GCCTGTCACTATGTCAGTAAAGGAGAAGAACTTTTCAC
	gfp Outer R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATTG
Δ*DSL apx-1::gfp ::unc-543' utr*	Δ*DSL* Outer F	GAGCGGCCGCATGCTCAAATATCTTATATTCCTTGCTATAT
	Δ*DSL* Inner R	GTTCTTCTCCTTTACTCAATCGAGGAAGCGTGGCCAAC
	gfp Inner F	CACGCTTCCTCGATTGAGTAAAGGAGAAGAACTTTTCACTG
	gfp Outer R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATTG
Δ*EGFL-1 apx-1::gfp ::unc-54 3' utr*	Δ*EGFL-1* Outer F	GAGCGGCCGCATGCTCAAATATCTTATATTCCTTGCTATAT
	Δ*EGFL-1* Inner R	GCGTGGTAGGCACTGACAATCTTCACCAGACCATCCG
	gfp Inner F	TCTGGTGAAGATTGTCAGTGCCTACCACGCGCGG
	gfp Outer R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATTG

Table 2.6: List of oligonucleotides used to generate deletion/substitution constructs for insertion into *lag-2* promoter/TOPO-XL backbone. Complementary overlapping region are underlined. *Not*I restriction sites shown in blue.

Target	Primer	Sequence (5'→3')
DSL(ETSD) apx-1::gfp ::unc-54 3'utr	DSL(ETSD) Outer F	GAGCGGCCGCATGATCGCTTACTTCCTCTTACTC
	DSL(ETSD) Inner R	GTGAAGATCAGACGTTTCGATACAATATCGATTGCAG CGTTTTCCG
	gfp Inner F	TATTGTATCGAAACGTCTGATCTTCACTGGGAGTGCTC AACACACG
	gfp Out R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATT G
I80-K86/SS apx-1::gfp ::unc-54 3' utr	ΔI80-K86(SS) Outer F	GAGCGGCCGCATGACTAACTTCTCTTCCCTTCTC
	ΔI80-K86(SS) Inner R	GTTCTTCTCCTTTACTGACTCGAAACGACGACTCGTCA TC
	gfp Inner F	GTCGTCGTTTCGAGTCAGTAAAGGAGAAGAACTTTTC ACTG
	gfp Outer R	GAGCGGCCGCCTAAGGAAACAGTTATGTTTGGTATATT G

2.1.4.4 Assembly of rescue constructs for *mos-1* mediated single copy insertion (mosSCI)

To generate constructs compatible with plasmids most commonly used for mosSCI, sequences were assembled in the pCFJ150 backbone using three-fragment gateway cloning. The general scheme is outlined in Figure 2.2, and primers used for generating att- flanked sequences are listed (Table 2.7). *lag-2p::ligand::gfp::unc-54 3' utr* constructs generated as described above were used as templates to amplify DNA sequences. Both *lag-2p::lag-2::gfp::unc-54 3' utr* and *lag-2p::lag-2 only:: unc-54 3' utr* in pCFJ150 were generated and verified by sequencing.

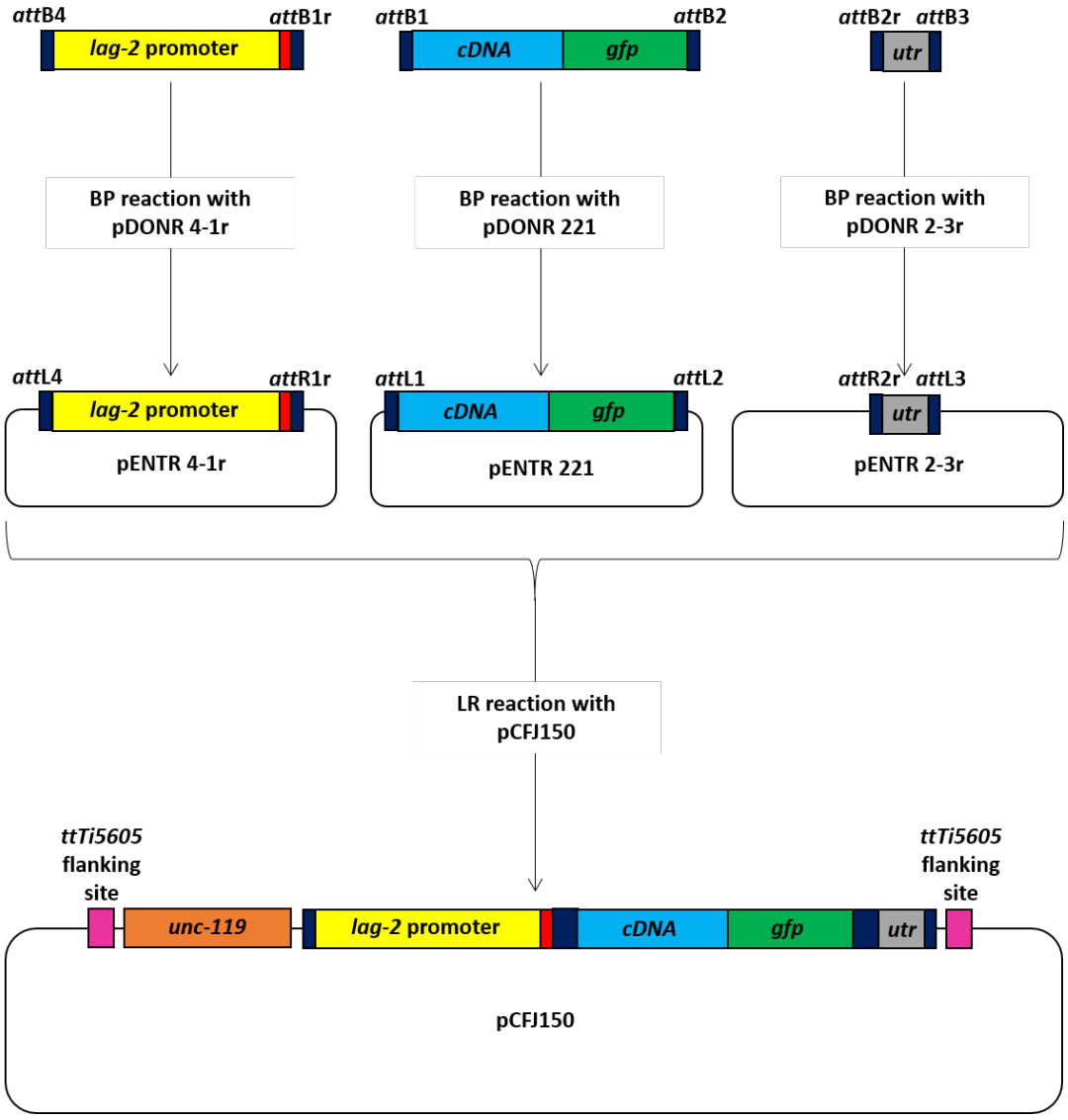


Figure 2.2: Scheme for generation of mosSCI constructs for microinjection - Sequences were amplified to contain appropriate flanking *att*B sites, allowing their recombination with appropriate pDONR plasmids via the BP reaction to give entry clones (pENTR) with compatible *att*L and *att*R sites. These allowed final assembly into the pCFJ150 vector via the LR reaction. pCFJ150 contains a wild-type copy of the *unc-119*$^{+}$ gene from *C. briggsae* and *ttTi5605* flanking sites which are required for *mos-1* mediated insertion of the DNA sequence at the target locus of *ttTi5605* worms.

Table 2.7: List of oligonucleotides used to generate full length ligand cDNA sequences for rescue constructs. All ligand sequences amplified were full length. *att* sites are shown in magenta.

Target	Primer	Sequence (5'→3')
lag-2p	*att*B4 *lag2p* F	GGGGACAACTTTGTATAGAAAAGTTGCGGGCAATTTG CCAAAACTTTGAC
	*att*B1r *lag-2p* R	GGGGACTGCTTTTTTGTACAAACTTGTCTGAAAATGT TCTATGTTATGTTAG
lag-2::gfp	*att*1 *lag-2gfp* F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGATC GCTTACTTCCTCTTACT
	*att*B2 *lag-2gfp* R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATTTG TATAGTTCATCCATG
lag-2 only	*att*B1 *lag-2* F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGATC GCTTACTTCCTCTTACT
	*att*B2 *lag-2* R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGACA TAGTGACAGGCTGG
apx-1::gfp	*att*B1 *apx-1gfp* F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGACT AACTTCTCTTCCCTTC
	*att*B2 *apx-1gfp* R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGAC TCGAAACGACGACTC
apx-1 only	*att*B1 *apx-1* F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGACT AACTTCTCTTCCCTTC
	*att*B2 *apx-1* R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTATTTG TATAGTTCATCCATG
unc-54 3' utr	*att*B2r *utr* F	GGGGACAGCTTTCTTGTACAAAGTGGCGCATTCGTAG AATTCCAACTGAG
	*att*B3 *utr* R	GGGGACAACTTTGTATAATAAAGTTGGAGGAAACAGT TATGTTTGGTATATTGGG

2.2 RNA methods

2.2.1 Semi-quantitative gene expression analysis

HEK293T cells were grown to 80% confluency in T25 flasks containing complete Dulbeccos Modified Eagle Medium (DMEM) (10% FBS, 2 mM L -glutamine, 100 units/ml Penicillin G, 100 μg/ml Streptomycin). Cells were transfected in 10 ml serum free DMEM with 8 μg of DNA and 16 μg of PEI and incubated for 20 hours at 37°C. Cells were then removed from the flask and pelleted at 1000 rpm for 3 minutes. The supernatant was then removed, and the RNA extracted from the pellet using the RNeasy Mini kit (Qiagen) according to the manufacturers' protocol, including the recommended Qiashredder and optional DNAse steps. The RNA was eluted in 50 μl of RNAse free H_2O and the conentration measured using a Nanodrop 1000 spectrophotometer. 1 μg of RNA was added to 2 μl of 10 μM of stock pHLsec reverse oligo (CCGGGAGATCATGAGGGTGTC) in 15 μl H_2O and heated to 70°C for 10 minutes, before chilling for 60 seconds on ice. RNA was then reverse transcribed to generate cDNA using the Omniscript Reverse Transcription kit (Qiagen), according to the manufacturer's instructions. After a 37°C, 1 hour incubation followed by 15 minutes at 70°C, the cDNA samples were chilled on ice before being used in PCRs (as per section 2.1.1.9). 1 μl of cDNA was used per reaction, with appropriate primers (Table 2.1) for amplification. Conditions of denaturation (95°C for 1 minute), annealing (60°C for 1 minute) and extension (68°C for 2 minutes) were repeated for 15, 20 and 25 cycles. 20 μl of PCR product was then loaded onto a 1% agarose gel for analysis.

2.3 Protein methods

2.3.1 Expression and purification of recombinant proteins

2.3.1.1 HEK293T transfection for protein expression

Initially, HEK293T cells were cultured to 90% confluency in T75 flasks containing complete DMEM. Cells were transfected with 25 μg DNA and 50 μg polyethylenimine (PEI) in 15 ml serum free DMEM followed by incubation at 37°C for 3 days. APX-1_{NE4Fc} and APX-1_{DE4Fc} were used for transfection of larger scale cultures produced by transfection of T175 flasks of HEK293T cells with 50 μg DNA and 100 μg PEI in 25 ml serum free DMEM per flask. BD multilayer flasks were used as a space-saving measure for larger scale cultures, according to the manufacturers instructions.

2.3.1.2 S2 transfection for protein expression

Transfection of the APX-1_{NE3His} encoding construct in the pExpres2-2 backbone was conducted by M. Jore using reagents and protocols provided by Expres^{2}ion technologies. Transfected S2 cells were maintained in suspension at 25°C, 200rpm in Excell420 media (Sigma Aldrich) with 10% FBS. Stable cell lines generated in the presence of Geneticin over a period of 26 days after transfection. Stably transfected cultures were scaled up to two 500 ml cultures in Corning roller bottles. APX-1_{NE1His} was cloned (using APX-1_{NE3His} in pExres2-2 as a template) and transfected by R. Suckling.

2.3.1.3 HEK293T nickel purification

Protein constructs were expressed with a hexa-histidine tag allowing their purification using a Ni^{2+} column. The medium from transfected flasks was decanted and

centrifuged (4400 rpm, 10 minutes). The supernatant was then collected and diluted 1:3 in purification buffer (50 mM Tris pH 9.0, 200 mM NaCl). 2 ml of charged Ni^{2+} beads were added, and after shaking (125 rpm, 20°C, 1 hour) the suspension was briefly centrifuged and the supernatant removed. The beads were applied to a gravity column and washed with 30 ml of purification buffer. Protein was then eluted with elution buffer (50 mM Tris pH 9.0, 200 mM NaCl, 300 mM imidiazole), and concentrated using a 5K Vivaspin 500 (Sartorius Stedim) to 500 μl.

2.3.1.4 S2 nickel purification

1 L of supernatant was harvested from cells and passed through a 0.45 μm filter before concentration with a Vivaflow 200 (Sartorius Stedim) and subsequent buffer exchange with purification buffer (50 mM Tris pH 9.0, 200 mM NaCl) to remove the surfactant pluronic-F68, a component of Excell-420 media. A final 50 ml concentrated sample was then loaded onto a 1 ml Ni^{2+} NTA superflow cartridge (Qiagen), and eluted using and AKTA FPLC system in purification buffer containing 200 mM imidiazole. 1 ml fractions were collected during elution, and the AKTA elution profile used to select fractions to pool for concentration and further analysis.

2.3.1.5 Gel filtration

HEK293T and S2 derived proteins were further purified on a Superdex 200 gel filtration column using an AKTA FPLC system. Concentrated nickel purified protein samples were loaded into a 500 μl sample loop and eluted in purification buffer. Fractions were collected corresponding to the appropriate elution peak and concentrated as before. The protein concentration of the final sample was estimated using a Nanodrop Lite 1000 spectrophotometer (Thermo Scientific).

2.3.2 Analysis of recombinant proteins

2.3.2.1 SDS-Polyacrylamide gel electrophoresis (PAGE)

Proteins were analysed using a 12.5% SDS-PAGE resolving gel under non-reducing and reducing conditions. Stocks of non-reducing (200 mM Tris-HCl pH 6.8, 12.5% (w/v) SDS, 40% glycerol, 50 mM EDTA, 0.25% (w/v) bromophenol blue) and reducing (non-reducing with 20% DDT) loading dyes were diluted 1:4 in the sample, and reducing samples were subsequently denatured at 95°C for 5 minutes before loading. Gels were run at 150V for 1 hour in running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS) before washing in H_2O. Gels were then stained using coomassie blue/silver stain solution (10% phosphoric acid, 10% (w/v) ammonium sulfate, 20% Methanol, 0.25% coomassie blue G250) overnight and destained with H_2O the following day. 15μl of supernatant purified by Ni^{2+} beads and 15μl of cell lysate (from 10 ml purification buffer containing 2% SDS, sonicated and passed through a 0.45μm filter) were used for cell lysate analysis.

2.3.2.2 Western blotting

Protein transfer from SDS-PAGE gels to a nitrocellulose membrane was achieved using a Trans Blot SD Semi Dry Transfer Cell (Bio-Rad) over 1 hour at 15V. The nitrocellulose was blocked for 90 minutes using a 5% solution of milk powder in TBS, (50 mM Tris pH 7.5, 150 mM NaCl), washed in TBS, and incubated (4°C, overnight) in a 1:2500 diluted solution of rabbit antibody to the hexa-histidine tag conjugated to horseradish peroxidise (HRP) (Abcam). After washing with TBS and, the blot was developed with ECL Plus Detection reagent (Amersham) and hyperfilm (GE Healthcare).

2.3.2.3 Limited proteolysis

Purified APX-1_{NE1His} was subjected to limited proteolysis in an initial attempt to identify the presence of a calcium binding site in the fragment. In each 50 μl reaction, $CaCl_2$ or EDTA was added to a final concentration of 10 mM and a total of 10 μg of APX-1_{NE1His} was added. At time 0, prior to the addition of enzyme (Trypsin or Endoproteinase Glu-C, with final concentrations of 100 μg/ml and 30 μg/ml, respectively), a 10 μl sample was withdrawn from the reaction and mixed with a 10 μl sample of a 1:2 dilution of loading dye. The 20 μl sample was then boiled at 95°C for 5 minutes. After addition of enzyme, samples were withdrawn 10, 20, 30 and 60 minutes later, and treated as before. Samples were then anaylysed using a 16% SDS PAGE gel.

2.4 *C.elegans* methods

2.4.1 Generation and maintenance of strains

2.4.1.1 Growth and maintenance

Strains were supplied by the *Caenorhabditis* Genetics Centre (CGC) which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). Worms were cultured on NGM agar plates, seeded with *Escherischia coli* OP50, as first described by Sydney Brenner (Brenner, 1974). Strains grown at room temperature were maintained by moving 3-5 L4 hermaphrodites every 3-4 days onto a fresh NGM plate seeded with OP50. Temperature sensitive strains cultured at 15°C in an incubator were maintained by moving 3-5 L4 hermaphrodites every 6-7 days onto fresh plates.

2.4.1.2 Freezing and thawing

Worms were grown on plates to the point where they had begun to reach starvation and a high number of L1/ dauer worms were present. At this stage, 700 μl of M9 buffer and 700 μl of freezing solution (Brenner, 1974) were added to the plate and mixed by swirling. 600 μl aliquots were transferred into 1.8 ml cryovials, and frozen in freezer containers (Nalgene) surrounded by isopropanol to ensure a cooling rate of 1°C/min in a -80°C freezer. Once completely frozen, cryovials were then transferred to liquid nitrogen storage.

Worms were thawed completely from liquid nitrogen storage at room temperature, and the contents of the cryovial was transferred to a fresh plate seeded with OP50. After 3-4 days, gravid hermaphrodites were picked off this plate and into a 10 μl spot of 2x bleach solution (4 ml NaClO (4%), 2.5 ml 4M NaOH, 3.5 ml H_20 for 10 ml stock) on a fresh plate, to remove contamination from the freezing/thawing process.

Table 2.8: New strains generated by genetic crossing Starting strains and sexes used are shown above. Two L4 hermaphrodites were placed on a fresh plate with 7-9 young adult males. Multiple plates were set up in parallel, and heterzygotes from the F1 generation were picked when the population contained $\sim$ 25-50% males.

Hermaphrodites	Males	Final strain
lag-2(q420)V	*him-8(e1489)IV*	*him-8(e1489)IV; lag-2(q420)V*
unc-119(ed3)III	*him-8(e1489)IV; lag-2(q420)V*	*unc-119(ed3)III; lag-2(q420)V*
cdh-3>mCherry ::PLC$^{\delta PH}$II?	*him-8(e1489)IV; slag-2(q420)V*	*cdh-3>mCherry ::PLC$^{\delta PH}$II?; lag-2(q420)V*

2.4.1.3 Crosses

Several genetic crosses were carried out to introduce markers into the *lag-2q(420)* background. The temperature sensitivity of the *lag-2(q420)* allele prevented the generation of males by heat shock, therefore and initial cross was carried out to generate male *lag-2(q420)* homozygotes that could be used for subsequent crosses. Strains generated by crossing are listed (Table 2.8). All crosses involving the *lag-2(q420)* allele were performed at 15°C, with approximately 6 days between generations, whilst room temperature crosses had an interval of 3 days between generations. Successful crossing of *unc-119(ed3)III, him-8(e1489) IV* and *cdh-3>mCherry::PLC$^{\delta\ PH}$II?* alleles could be confirmed phenotypically using a light/fluorescence microscope. To confirm *lag-2(q420)* homozygosity, genotyping by single worm lysis was performed.

2.4.1.4 Genotyping by single worm lysis

Single worm lysis was performed to extract genomic DNA from worms that could subsequently be used in a genotyping PCR reaction. A single hermaphrodite was

allowed to lay eggs on a plate before it was placed in 2.5μl of worm lysis buffer (50 μM KCl, 2.5 μM $MgCl_2$,10 μM Tris-HCl pH 8.3, 0.45% NP40, 0.45% Tween20, 0.01% gelatine with 1 μl (10 mg/ml) proteinase K per 100 μl solution) in a thin walled PCR tube, and placed at -80°C to crack cells. The tubes were then transferred to a thermocycler and lysed at 60°C for 1 hour, followed by 95°C for 15 minutes. PCR was then performed using Thermopol Taq polymerase using primers flanking the *lag-2(q420)* allele. WT, *lag-2(q420)* homozygotes and heterozygotes were identified by digesting PCR products directly with 1 μl of *Kas*I and 5 μl of cutsmart buffer (NEB) and running the products on a 2% agarose gel (Figure 2.3).

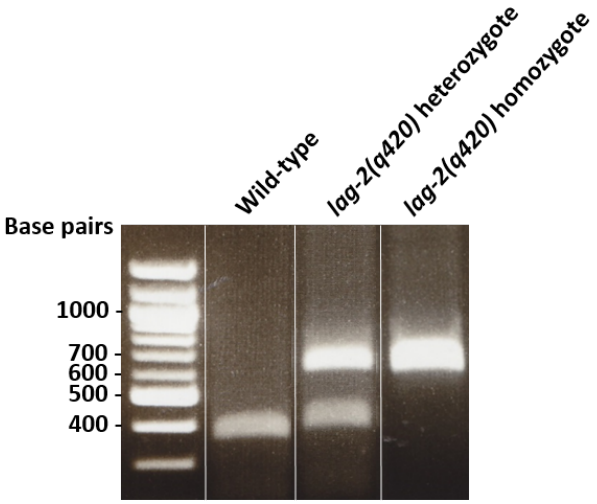


Figure 2.3: Example of digestion products for identification of the *lag-2(q420)* allele - In WT worms, *Kas*I digestion led to complete cleavage of the 586 bp PCR product, producing a 395 bp fragment and a 248 bp fragment (generally not visable on the gel). Heterzygotes were associated with two gel bands, with both full size and digestion products present. *lag-2(q420)* homozygotes remained undigested and full length.

2.4.2 Generation of transgenic lines

2.4.2.1 Generation of semi-stable extrachromosomal array transgenes

Semi-stable, heritable extrachromosomal arrays were generated by microinjection of plasmid DNA (Mello et al., 1991). Transformation markers used included 20 ng/μl *unc-119*$^{+}$(pDP#MM016β) (Maduro and Pilgrim, 1995) and 40 ng/μl *sur-5::gfp* (pTG96) injected at 40 ng/μl. Temperature sensitive worms were separated out onto fresh plates prior to injections, with 6-8 worms per plate. Plates were kept at 15°C until immediately before their injection, and were not at room temperature for more that 30 minutes. Young adults with a fully formed vulva were injected to ensure temperature shifting would not induce a protruding vulva.

All extrachromosomal transgenics referred to in this work have been defined in full (see Tables 2.9,2.10 and 2.11), and given an abbreviation that will be referred to in subsequent chapters. All Notch ligand constructs were injected at 2 ng/μl, except for diluted *lag-2::gfp* which was injected at 1 ng/μl, diluted in 100 ng/μl Bluescript plasmid DNA. In addition, to successfully generate *cdh-3>mCherry::PLC*$^{\delta\ PH}$*II?; lag-2(q420)V [lag-2p::lag-2::gfp::unc-54 3' utr]* lines, the concentration was lowered to 1 ng/μl. 2-3 independent lines were generated for each construct.

Table 2.9: Transgenic lines generated using the unc-119$^+$ transformation marker, and their abbreviations.

Strain	Array	Abbreviation
unc-119(ed3)III	*[unc-119$^+$]*	*unc-119* only rescued
unc-119(ed3)III	*[lag-2p:: lag-2:: gfp:: unc-54 3' utr]*	*unc-119[lag-2::gfp]*
unc-119(ed3)III; lag-2(q420)V	*[unc-119$^+$]*	*unc-119* only *(lag-2(q420))*
unc-119(ed3)III; lag-2(q420)V	[*WRM0620cE03 + unc-119$^+$*]	*lag-2* fosmid
unc-119(ed3)III; lag-2(q420)V	*[lag-2p:: lag-2:: gfp:: unc-54 3' utr + unc-119$^+$]*	*lag-2::gfp*
unc-119(ed3)III; lag-2(q420)V	*[lag-2p:: lag-2:: unc-54 3' utr + unc-119$^+$]*	*lag-2* only
unc-119(ed3)III; lag-2(q420)V	*[lag-2p:: lag-2:: gfp:: unc-54 3' utr + Bluescript + unc-119$^+$]*	diluted *lag-2::gfp*
unc-119(ed3)III; lag-2(q420)V	*[lag-2p:: dsl-1:: gfp:: unc-54 3' utr + unc-119$^+$]*	*dsl-1::gfp*
unc-119(ed3)III; lag-2(q420)V	*[lag-2p:: apx-1:: gfp:: unc-54 3' utr + unc-119$^+$]*	*apx-1::gfp*
unc-119(ed3)III; lag-2(q420)V	*[rlag-2p:: lag-2:: gfp:: unc-54 3' utr + unc-119$^+$]*	*rlag-2p:: lag-2::gfp*

Table 2.10: Transgenic lines generated using the unc-119$^+$ transformation marker, and their abbreviations.

Strain	Array	Abbreviation
unc-119(ed3)III	*[lag-2p::ΔMNNL apx-1::gfp::unc-54 3' utr + unc-119$^+$]*	*unc-119 [ΔMNNL apx-1::gfp]*
unc-119(ed3)III	*[lag-2p::ΔDSL apx-1::gfp::unc-54 3' utr + unc-119$^+$]*	*unc-119 [ΔDSL apx-1::gfp]*
unc-119(ed3)III	*[lag-2p::ΔEGF-like 1 apx-1::gfp::unc-54 3' utr + unc-119$^+$]*	*unc-119 [ΔEGF-like 1 apx-1::gfp]*
unc-119(ed3)III	*[lag-2p::DSL(ETSD) apx-1::gfp::unc-54 3' utr + unc-119$^+$]*	*unc-119 [ΔDSL(ETSD) apx-1::gfp]*
unc-119(ed3)III	*[lag-2p:: I80-K86/SS apx-1::gfp::unc-54 3' utr + unc-119$^+$]*	*unc-119 [I80-K86/SS apx-1::gfp]*

Table 2.11: Transgenic lines generated with the *sur-5::gfp* transformation marker, and their abbreviations.

Strain	Array	Abbreviation
N2	*[sur-5::gfp]*	WT*[sur-5::gfp]*
N2	*[lag-2p::lag-2::gfp::unc-54 3' utr + sur-5::gfp]*	*WT[lag-2::gfp]*
cdh-3>mCherry ::PLC$^{\delta PH}$II?;lag-2(q420)V	*[lag-2p::lag-2::gfp::unc-54 3' utr + sur-5::gfp]*	*lag-2::gfp*(AC)
cdh-3>mCherry ::PLC$^{\delta PH}$II?;lag-2(q420)V	*[lag-2p::apx-1::gfp::unc-54 3' utr + sur-5::gfp]*	*apx-1::gfp*(AC)
cdh-3>mCherry ::PLC$^{\delta PH}$II?;lag-2(q420)V	*[lag-2p::ΔMNNL apx-1::gfp::unc-54 3' utr + sur-5::gfp]*	*ΔMNNL apx-1::gfp*(AC)
cdh-3>mCherry ::PLC$^{\delta PH}$II?;lag-2(q420)V	*[lag-2p::ΔDSL apx-1::gfp::unc-54 3' utr + sur-5::gfp]*	*ΔDSL apx-1::gfp*(AC)
cdh-3>mCherry ::PLC$^{\delta PH}$II?;lag-2(q420)V	*[lag-2p::ΔEGF-like 1 apx-1::gfp::unc-54 3' utr + sur-5::gfp]*	*ΔEGF-like 1 apx-1::gfp*(AC)
cdh-3>mCherry ::PLC$^{\delta PH}$II?;lag-2(q420)V	*[lag-2p::DSL(ETSD)apx-1::gfp::unc-54 3' utr + sur-5::gfp]*	*DSL(ETSD) apx-1::gfp*(AC)
cdh-3>mCherry ::PLC$^{\delta PH}$II?;lag-2(q420)V	*[lag-2p::I80-K86/SS apx-1::gfp::unc-54 3' utr + sur-5::gfp]*	*ΔI80-K86(SS) apx-1::gfp*(AC)

2.4.2.2 Generation of mos-1 mediated single copy insertion transgenes (mosSCI)

The *lag-2p::lag-2::gfp::unc-54 3' utr* in pCFJ150 construct was initially microinjected at 10 ng/μl into *ttTi5605II; unc-119(ed3)III* worms (referred to as *EG6699*, with *oxEx1578* for maintenance of the strain), which was incrementally reduced to 1 ng/μl subsequently, along with transformation markers, according to the detailed protocol outlined by Christian Frokjaer-Jensen (www.wormbuilder.org). *EG6699* worms were cultured at 15°C using an HB101 lawn as a alternative to OP50, to reduce the likelihood of starvation. After injection, a maximum of 3 worms were placed on a fresh HB101 plate and incubated at 25°C until worms had starved. Plates were then heat shocked at 34°C for 2 hours to activate the PEEL-1 toxin, and stored at room temperature for a further 6 hours, before being screened for non-red-movers (NRM) using a dissecting fluorescence microscope (Leica).

NRMs were genotyped for insertions. Initially, single worm lysis was performed on non-red movers to isolate genomic DNA, before primers based on the Nonet lab protocol (http://thalamus.wustl.edu/nonetlab/ResourcesF/PCR %20of%20MosSCI %20transgenes.pdf) were used to amplify a region spanning the flanking region of the *ttTi5605II* locus, and the *unc-119*$^{+}$ gene associated with insertion of a pCFJ150 based DNA sequence (Table 2.12). An insertion could be detected by the presence of a 1.8 kbp PCR product absent in worms lacking an insertion.

To verify the insertion sequence, a single L4 hermaphrodite was moved to a fresh plate to lay eggs, and after 3 days the progeny were collected by washing with 1 ml of M9 buffer, and pelleted at 1400 rpm for 90 seconds. The supernatant was removed and the worms re-suspended on 1 ml of M9 buffer, before pelleting again, to remove excess bacteria. Genomic DNA was then extracted from the worms using a peqGOLD Tissue DNA Kit Mini (Peqlab), according to the manufacturers' instruc-

Table 2.12: List of oligonucleotides used to verify insertions at the *ttTi5605 mos-1* site. To verify full insertions, primers flanking the *ttTi5605II* site were used. Two reactions were performed to confirm accurate product size, both with the same *ttTi5605* left2 F forward primer, but with reaction 2 using a nested reverse primer, still right of the *mos-1* insertion site (*ttTi5605*right2).

Target	Primer	Sequence (5'→3')
unc-119 insertion	*ttTi5605*left1 F	TCTGGCTCTGCTTCTTCGTTGTAC
	unc-119 R	GCCAATTCATCCCGGTTTCTGTC
Full insertion (1)	*ttTi5605*left2 F	GAGGCAGAATGTGAACAAGACTCGAG
	*ttTi5605*right1 R	CACGCCCAGGAGAACACGTTAG
Full insertion (2)	*ttTi5605*left2 F	GAGGCAGAATGTGAACAAGACTCGAG
	*ttTi5605*right2 R	ATCGGGAGGCGAACCTAACTGTA

tions. After eluting the extracted genomic DNA in 200 μl of elution buffer, this was further concentrated to 20 μl using the protocol supplied. 10 μl of this product was then used per PCR. PCRs were performed using Primestar GXL polymerase (Clontech) using reagents and conditions exactly as recommended by the manufacturer. Primers flanking the mos-1 insertion sites of the ttTi5605 II locus were used (Table 2.12) to amplify the entire insertion region. PCR products were then run on a 0.7% agarose gel, and where required, excised and purified using a Qiaquick gel extraction kit, before being sent for DNA sequence analysis (Sourcebioscience).

2.4.3 *C. elegans* imaging and assays

2.4.3.1 Imaging

Worms to be imaged were mounted one of two ways. Small numbers of worms were picked directly into a 5 μl droplet of M9 buffer containing 0.5% 1-phenoxy-2-propanol (PP) on a glass coverslip, whilst for larger populations, worms were washed off a

plate in 1 ml M9 buffer containing 0.5% PP. The 1 ml suspension of worms was then pelleted by centrifuging at 1600 rpm for 90 seconds, and the excess supernatant removed. The final 10-15 μl pellet of worms was re-suspended mounted in 5 μl aliquots M9 onto a glass coverslip. The mounted worms were then place on an agarose pad (0.4% w/v agarose in M9 with 0.5% PP). Microscopy was performed using a 63x oil immersion objective lens of a fluoresence microscope (Zeiss). Different filters were used to collect images, including the Nomarski (DIC) filter (to take non- fluorescent images), as well as the FITC FS10 (GFP images), Rhodamine FS15 (mCherry images) and DAPI FS02 (DAPI images) filters. Exposures were varied according to intensity, since fluorescence was not compared quantitatively between images. Images were gathered, stored and exported using Axiovision software (Release 4.8).

2.4.3.2 Protruding vulva assay

A temperature shift assay was performed to calculate the penetrance of the protruding vulva phenoytype of *lag-2(q420)* worms compared with WT. At time 0, ∼50 gravid hermaphrodites were picked into 500 μl of M9 buffer with 0.5% Tween20. The worms were pelleted at 1400 rpm for 90 seconds, and the supernatant removed. The worms were then washed and pelleted from 500 μl of the same buffer 5 further times, until a small worm pellet of ∼50 μl remained. This was then rotated whilst incubated at 15°C for 18 hours, allowing progeny to hatch and arrest at the end of L1 due to a lack of food. The worm pellet was then pipetted onto a fresh plate, and adult hermaphrodites removed. The plate was placed at 25° for 40 hours, and then worms were scored for the absence or presence of a protruding vulva using a light microscope. Any visible protrusion that could be identified using a light microscrope (Nikon) was scored as a PVul.

2.4.3.3 DAPI staining

To visualise nuclei in the gonad, DAPI staining was performed. 10μl of M9 was pipetted into a clean watch glass, and 30-50 worms were picked into the drop. 200 μl of a 150 nM DAPI solution (in 100% ethanol) was added to the glass, and incubated at room temperature in a dark chamber for 1 hour. Worms were then rehydrated in 2 ml of M9 for 40 minutes, before the worms were transferred to a microcentrifuge tube with a glass pipette, and pelleted at 1400 rpm for 90 seconds. The supernatant was removed, and the remaining 10-20 μl mounted in 3 μl aliquots with 2 μl AF3 PBS antifade solution (Citifluor) for fluorescence microscopy.

2.4.3.4 Transgene transmission efficiency assay

To calculate the efficiency with which semi-stable extrachromosomal arrays were inherited, single fourth-larval stage (L4) hermaphrodites from lines grown at 15°C were placed on fresh plates at time 0. 48 hours later (day 3) and similarly on day 5 and 7, the hermaphrodite was moved to a fresh plate, and the progeny incubated at 15°C for a further 6 days until they had reached early adulthood, but before they had begun to lay eggs of their own. At this point, the number of worms containing the transgene (e.g *unc-119*$^{+}$) and those that did not were counted. The total number of transgenic and non-transgenic worms from days 1, 3, 5, 7 were added and the transmission efficiency calculated by expressing the total number of transgenics out of the total number of worms in the brood as a percentage. Results from worms that had either died, developed a protruding vulva, or silenced the transgene during the experiment were discarded when calculating a final average of efficiencies from multiple experiments.

2.4.3.5 Larval viability assay

To calculate the number of worms surviving beyond the first larval stage (L1), single fourth-larval stage (L4) hermaphrodites from lines grown at 15°C were placed on fresh plates at time 0, and incubated at 25°C. Worms were moved 24 and 48 hours later to fresh plates. A further 30 hours after worms were moved off the plate, the number of viable transgenic worms that were at the second larval stage (L2) or older were counted using a light microscope. To compare results from different lines, the results were subsequently normalised according to transmission efficiency as shown in Appendix 3 and 4.

2.4.3.6 Brood fate assay

To calculate the fates of the progeny from single worms, including the identification of dead embryos and L1 larvae displaying the Lag phenotype, it was necessary to prepare fresh OP50 seeded plates, to allow these specific fates to be followed clearly. A solution of OP50 was diluted to $O.D_{600}<0.1$, and a thin lawn spread on fresh NGM agar plates immediately prior to using them. Once the OP50 solution was dried, from lines grown at 15°C were placed on these plates at t=0. Worms were moved 24 and 48 hours later to fresh plates. The total number of embryos and larvae were counted at these point, followed by a second count a further 24 hours later. This allowed the fates of eggs laid on these plates to be followed, and at the end of the assay, the results from the 3 days were added together to give total numbers of dead embryos, and hatched larvae (both Lag and non-Lag). These fates were then expressed as a percentage of the total population, which was the sum of all fates recorded (Appendix 5).

2.4.3.7 Anchor cell duplication assay

Anchor cell duplication assays were performed across 5 days. At time 0, ∼100 gravid hermaphrodite worms grown at 15°C from a single strain/line were picked into 10μl spots of bleach on a fresh plate. The plate was then placed back at 15°C for 48 hours, when the progeny had hatched and reached the end of L1. At this stage, the plate was shifted to 25°C for 39 hours. The plate was then removed, and the worms from the plate washed off in M9 buffer containing 0.5% PP into a microcentrifuge tube. The worms were then pelleted by centrifuging at 1400 rpm for 90 seconds, and the excess supernatant removed. The final 10-15 μl pellet of worms was re-suspended and mounted for fluorescence microscopy allowing the frequency of single and duplicated anchor cells to be counted.

3

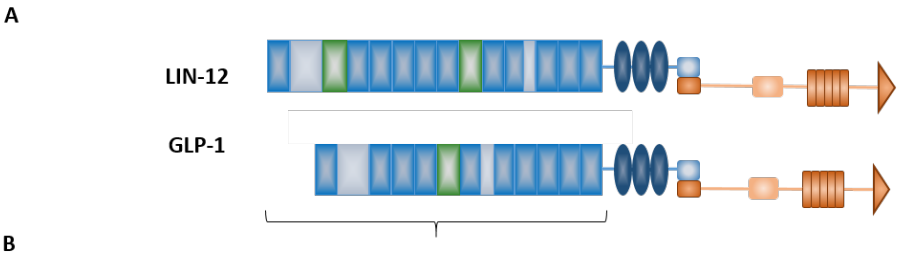
Expression and characterisation of *C. elegans* Notch proteins

3.1 Sequence analysis of *C. elegans* Notch proteins

3.1.1 Analysis of *C. elegans* Notch receptors

Whilst the major-ligand-binding domain (EGF-like 11-12) of dNotch and human Notch receptors has been well defined, the equivalent region has not yet been located in *C. elegans* receptors LIN-12 and GLP-1. Multiple sequence alignments were therefore carried out to search for homology between EGF-like regions of LIN-12 and GLP-1, and to compare their sequences with the known major ligand-binding domains of other species.

LIN-12 contains one EGF-like domain, followed by a region of 91 amino acids containing 14 cysteine residues (termed an EGF-derived region), a calcium-binding EGF-like domain, four EGF-like domains, followed by another calcium-binding domain. Within the remaining 5 EGF-like domains there is a region containing a pair of cysteine residues between EGF-like 8 and 9 (Figure 3.1A). In GLP-1, there is an



B

```
LIN-12(1-13)  MRIPTICFLFLLISLSKSLHIGSCLGLICGRNGHCHAGP---VNGTQTSYWCRCDEGFGG
GLP-1(1-11)   MRVLLILLAFFA-PI-----ASQLMGGECGREGACSVNGKCYNGKLIETYWCRCKKGFGG

LIN-12(1-13)  EYCEQQCDVSKCGADEKCVFDKDYRMETCVCKDCDINGNSL--LKPSCPSGYGGDDCKTQ
GLP-1(1-11)   AFCERECDLD-CKRGEKCIYDVYGENPTCICQDCEDETPPTERTQKGCEEGYGGPDCKT-

LIN-12(1-13)  GWCYPSVCMNGGQCIGAGNRAKCACPDGFKGERCELDVNECEENKNACGNRSTCMNTLGT
GLP-1(1-11)   ------------------------------------------------------------

LIN-12(1-13)  YICVCPQGFLPPDCLKPGNTSTVEFKQPVCFLEISADHPDGRSMYCQNGGFCDKASSKCQ
GLP-1(1-11)   ------------------------------------------------------------

LIN-12(1-13)  CPPGYHGSTCELLEKEDSCASNPCSHGVCISFSGGFQCICDDGYSGSYCQEGKDNCVNNK
GLP-1(1-11)   ----------PLFSGVNPCDSDPCNNGLCYPFYGGFQCICNNGYGGSYCEEGIDHCAQNE

LIN-12(1-13)  CEAGSKCINGVNSYFCDCPPERTGPYCEKMDCSAIPDICNHGTCIDSPLSEKAFECQCEP
GLP-1(1-11)   CAEGSTCVNSVYNYYCDCPIGKSGRYCERTECALMGNICNHGRCIPNRDEDKNFRCVCDS

LIN-12(1-13)  GYEGILCEQDKNECLSENMCLNNGTCVNLPGSFRCDCARGFGGKWCDEPLNMCQDFHCEN
GLP-1(1-11)   GYEGEFCNKDKNECLIEETCVNNSTCFNLHGDFTCTCKPGYAGKYCEEAIDMCKDYVCQN

LIN-12(1-13)  DGTCMHTSDHSPVCQCKNGFIGKRCEKECPIGFGGVRCDLRLEIGICSRQGGKCFNGGKC
GLP-1(1-11)   DGYCAHDSNQMPICYCEQGFTGQRCEIECPSGFGGIHCDLPLQRPHCSRSNGTCYNDGRC

LIN-12(1-13)  LSGFCVCPPDFTGNQCEVNRKNGKSSLSENLCLSDPCMNNATCIDVDAHIGYACICKQGF
GLP-1(1-11)   INGFCVCEPDYIGDRCEINRKDFKFP-DIQSCKYNPCVNNATCIDLK-NSGYSCHCPLGF

LIN-12(1-13)  EGDICERHKDLCLENPCSNGGVCHQHRESFSCDCPPGFYGNGCEQEKMFRCLKSTCQNGG
GLP-1(1-11)   YGLNCEQH-LLCTPTTCANGGTCEGVNGVIRCNCPNGFSGDYCEIKDRQLCSRHPCKNGG

LIN-12(1-13)  VCINEEEKGRKCECSYGFSGARCE
GLP-1(1-11)   VCKNT----GYCECQYGYTGPTCE
```

Figure 3.1: Primary sequence alignment of LIN-12 and GLP-1 EGF-like regions - (**A**) The architecture of LIN-12 and GLP-1 is generally well conserved, with and additional three EGF-like repeats (one of which binds calcium) present in LIN-12. Domains are coloured as in Figure 1.4. (**B**) Sequences are generally well conserved from EGF-like 5-13 of LIN-12 and 2-10 of GLP-1, with absolutely conserved residues highlighted in red. Regions predicted to align with major ligand-binding domains of human Notch receptors and dNotch are underlined.

EGF-like domain followed by an EGF-derived region containing 8 cysteine residues. The subsequent EGF-like repeats of GLP-1 appear highly conserved with EGF-like 5-13 of LIN-12 (Figure 3.1B), with a high level of sequence conservation supporting the proposal that *lin-12* and *glp-1* arose by gene duplication (Lambie and Kimble, 1991).

Prior to alignment of *C. elegans* Notch receptors to their human and *Drosophila* counterparts, EGF-like 11-12 of hNotch-1 and *Drosphila* Notch was compared to look for conservation of residues confirmed to be key participants in the Notch receptor/ligand interaction (section 1.6.3) (Figure 3.2). Although several key residues involved in the interaction were actually identified in rNotch-1, hNotch-1 EGF-like 11-12 varies from rNotch-1 by only two residues, neither of which are predicted to be involved in the interaction, therefore discussion of hNotch-1 will be representative of both homologues.

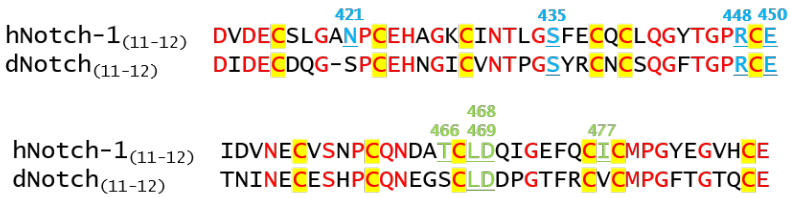


Figure 3.2: Conservation of residues in the major-ligand-binding region of hNotch-1 and dNotch - Alignment of EGF-like 11-12 of hNotch-1 and dNotch, with conserved residues highlighted in red, and cysteine residues that define the EGF-like fold highlighted in yellow. Key residues identified in hNotch-1 to participate in ligand binding are shown in blue (EGF-like 11) and green (EGF-like-12). Note S435 and T466 are both glycosylated, and sugar moieties mediate ligand interactions.

Amongst the multiple residues in EGF-like 11 of hNotch-1 predicted to contribute to interaction with the DSL domain of ligands (specifically, DLL-4), several key residues including N421, R448 and E450 forming H-bonds with ligand residues were identified (Luca et al., 2015). Of these residues, R448 and E450 are both conserved in dNotch, but not N421. In addition, the S435 which contributes an O-linked glycan to

the interaction is conserved in dNotch. In EGF-like 12, T466 is not fully conserved, however the serine residue at this site is still amenable to fucosylation. Whilst residues L468 and D469, which are involved in MNNL binding are also present in dNotch, the interaction residue I477 is replaced with valine in dNotch. The important conclusion to draw from this comparison is that even between human and *Drosophila* Notch receptors, which have shown so many similarities in biochemical studies, there are differences at the primary sequence level, even in key residues within the major ligand-binding region.

LIN-12$_{(8\text{-}9)}$ and GLP-1$_{(5\text{-}6)}$ aligned with the major ligand-binding region of both *Drosophila* Notch and human Notch-1 due to the presence of the calcium-binding EGF consensus motif (Figure 3.3). Individual amino acids are generally not widely conserved between the sequences, and those that are tend to be required to maintain the EGF-like fold (e.g cysteine and glycine). Of the key residues discussed above, N221 and S435 of EGF-like 11 and T466 of EGF-like 12 are conserved in LIN-12 EGF-like 8 and 9, respectively, supporting that this region within LIN-12 may be involved in ligand binding. On the other hand key residues do not appear to be conserved within GLP-1, except for E450. An additional consideration is that human and *Drosophila* Notch receptors form their major ligand-binding region across two tandem calcium-binding EGF-like repeats, which confers greater rigidity in the region, although EGF-like repeats can adopt a rod-like structure in the absence of calcium (Hambleton et al., 2004). There are no tandem calcium binding EGF-like repeats in LIN-12 and GLP-1, therefore the nature of the receptor/ligand interaction is likely to differ to some extent to other species.

Taken as a whole, alignment of the human Notch-1 and *Drosphila* Notch major ligand-binding regions with LIN-12 and GLP-1 did not allow clear definition of this site on the basis of sequence homology alone, although a comparison between human

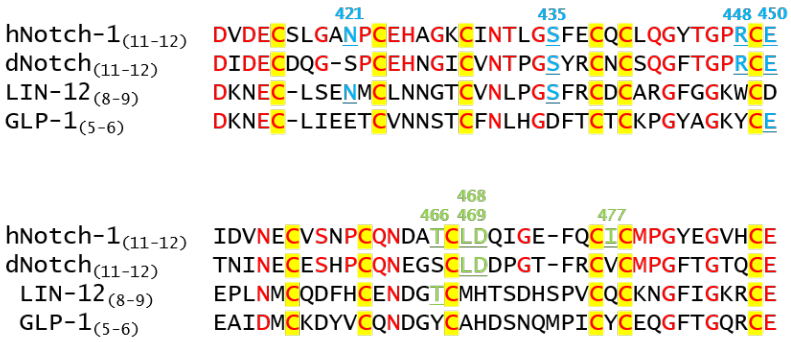


Figure 3.3: Primary sequence alignment of LIN-12 and GLP-1 EGF-like regions with human and *Drosophila* Notch major ligand-binding domains - LIN-12 EGF-like 8-9 and GLP-1 EGF-like 5-6 align with dNotch and human Notch-1 EGF-like 11-12 when inputted to the multiple sequence alignment server Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). Conserved residues are highlighted in red, and cysteine residues that define the EGF-like fold highlighted in yellow. Key residues identified in hNotch-1 to participate in ligand binding are shown in blue (EGF-like 11) and green (EGF-like 12).

and *Drosophila* EGF-like 11-12 highlighted that differences at the primary sequence level are to be anticipated between species. Conservation of some key residues involved in ligand engagement, particularly in LIN-12 was interesting, and may serve as a starting point for exploring the functionality of this receptor further. Since functional redundancy of *C. elegans* Notch proteins LIN-12 and GLP-1 has been observed in several developmental processes (Lambie and Kimble, 1991) (Fitzgerald et al., 1993), it is highly likely they share a common mechanism. There are currently no data reporting the relative ligand-binding affinities of these two receptors, although differences at the sequence level may be reconciled by this.

3.1.2 Analysis of the DSL domain of *C. elegans* Notch ligands

The presence of a DSL domain defines canonical Notch ligands, and this allowed the soluble *C. elegans* Notch ligands DSL-1 to DSL-7 to be identified computationally

(Chen and Greenwald, 2004). Whilst human and *Drosophila* ligands share a well conserved series of residues that contribute to the receptor interaction (section 1.6.1) (Figure 3.4), alignment of the 10 canonical *C. elegans* Notch ligands (Figure 3.5) indicated the lack of a strictly conserved motif. Interestingly, *in vivo* evidence demonstrated the ability for LAG-2, APX-1 and DSL-1 to function interchangeably as ligands for LIN-12 (Gao and Kimble, 1995) (Fitzgerald and Greenwald, 1995) (Chen and Greenwald, 2004), suggesting a shared mechanism, despite the lack of sequence conservation.



Figure 3.4: Primary sequence alignment of human and *Drosophila* Notch ligand DSL domains - Key residues forming the interface with Notch receptors are highlighted in blue, and those particularly in the C_3-C_4 loop are generally well conserved among ligands, excepting human DLL-3. The key phenylalanine residue (207 in hJag-1, 195 in rat and hDLL-4, 257 in dSerrate), which acts as a 'linchpin' for the ligand/receptor interaction is labelled. Other conserved residues are shown in red, cysteines are highlighted in yellow. Adapted from Cordle et al., 2008a.

Comparison of the DSL domains of *C. elegans* Notch ligands and their human and *Drosophila* counterparts allowed some similarities to be identified. The residue prior to the third cysteine residue (C_3) of the DSL domain is an aromatic residue, as in most human ligands, however, it was previously indicated this may be of structural significance in the case of Jagged-1 (Cordle et al., 2008a). The presence of some aspartate and arginine residues within the subsequent C_3-C_4 loop of the DSL



Figure 3.5: Primary sequence alignment of *C. elegans* Notch ligand DSL domains - Residues between the third and fourth cysteine residues of the motif (C_3-C_4 loop) are variable, although some contain arginine, aspartates and phenylalanines, which have been shown to contribute to human and *Drosophila* Notch receptor/ligand interactions. Conserved residues are shown in red, cysteines are highlighted in yellow.

domain was also conserved between humans and *C. elegans*, however these residues occur at different positions within the loop. A conserved phenylalanine within the human (F207 in Jagged-1, F195 in DLL-4) and *Drosophila* (F257 in dSerrate) C_3-C_4 loop which is essential for receptor/ligand binding was not present in any of the *C. elegans* ligands at this position. These variations within the DSL region of *C. elegans* were consistent with the variability of the Notch receptor sequences, and suggests that whilst the architectures of *C. elegans* Notch proteins are similar to human and *Drosophila*, the precise molecular mechanisms of the receptor/ligand interaction will vary. Structural information about the *C. elegans* proteins may reconcile these differences in primary sequences at a tertiary level. The differences in sequence lengths of the MNNL between species was an additional limitation to identification of residues on the basis of primary sequence alignment. The C2 of Jagged-1 spans 152 residues, whilst the C2 of Serrate comprises 122. Superposition of the recently obtained structure of a fragment of Serrate (Figure 3.7) indicated that the β strands are well conserved with the C2 domain of Jagged-1, however the loop regions (specifically β1-2 and β5-6) varied. *C. elegans* MNNL regions are shorter than the Serrate C2, spanning 103 (DSL-1), 105 (LAG-2) and 109 (APX-1)

residues. Whilst the lengths of these regions are sufficiently long to accommodate the multiple β strands of a C2 fold, it is likely that inter-strand loop regions would be shorter than those found in human ligands.

3.1.3 Analysis of the MNNL region of *C. elegans* Notch ligands

The presence of a C2 domain at the MNNL of human and *Drosophila* Notch ligands remained unknown until the crystal structure of Jagged-1 was solved, as there was little primary sequence conservation between ligands to aid the prediction of a common structure (Chillakuri et al., 2013) . However, additional crystallography studies allowed the presence of a C2 domain to be confirmed at the N-terminus of all human Notch ligands, as well as *Drosphila* Serrate, showing a high level of structural conservation (Luca et al., 2015), (Kershaw et al., 2015), (R.Suckling, personal communication). Alignment of the primary sequences of the N-terminus of human Notch ligands highlighted little sequence similarity (Chillakuri et al., 2013) and when *C. elegans* the MNNL regions of Notch ligands are were aligned with these sequences (Figure 3.6), no further common residues could be identified.

An alternative method of sequence alignment was used to identify regions of similarity between human and *C. elegans* MNNL domains using the Fold and Function Alignment System (FFAS) (Jaroszewski et al., 2005) (Xu et al., 2014). This server allows a multiple sequence alignment to be generated by PSI-BLAST, using an input sequence and a specified database (PDB), and a subsequent 3D model of the alignment using the coordinates of the database structure to be generated. The three full length *C. elegans* Notch ligands LAG-2, DSL-1 and APX-1 were used as input sequences into the FFAS server. Shorter fragments including the MNNL only, and MNNL-DSL were also inputted, but this failed to generate models using

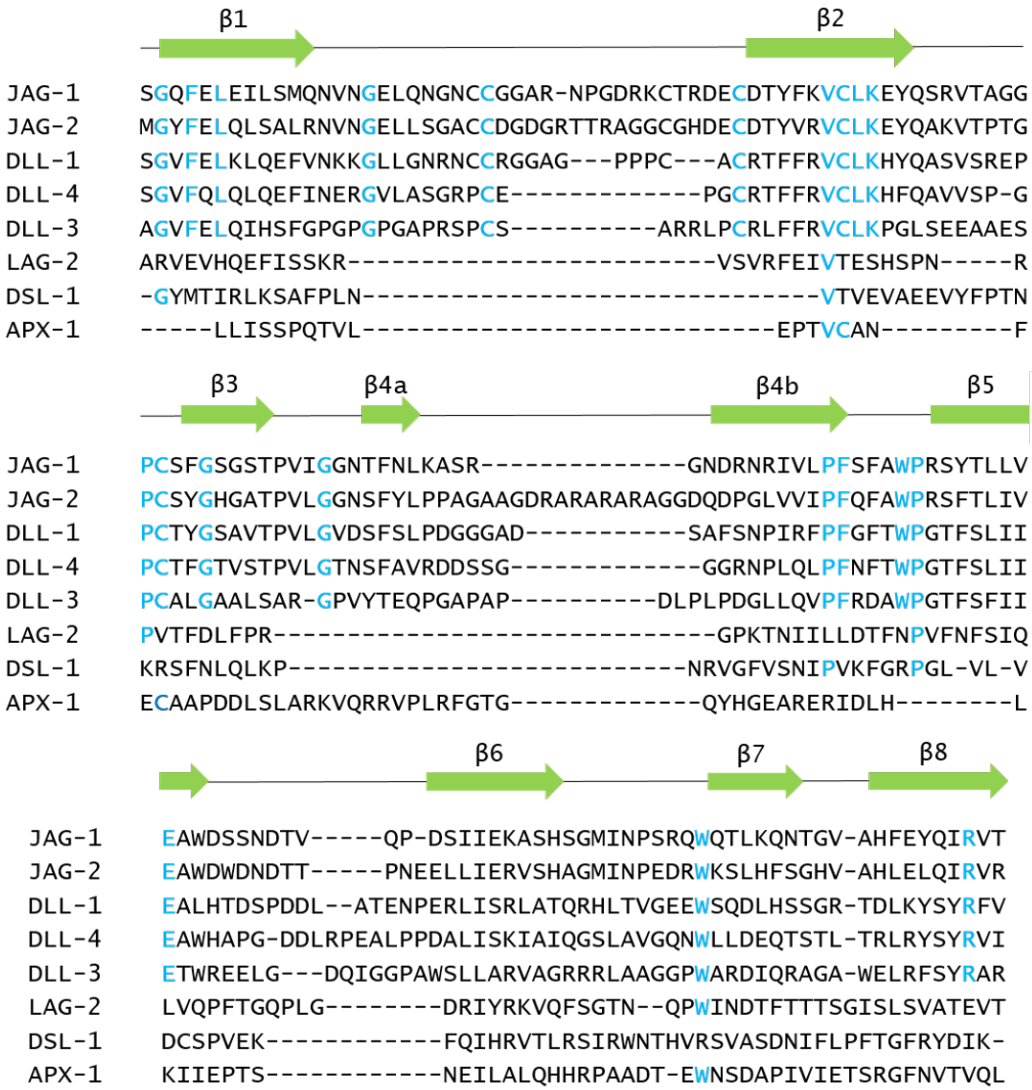


Figure 3.6: Primary sequence alignment of human and *C. elegans* Notch ligand C2/MNNL domains - MNNL sequences are shown for human ligands, as well as three of the ten *C. elegans* ligands. Residues conserved in all human ligands are shown in blue. The β strand regions of Jagged-1 are shown above the alignment in green.

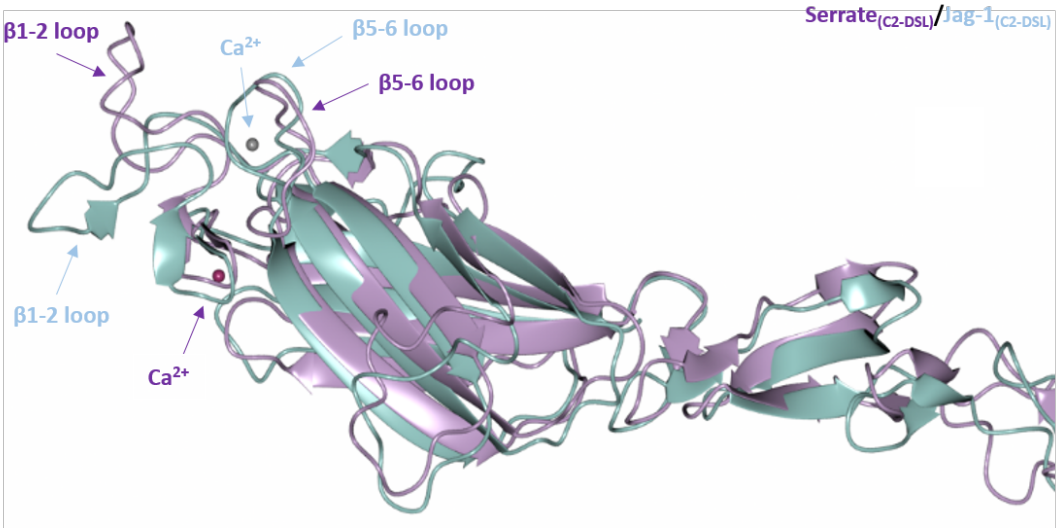


Figure 3.7: Superimposition of Serrate$_{NE2}$ and J1$_{NE3}$ (4CC1) crystal structures - The C2-DSL regions of Serrate (purple) and Jagged-1 (blue) are shown. Both were solved in the presence of calcium, and contain coordinate a single Ca^{2+} ion in their C2 domains. Serrate$_{NE2}$ coordinates kindly provided by R. Suckling. Models were visualised and superimposed using CCP4MG.

known Notch ligand structures. Therefore, the full length ligands were used, and the presence of the DSL and EGF-like domain together were able to form a sufficient alignment with the published structure of J1$_{NE3}$ (4CBZ) to generate a 3D prediction of the MNNL regions (Figure 3.8). It is important to note that the 4CBZ structure was obtained in the absence of calcium, and therefore has some incomplete loop regions. However, the FFAS server was unable to generate structures based on the calcium bound form of J1$_{NE3}$ (4CC1), and it may be that the absence of these loops was necessary for sufficient structural homology to be detected by the program.

Analysis of these models contributed little information regarding potential *C. elegans* MNNL regions that may be equivalent to the Notch- and phospholipid-binding sites of mammalian C2s. The C2 of Jagged-1 coordinates a single calcium with contributions from the β1-2 and β5-6 loops (Chillakuri et al., 2013), although residues from these loops were absent from the crystal structure of J1$_{NE3}$ generated in the absence of calcium (4CBZ). The β5-6 loop of J1$_{NE3}$ includes two residues

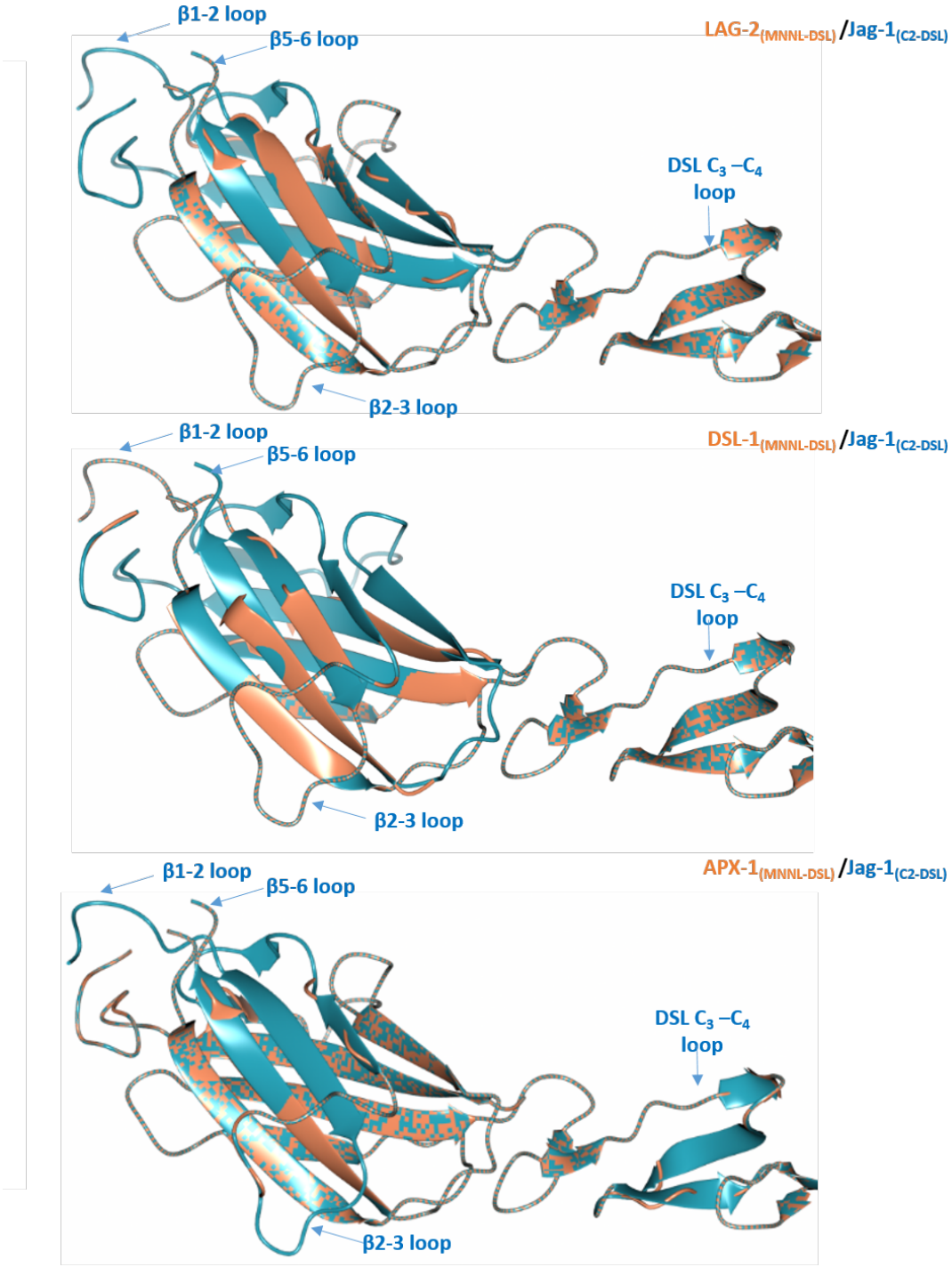


Figure 3.8: *C. elegans* Notch ligand models generated using alignments to $J1_{NE3}$ (4CBZ) - Full length ligand sequences were entered into the FFAS server, and FFAS-3D ouput is shown above, focussing on the MNNL-DSL region. Ligands are shown in orange, the 4CBZ structure of $J1_{NE3}$ is shown in blue. Models were visualised using CCP4MG. Primary sequence alignments are shown in Appendix 2.

D140 and D144, required for the coordination of calcium (section 1.6.2, Figure 1.11). Examination of the APX-1 residues that aligned to the absent region of the β5-6 loop of Jagged-1 in 4CBZ model allowed 6 amino acids (I80-K86) to be identified, including an aspartate (D81). This observation may be coincidental, and only high resolution structural information will be able to indicate the true significance of this region.

An additional consideration is whether a calcium-binding site is present in *C. elegans* MNNL regions, if they do indeed adopt an C2 fold. It has been suggested that calcium dependence of mammalian Notch ligand C2 domains is a feature of the Jagged/Serrate class of ligands, and that calcium is not bound in the C2 domains of Delta class ligands (Luca et al., 2015) (R. Suckling, personal communication). It has not been confirmed to which class the various *C. elegans* Notch ligands belong, although the presence of a region akin to a Cys-rich domain prior to the transmembrane region of APX-1 suggests that this at least may belong to the Jagged/Serrate class of ligands.

The β2-3 Notch binding loop initially identified in rDLL-4 is conserved in Jagged-1, however, the residues of the *C. elegans* MNNL regions that modelled to this area in the 4CBZ structure appeared to be an artefact of the FFAS alignment, consisting of regions either in or adjacent to the signal peptide region. Removal of the signal peptide prevented the generation of a model, and this indicated that discrepancies between the sizes of the *C. elegans* and Jagged-1 MNNL regions limited the use of this software for these particular structural predictions.

The computational methods used above to align primary sequences and predict structural similarities suggested that whilst it is possible that the N-terminal region of *C. elegans* Notch ligands adopt a C2 domain fold, structural information would be necessary to confirm this, as there is a high level of variability at the primary

sequence level, which is a common feature among human and *Drosophila* Notch ligand C2 domains. Structural information would also be essential in order to design specific point mutations to probe the *in vivo* functionality of a C2 domain in these ligands, therefore, the subsequent focus of this work was to express and purify fragments of *C. elegans* Notch ligands, to allow them to be studied at a structural and biochemical level.

3.2 Expression of Notch protein fragments using eukaryotic cells

3.2.1 Overview of target constructs

Several different *C. elegans* Notch protein fragments were targeted for expression, including the entire EGF-like region (1-13) of LIN-12, as well as the MNNL-EGF-like regions of the transmembrane Notch ligands LAG-2, APX-1 (Figure 3.9, Table 3.1). In LAG-2, the EGF-like region comprises the MNNL and DSL domain followed by two half EGF-like repeat and a whole EGF-like repeat, whilst in APX-1, 4 EGF-like repeats are present prior to a cysteine rich domain which was excluded from the recombinant fragment. In the case of the soluble Notch ligand DSL-1, which contains a single EGF-like region, the full length protein was targeted. All constructs were initially engineered to contain a C-terminal Fc-His tag, to facilitate purification, and allow characterisation using established liposome-binding and ELISA assays (Chillakuri et al., 2013). The construct $J1_{NE3FcHis}$, previously designed and cloned by C. Chillakuri, was also included as a positive control for expression from the system, during later experiments.

Unfortunately, no *C. elegans* cell line capable of high levels of recombinant protein production has been established. Therefore, transient transfection of HEK293T

Table 3.1: Properties of constructs designed for HEK293T expression - The number of amino acid residues and molecular weights of dimeric (in non-reducing conditions, due to the presence of the Fc tag) and monomer (in reducing conditions) constructs and the number of predicted N-linked glycosylation sites for each construct are shown. Additional constructs APX-1 $_{\text{DE4FcHis}}$ and J1$_{\text{DE3FcHis}}$ which are discussed in section 3.2.3 are also included.

Construct name	No. of residues	M_w of dimer (kDa)	M_w of monomer (kDa)	No. of N-linked glycan sites
LIN-12 E1-13FcHis	847	186.6	92.8	6
LAG-2 NE2FcHis	504	112.4	56.2	4
DSL-1$_{\text{FullFcHis}}$	480	108.8	54.4	4
APX-1 NE4FcHis	542	120.0	60.0	4
J1$_{\text{NE3FcHis}}$	550	122.4	61.4	3
APX-1 DE4FcHis	443	97.8	48.9	2
J1$_{\text{DE3FcHis}}$	397	89.0	44.5	1

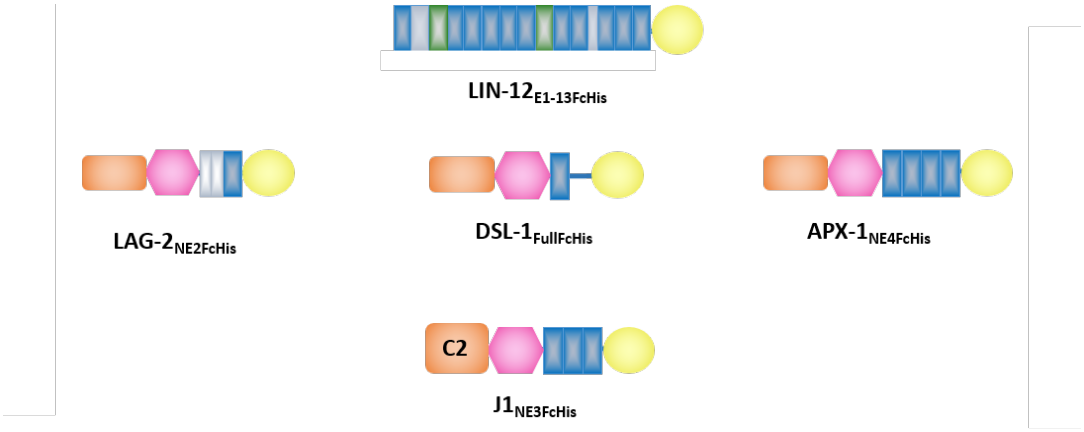


Figure 3.9: Architecture of constructs designed for HEK293T expression - Domains are coloured as in Figure 1.2. All constructs were designed to express as C-terminal Fc-His (yellow) fusion proteins. All were truncated extracellular fragments of transmembrane proteins except DSL-$1_{\text{FullFcHis}}$, which was a full length soluble ligand. Full architectures of the proteins are shown in Figure 1.4.

cells was initially chosen for expression, as successful production of recombinant fragments of human Jagged-1 and Notch-1, as well as *Drosophila* Serrate had previously been achieved with this system (Chillakuri et al., 2013). HEK293T cells facilitate folding of disulphide-rich EGF repeats, and post-translational modifications such as glycosylation that cannot be achieved in all expression systems (e.g *E. coli*).

3.2.2 Low levels of recombinant *C. elegans* Notch proteins are secreted by HEK293T cells

Proteins were affinity purified from the medium of transiently transfected HEK293T cells and an equal volume loaded for analysis by reducing SDS PAGE followed by Western blotting for each of the *C. elegans* constructs (Figure 3.10A), and, where present, their identities were confirmed by both their molecular weight (monomeric) and the identification of the His tag using a specific antibody. APX-1_{NE4FcHis} was detectable in the media, whilst LIN-$12_{\text{1-13FcHis}}$ and DSL-$1_{\text{FullFcHis}}$ were absent, and negligible LAG-2_{NE2FcHis} was detected in the affinity purified medium. To ascer-

tain whether the lack of protein detected in the medium was a consequence of low levels of gene expression for the *C. elegans* constructs, RNA was isolated from transfected cells and reverse transcribed using an Fc tag-specific primer. This was also carried out for cells transfected with $J1_{NE3FcHis}$, a human construct that is known to be secreted from transfected HEK293T cells. The cDNA was amplified using construct-specific primers for a range of PCR cycles, and the products run on a 1% agarose gel (Figure 3.10B). The results suggest that there are no large differences in the levels of $lag\text{-}2_{NE2FcHis}$, $dsl\text{-}1_{FullFcHis}$, $apx\text{-}1_{NE4FcHis}$ and $Jag\text{-}1_{NE3FcHis}$ mRNA, although there does appear to be less $lin\text{-}12_{1\text{-}13FcHis}$ mRNA compared with the other constructs present. Whilst this could either be due to reduced gene expression or high turnover/instability of the mRNA it is possible that the lower level is caused by the increased size of the amplification region, compared to the other mRNA sequences. To assess whether the *C. elegans* proteins were expressed within the transfected HEK293T cells, despite their absence/low levels within the purified medium, cell lysates were analysed using reducing SDS PAGE and Western blotting (Figure 3.10C). The results indicated that all the *C. elegans* protein constructs were present within the cell. It was noted that in the case of both the purified media and cell lysate Western blots, the monomeric constructs generally appeared to run at slightly higher molecular weight than had been predicted during the design of the constructs (Table 3.1), which may be caused by discrepancies between the predicted and experimental amount of detergent binding, or the presence of N-linked glycans, which might retard the movement of the protein through the SDS-PAGE gel.

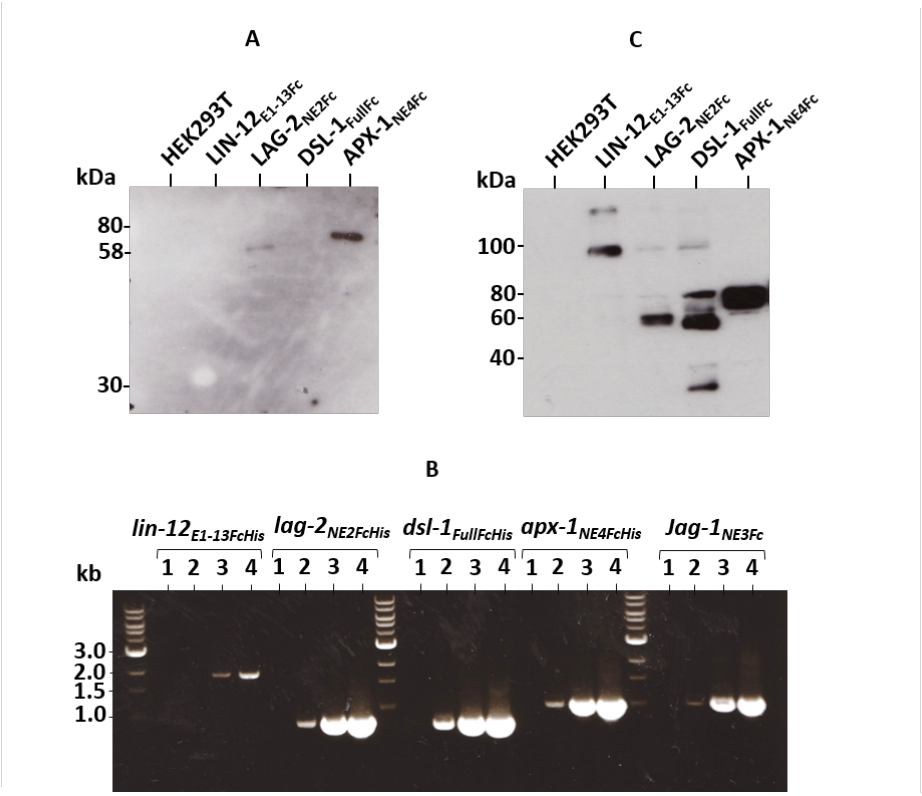


Figure 3.10: Expression of *C. elegans* Notch proteins in HEK293T cells - (**A**) Equal volumes of purified supernatant samples were run on an 12.5% SDS-PAGE gel, followed by Western blotting using an anti-His antibody for detection, and the results indicated an absence or low levels of protein. (**B**) Reverse transcription of mRNA followed by amplification of products for 15 (2), 20 (3), 25 (4) cycles of PCR suggested similar levels of mRNA between *C. elegans* ligands and the human ligand ($J1_{NE3FcHis}$). A negative control omitting reverse transcriptase was also included (1). (**C**) Western blotting of equal volumes of cell lysate samples indicated all constructs are present within the cells at the protein level.

3.2.3 Larger scale purifications of APX-$1_{NE3FcHis}$ yield insufficient material for further characterisation

The detectable quantity of APX-$1_{NE4FcHis}$ in the affinity purified medium of HEK293T cells during initial expression experiments led to subsequent efforts to scale up expression to purify a greater quantity. An additional APX-1 recombinant construct was included in these further purifications, APX-$1_{DE4FcHis}$, which lacked the MNNL region, in preparation for liposome-binding and interaction assays in which a control protein construct would be required. Both constructs were subjected to larger scale purifications (0.5 L of HEK293T media for a single purification). Following gel filtration, fractions were analysed using SDS PAGE followed by coomassie staining and Western blotting (Figure 3.11A and B), confirming dimerization of constructs due to the presence of the Fc tag under non-reducing condtions, which was lost under reducing conditions. Unfortunately, this approach was unable to yield more than 10 μg of protein/L for either construct, which was insufficient for additional characterisation experiments beyond SDS-PAGE and Western blotting. Successful purification of control proteins J$1_{NE3FcHis}$ and J$1_{DE3FcHis}$ indicated that the expression system is capable of yielding >50 μg of purified protein per two T175 flasks (50ml media) (Figure 3.11C).

Factors that may have contributed to the low quantity of APX-1 recombinant proteins expressed by HEK293T cells for different constructs include the presence of motifs encoding mRNA secondary structures that may have interfered with translation and codon usage bias of the host cells. Since codon usage appears to differ between *C. elegans* and humans (Gustafsson et al., 2004), the APX-1_{NE3} sequence was submitted to Genescript for human codon usage optimisation and synthesis. The construct was subsequently cloned into pHLsec, and the crude media of cells transfected with non-optimised and optimised contructs were compared (Figure 3.12).

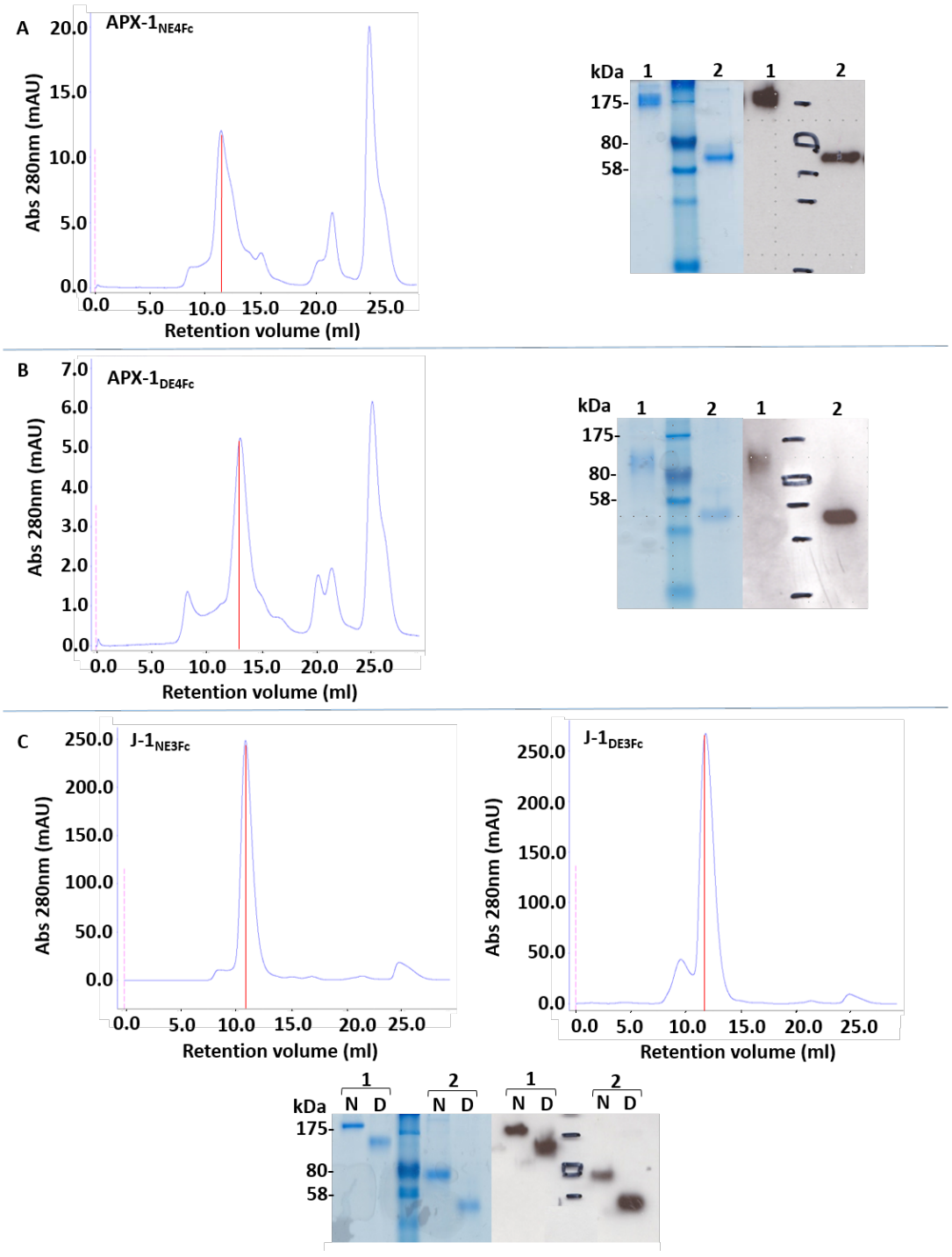


Figure 3.11: Analysis of purified APX-1 and Jag-1 recombinant proteins - Superdex 200 column gel filtration chromatography traces for each construct are shown, with peaks collected for analysis marked in red. 500 ng and 200 ng of purified protein were loaded on a 12.5% SDS PAGE gel for coomassie staining and Western blotting, respectively under non-reducing (1) and reducing (2) conditions. APX-1_{NE4FcHis} (**A**) and APX-1_{DE4FcHis} (**B**) both appeared to be successfully purified in dimeric form. (**C**) $\mathrm{J1}_{\mathrm{NE3FcHis}}$ (N) and $\mathrm{J1}_{\mathrm{DE3FcHis}}$ (D) were also purified.

Optimisation of the APX-1_{NE3} sequence did not appear to increase the quantity detected in the media, and it was concluded that in the case of this specific construct, codon usage bias was not the main factor contributing to the lack of protein present.

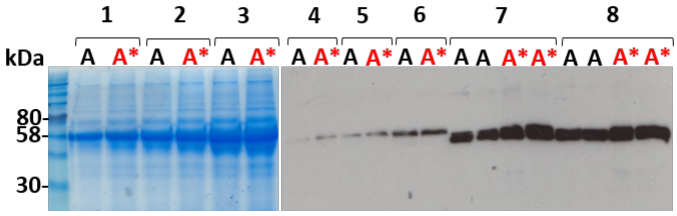


Figure 3.12: Comparison of levels of APX-$1_{NE4FcHis}$ before and after codon optimisation - Coomassie staining and western blotting of crude media from HEK293T transfection are shown (1-6). 2.5 μl, 5.0 μl, 7.5 μl, 1.0 μl, 2.0 μl and 3.0 μl samples were loaded in lanes 1, 2 and 3, 4, 5 and 6 respectively. 8 μl (lane 7) and 16 μl (lane 8) of a 10ml preparation of cell lystates were also analysed using western blotting. APX-$1_{NE3FcHis}$ prior to optimisation is denoted A, codon optimised samples are shown in red as A*. All samples were run under reducing conditions on a 12.5% SDS PAGE gel.

3.2.4 APX-$_{NE3His}$ can be expressed from S2 cells

In light of initial attempts to express recombinant *C. elegans* Notch fragments from HEK293T cells, an alternative eukaryotic expression system was sought. *Drosophila* S2 cells obtained through Expres^{2}ion technologies were considered to have several advantages that could potentially address the limitations of transient HEK293T cell expression. Expres^{2}ion S2 cells are stably transfected, and therefore after an initial investment of one month to generate cell lines, no further transfections are necessary. Once identified, stably transfected cells can be expanded in suspension, as opposed to a monolayer culture, allowing large quantities of media to be purified at once. Finally, the post translational modifications performed by S2 cells are considered simpler than in many mammalian cell lines. Specifically, N-linked glycosylation is less complex and sugar modifications are generally of the high-mannose type, and perhaps less likely to make non-native PTMs to *C. elegans* Notch proteins than

human cells (http://www.expres2ionbio.com/).

3.2.5 APX-1_{NE3His} can be expressed by S2 cells

APX-1 was selected as an initial target for expression using S2 cells. The region spanning the MNNL-EGF-like-3 was optimised and synthesised using Geneart services (Life technologies). This fragment (APX-1_{NE3His}) omitted the Fc tag used in previous constructs to rule out its interference in protein production, and to allow direct use of this construct in structural studies.

Transfection and generation of cell cultures stably expressing APX-1_{NE3His} was carried out with M. Jore. 1 L of crude media was then collected and purified. Multiple peaks were present throughout the gel filtration profile of this construct, and fractions were collected for the largest peaks and analysed using SDS-PAGE followed by Coomassie staining and Western blotting (Figure 3.13). These data indicated that the first two peaks of the gel filtration contained APX-1_{NE3His} (confirmed by His tag specific blotting), however this material appeared to form high molecular weight, non-covalent aggregates that were lost under reducing SDS-PAGE conditions. Additional high molecular weight impurities also appeared to be present. The third peak of the gel filtration profile contained monomeric APX-1_{NE3His}, as well as an additional protein species at a slightly higher molecular weight. Therefore, although APX-1_{NE3His} was successfully affinity purified from the media of S2 cells, this construct was not considered sufficiently homogeneous for structural studies. However, these experiments established S2 expression as a feasible route for expression recombinant fragments of *C. elegans* Notch ligands.

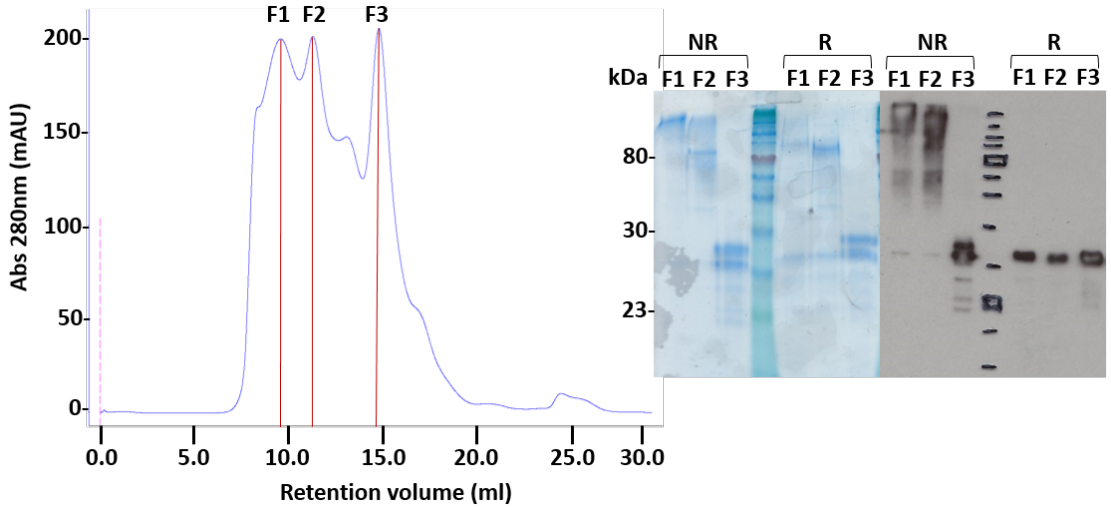


Figure 3.13: Analysis of purified APX-1NE3His from S2 cells - The gel filtration chromatography trace for APX-1_{NE3His} is shown. 500 ng and 200 ng of purified protein collected from peak fractions (F1, F2, F3) were loaded on a 12.5% SDS PAGE gel for coomassie staining and Western blotting, respectively, under non-reducing (NR) and reducing (R) conditions.

3.2.6 APX-1_{NE1His} can be purified from S2 cells with sufficiently high yield and quality for structural studies

R. Suckling generated the DNA sequence for the fragment APX-1_{NE1His}, (encoding the region spanning the MNNL-DSL-EGF-like 1 of APX-1) and transfected S2 cells to establish stable expression of the protein fragment. Subsequently, the construct was successfully purified (Figure 3.14) to yield 8 mg of APX-1_{NE1His} from 1 L S2 media. This material was used in crystal trials, also performed by R. Suckling, but optimal conditions have yet to be found to grow crystals of sufficient homogeneity to allow a high resolution crystal to be generated.

3.2.7 Limited proteolysis

Limited proteolysis was performed to look for the presence of a calcium-binding site in the MNNL region APX-1, as might be anticipated if this ligand belonged to the

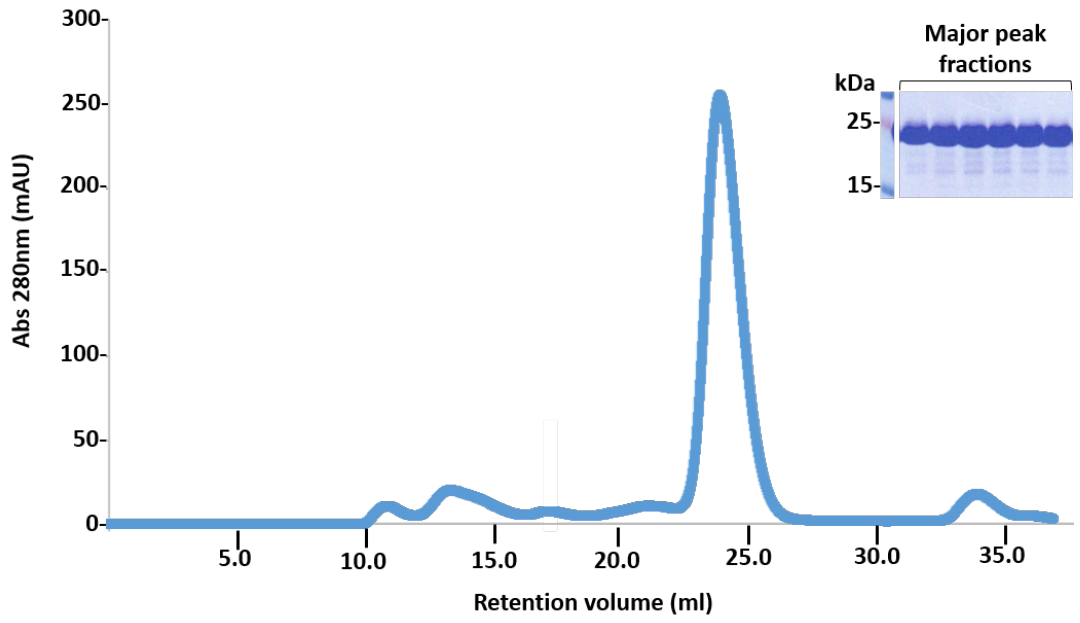


Figure 3.14: Analysis of purified APX-1NE1His from S2 cells - The gel filtration chromatography trace for APX-1_{NE1His} is shown, with a coomassie stained SDS-PAGE gel. 10 μl samples of 500 μl fractions collected from the major peak are shown to confirm purity. Data provided by R. Suckling.

Jagged/Serrate class of Notch ligands. Two serine proteinase enzymes with different specificities were used to digest APX-1_{NE1His} over time (Figure 3.15) . One small difference in the trypsin digestion pattern was observed, including an additional fragment of <10 KDa in the presence of EDTA, through out samples at all timepoints after enzyme addition. Overall, however, the results indicated that the presence of calcium offered no protection from digestion with either Trypsin or Endoproteinase Glu-C.

Previously, limited proteolysis experiments have supported the presence of a calcium-binding site in the C2 domain of Jagged-1 (Chillakuri et al., 2013), and human Notch-1 (11-13) (Cordle et al., 2008b), although it is important to note that whilst differences in cleavage patterns can support the presence of a calcium-binding site, the absence of these differences doesn't exclude the presence of such a site. Therefore, these data indicated that whilst it is possible that APX-1_{NE1His}

does not contain a calcium-binding site, it is also possible that calcium is present, but its position offers no additional protection from proteolysis. Trypsin cleaves at arginine and lysine residues (Olsen et al., 2004), with 10 potential sites in the MNNL of APX-1, whilst Endoproteinase Glu-C cleaves at glutamic acid residues (Drapeau et al., 1972), with 6 potential sites in the MNNL of APX-1. It is possible that access to these particular sites by the enzymes was not altered by the presence or absence of calcium, or that the concentrations of enzymes was not optimal to capture subtle protection effects. Additionally, if a low affinity calcium binding site were present, this may not be sufficient to offer protection from degradation. Therefore, these data did not allow conclusions to be drawn with confidence regarding the calcium dependence of the MNNL of APX-1.

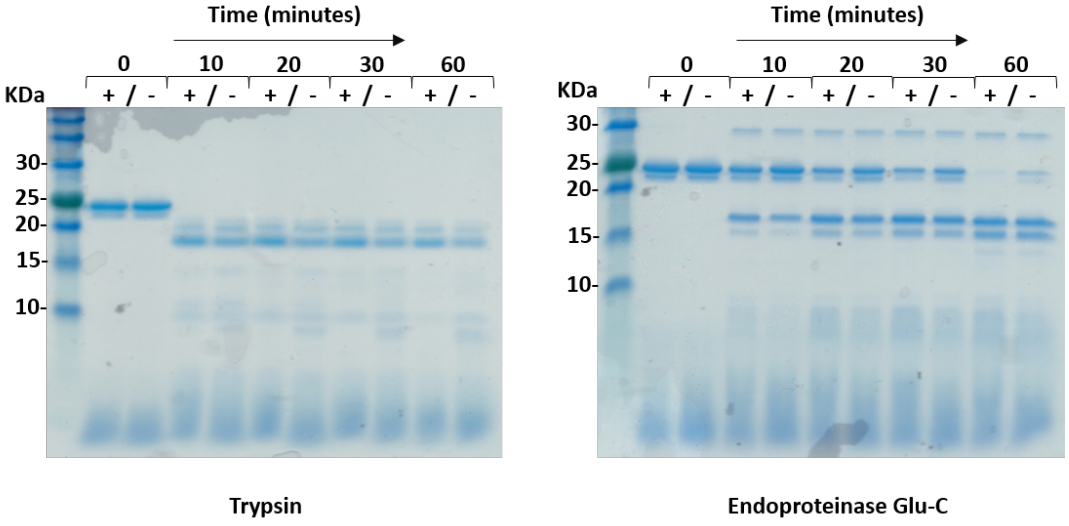


Figure 3.15: Tryspin and Endoproteinase Glu-C digestion time courses of APX-1$_{\mathbf{NE1His}}$ - 2 μg samples of APX-1$_{\mathrm{NE1His}}$ were withdrawn before digestion (0 min) and at subsequent timepoints, in the presence of 10 mM $CaCl_2$ (+) or 10 mM EDTA (-), and analysed on a 16% SDS-PAGE gel under non-reducing conditions.

3.3 Discussion

Analysis of *C. elegans* Notch protein primary sequences indicated a small number of similarities, as well as several differences compared with *Drosophila* and humans that suggested the precise molecular interactions occurring in the receptor/ligand interaction of *C. elegans* may differ to some extent from the current understanding in *Drosophila* and mammals. This highlighted the need for structural information, and led to attempts to establish an expression system for recombinant fragments of *C. elegans* Notch receptor and ligand proteins. Initially, a human cell line was used for expression, however this was unsuccessful, possibly due to inappropriate post-translational modification of the *C. elegans* proteins. One feature of HEK293T cells that is considered advantageous is their ability to perform complex modifications at N-linked glycosylation sites that can be vital for correct protein folding (Aricescu et al., 2006). However, it is possible that in the case of *C. elegans* proteins, this may have had deleterious effects on protein expression.

In both humans and *C. elegans*, proteins are transported into the endoplasmic reticulum (ER) during synthesis, where oligosaccharyl transferase (OST) catalyses the transfer of high-mannose glycans via amide linkage at the asparagine residue of the conserved N-XXX-S/T motif. These increase solubility, hence preventing aggregation of the polypeptide as it emerges into the lumen unfolded (Moremen and Molinari, 2006). Natively folded proteins proceed through the ER and into the Golgi apparatus. In humans, this series of compartments is the site of many modifications to the sugar moiety, including the addition of GlcNAc, galactose and sialic acid residues, forming a hybrid and subsequently a complex oligosaccharide as it moves towards its destination (e.g a secretory vesicle). Despite an abundance of high-mannose glycans in *C. elegans*, hybrid and complex N-linked glycans are not generally found. Instead, N-glycans appeared to be modified to paucimannose,

which is rarely found in vertebrates (Zhang et al., 2003).

The APX-1_{NE4FcHis} sequence contains three N-linked sites that could potentially be glycosylated by HEK293T cells: one near the C-terminus of the MNNL, one between EGF-like 1 and 2 and one in the human derived FcHis tag. The post translational modification requirements of *C. elegans* for successful recombinant Notch protein expression are still not fully understood, but differences in N-linked glycosylation may have contributed to the lack of *C. elegans* Notch protein expression using HEK293T cells.

The use of *Drosophila* S2 cells combined with *Drosophila* codon optimisation of the expression construct appeared to substantially improve recombinant protein secretion. Since initial attempts to purify APX-1_{NE3His}, many human and *Drosophila* Notch protein fragments of varying length have been successfully purified using the S2 expression system, giving insights into optimisation of the purification process and the range of conditions in which these types of constructs are most soluble and stable (R. Suckling, personal communication). In addition, the reduced size of APX-1_{NE1His} may have contributed to an increase in the solubility of the protein, compared with APX-1_{NE3His}. For structural studies this shorter fragment may be advantageous, as demonstrated by the high resolution co-crystal structure of rDLL-4_{SLPNE1}/r$N1_{\text{11-13}}$. Reducing the length of the fragment of APX-1 also reduced the number of potential N-linked glycosylation sites present to one, and this may have also increased the ability to produce homogeneous material. Although it was not possible to gain insights into APX-1 at the structural level, the work in this chapter has highlighted the power of the S2 expression system for recombinant protein production. This approach could facilitate the production of a variety of Notch ligand fragments in the future, which could ultimately allow structural information to be obtained.

4

Developing an *in vivo* assay to study *C. elegans* Notch ligand function

4.1 Introduction

4.1.1 *C. elegans* as a model to study Notch signalling

4.1.1.1 The anatomy and life cycle of *C. elegans*

Caenorhabditis elegans is a small, free-living nematode found in soil (particularly, rotting vegetation) that was first isolated as a model organism for genetic research by Sydney Brenner (Brenner, 1974). Adult *C. elegans* grow to ∼1 mm in length and *C. elegans* exist as either self-fertilising hermaphrodites with two X (XX) chromosomes, or males with a single X chromosome (XO), in addition to the 5 autosomes shared by both sexes. *C. elegans* have cylindrical bodies, tapering at each end, encapsulated in a collagenous cuticle (Figure 4.1). Like other nematodes, *C. elegans* anatomy is described as a 'tube within a tube', as the digestive system runs from the mouth to the anus of the worm via the pharynx, intestine and rectum. In adult hermaphrodites, the gonad occupies much of the internal space of the worm,

extending two U-shaped arms (proximal and distal) around nearly the full length of the worm, whilst males have only a single distal gonad arm. An additional anatomical feature found in only males is the tail, which contains all specialised structures required for mating.

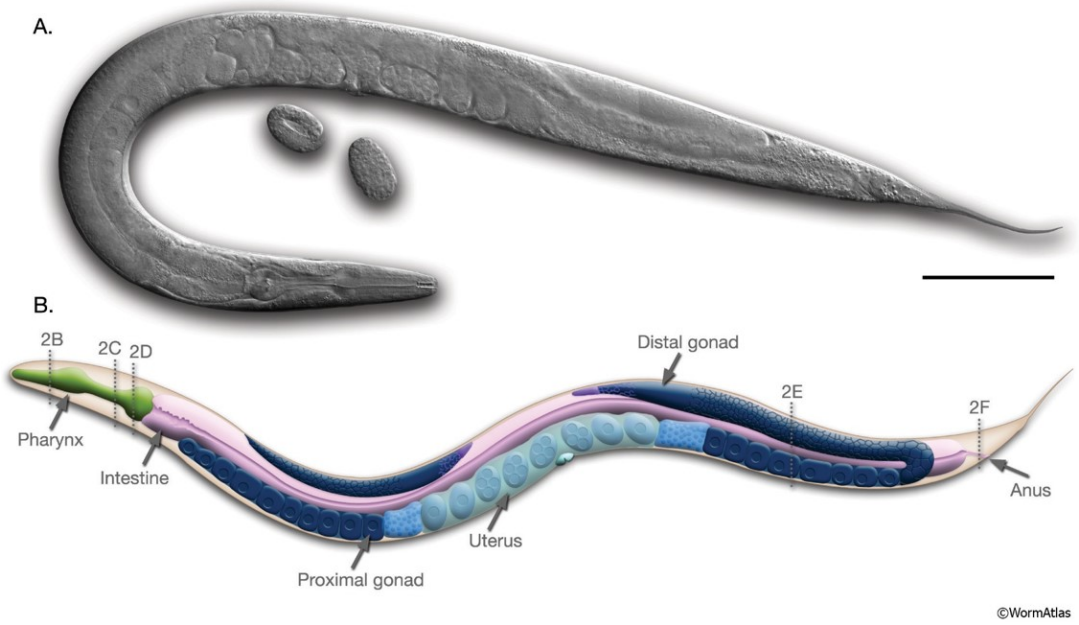


Figure 4.1: Anatomy of an adult hermaphrodite - (**A**) DIC image of an adult hermaphrodite, left lateral side. Scale bar 0.1 mm. (**B**) Schematic drawing of anatomical structures, left lateral side. Image and legend from http://www.wormatlas.org.

The life-cycle of *C. elegans* is temperature dependent (Byerly et al., 1976), taking approximately 3 days to develop from embryo to adult at 20°C (Figure 4.2). After fertilisation, *C. elegans* embryos develop to the ∼30 cell stage (gastrulation), before they are laid. The embryos then continue to develop until the 558 cell stage, when they hatch into larvae. In the presence of food (e.g *E. coli* or other bacteria) *C. elegans* develop post embryonically, undergoing four larval molts (L1-L4) in which they shed the cuticle associated with the previous larval stage, before reaching maturity. Adult hermaphrodites lay up to 300 eggs across a period of ∼3 days, and can live for approximately two weeks. When embryos hatch into an environment with

scant food supply, the larvae enter dauer, and can remain in this stage for weeks until food is once again available. Fed dauer re-enter post-embryonic development at L4.

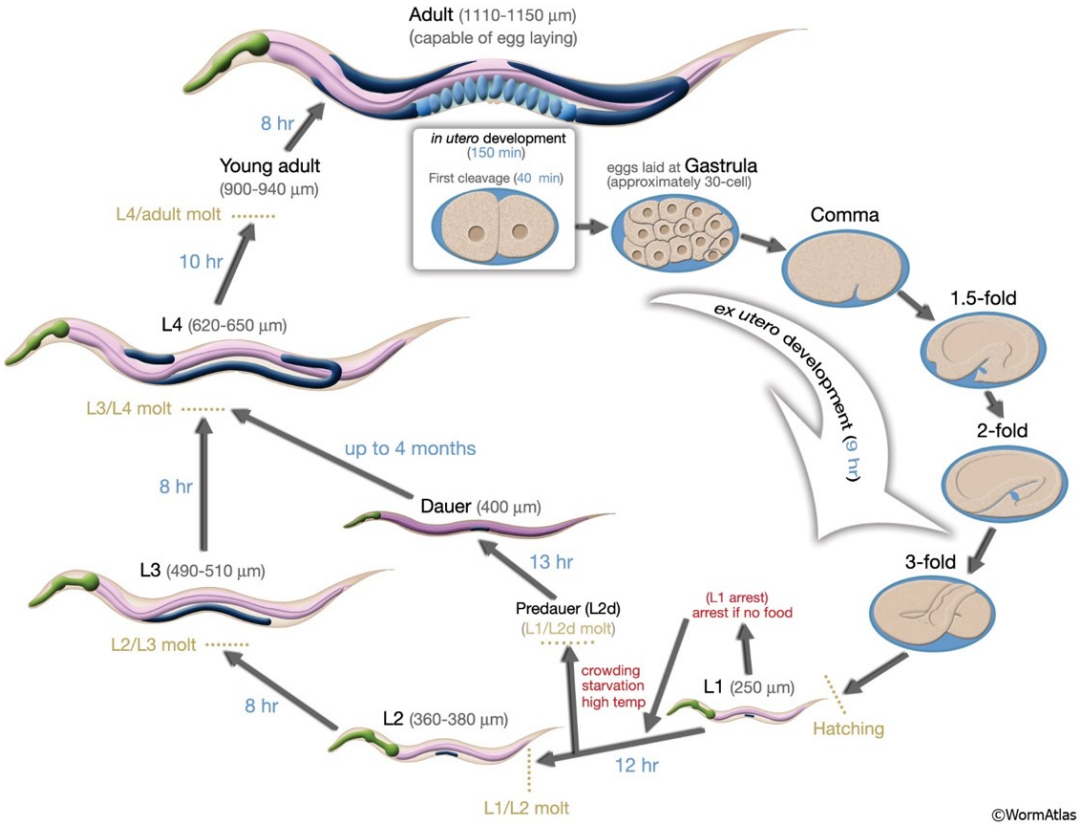


Figure 4.2: Life cycle of *C. elegans* at 22°C - 0 min is fertilization. Numbers in blue along the arrows indicate the length of time the animal spends at a certain stage. First cleavage occurs at about 40 minutes post-fertilization. Eggs are laid outside at about 150 minutes post-fertilization and during the gastrula stage. The length of the animal at each stage is marked next to the stage name in micrometers (μm).Image and legend from http://www.wormatlas.org.

The lineages of every single one of the 959 somatic cells in an adult hermaphrodite and the 1033 cells of adult males have been mapped through embryonic and post embryonic development (Sulston and Horvitz, 1977) (Sulston et al., 1983). This intricate knowledge of cell divisions and fates throughout the entire organism has provided a strong foundation for the use of *C. elegans* as a tool to study development. Other advantages of *C. elegans* as a model for developmental processes

includes their relative ease of manipulation, including their short generation time and transparency, making them suitable for simple light microscopy studies. The ability for *C. elegans* hermaphrodites to reproduce by themselves, or be mated with males also makes them extremely amenable for genetic studies.

4.1.2 Identification and characterisation of Notch receptors and ligands in *C. elegans*

Prior to it's identification as a Notch receptor gene, *lin-12* (identified in a screen of for **lin**eage-defective mutants) was characterised as a developmental gene acting as a binary switch, based on the differences between phenotypes of *lin-12* mutants associated with higher and lower gene dosage (Horvitz et al., 1983) (Greenwald et al., 1983). *lin-12* null mutants survive to adulthood, despite developing abnormalities such as a second anchor cell, whilst *lin-12 (gf)* dominant mutants display a multivulva phenotype, with up to six pseudovulvae present. *glp-1* (**g**erm**l**ine **p**roliferation defective) was first identified as a key component required for maintenance of the germline, since *glp-1(lf)* (loss-of-function) adult mutants are sterile (Austin and Kimble, 1987).

The two genes were later discovered to be Notch receptors (Greenwald, 1985) (Yochem et al., 1988), (Yochem and Greenwald, 1989), (Austin and Kimble, 1989), and their functions shown to be overlapping during embryogenesis (Lambie and Kimble, 1991), although they are not redundant during post-embryonic development, since single mutants have differing phenotypes, as described above. *lin-12/glp-1* double mutants are associated with a distinct and novel 'Lag' (***l****in-12* **a**nd ***g****lp-1*) phenotype. The most common and striking feature of this phenotype is the absence of a rectum and anus in larvae, causing development to arrest during L1. Other defects that are associated with the Lag phenotype include a twisted nose, and the lack

of the excretory cell. Additional evidence that *lin-12* and *glp-1* share functionality was provided by experiments studying the expression of *glp-1* under the control of the 5' flanking region of the *lin-12* gene (a *lin-12/glp-1* chimera) (Fitzgerald et al., 1993). In wild type animals, the adoption of secondary cell fates by vulval precursor cells is dependent on *lin-12* expression, however, the *lin-12/glp-1* chimera could substitute for *lin-12* to achieve correct specification, suggesting a common biochemical functionality between the two receptors. The discovery of the Lag phenotype in *lin-12/glp-1* double mutants formed the basis of a screen to identify related genes. WT hermaphrodites were treated with EMS mutagen and mutants whose progeny displayed the Lag phenotype were isolated and analysed, which allowed the *lag-1* and *lag-2* genes to be defined (Lambie and Kimble, 1991). Subsequent work allowed these components to be placed in the Notch pathway as the CSL (LAG-1) (Roehl et al., 1996) and the Notch ligand LAG-2 (Tax et al., 1994)(Henderson et al., 1994).

The *apx-1* gene was initially identified as a maternal gene required for the development of the ABp blastomere at the four-cell embryo stage, and APX-1 was subsequently classified as a ligand of GLP-1 in this context (Mango et al., 1994) (Mello et al., 1994). A third membrane bound Notch ligand, ARG-1, was also identifed (Mello et al., 1994). Later, computational analysis predicted the presence of seven additional *C. elegans* Notch ligands (DSL-1-7) (Chen and Greenwald, 2004). These were classed as canonical Notch ligands, due to the presence of a DSL motif, but surprisingly, appeared to be soluble, unlike all known mammalian and *Drosophila* ligands. The fact that these Notch ligand genes were not discovered previously through mutagenic screens implied functional redundancy of soluble ligands with their membrane-bound counterparts.

4.1.2.1 The *lag-2(q420)* allele is a temperature sensitive splice site mutation

Several different *lag-2* alleles associated with a reduction or loss of Notch signalling were isolated in an EMS screen (Lambie and Kimble, 1991) and analysed (Henderson et al., 1994). These mutants ranged from the *lag-2(q411)* null allele that must be maintained as a heterozygote to generate viable progeny, to the weaker *lag-2(q420)* temperature sensitive allele that can be maintained as a homozygote at the permissive temperature 15°C. Both of these alleles were associated with the Lag phenotype, and both alleles were subsequently used for transgenic rescue studies by the authors (Henderson et al., 1994). For rescue studies presented in this thesis, the *lag-2(q420)* allele was selected, due to the advantage of being able to maintain homozygous progeny and the ability to observe post embryonic developmental defects associated with *lag-2(q420)* mutants in addition to the Lag phenotype.

The precise mechanism by which the *lag-2(q420)* allele exerts its effect is unknown at present. *lag-2(q420)* worms have a single point mutation at an authentic splice acceptor site (AG to AA) that occurs at the 3' splice site of the second intron/ third exon of *lag-2* (Figure 4.3). The consequence of this mutation has not been studied at an mRNA level, but by removal of the splice acceptor site, it may be predicted that splicing is disrupted as a result. However, there have been previous studies to show that *C. elegans* do not require the AG site for correct splicing. Similar mutations (AG to AA) isolated in the unrelated genes *let-23* and *dpy-10* were studied for their effects on the production of mRNA transcripts (Aroian et al., 1993). Both mutations occur at the 3' splice site at the end of a short intron (∼50 nucleotides), and analysis of the mRNA products indicated the presence of multiple transcripts, over 50% of which are spliced at the AA mutated site, or at a cryptic splice site close by, whilst the remaining transcripts result from exon skipping, and

a lack of intron splicing. It may be that the *q420* allele also results in this variety of transcripts, and that in the case of this particular mutant, correct transcripts are generated at a much greater frequency at the permissive temperature of 15°C compared with the non-permissive temperature of 25°C.

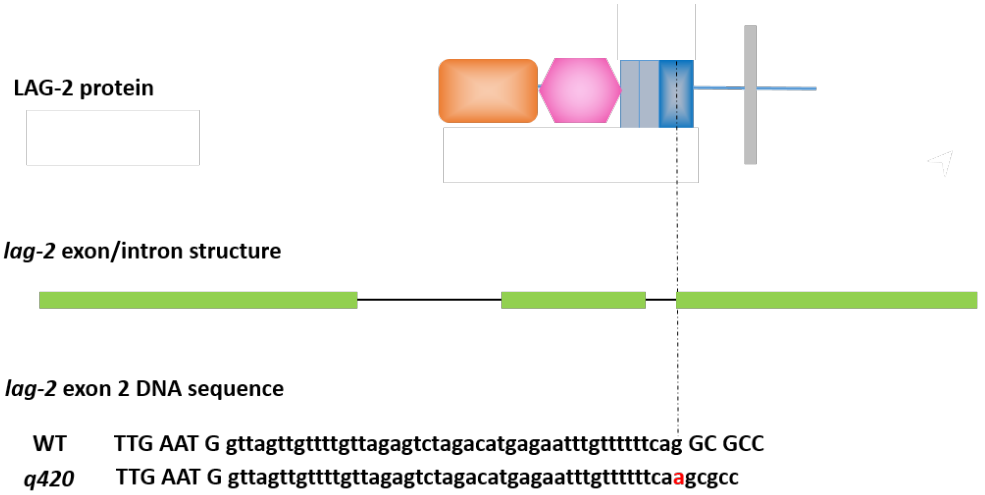


Figure 4.3: The site of the *q420* allele in *lag-2* - An A to G mutation potentially disrupts the 3' splice acceptor site of intron 2/exon 3 in the *lag-2* gene occurs within the region encoding EGF-like 2 of the protein LAG-2. Exons shown in green, introns as black lines.

4.1.3 Post embryonic developmental defects in *lag-2(q420)* mutants highlight the roles of Notch signalling in *C. elegans*

In addition to the Lag phenotype displayed by *lag-2(q420)* worms hatched at the restrictive temperature of 25°C, temperature shifting worms hatched at 15°C to 25°C at the end of L1 allows a range of post-embryonic developmental defects to be observed, which are generally also associated with *lin-12* and *glp-1* single mutants. This range of defects indicates the importance of Notch signalling throughout multiple developmental processes in the worm, and below, the role of Notch signalling

in the context of these developmental processes is briefly summarised.

4.1.3.1 Anchor cell specification

The anchor cell is specified at the end of L2, when lateral Notch signalling between two equivalent cells (Z1.ppp and Z4.aaa) determines their subsequent fates (Seydoux and Greenwald, 1989) (section 1.2). In wild-type worms, both Z1.ppp and Z4.aaa express LIN-12 and LAG-2 at their cell surface, however an unknown stochastic event leads to an increase in *lin-12* gene expression in one of the two cells. This cell, the pre-VU cell, is subsequently associated with increased *lin-12* expression and decreased lag-2 expression, whilst the neighbouring cell (consequently defined as the pre-AC cell) has restricted *lin-12* expression and an increase in *lag-2* expression (Wilkinson et al., 1994) (Figure 4.4). The precise mechanism of this feedback between the two cells is not fully understood, however there is evidence to suggest that LAG-1 (the CSL) is required for regulation of both the *lin-12* and *lag-1* genes within the VU cell, and for the regulation of *lag-2* in the AC cell (Park et al., 2013).

4.1.3.2 Vulval precursor cell patterning

The complex process of vulval development has been well studied, and takes place in many stages (http://www.wormbook.org/chapters/www_vulvaldev/vulvaldev.html). At one particular stage of vulva formation, vulval precursor cell (VPC) fates are specified by the interplay of multiple signalling pathways, including Notch (Figure 4.5) which limits inductive Ras signalling by lateral inhibition. Initially, six VPCs (P3.p-P8.p) are equivalent. The nearby AC releases an inductive signal, comprising LIN-3, a soluble ligand of the *C. elegans* EGF signalling pathway. LIN-3 binds to and activates LET-23 (an EGF receptor homologue) present on VPCs, activating Ras-MAPK signalling in the target cells. The extent of this activation if spatially

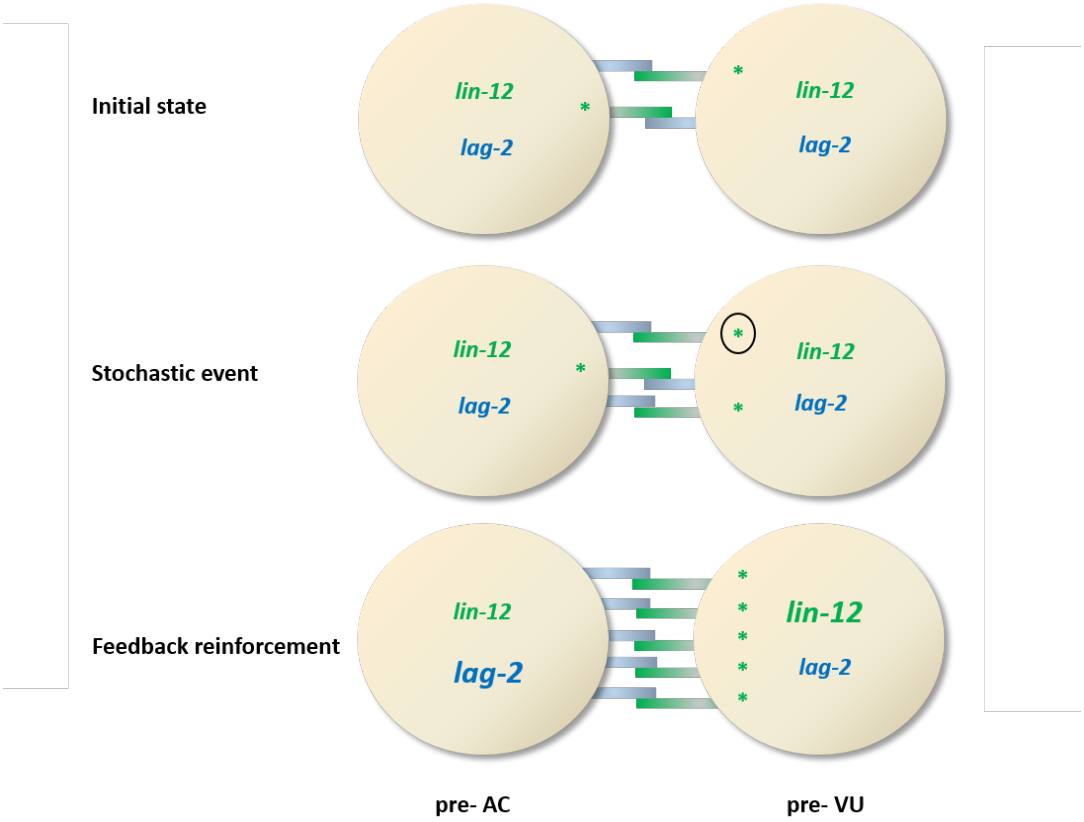


Figure 4.4: Specification of the AC/VU cells requires Notch signalling - In the initial state, Z1.ppp and Z4.aaa are equivalent, and both express both *lin-12* and *lag-2*. However, a stochastic event leads to and increase in *lin-12* expression in cell, defining it as the pre-VU cell. A positve feedback loop of Notch signalling escalates the expression of *lin-12* in the pre-VU cell, and *lag-2* in the pre-AC cell, whilst lowering expression levels of *lag-2* in the pre-VU cell and *lin-12* in the pre-AC cell. Adapted from (Riddle et al., 1997).

dependent, and the VPC in closest proximity to the AC (P6.p) is therefore activated to a greater extent than its neighbours (P5.p and P7.p) (Yoo et al., 2004). Ras mediated induction of P6.p leads to endocytosis of LIN-12 receptors present at the cell surface (Shaye and Greenwald, 2002) as well as increased expression of the Notch ligand genes *lag-2*, *apx-1* and *dsl-1* (Chen and Greenwald, 2004), which increases the Notch signal-sending capacity of P6.p. P5.p and P7.p receive Notch signals via LIN-12 at their cell surface, and the activation of the Notch pathway in these cells inhibits the effects of Ras/MAPK signalling in these cells (Berset et al., 2001). Therefore, Notch signalling prevents P5.p and P7.p adopting the primary fate as a result of induction by the AC, and instead, specifies the secondary fate in these cells.

4.1.3.3 Germline proliferation

Notch signalling plays a key role in germline proliferation. The anatomy of an adult gonad is summarised in Figure 4.6. Notch signalling between the somatic distal tip cell (DTC) and the distal region of the germ line (DG), a syncytium containing partially enclosed germ cells, is necessary to maintain the germ cells as a pool of mitotic stem cells, which enter meiosis further down the germ line, eventually maturing into oocytes prior to fertilisation. Loss of Notch signalling by either ablation of the DTC, or loss of GLP-1 in *glp-1(lf)* mutants results in the failure for the germline to proliferate, and for all mitotic germ cells to enter meiosis prematurely (Austin and Kimble, 1987). Conversely, *glp-1(gf)* mutants give rise to a 'tumorous' germline phenotype, in which a vast excess of mitotic germ cells are produced (Berry et al., 1997).

The distal tip cell (DTC) is a large somatic cell, and one DTC is associated with each tip of the DG (proximal and distal) in hermaphrodites. DTCs have a large cell

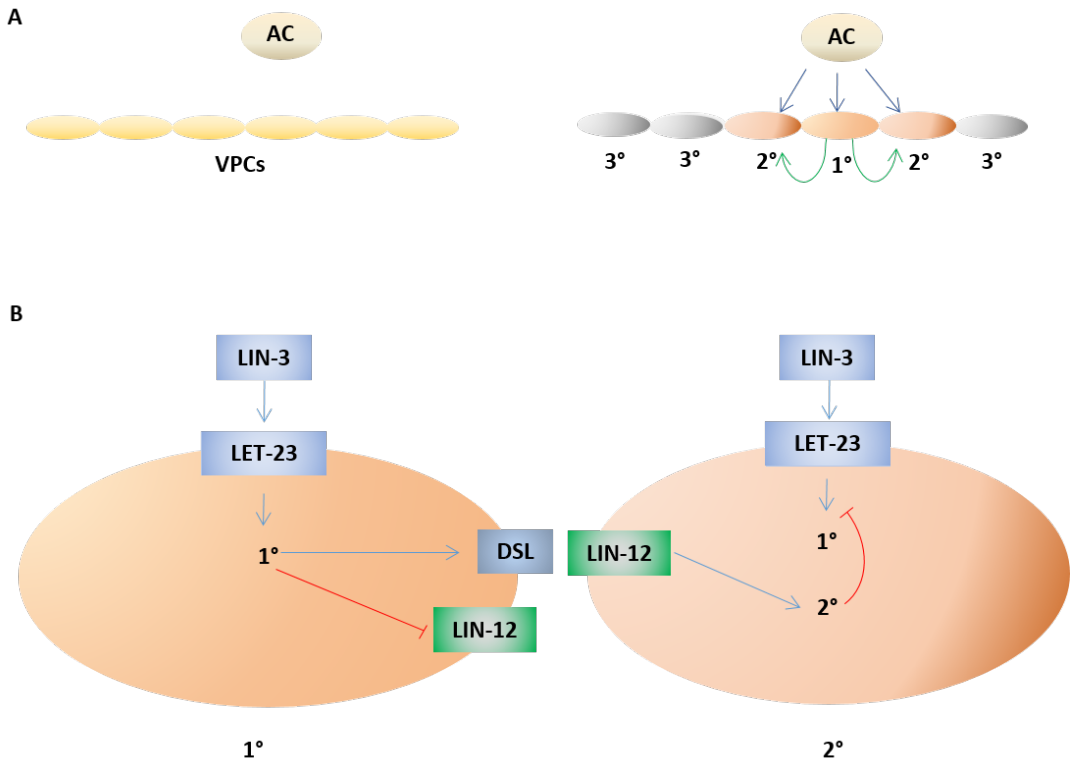


Figure 4.5: Notch provides the lateral signal for secondary cell fate specification during VPC patterning - (**A**) The VPCs receive an inductive EGF signal from the AC, which induces the primary fate in P6.p, the closest VPC. This then induces P6.p to signal to it's neighbours P5.p and P7.p via a lateral signal, which inhibits the primary fate in these cells, and causes them to adopt the secondary fate. (**B**) The lateral signal used the Notch pathway, to inhibit the LIN-3/LET-23 Ras-MAPK pathway in cells destined for the secondary fate. Adapted and modified from Sternberg and Horvitz 1989.

body 'cap' that covers the end of the DG, and multiple processes that extend down the side of the DG (Byrd et al., 2014). LAG-2 is expressed at the surface of both the cap region and the processes of the DTC, and activates GLP-1 receptors, which are restricted to germ cells found in the mitotic region of the germline (Crittenden et al., 1994). Activation of GLP-1 leads to the expression of target genes such as *lst-1* and *sygl-1*, which are necessary to maintain the stem cell pool (Kershner et al., 2014).

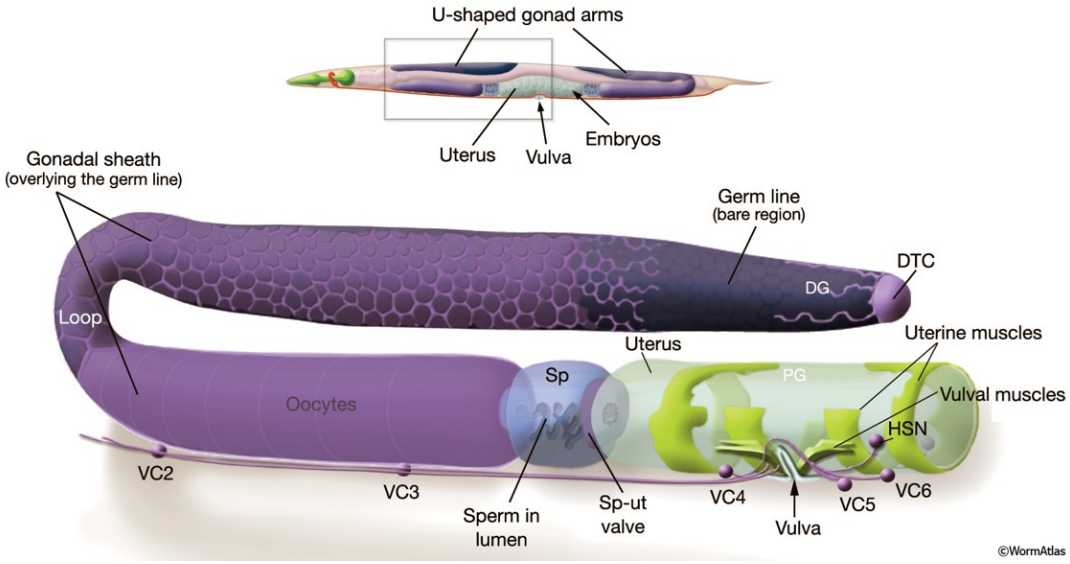


Figure 4.6: Anatomy of the adult hermaphrodite gonad - (Above, overview of the hermaphrodite reproductive system. Below, one half of the reproductive system, enlarged and separated from other body parts. (DTC) Distal tip cell; (DG) distal gonad; (PG) proximal gonad; (Sp) spermatheca; (Sp-ut) spermathecal-uterine valve; VC1-6 and HSNL/R are motor neurons that control egg-laying. The regions discussed in the main text include the DTC and DG. Image and legend from http://www.wormatlas.org.

4.1.3.4 Summary of post embryonic developmental defects associated with *lag-2(q420)* mutants

The consequence of reduced Notch signalling in *lag-2(q420)* mutants is the specification of two AC cells (as in *lin-12(lf)* mutants), as opposed to a single AC and VU

cell (Greenwald et al., 1983) (Lambie and Kimble, 1991). This presumably occurs since the pre-VU cell is unable to receive the sufficient levels of Notch signals required to adopt the VU fate, and subsequently adopts the AC fate, by default. It is highly likely that it is this duplication of the AC that contributes to the occurrence of vulval defects in adult *lag-2(q420)* worms. The duplication of the AC in *lin-12* mutants has been shown to potentially cause P5.p, P6.p and P7.p (and sometimes even P4.p and P8.p) to adopt the primary fate (Sternberg and Horvitz, 1989), and it is likely that this is also the case in *lag-2(q420)* mutants. It is unlikely that the protruding vulva phenotype is a direct consequence of the reduction of LAG-2 present in VPCs during lateral signalling, since it has been shown that other ligands also participate in lateral inhibition by Notch signlling between P6.p, P5.p and P7.p (Chen and Greenwald, 2004), although it is possible that this may exacerbate the abnormalities in vulval development. It has also been demonstrated that the germlines of *lag-2(q420)* mutants can fail to proliferate (Fitzgerald and Greenwald, 1995).

4.2 Results

4.2.1 Quantification of *lag-2(q420)* mutant phenotypes

The penetrance of the Lag phenotype in *lag-2(q420)* worms was studied at the restrictive temperature of 25°C. L4 hermaphrodites were shifted from 15° to 25°C, and the number of their progeny surviving to L2, past the L1 larval arrest point associated with the Lag phenotype, was counted. WT hermaphrodites generally produced over 200 progeny able to develop beyond L1, whereas progeny of *lag-2(q420)* almost invariably arrested at the L1 stage (Figure 4.7). The Lag phenotype was therefore confirmed to be highly penetrant in *lag-2(q420)* mutants.

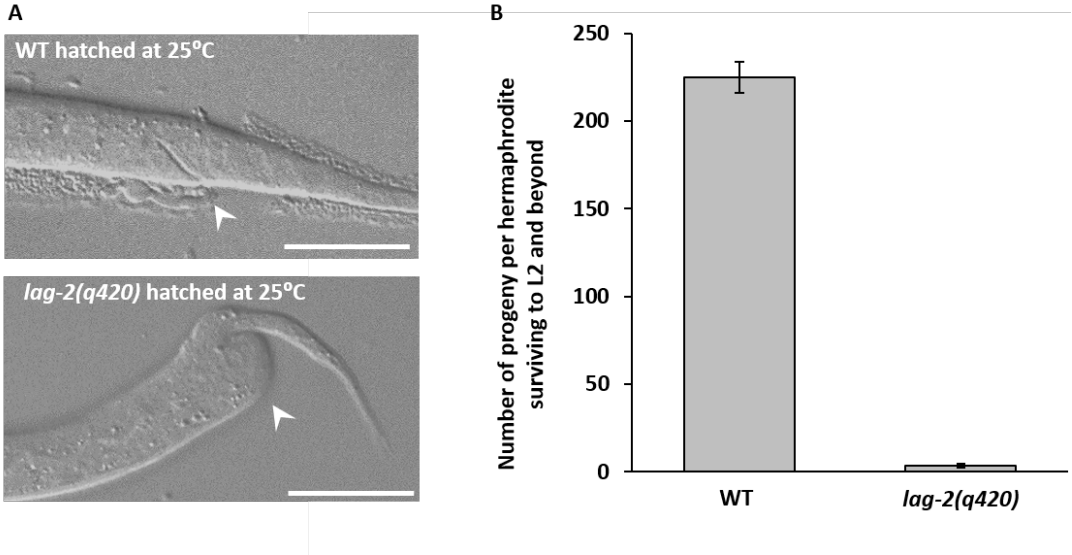


Figure 4.7: The Lag phenotype is highly penetrant in *lag-2(q420)* mutants - (**A**) DIC images of L1 larvae hatched at 25°C with either the presence (WT) or absence (*lag-2(q420)*) of a rectum/anus, as indicated. The absence of an anus in *lag-2(q420)* mutants consequently can result in swelling, as shown. Scalebar = 20 μm. (**B**) The mean number of progeny surviving to past the L1 arrest stage associated with the Lag phenotype is shown for WT and *lag-2(q420)*, where n$\geq$10. Experimental error indicated as $\pm$SEM

Experiments were also carried out to quantify the post embryonic defects, shifting *lag-2(q420)* L1s from 15°C to 25°C (Figure 4.8). In *lag-2(q420)* mutants containing an mCherry AC marker, duplication of the AC occurs in 57.6% of mutant animals (n=168) and 0% (n=50) of WT worms in the same AC marker background. The protruding vulva phenotype occurred in 26.7% of *lag-2(q420)* adults (n=324) compared to 0% of WT worms (n=100). This suggested that not all AC duplications resulted in visibly protruding vulva, although they may still have subtle vulval defects that were not quantified in this assay.

The quantification of germline development defects proved more challenging. Worms were examined for gross defects in germline proliferation. The extent of these defects were variable, and a sample of 50 *lag-2(q420)* mutants included the complete absence of any gonad (indicating failure of the germline to proliferate) in 2 worms, as well as several mutants with shortened gonad arms, however, these could not all be quantified with confidence as the extent of migration in WT worms also varies. Although these defects were dramatic, they did not appear highly penetrant, therefore this particular phenotype was not further pursued.

4.2.2 Rescue of the Lag phenotype of *lag-2(q420)* mutants

The interchangeability of the various Notch ligands during Notch-dependent developmental processes has previously been demonstrated in several rescue studies. It was initially shown that expression of *lag-2* genomic DNA under the control of the *lag-2* 5' flanking region (termed the *lag-2* promoter) was able to rescue the Lag phenotype associated with temperature sensitive *lag-2(q420)* worms. Subsequently, expression of *apx-1* cDNA under the control of the *lag-2* promoter was also shown to be sufficient to rescue the Lag phenotype of *lag-2(q411)* null mutants, as well as the anchor cell duplication phenotype of *lag-2(q420)* temperature sensitive mutants

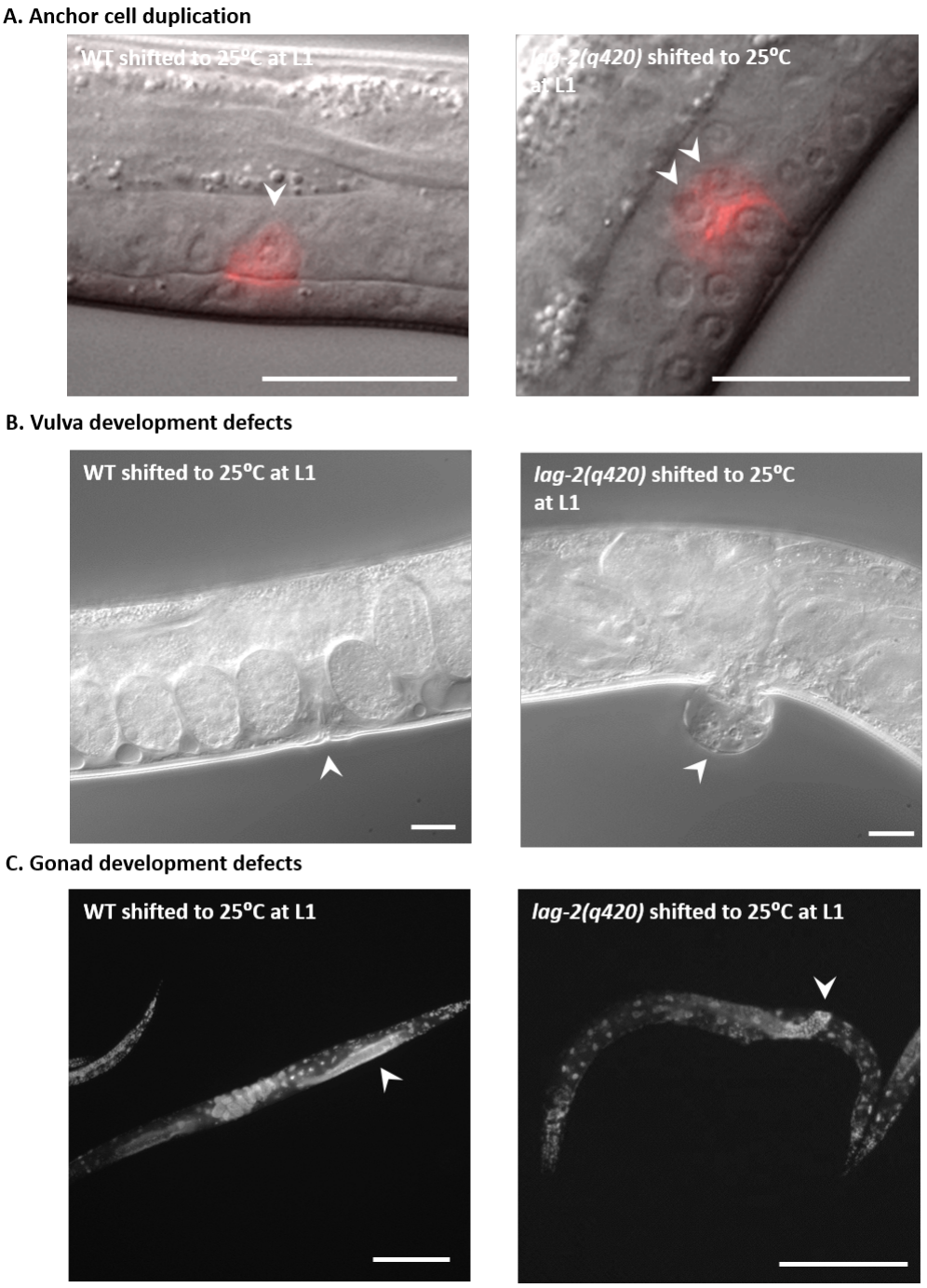


Figure 4.8: Post embryonic developmental defects associated with *lag-2(q420)* mutants - (**A**) *lag-2(q420)* mutants are associated with duplication of the AC, which is visible through the third larval stage (L3), and can be followed in *cdh-3>mCherry::PLC$^{\delta PH}$II?*; *lag-2(q420)V* worms by the expression of the mCherry. (**B**) A consequence of duplication of the AC is the protruding vulval phenotype, due to excessive induction of VPCs. (**C**) In addition, *lag-2(q420)* mutants often have abnormal germline proliferation, including the lack of proliferation of one gonad arm/ shortening of the remaining gonad arm as shown above. Scale bars for A and B = 20 μm, 200 μm for C.

(Gao and Kimble, 1995) (Fitzgerald and Greenwald, 1995). It was also proposed that expression of *dsl-1* cDNA under the *lag-2* promoter could also rescue these defects in the *lag-2(q420)* mutant (Chen and Greenwald, 2004). Due to this evidence demonstrating the ability for these WT ligands to rescue *lag-2(q420)* mutant phenotypes, they were selected for use in this thesis to establish an assay in which the effects of mutations within the ligands could be tested *in vivo*.

4.2.2.1 Generation of WT ligand::gfp constructs for rescue of the Lag phenotype of *lag-2(q420)* mutants

Plasmid rescue constructs included the *lag-2* promoter sequence (previously reported to be sufficient for expression of *lag-2*, *apx-1* and *dsl-1* cDNAs (Gao and Kimble, 1995) (Chen and Greenwald, 2004)) fused to a 42bp synthetic intron at the 5' end of *lag-2*, *dsl-1* or *apx-1* cDNA. cDNA was used since the large intronic regions in the *apx-1* genomic sequence (∼26Kb) make it impractical for use, and ligand cDNA had previously been shown to be sufficient to achieve rescue. Full length cDNAs were fused to the *gfp* gene at their 3' end, to encode C-terminal ligand::GFP fusion proteins, allowing expression of the protein to be verified in transgenic lines by the presence of a green signal in the DTCs of young adult hermaphrodite worms (Figure 4.9). In addition, a ∼30kb fosmid spanning the *lag-2* gene and its flanking regions was also used as a rescue construct, with the anticipation that this genomic fragment (including regulatory sequences) could act as a positive control for rescue. Rescue constructs were injected into young adult *lag-2(q420)* hermaphrodites at low concentration (2 ng/μl), to generate semi-stable extrachromosomal arrays, containing multiple copies of rescue plasmids, using the *unc-119(ed3)III[unc-119$^+$]* (*unc-119* only rescued) background. Three independent transgenic lines, each containing unique extrachromosomal arrays, were isolated for each *ligand::gfp* construct,

to account for variations in array composition. *lag-2::gfp* and *apx-1::gfp* appeared to be expressed in the DTCs from all three transgenic lines of for each construct, however, no green signal could be detected in any transgenic lines generated from injection of the *dsl-1::gfp* rescue construct.

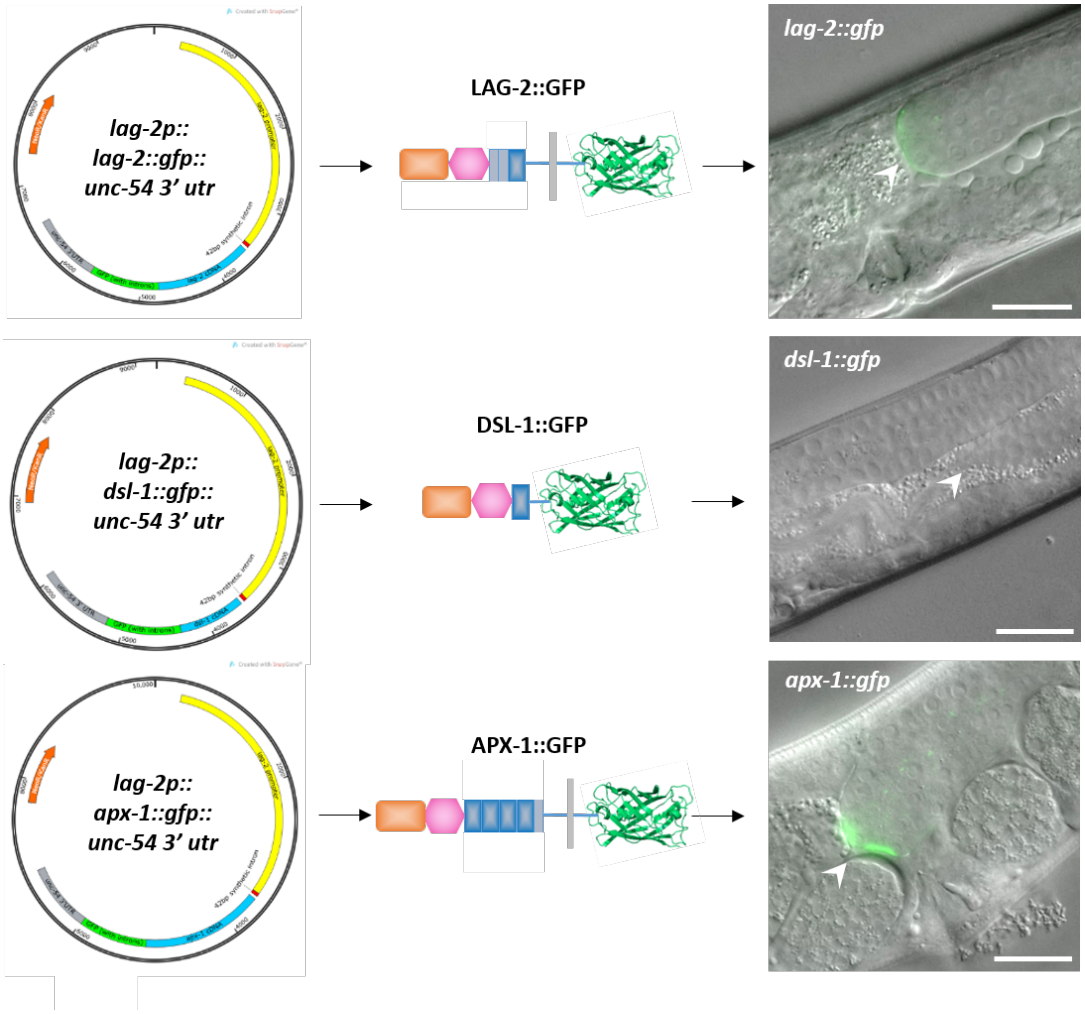


Figure 4.9: Schematic of constructs generated to rescue *lag-2(q420)* mutants - Plasmids containing the *lag-2* promoter and *unc-54* 3'UTR regulatory sequences were used for expression of full length *ligand::gfp* fusions (left), encoding ligand::GFP fusion proteins as shown (centre). Protein domains coloured as shown in Figure 1.4, with GFP structure in green. Right, DIC with FITC FS10 overlay images showed expression of LAG-2::GFP and APX-1::GFP in transgenic lines, but not DSL-1::GFP.

4.2.2.2 WT *lag-2::gfp* and *apx-1::gfp* can partially rescue the Lag phenotype in *lag-2(q420)* mutants

Transgenic *ligand::gfp* lines were studied using a larval viability assay, to detect rescue of the Lag phenotype, the most penetrant phenotype associated with *lag-2(q420)* mutants. Single L4 hermaphrodites were moved from 15°C to 25°C, and the number of their progeny surviving to L2 and beyond (i.e past the point of L1 arrest associated with the Lag worms) was counted. In order to compare independent lines, the transmission efficiency of the extrachromosomal array associated with each line was quantified, and the maximum potential transgenic brood for each line calculated accordingly (Appendix 3). The number of transgenic worms surviving to L2 and beyond within a brood was then expressed as a percentage of this maximum potential transgenic brood. It was noted that *unc-119(ed3)III [unc-119^{+}]* transgenic lines ('*unc-119* only rescued' worms) gave rise to fewer surviving progeny than WT. Therefore, the average percentage of *unc-119* only rescued progeny surviving to L2 and beyond was considered maximum rescue (100%), to account for loss of progeny solely due to the *unc-119* rescue background, and values for all other transgenics were calculated relative to this (Appendix 4).

The ability of transgenic *ligand::gfp* lines to rescue the Lag phenotype in *lag-2(q420)* worms is shown (Figure 4.10). A maximum of 12% of the potential transgenic brood survived to L2 or beyond in lines containing the *unc-119* marker in the *lag-2(q420)* background (*unc-119* only (*lag-2(q420)*)), however, lines expressing *lag-2::gfp*, and to a lesser extent, the *lag-2* fosmid and *apx-1::gfp* all showed an increase in the percentage of progeny surviving. Percentage of progeny surviving ranged from 42-58% in *lag-2::gfp* lines, 30-35% in *lag-2* fosmid lines and 21-44% in *apx-1::gfp* lines. The variability between transgenic lines confirmed the necessity to assay multiple independent lines for each *ligand::gfp* construct. In addition, consis-

tent with the absence of a visible GFP signal, *dsl-1::gfp* did not appear to rescue, with the percentage of the progeny surviving to L2 and beyond comparable to that on the *unc-119* only (*lag-2(q420)*) control.

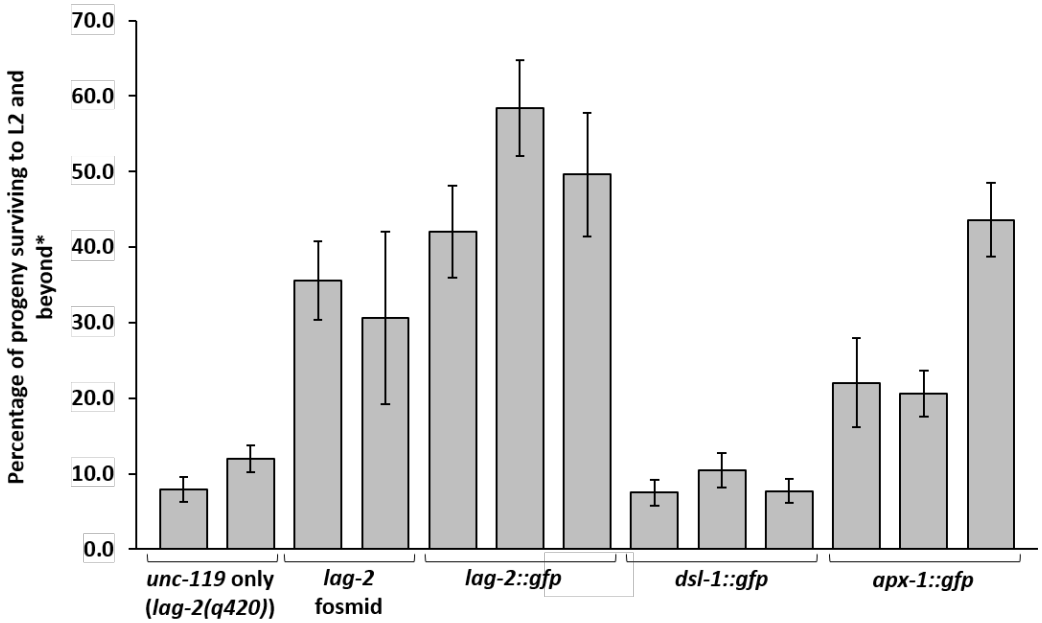


Figure 4.10: Rescue of the Lag phenotype in *lag-2(q420)* mutants by expression of WT Notch ligands - *A larval viability assay was used to count the number of progeny surviving to L2 and beyond, and this data was transformed to account for variations in transmission efficiency and the contribution of the *unc-119* rescue background as described in Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for a single independent transgenic line, for which n≥10. Experimental error indicated as ±SEM

To facilitate an overall comparison between transgenic constructs, the values from each independent line of a single construct (e.g *lag-2::gfp*) were pooled to give a mean of all values for that construct, and Mann-Whitney-U tests performed to determine statistically significant differences between the percentage of progeny surviving to L2 and beyond between the different trangenic constructs (Figure 4.11). Using this particular statistical test, values for *unc-119* only (*lag-2(q420)*) were considered significantly different from the *lag-2* fosmid, *lag-2::gfp* and *apx-1::gfp* with $p \leq 0.001$, supporting the conclusion that these three constructs are expressed

in the transgenic worms and are able to rescue the Lag phenotype to some extent with a high level of confidence. Interestingly, *lag-2::gfp* values were also significantly different from both the *lag-2* fosmid and *apx-1::gfp* values (p=0.007 and p≤0.001, respectively), which suggests that these latter two constructs do not rescue the Lag phenotype as well as *lag-2::gfp*. In addition, the percentage of progeny surviving was not significantly different between *unc-119* only (*lag-2(q420)*) and *dsl-1::gfp* constructs, consistent with a lack of rescue of by *dsl-1::gfp* (p=0.052). Therefore, DSL-1 was not pursued further as a target ligand.

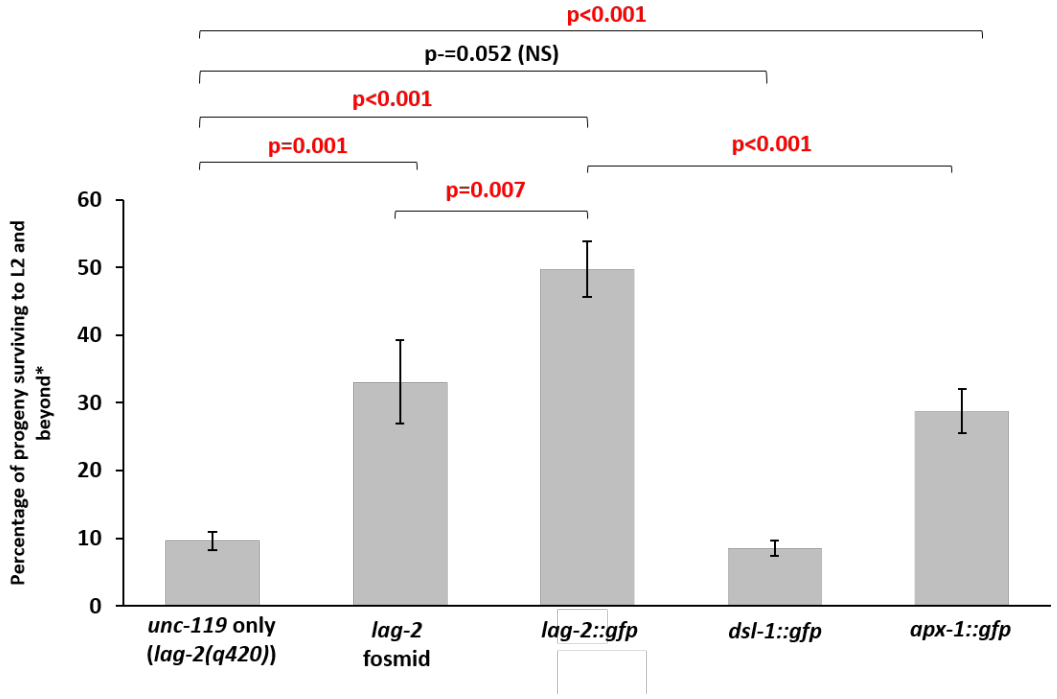


Figure 4.11: Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by expression of WT Notch ligands - *Data was normalised as discussed in section 4.2.2.2 and Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for multiple independent transgenic lines generated for a single construct, for which n≥20. p-values were calculated using a Mann-Whitney-U test, and p≤0.050 was considered statistically significant. Experimental error indicated as ±SEM.

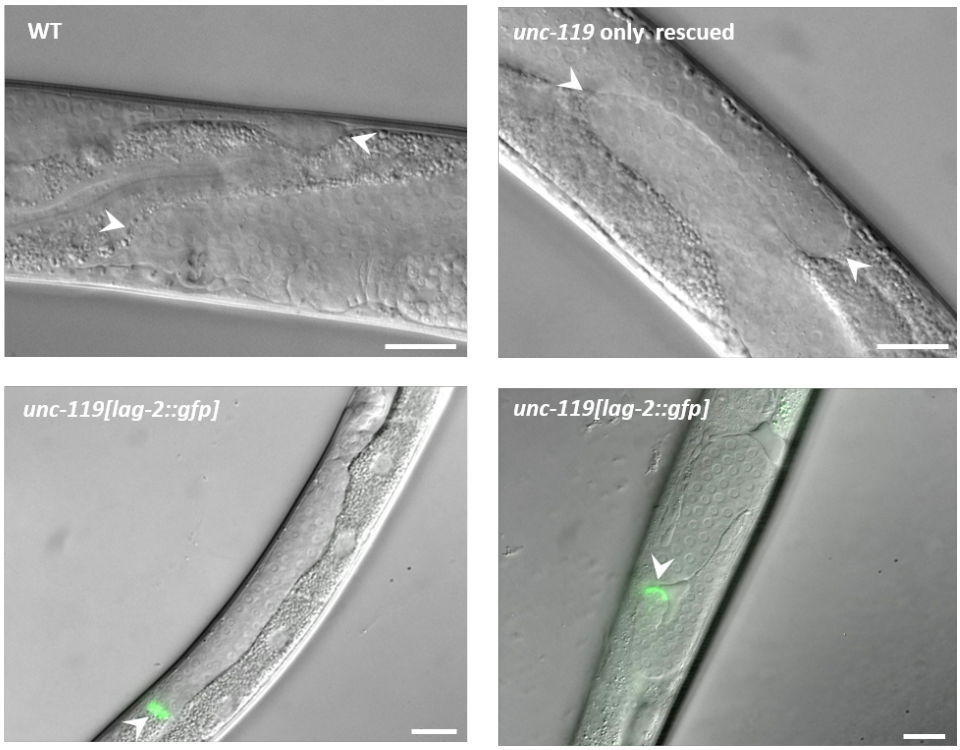
4.2.3 Extrachromosomal *lag-2::gfp* expression can cause gonad and vulval defects

Interestingly, the WT LAG-2 and APX-1 constructs were only able to partially rescue the Lag phenotype in *lag-2(q420)* mutants. Whilst the *lag-2::gfp* construct appeared to be associated with the highest percentage of surviving progeny (42-58%), this was still not the maximum potential progeny anticipated to survive in the case of full rescue of the Lag phenotype (100%). During the generation of transgenic lines, it was observed that a low concentration of *lag-2:gfp* rescue plasmid was necessary to generate transgenic lines. Injections of *ligand::gfp* constructs at 10 ng/μl and 5 ng/μl failed to give rise to transgenic lines, however, lines were successfully isolated at 2 ng/μl. The necessity of reducing the concentration suggested that increased concentrations of *lag-2::gfp* may be toxic, and that perhaps overexpression of this construct had a deleterious effect.

To explore this in further detail, *lag-2::gfp* was co-injected at 2 ng/μl into *unc-119(ed3)III* worms with the *unc-119*$^{+}$ rescue plasmid. This allowed the effects of *lag-2::gfp* expression in the absence of the *lag-2(q420)* background to be observed. Strikingly, despite the lack of the *lag-2(q420)* background, gonad migration defects were observed that appeared wholly dependent on the expression of *lag-2::gfp* (Figure 4.12). In both WT and *unc-119* only rescue worms, the promiximal and distal gonad arms migrated fully by adulthood. However, in *unc-119(ed3)III* worms expressing *lag-2::gfp*, 27% of worms displayed migration defects in the distal gonad arm, including shortening and failure to turn, as well as inappropriate turning/twisting (n=30). All distal gonad arms with defects were also associated with a strong GFP signal at the DTC of the distal gonad arm. The molecular mechanism by which *lag-2::gfp* expression acts to cause these defects is unclear, as LAG-2 has not been shown to be among factors directly involved in the mechanism of gonad

migration, but since the DTCs are specified at the end of L1 and drive gonadogenesis through to adulthood, it is possible that the overexpression of *lag-2::gfp* interferes with the function of the known migration factors in the DTC.

A



B

WT-like gonads (73.3%)		Abnormal gonads (26.7%)	
GFP visible	GFP not visible	GFP visible	GFP not visible
33.3%	40%	26.7%	0%

Figure 4.12: Extrachromosomal *lag-2::gfp* expression can cause gonad migration defects - (**A**) DIC with FITC FS10 overlay. Both the proximal and distal gonad arms migrated normally in WT (top left) and *unc-119* only worms (top right) (n=30 for each), however some *unc-119(ed3)III* worms expressing *lag-2::gfp* arrays displayed a shortened (bottom left) or twisted (bottom right) distal gonad arm. Arrowheads indicated the location of DTCs. All worms cultured at 20°C. Scale bar = 20 μm. (**B**) Summary of presence/absences of visible GFP and gonad defects in *unc-119(ed3)III* worms expressing *lag-2::gfp* arrays (n=30).

Of the 73% of gonad arms that migrated correctly, the GFP signal was not visible in 40% of DTCs observed. This indicated that expression of the extrachromosomal

array was variable between animals, consistent with the variability observed in the larval viability assay. Mosaic expression of extrachromosomal arrays from cell to cell within single animals was identified in early microinjection experiments (Stinchcomb et al., 1985), and is a limitation associated with this method of generating transgenic lines.

Although occurring infrequently, an additional phenotype was observed (Figure 4.13), that supported the hypothesis that *lag-2::gfp* extrachromosomal array expression could be deleterious. The presence of pseudovulvae, ventral protrusions arising from vulval tissue, in addition to the true vulva, was first observed in *lin-12(gf)* mutants (Greenwald et al., 1983). That this phenotype can also be induced by *lag-2::gfp* array expression suggested that *lag-2::gfp* extrachromosomal array expression could lead to an over-activation of the Notch pathway during vulval development.

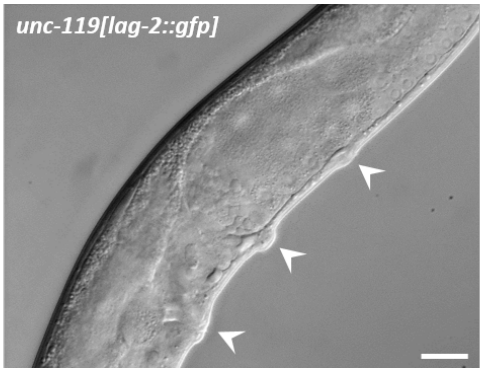


Figure 4.13: Extrachromosomal *lag-2::gfp* expression can lead to formation of pseudovulvae - Two additional ventral protrusions of vulval tissue (pseudovulvae) flanking the true vulva (centre arrow) were observed in 3% *unc-119(ed3)III* worms expressing *lag-2::gfp* arrays (n=30). Scale bar = 20 μm.

4.2.4 Presence of the C-terminal GFP tag does not appear to alter the extent of rescue LAG-2

In light of the observation that *lag-2::gfp* array expression appeared to be able to contribute to developmental defects, the effects of the presence of the C-terminal

GFP tag were studied to exclude the possibility that this had caused the reduction in larval viability. All three ligands in this study were expressed as full length cDNAs, directly fused to *gfp* in place of their stop codon. Full length cDNAs were used to ensure that lysine residues present at the intracellular C-terminus of Notch ligands were available for ubiquitination and subsequent recycling (Parks et al., 2000) (Heuss et al., 2008), to facilitate endogenous protein regulation. The direct fusion was chosen as it had been previously demonstrated that fusion of DLL-4 and Jag-1 to CFP both directly and with a short linker did not appear to alter its functionality *in vivo* (LeBon et al., 2014).

The full length *lag-2* cDNA sequence (including the native stop codon) was placed under the control of the *lag-2* promoter, and *unc-54* 3' UTR regulatory sequences, in the absence of the *gfp* sequence. Transgenic lines were then generated as before, and three independent lines isolated and assessed for rescue of the Lag phenotype in *lag-2(q420)* mutants using the larval viability assay. The results indicated that the extent of rescue of the Lag phenotype was not dependent on the presence or absence of the C-terminal GFP tag (Figure 4.14). Analysis of the combined values from all *lag-2::gfp* and *lag-2* only lines indicated that there was no statistically significant difference (p=0.264) between the percentage of progeny surviving to L2 and beyond for the two constructs (Figure 4.15), confirming that the inclusion of the GFP tag was not the source of the overexpression effects observed for *lag-2::gfp* expression.

4.2.5 Dilution of *lag-2::gfp* is not sufficient to enhance rescue of the Lag phenotpye in *lag-2(q420)* mutants

It has been demonstrated that extrachromosomal arrays can carry 10s-100s of copies of an injected DNA construct (Stinchcomb et al., 1985) (Mello et al., 1991), and

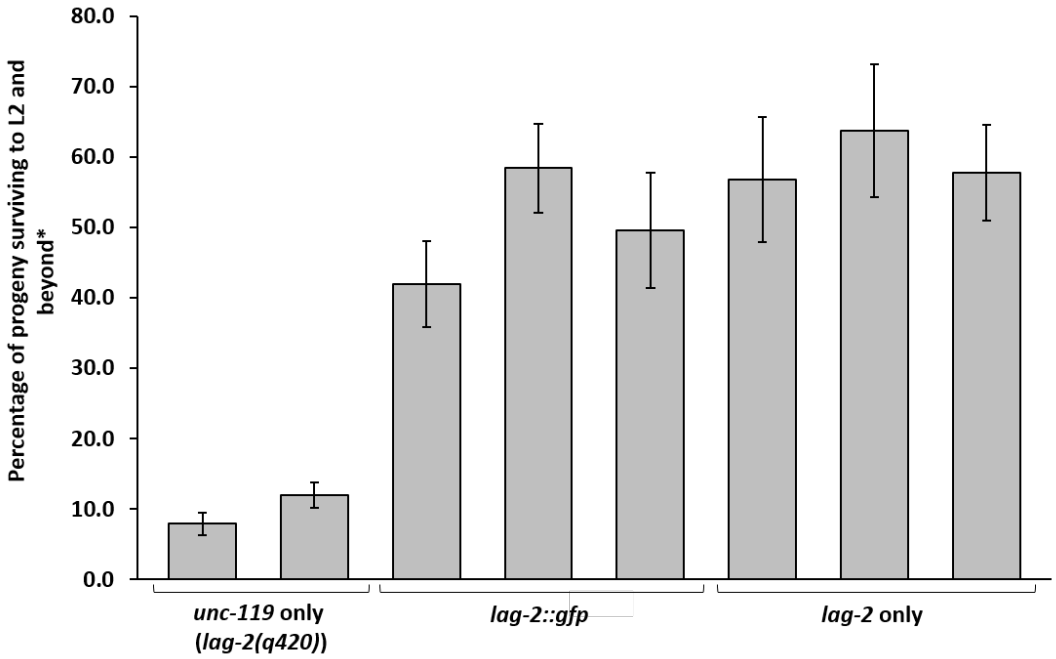


Figure 4.14: Absence of the GFP tag does not affect the extent of rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 - *Data was normalised as discussed in section 4.2.2.2 and Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for a single independent transgenic line, for which n≥10. Experimental error indicated as ±SEM.

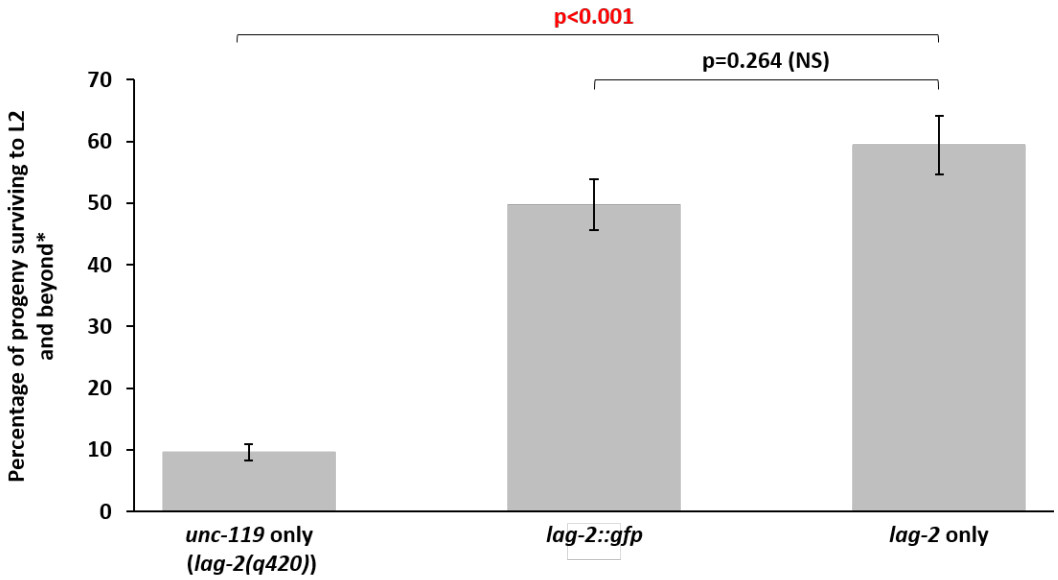


Figure 4.15: Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 with and without the GFP tag - *Data was normalised as discussed in section 4.2.2.2 and Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for multiple independent transgenic lines generated for a single construct, for which n≥20. p-values were calculated using a Mann-Whitney-U test, and p≤0.050 was considered statistically significant. Experimental error indicated as ±SEM.

whilst in many cases the elevated expression of genes above the endogenous level can still permit rescue, overexpression of genes in extrachromosomal arrays has also been shown to have deleterious effects (Fire and Waterston, 1989). One method to reduce copy number of a given construct is to dilute it with other DNA, so that fewer copies are likely to incorporate into the final array (Mello et al., 1991). Therefore, the *lag-2::gfp* construct was diluted 100 fold with Bluescript plasmid DNA, and injected at a final concentration of 1 ng/μl. Three independent lines were isolated, two of which were confirmed to express *lag-2::gfp* by the presence of GFP in the DTC. These two lines were assessed for their ability to rescue the Lag phenotype as before (Figure 4.16). The percentage of progeny surviving to L2 and beyond for these transgenic lines of diluted *lag-2::gfp* and injected at 1 ng/μl varied from 37-45% , compared to 42-58% in *lag-2::gfp* undiluted and injected at 2 ng/μl . This would suggest that attempts to lower the copy number of *lag-2::gfp* were unsuccessful in overcoming the effects of expression of multiple copies of this construct from an extrachromosomal array. No significant difference was detected between the combined the percentage of progeny surviving to L2 and beyond for the 2 ng/μl *lag-2::gfp* and 1 ng/μl diluted *lag-2::gfp* lines (p=0.056)(Figure 4.17), therefore, dilution of *lag-2::gfp* did not improve the extent of rescue of the Lag phenotype in transgenic worms.

4.2.6 *lag-2::gfp* expression is associated with increased embryonic lethality in *lag-2(q420)* mutants

To understand more fully why only partial rescue of the Lag phenotype by *lag-2::gfp* expression was observed, the fates of all progeny within a single brood were observed for transgenic lines expressing *lag-2::gfp* in different genetic backgrounds. L4 hermaphrodites were shifted from 15°C to 25°C and the number of progeny

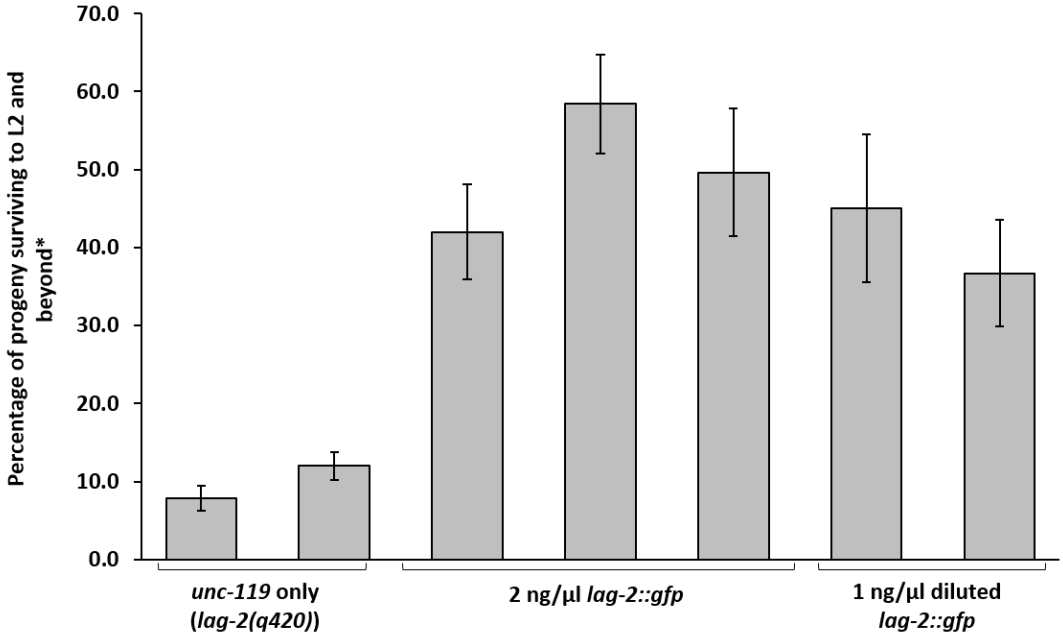


Figure 4.16: Dilution of *lag-2::gfp* does not increase the extent of rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 - *Data was normalised as discussed in section 4.2.2.2 and Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for a single independent transgenic line, for which n≥10. Experimental error indicated as ±SEM.

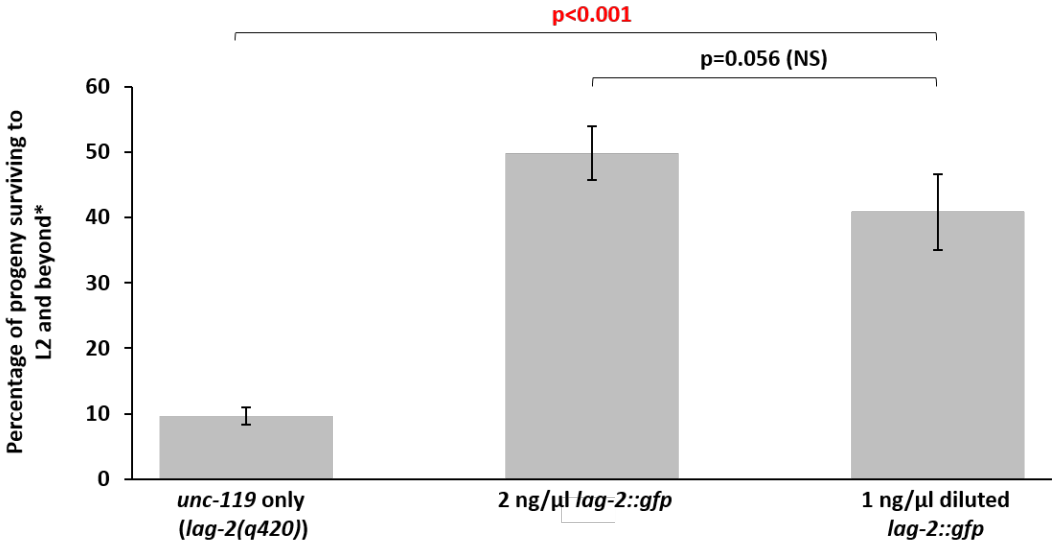


Figure 4.17: Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by LAG-2 with and without the GFP tag - *Data was normalised as discussed in section 4.2.2.2 and Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for multiple independent transgenic lines generated for a single construct, for which n≥20. p-values were calculated using a Mann-Whitney-U test, and p≤0.050 was considered statistically significant. Experimental error indicated as ±SEM.

that hatched into larvae, as well as the number that died as eggs was recorded, and expressed as a percentage of the total number of eggs laid. The results were surprising, and indicated that expression of *lag-2::gfp* arrays was associated with an increase in embryonic lethality, that could not be fully attributed to the *lag-2(q420)* allele or the *lag-2(q420)* allele in combination with the *unc-119* rescue background (Figure 4.18). This effect seemed to be particularly striking in *lag-2::gfp* transgenic line 3. In addition, the total number of eggs laid was fewer (100-162) compared to the *lag-2(420)* mutants (197-218) and the *unc-119* only (*lag-2(q420)*) control line (180-246) (Appendix 5).

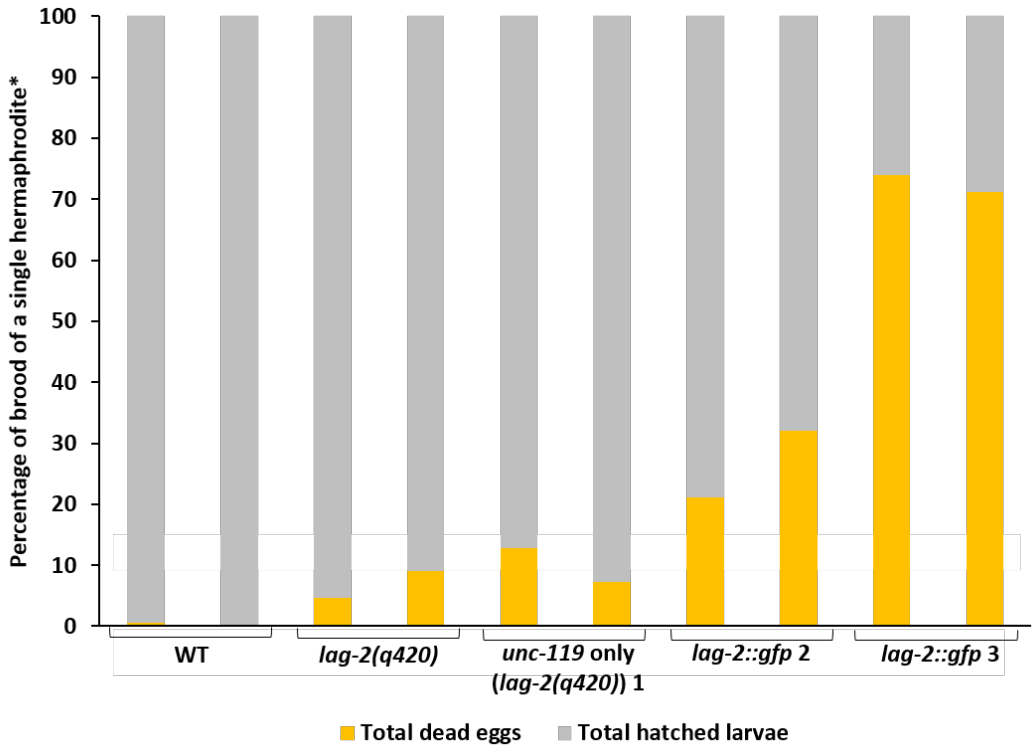


Figure 4.18: *lag-2::gfp* extrachromosomal array expression is associated with increased embryonic lethality in *lag-2(q420)* mutants - *Percentage of the brood of a single hermphrodite was calculated by expressing the total number of dead eggs (orange) or total number of hatched larvae (grey) out of the total number of eggs laid (Appendix 5). Each bar represents the brood of a single hermaphrodite.

Additional brood fate assays were carried out using lines in the absence of the

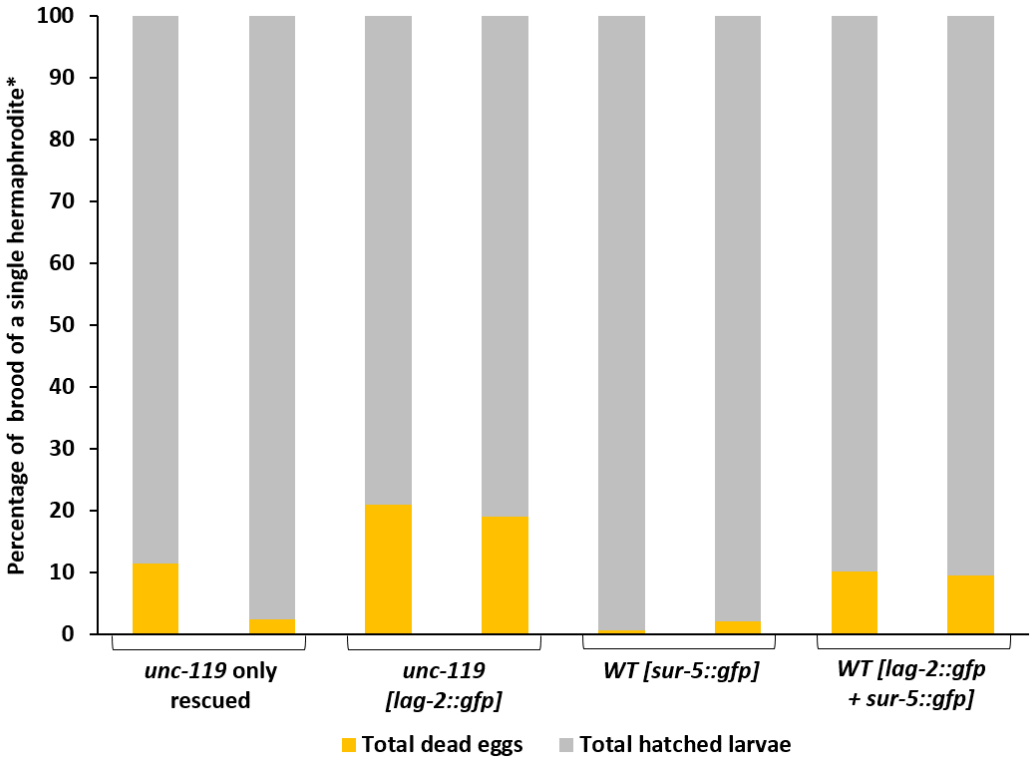


Figure 4.19: The effects of *lag-2::gfp* extrachromosomal array expression in the absence of the *lag-2(q420)* allele - *Percentage of the brood of a single hermphrodite was calculated by expressing the total number of dead eggs (orange) or total number of hatched larvae (grey) out of the total number of eggs laid (Appendix 5). Each bar represents the brood of a single hermaphrodite.

lag-2(q420) allele (Figure 4.19). Low levels of embryonic lethality were detected in the *unc-119* only (*lag-2(q420)*) rescue background control line, however an increase in lethality was observed in *unc-119(ed3)III* worms that expressed *lag-2::gfp (unc-119 [lag-2::gfp])*. When *lag-2::gfp* was expressed in WT worms with a *sur-5::gfp* co-injection marker, a small increase in embryonic lethality was also observed, compared to the absence of *lag-2::gfp* expression. Taken as a whole, these brood fate data suggested that *lag-2::gfp* array expression can lead to an increase in embryonic lethality, and this may be exacerbated in the presence of the *lag-2(q420)* allele.

4.3 Discussion

The results of these experiments demonstrated that WT LAG-2 and APX-1 Notch ligands were able to rescue the Lag phenotype, consistent with published studies. Surprisingly, a full length *dsl-1::gfp* failed to rescue the Lag phenotype in the larval viability assay. It had been previously reported by Chen and Greenwald that *dsl-1* was able to rescue the Lag phenotype in *lag-2(q420)* mutants, however, of the four transgenic *dsl-1* lines generated, only one was able to rescue the Lag and duplicated AC phenotypes associated with *lag-2(q420)* mutants (Chen and Greenwald, 2004). In addition, the authors demonstrated rescue of the phenotypes only in transgenics containing a *smg-1* mutant genetic background. *smg-1* encodes a phosphatidylinositol kinase-related protein involved in NMD of mRNA transcripts, and *smg-1* mutants are used to generate transcriptional reporter constructs, by inhibiting NMD of the mRNA of the inserted reporter gene when the native stop codon is included. It is not clear how the presence of the *smg-1* background could effect the expression of *dsl-1* cDNA from an extrachromosomal array, athough it is important to note differences in genetic backgrounds between previous studies and the data presented in this thesis, to identify potential sources of discrepancies between the

results.

Additional interesting observations from the data presented in this chapter were significant differences between the extent of rescue of the Lag phenotype by *lag-2* fosmid and *apx-1::gfp* compared to *lag-2::gfp*. Rescue of the Lag phenotype was reduced in both constructs compared with *lag-2::gfp*. In the case of *apx-1::gfp*, this was less surprising, considering that whilst LAG-2 and APX-1, share some functional overlap, they participate individually in distinct cell fate decisions and developmental processes (section 4.1.2). Alignment of the DSL domain of the two ligands also indicates the lack of a strictly conserved motif, despite some common residues (section 3.1.2), and this would also support the hypothesis that residues involved in binding of the two ligands to LIN-12/GLP-1 may vary to some extent. It may be that the affinity of the interaction between LAG-2 and APX-1 with Notch receptors differs, and this may influence the extent to which APX-1 can substitute for LAG-2 in the developmental processes associated with the Lag phenotype. Expression of recombinant fragments of both the ligands and LIN-12 and GLP-1 followed by biophysical studies such as SPR would be necessary to detect differences in binding affinities of the ligand/receptor interactions, and quantify them accordingly.

The observation that the *lag-2* fosmid was able to rescue the Lag phenotype in *lag-2(q420)* worms to a lesser extent than a *lag-2* cDNA construct is harder to rationalise. The *lag-2* fosmid had been included in the generation of transgenics to provide a construct in which *lag-2* was in it's native environment (including full 5' and 3' regulatory regions), and it was anticipated that this would result in the maximum extent of rescue of the Lag phenotype in *lag-2(q420)* mutants. However, this was not the case. It is possible that by including the full regulatory regions, sequences were included that prevented expression of genomic *lag-2* gene to the same extent as the shorter *lag-2::gfp* construct expression. It had been shown that

regions up to ∼7kb upstream of the *lag-2* gene can influence its expression in specific cell fate decisions, and it is likely that there are additional regions of regulatory significance that have yet to be identified (Zhang et al., 2011). In addition, the *lag-2* fosmid spans a ∼30kb region of the genome, and includes several other genes, generally encoding GPCRs (e.g *srab-5*, *srab-8*, *srbc-6*). Since multiple copies of all the genes associated with the fosmid will be incorporated into the extrachromosomal arrays, it cannot be ruled out that overexpression of additional fosmid genes may have an affect on viability. Another consideration of the differences between the genomic *lag-2* fosmid and *lag-2::gfp* constructs is the presence of a synthetic intron immediately upstream of the start site of *lag-2::gfp*. Synthetic introns have been previously shown to substantially increase expression of transgenes (Okkema et al., 1993) and it is possible that the inclusion of the 42bp synthetic intron prior to the *lag-2* cDNA sequence increases expression compared with *lag-2* fosmid expression, although this seems paradoxical considering subsequent observations that increased *lag-2* expression appears to lead to overexpression effects and embryonic lethality, impacting the larval viability and hence rescue of the Lag phenotype. It would be interesting to compare the copy numbers of the *lag-2* in the extrachromosomal arrays of both the *lag-2* fosmid and *lag-2::gfp* lines using a quantitative PCR approach, to help reconcile these data.

It was striking that expression of *lag-2::gfp*, was only able to partially rescue the Lag phenotype associated with *lag-2(q420)* mutants in the larval viability assay, and subsequent experiments indicated that embryonic lethality associated with the *lag-2(q420)* allele may be contributing to this. It should be noted that a limitation of the brood fate assay was that multiple samples could not easily be assessed in parallel, therefore only a small sample number (n=2) was obtained for each line, however, the percentage of embryonic lethality associated with each transgenic appeared to

be consistent between samples. The previous studies demonstrating rescue of the Lag phenotype in *lag-2(q420)* (Gao and Kimble, 1995) and *lag-2(q411)* (Fitzgerald and Greenwald, 1995) mutants did not report full quantitative data for the extent of this rescue, and therefore, cannot be compared to the data collected in this work. Whilst neither study reported the occurrence of overexpression effects, including embryonic lethality associated with ligand expression from extrachromosomal arrays, an increase in embryonic lethality had been observed in *lin-12(lf)* mutants (Lambie and Kimble, 1991), highlighting the importance of Notch signalling in embryonic development. Target genes of Notch signalling in the embryo include members of the *ref-1* gene family, which share homology with HES genes, the Notch targets of *Drosophila* and mammals (Neves and Priess, 2005). It is likely that the consequence of too much, or too little of these genes during development could both be deleterious for the embryo.

Interestingly, in addition to the increase in embryonic lethality observed in *lag-2::gfp* transgenic worms, there appeared to be a decrease in the total number of eggs laid in transgenic lines compared with WT or *lag-2(q420)* worms. One factor that may have contributed to this is germline proliferation defects, which could occur as a consequence of either an increase or decrease of Notch signalling, and which would be consistent with the presence of abnormal gonads observed in *unc-119[lag-2::gfp]* worms. Detailed studies of the germ lines of the transgenic worms to look for either a reduction or increase in the number of mitotic cells in the distal germline would need to be carried out to determine if this was indeed the case.

When normalising the larval viability data to account for the effects of transmission efficiency and the *unc-119* rescue background, the decrease in eggs laid by *unc-119* only rescued worms (an average of 142) was accounted for. A subsequent assumption made when representing the larval viability data was that all transgenic

brood sizes were considered equivalent to this, however, data from the transmission efficiency and brood fate assays suggested that the brood size was generally reduced even further in worms expressing the *ligand::gfp* rescue constructs. This would therefore make the percentage of progeny surviving to L2 and beyond an underestimate for these lines. A more accurate measure of rescue by each individual line would be obtained by accounting for the variations in brood sizes, therefore, the brood fate assay, whilst more labour-intensive than the larval viability assay would have given a more accurate measure of the percentage of progeny surviving to L2 and beyond, and thus, the extent of rescue by each construct.

The aim of these experiments was to establish an *in vivo* assay in which the effects of mutations in Notch ligands could be explored. However, the limitations found when assaying the rescue of the Lag phenotype in *lag-2(q420)* mutants, was the presence of some line-to-line variability between transgenics, and the apparent overexpression effects of Notch ligand rescue constructs from extrachromosomal arrays. For these reasons, it was felt that any subtle effects of mutations would not be detected using rescue of the Lag phenotype in *lag-2(q420)* mutants via extrachromosomal arrays expression. Whilst the results posed additional questions as to the exact nature of the overexpression effects, these were not pursued since experimental efforts were concentrated on developing an alternative approach that could avoid these limitations and provide a more sensitive system in which to test Notch ligand mutants.

5

Single copy insertion and promoter considerations

5.1 *Mos-1* mediated single copy insertion as an alternative to extrachromosomal arrays

5.1.1 Introduction

Mos-1 mediated single copy insertion (mosSCI) is a method for generating worms in which a single copy of a given transgene is inserted onto a chromosome, as opposed to being maintained as an extrachromosomal array (Frø kjaer Jensen et al., 2008) (Frøkjæ r Jensen et al., 2012). The advantages of mosSCI over the approach of generating extrachromosomal arrays is that it removes the concern of transgene overexpression due to high copy number, it is not associated with mosaic expression, and it is inherited by 100% of progeny. Therefore mosSCI was considered highly suitable to establish an *in vivo* assay to probe Notch ligand function, by assessing rescue of the Lag phenotype in *lag-2(q420)* worms.

The general scheme for mosSCI, using insertion of the *lag-2::gfp* rescue construct into *ttTi5605 II*; *unc-119(ed3)III* worms as an example, is summarised in Figure

5.1. The initial step of mosSCI relies on the generation of an extrachromsomal array, by microinjection of the insertion construct plasmid along with additional plasmids including a transposase under the control of the heat shock promoter, and two selection markers, to aid screening at a later step in the protocol. Injected worms are placed at 25°C to activate *hsp::transposase* expression which leads to the production of transposase enzyme. This excises the *mos-1* element on chromosome II of *ttTi5605* worms, generating a double strand break. This is then repaired using the insertion construct present in the extrachromasomal array as a template, by virtue of homologous regions flanking both the *mos-1* site of *ttTi5605* worms and the rescue construct of the plasmid. Once the site is repaired with the homologous region, worms carrying the insert can be identified by their WT movement, due to the co-insertion of the *unc-119*$^{+}$ also present in the rescue construct. Worms which have not lost the redundant extrachromosomal array can be eliminated by heat shock at 30° to activate expression of *peel-1*, encoding the PEEL-1 toxin, and the surviving worms can also be screened for the absence of the mCherry selection markers included in the initial injection. Non-red movers (NRM) can be subsequently analysed using genomic extraction and PCR, followed by sequencing to verify the insertion. Ultimately, worms carrying single *lag-2::gfp* insertions could then be crossed to introduce the *lag-2(q420)* allele, and the larval viability assay used to assess rescue of the Lag phenotype.

5.1.2 Results

The *lag-2::gfp* mosSCI rescue construct was generated using Gateway cloning, in order to insert the region into the pCFJ150 vector, which includes a 5' *unc-119*$^{+}$ marker gene, and right/left *mos-1* flanking regions that are homologous to the flanking regions at the *mos-1* insert site on chromosome II of *ttTi5605* worms. Although

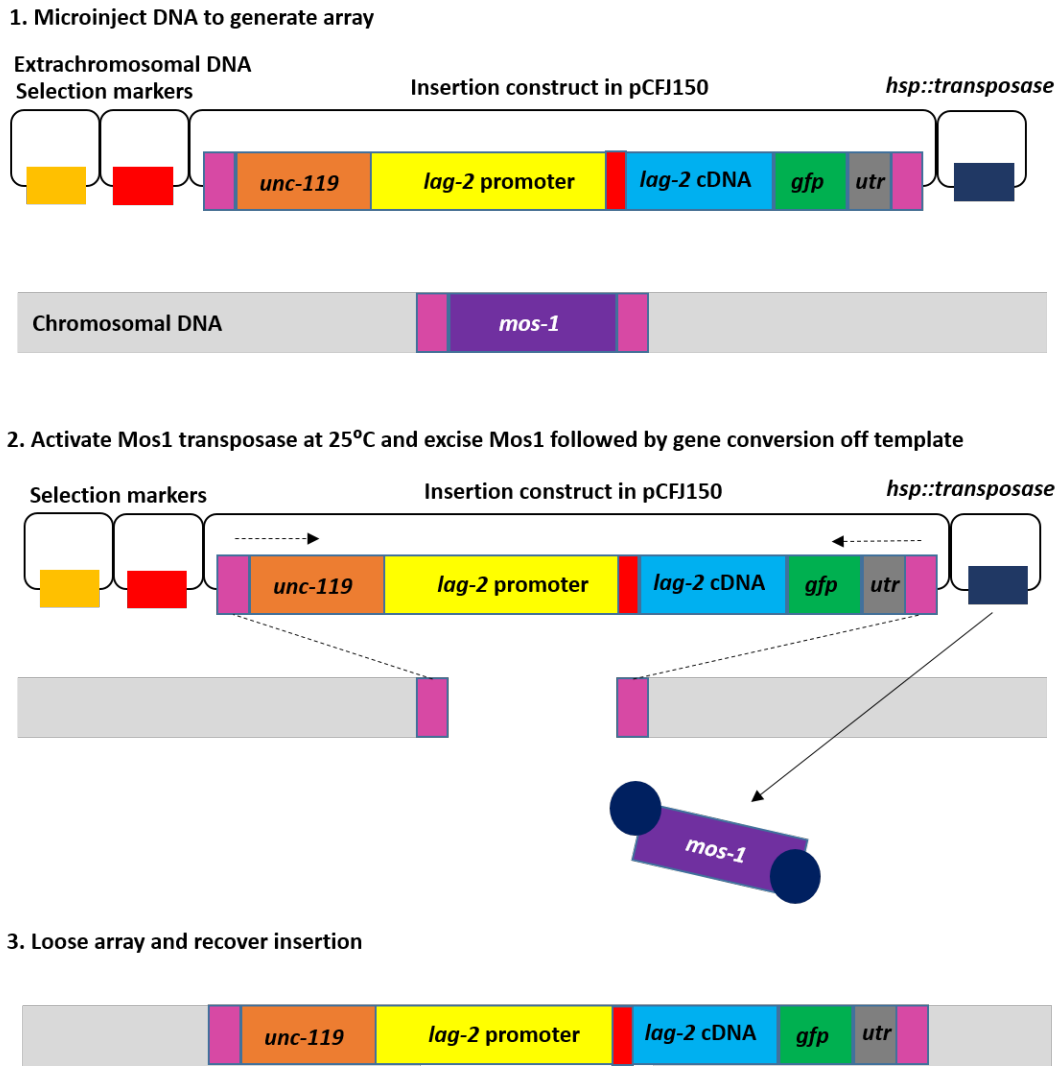


Figure 5.1: Schematic of the generation of single copy insertions using mosSCI - (**1**) Multiple plasmids, including the plasmid bearing the region targeted for insertion, as well as selection markers and heat shock-induced transposase are mircoinjected to form an extrachromosomal array. (**2**) Activation of the transposase transgene at 25° leads to the expression of transposase enzyme, which removes the *mos-1* from the chromosomal DNA of *ttTi5605 II*;*unc-119(ed3)III* worms, and regions of homology (pink) allow the site to be repaired using the insertion construct. (**3**) Loss of selection markers associated with the extrachromsomal array allows the worms carrying the insertion to be identified. Adapted from www.wormbuilder.org.

entry clones were successfully generated for the equivalent *apx-1::gfp* construct, these could not be recombined into pCFJ150. Sequence analysis revealed a 12 bp region (TGTACAAAAAAG) occurring in the native sequence of *apx-1* cDNA that is also present as part of the *attL1* recombination site. It is likely that this homologous region interfered with the success of the LR reaction, and therefore only *lag-2::gfp* in pCFJ150 was taken forward for injection into *ttTi5605* worms.

To generate extrachromosomal arrays for the first step of the scheme, the *lag-2::gfp* construct was injected at 10 ng/μl, after injection of $\sim$90 animals failed to generate insertions, the concentration was lowered to 2 ng/μl and 1 ng/μl, injecting $\sim$140 and $\sim$80 animals, respectively. However, only a small number of transformants could be detected in the F1 progeny of any of these animals, none of which ultimately yielded insertions. After purifying fresh aliquots of all DNA constructs, a further $\sim$90 worms were injected at the original concentration of 10 ng/μl *lag-2::gfp*, and a single non-red mover (NRM) identified from these animals. The progeny of this NRM were then analysed. Single worm lysis followed by PCR amplification of the *unc-119*$^{+}$ gene confirmed the presence of an insertion in the NRMs (Figure 5.2A), however, genomic extraction followed by PCR amplification of the entire insertion region indicated that the insertion was incomplete (Figure 5.2B). Two primer pairs were used to amplify the *lag-2::gfp* insertion and flanking chromosomal regions, that would give PCR products of 10 and 11.4 kb, however, the observed PCR products appeared to be significantly shorter ($\sim$5 and 7 kb, respectively). Excision of the DNA bands and subsequent sequencing revealed that the insertion present was truncated, and consisted of the *unc-119*$^{+}$ gene, the first 229 bp of the *lag-2* promoter sequence, and a short upstream region of the *mos-1* element (Figure 5.2C).

To test the validity of the methods and reagents used to attempt to generate a

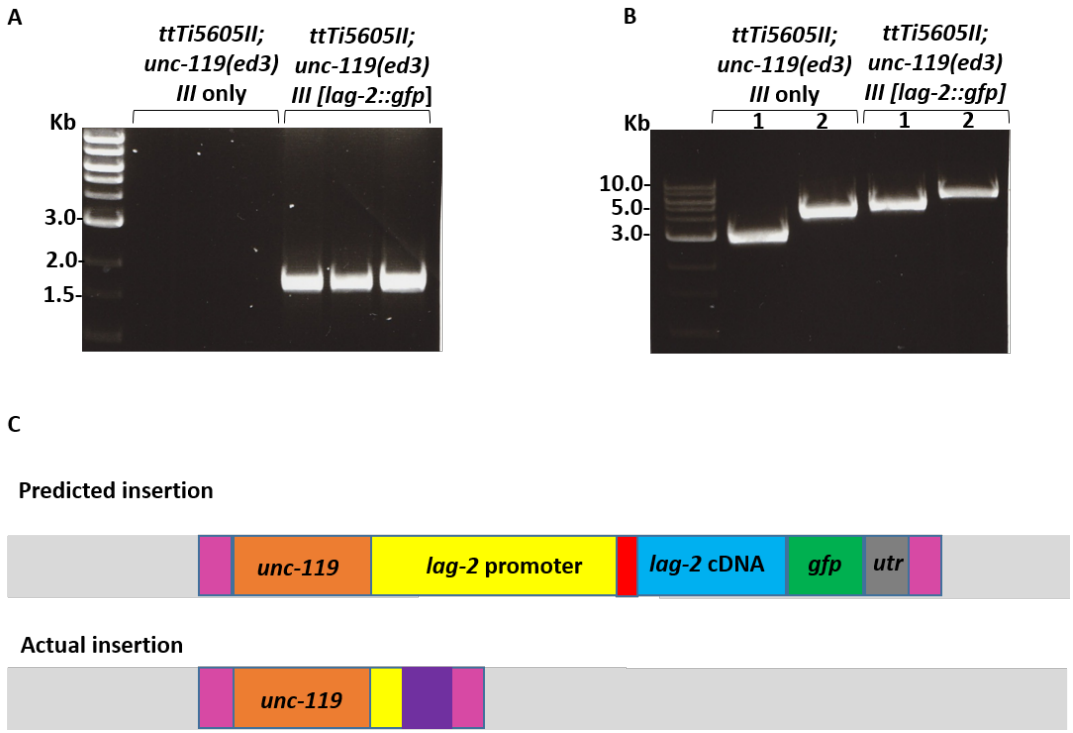


Figure 5.2: Analysis of the *ttTi5605 lag-2::gfp* insertion - (**A**) PCR amplification using *unc-119*$^{+}$ primers allows the presence of an insertion in *ttTi5605* worms to be detected by generating a 1.8 kb product that is absent in *ttTi5605* worms with the *mos-1* transposon intact. Triplicates are shown. (**B**) PCR amplification of the insertion region using primers from outside the homologous *mos-1* flanking region results in a PCR product of 3 and 4.2 kb (primer pairs 1 and 2, respectively), and PCR products of 10 and 11.4kb in *lag-2::gfp*, however, the *lag-2::gfp* products are approximately 5 kb smaller than anticipated. (**C**) Cartoon representation of the anticipated and actual insertion sequences, with the actual sequence including a small fragment of the promoter followed by a short region of the *mos-1* element.

lag-2::gfp insertion, an insertion construct that had previously been successfully used to generate single copy insertions was used. The *pSeam::unc-62cDNA::gfp::unc-54 3'utr* in pCFJ150 (pAW794, a gift from S. Hughes) plasmid was amplified and purified, and injected into ∼80 *ttTi5605* animals at 10 ng/μl, yielding one NRM. Single worm lysis followed by PCR of the *unc-119*$^{+}$ gene confirmed the presence of an insertion (Figure 5.3A), and genomic extraction followed by PCR amplification of the insertion using two primer pairs (to give 8.1 and 9.2 kb products) confirmed that the full length insertion was present (Figure 5.3B and C). This was verified by excision of the DNA bands, and sequencing. Expression of the insertion was confirmed by the presence of GFP in the seam cells (Figure 5.3D). This demonstrated that the methods and reagents (i.e, the worm strain and co-injected plasmids) were not the cause of the inability to generate a full length single copy insertion of *lag-2::gfp*, and suggested that the issue lay within the *lag-2::gfp* construct itself. Injection of the *lag-2::gfp* construct into *unc-119(ed3)III* worms allowed transgenic worms to be isolated, in which GFP was present at the DTC, therefore the *lag-2::gfp* construct was supported in an extrachromosomal array (Figure 5.3E).

5.2 The *lag-2* promoter - analysis and outcomes

The truncation of the *lag-2* promoter in the only *lag-2::gfp* insert isolated in the above experiments suggested that the promoter region may be contributing to the inability to generate full length insertions. The reasons for this were not clear, although one possibility was that sequences present in the *lag-2* promoter region may interfere with the gene conversion event. It was previously noted that the *pie-1* promoter yielded mosSCI insertions with reduced frequency compared to other promoter regions, and this was attributed to the presence of several inverted repeats which could interfere with synthesis-dependent strand annealing during gene con-

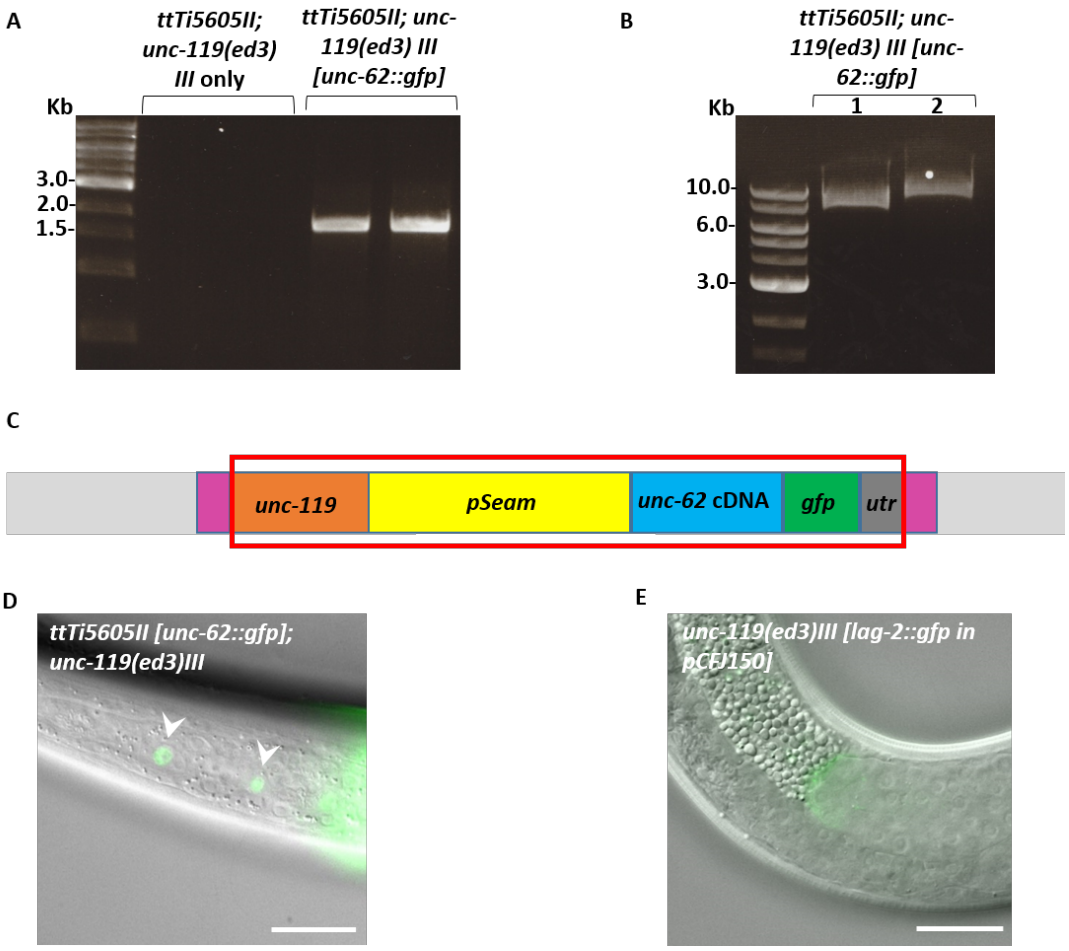


Figure 5.3: Analysis of the *ttTi5605 unc-62::gfp* insertion - (**A**) PCR amplification using *unc-119*$^{+}$ primers allows the presence of an insertion in *ttTi5605* worms to be detected by generating a 1.8 kb product that is absent in *ttTi5605* worms with the *mos-1* transposon intact. Duplicates are shown. (**B**) PCR amplification of the insertion region using primers from outside the homologous *mos-1* flanking region results in a PCR product of 8 and 9.2 kb in *unc-62::gfp*. (**C**) Cartoon representation of the successfully inserted *unc-62::gfp* sequence (**D**) DIC with FITC FS10 overlay, demonstrating expression of *unc-62::gfp* single copy insertions in the seam cells of an L3 worm. (**E**) DIC with FITC overlay expression of *lag-2::gfp* in pCFJ150 in *unc-119(ed3)III* worms. Scale bar = 20 μm.

version (Frø kjaer Jensen et al., 2008). An online inverted repeat prediction tool was used to analyse the *lag-2* promoter, and predicted the presence of an inverted repeat of 37 nucleotides (sequence aaattatgaaaaatctaaaatttttggagtttttta) present 3.1 and 1.2 kb upstream of the *lag-2* gene. The 3.1 kb site was within 100 bp of the point at which the *lag-2* promoter was truncated in the one incomplete insertion obtained, suggesting that this inverted repeat region may indeed have impaired successful gene conversion. Subsequent experiments were designed based on the assumption that this repeat region was the limiting factor to generating full length insertions.

To circumvent potential issues caused by the presence of the inverted repeat, a reduced fragment of the promoter was designed that excluded this region. Comparison of the *lag-2* promoter region in different *Caenorhabditis* species did not reveal clear regions of conserved regulatory motifs (Appendix 6). Therefore, Wormbase (www.wormbase.org) was used to highlight the known transcription factor binding regions within the full 3.2 kb *lag-2* promoter fragment (Figure 5.4). This suggested that the transcription factor binding regions of the promoter were clustered within 1 kb upstream of the *lag-2* gene start site. Additional transcription factor binding sites, mainly for the forkhead transcription factor UNC-130 and HLH-2, a basic-helix-loop-helix (bHLH) transcription factor were clustered within ∼1.6 kp upstream of the start site. Further HLH-2 binding sites, a were clustered 2.3 to 3.1 kb upstream of the lag-2 gene. Whilst the binding of HLH-2 to E-box motifs (CAANTG) has been shown to directly regulate the AC/VU decision in *lag-2(q420)* worms, it was proposed that HLH-2 does not act as a general factor for *lag-2* transcription regulation, since rescue of the Lag phenotype was not impaired by mutation of the E-box motifs (Karp and Greenwald, 2003) (Karp and Greenwald, 2004). Therefore, a fragment spanning 1.6 kb upstream of the lag-2 gene start site was targeted, to eliminate the inverted repeat region present at 3.1 kb upstream, whilst maintaining

the major transcription factor binding regions.

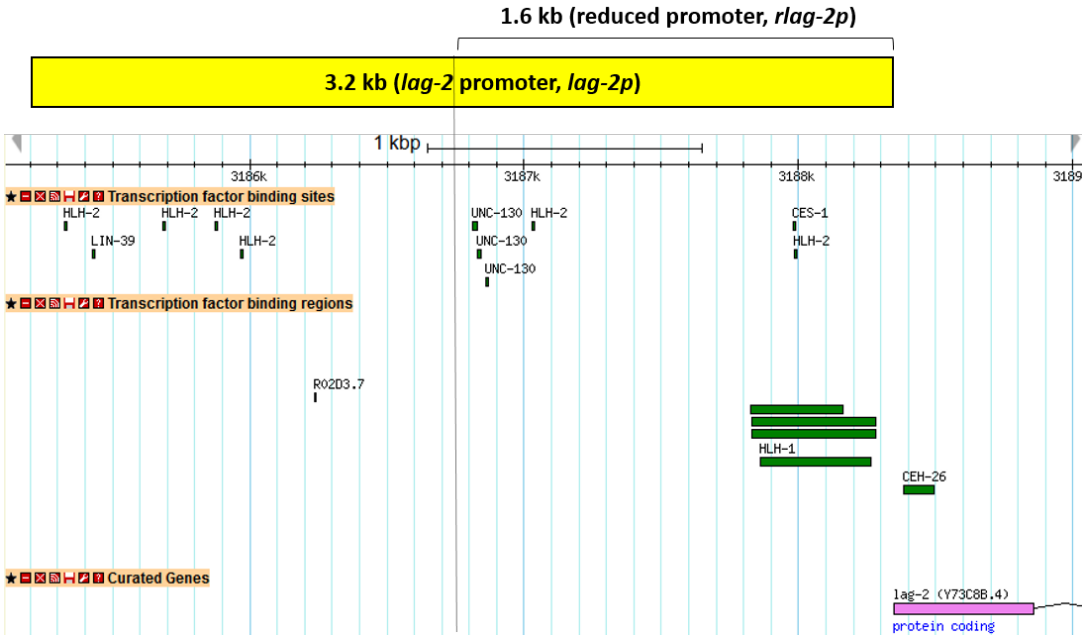


Figure 5.4: Schematic of the *lag-2* promoter region - Cartoon representation of the *lag-2* promoter region, annotated with transcription factor binding sites/regions from Wormbase. The region corresponding to the reduced promoter fragment, *rlag-2p*, is shown.

To test whether this promoter fragment was capable of driving expression, prior to its use to generate single copy insertion, extrachromosomal arrays were generated. The reduced promoter fragment spanning 1.6 kp upstream of the *lag-2* gene was used to drive expression of the *lag-2::gfp* rescue construct (*rlag-2p::lag-2::gfp*). *lag-2(q420)* mutants in an *unc-119*$^{+}$ rescue background were injected with 2 ng/μl of *rlag-2p::lag-2::gfp* to generate three independent transgenic lines, none of which were associated with a visable GFP signal in the DTC. The ability of *rlag-2p::lag-2::gfp* to rescue the Lag phenotype in *lag-2(q420)* was compared to *lag-2::gfp* using the larval viability assay (Figure 5.5). The percentage of progeny surviving to L2 and beyond was greatly reduced in *rlag-2p::lag-2::gfp* lines (13%-23%) compared with *lag-2::gfp* (42%-58%).

Statistical analysis of the combined values from all *rlag-2p::lag-2::gfp*, *lag-2::gfp*

and *unc-119* only (*lag-2(q420)*) lines indicated significant differences in the percentage of progeny surviving to L2 and beyond for *rlag-2p::lag-2::gfp*, *lag-2::gfp* ($p \leq 0.001$) as well as *rlag-2p::lag-2::gfp* and *unc-119* only (*lag-2(q420)*) ($p = 0.002$), supporting the observation that the reduced fragment of the *lag-2* promoter was not able to drive expression to the same extent of the full length promoter (Figure 5.6). This did however, suggest that *rlag-2p::lag-2::gfp* might have resulted in a low level of expression that was not visable during GFP imaging. It was concluded from these data that the reduced *lag-2* promoter fragment would not be a suitable substitute for the full length *lag-2* promoter fragment to generate single copy insertions for subsequent rescue assays.

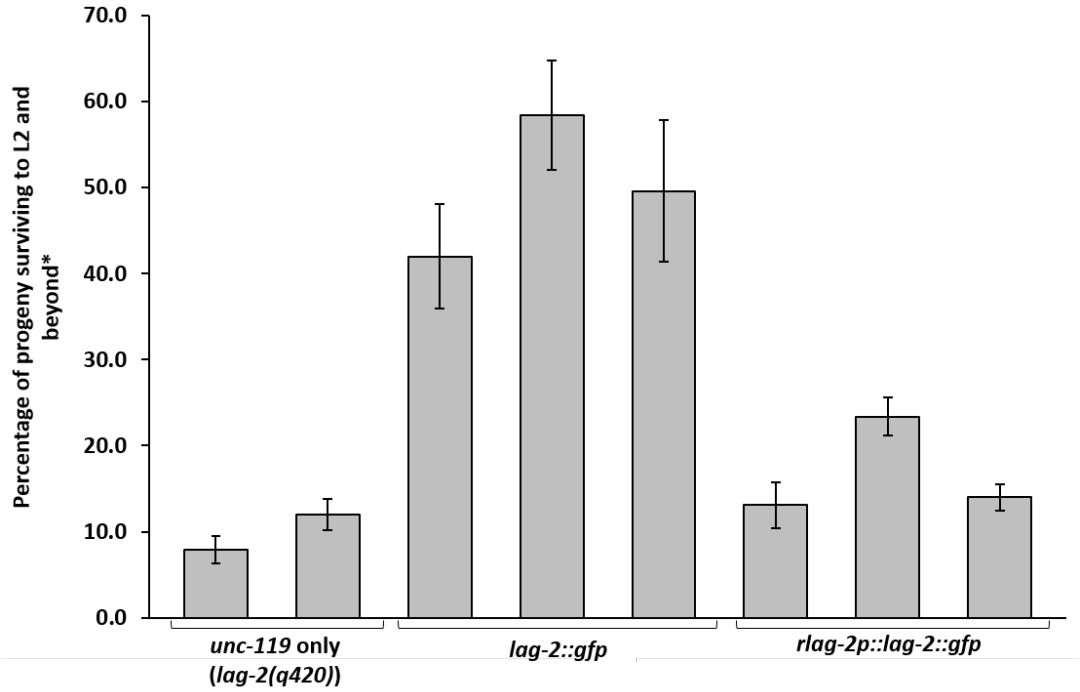


Figure 5.5: A reduced fragment of the *lag-2* promoter region cannot rescue the Lag phenotype - *Data was normalised as discussed in section 4.3.2 and Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for a single independent transgenic line, for which $n \geq 10$. Experimental error indicated as ±SEM.

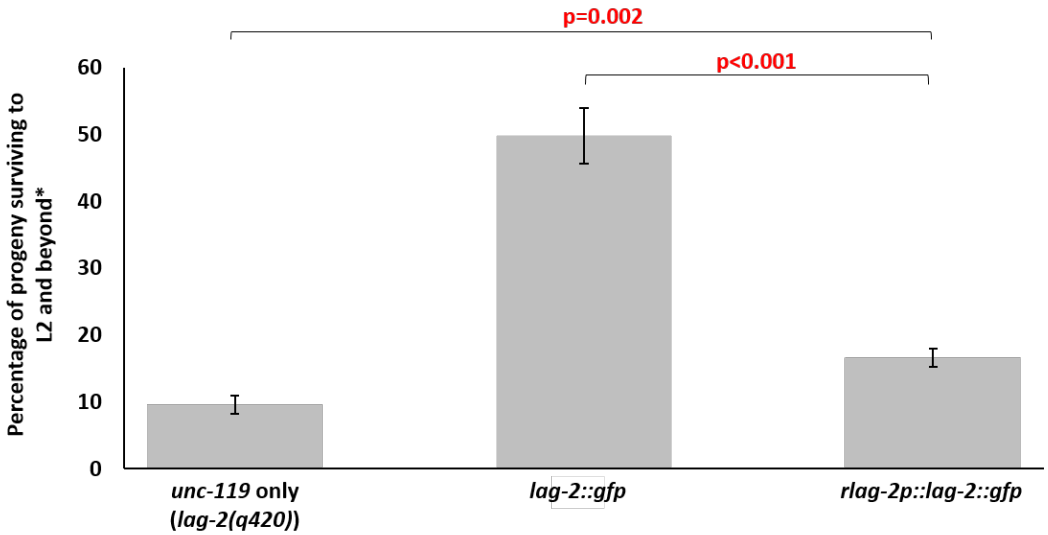


Figure 5.6: Statistical analysis for rescue of the Lag phenotype in *lag-2(q420)* mutants by a reduced *lag-2* promoter fragment - *Data was normalised as discussed in section 4.3.2 and Appendix 3 and 4, to give the percentage shown. Each bar corresponds to the results for multiple independent transgenic lines generated for a single construct, for which n≥20. p-values were calculated using a Mann-Whitney-U test, and p≤0.050 was considered statistically significant. Experimental error indicated as ±SEM.

5.3 Discussion

mosSCI was pursued to overcome some of the limitations associated with extrachromosomal array expression of transgenes that posed a challenge when establishing an *in vivo* assay to explore the function of Notch ligands. It was not possible to establish full length single copy inserts of the desired rescue constructs, and subsequently, it was discovered that this difficulty was also encountered by other members of the field who attempted to generate mosSCI lines with the *lag-2* promoter sequence (E. Sorensen and S. Crittenden, personal communication), suggesting that an essential motif within the *lag-2* promoter sequence may have prevented efficient insertion of the transgene. Analysis of the promoter region suggested that an inverted repeat present may have contributed to problems with insertion of the gene.

A shorter region of the *lag-2* promoter, excluding the upstream inverted repeat

region was tested in the larval viability assay for its ability to express *lag::gfp* to rescue the Lag phenotype in *lag-2(q420)* mutants. The results indicated that the reduced *lag-2* promoter fragment was insufficient to drive substantial expression to rescue the phenotype, suggesting that regions required for regulation of *lag-2* at some steps during embryogenesis are present 3.2 to 1.6 kb upstream of the *lag-2* gene. It may be that the inverted repeat is required, as these motifs can contribute to transcription regulation. Alternatively, the LIN-39 and HLH-2 sites in the 2.3 to 3.1 kb upstream region may be necessary to rescue the Lag phenotype in this particular context. These regions have not previously been implicated as necessary for regulation of *lag-2* transcription during embryogenesis, although it is interesting that deletion analysis of the *lag-2* promoter indicated that the requirement for specific regulatory regions of the promoter is context-dependent, when comparing expression in the AC/VU decision, vulval cell patterning and within the DTC (Zhang et al., 2011).

In conclusion, whilst there were many advantages of generating single copy insertion lines in which to assay the effects of mutations in Notch ligands, the limitations associated with finding a suitable fragment of the *lag-2* promoter region that is both efficiently inserted, whilst maintaining its functionality could not be overcome in the time-scale of this project. It would be interesting to make multiple reduced *lag-2* promoter fragments varying from 1.6-3.2 kb upstream of the *lag-2* gene, to find a region that might satisfy both criteria. However, in light of this data, the generation of mosSCI lines was not further pursued, and it was decided that rescue of the Lag phenotype in *lag-2(q420)* mutants may not be the most appropriate method of assaying rescue. Therefore, the post embryonic AC duplication phenotype associated with *lag-2(q420)* mutants was re-visited to establish an *in vivo* assay to explore the effect of mutations in Notch ligands in which extrachromosomal transgenic lines

might be used.

6

Rescue of the AC duplication phenotype to probe *C. elegans* Notch ligand function

6.1 Anchor cell duplication in *lag-2(q420)* mutants can be used to assay rescue constructs

As the second most penetrant phenotype after the Lag phenotype in *lag-2(q420)* mutants (and the foremost penetrant of the observed post embryonic phenotypes associated with these mutants, (section 4.2)), duplication of the AC was pursued to establish an *in vivo* system in which to probe *C. elegans* Notch ligand function. It was anticipated that by scoring one very specific cell fate decision, as opposed to the more general phenotype of viability (which requires Notch signalling at multiple stages of development), an assay with sufficient sensitivity to test the effects of Notch ligand mutants in could be developed.

To establish the assay, *lag-2(q420)* mutants were crossed to introduce an AC marker (*cdh-3>mCherry ::PLC$^{\delta PH}$II?*, generated by (Hagedorn et al., 2009)), to aid visualisation of the cell. Transgenic lines were generated by injecting rescue

construct DNA, as previously described (section 2.4.2.1) for *lag-2::gfp* and *apx-1::gfp*, with a *sur-5::gfp* co-injection marker, which allowed transgenic animals to be identified throughout the assay, and which eliminate the need to use the *unc-119* rescue marker background. Synchronised L1 larvae hatched at 15°C were moved to 25°, and imaged during L3, when the AC was fully specified and the mCherry marker was prominent. The total number of worms with single and duplicated anchor cells were scored.

Initially, a transgenic line containing only the *sur-5::gfp* marker in the *cdh-3>mCherry ::PLC$^{\delta PH}$II?; lag-2(q420)V* background was scored for the penetrance of the AC duplication phenotype, which was 52.2% ±SEM of 4.79 (Appendix 7). This was considered to be the maximum frequency of AC duplication in transgenic lines used in this assay. Therefore all data scored for other transgenic lines were represented as the percentage reduction in this phenotype, where 52.2% equalled the minimum value (0%) and 0% (i.e a control strain including the AC marker background and *sur-5::gfp* co-injection marker but not the *lag-2(q420)* allele, in which all animals had a single AC, as in WT) represented the maximum penetrance of rescue (100%). Transgenic lines expressing both *lag-2::gfp*(AC) and *apx-1::gfp*(AC) from extrachromosomal arrays reduced the penetrance of the mutant AC duplication phenotype by 25-33% (Figure 6.1). When data from independent transgenic lines were pooled, no significant difference was observed between the ability for *lag-2::gfp*(AC) and *apx-1::gfp*(AC) to rescue the AC duplication phenotype of *lag-2(q420)* mutants in this assay (p=0.372) (Figure 6.2). Therefore, in this specific context, LAG-2 and APX-1 were considered to be fully interchangeable, functionally. This AC duplication assay taken forward to measure the effect of mutations in the *C. elegans* Notch ligand APX-1 to determine whether mutant APX-1 constructs could rescue the AC duplication phenotype in *lag-2(q420)* worms, compared to WT APX-1.

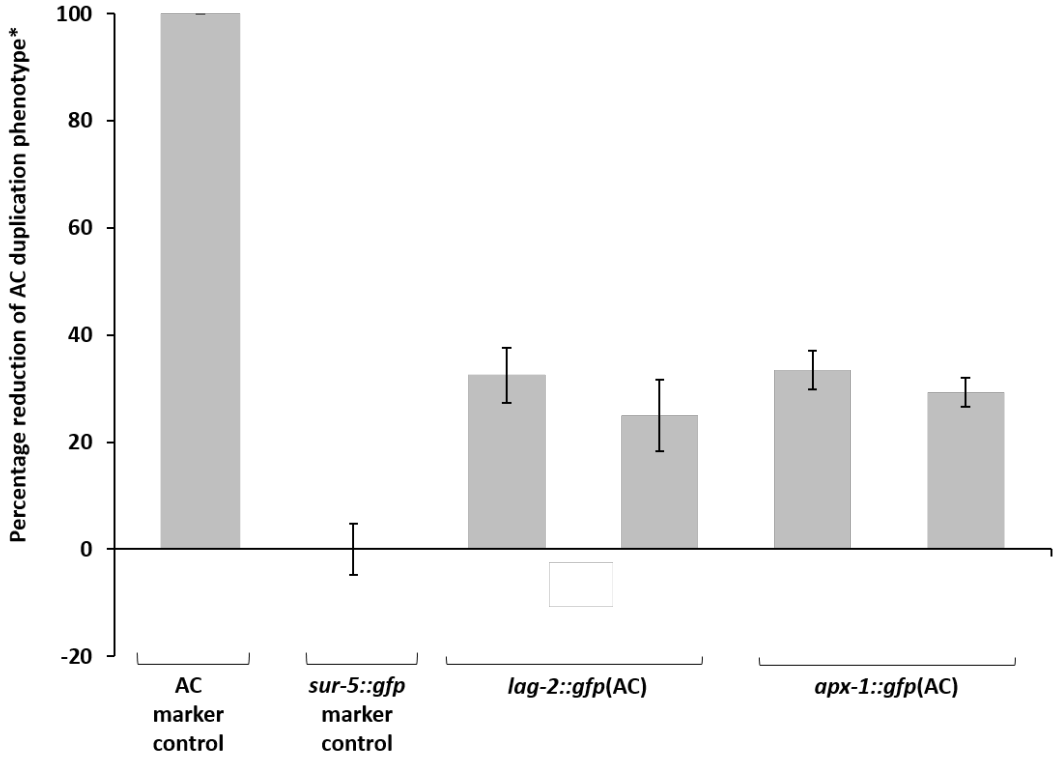


Figure 6.1: Expression of *lag-2::gfp*(AC) and *apx-1::gfp*(AC) can rescue the AC duplication phenotype in *lag-2(q420)* mutants - *Percentage reduction of phenotype was calculated as described in Appendix 7. Control lines shown demonstrate that the AC marker background (*sur-5::gfp* expression in the *cdh-3>mCherry ::PLC$^{\delta PH}$II?* background) does not itself cause AC duplication events, as all animals scored had a single AC, as in WT, and is therefore associated with '100%' reduction of the mutant duplication phenotype. The *sur-5::gfp* marker control (*sur-5::gfp* expression in the *cdh-3>mCherry ::PLC$^{\delta PH}$II?; lag-2(q420)V* background)) is considered to demonstrate maxiumum AC duplication in this assay, and therefore is associated with 0% reduction in the AC duplication phenotype. Each bar corresponds to the mean of 2-4 independent experiments in which $18 \leq n \leq 50$ for a single line. Experimental error is indicated as ±SEM.

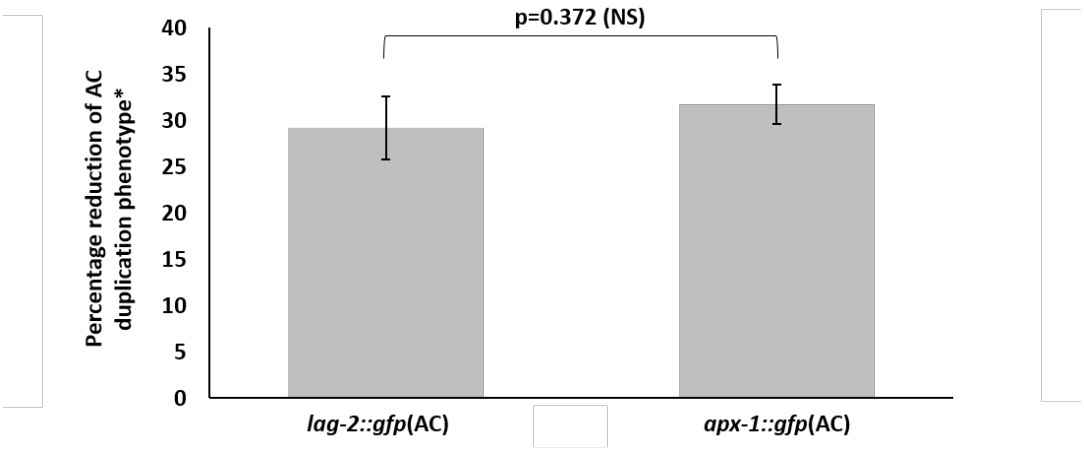


Figure 6.2: Statistical analysis suggests expression of *lag-2::gfp*(AC) and *apx-1::gfp*(AC) can rescue AC duplication phenotype to a similar extent - *Percentage reduction of phenotype was calculated as described in Appendix 7. Each bar corresponds to the mean of the combined values of 3-4 independent experiments in which $18 \leq n \leq 29$ for both independent transgenic lines obtained. p-values were calculated using a Mann-Whitney-U test, and $p \leq 0.050$ was considered statistically significant. Experimental error is indicated as ±SEM.

6.2 The effects of domain deletions in APX-1

Since LAG-2 and APX-1 were considered functionally fully interchangeable in this assay, APX-1 was targeted for mutagenic studies, to complement the ongoing efforts to gain structural information for APX-1 described in Chapter 3. The effects of whole domain deletions in LAG-2 have previously been explored by assaying rescue of the *lag-2(q411)* null allele, and demonstrated that deletion of either the MNNL or DSL domain of LAG-2 lead to a loss of the ability of extrachromosomally expressed *lag-2* cDNA to rescue the Lag phenotype in these mutants (Henderson et al., 1997). On the other hand, deletion of the EGF-like domains, or the intracellular domain of LAG-2 did not prevent rescue of the Lag phenotype, which suggested they are not essential for LAG-2 function in this context.

It was hypothesised that deletion of the MNNL or DSL domains of APX-1 would lead to a loss of functionality, and a failure to rescue the AC duplication phenotype in

lag-2(q420) mutants due to their known roles in Notch binding. Three mutant *apx-1::gfp*(AC) constructs were generated, by precisely deleting the MNNL, (Δ*MNNL apx-1::gfp*(AC)) the DSL (Δ*DSL apx-1::gfp*(AC)), and EGF-like 1 (Δ*EGF-like 1 apx-1::gfp*(AC)) domains of APX-1 (Figure 6.3A), based on primary sequence analysis. Due to the presence of the *sur-5::gfp* marker obscuring GFP expression at the DTC, expression of these constructs could not be confirmed in the transgenic lines that were directly assayed for rescue of the AC duplication phenotype. However, the same DNA constructs were injected into *unc-119(ed3)III* worms with the *unc-119*$^{+}$ rescue marker, confirming all three constructs were capable of being expressed at the protein level (Figure 6.3B).

The three deletion constructs were assayed for their ability to rescue the AC duplication phenotype (Figure 6.4). The penetrance of the AC duplication phenotype was not reduced by expression of either Δ*MNNL apx-1::gfp*(AC) or Δ*DSL apx-1::gfp*(AC), which was consistent with the hypothesis that these domains are important for functionality of APX-1. Conversely, Δ*EGF-like 1 apx-1::gfp*(AC) was still able to rescue the AC duplication phenotype to some extent, indicating that this region is not essential for APX-1 function. When the values from the individual transgenic lines were pooled for each construct, and statistical analysis performed, this clearly supported a significant difference between the ΔMNNL and ΔDSL constructs compared to the WT ($p=0.001$ for both). Interestingly, although ΔEGF-like 1 was able to rescue the AC duplication, this was not to the same level as WT APX-1 ($p=0.023$), suggesting that EGF-like 1 may have a functional role in Notch signalling (Figure 6.5).

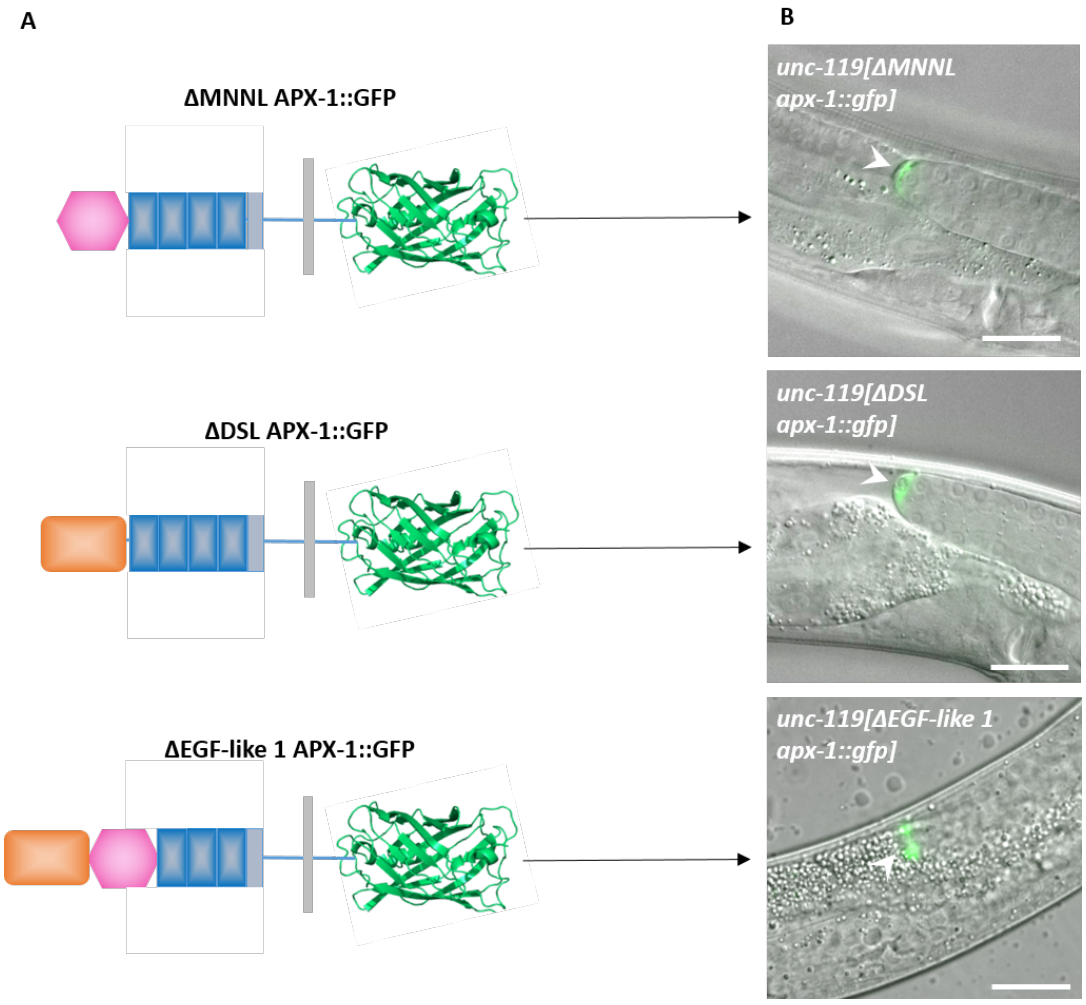


Figure 6.3: Schematic of APX-1 whole domain deletion constructs and their expression - (**A**) Cartoon representations of the whole-domain deletion constructs, coloured as in Figure 1.2, with GFP shown in green. (**B**) DIC images of trangenic lines with FITC FS10 overlay. All images shown are for worms in an *unc-119(ed3) III [unc-119^{+}]* rescue background, in the absence of the *lag-2(q420)* allele.

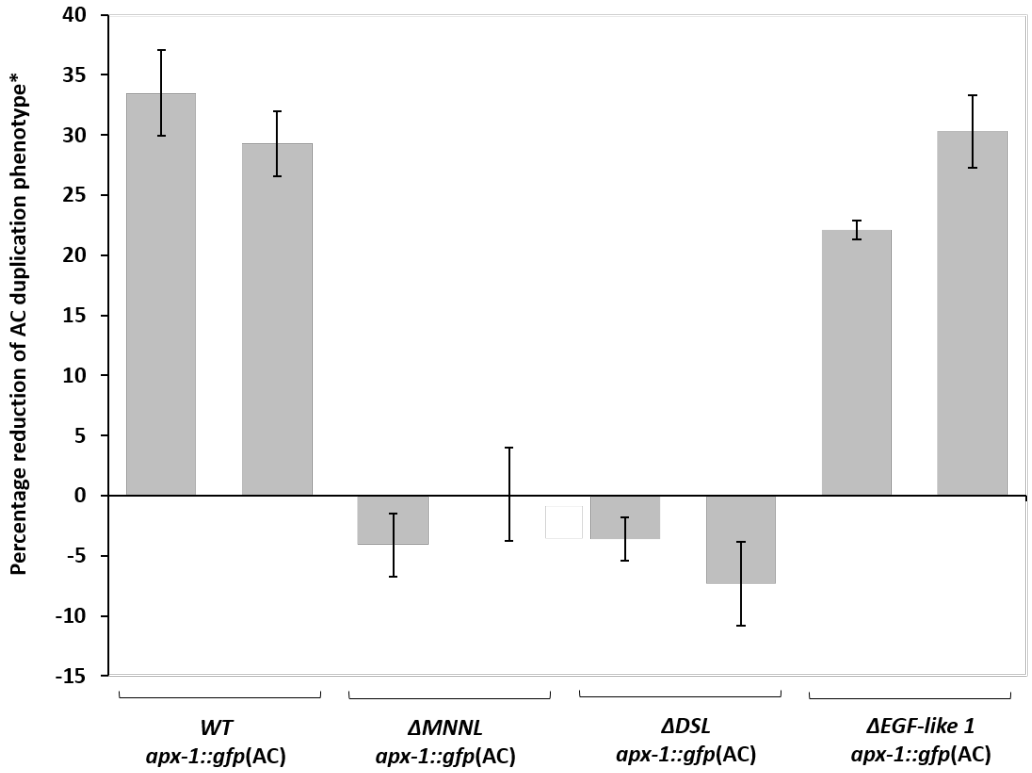


Figure 6.4: Deletion of the MNNL or DSL in APX-1 is associated with loss of AC duplication rescue in *lag-2(q420)* mutants - *Percentage reduction of phenotype was calculated as described in Appendix 7. Each bar corresponds to the mean of 3-4 independent experiments in which $18 \leq n \leq 25$ for a single line. Experimental error is indicated as ±SEM.

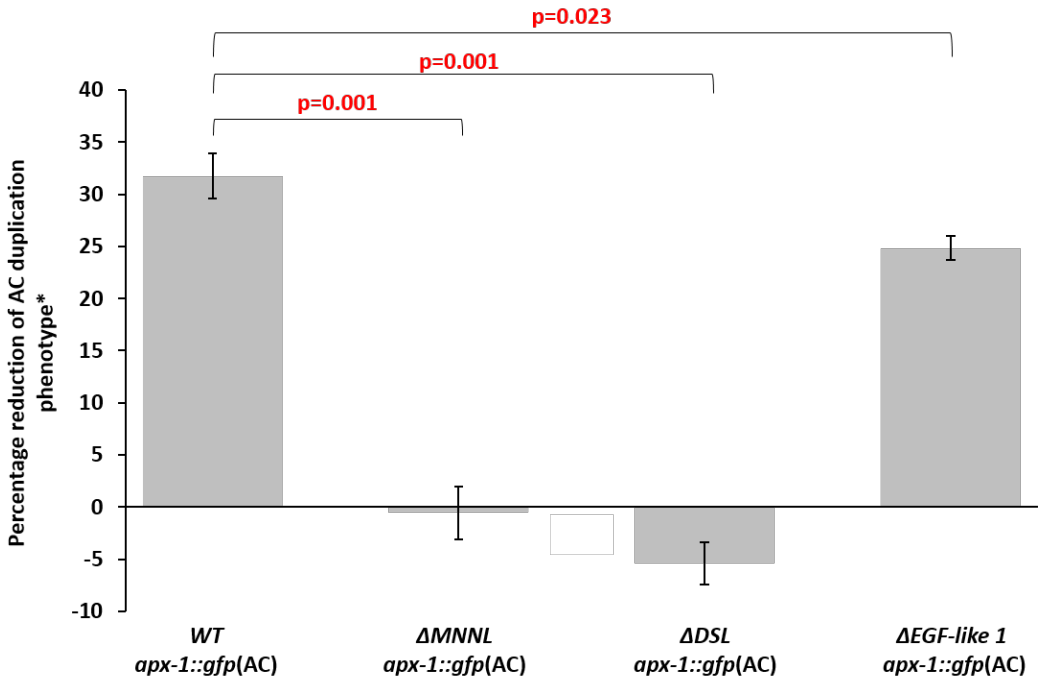


Figure 6.5: Statistical analysis of AC duplication rescue by domain deletion mutants - *Percentage reduction of phenotype was calculated as described in Appendix 7. Each bar corresponds to the mean of the combined values of 3-4 independent experiments in which $18 \leq n \leq 29$ for both independent transgenic lines obtained. p-values were calculated using a Mann-Whitney-U test, and $p \leq 0.050$ was considered statistically significant. Experimental error is indicated as ±SEM.

6.3 Targeted mutagenesis of APX-1

Having demonstrated the effects of whole domain deletions of APX-1 using the AC duplication phenotype assay, a more precise approach for generating mutants to test domain function was pursued. As previously discussed in Chapter 3, the lack of structural knowledge regarding APX-1 was a limitation in the ability to target specific sites for mutagenesis. Nevertheless, with the aim that a crystal structure would be forthcoming, two sites were targeted for mutagenesis based on sequence and structural predictions to demonstrate the power of the AC duplication assay to test functionality of Notch ligand mutants.

The two sites targeted for mutagenesis correspond to the predicted C_3-C_4 loop of the DSL domain, and the β5-6 loop of the MNNL domain of APX-1 (Figure 6.6A). The DSL domain loop mutant consisted of a substitution of four amino acid residues spanning A147-K150 (A N A K to E T S D). Whilst there is no highly conserved motif in the C_3-C_4 DSL domain loop of *C. elegans* membrane-bound Notch ligands compared with mammals and flies (section 3.1.2), it was noted that positively charged residues (K, R or H) were present in both LAG-2 and APX-1, although not at the same positions. Therefore, by removing the positively charged K150, and replacing it with the two negatively charged amino acids (E and D), this DSL mutant (*DSL(ETSD) apx-1::gfp*(AC)) was designed to probe whether these residues of the DSL was necessary for an interaction with Notch receptors.

The β5-6 loop was targeted with less confidence than the DSL loop, as this loop was identified on the basis of structural predictions that were known to be associated with some limitations (section 3.1.3). A region of 6 amino acids corresponding to part of the β5-6 loop of Jagged-1, was substituted with two polar serine residues (*I80-K86/SS apx-1::gfp*(AC)), to probe whether this region was necessary for functionality in APX-1. Both loop mutant constructs were successfully expressed in an

unc-119^{+} rescue background (Figure 6.6B).

Transgenic lines containing *DSL(ETSD)apx-1::gfp*(AC) and *I80-K86/SS apx-1::gfp*(AC) rescue constructs were tested for their ability to reduce the penetrance of the AC duplication phenotype in *lag-2(q420)* mutants (Figure 6.7). Both mutants were able to rescue this phenotype to some degree compared with WT APX-1. In the case of *DSL(ETSD)apx-1::gfp*(AC), the percentage reduction in AC duplication ranged from 20.5-21.7%, whilst *I80-K86/SS apx-1::gfp*(AC) ranged from 22.2-26.3%. Statisical analysis of the combined values from individual lines for each of the constructs indicated significant differences between the DSL(ETSD) and I80-K86/SS mutants compared with WT APX-1 (p=0.006 and p=0.031, respectively) (Figure 6.8). These data indicated that the loop region mutants have reduced ability to rescue the AC duplication phenotype in *lag-2(q420)* mutants compared with WT APX-1, suggesting a functional role for the loop regions targeted in both the DSL and MNNL domains.

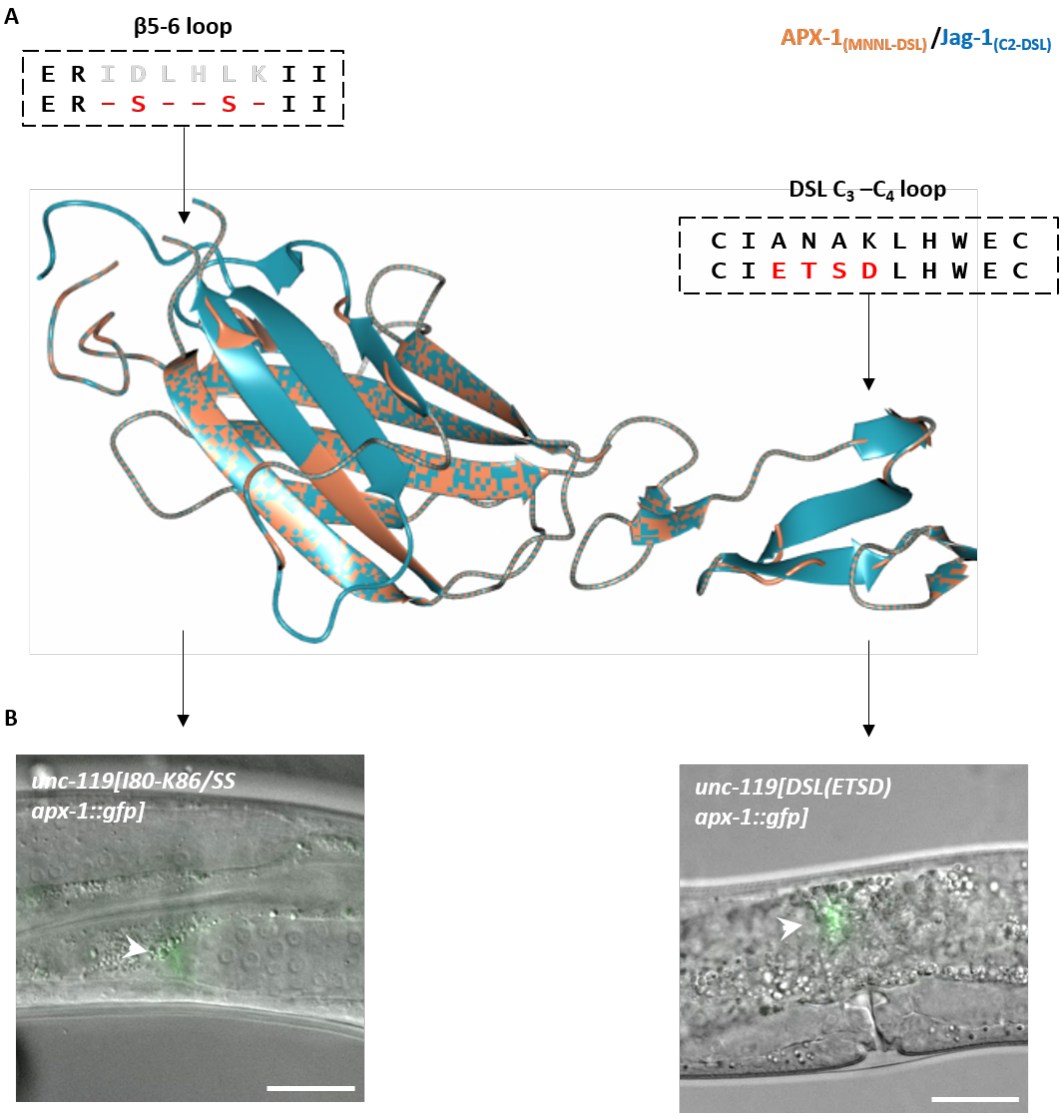


Figure 6.6: Generation of targeted loop region substitutions in the MNNL and DSL domain of APX-1 - (A) FFAS predicted structure of APX-1 is shown in orange, superimposed on the 4CBZ structure of $J1_{NE3}$ is shown in blue, focussing on the MNNL-DSL region of the structures. Loop regions are annotated with mutated sites, shown in red. The 6 residues in the β5-6 loop substituted with serines are shown in grey. (**B**) DIC images of trangenic lines with FITC FS10 overlay. All images shown are for worms in an *unc-119*$^{+}$ rescue background, in the absence of the *lag-2(q420)* allele. Scale bar = 20 μm

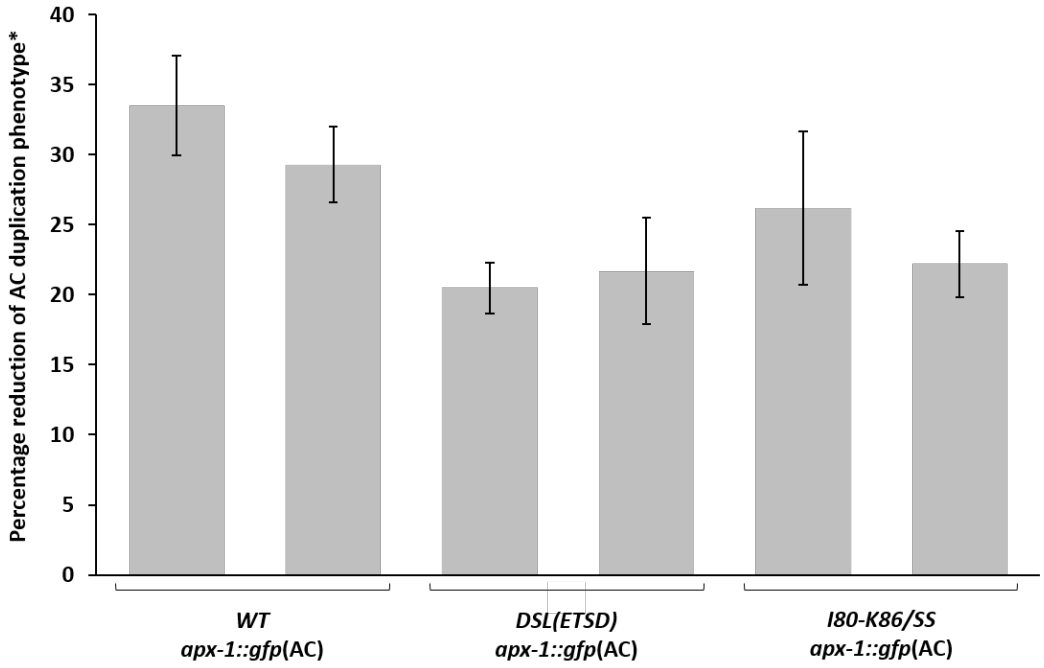


Figure 6.7: Specific mutations affecting the MNNL and DSL domain do allow some rescue of the AC duplication phenotype in *lag-2(q420)* mutants - *Percentage reduction in phenotype was calculated as described in Appendix 7. Each bar corresponds to the mean of 3-4 independent experiments in which $18 \leq n \leq 25$ for a single line. Experimental error is indicated as ±SEM.

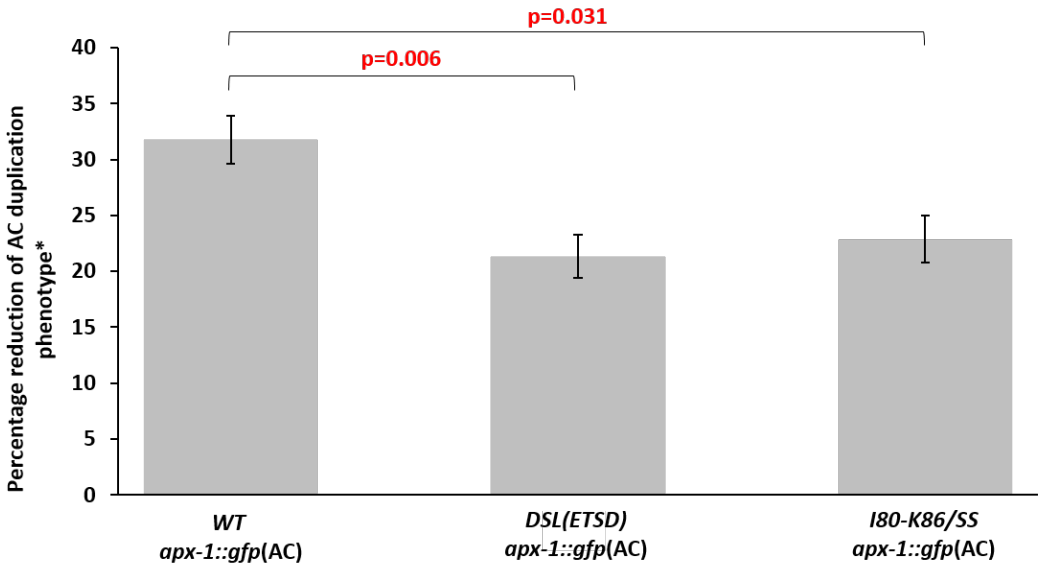


Figure 6.8: Statistical analysis of AC duplication rescue by loop region mutants - *Percentage reduction in phenotype was calculated as described in Appendix 7. Each bar corresponds to the mean of the combined values of 3-4 independent experiments in which $18 \leq n \leq 29$ for both independent transgenic lines obtained. p-values were calculated using a Mann-Whitney-U test, and $p \leq 0.050$ was considered statistically significant. Experimental error is indicated as ±SEM.

6.4 Discussion

A key aim of this thesis was to establish an *in vivo* assay in which the effects of mutations in *C. elegans* Notch ligands could be tested. Both whole domain deletions and amino acid substitutions in suspected loop regions of APX-1 were carried out, and the resulting constructs were assayed for their ability to rescue the AC duplication phenotype associated with *lag-2(q420)* mutants, allowing insights into the functionality of these regions.

Initially, rescue of the AC duplication phenotype was demonstrated in this assay by expression of WT *lag-2::gfp*(AC) and *apx-1::gfp*(AC). As observed with the Lag phenotype of *lag-2(q420)* mutants, expression of *lag-2::gfp*(AC) and *apx-1::gfp*(AC) appeared to allow partial rescue of the AC duplication phenotype. The cause of this was not explored in depth, however anecdotally, it was observed that both the

lag-2::gfp(AC) and *apx-1::gfp*(AC) transgenic lines were associated with embryonic lethality, as had been the case with *lag-2(q420)* mutants in the *unc-119*$^{+}$ rescue background. Since the AC duplication assay allowed observation of a phenotype in worms that had already survived into L3, it was considered less of a limitation in the context of this assay compared with the larval viability assay, in which survival was directly assayed. It is worth noting that *lag-2::gfp*(AC) lines could not be generated by injecting DNA at 2 ng/μl, and the concentration had to be lowered to 1 ng/μl to obtain lines.

The effects of whole domain deletions in APX-1 demonstrated the power of this assay, by showing a complete loss of rescue by expression of both Δ*MNNL apx-1::gfp*(AC) and Δ*DSL apx-1::gfp*(AC) compared to *WT apx-1::gfp*(AC). It was not possible to confirm whether the loss of functionality associated with deletion of either of these domains was due to structural or functional defects. If the deletion of either of these domains had introduced a misfold, it is possible that the protein was retained in the cells, and rescue was not observed since there was no protein targeted to the surface. However, in this circumstance, it would be anticipated that the misfolded protein would be subjected to endoplasmic reticulum associated protein degradation (ERAD) (Hampton, 2002), which would also result in loss of the GFP signal. Therefore, it is more likely that the mutant protein was able to be transported to the surface, and that the loss of either domain abolished the ability for APX-1::GFP to bind to LIN-12. To support this proposal experimentally, a specific antibody for the extracellular region of each construct could be generated and used to stain the DTC of a dissected gonad. If the material were present at the surface, the antibody would bind, and could be detected, however, if the protein constructs remained misfolded within the cell, the antibody would be unable to bind. Assuming the deletion of the MNNL or the DSL of APX-1 did not cause

retention of the protein within the cell, the loss of AC duplication phenotype rescue (and thus, loss of Notch signalling activation) is consistent with previous studies implicating both domains in Notch receptor binding in mammals and flies (Cordle et al., 2008a) (Luca et al., 2015), and in addition, the importance of the C2 domain in phospholipid-binding and Notch activation (Chillakuri et al., 2013).

Deletion of EGF-like 1 of APX-1 led to a reduction rescue of the AC duplication phenotype in *lag-2(q420)* mutants compared with WT APX-1, which suggested some impairment of ligand function by loss of this region. The EGF-like 1 of mammalian Jagged-1 has been implicated in enhancing Notch receptor binding (Shimizu et al., 2000), and in addition, it has been demonstrated that in the case of rDLL-4, the D218 residue at the beginning of EGF-like 1 participates at the Notch receptor/ DSL ligand interface site (Luca et al., 2015). The observation that the Δ*EGF-like 1 apx-1::gfp*(AC) construct was able to rescue the AC duplication phenotype to some extent was consistent with the observation that EGF-like domains of LAG-2 were not critical for signalling (Henderson et al., 1997). The EGF-like domain regions are very diverse between LAG-2 and APX-1, despite the two being able to both function as ligands for LIN-12 in the AC fate decision. In addition, deletion of this entire domain does not appear to prevent the protein from folding and being targeted to the surface of the cells. Therefore, whilst the EGF-like 1 domain of Notch ligands may not be absolutely essential, it appears to be required to generate an optimal interaction, possible through forming additional contacts with the Notch receptor, and/or contributing to the stability of the DSL domain.

The AC duplication assay data suggested that targeted substitutions of DSL residues and I80-K86 region of APX-1 led to some impairment of ligand function. The observation that expression of *DSL(ETSD)apx-1::gfp*(AC) and *I80-K86/SS apx-1::gfp*(AC) rescued the AC duplication to some extent suggested that the amino acid

substitutions associated with the DSL(ETSD) and I80-K86/SS had not targeted residues critical for the fold of either the DSL or MNNL domains, as functional APX-1 was able to reach the cell surface to participate in signalling. This was particularly interesting in the case of the I80-K86/SS mutant, which was based on a structural prediction that was associated with some limitations. If it is assumed that the MNNL comprises a C2 domain in APX-1, the ability of the I80-K86/SS mutant to rescue the APX-1 phenotype supports the prediction that these substitutions target a loop region, since disruption of residues associated with the β strands forming the hydrophobic core of the C2 domain have been demonstrated to generally result in misfolding and retention of the ligand Jagged-1 (Chillakuri et al., 2013).

Replacement of the A N A K region (amide, basic and hydrophobic residues) of the C_3-C_4 loop with E T S D residues (acidic and polar residues) in the DSL(ETSD) mutant was also associated with a reduction in its ability to rescue the AC duplication phenotype, compared to WT APX-1. It is likely that amino acids substituted in this mutant contribute to the interaction of the DSL domain with LIN-12 and GLP-1 in WT, but are not essential for binding. An additional positively charged histidine (H152) residue in the C_3-C_4 loop (Figure 3.5), that is conserved between LAG-2 and APX-1 may offer a second site for future mutagenesis, to further probe the role of the positively charged residues in this region.

7

General discussion and future directions

The core aim of this project was to use the nematode worm *C. elegans* to probe the function of the N-terminus of Notch ligands using complementary molecular and *in vivo* approaches. The Notch signalling pathway is utilised throughout metazoans during development and adult tissue maintenance, and the presence of Notch receptors and their canonical DSL ligands is conserved amongst humans, *Drosophila* and *C. elegans*. *C. elegans* Notch ligands are significantly smaller than human and *Drosophila* ligands, containing fewer EGF-like domains, whilst retaining a DSL domain and an N-terminal MNNL domain, both of which have been shown to participate in the Notch receptor/ligand interaction in mammals and flies. In addition, Notch signalling has previously been studied in *C. elegans* from an *in vivo* perspective, and therefore several mutant alleles of the Notch signalling pathway have been characterised, providing the foundations to design assays in which to test mutant Notch ligand constructs. Therefore, *C. elegans* presented several advantages to be pursued as a model for Notch signalling in which structural and functional information could be obtained and integrated.

An initial aim at the start of this project was to compare *C. elegans* Notch

receptors and ligands to their *Drosophila* and mammalian counterparts, to assess the extent to which they would provide a suitable model for Notch signalling. Primary sequence analysis was carried out to look for the conservation of key residues in the DSL domain of Notch ligands. Whilst *C. elegans* lacked the well-conserved motif found in the C_3-C_4 loop of the DSL domain associated with mammals and flies, conservation of some key residues between the suspected major ligand-binding region of *C. elegans* Notch receptors LIN-12 and GLP-1 and the major ligand-binding region of human Notch-1 suggested that the interaction interface may share some similarities between species. It was also noted that the same key resides were not fully conserved in the major ligand-binding region of *Drosophila* Notch, indicating the presence of some species-species variability was to be anticipated. In addition, the N-terminal MNNL domain was analysed, using the structural prediction program FFAS. The outcome of these predictions suggested that whilst the MNNL region of Notch ligands was capable of adopting a C2 domain fold, the predictions were unlikely to show accurate folding of the MNNL region, as *C. elegans* Notch ligand MNNL domains are shorter than the published mammalian C2/MNNL regions on which the predictions were based. Taken as a whole, sequence analysis and structural predictions supported the view that *C. elegans* could potentially provide interesting insights into the Notch signalling pathway that may be related to other species, if structural information could be obtained to reconcile the inter-species differences observed at the primary sequence level.

In light of this conclusion, the subsequent aim of this project was to produce recombinant fragments of *C. elegans* Notch ligands, for biochemical and structural characterisation, allowing functionally significant residues in the MNNL to be identified. Initial attempts to express several different recombinant fragments of *C. elegans* Notch ligands allowed small quantities of the ligand fragment APX-1_{NE4FcHis} to be

expressed from HEK293T cells, and subsequently purified, however insufficient material was yielded for further analysis. An S2 cell expression system allowed a large quantity of the fragment APX-1_{NE3His} to be expressed, although it was not possible to purify this fragment to the standard necessary for use in crystal studies. Eventually, R. Suckling successfully purified a smaller fragment, APX-1_{NE1His} from the S2 expression system, and subsequently used this material in crystal trials. An effort to ascertain whether a calcium-binding site was present in this fragment of APX-1 by limited proteolysis was inconclusive, therefore, without structural information, further characterisation of this protein could not be performed with confidence. However, this work did highlight the power of the S2 expression system to yield high quantities of recombinant proteins, and this would be used as the expression system of choice for future experiments.

Further experiments that could be conducted would require additional fragments to be made using the S2 system, allowing comparative studies to be performed. For example, an APX-1_{DE1His} fragment lacking the MNNL domain could be produced, and both APX-1_{NE1His} and APX-1_{DE1His} could be tested for phospholipid-binding properties associated with the MNNL domain in mammals and flies. This could contribute evidence for the presence of a C2 domain in APX-1, and would allow calcium dependence to be further studied in the context of this assay. It would also be interesting to generate recombinant fragments of the LIN-12 and GLP-1 Notch receptors for interaction assays with APX-1 fragments. By making multiple truncated fragments of LIN-12 and GLP-1, the minimal region required for binding APX-1 could be identified using assays such as ELISA or SPR. To conclude, although it was not possible to gain a structure of APX-1 during the time-scale of this project, a positive outcome of this work was the identification of S2 cells as a highly suitable method for the expression of recombinant *C. elegans* Notch proteins, which opens

up many avenues for future work.

Another key aim of this project was to establish an *in vivo* assay in which the effects of mutations in *C. elegans* Notch ligands could be explored at the physiological level. The *lag-2(420)* mutant allele was used in subsequent efforts to generate an appropriate assay, for advantages including the ability to maintain animals as homozygotes, and the variety of post embryonic developmental phenotypes, which could be observed in addition to the Lag phenotype, due to the temperature sensitive nature of the mutation. The temperature sensitivity of this allele did also have the disadvantage of requiring that all *lag-2(q420)* mutant worms had to be maintained at the permissive temperature of 15°C, and all genetic crosses and transgenic lines also had to be generated at 15°C. The life cycle of *C. elegans* is temperature dependant, and the generation time from egg to adult increases from ∼3 to ∼6 days from 20°C to 15°C, respectively, therefore this reduced the speed at which data could be collected and analysed before planning subsequent experiments. Initially, the *lag-2(q420)* allele was used to establish an assay to determine rescue of the Lag phenotype by assessing larval viability. In this assay, expression of WT *cDNA::gfp* fusion constructs encoding LAG-2 and APX-1 from extrachromosomal arrays was able to rescue the Lag phenotype to some extent, although in this assay APX-1 rescued the Lag phenotype significantly less than LAG-2. Interestingly, both constructs were associated with only partial rescue of the Lag phenotype, and further investigation suggested this was likely to be a consequence of overexpression of the ligand from the extrachromosomal array and associated embryonic lethality.

It cannot be ruled out that additional limitations of the larval viability assay itself were also contributing to partial rescue of the Lag phenotype by the ligand constructs. During the assay of larval viability in *unc-119* only rescued worms it was observed that after the transgene transmission efficiency had been accounted for,

substantially less than 100% of progeny were surviving to L2 and beyond. Whilst final values for other transgenic lines were normalised using the *unc-119* only rescued progeny data, the variability of the effects of the *unc-119* rescue background on viability had from animal to animal due to extrachromosomal expression could not be determined. The combination of limitations associated with the larval viability assay itself as well as the presence of overexpression effects and embryonic lethality indicated that this assay would not be appropriate to probe the effects of mutations in Notch ligands, and therefore an alternative approach circumventing these issues was sought.

Initially, it was believed that the challenges associated with extrachromosomal transgene expression could be overcome using mosSCI. In theory, this approach had the potential to eliminate variability caused by differences in transgene copy number and mosaicism between cells that are limitations of extrachromosomal array expression since a single copy of the transgene would be inserted into the genome. In practice, it was not possible to generate insertions of *lag-2::gfp*, and this was attributed to the presence of inverted repeat sequences in the *lag-2* promoter region that may disrupt template directed repair during gene conversion. Although inefficient transgene insertion has previously been considered a consequence of the presence of inverted repeats such as in the *pie-1* promoter, transgenes had still successfully been generated, although in some cases, the insertion frequency was as little as 0.20%, with one transgenic mosSCI line yielded from injection of 500 animals (Frø kjaer Jensen et al., 2008). Even if it had been possible to yield a single successful mosSCI insertion of the *lag-2::gfp* construct by injecting many hundreds, or up to 1000 animals, this would have been impractical in the remaining time-scale of the project, as it would have required the injection and screening of several thousand worms. To improve the transgene insertion frequency, the inverted repeat sequence

was removed, and a reduced fragment of the *lag-2* promoter was generated. Unfortunately, when this fragment was tested in the larval viability assay for it's ability to drive *lag-2::gfp* expression to rescue the Lag phenotype, it became apparent that the function of this reduced fragment was compromised compared to the full length promoter fragment. Therefore, it was concluded that mosSCI could not be further pursued as the issues surrounding the inefficiency of transgene insertion could not be adequately resolved within the time-scale of this project.

In light of the complications encountered when attempting to generate transgenic lines by semi-stable extrachromosomal arrays and mosSCI to assess rescue of the Lag phenotype, an alternative phenotype was used to establish an *in vivo* assay in which Notch ligand mutants could be tested. The AC duplication phenotype associated with the *lag-2(q420)* mutant represented the aberration of a single, well-defined cell fate decision (i.e specification of the AC), as opposed the assessment of viability of the worm (as a measure of the Lag phenotype), which could be influenced by additional factors such as genetic background, and the loss of worms due to embryonic lethality. By assaying the AC duplication phenotype in L3 worms, dead embryos were by definition excluded from the final count, and whilst this may have reduced the transgenic pool to some extent, it was assumed that surviving L3 worms were representative of the population as a whole. In addition, the *unc-119* marker background that had previously been used in the larval viability assay, and which was identified as a limitation of the system (as it was associated with a loss of viability) was not used in this assay. Therefore, the AC duplication assay was considered to provide a more reliable and sensitive system in which Notch ligand mutants could be tested compared to the larval viability assay, despite both utilising extrachromosomal transgene expression.

The AC duplication assay was successfully used to explore the effects of mu-

tations in *C. elegans* Notch ligands, by assessing their ability to rescue the AC duplication phenotype in *lag-2(q420)* mutants. Having established the extent to which WT APX-1 could rescue the phenotype, whole domain deletions of APX-1 were performed, and their effects demonstrated *in vivo*. Deletion of either the MNNL or DSL domain of APX-1 abolished rescue of the AC duplication phenotype, supporting the observation that these domains play a key role in Notch ligand function. On the other hand, rescue was not as drastically reduced by deletion of the EGF-like 1 domain, suggesting that whilst EGF-like 1 of APX-1 might not be critical for signalling, loss of this domain impairs the function of the ligand. In addition, primary sequence analysis and structural predictions were used to generate targeted substitutions two potential loop regions that may be involved in ligand binding (DSL(ETSD)) and phosopholipid-binding (I80-K86/SS) in the DSL and MNNL domains, respectively. Interestingly, rescue of the AC phenotype was significantly reduced compared to WT APX-1 for both DSL(ETSD) and I80-K86/SS, suggesting some functional impairment by substitution of these loop regions, as with the EGF-like 1 deletion mutant. These novel mutants demonstrated that the AC duplication assay has the potential to provide functional insights into *C. elegans* Notch ligands, and with the contribution of structural information, this assay could be applied to study the physiological role of the MNNL in greater detail.

During the course of this project, an alternative technology arose in the field as a method to directly study the consequences of mutations *in vivo*. CRISPR-Cas9 is a bacterial adaptive immune system that has been developed into a genome editing tool and its applications have been adapted for use in several organisms, including *C. elegans* (Friedland et al., 2013) (Dickinson et al., 2013). A guide RNA containing regions of specificity to both the target genome sequence and Cas-9 allows recruitment of this nuclease to the target site and the generation of a double strand break

(DSB), which can be repaired by non-homologous end joining (NHEJ) or homology-directed repair (HDR) mechanisms. The consequence of repair by NHEJ is generally to generate random insertions or deletions within a target gene, and therefore most often generating a knock-out allele. On the other hand, by introducing a template with flanking regions of homology to the site that is to be repaired, HDR can allow the incorporation of transgenes, or specific gene-edits at the DSB site by gene conversion. In theory, this opens up an almost limitless number of possibilities for mutagenic studies. In practice, current literature indicates that whilst generation of gene knock-outs by NHEJ is relatively robust, template driven repair by HDR tends to occur with lower efficiency, although protocols are emerging to improve this (Dickinson et al., 2015) (Paix et al., 2015). Although it was not practical to embark on this approach during the time-scale of this project, the reagents and protocols have been refined over the past two years and are now widely available, and it would be an attractive method to use from the outset in a future study into the effects of mutations in *C. elegans* Notch ligands.

The final aim of this project was to bring the complementary biochemical/ structural and *in vivo* data together to gain novel insights into the functionality of the N-terminal region of Notch ligands. Despite several challenges, an *in vivo* assay was established that could be used to probe the effects of mutations in *C. elegans* Notch ligands, and whilst mutants were designed based on structural predictions, rather than an X-ray structure of APX-1, these have offered new information regarding both the MNNL and DSL domains of APX-1. The ability of the I80-K86/SS to rescue the AC duplication phenotype to some extent suggests that the substitution had indeed targeted a loop region rather than a region required to maintain structural integrity of a C2 domain. Without high resolution structural data, the precise location of β5-6 loop region of a potential C2 domain of APX-1 could not be iden-

tified. However, one reason why this region was targeted for mutagenesis was the presence of an aspartate within the loop, that could potentially contribute to the co-ordination calcium, if the potential C2 domain of APX-1 was calcium-dependent. Limited proteolysis of a fragment of APX-1 did not reveal a calcium-binding site, although neither exclude such a site. Previously, mutation of the β5-6 loop of human Jagged-1 suggested that single aspartate mutants were insufficient to substantially reduce activation of Notch-1, and that a double mutant (D140/144) was necessary achieve this (Chillakuri et al., 2013). Whilst the I80-K86/SS AC duplication data could not further determine the presence/absence of calcium co-ordination in APX-1, this mutant provides an example which highlights how information could potentially be integrated using the diverse approaches described in this thesis.

Study of the DSL(ETSD) and I80-K86/SS mutants demonstrated the power of the AC duplication assay to facilitate the integration of biochemical/structural and whole-organism/functional data, and illustrates the use of complementary methods to provide a powerful approach to advancing our understanding of the role of the C2 domain/MNNL of Notch ligands. *C. elegans* has the potential to provide novel insights into functionality of the C2 domain, and assimilating information from similar studies in multiple model organisms could contribute to give a global understanding of its physiological role.

8

Appendices

8.1 Appendix 1 - Codon usage graphs

Codon usage graphs are shown for the two codon optimised constructs APX-1_{NE4His} (Figure 8.1) and APX-1_{NE3His} (Figure 8.2).

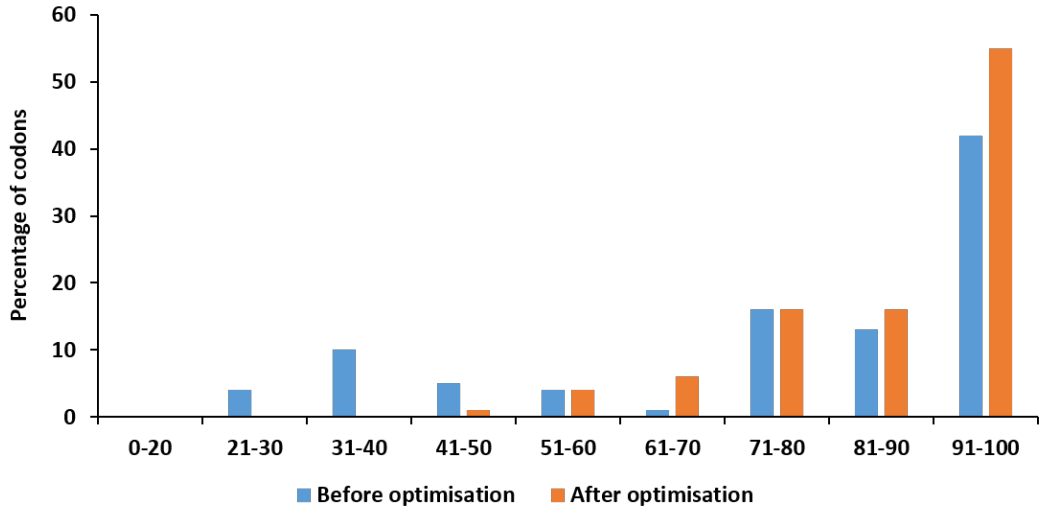


Figure 8.1: Frequency of optimal codons in APX-1_{NE4} sequence for human expression before and after optimisation by GenScript - Where a value of 100 is set for the codon with the highest frequency of usage for a given amino acid in humans. Adapted from information provided by Genscript.

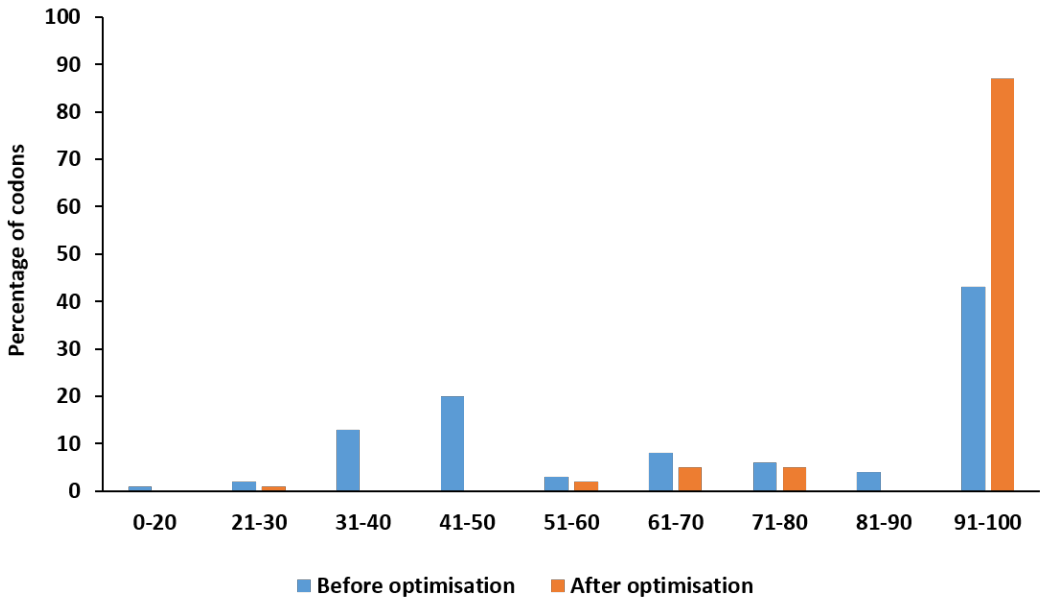


Figure 8.2: Frequency of optimal codons in APX-1_{NE3His} sequence for *Drosophila* expression before and after optimisation by Geneart - Where a value of 100 is set for the codon with the highest frequency of usage for a given amino acid in *Drosophila*. Adapted from information provided by Geneart.

8.2 Appendix 2 - FFAS primary sequence output

Primary sequence outputs used to generate the 3D superposed models in the main text are shown below, for the full length inputs of LAG-2 (Figure 8.3), DSL-1 (Figure 8.4) and APX-1 (Figure 8.5). All three are shown aligned to the sequence corresponding to the 4CBZ structure of $J1_{NE3}$.

```
  6 LLLLTCLPVLQARVEVHQEFISSKRVS-----VRFEIVTESHSPNRPVTF
 42 TYFKVCLKEYQSRVTAGGPCSFGSGSTPVIGGNTFNLKASRGNDRNRIVL

 51 DLFPRGPKTNIILLDTFNPVFNFSIQLVQPFTGQPLGDRIYRKVQFSGTN
 92 PFSFAWPRSYTLLVEAWDSSNDTVQPDSIIEKASHS---------GMINP

101 QPWINDTFTTTSGISLSVATEVTCARNYFGNRCENFCDAHLAKAARKRCD
133 SRQWQTLKQNTGVAHFEYQIRVTCDDYYYGFGCNKFCRPRDDFFGHYACD

151 AMGRLRCDIGWMGPHCGQAVDPRKCSCENDGICVS---SMIHPSQPNQTS
183 QNGNKTCMEGWMGPECNRAICRQGC-SPKHGSCKLPGDCRCQYGWQGLYC

198 SN-------------EQLICECTNGFTGTRCEIFGFNQFQLTAPRPDACS
232 DKCIPHPGCVHGICNEPWQCLCETNWGGQLCD-----------KDLNYCG

235 VKDACLNGAKCFP-NGPKVFCSCAVGFIGEFCEIS    268
271 THQPCLNGGTCSNTGPDKYQCSCPEGYSGPNCEIV    305
```

Figure 8.3: Full LAG-2/4CBZ FFAS alignment - conserved residues shown in red. Alignment score of -66.40.

```
 18 NGYMTIRLKS-----------AFPLNVTVEVAEEVYFPTNKRSFNLQLKP
  1 SGQFELEILSMQNVNGELQNGNCCGGARNPGDRKCTRDECDTYFKVCLKE

 57 NRVGFVSNIPVKFG----------------------------------RP
 51 YQSRVTAGGPCSFGSGSTPVIGGNTFNLKASRGNDRNRIVLPFSFAWPRS

 73 GLVLVDC-------SPVEKFQIHRVTLRSIRWNTHVRSVASDNIFLPFTG
101 YTLLVEAWDSSNDTVQPDSIIEKASHSGMINPSRQWQTLKQNTGVAHFEY

116 FRYDIKCNRYWHGLHCDHFCNDDFARTINRRCTQNGTLGCLEGFHGPNCE
151 -QIRVTCDDYYYGFGCNKFCRPRDDFFGHYACDQNGNKTCMEGWMGPECN

166 LPVPADSCKCQNGGKCVSSLENTWAQN---------------------GS
200 RAI-CRQGCSPKHGSCKLPGDCRCQYGWQGLYCDKCIPHPGCVHGICNEP

195 LICECRLGHFEGKHCEKKSFNYFPKIEATTYATKDSHL-ARQFYNNSRVP
249 WQCLCETN-WGGQLCDKDLNYCGTH-QPCLNGGTCSNTGPDKYQCSCPEG

244 NELATLPRL
297 YSGPNCEIV
```

Figure 8.4: Full DSL-1/4CBZ FFAS alignment - Conserved residues shown in red. Alignment score of -49.400.

```
  2 TNFSSLLTTIFLCIISSATGSGTIELLISSP-QTVLVEPTVCA------N
 26 GARNPGDRKCTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGSTPVIGGNT

 45 FECAAPDDLSLARKVQRRVPLRFGTGQYHGEARERIDLHLKIIEPTSNEI
 76 FNLKASRGNDRNRIVLPFSFAWPRSYTLLVEAWDSSNDTVQPDSIIEKAS

 95 LALQHHRPAADTEWNSDAPIVIETSRGFNVTVQLRNLCSSNYHGKRCNRY
126 HSGMINPSRQWQT-------LKQNTGVAHFEYQIRVTCDDYYYGFGCNKF

145 CIANAKLHWE--CSTHGVRRCSAGWSGEDCSNPIC-AGGCSNRGRCVAPN
169 CRPRDDFFGHYACDQNGNKTCMEGWMGPECNRAICRQGCSPKHGSCKLPG

192 QCSCADGFNGTRCEQCLPRAGCVNGDCVNETPNTCKCRDGFIGDRCDIDI
219 DCRCQYGWQGLYCDKCIPHPGCVHGICNE--PWQCLCETNWGGQLCDKDL

242 KICSLEKPCANGGICSIDSSSSTGYKCHCPFEFVGSQCKTP
267 NYCGTHQPCLNGGTCS--NTGPDKYQCSCPEGYSGPNCEIV
```

Figure 8.5: Full APX-1/4CBZ FFAS alignment - Conserved residues shown in red. Alignment score of -43.008.

8.3 Appendix 3 - Transgene transmission efficiency data

Transgene transmission effficiency was calculated by expressing the number of progeny carrying the transgene as a percentage of the total number of progeny in the brood. Full lists of transmission efficiencies of the transgenic lines generated in this work are shown in Figures 8.6, 8.7 and 8.8.

	Number of progeny per hermaphrodite					
	WT	*unc-119*	Total	% Progeny with transgene	Mean % transmission efficiency	STDEV
lag-2 fosmid 1	70	35	105	66.7	**69.1**	3.6
	66	32	98	67.3		
	93	34	127	73.2		
lag-2 fosmid 2	43	34	77	55.8	**56.5**	3.9
	53	47	100	53.0		
	76	49	125	60.8		
lag-2::gfp 1	76	53	129	58.9	**55.0**	3.6
	50	42	92	54.3		
	60	56	116	51.7		
lag-2::gfp 2	46	22	68	67.6	**68.0**	1.6
	74	32	106	69.8		
	70	35	105	66.7		
lag-2::gfp 3	49	30	79	62.0	**58.9**	2.9
	50	36	86	58.1		
	44	34	78	56.4		
dsl-1::gfp 1	36	24	60	60.0	**58.3**	2.3
	35	24	59	59.3		
	39	31	70	55.7		
dsl-1::gfp 2	48	41	89	53.9	**55.4**	1.4
	50	38	88	56.8		
	35	28	63	55.6		
dsl-1::gfp 3	36	35	71	50.7	**48.9**	2.9
	53	52	105	50.5		
	36	43	79	45.6		
apx-1::gfp 1	40	20	60	66.7	**66.8**	1.7
	48	22	70	68.6		
	47	25	72	65.3		
apx-1::gfp 2	81	34	115	70.4	**67.9**	4.8
	61	25	86	70.9		
	76	46	122	62.3		
apx-1::gfp 3	64	31	95	67.4	**71.8**	4.4
	48	15	63	76.2		
	56	22	78	71.8		

Figure 8.6: Transgene transmission efficiency data - For each line, the broods of three individual worms were scored to count the total number of transgenic (WT) and uncoordinated (*unc-119* worms were counted, and the number of progeny with the transgene was expressed as a percentage. The three percentages were then used to calculate a mean value, which was used for subsequent normalisation of data. Standard deviation was included to indicate experimental error.)

	Number of progeny per hermaphrodite					
	WT	*unc-119*	Total	%Progeny with transgene	Mean % transmission efficiency	STDEV
unc-119 only (lag-2(q420) 1	86	62	148	58.1	**55.9**	2.5
	93	82	175	53.1		
	84	65	149	56.4		
unc-119 only (lag-2(q420) 2	99	73	172	57.6	**61.1**	3.1
	95	55	150	63.3		
	88	52	141	62.4		
unc-119 only rescued 1	68	59	127	53.5	**53.7**	0.8
	72	64	136	52.9		
	89	74	163	54.6		
unc-119 only rescued 2	50	91	141	35.5	**34.3**	2.6
	49	107	156	31.4		
	51	90	141	36.2		
lag-2 only 1	72	69	141	51.1	**51.1**	2.6
	72	62	134	53.7		
	50	53	103	48.5		
lag-2 only 2	47	56	103	45.6	**42.7**	2.7
	48	66	114	42.1		
	37	55	92	40.2		
lag-2 only 3	66	39	105	62.9	**60.5**	2.3
	98	70	168	58.3		
	62	41	103	60.2		
rlag2p::lag-2::gfp 1	49	51	100	46	**54.0**	4.73
	66	47	113	58.4		
	90	75	165	54.5		
rlag2p::lag-2::gfp 2	69	82	151	45.7	**45.0**	3.52
	62	67	129	48.1		
	51	73	124	41.1		
rlag2p::lag-2::gfp 3	65	80	145	44.8	**42.0**	3.48
	43	70	113	38.1		
	50	66	116	43.1		

Figure 8.7: Transgene transmission efficiency data - For each line, the broods of three individual worms were scored to count the total number of transgenic (WT) and uncoordinated (*unc-119* worms were counted, and the number of progeny with the transgene was expressed as a percentage. The three percentages were then used to calculate a mean value, which was used for subsequent normalisation of data. Standard deviation was included to indicate experimental error.)

	Number of progeny per hermaphrodite					
	WT	*unc-119*	Total	%Progeny with transgene	Mean % transmission efficiency	STDEV
diluted *lag-2::gfp* 1	133	40	173	76.9	**72.0**	5.8
	82	43	125	65.6		
	95	34	129	73.6		
diluted *lag-2::gfp* 2	60	45	105	57.1	**59.3**	2.9
	82	49	131	62.6		
	64	46	110	58.2		

Figure 8.8: Transgene transmission efficiency data - For each line, the broods of three individual worms were scored to count the total number of transgenic (WT) and uncoordinated (*unc-119* worms were counted, and the number of progeny with the transgene was expressed as a percentage. The three percentages were then used to calculate a mean value, which was used for subsequent normalisation of data. Standard deviation was included to indicate experimental error.)

8.4 Appendix 4 - Raw larval viability data and its normalisation

All transgenic brood sizes were adjusted as a percentage of transgenics from the maximum potential brood size. The value for maximum total brood size was the average taken from '*unc-119* only rescued line 1' average brood size (142, from data above) to account for the presence of the *unc-119* rescue background. The transmission efficiency (TE) adjusted brood size was calculated by multiplying the average TE for a given line by the maximum brood size value (142). This gave the maxiumum amount of worms that would be anticipated to be transgenic within a line. The values calculated are shown in Figure 8.9.

In addition to accounting for the variation in transgene transmission efficiency, a second factor was accounted for when representing larval viability data. It was observed that *unc-119* only rescued worms, in the absence of the *lag-2(q420)* background, were not achieving full rescue (i.e observed number of progeny surviving to L2 and beyond over the maximum potential number of progeny was not 100%).

Strain/Line	Transgene transmission efficiency (%)	Maximum total brood	Transmission efficiency adjusted maximum transgenic brood
lag-2 fosmid 1	69.1	142	98.1
lag-2 fosmid 2	56.5	142	80.2
lag-2::gfp 1	55	142	78.1
lag-2::gfp 2	68	142	96.6
lag-2::gfp 3	58.9	142	83.6
dsl-1::gfp 1	58.3	142	82.8
dsl-1::gfp 2	55.4	142	78.7
dsl-1::gfp 3	48.9	142	69.4
apx-1::gfp 1	66.8	142	94.9
apx-1::gfp 2	67.9	142	96.4
apx-1::gfp 3	71.8	142	102.0
unc-119 only (*lag-2(q420)*) 1	55.9	142	79.4
unc-119 only (*lag-2(q420)*) 2	61.1	142	86.8
unc-119 only rescued 1	53.7	142	76.3
unc-119 only rescued 2	34.3	142	48.7
l*ag-2* only 1	51.1	142	72.6
lag-2 only 2	42.7	142	60.6
lag-2 only 3	60.5	142	85.9
rlag2p::lag-2::gfp 1	54.0	142	76.7
rlag2p::lag-2::gfp 2	45.0	142	63.9
rlag2p::lag-2::gfp 3	42.0	142	59.6
diluted *lag-2::gfp* 1	72	142	102.2
diluted *lag-2::gfp* 2	59.3	142	84.2

Figure 8.9: Maximum transgenic broods for each transgenic line calculated using transmission efficiency data - The maximum total brood was 142 for all lines, as this accounted for effects the *unc-119* rescue background. The maxiumum number of transgenic worms could be calculated by applying the percentage transgene transmission efficiency value for a given line to 142, to give the transmission efficiency adjusted brood

It was anticipated that transgenic lines with the *unc-119* background and the *lag-2(q420)* allele present would not rescue to a greater extent than the *unc-119* only rescued worms. Therefore, the larval viability data would needed to be normalised to the *unc-119* only rescued worms, defining this control line as representative of 100% rescue of the Lag phenotype.

Larval viability was measured in two independent lines of *unc-119* only rescued worms, as shown in Figure 8.10, and the average value of 76.5% from the *unc-119* only rescued line 1 was chosen as it was associated with the smaller error of the two lines.

	***unc-119* only rescued line 1**		***unc-119* only rescued line 2**	
	Raw larval viability data (Number of progeny surviving to L2 and beyond)	**Transformed larval viability data accounting for TE**	**Raw larval viability data**	**Transformed larval viability data accounting for TE**
	37	48.5	17	34.9
	72	94.4	16	32.9
	57	74.7	19	39.0
	68	89.1	35	71.9
	37	48.5	22	45.2
	45	59.0	71	145.8
	59	77.3	49	100.6
	51	66.8	32	65.7
	84	110.1	36	73.9
	43	56.4	55	112.9
	32	41.9		
	92	120.6		
	82	107.5		
Mean	58.4	**76.5**	35.2	72.3
SEM		7.2		11.9

Figure 8.10: *unc-119* only rescued lines were not associated with 100% progeny survival after transgene transmission efficiency is accounted for - Values are as follows: Raw larval viability data represents the raw score of the number of progeny from a single adult hermaphrodite surviving to L2 and beyond. Transformed larval viability data takes the raw data in the previous column, and expresses it as a percentage of the maximum transgenic brood size calculated for that particular line (values shown in final column of Figure 8.9). Note that for these control lines, only 72.3-77.5% of the maximum transgenic brood are survived, indicating that the *unc-119* rescue background contributed to a reduction in survival, irrespective of the *lag-2(q420)* allele.

Thus, the percentage of progeny surviving to L2 and beyond was calculated for each line by first adjusting values for transmission efficiency, and subsequently, relative to the *unc-119* only rescued line 1 (which was assumed to represent 100% rescue of the Lag phenotype). The raw data, T.E adjusted data and final *unc-119* data for each transgenic line is summarised in Figures 8.11, 8.12, 8.13, 8.14, 8.15 and 8.16.

Graphs representing the results from multiple independent lines shown the mean and ±SEM for all the values recorded for that particular line. For statistical analysis, the individual values from multiple transgenic lines of a single construct were pooled, and the mean and ±SEM indicated for the entire data set.

	lag-2 fosmid 1			*lag-2* fosmid 2		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	23	23.4	30.6	6	7.5	9.8
	50	51.0	66.6	6	7.5	9.8
	22	22.4	29.3	12	15.0	19.6
	21	21.4	28.0	8	10.0	13.0
	54	55.0	72.0	2	2.5	3.3
	23	23.4	30.6	76	94.8	123.9
	12	12.2	16.0	4	5.0	6.5
	22	22.4	29.3	5	6.2	8.1
	32	32.6	42.6	3	3.7	4.9
	17	17.3	22.7	8	10.0	13.0
	34	34.7	45.3	52	64.8	84.8
	11	11.2	14.7	43	53.6	70.1
Mean			**35.6**			**30.6**
St Error			5.2			11.4

	lag-2::gfp 1			*lag-2::gfp* 2		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	14	17.9	23.4	49	50.7	66.3
	43	55.1	72.0	67	69.4	90.7
	26	33.3	43.5	43	44.5	58.2
	45	57.6	75.3	73	75.6	98.8
	30	38.4	50.2	43	44.5	58.2
	30	38.4	50.2	39	40.4	52.8
	32	41.0	53.6	32	33.1	43.3
	16	20.5	26.8	23	23.8	31.1
	13	16.6	21.8	40	41.4	54.1
	15	19.2	25.1	40	41.4	54.1
	12	15.4	20.1	26	26.9	35.2
Mean			**42.0**			**58.4**
St Error			6.1			6.3

Figure 8.11: Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part I - Raw larval viability data represents the raw score of the number of progeny from a single adult hermaphrodite surviving to L2 and beyond. Transformed larval viability data takes the raw data in the previous column, and expresses it as a percentage of the maximum transgenic brood size calculated for that particular line (values shown in final column of Figure 8.9). This percentage was then divided by 76.5, the average transgenic brood of the *unc-119* only rescued control line (Figure 8.10) to account for the effect on larval viability contributed by the *unc-119* rescue background (*unc-119* adjusted). This final value was termed the 'percentage of progeny surviving to L2 and beyond" and these final values were used when plotting graphs shown in the main text.

	lag-2::gfp 3			*dsl-1::gfp* 1		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	46	55.0	71.9	0	0.0	0.0
	5	6.0	7.8	9	10.9	14.2
	34	40.7	53.2	3	3.6	4.7
	46	55.0	71.9	4	4.8	6.3
	4	4.8	6.3	4	4.8	6.3
	59	70.6	92.3	10	12.1	15.8
	35	41.9	54.7	9	10.9	14.2
	43	51.4	67.2	2	2.4	3.2
	34	40.7	53.2	9	10.9	14.2
	25	29.9	39.1	0	0.0	0.0
	18	21.5	28.1	6	7.3	9.5
				1	1.2	1.6
Mean			**49.6**			**7.5**
St Error			8.2			1.7

	dsl-1::gfp 2			*dsl-1::gfp* 3		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	1	1.3	1.7	2	2.9	3.8
	4	5.1	6.6	0	0.0	0.0
	12	15.2	19.9	4	5.8	7.5
	12	15.2	19.9	7	10.1	13.2
	3	3.8	5.0	9	13.0	17.0
	15	19.1	24.9	4	5.8	7.5
	3	3.8	5.0	4	5.8	7.5
	5	6.4	8.3	2	2.9	3.8
	11	14.0	18.3	6	8.6	11.3
	3	3.8	5.0	1	1.4	1.9
	2	2.5	3.3	8	11.5	15.1
	4	5.1	6.6	2	2.9	3.8
Mean			**10.4**			**7.7**
St Error			2.3			1.6

Figure 8.12: Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part II - Raw larval viability data represents the raw score of the number of progeny from a single adult hermaphrodite surviving to L2 and beyond. Transformed larval viability data takes the raw data in the previous column, and expresses it as a percentage of the maximum transgenic brood size calculated for that particular line (values shown in final column of Figure 8.9). This percentage was then divided by 76.5, the average transgenic brood of the *unc-119* only rescued control line (Figure 8.10) to account for the effect on larval viability contributed by the *unc-119* rescue background (*unc-119* adjusted). This final value was termed the 'percentage of progeny surviving to L2 and beyond" and these final values were used when plotting graphs shown in the main text.

	apx-1::gfp 1			*apx-1::gfp* 2		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	4	4.2	5.5	14	14.5	19.0
	27	28.5	37.2	23	23.9	31.2
	5	5.3	6.9	23	23.9	31.2
	6	6.3	8.3	10	10.4	13.6
	17	17.9	23.4	6	6.2	8.1
	49	51.6	67.5	25	25.9	33.9
	10	10.5	13.8	10	10.4	13.6
	8	8.4	11.0	21	21.8	28.5
	7	7.4	9.6	6	6.2	8.1
	32	33.7	44.1	8	8.3	10.8
	11	11.6	15.2	21	21.8	28.5
Mean			**22.0**			**20.6**
St Error			5.9			3.0

	apx-1::gfp 3			*unc-119* only (*lag-2(q420)*) 1		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	47	46.1	60.2	2	2.5	3.3
	29	28.4	37.2	9	11.3	14.8
	30	29.4	38.4	6	7.6	9.9
	56	54.9	71.8	11	13.9	18.1
	49	48.0	62.8	4	5.0	6.6
	39	38.2	50.0	2	2.5	3.3
	12	11.8	15.4	5	6.3	8.2
	32	31.4	41.0	4	5.0	6.6
	29	28.4	37.2	3	3.8	4.9
	22	21.6	28.2	2	2.5	3.3
	29	28.4	37.2			
Mean			**43.6**			**7.9**
St Error			4.9			1.6

Figure 8.13: Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part III - Raw larval viability data represents the raw score of the number of progeny from a single adult hermaphrodite surviving to L2 and beyond. Transformed larval viability data takes the raw data in the previous column, and expresses it as a percentage of the maximum transgenic brood size calculated for that particular line (values shown in final column of Figure 8.9). This percentage was then divided by 76.5, the average transgenic brood of the *unc-119* only rescued control line (Figure 8.10) to account for the effect on larval viability contributed by the *unc-119* rescue background (*unc-119* adjusted). This final value was termed the 'percentage of progeny surviving to L2 and beyond" and these final values were used when plotting graphs shown in the main text.

	***unc-119* only (*lag-2(q420)*) 2**			***lag-2* only 1**		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	2	2.3	3.0	66	90.9	118.8
	8	9.2	12.0	44	60.6	79.2
	9	10.4	13.6	28	38.6	50.4
	7	8.1	10.5	33	45.5	59.4
	11	12.7	16.6	46	63.4	82.8
	12	13.8	18.1	14	19.3	25.2
	13	15.0	19.6	30	41.3	54.0
	5	5.8	7.5	15	20.7	27.0
	10	11.5	15.1	38	52.3	68.4
	3	3.5	4.5	22	30.3	39.6
				11	15.2	19.8
Mean			**12.0**			**56.8**
St Error			1.8			8.9

	***lag-2* only 2**			***lag-2* only 3**		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	34	56.1	73.3	21	24.4	32.0
	61	100.7	131.6	69	80.3	105.0
	42	69.3	90.6	30	34.9	45.7
	23	38.0	49.6	22	25.6	33.5
	40	66.0	86.3	57	66.4	86.7
	27	44.6	58.2	40	46.6	60.9
	23	38.0	49.6	31	36.1	47.2
	29	47.9	62.6	37	43.1	56.3
	13	21.5	28.0	34	39.6	51.7
	9	14.9	19.4	29	33.8	44.1
	24	39.6	51.8	48	55.9	73.0
Mean			**63.7**			**57.8**
St Error			9.4			6.8

Figure 8.14: Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part IV - Raw larval viability data represents the raw score of the number of progeny from a single adult hermaphrodite surviving to L2 and beyond. Transformed larval viability data takes the raw data in the previous column, and expresses it as a percentage of the maximum transgenic brood size calculated for that particular line (values shown in final column of Figure 8.9). This percentage was then divided by 76.5, the average transgenic brood of the *unc-119* only rescued control line (Figure 8.10) to account for the effect on larval viability contributed by the *unc-119* rescue background (*unc-119* adjusted). This final value was termed the 'percentage of progeny surviving to L2 and beyond" and these final values were used when plotting graphs shown in the main text.

	diluted *lag-2::gfp* 1			diluted *lag-2::gfp* 2		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	13	12.7	16.6	23	27.3	35.7
	18	17.6	23.0	31	36.8	48.1
	56	54.8	71.6	19	22.6	29.5
	39	38.2	49.9	27	32.1	41.9
	90	88.1	115.1	20	23.8	31.0
	20	19.6	25.6	13	15.4	20.2
	28	27.4	35.8	8	9.5	12.4
	9	8.8	11.5	60	71.3	93.1
	61	59.7	78.0	13	15.4	20.2
	27	26.4	34.5	34	40.4	52.8
	26	25.4	33.3	12	14.3	18.6
Mean			**45.0**			**36.7**
St Error			9.5			6.8

Figure 8.15: Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part V - Raw larval viability data represents the raw score of the number of progeny from a single adult hermaphrodite surviving to L2 and beyond. Transformed larval viability data takes the raw data in the previous column, and expresses it as a percentage of the maximum transgenic brood size calculated for that particular line (values shown in final column of Figure 8.9). This percentage was then divided by 76.5, the average transgenic brood of the *unc-119* only rescued control line (Figure 8.10) to account for the effect on larval viability contributed by the *unc-119* rescue background (*unc-119* adjusted). This final value was termed the 'percentage of progeny surviving to L2 and beyond" and these final values were used when plotting graphs shown in the main text.

	rlag-2p::lag-2::gfp 1			*rlag-2p::lag-2::gfp* 2		
	Raw	TE adjusted	*unc-119* adjusted	Raw	TE adjusted	*unc-119* adjusted
	14	21.9	28.6	7	11.7	15.4
	14	21.9	28.6	6	10.1	13.2
	13	20.3	26.6	10	16.8	21.9
	11	17.2	22.5	7	11.7	15.4
	11	17.2	22.5	2	3.4	4.4
	6	9.4	12.3	8	13.4	17.5
	11	17.2	22.5	4	6.7	8.8
	8	12.5	16.4	8	13.4	17.5
	7	11.0	14.3	4	6.7	8.8
	17	26.6	34.8	5	8.4	11.0
	18	28.2	36.8	6	10.1	13.2
	13	20.3	26.6	9	15.1	19.7
	6	9.4	12.3	7	11.7	15.4
Mean			**23.4**			**14.0**
St Error			2.2			1.3

	rlag-2p::lag-2::gfp 3		
	Raw	TE adjusted	*unc-119* adjusted
	4	5.2	6.8
	15	19.6	25.6
	13	16.9	22.2
	4	5.2	6.8
	13	16.9	22.2
	1	1.3	1.7
	3	3.9	5.1
	3	3.9	30.7
	18	23.5	18.7
	11	14.3	8.5
	5	6.5	1.7
	1	1.3	15.3
	9	11.7	12.3
Mean			**13.1**
St Error			2.7

Figure 8.16: Normalisation of larval viability data to account for transmission efficiency and effects of the *unc-119* rescue background part VI - Raw larval viability data represents the raw score of the number of progeny from a single adult hermaphrodite surviving to L2 and beyond. Transformed larval viability data takes the raw data in the previous column, and expresses it as a percentage of the maximum transgenic brood size calculated for that particular line (values shown in final column of Figure 8.9). This percentage was then divided by 76.5, the average transgenic brood of the *unc-119* only rescued control line (Figure 8.10) to account for the effect on larval viability contributed by the *unc-119* rescue background (*unc-119* adjusted). This final value was termed the 'percentage of progeny surviving to L2 and beyond" and these final values were used when plotting graphs shown in the main text.

8.5 Appendix 5 - Raw brood fate assay data and conversion to percentages

Raw data is shown for the brood fate assays presented in the main text (Figure 8.17). Initially the number of Lag and non-Lag worms was quantified, however, these data were combined so as not to distract from the lethality data. Percentages of the brood were directly calculated using the total number of dead eggs and the total number of hatched larvae over the total number of eggs laid (Figure 8.18).

8.6 Appendix 6 - *lag-2* promoter considerations

The lag-2 promoter region of three *Caenorhabditis* species (*elegans*, *briggsae* and *remanei*) were aligned, the aim of which was two-fold: First, it was anticipated that common regulatory motifs would be conserved between species, and this might identify fragments of the promoter region most likely to be important for regulation of lag-2 expression. Secondly, that in the event that a high level of sequence conservation was conserved, the lag-2 promoter region of either C. briggsae or C. remanei might be explored as a substitute for the C. elegans lag-2 promoter, allowing a fully length promoter fragment for mosSCI. Alignment was carried out using CLC sequence viewer, and results (Figures 8.19, 8.20, 8.21, 8.22, 8.23, 8.24 and 8.25) suggested that there are very few regions of conservation between the three species. Therefore, this approach was not pursued further.

Raw data of total dead eggs and hatched larvae for given strains/lines

Strains	Total dead eggs	Total Lags	Total non-Lags	Total progeny
WT	1	0	207	208
WT	0	0	221	221
lag-2(q420)	10	207	1	218
lag-2(q420)	18	179	0	197
Transgenics	**Total dead eggs**	**Total Lags**	**Total non-Lags**	**Total progeny**
unc-119 only Lag 1	23	151	6	180
unc-119 only Lag 1	18	223	5	246
lag-2::gfp 2	30	96	16	142
lag-2::gfp 2	52	80	30	162
lag-2::gfp 3	74	20	6	100
lag-2::gfp 3	106	29	14	149
unc-119 only 1	18	0	138	156
unc-119 only 1	4	0	165	169
unc-119[lag-2::gfp] 1	46	173		219
unc-119[lag-2::gfp] 1	34	144		178

Figure 8.17: Raw data for brood fate assays - Each row represents the fates of the progeny from a single hermaphrodite of the named strain/transgenic line. The number of dead eggs, Lag and non-Lag larvae were added to give the total number of progeny within that brood.

Percentages of total dead eggs and hatched larvae for given strains/lines				
	Total dead eggs	Total Lags	Total non-Lags	Total hatched larva
WT	0.5	0.0	99.5	99.5
WT	0.0	0.0	100.0	100.0
lag-2(q420)	4.6	95.0	0.5	95.4
lag-2(q420)	9.1	90.9	0.0	90.9
Transgenics				
	Total dead eggs	Total Lags	Total non-Lags	Total hatched larva
unc-119 only (*lag-2(q420)*) 1	12.8	83.9	3.3	87.2
unc-119 only (*lag-2(q420)*) 1	7.3	90.7	2.0	92.7
lag-2::gfp 2	21.1	67.6	11.3	78.9
lag-2::gfp 2	32.1	49.4	18.5	67.9
lag-2::gfp 3	74.0	20.0	6.0	26.0
lag-2::gfp 3	71.1	19.5	9.4	28.9
unc-119 only 1	11.5	0.0	88.5	88.5
unc-119 only 1	2.4	0.0	97.6	97.6
unc-119[lag-2::gfp] 1	21.0			79.0
unc-119[lag-2::gfp] 1	19.1			80.9

Figure 8.18: Brood fates for strains and transgenic lines expressed as percentages of the total brood - Each row represents the fates of the progeny from a single hermaphrodite of the named strain/transgenic line. Percentages were calculated using the raw data shown in Figure 8.17, by expressing the total number of dead eggs and hatched larvae as a percentage of the total brood.

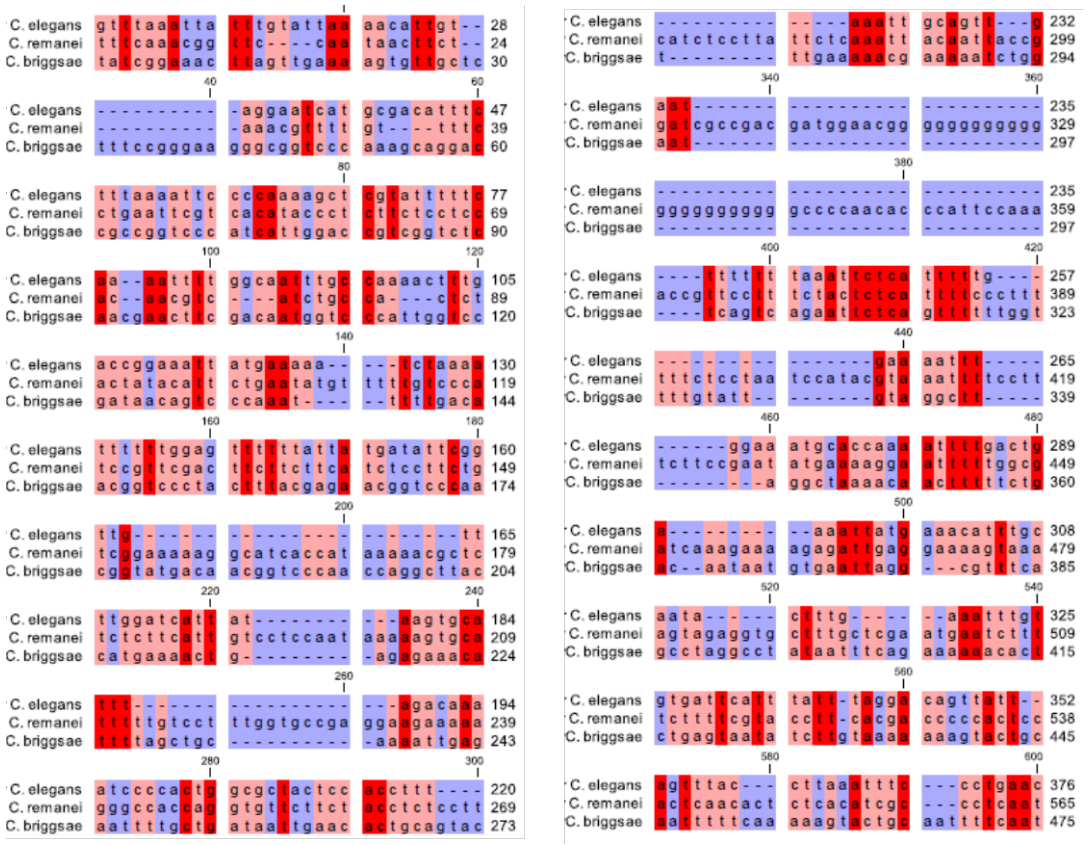


Figure 8.19: Sequence alignment of *Caenorhabditis lag-2* promoter regions part I - A 3.2 kb fragment of the 5' flanking region of *lag-2* was used for alignment. Bases conserved in three/two and one species are highlighted in red, pink and blue, respectively.

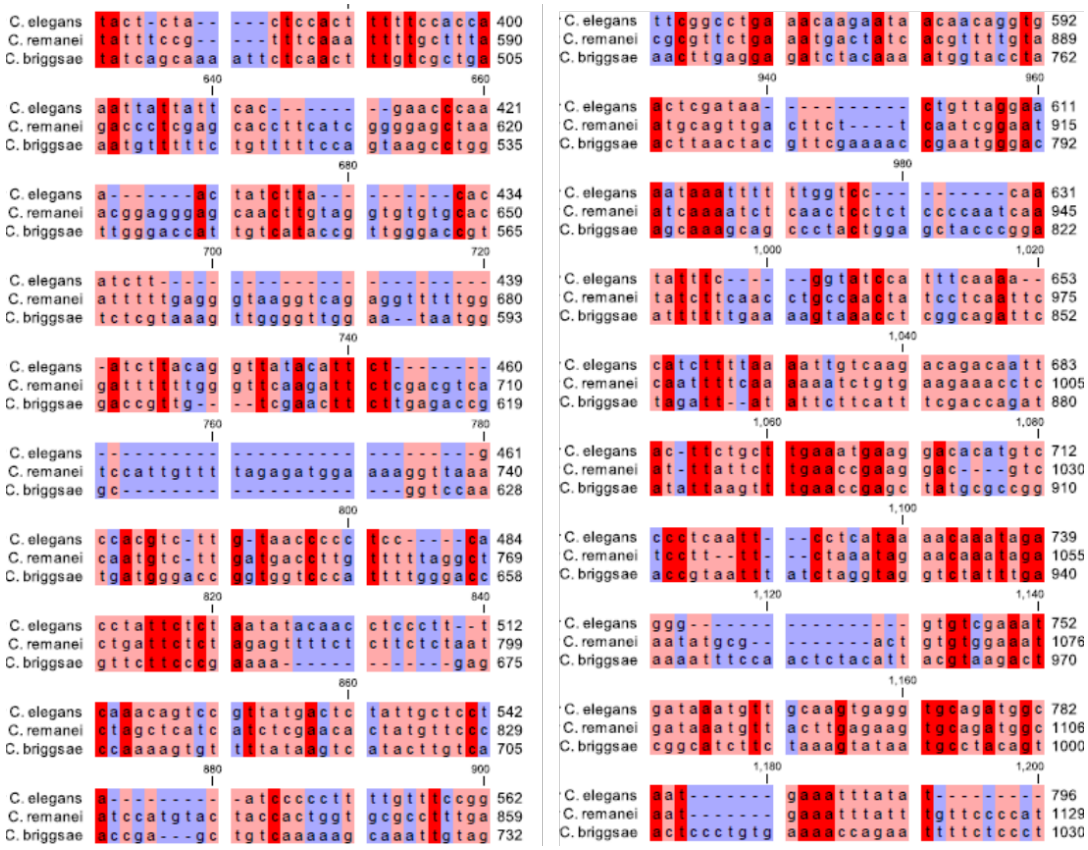


Figure 8.20: Sequence alignment of *Caenorhabditis lag-2* promoter regions part II - A 3.2 kb fragment of the 5' flanking region of *lag-2* was used for alignment. Bases conserved in three/two and one species are highlighted in red, pink and blue, respectively.

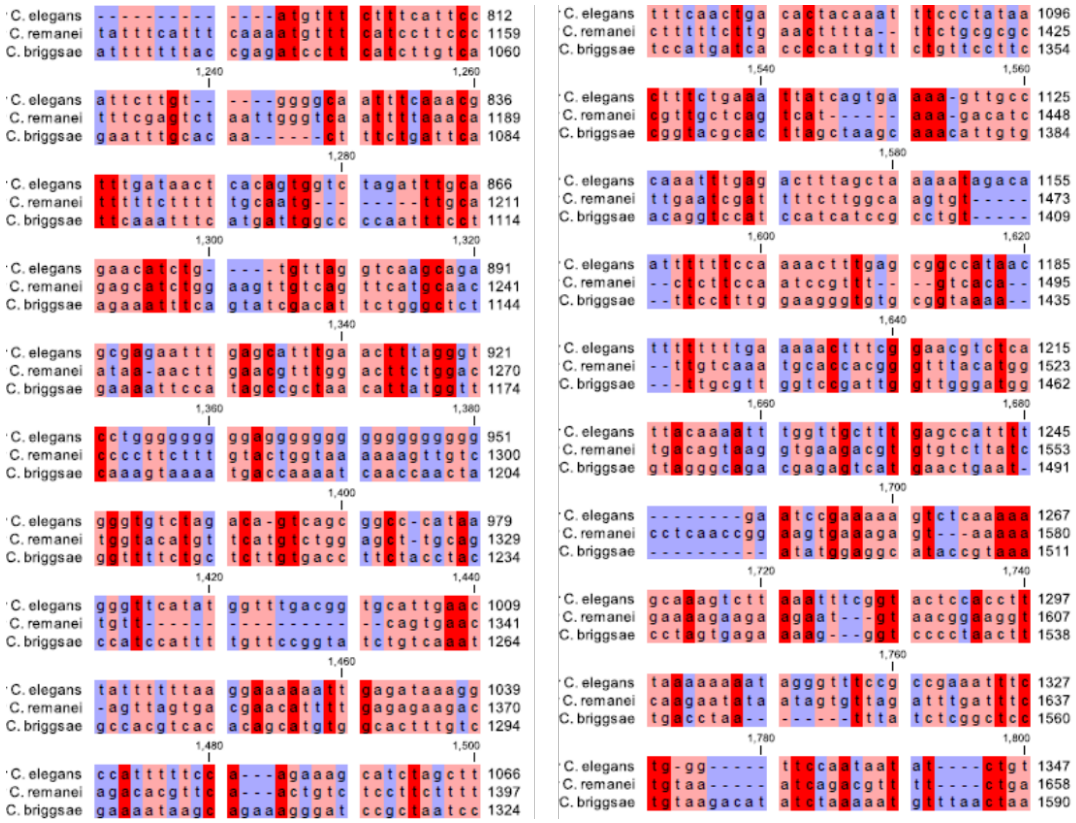


Figure 8.21: Sequence alignment of *Caenorhabditis lag-2* promoter regions part II - A 3.2 kb fragment of the 5' flanking region of *lag-2* was used for alignment. Bases conserved in three/two and one species are highlighted in red, pink and blue, respectively.

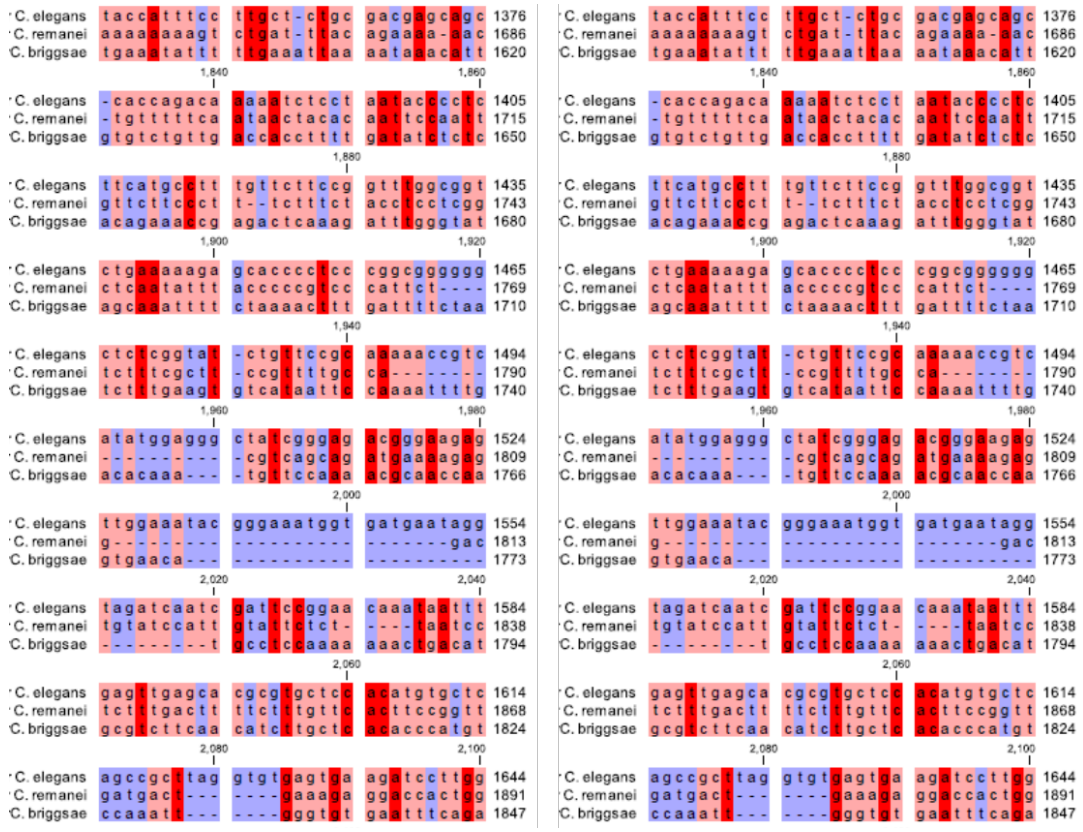


Figure 8.22: Sequence alignment of *Caenorhabditis lag-2* promoter regions part IV - A 3.2 kb fragment of the 5' flanking region of *lag-2* was used for alignment. Bases conserved in three/two and one species are highlighted in red, pink and blue, respectively.

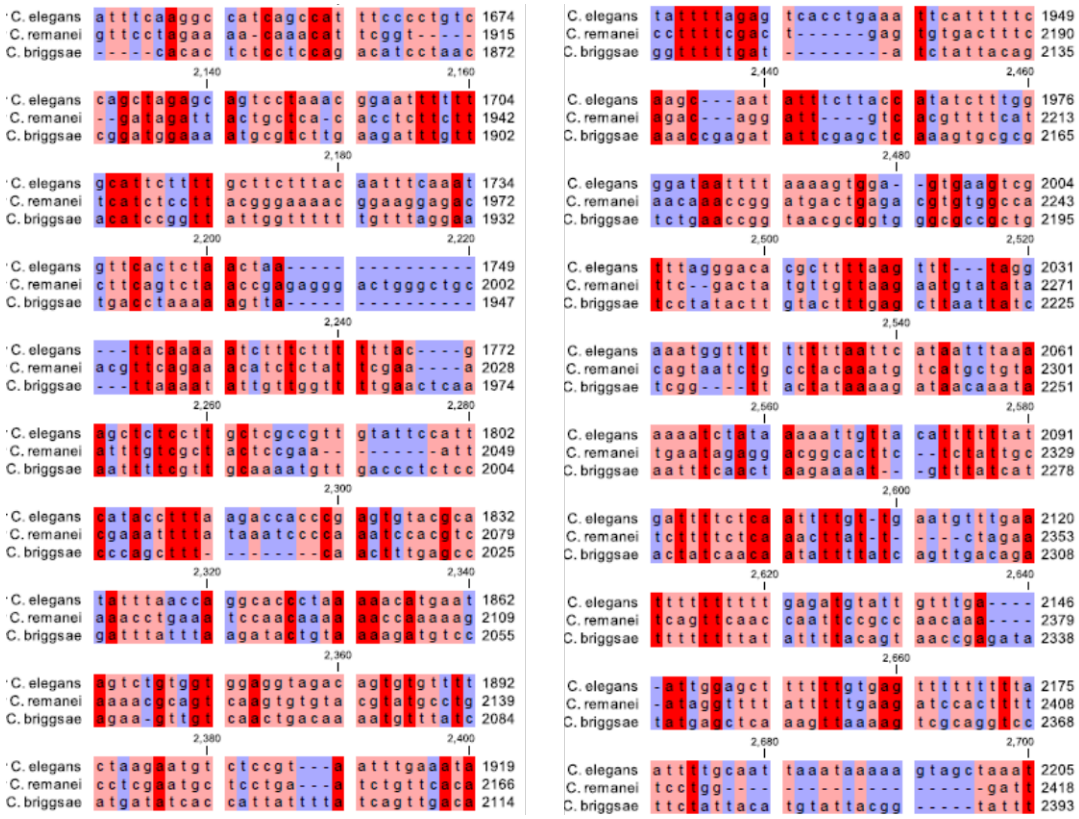


Figure 8.23: Sequence alignment of *Caenorhabditis lag-2* promoter regions part V - A 3.2 kb fragment of the 5' flanking region of *lag-2* was used for alignment. Bases conserved in three/two and one species are highlighted in red, pink and blue, respectively.

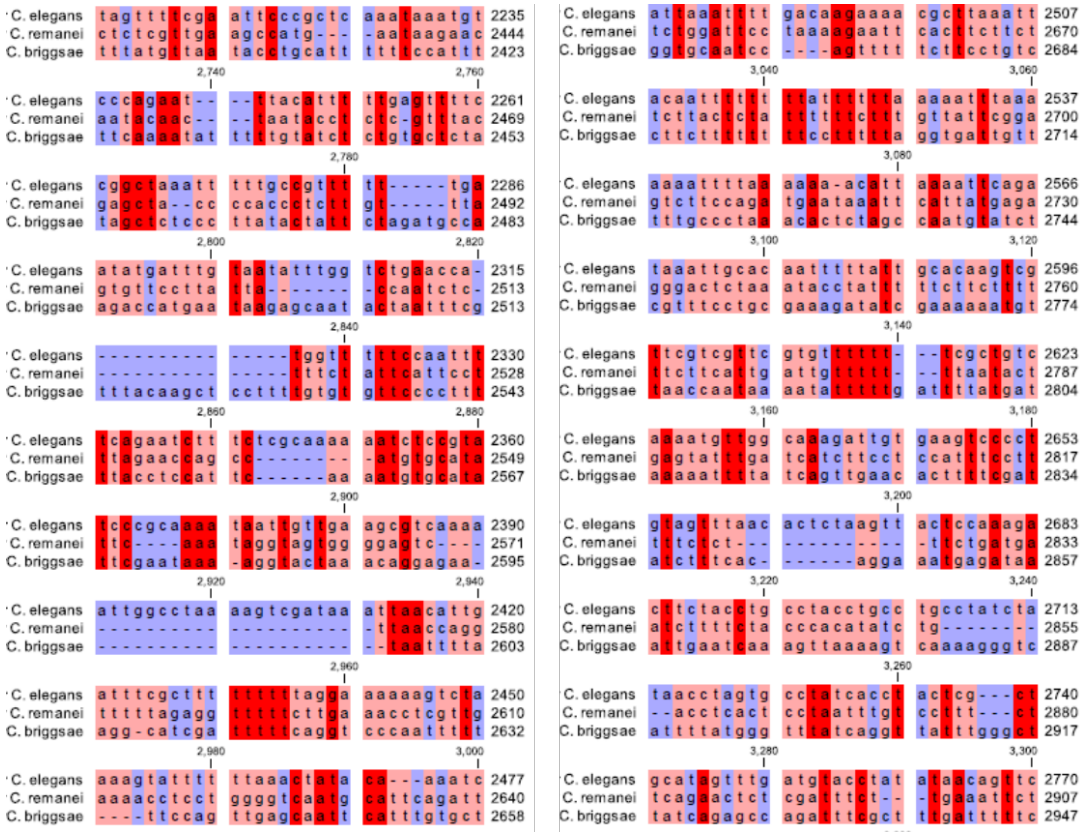


Figure 8.24: Sequence alignment of *Caenorhabditis lag-2* promoter regions part VI - A 3.2 kb fragment of the 5' flanking region of *lag-2* was used for alignment. Bases conserved in three/two and one species are highlighted in red, pink and blue, respectively.

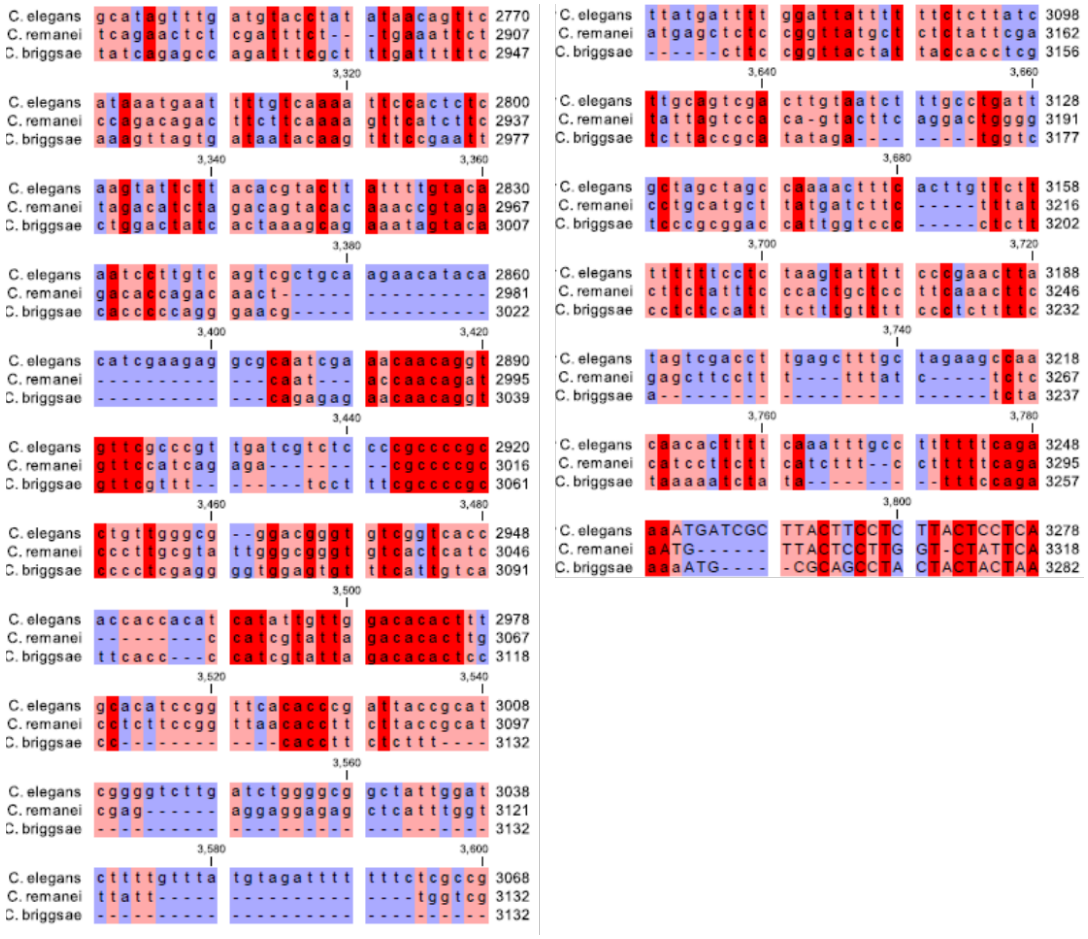


Figure 8.25: Sequence alignment of *Caenorhabditis lag-2* promoter regions part VII - A 3.2 kb fragment of the 5' flanking region of *lag-2* was used for alignment. Bases conserved in three/two and one species are highlighted in red, pink and blue, respectively..Uppercase characters represent the start of the ORF of *lag-2*.

8.7 Appendix 7 - Raw AC duplication assay data and conversion to percentage reduction in phenotype

The raw data recorded for AC duplication assays and the conversion of this data to give the final value shown in the main text is summarised in Figures 8.26, 8.27 and 8.28. The number of worms with a single (1AC) or duplicated (2AC) anchor cell were scored, and the sum of the worms was the total. The number of worms with 2ACs was expressed as a percentage of the total worms observed, to give the percentage of worms with duplicated AC. This value was then divided by 52.2 (the percentage of duplication associated with the *sur-5::gfp* marker control strain, and thus the maximum penetrance of the phenotype in this assay), multiplied by 100 (to give a percentage) and this figure subtracted from 100% to give the percentage reduction of the AC duplication phenotype. For each line, 2-4 independent experiments (where $16 \leq n \leq 50$ worms per experiment) were recorded and the mean percentage reduction of AC duplication plotted on graphs for each line in the main text. For statistical analysis, the values from each individual experiment of the multiple lines for a construct were pooled.

8.7 Appendix 7 - Raw AC duplication assay data and conversion to percentage reduction in phenotype

	1AC	2AC	Total	Percentage of worms with duplicated AC	Percentage reduction of AC duplication	SEM
AC marker control	30	0	30	0	100	
	30	0	30	0	100	
Total	60	0	60	0	**100**	0
***sur-5::gfp* marker control**	25	25	50	50.0	4.2	
	18	22	40	55.0	-5.4	
Total	43	47	90	52.2	**0.0**	4.79
***lag-2::gfp*(AC) 1**	19	10	29	34.5	33.9	
	10	7	17	41.2	21.1	
	15	6	21	28.6	45.3	
	13	8	21	38.1	27.0	
Total	57	31	88	35.2	**32.5**	5.19
***lag-2::gfp*(AC) 2**	15	10	25	40.0	23.4	
	11	5	16	31.3	40.1	
	16	12	28	42.9	17.9	
Total	42	27	69	39.1	**25.0**	6.69
***apx-1::gfp*(AC) 1**	16	7	23	30.4	41.7	
	13	6	19	31.6	39.5	
	14	8	22	36.4	30.3	
	21	13	34	38.2	26.8	
Total	64	34	98	34.7	**33.5**	3.58
***apx-1::gfp*(AC) 2**	19	11	32	34.4	34.1	
	22	13	34	38.2	26.8	
	12	7	18	38.9	25.5	
Total	53	31	84	36.9	**29.3**	2.70

Figure 8.26: AC duplication assay data summary for control lines and WT ligand constructs - Raw data is shown as the number of worms with 1 or 2 ACs, and the total of worms observed in a given experiment. From this, the percentage of the total worms with 2ACs was calculated. This was then converted to the percentage reduction of AC duplication relative to the *sur-5::gfp* marker strain (shown) by dividing by 52.2 and subtracting the value from 100. Bold underline values were the mean value calculated for each line and used in graphs shown in the main text.

8.7 Appendix 7 - Raw AC duplication assay data and conversion to percentage reduction in phenotype

	1AC	2AC	Total	Percentage of worms with duplicated AC	Percentage reduction of AC duplication	SEM
***ΔMNNL apx-1::gfp*(AC) 1**	10	14	24	58.3	-11.7	
	14	16	30	53.3	-2.2	
	10	11	21	52.4	-0.3	
	8	9	17	52.9	-1.4	
Total	42	50	92	54.3	**-4.1**	2.64
***ΔMNNL apx-1::gfp*(AC) 2**	8	7	15	46.7	10.6	
	17	22	39	56.4	-8.1	
	12	12	24	50.0	4.2	
	8	8	16	50.0	4.2	
Total	45	49	94	52.1	**0.1**	3.904
***ΔDSL apx-1::gfp*(AC) 1**	8	10	18	55.6	-6.4	
	16	20	36	55.6	-6.4	
	10	11	21	52.4	-0.3	
	11	12	23	52.2	0.0	
Total	45	53	98	54.1	**-3.6**	1.815
***ΔDSL apx-1::gfp*(AC) 2**	8	10	18	55.6	-6.4	
	7	11	18	61.1	-17.1	
	8	10	18	55.6	-6.4	
	10	11	21	52.4	-0.3	
Total	33	42	75	56.0	**-7.3**	3.477
***ΔEGF-like 1 apx-1::gfp*(AC) 1**	14	10	24	41.7	20.2	
	12	8	20	40.0	23.4	
	13	9	22	40.9	21.6	
	15	10	25	40.0	23.4	
Total	54	37	91	40.7	**22.1**	0.77
***ΔEGF-like 1 apx-1::gfp*(AC) 2**	13	8	21	38.1	27.0	
	12	7	19	36.8	29.4	
	14	9	23	39.1	25.0	
	17	8	25	32.0	38.7	
Total	56	32	88	36.4	**30.3**	3.02

Figure 8.27: AC duplication assay data summary for whole domain deletion constructs - Raw data is shown as the number of worms with 1 or 2 ACs, and the total of worms observed in a given experiment. From this, the percentage of the total worms with 2ACs was calculated. This was then converted to the percentage reduction of AC duplication relative to the *sur-5::gfp* marker strain (Figure 8.26) by dividing by 52.2 and subtracting the value from 100. Bold underline values were the mean value calculated for each line and used in graphs shown in the main text.

	1AC	2AC	Total	Percentage of worms with duplicated AC	Percentage reduction of AC duplication	SEM
***DSL(ETSD) apx-1::gfp*(AC) 1**	14	11	25	44.0	15.7	
	14	10	24	41.7	20.2	
	12	8	20	40.0	23.4	
	15	10	25	40.0	23.4	
Total	55	39	94	41.5	**20.5**	1.813
***DSL(ETSD) apx-1::gfp*(AC) 2**	14	11	25	44.0	15.7	
	16	9	25	36.0	31.0	
	11	7	18	38.9	25.5	
	14	11	25	44.0	15.7	
Total	55	38	93	40.9	**21.7**	3.797
***I80-K86/SS apx-1::gfp*(AC) 1**	14	11	25	44.0	15.7	
	17	8	25	32.0	38.7	
	16	9	25	36.0	31.0	
	12	9	21	42.9	17.9	
Total	59	37	96	38.5	**26.2**	5.462
***I80-K86/SS apx-1::gfp*(AC) 2**	15	10	25	40.0	23.4	
	14	11	25	44.0	15.7	
	15	10	25	40.0	23.4	
	13	8	21	38.1	27.0	
Total	59	37	96	40.6	**22.2**	2.381

Figure 8.28: AC duplication assay data summary for loop mutant constructs - Raw data is shown as the number of worms with 1 or 2 ACs, and the total of worms observed in a given experiment. From this, the percentage of the total worms with 2ACs was calculated. This was then converted to the percentage reduction of AC duplication relative to the *sur-5::gfp* marker strain (Figure 8.27) by dividing by 52.2 and subtracting the value from 100. Bold underline values were the mean value calculated for each line and used in graphs shown in the main text.

Bibliography

Acar, M., Jafar-Nejad, H., Takeuchi, H., Rajan, A., Ibrani, D., Rana, N. A., Pan, H., Haltiwanger, R. S., and Bellen, H. J. (2008). Rumi is a CAP10 domain glycosyltransferase that modifies Notch and is required for Notch signaling. *Cell*, 132(2):247–258. 20

Alagille, D., Odièvre, M., Gautier, M., and Dommergues, J. P. (1975). Hepatic ductular hypoplasia associated with characteristic facies, vertebral malformations, retarded physical, mental, and sexual development, and cardiac murmur. *The Journal of Pediatrics*, 86(1):63–71. 12

Andrawes, M. B., Xu, X., Liu, H., Ficarro, S. B., Marto, J. A., Aster, J. C., and Blacklow, S. C. (2013). Intrinsic Selectivity of Notch 1 for Delta-like 4 Over Delta-like 1. *Journal of Biological Chemistry*, 288(35):25477–25489. 34, 38

Aricescu, A. R., Lu, W., and Jones, E. Y. (2006). A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta crystallographica. Section D, Biological crystallography*, 62(Pt 10):1243–1250. 46, 101

Arnett, K. L., Hass, M., McArthur, D. G., Ilagan, M. X. G., Aster, J. C., Kopan, R., and Blacklow, S. C. (2010). Structural and mechanistic insights into cooperative assembly of dimeric Notch transcription complexes. *Nature structural & molecular biology*, 17(11):1312–7. 26

Aroian, R. V., Levy, A. D., Koga, M., Ohshima, Y., Kramer, J. M., and Sternberg, P. W. (1993). Splicing in Caenorhabditis elegans does not require an AG at the 3' splice acceptor site. *Molecular and Cellular Biology*, 13(1):626–637. 108

Austin, J. and Kimble, J. (1987). glp-1 is required in the germ line for regulation of the decision between mitosis and meiosis in C. elegans. *Cell*, 51(4):589–599. 106, 112

Austin, J. and Kimble, J. (1989). Transcript analysis of glp-1 and lin-12, homologous genes required for cell interactions during development of C. elegans. *Cell*, 58(3):565–571. 106

Bauer, R. C., Laney, A. O., Smith, R., Gerfen, J., Morrissette, J. J. D., Woyciechowski, S., Garbarini, J., Loomes, K. M., Krantz, I. D., Urban, Z., Gelb, B. D., Goldmuntz, E., and Spinner, N. B. (2010). Jagged1 (JAG1) mutations in patients with tetralogy of fallot or pulmonic stenosis. *Human Mutation*, 31(5):594–601. 12

Becam, I., Fiuza, U.-M., Arias, A. M., and Milán, M. (2010). A role of receptor Notch in ligand cis-inhibition in Drosophila. *Current Biology: CB*, 20(6):554–560. 7

Berry, L. W., Westlund, B., and Schedl, T. (1997). Germ-line tumor formation caused by activation of glp-1, a Caenorhabditis elegans member of the Notch family of receptors. *Development (Cambridge, England)*, 124(4):925–936. 24, 112

Berset, T., Hoier, E. F., Battu, G., Canevascini, S., and Hajnal, A. (2001). Notch inhibition of RAS signaling through MAP kinase phosphatase LIP-1 during C. elegans vulval development. *Science (New York, N.Y.)*, 291(5506):1055–8. 112

Bettenhausen, B., Hrab de Angelis, M., Simon, D., Guénet, J. L., and Gossler, A. (1995). Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta. *Development (Cambridge, England)*, 121(8):2407–2418. 29

Bhanushali, A. A., Babu, S., Thangapandi, V. R., Pillai, R., Chheda, P., and Das, B. R. (2010). Mutations in the HD and PEST domain of Notch-1 receptor in T-cell acute lymphoblastic leukemia: report of novel mutations from Indian population. *Oncology Research*, 19(2):99–104. 15

Blaumueller, C. M., Qi, H., Zagouras, P., and Artavanis-Tsakonas, S. (1997). Intracellular cleavage of Notch leads to a heterodimeric receptor on the plasma membrane. *Cell*, 90(2):281–291. 23

Bray, S. J. (2006). Notch signalling: a simple pathway becomes complex. *Nature reviews. Molecular cell biology*, 7(9):678–89. 2

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics*, 77(1):71–94. 64, 103

Brou, C., Logeat, F., Gupta, N., Bessia, C., LeBail, O., Doedens, J. R., Cumano, A., Roux, P., Black, R. A., and Israël, A. (2000). A novel proteolytic cleavage involved in Notch signaling: the role of the disintegrin-metalloprotease TACE. *Molecular Cell*, 5(2):207–216. 5

Bulman, M. P., Kusumi, K., Frayling, T. M., McKeown, C., Garrett, C., Lander, E. S., Krumlauf, R., Hattersley, A. T., Ellard, S., and Turnpenny, P. D. (2000). Mutations in the human delta homologue, DLL3, cause axial skeletal defects in spondylocostal dysostosis. *Nature Genetics*, 24(4):438–441. 11

Byerly, L., Cassada, R. C., and Russell, R. L. (1976). The life cycle of the nematode Caenorhabditis elegans. I. Wild-type growth and reproduction. *Developmental biology*, 51(1):23–33. 104

Byrd, D. T., Knobel, K., Affeldt, K., Crittenden, S. L., and Kimble, J. (2014). A DTC Niche Plexus Surrounds the Germline Stem Cell Pool in Caenorhabditis elegans. *PLoS ONE*, 9(2):e88372. 114

Chandu, D., Huppert, S. S., and Kopan, R. (2006). Analysis of transmembrane domain mutants is consistent with sequential cleavage of Notch by gamma-secretase. *Journal of Neurochemistry*, 96(1):228–235. 7

Chen, N. and Greenwald, I. (2004). The lateral signal for LIN-12/Notch in C. elegans vulval development comprises redundant secreted and transmembrane DSL proteins. *Developmental Cell*, 6(2):183–192. 9, 82, 107, 112, 115, 119, 135

Chillakuri, C. R., Sheppard, D., Ilagan, M. X. G., Holt, L. R., Abbott, F., Liang, S., Kopan, R., Handford, P. A., and Lea, S. M. (2013). Structural Analysis Uncovers Lipid-Binding Properties of Notch Ligands. *Cell Reports*, 5(4):861–867. 9, 13, 34, 38, 84, 99, 167, 168, 177

Coffman, C. R., Skoglund, P., Harris, W. A., and Kintner, C. R. (1993). Expression of an extracellular deletion of Xotch diverts cell fate in Xenopus embryos. *Cell*, 73(4):659–671. 23

Cordle, J., Johnson, S., Zi Yan Tay, J., Roversi, P., Wilkin, M. B., de Madrid, B. H., Shimizu, H., Jensen, S., Whiteman, P., Jin, B., Redfield, C., Baron, M., Lea, S. M., and Handford, P. A. (2008a). A conserved face of the Jagged/Serrate DSL domain is involved in Notch trans-activation and cis-inhibition. *Nat Struct Mol Biol*, 15(8):849–857. 7, 17, 18, 29, 30, 38, 82, 167

Cordle, J., RedfieldZ, C., Stacey, M., van der Merwe, P. A., Willis, A. C., Champion, B. R., Hambleton, S., and Handford, P. A. (2008b). Localization of the Delta-like-1-binding Site in Human Notch-1 and Its Modulation by Calcium Affinity. *Journal of Biological Chemistry*, 283(17):11785–11793. 19, 99

Crittenden, S. L., Troemel, E. R., Evans, T. C., and Kimble, J. (1994). GLP-1 is localized to the mitotic region of the C. elegans germ line. *Development (Cambridge, England)*, 120(10):2901–2911. 23, 114

Crosnier*, C., Driancourt*, C., Raynaud*, N., DhornePollet*, S., Pollet*, N., Bernard, O., Hadchouel*, ., and Meunier-Rotival*, M. (1999). Mutations in JAGGED1 gene are predominantly sporadic in Alagille syndrome. *Gastroenterology*, 116(5):1141–1148. 13

de Celis, J. F. and Bray, S. (1997). Feed-back mechanisms affecting Notch activation at the dorsoventral boundary in the Drosophila wing. *Development (Cambridge, England)*, 124(17):3241–3251. 3, 7

de Celis, J. F. and Bray, S. J. (2000). The Abruptex domain of Notch regulates negative interactions between Notch, its ligands and Fringe. *Development (Cambridge, England)*, 127(6):1291–1302. 19

de la Pompa, J. L., Wakeham, A., Correia, K. M., Samper, E., Brown, S., Aguilera, R. J., Nakano, T., Honjo, T., Mak, T. W., Rossant, J., and Conlon, R. A. (1997). Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development (Cambridge, England)*, 124(6):1139–1148. 2

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., and Kopan, R. (1999). A presenilin-1-dependent gamma-secretase-like protease mediates release of Notch intracellular domain. *Nature*, 398(6727):518–522. 7

Dexter, J. (1914). The analysis of a case of continuous variation in Drosophila by a study of its linkage relations. *American Naturalist*, 48:712–758. 1

Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D., and Goldstein, B. (2015). Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. *Genetics*, 200(4):1035–49. 176

Dickinson, D. J., Ward, J. D., Reiner, D. J., and Goldstein, B. (2013). Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. *Nature Methods*, 10(10):1028–1034. 175

Downing, A. K., Knott, V., Werner, J. M., Cardy, C. M., Campbell, I. D., and 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. 17

Drapeau, G. R., Boily, Y., and Houmard, J. (1972). Purification and properties of an extracellular protease of Staphylococcus aureus. *The Journal of biological chemistry*, 247(20):6720–6. 100

Eldadah, Z. A., Hamosh, A., Biery, N. J., Montgomery, R. A., Duke, M., Elkins, R., and Dietz, H. C. (2001). Familial Tetralogy of Fallot caused by mutation in the jagged1 gene. *Human Molecular Genetics*, 10(2):163–169. 11

Ellisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D., and 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. 9, 15, 17

Fehon, R. G., Kooh, P. J., Rebay, I., Regan, C. L., Xu, T., Muskavitch, M. A. T., and 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. 17

Fernandez-Valdivia, R., Takeuchi, H., Samarghandi, A., Lopez, M., Leonardi, J., Haltiwanger, R. S., and Jafar-Nejad, H. (2011). Regulation of mammalian Notch signaling and embryonic development by the protein O-glucosyltransferase Rumi. *Development (Cambridge, England)*, 138(10):1925–1934. 20

Fire, A. and Waterston, R. H. (1989). Proper expression of myosin genes in transgenic nematodes. *The EMBO journal*, 8(11):3419–28. 130

Fitzgerald, K. and Greenwald, I. (1995). Interchangeability of Caenorhabditis elegans DSL proteins and intrinsic signalling activity of their extracellular domains in vivo. *Development*, 121(12):4275–4282. 82, 115, 119, 138

Fitzgerald, K., Wilkinson, H. A., and Greenwald, I. (1993). glp-1 can substitute for lin-12 in specifying cell fate decisions in Caenorhabditis elegans. *Development (Cambridge, England)*, 119(4):1019–1027. 18, 81, 107

Fleming, R. J., Hori, K., Sen, A., Filloramo, G. V., Langer, J. M., Obar, R. A., Artavanis-Tsakonas, S., and Maharaj-Best, A. C. (2013). An extracellular region of Serrate is essential for ligand-induced cis-inhibition of Notch signaling. *Development*, 140(9):2039–2049. 7, 29

Fortini, M. E. and Artavanis-Tsakonas, S. (1994). The suppressor of hairless protein participates in notch receptor signaling. *Cell*, 79(2):273–282. 25

Foster, G. G. (1975). Negative complementation at the notch locus of Drosophila melanogaster. *Genetics*, 81(1):99–120. 19

Friedland, A. E., Tzur, Y. B., Esvelt, K. M., Colaiácovo, M. P., Church, G. M., and Calarco, J. A. (2013). Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nature methods*, 10(8):741–3. 175

Frischmeyer, P. A. and Dietz, H. C. (1999). Nonsense-Mediated mRNA Decay in Health and Disease. *Human Molecular Genetics*, 8(10):1893–1900. 13

Frøkjæ r Jensen, C., Davis, M. W., Ailion, M., and Jorgensen, E. M. (2012). Improved Mos1-mediated transgenesis in C. elegans. *Nature methods*, 9(2):117–8. 140

Frø kjaer Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. r.-P., Grunnet, M., and Jorgensen, E. M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. *Nature genetics*, 40(11):1375–83. 140, 147, 173

Gao, D. and Kimble, J. (1995). APX-1 can substitute for its homolog LAG-2 to direct cell interactions throughout Caenorhabditis elegans development. *Proceedings of the National Academy of Sciences of the United States of America*, 92(21):9839–9842. 82, 119, 138

Gordon, W. R., Roy, M., Vardar-Ulu, D., Garfinkel, M., Mansour, M. R., Aster, J. C., and 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. 24

Gordon, W. R., Vardar-Ulu, D., Histen, G., Sanchez-Irizarry, C., Aster, J. C., and Blacklow, S. C. (2007). Structural basis for autoinhibition of Notch. *Nature Structural & Molecular Biology*, 14(4):295–300. 5, 24

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., and Blacklow, S. C. (2015). Mechanical Allostery: Evidence for a Force Requirement in the Proteolytic Activation of Notch. *Developmental cell*, 33(6):729–36. 5

Gray, A., Dull, T. J., and Ullrich, A. (1983). Nucleotide sequence of epidermal growth factor cDNA predicts a 128,000-molecular weight protein precursor. *Nature*, 303(5919):722–725. 17

Greenwald, I. (1985). lin-12, a nematode homeotic gene, is homologous to a set of mammalian proteins that includes epidermal growth factor. *Cell*, 43(3 Pt 2):583–590. 2, 106

Greenwald, I. and Seydoux, G. (1990). Analysis of gain-of-function mutations of the lin-12 gene of Caenorhabditis elegans. *Nature*, 346(6280):197–199. 24

Greenwald, I. S., Sternberg, P. W., and Horvitz, H. R. (1983). The lin-12 locus specifies cell fates in Caenorhabditis elegans. *Cell*, 34(2):435–444. 106, 115, 126

Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends in biotechnology*, 22(7):346–353. 94

Hagedorn, E. J., Yashiro, H., Ziel, J. W., Ihara, S., Wang, Z., and Sherwood, D. R. (2009). Integrin Acts Upstream of Netrin Signaling to Regulate Formation of the Anchor Cell's Invasive Membrane in C. elegans. *Developmental Cell*, 17(2):187–198. 153

Hambleton, S., Valeyev, N. V., Muranyi, A., Knott, V., Werner, J. M., McMichael, A. J., Handford, P. A., and Downing, A. K. (2004). Structural and Functional Properties of the Human Notch-1 Ligand Binding Region. *Structure*, 12:2173–2183. 17, 18, 80

Hampton, R. Y. (2002). ER-associated degradation in protein quality control and cellular regulation. *Current opinion in cell biology*, 14(4):476–82. 166

Harris, R. J. and Spellman, M. W. (1993). O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology*, 3(3):219–224. 21

Henderson, S. T., Gao, D., Christensen, S., and 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. 29, 33, 156, 167

Henderson, S. T., Gao, D., Lambie, E. J., and Kimble, J. (1994). lag-2 may encode a signaling ligand for the GLP-1 and LIN-12 receptors of C. elegans. *Development*, 120(10):2913–2924. 107, 108

Heuss, S. F., Ndiaye-Lobry, D., Six, E. M., Israël, A., and Logeat, F. (2008). The intracellular region of Notch ligands Dll1 and Dll3 regulates their trafficking and signaling activity. *Proceedings of the National Academy of Sciences of the United States of America*, 105(32):11212–7. 127

Horvitz, H. R., Sternberg, P. W., Greenwald, I. S., Fixsen, W., and Ellis, H. M. (1983). Mutations that affect neural cell lineages and cell fates during the development of the nematode Caenorhabditis elegans. *Cold Spring Harbor symposia on quantitative biology*, 48 Pt 2:453–463. 106

Hubbard, E. J., Wu, G., Kitajewski, J., and Greenwald, I. (1997). sel-10, a negative regulator of lin-12 activity in Caenorhabditis elegans, encodes a member of the CDC4 family of proteins. *Genes & Development*, 11(23):3182–3193. 26

Hutterer, A. and Knoblich, J. A. (2005). Numb and alpha-Adaptin regulate Sanpodo endocytosis to specify cell fate in Drosophila external sensory organs. *EMBO reports*, 6(9):836–42. 3

Ilagan, M. X. G., Lim, S., Fulbright, M., Piwnica-Worms, D., and Kopan, R. (2011). Real-Time Imaging of Notch Activation with a Luciferase Complementation-Based Reporter. *Science Signaling*, 4(181). 36

Jaroszewski, L., Rychlewski, L., Li, Z., Li, W., and Godzik, A. (2005). FFAS03: a server for profile–profile sequence alignments. *Nucleic acids research*, 33(Web Server issue):W284–8. 84

Jarriault, S., Brou, C., Logeat, F., Schroeter, E. H., Kopan, R., and Israel, A. (1995). Signalling downstream of activated mammalian Notch. *Nature*, 377(6547):355–358. 7

Johnston, S., Rauskolb, C., Wilson, R., Prabhakaran, B., Irvine, K., and Vogt, T. (1997). A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development*, 124(11):2245–2254. 22

Joutel, A., Corpechot, C., Ducros, A., Vahedi, K., Chabriat, H., Mouton, P., Alamowitch, S., Domenga, V., Cecillion, M., Marechal, E., Maciazek, J., Vayssiere, C., Cruaud, C., Cabanis, E.-A., Ruchoux, M. M., Weissenbach, J., Bach, J. F., Bousser, M. G., and Tournier-Lasserve, E. (1996). Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature*, 383(6602):707–710. 11, 14

Joutel, A., Monet, M., Domenga, V., Riant, F., and Tournier-Lasserve, E. (2004). Pathogenic mutations associated with cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy differently affect Jagged1 binding and Notch3 activity via the RBP/JK signaling Pathway. *American journal of human genetics*, 74(2):338–47. 14

Joutel, A. and Tournier-Lasserve, E. (2002). [Molecular basis and physiopathogenic mechanisms of CADASIL: a model of small vessel diseases of the brain]. *Journal de la Societe de biologie*, 196(1):109–15. 14

Kamath, B. M., Bauer, R. C., Loomes, K. M., Chao, G., Gerfen, J., Hutchinson, A., Hardikar, W., Hirschfield, G., Jara, P., Krantz, I. D., Lapunzina, P., Leonard, L., Ling, S., Ng, V. L., Hoang, P. L., Piccoli, D. A., and Spinner, N. B. (2011). NOTCH2 mutations in Alagille syndrome. *Journal of Medical Genetics*. 13

Karp, X. and Greenwald, I. (2003). Post-transcriptional regulation of the E/Daughterless ortholog HLH-2, negative feedback, and birth order bias during the AC/VU decision in C. elegans. *Genes & development*, 17(24):3100–11. 147

Karp, X. and Greenwald, I. (2004). Multiple roles for the E/Daughterless ortholog HLH-2 during C. elegans gonadogenesis. *Developmental biology*, 272(2):460–9. 147

Kershaw, N. J., Church, N. L., Griffin, M. D. W., Luo, C. S., Adams, T. E., and Burgess, A. W. (2015). Notch ligand delta-like1: X-ray crystal structure and binding affinity. *The Biochemical Journal*, 468(1):159–166. 9, 36, 84

Kershner, A. M., Shin, H., Hansen, T. J., and Kimble, J. (2014). Discovery of two GLP-1/Notch target genes that account for the role of GLP-1/Notch signaling in stem cell maintenance. *Proceedings of the National Academy of Sciences*. 114

Kohsaka, T., Yuan, Z., Guo, S., Tagawa, M., Nakamura, A., Nakano, M., Kawasasaki, H., Inomata, Y., Tanaka, K., and Miyauchi, J. (2002). The significance of human jagged 1 mutations detected in severe cases of extrahepatic biliary atresia. *Hepatology*, 36(4):904–912. 12

Kopan, R., Schroeter, E. H., Weintraub, H., and Nye, J. S. (1996). Signal transduction by activated mNotch: importance of proteolytic processing and its regulation by the extracellular domain. *Proceedings of the National Academy of Sciences*, 93(4):1683–1688. 23

Krantz, I. D., Smith, R., Colliton, R. P., Tinkel, H., Zackai, E. H., Piccoli, D. A., Goldmuntz, E., and Spinner, N. B. (1999). Jagged1 mutations in patients ascertained with isolated congenital heart defects. *American Journal of Medical Genetics*, 84(1):56–60. 12

Lambie, E. J. and Kimble, J. (1991). Two homologous regulatory genes, lin-12 and glp-1, have overlapping functions. *Development (Cambridge, England)*, 112(1):231–240. 79, 81, 106, 107, 108, 115, 138

LeBon, L., Lee, T. V., Sprinzak, D., Jafar-Nejad, H., and Elowitz, M. B. (2014). Fringe proteins modulate Notch-ligand cis and trans interactions to specify signaling states. *eLife*, 3:e02950. 127

Lemmon, M. A. (2008). Membrane recognition by phospholipid-binding domains. *Nature Reviews Molecular Cell Biology*, 9(2):99–111. 34

Li, L., Krantz, I. D., Deng, Y., Genin, A., Banta, A. B., Collins, C. C., Qi, M., Trask, B. J., Kuo, W. L., Cochran, J., Costa, T., Pierpont, M. E. M., Rand, E. B., Piccoli, D. A., Hood, L., and Spinner, N. B. (1997). Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nature Genetics*, 16(3):243–251. 11

Logeat, F., Bessia, C., Brou, C., LeBail, O., Jarriault, S., Seidah, N. G., and 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. 5

Lu, F., Jennifer Morrissette, J. D., and Spinner, N. B. (2003). Conditional JAG1 Mutation Shows the Developing Heart Is More Sensitive Than Developing Liver to JAG1 Dosage. *The American Journal of Human Genetics*, 72(4):1065–1070. 13

Luca, V. C., Jude, K. M., Pierce, N. W., Nachury, M. V., Fischer, S., and Garcia, K. C. (2015). Structural basis for Notch1 engagement of Delta-like 4. *Science*, 347(6224):847–853. 9, 37, 79, 84, 88, 167

Maduro, M. and Pilgrim, D. (1995). Identification and Cloning of Unc-119, a Gene Expressed in the Caenorhabditis Elegans Nervous System. *Genetics*, 141(3):977–988. 67

Mango, S. E., Thorpe, C. J., Martin, P. R., Chamberlain, S. H., and Bowerman, B. (1994). Two maternal genes, apx-1 and pie-1, are required to distinguish the fates of equivalent blastomeres in the early Caenorhabditis elegans embryo. *Development (Cambridge, England)*, 120(8):2305–2315. 107

McCright, B., Lozier, J., and Gridley, T. (2002). A mouse model of Alagille syndrome: Notch2 as a genetic modifier of Jag1 haploinsufficiency. *Development*, 129(4):1075–1082. 13

McDaniell, R., Warthen, D. M., Sanchez-Lara, P. A., Pai, A., Krantz, I. D., Piccoli, D. A., and Spinner, N. B. (2006). NOTCH2 mutations cause Alagille syndrome, a heterogeneous disorder of the notch signaling pathway. *American Journal of Human Genetics*, 79(1):169–173. 13

McKellar, S. H., Tester, D. J., Yagubyan, M., Majumdar, R., Ackerman, M. J., and Sundt, T. M. (2007). Novel NOTCH1 mutations in patients with bicuspid aortic valve disease and thoracic aortic aneurysms. *The Journal of thoracic and cardiovascular surgery*, 134(2):290–6. 9

Meester, J. A. N., Southgate, L., Stittrich, A.-B., Venselaar, H., Beekmans, S. J. A., den Hollander, N., Bijlsma, E. K., Helderman-van den Enden, A., Verheij, J. B. G. M., Glusman, G., Roach, J. C., Lehman, A., Patel, M. S., de Vries, B. B. A., Ruivenkamp, C., Itin, P., Prescott, K., Clarke, S., Trembath, R., Zenker, M., Sukalo, M., Van Laer, L., Loeys, B., and Wuyts, W. (2015). Heterozygous Loss-of-Function Mutations in DLL4 Cause Adams-Oliver Syndrome. *American journal of human genetics*, 97(3):475–82. 11

Mello, C. C., Draper, B. W., and Prless, J. R. (1994). The maternal genes apx-1 and glp-1 and establishment of dorsal-ventral polarity in the early C. elegans embryo. *Cell*, 77(1):95–106. 107

Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *The EMBO Journal*, 10(12):3959–3970. 67, 127, 130

Moloney, D. J., Shair, L. H., Lu, F. M., Xia, J., Locke, R., Matta, K. L., and Haltiwanger, R. S. (2000). Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *The Journal of biological chemistry*, 275(13):9604–9611. 20, 21

Monet-Leprêtre, M., Bardot, B., Lemaire, B., Domenga, V., Godin, O., Dichgans, M., Tournier-Lasserve, E., Cohen-Tannoudji, M., Chabriat, H., and Joutel, A. (2009). Distinct phenotypic and functional features of CADASIL mutations in the Notch3 ligand binding domain. *Brain*, 132(6):1601–1612. 14

Moremen, K. W. and Molinari, M. (2006). N-linked glycan recognition and processing: the molecular basis of endoplasmic reticulum quality control. *Current Opinion in Structural Biology*, 16(5):592–599. 101

Morgan, T. and Bridges, C. (1916). Sex-linked inheritance in Drosophila. *Carnegie Institute of Washington Publication*, 237:1–88. 1

Morrissette, J. J. D., Colliton, R. P., and Spinner, N. B. (2001). Defective intracellular transport and processing of JAG1 missense mutations in Alagille syndrome. *Human Molecular Genetics*, 10(4):405–413. 13

Nam, Y., Sliz, P., Pear, W. S., Aster, J. C., and Blacklow, S. C. (2007). Cooperative assembly of higher-order Notch complexes functions as a switch to induce transcription. *Proceedings of the National Academy of Sciences*, 104(7):2103–2108. 26

Nam, Y., Sliz, P., Song, L., Aster, J. C., and Blacklow, S. C. (2006). Structural Basis for Cooperativity in Recruitment of MAML Coactivators to Notch Transcription Complexes. *Cell*, 124(5):973–983. 26

Neves, A. and Priess, J. R. (2005). The REF-1 family of bHLH transcription factors pattern C. elegans embryos through Notch-dependent and Notch-independent pathways. *Developmental cell*, 8(6):867–79. 138

Ntziachristos, P., Lim, J. S., Sage, J., and Aifantis, I. (2014). From fly wings to targeted cancer therapies: a centennial for notch signaling. *Cancer Cell*, 25(3):318–334. 16

Oberg, C., Li, J., Pauley, A., Wolf, E., Gurney, M., and Lendahl, U. (2001). The Notch intracellular domain is ubiquitinated and negatively regulated by the mammalian Sel-10 homolog. *The Journal of Biological Chemistry*, 276(38):35847–35853. 28

Okajima, T., Xu, A., Lei, L., and Irvine, K. D. (2005). Chaperone activity of protein O-fucosyltransferase 1 promotes notch receptor folding. *Science (New York, N.Y.)*, 307(5715):1599–1603. 21

Okkema, P. G., Harrison, S. W., Plunger, V., Aryana, A., and Fire, A. (1993). Sequence requirements for myosin gene expression and regulation in Caenorhabditis elegans. *Genetics*, 135(2):385–404. 137

Olsen, J. V., Ong, S.-E., and Mann, M. (2004). Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & cellular proteomics : MCP*, 3(6):608–14. 100

O'Neil, J., Grim, J., Strack, P., Rao, S., Tibbitts, D., Winter, C., Hardwick, J., Welcker, M., Meijerink, J. P., Pieters, R., Draetta, G., Sears, R., Clurman, B. E., and Look, A. T. (2007). FBW7 mutations in leukemic cells mediate NOTCH pathway activation and resistance to gamma-secretase inhibitors. *The Journal of Experimental Medicine*, 204(8):1813–1824. 15

Oswald, F., Täuber, B., Dobner, T., Bourteele, S., Kostezka, U., Adler, G., Liptay, S., and Schmid, R. M. (2001). p300 Acts as a Transcriptional Coactivator for Mammalian Notch-1. *Molecular and Cellular Biology*, 21(22):7761–7774. 7

Paix, A., Folkmann, A., Rasoloson, D., and Seydoux, G. (2015). High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR/Cas9 Ribonucleoprotein Complexes. *Genetics*, 201(1):47–54. 176

Panin, V. M., Papayannopoulos, V., Wilson, R., and Irvine, K. D. (1997). Fringe modulates Notch-ligand interactions. *Nature*, 387(6636):908–912. 3, 22, 30

Park, S. K., Choi, V. N., and Hwang, B. J. (2013). LIN-12/Notch Regulates lag-1 and lin-12 Expression during Anchor Cell/Ventral Uterine Precursor Cell Fate Specification. *Molecules and Cells*, 35(3):249–254. 110

Parks, A. L., Klueg, K. M., Stout, J. R., and Muskavitch, M. A. (2000). Ligand endocytosis drives receptor dissociation and activation in the Notch pathway. *Development (Cambridge, England)*, 127(7):1373–1385. 5, 24, 127

Parks, A. L., Stout, J. R., Shepard, S. B., Klueg, K. M., Dos Santos, A. A., Parody, T. R., Vaskova, M., and Muskavitch, M. A. T. (2006). Structure-function analysis of delta trafficking, receptor binding and signaling in Drosophila. *Genetics*, 174(4):1947–1961. 33

Pei, Z. and Baker, N. E. (2008). Competition between Delta and the Abruptex domain of Notch. *BMC Developmental Biology*, 8(Figure 1). 19

Petersen, T. N., Brunak, S. r., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10):785–786. 48

Poulson, D. (1939). Effects of Notch deficiencies. *Drosophila Information Service*, 12:64. 1

Rampal, R., Li, A. S. Y., Moloney, D. J., Georgiou, S. A., Luther, K. B., Nita-Lazar, A., and Haltiwanger, R. S. (2005). Lunatic fringe, manic fringe, and radical fringe recognize similar specificity determinants in O-fucosylated epidermal growth factor-like repeats. *The Journal of Biological Chemistry*, 280(51):42454–42463. 22

Rana, N. A., Nita-Lazar, A., Takeuchi, H., Kakuda, S., Luther, K. B., and Haltiwanger, R. S. (2011). O-glucose trisaccharide is present at high but variable stoichiometry at multiple sites on mouse Notch1. *The Journal of biological chemistry*, 286(36):31623–37. 20

Rebay, I., Fehon, R. G., and Artavanis-Tsakonas, S. (1993). Specific truncations of Drosophila Notch define dominant activated and dominant negative forms of the receptor. *Cell*, 74(2):319–329. 23

Rebay, I., Fleming, R. J., Fehon, R. G., Cherbas, L., Cherbas, P., and 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. 17

Riddle, D. L., Blumenthal, T., Meyer, B. J., and Priess, J. R. (1997). The Anchor Cell. 111

Roehl, H., Bosenberg, M., Blelloch, R., and Kimble, J. (1996). Roles of the RAM and ANK domains in signaling by the C. elegans GLP-1 receptor. *The EMBO Journal*, 15(24):7002–7012. 25, 107

Sanchez-Irizarry, C., Carpenter, A. C., Weng, A. P., Pear, W. S., Aster, J. C., and 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. 23

Schrijver, I., Liu, W., Brenn, T., Furthmayr, H., and Francke, U. (1999). Cysteine Substitutions in Epidermal Growth Factor-Like Domains of Fibrillin-1: Distinct Effects on Biochemical and Clinical Phenotypes. *The American Journal of Human Genetics*, 65(4):1007–1020. 14

Seydoux, G. and Greenwald, I. (1989). Cell autonomy of lin-12 function in a cell fate decision in C. elegans. *Cell*, 57(7):1237–45. 110

Sharma, A., Rangarajan, A., and Dighe, R. (2013). Antibodies against the extracellular domain of human Notch1 receptor reveal the critical role of epidermal-growth-factor-like repeats 2526in ligand binding and receptor activation. *Biochemical Journal*, 449(2):519–530. 19

Shaye, D. D. and Greenwald, I. (2002). Endocytosis-mediated downregulation of LIN-12/Notch upon Ras activation in Caenorhabditis elegans. *Nature*, 420(6916):686–690. 112

Shimizu, K., Chiba, S., Kumano, K., Hosoya, N., Takahashi, T., Kanda, Y., Hamada, Y., Yazaki, Y., and Hirai, H. (1999). Mouse Jagged1 Physically Interacts with Notch2 and Other Notch Receptors. *Journal of Biological Chemistry*, 274(46):32961–32969. 29

Shimizu, K., Chiba, S., Saito, T., Kumano, K., and Hirai, H. (2000). Physical interaction of Delta1, Jagged1, and Jagged2 with Notch1 and Notch3 receptors. *Biochemical and Biophysical Research Communications*, 276(1):385–389. 5, 9, 167

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., Pauli, R. M., Aftimos, S., Stewart, H., Kim, C. A., Holder-Espinasse, M., Robertson, S. P., Drake, W. M., and Trembath, R. C. (2011). Mutations in NOTCH2 cause Hajdu-Cheney syndrome, a disorder of severe and progressive bone loss. *Nature Genetics*, 43(4):303–305. 11

Stephenson, N. L. and 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 of the United States of America*, 109(41):E2757–65. 5

Sternberg, P. W. and Horvitz, H. R. (1989). The combined action of two intercellular signaling pathways specifies three cell fates during vulval induction in C. elegans. *Cell*, 58(4):679–93. 115

Stinchcomb, D. T., Shaw, J. E., Carr, S. H., and Hirsh, D. (1985). Extrachromosomal DNA transformation of Caenorhabditis elegans. *Molecular and Cellular Biology*, 5(12):3484–3496. 126, 127

Struhl, G. and Adachi, A. (1998). Nuclear access and action of notch in vivo. *Cell*, 93(4):649–660. 7

Sulston, J. E. and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental biology*, 56(1):110–56. 105

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental biology*, 100(1):64–119. 105

Takeuchi, H., Fernandez-Valdivia, R. C., Caswell, D. S., Nita-Lazar, A., Rana, N. A., Garner, T. P., Weldeghiorghis, T. K., Macnaughtan, M. A., Jafar-Nejad, H., and Haltiwanger, R. S. (2011). Rumi functions as both a protein O-glucosyltransferase and a protein O-xylosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 108(40):16600–16605. 20

Takeuchi, H. and Haltiwanger, R. S. (2014). Significance of glycosylation in Notch signaling. *Biochemical and Biophysical Research Communications*. 20

Tamura, K., Taniguchi, Y., Minoguchi, S., Sakai, T., Tun, T., Furukawa, T., and Honjo, T. (1995). Physical interaction between a novel domain of the receptor Notch and the transcription factor RBP-J kappa/Su(H). *Current biology: CB*, 5(12):1416–1423. 7, 25

Tax, F. E., Yeargers, J. J., and Thomas, J. H. (1994). Sequence of C. elegans lag-2 reveals a cell-signalling domain shared with Delta and Serrate of Drosophila. *Nature*, 368(6467):150–154. 28, 107

Taylor, P., Takeuchi, H., Sheppard, D., Chillakuri, C., Lea, S. M., Haltiwanger, R. S., and 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*, 111(20):7290–7295. 22, 39

Thomas, U., Speicher, S. A., and Knust, E. (1991). The Drosophila gene Serrate encodes an EGF-like transmembrane protein with a complex expression pattern in embryos and wing discs. *Development (Cambridge, England)*, 111(3):749–761. 28

Turnpenny, P. D., Whittock, N., Duncan, J., Dunwoodie, S., Kusumi, K., and Ellard, S. (2003). Novel mutations in DLL3, a somitogenesis gene encoding a ligand for the Notch signalling pathway, cause a consistent pattern of abnormal vertebral segmentation in spondylocostal dysostosis. *Journal of Medical Genetics*, 40(5):333–339. 11

van Es, J. H., van Gijn, M. E., Riccio, O., van den Born, M., Vooijs, M., Begthel, H., Cozijnsen, M., Robine, S., Winton, D. J., Radtke, F., and Clevers, H. (2005). Notch/gamma-secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature*, 435(7044):959–963. 16

van Tetering, G., van Diest, P., Verlaan, I., van der Wall, E., Kopan, R., and Vooijs, M. (2009). Metalloprotease ADAM10 Is Required for Notch1 Site 2 Cleavage. *The Journal of Biological Chemistry*, 284(45):31018–31027. 5

Vardar, D., North, C. L., Sanchez-Irizarry, C., Aster, J. C., and Blacklow, S. C. (2003). Nuclear magnetic resonance structure of a prototype Lin12-Notch repeat module from human Notch1. *Biochemistry*, 42(23):7061–7067. 23, 26

Wang, M. M. (2011). Notch signaling and Notch signaling modifiers. *The international journal of biochemistry & cell biology*, 43(11):1550–1562. 16

Wang, Y., Shao, L., Shi, S., Harris, R. J., Spellman, M. W., Stanley, P., and Haltiwanger, R. S. (2001). Modification of epidermal growth factor-like repeats with O-fucose. Molecular cloning and expression of a novel GDP-fucose protein O-fucosyltransferase. *The Journal of Biological Chemistry*, 276(43):40338–40345. 21

Weinmaster, G., Roberts, V. J., and Lemke, G. (1991). A homolog of Drosophila Notch expressed during mammalian development. *Development (Cambridge, England)*, 113(1):199–205. 2

Weng, A. P., Ferrando, A. A., Lee, W., Morris, J. P., Silverman, L. B., Sanchez-Irizarry, C., Blacklow, S. C., Look, A. T., and Aster, J. C. (2004). Activating Mutations of NOTCH1 in Human T Cell Acute Lymphoblastic Leukemia. *Science*, 306(5694):269–271. 15

Wharton, K. A., Johansen, K. M., Xu, T., and Artavanis-Tsakonas, S. (1985). Nucleotide sequence from the neurogenic locus notch implies a gene product that shares homology with proteins containing EGF-like repeats. *Cell*, 43(3 Pt 2):567–581. 1

Whiteman, P., Hernandez de Madrid, B., Taylor, P., Li, D., Heslop, R., Viticheep, N., Tan, J. Z., Shimizu, H., Callaghan, J., Masiero, M., Li, J. L., Banham, A. H., Harris, A. L., Lea, S. M., Redfield, C., Baron, M., and Handford, P. A. (2013). Molecular basis for Jagged-1/Serrate recognition by Notch. *The Journal of biological chemistry*. 19, 38, 39

Wilkinson, H. A., Fitzgerald, K., and Greenwald, I. (1994). Reciprocal changes in expression of the receptor lin-12 and its ligand lag-2 prior to commitment in a C. elegans cell fate decision. *Cell*, 79(7):1187–1198. 3, 110

Wilson, J. J. and Kovall, R. A. (2006). Crystal structure of the CSL-Notch-Mastermind ternary complex bound to DNA. *Cell*, 124(5):985–996. 26

Wong, G. T., Manfra, D., Poulet, F. M., Zhang, Q., Josien, H., Bara, T., Engstrom, L., Pinzon-Ortiz, M., Fine, J. S., Lee, H.-J. J., Zhang, L., Higgins, G. A., and Parker, E. M. (2004). Chronic Treatment with the γ-Secretase Inhibitor LY-411,575 Inhibits β-Amyloid Peptide Production and Alters Lymphopoiesis and Intestinal Cell Differentiation. *Journal of Biological Chemistry*, 279(13):12876–12882. 16

Wu, L., Aster, J. C., Blacklow, S. C., Lake, R., Artavanis-Tsakonas, S., and Griffin, J. D. (2000). MAML1, a human homologue of Drosophila mastermind, is a transcriptional co-activator for NOTCH receptors. *Nature Genetics*, 26(4):484–489. 26

Wu, Y., Cain-Hom, C., Choy, L., Hagenbeek, T. J., de Leon, G. P., Chen, Y., Finkle, D., Venook, R., Wu, X., Ridgway, J., Schahin-Reed, D., Dow, G. J., Shelton, A., Stawicki, S., Watts, R. J., Zhang, J., Choy, R., Howard, P., Kadyk, L., Yan, M., Zha, J., Callahan, C. A., Hymowitz, S. G., and Siebel, C. W. (2010). Therapeutic antibody targeting of individual Notch receptors. *Nature*, 464(7291):1052–1057. 16

Xu, D., Jaroszewski, L., Li, Z., and Godzik, A. (2014). FFAS-3D: improving fold recognition by including optimized structural features and template re-ranking. *Bioinformatics (Oxford, England)*, 30(5):660–7. 84

Xue, Y., Gao, X., Lindsell, C. E., Norton, C. R., Chang, B., Hicks, C., Gendron-Maguire, M., Rand, E. B., Weinmaster, G., and Gridley, T. (1999). Embryonic lethality and vascular defects in mice lacking the Notch ligand Jagged1. *Human Molecular Genetics*, 8(5):723–730. 13

Yochem, J. and Greenwald, I. (1989). glp-1 and lin-12, genes implicated in distinct cell-cell interactions in C. elegans, encode similar transmembrane proteins. *Cell*, 58(3):553–563. 106

Yochem, J., Weston, K., and Greenwald, I. (1988). The Caenorhabditis elegans lin-12 gene encodes a transmembrane protein with overall similarity to Drosophila Notch. *Nature*, 335(6190):547–550. 106

Yoo, A. S., Bais, C., and Greenwald, I. (2004). Crosstalk between the EGFR and LIN-12/Notch pathways in C. elegans vulval development. *Science (New York, N.Y.)*, 303(5658):663–6. 112

Zhang, H., Constantine, R., Vorobiev, S., Chen, Y., Seetharaman, J., Huang, Y. J., Xiao, R., Montelione, G. T., Gerstner, C. D., Davis, M. W., Inana, G., Whitby, F. G., Jorgensen, E. M., Hill, C. P., Tong, L., and Baehr, W. (2011). UNC119 is required for G protein trafficking in sensory neurons. *Nature Neuroscience*, 14(7):874–880. 137, 151

Zhang, W., Cao, P., Chen, S., Spence, A. M., Zhu, S., Staudacher, E., and Schachter, H. (2003). Synthesis of paucimannose N-glycans by Caenorhabditis elegans requires prior actions of UDP-N-acetyl-D-glucosamine:alpha-3-D-mannoside beta1,2-N-acetylglucosaminyltransferase I, alpha3,6-mannosidase II and a specific membrane-bound beta-N-acetylglucosamin. *The Biochemical journal*, 372(Pt 1):53–64. 102