

STRUCTURAL, BIOPHYSICAL AND CELLULAR ANALYSIS OF SYNAPTIC RECEPTORS AND ORGANIZERS

Amber Jayne Clayton

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy



Oriel College

Nuffield Department of Clinical Medicine

Division of Structural Biology

University of Oxford

Trinity Term 2013

Structural, Biophysical and Cellular Analysis of Synaptic Receptors and Organizers

Amber Jayne Clayton

A thesis submitted in partial fulfilment of the requirements for the degree of *Doctor of Philosophy*
Oriol College, Nuffield Department of Clinical Medicine and Division of Structural Biology
University of Oxford, Trinity Term 2013

Abstract

Synapse formation, differentiation and plasticity involve modulation of pre- and post-synaptic elements. During synaptogenesis synaptic compartments differentiate in response to molecular cues. Once contact is made, both dendritic and axonal cells undergo further morphological changes resulting in the formation of a synapse which then matures in response to activity. The α -amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid (AMPA) receptors are responsible for fast excitatory neurotransmission in the central nervous system. Here the AMPA receptor GluA2 amino terminal domain (ATD) structure was solved to 1.8Å resolution and revealed important insights into the domain's involvement in subunit specific assembly and the presence of a putative ligand binding site.

In a physiological context, the AMPA receptor spends little time alone. Shuttle proteins control its transport to the cell surface, intracellular scaffolding molecules synaptically anchor the receptor and auxiliary proteins modulate channel conductance but all interactions are structurally uncharacterised. This thesis aimed to give structural and biophysical insights into protein-protein interactions involving the AMPA receptor ectodomain. Biophysical analysis confirmed that neuronal pentraxins (NPs) interact with AMPA receptor ATDs (all GluA2, 3 & 4 ATDs and NP1, NP2 and NP receptor-pentraxin (PTX) domain combinations tested), where a preferential interaction occurred between GluA4_ATD and NPI_PTX. Stable homogenous oligomeric neuronal pentraxins were developed to investigate their physiological effects on AMPA receptors where receptor clustering had previously been observed. Cellular analysis indicated that AMPA receptors are internalised when stimulated with trimeric NP1. Additional AMPA receptor-interacting proteins were also investigated where biophysical analysis found that Brorin, a BMP antagonist, interacts with the LBC of GluA2. Attempts were also made to optimise the expression of the AMPA receptor auxiliary protein cystine-knot AMPA receptor modulating protein-44 (CKAMP44) and transmembrane AMPA regulatory protein-2 (TARP2) that are also postulated to interact with the AMPA receptor ectodomain but remain structurally uncharacterised.

Regulation of synapse formation and function can also be achieved through adhesive mechanisms. The recent identification of meprin/A5 protein/ receptor protein tyrosine phosphatase mu (MAM)-domain containing glycosylphosphatidylinositol (GPI)- anchor protein 1 (MDGA1) as a key modulator of the neuroligin-neurexin interaction in inhibitory synapse formation suggested its potential role as the first identified negative regulator of inhibition. MDGA1 was structurally uncharacterised but here the structure of the secreted MDGA1 ectodomain was solved to 3.4Å and presented a highly unusual interlocked dimer in a 'tubular bells' formation. Biophysical analysis determined that MDGA1 interacts with both neuroligin1, usually post-synaptically expressed at glutamatergic synapses, and neuroligin-2, expressed at γ -aminobutyric acid (GABA)-ergic synapses. MDGA1 may therefore be a key regulator of both excitatory and inhibitory synapse formation and function.

Acknowledgments

I first want to acknowledge my supervisor Radu Aricescu. His unfaltering sense of optimism amongst the constant chaos certainly rubs off and it has been an honour and a privilege to have worked for him. I will never forget his inadvertently humorous 'Radu-isms' including 'I'll be back in a picosecond' that has given a whole new meaning to being late.

I also want to thank my many STRUBI friends past and present especially my Yorkshire lab buddy Charlotte Coles for establishing my running-bug, for agreeing that there is no R in bath and proving that if you push yourself hard enough you can achieve anything. Thanks also to Benjamin Bishop, Ross Robinson, Ellie Healey, Sam Griffiths and Jordan Clay, for expanding my vocabulary- 'excellent badinage from the cruciverbalists' and also not forgetting Aletheia Lee, Wun Kwok, Paul Miller, Tomas Malinauskas, Luke Yates, Lexy Smith and all of the new friendly faces that have made STRUBI such an enjoyable place to work.

Thanks to Yvonne Jones for support and encouragement when my turn to present at lab meetings inevitably came around again and to Margaret Jones who always goes above and beyond to keep the STRUBI wet lab going. Thanks to Weixian Lu for teaching me mammalian expression techniques and to Yuguang Zhao for always providing help with the tissue culture robot. Many thanks also to Tom Walter and Karl Harlos for providing crystallization expertise. Thanks to Christian Siebold for his guidance in structural refinement and analysis and for setting the bar for beautiful figure making. Thanks to my newly adoptive supervisor Geoff Sutton, and Charlotte Coles, Benjamin Bishop and Jack Alden for extensively proofreading this thesis and providing constant encouragement throughout my time in STRUBI. Thanks to Veronica Chang for all of her help and expert advice about FACS analysis. I am also very grateful to the MRC for funding the majority of this work and to Oriel College for their support.

I also want to thank my mama and papa for believing in me and providing a homely welcoming place to escape to for respite, also not forgetting my sister and brother, Abby and Jack, for telling people that I make agar plates and orange juice for a living.

Finally I want to say a huge thank you to my fiancé Jack. Thank you for putting up with me working late, working weekends and working night shifts at the synchrotron, I really couldn't have done this without you, thank you so much. Now on to the next adventure....

"If you can't explain it simply, you don't understand it well enough."

Albert Einstein

Contents	Page
Abstract	i
Acknowledgements	ii
Contents	iii-vii
Abbreviations	viii-xiii
Chapter 1 Introduction	1-43
1.1 General overview.....	2-9
1.1.1 The synapse	2
1.1.2 Synaptic components	3
1.1.3 Types of synapses	3-4
1.1.4 Neurotransmitters	4-5
1.1.5 Long term potentiation (LTP)	5-6
1.1.6 AMPA receptor trafficking in LTP	7-8
1.1.7 Long term depression (LTD)	8
1.1.8 Negative effects of synaptic alterations	8-9
1.2 The iGluRs	10-12
1.2.1 Overview	10
1.2.2 Origins of the iGluRs	10-12
1.2.3 The iGluRs in the developing brain	12
1.3 The AMPA receptor.....	13-18
1.3.1 Overview	13
1.3.2 AMPA receptor subunit organisation	13-16
1.3.3 Structural information timeline	16-18
1.3.4 Examples of structural information aiding medical intervention	18-20
1.4 AMPA receptor interacting proteins	20-21
1.4.1 An overview	20-21
1.5 The neuronal pentraxins.....	21-26
1.5.1 The pentraxin family	21-22
1.5.2 Neuronal pentraxins	22-23
1.5.3 Interactions of AMPA receptors with neuronal pentraxins	24
1.5.4 Diseases/disorders implicating the neuronal pentraxins	24-25
1.5.5 Neuronal pentraxins at parvalbumin-expressing interneurons (PV-INs)	25-26
1.5.6 The neuronal pentraxins in learning and memory	26
1.6 Other AMPA receptor interactions.....	27-29
1.6.1 Cystine-knot AMPA receptor modulating protein 44 (CKAMP44)	27
1.6.2 Transmembrane AMPA regulatory protein-2 (TARP2)	27-28
1.6.2.1 TARP involvement in AMPA receptor trafficking	28
1.6.2.2 The effect of TARPs on AMPA receptor gating	28-29
1.6.3 Analysis of putative AMPA receptor interaction partners	29
1.7 MDGA.....	30-33
1.7.1 MDGA1 in development	30-31

1.7.2	The neurexin- neuroligin interaction	31-32
1.7.3	Neuroligins and neurexins in disease	32
1.7.4	Interactions of MDGA1 with neuroligins	33
1.8	Research aims.....	34
1.9	References.....	35-43

Chapter 2 General materials and methods.....44-75

2.1	Molecular biology.....	45-53
2.1.1	cDNA sources	45-46
2.1.2	Construct design	46-48
2.1.3	Polymerase chain reaction (PCR)	48
2.1.4	Splice overlap extension (SOE)-PCR	48-49
2.1.5	Expression vectors for mammalian expression	49-50
2.1.6	Generation of pHLsec_monoVenus vector	49-50
2.1.7	Additional protein tags	51
2.1.8	Restriction digest/cloning	51-52
2.1.9	Plasmid purification	52
2.1.10	Cloning for generation of stable cell lines	52-53
2.2	Protein expression in mammalian cell systems.....	53-56
2.2.1	Small-scale expression trials	53-54
2.2.2	Large-scale protein expression	54
2.2.3	<i>In vivo</i> selenomethionine-labelling	55
2.2.4	Generation of stable cell lines	55-56
2.2.5	Cell preparation for Immunocytochemistry (ICC)	56
2.3	Protein detection methods.....	57-58
2.3.1	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis	57
2.3.2	Western blot analysis	57-58
2.4	Protein purification from mammalian expression systems.....	59-60
2.4.1	Media harvest and dialysis	59
2.4.2	Immobilized metal ion affinity chromatography (IMAC)	59-60
2.4.3	Size exclusion chromatography (SEC)	60
2.4.4	Protein deglycosylation	60
2.5	Protein expression and purification from <i>E. coli</i>	61-64
2.5.1	Sequencing in pET vector	61
2.5.2	Transformation of Origami <i>E. coli</i> cells and protein expression	61-62
2.5.3	Protein purification from Origami cells	62
2.5.4	Transformation of Rosetta cells and protein expression in inclusion bodies	62-63
2.5.5	Protein extraction, refold and purification from inclusion bodies	63-64
2.6	Crystallization.....	65-66
2.7	Data collection.....	67
2.8	Biophysical analysis.....	68
2.8.1	Multi-angle light scattering (MALS) analysis	68-69
2.8.2	Surface plasmon resonance (SPR)	69-71
2.8.3	Analytical ultracentrifugation (AUC)	71
2.9	Fluorescence activated cell sorting (FACS).....	71-72

2.10	Electron microscopy.....	72
2.11	Data analysis and generation of figures.....	73
2.12	References.....	74-75

Chapter 3 Structural and functional analysis of the AMPA receptor GluA2 amino terminal domain 76-100

3.1	Introduction.....	77-78
3.2	Aims.....	79
3.3	Materials and methods.....	80-84
3.3.1	Construct design, PCR and restriction-based cloning overview	80
3.3.2	Generation of constructs	80-81
3.3.3	Large-scale protein expression and purification	82
3.3.4	Crystallization, data collection and processing	82-83
3.3.5	Structure determination and refinement	83-84
3.3.6	Thermal shift assay (thermofluor)	84
3.3.7	FACS analysis	84
3.4	Results.....	85-93
3.4.1	Structural analysis	85-87
3.4.2	Investigation of the putative small molecule binding site	88
3.4.3	Attempts to crystallize larger ectodomain constructs	89
3.4.4	Understanding the ATDs role in AMPA receptor assembly	89-92
3.4.5	Understanding the ATDs role in AMPA receptor trafficking	92-93
3.5	Discussion.....	94-97
3.6	Future work.....	98
3.7	References.....	99-100

Chapter 4 Structural and functional analysis of the Neuronal Pentraxins 101-141

4.1	Introduction.....	102-105
4.2	Aims.....	105
4.3	Materials and methods.....	106-112
4.3.1	Construct design, PCR and restriction-based cloning	106
4.3.2	Generation of neuronal pentraxins 1 constructs	106-108
4.3.3	Generation of neuronal pentraxins 2 constructs	108-109
4.3.4	Generation of neuronal pentraxins receptor constructs	109
4.3.5	Protein expression and purification	109
4.3.6	Crystallization, data collection and processing	110-111
4.3.7	Refinement and structural validation	111-112
4.3.8	Biophysical analysis	112
4.4	Results.....	113-135
4.4.1	Protein expression and purification	113-117
4.4.2	Structural analysis of neuronal pentraxin PTX domains	117-123
4.4.3	Mapping the AMPA receptor - neuronal pentraxin interaction site	124-126
4.4.4	Effect of NP1_PTX calcium binding on interactions with GluA4_ATD	126-127
4.4.5	Introduction of GluA4_ATD mutations to impede neuronal pentraxin	

binding	127-128
4.4.6 Effect of neuronal pentraxin multimerization on GluA4 ATD binding	128-131
4.4.7 Determining the effect of the AMPA receptor- neuronal pentraxin interaction	131-135
4.5 Discussion.....	136-137
4.6 Future work/ Future perspectives.....	138-139
4.7 References.....	140-141
 Chapter 5 Analysis of other AMPA receptor ectodomain-interacting proteins.....	 142-167
5.1 Introduction.....	143-144
5.2 Aims.....	145
5.3 Materials and methods.....	146-149
5.3.1 Construct design, PCR and cloning overview	146
5.3.2 Generation of the GluA2_LBC construct	146
5.3.3 TARP2 (Stargazin) construct optimisation	146-147
5.3.4 CKAMP44 construct design	147-148
5.3.5 Other putative interactions	148-149
5.4 Results.....	150-161
5.4.1 Expression of GluA2_LBC	150
5.4.2 TARP2_Trx expression/purification/ crystallization attempts	151-153
5.4.3 CKAMP44 expression and purification attempts	154-155
5.4.4 Biophysical analysis of GluA2_LBC and CKAMP44	156-157
5.4.5 Analysis of other putative AMPA receptor interactions	157-159
5.4.6 Brorin crystallization trials	159-161
5.5 Discussion.....	162-164
5.5.1 TARP2 analysis	162
5.5.2 CKAMP44 analysis	162-163
5.5.3 Analysis of other putative interactions	163-164
5.6 Future work and perspectives.....	165
5.6.1 TARP2	165
5.6.2 The AMPA receptor interactome	165
5.7 References.....	166-167
 Chapter 6 MDGA1.....	 168-195
6.1 Introduction.....	169-170
6.2 Aims.....	171
6.3 Materials and methods.....	172-178
6.3.1 Construct design, PCR and cloning	172-175
6.3.2 Protein expression and purification	175
6.3.3 Protein crystallization, data collection and processing	175-177
6.3.4 SPR analysis	177
6.3.5 SAXS analysis	178
6.4 Results.....	179-190
6.4.1 Structural analysis of MDGA1	179-184
6.4.2 Mapping the Neuroligin –MDGA1 interaction	185-187

6.4.3 MDGA1- NL2 complex crystallization attempts	188-189
6.4.4 SAX analysis	189-190
6.5 Discussion.....	191-192
6.6 Future work/ perspectives.....	193
6.7 References.....	194-195
Chapter 7 Further discussion.....	196-205
7.1 Summary of findings.....	197-198
7.2 Physiological insights gained from this research.....	198-201
7.3 Structural analysis to aid biophysical and functional experiments.....	202-203
7.4 Structural analysis providing insights for rational drug design.....	203
7.5 References.....	204-205
Appendices.....	206-247
A AMPA receptor constructs and oligonucleotides	207-209
B Neuronal pentraxin constructs and oligos	210-212
C Other AMPA receptor interacting proteins- constructs and oligos	213-222
D MDGA1 constructs and oligos	223
E monoVenus (mV) and pNeo-Sec vector map and sequences	224-228
F Published work Clayton <i>et al</i> , 2009 with supplementary material	229-242
G Trx scaffold generation	243-244
H Perineuronal nets (PNNs) constructs and oligos	245-247

Abbreviations

Å	Ångström
ALS	Amyotrophic lateral sclerosis
AMPA	α -amino-3-hydroxy-5-methyl-4-isoxazole-propionic acid
AP	Alkaline phosphate
AP-2	Adaptor protein complex 2
APBS	Adaptive Poisson-Boltzmann solver
ASD	Autism spectrum disorders
ASU	Asymmetric unit
ATD	Amino terminal domain
Brorin	Brain-specific chordin-like protein
BSA	Bovine serum albumin
BV	Bed volume
CAMKII	Calcium/calmodulin dependent protein kinase II
cAMP	Cyclic adenosine monophosphate
CC	Coiled-coil
CCP4	Collaborative computational project, number 4
cDNA	Complimentary deoxyribonucleic acid
CHO-K1	Chinese hamster ovary derived cell line
CKAMP44	Cytosine-knot AMPA receptor modulating protein of 44 kDa
CNS	Central nervous system
COS	CV-1 (simian) in Origin, carrying the SV40 genetic material
CREB	cAMP response element-binding protein
CRP	C-reactive protein
CSF	Cerebrospinal fluid
CSPGs	Chondroitin sulphate proteoglycans
CV	Column volume
°C	Degrees centigrade
dLGN	Dorsal lateral geniculate nucleus
DLS	Diamond light source
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribonucleic acid
dNTP	Dexoyribonucleate triphosphates

DRG neuron	Dorsal root ganglion neurons
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
ECL	Enhanced chemilluminescence
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
EndoF1	Endonuclease F1
EPSPs	Excitatory post-synaptic potentials
ESRF	European synchrotron radiation facility
ExPASy	Expert protein analysis system
FACS	Fluorescence-activated cell sorting
Fab	fragment antigen-binding
FBS	Foetal bovine serum
FL	Full length
FN3	Fibronectin type III
FPLC	Fast protein liquid chromatography
FRET	Förster resonance energy transfer
GABA	Gamma-aminobutyric acid
GFP	Green fluorescence protein
GlcNAc	N-acetylglucosamine
GnTI	N-acetylglucosaminyltransferase I
GPCR	G-protein-coupled receptor
GPI	Glycosylphosphatidylinositol
GRIP	Glutamate receptor interacting protein
GSH	Reduced glutathione
GSSG	Oxidised glutathione
GUI	Graphical user interface
HA	Haemagglutinin
HEK293T	Human embryonic kidney 293 cells expressing SV40 type T antigen
HEK293S	Human embryonic kidney 293 cells GnTI- deficient
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulphonic acid]
hr	Hour

HSPG	Heparan sulphate proteoglycan
ICC	Immunocytochemistry
IEG	Immediate early gene
Ig	Immunoglobulin
iGluR	Ionotropic glutamate receptor
IMAC	Immobilised metal-ion affinity chromatography
IPTG	Isopropylthio- β -galactoside
K	Kelvin
kDa	Kilodalton
KO	Knock-out
LB	Luria broth
LBC/LBD	Ligand binding core/ ligand binding domain
LIVBP	Leucine isoleucine valine binding protein
LLG	Log likelihood gain
LTD	Long term depression
LTP	Long term potentiation
MAD	Multi-wavelength anomalous dispersion
MAM	Meprin/A5-protein/PTPmu
MALS	Multi-angle light scattering
mCh	MonoCherry (fluorescent protein)
MCS	Multiple cloning site
MDCK	Madin-darby canine kidney epithelial cells
MDGA1	MAM domain-containing GPI anchor protein 1
mGluR	Metabotropic glutamate receptor
MIB	Malonic acid, Imidazole and Boric acid (crystallization screen)
MMP	Matrix metalloproteases
MOI	Multiplicity of infection
mRNA	Messenger ribonucleic acid
mV	MonoVenus (fluorescent protein)
MWCO	Molecular weight cut off
NEAA	Non-essential amino-acids
NCBI	National Center for Biotechnology Information
NEB	New England biolabs
NHS/ EDC	N-Hydroxysuccinimide/ ethyl(dimethylaminopropyl) carbodiimide

NMDA	n-methyl-d-aspartate
NP	Neuronal pentraxin
NP1	Neuronal pentraxin 1
NP2/NARP	Neuronal pentraxin 2/ Neuronal activity regulated protein
NPR	Neuronal pentraxin receptor
OD	Optical density
Oligo	Oligonucleotide
OMIM	Online Mendelian Inheritance in Man
O/N	Over night
OPPF	Oxford Protein Production Facility
PBS	Phosphate buffered saline
PCCD	Protein-complex crystallization database
PCR	Polymerase chain reaction
pdb	Protein data bank
PDZ	Domain present in PSD-95, Dlg, and ZO-1/2.
PEG	Polyethylene glycol
PEI	Polyethylenimine
PF-PN	Parallel fiber – Purkinje neuron
PFA	Paraformaldehyde
PG	Proteoglycans
pH	Power of hydrogen
pI	Isoelectric point
PI	Protease inhibitor
PICK	Protein interacting with C-kinase
PLL	Poly L-lysine
PMSF	Phenylmethylsulfonyl fluoride
PNN	Perineuronal net
PSD	Post synaptic density
PTX	Pentraxin domain
PV-INs	Parvalbumin-expressing interneurons
Rmsd	Root mean square deviation
RNA	Ribonucleic acid
RNAi	RNA interference
rpm	Revolutions per minute

RT	Room temperature
RU	Response unit
SAD	Single anomalous dispersion
SAP	Serum amyloid-P
SAXS	Small-angle X-ray scattering
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SE	Standard error
sec	Seconds
SEC	Size exclusion chromatography
SEC-MALS	Size exclusion chromatography-multi-angle light scattering
SeMet	Seleno-methionine
Sf9	<i>Spodoptera frugiperda</i> cell line 9
SFM	Serum free media
SHP	Structural homology program
shRNA	short hairpin RNA
SNP	Single nucleotide polymorphisms
SOE-PCR	Splice overlap extension polymerase chain reaction
SPR	Surface plasmon resonance
SS	Signal sequence
STRUBI	Division of Structural Biology
SV40	Simian vacuolating virus 40
TACE	Tumour necrosis factor- α converting enzyme
TAE	Tris base, acetic acid and EDTA
TAPI-2	N-(R)-[2-(Hydroxyaminocarbonyl)methyl]-4-methylpentanoyl-L-t-butyl-alanyl-L-alanine, 2-aminoethyl Amide, TNF- α Protease Inhibitor-2
TARP	Transmembrane AMPA receptor regulatory protein
TM	Trans-membrane domain
UV	Ultraviolet
V	Volts
VPA	Valproic acid
v/v	Volume by volume
w/v	Weight by volume
WIMM	Weatherall Institute of Molecular Medicine

WTCHG	Wellcome Trust Centre for Human Genetics
YFP	Yellow fluorescence protein

Chapter 1

General Introduction

1.1 General overview

1.1.1 The synapse

The human brain is the most complex organ known to mankind. It contains a vast and highly specialised network of cells which form connections termed synapses that evolve in an activity dependant manner. The term synapse, from the Greek synaptein “to clasp” was coined by Sir Charles Sherrington at the time research was beginning to disprove the ‘reticular theory’ which assumed that the axon of one cell was continuous with the dendrite of another and gave strength to the ‘neuron theory’ which hypothesized that nerve cells are indeed segregated bodies, contacting other nerve cells at specialised regions termed synapses (Sherrington 1897) (figure 1.1A).

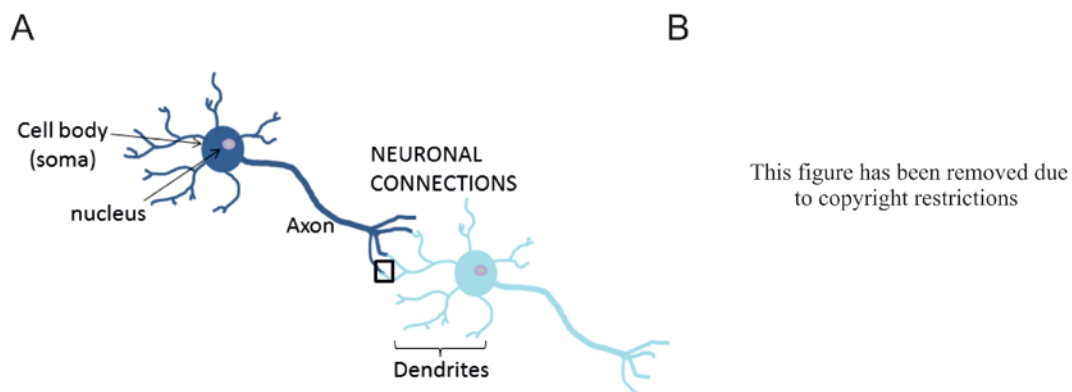


Figure 1.1 Neuronal synaptic connections.

A) Schematic representation of neuronal cell contacts highlighting key features including the cell body (soma), axons and dendrites. B) This figure has been removed due to copyright restrictions

There are many types of neuronal cells, which in turn appear to be involved in a variety of synaptic connections that can be characterised according to their structural morphology. They vary in size, shape and composition in their pre- and post-synaptic organisations, some key features of which are discussed here.

1.1.2 Synaptic components

The pre-synaptic compartment of the axon is typified by the presence of a large number of vesicles that contain neurotransmitters (figure 1.1B). The vesicle membranes are packed full of proteins involved in exo- and endocytosis (Takamori *et al.* 2006; Gordon *et al.* 2011). Vesicles vary in size but are typically spherical and the majority are between 20-40nm in diameter (Cowan *et al.* 2003; Takamori *et al.* 2006), although additional forms termed dense core vesicles have been observed up to 150nm in diameter and contain secreted proteins with a multitude of functions (Wegrzyn *et al.* 2010).

The synaptic cleft (figure 1.1B) is often oversimplified and wrongly depicted to be a void in which neurotransmitters diffuse across to reach targets on the post synapse. Close analysis reveals that it is more like an intercellular plaque containing a dense array of glycoproteins and glycolipids (Bourgeois *et al.* 1994; Lucic *et al.* 2005; Zuber *et al.* 2005).

The dendrite completes the connection and is typified by the presence of a post-synaptic density (PSD) (figure 1.1B). Sometimes the connection between the two cells appears symmetrical i.e. the pre- and post-synaptic specialisations are roughly the same size (as frequently observed at inhibitory synapses), however the dendritic PSD is often more substantial in size and can increase in response to excitatory activity (Cathala *et al.* 2005).

1.1.3 Types of synapses

Synapses do not always form between the axon of one neuron and the dendrite of another (i.e. axodendritic) (as depicted in figure 1.1A) but can form many connection types including axodendritic, axoaxonal and dendrodendritic (figure 1.2A-C) (Cowan and Kandel 2003). Synapses do not only occur at neuronal-neuronal contacts but are also present at neuromuscular junctions between muscle cells and motor neurons (Hughes *et al.* 2006). Generally speaking, all synapses can also be classified into two types, chemical and electrical. Throughout history there has been speculation as to whether neuronal synapses are chemical or electrical in nature. It is

now widely accepted that both types occur *in vivo*. The majority of synaptic connections are chemical, where neurotransmitter release triggers an action potential (Bliss *et al.* 1993). In rare cases, two cells may form an electrical connection, for example the dendrodendritic contacts at gap junctions that act to synchronise cell firing in inhibitory interneurons (Gibson *et al.* 1999).

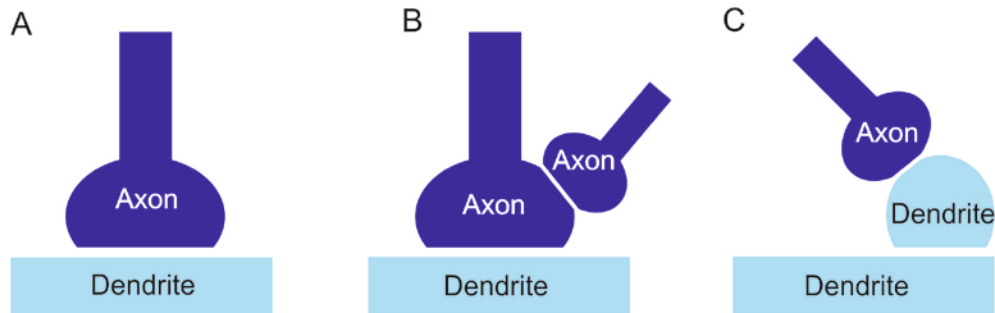


Figure 1.2 Schematic representation of three synaptic types.

A) The 'classic' axon (navy blue) and dendrite (light blue) axodendritic interaction. B) An axodendritic plus an axoaxonal interaction. C) An axodendritic synaptic connection plus a dendrodendritic interaction.

At synapses, a bidirectional flow of information occurs between cells and is modulated in response to activity. In addition to the different synaptic morphologies outlined above, neuronal synapses are either inhibitory or excitatory in nature and differ hugely in their molecular content. The presence of particular neuroreceptors indicates what type of signalling mechanisms occur at that specific synapse.

1.1.4 Neurotransmitters

Neurotransmitters are amino acids, peptides or small molecules that must meet a certain criteria. They must be synthesized by the pre-synaptic neuron, released upon stimulation (membrane depolarisation), produce rapid-onset reversible responses in the target cell and they must have a specific receptor (Cowan and Kandel 2003). Neurotransmitters are released from vesicular compartments within the pre-synaptic density in response to neuronal activity (Cowan and Kandel 2003) (figure 1.3). Glutamate is considered to be the major excitatory neurotransmitter of the central nervous system (CNS) (Danbolt 2001). Many other excitatory and inhibitory neurotransmitters have been discovered over the years including acetylcholine,

dopamine, serotonin, glycine, noradrenaline and gamma-aminobutyric acid (GABA) and all have well characterised effects (Nicoll *et al.* 1990). Development of the glutamate-derivative synthetic n-methyl-d-aspartate (NMDA) led to the NMDA and non-NMDA initial classification of glutamate receptors (Davies *et al.* 1979). The non-NMDA group was then divided into the kainate (Davies *et al.* 1979) and alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) subtypes (Honore *et al.* 1981). Most recently, D-serine was shown to modulate a subgroup of glutamate receptors, the delta receptors (Naur *et al.* 2007; Hansen *et al.* 2009).

Neurotransmitter release ultimately results in morphological remodelling of the synapse to alter the strength of the connection in response to activity. Hebbian theory aims to describe neurological changes associated with learning. It is often summarised by ‘cells that fire together wire together’ (Hebb 1949) and forms the basis of the modern view of a plastic synapse i.e. the ability of the synapse to alter its strength, in an activity dependant manner (Martin *et al.* 2000). Synaptic plasticity resulting in an increase in synaptic strength through high frequency stimulation is known as long term potentiation (LTP) (Bliss *et al.* 1973). LTP has been extensively studied *in vitro* in CA1 hippocampal pyramidal cells, where long-lasting morphological changes, resulting from molecular processes outlined below, have been observed (Malenka *et al.* 1999).

1.1.5 Long term potentiation (LTP)

In CA1 hippocampal pyramidal cells, LTP requires the activation of NMDA receptors. At low levels of neuronal activity NMDA receptors contain a Mg^{2+} plug that blocks the ion channel (figure 1.3). When a nerve impulse (or wave of depolarisation) sweeps down the axon into the pre-synaptic terminal it triggers the release of glutamate into the synaptic cleft (Meldrum 2000). Glutamate activates many pre- and post-synaptically located receptors including the post-synaptic α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors, where glutamate binding causes conformational changes to open their ion channels. An influx of Na^{2+} ions then ensues, causing the depolarisation of the post-synaptic membrane and release of the NMDA

receptor Mg^{2+} plug. The NMDA receptor GluN2 subunits bind glutamate, but receptor activation is further potentiated by the ability of the GluN1 and GluN3 subunits to bind glycine (Johnson *et al.* 1987; Furukawa *et al.* 2005). An increased influx of Ca^{2+} ions through NMDA receptors then initiates a complex array of downstream signalling pathways including AMPA receptor phosphorylation and receptor trafficking (Berridge *et al.* 2003). The influx of Ca^{2+} ions also initiates the transcription of a vast number of genes and has been shown to increase cAMP response element-binding protein (CREB) activity which regulates the transcription of multiple genes containing the CRE consensus sequence in their promoter region (Lonze *et al.* 2002). Calcium also induces transcription of immediate early genes (IEGs), for example the neuronal activity dependant protein (NARP) (also known as neuronal pentraxin 2, NP2) which potentially acts to cluster AMPA receptors (O'Brien *et al.* 1999) (discussed in section 1.5).

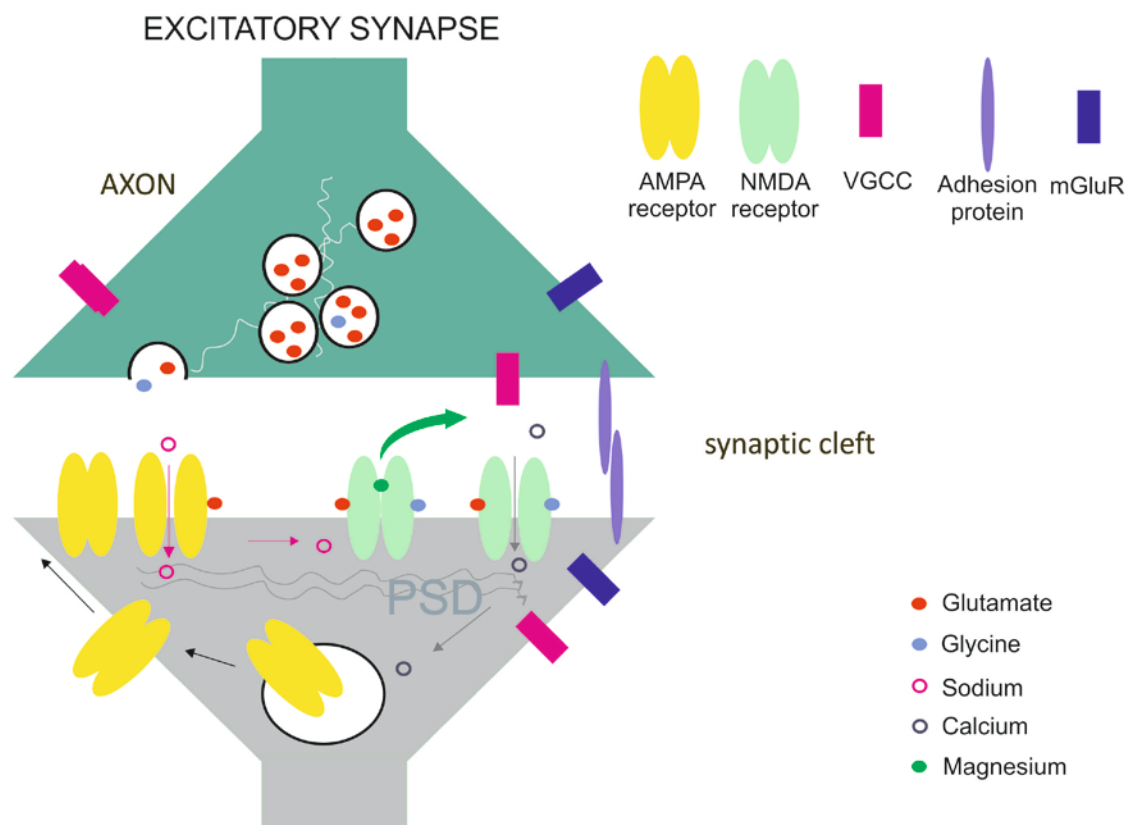


Figure 1.3 Excitatory synaptic transmission.

Schematic representation of an excitatory synaptic connection to depict a select number of cellular processes occurring as a result of long term potentiation (LTP). Release of glutamate from pre-synaptic compartment triggers activation of AMPA receptors on the post synaptic membrane causing an influx of Na^{+} ions. This depolarises the post synaptic compartment enabling the release of NMDA's Mg^{2+} plug. An influx of calcium ions ensues which have many downstream effects including an increase in AMPA receptor trafficking.

1.1.6 AMPA receptor trafficking in LTP

AMPA receptor synthesis, subunit assembly and trafficking mechanisms have been closely dissected to reveal important key events. The AMPA receptors are synthesised and assembled in the endoplasmic reticulum (ER). Exit from the ER is dependent on the subunit composition and ribonucleic acid (RNA) processing. Importantly for this discussion, AMPA receptor trafficking is implicated in LTP (Malinow *et al.* 2002) where its movements into the PSD leads to an increase in synaptic strength.

Mobile pools of AMPA receptors diffuse laterally in and out of the synapse by Brownian motion and can be recycled via clathrin-dependant endocytosis and exocytosis at distinct regions/zones (Petrini *et al.* 2009). LTP results in the synaptic recruitment of AMPA receptors from extra-synaptic and intracellular pools (figure 1.3) (Hayashi *et al.* 2000; Malinow and Malenka 2002; Kessels *et al.* 2009) and this facilitates rapid receptor trafficking, enabling a non *de novo* transcriptional response to synaptic activity. AMPA receptors are then anchored at the synapse by their intracellular interactions with PSD scaffolding proteins such as glutamate receptor-interacting protein (GRIP) and protein interacting with C-kinase (PICK) as well as through interactions with auxiliary proteins such as transmembrane AMPA regulatory protein-2 (TARP2) (discussed in section 1.6.2) to PSD-95.

AMPA receptor phosphorylation, specifically an increased phosphorylation of the GluA1 subunit, is also involved in the activity-dependant remodelling that occurs in LTP (Lee *et al.* 2000). A knock-out (KO) mouse model demonstrated that the presence of the AMPA receptor subunit GluA1 is vital for LTP to occur in the hippocampus (Zamanillo *et al.* 1999). The GluA1 subunit contains several phosphorylation sites at its intracellular C-termini. GluA1 phosphorylation by protein kinase A (PKA) and protein kinase C (PKC) was mapped to residue S845 and S831 respectively (Roche *et al.* 1996) and phosphorylation at S845 is postulated to cause AMPA receptor trafficking to extra-synaptic locations, ready for insertion into the synaptic PSD (Oh *et al.* 2006). Phosphorylation of GluA1 also enhances binding to an actin localisation

protein termed 4.1N, which enables the receptors to be inserted into the membrane (Lin *et al.* 2009).

1.1.7 Long term depression (LTD)

A decrease in synaptic strength induced by prolonged repetitive stimulation is a phenomenon known as long-term depression (LTD). Mechanisms of LTD are still not fully understood but are thought to involve the endocytosis of receptors (Collingridge *et al.* 2004), because inhibition of AMPA receptor endocytosis was shown to prevent LTD (Lüscher *et al.* 1999). In LTD, an increase in intracellular calcium levels activates the calcium sensor hippocalcin which then binds to the adaptor protein complex 2 (AP-2) causing its translocation to the plasma membrane. AP-2 interacts with AMPA receptors and recruits clathrin, resulting in AMPA receptor internalisation into clathrin coated vesicles (Palmer *et al.* 2005). The influx of calcium also triggers the phosphorylation of the AMPA receptor GluA2 subunit by PKC at residue S880 which affects its binding to the domain present in PSD-95, Dlg, and ZO-1/2 (PDZ) -containing proteins GRIP1 and PICK1, resulting in AMPA receptor internalisation (Hee Jung *et al.* 2000).

In NMDA-independent forms of LTD, activation of pre- and post-synaptic metabotropic glutamate receptors (mGluRs) triggers a cascade of intracellular biochemical events to modulate synaptic activity (Ozawa *et al.* 1998). The mGluRs are members of the G-protein-coupled receptor (GPCR) family of proteins and their activity results in the phospholipase C (PLC) controlled release of calcium from intracellular stores. This mGluR-associated release of calcium has been stated to result in the activation of tumor necrosis factor- α converting enzyme (TACE) (a matrix metallo-protease (MMP)) that has been postulated to cleave pre-synaptically attached neuronal pentraxin receptor (NPR) and results in AMPA receptor endocytosis (Cho *et al.* 2008).

1.1.8 Negative effects of synaptic alterations

Alterations and deficiencies in synaptic signalling components are involved in a vast number of neurological disorders. In fragile X syndrome, synapses appear unable to mature from

filopodia (Fiala *et al.* 2002) potentially as a result of a mutation in Fragile X mental retardation protein (FMRP) which is thought to be involved in mRNA transport and synaptic development (Bassell *et al.* 2008). Synaptic morphological changes, resulting in cognitive and motor decline, occur in neurodegenerative diseases such as Alzheimer's disease (AD), Parkinson's disease and amyotrophic lateral sclerosis (ALS) (Bossy-Wetzel *et al.* 2004; Kessels and Malinow 2009). An increase in surface expression of synaptic receptors involved in LTP is also observed in addiction (Wolf *et al.* 2004).

Diseases implicating AMPA receptors often actually result from dysregulation of proteins involved in AMPA receptor trafficking, synaptic surface expression and intracellular signalling mechanisms, which act to impair or enhance receptor signalling, rather than from mutations in the AMPA receptor themselves. In AD, soluble amyloid Beta ($A\beta$) protein was shown to impair AMPA receptor trafficking mechanisms, resulting in a reduction of synaptic AMPA receptor levels that is implicated in the AD-associated spine loss and subsequent neurodegeneration (Ting *et al.* 2007; Henley *et al.* 2011). Transmembrane AMPA receptor interacting proteins (TARPs), involved in the control of AMPA receptor synaptic surface expression and intracellular signalling, have also been associated with schizophrenia. TARPs modulate AMPA receptors by several methods (discussed in Chapter 1, section 1.6.2). They bind to the intracellular C-terminal of AMPA receptors, anchoring them to the synapse by bridging an interaction to PSD-95 protein. They also modulate AMPA receptor gating mechanisms to slow recovery from desensitization (Cho *et al.* 2007). Using *in-situ* hybridization analysis of brain sections from schizophrenic patients at post-mortem, led to the observation that TARP2 mRNA levels were increased in the dorsolateral prefrontal cortex and was postulated to be the cause of the AMPA receptor dysregulation associated with schizophrenia (Beneyto *et al.* 2006). It is rapidly becoming apparent that pathological changes in cellular signalling pathways, which are implicated in synaptic plasticity, form the basis of many neurological conditions. Therefore unravelling the complex web of signalling mechanisms is of huge medical importance.

1.2 The iGluRs

1.2.1 Overview

Receptors for glutamate can be broadly categorised into two groups; ionotropic (iGluR) and metabotropic (mGluR). The iGluRs were classified according to the synthetic agonist that they were first found to bind and are referred to as AMPA, NMDA, kainate and delta receptors (figure 1.4A). The iGluRs are ligand-gated ion channels that bind the neurotransmitter glutamate. All iGluR subunits share a common domain organisation consisting of an extracellular amino terminal domain (ATD) and ligand-binding core/domain (LBC/LBD) as well as a transmembrane region and an intracellular C-terminus (figure 1.4B).

A

This figure has been removed due to
copyright restrictions

B

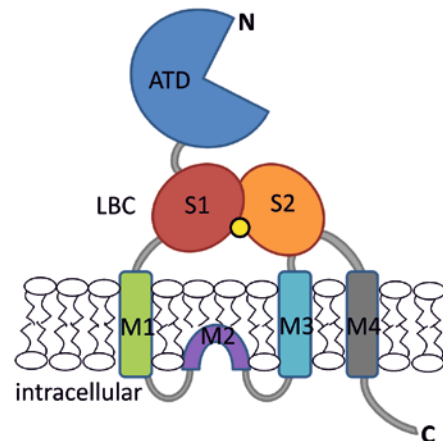


Figure 1.4 iGluR nomenclature.

A) This figure has been removed due to copyright restrictions.

B) Conserved iGluR domain organisation including amino terminal domain (ATD), ligand binding core (LBC) and transmembrane domain containing helices termed M1, 2, 3 & 4 (where M2 forms the receptor pore), each subunit also contains an intracellular C-terminal.

1.2.2 Origins of the iGluRs

Eukaryotic iGluRs have a highly conserved modular domain organisation, suggesting a common evolutionary ancestral path, postulated to have formed from the fusion of three separate genetically dissimilar gene segments (Mayer 2006). Phylogenetic analysis demonstrates that the iGluRs are an evolutionarily ancient class of proteins, predating the plant-animal divergence.

Glutamate-like receptors (GLRs) are present in plants (Lam *et al.* 1998) and contain the distinctive LBC and transmembrane domains. The functionally important pore-forming portion of the plant GLR transmembrane domain contains the highest sequence conservation to vertebrate iGluRs (Chiu *et al.* 1999). An iGluR receptor (termed AvGluR) has been identified in the ancient eukaryote *Adineta vaga* whose sequence alignment predicts a similar domain organisation to higher order eukaryotic iGluRs and the protein is thought to be an evolutionary link between prokaryotic and eukaryotic glutamate receptors (Janovjak *et al.* 2011). Recent structural analysis of the AvGluR LBD shows that it binds glutamate in the same conformation as eukaryotic LBDs, but like prokaryotic LBDs, it can also bind other amino acids (Lomash *et al.* 2013).

Eukaryotic iGluR ATDs show fold homology to bacterial leucine/isoleucine/valine-binding proteins (LIVBP) (figure 1.5) which are involved in the active transport of sugars and amino acids (Amanuma *et al.* 1974). The ATD has a clam-shell fold that contains two lobes (termed D1 and D2) with a central β -sheet surrounded by α -helices and forms a dimer in solution. The LBD is much smaller than the ATD and shows homology to bacterial potassium channels (Chen *et al.* 1999). The LBD is formed by two unconnected polypeptide chains termed S1 and S2 and like the ATD has a clam-shell topology also with two lobes (D1 and D2). The LBD is responsible for binding glutamate and has a flexible hinge region which acts to pull apart the transmembrane region, allowing the iGluR channel to open and close (Armstrong *et al.* 2000). Finally the transmembrane domain of each AMPA receptor subunit comprises four hydrophobic helices termed M1-4 (I-IV) with the M2 portion forming a central pore-like helix (figure 1.4) and shows sequence homology to an inverted form of the bacterial potassium channels (Panchenko *et al.* 2001; Kuner *et al.* 2003).

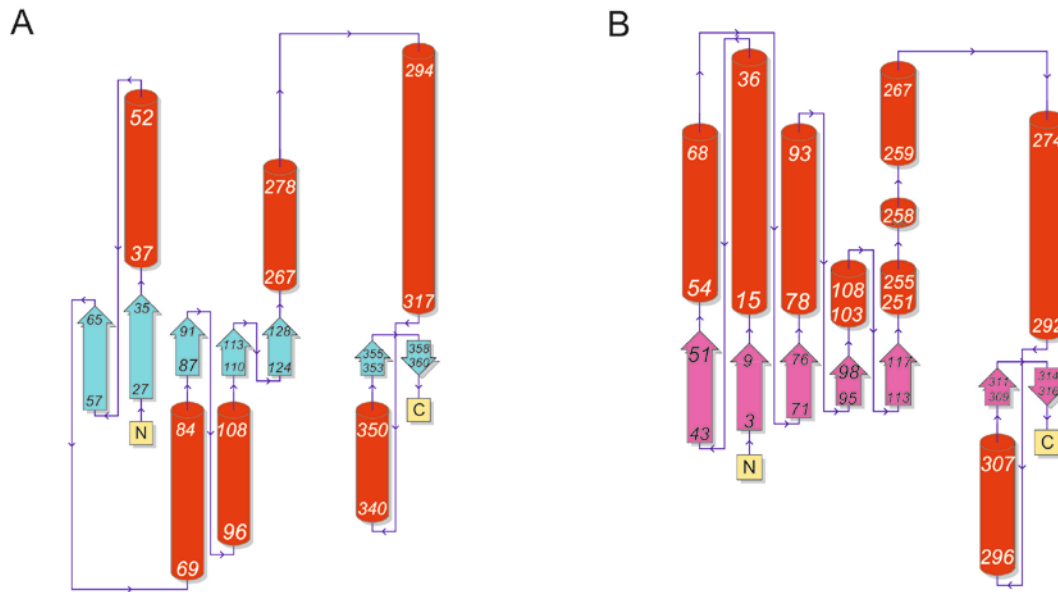


Figure 1.5 Structural homology conservation.

A) AMPA receptor GluA2_ATD (PDB entry 2WJW) and B) bacterial leucine isoleucine valine binding protein (LIVBP) (PDB entry 1Z17). Diagrams were generated using PDBsum (<http://www.ebi.ac.uk/pdbsum/>) to highlight overall fold conservation.

1.2.3 The iGluRs in the developing brain

During development neurons migrate, form branches (arborizations) and then go on to make synaptic connections that are tightly regulated by neuronal activity. In early embryonic stages of development NMDA receptors are expressed but are silent because the AMPA receptors present are non-functional, impeding the membrane depolarisation required for release of the NMDA Mg^{2+} channel plug (Ben-Ari *et al.* 1997).

AMPA receptor subunits contain alternative splice variations and RNA editing sites. The flip/flop variants are created by alternative splicing of an interchangeable exon in a 115bp region preceding the fourth transmembrane helix (M4) (figure 1.6) (Sommer *et al.* 1990). A higher level of GluA2 flip, associated with slower receptor desensitisation kinetics, is observed before birth while in adults, levels of the flop form (whose kinetics are much faster) become dominant (Seifert *et al.* 2000). A codon switch, controlled by intronic elements, edits an R/G site preceding the flip/flop variable region (figure 1.6) and results in faster AMPA receptor recovery following desensitization (Lomeli *et al.* 1994). Finally the Q/R RNA editing site is responsible for controlling calcium permeability in the GluA2 subunit (figure 1.6) (Swanson *et al.* 1997).

1.3 The AMPA receptor

1.3.1 Overview

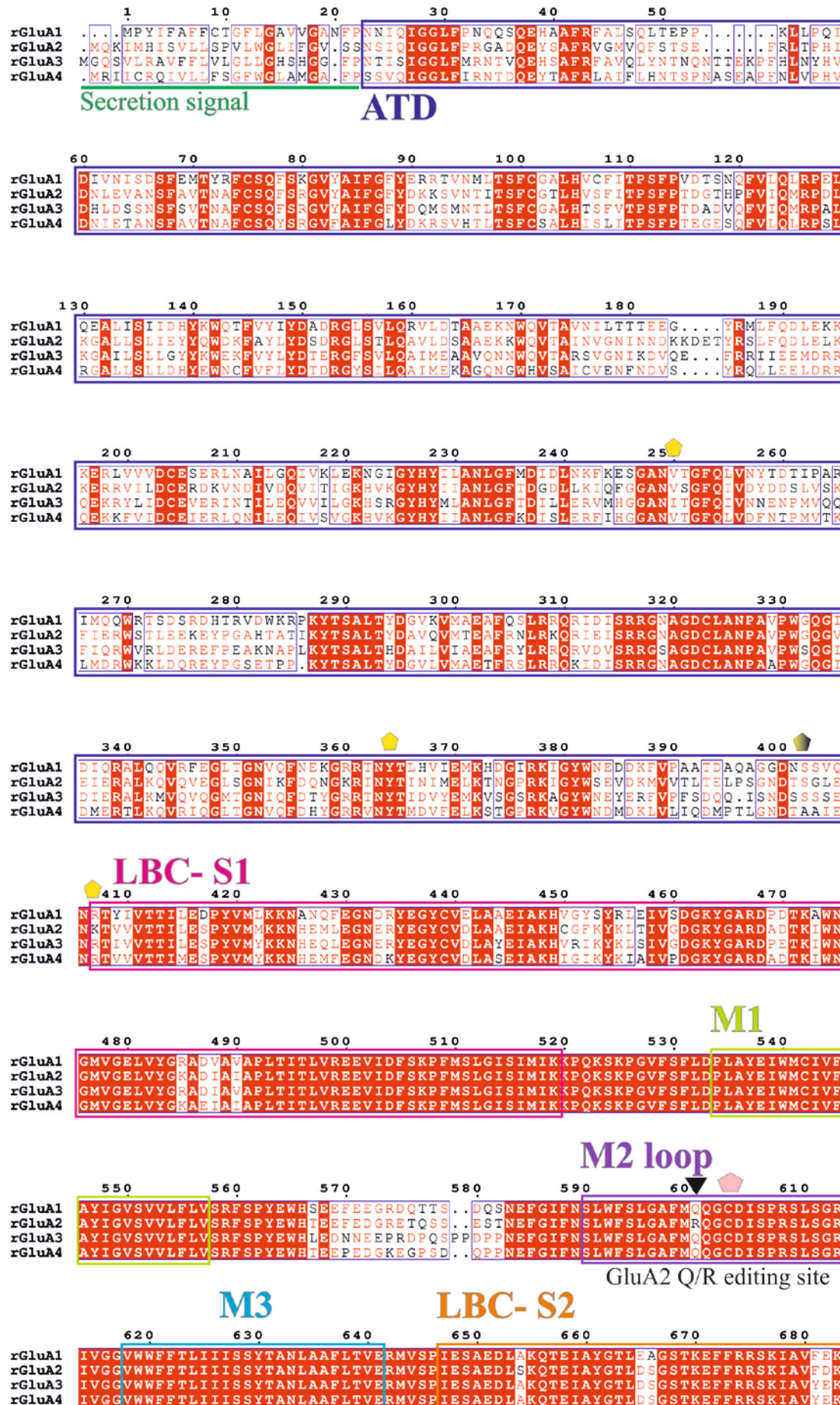
The AMPA receptors are responsible for mediating the majority of fast excitatory neurotransmission in the CNS (Ozawa 1996). AMPA receptors contain combinations of four subunits termed GluA1-4 (Collingridge *et al.* 2009) (previously termed GluR1-4). These subunits exist in hetero-tetrameric, dimer of dimers assemblies, encoded for by four genes; *GRIA1-4*.

AMPA receptors have been linked to several disease states. The flip/flop region been implicated in ALS, a form of motor neuron disease associated with a weakening of muscle, eventually leading to death. ALS may be caused by a variety of genetic mutations but patients suffering from this disease have an abundance of the flip variant (Tomiya *et al.* 2002). Flip-form AMPA receptors have lower desensitisation kinetics resulting in glutamatergic excitotoxicity-associated motor neuron death. The GluA2 subunit is responsible for rendering the AMPA receptor impermeable to calcium by its glutamine to arginine (Q/R) editing site. The positively charged arginine prevents calcium ions from passing through the receptor's pore. In cases of epilepsy and forebrain ischemia, GluA2 mRNA is down-regulated causing an increased expression of AMPA receptors that are permeable to calcium and thus enhances excitotoxicity.

1.3.2 AMPA receptor subunit organisation

Sequence alignment of GluA1-4 (figure 1.6) demonstrates that regions of both high and low sequence conservation are present across subunits. The transmembrane domains are the most conserved regions, whereas highest levels of variability are observed in the ATD and intracellular C-terminal. The intracellular C-terminal is involved in AMPA receptor subunit-specific protein-protein interactions (Ayalon *et al.* 2001). Each subunits C-terminal is also differentially phosphorylated and interacts with a host of PSD proteins (Santos *et al.* 2009). The ATD is postulated to have a role in subunit-specific receptor assembly and trafficking (Jin *et al.* 2009; Sukumaran *et al.* 2011), where the sequence divergence enables subunit-specific protein-protein

interactions to occur (Ayalon and Stern-Bach 2001). Truncations of the ATD have no effect on channel conductance, since AMPA receptors lacking their ATDs were shown to gate normally (Pasternack *et al.* 2002). Despite homology with proteins known to bind small molecules at the ATD inter-domain cleft between D1 and D2, at the time this research commenced there were no published data describing the AMPA receptor ATDs in such an interaction. AMPA receptor ATDs are involved in interactions with the neuronal pentraxins, postulated to cause receptor clustering (O'Brien *et al.* 1999; O'Brien *et al.* 2002; Sia *et al.* 2007; Chang *et al.* 2010). Clustering of AMPA receptor by neuronal pentraxins was shown to be impeded in ATD truncated AMPA receptors (O'Brien *et al.* 1999; Sia *et al.* 2007) (discussed in Chapter 4).



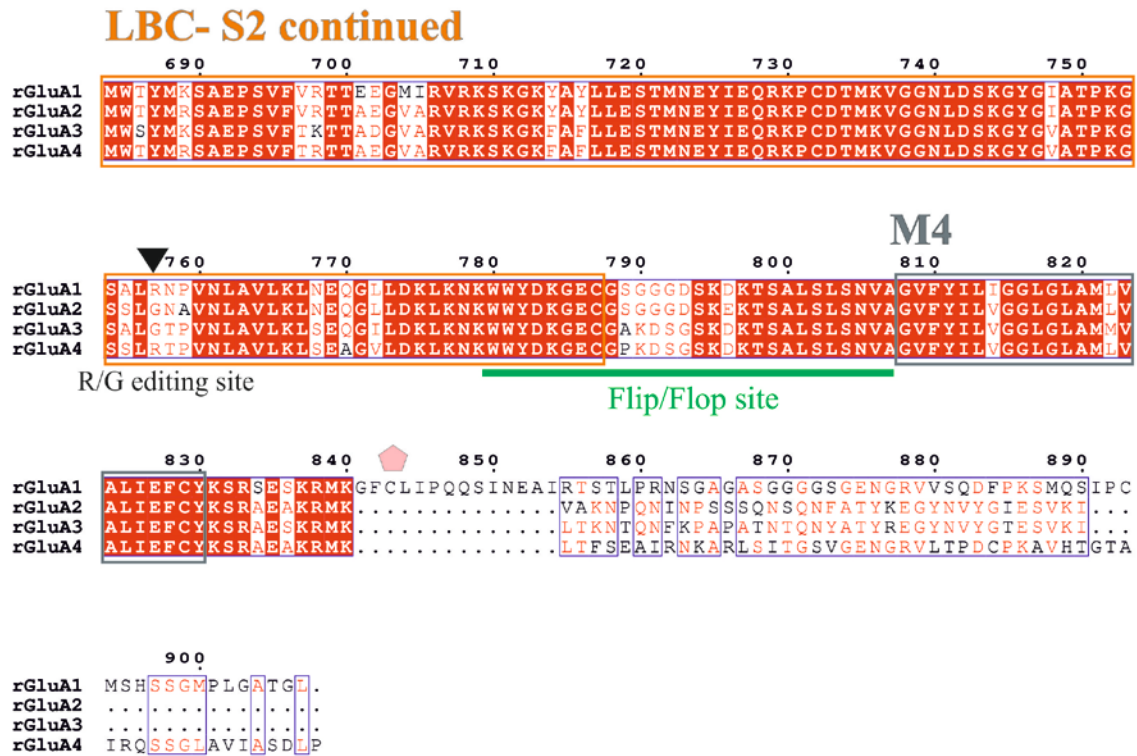


Figure 1.6 AMPA receptor subunits GluA1-4 alignment.

Rat GluA1-4 alignment generated in Multalin (<http://multalin.toulouse.inra.fr/multalin/>). Includes notations from SignalP (<http://www.cbs.dtu.dk/services/SignalP/>) for a predicted secretion signal. HHPred (<http://toolkit.tuebingen.mpg.de/hhpred>) domain predictions for ATD, LBD and TM regions. Additional trans-membrane domains prediction (based on TMHMM Server v. 2.0 <http://www.cbs.dtu.dk/services/TMHMM/>). N-linked glycosylation sites are indicated by yellow hexagons (yellow/black when subunit specific) as predicted by NetNGlyc (<http://www.cbs.dtu.dk/services/NetNGlyc/>). Pink hexagons indicate palmitoylation sites (Hayashi *et al* 2005). The GluA2 specific editing site (Q/R) also indicated (black triangle).

1.3.3 Structural information timeline

When this research project began, only the structure of the AMPA receptor LBC was known. There are now more than 40 GluA2 LBC structures deposited in the protein database (PDB) either in complex with the neurotransmitter glutamate, or with other agonist or antagonist ligands. These structures give atomic resolution knowledge to explain how the LBC controls AMPA receptor gating. A tight association of the LBC top lobes (D1-D1) form a dimer and disruption of inter-dimer salt bridges and hydrogen bonds of the D1 lobe result in rapid AMPA receptor desensitization (Horning *et al.* 2004). The range of domain closure upon ligand binding is attributed to the flexibility in the D2 lobes and differentially effects the receptor responses (Mayer 2006). The LBC is involved in protein-protein interactions with cystine-knot AMPA receptor modulating protein-44 (CKAMP44) (Von Engelhardt *et al.* 2010) and TARPs (Tomita *et*

al. 2003) (discussed in Chapter 5). These auxiliary proteins act to modulate the channel conductance, but there is currently no structural evidence to show how this occurs.

As already eluded to, the AMPA receptor ATD was structurally uncharacterised when this project was in its infancy, therefore crystallization of the GluA2 ATD was attempted and proved successful (Clayton *et al.* 2009) (results of which can be found detailed in Chapter 3). Since then, the structures for all AMPA receptor ATDs have been solved (GluA1 (Yao *et al.* 2011), GluA3 (Sukumaran *et al.* 2011) and GluA4 (Dutta *et al.* 2012), a comparative analysis of which can be found in the discussion section of Chapter 3. Following the publication of the GluA2 ATD structure detailed in Chapter 3, the full length homo-tetrameric rat GluA2 was solved to 3.6 Å (Sobolevsky *et al.* 2009) (figure 1.7). The protein was shown to be functionally active with gating properties similar to the wild type (WT) protein but may not be physiological because AMPA receptors do not typically form homo-tetramers (Greger *et al.* 2007). The structure confirms years of analysis of isolated individual receptor components and enables a vast amount of data to be interpreted, including how receptor desensitisation is caused by a disruption of the LBC D1-D1 interface that results in a rigid body rotation to a closed conformation (Sobolevsky *et al.* 2009). An unexpected feature was the proteins' mismatched 2- to 4-fold symmetry caused by an ATD 'domain swap', the authors postulating that this subunit crossover increases receptor stability *in vivo*. Extensive subunit-subunit interactions were also observed in the transmembrane domain which confirms the previous observation that the M4 helix is involved in receptor assembly (Wollmuth *et al.* 2004).

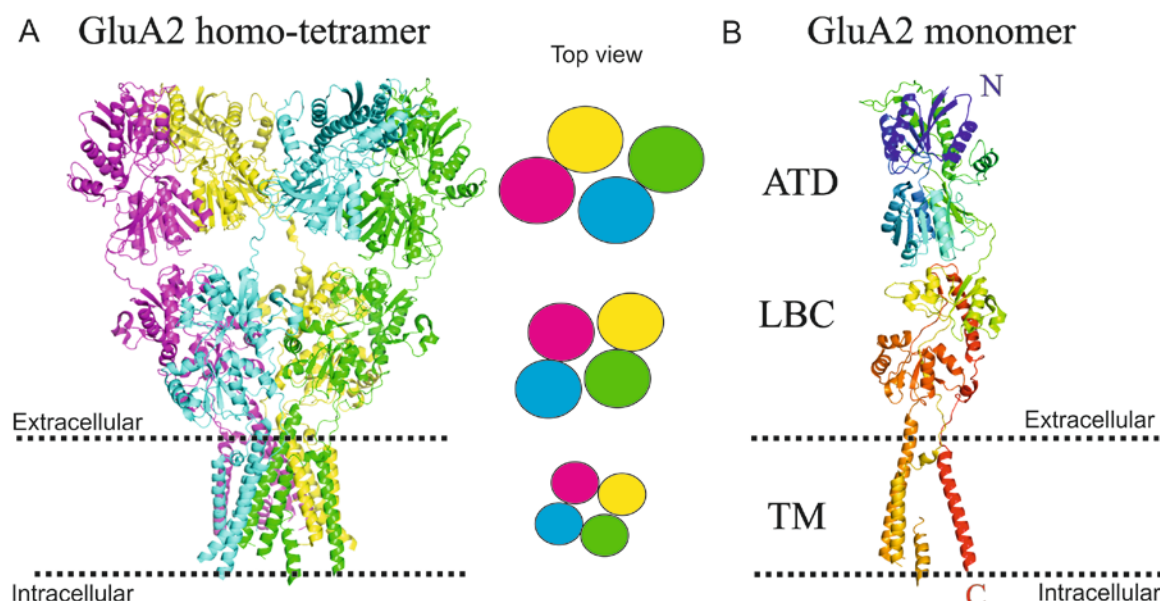


Figure 1.7 Structural analysis of the AMPA receptor GluA2 homotetramer.

A) Cartoon rendering of the GluA2 homo-tetramer with the four chains (subunits) individually coloured (pink, yellow, blue and green) including schematic 'top-down' view to highlight the mismatch in symmetry caused by the ATD. For example in the LBC region the pink and blue chains form a dimer whereas in the ATD the pink and the yellow chains form a dimer. B) Single GluA2 subunit rendered in cartoon format with rainbow colouring from N-terminal (blue) to C-terminal (red) to highlight convoluted chain organisation through the molecule. This figure was generated in PyMOL (DeLano, 2002) using PDB ID 3KG2 (Sobolevsky *et al.*, 2009).

1.3.4 Examples of structural information aiding medical intervention

Ampakines are a class of cognitive enhancement drugs (nootropic agents) which act as positive allosteric modulators of AMPA receptors to enhance glutamatergic synaptic transmission (Lynch 2006). Their ability to induce LTP without triggering seizures (Staubli *et al.* 1994) has made them a potential drug target for treating neurodegenerative and psychiatric diseases. Aniracetam was one of the first ampakine compounds noted for its ability to reversibly enhance excitatory post-synaptic potentials (EPSPs) (Ito *et al.* 1990), although its binding efficiency was considered to be relatively weak (5mM) (Tang *et al.* 1991). Many other structurally-related benzamide compounds have since been developed e.g. CX614 (Cortex Pharmaceuticals) (Arai *et al.* 2000). Protein structures solved using X-ray crystallography have provided a detailed understanding of how ampakines interact with AMPA receptors to stabilise the LBC clamshell in the closed (glutamate bound) conformation, slowing receptor desensitisation kinetics (Jin *et al.*

2005). Structural information can therefore enable target compounds to be modified to improve their binding specificity and drug potency and reduce side effects (Jin *et al.* 2005).

Structural information can also give insights into how disease states involving AMPA receptor mutations manifest at the protein level. The Online Mendelian Inheritance in Man (OMIM) database found at <http://www.ncbi.nlm.nih.gov/omim> provides information on all currently known genetic phenotypes. In one case X-linked mental retardation was linked to the AMPA receptor subunit GluA3, where several missense variants produced the mutations; G833R (causing rapid degradation of abnormally folded protein), R631S and M706T (which both altered desensitization kinetics) (Wu *et al.* 2007). Sequence alignment and structural-guided analysis using the solved GluA2 homotetramer (Sobolevsky *et al.* 2009) can reveal how these mutations might impact receptor function. Of the mutants which affect the receptors desensitisation kinetics, the R631S mutation would reside in the transmembrane M2 re-entrant loop and M706T would be found within the S2 domain of the LBC (located at R613 and M688 respectively in figure 1.6). Both of these mutants were stated to cause faster desensitization and deactivation kinetics (Wu *et al.* 2007).

Another *de novo* mutation has been described for the GluA1 subunit at A636T (de Ligt *et al.* 2012) which can be mapped to the transmembrane M3 pore-lining helix. An M3 mutation was also discovered during analysis conducted at the Wellcome Trust Centre for Human Genetics (WTCHG). Here a GluA3 (A653T) mutation (A635 in M3 figure 1.6) was found in twin brothers with mental retardation and sleep-wake cycle deficits (unpublished data). Sequence alignment suggests that these M3 mutations align to the region where the well-studied Lurcher mouse delta (iGluR) receptor mutation occurs (Zuo *et al.* 1997) (aligned to A636 in M3 figure 1.6). Heterozygous Lurcher mice display ataxia due to a gain of function where high inwardly rectifying membrane conductance results in apoptotic neural death and neurodegeneration (Zuo *et al.* 1997).

Finally, a small subset of patients with epilepsy and ataxia display autoimmunity to a host of synaptic proteins (Irani *et al.* 2011) and in certain types of epilepsy, autoantibodies specific to iGluRs are routinely screened for as a probable cause of this debilitating disease (Irani *et al.* 2011). Clinical cases implicating iGluR receptor mutations present an exciting opportunity to use structural analysis to understand clinical phenotypes. It is anticipated that observed phenotypes may be directly explained by alterations to receptor gating properties so hopefully, in the near future, structural-guided research will routinely enable medicinal control of protein dysregulation to prevent or reduce disease states.

1.4 AMPA receptor interacting proteins

1.4.1 An overview

AMPA receptor complexes have emerged as a field of interest in their own right since the full length GluA2 homo-tetrameric structure was solved (Sobolevsky *et al.* 2009). The view of what exactly constitutes a physiological AMPA receptor is rapidly evolving, as the tetrameric core forming the actual ion channel seems merely to represent the central component of a larger supra-molecular assembly (Rouach *et al.* 2005; Von Engelhardt *et al.* 2010; Schwenk *et al.* 2012; Shanks *et al.* 2012). Multiple protein partners are thought to surround and modulate this core (depicted in figure 1.8) but the stoichiometry, structural arrangement and functional consequences of such interactions are yet to be fully defined. In this thesis, research focused on several such molecules, putative interactors of the 700 amino-acid AMPA receptor extracellular region these being the neuronal pentraxins, CKAMP44, TARP2 (also known as Stargazin) and (brain-specific chordin-like protein) Brorin (discussed in sections 1.5 and 1.6). Recent research has also identified a host of other putative AMPA receptor-interacting proteins (Schwenk *et al.* 2012; Shanks *et al.* 2012), that will only be briefly described here.

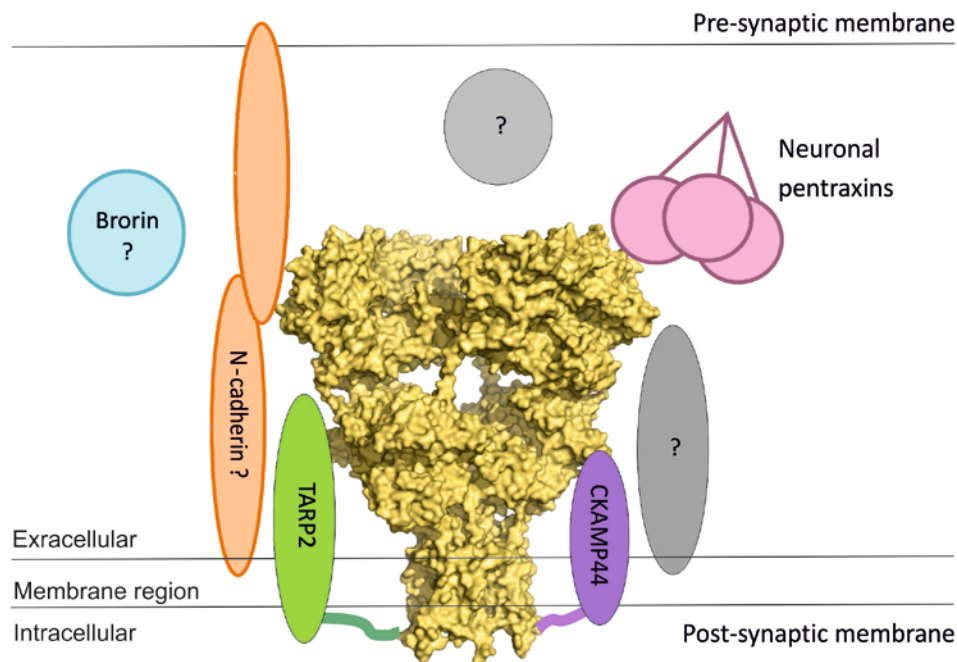


Figure 1.8 Deciphering the AMPA receptor interactome.

Schematic representation of an AMPA receptor (based on structure published by Sobolevsky *et al.*, 2004) surrounded by a host of auxiliary proteins. The AMPA receptor is known to interact with the membrane attached protein TARP2 (green) and CKAMP44 (purple). The AMPA receptor is also known to interact with other proteins such as the neuronal pentraxins (pink) and are postulated to interact with a host of other proteins including Brorin (blue) and N-cadherin (orange) and potentially a number of unknowns.

1.5 The neuronal pentraxins

1.5.1 The pentraxin family

Pentraxins can be divided into two groups simply termed long and short, which identifies the presence or absence of an N-terminal extension. The short pentraxins form multimeric peptide chains of pentraxin (PTX) domains which contains the ‘pentraxin motif’ HxCxS/TWxS where x is any amino acid (figure 1.9) (Garlanda *et al.* 2005) and the long pentraxins contain a PTX domain and an additional N-terminal region. Both types of pentraxin are involved in innate immunity, recognising ‘non-self’ to mark target cells for phagocytosis and degradation (Lu *et al.* 2008). The pentraxin family of proteins are evolutionarily conserved and have been found in the phylogenetically ancient horseshoe crab *Limulus polyphemus* (Shrive *et al.* 1999). The first innate immunity receptor identified was the short pentraxin C-reactive protein (CRP) where its pentameric arrangement (that gave the protein its name) was revealed by negative stain electron

microscopy (EM) (Osmand *et al.* 1977). The first pentraxin crystal structure solved was of another short pentraxin, serum amyloid P (SAP) and this also had a pentameric arrangement (Emsley *et al.* 1994). Since then, other SAP crystal structures have revealed that pentraxins can form higher order oligomers, such as octamers (Shrive *et al.* 1999). The pentraxin C-terminal PTX domain is known to bind to carbohydrates in a calcium-dependant manner (Garlanda *et al.* 2005), whereas the N-terminal region of all long pentraxins remains structurally and functionally uncharacterised.

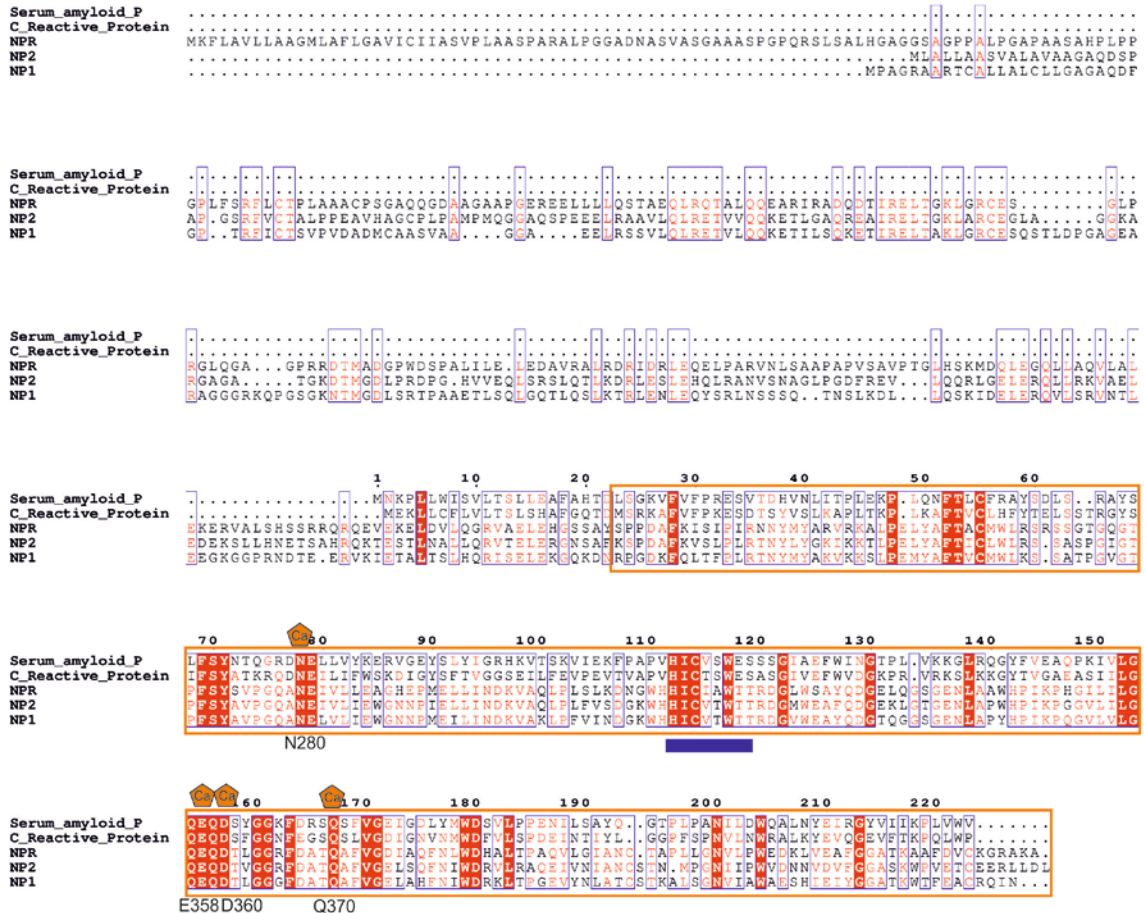
1.5.2 Neuronal pentraxins

The neuronal pentraxins are long pentraxins and are selectively expressed in various regions of the brain as well as the testis (Hsu *et al.* 1995). The neuronal pentraxins are implicated in the developing visual system. Neuronal pentraxin KO mice show impaired visual development, where elimination and eye-specific pruning of synapses in the dorsal lateral geniculate nucleus (dLGN) were found to be greatly reduced (Bjartmar *et al.* 2006). There are three known neuronal pentraxins, neuronal pentraxin 1 (NP1), neuronal pentraxin 2 (NP2) (also known as NARP) and neuronal pentraxin receptor (NPR). NP1 is secreted and stated to be constitutively expressed (Schlimgen *et al.* 1995) while NP2 is a secreted immediate-early response gene (IEG) exhibiting activity-regulated expression (Hsu and Perin 1995). NPR is N-terminally membrane attached and contains a predicted protease cleavage site, enabling release of a secreted form (Dodds *et al.* 1997).

The neuronal pentraxins are predicted to contain two N-terminally located coiled-coil domains, stated to have a role in self multimerization, either as homo- or hetero-complexes (Dodds *et al.* 1997; Kirkpatrick *et al.* 2000; Xu *et al.* 2003). Neuronal pentraxins are also predicted to contain a C-terminal globular PTX domain which has a ‘jelly-roll’ topology (Emsley *et al.* 1994; Shrive *et al.* 1999) with NPR containing a single residue alteration in its ‘pentraxin motif’ region (figure 1.9). Solved pentraxin crystal structures confirm that each PTX domain contains a binding site for two calcium ions, binding of which is required for the pentraxins to act

as lectins and bind carbohydrates (Kilpatrick 2002). The neuronal pentraxins are known to share the calcium-dependent lectin properties (Tsui *et al.* 1996) and alignment of neuronal pentraxin sequences to solved pentraxin structure sequences shows that the calcium co-ordinating residues are conserved (figure 1.9).

A



B

	NP2	NP1	NPR	Serum amyloid P	C reactive protein
NP2	100	67.46	63.51	27.14	28.36
NP1	67.46	100	62.2	26.5	24.75
NPR	63.51	62.2	100	25	25.74
Serum amyloid P	27.14	26.5	25	100	51.24
C Reactive protein	28.36	24.75	25.74	51.24	100

Figure 1.9 Pentraxin sequence alignment.

A) Alignment of *Homo sapiens* pentraxins to highlight pentraxin domain (PTX) conservation. PTX domain (orange boxes) with conserved pentraxin motif HxCxS/TWxS (blue underlined) are indicated. Conserved calcium co-ordination residues are indicated by orange hexagons. B) Scores table for PTX sequence homology calculated using (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) ClustalW2 shows the pair-wise scores calculated for every pair of aligned sequences i.e. the number of identities between the two sequences, divided by the length of the alignment, and represented as a percentage. Values higher than 50% are shown in pink, lower ones shown in blue.

1.5.3 Interactions of AMPA receptors with neuronal pentraxins

The neuronal pentraxins are stated to co-localise with and cluster AMPA receptors (O'Brien *et al.* 2002; Sia *et al.* 2007). This clustering was shown to induce synapse formation by accumulating AMPA receptors at the synapse (O'Brien *et al.* 1999). Some experimental data suggest that receptor clustering requires the AMPA subunit GluA2 to interact with NP2 (O'Brien *et al.* 1999), whilst other research states that all AMPA subunits interact with all neuronal pentraxins. The strongest interaction deduced by the co-localisation of pentraxins and AMPA receptors at synaptic contacts in artificial synapses was shown to occur between GluA4 and NP1 (Sia *et al.* 2007). Research has also shown that receptor clustering requires the presence of the AMPA receptor ATD, where AMPA receptors formed from GluA4 subunits lacking the ATD were not recruited to artificial or endogenous neuronal synapses (Sia *et al.* 2007). The mechanisms by which these proteins interact are yet to be structurally or biophysically characterised.

1.5.4 Disease/disorders implicating the neuronal pentraxins

The neuronal pentraxins have been investigated in a diverse number of systems and a multitude of phenotypes have emerged which link these proteins to diseases and disorder states, some of which are briefly described below.

NP2 is implicated in regulation of the circadian clock where over-expression of NP2 protein blocked the melatonin-induced sleep promotion in live zebrafish larvae (Appelbaum *et al.* 2010). In another study, post-mortem brain sections from narcoleptic patients had an 90% decrease in NP2 positive neurons in areas of the hypothalamus associated with narcolepsy (Blouin *et al.* 2005). Both studies concur that changes in NP2 levels may affect sleep patterns.

The neuronal pentraxins are also potential targets in the treatment of Alzheimer's disease where accumulation of amyloid- β , associated with neuronal cell death, occurs in the brain. An increased level of NP1 was observed in cortical neurons grown in the presence of amyloid- β and

preceded cell apoptosis and neuronal cell death. Promisingly, NP1 gene silencing by RNA interference (RNAi) was shown to prevent this synaptic loss (Abad *et al.* 2006). NPR is also a potential biomarker for neurodegenerative disease because NPR levels are increased in the cerebral spinal fluid (CSF) of patients with Alzheimer's disease (Yin *et al.* 2009).

The neuronal pentraxins have been implicated in certain types of drug addiction. When NP1 and NP2 KO mice were given cocaine, they developed more of a preference for the location where the drug was administered than WT animals. This conditioned place preference was distinct from that observed for NPR. But important to note is the fact that these mice spent a reduced amount of time at the centre of a novel environment, which may indicate that the place preference for the known environment may have been anxiety mediated (Pacchioni *et al.* 2009). NP2 protein levels are also increased in the amygdala following opiate withdrawal, implying that neuronal pentraxins may be associated with molecular adaptations involving AMPA receptors, that result in long lasting behavioural changes (Reti *et al.* 2003).

Research has also shown that inhibition of NP1 may reduce pain sensitivity in rostral ventromedial medulla (Zapata *et al.* 2012). Peripheral nerve injury can cause an increased firing of excitatory receptors and result in hypersensitivity to pain. AMPA receptors have a known involvement in this hypersensitivity because AMPA receptor blockade has been shown to reduce this type of pain. Finally, recently published data suggests a potential association of NP2 with mood disorders. Single nucleotide polymorphisms (SNP) linkage association markers linked NP2 to suicide attempt and major depression disorder (Strauss *et al.* 2012).

1.5.5 *Neuronal pentraxins at parvalbumin-expressing interneurons (PV-INs)*

Neuronal pentraxins are enriched at a class of GABAergic inhibitory neurons termed the parvalbumin-expressing interneurons (PV-INs) (Chang *et al.* 2010). PV-INs are coated in a specialised extracellular matrix termed the perineuronal net (PNN). In the 1880s, Golgi first described an extracellular pericellular net surrounding a select number of neuronal cells and noted

its ability to impede the spread of neural currents (Celio *et al.* 1994). Scientific curiosity in these structures diminished when it was thought that this mesh formed as a by-product of Golgi's 'black reaction' cell-staining technique. Interest in these honeycomb structures, now widely referred to as PNNs, was reignited when lectins were found to bind to them, establishing that PNNs are real and contain negatively charged proteoglycans (PGs) such as chondroitin sulphate proteoglycans (CSPGs) and heparan sulphate proteoglycans (HSPGs) as well as glycoproteins (Celio and Blümcke 1994). As already stated the neuronal pentraxins bind to carbohydrates in a calcium dependant manner, it is therefore perhaps unsurprising that the neuronal pentraxins have been shown to bind to the PNN (Chang *et al.* 2010).

1.5.6 The neuronal pentraxins in learning and memory

Research aiming to unravel the complexities of learning and memory is motivated by the observation of critical periods or 'windows of learning' that occur during development. These critical periods were noted to be reduced around the time that PNNs appear around PV-INs and dissolving PNNs was stated to improve memory (Kwok *et al.* 2011). For PV-INs to dampen excitation through release of GABA, they must be activated through an excitatory pathway. The neuronal pentraxins were shown to localise to excitatory inputs on PV-INs, and AMPA receptors were shown to localise to, but not directly interact with PNNs (Chang *et al.* 2010). In the NARP KO mouse model, the net excitatory drive onto these inhibitory synapses is reduced, resulting in hyper-excitability in the developing visual system (Gu *et al.* 2013).

1.6 Other AMPA receptor interactions

1.6.1 Cystine-knot AMPA receptor modulating protein 44 (CKAMP44)

CKAMP44 was identified as a novel AMPA receptor-associating protein by immunoprecipitation of AMPA receptor complexes (Von Engelhardt *et al.* 2010). This brain specific 44 kilodalton (kDa) (424 amino acid) protein consists of an intracellular region with a PDZ-like motif, a single-pass trans-membrane domain and an extracellular N-terminal region predicted to contain a cystine (cys)-knot (Von Engelhardt *et al.* 2010). Homologues of the *CKAMP44* (*sisha homolog 9*) gene show its conservation in chimpanzee, mouse, chicken, and zebrafish. The structurally undefined cys-knot region shows a cysteine pattern with homology to a fish-hunting snail neurotoxin termed con-ikot-ikot that is known to block AMPA receptor desensitisation (Walker *et al.* 2009; Von Engelhardt *et al.* 2010).

Recent studies show that over-expression of CKAMP44 accelerates AMPA receptor desensitisation and slows recovery rates after desensitisation (figure 1.10). The authors postulate that the slow recovery is due to CKAMP44 holding the AMPA receptor LBC cleft in a closed conformation (Von Engelhardt *et al.* 2010). Contrary to this, the con-ikot-ikot neurotoxin blocks desensitisation by mechanisms predicted to involve locking the LBC in an open conformation (Walker *et al.* 2009).

1.6.2 Transmembrane AMPA receptor regulatory protein-2 (TARP2/ Stargazin)

Lateral movement of AMPA receptors into synaptic density is modulated by TARPs (Straub *et al.* 2012). The TARP family of proteins are split into two classes (based on differences in their PDZ binding motifs), class I (γ -2 (TARP2, stargazin) γ -3, γ -4, and γ -8) and class II (γ -5 and γ -7) (Tomita *et al.* 2003; Kato *et al.* 2007), and have varied effects on AMPA receptors (Kato *et al.* 2008). The TARP2 protein was initially named Stargazin after a phenotype in the A/J inbred strain of mouse Stargazer which lacks the prototypical TARP Stargazin (Noebels *et al.*

1990). The mouse has an epileptic phenotype as well as a tendency to gaze upwards for several seconds, this vertical elevation of the head giving the mouse strain its name.

1.6.2.1 TARP involvement in AMPA receptor trafficking

TARP2 was first postulated to interact with the AMPA receptor via its intracellular C-terminal PDZ binding domain (Chen *et al.* 2000). TARP2 then binds directly to the PSD-95 protein, anchoring the receptor at the synapse (Schnell *et al.* 2002). Mature AMPA receptors cannot reach the cell surface (in neuronal or glial cells) if a TARP is not present (Tomita *et al.* 2003). GluA1 homo-tetramers in cells lacking TARPs have been shown to aggregate in the ER, but their surface presentation is enhanced by co-expression with TARP2 (Vandenberghe *et al.* 2005). In the hippocampus, TARP γ -8 was shown to modulate LTP by modulating cell surface receptor number (Rouach *et al.* 2005).

1.6.2.2 The effect of TARPs on AMPA receptor gating

TARPs can also modulate AMPA receptor gating, acting to slow the receptors rate of desensitisation (Cho *et al.* 2007) (figure 1.10). TARPs also increase GluA1s response to a potentiator (Tomita *et al.* 2006), although the efficacy is TARP dependent (Körber *et al.* 2007; Kott *et al.* 2007; Milstein *et al.* 2007). Modulation of AMPA receptors is controlled by the TARP present and TARP expression patterns are strictly segregated, thus having dramatic effects on systemic AMPA receptor activity (Tomita *et al.* 2003).

This figure has been removed due to
copyright restrictions.

Figure 1.10 Modulation of AMPA receptor kinetics by auxiliary subunits CKAMP44 and TARP2.
This figure has been removed due to copyright restrictions.

1.6.3 Analysis of putative AMPA receptor interaction partners

Recent publications have identified several putative novel AMPA receptor-interacting proteins by immuno-precipitation and mass spectrometry (Schwenk *et al.* 2012; Shanks *et al.* 2012). A novel interaction was claimed to occur between the AMPA receptor and the secreted protein Brorin. Prior to these publications, the neuronal pentraxins were the only secreted proteins known to interact with the AMPA receptor ectodomain and so another such interaction, with unknown effects on AMPA receptors required investigating further. Literature on Brorin is quite limited but upon its discovery in 2007 it was classified as a member of the chordin family of proteins (Koike *et al.* 2007). Chordin domains (also known as von Willebrand factor type C modules) are cysteine rich domains which typically contain 10 cysteine residues. Brorin was found to be a bone morphogenic protein (BMP) antagonist (Miwa *et al.* 2009). BMPs are secreted signalling molecules that are heavily involved in skeletal morphogenesis (Wan *et al.* 2005) and embryonic development (Von Bubnoff *et al.* 2001). Brorin gene expression is predominant in the brain and its ability to act as a promoter of neurogenesis was demonstrated in zebrafish knockdown experiments which impaired neurodevelopment (Koike *et al.* 2007). The potential physiological effects that Brorin may have on AMPA receptors are currently unknown.

1.7 MDGA1

This chapter has so far focussed on proteins expressed at excitatory synapses but one must appreciate that excitatory neurotransmission is modulated by inhibitory synapses. Inhibitory synaptic plasticity alters the excitatory/inhibitory balance and is postulated to have an important role in neural circuit refinement involved in experience-dependent learning (Castillo *et al.* 2011). A structurally uncharacterised protein known as meprin/A5-protein/PTPmu (MAM) domain-containing glycosylphosphatidylinositol (GPI) anchor protein 1 (MDGA1) was recently proposed to be the first identified negative regulator of inhibitory synapse formation. The mechanism of which was shown to involve an interaction of MDGA1 with neuroligin to inhibit the neuroligin-neuroligin trans-synaptic adhesion (Pettem *et al.* 2013), although the exact modes of action are difficult to predict in the absence of a solved protein structure.

MDGA1 was originally implicated in tumour progression when it was noted that the *MDGA1* gene (previously known as *GPIM*) was up-regulated in primary tumours of the lung, colon, breast, uterus and stomach (De Juan *et al.* 2002). The MDGA1 protein is tethered to the cell membrane by a GPI anchor. Its extracellular region, predicted to be highly glycosylated, contains a MAM domain and six immunoglobulin (Ig) domains, identifying it as a member of the immunoglobulin superfamily (IgSF) (De Juan *et al.* 2002).

1.7.1 MDGA1 in development

In chick development, MDGA1 was found to be up-regulated in spinal interneurons (Litwack *et al.* 2004) and dorsal root ganglia (DRG) in spinal cord dorsal (D1) regions (Fujimura *et al.* 2006). The MDGA1 interaction with these axon rich regions was predicted to involve its MAM domain (Fujimura *et al.* 2006).

MDGA1 and its protein homologue MDGA2 are noted to have restricted expression patterns throughout the CNS, and are postulated to have a role in neuronal cell migration

(Litwack *et al.* 2004). In mice, MDGA1 expression is detectable at postnatal stage P0, where expression seems specific to the projection neurons of the cortical plate (CP) which later differentiate into layers II and III of the neocortex (Takeuchi *et al.* 2006). MDGA1 RNAi was shown to prevent migration of neurons to the neocortical layers II and III (Takeuchi and O'Leary 2006). In another study MDGA1 deficient mice had no obvious phenotype but the radial migration of neurons was delayed. At P0, MDGA1 mutant neurons were undetectable in upper cortical layers, but by P7, mutant cell migration was comparable to WT controls (Ishikawa *et al.* 2011). Analysis *in vitro* suggested that the motility of cells expressing MDGA1 was increased and prevented cells from binding to collagen IV present in the basal lamina layer (Diaz-Lopez *et al.* 2010).

1.7.2 The neurexin- neuroligin interaction

During synaptogenesis post-synaptically located neuroligins mediate a trans-synaptic cell adhesion with pre-synaptic neurexins (Song *et al.* 1999; Scheiffele *et al.* 2000) (figure 1.11A). Neuroligins are type I transmembrane proteins with large ectodomains that contain a cholinesterase-like domain that is responsible for the trans-synaptic cell adhesion with neurexins (Südhof 2008). The cholinesterase domain also contains two alternative splice sites (SS) termed A and B and two EF-hand metal binding motifs which implicate the requirement for calcium in the heterologous interaction of neuroligins with neurexins (Tsigelny *et al.* 2000). The ectodomain also contains post-translational modification sites in the form of predicted multiple N-linked glycosylation sites and a juxtamembrane O-linked glycosylation regions (Lise *et al.* 2006). In addition to their adhesive roles in cell-cell contacts, the neuroligins are also postulated to have a role in cell signalling mechanisms via interactions of their intracellular COOH termini with PSD-95 (Irie *et al.* 1997). This potentially enables neuroligin to cluster with other PSD-95-interacting proteins such as NMDA receptors (Kornau *et al.* 1995) and TARP2 (Stargazin) (Bats *et al.* 2007) to form supra-molecular assemblies, thus implicating a role for neuroligins in synaptic signalling mechanisms (Sheng *et al.* 2001).

The neurexins are encoded for by three genes (*NRXN1*, 2 and 3) that are each controlled by two independent promoters which generate the longer (α) and shorter (β) isoforms (Ushkaryov *et al.* 1992; Missler *et al.* 1998). The α forms contain three repeats of a central epidermal growth factor (EGF)-like sequence that is flanked on either side by a laminin A, neurexin, sex-hormone-binding globulin (LNS) domain whereas the β form contains a single LNS domain (Ushkaryov *et al.* 1992). There are also a multitude of splice insertions, all which contributes to the vast number of neurexin variants currently known (Missler *et al.* 1998; Südhof 2008; Koehnke *et al.* 2010) that are postulated to generate specificity during synaptogenesis (Scheiffele *et al.* 2000).

Alternative splicing of neuroligins has been shown to affect their interactions with neurexins (Chih *et al.* 2006) and solved protein structures of a variety of splice isoforms and variants of neuroligins and neurexins reveal that the interaction is highly tuneable, dependant on splice variant present (Chen *et al.* 2008). Protein-protein interactions detected by surface plasmon resonance (SPR) analysis have demonstrated a 2-fold reduced affinity of neurexin-2 for NL1 containing a splice insertion at its 'site B' (Koehnke *et al.* 2010).

1.7.3 Neuroligins and neurexins in disease

A plethora of clinical cases are emerging that link both the neurexins and neuroligins to neurological diseases and disorders. These include the autism spectrum disorders, mental retardation and schizophrenia (Südhof 2008). In several cases, mutations at the genomic level can lead to translation of a truncated or functionally disadvantaged protein form. In one example, a mutation that introduced a stop codon in the ectodomain of neuroligin 4 was found in a family with X-linked mental retardation with or without autism. Analysis reveals that this mutation would result in the production of a truncated protein form that lacked the transmembrane domain and the protein would consequently be unable to dimerise (Laumonnier *et al.* 2004).

1.7.4 Interactions of MDGA1 with neuroligins

Recent studies have implied that MDGA1 can reduce inhibitory synaptic development through direct interactions with neuroligins (NLs) (figure 1.11B). MDGA1 knockdown experiments using short hairpin RNA (shRNA) show an increase in inhibitory synaptic density (Pettem *et al.* 2013) whereas over-expression of MDGA1 had the opposite effect. Addition of soluble neurexin had been previously shown to impair cells expressing neuroligin from undergoing pre-synaptic differentiation (Scheiffele *et al.* 2000).

MDGA1 was shown to interact with neuroligin and inhibit the interaction between neuroligin and neurexin (figure 1.11) (Lee *et al.* 2013; Pettem *et al.* 2013). The tightest interaction was found to be between the neuroligin NL2 isoform and MDGA1 (Pettem *et al.* 2013). It is not currently clear how the various neuroligin splice variants affect the binding to MDGA1. If the interaction of MDGA1 was affected by the neuroligin splice variant present it could be the mechanism to regulate the suppression of inhibitory synaptogenesis.

This figure has been removed due to
copyright restrictions.

Figure 1.11 Effects of MDGA1 on synaptogenesis.
This figure has been removed due to copyright restrictions.

1.8 Research aims

Synapse formation, differentiation and plasticity all involve modulation of pre- and post-synaptic elements. In synaptogenesis the pre- and post-synaptic compartments are initially differentiated in response to guidance cues present in their immediate environment (Sanes *et al.* 1999). Once dendritic and axonal contact is made, both cells undergo further morphological changes resulting in the formation of a synapse which then matures in response to activity.

The AMPA receptors are a well-studied subset of the iGluR family of proteins and a wealth of information on many aspects of their life cycle is available. A recent realisation important to appreciate is that the physiological AMPA receptor spends little time alone. Shuttle proteins are involved in transport to the cell surface, intracellular scaffold proteins anchor the receptor to the synapse and auxiliary proteins act to modulate the channel conductance. AMPA receptors are therefore involved in a large number of intra- and extra-cellular interactions, none of which are currently structurally characterised. The research presented in this thesis aims to give structural and biophysical insights into these interactions in the context of such synaptic events.

MDGA1 is postulated to have roles in cell migration, adhesion and synaptogenesis but little is currently known about its mechanisms of action specifically involving the neuroligins. Here, research aims to provide structural and biophysical analysis of such protein-protein interactions to gain insights into its potential role as the first identified negative regulator of inhibition (Pettem *et al.* 2013).

1.9 References

- Abad, M. A., M. Enguita, N. DeGregorio-Rocasolano, I. Ferrer and R. Trullas (2006). "Neuronal pentraxin 1 contributes to the neuronal damage evoked by amyloid-beta and is overexpressed in dystrophic neurites in Alzheimer's brain." *J Neurosci* **26**(49): 12735-12747.
- Amanuma, H. and Y. Anraku (1974). "Transport of sugars and amino acids in bacteria. XII. Substrate specificities of the branched chain amino acid binding proteins of *Escherichia coli*." *Journal of Biochemistry* **76**(6): 1165-1173.
- Appelbaum, L., G. Wang, T. Yokogawa, G. M. Skariah, S. J. Smith, P. Mourrain and E. Mignot (2010). "Circadian and Homeostatic Regulation of Structural Synaptic Plasticity in Hypocretin Neurons." *Neuron* **68**(1): 87-98.
- Arai, A. C., M. Kessler, G. Rogers and G. Lynch (2000). "Effects of the potent ampakine CX614 on hippocampal and recombinant AMPA receptors: interactions with cyclothiazide and GYKI 52466." *Mol Pharmacol* **58**(4): 802-813.
- Armstrong, N. and E. Gouaux (2000). "Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor: Crystal Structures of the GluR2 Ligand Binding Core." *Neuron* **28**(1): 165-181.
- Ayalon, G. and Y. Stern-Bach (2001). "Functional assembly of AMPA and kainate receptors is mediated by several discrete protein-protein interactions." *Neuron* **31**(1): 103-113.
- Bassell, G. J. and S. T. Warren (2008). "Fragile X Syndrome: Loss of Local mRNA Regulation Alters Synaptic Development and Function." *Neuron* **60**(2): 201-214.
- Bats, C., L. Groc and D. Choquet (2007). "The Interaction between Stargazin and PSD-95 Regulates AMPA Receptor Surface Trafficking." *Neuron* **53**(5): 719-734.
- Ben-Ari, Y., R. Khazipov, X. Leinekugel, O. Caillard and J. L. Gaiarsa (1997). "GABA(A), NMDA and AMPA receptors: A developmentally regulated 'menage a trois'." *Trends in Neurosciences* **20**(11): 523-529.
- Beneyto, M. and J. H. Meador-Woodruff (2006). "Lamina-specific abnormalities of AMPA receptor trafficking and signaling molecule transcripts in the prefrontal cortex in schizophrenia." *Synapse* **60**(8): 585-598.
- Berridge, M. J., M. D. Bootman and H. L. Roderick (2003). "Calcium signalling: Dynamics, homeostasis and remodelling." *Nature Reviews Molecular Cell Biology* **4**(7): 517-529.
- Bjartmar, L., A. D. Huberman, E. M. Ullian, R. C. Renteria, X. Liu, W. Xu, J. Prezioso, M. W. Susman, D. Stellwagen, C. C. Stokes, R. Cho, P. Worley, R. C. Malenka, S. Ball, N. S. Peachey, D. Copenhagen, B. Chapman, M. Nakamoto, B. A. Barres and M. S. Perin (2006). "Neuronal pentraxins mediate synaptic refinement in the developing visual system." *J Neurosci* **26**(23): 6269-6281.
- Bliss, T. V. and T. Lomo (1973). "Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path." *J Physiol* **232**(2): 331-356.
- Bliss, T. V. P. and G. L. Collingridge (1993). "A synaptic model of memory: Long-term potentiation in the hippocampus." *Nature* **361**(6407): 31-39.
- Blouin, A. M., T. C. Thannickal, P. F. Worley, J. M. Baraban, I. M. Reti and J. M. Siegel (2005). "Narp immunostaining of human hypocretin (orexin) neurons: loss in narcolepsy." *Neurology* **65**(8): 1189-1192.
- Bossy-Wetzell, E., R. Schwarzenbacher and S. A. Lipton (2004). "Molecular pathways to neurodegeneration." *Nat Med* **10**(9): S2-9.
- Bourgeois, J. P., P. S. Goldman-Rakic and P. Rakic (1994). "Synaptogenesis in the prefrontal cortex of rhesus monkeys." *Cerebral Cortex* **4**(1): 78-96.
- Castillo, P. E., C. Q. Chiu and R. C. Carroll (2011). "Long-term plasticity at inhibitory synapses." *Current Opinion in Neurobiology* **21**(2): 328-338.

- Cathala, L., N. B. Holderith, Z. Nusser, D. A. DiGregorio and S. G. Cull-Candy (2005). "Changes in synaptic structure underlie the developmental speeding of AMPA receptor-mediated EPSCs." Nature Neuroscience **8**(10): 1310-1318.
- Celio, M. R. and I. Blümcke (1994). "Perineuronal nets - A specialized form of extracellular matrix in the adult nervous system." Brain Research Reviews **19**(1): 128-145.
- Chang, M. C., J. M. Park, K. A. Pelkey, H. L. Grabenstatter, D. Xu, D. J. Linden, T. P. Sutula, C. J. McBain and P. F. Worley (2010). "Narp regulates homeostatic scaling of excitatory synapses on parvalbumin-expressing interneurons." Nat Neurosci **13**(9): 1090-1097.
- Chen, G. Q., C. Cul, M. L. Mayer and E. Gouaux (1999). "Functional characterization of a potassium-selective prokaryotic glutamate receptor." Nature **402**(6763): 817-821.
- Chen, L., D. M. Chetkovich, R. S. Petralia, N. T. Sweeney, Y. Kawasaki, R. J. Wenthold, D. S. Brecht and R. A. Nicoll (2000). "Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms." Nature **408**(6815): 936-943.
- Chen, X., H. Liu, A. H. Shim, P. J. Focia and X. He (2008). "Structural basis for synaptic adhesion mediated by neuroligin-neurexin interactions." Nat Struct Mol Biol **15**(1): 50-56.
- Chih, B., L. Gollan and P. Scheiffele (2006). "Alternative Splicing Controls Selective Trans-Synaptic Interactions of the Neuroligin-Neurexin Complex." Neuron **51**(2): 171-178.
- Chiu, J., R. DeSalle, H. M. Lam, L. Meisel and G. Coruzzi (1999). "Molecular evolution of glutamate receptors: A primitive signaling mechanism that existed before plants and animals diverged." Molecular Biology and Evolution **16**(6): 826-838.
- Cho, C. H., F. St-Gelais, W. Zhang, S. Tomita and J. R. Howe (2007). "Two Families of TARP Isoforms that Have Distinct Effects on the Kinetic Properties of AMPA Receptors and Synaptic Currents." Neuron **55**(6): 890-904.
- Cho, R. W., J. M. Park, S. B. E. Wolff, D. Xu, C. Hopf, J.-a. Kim, R. C. Reddy, R. S. Petralia, M. S. Perin, D. J. Linden and P. F. Worley (2008). "mGluR1/5-Dependent Long-Term Depression Requires the Regulated Ectodomain Cleavage of Neuronal Pentraxin NPR by TACE." Neuron **57**(6): 858-871.
- Clayton, A., C. Siebold, R. J. C. Gilbert, G. C. Sutton, K. Harlos, R. A. J. McIlhinney, E. Y. Jones and A. R. Aricescu (2009). "Crystal Structure of the GluR2 Amino-Terminal Domain Provides Insights into the Architecture and Assembly of Ionotropic Glutamate Receptors." Journal of Molecular Biology **392**(5): 1125-1132.
- Collingridge, G. L., J. T. R. Isaac and T. W. Yu (2004). "Receptor trafficking and synaptic plasticity." Nature Reviews Neuroscience **5**(12): 952-962.
- Collingridge, G. L., R. W. Olsen, J. Peters and M. Spedding (2009). "A nomenclature for ligand-gated ion channels." Neuropharmacology **56**(1): 2-5.
- Cowan, W. M. and E. R. Kandel (2003). Synapses: Chapter 1. W. M. S. Cowan, T.C.; Stevens, C.F., The Johns Hopkins University Press 792.
- Danbolt, N. C. (2001). "Glutamate uptake." Progress in Neurobiology **65**(1): 1-105.
- Davies, J., R. H. Evans, A. A. Francis and J. C. Watkins (1979). "Excitatory amino acid receptors and synaptic excitation in the mammalian central nervous system." Journal de Physiologie **75**(6): 641-644.
- De Juan, C., P. Iniesta, R. Gonzalez-Quevedo, A. Moran, A. Sanchez-Pernaute, A. J. Torres, J. L. Balibrea, E. Diaz-Rubio, J. Cruces and M. Benito (2002). "Genomic organization of a novel glycosylphosphatidylinositol MAM gene expressed in human tissues and tumors." Oncogene **21**(19): 3089-3094.
- de Ligt, J., M. H. Willemsen, B. W. van Bon, T. Kleefstra, H. G. Yntema, T. Kroes, A. T. Vulto-van Silfhout, D. A. Koolen, P. de Vries, C. Gilissen, M. del Rosario, A. Hoischen, H. Scheffer, B. B. de Vries, H. G. Brunner, J. A. Veltman and L. E. Vissers (2012). "Diagnostic exome sequencing in persons with severe intellectual disability." N Engl J Med **367**(20): 1921-1929.
- Diaz-Lopez, A., P. Iniesta, A. Moran, P. Ortega, T. Fernandez-Marcelo, A. Sanchez-Pernaute, A. J. Torres, M. Benito and C. De Juan (2010). "Expression of Human MDGA1 Increases

- Cell Motility and Cell-Cell Adhesion and Reduces Adhesion to Extracellular Matrix Proteins in MDCK Cells." Cancer Microenviron **4**(1): 23-32.
- Dodds, D. N. C., I. A. Omeis, S. J. Cushman, J. A. Helms and M. S. Perin (1997). "Neuronal Pentraxin Receptor, a Novel Putative Integral Membrane Pentraxin That Interacts with Neuronal Pentraxin 1 and 2 and Taipoxin-associated Calcium-binding Protein 49." Journal of Biological Chemistry **272**(34): 21488-21494.
- Dutta, A., I. H. Shrivastava, M. Sukumaran, I. H. Greger and I. Bahar (2012). "Comparative dynamics of NMDA- and AMPA-glutamate receptor N-terminal domains." Structure **20**(11): 1838-1849.
- Emsley, J., H. E. White, B. P. O'Hara, G. Oliva, N. Srinivasan, I. J. Tickle, T. L. Blundell, M. B. Pepys and S. P. Wood (1994). "Structure of pentameric human serum amyloid P component." Nature **367**(6461): 338-345.
- Fiala, J. C., J. Spacek and K. M. Harris (2002). "Dendritic spine pathology: Cause or consequence of neurological disorders?" Brain Research Reviews **39**(1): 29-54.
- Fujimura, Y., M. Iwashita, F. Matsuzaki and T. Yamamoto (2006). "MDGA1, an IgSF molecule containing a MAM domain, heterophilically associates with axon- and muscle-associated binding partners through distinct structural domains." Brain Research **1101**(1): 12-19.
- Furukawa, H., S. K. Singh, R. Mancusso and E. Gouaux (2005). "Subunit arrangement and function in NMDA receptors." Nature **438**(7065): 185-192.
- Garlanda, C., B. Bottazzi, A. Bastone and A. Mantovani (2005). Pentraxins at the crossroads between innate immunity, inflammation, matrix deposition, and female fertility. **23**: 337-366.
- Gibson, J. R., M. Belerlein and B. W. Connors (1999). "Two networks of electrically coupled inhibitory neurons in neocortex." Nature **402**(6757): 75-79.
- Gordon, S. L., R. E. Leube and M. A. Cousin (2011). "Synaptophysin Is Required for Synaptobrevin Retrieval during Synaptic Vesicle Endocytosis." Journal of Neuroscience **31**(39): 14032-14036.
- Greger, I. H. and J. A. Esteban (2007). "AMPA receptor biogenesis and trafficking." Current Opinion in Neurobiology **17**(3): 289-297.
- Gu, Y., S. Huang, Michael C. Chang, P. Worley, A. Kirkwood and Elizabeth M. Quinlan (2013). "Obligatory Role for the Immediate Early Gene NARP in Critical Period Plasticity." Neuron **79**(2): 335-346.
- Hansen, K. B., P. Naur, N. L. Kurtkaya, A. S. Kristensen, M. Gajhede, J. S. Kastrup and S. F. Traynelis (2009). "Modulation of the dimer interface at ionotropic glutamate-like receptor $\delta 2$ by D-serine and extracellular calcium." Journal of Neuroscience **29**(4): 907-917.
- Hayashi, Y., S. H. Shi, J. A. Esteban, A. Piccini, J. C. Poncer and R. Malinow (2000). "Driving AMPA receptors into synapses by LTP and CaMKII: Requirement for GluR1 and PDZ domain interaction." Science **287**(5461): 2262-2267.
- Hebb, D. O. (1949). *The Organization of Behavior*, New York: Wiley & Sons.
- Hee Jung, C., J. Xia, R. H. Scannevin, X. Zhang and R. L. Huganir (2000). "Phosphorylation of the AMPA receptor subunit GluR2 differentially regulates its interaction with PDZ domain-containing proteins." Journal of Neuroscience **20**(19): 7258-7267.
- Henley, J. M., E. A. Barker and O. O. Glebov (2011). "Routes, destinations and delays: Recent advances in AMPA receptor trafficking." Trends Neurosci **34**(5): 258-268.
- Honore, T., P. Krogsgaard-Larsen, J. J. Hansen and J. Lauridsen (1981). "Glutamate and aspartate agonists structurally related to ibotenic acid." Molecular and Cellular Biochemistry **38**(1): 123-128.
- Horning, M. S. and M. L. Mayer (2004). "Regulation of AMPA Receptor Gating by Ligand Binding Core Dimers." Neuron **41**(3): 379-388.
- Hsu, Y.-C. and M. S. Perin (1995). "Human Neuronal Pentraxin II (NPTX2): Conservation, Genomic Structure, and Chromosomal Localization." Genomics **28**(2): 220-227.
- Hughes, B. W., L. L. Kusner and H. J. Kaminski (2006). "Molecular architecture of the neuromuscular junction." Muscle and Nerve **33**(4): 445-461.

- Irani, S. R., C. G. Bien and B. Lang (2011). "Autoimmune epilepsies." Current Opinion in Neurology **24**(2): 146-153.
- Irie, M., Y. Hata, M. Takeuchi, K. Ichtchenko, A. Toyoda, K. Hirao, Y. Takai, T. W. Rosahl and T. C. Südhof (1997). "Binding of neuroligins to PSD-95." Science **277**(5331): 1511-1515.
- Ishikawa, T., N. Gotoh, C. Murayama, T. Abe, M. Iwashita, F. Matsuzaki, T. Suzuki and T. Yamamoto (2011). "IgSF molecule MDGA1 is involved in radial migration and positioning of a subset of cortical upper-layer neurons." Dev Dyn **240**(1): 96-107.
- Ito, I., S. Tanabe, A. Kohda and H. Sugiyama (1990). "Allosteric potentiation of quisqualate receptors by a nootropic drug aniracetam." Journal of Physiology **424**: 533-543.
- Janovjak, H., G. Sandoz and E. Y. Isacoff (2011). "A modern ionotropic glutamate receptor with a K(+) selectivity signature sequence." Nat Commun **2**(232).
- Jin, R., S. Clark, A. M. Weeks, J. T. Dudman, E. Gouaux and K. M. Partin (2005). "Mechanism of positive allosteric modulators acting on AMPA receptors." J Neurosci **25**(39): 9027-9036.
- Jin, R., S. K. Singh, S. Gu, H. Furukawa, A. I. Sobolevsky, J. Zhou, Y. Jin and E. Gouaux (2009). "Crystal structure and association behaviour of the GluR2 amino-terminal domain." EMBO Journal **28**(12): 1812-1823.
- Johnson, J. W. and P. Ascher (1987). "Glycine potentiates the NMDA response in cultured mouse brain neurons." Nature **325**(6104): 529-531.
- Kato, A. S., E. R. Siuda, E. S. Nisenbaum and D. S. Brecht (2008). "AMPA Receptor Subunit-Specific Regulation by a Distinct Family of Type II TARPs." Neuron **59**(6): 986-996.
- Kato, A. S., W. Zhou, A. D. Milstein, M. D. Knierman, E. R. Siuda, J. E. Dotzlaw, H. Yu, J. E. Hale, E. S. Nisenbaum, R. A. Nicoll and D. S. Brecht (2007). "New transmembrane AMPA receptor regulatory protein isoform, γ -7, differentially regulates AMPA receptors." Journal of Neuroscience **27**(18): 4969-4977.
- Kessels, H. W. and R. Malinow (2009). "Synaptic AMPA receptor plasticity and behavior." Neuron **61**(3): 340-350.
- Kilpatrick, D. C. (2002). "Animal lectins: a historical introduction and overview." Biochimica et Biophysica Acta (BBA) - General Subjects **1572**(2-3): 187-197.
- Kirkpatrick, L. L., M. M. Matzuk, D. C. Dodds and M. S. Perin (2000). "Biochemical interactions of the neuronal pentraxins. Neuronal pentraxin (NP) receptor binds to taipoxin and taipoxin-associated calcium-binding protein 49 via NP1 and NP2." J Biol Chem **275**(23): 17786-17792.
- Koehnke, J., P. S. Katsamba, G. Ahlsen, F. Bahna, J. Vendome, B. Honig, L. Shapiro and X. Jin (2010). "Splice Form Dependence of β -Neurexin/Neuroligin Binding Interactions." Neuron **67**(1): 61-74.
- Koiike, N., Y. Kassai, Y. Kouta, H. Miwa, M. Konishi and N. Itoh (2007). "Brorin, a Novel Secreted Bone Morphogenetic Protein Antagonist, Promotes Neurogenesis in Mouse Neural Precursor Cells." Journal of Biological Chemistry **282**(21): 15843-15850.
- Körber, C., M. Werner, S. Kott, Z. L. Ma and M. Hollmann (2007). "The transmembrane AMPA receptor regulatory protein γ 4 is a more effective modulator of AMPA receptor function than stargazin (γ 2)." Journal of Neuroscience **27**(31): 8442-8447.
- Kornau, H. C., L. T. Schenker, M. B. Kennedy and P. H. Seeburg (1995). "Domain interaction between NMDA receptor subunits and the postsynaptic density protein PSD-95." Science **269**(5231): 1737-1740.
- Kott, S., M. Werner, C. Körber and M. Hollmann (2007). "Electrophysiological properties of AMPA receptors are differentially modulated depending on the associated member of the TARP family." Journal of Neuroscience **27**(14): 3780-3789.
- Kuner, T., P. H. Seeburg and H. R. Guy (2003). "A common architecture for K⁺ channels and ionotropic glutamate receptors?" Trends Neurosci **26**(1): 27-32.
- Kwok, J. C., G. Dick, D. Wang and J. W. Fawcett (2011). "Extracellular matrix and perineuronal nets in CNS repair." Developmental Neurobiology **71**(11): 1073-1089.
- Lam, H. M., J. Chiu, M. H. Hsieh, L. Meisel, I. C. Oliveira, M. Shin and G. Coruzzi (1998). "Glutamate-receptor genes in plants [5]." Nature **396**(6707): 125-126.

- Laumonnier, F., F. Bonnet-Brilhault, M. Gomot, R. Blanc, A. David, M. P. Moizard, M. Raynaud, N. Ronce, E. Lemonnier, P. Calvas, B. Laudier, J. Chelly, J. P. Fryns, H. H. Ropers, B. C. J. Hamel, C. Andres, C. Barthélémy, C. Moraine and S. Briault (2004). "X-Linked Mental Retardation and Autism Are Associated with a Mutation in the NLGN4 Gene, a Member of the Neuroligin Family." American Journal of Human Genetics **74**(3): 552-557.
- Lee, H. K., M. Barbarosie, K. Kameyama, M. F. Bear and R. L. Huganir (2000). "Regulation of distinct AMPA receptor phosphorylation sites during bidirectional synaptic plasticity." Nature **405**(6789): 955-959.
- Lee, K., Y. Kim, S.-J. Lee, Y. Qiang, D. Lee, H. W. Lee, H. Kim, H. S. Je, T. C. Südhof and J. Ko (2013). "MDGAs interact selectively with neuroligin-2 but not other neuroligins to regulate inhibitory synapse development." Proceedings of the National Academy of Sciences **110**(1): 336-341.
- Lin, D. T., Y. Makino, K. Sharma, T. Hayashi, R. Neve, K. Takamiya and R. L. Huganir (2009). "Regulation of AMPA receptor extrasynaptic insertion by 4.1N, phosphorylation and palmitoylation." Nature Neuroscience **12**(7): 879-887.
- Lise, M. F. and A. El-Husseini (2006). "The neuroligin and neurexin families: from structure to function at the synapse." Cell Mol Life Sci **63**(16): 1833-1849.
- Litwack, E. D., R. Babey, R. Buser, M. Gesemann and D. D. M. O'Leary (2004). "Identification and characterization of two novel brain-derived immunoglobulin superfamily members with a unique structural organization." Molecular and Cellular Neuroscience **25**(2): 263-274.
- Lomash, S., S. Chittori, P. Brown and Mark L. Mayer (2013). "Anions Mediate Ligand Binding in Adineta vaga Glutamate Receptor Ion Channels." Structure **21**(3): 414-425.
- Lomeli, H., J. Mosbacher, T. Melcher, T. Hoyer, J. R. P. Geiger, T. Kuner, H. Monyer, M. Higuchi, A. Bach and P. H. Seeburg (1994). "Control of kinetic properties of AMPA receptor channels by nuclear RNA editing." Science **266**(5191): 1709-1713.
- Lonze, B. E. and D. D. Ginty (2002). "Function and regulation of CREB family transcription factors in the nervous system." Neuron **35**(4): 605-623.
- Lu, J., L. L. Marnell, K. D. Marjon, C. Mold, T. W. Du Clos and P. D. Sun (2008). "Structural recognition and functional activation of FcγR by innate pentraxins." Nature **456**(7224): 989-992.
- Lucic, V., T. Yang, G. Schweikert, F. Forster and W. Baumeister (2005). "Morphological characterization of molecular complexes present in the synaptic cleft." Structure **13**(3): 423-434.
- Lüscher, C., H. Xia, E. C. Beattie, R. C. Carroll, M. Von Zastrow, R. C. Malenka and R. A. Nicoll (1999). "Role of AMPA receptor cycling in synaptic transmission and plasticity." Neuron **24**(3): 649-658.
- Lynch, G. (2006). "Glutamate-based therapeutic approaches: Ampakines." Current Opinion in Pharmacology **6**(1 SPEC. ISS.): 82-88.
- Malenka, R. C. and R. A. Nicoll (1999). "Long-term potentiation - A decade of progress?" Science **285**(5435): 1870-1874.
- Malinow, R. and R. C. Malenka (2002). AMPA receptor trafficking and synaptic plasticity. **25**: 103-126.
- Martin, S. J., P. D. Grimwood and R. G. M. Morris (2000). Synaptic plasticity and memory: An evaluation of the hypothesis. **23**: 649-711.
- Mayer, M. L. (2006). "Glutamate receptors at atomic resolution." Nature **440**(7083): 456-462.
- Milstein, A. D., W. Zhou, S. Karimzadegan, D. S. Bredt and R. A. Nicoll (2007). "TARP Subtypes Differentially and Dose-Dependently Control Synaptic AMPA Receptor Gating." Neuron **55**(6): 905-918.
- Missler, M., R. Fernandez-Chacon and T. C. Südhof (1998). "The making of neurexins." Journal of Neurochemistry **71**(4): 1339-1347.

- Miwa, H., A. Miyake, Y. Kouta, A. Shimada, Y. Yamashita, Y. Nakayama, H. Yamauchi, M. Konishi and N. Itoh (2009). "A novel neural-specific BMP antagonist, Brorin-like, of the Chordin family." FEBS Letters **583**(22): 3643-3648.
- Naur, P., K. B. Hansen, A. S. Kristensen, S. M. Dravid, D. S. Pickering, L. Olsen, B. Vestergaard, J. Egebjerg, M. Gajhede, S. F. Traynelis and J. S. Kastrup (2007). "Ionotropic glutamate-like receptor $\delta 2$ binds D-serine and glycine." Proceedings of the National Academy of Sciences of the United States of America **104**(35): 14116-14121.
- Nicoll, R. A., R. C. Malenka and J. A. Kauer (1990). "Functional comparison of neurotransmitter receptor subtypes in mammalian central nervous system." Physiological Reviews **70**(2): 513-565.
- Noebels, J. L., X. Qiao, R. T. Bronson, C. Spencer and M. T. Davisson (1990). "Stargazer: A new neurological mutant on chromosome 15 in the mouse with prolonged cortical seizures." Epilepsy Research **7**(2): 129-135.
- O'Brien, R., D. Xu, R. Mi, X. Tang, C. Hopf and P. Worley (2002). "Synaptically targeted narp plays an essential role in the aggregation of AMPA receptors at excitatory synapses in cultured spinal neurons." J Neurosci **22**(11): 4487-4498.
- O'Brien, R. J., D. Xu, R. S. Petralia, O. Steward, R. L. Huganir and P. Worley (1999). "Synaptic clustering of AMPA receptors by the extracellular immediate-early gene product Narp." Neuron **23**(2): 309-323.
- Oh, M. C., V. A. Derkach, E. S. Guire and T. R. Soderling (2006). "Extrasynaptic membrane trafficking regulated by GluR1 serine 845 phosphorylation primes AMPA receptors for long-term potentiation." Journal of Biological Chemistry **281**(2): 752-758.
- Osmand, A. P., B. Friedenson, H. Gewurz, R. H. Painter, T. Hofmann and E. Shelton (1977). "Characterization of C reactive protein and the complement subcomponent C1t at homologous proteins displaying cyclic pentameric symmetry (pentraxins)." Proceedings of the National Academy of Sciences of the United States of America **74**(2): 739-743.
- Ozawa, S. (1996). "Permeation of calcium through glutamate receptor channels." Seminars in the Neurosciences **8**(5): 261-269.
- Ozawa, S., H. Kamiya and K. Tsuzuki (1998). "Glutamate receptors in the mammalian central nervous system." Progress in Neurobiology **54**(5): 581-618.
- Pacchioni, A. M., J. Vallone, P. F. Worley and P. W. Kalivas (2009). "Neuronal pentraxins modulate cocaine-induced neuroadaptations." J Pharmacol Exp Ther **328**(1): 183-192.
- Palmer, C. L., W. Lim, P. G. Hastie, M. Toward, V. I. Korolchuk, S. A. Burbidge, G. Banting, G. L. Collingridge, J. T. Isaac and J. M. Henley (2005). "Hippocalcin functions as a calcium sensor in hippocampal LTD." Neuron **47**(4): 487-494.
- Panchenko, V. A., C. R. Glasser and M. L. Mayer (2001). "Structural similarities between glutamate receptor channels and K⁺ channels examined by scanning mutagenesis." Journal of General Physiology **117**(4): 345-359.
- Pasternack, A., S. K. Coleman, A. Joupila, D. G. Mottershead, M. Lindfors, M. Pasternack and K. Keinänen (2002). "Alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor channels lacking the N-terminal domain." J Biol Chem **277**(51): 49662-49667.
- Petrini, E. M., J. Lu, L. Cognet, B. Lounis, M. D. Ehlers and D. Choquet (2009). "Endocytic trafficking and recycling maintain a pool of mobile surface AMPA receptors required for synaptic potentiation." Neuron **63**(1): 92-105.
- Pettem, K. L., D. Yokomaku, H. Takahashi, Y. Ge and A. M. Craig (2013). "Interaction between autism-linked MDGAs and neuroligins suppresses inhibitory synapse development." The Journal of Cell Biology **200**(3): 321-336.
- Reti, I. M. and J. M. Baraban (2003). "Opiate withdrawal induces narp in the extended amygdala." Neuropsychopharmacology **28**(9): 1606-1613.
- Roche, K. W., R. J. O'Brien, A. L. Mammen, J. Bernhardt and R. L. Huganir (1996). "Characterization of multiple phosphorylation sites on the AMPA receptor GluR1 subunit." Neuron **16**(6): 1179-1188.

- Rouach, N., K. Byrd, R. S. Petralia, G. M. Elias, H. Adesnik, S. Tomita, S. Karimzadegan, C. Kealey, D. S. Bredt and R. A. Nicoll (2005). "TARP γ -8 controls hippocampal AMPA receptor number, distribution and synaptic plasticity." Nature Neuroscience **8**(11): 1525-1533.
- Sanes, J. R. and M. Yamagata (1999). "Formation of lamina-specific synaptic connections." Current Opinion in Neurobiology **9**(1): 79-87.
- Santos, S. D., A. L. Carvalho, M. V. Caldeira and C. B. Duarte (2009). "Regulation of AMPA receptors and synaptic plasticity." Neuroscience **158**(1): 105-125.
- Scheiffele, P., J. Fan, J. Choih, R. Fetter and T. Serafini (2000). "Neuroigin Expressed in Nonneuronal Cells Triggers Presynaptic Development in Contacting Axons." Cell **101**(6): 657-669.
- Schlimgen, A. K., J. A. Helms, H. Vogel and M. S. Perin (1995). "Neuronal pentraxin, a secreted protein with homology to acute phase proteins of the immune system." Neuron **14**(3): 519-526.
- Schnell, E., M. Sizemore, S. Karimzadegan, L. Chen, D. S. Bredt and R. A. Nicoll (2002). "Direct interactions between PSD-95 and stargazin control synaptic AMPA receptor number." Proceedings of the National Academy of Sciences of the United States of America **99**(21): 13902-13907.
- Schwenk, J., N. Harmel, A. Brechet, G. Zolles, H. Berkefeld, C. S. Muller, W. Bildl, D. Baehrens, B. Huber, A. Kulik, N. Klocker, U. Schulte and B. Fakler (2012). "High-resolution proteomics unravel architecture and molecular diversity of native AMPA receptor complexes." Neuron **74**(4): 621-633.
- Seifert, G., M. Zhou, D. Dietrich, T. B. Schumacher, A. Dybek, T. Weiser, M. Wienrich, D. Wilhelm and C. Steinhäuser (2000). "Developmental regulation of AMPA-receptor properties in CA1 pyramidal neurons of rat hippocampus." Neuropharmacology **39**(6): 931-942.
- Shanks, N. F., J. N. Savas, T. Maruo, O. Cais, A. Hirao, S. Oe, A. Ghosh, Y. Noda, I. H. Greger, J. R. Yates and T. Nakagawa (2012). "Differences in AMPA and Kainate Receptor Interactomes Facilitate Identification of AMPA Receptor Auxiliary Subunit GSG1L." Cell Reports **1**(6): 590-598.
- Sheng, M. and C. Sala (2001). PDZ domains and the organization of supramolecular complexes. **24**: 1-29.
- Sherrington, C. S. (1897). A Textbook of Physiology. London, Macmillan.
- Shrive, A. K., A. M. Metcalfe, J. R. Cartwright and T. J. Greenhough (1999). "C-reactive protein and SAP-like pentraxin are both present in Limulus polyphemus haemolymph: crystal structure of Limulus SAP." J Mol Biol **290**(5): 997-1008.
- Sia, G. M., J. C. Beique, G. Rumbaugh, R. Cho, P. F. Worley and R. L. Huganir (2007). "Interaction of the N-terminal domain of the AMPA receptor GluR4 subunit with the neuronal pentraxin NP1 mediates GluR4 synaptic recruitment." Neuron **55**(1): 87-102.
- Sobolevsky, A. I., M. P. Rosconi and E. Gouaux (2009). "X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor." Nature **462**(7274): 745-756.
- Sommer, B., K. Keinänen, T. A. Verdoorn, W. Wisden, N. Burnashev, A. Herb, M. Kohler, T. Takagi, B. Sakmann and P. H. Seeburg (1990). "Flip and flop: A cell-specific functional switch in glutamate-operated channels of the CNS." Science **249**(4976): 1580-1585.
- Song, J. Y., K. Ichchenko, T. C. Südhof and N. Brose (1999). "Neuroigin 1 is a postsynaptic cell-adhesion molecule of excitatory synapses." Proceedings of the National Academy of Sciences of the United States of America **96**(3): 1100-1105.
- Staubli, U., G. Rogers and G. Lynch (1994). "Facilitation of glutamate receptors enhances memory." Proceedings of the National Academy of Sciences **91**(2): 777-781.
- Straub, C. and S. Tomita (2012). "The regulation of glutamate receptor trafficking and function by TARPs and other transmembrane auxiliary subunits." Curr Opin Neurobiol **22**(3): 488-495.

- Strauss, J., S. McGregor, N. Freeman, A. Tiwari, C. J. George, M. Kovacs and J. L. Kennedy (2012). "Association study of early-onset mood disorders and suicide attempt." Psychiatry Res **197**(1-2): 49-54.
- Südhof, T. C. (2008). "Neurologins and neuroligins link synaptic function to cognitive disease." Nature **455**(7215): 903-911.
- Sukumaran, M., M. Rossmann, I. Shrivastava, A. Dutta, I. Bahar and I. H. Greger (2011). "Dynamics and allosteric potential of the AMPA receptor N-terminal domain." EMBO J **30**(5): 972-982.
- Swanson, G. T., S. K. Kamboj and S. G. Cull-Candy (1997). "Single-channel properties of recombinant AMPA receptors depend on RNA editing, splice variation, and subunit composition." Journal of Neuroscience **17**(1): 58-69.
- Takamori, S., M. Holt, K. Stenius, E. A. Lemke, M. Grønborg, D. Riedel, H. Urlaub, S. Schenck, B. Brügger, P. Ringler, S. A. Müller, B. Rammner, F. Gräter, J. S. Hub, B. L. De Groot, G. Mieskes, Y. Moriyama, J. Klingauf, H. Grubmüller, J. Heuser, F. Wieland and R. Jahn (2006). "Molecular Anatomy of a Trafficking Organelle." Cell **127**(4): 831-846.
- Takeuchi, A. and D. D. O'Leary (2006). "Radial migration of superficial layer cortical neurons controlled by novel Ig cell adhesion molecule MDGA1." J Neurosci **26**(17): 4460-4464.
- Tang, C., Q. Shi, A. Katchman and G. Lynch (1991). "Modulation of the time course of fast EPSCs and glutamate channel kinetics by aniracetam." Science **254**(5029): 288-290.
- Ting, J. T., B. G. Kelley, T. J. Lambert, D. G. Cook and J. M. Sullivan (2007). "Amyloid precursor protein overexpression depresses excitatory transmission through both presynaptic and postsynaptic mechanisms." Proceedings of the National Academy of Sciences of the United States of America **104**(1): 353-358.
- Tomita, S., L. Chen, Y. Kawasaki, R. S. Petralia, R. J. Wenthold, R. A. Nicoll and D. S. Brecht (2003). "Functional studies and distribution define a family of transmembrane AMPA receptor regulatory proteins." Journal of Cell Biology **161**(4): 805-816.
- Tomita, S., M. Sekiguchi, K. Wada, R. A. Nicoll and D. S. Brecht (2006). "Stargazin controls the pharmacology of AMPA receptor potentiators." Proceedings of the National Academy of Sciences of the United States of America **103**(26): 10064-10067.
- Tomiyama, M., R. Rodríguez-Puertas, R. Cortés, A. Pazos, J. M. Palacios and G. Mengod (2002). "Flip and flop splice variants of AMPA receptor subunits in the spinal cord of amyotrophic lateral sclerosis." Synapse **45**(4): 245-249.
- Tsigelny, I., I. N. Shindyalov, P. E. Bourne, T. C. Südhof and P. Taylor (2000). "Common EF-hand motifs in cholinesterases and neuroligins suggest a role for Ca²⁺ binding in cell surface associations." Protein Sci **9**(1): 180-185.
- Tsui, C. C., N. G. Copeland, D. J. Gilbert, N. A. Jenkins, C. Barnes and P. F. Worley (1996). "Narp, a novel member of the pentraxin family, promotes neurite outgrowth and is dynamically regulated by neuronal activity." Journal of Neuroscience **16**(8): 2463-2478.
- Ushkaryov, Y. A., A. G. Petrenko, M. Geppert and T. C. Südhof (1992). "Neurexins: Synaptic cell surface proteins related to the α -latrotoxin receptor and laminin." Science **257**(5066): 50-56.
- Vandenberghe, W., R. A. Nicoll and D. S. Brecht (2005). "Interaction with the unfolded protein response reveals a role for stargazin in biosynthetic AMPA receptor transport." Journal of Neuroscience **25**(5): 1095-1102.
- Von Bubnoff, A. and K. W. Y. Cho (2001). "Intracellular BMP signaling regulation in vertebrates: Pathway or network?" Developmental Biology **239**(1): 1-14.
- Von Engelhardt, J., V. Mack, R. Sprengel, N. Kavenstock, K. W. Li, Y. Stern-Bach, A. B. Smit, P. H. Seeburg and H. Monyer (2010). "CKAMP44: A brain-specific protein attenuating short-term synaptic plasticity in the dentate gyrus." Science **327**(5972): 1518-1522.
- Walker, C. S., S. Jensen, M. Ellison, J. A. Matta, W. Y. Lee, J. S. Imperial, N. Duclos, P. J. Brockie, D. M. Madsen, J. T. Isaac, B. Olivera and A. V. Maricq (2009). "A novel Conus snail polypeptide causes excitotoxicity by blocking desensitization of AMPA receptors." Curr Biol **19**(11): 900-908.

- Wan, M. and X. Cao (2005). "BMP signaling in skeletal development." Biochemical and Biophysical Research Communications **328**(3): 651-657.
- Wegrzyn, J. L., S. J. Bark, L. Funkelstein, C. Mosier, A. Yap, P. Kazemi-Esfarjani, A. R. La Spada, C. Sigurdson, D. T. O'Connor and V. Hook (2010). "Proteomics of dense core secretory vesicles reveal distinct protein categories for secretion of neuroeffectors for cell-cell communication." J Proteome Res **9**(10): 5002-5024.
- Wolf, M. E., X. Sun, S. Mangiavacchi and S. Z. Chao (2004). "Psychomotor stimulants and neuronal plasticity." Neuropharmacology **1**: 61-79.
- Wollmuth, L. P. and A. I. Sobolevsky (2004). "Structure and gating of the glutamate receptor ion channel." Trends Neurosci **27**(6): 321-328.
- Wu, Y., A. C. Arai, G. Rumbaugh, A. K. Srivastava, G. Turner, T. Hayashi, E. Suzuki, Y. Jiang, L. Zhang, J. Rodriguez, J. Boyle, P. Tarpey, F. L. Raymond, J. Nevelsteen, G. Froyen, M. Stratton, A. Futreal, J. Gecz, R. Stevenson, C. E. Schwartz, D. Valle, R. L. Haganir and T. Wang (2007). "Mutations in ionotropic AMPA receptor 3 alter channel properties and are associated with moderate cognitive impairment in humans." Proc Natl Acad Sci U S A **104**(46): 18163-18168.
- Xu, D., C. Hopf, R. Reddy, R. W. Cho, L. Guo, A. Lanahan, R. S. Petralia, R. J. Wenthold, R. J. O'Brien and P. Worley (2003). "Narp and NP1 form heterocomplexes that function in developmental and activity-dependent synaptic plasticity." Neuron **39**(3): 513-528.
- Yao, G., Y. Zong, S. Gu, J. Zhou, H. Xu, I. I. Mathews and R. Jin (2011). "Crystal structure of the glutamate receptor GluA1 N-terminal domain." Biochemical Journal **438**(2): 255-263.
- Yin, G. N., H. W. Lee, J. Y. Cho and K. Suk (2009). "Neuronal pentraxin receptor in cerebrospinal fluid as a potential biomarker for neurodegenerative diseases." Brain Res **10**: 158-170.
- Zamanillo, D., R. Sprengel, Ø. Hvalby, V. Jensen, N. Burnashev, A. Rozov, K. M. M. Kaiser, H. J. Köster, T. Borchardt, P. Worley, J. Lübke, M. Frotscher, P. H. Kelly, B. Sommer, P. Andersen, P. H. Seeburg and B. Sakmann (1999). "Importance of AMPA receptors for hippocampal synaptic plasticity but not for spatial learning." Science **284**(5421): 1805-1811.
- Zapata, A., S. Pontis, R. J. Schepers, R. Wang, E. Oh, A. Stein, C. M. Backman, P. Worley, M. Enguita, M. A. Abad, R. Trullas and T. S. Shippenberg (2012). "Alleviation of neuropathic pain hypersensitivity by inhibiting neuronal pentraxin 1 in the rostral ventromedial medulla." J Neurosci **32**(36): 12431-12436.
- Zuber, B., I. Nikonenko, P. Klauser, D. Muller and J. Dubochet (2005). "The mammalian central nervous synaptic cleft contains a high density of periodically organized complexes." Proc Natl Acad Sci U S A **102**(52): 19192-19197.
- Zuo, J., P. L. De Jager, K. A. Takahashi, W. Jiang, D. J. Linden and N. Heintz (1997). "Neurodegeneration in Lurcher mice caused by mutation in delta2 glutamate receptor gene." Nature **388**(6644): 769-773.

Chapter 2

General Materials and Methods

2.1 Molecular Biology

The starting point for the production of every protein detailed in this thesis was a cDNA template. From these double stranded sequences of coded information, all of the required constructs were generated. Details of the origin of each template, how each protein construct was designed and the vectors (plasmids) they were cloned into are described in depth in this section.

2.1.1 cDNA sources

All constructs of interest can be identified by their pubmed accession numbers. Constructs available as Integrated Molecular Analysis of Genomes and their Expression (IMAGE) clones were supplied by SourceBioscience (formally Geneservice). Constructs already in use by collaborators, were donated as gifts. Sources selected in the first instance were of human origin, although mouse and rat constructs were also frequently used. Information detailing construct source and accession number can be found in table 2.1.

Table 2.1. Accession numbers for all constructs of interest.

Information for where the cDNA was obtained and GenBank accession numbers for each construct. IMAGE clones are non-highlighted, gifts from collaborator are highlighted in yellow and synthetically generated constructs are highlighted in blue, all constructs are then grouped according to project.

Construct of interest	Accession number	Source
GRIA1_MOUSE (GluA1)	BC056397	IMAGE clone ID 6842391 (AV132-g6)
GRIA2_MOUSE (GluA2)	BC048248	IMAGE clone ID 6494013 (AV91-e6) GenBank 6494013
GRIA2_HUMAN (GluA2)	BC010574	IMAGE clone ID GenBank NM_000826
GRIA3 GluA3ecto domain (R3T)	P19492-2	Ingo Gregor, LMB Cambridge
GRIA4_MOUSE (GluA4)	BC060697	IMAGE clone ID 6409918 (AV115-e9)
GRIA4_HUMAN (GluA4)	BC167786	IMAGE clone ID 100068176
GRIA4_RAT (GluA4FL)	NM_017263.2	Ingo Gregor, LMB Cambridge
Neuronal Pentraxin 1_HUMAN (NPTX1)	BC089441	IMAGE clone ID 30555728 (AK201-a10)
Neuronal Pentraxin 2_HUMAN, NARP	BC009924	IMAGE clone ID 3508049 (AU48-a12)
Neuronal Pentraxin Receptor_HUMAN (NPTXR)	BC153027	IMAGE clone ID 100016404 (CV11_610)
Sisha9_pCR2.1	Q9CZN4.2	Gene synthesis -see Appendix C
Cys_knot_pCR2.1	Q9CZN4.2	Gene synthesis- see Appendix C
TARP2 (stargazin) hTrx TARP2	O88602.1	Gene synthesis- see Appendix C
Brorin (rat)	NP001102782.1	Nadine Harmel (Düsseldorf, Germany)
Noelin-1 (mouse)	NP062371.2	Nadine Harmel (Düsseldorf, Germany)
Noelin-2 (rat)	NP001015017.1	Nadine Harmel (Düsseldorf, Germany)
Noelin-3 (mouse)	NP703188.1	Nadine Harmel (Düsseldorf, Germany)
Neuritin (rat)	NP445798.1	Nadine Harmel (Düsseldorf, Germany)

2.1.2 Construct design

Nucleotide, cDNA and protein sequence information was provided by the National Centre for Biotechnology Information (NCBI) Genbank Database and the Universal Protein Resource Knowledgebase (UniProtKB). Constructs not generated synthetically were produced by PCR amplification of the IMAGE clone. When designing new constructs it was important to get the domain boundaries correct. If boundaries were cut too short, it could have impeded protein folding and therefore its secretion. Protein domain boundaries can now be predicted using a variety of open access online tools. For this thesis, domain boundaries were calculated using

HHpred (<http://toolkit.tuebingen.mpg.de/hhpred>). When provided with an input search sequence or multiple alignment file, HHpred searches a wide number of sequence alignment databases and promptly provides information about protein homology and structure prediction. HHpred can also generate similarity scores based on global or local alignment requirements as specified by the user (Soding *et al.* 2005). This was therefore a vital tool for the design of domain-truncated constructs when structural information was not available. When designing constructs it was also vital to know where, if present, a protein's secretion signal was located. The native secretion signal was not routinely included in constructs because of the high level of protein obtained when the pHLsec secretion signal was used (Aricescu *et al.* 2006). However in some cases, the protein was only soluble when expressed with its native signal. Signal P (Nielsen *et al.* 1997) enabled signal sequences and their cleavage sites to be identified so that the sequence could be removed or included as required.

Post-translational modifications including the proteins glycosylation state was also researched. Asparagine (Asn, N) -linked glycosylation sites have the consensus sequence Asn-Xaa-Ser/Thr (where Xaa is not a proline) and can be predicted robustly by NetNGlyc (<http://www.cbs.dtu.dk/services/NetNGlyc/>). If required, these sites were removed from the construct of interest at the cloning stage or after protein purification by EndoF1 deglycosylation. There are currently no methods available to remove O-linked sugars from proteins and due to their large, flexible nature they can drastically reduce a proteins chance of crystallization. It was therefore sometimes necessary to remove O-linked glycosylation sites by mutation; NetOGlyc (<http://www.cbs.dtu.dk/services/NetOGlyc-2.0/abstract.php>) was used to predict potential glycosylation at serine and threonine residues.

Oligonucleotides were supplied by Eurofins (MWG) Operon and were designed using the Java-based Molecular Biologist's Workbench (Toldo 1997) (JaMBW) (<http://www.bioinformatics.org/JaMBW/>). Appropriate restriction sites, absent in the target gene, were incorporated in the forward and reverse primers so that they flanked the target gene

sequence generated by PCR. In general, around 15 base pairs (bp) of target gene sequence was incorporated into the primer design. A random 5' GCG cap ensured efficient DNA cleavage by the restriction enzyme, although longer caps were required for certain enzymes (details on New England Biolabs (NEB) website). In some cases, restriction enzyme cleavage sites contained within the gene of interest meant that alternative restriction enzymes, present in the pHLsec multiple cloning site (MCS), had to be selected. Where this was not possible, internal cleavage sites were removed by introduction of silent mutations. Oligonucleotides were designed to have ~50% GC content and to have an annealing temperature of around 60°C, to enable amplification with standard thermocycler PCR parameters outlined in table 2.2.

2.1.3 Polymerase Chain Reaction (PCR)

PCR reactions contained 1µl (10pmol/µl) of each oligonucleotide, 1µl template (~1-100µg), 5µl 10x buffer, 8µl deoxyribonucleotide triphosphates (dNTPs), 33.75µl ultrapure (MilliQ) water and 0.25µl Pyrobest polymerase (Takara) in a 50µl final volume.

Table 2.2 General thermocycler parameters

Step	Temperature °C	Time (s)
1	96	120
2	96	50
3	55	50
4	72	60 per Kb
5	Repeat steps 2-5 for 30 cycles	
6	72	600
7	10	∞

2.1.4 Splice Overlap Extension (SOE)-PCR

SOE-PCR was used to generate constructs requiring site-directed mutagenesis. This required two-rounds of PCR amplification, the first round outlined above (section 2.1.3) was used to generate two fragments (routinely termed A and B) with a 15-20 residue overlap. In the second

round of PCR amplification 3 μ l of fragments A and B were used together as ‘templates’ and primers were added to the reaction after three cycles of amplification (table 2.2). The volume of ultrapure (MilliQ) water was adjusted to retain a 50 μ l final reaction volume.

2.1.5 Expression vectors for mammalian expression

Soluble proteins, expressed in the mammalian cell expression system, were cloned into the pHLsec vector (Aricescu *et al.* 2006) via AgeI/KpnI restriction sites unless otherwise stated. This high copy number pHLsec plasmid contains a C-terminal hexa-histidine (6xHis) tag, an optimized Kozak sequence to initiate translation, a secretion signal sequence and an Ampicillin (Amp) resistance gene (Figure 2.1A & B). Cell membrane attached proteins e.g. those used in cellular assays and for attempted membrane purification, were cloned into a modified version of pHLsec termed pHLsec_monoVenus (mV) containing a C-terminal yellow fluorescent protein (YFP) (figure 2.1C) fluorescence tag for which the design and generation is detailed below.

2.1.6 Generation of pHLsec_monoVenus vector

A dimeric YFP termed Venus, was previously developed to have a faster maturation rate than that of green fluorescent protein (GFP) (Nagai *et al.* 2002). To create a monomeric form of Venus (mV), a lysine mutation was introduced at residue A206 by site-directed mutagenesis (SDM) and PCR. The neurospora expression vector pYFP (accession number EF661030) was used as a template in reactions to generate the two fragments which were subsequently used as templates in a splice overlap extension (SOE) PCR to create the mV construct (sequence in Appendix E). This construct was then cloned into the pHLsec Kpn1/XhoI to lie downstream of the pHLsec multiple cloning site (MCS) (Figure 2.1C).

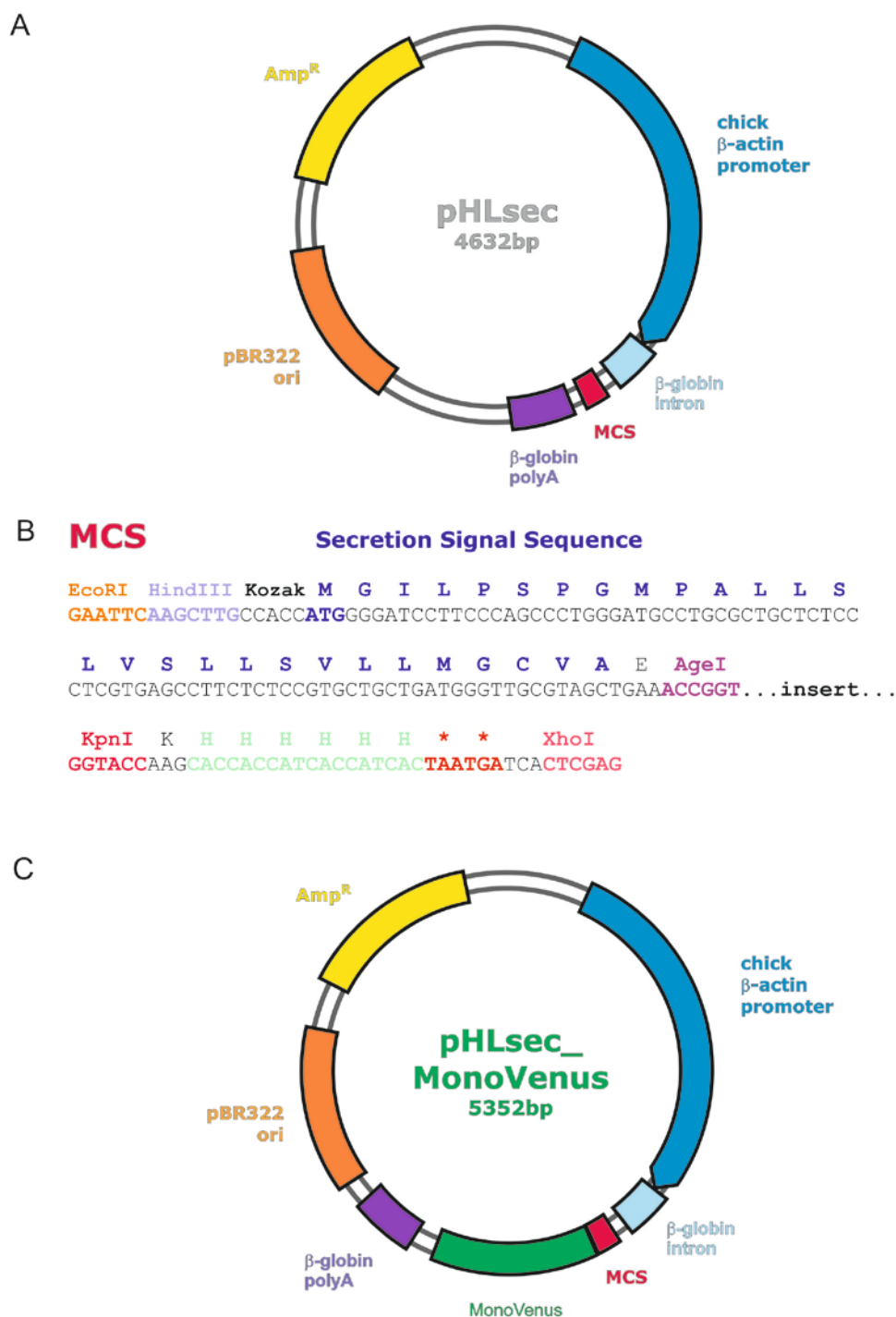


Figure 2.1 Cloning Vectors.

A) pHLsec vector map, B) pHLsec multiple cloning site (MCS) indicating key restriction sites (EcoRI, AgeI, KpnI and XhoI) as well as the secretion signal (navy blue) and hexa-histidine tag (mint green) highlighted. Stop codons are indicated by red * and C) pHLsec_monoVenus (mV) vector map.

2.1.7 Additional protein tags

The hexa-His tag, although excellent for protein purification techniques, is not optimal for a variety of other cellular analysis methods. Additional protein tags, more suitable to the techniques being employed, were therefore generated. For immunocytochemistry (ICC) and for analysis by fluorescence activated cell sorting (FACS) analysis, constructs were cloned to include a Haemagglutinin (HA) (peptide sequence YPYDVPDYA) or FLAG tag (peptide sequence DYKDDDDK) (Hopp *et al.* 1988). Constructs for surface plasmon resonance (SPR) analysis were cloned into the pHLsec_Avi (pHL_Avitag3) vector (Aricescu *et al.* 2006) to enable proteins to be biotinylated. In pHLsec_Avi the Avitag sequence (GLNDIFEAQKIEWHE) is located downstream of the insert but spaced away from the protein by a linker region (GSGGS) with the hexa-His tag remaining C-terminally located.

2.1.8 Restriction digest/ cloning

Following PCR amplification, a 2 µl sample of the PCR product was visualised on a 1% agarose/TAE gel, and the remaining 48µl was purified (standard Qiagen PCR clean up method). DNA was eluted in 30µl of 10mM Tris-HCl (pH8.0) and then enzymatically digested to produce cohesive ends, compatible with those generated by restriction digest of plasmid of choice. Each 35µl reaction typically contained 0.75µl of each enzyme (NEB), 3.5µl of appropriate buffer and 0.35µl BSA (if required). Reactions were then incubated according to conditions specified by NEB. Vector DNA (2-4µg) was digested in a 30µl reaction volume. DNA was then excised from a 1% agarose gel after electrophoresis and purified by gel purification (Qiagen) following manufacturer's standard protocol.

A standard ligation reaction was then established and contained 2.5µl of cut vector, 7.5µl of 'insert', 10µl of the quick ligase 2x reaction buffer and 1µl of the quick ligase enzyme (NEB) in a total volume of 20µl. Reactions were incubated at room temperature (RT) for 20 minutes (mins) and then on ice for a further 20 mins. Following this incubation, 1.5µl of the ligation

reaction was used to transform 35µl of library efficiency DH5α competent *Escherichia coli* (*E. coli*) cells (Invitrogen). The cell/ligation mix was incubated on ice for 30 mins, before a 45 second (s) heat shock at 42°C and 1 hr recovery at 37°C with 80µl SOC added at 200rpm shaking. Transformed cells were plated onto Luria Broth (LB)-agar plates containing 100µg/ml Amp and grown overnight at 37°C. Individual colonies were picked and grown in 3ml LB-Amp cultures overnight at 37°C, with 200rpm shaking. Cultures were centrifuged and plasmid DNA isolated (Qiagen Miniprep, standard protocol). DNA (1µg) was then enzymatically digested and analysed by agarose gel electrophoresis to identify plasmids with inserts of the correct size. Sequence verification then confirmed that constructs were correct (SourceBioscience, Oxford).

The following sequencing primers were used for constructs cloned into pHLsec and pHLsec_mV unless otherwise stated:

pHLsec_fwd 5'-GCTGGTTGTTGTGCTGTCTCATC-3'

pHLsec_rev 5' -CACCAGCCACCACCTTCTGATAG -3'

monoVenus_rev 5' -GCCCTGGAACAGCACCTCCAG-3'

2.1.9 Plasmid purification

Individual colonies were picked from freshly transformed LB-(Amp) agar plates and cells were grown overnight at 37°C with 200rpm shaking. The purification protocol was as stated in the manufacturer's instructions. The amount of DNA required dictated the size of the *E. coli* culture to grow. Depending on the requirements, a variety of plasmid purification kits were used. Typical yields expected for each Qiagen kit was as follows; Miniprep 25µg, Maxiprep 500µg, Megaprep 3mg and Gigaprep 15mg.

2.1.10 Cloning for generation of stable cell lines

The pNeo-Sec vector (Yuguang Zhao) (Appendix E vector map and sequence) is a newly developed pHLsec-based vector that enables the production of a stable polyclonal cell line in

Human Embryonic Kidney (HEK) 293S cells. The neo indicates its Neomycin (G418) selection requirement. Constructs to be expressed by stable cell line were cloned into pNeo-Sec by AgeI/KpnI.

2.2 Protein expression in mammalian cell systems

The next stage of the production line involved taking the clones constructs of interest and expressing them in a suitable system. Eukaryotic expression systems including mammalian expression are typically more favourable to a crystallographer than prokaryotic methods (e.g. *E. coli*) because the in-cell production mechanisms express correctly folded proteins, complete with post-translational modifications (Aricescu *et al.* 2006). Within the Division of Structural Biology (STRUBI) the HEK cell systems has been highly tuned to express soluble secreted proteins (Aricescu *et al.* 2006). Methodology for how this system was utilised is detailed below. Occasionally, when a large amount of protein was required, it was deemed more cost and time effective to generate a stable cell line, details of which are also below. For certain techniques such as microscopy work, the HEK cells lines were not suitable due to their shape. Where this was the case, details of all other cellular techniques with various cell lines are also described.

2.2.1 Small-scale expression trials

Initial expression of membrane-attached and secreted protein was tested by small-scale transient transfection of HEK293T cells. Flat-bottomed 35mm multi-well dishes (Corning) were seeded with 2ml of cell suspension in Dulbecco's Modified Eagle Media (DMEM) (Sigma) supplemented with 10% Foetal Bovine Serum (FBS) (Invitrogen), 5µg/ml L-glutamine (L-Glu) (Sigma) and 5µg/ml non-essential amino acids (NEAA) (Invitrogen). Cells were seeded with 0.5×10^6 cells per well and incubated for 24 hrs with 5% CO₂ at 37°C until a final density of 80-90% was achieved (where 100% confluency is $\sim 1 \times 10^6$ cells/well). A transfection 'mastermix' was then prepared as follows: 4µg of DNA (Qiagen Miniprep grade) was added to 250µl serum free

DMEM media (SFM) and vortexed briefly; 8µg of Lipofectamine™ 2000 (Invitrogen) was then added to a second tube containing 250µl of SFM before being briefly vortexed; the Lipofectamine-containing tube was then added to the DNA-containing tube and vortexed again; samples were incubated for 20 mins at RT. Cell media was exchanged for 1.5mls fresh DMEM supplemented with 2% FBS and 5µg/ml NEAA/L-Glu. The transfection mastermix was then gently added to each well, so as not to disrupt the monolayer, before being incubated for 48 hrs at 37°C, 5% CO₂.

2.2.2 Large-scale protein expression

Constructs considered to have good secretion levels in initial small-scale trials were scaled up for large-scale transient expression in expanded surface 2125cm² roller bottles (Greiner Bio-one). HEK293T (+/-kifunensine) or HEK N-Acetylglucosaminyltransferase I-negative GnTI cells (HEK293S) (Reeves *et al.* 2002) were universally used because both of these cell lines can express a high level of protein when used in conjunction with the transfection reagent polyethylenimine (PEI). A transfection cocktail containing 0.5mg of chloroform-extracted DNA (Qiagen Gigaprep) and 1mg of PEI was added to 50mls of SFM. The cocktail was mixed and incubated for 10 mins at RT. Concurrently the cell media FBS levels were reduced from 10% to 2% by media exchange. The transfection ‘mastermix’ was added to each roller bottle before being incubated at 37°C with gentle rotation.

With the exception of NP1_PTX domain construct, all other constructs contained one or more N-linked glycosylation sites. Therefore transfection of 293T cells required the addition of 1mg/litre kifunensine (Toronto Research Chemicals). This α -mannosidase inhibitor enables Endoglycosidase F1 (EndoF1) cleavage of N-linked glycosylation to leave a single N-acetylglucosamine (GlcNAc) residue. This potentially increases crystallization probability and in some cases might also improve dataset resolution by allowing tighter crystal packing (Chang *et al.* 2007). Where required, kifunensine (1mg/L) was added to the transfection mastermix after the 10 min incubation period.

2.2.3 In vivo selenomethionine (SeMet)-labelling

Where molecular replacement models are not available, one method to solve the crystallographic ‘phase problem’ is heavy metal phasing by selenium atoms that have been incorporated into the protein during its expression. SeMet incorporation was first used in *E. coli* systems (Hendrickson *et al.* 1990) but later adapted for use in mammalian systems (Wu *et al.* 1994; Lustbader *et al.* 1995). It works on the principle that all methionine residues (containing sulphur atoms) get substituted by selenomethionine, which contain the heavier selenium atom. Large-scale transient transfection was carried out as outlined above (2.2.2). After 24 hrs at 37°C, media was discarded and cells were washed twice with 50ml PBS/bottle to remove L-methionine (L-Met). For the media exchange, SeMet (final concentration 30mg/L) was added to 500ml DMEM media (L-Met and L-Glu free) containing 10ml dialysed FBS and 5µg/ml NEAA/L-glutamine. Where required, kifunensine (1mg/L) was also added at the media exchange stage. Due to a faster rate of cell death caused by the toxicity of SeMet, media was harvested typically 48 hrs post media exchange.

2.2.4 Generation of stable cell lines

Flat-bottomed 35mm multi-well dishes (Corning) were seeded with 0.5×10^6 HEK293S cells per well and incubated for 24 hrs with 5% CO₂ at 37°C until a final density of 80-90% was achieved. Cells were then transfected (standard protocol, section 2.2.1) with 1µg chloroform extracted target construct in the pNeo-Sec vector and 3µg of phiC31 integrase plasmid (stock concentration 1mg/ml). Cells were transferred 48 hrs post transfection from each well into a T75 flask in standard 10% DMEM medium. After a further 24 hrs, G418 (final concentration 0.5mg/ml) was added directly to the flask. Addition of G418 prevented the survival of non-transfected cells. Cells were transferred (48 hrs after the addition of G418), to a T175 flask with media supplemented with 0.5mg/ml G418. Many cells died, but after 14 days, islands of cells had grown large enough to be visible by eye. At this point the cells were transferred to fresh T175 flask in media supplemented with G418, to enable cells to grow in a monolayer. Cells were

confluent around 21 days after the protocol was started. The media was tested for secreted target protein by western blot (section 2.3.2) and positive cells were expanded as required.

2.2.5 Cell preparation for Immunocytochemistry (ICC)

Chinese hamster ovary (CHO)-K1 or COS7 (CV-1 (simian) in Origin, carrying SV40 genetic material) cells were transfected in a similar way to that described for small-scale transfections (section 2.2.1), except that four-well glass chamber slides (measuring 25 x 75mm) (Nunc), pre-coated with poly L-lysine were used. Cells were seeded 24 hrs preceding addition of the transfection cocktail containing 0.5µg DNA and 1µl (1mg/ml) Lipofectamine. Cells were prepared for microscopic analysis 48 hrs post transfection using the immunocytochemistry protocol outline below.

Cells were washed with PBS and fixed with 4% paraformaldehyde/PBS for 15 mins. The cross-linking reaction was then quenched with 50mM Tris pH7.5, 150mM NaCl for 10 mins before the cells were rinsed twice with PBS. A 15 min pre-block of 1% w/v bovine serum albumin (BSA)/PBS was then added to cells before incubation with a primary antibody mouse anti-HA antibody diluted 1:1000 in 3% BSA/PBS for 1 hr. Three additional 5 min washes with 0.1% BSA/PBS were then followed by the addition of fluorescently labelled secondary antibody anti mouse_568 diluted 1:400 in 3% BSA/PBS for 1 hr whilst protected from light. After 3 more washes, cells were rinsed briefly in water and mounted with Fluorosave (Millipore) containing a nuclear DAPI stain. Cells were visualised using a Zeiss 510 MetaHead confocal fluorescence microscope (Wellcome Trust Centre for Human Genetics, WTCHG).

2.3 Protein detection methods

2.3.1 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is one of many techniques used to separate proteins. The basic principle being that the SDS gives proteins a comparable negative charge based on a known charge to length ratio, the charge to mass ratio is therefore fixed, enabling proteins to migrate through the polyacrylamide gel based on their size alone. Smaller proteins are able to migrate through the acrylamide matrix more quickly, thus generating a gradient where proteins are separated according to size. SDS-PAGE gels with 12% acrylamide content were typically used but adjustments to the levels of acrylamide were made according to the range in size of the proteins requiring separation. Samples (15µl) were boiled with a 2x loading buffer containing beta-mercaptoethanol and SDS for 5 mins before being briefly centrifuged prior to gel loading. Novel Minicell (Invitrogen) gel running apparatus was used and 7.5µl of BenchMark™ (Invitrogen) protein molecular weight marker included in one lane of the gel to enable protein molecular weights to be determined. 180 Volts (V) was applied to the gel for ~50 mins and gels were then washed three times in MilliQ water before proteins were visualised by a Coomassie derived stain SimplyBlue™ SafeStain (Invitrogen).

2.3.2 Western blot analysis

In small-scale transient expression trials the quantity of proteins produced can be low and additionally the aim is to obtain fast answers concerning the level of protein expression and secretion, so the complex and time consuming protein purification methods are not employed. Instead expression and secretion is detected by western blot which for this thesis typically selected for proteins containing a hexa-His tag. For secreted proteins, cell culture media was centrifuged for 5 mins at 16'000 xg to remove cellular debris and adherent cells were re-suspended in 1ml PBS to be used for cellular fraction analysis to determine overall protein expression levels. Samples (15µl) were boiled with a 2x loading buffer containing beta-

mercaptoethanol for 5 mins and briefly centrifuged prior to gel loading. A mix of BenchMark Prestained™ (7.5µl) and the hexa-His tagged BenchMark™ (0.5µl) (Invitrogen) markers was also loaded into one lane of the gel to enable protein size determination. Expression levels were determined by SDS-PAGE, transferred for 70 mins at 30V onto nitrocellulose membrane (Invitrogen) for western blot analysis.

Cellular and secreted protein fractions were analysed using anti-penta-His (Qiagen) primary antibody diluted 1:1000 in a blocking solution of 5% w/v non-fat milk (Marvel milk powder) in phosphate buffered saline (PBS) in the protocol as follows. Membranes were incubated in the blocking solution for 1 hr before incubation with primary antibody at RT for 1 hr with gentle shaking. The membrane was then washed three times with PBS-Tween-20 (0.05%) (5 mins per wash) then incubated in the secondary antibody (goat anti-mouse IgG (Fc specific) – peroxidase (Sigma, A0168)) diluted 1:2000 in blocking solution for 1 hr. After three more washes with PBS-Tween-20 (5 mins each) the signal was detected by Enhanced Chemilluminescence (ECL) (GE Healthcare).

2.4 Protein purification from mammalian expression systems

Prior to protein purification, primary structural analysis including molecular weight, theoretical isoelectrical point (pI), amino acid composition, atomic composition and extinction coefficient were computed by ProtParam bioinformatics tools provided by Expert Protein Analysis System (ExPASy) (<http://expasy.org/>) (Gasteiger *et al.* 2003).

2.4.1 Media harvest and dialysis

For transiently transfected secreted proteins, media was collected 7-10 days (for HEK293T) and 10-18days (for HEK293S) post transfection. General cell health was indicated by the pH indicator present in DMEM media and also the level of cellular detachment from the roller bottle surface. Centrifugation at 4000 xg for 20 mins removed cellular debris. Media was then sterilized by filtration through a 0.2µm filter (Millipore) and dialysed against 3x 15litres exchanges of PBS for 48-72 hrs at 4°C. Dialysis membrane (Visking) MWCO 12-14kDa (Medicell International Ltd, DTV12000.13.000) was activated in 2% sodium bicarbonate, 2mM EDTA (pH8.0) for 10 mins at 70°C then washed with distilled water prior to use.

2.4.2 Immobilized metal ion affinity chromatography (IMAC)

Poly-histidine tagged proteins were initially purified by immobilised metal affinity chromatography (IMAC) using TALON[®] beads (Clontech) which utilises the cobalt ions (Co²⁺) affinity for poly histidine residues. Beads were prepared as per the listed manufacturer's procedures prior to the media being incubated with beads for 2 hrs at 16°C. This slurry was mixed by gentle rotation at 100rpm to enable a constant distribution of the beads within the media, without causing too much agitation. Proteins may naturally contain poly histidine sites, and so will weakly co-ordinate Co²⁺ ions. The beads were therefore pooled and washed with 250-500ml of 5mM imidazole pH8.0 to remove non-specifically bound proteins. The target protein was then eluted in 1ml fractions in 10-15 bed volumes (BV) of 300mM imidazole pH8.0 which acts by competing with the histidine residues in binding to the resin. A small-scale Coomassie test based

on the Bradford assay was used as a quick determinant for when all protein had been eluted from the beads. This involved samples from eluted protein fractions (10µl) being mixed with 100µl of SimplyBlue™ SafeStain (Life technologies) and if the blue colour persisted, a longer elution was required.

2.4.3 Size exclusion chromatography (SEC)

Size exclusion chromatography (SEC) was performed using an AKTA™ system. Protein eluted from IMAC was concentrated to a required volume, dependant on the column size used (typically 250µl for use with an S200 10/30 column, and up to 2.5mls onto a larger S200 16/60 column). Running buffer was routinely 10mM HEPES pH7.5-8.0, 150mM NaCl, filtered and degassed prior to use. AKTA system running parameters were as in the column manufacturer's instructions (typically 0.5ml/min for an S200 10/30 and 1ml/min for an S200 16/60). SDS-PAGE analysis was used to determine protein purity and confirm glycosylation state(s).

2.4.4 Protein deglycosylation

The complexity of N-linked glycosylation branching is reduced in proteins expressed in HEK293S (or HEK293T if kifunensine has been added) (described in section 2.2.2). Concentrated purified protein was incubated at room temperature for 4-7 hrs at RT or at 37°C for 1 hr with 10µl of EndoF1 (1mg/ml). If the protein was prone to precipitation upon deglycosylation then the procedure was performed on the protein as it was being set up for crystallization. Briefly 1µl of EndoF1 was added to 100µl of concentrated protein which was then loaded onto the crystallization system.

2.5 Protein expression and purification from *E. coli*

Several constructs were cloned into the pET22b(+) vector (Novagen) via NdeI/XhoI restriction sites and expressed in *E. coli*. The pET vector cloning system also enabled proteins to be expressed with a hexa-His purification tag. In this system protein expression is controlled by a promoter that is inducible by addition of isopropylthio- β -galactoside (IPTG). Concentrations of IPTG were optimized for each protein as required.

2.5.1 Sequencing in pET vector

Constructs were cloned into pET22b and sequence verified using standard T7F and T7R primers:

T7 promoter (F) 5'-TAATACGACTCACTATAGGG-3'

T7 terminator (R) 5'-GCTAGTTATTGCTCAGCGG-3'

*2.5.2 Transformation of Origami *E. coli* cells and protein expression*

Origami *E. coli* cells (Novagen, Cat. No. 71346-30) have mutated glutathione reductase genes and so these cells enable disulphide bond formation, that is required for protein folding, to occur during protein expression. A 50 μ l aliquot of Origami cells were transformed with 1.5 μ l pET22b vector following the Origami transformation protocol stated by the manufacturer. Cells were then plated onto LB agar plates with Amp (100 μ g/ml), Chloramphenicol (Chloramp) (34 μ g/ml), Tetracyclin (Tet) (12.5 μ g/ml) and Streptomycin (Strep) (100 μ g/ml). Colonies were grown overnight in a static 37°C incubator. Single colonies were used to inoculate starter cultures of 10ml LB + Amp, Chloramp, Tet and Strep (concentrations as above) grown for 12 hrs (or overnight) at 37 °C, shaking at 200rpm. Starter cultures were used to inoculate 1L cultures of LB plus all antibiotics as above and grown at 30°C shaking at 200rpm to an optical density (OD)₆₀₀ ~0.6 – 0.7. Cultures were cooled to 18°C (RT) for 1 hr and were then induced with 0.4mM IPTG

and grown for a further~20-24 hrs with 200rpm shaking at 18°C. Note this lower temperature acts to reduce the formation of inclusion bodies.

2.5.3 Protein purification from Origami cells

After incubation, cells were harvested by centrifugation (20 min at 4000 $\times g$). Pelleted cells from 1L of culture were lysed in 25ml buffer containing 150mM NaCl, 10mM Tris pH7.5, 50 μ l of 50mg/ml DNase1, 0.1g lysozyme and 200 μ l protease inhibitor cocktail (PI) (Sigma) and incubated with gentle mixing for 30 mins at 4°C. Cells were then sonicated with 10s pulses, followed by 10s off for 4 mins on ice and then centrifuged for 40 mins, at 108,000 $\times g$. For each litre of original cell culture, supernatant was diluted to 0.5L in a PBS buffer pH8.0, 500mM NaCl plus PI cocktail then incubated with 5ml freshly prepared Talon beads at 4°C for 90 mins with gentle mixing. Beads were washed with 200ml buffer (as above) containing 0.1% Tween-20, then with 200ml buffer alone, followed by a third wash with 200ml buffer containing 5mM imidazole. Protein was eluted in 5x10ml aliquots of 10mM HEPES pH8.0, 500mM NaCl, 500mM imidazole. Size exclusion chromatography (SEC) was performed with running buffer 10mM HEPES pH8.0, 500mM NaCl. In the case of GluA2_LBC, 1mM L-glutamine was also included in the IMAC elution buffer and the SEC running buffer to aid correct protein folding and LBC stabilisation.

2.5.4 Transformation of Rosetta cells and protein expression in inclusion bodies

When production of soluble folded proteins failed in the Origami *E. coli* cell system, proteins were expressed in RosettaTM *E. coli* at high temperatures to force proteins into inclusion bodies, from which the protein could then be isolated, purified and refolded.

A 30 μ l aliquot of Rosetta cells were transformed with 1 μ l ‘construct’_pET vector. They were then plated onto LB agar plates with Amp (100 μ g/ml) and Chloramp (34 μ g/ml) and grown overnight at 37°C. A single colony was used to inoculate 10ml LB with Amp & Chloramp (concentrations as above) for 12 hrs (or overnight) at 37 °C, shaking at 200rpm. Each culture was then expanded into two conical Erlenmeyer flasks containing 500ml LB (plus antibiotics as

above) and grown at 37°C with 200rpm. The culture was induced with 0.8mM IPTG once the culture density had reached OD₆₀₀= 0.6-0.7). Cells were then grown for 4-12 hrs at 37°C and harvested by centrifugation (20 mins at 4000 xg).

2.5.5 Protein extraction, refold and purification from inclusion bodies

A lysis buffer containing 150mM NaCl, 10mM Tris pH 7.5, 1mM phenylmethylsulfonyl fluoride (PMSF), 50µl of 50mg/ml DNase 1, with 0.1g lysozyme and 50µl PI cocktail (Sigma) was freshly prepared for immediate use. Cell pellets were re-suspended in 25ml lysis buffer and incubated with rocking for 30 min at 4°C. Cells were then lysed by sonication with 10s pulses, followed by 10s off for 5 mins and centrifuged for 30 mins at 108,000 xg. Supernatant was discarded and the remaining pellet dislodged and washed three times in 50mM Tris pH8.0, 0.5% Triton, 100mM NaCl, 1mM EDTA with 1mM of the reducing agent dithiothreitol (DTT) and re-centrifuged 108,000 xg for 10 mins. The cell pellet became whiter with each wash, indicating that the inclusion bodies were being isolated.

Pellets were then re-suspended in 5ml of the denaturing solution (6M Guanidinium-HCl) and incubated overnight at 4°C. The denatured protein solution was centrifuged at 108,000 xg for 20 mins to remove any aggregated protein. Concurrently a refold buffer (containing 10mM Tris pH 8.0, 0.5M L-Arginine, 5mM reduced glutathione (GSH), 0.5mM, oxidised glutathione (GSSG), 5mM EDTA pH8.0, 150mM NaCl (250ml per litre of culture)) was chilled to 4°C and then the protein solution slowly added to the refold buffer in drop-wise fashion and left overnight at 4°C to refold (minimum 24 hrs).

Refolded protein was dialysed three times against 15L PBS, 0.5M NaCl for 72 hrs at 4°C. The salt content of the dialysed protein solution was then reduced to 50mM via a 1 in 10 dilution with 10mM Tris pH8.0. Protein was then bound to a 5ml HiTrap Q HP anion exchange column (GE Healthcare) using a peristaltic pump (5ml/min). Separation was achieved according to the net charge of the proteins by ion exchange chromatography (IEC) which was carried out using a

constant gradient increase from a low salt 'Buffer A' 10mM Tris pH 8.0, 50mM NaCl, into a high salt 'Buffer B' 10mM Tris pH8.0, 500mM NaCl. Further purification utilised SEC on either an S200 or S75 column in 10mM HEPES pH8.0, 0.5M NaCl.

2.6 Crystallization

High-throughput nano-litre sized crystallization trials were carried out using the vapour diffusion sitting drop method. (Brown *et al.* 2003; Walter *et al.* 2005; Walter *et al.* 2007). The crystallization process involves the protein-containing drop becoming concentrated over time by dehydration and if conditions are optimal, will result in spontaneous protein nucleation and crystal growth. The purified proteins were typically concentrated to around 8-10mg/ml in the first instance. A pre-crystallization trial (PCT) (Hampton) (table 2.3) was then performed to ensure that the protein concentration was suitable for a full screen of crystallization conditions (table 2.4).

Table 2.3 Commercial Pre-Crystallization Trial (PCT) formulation.

This figure has been removed due to copyright restrictions.
--

The variety of crystallization screens tested (table 2.4) are based on commercially available screens that have been reformatted for 96 well plates (Walter *et al.* 2005). Precipitant solution was added to each reservoir of a 96 well barcoded plate (Greiner), using an automated Hydra system (Matrix Technologies). A Cartesian Microsys dispenser system (Brown *et al.* 2003; Walter *et al.* 2005) was used to first dispense 100nl of protein then 100nl reservoir solution to wells in the Greiner plate. Room temperature plates were then transferred to an automated storage vault (The Automated Partnership, TAP). An automated Oasis 1700 (Veeco) imaging system enabled remote monitoring of potential crystal growth. For 4°C crystallization trials the

temperature of the Greiner plates was gradually reduced to prevent condensation on the lid seal, which otherwise would have detrimental effects on image quality, before being transferred to a Formulatrix storage and imaging system maintained in a cold room.

Table 2.4 Range of crystallization screens utilised.

Screen Name	Manufacturer	Screen Type	Commercial screen contents
Block 1	Hampton Research	Sparse matrix screen	Crystal screen 1/2
Block 2	Emerald BioSystems	Sparse matrix screen	Emerald Wizard I/II
Block 3	Hampton Research	Sparse matrix screen	Hampton PEG/Ion/APS screen
Block 4	Hampton Research	Sparse matrix screen	Natrix/ crystal screen cryo/ Membfac.
Block 5	Hampton Research	Sparse matrix screen	Grid screen PEG/LiCl/ NaCl &MPD/Quik
Index	Hampton Research	Sparse matrix screen	Hampton Index
SaltRx	Hampton Research	Sparse matrix screen	Hampton SaltRx Salt/ versus pH matrix crystallization screen
PEGRx	Hampton Research	PEGRx screen	polymer and pH based crystallization screen
Morpheus	Molecular Dimensions	Morpheus screen	3D screen of low MW ligands
PACTpremier	Molecular Dimensions	Sparse matrix screen	pH, cation & anion screen with PEG precipitant
Membrane protein screen	Molecular Dimensions	Sparse matrix screen	Optimized pH, precipitant and salt concentrations
Proplex	Molecular Dimensions	Sparse matrix screen	Targeted to protein-complex crystallization database (PCCD)
Additive	Hampton Research	Additive	Volatile/ non-volatile compounds
Limited	In-house development based on crystal hits	Sparse matrix screen	Hits screen

Not all of the initial crystals obtained were suitable for data collection and needed optimisation. One method of optimisation employed was to vary both the protein to precipitant ratio and the precipitant concentration in a small number of drops. This was achieved by creating a dilution series (using a RoboGo robot (MWG) of precipitant across 12 drops (1 row on a 96 well plate) typically with a 3% gradient. For each row the protein to precipitant ratio was either 1:1, 1:2 or 2:1 through rows A-C and in duplicate through rows E-G.

2.7 Data collection

All datasets produced for this thesis were either collected at the Diamond Light Source (DLS), Harwell in Oxfordshire on the IO3, IO4 and IO4-1 beamlines or at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Prior to testing, frozen crystals were prepared for shipment as described below.

Cryo-protection attempts to protect protein crystals from the freezing damage associated with the formation of ice crystals. In most cases the ‘mother liquor’ did not provide any cryo-protection so either ethylene glycol (30%) or glycerol (25-30%) was added directly to crystals in the drops. Mother liquors that contained small molecule PEGs or had a high enough percentage of glycerol did not require any further cryo-protectant (Garman *et al.* 1997). The crystals were then picked using a suitable sized loop, either the nylon CryoLoopTM (Hampton Research) or the mesh type LithoLoop (Molecular Dimensions) which is more suitable for fragile crystals. The crystals could then be flash frozen at 100 Kelvin (K) in liquid nitrogen awaiting data collection. At the DLS, a robotic sample changer mounts the crystal on the beamline and software allows crystals to be orientated remotely. Details of data collection, processing, refinement and structural validation specific to each solved protein crystal structure can be found in subsequent chapters.

2.8 Biophysical analysis

It is important to support crystallographic observations with biophysical analysis. Such techniques provide information on a protein's oligomeric state(s) as well as giving further evidence that the occurrence of any protein-protein interactions are not artificial products of protein crystallography. A range of biophysical techniques were therefore employed to complement the structural analysis of every protein. Details of specific techniques utilised during this thesis are described here.

2.8.1 Multi-angle light scattering (MALS) analysis

MALS is a technique that enables the absolute molecular mass of a protein to be calculated. In an ideal world one would calculate the absolute molecular mass of a protein by measuring the intensity of light scattered at zero degrees (0°) from the laser source. However in practice this cannot be calculated because the intensity of the source will always be greater than the light scattered. In MALS, measurements are made at multiple angles to improve accuracy and the signal to noise ratio where low angle measurements are associated with a higher level of noise. MALS then uses a model (e.g. the Raleigh ratio) to extrapolate the value to 0° . In all, three values must be known; the specific refractive index (dn/dc), the concentration of the protein (calculated from the UV absorbance) and the amount of light scattered (Wen *et al.* 1996). All protein dn/dc values fit within a narrow range of values (Zhao *et al.* 2011) where a value of 0.186mL/g is typically used (Wen *et al.* 1996). However the dn/dc value must be adjusted for glycoproteins because post-translational modifications in the form of carbohydrate additions cause excess light scattering (Arakawa *et al.* 2001).

In practice, the protein sample is first separated according to size by SEC and then pumped through a series of detectors for light scattering, absorbance and refractive index. MALS coupled with SEC analysis required a high level of protein purity therefore could only be carried out after large-scale transient expression and protein purification. All proteins were concentrated

to 1mg/ml in 110µl unless otherwise stated. Running buffer (10mM HEPES pH7.5-8.0, 150mM NaCl) and MilliQ H₂O were filtered and extensively degassed before use. The Wyatt MALS system and sample loading loop were rinsed with water prior to use. With the pressure limit set to 0.4bar, a superdex S200 HR10/30 column was equilibrated at 0.5ml/min into MilliQ H₂O and then into running buffer. A 100µl sample aliquot was injected onto the column. Light scattering and refractive index data were collected and analysed using the Astra software package (Wyatt). During data analysis the baseline levels were first set and then the peaks of interest selected. This software also enabled individual light scattering detector measurements to be removed from the analysis if the signal to noise ratio was too low. The data was then plotted as refractive index (y1 axis) and molar weight (kDa) (y2 axis) against elution volume or time (x axis).

2.8.2 Surface Plasmon Resonance (SPR)

SPR is an optical detection technique to measure biomolecular interactions in real time. Many systems, including the Biacore system utilised here, employ a detection scheme known as the Kretschmann configuration. This involves polarized light passing through a prism and being reflected from the bottom of a thin layer of gold into a detector. At an incidence angle known as the resonance angle, light is absorbed by the electrons within the gold layer, causing them to resonate. Any absorbed light results in the reflected beam losing intensity which can be measured by the detector. Upon ligand binding, the refractive index close to the surface of the chip changes, where any change in mass at the sensor surface shifts the angular position of reflected light (Homola *et al.* 1999). In the Biacore system the resonance unit (RU) is the unit of measure where 1 RU approximately = 1pg protein/mm² of the sensor surface.

Protein ligands were coupled to the chip utilising the affinity of streptavidin for biotin. Prior to SPR, proteins were expressed with an Avitag in small-scale transient HEK293T cell transfections (section 2.2.1) ready for biotinylation. The biotinylation protocol required 4ml of media to be diluted 1:1 in PBS containing 10mM Tris pH8.0. Charged Talon resin (Clontech) (200µl) was then added and incubated for 2 hrs at room temperature with gentle mixing. Beads

were washed with 4 column volumes (CV) of 10mM Tris pH8.0 to reduce salt concentration which inhibits the biotinylation reaction, before beads were re-suspended in 500µl of 10mM Tris pH8.0. This reduced volume was then transferred to a fresh 1.5ml tube before 100µl of 10x buffer A (0.5M Bicine pH8.3), 100µl 10x Buffer B (100mM ATP, 100mM MgOAc, 500µl d-biotin) and 25µl BirA (biotin ligase) were added to the sample and incubated for 2 hrs at RT with gentle mixing. Beads were washed with 4 CV of PBS containing 10mM Tris pH8.0, and protein eluted with 5x 1ml fractions of 300mM imidazole. Protein was then buffer exchanged into 10mM HEPES pH7.4, 150mM NaCl by overnight dialysis at 4°C and then concentrated to 0.5ml, ready for use.

The CM5 sensor chip (Biacore) surface was primed with 10mM HEPES pH7.4, 150mM NaCl, 0.05% Tween-20 at a 5µl/min flow rate. A 'normalisation solution' (0.1M glycine, 50mM NaOH, 0.3% Triton X100) was then injected over all four flow cells (FCs) at 5µl/min for 5 mins. The chip was then activated with a coupling reaction of *N*-hydroxysuccinimide/ethyl(dimethylaminopropyl) carbodiimide (NHS/EDC) for 7 mins at 5µl/min. Streptavidin (Sigma) in 10mM sodium acetate pH4.5 was then injected over all four FCs for 14 mins at 5µl/min (to bind around 5000units). The chip surface was then blocked to inactivate any remaining sites with 1M ethanolamine hydrochloride pH8.5 at 5µl/min for 7 mins. The amount of ligand immobilized was varied as required. CM5 chips contain 4 FCs and the ligand was initially only injected over FC-4 for 420seconds at 5µl/min to gauge maximum protein levels obtainable. Units of protein ligand were then bound as required onto FC- 2 and 3 and sometimes required the protein to be diluted. The FC-1 was always the control lane with no ligand bound.

Once ligands were bound, 30s test analyte injections were performed at 10µl/min. Where regeneration was required, 1.5M NaCl, 10mM HEPES pH7.5, 0.05% Tween-20 was typically applied and the flow rate increased to 50µl/min for 30s bursts until baseline response units (RU) levels returned. For any analyte showing a potential interaction with a ligand, a serial dilution was set up and tested in duplicate. After each injection 600s buffer then 30s regeneration buffer (if

required), followed by another 100s of buffer then 30s of regeneration buffer and then finally another 600s of buffer was flowed over the chip between samples at 10 μ l/min.

2.8.3 Analytical ultracentrifugation (AUC)

AUC was performed by Dr Robert Gilbert (STRUBI) using an Optima XL-I analytical ultracentrifuge (Beckman Coulter). Sedimentation equilibrium experiments were conducted as described previously (Clayton *et al.* 2009). Protein was applied at set concentrations at speeds of 8,000 rpm, 14,000 rpm and 18,000 rpm to determine the oligomeric state of the protein. The technique measured any reversible self-associated oligomers that may have occurred at higher concentrations by calculating the K_D (equilibrium dissociation constant).

2.9 Fluorescence activated cell sorting (FACS)

FACS is a technique that enables cell populations to be sorted in a variety of ways. The principle technique uses a laser, focussed on a stream of cells in solution, set to the required wavelength to activate the fluorophore present. An emission is then detected which is highly sensitive and can detect changes in the levels of protein on the surface of the cell.

For the sake of this thesis, HEK293T cells in flat-bottomed 35mm multi-well dishes (Corning) were transfected with HA-tagged proteins either with or without the monoVenus tag. To reduce variations in transfection efficiency, a single ‘mastermix’ transfection cocktail with 4 μ g of DNA/well was used to transfect all wells. This ‘mastermix’ was added 26 hrs prior to cell staining and FACS analysis. Then if required, 16 hrs pre-cell staining, the potential interaction partner was applied in purified form at or above the K_D as determined by SPR.

Cells were washed and re-suspended in 1ml ice cold PBS before being centrifuged at 600 $\times g$ for 2 mins. The PBS was removed and cells re-suspended by agitation in the remaining liquid. Cells were then washed with 3ml cold PBS with 0.05% sodium azide (NaN₃) and centrifuged

again as before. The process was repeated twice more before cells were incubated for 45 mins at 4 °C with 1µl of primary antibody solution (mouse anti-HA at 0.5ng/µl) diluted in 100µl PBS NaN₃, cells were then washed three times as before. A 1µl aliquot of secondary antibody (either donkey α-mouse APC (Invitrogen) or Alexa488 Fluor® conjugated secondary antibody (Abcam) both at 0.5µg/ml) was used per 100µl PBS NaN₃ and cells were incubated for 30 mins at 4 °C protected from light using aluminium foil to prevent fluorophore bleaching. Cells were again washed three times as before, before being re-suspended in 2% paraformaldehyde (PFA)/PBS + 0.05% NaN₃.

Cells were sorted using a CyAn cell sorter (Beckmann Coulter). Several parameters in the form of 'gates' were applied to ensure that the cells were live (gate 1) and singular (gate 2). Occasionally, with low transfection efficiency rates, a third parameter was also included to identify only those cells that had been successfully transfected (gate 3).

2.10 Electron Microscopy

Electron microscopy (EM) was performed by Terunaga Nakagawa in San Diego. Samples were prepared as previously described (Ohi *et al.* 2004) and diluted to obtain optimal number of particles on each grid. Uranyl formate was used as a stain due to its fine grain structure. Protein samples were less prone to aggregation and high levels of background when they were transported in a buffer containing 10mM HEPES pH7.5, 150mM NaCl. It was also noted that when glycerol was added to the protein in an attempt to improve stability during shipment, the background staining was increased. Additionally, proteins shipped in the presence of PI (to reduce protein degradation by protease activity) did not withstand the sample preparation methods and aggregated.

2.11 Data analysis and generation of figures

Analysis of plasmid sequencing data was performed using MultAlin (Corpet 1988) after the nucleotide sequences had been translated to peptide sequences using ExPASy (<http://web.expasy.org/translate/>).

Raw sensogram curves, generated during each SPR Biacore run, were used to create a concentration curve in the Scrubber2 evaluation software. A detailed guide to the scrubber data analysis procedure can be found in Tutorial 0: Introduction to scrubber (2006) supplied with Scrubber2 program. Scrubber2 provided a step by step method to prepare data for presentation. Raw data curves were first loaded from the Biacore results output file. This data was then displayed as an initial raw data plot with Time (x axis) plotted against Response units (RU) (y axis). The baseline (zero) for RU was then manually set and data cropped to remove any regeneration stage etc. not required for the calculation. All curves were then aligned to zero on the x axis and blank (buffer only) runs were then selected for reference. Data generated for this thesis utilised the equilibrium analysis SPR method where the equilibrium position for each run was selected manually and then used to automatically generate a plot of Response (RU) against analyte concentration. These data were then exported via Microsoft Excel into Sigmaplot where simple ligand binding analysis was performed and a dissociation constant K_D calculated. Data were then transferred to CoralDraw® for graphical representation.

FACS data analysis and generation of raw data figures was performed in FlowJo and then exported to CorelDraw® to make the final figures.

Finally, plasmid maps were generated in CorelDraw®. Agarose DNA gels were visualised using ultraviolet (UV) illumination. SDS-PAGE gels and western blot films were scanned and figures labelled in CoralDraw® and Powerpoint. SEC data was visualised using the UNICORN™ software and then exported to Microsoft Excel to generate thesis worthy chromatograms.

2.12 References

- Arakawa, T. and J. Wen (2001). "Determination of carbohydrate contents from excess light scattering." Anal Biochem **299**(2): 158-161.
- Aricescu, A. R., W. Lu and E. Y. Jones (2006). "A time- and cost-efficient system for high-level protein production in mammalian cells." Acta Crystallogr D Biol Crystallogr **62**(Pt 10): 1243-1250.
- Brown, J., T. S. Walter, L. Carter, N. G. A. Abrescia, A. R. Aricescu, T. D. Batuwangala, L. E. Bird, N. Brown, P. P. Chamberlain, S. J. Davis, E. Dubinina, J. Endicott, J. A. Fennelly, R. J. C. Gilbert, M. Harkiolaki, W. C. Hon, F. Kimberley, C. A. Love, E. J. Mancini, R. Manso-Sancho, C. E. Nichols, R. A. Robinson, G. C. Sutton, N. Schueller, M. C. Sleeman, G. B. Stewart-Jones, M. Vuong, J. Welburn, Z. Zhang, D. K. Stammers, R. J. Owens, E. Y. Jones, K. Harlos and D. I. Stuart (2003). "A procedure for setting up high-throughput nanolitre crystallization experiments. II. Crystallization results." Journal of Applied Crystallography **36**(2): 315-318.
- Chang, V. T., M. Crispin, A. R. Aricescu, D. J. Harvey, J. E. Nettleship, J. A. Fennelly, C. Yu, K. S. Boles, E. J. Evans, D. I. Stuart, R. A. Dwek, E. Y. Jones, R. J. Owens and S. J. Davis (2007). "Glycoprotein Structural Genomics: Solving the Glycosylation Problem." Structure **15**(3): 267-273.
- Clayton, A., C. Siebold, R. J. C. Gilbert, G. C. Sutton, K. Harlos, R. A. J. McIlhinney, E. Y. Jones and A. R. Aricescu (2009). "Crystal Structure of the GluR2 Amino-Terminal Domain Provides Insights into the Architecture and Assembly of Ionotropic Glutamate Receptors." Journal of Molecular Biology **392**(5): 1125-1132.
- Corpet, F. (1988). "Multiple sequence alignment with hierarchical clustering." Nucleic Acids Research **16**(22): 10881-10890.
- Garman, E. F. and T. R. Schneider (1997). "Macromolecular Cryocrystallography." Journal of Applied Crystallography **30**(3): 211-237.
- Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch (2003). "ExPASy: The proteomics server for in-depth protein knowledge and analysis." Nucleic Acids Research **31**(13): 3784-3788.
- Hendrickson, W. A., J. R. Horton and D. M. LeMaster (1990). "Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): A vehicle for direct determination of three dimensional structure." EMBO Journal **9**(5): 1665-1672.
- Homola, J., S. S. Yee and G. Gauglitz (1999). "Surface plasmon resonance sensors: review." Sensors and Actuators, B: Chemical **54**(1): 3-15.
- Hopp, T. P., K. S. Prickett, V. L. Price, R. T. Libby, C. J. March, D. P. Cerretti, D. L. Urdal and P. J. Conlon (1988). "A short polypeptide marker sequence useful for recombinant protein identification and purification." Bio/Technology **6**(10): 1204-1210.
- Lustbader, J. W., H. Wu, S. Birken, S. Pollak, M. A. G. Kolks, A. M. Pound, D. Austen, W. A. Hendrickson and R. E. Canfield (1995). "The expression, characterization, and crystallization of wild-type and selenomethionyl human chorionic gonadotropin." Endocrinology **136**(2): 640-650.
- Nagai, T., K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba and A. Miyawaki (2002). "A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications." Nature Biotechnology **20**(1): 87-90.
- Nielsen, H., J. Engelbrecht, S. Brunak and G. Von Heijne (1997). "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Protein Engineering **10**(1): 1-6.
- Reeves, P. J., N. Callewaert, R. Contreras and H. G. Khorana (2002). "Structure and function in rhodopsin: High-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line." Proceedings of the National Academy of Sciences of the United States of America **99**(21): 13419-13424.

- Soding, J., A. Biegert and A. N. Lupas (2005). "The HHpred interactive server for protein homology detection and structure prediction." Nucleic Acids Res **33**(Web Server issue): W244-248.
- Toldo, L. I. G. (1997). "JaMBW 1.1: Java-based Molecular Biologists' Workbench." Computer Applications in the Biosciences **13**(4): 475-476.
- Walter, T. S., J. M. Diprose, C. J. Mayo, C. Siebold, M. G. Pickford, L. Carter, G. C. Sutton, N. S. Berrow, J. Brown, I. M. Berry, G. B. E. Stewart-Jones, J. M. Grimes, D. K. Stammers, R. M. Esnouf, E. Y. Jones, R. J. Owens, D. I. Stuart and K. Harlos (2005). "A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization." Acta Crystallographica Section D **61**(6): 651-657.
- Walter, T. S., E. J. Mancini, J. Kadlec, S. C. Graham, R. Assenberg, J. Ren, S. Sainsbury, R. J. Owens, D. I. Stuart, J. M. Grimes and K. Harlos (2007). "Semi-automated microseeding of nanolitre crystallization experiments." Acta Crystallographica Section F: Structural Biology and Crystallization Communications **64**(1): 14-18.
- Wen, J., T. Arakawa and J. S. Philo (1996). "Size-Exclusion Chromatography with On-Line Light-Scattering, Absorbance, and Refractive Index Detectors for Studying Proteins and Their Interactions." Analytical Biochemistry **240**(2): 155-166.
- Wu, H., J. W. Lustbader, Y. Liu, R. E. Canfield and W. A. Hendrickson (1994). "Structure of human chorionic gonadotropin at 2.6 Å resolution from MAD analysis of the selenomethionyl protein." Structure **2**(6): 545-558.
- Zhao, H., P. H. Brown and P. Schuck (2011). "On the distribution of protein refractive index increments." Biophys J **100**(9): 2309-2317.

Chapter 3

Structural and functional analysis of the AMPA receptor GluA2 amino terminal domain

3.1 Introduction

Alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors share a common domain organisation that is conserved amongst all ionotropic glutamate receptors (iGluRs). Their large ectodomain region contains a substantial amino terminal domain (ATD) (figure 3.1) of around 380 residues that shows sequence homology to a family of proteins that includes the bacterial leucine/isoleucine/valine-binding protein (LIVBP) (Matsuda *et al.* 2005) and the ligand binding domain of metabotropic glutamate receptor 1 (mGluR1) (Nakagawa *et al.* 2005) both noted to form dimers in solution (Kunishima *et al.* 2000; Matsuda *et al.* 2005). The iGluR ligand binding core (LBC) has a clam-shell topology containing two lobes termed S1 and S2. The peptide sequence forming the LBC is interjected by transmembrane domains M1, M3 and the M2 re-entrant loop, in which the Q/R RNA editing site (controlling AMPA receptor calcium permeability) is located (figure 3.1) (Ozawa 1996). This is followed by the final transmembrane helix (M4) into the intracellular C-terminal.

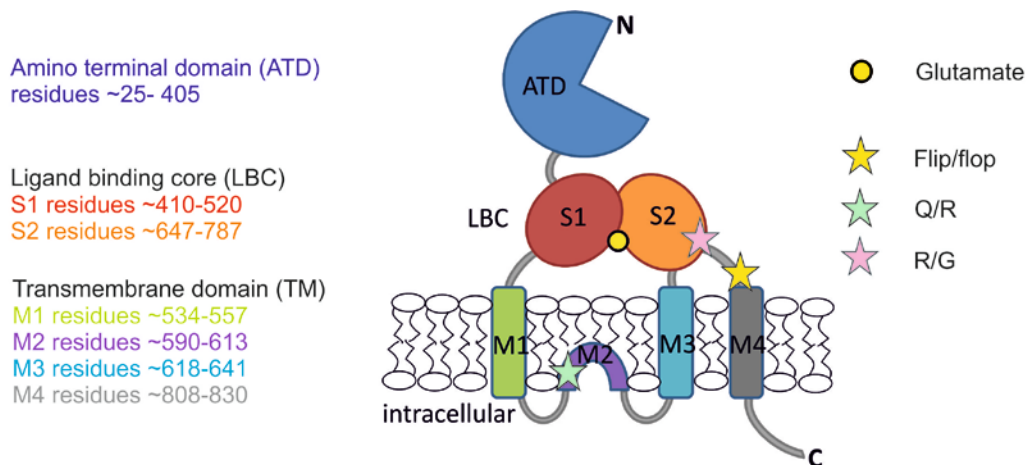


Figure 3.1 AMPA receptor schematic.

The AMPA receptor domain organisation includes the amino terminal domain (ATD), ligand binding core (LBC) with glutamate binding site indicated and transmembrane domain containing helices termed M1, 2, 3 & 4 (where M2 forms the receptor pore), each subunit also contains an intracellular C-terminal. The flip/flop, Q/R and R/G modification sites are also indicated.

The ATD is implicated in receptor trafficking (Xia *et al.* 2002), subunit assembly (Ayalon *et al.* 2001; Rossmann *et al.* 2011) and protein-protein interactions (discussed in Chapter 1). Examples of protein-protein interactions includes the postulated interaction with N-cadherin

stated to induce dendritic spine formation and growth in hippocampal neuronal cultures (Saglietti *et al.* 2007). The ATD also interacts with neuronal pentraxins to induce synapse formation (Passafaro *et al.* 2003; Sia *et al.* 2007) (discussed in Chapter 4) and other potential protein-protein interactions are currently under investigation (discussed in Chapter 5). In addition to these roles, the ATD of other iGluRs can bind small molecules which allosterically modulate channel conductance. For instance the ATD of the GluN2 subunit of NMDA receptors is known to bind Zn^{2+} at its interlobe cleft. Zn^{2+} binding locks the ATD in a closed conformation which exerts tension onto the linker connecting the ATD to the LBC and destabilises the LBC dimer interface. This relieves the strain that the LBC exerts onto the TM domain and allows the channel to close (Gielen *et al.* 2008). The NMDA receptor channel activity is therefore reduced and as such the ATD of the GluN2 subunit has been targeted as a binding site in the development of neuroprotective drugs (Karakas *et al.* 2009).

Whilst binding of glutamate to the LBC of iGluRs results in ion channel opening, glutamate binding to mGluRs results in conformational changes which activate their associated GTP-binding proteins (G-proteins) to regulate intracellular secondary messengers (Ozawa *et al.* 1998). Glutamate binds at the bi-lobed ligand binding region of the mGluR homodimer. This causes the two lobes to come together and rotate into a closed state and this architectural flexibility translocates through the protein and acts to stabilise the TM domain in an active conformation (Kunishima *et al.* 2000; Kew *et al.* 2005).

Despite domain deletion and mutagenesis studies, the full involvement of the AMPA receptor ATD in receptor modulation remains to be elucidated. The structural homology to bacterial and mGluR ligand binding domains and the fact that some NMDA receptor ATDs can bind known ligands which allosterically modulate the receptor led to the hypothesis that the AMPA receptor ATD might also share the ability to translate structural rearrangements into receptor modulation. High resolution X-ray crystallographic protein structures are expected to contribute considerable insights into this hypothesis.

3.2 Aims

The evolutionarily related bacterial LIVBP, mGluR1 and some iGluR ATDs have the ability to bind small molecules (ligands) at their interdomain cleft, so the hypothesis that the AMPA receptor ATD shares this capability will be investigated. The ATD is also postulated to have roles in receptor assembly, trafficking and in protein-protein interactions. Once structural insights are obtained here the project aims to characterise the AMPA receptor and investigate potential roles of AMPA receptor ATDs (their role in protein-protein interactions forms the focus of Chapters 4 and 5). The aims of this chapter are summarised below:

- 1) Structurally characterise the AMPA receptor GluA2_ATD using X-ray crystallography.
- 2) Examine whether the interdomain cleft of GluA2 ATD is an allosteric modulation site using biophysical analysis.
- 3) Determine the role of the AMPA receptor ATD in subunit specific assembly using MALS analysis.
- 4) Explain the role of the ATD in receptor trafficking using FACS and cellular analysis to compare surface expression levels of both full length and ATD truncated proteins.

3.3 Materials and methods

Information regarding accession numbers and cDNA sources for all AMPA receptor constructs of interest is detailed in Chapter 2, table 2.1. Information outlining all oligonucleotides and constructs amplified by PCR can be found in Appendix A. Amino acid residue numbering commences from the start of each protein's native secretion signal. Cloning details are split into sections below, where constructs received from collaborators are also indicated.

3.3.1 Construct design, PCR, and restriction-based cloning overview

Typically all AMPA receptor constructs were cloned into the pHLsec family of vectors (vector map Chapter 2, figure 2.1) by restriction-based cloning methods. Unless otherwise stated, all PCR amplified inserts were digested with restriction enzymes AgeI/BsiWI and all vectors with AgeI/Acc65I, thus generating compatible cohesive ends for ligation. Protocols for construct design, PCR, restriction digestion, ligation and transformation were as described in Chapter 2, section 2.1.

3.3.2 Generation of constructs

The mouse GluA1 ATD (GluA1_ATD) (GenBank accession number BC056397) construct (residues N23-T393) in the pHLsec vector was a gift from Ingo Greger (LMB, Cambridge). When problems with protein expression and purification arose, this construct was also cloned into pHLsec with a 1D4 tag (to improve protein purity), and in an attempt to improve protein yield it was cloned into the pBacPAK9 vector (Clontech) for expression in baculovirus systems.

A human GluA2 ATD (GenBank accession number NM_000826) construct, termed from hereon in as GluA2_ATD (residues N25-E412), was cloned into pHLsec and pHLsec_Avi vectors. To produce a monomeric form of GluA2_ATD (GluA2_M1), an F71D mutation was introduced into GluA2_ATD by site-directed mutagenesis. Additional GluA2_ATD mutant

constructs were also generated. These also targeted the dimer interface and aimed to disrupt hydrophobic interactions via the introduction of charged residues: GluA2_M2 (F103D), GluA2_M3 (L161E) and GluA2_M5 (L165E). An N-linked glycosylation site GluA2_M4 (T74N-A76S) was also introduced (see Appendix F- Supplementary figure 4 for mutant expression analysis). GRIA2_mouse cDNA (GenBank accession number BC048248) was used to generate a GluA2 full length Q-form (calcium permeable) construct termed GluA2_FL (residues N25-I883) with either an N-terminal haemagglutinin (HA) tag (construct L9) or FLAG tag (construct L11). A GluA2 Q-form construct with an ATD truncation termed GluA2_ΔATD (residues N406- I883) was also PCR amplified to incorporate either an N-terminal HA (construct L13) or FLAG tag (construct L15). Constructs were then ligated into pHLsec and pHLsec_mV vectors for FACS and cellular analysis. The ATD-D1 construct (AM212) was based on GenBank accession number BC010574. The construct contained residues N25-S796 where residues G520-D757 were replaced by one of two linker regions (DNSS or DDSS).

A mouse GluA3_ATD cDNA (GenBank accession number NM_016886) construct (G22-S403) in the pHLsec vector was a gift from Ingo Greger (LMB, Cambridge).

Initial attempts to generate the human GluA4_ATD construct (GenBank accession number BC167786) containing residues S25-T416 with an N414R mutation to remove predicted N-linked glycosylation site produced only small amounts of secreted proteins. A HEK293S G418-resistant stable cell line expressing GluA4_ATD was created by Ingo Greger (LMB, Cambridge). An aliquot of cells were kindly donated, which were grown and expanded to produce all GluA4_ATD protein required for experiments described here. RatGluA4_FL (residues S25-G893) and ratGluA4_ΔATD (residues N407-G893) were cloned into the pHLsec_mV vector for use in FACS and cellular analysis. All clones described here were sequence verified before initial small-scale transient expression attempts in HEK293T cells and analysis by SDS-PAGE/western blot as described in Chapter 2, section 2.3.2.

3.3.3 Large-scale protein expression and purification

Constructs were expressed by transient transfection of HEK293T and HEK293S cells as described in Chapter 2, section 2.2.2 unless otherwise stated. Proteins were purified from media after dialysis into PBS to prevent cobalt leaching from the TALON beads (Clontech). The proteins were purified by immobilised metal affinity chromatography (IMAC) and size exclusion chromatography (SEC) (described in Chapter 2, section 2.4). All SEC peak fractions were analysed by SDS-PAGE as described in Chapter 2, section 2.4. Because a suitable molecular replacement model was not available for GluA2_ATD, selenomethionine (SeMet) was incorporated into GluA2_ATD protein using the adapted protocol of HEK293T cell transient transfection (described in Chapter 2, section 2.2.3).

GluA4_ATD protein was expressed in the HEK293S GluA4_ATD stable cell line (G418 selective) and purified from dialysed conditioned medium using the same IMAC and SEC methods utilised for all other mammalian expressed secreted proteins (described in Chapter 2, section 2.4).

3.3.4 Crystallization, data collection and processing

This text has been removed due to copyright restrictions.

3.3.5 Structure determination and refinement

This text has been removed due to copyright restrictions.

3.3.6 Thermal shift assay (*thermofluor*)

Thermofluor was used to screen potential interaction ligands of the GluA2_ATD. Sypro orange dye (Sigma) (4µl of a 5000x stock) was diluted in 996µl HEPES buffer. The dye solution was added to the protein and the temperature increased from 0 °C to 100°C, the general principal being that as the protein is denatured its hydrophobic residues are exposed, which the dye then binds to resulting in an increased fluorescence (Pantoliano *et al.* 2001). Fluorescence was detected by an RT-PCR machine and the protein's melting temperature analysed. Protein stabilization was attempted using amino acids (10µM final concentration) and additives from the 'Silver bullet' crystallization screen (Hampton Research). Stabilisation was visualised as an increase in melting temperature (T_m) where 63 °C was obtained for the native protein.

3.3.7 FACS analysis

FACS was used to determine the role of the ATD in receptor trafficking. GluA2_FL, GluA2_ΔATD, GluA4_FL and GluA4_ ΔATD in pHLsec and pHLsec_mV (all including a HA tag) were expressed in HEK293T cells. Cell preparation and staining was as described in Chapter 2, section 2.9. Cell surface expression was detected by cell staining and FACS analysis carried out at the Weatherall Institute of Molecular Medicine (WIMM) in Oxford with assistance from Veronica Chang. Geometric mean values were analysed to enable the comparison of fluorescence intensities with the control data to give a degree of positivity measure.

3.4 Results

3.4.1 Structural analysis

Results describing the solved structure of the AMPA receptor GluA2 ATD have been published (Clayton *et al* 2009). Key features of the published work are summarised below with reference to Appendix F (published article with supplementary figures), to give context to the further developments that have since been made.

This text has been removed due to copyright restrictions.

**This figure has been removed due to
copyright restrictions**

Figure 3.2 This text has been removed due to copyright restrictions.

This text has been removed due to copyright restrictions.

This figure has been removed due to
copyright restrictions

Figure 3.3 This text has been removed due to copyright restrictions.

3.4.2 Investigation of the putative small molecule binding site

This text has been removed due to copyright restrictions.

This figure has been removed
due to copyright restrictions

Figure 3.4 This text has been removed due to copyright restrictions.

3.4.3 Attempts to crystallize larger ectodomain constructs

In addition to the GluA2_ATD and SeMet-labelled GluA2_ATD proteins crystallized, attempts were made to crystallize larger GluA2 ectodomain constructs for structural and functional analysis. A construct was designed to incorporate the S1 domain of the LBC and the ATD (construct ID AM212) (figure 3.5B) and was transiently expressed and purified from mammalian expression systems (figure 3.5C). The GluA2_ATD_D1 protein failed to crystallize but MALS analysis reveals masses corresponding to a dimer (130.14kDa), with indications that this protein can form higher order oligomers (242.77kDa), possibly tetramers if at a suitable concentration (figure 3.8B).

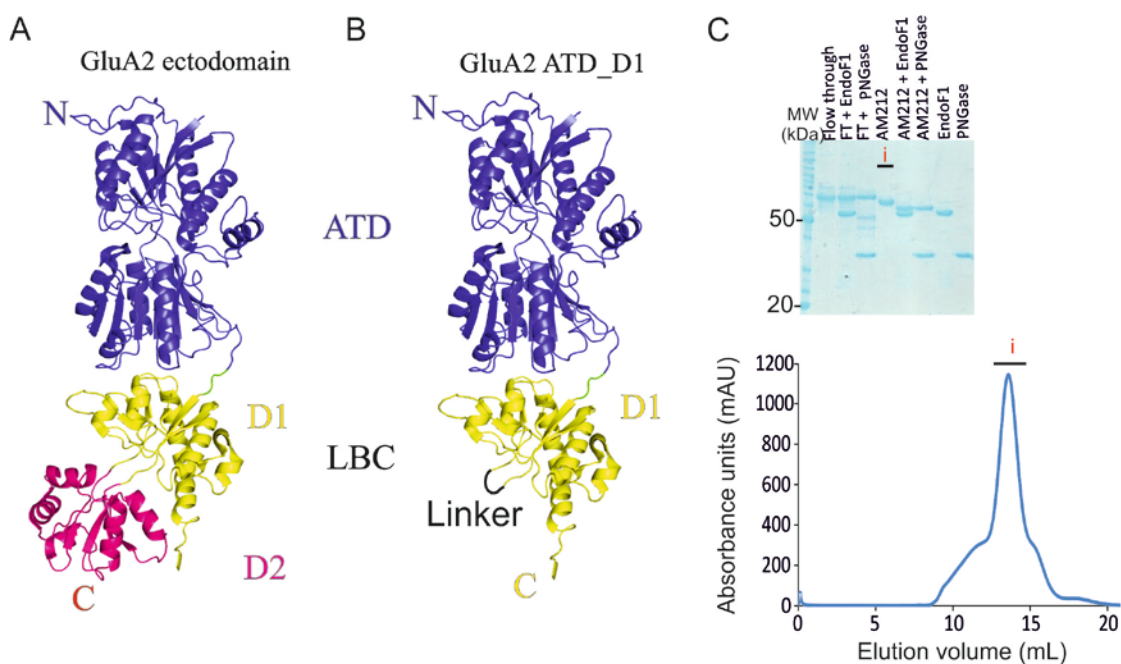


Figure 3.5 Generation of GluA2_ATD_D1.

A) The GluA2 ectodomain with the ATD (blue) and lobes D1 (yellow) and D2 (pink) of the LBC highlighted. B) Generation of GluA2_ATD_D1 protein with linker region (DDSS or DNSS) highlighted. Construct mapped to published structure (3KG2) (Sobolevsky *et al*, 2009) to show ATD in blue and two lobes D1 and D2 of LBC shown in yellow and pink respectively. C) Purification of GluA2_ATD_D1 from HEK293T cells +kifunensine by SEC. An S200 10/30 column and 10mM HEPES, pH7.5, 150mM NaCl running buffer were used. 12% SDS-PAGE analysis of the peak fraction with attempts to deglycosylate protein by EndoF1 for crystallization. Protein was also treated with PNGase to confirm EndoF1 deglycosylation had been successful. Molecular weight marker (MW) Benchmark (Invitrogen).

3.4.4 Understanding the ATDs role in AMPA receptor assembly

To investigate the role of the ATD in receptor assembly (and protein-protein interactions in Chapter 4 and 5) the GluA1-4 ATDs were transiently expressed in mammalian expression systems. Purification of GluA2, GluA3 and GluA4 ATDs generated good amounts of protein (>1mg/L) and SDS-PAGE confirmed the purity obtained was suitable for use in structural, biophysical and cellular analysis (figure 3.6B-D).

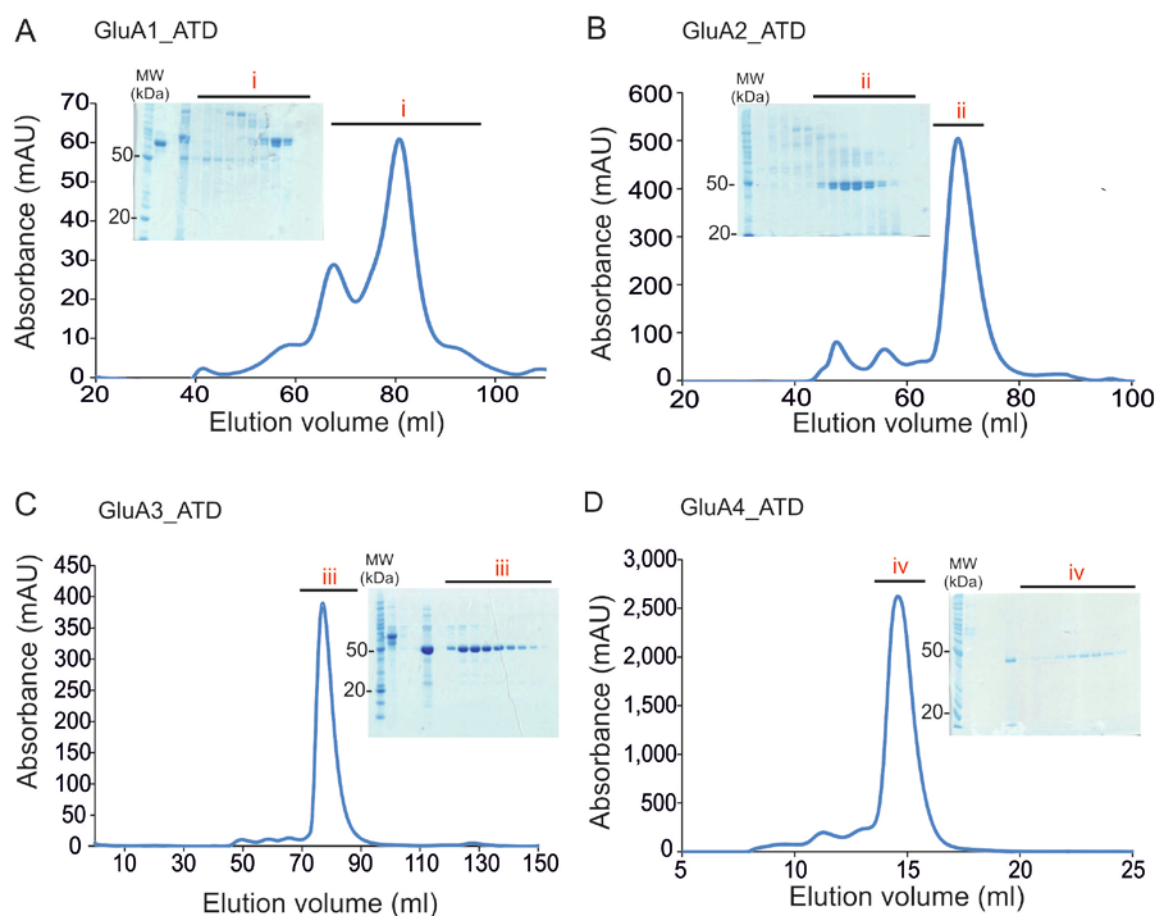


Figure 3.6 SEC and SDS-PAGE analysis of AMPA receptor ATD proteins purified from HEK293S cells.

A) GluA1_ATD, B) GluA2_ATD, C) GluA3_ATD and D) GluA4_ATD (from HEK293S stable cell line). SEC purification of GluA1, 2 and 3_ATDs were performed on an S200 16/60 column and an S200 10/30 column for GluA4_ATD. Gel filtration running buffer was 10mM HEPES pH 7.5, 150mM NaCl. The observed SEC peak fractions were then analysed by 12% SDS-PAGE gels where the gel loading order was as follows; Benchmark (Invitrogen) protein molecular weight marker (MW), Talon IMAC flow through, 5mM Imidazole wash from Talon IMAC, 1µl of SEC injected sample and then the peak SEC fractions as indicated by Roman numerals.

Expression and purification of GluA1_ATD proved problematic and purified proteins were unstable, weakly expressed and prone to aggregation (figure 3.6A). The addition of a 1D4

tag did not improve protein purity or yield, so expression was attempted in the baculovirus system which was used to produce secreted GluA1_ATD protein (from *Spodoptera frugiperda* Sf9 cells) which had been used to crystallize and solve a GluA1_ATD structure (Yao *et al.* 2011). Here, western blot analysis of small-scale expression trials in sf9 cells revealed a secreted band present in the media fraction (figure 3.7). A large-scale sf9 cell infection was therefore carried out (by Yuguang Zhao and Weixian Lu, STRUBI) (uncalculated multiplicity of infection (MOI)) but unfortunately the result was not reproducible in large-scale where only a 20mAU peak for purified secreted protein was observed in SEC (not shown). Infection of insect cells with baculovirus can require extensive optimisation and so future work to improve secretion levels will be attempted.

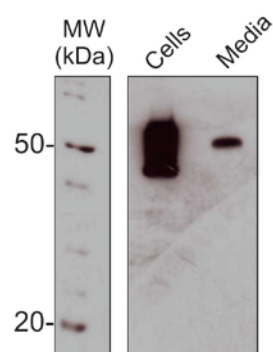
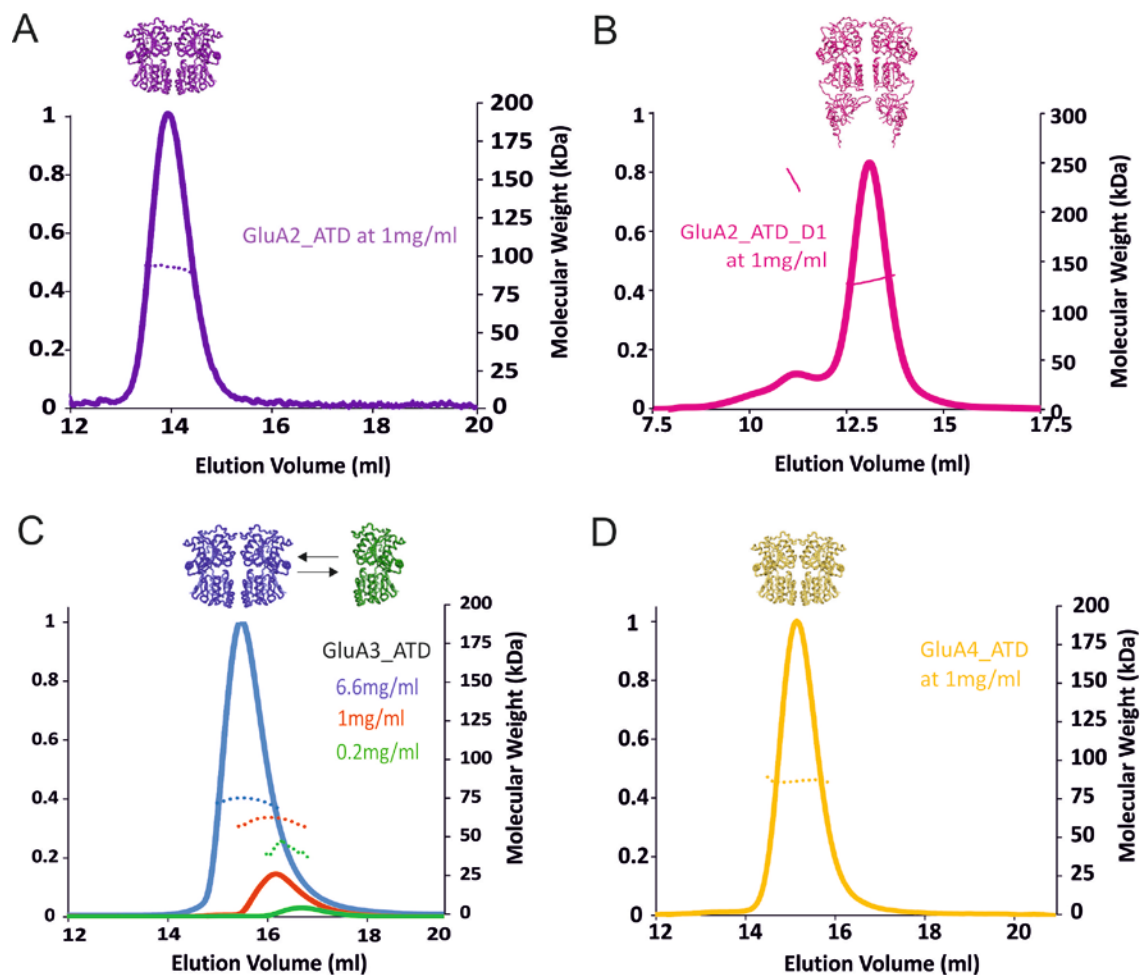


Figure 3.7 R1_ATD baculovirus small-scale expression test in sf9 cells.

Sf9 insect cells were infected with baculovirus expressing GluA1_ATD in the pBacPAK9 vector. At 48hours post infection the cell suspension was harvested and centrifuged for 5 mins at 1000 xg. Cell pellet and secreted (media) fractions were analysed by 12 % SDS-PAGE/ anti-His western blot. Molecular weight marker (MW) Benchmark (Invitrogen) protein ladder.

To investigate the role of the AMPA receptor ATD in subunit specific assembly, the oligomeric states of several AMPA receptor ectodomain constructs were determined by MALS analysis (figure 3.8). GluA2_ATD (figure 3.8A) and GluA4_ATD (figure 3.8D) were shown to have masses of 92.18kDa and 87.01kDa respectively, indicating that these proteins form dimers in solution (at 1mg/ml). The GluA3_ATD (figure 3.8C) was also analysed at 1mg/ml and a mass of 62.41 kDa was obtained. Because this mass accounted neither for a monomeric or dimeric form it was postulated that the protein was in equilibrium between the two states. To confirm this, MALS was repeated with the protein at 6.6mg/ml and 0.2mg/ml which indicated masses of

75.43kDa and 43.54kDa respectively and the lower concentration confirmed that a monomeric form of the protein was present.



Protein	Injected conc.	Polydispersity (Mw/Mn)	Molar mass Mn (kDa)
GluA2_ATD	1mg/ml	1.000 (1%)	92.18 (0.9%)
GluA2_ATD_D1	1mg/ml	1.001 (0.3%)	130.14 (0.2%)
GluA3_ATD	0.2mg/ml	1.005 (3%)	43.54 (2%)
GluA3_ATD	1mg/ml	1.001 (2%)	62.41 (1%)
GluA3_ATD	6.6mg/ml	1.001 (1%)	75.43 (0.9%)
GluA4_ATD	1.0mg/ml	1.000 (1%)	87.01 (0.8%)

Figure 3.8 MALS analysis of AMPA receptor ectodomain constructs.

MALS analysis of the various AMPA receptor proteins: A) GluA2_ATD (AM21) at 1.0 mg/ml, B) GluA2_ATD_D1 (AM212) injected at 1.0 mg/ml, C) GluA3_ATD injected at 0.2, 1.0 and 6.6 mg/ml and D) GluA4_ATD injected at 1mg/ml. Proteins were analysed on an S200 10/30 column in a running buffer containing 10mM HEPES pH 7.4, 150mM NaCl. Schematic representation of the corresponding protein oligomeric state is depicted above each curve. Table indicates MALS molar mass values calculated using the ASTRA software package and the Mn is the number average molar mass. The % value in paratheses is the uncertainty in this value as calculated by ASTRA. Where the uncertainty is a measure of the statistical consistency of the data.

3.4.5 Understanding the ATDs role in AMPA receptor trafficking

To determine whether the ATD is required for AMPA receptor trafficking FACS analysis was performed and cell surface levels of GluA2_FL and GluA2_ΔATD proteins were examined. Analysis indicates that the cell surface expression levels for GluA2_ΔATD are higher in both GluA2 Q (calcium permeable) and R forms (figure 3.9A). This was also confirmed by cellular analysis where COS7 cell surface detection clearly demonstrates that both proteins were able traffic to the surface (figure 3.9B). FACS analysis with GluA4_FL and GluA4_ΔATD also provided the same result where a larger amount of GluA4_ΔATD protein was observed at the cell surface compared to GluA4_FL levels (Chapter 4, figure 4.19C & D).

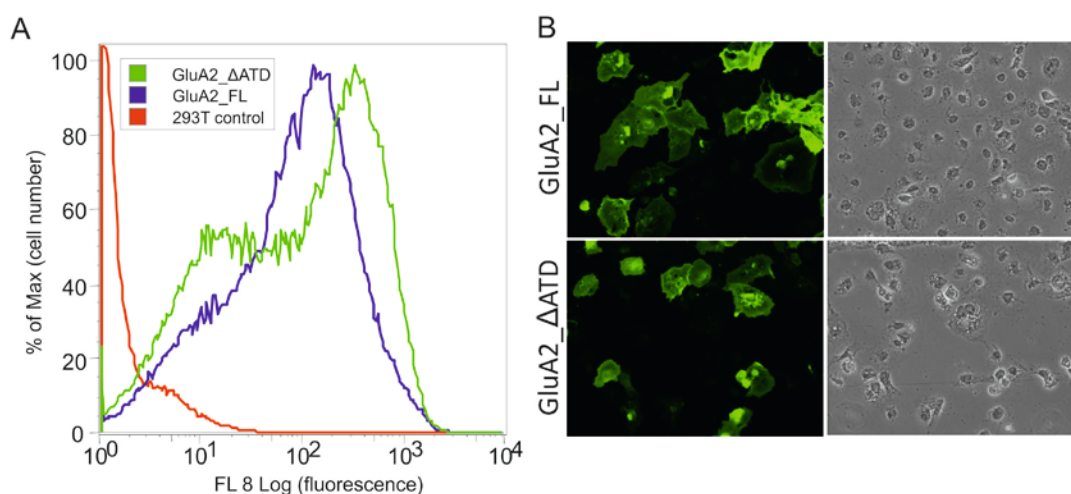


Figure 3.9 FACS analysis of GluA2.

A) HEK293T cells were transiently transfected with GluA2_FL and GluA2_ΔATD. Proteins were expressed with an N-terminal HA-tag and FACS analysis enabled anti-HA detection of surface protein levels. Data was processed using FloJo software. B) Complementary cellular analysis using anti-HA antibody surface detection of GluA2_FL and GluA2_ΔATD that were transiently expressed in COS7 cells. Data shown is representative of data obtained for the GluA2 Q form (calcium permeable), but the GluA2 R forms exhibits the same trend.

3.5 Discussion

This text has been removed due to copyright restrictions.

Since 2009, structures for all AMPA receptor ATDs have been solved. When the first 100 residues of GluA2_ATD (PDB ID: 2WJW) (Clayton *et al.* 2009), GluA3_ATD (PDB ID: 3O21) (Sukumaran *et al.* 2011) and GluA4_ATD (PDB ID: 4GPA) (Dutta *et al.* 2012) are aligned to the GluA1_ATD (PDB ID: 3SAJ) (Yao *et al.* 2011) using SHP (Stuart *et al.* 1979) and the rest of each molecule aligned to this matrix, a movement or potential flexibility is observed in the D2 domain (figure 3.10).

The GluA3_NTD structure reveals a large degree of previously unseen flexibility in the lower (D2) lobes of the ATD interlobe cleft, where movements of up to 8Å were observed (Sukumaran *et al.* 2011) which is more reminiscent of the mGluR1 domain arrangement (Kunishima *et al.* 2000).

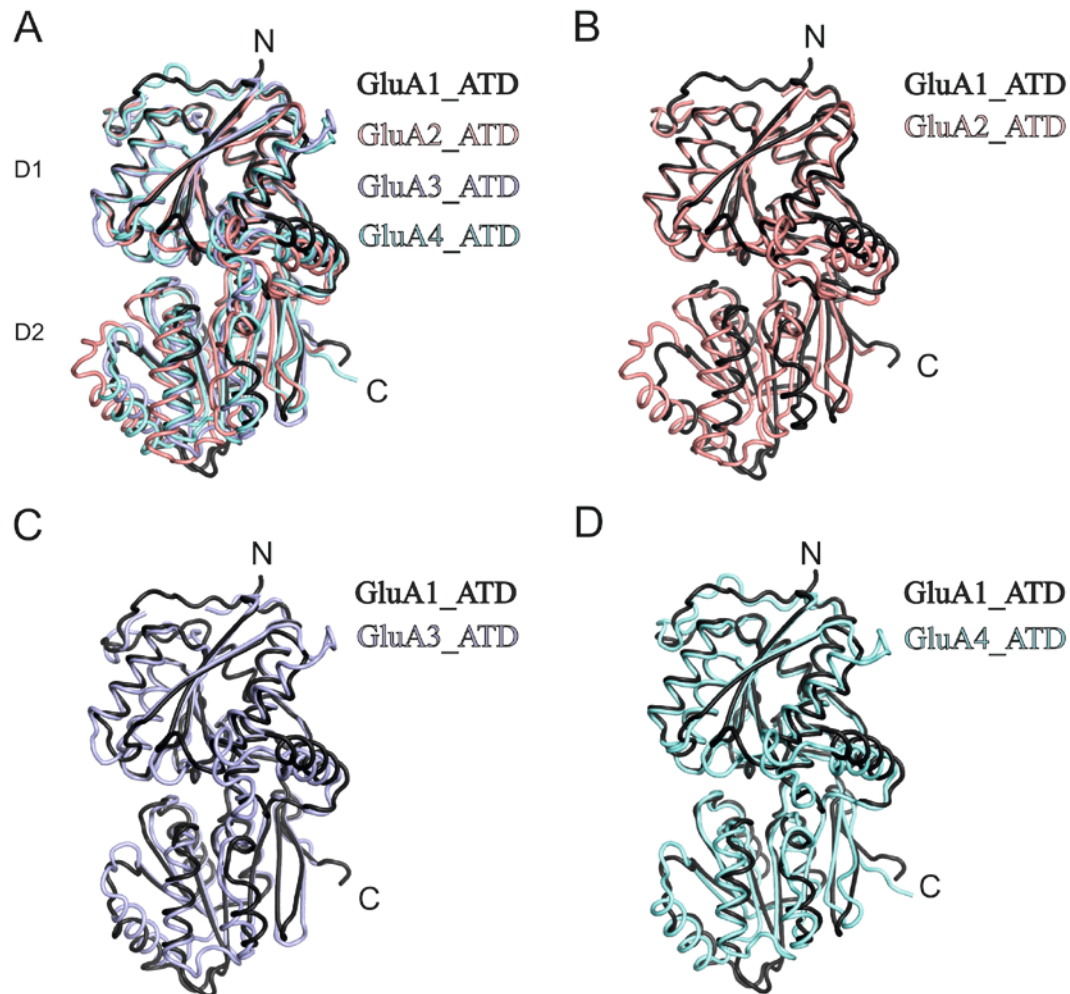


Figure 3.10 AMPA GluA1-4_ATD structural alignment.

Amino terminal domain (ATD) of AMPA receptors GluA1 (grey), GluA2 (pink), GluA3 (lilac) and GluA4 (blue) subunits. The proteins were aligned to GluA1 ATD using SHP (Stuart *et al.* 1979) based on the first 100 residues of the D1 lobe. The rest of the domain was then orientated based on this matrix. Structures depicted as cartoon putty in pymol (DeLano, 2002). N- and C-terminal residue locations indicated.

Recent advancements have demonstrated that the AMPA receptor ATDs act as either obligate or preferential hetero-tetramers (Sukumaran *et al.* 2012) and although these remain to be structurally characterised, many iGluR ATD hetero-complex structures have recently been solved. The heterodimer complex of Kainate receptor GluK2/GluK5 subunits reveals that the GluK5 ATD contains a glycan moiety at the interdomain cleft which acts as a steric block to prevent full domain closure (Kumar *et al.* 2011). The structure of the NMDA receptor GluN1/GluN2B heterodimer identified the molecular requirements, present only in heterodimers, that are compulsory for binding of the neuroprotective drug phenylethanolamine and demonstrated that conformational freedom in the ATD is essential (Karakas *et al.* 2011).

This text has been removed due to copyright restrictions.

The GluA2 ATD has since been solved again to a higher resolution and was also initially thought to contain unidentified density at the interdomain cleft, (Sukumaran *et al.*, 2011). However the structure has since been reanalysed and suggests that the density can be accounted for by a phosphate ion and water molecules (Mayer 2011).

Additional putative binding sites involving the AMPA receptor ATD should also be considered since more binding sites in other iGluRs have recently been identified. For example it was initially proposed that Zn^{2+} bound in the hydrophobic pocket of the NMDA GluN2 ATD interdomain cleft but the structure of GluN2 in complex with zinc has since revealed that Zn^{2+} does not bind quite as expected and actually clamps the outer edge of the interdomain cleft, locking the domain in a closed conformation (Karakas *et al.* 2009). In addition to this site, the NMDA ATD can be allosterically modulated by the NMDA receptor inhibitor termed Ifenprodil and the related phenylethanolamine compounds which were known to inhibit GluN1 and GluN2B. The cleft had again been suggested to be the binding site (Karakas *et al.* 2009) but the compounds have since been shown to bind to a novel site at the hetero-dimer interface between the GluN1 and GluN2 ATDs (Karakas *et al.* 2011). Whether the AMPA receptor ATD is able to bind small molecules and allosterically modulate the receptor is still to be defined. Work to identify the potential ATD binding molecules was continued in an MRC-tech grant by collaborators (Ingo Greger, LMB Cambridge) and has since expanded to examine GluA3_ATD as a potential target.

The 4.1Å structure presented here confirms the hypothesis that the ATD is involved in receptor assembly by indicating that the ATD can form higher lever oligomers and was also confirmed by AUC analysis (Appendix F). The AMPA receptor GluA2, GluA3 and GluA4 ATDs were analysed by MALS and confirmed the ability to form homodimers in solution although the affinity for GluA3_ATD homodimers was weaker in agreement with previously published data (Rossmann *et al.* 2011). The ATDs of NMDA and kainate receptors have also been solved. These iGluRs function as obligate heterodimers which can be elegantly explained by the fact that the equivalent ATDs interact very weakly whereas the binding affinity for heterodimer assembly is 1000 fold stronger (Kumar *et al.* 2013).

Finally, the ATD had previously been stated to be involved in receptor trafficking (Xia *et al.* 2002), where previous research has demonstrated that reduced GluA1 surface expression levels were detectable when N-terminal truncations were made to replace residues ANFPNNIQI (residues A19-I27) with a FLAG tag. Authors implied that the receptors do not leave the endoplasmic reticulum (ER) because the ATD is required for trafficking (Xia *et al.* 2002). Structural analysis provided here would suggest that truncating 9 residues from the N-terminal would encroach into the ATD so it is perhaps not surprising the protein was trapped in the ER, as would be the case for a protein that had not correctly folded or had become aggregated. Here AMPA receptor trafficking was actually increased for proteins lacking their ATDs. The increased cell surface expression detected might simply be attributed to the ΔATD proteins being much smaller than the FL versions, so more protein can be made in the time frames observed.

3.6 Future work

Several attempts to purify GluA1_ATD with a variety of purification tags and mammalian and baculovirus expression methods were unfortunately unsuccessful. A structure of GluA1_ATD expressed in an insect cell line was able to generate enough protein to crystallize and solve the structure (Yao *et al.* 2011) but >90% of the protein was lost during the purification procedure (personal communication), implying that the protein is inherently unstable. Because many of the AMPA receptor ATD crystals grew in low pH conditions (GluA1_ATD pH 5.0 (Yao *et al.* 2011), GluA2 data presented here for crystal form 1 pH4.6, GluA3_ATD pH4.6 (Sukumaran *et al.* 2011) and GluA4_ATD pH5.0-7.0 (Dutta *et al.* 2012)) one might hypothesize that purification attempts using buffer solutions in this range (and not the standard pH7.4-8.0 routinely utilised) might stabilise the protein and reduce the observed precipitation and aggregation.

3.7 References

- Adams, P. D., R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger (2002). "PHENIX: building new software for automated crystallographic structure determination." Acta Crystallographica Section D **58**(11): 1948-1954.
- Ayalon, G. and Y. Stern-Bach (2001). "Functional assembly of AMPA and kainate receptors is mediated by several discrete protein-protein interactions." Neuron **31**(1): 103-113.
- Baker, N. A., D. Sept, S. Joseph, M. J. Holst and J. A. McCammon (2001). "Electrostatics of nanosystems: Application to microtubules and the ribosome." Proceedings of the National Academy of Sciences of the United States of America **98**(18): 10037-10041.
- Clayton, A., C. Siebold, R. J. C. Gilbert, G. C. Sutton, K. Harlos, R. A. J. McIlhinney, E. Y. Jones and A. R. Aricescu (2009). "Crystal Structure of the GluR2 Amino-Terminal Domain Provides Insights into the Architecture and Assembly of Ionotropic Glutamate Receptors." Journal of Molecular Biology **392**(5): 1125-1132.
- Dutta, A., I. H. Shrivastava, M. Sukumaran, I. H. Greger and I. Bahar (2012). "Comparative dynamics of NMDA- and AMPA-glutamate receptor N-terminal domains." Structure **20**(11): 1838-1849.
- Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallographica Section D **60**(12 Part 1): 2126-2132.
- Gielen, M., A. Le Goff, D. Stroebel, J. W. Johnson, J. Neyton and P. Paoletti (2008). "Structural Rearrangements of NR1/NR2A NMDA Receptors during Allosteric Inhibition." Neuron **57**(1): 80-93.
- Hooft, R. W. W., G. Vriend, C. Sander and E. E. Abola (1996). "Errors in protein structures [3]." Nature **381**(6580): 272.
- Karakas, E., N. Simorowski and H. Furukawa (2009). "Structure of the zinc-bound amino-terminal domain of the NMDA receptor NR2B subunit." EMBO Journal **28**(24): 3910-3920.
- Karakas, E., N. Simorowski and H. Furukawa (2011). "Subunit arrangement and phenylethanolamine binding in GluN1/GluN2B NMDA receptors." Nature **475**(7355): 249-253.
- Kew, J. N. C. and J. A. Kemp (2005). "Ionotropic and metabotropic glutamate receptor structure and pharmacology." Psychopharmacology **179**(1): 4-29.
- Kumar, J. and M. L. Mayer (2013). Functional insights from glutamate receptor ion channel structures. **75**: 313-337.
- Kumar, J., P. Schuck and M. Mayer (2011). "Structure and Assembly Mechanism for Heteromeric Kainate Receptors." Neuron **71**(2): 319-331.
- Kunishima, N., Y. Shimada, Y. Tsuji, T. Sato, M. Yamamoto, T. Kumasaka, S. Nakanishi, H. Jingami and K. Morikawa (2000). "Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor." Nature **407**(6807): 971-977.
- Lovell, S. C., I. W. Davis, W. B. Arendall Iii, P. I. W. De Bakker, J. M. Word, M. G. Prisant, J. S. Richardson and D. C. Richardson (2003). "Structure validation by α geometry: ϕ , ψ and $C\beta$ deviation." Proteins: Structure, Function and Genetics **50**(3): 437-450.
- Matsuda, S., Y. Kamiya and M. Yuzaki (2005). "Roles of the N-terminal domain on the function and quaternary structure of the ionotropic glutamate receptor." Journal of Biological Chemistry **280**(20): 20021-20029.
- Mayer, M. L. (2011). "Emerging models of glutamate receptor ion channel structure and function." Structure **19**(10): 1370-1380.
- McCoy, A. J., R. W. Grosse-Kunstleve, L. C. Storoni and R. J. Read (2005). "Likelihood-enhanced fast translation functions." Acta Crystallographica Section D: Biological Crystallography **61**(4): 458-464.
- Nakagawa, T., Y. Cheng, E. Ramm, M. Sheng and T. Walz (2005). "Structure and different conformational states of native AMPA receptor complexes." Nature **433**(7025): 545-549.

- Otwinowski, Z. and W. Minor (1997). Processing of X-ray diffraction data collected in oscillation mode. **276**: 307-326.
- Ozawa, S. (1996). "Permeation of calcium through glutamate receptor channels." Seminars in the Neurosciences **8**(5): 261-269.
- Ozawa, S., H. Kamiya and K. Tsuzuki (1998). "Glutamate receptors in the mammalian central nervous system." Progress in Neurobiology **54**(5): 581-618.
- Pantoliano, M. W., E. C. Petrella, J. D. Kwasnoski, V. S. Lobanov, J. Myslik, E. Graf, T. Carver, E. Asel, B. A. Springer, P. Lane and F. R. Salemme (2001). "High-density miniaturized thermal shift assays as a general strategy for drug discovery." Journal of Biomolecular Screening **6**(6): 429-440.
- Passafaro, M., T. Nakagawa, C. Sala and M. Sheng (2003). "Induction of dendritic spines by an extracellular domain of AMPA receptor subunit GluR2." Nature **424**(6949): 677-681.
- Perrakis, A., R. Morris and V. S. Lamzin (1999). "Automated protein model building combined with iterative structure refinement." Nature Structural Biology **6**(5): 458-463.
- Rossmann, M., M. Sukumaran, A. C. Penn, D. B. Veprintsev, M. M. Babu and I. H. Greger (2011). "Subunit-selective N-terminal domain associations organize the formation of AMPA receptor heteromers." EMBO J **30**(5): 959-971.
- Saglietti, L., C. Dequidt, K. Kamieniarz, M. C. Rousset, P. Valnegri, O. Thoumine, F. Beretta, L. Fagni, D. Choquet, C. Sala, M. Sheng and M. Passafaro (2007). "Extracellular Interactions between GluR2 and N-Cadherin in Spine Regulation." Neuron **54**(3): 461-477.
- Schneider, T. R. and G. M. Sheldrick (2002). "Substructure solution with SHELXD." Acta Crystallographica Section D: Biological Crystallography **58**(10 I): 1772-1779.
- Sia, G. M., J. C. Beique, G. Rumbaugh, R. Cho, P. F. Worley and R. L. Huganir (2007). "Interaction of the N-terminal domain of the AMPA receptor GluR4 subunit with the neuronal pentraxin NP1 mediates GluR4 synaptic recruitment." Neuron **55**(1): 87-102.
- Stuart, D. I., M. Levine, H. Muirhead and D. K. Stammers (1979). "Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å." Journal of Molecular Biology **134**(1): 109-142.
- Sukumaran, M., A. Penn and I. Greger (2012). AMPA Receptor Assembly: Atomic Determinants and Built-In Modulators. Synaptic Plasticity. M. R. Kreutz and C. Sala, Springer Vienna. **970**: 241-264.
- Sukumaran, M., M. Rossmann, I. Shrivastava, A. Dutta, I. Bahar and I. H. Greger (2011). "Dynamics and allosteric potential of the AMPA receptor N-terminal domain." EMBO J **30**(5): 972-982.
- Vonrhein, C., E. Blanc, P. Roversi and G. Bricogne (2007). Automated structure solution with autoSHARP. **364**: 215-230.
- Xia, H., M. Von Zastrow and R. C. Malenka (2002). "A novel anterograde trafficking signal present in the N-terminal extracellular domain of ionotropic glutamate receptors." Journal of Biological Chemistry **277**(49): 47765-47769.
- Yao, G., Y. Zong, S. Gu, J. Zhou, H. Xu, I. I. Mathews and R. Jin (2011). "Crystal structure of the glutamate receptor GluA1 N-terminal domain." Biochemical Journal **438**(2): 255-263.

Chapter 4

Structural and functional analysis of the Neuronal Pentraxins

4.1 Introduction

Neuronal pentraxins are released from the pre-synaptic neuronal compartment and were previously shown to cluster post-synaptic AMPA receptors (figure 4.1) (O'Brien *et al.* 1999; O'Brien *et al.* 2002) (see Chapter 1). The three neuronal pentraxins are termed neuronal pentraxin 1 (NP1), neuronal pentraxin 2 (NP2) (also known as neuronal activity regulated protein (NARP)) and neuronal pentraxin receptor (NPR). NP1 is secreted and stated to be constitutively expressed (Schlimgen *et al.* 1995), while NP2 is a secreted immediate-early response gene (IEG) that exhibits activity-regulated expression (Hsu *et al.* 1995). NPR is N-terminally attached to the pre-synaptic membrane but its extracellular region is predicted to contain a protease cleavage site to enable release of a secreted form (Dodds *et al.* 1997). Because the expression of these structurally similar proteins is regulated by different mechanisms, it alludes to them having specialised, independent roles at the synapse.

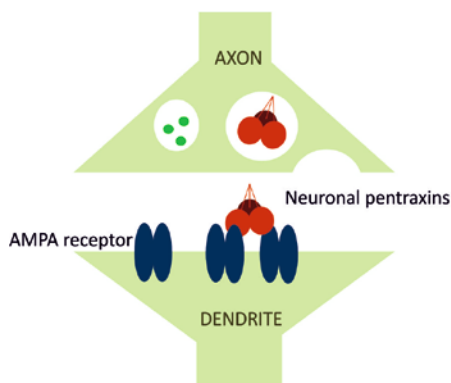


Figure 4.1 AMPA receptor clustering by neuronal pentraxins. The oligomeric neuronal pentraxins (red) are axonally expressed and cluster AMPA receptors (blue) on the post-synaptic membrane.

The neuronal pentraxin C-terminal pentraxin (PTX) domain shows a high level of sequence homology to other members of the pentraxin family such as C-reactive protein (CRP) and serum amyloid-P (SAP) (figure 1.10) that are both involved in innate immunity clearance processes (Garlanda *et al.* 2005). Like all pentraxins, the neuronal pentraxins have lectin-like properties and bind carbohydrates in a calcium-dependant manner (Garlanda *et al.* 2005). Sequence alignment indicates that neuronal pentraxins are able to act as lectins because the key calcium co-ordinating residues present in CRP and SAP protein structures are conserved (blue hexagons, figure 4.2). In addition to the PTX domain, all neuronal pentraxins have an N-terminal

extension, predicted to contain coiled-coil regions (green boxes figure 4.2) that enable the protein to form oligomeric homo- or hetero-complexes which then cluster the AMPA receptors (figure 4.1) (Dodds *et al.* 1997; Kirkpatrick *et al.* 2000; Xu *et al.* 2003). Previous neuronal pentraxin expression attempts have not been able to provide any structural information (Xu *et al.* 2003) (figure 4.3A) but we hypothesize that the PTX domain will show structural homology to other pentraxin proteins. To obtain higher order oligomeric structural information, the expression and purification of homogenous full length neuronal pentraxins is required.

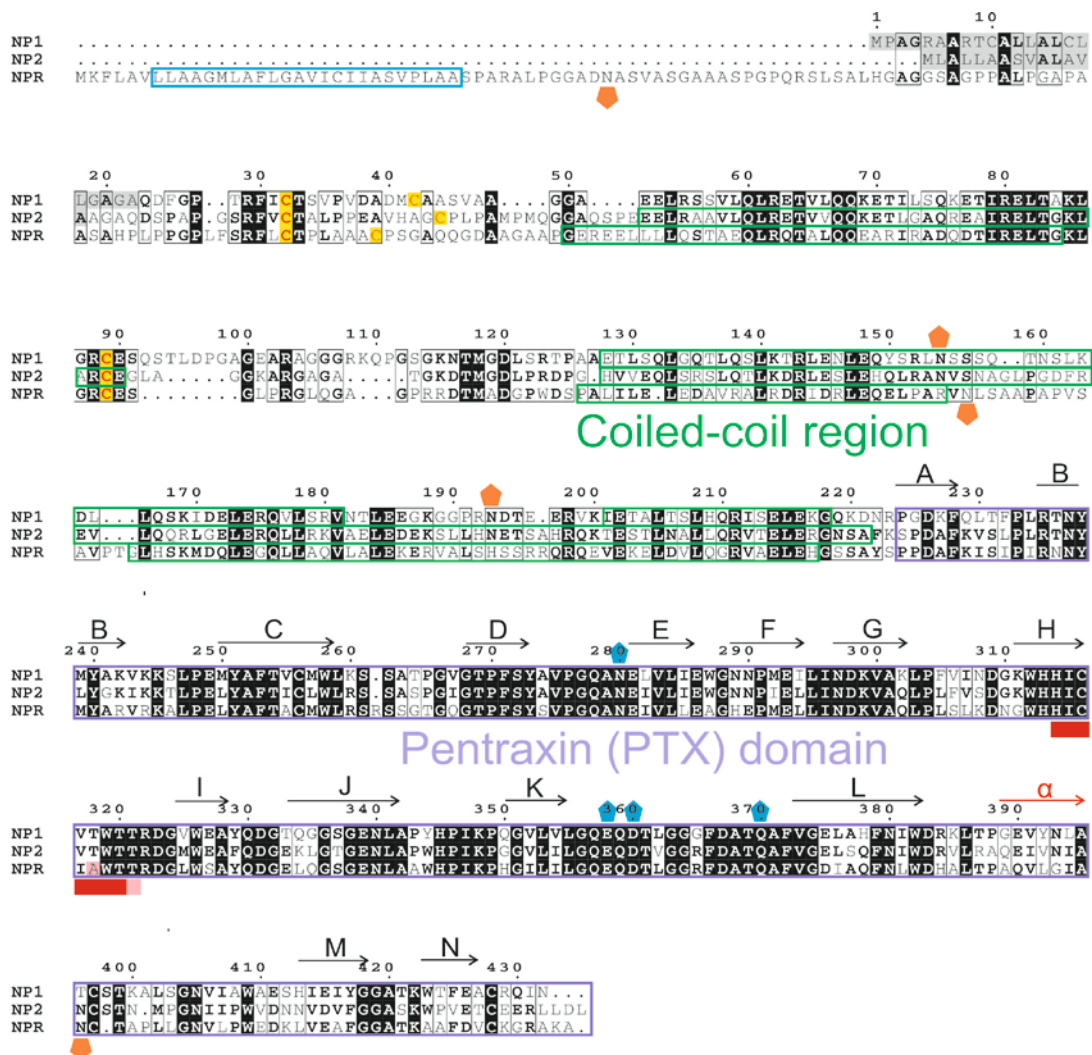


Figure 4.2 Sequence alignment of human neuronal pentraxins NP1, NP2 and NPR.

Sequences were aligned using MULTALIN (<http://multalin.toulouse.inra.fr/multalin/>). Predicted pentraxin domains (PTX) are highlighted by lilac boxes. NetNGlyc (<http://www.cbs.dtu.dk/services/NetNGlyc/>) -predicted N-linked N-acetylglucosamine (GlcNAc) glycosylation sites are highlighted with orange pentagons. Cysteine residues that potentially form disulphide bridges are highlighted in yellow with red text. Predictions of potential coiled-coil regions by COILS software (http://www.ch.embnet.org/software/COILS_form.html) are highlighted by green boxes. Signal sequence predictions (http://bmbpcu36.leeds.ac.uk/cgi-bin/sig_pred/Signal.cgi) are highlighted in grey. The NPR trans-membrane region is indicated by the light blue box. Conserved calcium co-ordinating residues are highlighted by blue pentagons. The red underlined region indicates the 'pentraxin motif' where pink indicates variation from motif. Secondary structure prediction based on C-reactive protein (Shrive *et al.* 1996) is indicated where Beta sheets are labelled A-N (black arrows) and the alpha helix labelled α (red arrow).

In hippocampal neurons, post-synaptic spine size and density can be increased by over-expression of AMPA receptors and is postulated to be triggered by a receptor-ligand interaction involving the AMPA receptor GluA2 amino-terminal domain (ATD) (Passafaro *et al.* 2003). Additional research identified that these ligands were in fact the neuronal pentraxins which were shown to interact with and cluster the AMPA receptor ATDs to induce synapse formation (O'Brien *et al.* 1999; Sia *et al.* 2007). Interactions between AMPA receptors and neuronal pentraxins are postulated to occur at perineuronal net (PNN)-coated GABAergic parvalbumin expressing interneurons (PV-INs), that act to dampen excitation of hippocampal pyramidal cells. Developmental 'critical periods' of learning are reduced around the time that PNNs form (Kwok *et al.* 2011) (see Chapter 1). Neuronal pentraxins were shown to localise to PNN-coated regions at excitatory inputs on PV-INs whereas AMPA receptors were shown to localise to but do not directly interact with PNNs (Chang *et al.* 2010). In NARP (NP2) knock out (KO) mice, the net excitatory drive onto PV-INs is reduced, resulting in hyper-excitability in the developing visual system (Gu *et al.* 2013). We propose that neuronal pentraxin-induced clustering of AMPA receptor may have one of two effects:

- 1) Synaptically localise and cluster the AMPA receptor which might trigger receptor internalisation.
- 2) Synaptically localise and cluster the AMPA receptor, and enhance activation of the mature GABAergic PV-IN, suppressing hippocampal excitation (figure 4.3C).

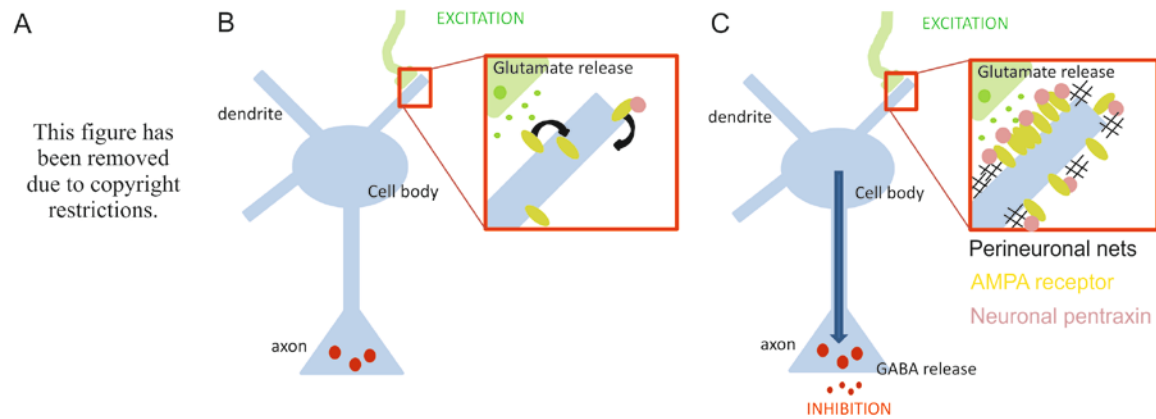


Figure 4.3 The interaction between AMPA receptors and neuronal pentraxins- a hypothetical mechanism.

A) This text has been removed due to copyright restrictions. B) A parvalbumin-expressing interneuron (PV-IN) during a 'window of learning' (critical period) whereby excitation results in AMPA receptor (yellow) internalisation and no activation of the neuron. C) A mature inhibitory neuron where the presence of perineuronal nets (PNNs) enables the neuronal pentraxins (pink) to anchor the AMPA receptor to the synapse, leading to activation of the neuron, the release of the neurotransmitter GABA and subsequent depression of any downstream excitatory pyramidal neuronal contacts in the hippocampus.

4.2 Aims

The structure of the neuronal pentraxins, their physiological oligomeric state(s), the purpose of their self-multimerization, how they interact with the AMPA receptors and why the neuronal pentraxins have retained the ability to act as lectins within a neuronal system are all currently unanswered questions. The aims of this thesis are to address these questions using methods summarised below.

- 1) Investigate the hypothesis that the neuronal pentraxin PTX domain will show homology to other pentraxins by solving protein structures using X-ray crystallography.
- 2) Determine the oligomeric states of the neuronal pentraxins by multi-angle light scattering (MALS) and electron microscopy (EM).
- 3) Analyse how AMPA receptors interact with the neuronal pentraxins by biophysical analysis (surface plasmon resonance (SPR)) and the effect of this interaction by fluorescence activated cell sorting (FACS).

4.3 Materials and methods

4.3.1 Construct design, PCR amplification and restriction-based cloning

All neuronal pentraxin constructs were cloned into the pHLsec and pHLsec_Avi vectors between the AgeI and KpnI restriction sites to utilise the pHLsec secretion signal (described in Chapter 2, section 2.1). All amino acid residue numbering commences from the start of each protein's native secretion signal. Oligonucleotides (Appendix B, Table B2) were designed to PCR amplify constructs of interest as described below and are outlined in Appendix B Table B1.

4.3.2 Generation of neuronal pentraxin 1 constructs

All human neuronal pentraxin 1 (NP1) constructs were based on GenBank entry BC089441. In summary, the pentraxin (PTX) domain construct NP1_PTX (P224-I431) was based on sequence similarity to other known pentraxin PTX structures (figure 1.10). A full length (FL) construct NP1_FL (Q23-I431) was also generated but its proteolysis led to the generation of an N-terminally truncated construct NP1_CC2 (G112-I431), design of which was based on mass spectrometry analysis of a cleavage product frequently occurring during NP2_FL purification.

Sequence information for NP1, highlighting construct boundary residues, N-linked glycosylation sites and targeted mutagenesis residues is outlined in figure 4.4. The C-terminal residue (N432) was not included in the construct because ligation into the pHLsec vector via the KpnI restriction site would generate an additional N-linked glycosylation motif (NGT).

MPAGRAARTCALLALCLLGAGADFGPTRFICTSVVPVDADMCAASVAAGGAELRSSVLQLRETV
 LQQKETILSQKETIRELTAKLGRCESQSTLDPGAGEARAGGGRKQPGSGKNTMGDLRSRTAAETL
 SQLGQTLQSLKTRLENLEQYSRLSSSQTNLSKDLLQSKIDELERQVLSRVNTLEEGKGGPRNDT
 EERVKIETALTSLHQRISELEKQKQDNREGDKFQLTFPLRTNYMYAKVKKSLPEMYAFTVCMWLK
 SSATPGVGTFFSYAVPGQANELVLI EWGNNPMEILINDKVAKLPFVINDGKWHHICVTWTTRDGV
 WEAYQDGTQGGSGENLAPYHPIKPQGVVLVGOEDTLGGGFDAQAFVGE LAHFNIWDRKLTPGE
 VYNLATCSTKALSGNVIAWAESHIEIYGGATKWTFEACROIN

Figure 4.4 Neuronal pentraxin 1 sequence information and construct design.

The NP1 secretion signal is indicated by the yellow highlighted region. The start positions for the PTX, FL and CC2 constructs are indicated by green, magenta and cyan highlighting respectively. The construct's end residue is indicated by red font, in bold. Site-directed mutagenesis locations are highlighted in grey and the two N-linked glycosylation sites are highlighted in blue bold font and are underlined.

To produce artificial dimeric, trimeric and tetrameric forms of NP1_PTX, hereafter referred to as NP1_PTX-2coil, -3coil and -4coil, N-terminal coil modifications were introduced by addition of short peptide sequences (figure 4.5A). These sequences were based on previously engineered peptide sequences used to generate dimeric, trimeric and tetrameric coiled regions by altering leucine and isoleucine hydrophobic residues at key positions (highlighted in yellow and green, figure 4.5A) (Harbury *et al.* 1993).

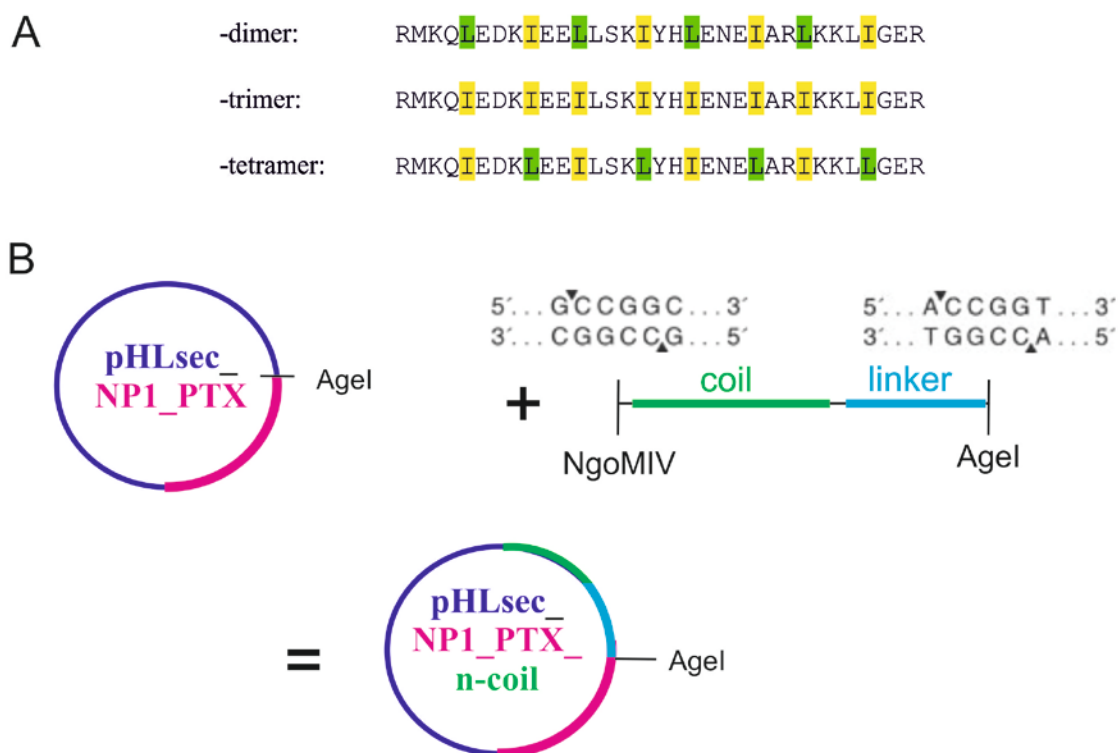


Figure 4.5 Generation of multimeric PTX domains.

A) Artificial coiled-coil modifications (Harbury *et al.* 1993) to form coils in a dimeric, trimeric and tetrameric arrangement by alteration of leucine and isoleucine hydrophobic residues. B) Engineering the coiled-coil modification into pHLsec vector containing various PTX domain constructs.

DNA fragments encoding the peptides (figure 4.5A) were synthetically generated by Eurofins (MWG), supplied in the pCR2.1 vector and were engineered to contain an N-terminal NgoMIV site to produce complementary restriction ends to an AgeI cleaved vector. An additional C-terminal AgeI restriction site was also included after a C-terminal GGSGGSGGS ‘linker’ region to space out the pentraxin domain from the clustering peptide (figure 4.5B). AgeI restriction digested pHLsec and pHLsec_Avi vectors containing NP1_PTX were dephosphorylated with alkaline phosphatase to remove 5′ phosphate groups. The treated vector therefore lacked the 5′ phosphoryl termini, preventing self-ligation. The synthetic peptide-containing vectors were digested with NgoMIV and AgeI restriction enzymes and the gel purified peptide fragments were ligated into the pHLsec vectors N-terminally to the PTX domain (figure 4.5B). Sequence verification was performed to ensure the inserts were in the correct orientation.

The NP1_PTX-2coil, -3coil and -4coil constructs were also cloned into the pNeo_Sec vector (Chapter 2, section 2.1.10) (Appendix E- pNeo_Sec vector map) for expression in the polyclonal HEK293S stable cell line system (Chapter 2, section 2.2.4).

4.3.3 Generation of neuronal pentraxin 2 constructs

For human neuronal pentraxin 2 (NP2) (GenBank accession number BC009924) NP2_PTX (S221-L431), NP2_FL (Q18-L431) and NP2_CC2 (G89-L431) constructs were generated (figure 4.6). Site-directed mutagenesis was performed in an attempt to remove a potential tumour necrosis factor- α converting enzyme (TACE) cleavage site (C94S+K101Q+R103Q) (figure 4.6 grey highlights). A ‘sugarless’ NP2_PTX (S221-L431) was created by introduction of the N393Q mutation to remove the predicted N-linked glycosylation site within the PTX domain, highlighted below in figure 4.6 with bold blue underlined font.

```

MLALLAASVALAVAAGACDS PAPGSRFVCTALPPEAVHAGCPLPAMPMQGGAQSP EEEELRAAVLQ
LRET VVQKETLGAQREAIRELTKLARCEGLAGGKARGAGATGKDTMGDLPRDPGHVVEQLSRS
LQTLKDRLESLEHQLRANVS NAGLP GDFREVLQQR LGELERQLLRKVAELED EKSLLHN ETS AHR
QKTESTLNALLQRVTELERGNSAFKSPDAFKVSLPLRTNYLYGKIKKTLPELYAFTICLWLRSSA
SPGIGTPFSYAVPGQANEIVLIEWGN NPIELLINDKVAQLPLFVSDGKWHHICVTWTTTRDGMWEA
FQDGEKLG TGENLAPWHPIKPGGV LILGQE QD TVGGRFDATQAFVGELS QFNIWDRVLRAQEIVN
IA CSTNMPGNIIPWVDNNVDVFGGASKWPVETCEERLLDL

```

Figure 4.6 Neuronal pentraxin 2 sequence information and construct design.

The NP2 secretion signal is indicated by the yellow highlighted region. The start positions for the PTX, FL and CC2 constructs are indicated by green, magenta and cyan highlighting respectively. The construct's end residue is indicated by red bold font. Site-directed mutagenesis locations are highlighted in grey and the N-linked glycosylation sites are highlighted in blue bold underlined font.

4.3.4 Generation of neuronal pentraxin receptor constructs

The membrane bound neuronal pentraxin receptor (NPR) (GenBank accession number BC153027) constructs NPR_PT X (S289-A500) and NPR_CC2 (G118-A500) were based on known sequence information (figure 4.7).

```

MKFLAVLLAAGMLAFLGAVICIIASVPLAASPARALPGGADNASVASGAAASPGPQRSLSALHGA
GGSAGPPALPGAPAASAHPLPPGPLFSRFLCTPLAAACPSGAQQGDAGAAPGEREELLLLQSTA
EQLRQTALQQEARIRADQDTIRELTGKLGRCE SGLPRGLQGAGPRRDTMADGPWDSPALILELED
AVRALRDRIDRLEQELPARVNSAAPAPVSAVPTGLHSKMDQLEGQLLAQVLALEKERV ALSHSS
RRQRQEVEKELDVLQGRVAELEHGSSAYSPDAFKISIPIRNNMYARVRKALPELYAFTACMWL
RSRSSGTGQGTFFSYSVPGQANEIVLLEAGHEPME LLINDKVAQLPLSLKDN GWHHICIAWTTRD
GLWSAYQDGE LQSGENLAAWHP IKPHGILILGQE QDTLGGRFDATQAFVGDIAQFNLWDHALTP
AOVLGIANCTAPLLGNVLPWEDKLVEAFGGATKA AFDVCKGRAKA

```

Figure 4.7 Neuronal pentraxin receptor sequence information and construct design.

NPR sequence information highlighting predicted cell adhesion region (black box), start positions for PTX and CC2 constructs are highlighted in green and blue respectively. Predicted N-linked glycosylation sites are highlighted in blue bold underlined font and the construct's end residue is indicated by red bold font. Predicted TACE cleavage sites are indicated by black arrows (Cho *et al.* 2008).

4.3.5 Protein expression and purification

Following PCR amplification, restriction digest, cloning and sequence verification, protein secretion levels were initially detected by western blot analysis of media harvested from small-scale transiently transfected human embryonic kidney (HEK) 293T cells (Chapter 2, section 2.3.2). For large-scale expression proteins were expressed in 3L of HEK293T or HEK293S and purified using IMAC and SEC as outlined in Chapter 2, section 2.4.

4.3.6 Crystallization, data collection and processing

NP1_PTX crystals typically grew in conditions that contained polyethylene glycol (PEG). The NP1_PTX crystal used to solve the structure grew from a three-row optimization with a 3% gradient of a reservoir containing PEG 8000 at pH 6.5, details of which are outlined in Table 4.1. Crystals from this optimization plate were soaked with CaCl₂ (final concentration 10mM) before freezing. NP2_PTX crystals grew in another PEG 8000-containing condition, this time at pH 8.5. Some needle-like crystals were obtained for an NP2_CC2 construct, later determined to contain only three extra residues at the N-terminus of the PTX domain and did not extend into the coiled-coil region. NPR_PTX crystals grew in an optimization plate containing PEG 1500 at pH 9.0. All other crystals for longer protein constructs (CC2, FL etc.) were found to contain only the PTX domain so it was postulated that these were being cleaved in the drop.

Table 4.1 Summary of all neuronal pentraxin crystals tested.

Crystallization conditions for all NP1_PTX, NP2_PTX and NPR_PTX crystals Details of crystal screen contents (column 3) can be found in Chapter 2, Table 2.4.

Crystal	Plate ID_well	Crystal Screen	Condition	Date set	Comments	Temp
NP1_PTX	391731_F08	Optimisation of Block 4 F03 plate ID 391557	15 %v/v Glycerol 25.5 %w/v Polyethylene Glycol (PEG) 8000, 0.17M Ammonium Sulphate 0.085M Sodium Cacodylate pH 6.5	25/10/2010	6.5mg/ml HEK293T cells	4°C
NP1_PTX +2mM CaCl ₂	391731_F07	Optimisation as above	Condition as above	25/10/2010	As above + CaCl ₂ soak	4°C
NP1_PTX +20mM Sucrose Octa sulphate (SOS), 2mM CaCl ₂	435046_C06	Optimisation of Block 4_F03 Plate ID 391557	15 %v/v Glycerol 25.5 %w/v PEG 8000, 0.17M Ammonium Sulphate 0.085M Sodium Cacodylate pH 6.5	07/03/2012	6.7mg/ml HEK293T cells	4°C
NP2_CC2_NHis	366081_G09	Block 5	0.35M di-Potassium Hydrogen Phosphate 0.65M Sodium di-Hydrogen Phosphate pH 6.3	15/02/2010	2.8mg/ml HEK293S triple flask EndoF1 deglyc.	RT
NP2_PTX	296708_G12	Block 4	35 %v/v Glycerol 5.2 %w/v PEG 8000 0.065M Tris-HCl pH 8.5	22/07/2009	8.9mg/ml HEK293T +kif EndoF1 deglyc.	RT
NP2_PTX 2 nd crystal form	297347_F09	Optimisation of Block 5_F10 Plate ID 296661	0.1M HEPES pH 7.0 , 65%v/v 2-Methyl-2,4-pentanediol	30/07/2009	8.9mg/ml HEK293T +kif EndoF1 deglyc	4°C
NPR_PTX	418445_B10	Optimisation of PACTPremier_B6 Plate ID 418520	0.1M MIB system pH 9.0 22.5%w/v PEG 1500	22/08/2011	13mg/ml HEK293T +kif EndoF1 deglyc	RT

All crystals were cryo-protected with 30% ethylene glycol prior to data collection at the Diamond Light Source in Harwell in Oxfordshire. NP2_PTX data was collected at Diamond beamline IO3 and solved using a molecular replacement model from SAP (PDB ID: 1SAC (Emsley *et al.* 1994)) to a resolution of 1.8Å (figure 4.11). NP2_PTX data were indexed to space group P3₂21 (table 4.2).

NP1_PTX datasets were collected at Diamond beamlines IO3 and indexed to space group C2 (table 4.2). The NP1_PTX structure was then solved to 1.45Å using NP2_PTX as a search model in Phaser (McCoy *et al.* 2005). Automatic model building using the NP1_PTX peptide sequence was then carried out by ARP/wARP (<http://www.embl-hamburg.de/ARP/>).

NPR_PTX data was collected at the Diamond IO4-1 beamline. Several datasets were collected and the best crystal diffracted to 1.82Å with a P1 crystal lattice (Table 4.2) and two molecules per asymmetric unit (ASU). Matthew's coefficient analysis confirmed that two molecules of 24530Da were present in the ASU with 40% solvent. The NPR_PTX structure was solved using NP1_PTX as a search model in Phaser (McCoy *et al.* 2005). Automatic model building using the existing model and the NPR_PTX peptide sequence was then carried out by ARP/wARP (<http://www.embl-hamburg.de/ARP/>).

4.3.7 Refinement and structural validation

Refinement was carried out using Phenix (Adams *et al.* 2002), Buster (Gilmore *et al.* 1993) and Refmac (Murshudov *et al.* 1997) programs. Manual building in Coot (Emsley *et al.* 2004) and iterative cycles of refinement were then used to generate the final models.

A Ramachandran plot was performed in MOLPROBITY (Lovell *et al.* 2003) to assess the stereochemical properties of the built model. This type of plot identifies unfavourable conformations brought about by a steric hindrance when the dihedral angles Phi ϕ and/ or Psi ψ angles are present in forbidden or unstable orientations. Angles that did not adhere to expected backbone conformations were termed outliers, indicating that further refinement was required to

generate the structural final statistics (table 4.2). Structural analysis involved model superposition and enabled a direct comparison of domain organisation between molecules. The root mean square deviation (RMSD) values (i.e. the measure of difference) were calculated using the structural homology program (SHP) (Stuart *et al.* 1979) or the PyMOL align tool (DeLano 2002).

4.3.8 Biophysical analysis

MALS analysis was performed as described in Chapter 2, Section 2.8.1. All proteins were analysed at 1mg/ml unless otherwise stated. Analysis was performed on a Superdex S200 10/300 (GE Healthcare) in 10mM HEPES pH 7.5, 150mM NaCl running buffer.

SPR was performed as described in Chapter 2 section 2.8.2. For SPR analysis, purified AMPA receptor ATD proteins (GluA2, A3 and A4 ATDs) (described in Chapter 3, section 3.3.2) were concentrated to between 30-50 μ M and a dilution series established for use as the 'analyte'. In initial experiments the standard running buffer was 10mM HEPES (pH 7.4), 150mM NaCl and 0.05% w/v Tween-20 but it was noticed that binding was enhanced (10 fold) in the presence of calcium (data not shown), therefore the running buffer in all further SPR experiments also included CaCl₂ and MgCl₂ (both at final concentration 2mM). The NP1_PTX NP2_PTX, NPR_PTX, NP1_PTX_2coil and NP1_PTX_3coil constructs were cloned into pHlsec_Avi and expressed in HEK293T cells and biotinylated for biophysical analysis by SPR using methods outlined in Chapter 2 section 2.8.2.

FACS was performed as described in Chapter 2, section 2.9. Analysis was performed with GluA4_FL protein because SPR analysis revealed that of all interaction pairs tested, GluA4_ATD formed the tightest interaction with NP1_PTX (~5 μ M). HEK293T cells were transfected with GluA4_FL_HA_monoVenus (mV) and GluA4_ Δ ATD_HA_mV as well as GluA1 constructs GluA1_FL_HA_mV, GluA1_ Δ ATD_HA_mV (cloned by Aletheia Lee, DPhil student in our laboratory). Cells were stimulated with purified NP1_PTX and NP1_PTX_3coil protein either 1 hour or 12 hours prior to cell surface staining and protein surface levels analysed by anti-HA antibody detection and FACS analysis.

4.4 Results

4.4.1 Protein expression and purification

The 24.1kDa NP1_PTX protein, expressed in HEK293T cells, yielded 5-8mg/L and was determined by MALS to be monomeric in solution (figure 4.17E). The protein was prone to precipitation when concentrated, therefore to overcome this problem 1ml elution fractions were collected from IMAC. These peak protein fractions were pooled and dialysed into an imidazole free buffer, which successfully reduced the aggregation problem and removed any metal ions leached from the IMAC resin. SDS-PAGE analysis of the purified NP1_PTX protein revealed it to be >90% pure (figure 4.8A). NP1_CC2 (36.6kDa) was expressed to 2-5mg/L where a single peak at 15.8mls was observed on SEC (figure 4.8B). NP1_FL (46.0kDa) was expressed to <0.5mg/L where SEC presented several protein peaks. The protein determined to be full length (by SDS-PAGE analysis) was immediately pooled and protease inhibitors (PI) (Sigma) added to prevent any further proteolysis. Despite attempts to optimise protein expression, small amounts of NP1_FL protein were always obtained. In an attempt to boost expression, transient transfection of HEK293T cells were repeated with the addition of valproic acid (VPA) and sodium butyrate which act as histone deacetylase inhibitors (Backliwal *et al.* 2008). VPA (1mM final concentration) was found to improve protein yields. Analysis of multiple peaks from SEC by SDS-PAGE suggesting proteolysis had occurred (figure 4.8C). The oligomeric state(s) of the full length protein remains undetermined despite extensive attempts to reduce protease activity and produce protein stable enough for use in negative stain EM and MALS analysis. The protease degradation was rapid, for instance the same protein sample analysed twice on MALS immediately after purification but just 1 hour apart produced very different mass profiles.

NP2_PTX (24.6kDa) yields were 5-8mg/L and did not present any issues during purification. Partial occupation of the N-linked glycosylation site was postulated to be the cause of the two protein bands on SDS-PAGE. The SEC elution volume (figure 4.9A) indicated that NP2_PTX was monomeric (as observed for NP1_PTX). NP2_CC2 (39.1kDa) typically yielded 2-5mg/L from a single peak on SEC (figure 4.9B) A shift in the elution volume of NP2_CC2 to

14mls from that observed for NP1_CC2 (15.8mls) potentially indicated that a higher oligomeric form was present. SDS-PAGE analysis of peak SEC fractions of NP2_FL revealed the presence of several protein bands (figure 4.9C) and only a small fraction of protein was of the expected MW (45.7kDa) resulting in low yields of purified protein (<0.2mg/L). Previous research has stated that a TACE or TACE-like matrix metalloproteases (MMPs) acts to cleave the neuronal pentraxin receptor from the membrane (Cho *et al.* 2008), so TACE activity was considered to be a possible cause of NP2_FL proteolysis. Addition of a protease inhibitor cocktail (Sigma) did not prevent the proteolysis, therefore point mutations were designed to remove the potential TACE cleavage site (C K R residues highlighted in grey figure 4.6) but this did not prevent the protein cleavage occurring. The oligomeric state(s) for NP2_FL remain undetermined due to the protein's unstable nature.

Two secreted NPR constructs (NPR_PTX and NPR_CC2) were also cloned, expressed (with 1mM VPA added during NPR_CC2 expression) and purified, to produce 5-10 and 4-6mg/L of protein respectively (figure 4.10A &B). SEC analysis indicated that NPR_PTX was monomeric in solution and NPR_CC2 was likely to be multimeric because a larger than previously observed protein (compared to the NP1_CC2 and NP2_CC2 equivalent proteins) eluted at 12.5mls (if not cleaved). NPR_CC2 protein was degraded despite the addition of a PI cocktail. Protein bands observed on SDS-PAGE matched the MW expected for the full NPR_CC2 construct (39kDa) as well as the potential products if the protein was cleaved between the PTX and coiled-coil domains.

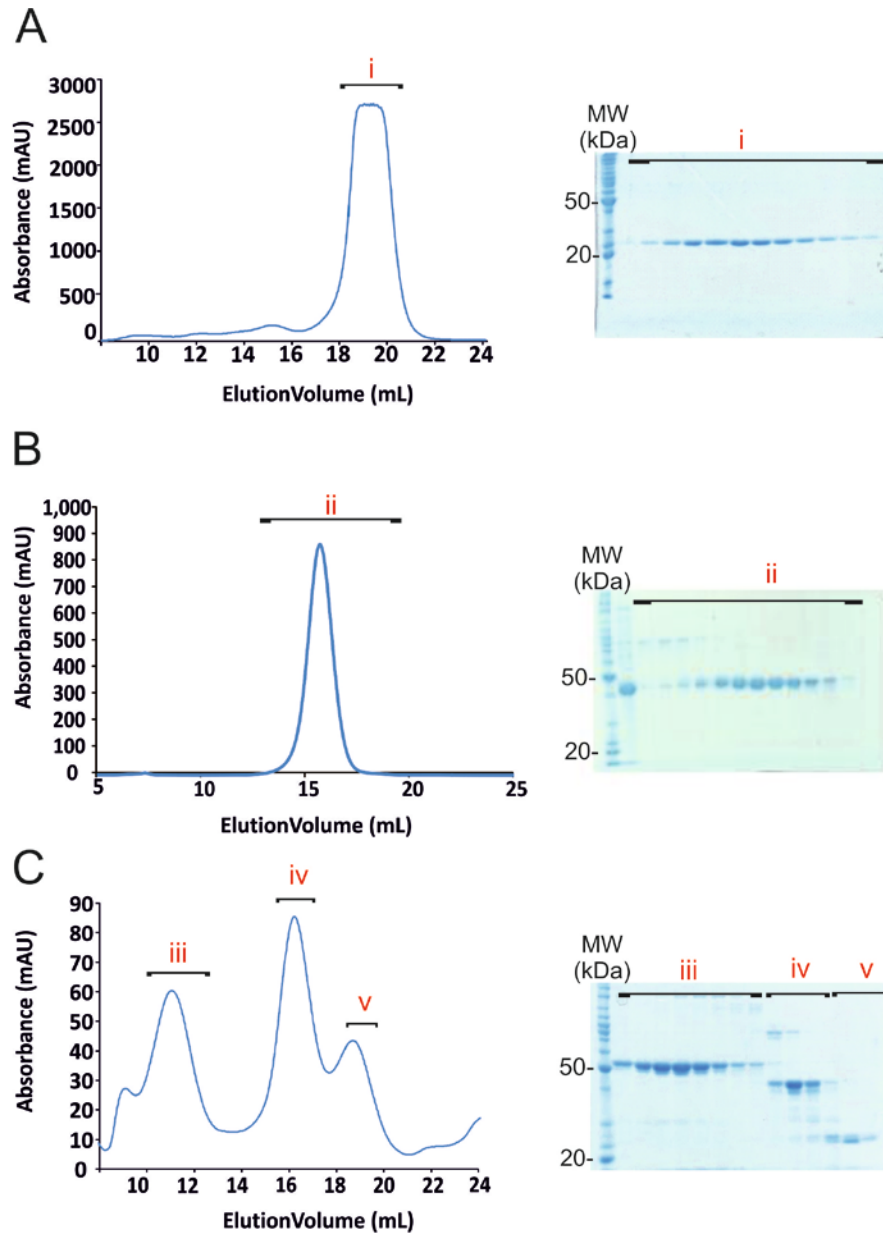


Figure 4.8 Neuronal Pentraxin 1 (NP1) protein purification analysis.

Size exclusion chromatography (SEC) was performed after IMAC (TALON) purification for NP1 proteins. SDS-PAGE analysis of SEC peak fractions shown here was performed using 15% (NP1PTX) and 12% (NP1_CC2 and NP1_FL) acrylamide gels. A) NP1_PTX, B) NP1_CC2, C) NP1_FL. SEC peaks are identified by red Roman numerals i-v to correspond to SDS-PAGE labels. Gel load order in A) and C) was as follows: Benchmark protein ladder (Invitrogen) then SEC peak fractions as indicated. Gel load order for B), Benchmark protein ladder (Invitrogen), 1 μ l of SEC injected sample then peak SEC fractions. SEC running buffer 10mM HEPES pH 7.4, 150mM NaCl and samples were analysed on an S200 10/30 column.

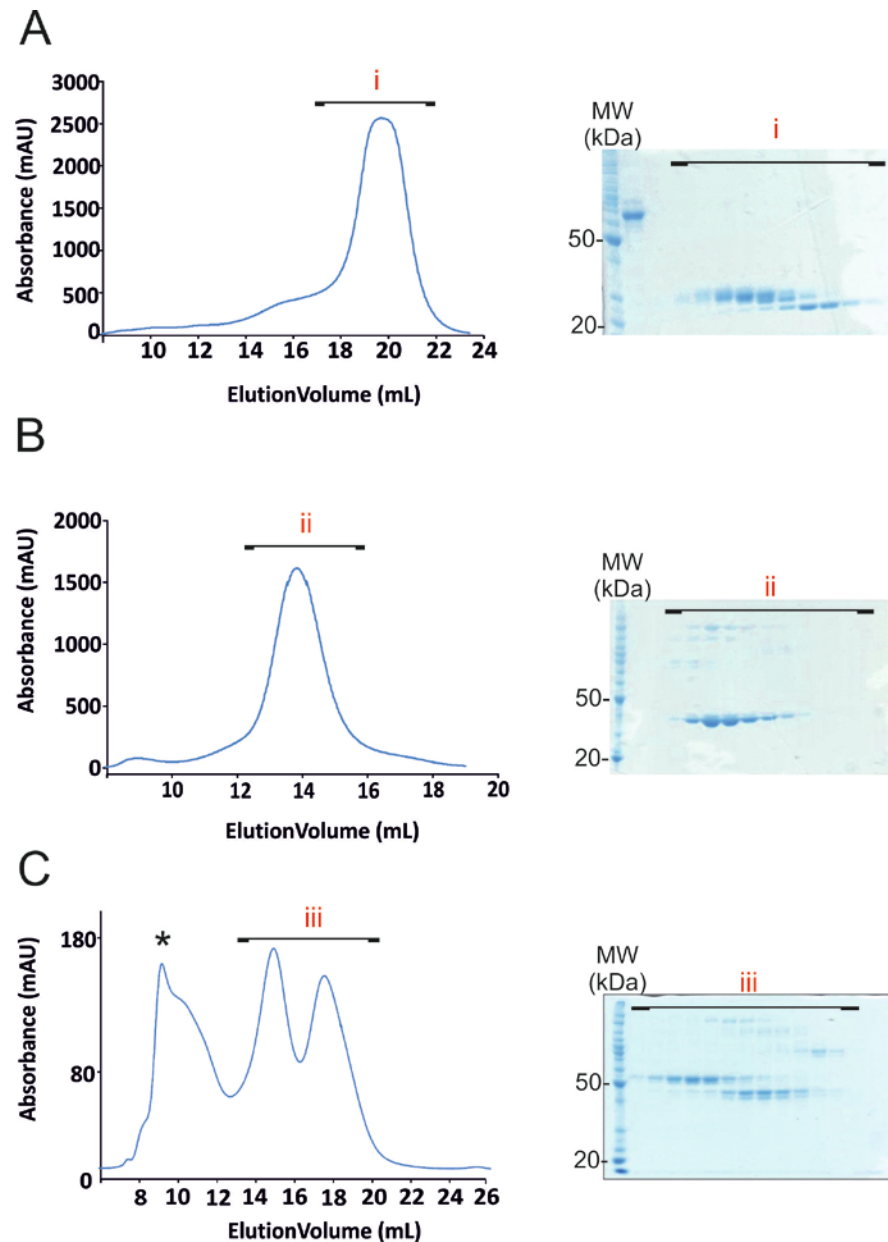


Figure 4.9 Neuronal pentraxin 2 (NP2) protein purification analysis.

SEC was performed after IMAC (TALON) purification and analysis of SEC peak fractions was performed using 12% acrylamide SDS-PAGE gels. A) NP2_PTX, B) NP2_CC2, C) NP2_FL. * indicates aggregate peak at column void volume. Gel load order was as follows: A) Molecular weight marker (MW) Benchmark (Invitrogen), IMAC flow-through sample, 5mM imidazole wash then peak SEC fractions. In B) and C) Benchmark protein ladder followed by peak SEC fractions. SEC peaks are identified by red Roman numerals i-iii and correspond to SDS-PAGE labels. SEC running buffer 10mM HEPES pH 7.4, 150mM NaCl. Samples were analysed on an S200 10/30 column.

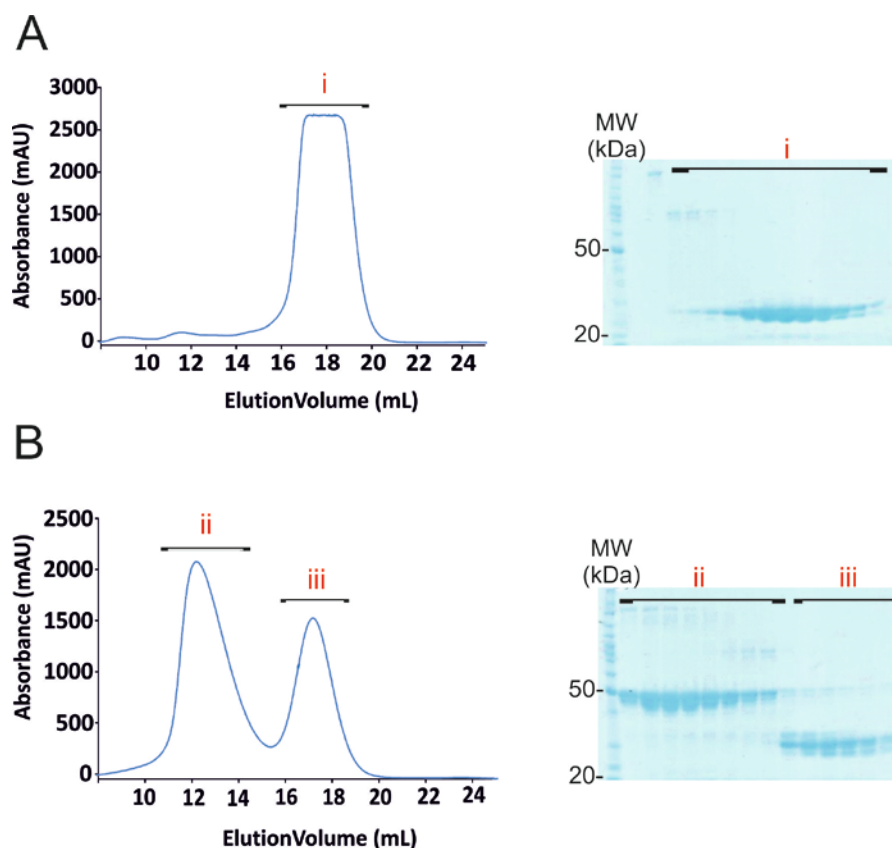


Figure 4.10 Neuronal pentraxin receptor (NPR) protein purification analysis.

A) NPR PTX and B) NPR CC2 size exclusion chromatography (SEC) analysis performed after IMAC (TALON) purification. Peak SEC fractions were analysed on 12% acrylamide SDS-PAGE gels. Gel load order: Molecular weight marker (MW) Benchmark (Invitrogen) followed by SEC peak fractions as identified by red Roman numerals i-iii to correspond to SEC profile labels. SEC running buffer: 10mM HEPES pH 7.4, 150mM NaCl. Samples analysed on an S200 10/30 column.

4.4.2 Structural analysis of neuronal pentraxin PTX domains

Information regarding crystallization conditions for NP1_PTX, NP2_PTX and NPR_PTX can be found in section 4.3.6. A summary of the crystals diffracting to the highest resolution for each protein, the diffraction patterns obtained and structural highlights observed in electron density can be found in figure 4.11.

In addition to the successful crystallization results summarised in table 4.1 other crystallization attempts are included. The neuronal pentraxins, like all pentraxins, act as lectins to bind carbohydrates in a calcium dependant manner (Tsui *et al.* 1996). Crystallization of NP2_PTX with the addition of 10mM SOS or heparan sulphate and 2mM CaCl_2 was therefore attempted but unfortunately these were not observed in any of the data collected.

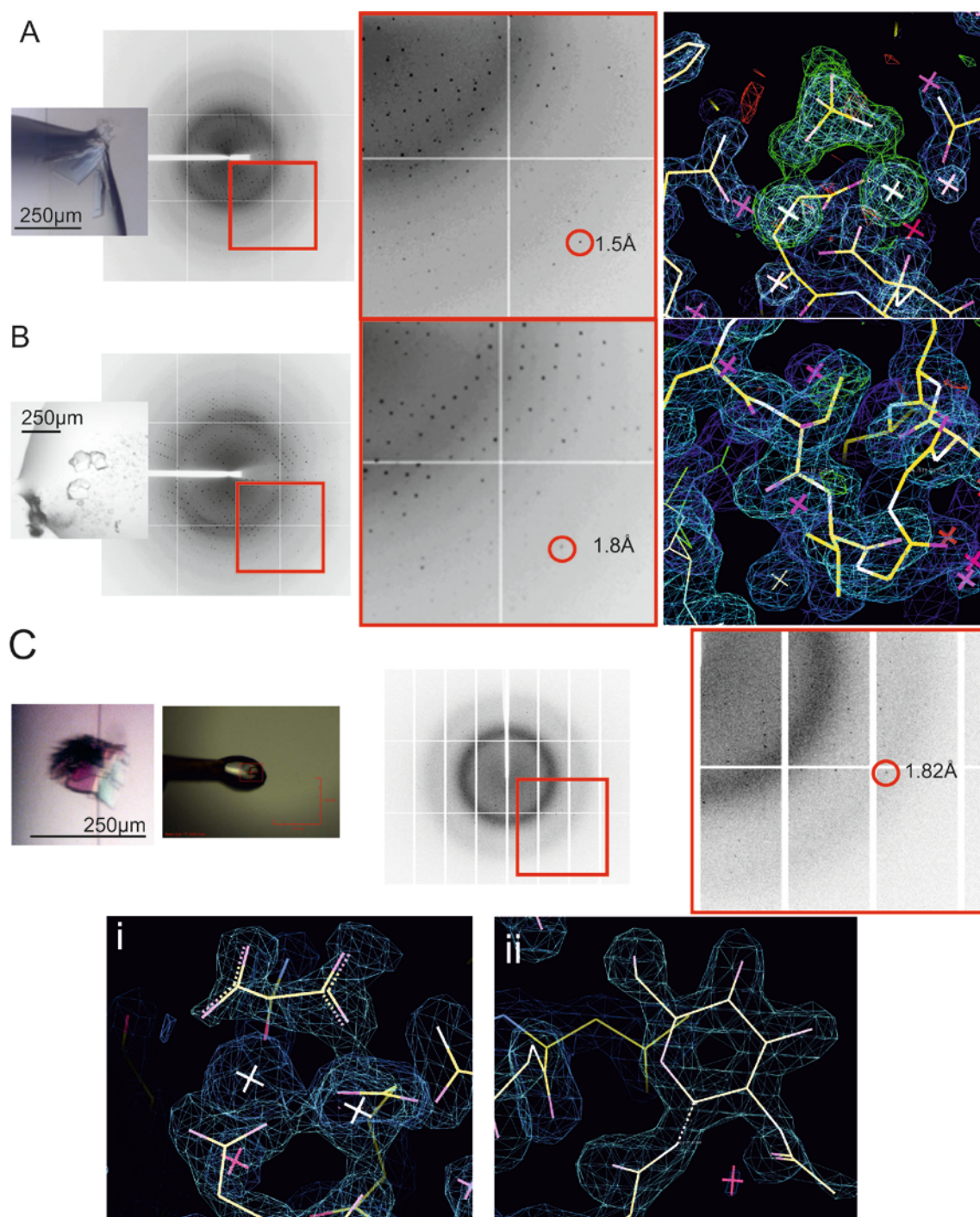


Figure 4.11 Diffraction data and key structural features of NP1_PTX, NP2_PTX and NPR_PTX.

A) X-ray diffraction data for NP1_PTX collected at Diamond IO-3 diffracted to 1.45 Å. An Fo-Fc map contoured to 1.7σ provides a snapshot of electron density for data refined without SO₄ and the two Ca²⁺ ions. B) X-ray diffraction data for NP2_PTX collected at Diamond IO3 diffracted to 1.8 Å. The Fo-Fc map contoured to 1.7σ provides a snapshot of electron density highlighting the loop region usually involved in calcium binding but in this calcium-free crystal the loop is involved in a crystal contact. C) X-ray diffraction data for NPR_PTX collected at Diamond IO4-I diffracted to 1.82 Å. The snapshots are i) Fo-Fc map contoured to 1.4σ to highlight calcium co-ordination this time by malonic acid, ii) N-acetyl glucosamine moiety observed in Fo-Fc map contoured to 1.7σ. Dotted line indicates distance of 1.43 Å.

Structures solved for all three neuronal pentraxin PTX domains demonstrate that these proteins are highly similar to each other (figure 4.12D) (RMSD values for NP1_PTX-NP2_PTX 0.475Å over 1174 atoms, NP1_PTX-NPR_PTX 0.492Å over 1059 atoms and NP2_PTX-NPR_PTX 0.45Å over 1087 atoms (Pymol align)). All PTX domains contain a central anti-parallel arrangement of β -sheets in a 'jelly-roll' topology with two faces termed here A and B that contain an α -helix and calcium binding site respectively. Overall, the fold showed similarities to other known pentraxin structures (figure 4.13). Alignment of NP2_PTX onto SAP (PDB: 1SAC (Emsley *et al.* 1994)) using the structural homology program (SHP) (Stuart *et al.* 1979) gave an RMSD value of 2.2Å over 200 equivalent residues with variations in loop regions (figure 4.13C). PISA (Protein Interfaces, Surfaces, Assemblies) [http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html] server results indicated that no higher order oligomerisation was present in the crystal packing, confirming the monomeric state of the NP1_PTX observed on MALS (figure 4.17E). The pentraxin β -sheets have been previously labelled A-N (Shrive *et al.* 1996), here the same numbering was employed and revealed that β -sheets B-M are conserved (figure 4.12A), as is the α -helix on the A-face containing residues P388-T396 (NP1 numbering) (figure 4.12A). Disulphide bonds formed between cysteine residues C428-C397 and C256-C316 are also conserved across NP1_PTX, NP2_PTX and NPR_PTX structures and act to stabilise β -sheets C & H and the distal end of the α -helix to the C-terminal region of the protein and suggesting that mutation of either C256 or C316 residues would disrupt correct protein folding.

Of particular note is the calcium binding site on the B-face (figure 4.12B) that had been previously predicted (Dodds *et al.* 1997). Based on sequence alignments (figure 4.2) all PTX domains have the ability to bind calcium but the NP2_PTX structure was solved without calcium. Despite the fact that NP2_PTX crystals did not contain calcium, the ability of all neuronal pentraxins to bind to calcium is evident from sequence conservation and their electrostatic potentials which indicate that a negatively charged pocket is present on the B face of all neuronal pentraxin PTX domains (figure 4.12C).

When comparing the calcium-bound (NP1_PTX and NPR_PTX) and calcium-free forms (NP2_PTX), an alternative conformation is evident in one loop containing residues D360-D367 (figure 4.12D). In the structures with calcium this loop is folded back onto the body of the protein away from the calcium ions. In the calcium free structure (NP2_PTX) this loop was able to flip out and form a crystal contact with another molecule and can explain why soaking these crystals with calcium caused them to dissolve. In the NP1_PTX structure the first calcium ion is coordinated by residues E358, D360 and Q370, the second by E358, N280 and D360 with the rest of the coordination being provided by the sulphate ion and water molecules present (figure 4.12E). In the NPR_PTX structure the additional calcium co-ordination is provided by malonic acid (figure 4.11Ci), present in the malonic acid, imidazole, boric acid (MIB) crystallization condition (table 4.1) that this crystal grew in. Crystallographic statistics for the NP1_PTX, NP2_PTX and NPR_PTX solved structures can be found in table 4.2.

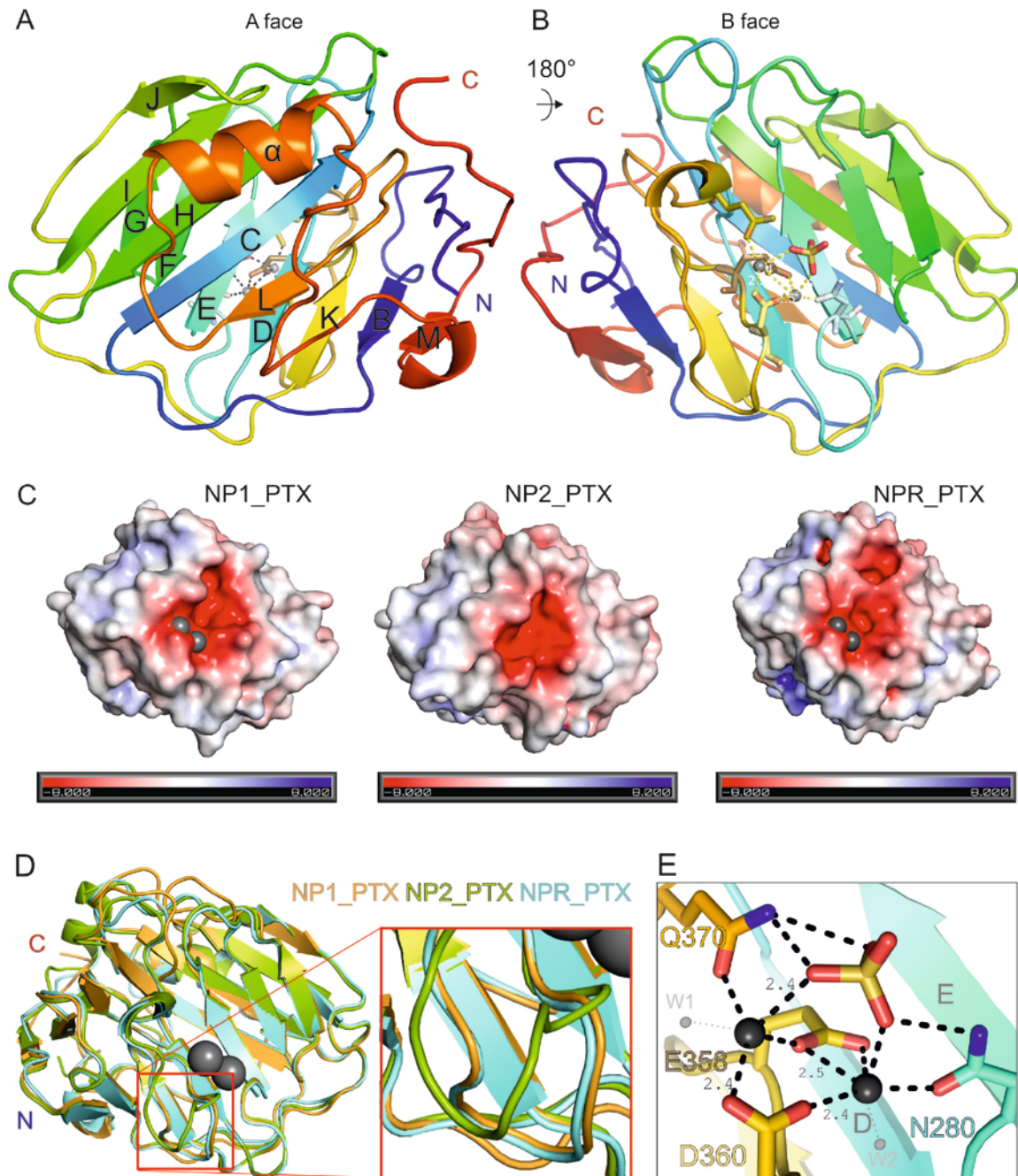


Figure 4.12 Neuronal pentraxin structural analysis.

A) The NP1_PTX structure represented in cartoon rainbow format where Beta sheets are labelled B-M and the prominent alpha helix of the A-face is labelled as α according to human C-reactive protein (PDB ID: 1GNH) from Shrive *et al.*, 1996. B) 180° rotation about y to highlight the calcium binding site on B-face. C) Electrostatic potential of NP1_PTX, NP2_PTX and NPR_PTX generated in APBS (Pymol) contoured to -8/+8 kT/e- where the NP2+PTX structure did not contain calcium. D) Overlay of NP1_PTX (orange), NP2_PTX (green) and NPR_PTX (blue) to highlight high level of structural homology. Calcium ions shown as grey spheres. Red box highlights calcium induced loop flip. E) Enlarged view of the calcium binding site of NP1_PTX co-ordinated by residues N280, E358, D360 and Q370 as well as water molecules and sulphate ion.

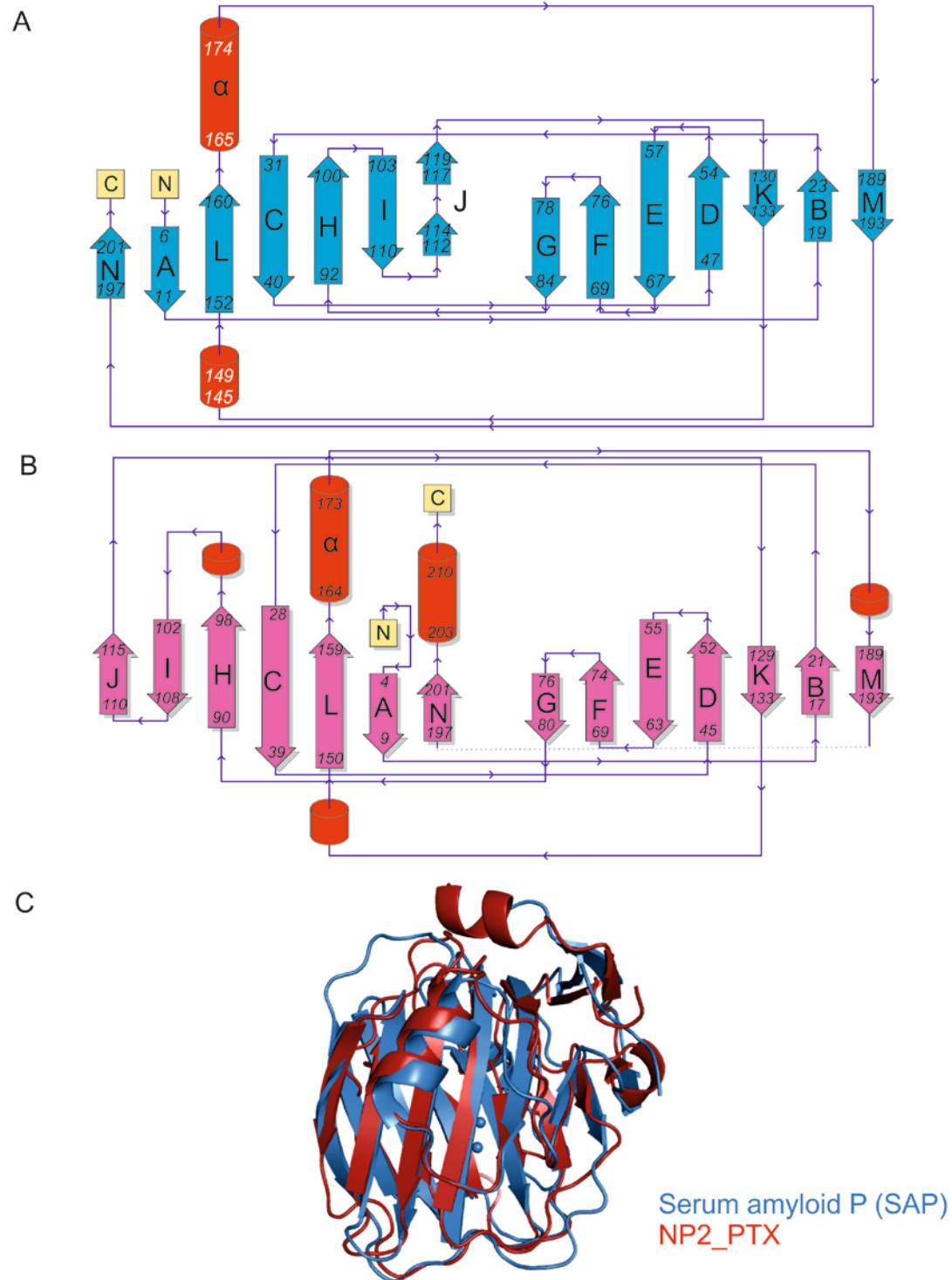


Figure 4.13 The conserved 'jelly-roll' fold of the pentraxins.

A) Secondary structure topology diagram to highlight pentraxin domain fold homology. A) Serum amyloid P (SAP) (PDB entry 1SAC (Emsley *et al.* 1994) and B) NP2_PTX. Figures generated in PDBsum (<http://www.ebi.ac.uk/pdbsum/>). Beta sheets and alpha helices are labelled according to Shrive *et al.* 1996. C) Structural superposition of SAP and NP2_PTX where alignment in SHP (Stuart *et al.*, 1979) gave an RMSD value of 2.21 Å over 200 equivalent residues.

Table 4.2 Neuronal pentraxin statistical analysis

	NP1-PTX +Ca ²⁺	NP2-PTX	NPR-PTX+Ca ²⁺
Data collection			
Beamline	Diamond I03	Diamond I03	Diamond I04-1
Space group	C2	P3 ₂ 21	P1
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	121.2, 36.1, 92.2	73.32, 73.32, 85.10	36.7, 41.5, 70.1
α , β , γ (°)	90.0, 97.5, 90.0	90, 90, 120	104, 96.5, 106.2
Wavelength (Å)	0.976	1.0714	0.917
Resolution (Å)	40.0-1.45 (1.5-1.45)	63.5-1.80 (1.87-1.80)	34.58-1.82 (1.87-1.82)
<i>R</i> _{merge}	0.086 (0.742)	0.080 (0.711)	0.082 (0.683)
<i>I</i> / σ <i>I</i>	12.8 (1.3)	28.5 (3.7)	9.7 (2.1)
Completeness (%)	98.6 (95.8)	100 (100)	91.4 (92.4)
Redundancy	3.4 (2.7)	10.6 (10.4)	3.7 (3.8)
Wilson B factor (Å ²)	9.27	24.4	20.214
Refinement			
Resolution (Å)	38.0-1.45 (1.49-1.45)	36.65-1.8 (1.87-1.80)	34.0-1/82 (1.88-1.82)
No. reflections	69597 (4537)	24906 (2592)	30942 (2823)
<i>R</i> _{work}	0.162 (0.246)	0.160 (0.212)	0.177 (0.216)
<i>R</i> _{free}	0.194 (0.294)	0.193 (0.269)	0.207 (0.218)
No. atoms			
Protein	3204	1670	3147
N-linked glycans	-	-	28
Glycerol	-	12	-
Calcium	4	-	4
Sulphate	10	-	-
Water	562	140	388
Malonate	-	-	14
B-factors (Å ²)			
Protein main/side chains	12.4/ 16.6	26.7/ 31.8	21.88/ 28.32
N-linked glycans	-	-	31.80
Glycerol	-	66.3	-
Calcium	10.86	-	21.66
Sulphate	17.25	-	-
Water	33.2	42.5	39.29
Malonate	-	-	47.07
R.m.s deviations			
Bond lengths (Å)	0.006	0.007	0.007
Bond angles (°)	1.145	1.123	1.033
Ramachandran plot			
Favoured (%)	96.79	97.17	98.75
Allowed (%)	3.21	2.83	1.25
Disallowed (%)	0	0	0
Molprobtity clash score	2.22 (99 th pc)	3.59 (98 th pc)	3.98 (98 th pc)
Molprobtity overall score	1.26 (96 th pc)	1.31 (98 th pc)	1.39 (97 th pc)

Numbers in brackets refer to the highest resolution shell.

*R*_{free} was calculated as per *R*_{work} for a 5% subset of reflections not used in the crystallographic refinement.

Ramachandran statistics and Molprobtity scores were calculated with Molprobtity (Lovell *et al*, 2003).

4.4.3 Mapping the AMPA receptor - neuronal pentraxin interaction site

Published research suggests that all neuronal pentraxins may interact with AMPA receptor ATDs (Xu *et al.* 2003; Sia *et al.* 2007; Al Rahim *et al.* 2013), but to increase the likelihood of a complex crystal forming it was important to biophysically characterise each of the interactions to enable crystallization attempts with the tightest binding partners. SPR analysis was performed as described in Chapter 2 section 2.8.2. Biophysical analysis of biotinylated NP1_PTX, NP2_PTX and NPR_PTX (pHLsec_Avi constructs) was carried out with the neuronal pentraxin construct bound to the chip as the 'ligand'. AMPA receptor GluA2_ATD and GluA3_ATD constructs were transiently expressed in HEK293S cells. GluA4_ATD was expressed from a HEK293S stable cell line. GluA1_ATD was expressed at low levels (<0.5mg/L (estimate)) in mammalian cells and could not be successfully purified as it was prone to aggregation during purification. Despite extensive optimisation attempts the GluA1_ATD protein could not be obtained at levels required for SPR, so was not included in this analysis. Unfortunately SPR analysis could also not be performed in the reverse orientation with PTX proteins as analyte (which would usually confirm the binding (dissociation) constant (K_D) and stoichiometry) because the presence of calcium in the buffer caused the PTX proteins to precipitate at high concentrations. Actually, in the case of NP1_PTX this 'precipitate' was later visually determined to be microscopic needle crystals, still therefore unsuitable for SPR but still useful to note that at high concentrations this protein can crystallize without the addition of a specific precipitant.

Figure 4.14 summarises all SPR (Biacore) experiments conducted. In all cases the neuronal pentraxin PTX domains were the ligands bound to a chip and AMPA receptor ATD proteins were applied as analyte at various concentrations to determine each binding affinity (K_D). All affinities are in the micromolar range, but one would expect this to be the case at a plastic synapse where AMPA receptor localisation is activity dependant. Of most importance, the tightest interaction was determined to be between GluA4 ATD and NP1_PTX where a K_D between 5-17 μ M was reproducibly observed (figure 4.14C).

Complex crystallization of NP1_PTX and GluA4_ATD was therefore attempted but trials are yet to yield complex crystals due to the strong tendency of NP1_PTX to crystallize alone.

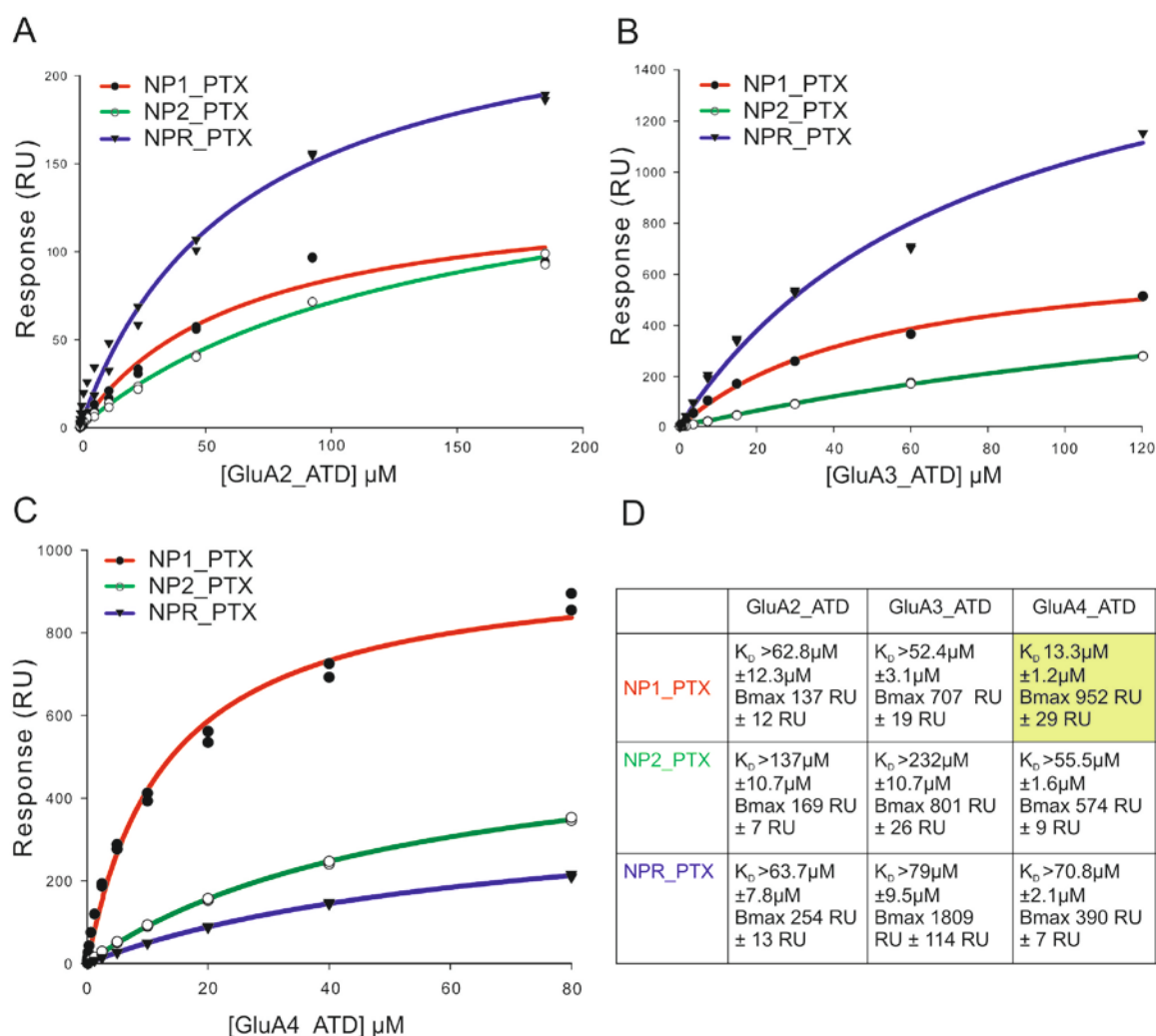


Figure 4.14 SPR analysis of AMPA receptor ATDs binding to neuronal pentraxin PTXs.

Neuronal pentraxin NP1, NP2 and NPR pentraxin (PTX) domains were coupled to the SPR CM5 chip as the 'ligand'. Purified AMPA receptor amino terminal domains (ATDs) were then applied as the 'analyte' in a dilution series and the K_D values calculated. Binding to NP1_PTX, NP2_PTX and NPR_PTX is shown by red, green and blue curves respectively for each AMPAATD analytes as follows:

A) GluA2_ATD and B) GluA3_ATD where 508 response units (RU), 557 RU and 1353 RU of NP1_PTX, NP2_PTX and NPR_PTX were bound to the chip. C) GluA4_ATD where 457 RU, 500 RU and 546 RU of NP1_PTX, NP2_PTX and NPR_PTX were bound to the chip. D) Table summary of all interactions analysed highlighting (in yellow) the tightest binding interaction between GluA4_ATD and NP1_PTX. All values were calculated to be in the microMolar range although exact K_D cannot be accurately determined because binding was not saturated. All error values were calculated in Sigmaplot.

4.4.4 Effect of NP1_PTX calcium binding on interactions with GluA4_ATD

Crystallographic information confirms the presence of a double calcium binding site within the neuronal pentraxin PTX domain and further structural analysis of NP1_PTX revealed several residues that were facilitating calcium binding (figure 4.12E).

To measure the effect that the NP1_PTX calcium binding has on the interaction with GluA4_ATD several point mutations were designed to alter the ability of NP1_PTX to bind calcium ions. Constructs termed 'M1' and 'M2' contained mutant sites N280D and Q370E respectively and aimed to enhance calcium binding by providing additional calcium co-ordinating residues. It was hoped that these mutants would act by locally increasing negative charge but the resulting proteins were unfortunately not secreted. An 'M3' mutant was designed to disrupt calcium binding and contained two mutations at residues E358S and D360S (mutant sites indicated in figure 4.15A). All mutants aimed to cause minimal structural disruption and their positions are also highlighted in grey in the NP1 sequence in figure 4.4.

SPR was performed with NP1_PTX and NP1_PTX_M3 (E358S + D360S) bound as ligands and a dilution series of GluA4_ATD with the highest concentration of 30 μ M applied as the analyte. Analysis of NP1_PTX_M3 revealed a marked reduction in binding to GluA4_ATD, where binding to NP1_PTX_M3 was not accurately determined because saturation could not be reached with the concentration range of analyte that could be tested (figure 4.15B).

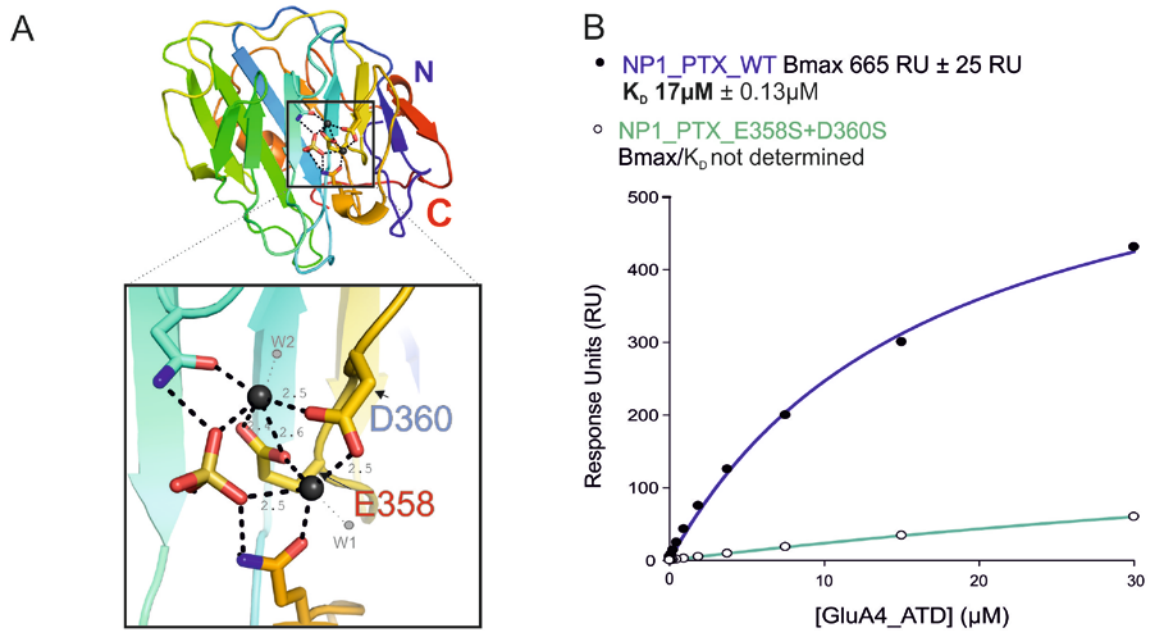


Figure 4.15 Structure-based mutagenesis to disrupt the neuronal pentraxin 1 calcium binding site.

A) Zoomed in view of the NP1_PTX calcium binding site to highlight residues E358 and D360 that were mutated to serine residues (M3) to impede binding to calcium. B) SPR analysis of GluA4_ATD binding to NP1_PTX and NP1_PTX_E358S_D360S calcium binding site mutant (M3) where 430 response units (RU) and 382 RU respectively were bound.

4.4.5 Introduction of GluA4_ATD mutations to impede neuronal pentraxin binding

Structure-guided mutagenesis to disrupt the calcium binding site within the PTX domain was shown to impede binding, but the binding site on the GluA4_ATD remained unknown. N-linked glycosylation residues were introduced at two sites on GluA4_ATD (figure 4.16A) to act as potential steric barriers to the interaction between GluA4_ATD and NP1_PTX. The sites selected were at the flap (K319N-D321S) and ‘wing’ (K274N-M276S) loops of the ATD (figure 4.16A), selected after examination of exposed loops observed in the GluA4_ATD published structure (Dutta *et al.* 2012). In small-scale expression trials the ‘wing’ mutant was not secreted, possibly due to the glycan interfering with protein folding, but the ‘flap’ mutant expressed to 1mg/L and protein expressed from HEK293S cells was purified for use in SPR analysis. In SPR addition of an N-linked glycosylation site in the GluA4_ATD flap loop was able to reduce binding to the PTX domain when compared to the wild type GluA4_ATD protein (figure 4.16B).

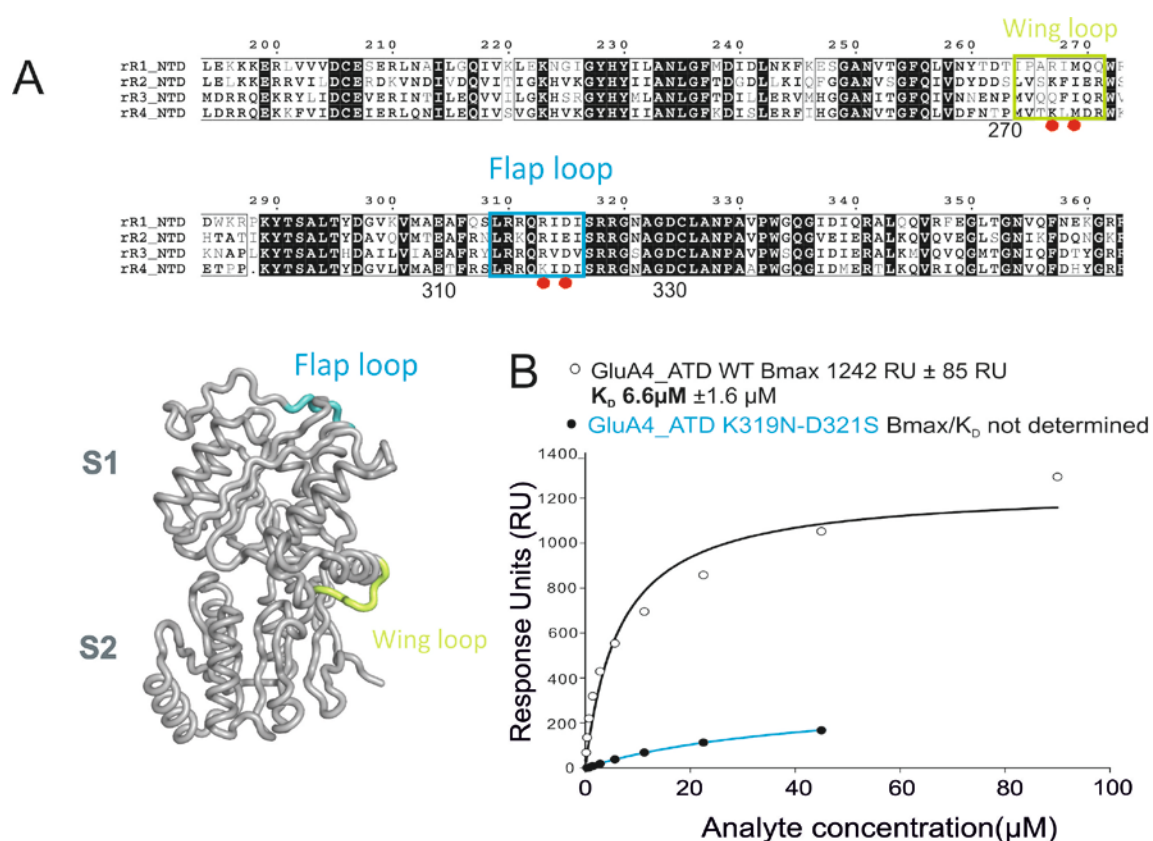


Figure 4.16 Disruption of the interaction between GluA4_ATD and NP1_PTX by structure-guided mutagenesis. A) Insertion of N-linked glycosylation sites at the wing loop (green box) (K274N-M276S) and flap loop (blue box) (K319N-D321S) of the ATD. Mutated residues are indicated by red dots on the AMPA receptor ATD sequence alignment where GluA4_ATD residue numbering is indicated underneath the sequence alignment. Mutated residues are mapped onto the previously published GluA4_ATD structure PDB-ID: 4GPA (Dutta *et al.* 2012). B) Surface plasmon resonance (SPR) analysis of the wild type (WT) GluA4_ATD analyte (at highest concentration 90 μ M) and the flap loop GluA4_ATD (K319N-D321S) mutant (at highest concentration 45 μ M) where 520 units of NP1PX were bound to chip.

4.4.6 The effect of neuronal pentraxin multimerization on GluA4_ATD binding

The structure of the neuronal pentraxin related SAP protein in complex with an Fc γ receptor (Fc γ R) fragment shows one Fc γ R being bound diagonally across the SAP pentamer (Lu *et al.* 2008). Since native neuronal pentraxins appear to be multimeric we sought to determine whether their multimeric forms may result in a binding site with a higher affinity than the monomeric PTX domain for binding to the GluA4_ATD.

Previously published negative stain EM analysis of native NP2_FL would suggest that full length proteins form large aggregates (figure 4.17A) and that these may act to cluster the AMPA receptor (Xu *et al.* 2003). It is highly unlikely that these aggregates are physiological and most likely result from the large amounts of protein generated from mammalian expression

systems utilizing the transfection reagent Lipofectamine 2000TM (Invitrogen) and the elution of purified protein into 0.15M acetate buffer (pH 4.0) (Xu *et al.* 2003). Here, attempts were made to generate full length NP protein (NP1_FL and NP2_FL) in a uniformly oligomeric form, but numerous issues arose. Protein precipitation occurred at all stages of purification, and then the small amount of remaining protein was highly susceptible to proteolysis (figure 4.8C and 4.9C). It was postulated that a TACE cleavage site was present, but attempts to mutate its potential sites of action were fruitless in preventing this degradation.

Negative stain EM data was obtained for NP2_CC2 protein purified in Oxford (figure 4.9B) and analysed by our collaborator Terunaga Nakagawa in San Diego (figure 4.17C) which showed a marked improvement to previous observations in terms of homogeneousness (Xu *et al.* 2003). A trimeric arrangement was consistently observed but the protein was only stable for only a short period of time. Because all native oligomeric forms of neuronal pentraxins were unstable when purified they were deemed unsuitable for cellular and biophysical analysis where result reproducibility is paramount. Artificial PTX-coiled constructs (NP1_PTX_2coil, NP1_PTX_3coil and NP1_PTX_4coil) were engineered and expressed by generation of a polyclonal cell line in the pNeo_Sec vector to enable a constant supply of protein for all future use. The coiled-coil system has previously proved to be a useful tool in deciphering the functional role of oligomerisation in AMPA receptors (Matsuda *et al.* 2005). MALS (figure 4.17E) confirmed masses for NP1_PTX_2coil and NP1_PTX_3coil to be 55.8kDa and 83.7kDa respectively. These were both as expected but the mass of NP1_PTX_4coil (expected mass 111.6kDa) was not and was postulated to be because of steric constraints preventing four globular PTX domains coming together. The NP1_PTX_3coil protein was analysed by negative stain EM (figure 4.17C) and had a similar trimeric ‘bunch of flowers’ arrangement to previously obtained EM images for the NP2_CC2 construct, therefore it was considered valid to use this protein for functional experiments to analyse interactions with AMPA receptors.

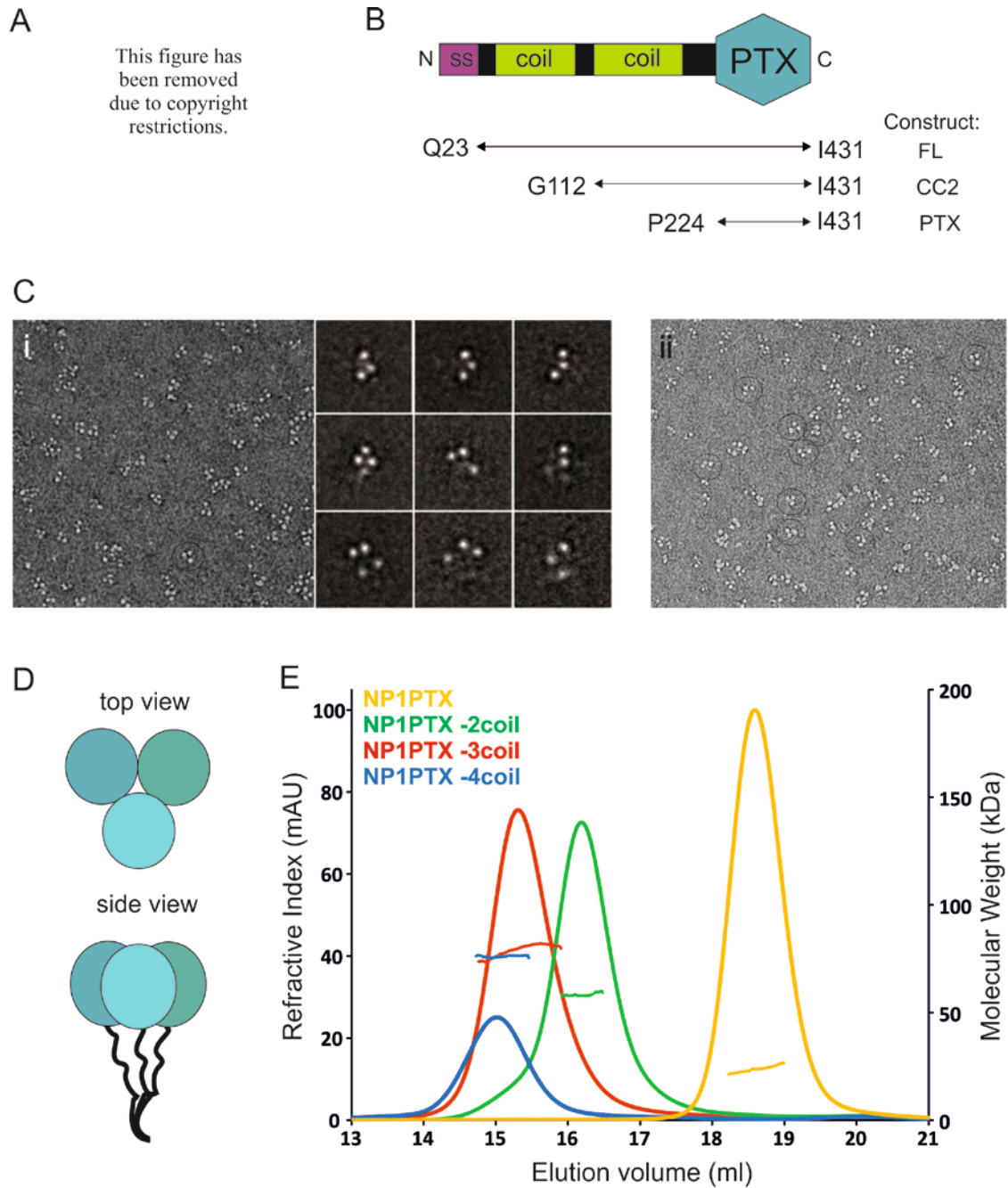


Figure 4.17 Structural and biophysical analysis of neuronal pentraxin 1.

A) Previously published EM analysis of purified NP1 protein shows large aggregates (Xu *et al*, 2003). B) Schematic representation of NP1's domain organisation with two N-terminal coiled-coil regions (lime green) and a large globular C-terminal pentraxin (PTX) domain (blue). C) EM negative stain analysis of i) NP1_CC2 with class averages and ii) NP1_PTX_3coil protein. D) Proposed structure based on class averages showing top and side view. E) MALS analysis of NP1_PTX (yellow) and artificial coil constructs: NP1_PTX_2coil (green), NP1_PTX_3coil (red) and NP1_PTX_4coil (blue).

The engineered oligomeric forms of NP1 were bound as ligands and SPR conducted against the GluA4_ATD analyte. Binding did not appear significantly enhanced as a K_D of $2.9\mu\text{M}$ was consistently observed whether the protein was monomeric, dimeric or trimeric (figure 4.18).

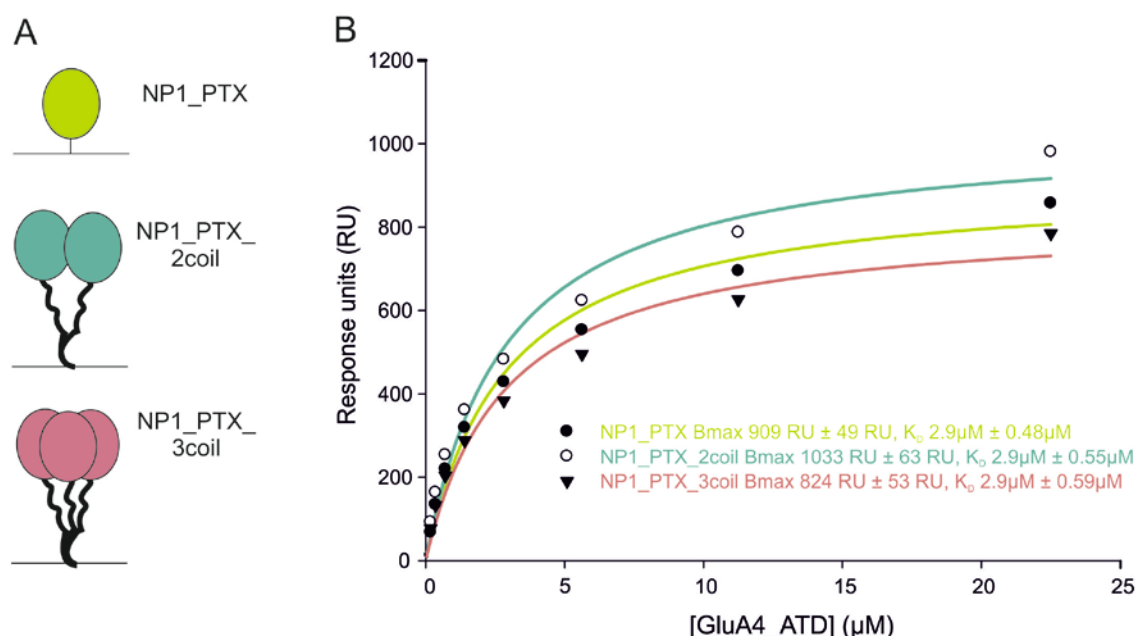


Figure 4.18 Effect of neuronal pentraxin multimerisation on GluA4_ATD binding.

A) Schematic representation of NP1_PTX, NP1_PTX_2coil and NP1_PTX_3coil that were used as ligands in SPR. B) Binding curves obtained when 497 RU, 623 RU and 520 RU of NP1_PTX, NP1_PTX_2coil and NP1_PTX_3coil respectively were bound as ligands. A binding constant (K_D) was calculated to be $2.9\mu\text{M}$ for monomeric, dimeric and trimeric forms of NP1_PTX.

4.4.7 Determining the effect of the AMPA receptor- neuronal pentraxin interaction

To investigate the effect that neuronal pentraxins have on AMPA receptors and investigate the previously observed AMPA receptor clustering (O'Brien *et al.* 1999; Xu *et al.* 2003) FACS analysis was attempted with the hypothesis that clustering of AMPA receptors would act to increase cell surface protein levels. Proteins were as described in section 4.3.8 (FACS analysis) and the anti-HA cell staining procedure enabled only surface expression levels to be detected. Analysis of results indicated a reduction in AMPA receptor (GluA4_FL_mV_HA) surface levels when NP1_PTX_3coil was added at both time points (1hr and 16hrs prior to FACS analysis) which based on previously published observations (O'Brien *et al.* 1999; Xu *et al.* 2003) was quite unexpected. When cells were incubated with the monomeric NP1_PTX, only a slight reduction in surface levels was observed, close to the experimental error for GluA4_FL_mV_HA

alone (figure 4.18A & B). The interaction between AMPA receptors and neuronal pentraxins was previously shown to require the AMPA receptor ATD (Sia *et al.* 2007). Results here indicate that when NP1_PTX_3coil was added to cells expressing the full length protein (GluA4_FL_HA_mV), a larger reduction in surface levels was observed when compared to the detection of protein lacking the ATD (GluA4 Δ ATD_HA_mV) (figure 4.19C & D) which would imply that the ATD was involved in the interaction as previously suggested (Passafaro *et al.* 2003; Sia *et al.* 2007).

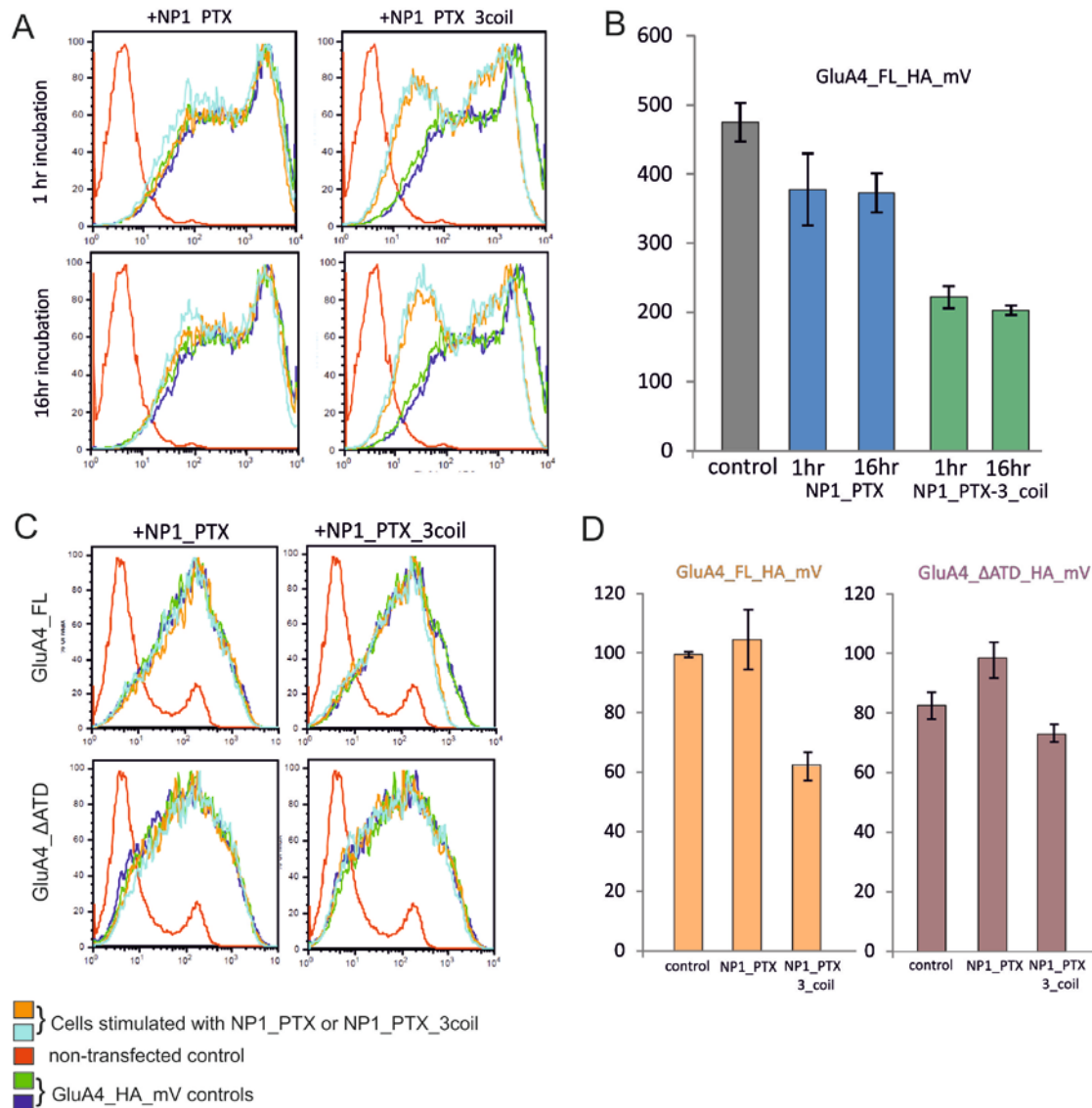


Figure 4.19 FACS analysis of surface detection of GluA4_FL_HA_mV and GluA4_ΔATD_HA_mV expressed in HEK293T cells stimulated with purified NP1_PTX or NP1_PTX_3coil protein at 10μM.

A) FloJo representation of raw geometric mean data for APC detection from cells transfected with GluA4_FL_HA_mV and incubated with NP1PTX or NP1_PTX_3coil for 1hr or 16hrs. The x and y axis labels are FL8 Log APC and % max respectively. B) Reference subtracted detection of anti-HA surface levels shown in A) where cells were stimulated with NP1_PTX (blue bars) (1hr: geometric mean 386 ± 22 , 16hrs: geometric mean 400.5 ± 27.5) or NP1_PTX_3coil (green bars) (1hr: geometric mean 179 ± 5 , 16hrs: geometric mean 175 ± 5) prior to cell staining. Geometric mean of control cells (grey bar): 479 ± 28 , $n=2$.

C) Raw data curves for geometric mean values of APC detection from cells transfected with GluA4_FL_HA_mV or GluA4 lacking the FL8 domain (GluA4_ΔATD_HA_mV), stimulated with NP1_PTX or NP1_PTX_3coil. X and y axis labels are FL8 Log APC and % max respectively. D) Reference subtracted surface detection of data shown in C) normalised to GluA4_FL_HA_mV detection levels (orange bars) (Control cells geometric mean: 100 ± 1), cells incubated with NP1_PTX (geometric mean 104.4 ± 10) or NP1_PTX_3coil (geometric mean 62.24 ± 4.7). GluA4_ΔATD_HA_mV (purple bars) (control 82.74 ± 4.5) when incubated with NP1_PTX (98.44 ± 6.0) or NP1_PTX_3coil (geometric mean 73.14 ± 3). Error bars represent standard error of the mean (SEM), $n=2$.

Colour key identifies histogram samples by colour, where orange and turquoise histograms represent the stimulated cells, specific to that experiment, blue and green histograms represent GluA4_FL_HA_mV positive control (i.e. no pentraxin added). The red histogram represents non-transfected cells, stained with anti-HA that form the negative control (reference).

As already stated, large-scale expression and purification of GluA1_ATD was problematic, where the protein became unstable when it was purified. However, small-scale

expression of GluA1 full length protein expressed with a monoCherry (mCh) and HA tag (GluA1_FL_HA_mCh) (created by Aletheia Lee) was confirmed to traffic to the cell surface by anti-HA cell surface staining (conducted by Aletheia Lee). FACS analysis was therefore attempted with HEK293T cells transfected with GluA1_FL_HA_mCh again stimulated with soluble purified NP1_PTX or NP1_PTX_3coil protein using methods described in Chapter 2, section 2.9.

As observed in FACS analysis of GluA4_FL_HA_mV (figure 4.19), the surface expression GluA1_FL_HA_mCh was reduced when the cells were stimulated with purified NP1_PTX_3coil (figure 4.20A & B) (final concentration 10 μ M). FACS analysis was also conducted with a fragment antigen-binding (Fab) fragment designed to bind specifically to the GluA1 subunit's wing loop (developed by Aletheia Lee). SPR analysis (conducted by Aletheia Lee) found that the Fab bound to GluA1 with 450-500nM affinity. Here FACS analysis revealed that the Fab appeared to enhance surface expression of GluA1_FL_HA_mCh (figure 4.20C & D). When cells expressing GluA1_FL_HA_mCh were stimulated with Fab together with NP1_PTX_3coil an overall reduction in surface expression was observed, but the reduction was not as great as that occurring in cells stimulated with NP1_PTX_3coil alone (figure 4.20 C & D).

Results would therefore indicate one of two scenarios: Either that the Fab fragment was able to partially block the internalisation of GluA1_FL_HA_mCh when co-stimulated with NP1_PTX_3coil by occupying same or similar binding site as NP1_PTX_3coil, or that the apparent reduced internalisation of GluA1_FL_HA_mCh could just be the combined effect of a Fab enhanced surface expression and a NP1_PTX_3coil internalisation, both proteins independently affecting GluA1_FL_HA_mCh surface expression.

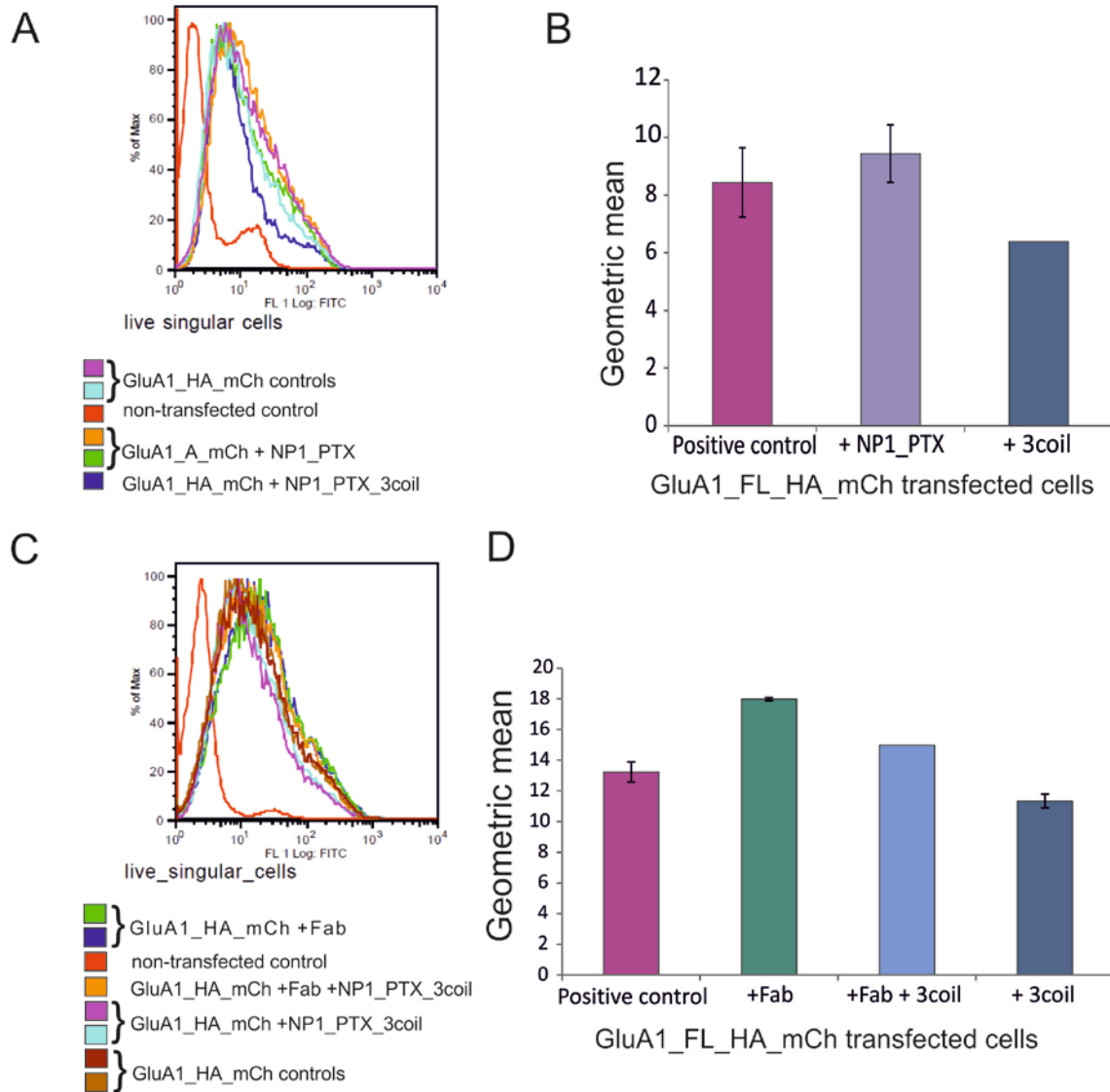


Figure 4.20 FACS analysis of GluA1_FL_HA_mCherry (mCh) transfected HEK293T cells stimulated with Fab and pentraxins.

A) Raw data histograms from anti-HA surface detection of GluA1_HA_mCh. B) Graphical representation of reference subtracted geometric mean values for anti-HA surface detection when cells were transfected with GluA1_FL_HA_mCh positive control (pink) (geometric mean 8.44 ± 1.22 , $n=2$), or incubated with purified NP1_PTX (lilac) (geometric mean 9.44 ± 1.0 , $n=2$) or NP1_PTX_3coil (navy blue) (geometric mean 6.39 , $n=1$) at $10\mu\text{M}$.

C) Preliminary data represented as a histogram (generated in FloJo) to show geometric mean values for Fab competition assay. D) Graphical representation of data in A) where cells transfected with GluA1_FL_HA_mCh = positive control (pink bar) (geometric mean 13.23 ± 0.65 , $n=2$), cells incubated with GluA1_ATD specific Fab fragment (green bar) (geometric mean 17.98 ± 0.1 , $n=2$), cells incubated with Fab and NP1_PTX_3coil (light blue bar) (geometric mean 14.98 , $n=1$) and cells incubated with NP1_PTX_3coil alone (navy blue bar) (geometric mean 11.33 ± 0.45 , $n=2$). Standard error of the mean (SEM) bars present where applicable.

4.5 Discussion

The first aim of this chapter was to structurally characterise the neuronal pentraxin PTX domains. Here structural information was obtained for all three pentraxin (NP1, NP2 and NPR) PTX domains where structural homology was high and only small differences in a 'B face' loop (implicated in calcium binding) was observed. SPR analysis then identified that the binding between the ATD of GluA4 and NP1_PTX was the tightest of all interaction pairs tested. In the absence of a protein complex structure, the interaction of NP1_PTX and GluA4_ATD was mapped by mutagenesis. Mutations were introduced to the PTX domain to disrupt the calcium binding site and to the GluA4_ATD to create a steric block using glycan moieties. SPR analysis conducted with an N-linked glycan mutation introduced at the GluA4_ATD flap loop reduced the binding affinity to NP1_PTX, suggesting that the flap loop could be involved in this interaction. The flap loop of other iGluR ATDs is known to be involved in protein-protein interactions, for example the delta receptor GluD2 interacts with Cbln1 at the corresponding GluD2 flap loop and when GluD2 ATD flap loop residues were mutated to alanine residues it abolished the interaction (Kuroyanagi *et al.* 2010). Disruption of the NP1_PTX calcium binding site also disrupted binding to GluA4_ATD but whether this is due to the site being directly involved in the interaction or acting by allosteric effects remains to be determined.

The oligomeric state of the neuronal pentraxins was previously unknown where EM analysis presented protein as large complexes of aggregated globules (Xu *et al.* 2003). EM analysis of NP2_CC2 protein (construct design based on an observed cleavage product occurring during purification) revealed that a trimeric 'bunch of flowers' form was present but was unstable. Here work to create a homogenous stable trimeric form successfully generated a protein (NP1_PTX_3coil) suitable for use in biophysical and cellular assays. Based on previously published work where AMPA receptor clustering was observed after the addition of neuronal pentraxins (O'Brien *et al.* 1999; O'Brien *et al.* 2002), one would have predicted that an increased level of AMPA receptor might have been observed when NP1_PTX_3coil was added to cells in FACS analysis performed here but a reduction in overall cell surface expression was observed

(figure 4.19). This does not rule out the requirement for an initial pre-clustering or rearrangement that results in the observed receptor internalisation. FACS analysis indicates that internalisation occurs in less than 1 hour (figure 4.19B) but further investigation is required to analyse the rapid processes resulting in AMPA receptor internalisation.

Due to poor expression and purification yields SPR analysis could not be performed with GluA1_ATD protein but FACS analysis was able to provide some insights into its interaction with NP1. Preliminary FACS analysis of cells surface levels of GluA1_FL_HA_mCh after incubation with NP1_PTX_3coil alone caused a reduction of detectable surface protein levels. When cells expressing GluA1_FL_HA_mCh were incubated with a GluA1 wing specific Fab fragment and NP1_PTX_3coil protein it also resulted in an overall, yet less marked, reduction in cell surface GluA1_FL_HA_mCh levels. Incubation with the Fab alone resulted in an increase in surface AMPA receptor levels. The Fab used here was targeted to the wing region of GluA1_ATD due to this regions divergence across iGluR ATDs so any Fab binding to this region would bind specifically to the GluA1 subunit. These current data do not confirm whether the Fab and NP1_PTX_3coil interact at the same site. It is also important to note that there are no known protein- protein interactions that occur at the wing loop of any other iGluR ATD domains, so if the neuronal pentraxins do indeed bind to the wing loop it would be a novel binding site. Most likely is that there are multiple points of contact between the AMPA receptor and the neuronal pentraxin proteins and may involve the wing, flap and other regions of the AMPA receptor ectodomain.

4.6 Future work/ future perspectives

The neuronal pentraxins are noted for their lectin-like ability to bind to proteoglycans in a calcium dependant manner (Tsui *et al.* 1996). More recently the neuronal pentraxins were found to interact directly with PNN surrounding GABAergic PV-INs and increase excitatory synaptic strength onto these inhibitory neurons (Chang *et al.* 2010). Here crystallization attempts failed to obtain a crystal structure of NP1_PTX bound to the heparin analogue sucrose octasulphate (SOS) but exploring other proteoglycans may prove more successful to analyse the lectin properties of neuronal pentraxins. A PNN model was recently used to demonstrate the role of the mesh-like PNNs in CNS repair mechanisms (Kwok *et al.* 2011) where the basic components required to form a PNN were shown to be hyaluronan (HA), chondroitin sulphate proteoglycans (CSPGs), link proteins (HAPLN1), and the extracellular matrix protein tenascin R. Work is currently underway to recreate these artificial PNNs in mammalian expression systems, where light microscopy will be used to examine the effects of their presence on AMPA clustering by neuronal pentraxins. One hypothesis is that these nets anchor the pentraxins to enable clustering of AMPA receptors which would, *in vivo*, result in an increased excitatory drive onto GABAergic PV-INs resulting in inhibition of pyramidal hippocampal neurons as occurs outside of the ‘critical periods’ of learning (figure 4.3).

Electrophysiology will be conducted to examine the effects on AMPA receptor gating activity upon addition of neuronal pentraxins. Work will be conducted by our collaborator Ingo Greger (LMB, Cambridge) and will require substantial amount of protein to be purified. Frozen stocks of NP1_PTX and NP1_PTX_3coil have been stored for such work, hopefully occurring in the near future.

Mouse model experiments will enable *in vivo* analysis of phenotype following addition of neuronal pentraxins. Purified neuronal pentraxin proteins NP1_PTX or NP1_PTX_3coil will be injected into mouse brains. Analysis of performance in memory and behavioural tests will then be

conducted by Aletheia Lee, supervised by David Bannerman (Oxford) but the exact experimental design remains to be finalised.

4.7 References

- Adams, P. D., R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger (2002). "PHENIX: building new software for automated crystallographic structure determination." Acta Crystallographica Section D **58**(11): 1948-1954.
- Al Rahim, M. and M. A. Hossain (2013). "Genetic deletion of NP1 prevents hypoxic-ischemic neuronal death via reducing AMPA receptor synaptic localization in hippocampal neurons." J Am Heart Assoc **2**(1): 006098.
- Backliwal, G., M. Hildinger, I. Kuettel, F. Delegrange, D. L. Hacker and F. M. Wurm (2008). "Valproic acid: A viable alternative to sodium butyrate for enhancing protein expression in mammalian cell cultures." Biotechnology and Bioengineering **101**(1): 182-189.
- Chang, M. C., J. M. Park, K. A. Pelkey, H. L. Grabenstatter, D. Xu, D. J. Linden, T. P. Sutula, C. J. McBain and P. F. Worley (2010). "Narp regulates homeostatic scaling of excitatory synapses on parvalbumin-expressing interneurons." Nat Neurosci **13**(9): 1090-1097.
- Cho, R. W., J. M. Park, S. B. E. Wolff, D. Xu, C. Hopf, J.-a. Kim, R. C. Reddy, R. S. Petralia, M. S. Perin, D. J. Linden and P. F. Worley (2008). "mGluR1/5-Dependent Long-Term Depression Requires the Regulated Ectodomain Cleavage of Neuronal Pentraxin NPR by TACE." Neuron **57**(6): 858-871.
- DeLano, W. L. (2002). "The PyMOL Molecular Graphics System " DeLano Scientific, San Carlos, CA, USA.
- Dodds, D. N. C., I. A. Omeis, S. J. Cushman, J. A. Helms and M. S. Perin (1997). "Neuronal Pentraxin Receptor, a Novel Putative Integral Membrane Pentraxin That Interacts with Neuronal Pentraxin 1 and 2 and Taipoxin-associated Calcium-binding Protein 49." Journal of Biological Chemistry **272**(34): 21488-21494.
- Dutta, A., I. H. Shrivastava, M. Sukumaran, I. H. Greger and I. Bahar (2012). "Comparative dynamics of NMDA- and AMPA-glutamate receptor N-terminal domains." Structure **20**(11): 1838-1849.
- Emsley, J., H. E. White, B. P. O'Hara, G. Oliva, N. Srinivasan, I. J. Tickle, T. L. Blundell, M. B. Pepys and S. P. Wood (1994). "Structure of pentameric human serum amyloid P component." Nature **367**(6461): 338-345.
- Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallographica Section D **60**(12 Part 1): 2126-2132.
- Garlanda, C., B. Bottazzi, A. Bastone and A. Mantovani (2005). Pentraxins at the crossroads between innate immunity, inflammation, matrix deposition, and female fertility. **23**: 337-366.
- Gilmore, C. J., K. Shankland and G. Bricogne (1993). "Applications of the maximum entropy method to powder diffraction and electron crystallography." Proceedings - Royal Society of London, A **442**(1914): 97-111.
- Gu, Y., S. Huang, Michael C. Chang, P. Worley, A. Kirkwood and Elizabeth M. Quinlan (2013). "Obligatory Role for the Immediate Early Gene NARP in Critical Period Plasticity." Neuron **79**(2): 335-346.
- Harbury, P. B., T. Zhang, P. S. Kim and T. Alber (1993). "A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants." Science **262**(5138): 1401-1407.
- Hsu, Y.-C. and M. S. Perin (1995). "Human Neuronal Pentraxin II (NPTX2): Conservation, Genomic Structure, and Chromosomal Localization." Genomics **28**(2): 220-227.
- Kirkpatrick, L. L., M. M. Matzuk, D. C. Dodds and M. S. Perin (2000). "Biochemical interactions of the neuronal pentraxins. Neuronal pentraxin (NP) receptor binds to taipoxin and taipoxin-associated calcium-binding protein 49 via NP1 and NP2." J Biol Chem **275**(23): 17786-17792.
- Kuroyanagi, T. and T. Hirano (2010). "Flap loop of GluD2 binds to Cbln1 and induces presynaptic differentiation." Biochem Biophys Res Commun **398**(3): 537-541.

- Kwok, J. C., G. Dick, D. Wang and J. W. Fawcett (2011). "Extracellular matrix and perineuronal nets in CNS repair." Developmental Neurobiology **71**(11): 1073-1089.
- Lovell, S. C., I. W. Davis, W. B. Arendall Iii, P. I. W. De Bakker, J. M. Word, M. G. Prisant, J. S. Richardson and D. C. Richardson (2003). "Structure validation by $C\alpha$ geometry: ϕ, ψ and $C\beta$ deviation." Proteins: Structure, Function and Genetics **50**(3): 437-450.
- Lu, J., L. L. Marnell, K. D. Marjon, C. Mold, T. W. Du Clos and P. D. Sun (2008). "Structural recognition and functional activation of Fc γ R by innate pentraxins." Nature **456**(7224): 989-992.
- Matsuda, S., Y. Kamiya and M. Yuzaki (2005). "Roles of the N-terminal domain on the function and quaternary structure of the ionotropic glutamate receptor." Journal of Biological Chemistry **280**(20): 20021-20029.
- McCoy, A. J., R. W. Grosse-Kunstleve, L. C. Storoni and R. J. Read (2005). "Likelihood-enhanced fast translation functions." Acta Crystallographica Section D: Biological Crystallography **61**(4): 458-464.
- Murshudov, G. N., A. A. Vagin and E. J. Dodson (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallographica Section D: Biological Crystallography **53**(3): 240-255.
- O'Brien, R., D. Xu, R. Mi, X. Tang, C. Hopf and P. Worley (2002). "Synaptically targeted narp plays an essential role in the aggregation of AMPA receptors at excitatory synapses in cultured spinal neurons." J Neurosci **22**(11): 4487-4498.
- O'Brien, R. J., D. Xu, R. S. Petralia, O. Steward, R. L. Huganir and P. Worley (1999). "Synaptic clustering of AMPA receptors by the extracellular immediate-early gene product Narp." Neuron **23**(2): 309-323.
- Passafaro, M., T. Nakagawa, C. Sala and M. Sheng (2003). "Induction of dendritic spines by an extracellular domain of AMPA receptor subunit GluR2." Nature **424**(6949): 677-681.
- Schlimgen, A. K., J. A. Helms, H. Vogel and M. S. Perin (1995). "Neuronal pentraxin, a secreted protein with homology to acute phase proteins of the immune system." Neuron **14**(3): 519-526.
- Shrive, A. K., G. M. T. Cheetham, D. Holden, D. A. A. Myles, W. G. Turnell, J. E. Volanakis, M. B. Pepys, A. C. Bloomer and T. J. Greenhough (1996). "Three dimensional structure of human C-reactive protein." Nature Structural Biology **3**(4): 346-354.
- Sia, G. M., J. C. Beique, G. Rumbaugh, R. Cho, P. F. Worley and R. L. Huganir (2007). "Interaction of the N-terminal domain of the AMPA receptor GluR4 subunit with the neuronal pentraxin NP1 mediates GluR4 synaptic recruitment." Neuron **55**(1): 87-102.
- Stuart, D. I., M. Levine, H. Muirhead and D. K. Stammers (1979). "Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å." Journal of Molecular Biology **134**(1): 109-142.
- Tsui, C. C., N. G. Copeland, D. J. Gilbert, N. A. Jenkins, C. Barnes and P. F. Worley (1996). "Narp, a novel member of the pentraxin family, promotes neurite outgrowth and is dynamically regulated by neuronal activity." Journal of Neuroscience **16**(8): 2463-2478.
- Xu, D., C. Hopf, R. Reddy, R. W. Cho, L. Guo, A. Lanahan, R. S. Petralia, R. J. Wenthold, R. J. O'Brien and P. Worley (2003). "Narp and NP1 form heterocomplexes that function in developmental and activity-dependent synaptic plasticity." Neuron **39**(3): 513-528.

Chapter 5

Analysis of other AMPA receptor ectodomain- interacting proteins

5.1 Introduction

The AMPA receptor interactome orchestrates many dynamic processes involved in receptor trafficking, recruitment and localisation that have been implicated in synaptic plasticity (described in detail in chapter 1) (Shepherd *et al.* 2007). AMPA receptor trafficking can be controlled by phosphorylation of the intracellular C-terminal domain of the GluA1 subunit by protein kinase A (PKA). This enables an interaction with the actin association protein 4.1, triggering receptor trafficking and membrane insertion at extra-synaptic locations, ready for insertion into the synapse (Oh *et al.* 2006; Lin *et al.* 2009). AMPA receptors are anchored at the synapse by PSD cytoplasmic scaffolding proteins and auxiliary proteins (such as transmembrane AMPA receptor regulatory protein-2 (TARP2)) which tether receptors to PSD-95 preventing their lateral diffusion and recycling (Hee Jung *et al.* 2000; Jin *et al.* 2006; Shepherd and Huganir 2007). Initiated by activity, calcium dependant GluA1 dephosphorylation disrupts intracellular protein-protein interactions to trigger AMPA receptor endocytosis (Ehlers 2000).

In addition to an intracellular interactome, the extracellular regions of other iGluRs (NMDA, kainate & Delta receptors) are known to be involved in a variety of ectodomain-specific interactions. These interactions may function either to locate the iGluRs to a required site, or to modulate receptor function. Synaptic recruitment and clustering of NMDA receptors is mediated through an interaction of the NMDA receptor ectodomain with the EphB receptor-EphrinB complex, acting to locate the NMDA receptor at the synapse which stimulates excitatory synapse formation (Dalva *et al.* 2000). In addition to this, the ATD flap loop, present in the extracellular domain of the Delta receptor GluD2 subunit, interacts with secreted Cbln1 at the synapse to induce pre-synaptic formation at cerebellar parallel-fiber (PF)-Purkinje neuron (PN) synapses (Kuroyanagi *et al.* 2010). Channel conductance properties of kainate receptors is modulated by extracellular interactions with their auxiliary proteins Neto1 and 2, where mutations in the extracellular domains of Neto1 were found to eliminate its effects on kainate receptor desensitization (Fisher *et al.* 2012).

AMPA receptors are also known to form extracellular protein-protein interactions and are increasingly seen at the core of larger supra-molecular assemblies (figure 1.9) The neuronal pentraxins are postulated to cluster the AMPA receptor, acting to enrich synaptic receptor numbers (Gu *et al.* 2013) (discussed in Chapter 4). Research suggests that the N-cadherins interact directly with the ATD of GluA2 to decrease the receptor's lateral movements that was shown to promote spine formation (Saglietti *et al.* 2007). In both of these examples, the interactions are postulated to have roles in locating the AMPA receptor. Extracellular AMPA receptor protein-protein interactions acting to modulate channel conductance have also been proposed. TARP2 (Stargazin), a trailblazer to the iGluR social network, has well documented intracellular interactions with the AMPA receptors with roles in AMPA receptor trafficking and synaptic targeting (Rouach *et al.* 2005). Of interest to this chapter is the structurally uncharacterised TARP2 ectodomain loop-2 that is postulated to regulate AMPA receptor gating via a novel interaction with the AMPA receptor ectodomain (Milstein *et al.* 2009). The extracellular cystine-knot region of cystine-knot AMPA receptor modulating protein-44 (CKAMP44) has also been shown to modulate AMPA receptor channel conductance (described in Chapter 1, section 1.6) (Von Engelhardt *et al.* 2010) but again the protein is structurally uncharacterised. Finally several secreted proteins that potentially interact directly with the AMPA receptor ectodomain were recently identified in pull-down assays and mass spectrometry analysis (Schwenk *et al.* 2012; Shanks *et al.* 2012) but the effect that these interactions have on AMPA receptors are currently unknown. One would hypothesize that novel AMPA receptor extracellular protein-protein interactions may have similar roles (in receptor localisation or channel modulation) as those previously described for other iGluRs.

5.2 Aims

Many auxiliary partners putatively interact with the AMPA receptor ectodomain, but the stoichiometry, structural arrangement and functional consequences of such interactions are yet to be fully defined. TARPs and CKAMP44 are known to modulate AMPA receptor gating mechanisms but their interactions with the AMPA receptor ectodomain remains structurally uncharacterised. The recently identified novel AMPA receptor interactome proteins have unknown effects on AMPA receptors and their binding sites are also uncharacterised. Therefore the aims of this chapter were to:

- 1) Express secreted ectodomain regions of AMPA receptor extracellular interactome proteins (TARP2, CKAMP44 and the novel proteins).
- 2) Conduct SPR analysis to biophysically characterise if and how these proteins interact with the AMPA receptor ectodomain (GluA2 subunit).
- 3) Use X-ray crystallography to give vital insights into the expanding AMPA receptor interactome.
 - a) Attempt crystallization of structurally uncharacterised proteins.
 - b) Attempt co-crystallization of protein-protein complexes to structurally characterise interactions.

5.3 Materials and Methods

5.3.1 Construct design, PCR and cloning overview

PCR amplification, restriction digest, plasmid ligation and *E. coli* transformation protocols were conducted as previously described (Chapter 2, section 2.1). Construct residue numbering commences from the start of each protein's native secretion signal. Accession numbers for all constructs are stated in Chapter 2 table 2.1. All oligonucleotide information for proteins described in this chapter can be found in Appendix C. Some cDNA constructs were gifts from collaborators, details of which can be found below.

5.3.2 Generation of the GluA2_LBC

GRIA2_human cDNA (GenBank accession number BC010574) was used in an SOE-PCR to generate an artificial ligand binding core (LBC) domain termed GluA2_LBC (K414-G795ΔP528-S652) in which residues P528- S652 (which form part of the transmembrane region) were replaced by a GT linker. After ligation into pHLsec, GluA2_LBC was sub-cloned into pET22b via NdeI/XhoI restriction sites for expression in *E. coli* systems. GluA2_LBC was then expressed in OrigamiTM cells (Novogen) as previously described (Chapter 2, section 2.5) for experiments involving protein complex formation, crystallization and biophysical analysis.

5.3.3 TARP2 (Stargazin) construct optimisation

The TARP2 protein (sequence in Appendix C) is predicted to be a member of the tetraspanin family of proteins which share a common domain organisation. Their N- and C-termini are both intracellular and they contain a large four-pass transmembrane domain and two extracellular loops with low secondary fold prediction (figure 5.4A) (Hemler 2005). Previous research suggested that the TARP2 ectodomain loop-2 was involved in the regulation of AMPA receptor gating (Milstein and Nicoll 2009), so a synthetic codon-optimised construct (produced by Geneservice) was designed to contain only the second loop of TARP2 (G29-S105) (Appendix

C for details). The TARP2 loop-2 (G29-S105) was inserted into the Trx scaffold (figure 5.4B) (sequence in Appendix G). The TARP2 loop-2 in the Trx scaffold, termed from hereon in as TARP2_Trx, was cloned into the pET22b vector via NdeI/XhoI restriction sites for expression in *E. coli* systems. Sequencing analysis confirmed correct sequence was present (sequencing analysis in Appendix C).

TARP2_Trx was purified from 3L of *E. coli* culture after expression that resulted in inclusion body formation (Chapter 2 section 2.5.4). Briefly RosettaTM cells were transformed with TARP2_Trx_pET vector. The cell culture was induced with 0.8mM IPTG and grown for 6 hours at 37°C to force proteins into inclusion bodies and cells were harvested, lysed and centrifuged. Inclusion body pellets were re-suspended and washed then the proteins were purified by anion exchange and SEC. During anion exchange the NaCl content of the protein solution was reduced to 50mM via a 1 in 10 dilution in 10mM Tris pH 8.0. The protein was bound to a 5ml Q column (GE Healthcare) using a peristaltic pump (5ml/min) and separated according to net charge by use of a constant gradient increase from a low salt 'Buffer A' (10mM Tris pH 8.0, 50mM NaCl), into a high salt 'Buffer B' (10mM Tris pH 8.0, 500mM NaCl) (figure 5.5C).

5.3.4 CKAMP44 construct design

The full length CKAMP44 (GenBank accession number NP_0011386762) contains 465 amino acids but only residues L193-T212 are stated to have a secondary structure and are predicted to form a transmembrane domain (HHpred) (figure 5.1 highlighted in blue). Three N-linked glycosylation sites are predicted at residues N86, N130 and N157 (figure 5.1 highlighted in purple bold font). The ectodomain contains multiple cysteine residues (red bold font underlined bold), 8 of which reside within a 66 residue stretch that has been predicted to form a cystine-knot (Von Engelhardt *et al.* 2010).

MCGPRRLAAPGSSSLGSFGRASRGRRGSSSGRAAEPGLGDTMRRVLRLLLGCFLELTCARVCRQAQERAGHGQLAQLGGVLLLAGGN
 RSGAASGEASEGAEASDAPPTRAPTPDFCRGYFDVMQWDPPFNCSSGDFIFCCGTGGRFCCTFKKRRLNQSTCTNYDTPLWLNT
 GKPPARKDDPLHDPTKDKTNLIVYIICGVAVMVLVGIFTKLGLEKAHRPQREHMSRALADVMRPQGHGHCNTDHMERDLNIVVHVQH
 YENMDTRTPINNHLATQMNNVPTSPLLQQMGHPSYPNLGQISNPYEQPPGKELNKYASLKAVGSSDGDWAVSTLKSPKADKVN
 DDFYTKRRHLAELAAGNPLPLHPVRVEDEPRAFSPEHGPAKQNGQKSRNTKMPHPLAYTSTNFKGWDPNESLRRQAYSNGKGL
 GTAETGSSDPLGTRPQHYPPQPYFITNSKTEVTV

Figure 5.1 CKAMP44 domain organisation.

Constructs designed for *Homo sapiens* CKAMP44 protein with GenBank accession number NP_0011386762. Secondary structure prediction (HHPred, <http://toolkit.tuebingen.mpg.de/hhpred>) indicates the protein's trans-membrane domain (L193-T212) (blue highlight). Predicted N-linked glycosylation sites at residues N86, N130 and N157 are indicated by purple bold font. The region predicted to form the ectodomain contains multiple cysteine residues (indicated by red bold underlined font).

A CKAMP44 ectodomain construct (M42-T191) and a cystine-knot (cys-knot) construct (D113-K174) were cloned into the pHLsec and pHLsec_Avi vectors. Additional constructs D113-H184 (CKAMP44_17) and D113-T191 (CKAMP44_18) to optimise the start and end residues around the cys-knot were also generated.

5.3.5 Other putative interactions

Brorin (GenBank accession number NP001102782.1), Noelin-1 (GenBank accession number NP062371.2), Noelin-2 (GenBank accession number NP001015017.1), Noelin-3 (GenBank accession number NP703188.1) and Neuritin (GenBank accession number NP445798.1) constructs were gifts from Nadine Harmel (Düsseldorf, Germany), who published the original paper proposing their interaction with AMPA receptor ectodomains (Schwenk *et al.* 2012). Constructs were designed to encompass only the secreted ectodomains for cloning into the pHLsec vector. Additional information detailing oligonucleotides for PCR amplification, cloning and sequence information can be found in Appendix C.

The full length Brorin protein is 324 amino acids in length with residues M1-C27 coding for a native secretion signal. Residues T117 and T137 were predicted to be O-linked glycosylation sites (NetOglyc) (Steentoft *et al.* 2013) and are depicted schematically in figure 5.8D. The protein contains two predicted von Willebrand C (VWC) cysteine rich domains. A full length Brorin (Br_FL) (M1-M324) construct was designed to utilise the protein's native secretion signal. Br_FL was ligated into pHLsec after PCR incorporation of EcoR1/KpnI restriction sites. Analysis of the Brorin amino acid sequence highlighted a region upstream of the VWC domain

that was predicted by regional order neural network (RONN) (<http://www.strubi.ox.ac.uk/RONN>) to be disordered (figure 5.2).

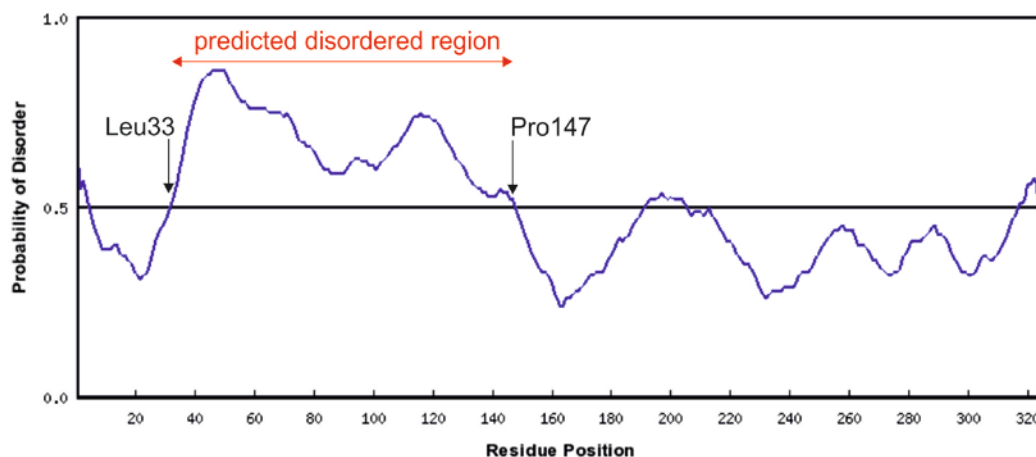


Figure 5.2 RONN disorder prediction for Brorin

Regional order neural network (RONN) <http://www.strubi.ox.ac.uk/RONN> predicts Brorin to be disordered from residue 33 (Leu) to residue 147 (Pro).

This disordered region might have hindered crystallization, so a new primer was designed to generate a new shorter Brorin construct. This shortened Brorin construct, (Br_short) (D148-M324) as well as other Brorin constructs S28-M324 (S1), R48-M324 (R21), D68-M324 (D41), E88-M324 (E61), D108-M324 (D81) and E128-M324 (E101) were ligated into the pHLsec vector to utilise the vectors secretion signal therefore AgeI/KpnI restriction sites were incorporated into oligo design. Some construct were initially numbered using a different system and so reference to the original oligo numbers (used in Appendix C) is indicated by the residue numbers in brackets.

Finally the Noelin-1 (EcoR1/KpnI), Noelin -2 (EcoR1/BsiWI), Noelin -3 (EcoR1/KpnI) and Neuritin (EcoR1/KpnI) constructs were ligated into pHLsec and pHLsec_Avi after PCR incorporation of suitable restriction sites indicated above in brackets.

5.4 Results

5.4.1 Expression of GluA2_LBC

The GluA2_LBC structure has been solved many times in complex with various small molecule agonists and antagonists (examples include PDB IDs: 3BFU, 3BBR and 1MS7) but never in complex with another protein. Work to crystallize GluA2_LBC in complex with the various auxiliary proteins described here was therefore undertaken. The GluA2_LBC construct was expressed in Origami *E. coli* cells and purified alone (Chapter 2, section 2.5). The LBC has a well-documented glutamate binding site (Armstrong *et al.* 2000) so it was postulated that addition of L-glutamine would stabilise the protein in a closed conformation. L-glutamine (2mM final concentration) was therefore included in IMAC elution and SEC buffers to prevent protein precipitation. Typical yields of 1mg/L of the 29.2kDa protein were obtained. A single peak corresponding to the elution volume expected for a monomeric form of the protein was observed on SEC (figure 5.3).

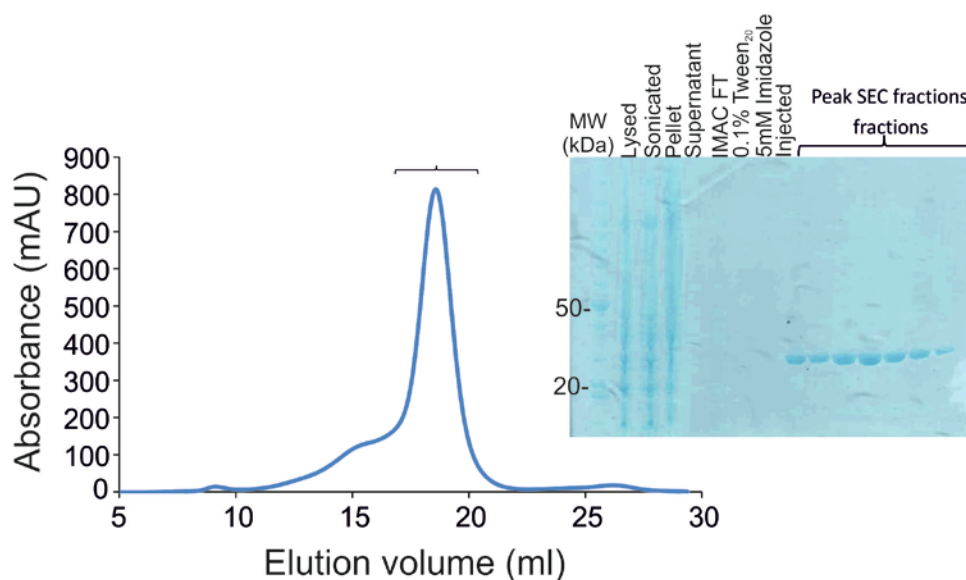


Figure 5.3 Purification of GluA2_LBC.

SEC analysis of GluA2_LBC after IMAC purification on Superdex S200 10/30 column with 10mM HEPES, pH7.5, 500mM NaCl, 2mM L-glutamate running buffer. 12% SDS-PAGE analysis including lysed, sonicated, cell pellet and supernatant fractions from *E. coli* expression. IMAC media flow through (FT), 0.1% w/v Tween-20 and imidazole washes also included alongside SEC peak fractions.

5.4.2 TARP2_Trx expression/ purification/ crystallization attempts

Previous research had identified involvement of the TARP2 ectodomain loop-2 in regulation of AMPA receptor gating (Milstein and Nicoll 2009), so the synthetic TARP2 loop-2 (G29-S105) (Appendix C) (schematic figure 5.4A) was inserted into a Trx scaffold (figure 5.4B). This scaffold had previously been used as a peptide aptamer scaffold and here allowed the region of interest (loop-2) to be presented on the surface of the well-ordered globular Trx protein (Borghouts *et al.* 2008; Brown *et al.* 2010) (Appendix G for Trx sequence information).

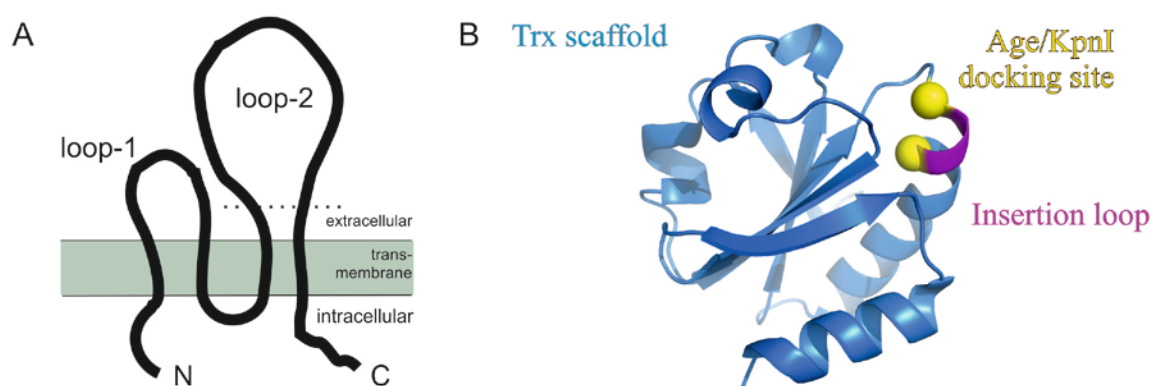


Figure 5.4 TARP2_Trx construct generation.

A) Schematic of a tetraspanin protein. The protein's N- and C- terminals are intracellular. There are four transmembrane domains and two extracellular loops with low secondary structure prediction. B) Trx scaffold structure (PDB ID 2TRX (Katti *et al.*, 1990)). Trx protein overall fold shown in blue. Residues highlighted in yellow are the modified glycine residue sites at which to ligate in target sequence (i.e. TARP2 loop-2) (removing purple loop (residues GTGPGGTG)).

TARP2_Trx expression was optimised in small-scale expression trials in RosettaTM and B834 (DE3) competent *E. coli* cells where SDS-PAGE analysis of induced verses non-induced cells shows a prominent band at 22kDa (figure 5.5B). A higher level of inducible protein is present in the RosettaTM cell fractions and expression leakage is evident in the non-induced cellular fraction.

Non-reducing gel analysis of peak fractions from large scale *E. coli* expression and inclusion body purification indicates that one peak at 60-75mls (at approximately 250mM NaCl) contained protein of a single molecular weight (22kDa) (figure 5.5C). All other peak fractions contained multiple bands of protein of the same molecular weight and higher, demonstrating the

nature of the refold protocol. The monomeric peak fractions were then pooled and concentrated for size exclusion chromatography (SEC) on a Superdex S200 10/30 column in 10mM HEPES pH 8.0, 0.5M NaCl (figure 5.5D). Protein was visually judged to be >95% pure and western blot analysis confirmed that a His-tagged protein has been purified (figure 5.5E).

The purified TARP2_Trx protein was mixed in a 1:1 ratio with GluA2_LBC and crystallization screens were attempted. Several crystals grew (figure 5.5F, G & H) but analysis of diffraction data concluded that the crystals only contained GluA2_LBC protein. Matthews's coefficient calculations found that the unit cell dimensions were incompatible with mass of a complex and when (for the sake of completeness) the previously solved GluA2_LBC was used as a molecular replacement model, it fit all available electron density, clearly leaving no space for the TARP2_Trx protein.

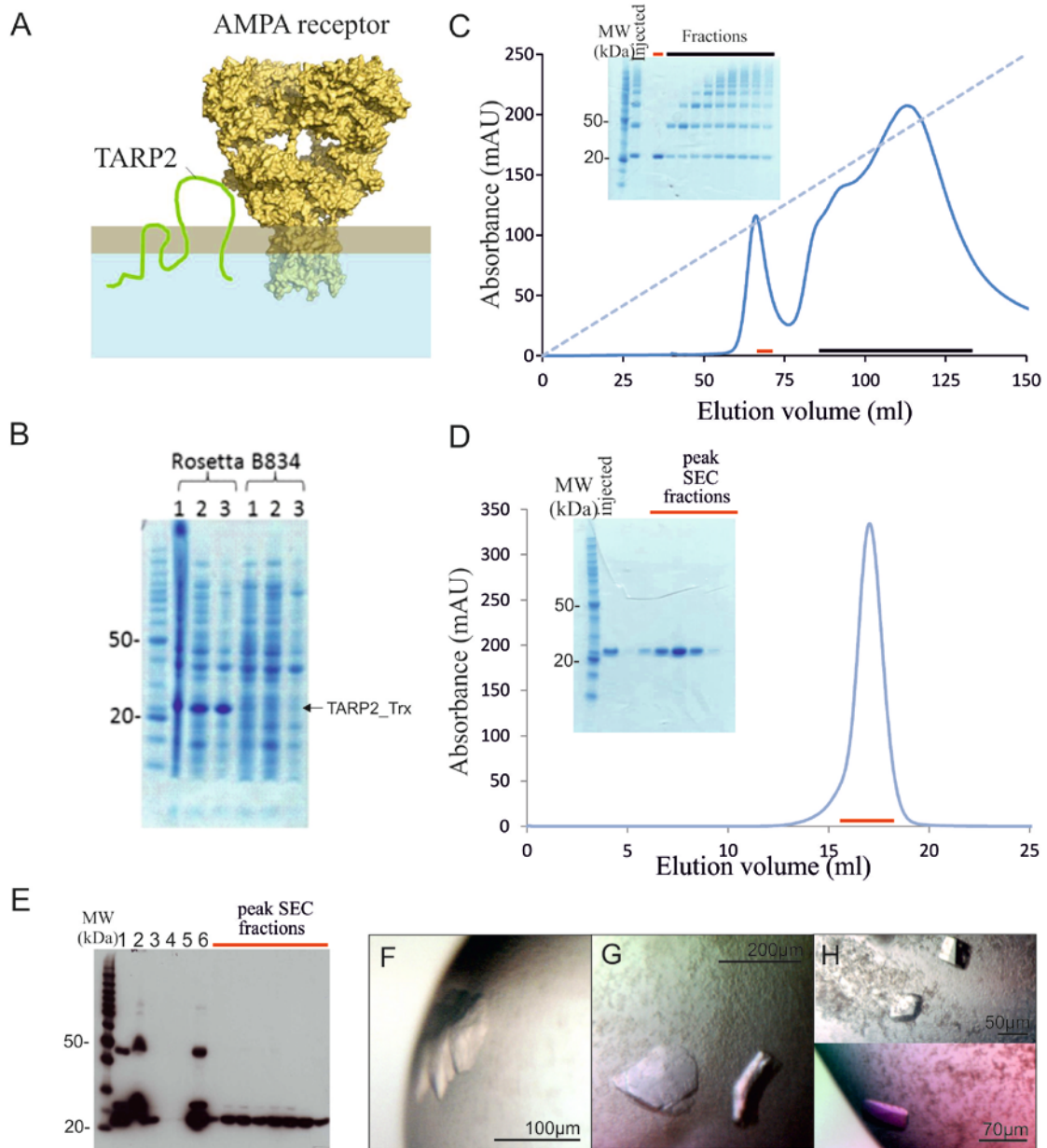


Figure 5.5 TARP2 expression, purification and co-crystallization attempts.

A) Postulated extracellular interaction between the AMPA receptor and TARP2 loop-2. B) TARP2-Trx small-scale expression trials in Rosetta™ and B834 (DE3) competent *E. coli* cells. Analysis of non-induced cells (lane 1) whole cell lysed pellet (lane 2) and inclusion bodies pre-wash (lane 3) were analysed by SDS-PAGE with molecular weight marker Benchmark protein ladder (Invitrogen). C) Anion exchange of TARP2-Trx purified from 3L expression in Rosetta *E. coli* cells and non-reduced gel analysis. D) SEC purification of TARP2-Trx including SDS-PAGE analysis with protein molecular weight marker Benchmark (Invitrogen), E) western blot analysis showing in lane 1) sonicated cells, 2) cell pellet, 3) IMAC flow through (FT) protein, 4) Tween₂₀ wash, 5) Imidazole wash, 6) protein injected on SEC and peak SEC fractions with molecular weight marker Benchmark prestained (Invitrogen), F, G and H) GluA2_LBC and TARP2-Trx co-crystallization attempts. Crystals grew in a variety of PEG containing conditions, examples include: F) 0.15M ammonium sulfate, 0.1M MES pH6.0, 15 % w/v PEG 4000, G) 0.2M lithium sulfate 0.1M MES pH6.0, 20 % w/v PEG 4000 and H) 0.1M sodium citrate tribasic dihydrate pH5.5, 16 % w/v PEG 8,000.

5.4.3 CKAMP44 expression and purification attempts

Expression of the CKAMP44 ectodomain construct (M42-T191) was initially tested in HEK293T cells and expressed protein was secreted but the SEC elution volume indicated that it was aggregated (data not shown). Because the interaction of CKAMP44 with the AMPA receptor LBC is thought to involve the cys-knot region of CKAMP44 (Von Engelhardt *et al.* 2010). A construct was developed to contain only the predicted cys-knot region (D113-K174) but in small-scale expression attempts it was not secreted. Ectodomain constructs to optimise the start and end residues around the predicted cys-knot were therefore designed. Of note here are two constructs containing residues D113-H184 (CKAMP44_17) and D113-T191 (CKAMP44_18) that were secreted in small-scale transient HEK293T expression trials. Both proteins were then expressed in 3L of HEK293T cells with kifunensine (1mg/L final concentration). The SEC elution profile of IMAC purified CKAMP44_17 contained a single peak with a slight shoulder (figure 5.6A) and CKAMP44_18 (figure 5.6B) revealed a single peak, around ~10mls elution volume. These peaks were not at the expected elution volume for a monomeric protein (molecular weight 9kDa) and potentially eluted at the column void volume. Attempts to deglycosylated CKAMP44_18 appeared successful where a shift in the SEC profile was observed (figure 5.6B). SDS-PAGE analysis confirmed that the protein dissociate into monomers (16.4kDa) (figure 5.6C gel 1 and 2), suggesting that the aggregation or higher oligomeric form was caused by disulphide bond formation which were reduced by β -mercaptoethanol. This was confirmed by non-reducing gel analysis where a protein of high undetermined molecular weight was observed (figure 5.6C gel 3).

Because of the success in obtaining monomeric TARP2_Trx from an *E. coli* refold purification (figure 5.5), the CKAMP44 construct (D113-T191) enveloping the predicted cys-knot ectodomain region, postulated to interact with the GluA2_LBC (Von Engelhardt *et al.* 2010), was cloned into the Trx scaffold. Small-scale expression trials were again attempted in RosettaTM and B834 *E. coli* cells, where an inducible band was more prominent in the RosettaTM cells (figure

5.6D). The refold protocol used to purify TARP2_Trx was reproduced, but this time unfortunately failed. SEC analysis revealed only very small amounts of aggregated protein was present (figure 5.6E). More work to optimise the refold protocol is still therefore required.

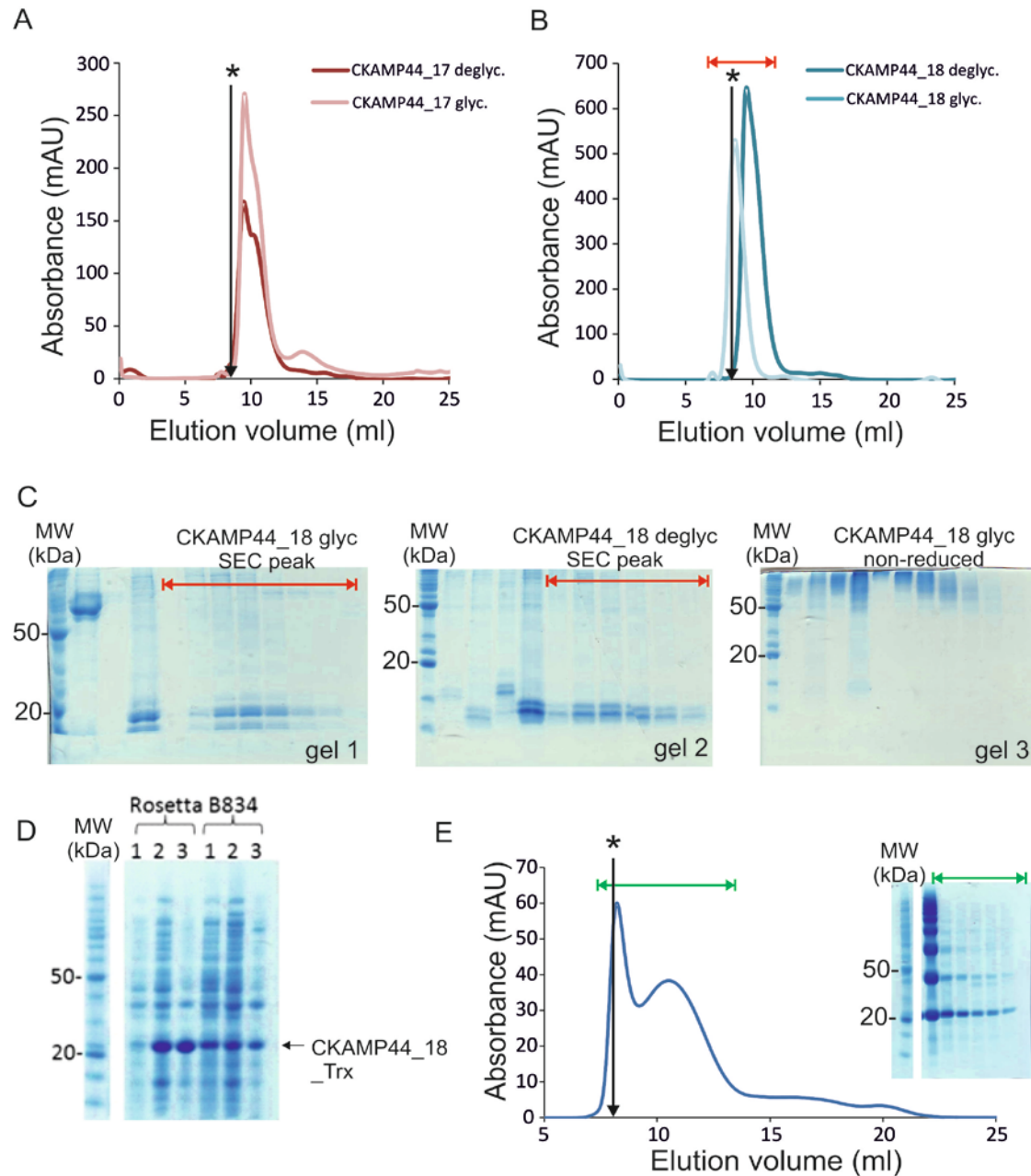


Figure 5.6 CKAMP44 Expression and purification attempts.

A) CKAMP44_17 large-scale transient expression in HEK293T cells with kifunensine (kif) prior to and post Endonuclease F1 (EndoF1) treatment (glycosylated (glyc) and deglycosylated (deglyc)). B) CKAMP44_18 large-scale transient expression in HEK293T cells +kif prior to and post EndoF1 treatment. In all cases SEC was performed on an S75 10/30 column in 10mM HEPES pH 7.5, 150mM NaCl. C) SDS-PAGE analysis of CKAMP44_18 glyc (gel 1, 12% acrylamide) CKAMP44_18 deglyc (gel 2, 15% acrylamide) and sample analysed on a non-reduced gel (CKAMP44_18 deglyc) (gel 3, 15% acrylamide). Gel load order in gel 1: MW marker, IMAC flow through, 5mM Imidazole wash, 1µl of sample injected for SEC then SEC peak fractions as indicated by red line. D) Expression attempts of CKAMP44_18_Trx in Rosetta and B834 *E. coli* cells. Lane 1) non-induced cells, lane 2) whole cell lysed pellet, 3) inclusion bodies pre-wash stage. Molecular weight marker Benchmark (Invitrogen) was used for all SDS-PAGE gels. E) SEC and SDS-PAGE analysis of large scale *E. coli* expression attempts. Black arrows with asterisks (*) indicate approximate column void volumes.

5.4.4 Biophysical analysis of GluA2_LBC and CKAMP44

The CKAMP44 ectodomain construct M42-T191 was ligated into pHLsec_Avi and protein expressed in small-scale transient transfection of HEK293T cells. The secreted protein was biotinylated to be used as a ligand in SPR by methods previously described (Chapter 2 section 2.8.2). Purified GluA2_LBC protein was concentrated to 80 μ M and a dilution series established. SPR was conducted in the presence of 2mM L-Glutamine. Analysis of data found that these two proteins do interact, with a binding constant of 6.2 μ M observed (figure 5.7).

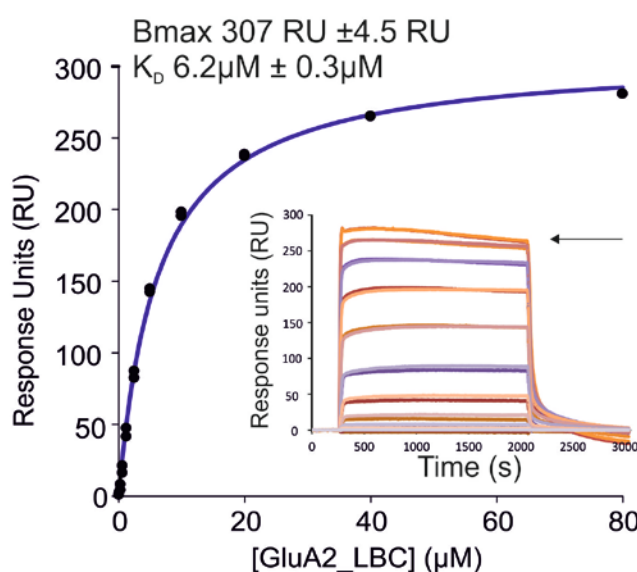


Figure 5.7. SPR analysis of GluA2_LBC binding to CKAMP44 ectodomain.

SPR analysis of the CKAMP44 ectodomain (bound as ligand) interacting with the AMPA receptor GluA2_LBC (analyte). Data analysis predicts a B_{max} of 307 response units (RU) and a binding constant (K_d) 6.2 μ M with 1770 units of CKAMP44 bound as ligand. Inlay shows raw data curves obtained for the dilution series of GluA2_LBC (80 μ M highest concentration). Running buffer 10mM HEPES pH 7.5, 150mM NaCl, 0.005% w/v Tween-20 and 2mM L-glutamine. Analysis performed on a Biacore T100.

At high concentrations the GluA2_LBC protein behaved anomalously, where binding to CKAMP44 seemed to decrease during the injection (figure 5.7 raw data curves, indicated by the arrow). There were also 1770 units of CKAMP44 ligand bound to the chip so although a binding curve was generated, one would have expected a larger B_{max} value. What cannot be fully determined is the oligomeric state of the CKAMP44 (M42-T191) protein on the chip. Oligomeric or aggregated protein may have reduced the number of binding sites available to GluA2_LBC.

These data would suggest that the majority of available binding sites were occupied at higher concentrations as near saturation was reached on the binding curve (figure 5.7).

5.4.5 Analysis of other putative AMPA receptor interactions

Noelin-1, Noelin-3, Brorin (full length) and Neuritin constructs were cloned and sequence verified before small-scale HEK293T expression trials were conducted. Western blot analysis confirmed the secretion of all proteins where bands were present in the media fraction (M lanes in figure 5.8A). Interestingly, it was frequently observed that constructs expressed in the pHLsec_Avi-tag vector were secreted to a higher level than those expressed in pHLsec, where Noelin-1 was actually only secreted when expressed with an Avi-tag. The proteins were then prepared for SPR analysis by standard protocol (Chapter 2 section 2.8.2).

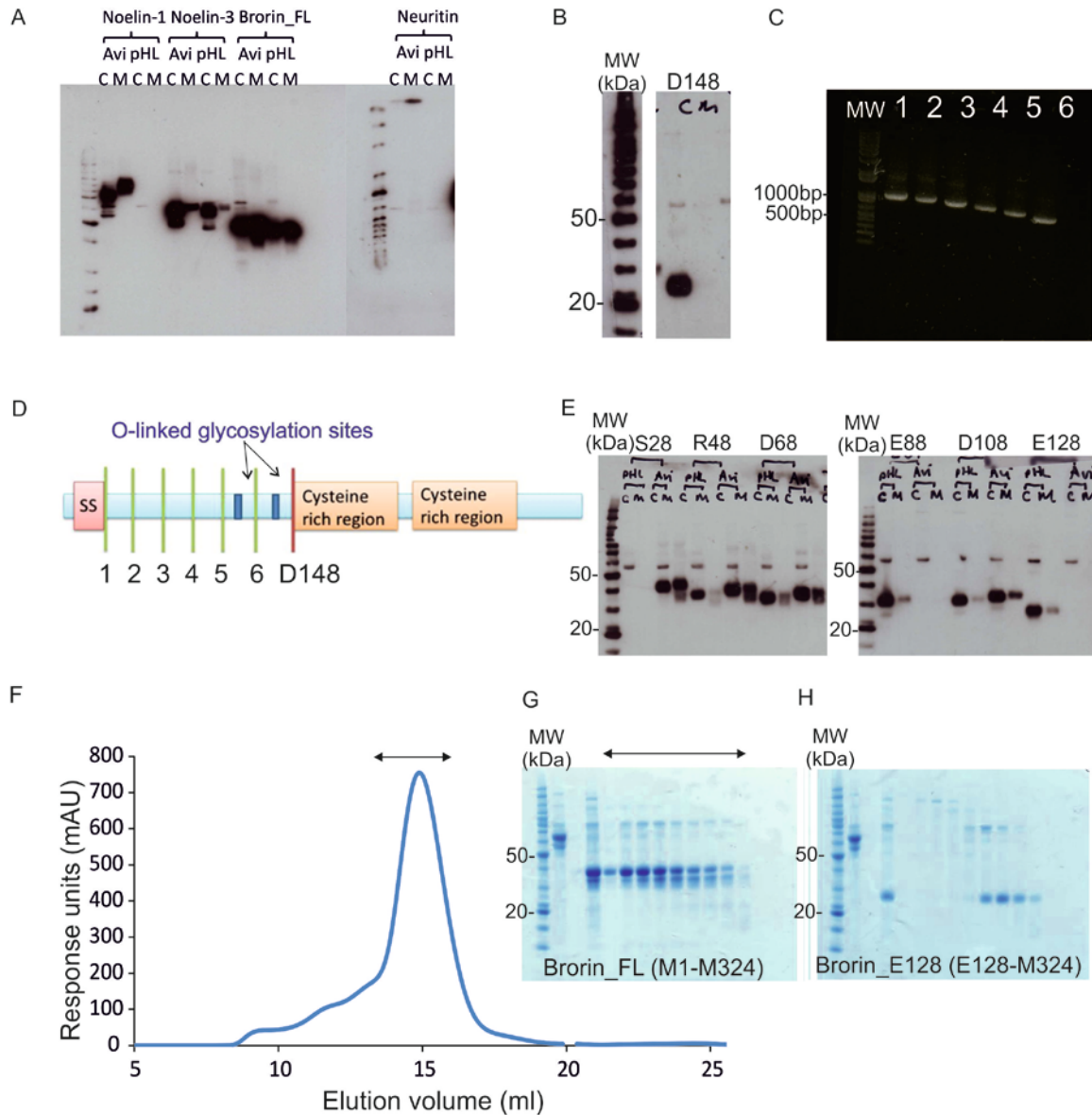


Figure 5.8 Expression and purification of novel AMPA receptor-interacting proteins.

A) Western blot analysis of Noelin-1, Noelin-3, Brorin full-length (FL) (M1-M324) and Neuritin-1 protein from small-scale HEK293T expression trials with cellular (C) and secreted (media, M) fractions analysed. B) Western blot analysis of Br_sh (D148-M324) cells and media fractions to show this construct was not secreted. C) PCR amplification of Brorin truncated constructs S28-M324 (1), R48-M324 (2), D68-M324 (3), E88-M324 (4), D108-M324 (5) and E128-M324 (6) as shown in (D) as green lines. D) Schematic representation of the full length Brorin construct (Br_FL) (M1-M324) with secretion signal (SS), and truncated constructs as in (C). Predicted O-linked glycosylation sites are indicated by blue boxes, the start residue for the Br_sh construct (D148-M324) that was not secreted is indicated by a red line, the two predicted cysteine rich regions are indicated by orange boxes. E) Western blot analysis of Brorin truncated proteins expressed in pHLsec and pHLsec_Avi with cellular (C) and media (M) fractions analysed. F) SEC analysis of Br_FL and G) SDS-PAGE analysis of Br_FL (containing two O-linked glycosylation sites) (predicted MW 32.8kDa without glycosylation) and H) Br_E128-M324 (predicted MW 22.1kDa without glycosylation) potentially containing only one O-linked glycosylation site (G and H).

SPR analysis of potential binding interactions of Noelin-1, Noelin-3, Brorin_FL and Neuritin to AMPA receptor was then conducted. The AMPA receptor constructs GluA2_ATD, AM212 (GluA2_ATD+D1 of LBC) (details in Chapter 2) and GluA2_LBC were purified from large-scale mammalian and *E. coli* expression by methods previously described. Proteins were

concentrated to high concentrations (typically 330 μ M) for SPR analysis. Noelin-1, Noelin-3, Neuritin and full length Brorin (Br_FL) were bound to the SPR chip as ligands and analysis of binding to the various GluA2 analytes was conducted as depicted in figure 5.9A. Initial binding tests found that only GluA2_LBC weakly bound to Br_FL, so a full dilution series experiment was conducted to analyse this interaction (figure 5.9B). An estimated binding constant (K_D) was obtained but was considered to be inconclusive because equilibrium had not been reached (figure 5.9C) and the injection curves did not plateau when the analyte was injected for 60 sec (figure 5.9B). Several regeneration attempts were made, but the procedure could not be optimised which also affected overall data quality and analysis.

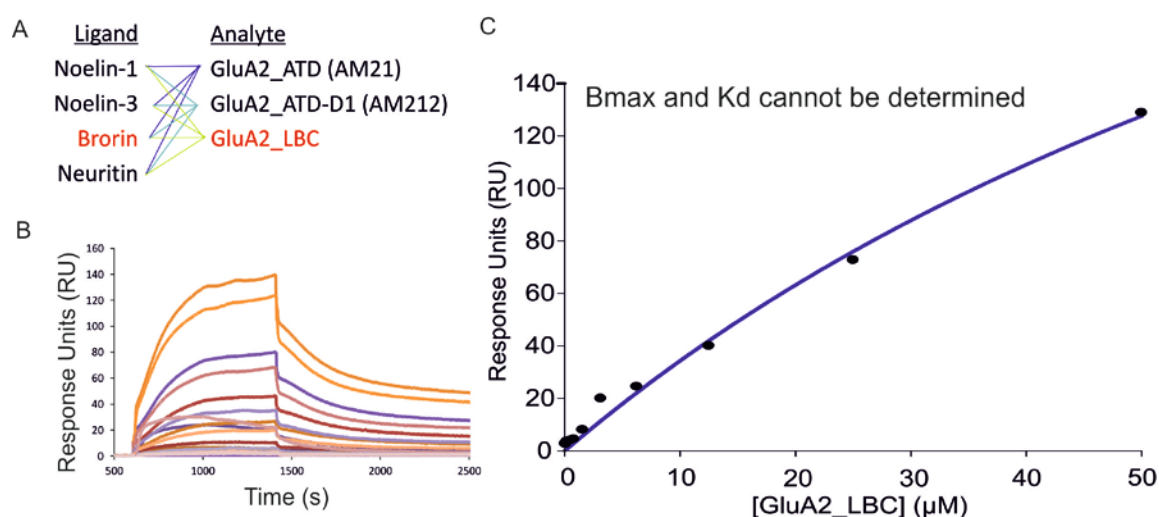


Figure 5.9 SPR analysis of putative AMPA receptor interactors.

A) Schematic representation of all constructs tested in SPR where Noelin-1, Noelin-3, Brorin and Neuritin were bound to the chip as ligands and AMPA receptor constructs GluA2_ATD, GluA2_ATD-D1 and GluA2_LBC were applied as analytes. Only Brorin with GluA2_LBC showed a potential interaction (red highlight) so was investigated further. B) Raw data curves for binding of a dilution series of GluA2_LBC (highest concentration 330 μ M) to Brorin full length (FL) protein to highlight the need for regeneration and C) SPR binding curve generated for GluA2_LBC to Brorin FL (data shown here cropped to 50 μ M) where Bmax and the binding constant (K_D) could not be accurately determined with 750 RU of Brorin bound to the chip.

5.4.6 Brorin crystallization trials

Western blot analysis of small-scale transient expression trials indicated good protein expression levels for Brorin full length (Brorin_FL) (M1-M324 including its native secretion signal) (figure 5.8A) and consequently crystallization was attempted. Brorin_FL protein expression was scaled up to 3L in HEK293S cells but only low levels of protein was purified

from media (<0.5mg/L). Large-scale expression was repeated with 2mM valproic acid (VPA) added to the media. A better yield was obtained (1mg/L), although SDS-PAGE analysis made protein levels appear deceptively high due to the smearing effect caused by O-linked glycosylation. After IMAC and SEC purification (Chapter 2 section 2.4.2), SDS-PAGE analysis of SEC peak fractions revealed several potential protein contamination bands. Therefore a second round of IMAC was carried out to produce crystallization-grade protein. Briefly this involved peak SEC fractions being pooled and bound to a 1ml Ni-NTA column, then washed with a 5-75mM gradient of imidazole. The protein was then eluted with 250mM imidazole and a second SDS-PAGE gel analysis deemed the protein to be of suitable purity for crystallization (figure 5.8G). Only limited crystallography conditions were screened (Block 1 and Block 4 (Chapter 2 table 2.4) at room temperature and 4°C) with protein at 4.8mg/ml due to limited availability. Non-detergents sulphobetaines-256 (NDSB-256) (final concentration 100mM) were also included because they reduced the propensity of the protein to precipitate at high concentrations. After several days, a high amount of phase contrast was observed in the drops, which can be indicative of a suitable protein concentration but the likelihood of crystal formation remained low due to the presence of several O-linked glycosylation sites which create flexible regions within the protein that can impede crystallization. The NDSB could also be preventing crystallization and crystals are yet to appear after more than 1 year.

The Br_short construct (D148-M324) was tested in small-scale mammalian expression trials, analysis of which found it to be expressed but not secreted (figure 5.8B). This secretion hindrance was thought to be because both O-linked glycosylation sites had been removed since removing a protein's O-linked glycosylation sites can prevent protein secretion (Strahl-Bolsinger *et al.* 1999). SDS-PAGE analysis of Brorin_FL (M1-M324) SEC (figure 5.8F and G) presented smeared multiple protein bands around 40kDa, indicating that the O-linked glycosylation sites may be occupied. Further construct design was required to generate secreted shortened Brorin protein but proved difficult because it was unclear where the N-terminal residue should be. Therefore N-terminal truncations in 20 residue steps were generated to optimise secretion with

minimal O-linked glycosylation sites (figure 5.8C, D and E).

The Brorin_E128-M324 protein (5.8F) was deduced to be more likely to crystallize because it lacked one of the two O-linked glycosylation sites and a large proportion of the N-terminal region predicted to be disordered (L33-P147) (figure 5.2). Expression of this construct was attempted in 3L of HEK293S cells with 2mM VPA added to the media and the expressed protein was purified and concentrated to 8.25mg/ml for crystallization. Small crystals grew after 15 days in 20%w/v PEG 3350, 0.2M potassium acetate, but these were unfortunately determined to be salt crystals.

5.5 Discussion

5.5.1 TARP2 analysis

An interaction between the ectodomain loop-2 of TARP2 and the LBC of GluA2 had been previously identified (Milstein and Nicoll 2009), so attempts were made to express this loop region as a peptide aptamer on the Trx scaffold (TARP2_Trx). Purification of TARP2_Trx from *E. coli* proved to be a success where a single protein species could be isolated from the refolded protein material. Co-crystallization of TARP_Trx with GluA2_LBC was then conducted but only crystals of GluA2_LBC were obtained. Different ratios of protein could be attempted to outcompete the tendency for GluA2_LBC to crystallize alone. TARP_Trx would also require MALS to determine if it was truly monomeric in solution and biophysical analysis (SPR) to determine whether the loop, whilst presented on the Trx scaffold, was still able to interact with the AMPA receptor LBC. SPR analysis could also determine if the interaction involves a larger proportion of the AMPA receptor ectodomain (detail given in 5.6.1). The TARP2 loop-2 may also be stabilised by a (currently unknown) cofactor or metal ion. Small molecule ligands that stabilise the loop could be identified using thermo-stabilisation assays (thermofluor) and might increase the likelihood of complex formation in co-crystallization attempts.

5.5.2 CKAMP44 analysis

Expression of the CKAMP44_18 ectodomain construct on the Trx scaffold protein (CKAMP44_Trx) from *E. coli* initially seemed promising but the protein could not be purified with the same success of the TARP2_Trx construct. Mammalian expression of the CKAMP44_18 construct seemed to generate a higher yield but the SEC elution volume would suggest that the protein was forming multimers or aggregations, predicted here to be caused by free cysteine residues. CKAMP44 belongs to the Shisa family of proteins, predicted to contain a cystine-knot motif required for correct protein folding (Muller *et al.* 2002). Mutagenesis of individual cysteine residues could determine whether a free cysteine residue was causing the protein to aggregate.

SEC analysis did reveal a single elution peak close to the column void volume, implying that a single species may be present but requires confirming by MALS analysis although it is likely that aggregation is occurring as a side effect of the over-expression systems utilised. SPR analysis of the CKAMP44 ectodomain construct (M42-T191) did suggest an interaction with GluA2_LBC where a binding curve could be generated. However, there were 3000 units of CKAMP44_M42-T191 bound to the chip as ligand but the $B_{\max}$ for GluA2_LBC binding was only calculated to be 307 RU. If the majority of the ligand's activity was impeded by aggregation it could explain the limited physiological activity observed.

5.5.3 Analysis of other putative interactions

The interaction between Brorin and the AMPA receptor, previously identified by pull-down assay and mass spectrometry analysis (Schwenk *et al.* 2012), was confirmed here by SPR analysis and mapped to the AMPA receptor GluA2_LBC. Further analysis of the binding curve revealed that saturation was not reached which prevented calculation of the true binding constant. In addition to this issue, once the Brorin_FL 'analyte' had bound to the GluA2_LBC 'ligand', it was difficult to detach even with extensive regeneration attempts. This affected overall data analysis where the binding constant suggested here might have been tighter than a true value. One suggestion for this interaction was that the Brorin protein has non-specific adhesive qualities, but binding to the GluA2_ATD or GluA2_ATD_D1 proteins was not observed, indicating a specific interaction with the LBC. Other methods, such as ELISA may enable the true K_D value to be calculated.

Other recently identified putative AMPA receptor ectodomain-interacting proteins (Schwenk *et al.* 2012) were analysed here by SPR. Unfortunately only one of five proteins analysed (Brorin) directly interacted with the AMPA receptor ectodomain (GluA2_LBC). What should be considered is that the SPR experiments described here only examined binding to the AMPA receptor GluA2 subunit. This therefore does not rule out a potential interaction of any of these novel partners with the other AMPA receptor subunits GluA1, 3 & 4. Also to consider is the

wider AMPA receptor interactome picture, in which auxiliary proteins (including TARPs and CKAMP44) feature prominently. These auxiliary proteins may therefore act as intermediary binding partners but further analysis will be required to determine how each protein fits into the complex interactome picture.

5.6 Future work and perspectives

5.6.1 TARP2

The TARP2_Trx_pET22b construct and its optimised protein purification protocol were sent to Ingo Greger (LMB, Cambridge) for electrophysiological analysis to examine the effects that the TARP2 ectodomain loop-2 has on AMPA receptor gating mechanisms. In Oxford, biophysical analysis will determine if the interaction between the TARP2 loop-2 (which is 77 residues long) and the AMPA receptor extends beyond the currently identified LBC region. Once the interaction site has been more precisely located, it will ensure that the tightest interacting partners are used in future co-crystallization attempts. Since initial crystallization attempts of TARP2_Trx in complex with GluA2_LBC were conducted, several new screens have been added to the repertoire of screens available and may contain the optimal conditions required for a complex to form so will also be included in any future co-crystallization attempts.

5.6.2 The AMPA receptor interactome

What has become clear whilst trying to map AMPA receptor protein-protein interactions is the difficulty to reproduce results observed by other research groups. Whilst SPR analysis can generate K_D values for true binding interactions, a pre-SPR screening stage may increase throughput when screening for novel partners. A high-throughput method to enable fast screening of potential interactions could be generated in an ELISA format. Put simply, the AMPA receptors could be coupled to a plate and incubated with constructs of interest, expressed as alkaline phosphate (AP)-fusion proteins. A simple wash stage may then remove any non-specific unbound proteins leaving positive interactions to be detected by their AP activity.

5.7 References

- Armstrong, N. and E. Gouaux (2000). "Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor: Crystal Structures of the GluR2 Ligand Binding Core." *Neuron* **28**(1): 165-181.
- Borghouts, C., C. Kunz, N. Delis and B. Groner (2008). "Monomeric recombinant peptide aptamers are required for efficient intracellular uptake and target inhibition." *Mol Cancer Res* **6**(2): 267-281.
- Brown, C. J., S. G. Dastidar, H. Y. See, D. W. Coomber, M. Ortiz-Lombardia, C. Verma and D. P. Lane (2010). "Rational design and biophysical characterization of thioredoxin-based aptamers: insights into peptide grafting." *J Mol Biol* **395**(4): 871-883.
- Dalva, M. B., M. A. Takasu, M. Z. Lin, S. M. Shaham, L. Hu, N. W. Gale and M. E. Greenberg (2000). "EphB Receptors Interact with NMDA Receptors and Regulate Excitatory Synapse Formation." *Cell* **103**(6): 945-956.
- Ehlers, M. D. (2000). "Reinsertion or degradation of AMPA receptors determined by activity-dependent endocytic sorting." *Neuron* **28**(2): 511-525.
- Fisher, J. L. and D. D. Mott (2012). "The auxiliary subunits Neto1 and Neto2 reduce voltage-dependent inhibition of recombinant kainate receptors." *Journal of Neuroscience* **32**(37): 12928-12933.
- Gu, Y., S. Huang, Michael C. Chang, P. Worley, A. Kirkwood and Elizabeth M. Quinlan (2013). "Obligatory Role for the Immediate Early Gene NARP in Critical Period Plasticity." *Neuron* **79**(2): 335-346.
- Hee Jung, C., J. Xia, R. H. Scannevin, X. Zhang and R. L. Huganir (2000). "Phosphorylation of the AMPA receptor subunit GluR2 differentially regulates its interaction with PDZ domain-containing proteins." *Journal of Neuroscience* **20**(19): 7258-7267.
- Hemler, M. E. (2005). "Tetraspanin functions and associated microdomains." *Nature Reviews Molecular Cell Biology* **6**(10): 801-811.
- Jin, W., W. P. Ge, J. Xu, M. Cao, L. Peng, W. Yung, D. Liao, S. Duan, M. Zhang and J. Xia (2006). "Lipid binding regulates synaptic targeting of PICK1, AMPA receptor trafficking, and synaptic plasticity." *Journal of Neuroscience* **26**(9): 2380-2390.
- Kuroyanagi, T. and T. Hirano (2010). "Flap loop of GluD2 binds to Cbln1 and induces presynaptic differentiation." *Biochem Biophys Res Commun* **398**(3): 537-541.
- Lin, D. T., Y. Makino, K. Sharma, T. Hayashi, R. Neve, K. Takamiya and R. L. Huganir (2009). "Regulation of AMPA receptor extrasynaptic insertion by 4.1N, phosphorylation and palmitoylation." *Nature Neuroscience* **12**(7): 879-887.
- Milstein, A. D. and R. A. Nicoll (2009). "TARP modulation of synaptic AMPA receptor trafficking and gating depends on multiple intracellular domains." *Proceedings of the National Academy of Sciences of the United States of America* **106**(27): 11348-11351.
- Muller, Y. A., C. Heiring, R. Misselwitz, K. Welfle and H. Welfle (2002). "The cystine knot promotes folding and not thermodynamic stability in vascular endothelial growth factor." *Journal of Biological Chemistry* **277**(45): 43410-43416.
- Oh, M. C., V. A. Derkach, E. S. Guire and T. R. Soderling (2006). "Extrasynaptic membrane trafficking regulated by GluR1 serine 845 phosphorylation primes AMPA receptors for long-term potentiation." *Journal of Biological Chemistry* **281**(2): 752-758.
- Rouach, N., K. Byrd, R. S. Petralia, G. M. Elias, H. Adesnik, S. Tomita, S. Karimzadegan, C. Kealey, D. S. Brecht and R. A. Nicoll (2005). "TARP γ -8 controls hippocampal AMPA receptor number, distribution and synaptic plasticity." *Nature Neuroscience* **8**(11): 1525-1533.
- Saglietti, L., C. Dequidt, K. Kamieniarz, M. C. Rousset, P. Valnegri, O. Thoumine, F. Beretta, L. Fagni, D. Choquet, C. Sala, M. Sheng and M. Passafaro (2007). "Extracellular Interactions between GluR2 and N-Cadherin in Spine Regulation." *Neuron* **54**(3): 461-477.
- Schwenk, J., N. Harmel, A. Brechet, G. Zolles, H. Berkefeld, C. S. Muller, W. Bildl, D. Baehrens, B. Huber, A. Kulik, N. Klocker, U. Schulte and B. Fakler (2012). "High-resolution

- proteomics unravel architecture and molecular diversity of native AMPA receptor complexes." Neuron **74**(4): 621-633.
- Shanks, N. F., J. N. Savas, T. Maruo, O. Cais, A. Hirao, S. Oe, A. Ghosh, Y. Noda, I. H. Greger, J. R. Yates and T. Nakagawa (2012). "Differences in AMPA and Kainate Receptor Interactomes Facilitate Identification of AMPA Receptor Auxiliary Subunit GSG1L." Cell Reports **1**(6): 590-598.
- Shepherd, J. D. and R. L. Huganir (2007). The cell biology of synaptic plasticity: AMPA receptor trafficking. **23**: 613-643.
- Steentoft, C., S. Y. Vakhrushev, H. J. Joshi, Y. Kong, M. B. Vester-Christensen, K. T. Schjoldager, K. Lavrsen, S. Dabelsteen, N. B. Pedersen, L. Marcos-Silva, R. Gupta, E. P. Bennett, U. Mandel, S. Brunak, H. H. Wandall, S. B. Levery and H. Clausen (2013). "Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology." EMBO J **32**(10): 1478-1488.
- Strahl-Bolsinger, S., M. Gentzsch and W. Tanner (1999). "Protein O-mannosylation." Biochimica et Biophysica Acta - General Subjects **1426**(2): 297-307.
- Von Engelhardt, J., V. Mack, R. Sprengel, N. Kavenstock, K. W. Li, Y. Stern-Bach, A. B. Smit, P. H. Seeburg and H. Monyer (2010). "CKAMP44: A brain-specific protein attenuating short-term synaptic plasticity in the dentate gyrus." Science **327**(5972): 1518-1522.

Chapter 6

Structural and biophysical analysis of MDGA1 and its interactions with neuroligins

6.1 Introduction

The meprin/A5 protein/ receptor protein tyrosine phosphatase mu (MAM)-domain containing glycosylphosphatidylinositol (GPI)- anchor protein 1 (MDGA1) is postulated to be a multifunctional protein with roles in cell adhesion, migration and axon guidance (Litwack *et al.* 2004) (introduced in Chapter 1). MDGA1 and its homologue MDGA2 show a high level of sequence homology across species and are both predicted to contain six immunoglobulin (Ig) domains, a fibronectin type 3 (FN3) domain and a MAM domain. Previous research has attempted to correlate the multiple MDGA1 functions to the domains it contains but a variety of outcomes have been observed. In one study, secreted AP-fused MDGA1 protein and domain deletion mutations predicted that association with axon rich regions of motor and dorsal root ganglion (DRG) neurons requires the MDGA1 MAM domain (Fujimura *et al.* 2006). AP-fused MDGA1 or MDGA1 ectodomain deletion mutations also identified that a heterophilic cell adhesion occurs between MDGA1 and an unidentified protein (present on differentiating muscle cells) and requires the MDGA1 Ig domains (Fujimura *et al.* 2006). When the adhesive properties of MDGA1 were analysed in Madin-Darby canine kidney epithelial cells (MDCK), which do not display endogenous MDGA1 expression, a heterophilic cell-cell adhesion did not appear to involve MDGA1 N-terminal regions but instead required the MDGA1 MAM domain (Diaz-Lopez *et al.* 2010). Together these results allude to MDGA1 having multiple roles dependant on where and potentially when it is expressed. Without structural information (which is currently unavailable) it is difficult to precisely map the postulated interaction sites and therefore to gain any further insights onto the protein's complex modes of action the structure of MDGA1 must be solved.

Recently a heterophilic *cis* interaction involving post-synaptically located MDGA1 and neuroligin-2 was shown to prevent the trans-synaptic adhesion properties of neuroligin-2 and neurexins at inhibitory neurons (Lee *et al.* 2013; Pettem *et al.* 2013) (figure 6.1). Over-expression of MDGA1 was shown to decrease inhibitory synapse density and MDGA1 was therefore

postulated to be the first identified negative regulator of inhibitory synapse formation (Pettem *et al.* 2013).

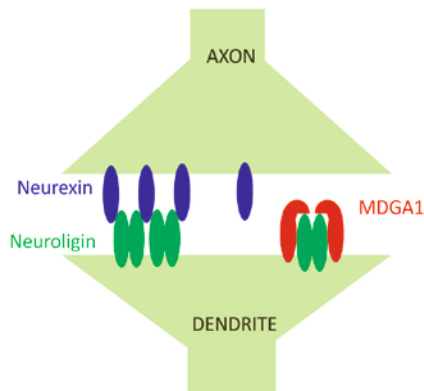


Figure 6.1 The neurexin-neuroigin interaction is inhibited by MDGA1.

Expression of post-synaptic MDGA1 interacts with post-synaptic neuroigin and acts to impede the trans-synaptic interaction of neuroigin with neurexin.

Neuroligins and neurexins are encoded for by multiple genes and alternative splicing of both proteins generates a vast number of isoforms (Chih *et al.* 2006). Rare genetic variants that affect the neuroigin-neurexin cell-cell adhesion properties have been implicated in neurodevelopmental cognitive diseases including the autism spectrum disorders (ASDs) (Südhof 2008). The presence of different neuroigin splice isoforms is known to alter their capacity to bind to neurexins (Chih *et al.* 2006). One would therefore hypothesize that different neuroigin splice isoforms may also affect the binding affinity to MDGA1 and therefore could be involved in a complex mechanism to regulate selective extracellular interactions.

6.2 Aims

Cell adhesion proteins are known to be involved in a variety of homophilic and heterophilic interactions and many such proteins contain multiple Ig and FN3 domains (Kamiguchi *et al.* 2000). Due to the predicted domain organisation one would hypothesise that MDGA1 would also be involved in cell adhesion (Diaz-Lopez *et al.* 2005). A heterophilic *cis* interaction with neuroligin-2 has been identified (Lee *et al.* 2013; Pettem *et al.* 2013). The aims of this project were to examine the properties of MDGA1 and are summarised below.

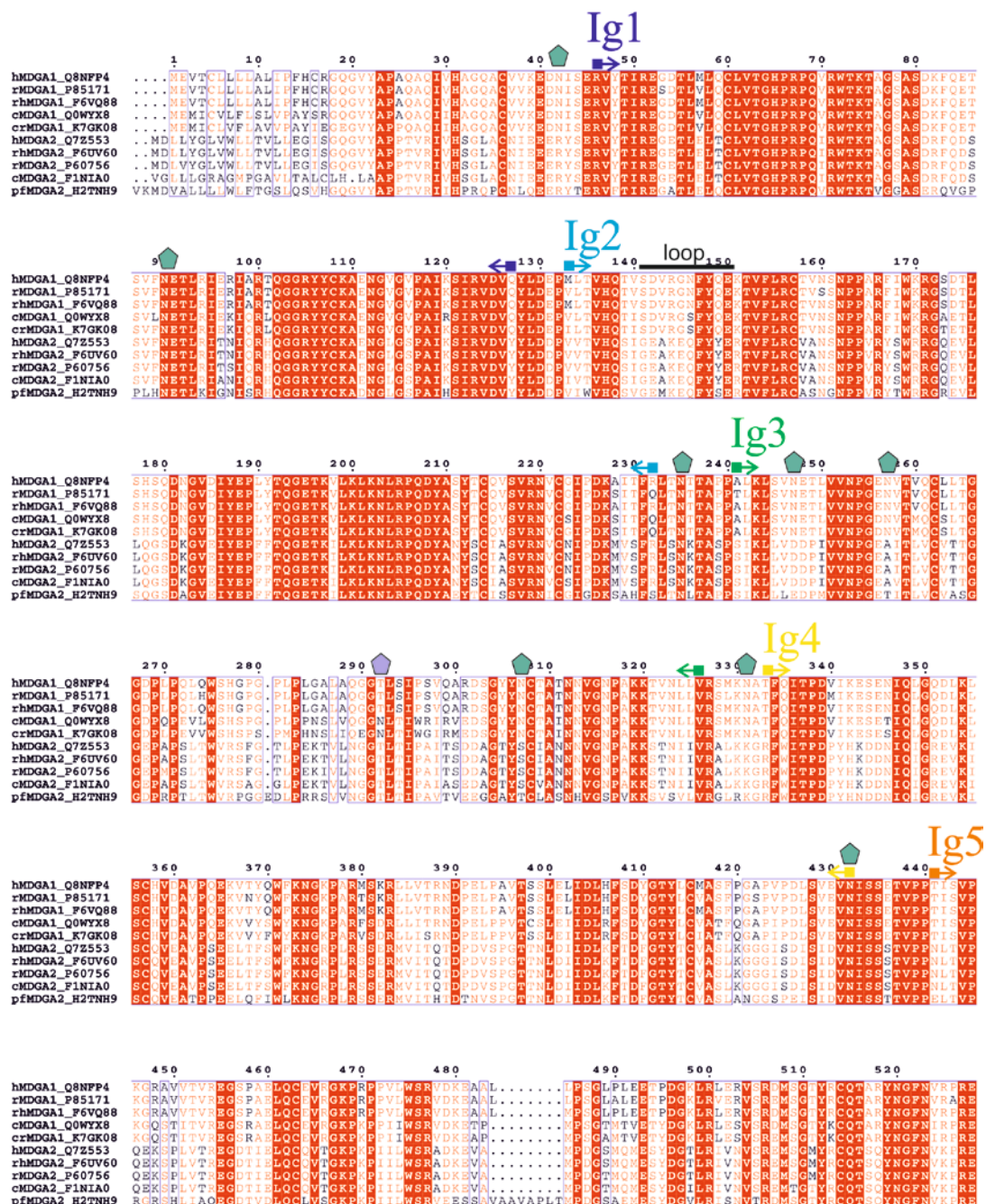
- 1) X-ray crystallography will be used to solve the structure of MDGA1.
- 2) The interactions of MDGA1 and the neuroligins 1-4 (various splice isoforms) will be characterised using biophysical analysis (SPR).
- 3) A longer term aim is to obtain a protein crystal structure of MDGA1 in complex with neuroligin-2.
- 4) The solved structure(s) will enable collaborative efforts to use structural-guided mutagenesis to target sites of interest within MDGA1 to;
 - a. Examine the functional roles of MDGA1 in chick embryogenesis (Collaborator: Dr Jo Begbie, Oxford).
 - b. Test the impact of the MDGA1-NL2 interaction in the modulation of synapse formation (Collaborator: A. M. Craig, University of British Columbia).

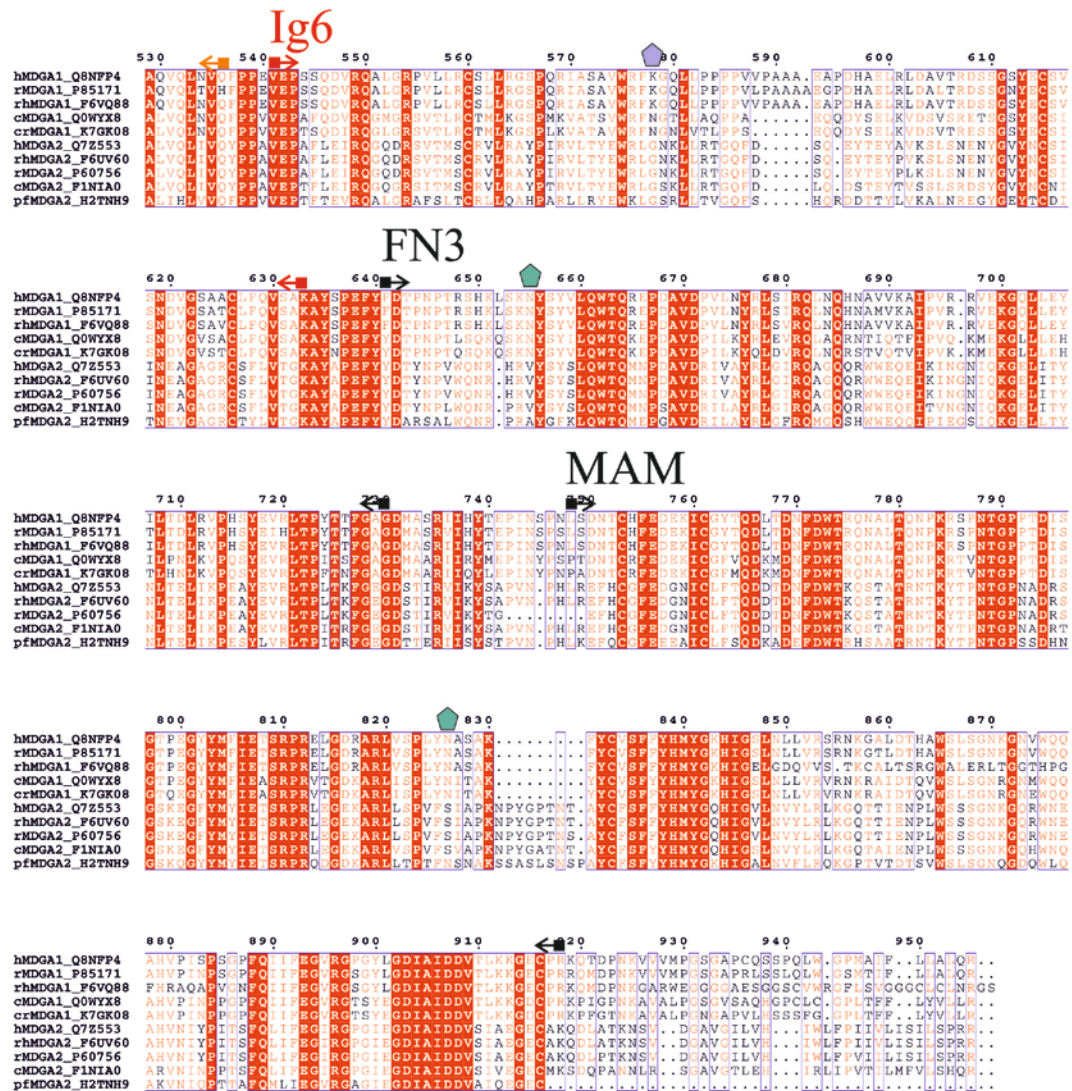
6.3 Materials and methods

6.3.1 Construct design and cloning

All oligonucleotides for PCR amplification can be found in Appendix D, table D2. Constructs are numbered according to the start of each protein's native secretion signal (summarised in Appendix D, table D1). Constructs were created using methods outlined in Chapter 2, section 2.1 and are briefly described here.

A





B

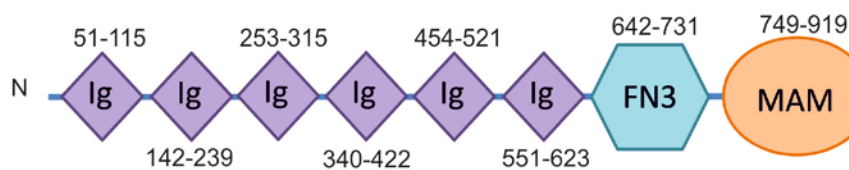


Figure 6.2 MDGA1 and MDGA2 species alignment and schematic.

A) MDGA1 sequence alignment of human (h) (*Homo sapiens*), rat (r) (*Rattus norvegicus*), chicken (c) (*Gallus gallus*), rhesus monkey (rh) (*Macaca mulatta*) and Chinese soft-shell crab (cr) (*Pelodiscus sinensis*). MDGA2 alignment of human (*Homo sapiens*), rat (*Rattus norvegicus*), chicken (*Gallus gallus*), rhesus monkey (*Macaca mulatta*) and the Chinese pufferfish (pf) (*Takifugu rubripes*). Alignment produced using Multalign (Corpet, 1988). Ig domain start and end residues are indicated above the sequence by arrows. N-linked glycosylation of MDGA1 is indicated by coloured hexagons. B) Schematic depicting the MDGA1 ectodomain construct domain organisation with 6 N-terminal immunoglobulin (Ig) domains, one fibronectin (FN3) and one C-terminal MAM domain. With domain boundaries predicted. The C-terminal GPI anchor is removed.

A chick MDGA1 ectodomain construct (Q19-N925), termed from hereon in as chMDGA1_FL (schematic figure 6.2B) was generated for GenBank accession number

NP_001038112 and cloned into the pHLsec vector by our collaborators (A. Smith, Oxford). A plasmid containing rat MDGA1 (HA-MDGA1 Kan^R) (GenBank accession number NM_001107618.1) was a gift from our collaborator Ann-Marie Craig (University of British Columbia, Canada). Based on the domain prediction (figure 6.2) a rat ectodomain construct (Q19-N925), termed rMDGA1_FL was designed and PCR amplified using oligos rMDGA1_Age_fwd and rMDGA1_BsiWI_re. A truncated construct (Q19-K330) containing a region that was predicted to encompass the first three Ig domains (rMDGA1_Ig1-3) (figure 6.2A) was generated by PCR amplification using oligos rMDGA1_Age_fwd and MDGA1Ig1-3BsiRe.

Rat neuroligin-2 (NL2) (NCBI accession number NM_053992.1) (Nlg2-FC Amp^R) was a gift from Ann-Marie Craig (University of British Columbia, Canada). The PCR amplification of the NL2 construct (residues E38-H612) was carried out using oligos NL2fwdBspE1_71 and NL2revBsiW_72 (incorporating BsrGI and BsiWI restriction sites). Amendments to the general thermocycler parameters stated in Chapter 2, table 2.2 step 3 (T_m here was 65 °C) resulted in amplification of a strong singular band of 1725bp. This was cloned into the pHLsec and pHLsec_Avi vectors via AgeI/Acc65I restriction sites.

The NL2 construct was sequenced and found to contain a splice insertion at splice site A (SSA) (figure 6.3). Because splice variants are known to dramatically affect binding of neuroligins to neurexins (Chih *et al.* 2006; Koehnke *et al.* 2010), a NL2 construct lacking the SSA (NL2-veSSA) was generated for SPR analysis. Oligos NL2-veSSA_fwd and NL2-veSSA_rev were designed to remove the splice insertion site by site-directed mutagenesis and SOE-PCR. All additional NL1, NL3 and NL4 constructs were cloned into the pHLsec and pHLsec-Avitag vectors by Kwok Wun (Oxford).

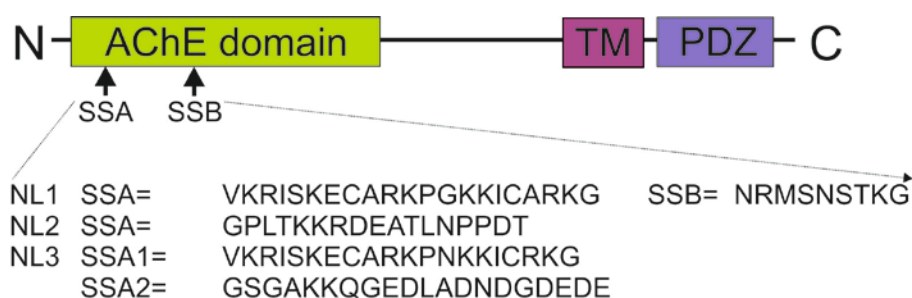


Figure 6.3 Neurologin domain organisation.

Neurologin domain prediction (HHpred) and splice isoforms present in neurologin 1, 2 and 3 (NL1, NL2 and NL3). Neurologins share a common domain organisation including an N-terminal acetylcholinesterase-like domain, in which both splice insertions are located (splice site A, SSA and splice site B, SSB). Neurologins also contain a transmembrane domain (S678-Y700) and an intracellular PZD domain. The larger SSA of NL3 is split into two insertions where SSA1 or SSA2 may be present.

6.3.2 Protein expression and purification

Western blot analysis (Chapter 2, section 2.3.2) was used to confirm the expression and secretion of all proteins from small-scale transient expression trails (Chapter 2, section 2.2.1). Secreted protein candidates were expressed in large-scale from HEK293T and 293S cells (3L) (Chapter 2, section 2.2.2) and purified as stated in Chapter 2 section 2.4.1-2.4.4, using IMAC and SEC. SDS-PAGE analysis of peak SEC fractions was conducted to determine protein purity (Chapter 2, section 2.3.1). Where deglycosylation was required, EndoF1 was added to concentrated protein and then incubated at room temperature for 10 mins immediately prior to crystallization attempts (amended from Chapter 2, section 2.4.4). Selenomethionine-incorporated cMDGA1 was produced as described in Chapter 2, section 2.2.3. Where proteins were to be co-expressed, the DNA was added in a 1:1 molar ratio. Proteins used in SPR and SAXS analysis were pooled and concentrated in their glycosylated form. Purified glycosylated chMDGA1_FLAG protein was also used by our collaborators for chick embryo work (Jo Begbie/Lexy Smith DPAG).

6.3.3 Protein crystallization, data collection and processing

A full crystallization screen was conducted for chMDGA1_FL in both glycosylated and deglycosylated states at 5mg/ml. Many crystals grew but they were fragile and/or small and so

required optimisation. The crystals diffracting to the highest resolution grew in an optimization of a PEGRx condition (Chapter 2, table 2.4) where the precipitant reservoir contained 0.1M HEPES pH7.5, 4% w/v polyethylene glycol 8,000. Once a dataset had been collected and it was established that molecular replacement models could not be used to solve the structure, selenomethionine-(SeMet) labelled chMDGA1_FL crystallization was also attempted. A full screen was conducted but the initial SeMet chMDGA1_FL (deglyc) crystallization screens produced only flaky-looking, multiple, thin crystals. Due to the nature of data collection required for a complete SeMet data set, these crystals required optimisation. Native chMDGA1_FL (deglyc) protein crystals had grown in a PEGRx condition (described above) so an optimisation of this condition with SeMet_chMDGA1_FL (deglyc) was established but failed to produce any crystals.

A seeding experiment was therefore attempted to improve the probability that any crystals that appeared were the same space group as the native dataset. A crystal of native chMDGA_FL (deglyc) was fragmented as previously described (Walter *et al.* 2007) and a ¼ serial dilution series of a seed solution was established through rows A-G of a 96 well crystallization plate (figure 6.4). Row H, functioned as a control, i.e. was without a crystal seed. The seed dilution series was screened against a 4% gradient of the original precipitant through columns 1-12 and chMDGA1_FL (SeMet deglyc) protein at 5mg/ml was applied to each drop (figure 6.4).

Several of the seeding-optimised SeMet chMDGA1_FL (deglyc) crystals were tested at the Diamond Light Source I03 beamline. Initial analysis of crystal quality came from four test images collected 45-90 degrees apart. Auto-processing software was then used to determine the spacegroup and produce a data collection strategy. SeMet chMDGA1_FL (deglyc) crystal produced diffraction spots to 3.4 Ångström (Å).

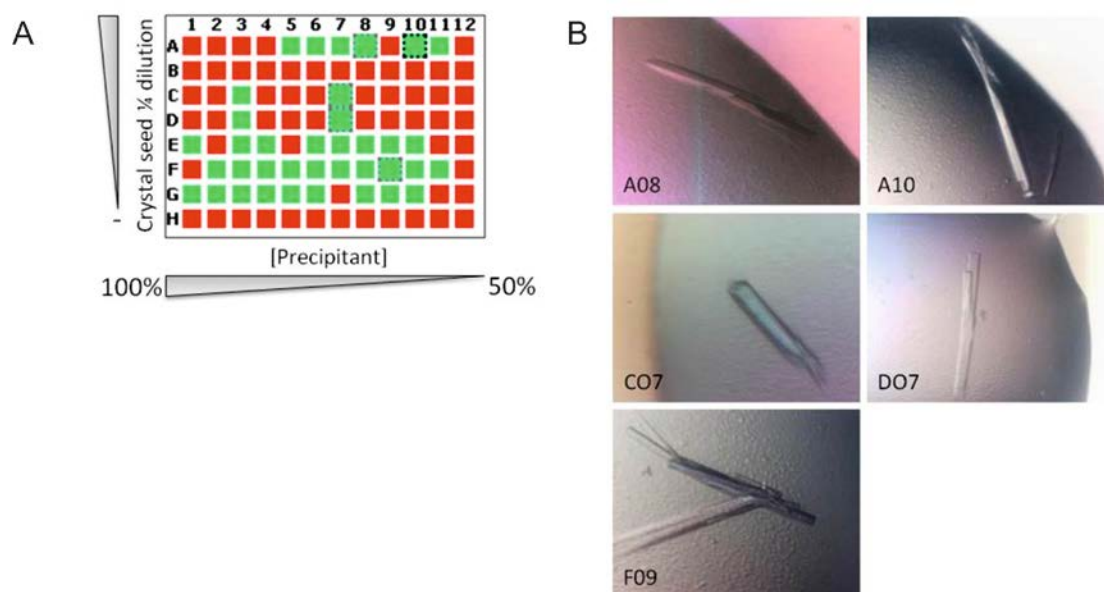


Figure 6.4 MDGA1 crystallization optimisation.

Selenomethionine (SeMet) chMDGA1 (deglycosylated) protein crystallization seeding experiment. A) Crystal hits indicated in green in a 1/4 seed dilution (well row A-G, H = seed free control) screened against precipitant 100%-50% (well column 1-12) 0.1M HEPES pH7.5, 4% w/v PEG 8000. B) Selection of optimised crystals obtained from wells highlighted in green with black dotted boxes in A.

The cMDGA1 structure was determined by single anomalous dispersion (SAD) analysis and the positions of selenium atoms were determined by SHELXD (Schneider *et al.* 2002). This solution was input into autoSHARP (Vonrhein *et al.* 2007) for phase calculation using the high-resolution native data (3.4 Å resolution). ARP/wARP (Perrakis *et al.* 1999) was used for initial model building but the majority of the peptide sequence had to be manually hand built in COOT (Emsley *et al.* 2004). Refinement was conducted in Phenix (Adams *et al.* 2002).

6.3.4 SPR analysis

SPR was conducted as described in Chapter 2, section 2.8.2 with the following amendments. After ligands were coupled to the chip, the running buffer was switched from the standard buffer to 10mM Tris pH7.4, 150mM NaCl, 2mM CaCl₂, 1g/L BSA and 0.005% Tween-20. The buffer flow speed was then increased to 20µl/min. SPR analysis was conducted with chMDGA1_FL and rMDGA1_FL purified protein as analyte sample and various neuroligins as ligands to calculate binding constants K_D for each interaction.

6.3.5 SAXS

Small angle X-ray scattering (SAXS) was used to determine the oligomeric state of MDGA1 in solution in their glycosylated states. Data collection and analysis was conducted by J. Elegheert, (STRUBI). SAXS data were collected at the BM29 beamline at the ESRF, Grenoble (Pernot *et al.* 2013). Measurements were carried out at 293 K, within a momentum transfer range of $0.01 \text{ \AA}^{-1} < s < 0.6 \text{ \AA}^{-1}$, where $s = 4\pi\sin(\theta)/\lambda$ and 2θ is the scattering angle. rMDGA1 was measured at concentrations of 0.85, 1.70, 3.20, 5.00 and 6.80 g/L in 10mM Tris HCl pH7.5, 150mM NaCl, 3mM CaCl₂, 0.005% Tween-20. The rMDGA1-NL2 complex was measured at concentrations of 1.00, 3.00 and 5.00 g/l with identical buffer composition. The cMDGA1 was measured at concentrations of 1.50 and 3.36 g/L in 10mM HEPES pH7.5, 150mM NaCl. No measurable radiation damage was detected on comparison of ten successive time frames 10 sec. exposures. Initial data processing was performed using automated procedures implemented in the ATSAS software suite (Petoukhov *et al.* 2012). Buffer subtraction was performed with the program PRIMUS (Konarev *et al.* 2003) in accordance with standard procedures. Merged curves were generated using the low angle part of the lower concentration curves and the high angle part of the higher concentration curves. The radius of gyration r_g and forward scattering I_0 were determined using PRIMUS through evaluation of the Guinier region. Maximum particle dimension D_{MAX} and the distance distribution function $P(r)$ were calculated using GNOM (Svergun 1992). Molecular masses of different constructs were calculated by comparison with the reference BSA sample, by evaluation of the Porod volume, and by using the volume of correlation invariant V_c (Rambo *et al.* 2013). Theoretical scattering patterns from available structures and generated models, together with their respective discrepancy to the experimental data, were calculated using CRY SOL (Svergun *et al.* 1995; Petoukhov *et al.* 2012).

6.4 Results

6.4.1 Structural analysis of MDGA1

Expression and secretion of chMDGA1_FL and rMDGA1_FL indicates that a single band of ~120kDa is present and correlated to a protein mass prediction of 108kDa plus 12 predicted N-linked glycosylation sites (example western blot analysis for chMDGA1_FL in figure 6.5A). A singular symmetrical peak was observed for both chick and rat proteins on SEC following their IMAC purification (example for chMDGA1_FL in figure 6.5B). SDS-PAGE analysis of the peak identified a potential contaminant or cleavage product so only selected fractions (identified by i in figure 6.5B) were pooled.

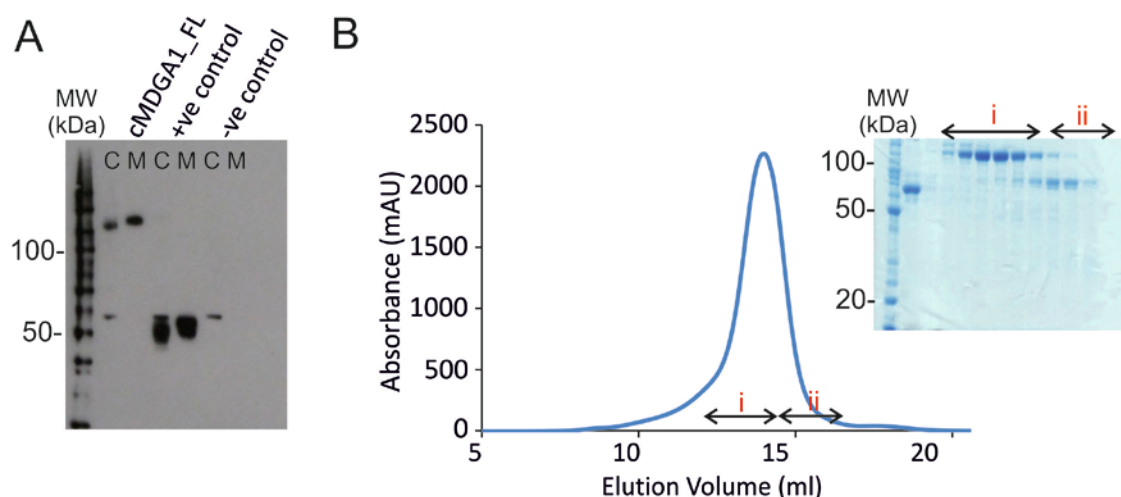


Figure 6.5 MDGA1 purification.

A) Western blot analysis of chMDGA1_FL with transfection positive and negative controls. Total expression (cellular fraction, C) and secreted fraction (media M) were analysed. B) SEC elution profile and SDS-PAGE analysis of chMDGA1_FL after IMAC purification. An S200 10/30 SEC column was run in 10mM HEPES pH7.5, 150mM NaCl. Peak SEC fractions were analysed on 12% SDS-PAGE gels with the following gel loading order: Benchmark (Invitrogen) protein molecular weight marker (MW), IMAC flow through, 5mM imidazole IMAC wash followed by peak SEC fractions as indicated on SEC elution profile by Roman numerals i and ii.

Crystallization screens were established for chMDGA1 and rMDGA1 protein. Initial crystallisation attempts required optimisation where a seeding experiment successfully increased the size (figure 6.4) and diffraction resolution (figure 6.6A) of the previously obtained crystals.

The chMDGA1_FL contains 16 methionine residues whereas the rMDGA1_FL contains only 7. The chMDGA1 crystal was confirmed to contain selenium atoms by a fluorescence energy scan (figure 6.6C) so the beam energy was shifted accordingly to the peak of 12659.5eV. The best crystal diffracted to 3.4Å (figure 6.6A) and Ig domains were clearly evident in the SAD phased 2Fo-Fc electron density map (contoured at 1.5 σ) prior to refinement (figure 6.6B).

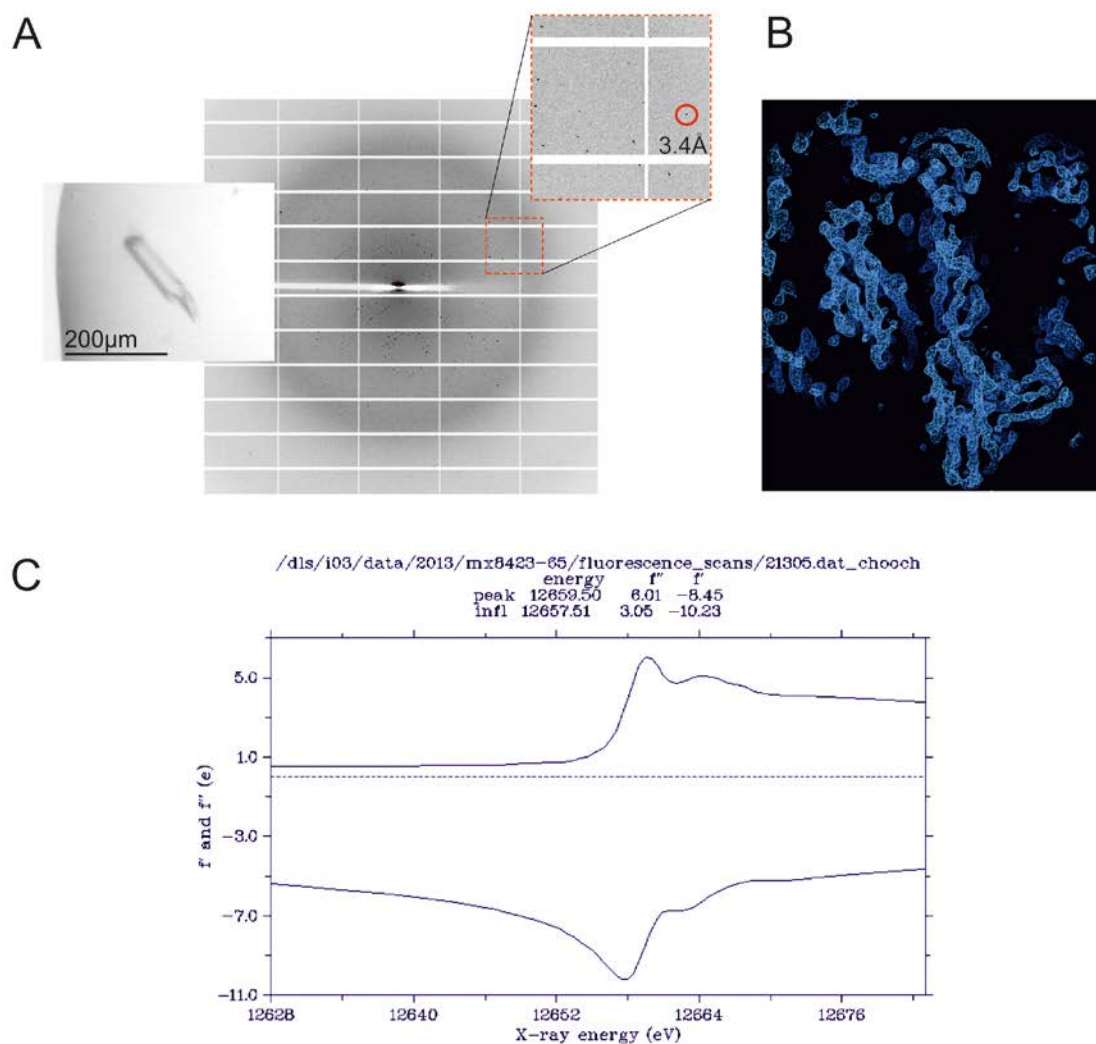


Figure 6.6 MDGA1 data collection.

A) MDGA1 selenomethionine-incorporated (SeMet) crystal and X-ray diffraction to 3.4Å collected at Diamond beamline IO3. B) Unbiased experimental SeMet SAD-phased electron density (contoured at 1.5 σ) with Ig-like domain features (Beta protein strands and sheets) clearly visible in electron density map. C) A fluorescence energy scan was performed to determine that anomalous-scattering selenium atoms were present within the crystal.

Xia2 (Winter 2010) automatically performed data reduction, merging structure factor amplitudes and processed the data to spacegroup $P2_12_12_1$ with unit cell dimensions 103.12, 109.09, 209.08 a b c (Å), 90, 90, 90 $\alpha\beta\gamma$ (°). Matthew's coefficient analysis (Matthews 1968;

Kantardjieff *et al.* 2003) estimated that for a 108kDa protein (excluding any glycosylation) with these unit cell dimensions, there was a 95% probability that there were two molecules per asymmetric unit (ASU) with a 54% solvent content.

The solved structure indicates that the six immunoglobulin (Ig) domains are organised in a novel arrangement (Figure 6.7 and 6.8 D) hence why molecular replacement models using Ig domain-containing proteins DSCAM (PDB ID: 3DMK, Axonin (PDB ID: 1CS6) and DCC (PDB ID: 1LAF) failed (figure 6.8A-C). These cell-adhesion protein models all contain Ig domains in a ‘horse-shoe’ arrangement. However in the MDGA1 structure the Ig domains are folded around in a triangular arrangement with two Ig domains forming each side of the triangle. Angles of approximately 60° are observed between domains Ig2 and Ig3 and Ig4 and Ig5. The MAM domain is not present in the model, and is probably disordered due to its flexibility within the crystal. Model building is currently on-going, and explains why only half of the FN3 domain has been presented here. Further model building may also enable the location of the MAM domain to be identified (if it has not been cleaved). Structural analysis reveals that there are two molecules locked together in a *cis*-arranged dimer with a twisted double ‘tubular bells’ conformation (figure 6.7). Interestingly a loop region of the Ig2 domain from each molecule extends to form a contact with a second protein dimer in a head to tail arrangement (figure 6.8E). Whether this architecture is physiological or an artefact of crystallization remains to be determined. The loop region incorporates residues SDVRGSFYQE (residues 141-150 in chicken) and sequence alignment (Figure 6.2) indicates that this region is specific to MDGA1 (60-70% different in MDGA2 dependent upon species) with divergence in MDGA1 at one residue (S146) in chicken (*Gallus gallus*) and Chinese soft-shell crab (*Pelodiscus sinensis*) where an asparagine residue is observed in other species.

Structural refinement of this 949 residue (per monomer) structure is currently on-going and is indicated by crystallographic statistics (Table 6.1). Once refinement is complete it will

reveal more sites of interest to target for structural-guided mutagenesis required for future functional analysis.

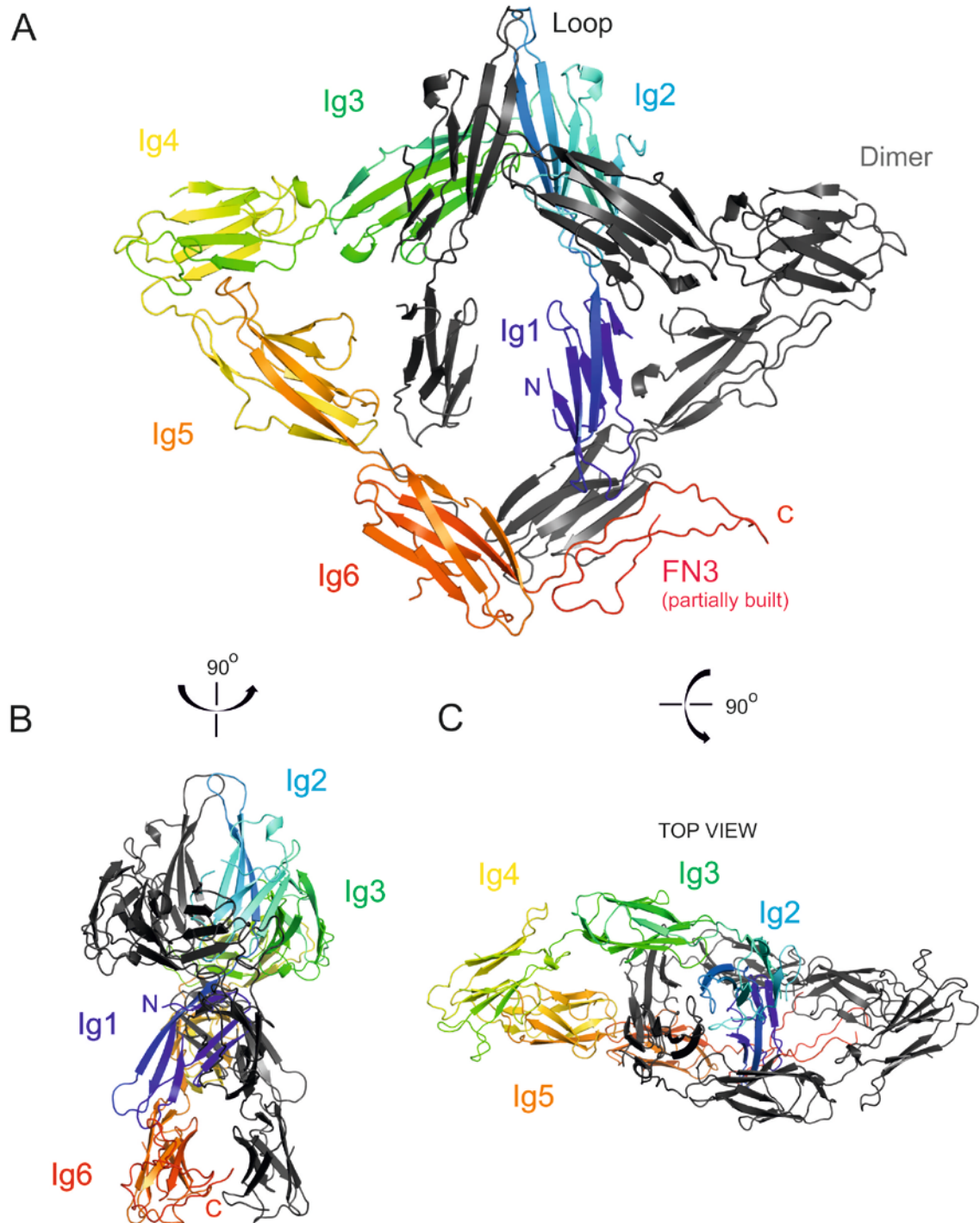


Figure 6.7 MDGA1 Structural analysis.

Domain organisation of MDGA1 in cartoon rainbow format (chain A) for six Ig domains and a fibronectin domain that have been coloured to match domain labelling in figure 6.2A. Chain B is coloured in dark grey. The position of the MAM domain is not yet determined. A) Protein face view, B) 90 degree rotation of (A) around y-axis to show the protein side view and C) 90 degree rotation of (A) around x-axis to show the top view of the crystallographic dimer.

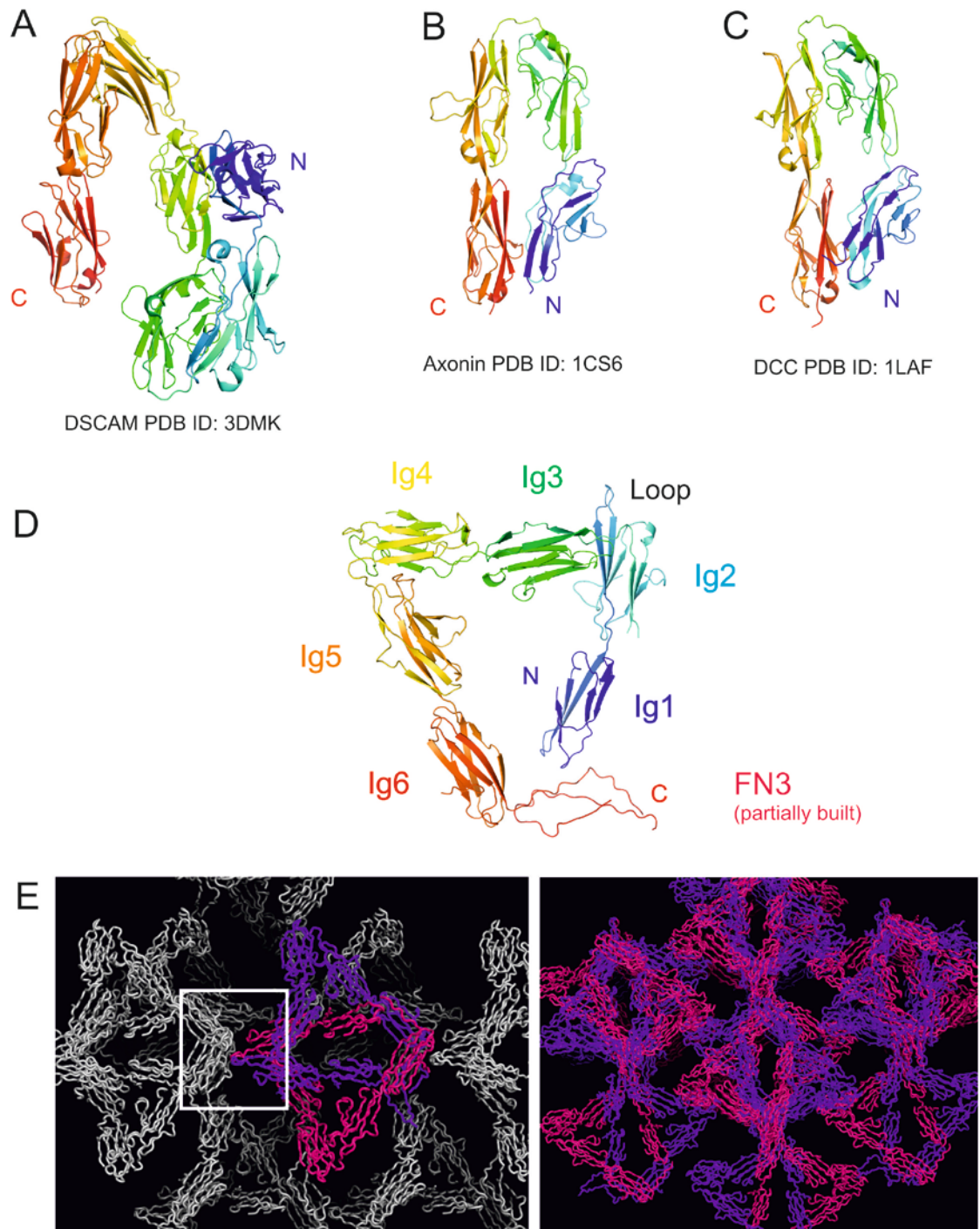


Figure 6.8 Further MDGA1 structural analysis.

Molecular replacement models A) DSCAM (PDB ID: 3DMK), B) Axonin (PDB ID: 1CS6) and C) DCC (PDB ID: 1LAF) represented in cartoon rainbow format which all have a 'horse-shoe' arrangement of four Ig domains and all failed to solve the MDGA1 structure. D) The structure of MDGA1 indicates a previously unseen arrangement of Ig domains. E) Crystallographic packing represented in cartoon tube format indicates head to tail pattern of MDGA1 in a crystallographic dimer (pink and purple chains) where the extended loops of the Ig2 and Ig3 dimer makes contacts with Ig6 and the partially built FN3 domain of another molecule (in white boxed region highlighted).

Table 6.1 Crystallographic data collection and structure refinement statistics.

	MDGA1_SeMet
Data collection	
Beamline	Diamond I03
Phi width	0.10°
Images	1-7200
Space group	P2 ₁ 2 ₁ 2 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	103.13, 109.10, 209.11
α , β , γ (°)	90, 90, 90
Wavelength (Å)	1.26595
Resolution (Å)	109.1 -3.1
	(3.18-3.1)
<i>R</i> _{merge}	0.184 (2.382)
<i>R</i> _{pim} (<i>I</i>)	0.038 (0.485)
<i>I</i> / σ <i>I</i>	17.3 (1.7)
Completeness (%)	100 (99.9)
Redundancy	26.5 (25.8)
Refinement	In progress

- Numbers in brackets refer to the highest resolution shell. *R*_{pim} = Precision-indicating merging R-value

6.4.2 Mapping the Neuroligin- MDGA1 interaction

Results have previously shown that MDGA1 binds the neuroligins (NLs) where analysis by cell surface binding assays suggested that the interaction between MDGA1-NL2 was in the nanoMolar (nM) range where affinities of 145nM (Lee *et al.* 2013) and 7.3nM (Pettem *et al.* 2013) were previously observed. To validate this nM affinity, SPR experiments were conducted as previously described (Chapter 2, section 2.8.2) with amendments as described in section 6.3.4. Results are summarised below in figure 6.9.

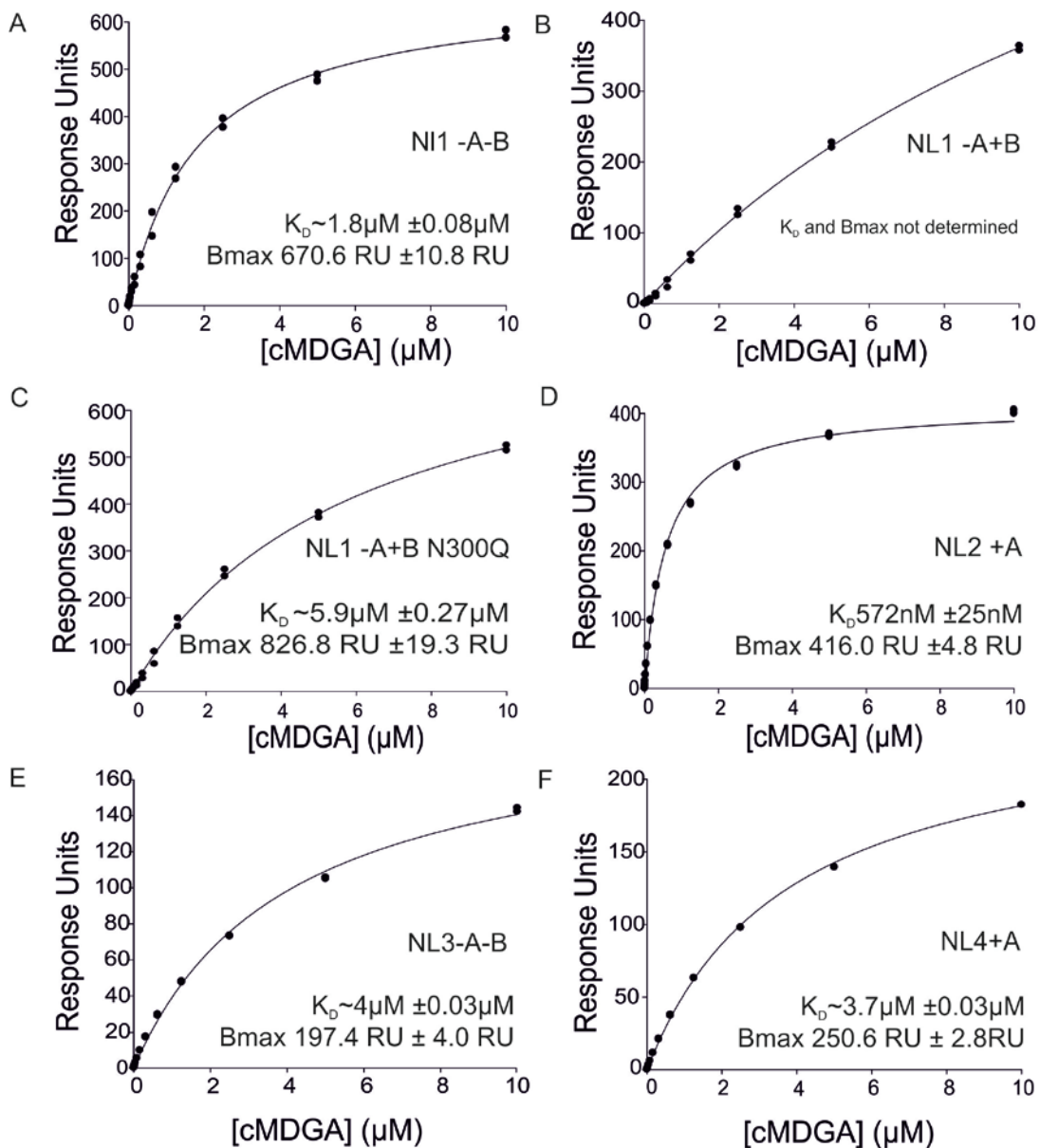


Figure 6.9 SPR analysis mapping the chick MDGA1-Neuroligin interaction.

Analysis of chMDGA1 binding to A) NL1 lacking splice inserts at sites A & B (-A-B), B) NL1 -A+B, C) NL1 -A+B N300Q, D) NL2 +A with 483 RU of ligand bound, E) NL3 -A-B with 458 RU of ligand bound and F) NL4 +A with 406 RU of ligand bound. SPR was performed on an T-200 Biacore machine.

SPR analysis reveals an interaction of MDGA1 with all neuroligins, where the strongest interaction of 572nM is observed for MDGA1 binding to NL2. Here SPR analysis also demonstrates that the interaction of the NL1 with MDGA1 can be dependent upon the splice isoform present (figure 6.9A & B). When the NL1 splice insertion at splice site B is not present, binding between MDGA1 and NL1 is enhanced 10 fold (figure 6.8A). Removal of an N-linked glycan present in NL1s splice insertion site B (previously shown to cause a steric hindrance to the interaction with neurexin (construct NL1 –A +B N300Q, created by J. Elegheert, Oxford)) indicated that binding is comparable to if the insertion at SSB was not present at all (figure 6.9A& C).

It was previously stated that binding of NL2 to MDGA1 was in the nM range, where assays were conducted with NL2 containing the insertion at SSA (Pettem *et al.* 2013). To determine whether removal of the splice insertion at SSA could increase the binding affinity for MDGA1 in the same way that removal of the splice insertion at SSB had in the NL1 interaction with MDGA1, a NL2 –SSA isoform was produced by site-directed mutagenesis. Initial SPR analysis revealed no difference in the binding affinity between the two isoforms (figure 6.10 A & B).

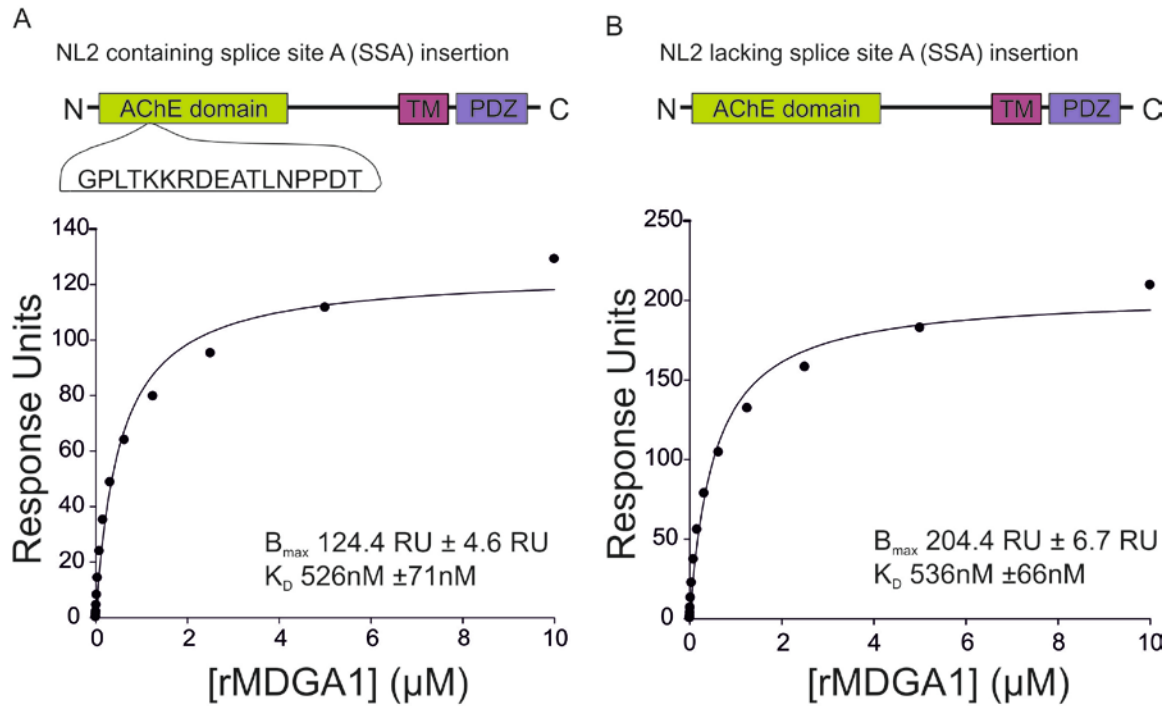


Figure 6.10 SPR analysis of rMDGA1 binding to NL2 with or without its insertion at splice site A (SSA).

A) analysis of rMDGA1 binding to NL2 +SSA where 275units of ligand were immobilised on the chip. B) Analysis of rMDGA1 binding to NL2 -SSA where 380units of ligand were immobilised on the chip. Analysis was conducted on a T-100 sensitivity enhanced (T-200) Biacore machine. Running buffer 10mM Tris pH7.5, 150mM NaCl, 2mM CaCl₂, 1g/L BSA, 0.005% w/v Tween-20 at 20 μ l/min flow speed. A ratMDGA1 (rMDGA1) 1:2 dilution series was established from 10 μ M (highest molar concentration).

6.4.3 MDGA1- NL2 complex crystallization attempts

SPR analysis suggested that the binding affinity of MDGA1 to NL2 is not affected by the presence or absence of the NL2 splice insertion at SSA. Complex crystallization was therefore attempted with rMDGA1 and both forms of NL2 (+/- SSA). The proteins were co-expressed from large-scale transient transfection of HEK293S cells. Initial SEC and SDS-PAGE analysis revealed different elution profiles (figure 6.11). MDGA1 in complex with NL2 +SSA presented one large peak, close to the column void volume and a small excess of NL2 +SSA (figure 6.11A) where both proteins within the larger peak appear to be present in an equal ratio. When MDGA1 was co-expressed with NL2 lacking the splice insertion at site A (-veSSA) a doublet peak was observed on SEC (figure 6.11B). This was considered to be caused by the different resolving ability of the columns used where an s200 1660 column was used in figure 6.11A and only a small S200 1030 column was used in figure 6.11B.

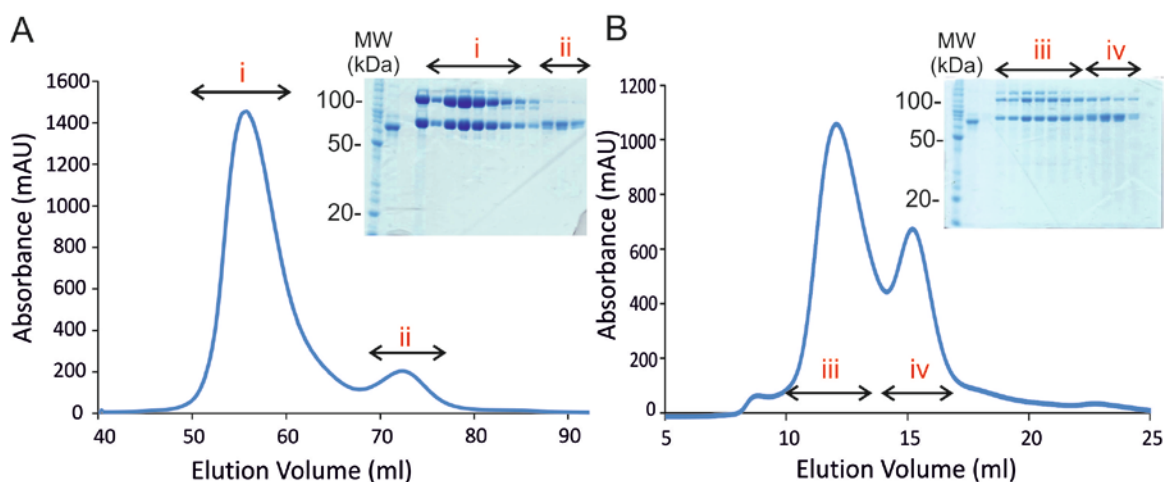


Figure 6.11 MDGA1-Neuroigin 2 complex formation.

A) Purification of rMDGA1_FL and NL2 (+SSA) on S200 16/60 column in 10mM HEPES pH7.5, 150mM NaCl. B) Purification of rMDGA1_FL and NL2 (-ve SSA) using S200 10/30 SEC columns in 10mM HEPES pH7.5, 150mM NaCl. Both co-expression attempts were expressed in HEK293S cells. Peak SEC fractions were analysed on 12% SDS-PAGE gels. Gel load order: Benchmark (Invitrogen) protein molecular weight marker (MW), IMAC flow-through, 5mM imidazole wash then peak SEC fractions identified by red Roman numerals i-iv.

Crystallization was attempted for both complexes from peaks i and iii (Figure 6.11A & B) and have yielded small needle crystals which form after 10 days at room temperature. These

crystals were tested at the Diamond Light Source, Harwell and diffracted to $>15\text{\AA}$, therefore further optimisation will be required.

6.4.4 *Small-angle X-ray scattering (SAXS) analysis*

SAXs analysis was performed to determine the oligomeric state of rMDGA1 alone and in complex with neuroligin 2 (+SSA). Preliminary results indicate that MDGA1 is monomeric in solution (132kDa) and has a maximum diameter of 17nm (figure 6.12A). This is similar to a predicted diameter based on the MDGA1 structure where $130\text{\AA}$ (13nm) was measured (figure 6.12B). When rMDGA1 is in complex with NL2, a mass of 344kDa is observed which indicates that a complex is present in a 2:2 ratio (figure 6.12A) which would be expected as the neuroligins are known to form dimers in solution (Südhof 2008).

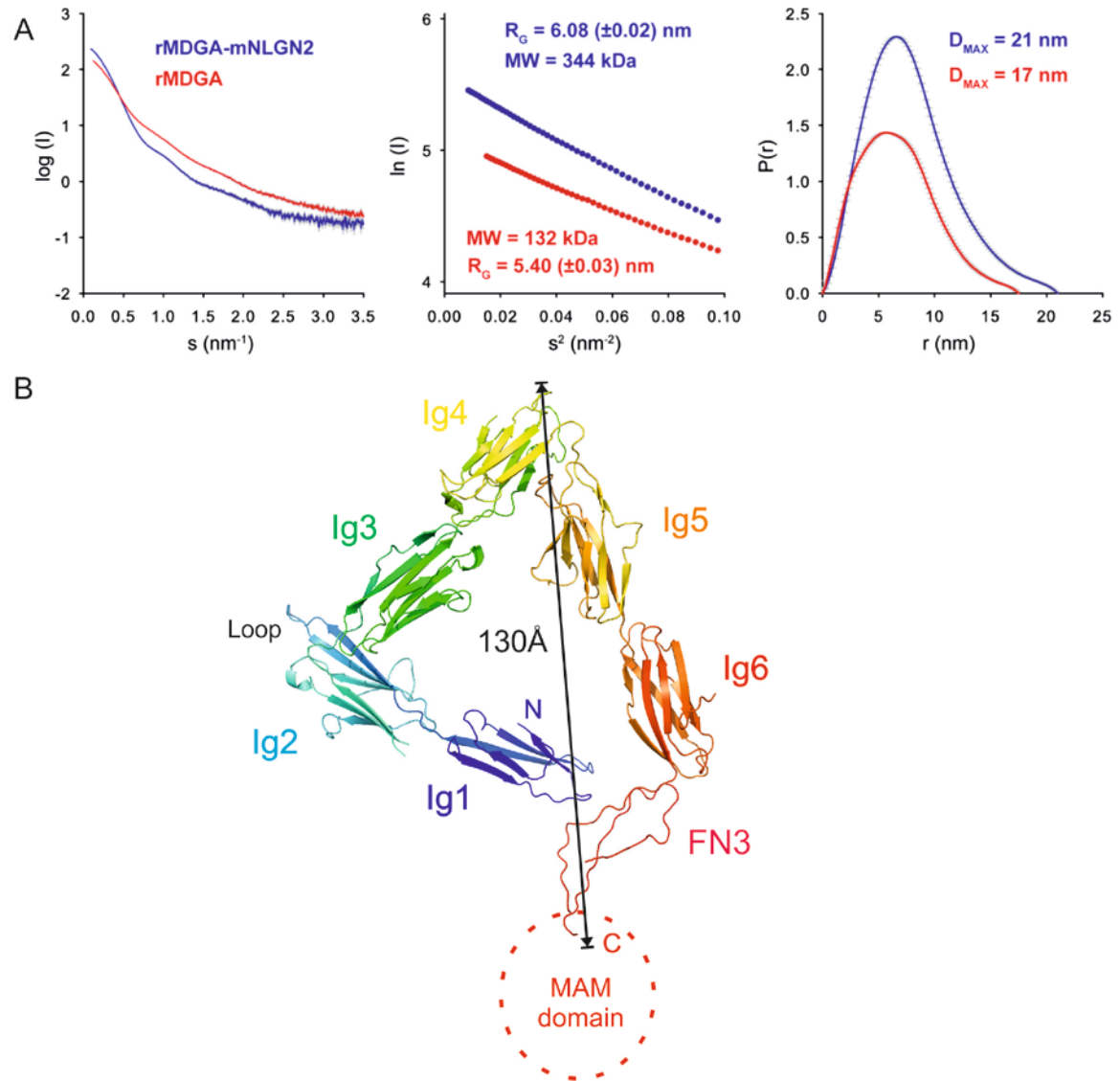


Figure 6.12 SAXS analysis of MDGA1 and MDGA1-NL2 complex.

A) SAXS analysis of MDGA1 (red trace) and MDGA1 in complex with NL2 (+SSA) (blue trace) B) MDGA1 monomer represented in cartoon rainbow format with distance measured across the molecule estimated to be $130\text{\AA}$ (using PyMol measure function).

6.5 Discussion

The first aim of the chapter was to structurally characterise MDGA1 and here the structure was solved to 3.4Å by selenomethionine SAD phasing. The structure revealed an unusual double ‘tubular bells’ conformation, never before observed in a multi-Ig domain-containing protein. An additional point of interest is the fact that many cell-adhesion proteins belonging to the Ig superfamily (IgSF) contain specialised intracellular domains required for downstream processes (Brümmendorf *et al.* 2001). However the MDGA1 protein is GPI anchored and for it to have any downstream effects they must occur via protein-protein interactions with its extracellular region. Further investigation is therefore required to determine why MDGA1 is structurally dissimilar to other IgSF protein members.

The SPR analysis conducted here confirmed that a direct interaction occurs between MDGA1 and all neuroligins to high μM and nM affinities. NL1 protein lacking the splice insertion at SSB bound with higher affinity than the protein containing an insertion at SSB, thus one would predict that the neuroligin binding site for MDGA1 involves the splice site B region. When the glycan moiety in the splice site B was removed by mutation, it was able to restore binding to levels obtained for the protein lacking an insertion at SSB. This implies that the insertion at splice site B infers a steric hindrance to the interaction with MDGA1. The presence of a neuroligin 1 splice insertion at site B has also been shown to reduce the binding affinity to β -neurexin (Koehnke *et al.* 2010). NL2 does not contain a splice insertion at SSB and inclusion or removal of the splice insertion at SSA did not affect binding which would infer that this site is not involved in the interaction with MDGA1.

MDGA1 crystallized as a dimer, but appears to be monomeric in solution (as determined by SAXS analysis). The interaction between MDGA1 and NL2 is able to withstand SEC purification and appears to be in complex in a 2:2 ratio. The neuroligins form dimers and are not trafficked to the cell surface if in a monomeric form (Poulopoulos *et al.* 2012). Two putative modes of binding are therefore proposed, the first involves a dimer of MDGA1 interacting head-

to-head in a 'scrum model' with the neuroligin dimer (figure 4.13B). The second model involves the MDGA1 molecules interacting with the neuroligin dimer to form a *cis* array on the post-synaptic membrane both models acting to inhibit the neuroligin- neurexin interaction. Also to consider is a possible MDGA1 structural rearrangement which may not be crystallisable alone. Optimisation of MDGA1- neuroligin2 complex crystallization trials will be required to confirm or disprove these models.

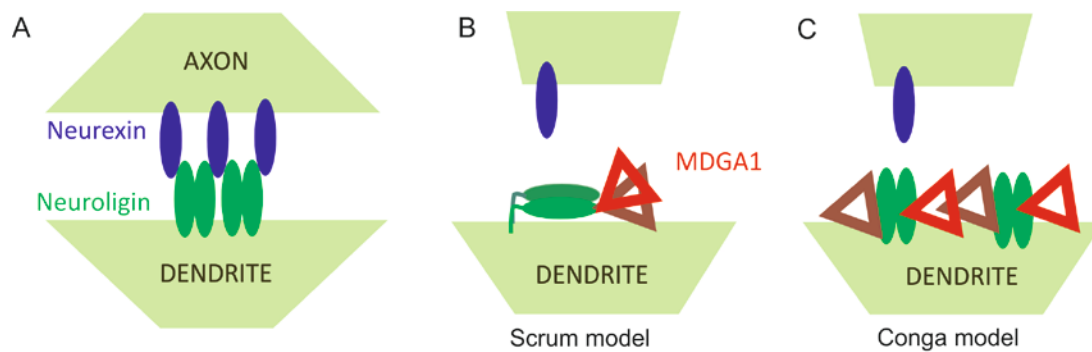


Figure 6.13 Proposed models for the MDGA1-neuroligin 2 interaction.

A) The trans-synaptic interaction of neuroligins with neurexins. B) The 'scrum model' where two MDGA1 molecules bind to each neuroligin dimer in a head-to-head arrangement. C) the 'Conga model' where an array of MDGA and neuroligins form on the post-synaptic membrane.

6.6 Future work/ perspectives

Due to time constraints, only a limited number of crystallization conditions for the rMDGA1 and NL2 (+/- splice site A) complex have been screened and so far only thin, needle like crystals have grown. Promisingly these crystals were deduced to be of protein origin although they only diffract to a low resolution ($>15\text{\AA}$). With more time, these crystals will be improved through optimisation and may enable a novel complex to be solved. This would shed light on how the MDGA1-NL2 interaction is able to out-compete binding of neuroligin to neurexin (Pettem *et al.* 2013).

Further investigation is also required to investigate whether the protein is a monomer or dimer in a cellular context and whether either of the models proposed here (the scrum or the conga) occur. All constructs that have been expressed here were membrane truncated forms and therefore soluble, although the full length protein is GPI-anchored to the post synaptic membrane. Förster resonance energy transfer (FRET) analysis could therefore be used to investigate the cell surface attached full length MDGA1 and determine whether dimers form.

6.7 References

- Adams, P. D., R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger (2002). "PHENIX: building new software for automated crystallographic structure determination." Acta Crystallographica Section D **58**(11): 1948-1954.
- Brümmendorf, T. and V. Lemmon (2001). "Immunoglobulin superfamily receptors: cis-interactions, intracellular adapters and alternative splicing regulate adhesion." Current Opinion in Cell Biology **13**(5): 611-618.
- Chih, B., L. Gollan and P. Scheiffele (2006). "Alternative Splicing Controls Selective Trans-Synaptic Interactions of the Neuroligin-Neurexin Complex." Neuron **51**(2): 171-178.
- Diaz-Lopez, A., P. Iniesta, A. Moran, P. Ortega, T. Fernandez-Marcelo, A. Sanchez-Pernaute, A. J. Torres, M. Benito and C. De Juan (2010). "Expression of Human MDGA1 Increases Cell Motility and Cell-Cell Adhesion and Reduces Adhesion to Extracellular Matrix Proteins in MDCK Cells." Cancer Microenviron **4**(1): 23-32.
- Diaz-Lopez, A., C. Rivas, P. Iniesta, A. Moran, C. Garcia-Aranda, D. Megias, A. Sanchez-Pernaute, A. Torres, E. Diaz-Rubio, M. Benito and C. De Juan (2005). "Characterization of MDGA1, a novel human glycosylphosphatidylinositol-anchored protein localized in lipid rafts." Exp Cell Res **307**(1): 91-99.
- Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallographica Section D **60**(12 Part 1): 2126-2132.
- Fujimura, Y., M. Iwashita, F. Matsuzaki and T. Yamamoto (2006). "MDGA1, an IgSF molecule containing a MAM domain, heterophilically associates with axon- and muscle-associated binding partners through distinct structural domains." Brain Research **1101**(1): 12-19.
- Kamiguchi, H. and V. Lemmon (2000). "IgCAMs: bidirectional signals underlying neurite growth." Current Opinion in Cell Biology **12**(5): 598-605.
- Kantardjieff, K. A. and B. Rupp (2003). "Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals." Protein Science **12**(9): 1865-1871.
- Koehnke, J., P. S. Katsamba, G. Ahlsen, F. Bahna, J. Vendome, B. Honig, L. Shapiro and X. Jin (2010). "Splice Form Dependence of β -Neurexin/Neuroligin Binding Interactions." Neuron **67**(1): 61-74.
- Konarev, P. V., V. V. Volkov, A. V. Sokolova, M. H. J. Koch and D. I. Svergun (2003). "PRIMUS: a Windows PC-based system for small-angle scattering data analysis." Journal of Applied Crystallography **36**: 1277-1282.
- Lee, K., Y. Kim, S.-J. Lee, Y. Qiang, D. Lee, H. W. Lee, H. Kim, H. S. Je, T. C. Südhof and J. Ko (2013). "MDGAs interact selectively with neuroligin-2 but not other neuroligins to regulate inhibitory synapse development." Proceedings of the National Academy of Sciences **110**(1): 336-341.
- Litwack, E. D., R. Babey, R. Buser, M. Gesemann and D. D. M. O'Leary (2004). "Identification and characterization of two novel brain-derived immunoglobulin superfamily members with a unique structural organization." Molecular and Cellular Neuroscience **25**(2): 263-274.
- Matthews, B. W. (1968). "Solvent content of protein crystals." Journal of Molecular Biology **33**(2): 491-497.
- Pernot, P., A. Round, R. Barrett, A. D. Antolinos, A. Gobbo, E. Gordon, J. Huet, J. Kieffer, M. Lentini, M. Mattenet, C. Morawe, C. Mueller-Dieckmann, S. Ohlsson, W. Schmid, J. Surr, P. Theveneau, L. Zerrad and S. McSweeney (2013). "Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution." Journal of Synchrotron Radiation **20**: 660-664.
- Perrakis, A., R. Morris and V. S. Lamzin (1999). "Automated protein model building combined with iterative structure refinement." Nature Structural Biology **6**(5): 458-463.
- Petoukhov, M. V., D. Franke, A. V. Shkumatov, G. Tria, A. G. Kikhney, M. Gajda, C. Gorba, H. D. T. Mertens, P. V. Konarev and D. I. Svergun (2012). "New developments in the

- ATSAS program package for small-angle scattering data analysis." Journal of Applied Crystallography **45**: 342-350.
- Pettem, K. L., D. Yokomaku, H. Takahashi, Y. Ge and A. M. Craig (2013). "Interaction between autism-linked MDGAs and neuroligins suppresses inhibitory synapse development." The Journal of Cell Biology **200**(3): 321-336.
- Poulopoulos, A., T. Soykan, L. P. Tuffy, M. Hammer, F. Varoqueaux and N. Brose (2012). "Homodimerization and isoform-specific heterodimerization of neuroligins." Biochemical Journal **446**(2): 321-330.
- Rambo, R. P. and J. A. Tainer (2013). "Accurate assessment of mass, models and resolution by small-angle scattering." Nature **496**(7446): 477-481.
- Schneider, T. R. and G. M. Sheldrick (2002). "Substructure solution with SHELXD." Acta Crystallographica Section D: Biological Crystallography **58**(10 I): 1772-1779.
- Südhof, T. C. (2008). "Neuroligins and neuroligins link synaptic function to cognitive disease." Nature **455**(7215): 903-911.
- Svergun, D., C. Barberato and M. H. J. Koch (1995). "CRY SOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates." Journal of Applied Crystallography **28**: 768-773.
- Svergun, D. I. (1992). "Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria." Journal of Applied Crystallography **25**: 495-503.
- Vonrhein, C., E. Blanc, P. Roversi and G. Bricogne (2007). Automated structure solution with autoSHARP. **364**: 215-230.
- Walter, T. S., E. J. Mancini, J. Kadlec, S. C. Graham, R. Assenberg, J. Ren, S. Sainsbury, R. J. Owens, D. I. Stuart, J. M. Grimes and K. Harlos (2007). "Semi-automated microseeding of nanolitre crystallization experiments." Acta Crystallographica Section F: Structural Biology and Crystallization Communications **64**(1): 14-18.
- Winter, G. (2010). "xia2: an expert system for macromolecular crystallography data reduction." Journal of Applied Crystallography **43**(1): 186-190.

Chapter 7

Further Discussion

7.1 Summary of findings

In 2009, only structures of the AMPA receptor ligand binding core (LBC) had been solved using proteins expressed in bacteria and refolded (GluA2_LBC (Armstrong *et al.* 2000), GluA3_LBC (Ahmed *et al.* 2009) and GluA4_LBC (Gill *et al.* 2008)). The last four years have witnessed huge progress in the structural characterisation of glutamate receptors, for which eukaryotic expression was essential. The GluA2_ATD structure solved here (Clayton *et al.* 2009) revealed a dimeric clam-shell fold with extensive dimer interfaces but exhibited flexibility in its lower D2 lobes. The second crystal form identified a putative ligand binding site with potential novel allosteric regulatory mechanisms. Structures for all AMPA receptor ATDs have since been solved (GluA1_ATD (Yao *et al.* 2011), GluA3_ATD (Sukumaran *et al.* 2011) and GluA4_ATD (Dutta *et al.* 2012) but the ground-breaking crystal structure of a full length GluA2 homotetramer (Sobolevsky *et al.* 2009) was a crowning moment for structural analysis of the AMPA receptor. Since this breakthrough, AMPA receptor complexes have emerged as a field of interest in their own right. The view of exactly what constitutes a physiological AMPA receptor is rapidly evolving, as the tetrameric core forming the actual ion channel seems merely to represent the central component of a larger supra-molecular assembly. The AMPA receptor's protein-protein interaction map is constantly expanding and we are still a long way from fully understanding all of its synaptic interactions at a molecular level.

This thesis aimed to expand the current knowledge on AMPA receptors and their multiple protein interactions, with an overall aim to deconstruct synaptic complexity. Structural and functional insights into the neuronal pentraxins and how they interact with the AMPA receptors presented many more research avenues to explore. For instance the neuronal pentraxins are either activity dependant or constitutively expressed but it is unclear why this is the case and what effect this has on AMPA receptors. Here all neuronal pentraxins were shown to weakly associate with AMPA receptor ATDs (although the GluA1_ATD remains to be investigated). From a structural point of view, why does the GluA4_ATD bind with highest affinity to the neuronal pentraxin 1 PTX domain? The PTX domains were shown here to share a high level of structural and sequence

homology, so a detailed analysis will be required to examine this. Further co-crystallization attempts of AMPA receptor ATDs with neuronal pentraxin PTX domains, as well as biophysical analysis and functional experiments will hopefully soon solve this conundrum.

This thesis also presented a crystal structure of MDGA1 in a novel dimeric tubular bells arrangement that provided a tantalising glimpse into this vast protein's potential mechanisms of action. Interestingly, results obtained here suggest that its roles are perhaps not limited solely to inhibitory synaptic interactions with neuroligin 2 (NL2) as previously proposed (Pettem *et al.* 2013). Biophysical analysis demonstrated that MDGA1 interacts with all NL 1, 2, 3 and 4, with NL1 expressed only at glutamatergic and not γ -aminobutyric acid (GABA)ergic synapses (Song *et al.* 1999). Due to the interactions with both NL1 and NL2, MDGA1 may therefore be a key regulator of both excitatory and inhibitory synapse formation and function. This hypothesis presents exciting research opportunities for further structural and functional investigations to understand the physiological implications of the MDGA1–neuroligin interaction.

7.2 Physiological insights gained from this research

A more in depth analysis of the proteins examined here finds a common theme emerging, that is, how each protein is involved in the post-synaptic clustering of neurotransmitter receptors. Physiologically, synaptic plasticity (Bliss *et al.* 1973) in terms of the activity dependant morphological changes and subsequent strengthening of the synapse, ultimately result from neuroreceptor clustering. Two methods of receptor clustering seem to be occurring, which show diversity in the proteins involved, dependent on the synapse type/ location in question. Neurotransmitter receptor clustering may result from one of two mechanisms:-

- 1) Cytoplasmic scaffold-independent mechanisms
- 2) Cytoplasmic scaffold-dependent mechanisms

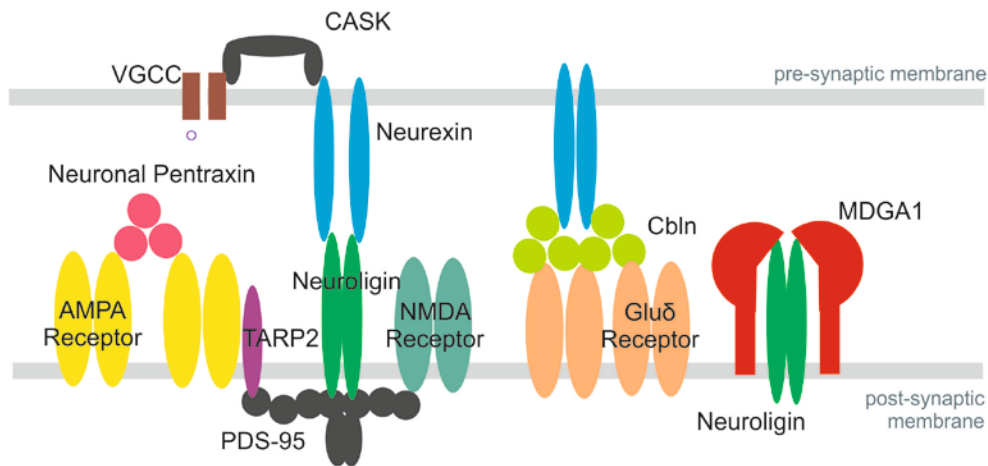
At excitatory synapses, cytoplasmic scaffold-independent mechanisms result from extracellular protein-protein interactions that have evolved to cluster the iGluRs (figure 1.7A).

Neuronal pentraxins have been shown to induce clustering of the AMPA receptors in the hippocampus (O'Brien *et al.* 1999) and GluD (Delta, δ) receptors are clustered by cerebellin (Cbln) complexed with neurexin in cerebellar parallel-fiber (PF)-Purkinje neurons (Kuroyanagi *et al.* 2010). Scaffold-dependant mechanisms of neurotransmitter clustering involve cell adhesion proteins and in some cases iGluR auxiliary proteins (eg TARPs) (figure 7.1A). The presynaptic adhesion protein neurexin is intracellularly anchored by its PDZ domain to a scaffolding complex containing CASK (Missler *et al.* 1998) and its trans-synaptic partner neuroligin binds to PSD-95 which is a major components of the glutamatergic scaffold (Craig *et al.* 2007). The neuroligin-neurexin interaction has been shown to regulate excitatory synaptic activity where clustering of NLs results in recruitment of scaffolding protein as well as NMDA iGluRs (figure 7.1A) (Chih *et al.* 2005).

At inhibitory synapses, accumulation of the neuroreceptor GABA_A subtype is dependent on its ability to interact with the post-synaptic scaffolding protein gephyrin (Tretter *et al.* 2008). Neuroligin 2 (NL2) aggregates gephyrin at inhibitory synapses and NL2 knock out reduces gephyrin accumulation and GABA_A receptor clustering. (Tyagarajan *et al.* 2014). NL2 is therefore in effect, acting as the foundation on which to build the cytoplasmic protein platform which clusters neuroreceptors at GABAergic synapses (Graf *et al.* 2004; Squire 2008) (figure 7.1B).

Inhibitory and excitatory cytoplasmic scaffolds are dynamic assemblies, held together and regulated by multiple protein-protein interactions, of which only a small handful of interaction partners are discussed here. Questions remain as to how the neuroligin -MDGA1 interaction reduces synapse formation but one could hypothesise that it prevents gephyrin (Pettem *et al.* 2013) and PSD-95 accumulation. Therefore synaptic adhesion proteins such as neuroligins and neurexins not only have roles in a trans-synaptic adhesion but also generate bi-directional, pre-and post-synaptic structural rearrangements that strengthen the synapse.

A EXCITATORY SYNAPSE



B INHIBITORY SYNAPSE

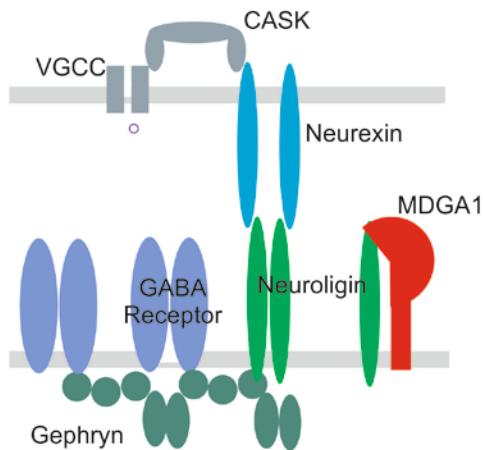


Figure 7.1 Neuroreceptor clustering at excitatory and inhibitory synapses.

A) AMPA (GluA) receptors are clustered by neuronal pentraxins and by via their auxiliary protein TARP2 to PSD-95. Neuroligins (NLs) also bind to PSD-95 and act as trans-synaptic cell adhesion molecules interacting with neurexins. NL clustering also recruitments scaffolding protein PSD-95 and NMDA iGluRs. Delta (GluD) receptors are clustered by neurexins via cerebellin (cbln). MDGA1 prevents trans-synaptic and intracellular neuroligin interactions.

B) Neurotransmitter receptor clustering at inhibitory synapses via scaffold dependant mechanisms. The interaction of NL2 with neurexin induces gephyrin mediated GABA_A receptor clustering and the interaction of MDGA1 with neuroligin inhibits this clustering.

The autism spectrum disorders (ASDs) are postulated to arise from disruptions in the overall synaptogenic excitatory-inhibitory (E/I) balance (Bourgeron 2009). A NL3 R451C mutation was identified in a boy with X-linked Asperger syndrome and in his brother with X-linked autism (Jamain *et al.* 2003). Initial analysis suggested that this mutation prevents NL3s dissociation from its chaperone prior to dimerization, thus impeding correct protein trafficking (De Jaco *et al.* 2006). However, when this mutation was introduced into a knock-in (KI) mouse model, a gain of function and increased inhibitory synaptic transmission was observed. Like patients with ASD, the R451C mice exhibited impaired social behaviour but enhanced spatial learning abilities not observed in the NL3 KO mouse (Tabuchi *et al.* 2007). In this instance it was noted that only a partial retention of NL3 was actually occurring (90% reduction to total surface

expression), with the remaining protein postulated to cause the gain of function effect. Because there are no structures of NL3 currently available, this mutation can be mapped to a NL3 homologue, NL4 (pdb ID: 2XB6 (Leone *et al.* 2010)) (figure 7.2). Analysis reveals that the mutation is located near the neuroligin dimer interface (figure 7.2). Several scenarios which could cause the gain of function effect are therefore proposed here:-

- 1) The mutation could result in an altered protein conformation. In this instance solving the structure of the NL3 R451C mutant could confirm or disprove this.
- 2) The mutation may result in a stabilisation to the dimer interface enhancing the likelihood of a neuroligin-neurexin interaction. Biophysical analysis could confirm or disprove this.
- 3) The mutation could affect the NL3 interaction with MDGA. An MDGA- NL complex structure would therefore provide a huge amount of information to resolve this problem.

This figure has been removed due to
copyright restrictions.

Figure 7.2 Structural analysis of the gain of function mutation R451.

The mutation was mapped to human neuroligin-3 residue R450 based on sequence information from Tabuchi *et al.* (2007). The mutation was then aligned to the solved structure of neuroligin-4 in complex with neurexin-1 (2XB6) (Leone *et al.* 2010).

7.3 Structural analysis to aid biophysical and functional experiments

Structural analysis of the neuronal pentraxin PTX domains found a high level of structural fold homology. Only a small loop region close to the calcium binding site in a calcium free NP2PTX crystal form presented an alternative conformation. Interestingly the ATD- PTX interaction was found to require the presence of calcium. If this loop is implicated in the PTX-ATD interaction, then any sequence differences found in the loop could account for the variations in K_D observed in SPR analysis. Further investigation reveals that within the loop sequence a charged to uncharged residue variation is present (Arg to Gly) where NP2 and NPR loops are more positively charged than the NP1 loop (figure 7.3). The loops of NP2 and NPR could therefore be mutated to the corresponding NP1 loop sequence, where SPR analysis could corroborate or contradict the hypothesis that this loop dictates binding affinity.



Figure 7.3 Sequence alignment of human neuronal pentraxin NP1, NP2 and NPR.

Sequences were aligned using MULTALIN (<http://multalin.toulouse.inra.fr/multalin/>). Conserved calcium co-ordinating residues are highlighted by blue pentagons. Beta sheets are labelled K and L based on nomenclature stated for the C-reactive protein structure (Shrive *et al*, 1996). The loop region exhibiting flexibility in the crystal structure is highlighted in the red box. Structural analysis of loop region of NP1_PTX (orange), NP2_PTX (green) and NPR_PTX (blue).

In addition to structural analysis of the PTX domain, electron microscopy provided information as to the oligomeric state of NP2_CC2 where a trimeric arrangement was observed. Stability of oligomeric protein was a major issue where rapid degradation often occurred. To overcome this and generate stable protein for use in functional experiments, artificial trimeric PTX domains were developed.

The MDGA1 structure presented a novel, dimeric tubular bells arrangement of Ig domains which was interesting for several reasons. Structural analysis found that an Ig 2 loop (SDVRGSFYQE residues 141-150 in chicken) extended to form a contact with a second protein dimer in a head to tail arrangement (figure 6.8E). The loop region is specific to MDGA1 and is 60-70% different to MDGA2 (dependent upon species). This Ig domain also appeared

fundamental to protein dimerization within the crystal structure, although biophysical analysis of native MDGA1 protein found it to be monomeric in solution. Structure-guided mutagenesis will help to determine whether this architecture is physiological or an artefact of crystallization.

7.4 Structural analysis providing insights for rational drug design

Structural analysis is a useful tool for rational drug design because it enables sites of interest to be specifically targeted by visualising potential drugs in complex with target proteins (Williams *et al.* 2005). Fine tuning of drug candidates can improve molecular characteristics to enhance drug target specificity and increase potency whilst reducing side effects.

During synapse formation and maturation inhibitory and excitatory synaptic connections are tightly regulated in an activity dependant manner (Dean *et al.* 2006). Atypical changes in synaptic strength can therefore dramatically affect the fine balance of inhibitory and excitatory circuitry, and such events are implicated in ASDs (Bourgeron 2009). Pettem *et al.* (2013) eloquently demonstrated that neuronal cells form fewer connections with cells expressing NL2 and MDGA1 than cells expressing NL2 alone, so in a physiological context MDGA1 acts to reduce net inhibitory drive (Lee *et al.* 2013; Pettem *et al.* 2013). Potential drugs to inhibit the NL-MDGA1 interaction whilst maintaining the NL-neurexin interaction could therefore be a potential target for ASD treatment. It would be difficult to disrupt a specific site without complex structural information. A NL2-MDGA1 complex structure will be therefore enable small molecules to be targeted to the MDGA1-NL interaction interface to aid in understanding of underlying causes of ASDs in which MDGA1, NLs and neurexins are implicated.

7.5 References

- Ahmed, A. H., Q. Wang, H. Sondermann and R. E. Oswald (2009). "Structure of the S1S2 glutamate binding domain of GLuR3." *Proteins* **75**(3): 628-637.
- Armstrong, N. and E. Gouaux (2000). "Mechanisms for Activation and Antagonism of an AMPA-Sensitive Glutamate Receptor: Crystal Structures of the GluR2 Ligand Binding Core." *Neuron* **28**(1): 165-181.
- Bliss, T. V. and T. Lomo (1973). "Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path." *J Physiol* **232**(2): 331-356.
- Bourgeron, T. (2009). "A synaptic trek to autism." *Current Opinion in Neurobiology* **19**(2): 231-234.
- Chih, B., H. Engelman and P. Scheiffele (2005). "Control of Excitatory and Inhibitory Synapse Formation by Neuroligins." *Science* **307**(5713): 1324-1328.
- Clayton, A., C. Siebold, R. J. C. Gilbert, G. C. Sutton, K. Harlos, R. A. J. McIlhinney, E. Y. Jones and A. R. Aricescu (2009). "Crystal Structure of the GluR2 Amino-Terminal Domain Provides Insights into the Architecture and Assembly of Ionotropic Glutamate Receptors." *Journal of Molecular Biology* **392**(5): 1125-1132.
- Craig, A. M. and Y. Kang (2007). "Neurexin-neuroligin signaling in synapse development." *Current Opinion in Neurobiology* **17**(1): 43-52.
- De Jaco, A., D. Comoletti, Z. Kovarik, G. Gaietta, Z. Radic, O. Lockridge, M. H. Ellisman and P. Taylor (2006). "A mutation linked with autism reveals a common mechanism of endoplasmic reticulum retention for the alpha,beta-hydrolase fold protein family." *J Biol Chem* **281**(14): 9667-9676.
- Dean, C. and T. Dresbach (2006). "Neuroligins and neurexins: linking cell adhesion, synapse formation and cognitive function." *Trends Neurosci* **29**(1): 21-29.
- Dutta, A., I. H. Shrivastava, M. Sukumaran, I. H. Greger and I. Bahar (2012). "Comparative dynamics of NMDA- and AMPA-glutamate receptor N-terminal domains." *Structure* **20**(11): 1838-1849.
- Gill, A., A. Birdsey-Benson, B. L. Jones, L. P. Henderson and D. R. Madden (2008). "Correlating AMPA receptor activation and cleft closure across subunits: crystal structures of the GluR4 ligand-binding domain in complex with full and partial agonists." *Biochemistry* **47**(52): 13831-13841.
- Graf, E. R., X. Zhang, S.-X. Jin, M. W. Linhoff and A. M. Craig (2004). "Neurexins Induce Differentiation of GABA and Glutamate Postsynaptic Specializations via Neuroligins." *Cell* **119**(7): 1013-1026.
- Jamain, S., H. Quach, C. Betancur, M. Råstam, C. Colineaux, I. C. Gillberg, H. Soderstrom, B. Giros, M. Leboyer, C. Gillberg, T. Bourgeron and null (2003). "Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism." *Nature Genetics* **34**(1): 27-29.
- Kuroyanagi, T. and T. Hirano (2010). "Flap loop of GluD2 binds to Cbln1 and induces presynaptic differentiation." *Biochem Biophys Res Commun* **398**(3): 537-541.
- Lee, K., Y. Kim, S.-J. Lee, Y. Qiang, D. Lee, H. W. Lee, H. Kim, H. S. Je, T. C. Südhof and J. Ko (2013). "MDGAs interact selectively with neuroligin-2 but not other neuroligins to regulate inhibitory synapse development." *Proceedings of the National Academy of Sciences* **110**(1): 336-341.
- Leone, P., D. Comoletti, G. Ferracci, S. Conrod, S. U. Garcia, P. Taylor, Y. Bourne and P. Marchot (2010). "Structural insights into the exquisite selectivity of neurexin/neuroligin synaptic interactions." *EMBO J* **29**(14): 2461-2471.
- Missler, M., R. Fernandez-Chacon and T. C. Südhof (1998). "The making of neurexins." *Journal of Neurochemistry* **71**(4): 1339-1347.
- O'Brien, R. J., D. Xu, R. S. Petralia, O. Steward, R. L. Huganir and P. Worley (1999). "Synaptic clustering of AMPA receptors by the extracellular immediate-early gene product Narf." *Neuron* **23**(2): 309-323.

- Pettem, K. L., D. Yokomaku, H. Takahashi, Y. Ge and A. M. Craig (2013). "Interaction between autism-linked MDGAs and neuroligins suppresses inhibitory synapse development." The Journal of Cell Biology **200**(3): 321-336.
- Sobolevsky, A. I., M. P. Rosconi and E. Gouaux (2009). "X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor." Nature **462**(7274): 745-756.
- Song, J. Y., K. Ichtchenko, T. C. Südhof and N. Brose (1999). "Neuroligin 1 is a postsynaptic cell-adhesion molecule of excitatory synapses." Proceedings of the National Academy of Sciences of the United States of America **96**(3): 1100-1105.
- Squire, L. R. (2008). *Fundamental Neuroscience*. L. R. Squire, Academic Press: 1152
- Sukumaran, M., M. Rossmann, I. Shrivastava, A. Dutta, I. Bahar and I. H. Greger (2011). "Dynamics and allosteric potential of the AMPA receptor N-terminal domain." EMBO J **30**(5): 972-982.
- Tabuchi, K., J. Blundell, M. R. Etherton, R. E. Hammer, X. Liu, C. M. Powell and T. C. Südhof (2007). "A Neuroligin-3 Mutation Implicated in Autism Increases Inhibitory Synaptic Transmission in Mice." Science **318**(5847): 71-76.
- Tretter, V., T. C. Jacob, J. Mukherjee, J.-M. Fritschy, M. N. Pangalos and S. J. Moss (2008). "The Clustering of GABAA Receptor Subtypes at Inhibitory Synapses is Facilitated via the Direct Binding of Receptor $\alpha 2$ Subunits to Gephyrin." The Journal of Neuroscience **28**(6): 1356-1365.
- Tyagarajan, S. K. and J. M. Fritschy (2014). "Gephyrin: a master regulator of neuronal function?" Nat Rev Neurosci **15**(3): 141-156.
- Williams, S. P., L. F. Kuyper and K. H. Pearce (2005). "Recent applications of protein crystallography and structure-guided drug design." Current Opinion in Chemical Biology **9**(4): 371-380.
- Yao, G., Y. Zong, S. Gu, J. Zhou, H. Xu, I. I. Mathews and R. Jin (2011). "Crystal structure of the glutamate receptor GluA1 N-terminal domain." Biochemical Journal **438**(2): 255-263.

Appendices

Appendix A- AMPA receptor constructs and oligonucleotides

Table A1- Information for all AMPA receptor constructs detailed in this thesis, including domain boundaries, oligonucleotides (oligos) reference numbers used for PCR amplification and cloning strategies. Presence of * indicates that these oligos were used to introduce a mutation. Oligos for the original domain boundary were then used to amplify constructs as required.

Construct ID	Domain boundaries	Forward	Reverse	Cloning
AM21	N25-E412	90	91	AgeI/BsiWI
AM21m1	F50D (F71D)	264*	265*	AgeI/BsiWI
AM21m2	F82D	266*	267*	AgeI/BsiWI
AM21m3	T53N-A55S	268*	269*	AgeI/BsiWI
AM21m4	L140E	270*	271*	AgeI/BsiWI
AM21m5	L144E	272*	273*	AgeI/BsiWI
GluR2 ATD_D1 DNSS linker at fusion point	N25-S796 ΔG520-D757	GluR2_DNSS_fwd	GluR2_DNSS_rev	AgeI/BsiWI
GluR2 ATD_D1 DDSS linker at fusion point	N25-S796 ΔG520-D757	GluR2_DDSS_fwd	GluR2_DDSS_rev	AgeI/BsiWI
GluA4_ATD_1	S25-G406	655	866	AgeI/BsiWI
GluA4_ATD_2	S25-E413	655	867	AgeI/BsiWI
GluA4_ATD_3	S25-T416	655	868	AgeI/BsiWI
GluA4_ATD_4	S25-T421	655	869	AgeI/BsiWI
GluA2_LBC_flop	K414-G795 ΔP528-S652	118	145	AgeI/BsiWI
GluA2_FL_HA	N25-I883	98	347	AgeI/BsiWI
GluA2_FL_FLAG	N25-I883	358	347	AgeI/BsiWI
GluA2_ΔATD_HA	N406-I883	360	347	AgeI/BsiWI
GluA2_ΔATD_FLAG	N406-I883	359	347	AgeI/BsiWI
RatGluA4FLflip	S25-G893	657	653	AgeI/BsiWI
RatGluA4 ΔATDflip	N407-G893	658	653	AgeI/BsiWI
rGluA3_FL_pHL	F24-I888	1203	1199	AgeI/KpnI
rGluA3_FL_mV	F24-I888	1203	1198	AgeI/KpnI
rGluA3_FL_A653T_pHL	F24-I888_ A653T	1196	1197	AgeI/KpnI
rGluA3_FL_A653T_mV	F24-I888	1196	1198	AgeI/KpnI

*oligo were ordered using old numbering system. New residue numbering is indicated in brackets.

Appendix A- AMPA receptor constructs and oligonucleotides

Table A2- Details all AMPA receptor oligonucleotides used for PCR amplification with reference to oligo numbering detailed in Table A1.

Oligo number	Oligo ID	Sequence 5'-3'
90	hGRIA2_fwd_Age	cgcaccggtAACAGCATACAGATAGGGGGGCTATTTC
91	hGRIA2atd_rev_Bsi	cgccgtacgcTCAAGCCCAGAGGTGTCATTTCAG
98	hGRIA2_HA_fwd_Age	cgcaccggtTACCCCTACGACGTGCCCCGACTACGCCAACAGCATACA GATAGGGGGGCTATTTC
118	hGRIA2_S1fwd_Age	CGCACC GGTAAGACTGTTGTTGTCAACCACAATTTGGAATCTC
N/A	GluR2_DNSS_rev	CTTTGGACGACGAGTTGTCGAGGCTCATGAAGGGCTTTGAGAAG TC
N/A	GluR2_DNSS_fwd	CATGAGCCTCGACAACCTCGTCGTCCAAAGGCTATGGCATCGCAA CACC
N/A	GluR2_DDSS_rev	CTTTGGACGACGAGTcGTCGAGGCTCATGAAGGGCTTTGAGAAGT C
N/A	GluR2_DDSS_fwd	CATGAGCCTCGACgACTCGTCGTCCAAAGGCTATGGCATCGCAAC ACC
264	F50D_fwd (*71)	GACAATTTGGAGGTGGCAAACAGCGACGCACTAATGCTTT CTGCTC
265	F50D_rev (*71)	GAGCAGAAAGCATTAGTGACTGCGTCGCTGTTTGCCACCTCCAA ATTGTC
266	F82D_fwd (*103)	GAAGTCTGTAAATACCATCACATCAGATTGCGGAACACTCCACG TCTCCT
267	F82D_rev (*103)	AGGAGACGTGGAGTGTTCCGCAATCTGATGTGATGGTATTTACA GACTTC
268	T53N_A55S_fwd (*74_76)	GTGGCAAACAGCTTCGCAGTCAATAATTCTTTCTGCTCCCAGTTT TCGAG
269	T53N_A55S_rev (*74_76)	CTCGAAAACCTGGGAGCAGAAAGAATTATTGACTGCGAAGCTGTT TGCCAC
270	L140E_fwd (161)	GACAGTGACAGAGGCTTATCAACAGAGCAAGCTGTGCTGGATTC TGCTG
271	L140E_rev (161)	CAGCAGAATCCAGCACAGCTTGCTCTGTTGATAAGCCTCTGTCA TGTC
272	L144N_fwd (165)	GCTTATCAACACTGCAAGCTGTGAACGATTCTGCTGTGAAAAG AAATGG
273	L144N_rev (165)	CCATTTCTTTTCAGCAGCAGAATCGTTCACAGCTTGCAAGTGTGA TAAGC
347	R2_rev*_BsiWI	CGCCGTACGCTAAATTTTAACACTCTCGATGCCATATACGTTG
358	R2fl_FLAG_fwd_Ag	CGCACC GGTAAGACTACAAGGACGACGACGACAAGAAGCAGCATAC AGATAGGGGGGCTATTTC
359	R2dATD_FLAG_fwd_	CGCACC GGTAAGACTACAAGGACGACGACGACAAGAATGACACCTC TGGGCTTGAGAATAAGACTG
360	R2dATD_HA_fwd_Ag	CGCACC GGTTACCCCTACGACGTGCCCCGACTACGCCAATGACAC CTCTGGGCTTGAGAATAAGACTG
652	rR2_fl_rev_Bsi	CGCCGTACGAATTTTAACACTCTCGATGCCATATACGTTGTAACC TTC
653	rR4_fl_rev_Bsi	CGCCGTACGTGGTAGGTCCGATGCAATGACAGCCA
654	rR4_fl**_rev_Bsi	CGCCGTACGTCAATTATGGTAGGTCCGATGCAATGACAGCCA
655	rR4_ATD_fwd	CGCACC GGTAAGCAGCGTTCAAATAGGTGGTCTCTTCATC
657	rR4_fl_HA_fwd	CGCACC GGTTACCCCTACGACGTGCCCCGACTACGCCAGCAGCGT TCAAATAGGTGGTCTCTTCATC
658	rR4dATD_HA_fwd	CGCACC GGTTACCCCTACGACGTGCCCCGACTACGCCAATGACAC AGCAGCTATTGAGAACAAGACAG
655	rR4_ATD_fwd	CGCACC GGTAAGCAGCGTTCAAATAGGTGGTCTCTTCATC
866	rR4_ATD_rev1_Bsi	cgccgtacgGCCCAGAGTAGGCATATCTTGAATCAAGA
867	rR4_ATD_rev2_Bsi	cgccgtacgCTCAATAGCTGCTGTGTCAATTGCCAG
868	rR4_ATD_rev3_Bsi	cgccgtacgTGTTCTGTCTCAATAGCTGCTGTGTCAATTG
869	rR4_ATD_rev4_Bsi	cgccgtacgTGTGGTTACAACCACTGTTCTGTTCTCAATAG

Appendix A- AMPA receptor constructs and oligonucleotides

118	hGRIA2_S1fwd_Age	CGCACCGGTAAGACTGTTGTTGTCACCACAATTTTGAATCTC
145	hGRIA2S2_rev_Bsi	CGCCGTACGGCCGCACTCTCCTTTGTCTGTACCA
1196	rGluA3_A653T_fwd	TCTCTTCCTACACTGCAAACCTTaCTGCTTTCCTGACTGTGGAGAG GATG
1197	rGluA3_A653T_rev	CATCCTCTCCACAGTCAGGAAAGCAGtAAGGTTTGCAGTGTAGGA AGAGA
1203	rGluA3_HA_fwd_Ag	cgcaccggtTACCCCTACGACGTGCCCCGACTACGCCTTCCCCAACACC ATCAGCATAGGTGGA
1199	rGluA3_**_rev_Kp	cgcggtaccCTACTAGATCTTAACACTTTCTGTTCCATACACGTTG
1198	rGluA3_rev_Kpn	cgcggtaccGATCTTAACACTTTCTGTTCCATACACGTTG

* oligo were ordered using old numbering system. New residue numbering is indicated in brackets.

Appendix B- Neuronal pentraxin constructs and oligonucleotides

Appendix B containing construct information, oligonucleotide numbering, N-linked glycosylation states and species sequence alignment for NP1

Table B1- All neuronal pentraxins construct info, including domain boundaries, oligonucleotides (oligos) used in PCR and cloning strategy. Presence of * indicates that these oligos were used to introduce the required mutation. Original domain boundary oligos were then used to amplify constructs as required.

Construct ID	Boundaries	Forward	Reverse	Cloning
NP1_PTX	P224-I431	253	254	Age1/Kpn1
NP1_FL	Q23-I431	252	254	Age1/Kpn1
NP1_CC2	G112-I431	705	254	Age1/Kpn1
NP1_PTX_N280D	P224-I431	1178*	1177*	Age1/Kpn1
NP1_PTX_Q370E	P224-I431	1183*	1182*	Age1/Kpn1
NP1PTX_E358S+D360S	P224-I431	1185*	1184*	Age1/Kpn1
NP2_PTX	S221-L431	256	257	Age1/Kpn1
NP2_FL	Q18-L431	255	257	Age1/Kpn1
NP2_CC2	G89-L431	379	257	Age1/Kpn1
NP2_PTX_N393Q	S221-L431	483*	484*	Age1/Kpn1
NPR_PTX	S289-A500	839	840	Age1/Kpn1
NPR_CC2	G118-A500	841	840	Age1/Kpn1

Table B2- Details all neuronal pentraxins oligonucleotides used for PCR amplification. Presence of ** indicates that these oligos were named according to an old numbering system which is suitable for reordering means only. All constructs described are correctly numbered starting from the first amino acid residue, typically the start of each proteins native secretion signal.

Oligo number	Oligo ID	Sequence 5'-3'
252	NP1fl_fwd_Age	cgcaccggtCAGGATTTCGGGCCGACGCGCT
253	NP1ptx_fwd_Age	cgcaccggtCCTGGAGACAAGTTCAGCTCACATTC
254	NP1fl_rev_Kpn	CGCGGTACCGATCTGGCGACAGGCCTCGAAGGT
255	NP2fl_fwd_Age	CGCACCGGTCAGGACAGCCCGGCGCCCGGTA
256	NP2ptx_fwd_Age	cgcaccggtTCACCAGATGCGTTCAAGGTGTCCCT
257	NP2fl_rev_Kpn	cgcgggtaccCAAGTCAAGGAGACGCTCCTCACAC
379	NP2cc2_Age_fwd	cgcaccggtACGGGCAAGGACACTATGGGCGA
705	NP1_CC2_fwd_Age	CGCACCGGTTCCGGCAAGAACACCATGGGCGAC
483	N393Q_fwd	CACAAGAAATTGTCAACATCGCCcAgTGCTCCACAAACATGCCGGGCAA
484	N393Q_rev	TTGCCCCGGCATGTTTGTGGAGCAcTgGGCGATGTTGACAATTTCTTGTG
839	NPR_PTX_fwd_Age	CGCACCGGTAGTCCTCCAGATGCCTTCAAGATCAGCA
840	NPR_rev_Kpn	CGCGGTACCTGCCTTGGCCCTCCCTTGCAGA
841	NPR_CC_fwd_Age	CGCACCGGTGAGCGCGAAGAGCTGCTGCTGCT(32bp)
1177	NP1PTX_N57D_fwd**	CTACGCTGTGCCCCGGCCAGGCCGACGAGCTGGTCCTCATTGAGTGGGGC
1178	NP1PTX_N57D_rev**	GCCCCACTCAATGAGGACCAGCTCGTCGGCCTGGCCGGGCACAGCGTAG
1182	NP1PTX_Q147E_fw	GGTGGTGGGTTTGATGCCACCGAGGCATTTGTGGGTGAGCTGGCCCCAC
1183	NP1PTX_Q147E_rv	GTGGGCCAGCTCACCCACAAATGCCTCGGTGGCATCAAACCCACCACC
1184	NP1PTX_SSmut_fw	CGTGCTGGTGTGGGCCAGTCCCAGTCCACTCTGGGTGGTGGGTTTGATGC
1185	NP1PTX_SSmut_rv	GCATCAAACCCACCACCCAGAGTGGACTGGGACTGGCCAGCACCAGCAGC
N/A	R4_ATD_wing_fwd	CACCCATGGTAACCAACCTATCGGATCGGTGGAAGAACTAGATC
N/A	R4_ATD_wing_rev	GATCTAGTTTCTTCCACCGATCCGATAGGTTGGTTACCATGGGTG
N/A	R4_ATD_flap_fwd	CCGAAGTCTCAGAAGACAGAACATTCTATTTCAGGAGAGGAAATGCTG
N/A	R4_ATD_flap_rev	CAGCATTCCTCTCCTTGAAATAGAAATGTTCTGTCTTCTGAGACTTCGG

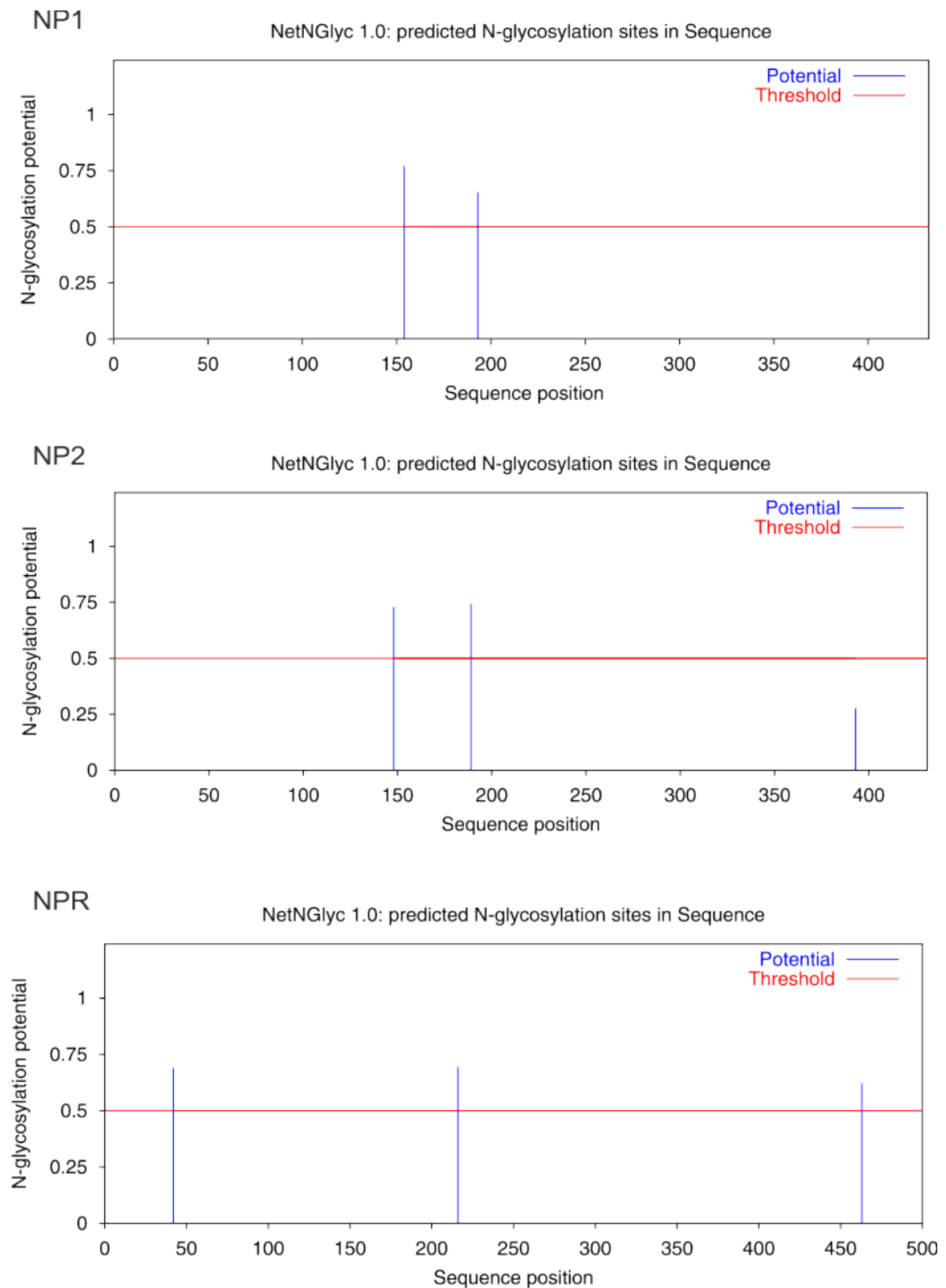


Figure B1 NetNGlyc prediction for NP1, NP2 and NPR full length protein.

Appendix B- Neuronal pentraxin constructs and oligonucleotides

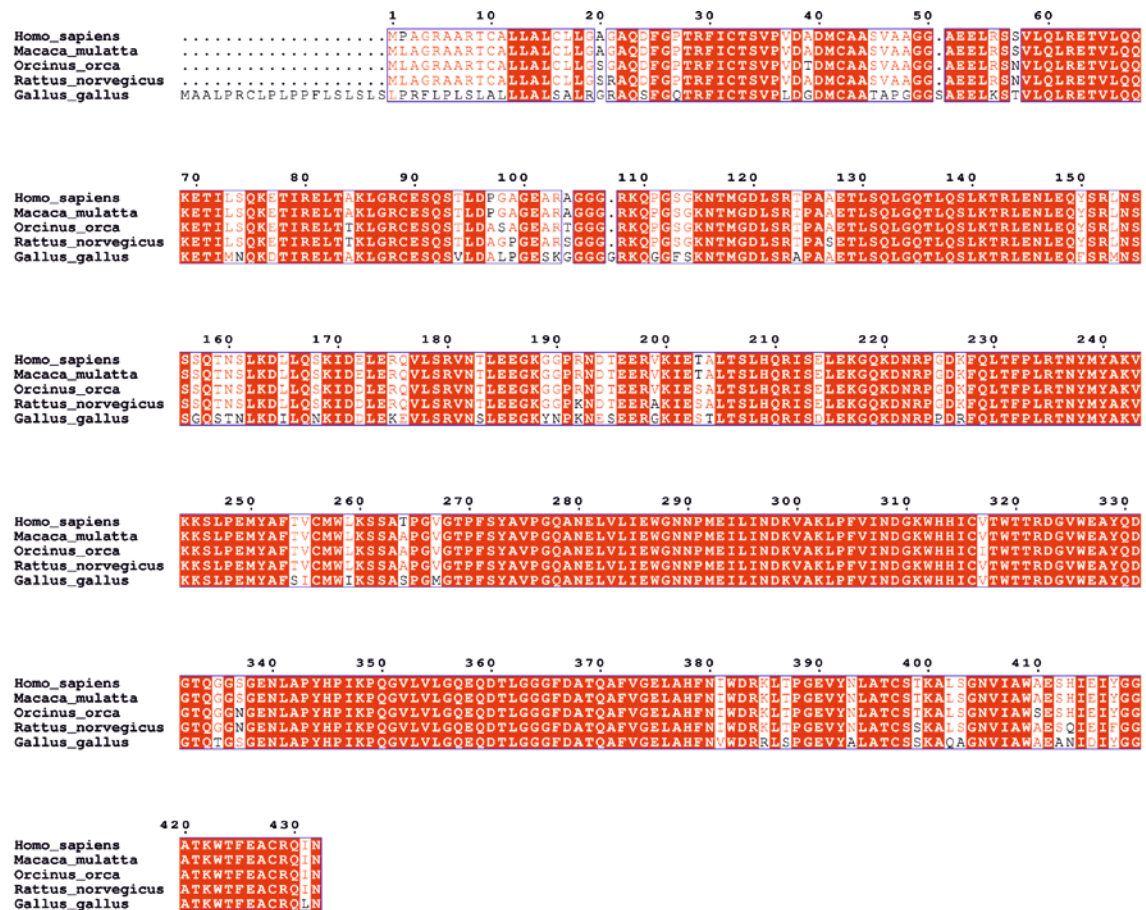


Figure B2 Neuronal pentraxin I is conserved amongst species.

Multialign showing high level of neuronal pentraxin I sequence homology across *Homo sapiens* (human), *Macaca mulatta* (rhesus macaque), *Orcinus orca* (killer whale), *Rattus norvegicus* (rat) and *Gallus gallus* (chicken). All sequences sourced from PubMed <http://www.ncbi.nlm.nih.gov/pubmed>.

Appendix C- AMPA receptor-interacting proteins oligos and constructs

Table C1- Construct info, including domain boundaries, forward and reverse oligo IDs used for PCR amplification and cloning strategies.

Construct ID	Boundaries	Forward	Reverse	Cloning
Brorin	M1-M324	1029	1030	EcoRI/KpnI
Brorin_S1*	S28-M324	1190	1030	Age1/KpnI
Brorin_R21*	R48-M324	1191	1030	Age1/KpnI
Brorin_D41*	D68-M324	1192	1030	Age1/KpnI
Brorin_E61*	E88-M324	1193	1030	Age1/KpnI
Brorin_D81*	D108-M324	1194	1030	Age1/KpnI
Br_E101	E128-M324	1195	1030	Age1/KpnI
Br_short	D148-M324	1176	1030	Age1/KpnI
Noelin_1	M1-L485	1023	1024	EcoRI/KpnI
Noelin_2	M1-P478	1025	1026	EcoRI/BsiWI
Noelin_3	M1-T458	1027	1028	EcoRI/KpnI
Neuritin	M1-G114	1031	1032	EcoRI/KpnI

*Old construct numbering system used to order oligos. True boundaries indicated in Table C1 column 2.

Table C2- Details all oligonucleotides used for PCR amplification. All constructs described are numbered starting from the first amino acid residue, typically the start of each proteins native secretion signal.

Oligo number	Oligo ID	Sequence 5'-3'
1023	Noel1_EcoR1_fwd	TCCTGAATTGCGCCACCatgTCGGTGCCGCTGCTGAAGATC
1024	Noel1_Kpn_rev	gcgGGTACCCAACTCATCGGAGCGGATGACATGGA
1025	Noel2_EcoR1_fwd	TCCTGAATTGCGCCACCatgtcggtgccgtgctcaagatc
1026	Noel2_BsiWI_rev	cgccgtacgTGGGTCTCCGGCAGTGCTGATGAC
1027	Noel3_EcoR1_fwd	TCCTGAATTGCGCCACCatgcaggcgagggtccagctttctc
1028	Noel3_Kpn_rev	gcgGGTACCTGTGTCTATCTTCAGTCTTAATGATGTGGAACAG
1029	Brorin_Eco_fwd	TCCTGAATTGCGCCACCatgCCTAGCTCCTCTGCGATGGCA
1030	Brorin_Kpn_rev	gcgGGTACCCATTTGCCTGCATTCATGTCTCGTGAC
1031	Neur_EcoRI_fwd	TCCTGAATTGCGCCACCatgGGAAGTAAAGTTGAACGGCAGATATATTCA
1032	Neur_Kpn_rev	gcgGGTACCGCTGCCGCAGAGTTCAATAAGCTG
1176	Bror_sh_Age_fwd	gcgACCGGTgactaccgaggttaagggtgcgtggac
1190	Br_S1_Age_fwd	gcgACCGGTAGTCCGAGCATCCCGCTGGAGAA
1191	Br_R21_Age_fwd	gcgACCGGTCCGAGCAGCGTCTCGG
1192	Br_D41_Age_fwd	gcgACCGGTGACGAGGGCAGCAGTGCCC
1193	Br_E61_Age_fwd	gcgACCGGTGAGGCGTGAGCAAGCAGAAGC
1194	Br_D81_Age_fwd	gcgACCGGTGACTGGCAGGTCCGGCCTC
1195	Br_E101_Age_fwd	gcgACCGGTGAAGCGATCAGTCTGGAAGTCATGC

The original Brorin construct had been inserted into the pcDNA3.1+ plasmid via HindIII/EcoRI. The full length protein was 324 amino acids in length, the molecular weight 36.57kDa, with a theoretical pI of 5.82 and the extinction coefficient 1.1.

Nucleotide sequence:

aagcttatgcctagctcctctgcatggcagttggcgcgctctccagttccctcttggtcacctg
ctgcctgatgggtggctctgtgcagtccgagcatcccgtggagaagctggcccaagctccagagc
aaccggccaggaaaagcgagcagcgcgtctcgggacagccgggaagggtgagcgagctcggg
cgcgctctagggacgagggcagcagtgcccggaactggaagagcaagggcagccgcgcgtctc
gggcccggaggcgtggagcaagcagaagcagcctgggctgcccagagcgggagcgccaaggccg
cggactggcaggtccggcctcgcgggatacccccagggggagccccggcgccgcccaggaa
gcgatcagctctggaactcatgcccacaccggagctgcctgaggaatacgcctatcccgactaccg
cggaagggtgctggagcagagtggttcgtgtatgcgattggggagaagttcgccccaggtc
cctcagcctgcccattgtctgtgtacggaggagggggccctttgcgcccagccagagtgcccgcg
cttcaccgcgctgcatccacgttgacaccagtcagtgctgtccgcagtgcaaagagaggaagaa
ctactgcgagttccggggcaagacctaccagactttggaggagttcgtggtgtctccatgtgaac
ggtgtcgtgtgaagccaatggtgaagtgtgtgcacagtgctcggtgtgtccccagacggaatgt
gtggaccccggtgtatgagcctgatcagtgctgcccgatctgcaaaaacggtccaaactgctttgc
agaaaccgctgtcatcccagctggtcagaggtgaagactgacgaatgcacatatgtcattgta
cttatgaggagggcagctggagaattgaacggcaagccatgtgcacgagacatgaatgcaggcaa
atgtaggaattc

Amino acid sequence:

MPSSAMAVGALSSLLVTCCLMVALCSPSIPLEKLAQAPEQPGQEKREHASRDSPGRVSELGRA
SRDEGSSARDWKSGRALSGREAWSKQKQAWAAQSGSAKAADWQVRPRGDT PQGEPPAAAQEI
SLELMPTPELPPEEYAPDYRGKGCVDSEGFVYAIGEFAPGPSACPCLCTEEGPLCAQPECPR LH
PRCIHVDTSQCCPQCKERNYCEFRGKTYQTLEEFVVSPECRCRCEANGEVLCTVSACPQTECVD
PVYEPDQCCPICKNGPNCFAETA VIPAGREVKTDECTICHCTYEEGTWRIERQAMCTRHECRQM-
EF

Alignment:

atgcctagctcctctgcatggcagttggcgcgctctccagttccctcttggtc
M P S S S A M A V G A L S S S L L V
acctgctgcctgatgggtggctctgtgcagtccgagcatcccgtggagaagctggcccaa
T C C L M V A L C S P S I P L E K L A Q
gctccagagcaaccggccaggaaaagcgcgagcagcgcgtctcgggacagccgggaagg
A P E Q P G Q E K R E H A S R D S P G R
gtgagcgagctcgggcgcgcctctagggagggcagcagtgcccggaactggaagagc
V S E L G R A S R D E G S S A R D W K S
Aagggcagccgcgcgtctcggggccgggagcgtggagcaagcagaagcaggcctgggct
K G S R A L S G R E A W S K Q K Q A W A
Gccagagcgggagcgccaaggccgcggactggcaggtccggcctcgcggggatacccca
A Q S G S A K A A D W Q V R P R G D T P
Cagggggagccccggcgccgcccaggaagcagatcagctctggaactcatgccacaccg
Q G E P P A A A Q E A I S L E L M P T P
gagctgcctgaggaatacgcctatcccagctaccgcccgaagggtgctgctggagcagagtg
E L P E E Y A Y P D Y R G K G C V D E S
Ggcttcgtgtatgcgattggggagaagttcgccccaggtccctcagcctgcccattgtctg
G F V Y A I G E K F A P G P S A C P C L
Tgtacggaggagggggccctttgcgcccagccagagtgcccgcggttcaccgcgcgtgc
C T E E G P L C A Q P E C P R L H P R C
atccacgttgacaccagtcagtgctgtccgcagtgcaaagagaggaagaactactgcgag
I H V D T S Q C C P Q C K E R K N Y C E
ttccggggcaagacctaccagactttggaggagttcgtggtgtctccatgtgaacggtgt

F R G K T Y Q T L E E F V V S P C E R C
 cgctgtgaagccaatggtgaagtgtgtgcacagtgtcggcgtgtccccagacggaatgt
 R C E A N G E V L C T V S A C P Q T E C
 gtggaccccggtgtatgagcctgatcagtgctgcccgatctgcaaaaacggtccaaactgc
 V D P V Y E P D Q C C P I C K N G P N C
 tttgcagaaaccgctgtcatcccagctggtcgagaggtgaagactgacgaatgcaccata
 F A E T A V I P A G R E V K T D E C T I
 tgtcattgtacttatgaggaggcacgtggagaattgaacggcaagccatgtgcacgaga
 C H C T Y E E G T W R I E R Q A M C T R
 catgaatgcaggcaaatgttaggaattc
 H E C R Q M - E F

The Brorin signal sequence (M1-C27) is highlighted in cyan blue, O-linked glycosylation sites are shown in red, the N-terminal truncation start residues are highlighted in magenta, a potential cell-adhesion motif shown in green. The start residue for the Brorin_short construct highlighted in yellow.

Brorin_short construct design:

Bror_sh_Age_fwd

DYRGKGCVDESGFVYAIGKEKFAPGPSACPLCTEEGPLCAQPECPR LHPRCIHVDTSQCCPQCKE
 RKNYCEFRGKTYQTLEEFVVS PCERCRC EANGEVLCTVSACPQTECVDPVYEPDQCCPICKNGPN
 CFAETAVIPAGREVKTDECTICHCTYE EGTWRIERQAMCTRHECRQM **GTKHHHHHH**
 21KDa

Ex Co 0.77

N-terminal truncations in 20 amino acid steps

mouse Noelin-1 (EcoRI/AgeI) in pEGFP-N1

secretion signal SignalP does not predict a strong signal sequence

8 N-linked glycosylation sites

potential O-linked site

Olfactomedin-like domains

stop codon

gaattcatgtcgggtgccgctgctgaagatcggggtcgtgctaagcaccatggccatgatc
 E F M S V P L L K I G V V L S T M A M I
 accaactggatgtcccagacgctgccctcgtggtgggcctcaataccacccggctttcg
 T N W M S Q T L P S L V G L N T T R L S
 gccgccagcggcgggacgctggaccgcagcaccggcgtgttgcccaccaaccccgaggag
 A A S G G T L D R S T G V L P T N P E E
 agctggcaggtgtacagctctgccaggacagtgagggcaggtgtatctgcacagtggtc
 S W Q V Y S S A Q D S E G R C I C T V V
 gccccccagcagaccatgtgttcggggatgcccgcacaaaacagctgaggcagctactg
 A P Q Q T M C S R D A R T K Q L R Q L L
 gagaaggtgcaaaacatgtctcaatccatagaagtcttggacaggcgaactcagagagac
 E K V Q N M S Q S I E V L D R R T Q R D
 ctgcagtacgtggagaagatggagaaccagatgaaaggcctggagaccaagttcaagcag
 L Q Y V E K M E N Q M K G L E T K F K Q
 gtggaggagagccataagcagcatctagccaggcagttcaaggcaataaaagcgaaaatg
 V E E S H K Q H L A R Q F K A I K A K M
 gatgaacttaggcctttgattcctgtgttgaagagtacaaggccgatgccaatttgga

D E L R P L I P V L E E Y K A D A K L V
 ttgcagtttaaggaggaggtccagaatctgacgtcagtgcttaacgaactgcaagaggag
 L Q F K E E V Q N L T S V L N E L Q E E
 attggcgctatgactacgatgaacttcagagcagagtgccaatcttgaagaaaggctc
 I G A Y D Y D E L Q S R V S N L E E R L
 cgtgcatgcatgcaaaaactagcctgcgggaagctgacaggcataagtgacccagtgact
 R A C M Q K L A C G K L T G I S D P V T
 gtcaagacctctgggtcaagggttcgggtcctggatgacagaccgctggcccagaaggc
 V K T S G S R F G S W M T D P L A P E G
 gataaccgggtctggtacatggatggttatcacaacaaccgctttgtccgtgagtacaag
 D N R V W Y M D G Y H N N R F V R E Y K
 tccatggtggacttcatgaacacggacaactttacctctcaccgtctccacaccctgg
 S M V D F M N T D N F T S H R L P H P W
 tcaggtacgggccaggtggtctacaatggctccatctatttcaaaaattccagagtcac
 S G T G Q V V Y N G S I Y F N K F Q S H
 attatcattaggtttgacctgaagacagagacgattctgaagacgcgcagcctggactat
 I I I R F A D L K T E T I L K T R S L D Y
 gctggtacaacaatgtatcactacgcctggggcgccactcagacatcgacctcatg
 A G Y N N M Y H Y A W G G H S D I D L M
 gtggacgagaatgggctctgggtgtctatgccacaaccagaacgcaggcaacattgtc
 V D E N G L W A V Y A T N Q N A G N I V
 atcagcaagctggatcctgtctcactgcagatcctccagacctggaacaccagctacccc
 I S K L D P V S L Q I L Q T W N T S Y P
 aagcgagtgccggtgagccttcacatctgcggcaccctctatgtcaccacgggttac
 K R S A G G E A F I I C G T L Y V T N G Y
 tcggggggcaccaggtccactatgcctaccaaacacgcctcgacctacgagtacatt
 S G G T K V H Y A Y Q T N A S T Y E Y I
 gacatcccttccagaacaatactctcacatctccatgctggactacaaccctaaagac
 D I P F Q N K Y S H I S M L D Y N P K D
 cgcgcgctgtacgcctggaacaatggccatcagaccctctataatgtcaccctcttccat
 R A L Y A C W N N G H Q T L Y N V T L F H
 gtcacccgctccgatgagttgtagagatctgcagctgggtaccacgggt
 V I R S D E L - R S A A G T T G

EFMSVPLLKIGVVLSTAMITNWSQTLPSLVGLNTRLSAASGGTLDRSTGVLPTNP
 SWQVYSSAQDSEGRICITVAPQQTMCSDARTKQLRQLLEKVQNSQSIEVLDRTQRD
 LQYVEKMENQMKGLETKFKQVEESHKQHLARQFKAIAKAKMDLRPLIPVLEEYKADAKLV
 LQFKEEVQNLTSVLNELQEEIGAYDYDELQSRVSNLEERLRACMQKLACGKLTGISDPVT
 VKTSGSRFGSWMTDPLAPEGDNRVWYMDGYHNNRFVREYKSMVDFMNTDNFTSHRLPHW
 SGTGQVYNGSIYFNKFQSHIIRFDLKTETILKTRSLDYAGYNNMYHYAWGGHSDIDL
 VDENGWLWAVYATNQAGNIVISKLDPVSLQILQTWNTSYPKRSAGEAFIICGTLYVTNGY
 SGGTKVHYAYQTNASTYEYIDIPFQNKYSHISMLDYNPKDRALYAWNNGHQTLYNVTLFH
 VIRSDEL-RSAAGTTG

Extinction coefficient 1.6
 Number of amino acids: 485
 Molecular weight: 55427.7
 Theoretical pI: 6.48

Noelin 2 rat Noelin-2 (EcoRI/AgeI) in pEGFP-N1

gaattcatgtcgggtgccgctgctcaagatcgggcggtgctgagcaccatggccatggct
 E F M S V P L L K I G A V L S T M A M V
 actaactggatgtcgcagacgctgccctcgtcgtgggactcaacagcaccgtgtccgc
 T N W M S Q T L P S L V G L N S T V S R
 gcgggctcctcagagaaaaactctcttccagagcccagaagagggtggcagctgtac
 A G S S E K I T L F Q S P E E G W Q L Y
 acgtcagcgcgagcgccggatggcaaatgcatctgcacagctgtgattcctgcacagac
 T S A Q A P D G K C I C T A V I P A Q S
 acctgtgcccagatggtcggagcagagagcttcgccagctcatggagaagggtccagaat
 T C A R D G R S R E L R Q L M E K V Q N
 gtgtctcagtcacatggaagtccttgagctaaggacataccgggatctccagtatgttcg
 V S Q S M E V L E L R T Y R D L Q Y V R

agcatgtgagaccctcatcgggagcctggatgcaaggctcaggacagcgatgggtcagtc
S M E T L M R S L D A R L R T A D G S V
tcagccaaaagccttcagggaactgaaggacaggatgacagagctgctgcccctgagttca
S A K S F Q E L K D R M T E L L P L S S S
gtgcttgagcagtagtcaaagcagacacggaccattgtgcgcctgcgggaggaggtgagg
V L E E Q Y K A D T R T I V R L R E E V R
aacctctcaggcaacctggctgccatccaggaggaaatgggtgcctacgggtacaggagac
N L S G N L A A I Q E E M G A Y G Y E D
ttgcagcagcgcgtgatggccctggaggcccgaactccatgcctgcgcgcagaagctgggc
L Q Q Q R V M A L E A R L H A C A Q K L G
tgcgggaagctgcacggcgctcagtaacccccattaccatccgagccatgggattcccgctt
C G K L T G G V S N P I T I R A M G S R F
ggctcctggatgactgacacgatggccccagtgacagacagtcgggtctggtacatggat
G S W M T D T M A P S A D S R V W Y M D
ggttattacaaaggccgcccagtgctggagttccgtactttgggagacttcatcaaaggc
G Y Y T K G R R V L E F R T L G D F I K G
cagaactcatccagcactgtgtccgcagccatgggcaggtagggccattgtggtatac
Q N F I Q H L L P Q P W A G T G H V V Y
aatggatctctgtttctacaacaagtaccagagcaatgtggtggttaagtaccacttcggg
N G S L F Y N K Y Q S N V V V K Y H F R
tcccgcctccgtgttggtgcagaggagcctcccgggggctggttacaataacaccttcccc
S R S V L V Q R S L P G A G Y N N T F P
tattcctggggcggttctcggacatggacttcatggtagacgagagcggtgtgtgggt
Y S W G G F S D M D F M V D E S G L W A
gtgtatacccaaccagcaacgcggcaacattgtgtgcagccggctggaccctcacac
V Y T T N Q N A G G N I V V S R L D P H T
ctggaggtcgtgaggtcctgggacactgggtaccctaagcgcagcgccggtgaggccttc
L E V V R S W D T G Y P K R S A G E A F
atgatctgtggtgtcctctacgtgaccaactctcacctggccggagccaaggtctacttt
M I C G V L Y V T N S H L A G A K V Y F
gcgtacttcaccaacacgtccagctatgagtacacggatgtgcccttcacaaccagtac
A Y F T N T S S Y E Y T D V P F H N Q Y
tcgcacatctccatgtctggattacaaccccaggggagcgggccctgtatacttggaacaac
S H I S M L D Y N P R E R A L Y T W N N
gggcaccaggtgtgttacaacgtcacccctcttcacgtcatcagcactgcccggggaccccc
G H Q V L Y N V T L F H V I S T A G D P
tagagatctgcagctggtaccgggt
- R S A A G T G

EFMSVPLLKIGAVLSTMAMVTNWSQTLPSVLGLNSTVSRAGSSEKITLFQSP EEGWQLY
TSAQAPDGKCICTAVIPAQSTCARDGRSRELRLMEKVQNVQSMEVLELR TYRDLQYVR
SMETLMRSLDARLRTADGSVSAKS FQELKDRMTELLPLSSVLEQYKADTRTIVRLREEVR
NLSGNLAATQEEMGAYGYEDLQQRVMAL EARLHACAQKLGCGKLTGVSNPITIRAMGSRF
GSWMTDTMAPSADSRVWYMDGYKGRRVLEFRTL GDFIKGQNF IQHLLPQPWAGTHGVVY
NGSLFYNKYQSNVVVYKHFRSRSVLVQRSLPGAGYNNTFPYSWGGFSDMDFMVDESGLWA
VYTTNNQNASNI VVSRLDPHLETVRSWDTGYPKRSAGEAFMICGVLYVTNSHLAGAKVYF
AYFTNTSSYEYTDVPFHNQYSHISMLDYNPRERALTWNNGHQVLYNVTLFHVISTAGD**P**
-RSAAGTG

Ex Co 1.62

Number of amino acids: 487

Molecular weight: 54747.0

Theoretical pI: 8.58

Noelin -3

mouse Noelin-3 (**XhoI/AgeI**) in pEGFP-N1

ctc**gag**atgcaggcgagggtccagctttctcaacctccttcttctgtctttacttgcgtgattagatccttccaagactcagattagccccaagaaggggtggcaagtgtacagctcagcacaggacctgatgggcgatgcatttgtacttgtgttgcgcagaacaaaacctgtgttccagagatgccaaaagcaggcagcttcgccagctactggaaaaggtacagaatatgtcacagtctatagaagtcctaaacttgagaactcagagagattttcagtatgtcttaaaaatggaacccaaatagaagggtgaaaggccaagtttcggcagattgaagatgatcggagacgcctaatagaccaagcattttcaagagttgaaagagaaaatggatagctatgtccctgatcccagttgtggagcagtacaacacgatgcgaagctcatcacagcttcaggagaagaattcggaaactgtcttcagtcctcactggaattcaggaggaataatggtgcctatgactatgaagaacttcataacgcggtcctcag

Appendix C- AMPA receptor-interacting proteins oligos and constructs

cttagagaccaggcttcgtgactgcatgaaaaagctaacatgtggcaaattgatgaaaatcacaggcccatcacagtca
agacatctggaactcgatttgggtgcttggatgacggatccttttagcatctgaaaagaataacagagtctggtatatggac
agttataccaacaataaaaatagttcgtgaatacaagtcattgcagactttgtgagtgaggctgaatcaaggacatacaa
cctccctttcaagtgggcaggaactaaccatgttgctctacaatggctctctctattttaacaaatatcagagtaatatca
tcatcaagtacagctttgacttagggagagtgctagcccaaagaagcctggagtagcgtgggtttcacaaatgtgtacccc
tatacatgggggtgattctctgacattgacctaatggctgatgaaattggctgtgtgggctgtatatgcaactaaccagaa
tgcaggcaatattgtcatcagccagcttaatacaagataccttggagggtgatgaaaagctggagcactggctatcccaaga
gaagtgcagggaatccttcatgatctgtgtggactctgtacgtcaccaattcccacttaacaggagccaaagtgtattat
tcctattccacaaaaacttccacatatgagtacacggacatccccttccataaccaatacttccacatatctatgcttga
ctacaatgcaagagatagagctctttatgcctggaacaatggccaccaggctcctgttcaatgttaccctgttccacatca
ttaagactgaagatgacaca **tagtctagactagttaaccgggtaccggt**

ctcgagatgcaggcgagggtccagctttctcaacctccttctctgtctttacttgcgtgga

L E M Q A R S F L N L L L L S L L A G
ttagatccttccaagactcagattagccccaagaagggtggcaagtgtacagctcagca
L D P S K T Q I S P K E G W Q V Y S S A
caggaccctgatgggcgatgcatttgtactgtggttgcctcagaacaaaacctgtgttcc
Q D P D G R C I C T V V A P E Q N L C S
agagtgccaaagcaggcagcttctcgagctactgaaaagggtacagaatatgtcacag
R D A K S R Q L R Q L L E K V Q N M S Q
tctatagaagtcttaaaacttgagaactcagagagattttcagtatgtcttaaaaatggaa
S I E V L N L R T Q R D F Q Y V L K M E
acccaaatgaaagggtgaaggccaagtttcggcagattgaagatgatcggaagacgcta
T Q M K G L K A K F R Q I E D D R K T L
atgaccaagcattttcaagagttgaaagagaaaatggatgagctactgccctgatccca
M T K H F Q E L K E K M D E L L P L I P
gtgttggagcagtacaaaaccgatgcaaagctcatcacacagttcaaggaagaaattcgg
V L E Q Y K T D A K L I T Q F K E E I R
aacctgtcttcagtcctcactgggattcaggaggaattgtgtcctatgactatgaagaa
N L S S V L T G I Q E E I G A Y D Y E E
cttcatcaacgggtcctcagcttagagaccaggcttcgtgactgc **atgaaaaagctaaca**
L H Q R V L S L E T R L R D C M K K L T
tgtggcaaattgatgaaaatcacaggcccatcacagtcaagacatctggaactcgattt
C G K L M K I T G P I T V K T S G T R F
ggtgcttggatgacggatccttttagcatctgaaaagaataacagagtctggtatatggac
G A W M T D P L A S E K N N R V W Y M D
agttataccaacaataaaaatagttcgtgaatacaagtcattgcagactttgtgagtgga
S Y T N N K I V R E Y K S I A D F V S G
gctgaatcaaggacatacaacctcccttcaagtgggcaggaactaaccatgttgtctac
A E S R T Y N L P F K W A G T N H V V Y
aatggctctctctattttaacaaatatcagagtaatatcatcatcaagtacagctttgac
N G S L Y F N K Y Q S N I I I K Y S F D
ttagggagagtgctagcccaaagaagcctggagtagcgtggttttcacaaatgtgtacccc
L G R V L A Q R S L E Y A G F H N V Y P
tatactgggggtgattctctgacattgacctaatggctgatgaaattggctgtgtgggt
Y T W G G F S D I D L M A D E I G L W A
gtatatgcaactaaccagaatgcaggcaatattgtcatcagccagcttaatacaagatacc
V Y A T N Q N A G N I V I S Q L N Q D T
ttggagggtgatgaaaagctggagcactggctatcccaagagaagtgcagggaatccttc
L E V M K S W S T G Y P K R S A G E S F
atgatctgtggactctgtacgtcaccaattcccacttaacaggagccaaagtgtattat
M I C G G T L Y V T N S H L T G A K V Y Y
tcctattccacaaaaacttccacatatgagtacacggacatccccttccataaccaatac
S Y S T K T S T Y E Y T D I P F H N Q Y
tttcacatatctatgcttgactacaatgaagagatagagctctttatgcctggaacaat
F H I S M L D Y N A R D R A L Y A W N N
ggccaccagggtcctgttcaatgttaccctgttccacatcat **taagactgaagatgacaca**
G H Q V L F N V T L F H I I K T E D D T
tagtctagactagttaaccgggtaccggt
- S R L V N P G T G

LEMQARSSFLNLLLLSLLAGLDPSKTQISPKEGWQVYSSAQDPDGRICITVVAPEQNLCS
RDAKSRLRQLLEKVQNMSQSIEVLNLRQDRFQYVLKMETQMKGLKAKFRQIEDDRKTL
MTKHFQELKEKMDLLPLIPVLEQYKTDALITQFKEEIRNLSSVLTGIQEEIGAYDYEE
LHQRVLSLETRLRDCMKKLTGKLMKITGPI TVKTS GTRFGAWMTDPLASEKNNRVWYMD
SYTNNKIVREYKSIAD FVSGAESRTYNLPFKWAGTNHVVYNGSLYFNKYQSNIIKYSFD
LGRVLAQRSLEYAGFHNVPYTWGGFSDIDLMADEIGLWAVYATNQNAGNIVISQLNQDT
LEVMSWSTGYPKRSAGESFMICGTLVYTNHSLTGAKVYYSYSTKTSTYETDIPFHNQY

Appendix C- AMPA receptor-interacting proteins oligos and constructs

FHISMLDYNARDRALYAWNNGHQVLFNVTLFHIIKTEDDT-SRLVNPGTG

Number of amino acids: 469

Molecular weight: 53965.4

Theoretical pI: 8.31

Ex.co 1.57

Neuritin

rat Neuritin (**EcoRI/XhoI**) in pcDNA3.1+

gaattcatgggacttaagttgaacggcagatatatcttactgatcctcgcggtgcaaatagcttacctgggtgcaggccgt
gagagcagcaggcaagtgcgatgcagtccttaagggctttcagactgtttgctcaagctgggtgacagcatggccaact
acccgcagggcctggacgacaagacgaacatcaagaccgtgtgcacatactgggaggatttccacagctgcacggtcaca
gctcttacggattgccaggaagggcgaaagatatgtgggataaactgagaaaagaatcgaaaaacctcaatatccaagg
cagcttattcgaactctgcggcagcggcaacggggcggggtccctgctcccgcgctttccgtgctcctgggtgtctc
tctcggcagcttttagcgacctggcctttccttc**cgagtcgacgagctcgag**

gaattc**atgggacttaagttgaacggcagatatatcttactgatcctcgcggtgcaaata**
E F M G L K L N G R Y I S L I L A V Q I
gcttacctgggtgcaggccgtgagagcagcaggcaagtgcgatgcagtccttaagggcttt
A Y L V Q A V R A A G K C D A V F K G F
tcagactgtttgctcaagctgggtgacagcatggccaactacccgcagggcctggacgac
S D C L L K L G D S M A N Y P Q G L D D
aagacgaacatcaagaccgtgtgcacatactgggaggatttccacagctgcacggtcaca
K T N I K T V C T Y W E D F H S C T V T
gctcttacggattgccaggaagggcgaaagatatgtgggataaactgagaaaagaatcg
A L T D C Q E G A K D M W D K L R K E S
aaaaacctcaatatccaagcagcttattcgaactctgcggcagcggcaacggggcgggcg
K N L N I Q G S L F E L C G S G N G A A
gggtccctgctcccgcgctttccgtgctcctgggtgtctctctcggcagcttttagcgacc
G S L L P A L S V L L V S L S A A L A T
tggcctttccttc**cgagtcgacgagctcgag**
W L S F - V D E L E

EF**M**GLKLNGRYISLILAVQIAYLVQAVRAAGKCDVFKGFSDCLLKLGD^{SMANYPQGLDD}
KTNIKTVCTYWEDFHSCTVTALTDCEGAKDMWDKLRKESKNLNIQGS**L**FELCGS**G**NGAA
GSLLPALSVLLVSLAALATWLSF-VDELE

Number of amino acids: 149

Molecular weight: 16150.6

Theoretical pI: 4.91

Ex. Co. 1.4

GPI anchor site (GPI SOM program <http://gpi.unibe.ch/>)

Other GPI predictions

361 w__ Spacer__**Hydrophobic domain** Probability
Wild type MDPSAA ¼CRTNYGYSAAPSLHLPPGS**LLASLVPLLLLSLP** 1.0

frag anchor

HMM

Score 1: 21.9999 *** 58

Structure: [GAA] GSLLPALSVLLVSLAALATWLSF

Classification: Highly probably

Score 2: 21.7937 *** 59
Structure: [AAG] SLLPALSVELLVSLAALATWLSF
Classification: Highly probably
Score 3: 20.3648 *** 60
Structure: [AGS] LLPALSVELLVSLAALATWLSF
Classification: Highly probably

TARP2 sequence:

MGLFDRGVQMLLTTVGAFAAFSLMTIAV**GT**DYWLYSRGVCKTKSVSENETSKKNEEVMTHSGLWRTCCLEG
NFKGLCKQIDHFPEDADYEADTAEYFLRAVRASSIFPILSVILLFMGGLCIAASEFYKTRHNIILSAGIFFVSAGLSN
IIGIIVYISANAGDPSKSDSKNSYSYGWSFYFGALSFIIAEMVGVLAHMFIDRHKQLRATARATDYLQASAITR
IPSYRYRYQRRSRSSSRSTEPHSRDASPVGVKGFNTLPSTEISMYTLSRDPLKAATTPTATYNSDRDNSFLQVH
NCIQKDSKDSLHANTANRRTPV

Trx scaffold: (see Chapter 5, figure 5.4)

MVKQIESKTA**FQ**EALDAAGDKLVV**VD**FSA**TW**GGTGGTGGKMIK**PF**HSLEKYSNVIFLEVDVDDSDQ**DA**
SESEVKS**MP**TFQ**FF**KKGQKVGEFSGANKEKLEATINELV

TARP2-loop-2:

GTDYWLYSRGVCKTKSVSENETSKKNEEVMTHSGLWRTCCLEG**NFKGLCKQIDHFPEDADYEADTA**EYFLRA
VRASS

Codon optimised TARP2-Trx construct:

TARP2 loop (yellow highlighted) cysteine residues (red highlighted)

MVKQIESKTA**FQ**EALDAAGDKLVV**VD**FSA**TW**GGT**GT**DYWLYSRGVCKTKSVSENETSKKNEEVMTHSGLW
RTCCLEG**NFKGLCKQIDHFPEDADYEADTA**EYFLRAVRASS**GT**GGKMIK**PF**HSLEKYSNVIFLEVDVDDSDQ**D**
VASESEVKS**MP**TFQ**FF**KKGQKVGEFSGANKEKLEATINELV**LE**HHHHHH

Sequencing analysis:

Blue box highlights TARP2 loop-2 sequence

Appendix C- AMPA receptor-interacting proteins oligos and constructs

```

TARP_synth
fwd .....
rev VRDLDALPYATPALGSSPVVGGRAPPPQGMVHARRWRPTVPRPRGLPPYPRRNKRARSARSGEPDLPHRCRRYRQQPHLWRRCRPRCVRRR
consensus>50 .....

TARP_synth
fwd ..... 1 10 20 30 40 50
rev ..... FVLEGGDIHMVKQIESKTAPQEQALDAAGDKLVVVDPSATWGGTGTDTYWLYSRGVCKTKSVSENE
consensus>50 ..... fvlegdiHMVKQIESKTAPQEQALDAAGDKLVVVDPSATWGGTGTDTYWLYSRGVCKTKSVSENE

TARP_synth
fwd ..... 60 70 80 90 100 110 120 130 140
rev ..... TSKKNBEEVMTHTSGLWRTCCLEGNFKGLCKQIDHFPEDADYEADTAEYFLRAVRASSGTGGKMIKPPFHSLSSEKYSNVI FLEVDDVDDSDV
consensus>50 ..... TSKKNBEEVMTHTSGLWRTCCLEGNFKGLCKQIDHFPEDADYEADTAEYFLRAVRASSGTGGKMIKPPFHSLSSEKYSNVI FLEVDDVDDSDV

TARP_synth
fwd ..... 150 160 170 180
rev ..... ASESEVKSMPTFPQFFKKGQKVGEFS GANKEKLEATINELVLEHHHHHDPAAANKARKBAELAAATAEQLAPLGASKRVLRGFLFKGGTIS
consensus>50 ..... ASESEVKSMPTFPQFFKKGQKVGEFS GANKEKLEATINELVLEHHHHHDPAAANKA.....

TARP_synth
fwd .....
rev GLANGTRPVAAHARRVWVWLRAPLHLPAPRPLLSLSPFSPRSPAPFPVKLI GGSLSGSDLVLYGTSTPKNLIRVMVHV
consensus>50 .....

```

CKAMP44

CKAMP44_18 construct D113-T191

```

MCGPRRLAAPGSSLGSFGRASRGRRGSSSGRAAEPGLGDTMRRVLRLLLGCFLTELCA
RVCRAQERAGHGQLAQLGGVLLLAGGNRSGAASGEASEGAEASDAPTRAPTPDFCRG
YFDVMGQWDPPFNCSGDFIFCCGTCGFRFCCTFKKRRLNQSTCTNYDTPLWLNTGKPP
ARKDDPLHDPTKDKTNLIVYIICGVVAVVMVLVGIFTKLGLEKAHRPQREHMSRALADVM
RPQGH CNTDHMERDLNIVVHVQHYENMDTRTPINN LHATQMNN AVPTSPLLQQM GHP
HSYPNLGQISNPYEQPPGKELNKYASLKAVGSSDGDWAVSTLKSPKADKVND DFYTK
RRHLAELAAKGNLPLHPVRVEDEPRAFSPEHGP AKQNGQKSRTNKMPPHPLAYTSTTNF
KGWDPNEQSLRRQAYS NKGKLGTAETGSSDPLGTRPQHYP PPQPYFITNSKTEVTV

```

Output for 'Sequence'

Name: Sequence Length: 429

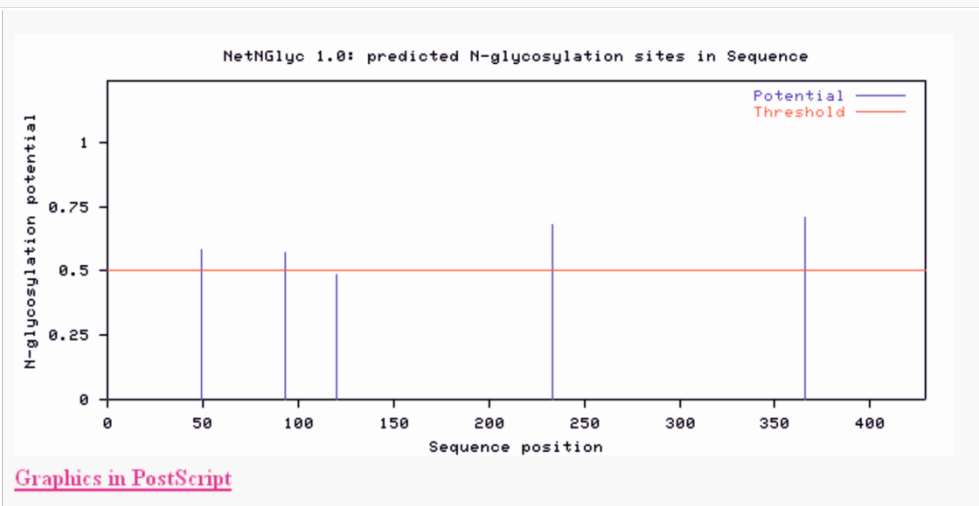
```

EFATHRRVRLRLLLGCFTELCAHRCRAQERSGHGQLAQLGGVLLLTGGNRS GAASGEAGEGVGGSDAPPTRAPTPDSCRG      80
YFDVMGQWDPFFNCSGDFIFCCGTCGFRFCCTFFKKRRLNQSTCTNYDTPLWLNTGKPPARKDDPLHDP TKDKTNLIVYI      160
ICGVVAVMVLVGIFTKLGLEKAHRPQREHMSRALADVMRPQGH CNTDHMERDLNIVVHVQHYENMDSRTPINNLTTQMNN      240
AVPTSPLLQQMGHPHSYPNLGQISNPYEQQPPGKELNKYASLKAVGNSDGDWAVATLKSPKADKVNDDFYAKRRHLAELA      320
VKGNLPLHPVRVEDEPRAFSPHGPQQNGQKSRTNKMPPHPLAYNSTANFKTWDPSDQSLRRQAYGNKGKLGIAESGSC      400
DPLGTRTQHFPPTQPYFITNSKTEVTGVT
.....N.....      80
.....N.....      160
.....N.....      240
.....N.....      320
.....N.....      400
.....      480

```

(Threshold=0.5)

SegName	Position	Potential	Jury agreement	N-Glyc result
Sequence	49 NRS	0.5814	(7/9)	+
Sequence	93 NCS	0.5738	(6/9)	+
Sequence	120 NQST	0.4839	(4/9)	-
Sequence	233 NLTT	0.6782	(8/9)	+
Sequence	366 NSTA	0.7064	(9/9)	++



```

          1      10      20      30      40      50      60      70      80
CKAMP44_18+Trx_pET .....MVKQIESKTAPEALDAAGDKLVVDFSATWGGTGDSCRGYFDVMGQWDPFFNCSSGDFIFCCGTCGFRFCCTFFKKRRLN
fwd_18 .....LVGGDIEMVQIESKTAPEALDAAGDKLVVDFSATWGGTGDSCRGYFDVMGQWDPFFNCSSGDFIFCCGTCGFRFCCTFFKKRRLN
rev_18 .....VLEGGDIEMVQIESKTAPEALDAAGDKLVVDFSATWGGTGDSCRGYFDVMGQWDPFFNCSSGDFIFCCGTCGFRFCCTFFKKRRLN

          90      100     110     120     130     140     150     160
CKAMP44_18+Trx_pET QSTCTNYDTPLWLNTGKPPARKDDPLHDP TKDKTGTGGKMIKPPFHSLSSEKYSNVI PLEVVDVDSQDVASESEVKSMPTFPFFKKGQ
fwd_18 QSTCTNYDTPLWLNTGKPPARKDDPLHDP TKDKTGTGGKMIKPPFHSLSSEKYSNVI PLEVVDVDSQDVASESEVKSMPTFPFFKKGQ
rev_18 QSTCTNYDTPLWLNTGKPPARKDDPLHDP TKDKTGTGGKMIKPPFHSLSSEKYSNVI PLEVVDVDSQDVASESEVKSMPTFPFFKKGQ

        170      180
CKAMP44_18+Trx_pET KYGQPSGANKKKLEATINELV.....KEAELAAATAEQALPLGASKRVLRGLLKGGTISQLANGTRPVAAH
fwd_18 KYGQPSGANKKKLEATINELVLEHHHHHHDPAAANKAR.....
rev_18 KYGQPSGANKKKLEATINELVLEHHHHHHDPAAANKAR.....

```

CKAMP44 Figure Appendix C

CKAMP44 N-linked glycosylation prediction (NetNGlyc) and sequencing analysis of CKAMP44 in Trx scaffold in pET22b

Appendix D- MDGA1 and neuroligin-2 constructs and oligonucleotides

Table D1- MDGA1 and Neuroligin (NL2) construct info, including domain boundaries, oligonucleotides (oligos) used in PCR and cloning strategy. Presence of * indicates that these oligos were used to introduce the required mutation by site directed mutagenesis. Original domain boundary oligos were then used to amplify constructs as required in a second round PCR.

Construct ID	Boundaries	Forward	Reverse	Cloning
chMDGA1	Q19-N925	N/A	N/A	N/A
ratMDGA1_FL	Q19-N925	1268	1272	AgeI/BsiWI
rMDGA1_Ig1-3	Q19-K330	1268	1276	AgeI/BsiWI
NL2	E38-H612	1322	1323	BsrGI/BsiWI
NL2-veSSA	E38-H612 ΔG157-T173	*fwd	*rev	BsrGI/BsiWI

Table D2- Details all MDGA1 and NL2 oligonucleotides used for PCR amplification. All constructs described are correctly numbered starting from the first amino acid residue, typically the start of each proteins native secretion signal.

Oligo number	Oligo ID	Sequence 5'-3'
1268	rMDGA1_Age_fwd	cgcaccgggCAAGGAGTCTACGCTCCAGCCCA
1272	rMDGA1_BsiWI_re	tctataCTCGAGTGATCATTAGTGATGGTGATGGTGCTTGGTACC TTTGTTGGGATCCATCTGCCTCCGGGGA
1276	MDGA1Ig1-3BsiRe	cgccgtacgCTTCAAGGATCGCACCAGCAGGTTGA
1269	Nlg2_AgeI_fwd	CgcaccgggGAGGAGCGCTTCCCGGTGGTGA
1270	Nlg2_KpnI_rev	cgcggtaccGTGCAGGTTGTGCAGGTGGGGCA
1281	Nlg2_fwd_BspE1	gctgcttcggaGAGGAGCGCTTCCCCGTGGTGA
1282	Nlg2_rev_BsiWI	cgccgtacgGTGCAGGTTGTGCAGGTGGGGCA
1322	NL2fwdBspE1_71	gctgcttcggaGAGGAGCGCTTCCCCGTGGTGAACACAGCCTAC
1323	NL2revBsiW_72	cgccgtacgGTGCAGGTTGTGCAGGTGGGGCACGAGCTCCAG
N/A *fwd	NL2-veSSA_fwd	AACCTCTACGTGCCCACTGAGGACGATATCCGGGACTCTGGGAA GAAACC
N/A *rev	NL2-veSSA_rev	GGTTTCTTCCCAGAGTCCCGGATATCGTCCTCAGTGGGCACGTAG AGGTT

Additional vector information for monoVenus generation and pNeo_Sec vector are contained within this appendix. The MonoVenus tag generated was based on the neurospora expression vector pYFP, GenBank accession number EF661030.1. The monoVenus cDNA and amino acid sequence information is detailed below. Site directed mutagenesis of the A206K to enable monomerisation is highlighted in yellow.

GenBank accession number EF661030.1:

ATGGTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGA
CGTAAACGGCCACAAGTTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGA
CCCTGAAGCTGATCTGCACCACCGGCAAGCTGCCCGTGGCCACCTCGTGACCACCCTG
GGCTACGGCCTGCAGTGCTTCGCCCCTACCCCGACCACATGAAGCAGCAGCACTTCTTCAAGTC
CGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGA
CCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGAC
TTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACACTACAACAGCCACAACGTCTA
TATCACCGCCGACAAGCAGAAGAACGGCATCAAGGCCAACTTCAAGATCCGCCACAACATCGAGG
ACGGCGGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTG
CTGCCCCGACAACCACTACCTGAGCTACCACTGCTCAAGCTGAGCAAGACCCCAACGAGAAGCGCGA
TCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGTAA

MVSKGEELFTGVVPIVELDGDVNGHKFSVSGEGEGDATYGKLTCLKICTTGKLPVPWPTLVTTL
GYGLQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKID
FKEDGNILGHKLEYNYSNHNVIYITADKQKNGIKANFKIRHNIEDGGVQLADHYQNTPIGDGPVL
LPDNHLYSYQSKLSKDPNEKRDHMLLEFVTAAGITLGMDELYK-

atggtgagcaagggcgaggagctgttcacgggggtgggtgcccacatcctggtcgagctggac
M V S K G E E L F T G V V P I L V E L D
ggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctac
G D V N G H K F S V S G E G E G D A T Y
ggcaagctgaccctgaagctgatctgcaccacggcaagctgcccgtgcccctggcccacc
G K L T L K L I C T T G K L P V P W P T
ctcgtgaccaccctgggctacggcctgcagtgttcgcccgtaccccgaccacatgaag
L V T T L G Y G L Q C F A R Y P D H M K
cagcagcacttcttcaagtccgcatgcccgaaggctacgtccaggagcgcaccatcttc
Q H D F F K S A M P E G Y V Q E R T I F
ttcaaggacgacggcaactacaagacccgcgagggtgaagttcgagggcgacaccctg
F K D D G N Y K T R A E V K F E G D T L
gtgaaccgcatcgagctgaagggcatcgacttcaaggaggacggcaacatcctggggcac
V N R I E L K G I D F K E D G N I L G H
aagctggagtacaactacaacagccacaacgtctatatcaccgcccagcagaagaac
K L E Y N Y N S H N V Y I T A D K Q K N
ggcatcaaggccaacttcaagatccgccacaacatcgaggacggcggtgagctcgcc
G I K A N F K I R H N I E D G G V Q L A
gaccactaccagcagaacacccccatcgggcgacggccccgtgctgctgcccgacaaccac
D H Y Q Q N T P I G D G P V L L P D N H
tacctgagctaccagttcaagctgagcaaaagaccccaacgagaagcgcgatcacatggtc
Y L S Y Q S K L S K D P N E K R D H M V
ctgctggagttcgtgaccgcccggggatcactctcggcacatggacgagctgtacaagtaa
L L E F V T A A G I T L G M D E L Y K -

pNeoSec vector

This figure has been removed due
to copyright restrictions

Appendix F

This appendix has been removed due to copyright restrictions

The following appendix contains sequence information for the design of the Trx scaffold to attempt expression and secretion of the TARP-2 loop-2 and CKAMP44 cysteine knot from *E. coli* cells.

The *E. coli* TrxA sequence:

ATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGG
GGCGATCCTCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGTGCAAAATGATCGCCCCGATTCTGG
ATGAAATCGCTGACGAATATCAGGGCAAAGTACCGTTGCAAACTGAACATCGATCAAAACCCT
GGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGT
GGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGACGCTAACCTGG
CC

Translation:

1	M	S	D	K	I	I	H	L	T	D	D	S	F	D	T	D	V	L	K	A
1	ATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCG																			
21	D	G	A	I	L	V	D	F	W	A	E	W	C	G	P	C	K	M	I	A
61	GACGGGGCGATCCTCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGTGCAAAATGATCGCC																			
41	P	I	L	D	E	I	A	D	E	Y	Q	G	K	L	T	V	A	K	L	N
121	CCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAAGTACCGTTGCAAACTGAAC																			
61	I	D	Q	N	P	G	T	A	P	K	Y	G	I	R	G	I	P	T	L	L
181	ATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGCTG																			
81	L	F	K	N	G	E	V	A	A	T	K	V	G	A	L	S	K	G	Q	L
241	CTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTG																			
101	K	E	F	L	D	A	N	L	A											
301	AAAGAGTTCCTCGACGCTAACCTGGCC																			

Human Trx (wt):

MVKQIESKTAFQEALDAAGDKLVVDFSATWCGPCKMIKPFHSLSEKYSNVIFLEVDVDDCQDVASECEVK
CMPTFQFFKKGQKVGEFSGANKEKLEATINELV

Human Trx (nocys):

MVKQIESKTAFQEALDAAGDKLVVDFSATWgGPgKMIKPFHSLSEKYSNVIFLEVDVDDsQDVASEsEVK
SMPTFQFFKKGQKVGEFSGANKEKLEATINELV

Human Trx_no_cys, *E. coli* codon optimised and AgeI/KpnI docking site inserted

ATGGTGAAACAGATTGAGAGCAAAACCGCGTTTCAAGAAGCGTTAGATGCAGCTGGCGATAAACT
CGTTGTGGTTGACTTTAGTGCGACTTGGGGTGGAAaccggtCCAGGagggtaccggaggcAAAATGA
TCAAACCGTTCTTTTATTCTTGTGCGAAAAGTATTGCAACGTGATCTTTCTGGAAGTTGACGTA
GACGATAGCCAAGATGTCGCAAGTGAATCAGAAGTCAAAAGCATGCCGACCTTCCAGTTCTTCAA

GAAAGGCCAGAAAGTAGGCGAGTTTTCTGGAGCCAATAAGGAGAACTGGAAGCCACGATTAACG
AACTGGTG

Translation:

```
1  M V K Q I E S K T A F Q E A L D A A G D
1  ATGGTGAAACAGATTGAGAGCAAAACCGCGTTTCAAGAAGCGTTAGATGCAGCTGGCGAT

21  K L V V V D F S A T W G G T G P G G T G
61  AAACGTTGTGGTTGACTTTAGTGCGACTTGGGGTGAACCGGTCCAGGAGGTACCGGA

41  G K M I K P F F H S L S E K Y S N V I F
121 GGCAAAATGATCAAACCGTTCTTTCATTCCTTGTCGGAAAAGTATTCGAACGTGATCTTT

61  L E V D V D D S Q D V A S E S E V K S M
181 CTGGAAGTTGACGTAGACGATAGCCAAGATGTCGCAAGTGAATCAGAAGTCAAAGCATG

81  P T F Q F F K K G Q K V G E F S G A N K
241 CCGACCTTCCAGTTCTTCAAGAAAGGCCAGAAAGTAGGCGAGTTTTCTGGAGCCAATAAG

101 E K L E A T I N E L V
301 GAGAAACTGGAAGCCACGATTAACGAACTGGTG
```

This appendix contains information for protein constructs required to generate artificial perineuronal nets (HAS3, HAPLN1, Brevican or GPC4) as described in Kwok *et al.* 2011 as well as the generation of a C-terminal AP-fusion tag.

Table H-1 Construct required to generate the perineuronal nets (PNN).

Domain boundaries, oligos used in PCR and their cloning strategies are included.

Construct ID	Boundaries	Forward	Reverse	Digest	Cloning
AP-tag	I42-Y531	1271	1262	AgeI/BsiWI	pHLsec AgeI/Acc65I
GPCR_WT	A23-K521	1263	1264	AgeI/BsiWI	pHLsec_AP AgeI/Acc65I
GPCR_sugarless	A23-K521A	1263	1265	AgeI/BsiWI	pHLsec_AP AgeI/Acc65I
Brevican	M1-P911	1266	1267	EcoR1/BsiWI	pHLsec_AP EcoRI/Acc65I
HAS3_veSS+VSV	M1-Q281	1273	1278	EcoR1/BsiWI	pHLsec_AP EcoRI/Acc65I
HAPLN1	D16-N354	1274	1275	AgeI/KpnI	pHLsec AgeI/KpnI

Table H-2 Details all oligos used to generate the various PNN components by PCR amplification.

All constructs described are correctly numbered starting from the first amino acid residue, typically the start of each proteins native secretion signal.

Oligo number	Oligo ID	Sequence 5'-3'
1262	BsiWI_AP_rev	cgccgtacgATAACCCGGGTGCGCGGCGTC
1263	GPC4_AgeI_fwd	Cgcaccggtgaggagcccaagtcgaaaagt
1264	GPC4_WT_BsiWI_rv	cgccgtacgCTTCCCAGAATGGTCAGTGGCGTTG
1265	GPC4_sug_BsiW_rv	cgccgtacgCTTCCCAGcATGGTCAGTGGCGTTG
1266	Brev_EcoR1_fwd	GCCGAATTCGCCACCATGGCCCAGCTGTTCTGCCCC T
1267	Brev_BsiWI_rev	cgccgtacgGGGACCTGGCATGGGGCTGGA
1271	Age_xx_KpnI_AP_f	cgcaccggtggaggctcaggtaccggaggctcaATCATCCAGTTGAG GAGGAGAACC
1273	HAS3EcoR1_fwd	GCCGAATTCGCCACCATGCCGGTGCAGCTGACGACA G
1274	HAPLN1_fwd_ageI	CgcaccggtGATCATCTTTCAGACAACTATACTCTGGATC ATGAC
1275	HAPLN1_KpnI_rev	cgcggtagcGTTGTATGCTCTGAAGCAGTAGACACCATAC A
1278	HAS3BsiVSVKpn_r	cgccgtacgcttcccagccggttcattcaatgtccgtgtaggtacctccCTGCTC TCTAGAAACGTGGCGTTGGT
1305	GPC_fwd_NheI	AAGCTGGCTAGCCACCATGGAGAC
1306	Rev_GPC 4wt_Hind	tagtagAAGCTTtCCCAGAATGGTCAGTGGCGTTG
1307	Rev_GPC4noHS_Hi	tagtagAAGCTTtCCCAGcATGGTCAGTGGCGTTG
1308	APTAG5_rev_seq	TGCCTCGCGGTTCCAGAAGTC
1309	LINK_fwd_Eco	TCCTGAATTCGCCACCatgAAGAGTCTACTTCTTCTGGT GCTGATTTC

Generation of the pHLsec-AP vector for standard use

The AP sequence was amplified from the Neurocan-AP plasmid by PCR. Oligonucleotide design enabled the AP-tag construct (I42-Y531) to be cut with AgeI/ BsiWI and ligated into the pHLsec plasmid cut with AgeI/Acc65I. Introduction of a KpnI cloning site N-terminally enabled its future use as a c-terminal tag to any construct inserted into pHLsec via AgeI/KpnI.

Forward (AgeI-xxx-KpnI-AP_fwd):

cgccacgggtggaggctcagggtaccggaggctcaATCATCCCAGTTGAGGAGGAGAACC

Reverse (BsiWI-AP_rev):

cgcggtacgATAACCCGGGTGCGCGGCGTC

Translation:

SWLATMETDTLLLWVLLLVPGSTGDAAQPARRARRTYEAYVRSSGIIPVEEENPDFWNR
EAAEALGAACKLQPAQTAAKNLIIFLGDMGVSTVTAARILKGQKKDKLGPEIPLAMDRF
PYVALSKTYNVDKHVPDSGATATAYLCGVKGNFQITIGLSAAARFNQCNTTRGNEVISVMN
RAKKAGKSVGVVTTTRVQHASPAGTYAHTVNRNWYSDADVPASARQEGCQDIATQLISM
DIDVILGGGRKYMFPMTGPDPEYPDDYSQGGTRLDGKNLVQEWLAKRQGARYVWNRTLM
QASLDPSVTHLMGLFEPGDMKYEIHRDSTLDPSLMEMTEAALRLLSRNPRGFFLFVEGGR
IDHGHESRAYRALTETIMFDDAIERAGQLTSEEDTSLSLVTADHSHVFSFGGYPLRGSSI
FGLAPGKARDRKAYTVLLYGNPGYVLKDGARPDVTESESGSPEYRQQSAVPLDEETHAG
EDVAVFARGPQAHLVHGVQEQTFAHVMAFAACLEPYTACDLAPPAGTTDAAHPGYLEEA
LSLEGPEQKLISEDELNSAVDHHHHHH - V - T

Hyaluronan synthase 3 (HAS3)

Translation (281aa)

MPVQLTTALRVVGTSLFALAVLGGILAAVVTGYQFIHTEKHYSFGLYGAILGLHLLIQS
LFAFLEHRRMRAGQALKLPSRRGSVALCIAAYQEDPDYLRKCLRSAQRISFPDLKVV
VVDGNRQEDAYMLDIFHEVLGGTEQAGFFVWRSNFHEAGEGETEASLQEGMDRVRDVRA
STFSCIMQKWGGKREVMYATFKAALGDSVDYIQVCDSDTVLDPACTIEMLRVLEEDPQVGG
VGGDVQPPGKMAVEDDQVQAAQVRATEAWSVHQRHVSREQ

Hyaluronan and proteoglycan link protein 1 (HAPLN1)

Translation (354aa):

MKSLLLVLISICWADHLSDNYTLHDRAIHIQAENGPHLLVEAEQAKVFSHRGGNVTL
CKFYRDPTAFSGSIHKIRIKWTKLTSYDLKEVDVFSVMGYHKKTYGGYQGRVFLKGGSDS
DASLVITDLTLEDYGRYKCEVIEGLEDDTVVVALDLQGVVFPYFRLGRYNLNFHEAQQ
CLDQDAVIASFDQLYDAWRGGLDWCNAGWLSGDSVQYPITKPREPCGGQNTVPGVRNYGF
WDKDKSRYDVFCFTSNFNGRFYLLIHTPKLTYDVAVQACLNDGAQIAKVGQIFAAWKILG
YDRCDAGWLADGSVRYPISRPRRRCSPTAAVRVFGFPDKKHKLGYVYCFRAYN

Brevican

The human Brevican construct (M1-P911) based on IRATp970D0944D was cloned into pHlsec-AP. PCR incorporation of EcoRI/BsiWI restriction sites allowed its native secretion signal to be utilized.

Translation:

MAQLFLPLLAALVLAQAPAALADVLEGDSSSEDRAFRVRIAGDAPLQGVLGALTIPCHVH
YLRPPPSRRAVLGSPRVKWTFLSRGREAEVLVARGVRVKVNEAYRFRVALPAYPASLTDV
SLALSELRPNDSGIYRCEVQHGIIDSSDAVEVKVKGVVFLYREGSARYAFSFGAQEACA
RIGAHIAITPEQLYAAYLGGYEQCDAGWLSQTVRYPIQTPREACYGDMDFGVRNYGVV
DPDDLVDVYCYAEDLNGELFLGDPPEKLTLEEARAYCQERGAETATTGQLYAAWDGGLDH
CSPGWLADGSVRYPITPSQRCGGGLPGVKTLFLFPNQTFGPNKHSRNFVYCFRDSAQPS
AIPASNPASNPASDGLEAIVTVTETLEELQLPQEATESERGAIIYSIPIMEDGGGGSST
PEDPAEAPRTLLEFETQSMVPPTGFSEEEGKALEEEEEKYEDEEEKEEEEEEEEEVEDEALW
AWPSELSSPGPEASLPTEPAAQEKSLSQAPARAVLQPGASPLPDGESEASRPPRVHGPPT
ETLPTPRERNLASPSPSTLVEAREVGEATGGPELSGVPRGESEETGSSEGAPELPLPATRA
PEGTRELEAPSEDNSGRTAPAGTSVQAQPVLPDTSASRGGVAVVPASGDCVPSPCHNGGT
CLEEEEGVRCLCLPGYGGDLCDVGLRFCNPGWDAFQGACYKHFSTRRSWEEAETQCRMYG
AHLASISTPEEQDFINNRYREYQWIGLNDRTIEGDFLWSDGVPLLYENWNPGQPDYSYFLS
GENCVVMVWHDQGGQWSDVPCNYHLSYTKMGLVSCGPPPELPLAQVFGRPRLRYEVDTVL
RYRCREGLAQRNLPLIRCQENGRWEAPQISCVPRRPARALHPEEDPEGRQGRLLGRWKAL
LIPPSSPMPPGP -

Rat glypican 4

Rat glypican 4 (GPC4) was a kind gift of Nicola Allen (location). A rat glypican 4 mutant with the C-terminal glycosylation sites mutated was also a gift of Nicola Allen.

Translation:

MARLGLLALLCTLAALSASLLAAEPKSKSCSEVRRLYVSKGFNKNDAPLYEINGDHLKIC
SQDYTCCSQEMEEKYSLQSKDDFQTVVSEQCNHLQAIASFYKKFDEFFKELLENAEKSL
NDMFVKTYGHLYMQNSELFKDLFVELKRYVAGNVNLEEMLNDFWARLLERMFRLVNSQY
HFTDEYLECVSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVARDVVRVSVVNP
TAQCTQALLKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLDLEWNNFIDAMLMVAE
RLEGPFNIESVMDPIDVKISDAIMNQDNSVQVSQKVFQGCPPKPLPAGRISRSISESA
FSARFRPYHPEQRPTTAAGTSLDRLVTDVKEKLQAKKFWSLPSTVCNDERMAAGNENE
DDCWNGKSKSRYLFAVTGNGLANQGNNEVQVDTSKPDILILRQIMALRVMTSKMKNAYN
GNDVDFFDISDESSGEGSGSGCEYQQCPSEFEYNATDHSGKSANEKADSAGGAPAEAKPY
LLAALCILFLAVQGEWR -