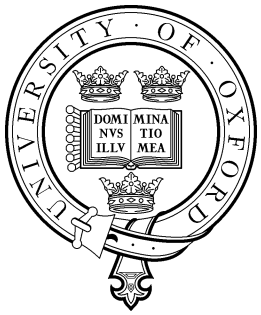
Structural Studies of Class A and D Plexins



Youxin Kong

Merton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2015

Acknowledgements

First and foremost, I would like to thank my supervisor Prof. E. Yvonne Jones for her guidance and support throughout the course my D.Phil. I would also like to thank Dr. Tomas Malinauskas who has been helping me since my first day in lab. Special thanks also go to Dr. Bert. J. C. Janssen, who started the semaphorin–plexin project, and Dr. Sergi Padilla-Parra for teaching me light microscopy.

I cannot thank everyone in STRUBI enough for their kindness and support. They have made my four years in Oxford exceptional. I would like to specially acknowledge Dr. Karl Harlos, Dr. Robert Gilbert, Dr. Yuguang Zhao, Ms. Weixian Lu, Dr. Jingshan Ren, Dr. Charlotte Coles, Dr. Kamel El Omari and Dr. Rainer Kauffman for their help. I would also like to thank all the invaluable fellow students in STRUBI, Santiago, Alex, Tao-hsin, Hanwen, Tao, Ling and Xiangxi, Max etc. I would also like to acknowledge Merton College for having me as part of its continuing legacy.

Finally, I want to thank my family who are the pillars of my life. Special thanks go to Lukas Witkowski, Arne Bruyneel and all my friends for helping with my work and for making these four years enjoyable.

Abstract

Semaphorin signalling via the plexin receptors, important in the nervous, cardiovascular, immune and skeletal systems, requires cytoplasmic domain dimerisation but its extracellular regulation and activation mechanisms remain unclear. In the thesis I conducted structural characterisation for two particular classes of verterbrate plexins, PlxnAs and PlxnD1.

Here I present crystal structures of PlxnA1, PlxnA2 and PlxnA4 full ectodomains. The first nine domains of PlxnAs form a 230 Å long ring-like stalk from which the tenth C-terminal domain points away. In agreement, negative stain electron microscopy of the PlxnA1 ectodomain reveals a predominant ring-like conformation and a minor twisted-open conformation of the stalk. All PlxnA crystal structures also make intermolecular "head-to-stalk" (domain 1 to domain 4-5) interactions, which are confirmed by structure-guided biophysical assays and live cell fluorescence microscopy. Functional assays performed in COS-7 cell and dentate gyrus growth cones revealed an autoinhibitory mechanism for PlxnAs based on this head-to-stalk interface. My results reveal a two-fold role for the PlxnA ectodomains: imposing a presignalling autoinhibitory separation of the PlxnA membrane proximal do-

mains via a head-to-stalk, intermolecular interaction and supporting PlxnA cytoplasmic domain dimerisation post semaphorin-binding.

For PlxnD1, I present high-resolution crystal structures of its semaphorin-binding segment (domains 1-2) as well as an initial structural model of its complete ectodomain (domains 1-10). The sema (domain 1) of PlxnD1 has the classic fold of a plexin sema domain as well as distinct insertions potentially important for its ligand-binding specificity. The PlxnD1 ectodomain adopts a closed-ring conformation in its crystal structure, in which the tail (domain 9) interacts with the sema domain head. In-solution experiments suggest structural variability and potential weak oligomerization for the PlxnD1 ectodomain. These new structures shed light on the unique and previously uncharacterised structural basis for PlxnD1 signalling.

CONTENTS

List of Figures

ABBREVIATIONS

AUC	analytical centrifugation.
BIS	2-bisamino]-2–1,3-propanediol.
cDNA	complimentary deoxyribonucleic acid.
CFP	cyan fluorescent protein.
CNS	central nervous system.
COS-7	CV-1 in Origin with SV40 genes-7.
CSPG	chondroitin sulfate proteoglycan.
DG	dentate gyrus.
EGFR	epidermal growth factor receptor.
EM	electron microscopy.
FLIM	fluorescence lifetime imaging microscopy.
FRET	Föster energy transfer.
GAP	GTPase-activating protein.
GDP	guanosine diphosphate.
GTP	guanosine triphosphate.
HEK	human embryonic kidney.
HSPG	heparin sulfate proteoglycan.
Ig	immunoglobin.
IPT	Ig-like domain shared by plexins and transcription factors.
Kd	dissociation equilibrium constant.
MALS	multi-angle light scattering.
MES	2-(N-morpholino)ethanesulfonic acid.
MR	molecular replacement.
mTFP1	monomeric teal fluorescent protein 1.

NA	numerical aperture.
NAG	N-Acetylglucosamine.
NDSB	non-detergent sulfobetaines.
Nrp	neuropilins.
PALM	photoactivated localization microscopy.
PCR	polymerase chain reaction.
PDB	protein data bank.
PDZ	post synaptic density protein, drosophila disc large tumor suppressor, and zonula occludens-1 protein.
PFA	paraformaldehyde.
Plxn	plexins.
PSI	plexin-semaphorin-integrin.
Pt	platinum.
RBD	Rho-GTPse binding domain.
ROI	region of interest.
rpm	revolutions per minute.
SAXS	small angle X-ray scattering.
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis.
Sema	semaphorins.
SPR	surface plasmon resonance.
STORM	Stochastic optical reconstruction microscopy.
TCSPC	time-correlated single photon counting.
TRIS	trisaminomethane.
VEGF	vascular endothelial growth factor.
YFP	yellow fluorescent protein.

CHAPTER 1

INTRODUCTION

1.1 Overview on Ligand–Receptor signalling

One major characteristics of living organisms is their ability to process and share information. From the simplest unicellular organisms to higher mammals, each individual cell making up the organism is able to respond to chemical or physical stimulants in the environment. For multicellular organisms, the ability of cells to work in synergy and orchestrate complex biological processes depends on cell–cell communication. In general, cells can communicate with each other directly through physical contacts or indirectly through exchanging signalling molecules distributed in the environment. These signalling molecules, encompassing, metal ions, chemical compounds, peptides and proteins, can be registered by receptors on the cell surface or transported across the

membrane to reach receptors in the cytosol .

The most common type of cell–cell communication is ligand–receptor signalling via the cell-surface receptors, in which molecular cues deployed in the environment (the ligands) by signal-generating cells, are recognized by their cognate transmembrane receptors and transduced into distinct functions in the signal-receiving cells. Typically, ligand-binding in the extracellular segments (ectodomains) of cell-surface receptors leads to rearrangements in the overall receptor architectures. These rearrangements, which are often conformational changes or alternations in the oligomerisation states, may trigger changes in the receptor-related downstream signalling pathways and eventually, initiate particular functional outcomes. Recent studies highlighted that the ligand–receptor signalling systems are subjected to multi-level and spatial-temporal regulations, in which the cell-surface receptors act as key control points (2–5). It has now become clear that the cell-surface receptors, diverse in structures and in functions, do not merely serve as simple on-off switches for certain ligand-elicited responses, but rather act as complex modules to gate, fine-tune, diversify and integrate signals (6–8).

1.2 The Single-Pass Transmembrane (TM1) Receptors

1.2.1 Overview

Cell-surface receptors can be categorised by the structures of their transmembrane regions. The single-pass transmembrane (TM1) receptors, characterised by a single trans-

membrane alpha-helix, constitute the most common and diverse category of receptors in vertebrates. signalling via TM1 receptors governs a wide variety of cellular functions, encompassing cell proliferation, differentiation, motility, and survival. Due to their important functions in the neural and immune systems, the TM1 receptors are particularly abundant in higher mammals, amounting to more than 1280 examples in humans (4). The most important and well-characterised TM1 receptor families include receptor tyrosine kinases and phosphatases, immune receptors, as well as receptors for cytokines, chemokines and cell guidance cues (8–11). The extracellular segments of TM1 receptors, though varied in architectures, are made up of multiple structural domains, many of which glycosylated and/or lipid-modified. The relatively shorter cytoplasmic segments of TM1 receptors, however, are more pleiotropic as they may comprise large structural motifs or short disordered loops. The binding between TM1 receptor ectodomains and their cognate ligands is highly specific, though binding affinities may be disparate. In contrast, the cytoplasmic segments of TM1 receptors are often promiscuous in binding. The transmembrane helices of TM1 receptors, have recently been found to not only provide connections for the extracellular and cytoplasmic regions, but also regulate receptor-association during signalling (8, 12).

1.2.2 The Signalling Mechanisms of TM1 Receptors

Due to their essential roles in human physiology and implications in various diseases, the TM1 receptors have been placed center stage for biomedical studies. The TM1 receptors are typically activated in an outside-in manner, in which extracellular ligand-

binding induces concurrent structural rearrangements in the extracellular and cytoplasmic regions and thus, triggers downstream signalling. However, several TM1 receptors such as the integrins may also signal in an inside-out manner, in which substrate-binding in the receptor cytoplasmic segments elicits structural and functional changes in the ectodomains (13, 14). The driving force for the common, outside-in receptor activation is most often ligand-binding, although for some receptors such as receptor tyrosine kinase HER2 (or ErbB2) ligand-independent activation may also occur (15). From the studies of many receptor systems a central paradigm of TM1 receptor signalling has emerged. In this paradigm, the activation of TM1 receptors requires ligand-induced receptor dimerisation followed by clustering (3, 9, 16). In some cases the dimerisation and clustering of receptors is aided by coreceptors or cofactors, forming hetero-multimeric complexes on the cell surface (17–19).

Signalling via TM1 receptors is tightly controlled and exquisitely balanced between inhibition and activation. For instance, the TM1 receptors can be regulated by post translational modifications, proteolytic cleavages, trafficking and compartmentalization pre-signalling (2, 4, 20). In particular, some TM1 receptors have been found to also to impose self-regulations, such as pre-associations or autoinhibition (13, 21–23). Post-signalling, the receptors and ligand-receptor complexes are also subjected to multifaceted regulations including endocytosis, degradation and recycling (2, 20, 24). The precise control of TM1 receptors can diverify and fine-tune signal transduction (20).

1.2.3 The Challenges in Studying TM1 Receptors

Although major progress has made regarding the activation of many TM1 receptors, gaining insights into their regulatory mechanisms poses major challenges even for relatively well-characterised systems such as the epidermal growth factor receptor (EGFR) family (5). The major obstacles include obtaining high-resolution structures for the often large and heavily glycosylated receptor ectodomains, let alone for the full-length transmembrane receptors. Another challenge is analysing the highly dynamic TM1 receptors on the cell surface, a process indispensable for scrutinizing any structure-related hypothesis. Thanks to recent advances in crystallography, transmission electron microscopy and fluorescence light microscopy, comprehensive structural studies on the regulatory mechanisms of TM1 receptors have been made possible (25–27). However, for most TM1 receptor systems we lack a complete picture of how signalling is activated, regulated and diversified.

1.3 The Axon Guidance Cue Receptors

1.3.1 Overview on Axon Guidance

During the lifespan of a higher organism, millions of cells need to migrate to appropriate locations and establish contacts with other cells. The movements of cells during the formation and maintenance of complex organ systems are dauntingly intricate and thus

could not be entirely pre-programed into intracellular machineries. Instead, the motility of cells are largely guided by external molecular cues deployed in the environment (28). These particular molecular cues, known as the cell guidance cues, can be diffusing in the environment or attaching to the surfaces of other cells and extracellular matrices. Upon registration of their cognate cues, the cell guidance cue receptors induce discrete and often localised rearrangements in the cytoskeleton and thus, direct the navigation of cells.

The cell guidance phenomenon is best exemplified by axon guidance, in which the growing axons migrate and assemble into a stereotypical pattern directed by guidance cues and their receptors. The wiring of axonal circuits can be broken down into an ensemble of individual guidance events, in which the axons steer towards more permissive sites according to the lay-out of attractive and repulsive cues (7, 10, 21). A growing axon probes for guidance cues with a highly dynamic structure on its tip, termed the growth cone, which is teeming with guidance cue receptors on its extensive surface. The dramatic network of axons, which is essentially an *in vivo* record of axon guidance, provides us with an excellent subject for the molecular studies of cell guidance.

1.3.2 The Four Classic Families of Axon Guidance Cue Receptors

Extensive studies in the past decades have uncovered a variety of guidance cues, among which four protein types, the semaphorins, netrins, slits and ephrins, stand out as the most abundant and essential (7, 21). The major receptors for these four guidance cues

are the plexins, DCC (Deleted in Colorectal Cancer), Robos (Roundabouts) and Ephs, respectively. These ligand–receptor signalling systems, which were first discovered from the genetic and molecular dissection of axon guidance in several invertebrates and later found to be also conserved through the vertebrates, are since known as the four classic axon guidance cue–receptor families. The four classic families of axon guidance receptors are typical TM1 receptors with large, multidomain extracellular segments and relatively smaller cytoplasmic domains. Although they vary in mechanisms, these four receptors translate guidance cue-binding in extracellular space into discrete changes in the cytoskeleton. The four classic axon guidance cue–receptor systems have also been found to govern many other important cellular functions such as cell morphogenesis, adhesion and survival (21). Axon guidance was later found to also involve many other signalling molecules and TM1 receptors, including morphogens such BMPs (bone morphogenic proteins), Wnts, Hedgehogs and neurogenins, several fibroblast and nerve growth factors, RPTPs (receptor protein tyrosine phosphatases), as well as CAMs (cell-adhesion molecules) (29–31). Although divergent in structures and mechanisms of action, these newly characterised proteins elicit similar functional outcomes, i.e. axon attraction or repulsion, as the four classic axon guidance cues.

1.3.3 The Activation of Axon Guidance Cue Receptors

Recent studies suggest that the four classic axon guidance cue–receptor systems, semaphorin–plexin, netrin–DCC, slit–Robo and ephrin–Eph, may follow similar modes of activation. The key to receptor activation appears to be an increase in their oligomeri-

sation states rather than substantial changes in their ectodomain structures. For the best-characterised ehprin–Eph system, it has been proposed that the initial formation of 2:2 ephrin–Eph heterotetramers may lead to higher-order clustering (32, 33), although recent discoveries indicate that the Eph receptors, pre-associating on the cell-surface via *cis*-interactions, nucleate into large signalling clusters upon the formation of a limited number of ephrin–Eph complexes (25, 34). Recent studies regarding the netrin–DCC system also indicated that the DCCs are activated via netrin-induced dimerisation and potential oligomerisation (31, 35). This monomer-to-dimer-oligomer paradigm of receptor activation is echoed by the recently published 2:2 semaphorin–plexin recognition complex (36, 37), which will be discussed in detail in later sections. Although the slit–Robo system is less well-characterised, it has been suggested that monomer-dimer transition is essential to Robo activation (38).

1.3.4 The Regulation of Axon Guidance Cue Receptors

Remarkably, the wiring of neural circuits is controlled by a disproportionally small number of guidance cues. These guidance cues are also repeatedly used during multiple stages of embryonic and adult tissue development in many other organ systems (39, 40). This observation indicates that the guidance cue receptors must diversify their repertoires and be subjected to spatial-temporal regulations (20, 39, 41). The four classic axon guidance cue receptors, and indeed most other guidance cue receptors, can form holoreceptors together with different cognate coreceptors or modulatory cofactors such as heparin sulfate proteoglycan (HSPG) and chondroitin sulfate proteoglycan (CSPG)

(10, 39). The plethora of holoreceptors greatly expanded the receptor repertoires and often have boosted specificity for cognate cues. All the pre- and post-signalling regulatory mechanisms for TM1 receptors (described in section 1.2.3) can also be applied to the axon guidance cue receptors. For a higher level of complexity enter the non-canonical signalling modes such as ligand–receptor *cis*-activation or ligand reverse signalling, which are exemplified by the semaphorin–plexin and ephrin–Eph systems(42, 43). Increasing evidence also suggests that different guidance cue–receptor systems can cross-interact, adding another dimension to the diversification of signalling.

One of the principal control points for guidance cue receptors is their ectodomains. One can envision that the large and multidomain extracellular segments may possess distinct structural variabilities and interaction sites serving as bases for regulatory mechanisms. Indeed, recent studies have revealed that the DCC and plexin ectodomains, in particular the non-ligand-binding segments, may adopt distinct conformations important for the inhibition and activation of receptor functions (22, 36). Moreover, the ectodomains of these receptors may harbour multiple binding sites for modulatory co-receptors and cofactors (44–46). In particular, for the ephrin–Eph system it has been demonstrated that the ectodomain-based receptor clustering pre- and post-signalling is essential for the subtype-specific regulation of signalling (47, 48). Novel regulatory mechanisms targeting the ectodomains of major axon guidance cue receptors are still surfacing.

1.4 The Semaphorin–Plexin Signalling System

1.4.1 Overview

The semaphorins (Sema), aptly named after their role as molecular semaphores in directing growing axons, constitute the largest and most diverse guidance cue family. The major receptors for semaphorins are the plexins (Plxn), which stand out amongst other guidance cue receptors for their diversity as well as wide-spread contributions in cell morphogenesis, motility, adhesion, and survival across organ systems. Several coreceptors and cofactors have also been found to modulate the semaphorin–plexin system, including coreceptor neuropilins (Nrp), cell adhesion molecule (CAM) L1, as well as proteoglycan HSPGs and CSPGs. The pleiotropic nature of the semaphorin–plexin system allow for specific and diversified ligand–receptor interactions well suited for complex guidance events.

1.4.2 The Semaphorin Family

The semaphorins are a family of secreted or cell surface-attached proteins encompassing more than twenty members (17). Based on structural homology the semaphorins are categorised into eight classes, class 1-7 and V (Sema1-7 and V). Class 1 and 2 semaphorins are found in invertebrates and class 3 to 7 in vertebrates (Fig. 1.1). A number of viral mimics for the sempahorins have also been identified and categorised into class

V. Viral semaphorins are mimics of Sema7A produced by pox- and herpesviruses which have been suggested to affect immune evasion and act on several immune cells. The most well-known viral semaphorin, A39R, has been linked to cytoskeleton rearrangement and phagocytosis of dendritic cells (DCs) upon binding its receptor, PlxnC1 (49). Among all vertebrate semaphorins only class 3 is secreted, while class 4, 5, and 6 are TM1 transmembrane proteins and class 7 is GPI-anchored to the cell membrane. The structural hallmark of the semaphorins is their ~500-residue, N-terminal sema (short for semaphorin) domain comprising seven β-strand blades arrayed consecutively around a central axis. In all semaphorins apart from the viral mimics the sema domain is immediately followed by a small cysteine-knot of ~50 amino acids, the plexin-semaphorin-integrin (PSI) domain. Although the sema-PSI combination constitutes the extracellular segments for class 1 and 6 semaphorins, in all other vertebrate classes the PSI domain is immediately followed by an immunoglobin (Ig)-like domain and in particular, an additional series of seven thrombospondin type 1 repeats for class 7 semaphorins. The viral mimic (A39R), however, consists of only the sema domain. The cytoplasmic regions of the transmembrane semaphorins vary in length and predicted structures. In particular for the vertebrate classes, the cytoplasmic segments of class 4 semaphorins contain a putative binding site for post synaptic density protein, drosophila disc large tumor suppressor, and zonula occludens-1 protein (PDZ) domains. For class 5 and 6 semaphorins the cytoplasmic segments are relatively short and predicted to be random coils.

1.4.3 The Plexin Family

Plexins are large TM1 receptors chiefly responsible for semaphorin recognition (17). Based on structural homology the plexins are categorised into six classes, including class a and b in invertebrates and class A-D in vertebrates (PlxnA-D) (Fig. 1.1). Among the four classes of vertebrate plexins, nine subtypes are found in humans, including PlxnA1-A4, PlxnB1-B3, PlxnC1, and PlxnD1. The most prominent structural component of the plexin extracellular segments is an N-terminal sema domain homologous to those in the semaphorins. In class A, B and D plexins the sema domain is followed by three pairs of alternating PSI (plexin-semaphorin-integrin) and Ig-like domain shared by plexins and transcription factors (IPT) domains. Subsequent to the three PSI-IPT repeats are three additional IPT domains (IPT4-IPT6). Uniquely, class B plexins harbour an additional mucin-like insertion between the second pair of PSI-IPT repeat (PSI2-IPT2) as well as a furin cleavage site at the C-terminal. Compared to the ten-domain class A and D and the eleven-domain class B plexins, the seven-domain class C plexin (PlxnC1) is much shorter, consisting of only two PSI-IPT repeats and two additional IPT domains. The plexin sema domain is generally referred to as the "head" and the string of PSI and IPT domains the "stalk". Little is known on the function of plexin transmembrane helices, although recent simulation-based studies suggest that the helices may have dimerization or even oligomerization tendency in different plexin families (63, 64). In contrast to the class-specific ectodomains, the plexin cytoplasmic domains are highly conserved. All plexin cytoplasmic segments consist of a juxtamembrane segment predicted to be

helical, and a GTPase-activating protein (GAP) with a Rho-GTPse binding domain (RBD) insert. The class B plexins contain an additional putative PDZ domain binding site at the C-terminal. The ability for plexins to directly hydrolyze small GTPases via their GAP domains is unique among all TM1 receptors. Recent studies indicate that despite sharing the topology of a Ras-GAP, the GAP domain in plexins preferentially activates Rap, a distinct sub-family of the Ras homologues (50, 51).

1.4.4 The Neuropilin Family

Since more than twenty semaphorins funnel down into interactions with eleven plexins, the plexin receptors must to broaden their repertoires to discern multiple ligands. For this purpose, certain types of plexins form holoreceptors with coreceptors deployed on the same cell surface. In vertebrates the main coreceptors for plexins are the neuropilins, which are TM1 proteins subdivided into class 1 and 2 (Nrp1-2) (17, 52). The extracellular segment of neuropilins consist of two N-terminal CUB (C1r/C1s, Uegf, Bmp1) domains (a1 and a2), two coagulation factor V/VIII homology domains (b1 and b2), followed by a long disordered region of ~60 amino acids and a membrane proximal MAM (meprin/A5/mu) domain (c1). The transmembrane segments of neuropilins are predicted to be helical and the cytoplasmic regions short (40-43 amino acids) and disordered. Proteolytic processing of both Nrp1 and Nrp2 can yield different variants, in which secreted forms of part of the ectodomains may act as natural inhibitors for plexins (53, 54). Signalling of all the secreted class 3 semaphorins except for Sema3E via the class A and D plexins requires neuropilins (17, 55).

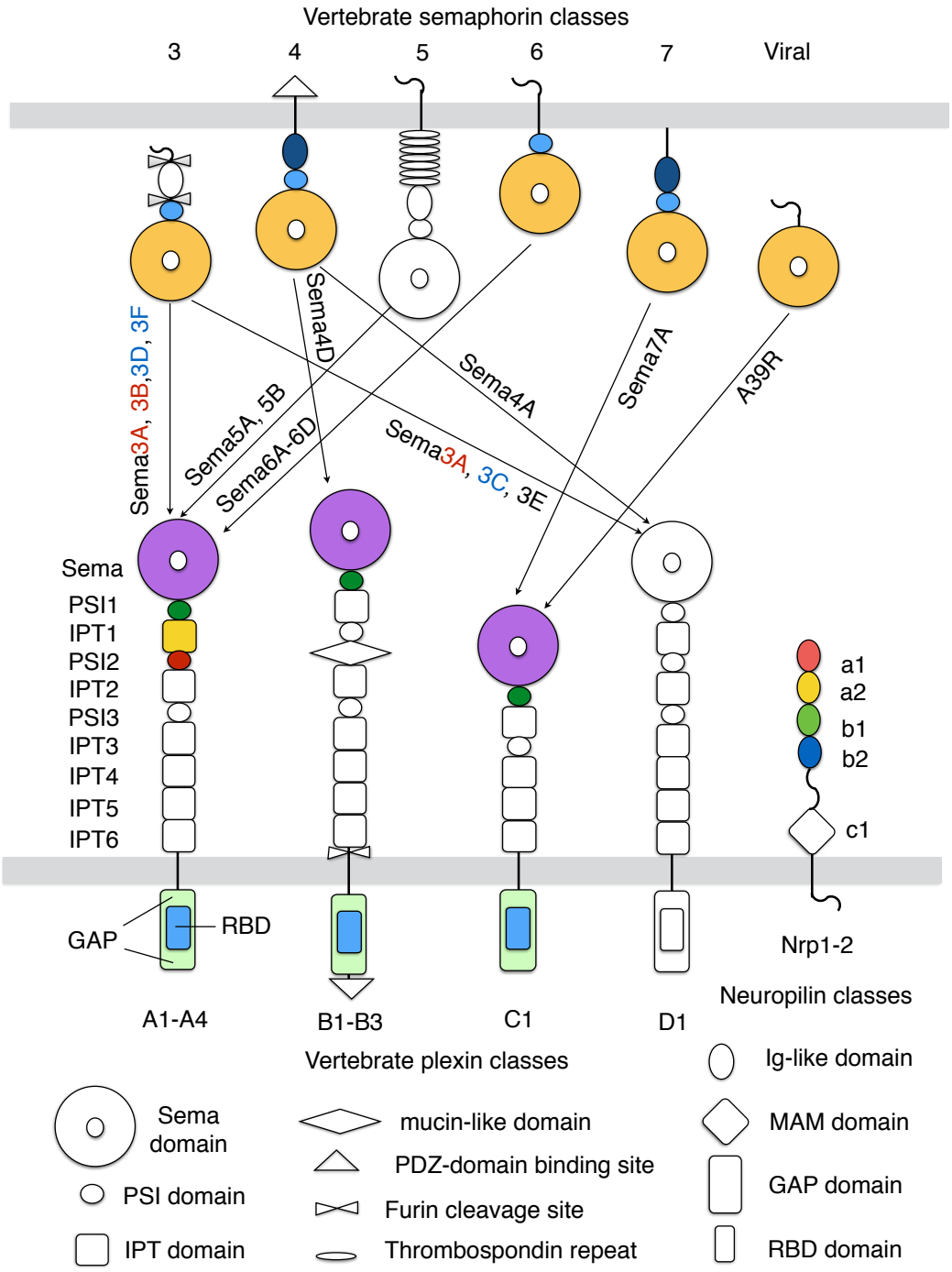


Figure 1.1: The semaphorin, plexin and neuropilin families. Semaphorin–plexin interactions are indicated with arrows (note that not all interactions are listed due to space limitation). The protein domains are shown as cartoons and labeled at the bottom. The coloured domains indicate for known crystal structures. The coloured labelling of Sema3s indicate the requirement of Nrp1 (red) and Nrp2 (blue) in signalling. Cell membrane is drawn in grey. Transmembrane helices are drawn as black lines and disordered loops as black curves.

1.5 Structure Basis of Semaphorin–Plexin Signalling

1.5.1 Structural Basis of Semaphorin–Plexin Recognition

In 2010, three independent studies published crystal structures encompassing members of the class A, B and C plexins in complex with their cognate semaphorin ligands (36, 49, 56). These complexes include two independently reported complexes for Sema6A–PlxnA2, Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$ at 2.19 Å resolution and Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}2}$ at 3.60 Å resolution, Sema4D$_{1\text{-}3}$–PlxnB1$_{1\text{-}2}$ at 2.99 Å resolution, Sema7A$_{1\text{-}3}$–PlxnC1$_{1\text{-}2}$ at 2.40 Å resolution as well as A39R (a Sema7A viral mimic)–PlxnC1$_{1\text{-}2}$ at 2.26 Å resolution, with protein data bank (PDB) accession numbers 3OKY, 3AL8, 3OL2, 3NVQ and 3NVN, respectively (Fig. 1.2.A-D). In all crystal structures of the semaphorin–plexin complexes, two copies of plexin N-terminal segments interact head on with the N-terminal segments of the dimeric semaphorins in their sema domains. Despite the multidomain organisations of both semaphorin and plexins, this head-to-head interaction is restricted to only the sema domains of both binding partners. The principal structural components contributing to this sema–sema interaction are distinctive insertions in or between blades 1 to 5 in the sema domains of both semaphorins and plexins, which are also sufficient to determine binding specificity between cognate semaphorin–receptor pairs (Fig. 1.2.E). These distinctive insertions include a loop of ~20 residues slotting between bladed 1 and 2 (β1D-β2A) and another loop of ~70 residues extending from blade 5 which is generally termed the extrusion. One should note that the same insertions also distinguish

the sema domains of plexins and semaphorins from other types of 7-bladed β-propeller domains such as in integrins and MET, suggesting that the sema–sema binding specificity of plexins and semaphorins can be attributed to ligand–receptor co-evolution (57). Only few residues from the blade edges in the standard 7-bladed β-propeller fold also contribute to this interface. The sequence alignment for the sema-PSI1 portion of class A-D plexins is shown in Appendix 2.

The most striking feature of the semaphorin–plexin complexes is that they share a generic, 2:2 ligand–receptor architecture. All semaphorins are intrinsically dimeric due to their extensive interface between the sema domains. In particular, for the transmembrane class 4 semaphorins, the dimeric form is further strengthened by interactions in the Ig-like domains. For the secreted class 3 semaphorins, disulfide bridges located in C-terminal regions in the uncleaved forms construct obligate homodimers. In contrast, the N-terminal segments of plexins up to the first four domains (sema-PSI2) are strictly monomeric, as revealed by the PlxnA2$_{1\text{-}4}$ crystal structures as well as in solution studies (36). In all the 2:2 semaphorin–plexin complexes, each half of the semaphorin dimers binds extensively to a plexin monomer (a total surface area of ~2000 Å^2) without bringing them into any direct contacts. In-solution binding assays revealed an avidity effect critical for high-affinity binding between semaphorins and plexins, as pre-dimerised semaphorins can boost the binding affinity to plexins over 100 fold compared to the monomeric forms (36). Despite making weak interactions with plexins, monomeric semaphorins failed to trigger any signalling (36, 56). Therefore, the establishment of a 2:2 semaphorin–plexin heterotetrameric complex is key to the activation of plexins, a

mechanism consistent with many TM1 receptors as introduced earlier.

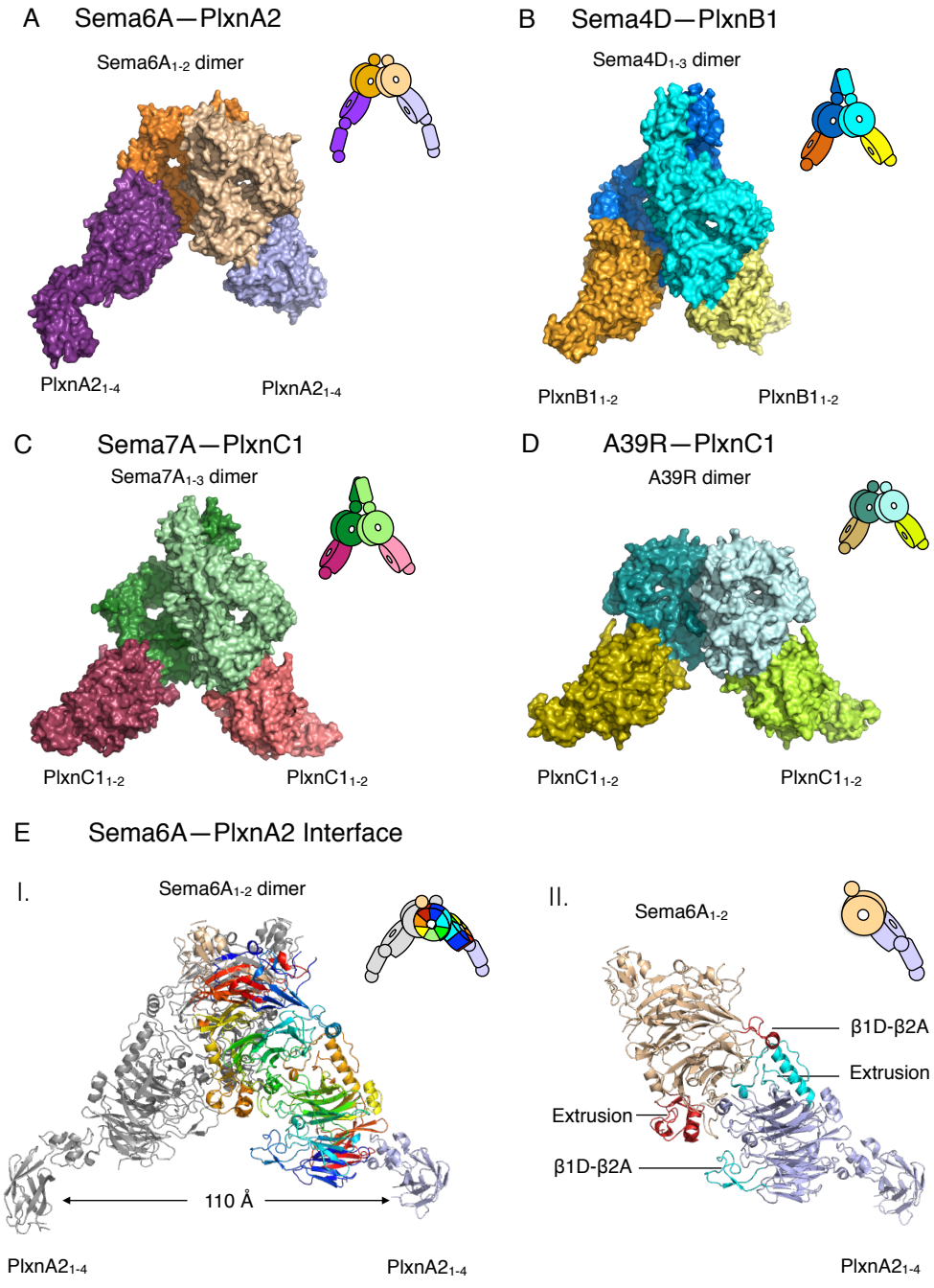


Figure 1.2: Structural basis of semaphorin–plexin recognition. Crystal structures of the Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$, (PDB 3OKY) (A), Sema4D$_{1\text{-}3}$–PlxnB1$_{1\text{-}2}$ (PDB 3OL2) (B), Sema7A$_{1\text{-}3}$–PlxnC1$_{1\text{-}2}$ (PDB 3NVQ) (C), and A39R–PlxnC1$_{1\text{-}2}$ (PDB 3NVN) (D) complexes. (E) The Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$ complex with colour coded β-propeller blades in the sema domains (I) and insertions β1D-β2A and extrusion indicated (II). Each complex is also represented by cartoons (top right).

1.5.2 Structural Basis of Semaphorin–Plexin–Neuropilin Recognition

For all secreted class 3 semaphorins except for Sema3E, signalling through their class A and D plexin receptors requires obligatory coreceptor Nrp1 or Nrp2. The structural basis of recognition between class 3 semaphorins, plexins and neuropilins is revealed by the crystal structure of Sema3A in complex with PlxnA2 and Nrp1 published by our group (45). In the 7.0 Å crystal structure of $Sema3A_{1\text{-}2}$–$PlxnA2_{1\text{-}4}$–$Nrp1_{1\text{-}4}$ complex (PDB 4GZA), the generic 2:2 binding mode of semaphorins and plexins is conserved (Fig. 1.3.A-B.). Two copies of the a1 domain of Nrp1 slot into the 2:2 Sema3A–PlxnA2 complex and cross-brace the sema domain of Sema3A in one interacting unit and the sema domain of PlxnA2 in another unit (Fig. 1.3.B). Structure-guided biophysical and functional assays revealed that $Nrp1_{1\text{-}4}$ cements the 2:2 Sema3A–PlxnA2 complex without disrupting the overall architecture. One should note that the first four domains of Nrp1, domains a1-b2 ($Nrp1_{1\text{-}4}$), are all potentially important for the formation of the Sema3A–PlxnA2–Nrp1 complex, despite only domain a1 being visible in the 7.0 Å crystal structure. How a2-b2 (domains 2-4) of Nrp1 buttress the entire Sema3A–PlxnA2–Nrp1 complex remains elusive. The C-terminal MAM domain of neuropilins (Nrp_5) is not likely to mediate semaphorin–plexin interaction but may be implicated in other neuropilin functions such as dimerisation or clustering (37, 45). Due to the high sequence conservation between different members of the class 3 semaphorins, we may surmise that all the cannonical Sema3s–PlxnAs–Nrps (potentially also PlxnD1) systems have the same 2:2:2 architecture during activation.

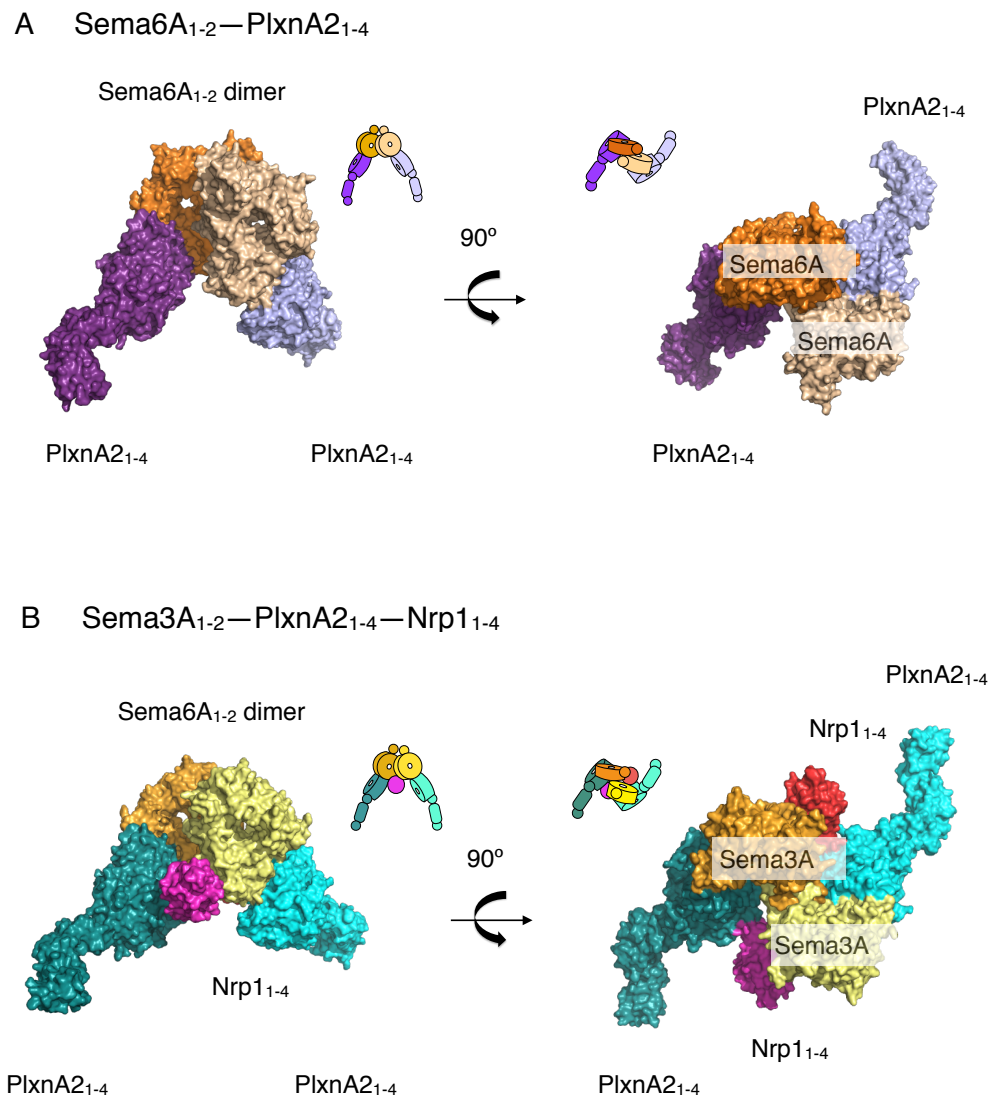


Figure 1.3: Structural basis of Sema3s–PlxnAs–Nrps recoginition. (A) Crystal structure of the Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$ recognition complex (PDB 3OKY) in two views. (B) Crystal structure of the Sema3A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$–Nrp1$_{1\text{-}4}$ recognition complex (PDB 4GZA) in two views. Orientations of each complex are represented by cartoons on the top right corner.

1.5.3 Plexin Signalling in the Cytoplasmic Region

More than 20 proteins have been reported to interact with the plexin cytoplasmic regions to transduce downstream signalling or cross-regulate other receptors (58). Albeit the substrates for plexin cytoplasmic region can be class- or even subtype- specific, the major binding partners of plexins are several members of the small GTPase families. Past studies revealed that the main portion of the plexin cytoplasmic segment constitutes a GAP domain, which directly hydrolyzes the activated guanosine triphosphate (GTP)-bound small GTPases to their guanosine diphosphate (GDP)-bound, inactive states. As these small GTPases are direct regulators of actin and microtubule polymerisation, the inclusion of the GAP domain makes plexins the only TM1 receptor to directly control cytoskeleton dynamics. Though traditionally considered a GAP domain for Ras due to structural homology with Ras-Gap proteins, the plexin GAP domain is in fact specialised for Rap binding (50, 58). While maintaining the general topology of a GAP protein, the plexin GAP domain is interrupted by an insertion of ~60 residues. This insertion constitutes a Rho-GTPase binding domain (RBD), which interact with the Rho family of small GTPases including RND1, Rac1 and RhoD (59). The GAP and RBD of plexins are likely to be independent of each other, as RBD is packed against the back of the GAP away from the Rap binding site (50, 59).

Recent studies revealed that the plexin Rap-GAP activity critically depends on the dimerisation of plexin cytoplasmic domains (50, 51). Dimerisation is likely to induce re-orientation and conformational changes of the plexin cytoplasmic region, opening up the Rap bind-

ing pocket and enabling GTP hydrolysis. This mechanism is best shown by comparing the crystal structures of an active dimer of PlxnC1 cytoplasmic segment with an inactive monomer of PlxnA3 cytoplasmic segment (Fig.1.4.B). An open-swing accompanied by a ~90° pivot of the juxtamembrane helix may serve to reorient and couple together the two copies of the PlxnC1 cytoplasmic region. Pulled by the juxtamembrane helix, an activation loop blocking the Rap entry site opens up, allowing the entry for a GTP-bound Rap and the hydrolysis of GTP into GDP. Should the predictions for the structures of the juxtamembrane and transmembrane residues missing from this model be correct, the Rap-bound PlxnC1 activation dimer could extend its newly aligned juxtamembrane helices, which connect into the transmembrane helices, with little freedom of movement at the membrane proximal region. Considering the bivalent binding mode of semaphorin–plexin signalling complexes in the extracellular space, it is possible that concomitant dimerisation in the plexin extracellular as well as cytoplasmic regions is key to the Rap-GAP activity.

Recent studies have also revealed that the relatively weak binding between RBD and RhoGTPases with a dissociation equilibrium constant (Kd) ranging from 3 to 40 μM does not significantly alter the overall conformation of the plexin cytoplasmic region (59, 60). Although oligomers of plexin cytoplasmic regions bridged by Rho-GTPase Rac have been observed in crystal structures (59), the existence of such oligomers on the cell surface remains elusive. How other substrates such tyrosine kinases and redox enzyme MICAL interact with the plexin cytoplasmic region awaits future characterisations (17).

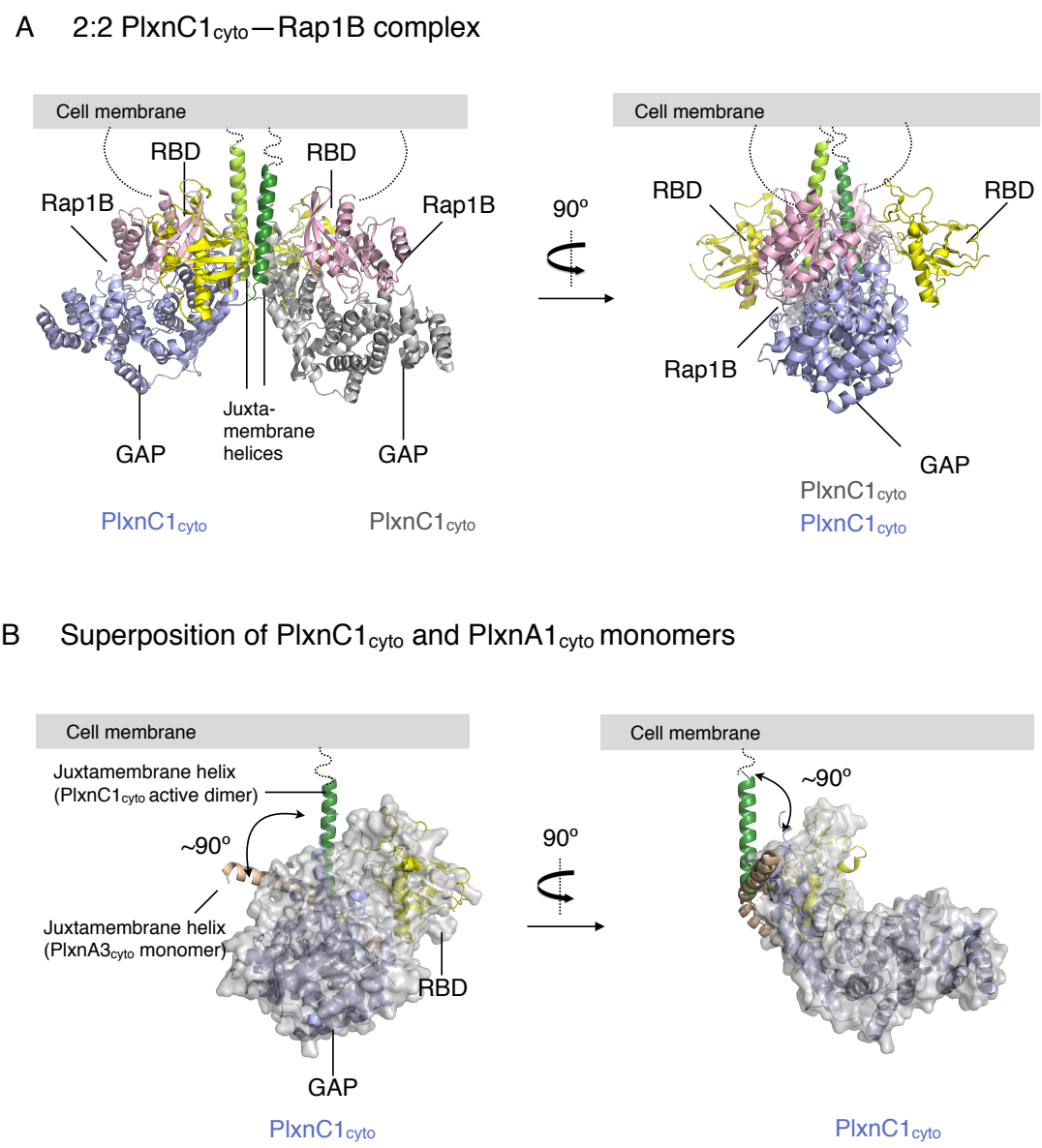


Figure 1.4: Structural basis of plexin signalling in the cytoplasmic region (A) Crystal structure of the zebra fish $PlxnC1_{cyto}$ dimer in complex with Rap1B (PDB 4M8N) shown in two views. The two GAP domains of $PlxnC1_{cyto}$ monomers are coloured in light blue and grey, respectively. (B) Superposition of a monomer of the $PlxnC1_{cyto}$–Rap1B dimer (light blue, ribbon representation) and a monomer of the unliganded $PlxnA3_{cyto}$ (3IG3) (grey surface) shown in two views. The juxtamembrane helices of $PlxnC1_{cyto}$ and $PlxnA3_{cyto}$ are shown in green and wheat, respectively.

1.6 Functions of Class A and D Plexins

1.6.1 Overview

Semaphorin–plexin signalling mediates a plethora of important biological processes, including neural circuit development, cardiovascular patterning, angiogenesis, immune response, and bone homeostasis (39, 41, 61). Malfunctions in plexin signalling are implicated in various human pathologies, ranging from neural and cardiovascular diseases, immune disorders, as well as cancer (62, 63). In this thesis, I focused on studying two particular classes of plexins, the Plexin As and Plexin D1 (PlxnA1-A4, and PlxnD1). The class A and D plexins, which share the same domain organization, are in many cases analogous in functions. Both PlxnAs and PlxnD1 interact with class 3 semaphorins aided by neuropilins except for the Sema3E–PlxnD1 system. The PlxnAs also signal independently of neuropilins with class 5 and 6 semaphorins while PlxnD1 with Sema4A (17, 62). The diversification and regulation of signalling is particularly important for PlxnAs and PlxnD1 in the neural and vascular systems, in which these two classes of plexins are also most abundant (39, 41, 55). Perturbation of the exquisitely controlled PlxnAs and PlxnD1 leads to many forms of cancer (62, 64). Thus, I will elaborate on the roles of PlxnAs and PlxnD1 in these aforementioned systems.

1.6.2 Class A and D Plexins in Neural Development

Both PlxnAs and PlxnD1, first discovered in axon guidance, were later found to regulate multiple stages during the development of the nervous systems ranging from the migration and polarization of neurons, the growth and pruning of axons, as well as synaptogensis (39, 41). To correctly establish and maintain the neural network, the diversification of PlxnA and PlxnD1 signalling based on multifaceted regulatory mechanisms is essential.

Typically, canonical semaphorin signalling via the class A plexins results in repulsive guidance in the wiring of neural circuits. However, it has recently been discovered that Sema6B–PlxnA2 interactions can result in reverse signalling in the developing central nervous system (CNS), in which the transmembrane Sema6B acts as the receptor and PlxnA2 as the ligand (65). It has also been found that Sema6B–PlxnA2 and Sema6A–PlxnA4 not only interact in *trans* but also in *cis*. The *cis*-interactions have been found to impose an inhibitory effect to both the Sema6s and the PlxnAs, resulting in ligand–receptor mutual attenuation or silencing (65–67). These *cis*-inhibitions of guidance cues and their receptors, which have been reported for other systems such as ephrin–Eph and Notch–Delta, may be important for the temporal control of signalling and axon self-avoidance (39). *Cis*-interactions between PlxnAs and transmembrane semaphorins can also lead to receptor activation, as in the case of Sema5A and PlxnA2 mediated by HSPGs and CSPGs (46). As PlxnA extracellular segments comprising of ten domains are much larger than the ectodomains of Sema5s and 6s, it remains puzzling how the head-to-

head, semaphorin–plexin interactions can be achieved in *cis*. Another important regulatory mechanism, which has so far only been found for the PlxnAs, is receptor autoinhibition (22). This pre-signalling autoinhibition was revealed to be imposed by the PlxnA ectodomains and in particular the sema domain, without which the PlxnAs are constitutively active. It has been hypothesized that the PlxnA sema domains are locked in an inhibitory state via distinct intramolecular interactions between the sema (head) and the rest of the PlxnA ectodomain (stalk) before semaphorin-binding (22, 36). However, further molecular evidence supporting this hypothesis has not been reported so far.

The sole member of the class D plexins, PlxnD1, triggers a repulsive response upon binding most of its cognate class 3 semaphorins and coreceptors Nrps, a process similar to the Sema3s–PlxnAs–Nrps systems (39, 55). However, a particular member of the secreted semaphorins, Sema3E, stands out as an irregularity. Sema3E can directly signal through PlxnD1 to initiate axon repulsion while the presence of the Nrp1 switches such repulsion to attraction (68). It has also been recently discovered that the Sema3E–PlxnD1 (and potentially Nrp1) system can cross-talk with the VEGF–VEGFR system, although the existence of a Sema3E–PlxnD1–VEGF–VEGFR multimeric complex remains elusive (55, 69). How Nrp1 or other receptors gate the bi-functionality of the Sema3E–PlxnD1 signalling axis, remains a fascinating question.

1.6.3 Class A and D Plexins in Cardiovascular Development

Many parallels can be drawn between the developing cardiovascular and neural systems. During embryotic stages of cardiovascular organogenesis, the cardiac and vascular cells need to travel to appropriate positions. At later stages, blood vessels migrate, branch and extend to form stereotypical patterns guided by external cues. The growing capillaries tipped by highly dynamic endothelial cells teeming with growth factor and guidance cue receptors much resemble the growing axons tipped by growth cones. The patterning of the cardiovascular system, which is analogous to that of the neural networks, is mainly governed by extrinsic growth factors such as vascular endothelial growth factor (VEGF), morphogens, and a number of cell guidance cues. signalling between PlxnAs and their cognate class 3 and 6 semaphorins mostly negatively regulate angiogenesis and trigger repulsions in the migrating cardiovascular cells. Analogous to the Sema6s–PlxnAs system in the nervous system, Sema6D has been found to act as both the ligand and the receptor in its interactions with PlxnA1 during the morphogenesis of the cardiovasculatures (70).

PlxnD1 is particularly important in the developing cardiovascular network. For instance, signalling between Sema3C, PlxnD1 and co-receptor Nrp1 leads to attractive guidance (71). Similar to the nervous system, the Sema3E–PlxnD1 signalling axis has been shown to be intertwined with the VEGF–VEGFR pathways (72, 73). Importantly, Sema3E can signal directly through PlxnD1 without Nrp1 and induce repulsion in growing vascular endothelial cells (72). It has also been shown that competition between Sema3E and

Sema4A when binding PlxnD1 is an important regulatory factor in suppressing angiogenesis (74). Whether the structural basis for the activation and regulation of PlxnAs and PlxnD1 are the same across multiple organ systems awaits further characterisation.

1.6.4 Class A and D Plexins in Cancer

The formation, angiogenesis, and metastasis of many forms of cancer involve the dysfunctioning PlxnAs and PlxnD1 (62). PlxnA signalling has been shown to negatively regulate tumor angiogenesis and metastasis (75). In particular, recent studies highlighted the deterministic roles of Sema3E–PlxnD1 signalling in the formation of cancer, in which PlxnD1 was identified as a dependence receptor analogous to another guidance cue receptor DCC (64, 76). It is recently discovered that PlxnD1 unliganded PlxnD1 can signal through the mitochondrial-associated proapototic pathways while Sema3E binding to PlxnD1 triggers the canonical, cytoskeleton-related PlxnD1 signalling pathways and abolishes its pro-apototic properties (76). This finding fits well with a dependence receptor paradigm, in which an unliganded transmembrane receptor actively induces apoptosis while binding of its cognate ligands switch the receptor to its canonical signalling function (77). The same studies also found that trapping the excessive amount of Sema3E in cancer cell microenvironment with soluble sema domain of PlxnD1 can effectively restore the restore PlxnD1-related apoptosis in breast cancer cells. PlxnD1 signalling has also been found to promote cancer cell metastasis and invasiveness potentially in conjunction with human epidermal growth factor (EGFR) HER2 (or ErbB2), in which the shortest, proteolytically-cleaved form of Sema3E consisting only of the sema

and the partial PSI domain (Sema3E p61) plays an essential role (77). In sum, structural insights into the signalling mechanism of PlxnAs and PlxnD1 may provide invaluable guidance to the design of cancer therapeutics.

1.7 The Missing Links in Plexin Signalling

Crystal structures of the semaphorin–plexin recognition complexes and the plexin cytoplasmic segments provided us with insightful yet fragmentary snapshots of plexin signalling. One of the ultimate goals in the structural studies of cell-surface receptors is the reconstruction of a complete and time-resolved picture of signal transduction. Based on our current structural knowledge, we can derive the hypothesis that the dimerisation of plexins induced by bivalent semaphorin-binding is central to receptor activation. In the semaphorin–plexin as well as semaphorin–plexin–neuropilin recognition complexes, the C-termini of two PlxnA2 monomer fragments, though contacting the dimeric semaphorins at the N-termini, are separated by more than 100 Å (Fig. 1.2). It has also been demonstrated that the main cytoplasmic activity of plexins, i.e. Rap-GAP function, is only triggered through dimerisation (Fig. 1.4). These observations raise an important question, how do the initially divergent plexin ectodomain allow their cytoplasmic regions to converge upon semaphorin-binding? Moreover, how do the full-length co-receptor neuropilins fit into the most probably conserved architecture of semaphorin–plexin signalling complexes?

As detailed in previous sections, the plexin receptors, in particular the PlxnAs, possess

versatile modes of interaction apart from the canonical, *trans*-interactions with dimeric semaphorins. How do the PlxnAs make inhibitory *cis*-interactions with transmembrane class 6 semaphorins, while the same *cis*-interactions with class 5 semaphorins result in receptor activation? What are the structural basis for PlxnA autoinhibition and potential clustering pre- and post-signalling? In particular for PlxnD1, what structural basis makes it stand out as the only plexin to achieve bi-functionality (attraction and repulsion) with Sema3E gated by coreceptor Nrp1?

To answer the above questions and eventually reconstruct a complete picture of plexin activation and regulation, structural insights into the entire plexin ectodomains are indispensable. Unfortunately, existing crystal structures for the plexin ectodomains are fragmented, encompassing at most four N-terminal domains in PlxnA2$_{1\text{-}4}$ (sema-PSI2) (36, 45). Structures for the large part of the plexin ectodomains remain missing. Moreover, PlxnD1 remains the only class of plexins for which structural information is completely unknown. Thus, in this thesis I set out to characterise the PlxnA and PlxnD1 ectodomains using a combination of structural, biophysical and functional analyses.

1.8 Current hypotheses of Plexin Signalling

Based on the available experimental data regarding the semaphorin–plexin system and many other TM1 receptors, the following hypotheses have been proposed for semaphorin–plexin signalling (36, 37, 45, 58). These hypotheses are mostly applicable to PlxnAs but can potentially be extended to other types of plexins particularly the PlxnD1, which

share similar domain organization with the PlxnAs (Fig. 1.5). Before encountering the semaphorins, the PlxnAs may reside in an autoinhibited state potentially via intramolecular interactions using the "head-down" conformation, in which sema domains are locked downwards by the remaining extracellular domains. One should note that this autoinihibited conformation is proposed based on the inactive conformation of integrins, in which the seven-β-propeller domain of the integrin α subunit is folded down close to the α subunit leg region (13). The existence of such intramolecular interactions remains elusive.

When meeting the cognate transmembrane class 6 semaphorins (Fig. 1.5.A), the PlxnAs may be released from an autoinhibited state and the high-affinity, 2:2 Sema6s-PlxnAs binding may take over. How the cytoplasmic segments of the PlxnAs are then induced to dimerise was hypothesized to be achieved in two ways. First, the PlxnA ectodomains may adopt a specific conformation allowing the initially far-separated structures to curve and converge. Close proximity of the two PlxnAs binding to one Sema6 dimer may allow the dimerisation of the cytoplasmic regions and even seed further clustering (Fig. 1.5.A.I). Alternatively, the PlxnA ectodomains may maintain the initial segregation and extend into the transmembrane and cytoplasmic regions. In this case, it is the clustering of the semaphorin-plexin signalling complexes that allows adjacent PlxnA cytoplasmic segments to dimerise and trigger signalling (Fig. 1.5.A.II).

Similar signalling mechanisms have been proposed for secreted class 3 semaphorins through PlxnAs and potentially PlxnD1 in the presence of neuropilins (Fig. 1.5.B). In these hypotheses, two copies of Nrp1 or Nrp2 slot into the conserved 2:2 Sema3s-PlxnAs

complex, buttressing the entire signalling complexes without altering the overall architectures. Nrps may engage in *cis*-interactions with each other via the MAM domains or the cytoplasmic segments, bringing signalling complexes together as clusters. One should note that in all hypotheses, the extracellular segments of plexins are speculated to be in relatively extended and rigid conformations, by drawing reference to many other large, multidomain TM1 receptors such as the Eph receptors, RPTPσ (receptor protein tyrosine phosphatase σ) and several cell adhesion molecules (CAMs) (47, 52, 78). However, whether any of these hypotheses should stand requires further experimentation particularly structural analysis of the PlxnA and PlxnD1 ectodomains.

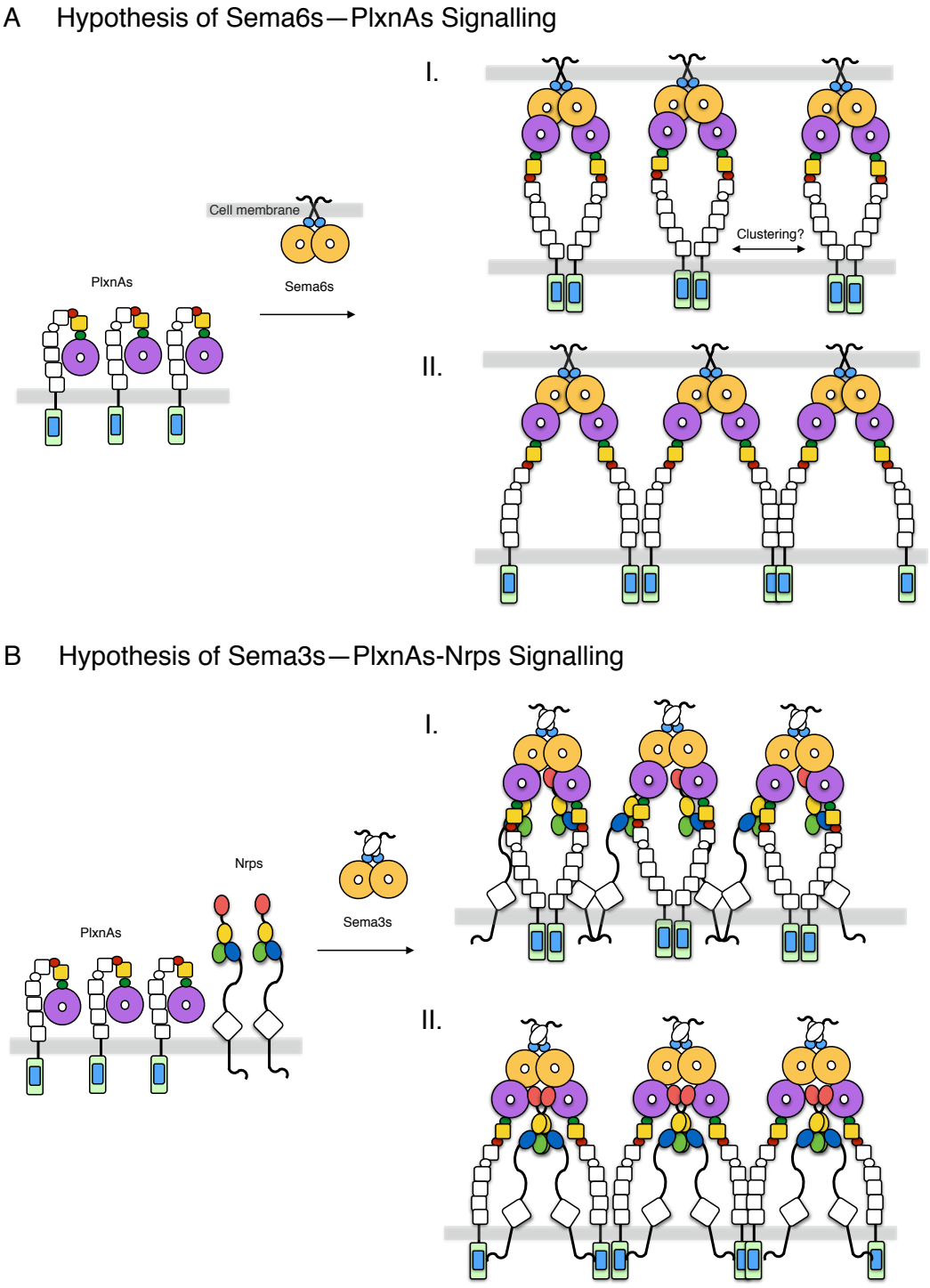


Figure 1.5: Current hypotheses of plexin signalling. (A) Hypothesis of the signalling mechanism for transmembrane semaphorin and plexins (illustrated by the Sema6s–PlxnAs system). (B) Hypothesis of the signalling mechanism for secreted class 3 semaphorins and class A and D plexins (illustrated by the Sema3s–PlxnAs–Nrps system).

1.9 Summary

The semaphorin–plexin signalling system, first discovered in axon guidance, governs a plethora of important biological processes. The plexins are typical TM1 receptors which exemplify the effectively activated and precisely controlled guidance cue receptor signalling. Despite our current knowledge on the semaphorin–plexin recognition mode and plexin Rap-GAP signalling in the cytoplasmic region, substantial structural segments connecting the ligand-binding end to the cytoplasmic-end of plexins remain missing. For the extensively studied PlxnAs as well as the least characterised PlxnD1, the structural basis for the activation and regulation of signalling remains unclear.

In this thesis, I first set out to structurally characterise the full extracellular segments of PlxnAs. Contrary to previous hypotheses, I uncovered a distinct ring-like conformation together with a minor, twisted-open conformation for the PlxnA ectodomains using protein crystallography and negative stain electron microscopy (EM) (Chapter 3). Superposing these two PlxnA ectodomain structures onto the Sema6s–PlxnAs as well as the Sema3s–PlxnAs–Nrps signalling complex provided an elegant model to answer the long-standing questions regarding the dimerisation-induced PlxnA activation. Observations in the crystal structures then lead me to investigate a head-to-stalk *cis*-interaction in PlxnA ectodomains using biophysical assays and state-of-the-art fluorescence microscopy techniques. My results in combination with functional assays revealed that this head-to-stalk *cis*-interaction is the basis for PlxnA autoinhibition and small-scale oligomerisation pre-signalling (Chapter 4). My findings uncovered a two-

fold role played by the ring-like PlxnA ectodomains, first, imposing receptor autoinhibition based on the head-to-stalk *cis*-interactions and second, enabling the dimerisation of plexin cytoplasmic domain upon semaphorin-binding. Although much remains to be done, in this thesis I also report the structures of the PlxnD1 N-terminal region together with a remarkable closed-ring architecture of the PlxnD1 full ectodomain (Chapter 5). These structures of the PlxnD1 extracellular segment tantalisingly suggest mechanisms which may underlie the uniquely diverse functions of PlxnD1. In sum, this thesis provides structural insights into the activation and regulation for two classes of plexins, PlxnAs and PlxnD1, shedding new light on the plexin as well as other guidance cue receptor signalling systems. The applications of my results as well as future outlooks are detailed in Chapter 6.

CHAPTER 2

MATERIALS AND METHODS

2.1 Molecular Biology

2.1.1 cDNAs and Constructs

Protein constructs used in this thesis were cloned from the mouse PlxnA1 (Gene ID 18844), PlxnA2 (Gene ID 18845), PlxnA4 (Gene ID 243743), Sema6A (Gene ID 20358), PlxnD1 (Gene ID 67784) and human PlxnD1 (Gene ID 23129) complimentary deoxyribonucleic acid (cDNA). All constructs were designed based on their structural homology to existing crystal structures together with the help of homology modeling server SwissModel (79), protein tertiary structure prediction program HHPred (80) and protein disorder prediction server RONN (81).

For crystallisation experiments and biophysical assays, mouse PlxnA1$_{1\text{-}10}$, PlxnA1$_{7\text{-}10}$, PlxnA2$_{1\text{-}10}$, PlxnA2$_{1\text{-}2}$, PlxnA2$_{4\text{-}5}$ and PlxnA4$_{1\text{-}10}$ (residues 37-1236, 861-1236, 35-1231, 35-561, 655-804, 36-1229, respectively), together with human PlxnD1$_{1\text{-}2}$, PlxnD1$_{1\text{-}10}$ and mouse PlxnD1$_{1\text{-}10}$ (residues 47-602, 49-1272 and 49-1271, respectively) were cloned into the pHLsec vector with a C-terminal hexa-histidine tag designed for large-scale protein expression in mammalian cell lines (1). Proteins used as ligands in SPR experiments, mouse PlxnA2$_{4\text{-}5}$ and PlxnA2$_{4\text{-}5}$ F690N/E692S were cloned into the pHLsec vector with a C-terminal BirA recognition sequence for biotinylation. PlxnA2_mVenus, PlxnA2_mTFP, and PlxnA2$_{\text{FL}}$-mVenus together with their mutant forms used in microscopy experiments were cloned into the pHLsec vector with a C-terminal mVenus or mTFP tag. For PlxnA functional assays, full-length PlxnA4$_{\text{FL}}$ (residues 36-1894) and its interface mutant PlxnA4$_{\text{FL}}$
F689N/E691S, and cytoplasmic-deleted version PlxnA4$_{\Delta\text{ecto}}$ were cloned into the pcDNA3.1(+) vector. A pre-dimerised version of mouse Sema6A used as ligands in COS-7 cell-based experiments, Sema6A-Fc (residues 19-571), was cloned into the pHLsec vector with a C-terminal Fc-tag covalently linking the sema-PSI domains of Sema6A.

A list of all constructs used in this thesis is shown in Appendix Fig. A.3 and Fig. A.4.

2.1.2 Molecular Cloning

In brief, recombinant protein constructs were created using polymerase chain reaction (PCR) followed by restriction enzyme digestion, ligation and transformation. A typical

PCR reaction contains 2-5 ng/l of template DNA, 0.2 pmol/µl of forward and reverse primer, 1× polymerase buffer, 0.4 mM DNTPs and 0.1 units of polymerase. PCR cycles were carried out according to the recommended protocols of the Pyrobest (Takara Co.) polymerase. PCR produces were purified using the DNA extraction kit (Qiagen), digested with appropriated restriction enzymes (NEB) and ligated using the the Quick Ligation Kit (NEB) before being heat-shock transformed into the DH5α library efficient competent cells from Invitrogen. Subsequent plasmid extraction were performed using the QIAprep Miniprep, Megaprep and Gigaprep kits. All mutant constructs in this thesis, including PlxnA1$_{1\text{-}10}$ F145N/L147S, PlxnA1$_{1\text{-}10}$ F693N/E695S, PlxnA2$_{4\text{-}5}$ F690N/E692S, PlxnA2$_{1\text{-}10}$ F690N/E692S-mTFP1 (and -mVenus), and PlxnA4$_{FL}$ F689N/E691S were produced via two-step site directed mutagenesis as previously described (36).

2.2 Protein Production and Purification

2.2.1 Protein Expression in Mammalian Cells

All recombinant proteins used in this thesis were produced in an expression system optimized for high-level protein production in mammalian cells (1). Typically, human embryonic kidney (HEK)293T and HEK293S cells were grown at 37°C and 5% CO_2 in Dulbecco's modified eagle's medium (DMEM, Sigma), supplemented with 10% fetal calf serum (FBS, Sigma), 1% non-essential amino acids (Gibco) and L-glutamine (Gibco). For all protein constructs a small-scale expression test was first carried out and the level

of recombinant proteins secreted into the medium was tested with Western blots. For large-scale protein expression, a total of 6 mg plasmids were transfected into 12 roller bottles (surface area of 2125 cm^2) of HEK293S (for crystallisation) or HEK 293T (for biophysical assays) cells grown to 90% confluence using transfection reagent polyethylenimine (PEI). The HEK293T cells contain all the necessary cellular machinery to ensure full glycosylation of secreted proteins, while the HEK293S cells lack the N-acetyl-glucosaminyltransfera I (GnTI) activity and therefore cannot synthesize complex N-linked glycans. The reduction of complex glycans is beneficial for lowering the surface entropy of proteins in solution and thus facilitating crystallisation. The transfected HEK293 cells were grown in DMEM supplied with 2% FCS at 37°C and 5% CO_2 for 7-10 days before harvesting.

2.2.2 Protein Purification

Typically for each protein produced from HEK293S or HEK293T cells, 3 litres of medium were harvested from 12 roller bottles. Residual cell debris in the medium were first spun down and then filtrated out with a 0.22 µm filter. The media was buffer exchanged into PBS buffer and concentrated to 200-500 ml before being pressed through an immobilized nickel metal-affinity chromatography (IMAC) column. The column was then washed with a gradient of 10-50 mM of imidazole, 2 × PBS and eluted with 350-400 mM of imidazole. 2 × PBS. Subsequently, size-exclusion chromatography (SEC) using a prepacked HiLoad 16/600 Superdex 200 or 75 prep grade column from GE Healthcare Life Sciences was conducted and the purity of each fraction was assesed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

2.3 Protein Crystallography

2.3.1 Basic Concepts of Crystal Geometry

If we treat the molecular contents of a protein crystal as lattice points, we can interpret the crystal as a periodic, finite repetition of unit cells which are only related to each other by translational symmetry. The smallest, necessary assembly of lattice points, upon which symmetry operations generate a unit cell, is termed the asymmetric unit. The combinations of intrinsic unit cell symmetry together with crystal lattice symmetry yield a total number of 14 Bravais lattices and 230 general space groups, detailed in the International Tables for Crystallography (82). However, due to the non-chiral property of protein molecules, the symmetry operations allowed in three-dimensional space for a protein crystal only include the 2-, 3-, 4- and 6-fold rotation axes and screw axes. Thus the allowable space groups for protein crystals are limited to 65 across 7 crystal systems.

In real space $\mathbb{R}$ represented by the vectors [$\mathbf{a}, \mathbf{b}, \mathbf{c}$], a unit cell is represented by three vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and three angles α, β, γ. From the same origin we can construct the reciprocal space $\mathbb{R}^*$, in which the reciprocal axes $\mathbf{a}^*$, $\mathbf{b}^*$, and $\mathbf{c}^*$ are normal to the real lattice planes (b,c), (a,c) and (a,b), respectively. A point in the reciprocal lattice can be conveniently represented by the Miller indices as (h,k,l) or by the reciprocal lattice vector $\mathbf{d}^*_{hkl} = (h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*)$. The Miller indices hkl can be visualized as a set of parallel, equidistant reciprocal lattice planes cutting through the unit cell an integral number of times. The perpendicular reciprocal lattice vector $\mathbf{d}^*_{hkl}$ of planes hkl points to the cor-

responding reciprocal lattice point (h,k,l) (Fig 2.1.A). The construction of the reciprocal lattice described by the Miller indices offers great advantage for crystal diffraction analysis.

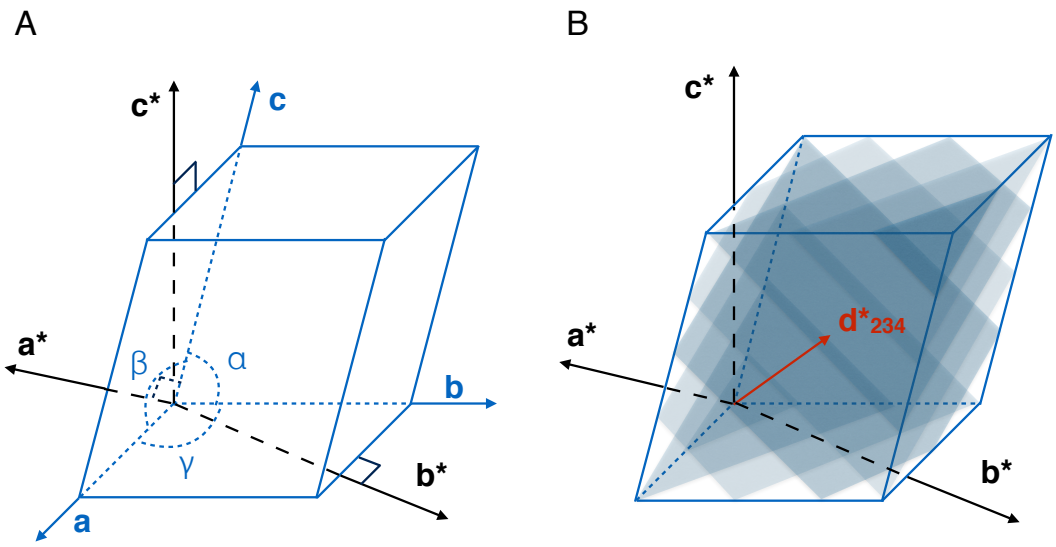


Figure 2.1: The unit cell and the reciprocal lattice. (A) A representative unit cell with unit cell axes **a**, **b**, **c** and the reciprocal lattice axes $\mathbf{a}^*$, $\mathbf{b}^*$, $\mathbf{c}^*$. (B) A set of reciprocol lattice planes with Miller indices 234 and reciprocal lattice vector $\mathbf{d}^*_{234}$ slicing through the unit cell.

2.3.2 Scattering of X-rays

In a complex number plane, a partial X-ray wave can be conveniently represented by the amplitude $|F|$ and phase angle φ :

$$F = A + iB = |F|(\cos\varphi + i\sin\varphi) \tag{2.1}$$

$$F = \sum_{j=1}^{n} |F_j| \cdot \exp(i\varphi_j) \tag{2.2}$$

When a system of electrons interact with incoming X-rays, they start to oscillate and emanate electromagnetic waves. If the oscillating electrons retain their ground state and emit radiation of the same energy as the incoming X-ray, elastic (Thomson) scattering occurs. These elastically scattered X-rays carry information on the scattering system and can be used to probe the internal structure of a protein crystal. When a beam of X-rays is elastically scattered by an atom, the emanating partial waves have the same wavelength but different phases. For the path difference of partial waves between two scattering elements in an atom (O and P), we can derive the phase difference $\Delta\varphi = 2\pi\mathbf{r} \cdot \mathbf{S}$ (Fig. 2.2.A). By integrating all the partial waves emanating from an entire atom with electron density $\rho(r)$, we obtain the atomic scattering factor $\mathbf{f}_s$:

$$\mathbf{f}_s = \int_V \rho(r) \exp(i\Delta\varphi) d\mathbf{r} = \int_V \rho(r) \exp(2\pi i \mathbf{r}\mathbf{S}) d\mathbf{r} \tag{2.3}$$

in which V and r represent is the volume and radius of the atom, respectively.

2.3.3 Diffraction of X-rays by a Protein Crystal and the Laue Condition

The scattering of X-ray by atoms in a single molecule is diminutive and practically impossible to measure. However, when a large number of molecules pack as periodic arrays in a crystal, their scattered waves can interfere constructively in certain directions and thus achieve systematic amplification. This constructive interference maximum, can experimentally recorded as reflections from a diffracting crystal. Similar to the scattering of x-rays by two neighbouring elements in an atom, The path difference between emanating waves from two scattering objects in a crystal (represented as O and P in Fig. 2.2.B) is $\Delta p\lambda\mathbf{r}\cdot\mathbf{S}$, and thus the phase difference is $\Delta\varphi = 2\pi\mathbf{r}\cdot\mathbf{S}$. From the properties of plane waves we know that maximal constructive interference occurs only when the waves are in phase, i.e. the phase difference is 2π or its integer multiples $n\cdot 2\pi$. Therefore, for maximum interference to occur we need $\Delta\varphi_{max} = 2\pi\mathbf{r}\cdot\mathbf{S} = n\cdot\ 2\pi$, and thus, $\mathbf{r}\cdot\mathbf{S} = n$. In three-dimensional space this requirement is written as three independent equations, termed the Laue condition:

$$\mathbf{a}\cdot\mathbf{S} = n_1, \mathbf{b}\cdot\mathbf{S} = n_2, \mathbf{c}\cdot\mathbf{S} = n_3 \tag{2.4}$$

in which **a**, **b**, **c** are the unit cell vectors and n_1, n_2, n_3 must be integers.

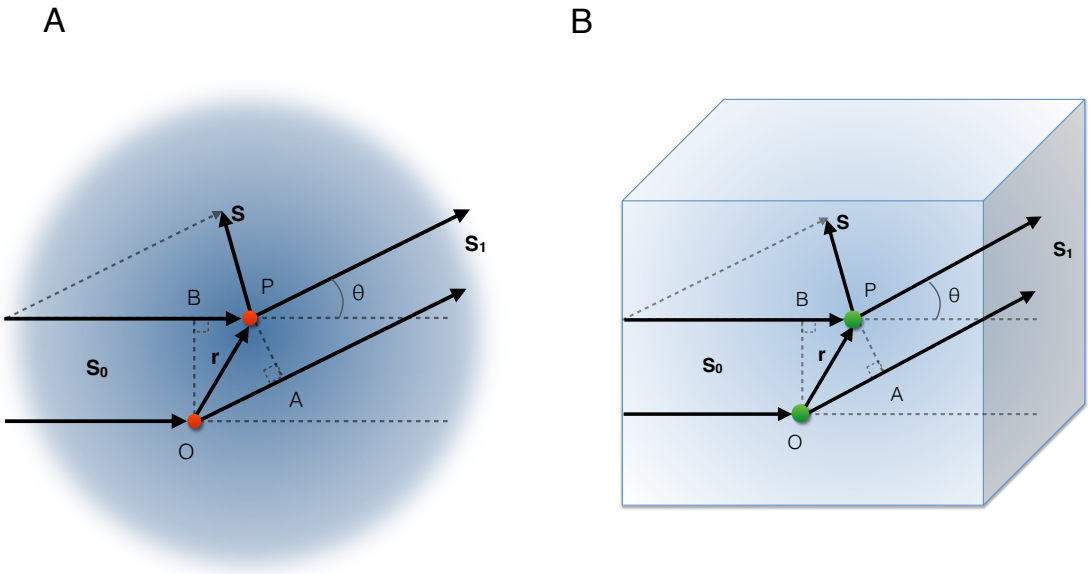


Figure 2.2: Thomson Scattering of X-rays by an atom and a protein crystal. (B) Scattering of x-rays by two elements (point *O* and *P*) in an atom. (B) Scattering of X-rays by two volume elements (*O* and *P*) in a protein crystal.

2.3.4 Bragg's Law and the Ewald Construction

In practice, working with the three separately expressed Laue equations is challenging. In 1912, Sir William Lawrence Bragg proposed a straightforward way to represent X-ray diffraction. If one interprets X-rays scattering off a crystal as electromagnetic waves being reflected by a set of parallel and equidistant lattice planes, one can aptly summarize the relationship between scattering angle θ and interplanar lattice spacing d_{hkl} when the constructive interference maxima in diffraction occurs. This relationship is termed Bragg's law, which can be visualized by simply rotating the scattering diagram of two scattering objects (as illustrated in Fig. 2.2) by a certain degree until the two objects each

sit perpendicularly on top of two parallel lattice planes (Fig. 2.3.A). The distance vector **r** in this case is collinear with the interplanar distance d_{hkl}. As when diffraction maxima occur the path difference between O and P is $\Delta p = n \cdot \lambda$, we arrive at the Bragg Equation:

$$n \cdot \lambda = 2 \cdot d_{hkl} \cdot \sin\theta \tag{2.5}$$

Bragg's law can be conveniently visualized using the Ewald construction. If we draw a sphere of radius $1/\lambda$, termed Ewald's sphere, and place the reflecting lattice planes from Bragg's diagram in the centre, we can derive from simple geometric knowledge that the scattering vector **S** is parallel and of equal magnitude with the reciprocal lattice vector $\mathbf{d}^*_{hkl}$ (Fig. 2.3.B). Since $\mathbf{d}^*_{hkl}$ extending from the lattice origin to point (h,k,l) is normal to lattice planes hkl and thus collinear with the interplanar distance vector $\mathbf{d}_{hkl}$, Bragg's equation can be written as:

$$\mathbf{S} = \mathbf{d}^*_{hkl} \quad \text{with} \quad d^*_{hkl} \equiv \frac{1}{d_{hkl}} = \frac{2\sin\theta}{n\lambda} \tag{2.6}$$

The Ewald construction demonstrates that if a reciprocal lattice point (h,k,l) lies on the Ewald's sphere, the conditions for constructive interference of X-rays reflecting off a set of lattice planes hkl are automatically fulfilled. As only a few reciprocal lattice points will lie on the Ewald sphere at a time in a certain direction, the reciprocal lattices must be rotated so that more reciprocal lattice points can intersect with the Ewald sphere. In practice, this is achieved by a continuous rotation of the crystal around one axis during

X-ray diffraction experiment.

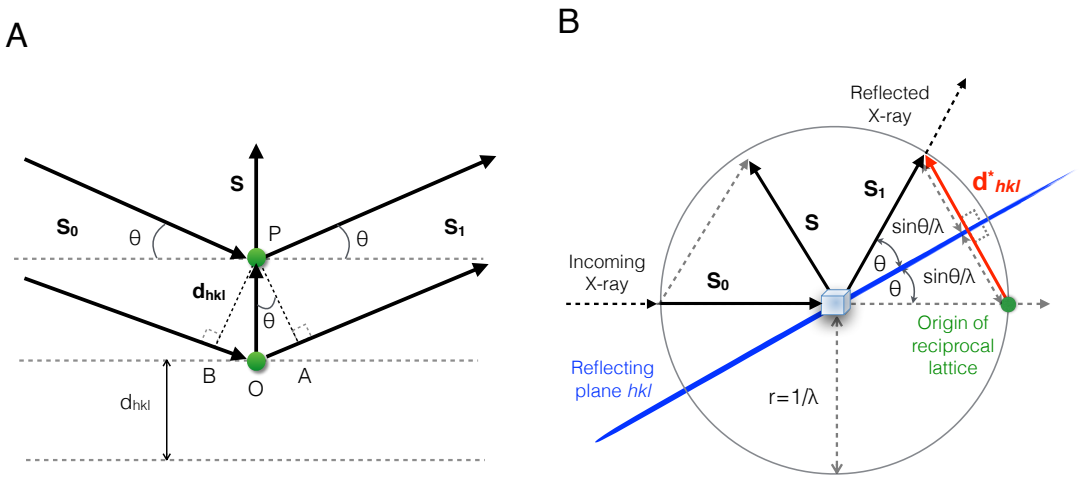


Figure 2.3: Bragg's law and the Ewald Construction. (A) Illustration of Bragg's law for a set of lattice plane hkl. (B) The Ewald construction for in which the reciprocal lattice vector is $\mathbf{d}^*_{hkl}$.

2.3.5 The Structure Factor and the Reconstruction of Electron Density

Bragg's law together with the Ewald construction indicate at which specific directions we can expect constructive interference, i.e. reflections. However, it is the intensities of such reflections that carry information on the scattering elements. For a set of reflections $\mathbf{h}$, a scattering function $\mathbf{F}_h$ is used to summarize the contributions of all individual X-ray scattering events from all scattering elements within the unit cell in that specific

direction:

$$\mathbf{F}_h = \sum_{j=1}^{\text{atoms}} \mathbf{f}^0_{\mathbf{S},j} \cdot \exp(2\pi i \mathbf{h}\mathbf{x}_j) \tag{2.7}$$

in which $\mathbf{x}_j$ is the coordinate vector for atom j in relation to the origin and $\mathbf{f}^0_{\mathbf{S},j}$ is the atomic scattering factor for each individual atom j. Structure determination of macromolecules via X-ray diffraction experiments is in essence the reconstruction of the electron density of a molecule from its structure factors. The structure factors ($\mathbf{F_h}$) in reciprocal space $\mathbb{R}^*$ can be converted to electron density ($\rho(r)$) in real space $\mathbb{R}$ without any loss in information and *vice versa* by a Fourier transform (FT). From Eq. 2.7 we can derive the complex structure factor $\mathbf{F}_{hkl}$ in relation to the electron density $\rho(x,y,z)$:

$$\mathbf{F}_{hkl} = V \int_{x=0}^{1} \int_{x=0}^{1} \int_{x=0}^{1} \rho(x,y,z) \cdot \exp[2\pi i(hx+ky+lz)] dx dy dz \,, \tag{2.8}$$

where V is the unit cell volume.

To obtain the electron density equation we apply a Fourier transform to the above structure factor equation. Since Miller indices constitute discrete data the integration is replaced by a discrete summation, we arrive at:

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l \mathbf{F}_{hkl} \cdot \exp[-2\pi i(hx+ky+lz)] \,. \tag{2.9}$$

The electron density of a protein can thus be calculated from the X-ray diffraction data

of its crystal according to the above equation.

2.3.6 The Structure Factor Amplitude and the Phase Problem

As the complex structure factor $\mathbf{F}_{hkl}$ can be separated into an amplitude term $|\mathbf{F}_{hkl}|$ and a phase term φ_{hkl}, Eq. 2.9 can be represented as:

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}|\mathbf{F}_{hkl}|\cdot\exp[-2\pi i(hx+ky+lz)+i\cdot\varphi_{hkl}] \tag{2.10}$$

The amplitude of the structure factor, $|\mathbf{F}_{hkl}|$, can be directly obtained experimentally from the diffraction intensity $I_{obs}(h,k,l)$, which is proportional to $|\mathbf{F}_{hkl}|^2$. However, in a X-ray diffraction experiment the phase information of the emanating waves reaching the detector is completely lost. As the phase term provides dominant information in Fourier synthesis, the measurement of phase angles becomes an essential challenge in crystallography. In general, the phase problem of protein crystallography can be solved by two methods: molecular replacement (MR) and experimental phasing.

2.3.7 Overview of MR Phasing

In brief, MR borrows the phases from an existing structure model which is homologous to the target protein as initial phases. Assuming the starting model is the same as the target protein, a MR program searches for correct orientation and location of the model based on the diffraction intensities recorded in diffraction experiments and the calcu-

lated phase angles of the input model. Reiterations of the searches eventually places the model structure at appropriate positions in the crystal lattice and generates an approximation of the real electron densities for the target protein. Based on the approximated electron density map and the known structural differences between the target protein and the initial model, one can further improve the model until a most probable match between the model and the diffraction data is obtained.

The majority of the crystal structures in this thesis were solved by MR using the PHASER program (83). PHASER generally performs a rotational search to find the appropriate orientation of the initial model followed by a translational search applied to the top solutions in the rotational search to identify the location. Using a maximum-likelihood method, PHASER explores and scores all the possible solutions until the calculated structure factor amplitude $|\mathbf{F}_{cal}|$ best matches the observed structure factor amplitude $|\mathbf{F}_{obs}|$. The solutions are then ranked according to these scores and clashes within the lattice, followed by several rounds of rigid-body refinement. The MR solutions are best evaluated by the log-likelihood gain (LLG) and the Z-score. LLG describes the differences between the probability of the model structure coinciding with experimental data and the probability of the model structure coinciding with the same set of atoms randomly positioned. As LLG represents the superiority of the solution over a random set of atoms, if the MR solution is correct the LLG should increase steadily through each search cycle. The Z-score is a multiple of the root-mean-square (r.m.s.) deviation for a specific set of LLGs in comparison to the mean LLG, and thus the higher the Z-score the better the solution. In general, a correct molecular replacement solution should have a a Z-score

of no less than 7, and no clashes between any structural elements. The electron density map should reveal positive density for missing features and subsequent refinement should improve the agreement between experimental and model structure factors.

2.3.8 Overview of Single Isomorphous Replacement with Anomalous Scattering (SIRAS) Phasing

The phase problem can also be solved by several experimental methods, as in the case of the mouse PlxnD1$_{1\text{-}10}$ crystal structure in this thesis (Chapter 5). As the Pt-SIRAS (platinum single-wavelegth isomorphous replacement anomalous scattering) method is used for PlxnD1$_{1\text{-}10}$, I will briefly introduce its basic principles.

When the incoming X-ray is tuned to be near the absorption edge of the anomalous marker atoms (in our case platinum (Pt) atoms), a portion of the X-rays is absorbed and reemitted as waves with altered phases. This process, termed anomalous scattering, gives rise to a different atomic scatter factor dependent on the wavelength ($\mathbf{f}_{(\theta,\lambda)}$) which is different from the conventional scatter factor for light atoms ($\mathbf{f}^{0}_{(\theta)}$) as described in Eq. 2.3:

$$\mathbf{f}_{(\theta,\lambda)} = \mathbf{f}^{0}_{(\theta)} + \mathbf{f}^{'}_{(\lambda)} + i \cdot \mathbf{f}^{''}_{(\lambda)} \tag{2.11}$$

in which the dispersive component ($\mathbf{f}^{'}_{(\lambda)}$) is phase shifted by 90 °to the absorption component ($\mathbf{f}^{''}_{(\lambda)}$).

Without the presence of anomalous scattering, the structure factors of reflections with conjugate phases, i.e. reflections of a set of reciprocal lattice planes measured from the exact opposite directions, follows Friedel's law, in which they have identical amplitude but opposite phase angles:

$$|\mathbf{F}_{hkl}| = |-\mathbf{F}_{\overline{hkl}}| \tag{2.12}$$

$$\alpha_{hkl} = -\alpha_{\overline{hkl}}$$

Friedel's law also indicates that a 180° rotation is sufficient to collect all reflections of a crystal without any additional symmetry. When anomalous scatterers are incorporated into the native crystal to form an isomorphous derivative crystal, Friedel's law breaks down. The structure factor of the derivative reflections can be represented as $\mathbf{F}_{PH}$, which is related to the structure factor of the native reflections ($\mathbf{F}_P$) and the structure factor of the anomalous scatter (heavy-metal such as platinum atoms) ($\mathbf{F}_H$):

$$\mathbf{F}_{PH} = \mathbf{F}_P + \mathbf{F}_{PH} \tag{2.13}$$

The above relation is best represented in the complex number plane, which also provides a visualization of the geometric solution to the phase problem (termed the Harker diagram (Fig. 2.4.A). One can see that once the amplitudes and phases of $\mathbf{F}_{PH}$ and $\mathbf{F}_H$ are obtained, together with the amplitude of $\mathbf{F}_P$ from the diffraction intensities of the

native crystal, the phase of $\mathbf{F}_P$ can be calculated ((Fig. 2.4.B).

In a typical SIRAS phasing experiment using Pt-derivative crystals, a set of reflections for the native crystal is first measured at an appropriate wavelength. Subsequently, multiple diffraction datasets of Pt-derivative crystals are measured at an optimal wavelength for the anomalous scattering of Pt atoms. From the disparity of the intensities of these two datasets,we can obtain the phase information for structure factor $\mathbf{F}_{PH}$:

$$\mathbf{F}_{PH}^{\lambda_2} = \mathbf{F}_{PH}^{\lambda_1} + \Delta\mathbf{F}_r + \Delta\mathbf{F}_i \tag{2.14}$$

in which $\Delta\mathbf{F}_r$ and $\Delta\mathbf{F}_i$ represent the real and imaginary vector component of anomalous scattering, respectively. $\mathbf{F}_{PH}^{\lambda_1}$ represents the structure factors of the derivative crystal measured at wavelength λ_1 and $\mathbf{F}_{PH}^{\lambda_2}$ represents the same structure factor measured at wavelength λ_2, where anomalous scattering occurs. Similarly, the geometric solution of $\mathbf{F}_{PA}$ can also be represented in the Harker diagram, which is omitted here.

In practice, SIRAS phasing is often used in combination with MR, as in the case for the crystal structure of mouse $\text{PlxnD1}_{1\text{-}10}$ in this thesis. Initial phases provided by a partial MR solution model may aid the determination of anomalous-scatterer positions as well as the generation of experimental phases with correct handedness. Experimental phases can in turn help prevent model bias in MR.

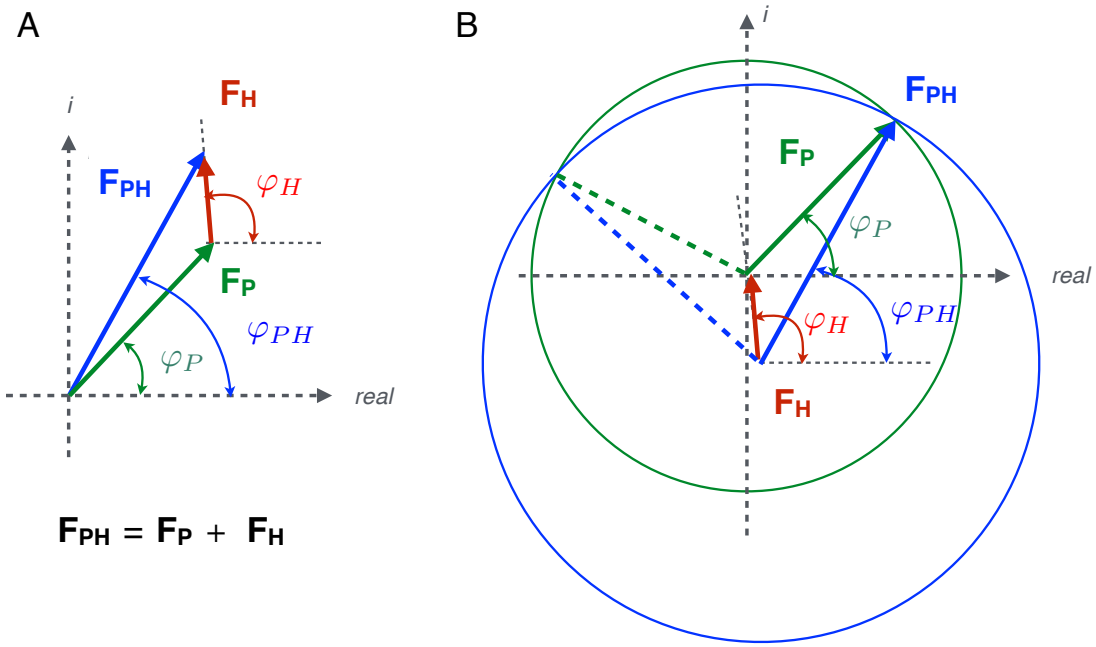


Figure 2.4: The geometric solution of the phase equation (the Harker diagram). (A) Vector representation of the structure factors $\mathbf{F}_{PH}$ (blue), $\mathbf{F}_P$ and $\mathbf{F}_H$ (red) in the complex number plane. (B) Solution of the phase equation in which the phase angle φ_P can be pinned down by the intersections made by a circle with a radius of $|\mathbf{F}_P|$ and an original center together with another circle with a radius of $|\mathbf{F}_{PH}|$ and a center shifted by $\mathbf{F}_H$.

2.3.9 Protein Crystallisation

Protein crystallisation experiments were conducted by mixing 100 nl (or 200 nl for PlxnA2$_{1\text{-}10}$) of protein solution with 100 nl reservoir solution using a Cartesian Technologies pipetting robot (84). Before crystallisation, the purified proteins were typically deglycosylated using endoglycosidase F1. crystallisation plates were maintained at 20.5°C and 6.5°C in a Formulatrix Homebase storage vault and imaged with a Veeco visualization system. Details of the crystallisation procedures for mouse PlxnA2$_{4\text{-}5}$ and PlxnA1$_{7\text{-}10}$ are listed in Chapter 3, and for human PlxnD1$_{1\text{-}2}$ and mouse PlxnD1$_{1\text{-}10}$ in

Chapter 5. Crystallisation procedures for full length PlxnA ectodomains are listed in Appendix 1.1 as they were performed by Dr. Bert J.C. Jassen before this thesis work.

2.3.10 Data Collection and Processing

A rotation method is most commonly used for protein crystallography data collection, in which each diffraction image is recorded when the crystal is rotated around one axis at usually 0.1-0.3 degrees increment (oscillation angle φ) at a time. A small oscillation angle allows for fine-slicing of the reciprocal lattice so that more reflections intersecting with the Ewald sphere can be sampled. After determining the space group and unit cell vectors based on indexing the first few diffraction images, a data collection strategy which allows for high completeness and redundancy while best-avoiding radiation damage is adopted. The spot intensities on each diffraction image are then integrated and scaled so the accurate intensities for each reflection are determined.

The overall data quality of diffraction experiments can be assessed by the following parameters: data completeness, redundancy, R_{merge}, R_{pim}, $I/\sigma I$, and $CC_{1/2}$. Data completeness denotes the how many unique reflections are recorded experimentally out of all unique reflections we theoretically expect. The unique reflections are the minimum number of symmetry-unrelated reflections which generate a complete dataset. Thus, the higher the completeness the better the dataset. Redundancy describes how many times a unique reflection is repeated measured. Generally speaking, the higher the redundancy the more accurate the calculated intensity for the recorded reflections. The

linear merging R-value, Rmerge, represents the extent of deviation between equivalent intensity measurements, which can be calculated as follows:

$$R_{merge} = \frac{\sum_{\mathbf{h}} \sum_{i=1}^{N} |\mathbf{I}_{(\mathbf{h})i} - \overline{\mathbf{I}_{(\mathbf{h})i}}|}{\sum_{\mathbf{h}} \sum_{i=1}^{N} \mathbf{I}_{(\mathbf{h})i}} \tag{2.15}$$

in which N is redundancy and $I/\sigma I$ is the average intensity of each reflection. Although Rmerge can describe the inherent consistency of a dataset, it does not account for the redundancy of repeatedly measured reflections as the value of Rmerge always increases with redundancy. To correct for this problem the precision-indicating merging value, R_{pim}, is introduced:

$$R_{pim} = \frac{\sum_{\mathbf{h}} \sqrt{\frac{1}{N-1}} \sum_{i=1}^{N} |\mathbf{I}_{(\mathbf{h})i} - \overline{\mathbf{I}_{(\mathbf{h})i}}|}{\sum_{\mathbf{h}} \sum_{i=1}^{N} \mathbf{I}_{(\mathbf{h})i}} \tag{2.16}$$

Rpim is particular important for assessing highly redundant datasets such as those derived from anomalous diffraction experiments. The signal to noise ratio of the dataset, $I/\sigma I$, can be used to assess general data quality and the resolution limit. Ideally, the highest resolution shell of a dataset should have an $I/\sigma I$ larger than 2.0. Calculated from the Pearson correlation coefficient (CC) of two random halves of a dataset, $CC_{1/2}$ is a parameter for assessing information content and data accuracy (94). Ideally, $CC_{1/2}$ should be close to 1.0 for an overall dataset and larger than 0.1 at the highest resolution shell.

X-ray diffraction data in this thesis were collected at 100 K at Diamond beamlines I02, I03, I04, I04-1 and I24 and European Synchrotron Radiation Facility beamline ID23-2. Typically, a rotation angle of 180-360° and an oscillation angle of 0.1° was used for all native crystals. In particular for the Pt-derivative crystals of mouse PlxnD1$_{1\text{-}10}$, multiple platinum-SAD (single-wavelength anomalous dispersion) datasets were collected with a rotation angle of 990° (oscillation angle 0.1°) at the platinum peak wavelength (1.0719 Å) based on a fluorescence scan. The collected reflections were indexed, integrated and scaled using iMOSFLM (85), SCALA (86) and multiple versions of Xia2 (87). A total of 5 datasets of mouse PlxnD1$_{1\text{-}10}$ platinum derivative crystals were merged to produce high anomalous signals using Xia2 3dii.The experimental details of each diffraction experiment are described in corresponding sections in Chapter 3 and 5 of this thesis.

2.3.11 Structure Refinement

In brief, structure refinement is the process of modifying a structural model until it best fits the crystallographic observations. In practice, refinement is evaluated by comparing the calculated structure factor amplitude of the model ($|\mathbf{F}_{cal}|$) and the experimentally observed structure factor amplitude ($|\mathbf{F}_{obs}|$). The parameter used to evaluate this agreement is the working R-factor, or R_{work}, which is defined as:

$$R_{work} = \frac{\Sigma||\mathbf{F}_{obs}| - |\mathbf{F}_{cal}||}{\Sigma|\mathbf{F}_{obs}|} \tag{2.17}$$

Another cross-validation parameter is R_{free}, which is calculated analogous to R_{work} from a randomly selected set of reflections (~ 5%) excluded from the refinement process. The reduction in R_{free} indicates unbiased model refinement.

During refinement, an electron density map based on the experimentally observed amplitude ($|\mathbf{F}_{obs}|$) and model phases (φ_{cal}) and a difference map ($|\mathbf{F}_{obs}| - |\mathbf{F}_{cal}|$, φ_{cal}) is used for manual corrections and building of missing features. These maps are used as basis for manual correction of the existing model. Interwoven rounds of manual and automated refinement are conducted, during which the stereochemical properties of the model assessed by the Ramachendran plot as well as the agreement between $|\mathbf{F}_{cal}|$ and $|\mathbf{F}_{obs}|$ should steadily improve. Structural refinement is generally complete when the structural model and experimental data reaches maximum fit. In this case R_{work} and R_{free} should have reached their lowest, appropriate values and the electron density difference map no longer reveals new interpretable features. It is important to note that at the end of refinement R_{free} should not be more than 7% higher than R_{work}, as a larger difference typically indicates model bias. The r.m.s. deviation values of bond lengths and bond angles, which compares the model geometry to reference structures of small molecules, should be reasonably small.

Structure solution and refinement of mouse PlxnA2$_{4\text{-}5}$ and PlxnA1$_{7\text{-}10}$ are detailed in Chapter 3, and for human PlxnD1$_{1\text{-}2}$ and mouse PlxnD1$_{1\text{-}10}$ in Chapter 5. Structure solution and refinement of full length PlxnA ectodomains are listed in Appendix 1.1 as they were performed by Dr. Bert J.C. Jassen before this thesis work. Automatic refinement was performed using COOT(88) and PHENIX (89).

2.3.12 Sequence Analysis and Figure Production

Electrostatic charge distribution for all crystal structures was calculated with PDB2PQR (90) and APBS (91), alignments were generated with Clustal Omega (Sievers et al., 2011), residue conservation was calculated with the Consurf Server (92) or MODELLER and buried surface areas were calculated with PISA (93). Figures were produced with The PyMOL Molecular Graphics System (Schrödinger, LLC).

2.4 Negative Stain Electron Microscopy

2.4.1 Basic Principles of Negative Stain Electron Microscopy (EM)

Negative stain electron microscopy (EM) is a fast and convenient transmission electron microscopy (TEM) technique to visualize the structures of proteins. During sample preparation, the target proteins, originally freely diffusing in a solution environment, are embedded into a layer of heavy-metal stain (in my case, uranyl formate). The cast of uranyl atoms surrounding the proteins can interact with electrons at much greater scales than the light atoms making up the proteins and thus, provide high image contrast while protecting the specimen from electronic bombardment. Negative stain EM is often used as a complimentary technique to protein crystallography, as it may reveal novel structural arrangements otherwise obstructed by constraints in crystal packing. However, one should note that the process of instant drying and flatting of the stain-

ingrained proteins onto the grid may induce artefacts.

The basic components of a transmission electron microscope include an electron source, a series of electromagnetic lenses, apertures, and an image detection system (Fig. 2.5). When electrons strike a negative stain EM sample, a small portion of the electrons are absorbed by the heavy metal atoms surrounding the specimen, resulting in reduced intensity at stain-ingrained regions in the projection image, i.e. amplitude contrast. Only a very small fraction of electrons collected by the detector are scattered elastically and interfere with the non-scattered electrons, resulting in phase contrast. In negative stain EM amplitude contrast account for the majority of image contrast, while phase contrast, especially in our case where the protein specimen is only the size of several nanometers, is diminutive. In practice, image contrast can be increased by adjusting the projector lenses to obtain larger defocus values. Although image defocus enhances contrast, it may sacrifice resolution and bring about complications in image interpretation and thus need to be carefully adjusted in a case-by-case manner.

Despite the low resolution and contrast of raw EM images, a large number of protein particles can be assembled and averaged to gain signal-to-noise ratio. Typically, an assembly of single particles are first aligned according to the correlations between all possible particle pairs in a large data set taking into account both translational and rotational shifts. Alignment of particles is performed until the contrast and visible details of the aligned image sets no longer increases and the maximum value of the CCF (cross-correlation function) between the references and the raw images reaches a maximum. The aligned particles are then classified and averaged to generate a representative image

per class. These class averages represent two-dimensional projections of single particles at very low resolutions.

2.4.2 Experimental Procedures

The preparation of carbon-coated grids followed the previously described protocols for negative stain EM (94). In brief, mesh Gilder Cu grids from Ted Pella Inc. (USA) were first coated with collodion. After the grids were air-dried they were placed into the vacuum chamber filled with vapourised carbon of the Cressington 208C High Vacuum Turbo Carbon Coater (Ted Pella) and coated with a thin layer of carbon. A conventional, two-drop method (95) was used to embed target proteins with uranyl formate onto the carbon-coated grids, in which a drop of 2.5 µl freshly gel-filtrated $PlxnA1_{1\text{-}10}$ at a concentration of 1-5 µg/ml in 10 mM HEPES, pH 7.5 and 150 mM sodium chloride was adsorbed to the newly glow-discharged carbon-coated copper grid, washed with two drops of 50 µl deionized water, and stained with two drops of 50 µl 0.75% uranyl formate. The excess stain on the grids was removed carefully with filter paper before air-drying. Samples were imaged at room temperature using an FEI Tecnai T12 electron microscope equipped with a LaB_6 filament operating at an acceleration voltage of 120 kV and a dose of 15 electrons per square Å. Images were taken using a 4k × 4k FEI Eagle^{TM} CCD camera at a magnification of 57,000 × with under-focus values ranging from 1.0 to 1.5 µm and a pixel size of 2.16 Å.

From 355 electron micrographs of $PlxnA1_{1\text{-}10}$, a total of 13,744 particles were manually

selected using EMAN2 (96) and framed into boxes with a size of 298 Å × 298 Å . The particle images were normalized to re-scale the grey values and filtered to remove noise and accentuate features for alignment. All images were then centered upon the prominent features aided by appropriate masking. Reference-free classification using IMAGIC (97) was first performed. 60 classes were generated and each averaged into a representative image. The structural models based on the 4 Å crystal structure of $PlxnA1_{1\text{-}10}$ were generated manually using The PyMOL Molecular Graphics System (Schrödinger, LLC). The two-dimensional projections of the $PlxnA4_{1\text{-}10}$ extracellular models were then subjected to automated correlation analysis with the class averages in which all the class averages were compared to all projections of all models. Thirteen models were tested and of these seven were found to be sufficient to represent the experimental class averages. Two-dimensional projections of the crystal structure and models were generated using SPIDER and WEB (98) and WellMAP (99) software and the correlations of the class averages were performed using scripts operating through SPIDER.

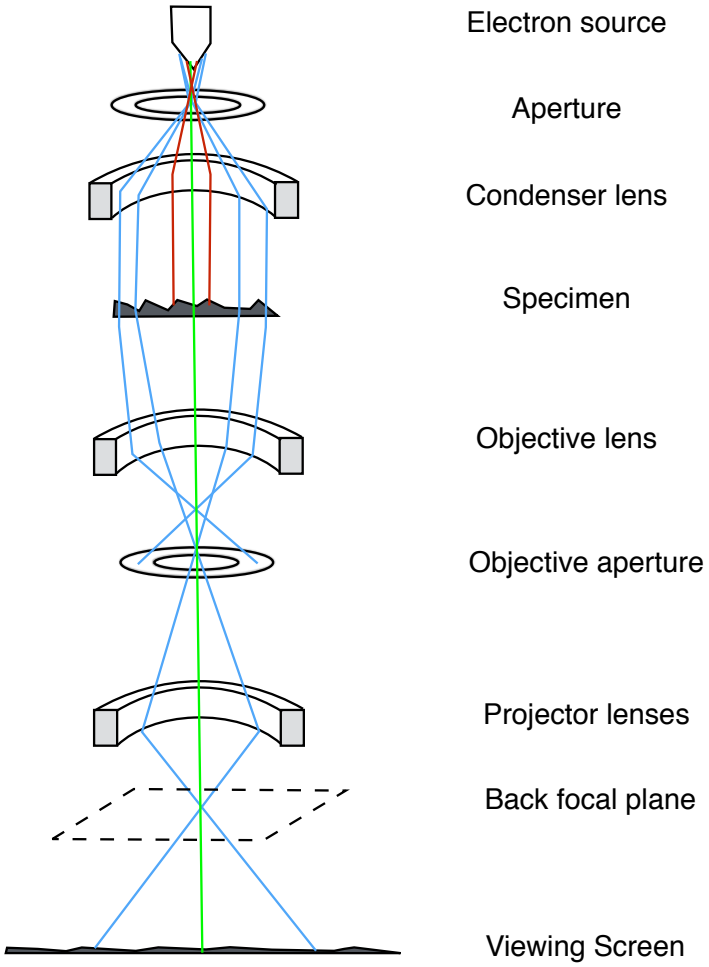


Figure 2.5: Schematic illustration of the negative stain EM microscope setup. The basic components of a FEI Tecnai T12 transmission electron microscope used in this thesis are illustrated as cartoons. The paths of representative electrons absorbed by the sample are colored in red, while the paths of non-scattered and elastically scattered electrons are coloured in green and blue, respectively.

2.5 Multi-angle Light Scattering (MALS)

2.5.1 Basic Principles of MALS

MALS is an optics-based, non-destructive method to measure the molecular weight and sizes of macromolecules in solution. In a typical multi-angle light scattering (MALS) experiment coupled to size exclusion chromatography (SEC-MALS), an injected protein solution is first separated into fractions according to the protein molar masses and sizes by a SEC column. The fractionated solutions are then subjected to a laser light to induce Rayleigh scattering in a MALS detector. Rayleigh scattering denotes a form of elastic scattering emitted by particles of sizes much smaller than the wavelength of incoming radiation. Since the intensity of a molecule's Rayleigh scattering power is directly proportional to its molecular weight, the absolute molar mass and the aggregation state of the proteins in each fraction can be calculated. SEC-MALS is a convenient method to evaluate the homogeneity, molecular weight, and oligomerisation states of proteins, despite the fact that retrievable information on the protein globular shapes or interaction modes is limited.

2.5.2 Experimental Procedures

All proteins used in MALS experiments were produced in HEK293T cells using standard protocols. SEC-MALS were performed with a system in which an analytical Su-

perdex S200 10/60 column (GE Healthcare) was coupled to static light-scattering (DAWN HELEOS II, Wyatt Technology), differential refractive index (Optilab rEX, Wyatt Technology) and Agilent 1200 UV (Agilent Techologies) detectors. Data analysis was done using the ASTRA software package (Wyatt Technology).

2.6 Analytical Ultra-centrifugation (AUC)

2.6.1 Basic Principles of AUC

AUC allows for the characterization of the protein sizes, shapes and interactions without the restrictions of any matrices or surfaces. In brief, analytical centrifugation (AUC) utilizes an ultra-high speed centrifuge coupled to an optical detection system to measure the migration properties of proteins under centrifugal force in solution (100). When a protein solution is subjected to high centrifugal force, all particles in the solution experience three forces, the gravitational force from the centrifugal field, the buoyant force and the frictional force related to surrounding environment. The particles of varied masses will redistribute until the three forces balance out. From this sedimenting process we can derive the sedimentation coefficient s (in unit Svedberg, S), which describes the velocity of a molecule migrating along the centrifugal field. For a particular protein diffusing in a solution of certain viscosity and temperature, s is an intrinsic value depending only on the protein's molar mass and shape.

In this thesis, AUC sedimentation velocity experiments were performed to assess the

globular shapes and oligomerisation states of proteins. In a typical AUC sedimentation velocity experiment, the separation of particles with different migration rates under centrifugal force causes a time-evolving redistribution of masses along the radial axis, which can be recorded as shifts at the sedimentation boundaries. These sigmoid-shaped boundary shifts can be modelled by the Lamm equation:

$$\frac{\partial c}{\partial t} = D\left[\frac{\partial^2 c}{\partial r^2} + \frac{1}{r}\frac{\partial c}{\partial r}\right] - s\omega^2\left[r\frac{\partial c}{\partial r} + 2c\right] \tag{2.18}$$

,

in which t is time, r is the distance to the center of rotation, and ω is the angular velocity of the rotor. The modeling of sedimentation boundary shifts using the Lamm equation yields the sedimentation coefficient s (extracted from the mid-points of shifts). The final result of an AUC velocity experiment is a continuous sedimentation coefficient distribution $c(s)$, in which each species of proteins shows a sedimentation coefficient peak (100).

2.6.2 Experimental Procedures for AUC Sedimentation Velocity Experiments

Mouse PlxnA1$_{1\text{-}10}$ and PlxnA1$_{1\text{-}10}$ F693N/E695S used in AUC experiments were produced in HEK293T cells, while human PlxnD1$_{1\text{-}10}$ was produced in HEK293S cells and separated into two species according to the size exclusion profile. AUC sedimentation veloc-

ity experiments were conducted using a Beckman Optima XL-I analytical ultracentrifuge in absorbance mode (280 nm incident light) at 40,000 revolutions per minute (rpm) rotor speed. A total number of 100 scans were collected with one scan every 6 minutes. Scans 6-100 were analysed with Sedfit (101) using a $c(s)$-based sedimentation profile analysis and plotted using the program ProFit (Martin, A.C.R.). Theoretical sedimentation coefficients of the structural models were predicted using PROHYDROUS-SOMO (102).

2.7 Surface Plasmon Resonance (SPR)

2.7.1 Basic Principles of SPR

SPR is a convenient technique to monitor binding events between label-free proteins in solution. In short, the SPR phenomenon causes the reduction of the intensity of light reflected off a surface coated with protein molecules at a certain angle. Accumulation of molecules on this surface leads to changes in the refractive index close to the surface, and thus alters the angle of minimum reflected intensity and the surface plasmon resonance (SPR) angle. One binding partner, termed the ligand, is typically immobilized on the surface of a sensor chip, while the other partner, termed the analyte, is injected in aqueous solution onto the chip via a microfluid system. As the analytes bind to the ligands on the sensor chip surface, the accumulation of overall mass leads to an increase in SPR angle and thus the refractive index. By comparing the refractive index of the binding

partners and that of the background, we can obtain a real-time sensorgram describing the changes in refractive indices, and thus interactions, in resonance units (R.U.). In a SPR equilibrium experiment, a series of analytes at different concentrations are sequentially injected onto the ligands. From the sensorgrams we can obtain the time for the binding events to reach equilibrium and thus, calculate the equilibrium dissociation constant (Kd) and the maximum analyte binding value (B_{max}).

2.7.2 Experimental Procedures

All proteins used in SPR experiments were produced in HEK293T cells using standard protocols. To biotinylate the ligands, HEK293T cell media containing the ligands with a C-terminal biotinylation tag (Avitag) was buffer exchanged into 10mM TRIS-HCl, pH 8.0 and concentrated to 250 µl. Biotinylation reactions were subsequently carried out by mixing the Biomix A and B buffer (Avidity) and BirA enzyme with the medium to reach final concentrations of 15 mM magnesium acetate, 15 mM ATP, 80 µM biotin, 80 mM bicine, pH 8.3 and approximately 30 µg/ml BirA. After incubation for 30 min at 30°C the media is buffer exchanged and further concentrated into the SPR running buffer (10mM HEPES, pH7.5, 150mM sodium chloride and 0.005% (v/v) Tween 20).

The SPR equilibrium experiments were performed with a Biacore^{TM} T200 equipment (GE Healthcare) at 25°C in SPR running buffer (150 mM sodium chloride, 10 mM HEPES pH 7.5, 2 mM calcium chloride, and 2 mM magnisum chloride and regeneration buffer (2 M magnesium chloride). The analytes were subjected to gel filtration in SPR running

buffer immediately before the experiments. The biotinylated ligands were coated onto the sensor chip via a two-step immobilization strategy based on amine-coupling between biotin and streptavidin (103). In-solution contact times of 10-15 min were used. Data analysis was done using Biacore^{TM} T200 v2.0 software with the nonlinear curve fitting of a 1:1 Langmuir binding model (104) to calculate the equilibrium dissociation constant (Kd) and the maximum analyte binding value (B_{max}) via the following equation:

$$R_{Bound} = \frac{C_A \times B_{max}}{C_A + K_d}, \tag{2.19}$$

where R_{Bound} is measured in response units (RU) and C_A is the concentration of analyte.

2.8 Förster Energy Transfer by Fluorescence Lifetime Imaging Microscopy (FRET-FLIM)

2.8.1 Basic Principles of FRET-FLIM

Föster energy transfer (FRET) describes the phenomenon in which a fluorophore in the electron-excited state (the donor) transfers energy via a non-radiative process to a non-excited fluorophore in its vicinity (the acceptor) (Fig. 2.6). As the extent of FRET depends on the nanometre-range distances between the donor and the acceptor, it can be used to probe the interactions between fluorophore-labelled proteins. Each fluorophore has

an intrinsic fluorescence lifetime, which describes the average time the fluorophore remains in its excited state before photon emission. The decay of fluorescence is generally a mono-exponential process:

$$[S_1] = [S_1]_0 e^{-\frac{t}{\tau_D}} \tag{2.20}$$

,

in which $[S_1]$ is the concentration of fluorophores at excited state S_1 and time t, $[S_1]_0$ is the initial concentration of excited fluorophores, and τ is the fluorescence lifetime. When FRET occurs the fluorescence decay of of the donor molecular accelerates, resulting in shortened lifetimes.

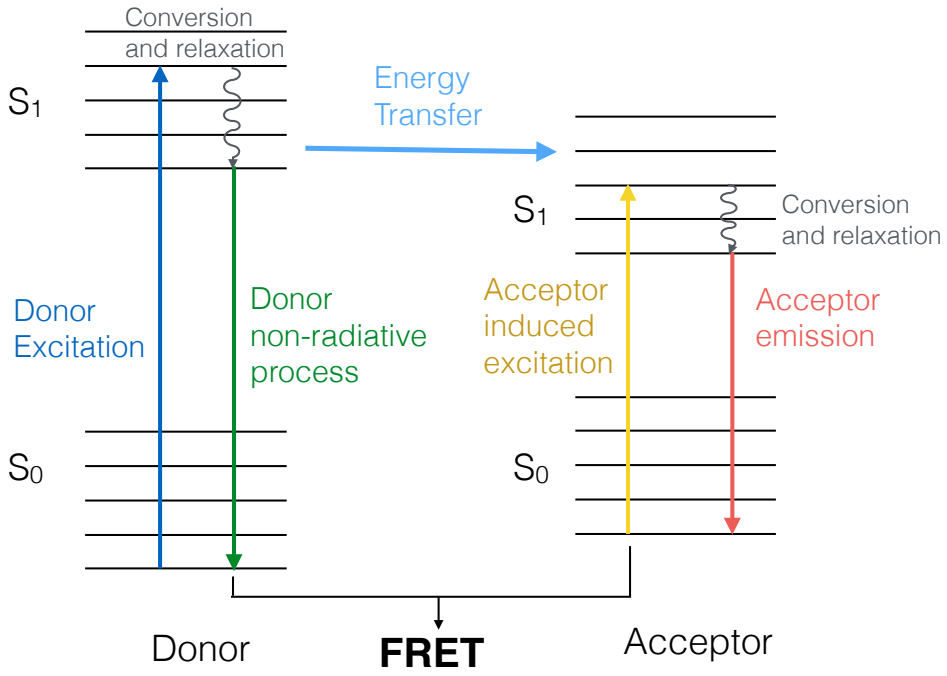


Figure 2.6: Illustration of the FRET process based on the Perrin-Jablonski diagram. S_0 and S_1 represent the ground state and excited state of fluorophores, respectively. the excitation and emission of electrons are represented by straight arrows. The non-fluorescent conversion of electrons within the same state are represented by curved arrows.

Traditionally, the extent of FRET (efficiency E) is calculated from the fluorescence intensity measured for the donor and the acceptor. The intensity-based methods are subjected to many artefacts such as photobleaching, donor-acceptor photoconversion, and the uneven distributions of fluorophores. These common pitfalls can be avoided by using the recently developed fluorescence lifetime imaging (FLIM) method for the measurement of FRET. In FRET-FLIM, the FRET efficiency E can be directly calculated from the fluorescence lifetime of the donor (τ_D) at FRET and non-FRET situations. Due to its high sensitivity and spatial-temporal resolution, FRET-FLIM is well-suited for probing

transient protein–protein interactions in live cells (105, 106).

In brief, FRET by fluorescence lifetime imaging microscopy (FLIM) calculates the extent of FRET from the discrepancy between the donor's fluorescence lifetime when existing alone (τ_D) and the lifetime in the presence of acceptors (τ_F). In a non-FRET situation, the donor's rate of fluorescence decay $i(t)$ can be expressed as:

$$i(t) = \alpha_D e^{-t/\tau_D} \tag{2.21}$$

,

where α_D is the amplitude of the donor, i.e. the initial fluorescence at $t = 0$. When FRET occurs, a portion of the donors undergo FRET resulting in a shortened lifetime while the other portion remain unaffected. This situation results in a bi-exponential decay:

$$i(t) = \alpha_D e^{-t/\tau_D} + \alpha_F e^{-t/\tau_F} \tag{2.22}$$

in which τ_F is the FRET lifetime of the donor and α_F is the FRET amplitude. When normalizing the amplitudes in Eq.2.22 to one, we arrive at the fraction of interacting donors, f_D:

$$i(t) = (1 - f_D) e^{-t/\tau_D} + f_D e^{-t/\tau_F} \tag{2.23}$$

in which:

$$f_D = \frac{\alpha_F}{\alpha_F + \alpha_D} \tag{2.24}$$

.

The fraction of interacting donors (f_D) is a convenient parameter to describe protein–protein interactions in a cellular context, as it directly quantifies the percentage of interacting donor-labelled proteins and can be attributed to each pixel in the FLIM image.

2.8.2 The Quantification of Protein–Protein Interactions by FRET-FLIM

One of the greatest advantages of FRET-FLIM is that the extent of protein–protein interactions can be measured within each pixel of an image. The level of FRET for each pixel is best described by two inter-related parameters, the mean fluorescence lifetime ($\langle\tau\rangle$) and the fraction of interacting donors (f_D). In a fluorescence decay $i(t)$, the mean fluorescence lifetime ($\langle\tau\rangle$) is defined by the following equation:

$$\langle\tau\rangle = \frac{\int t \cdot i(t)\, dt}{\int i(t)\, dt} \tag{2.25}$$

In practice, $\langle\tau\rangle$ can be obtained from a stack of images recorded through a period of time:

$$\langle \tau \rangle = \frac{\sum \Delta t_i \cdot I_i}{\sum I_i} \tag{2.26}$$

where Δt_i represents the time intervals between the ith acquired image and the laser pulse, and I_i represents the intensity of each pixel of the ith image.

Having obtained the mean fluorescence lifetime ($\langle \tau \rangle$) for each pixel, the fraction of interacting donor (f_D) for each pixel can be calculated from the following equation:

$$f_D = \frac{\left(1 - \frac{\langle \tau \rangle}{\tau_D}\right)}{1 - \frac{\langle \tau \rangle}{\tau_D} - \left(\frac{\tau_F}{\tau_D}\right)^2 + \left[\frac{\langle \tau \rangle}{\tau_D} + \frac{\langle \tau \rangle}{\tau_D}\frac{\tau_F}{\tau_D}\right]} \tag{2.27}$$

in which the donor lifetime when existing alone (τ_D) and the donor's FRET lifetime (τ_F) can be obtained by fitting the donor's fluorescence decay profile with a single and a double exponential decay model (Eq. 2.21 and Eq. 2.22).

2.8.3 Experimental Procedures

In this thesis, the cyan fluorescent protein (CFP) variant mTFP1 and the yellow fluorescent protein (YFP) variant mVenus were used as the donor and the acceptor, respectively. The fluorescent protein monomeric teal fluorescent protein 1 (mTFP1) is an ideal donor for FRET-FLIM measurements due to its mono-exponential fluorescence decay profile, a particularly long lifetime of 2.7 ns, and highly photostable properties (107). A flexible linker of 10 residues (RTLEVLFQGP) was used to connect the fluorophores to

the target proteins, allowing free rotations of the donor-acceptor pair to achieve optimal orientations for the FRET dipole–dipole coupling (Fig. 2.7.

COS-7 cells at 60-70% confluence grown in glass–bottom 35 mm Petri dishes (Mattek) were transiently transfected with PlxnA2-mTFP1, PlxnA2-mVenus, and the mTFP-mVenus tandem constructs using lipofectamine 2000. The COS-7 cells were then incubated in Dulbecco's modified eagle medium (DMEM) with 10% fetal calf serum (FCS). Before imaging, the medium was changed to PBS imaging buffer supplied with with 0.5% fetal bovine serum. A Leica SP8–SMD Leica microscope (Leica Microsystems, Germany) coupled with pulsed laser and time-correlated single photon counting (TCSPC) photon-counting devises was used for acquisition. A suitable image field focusing on the surface of a robust COS-7 cell was chosen with a 63×/1.4 numerical aperture (NA) oil immersion objective. To excite the donor (mTFP1), a 440 nm pulsed laser (Picoquant GmbH, Germany) tuned at 40 MHz and single photon counting electronics (PicoHarp 300) were used. Only the emitted photons from the donor were allowed to pass through a 460–500 nm filter and were subsequently detected by a HyD detector (Leica Microsystems). To rule out artefacts due to photo-bleaching and insufficient signal to noise, only cells with at least 700-1000 photons per pixel and negligible amount of bleaching were included in the analysis after a 2 × 2 image binning.

2.8.4 Data Processing

A time-correlated single-photon counting (TCSPC) method was employed to record the time intervals between each emitted photon and its pulsed excitation. By repeatedly exciting the donor with a pulsed laser and recording the lapse of time between laser excitation pulses and the arrival of photons, we generated a histogram of the number of photons allocated into different time windows. Over an appropriate course of time during which a desired number of emitted photons were collected, a histogram which best approximates the actual fluorescence decay curve of the donor was obtained. The acquired fluorescence decay, $i(t)$, of each pixel in image was deconvoluted with the instrument response function (IRF) and fitted by a Marquandt nonlinear least–square algorithm (108) with one- or two-exponential theoretical models using Symphotime software (Picoquant GmbH). The mean fluorescence lifetime ($\langle\tau\rangle$) and fraction of interacting donor (f_D) were calculated as described by Eq. 2.26 and Eq. 2.27 (106, 109). In the resulting image, the intensity of each pixel represents the fraction of interacting donors (f_D) or the mean fluorescence lifetime of all donors inside that pixel ($\langle\tau\rangle$). This method was only capable of detecting ~1/4 of the real interaction (1/2 because of the labelling strategy and 1/2 given the dynamic range of the FRET couple (107). This factor together with the spectral heterogeneity of the fluorophores and the temporal resolution of the technique determined that the f_D only represents the minimal fraction of the actual interaction (106). Statistical analysis is done using the Origin software (OriginLab, MA) and the ANOVA test was used to examine the discrepancy between each set of data.

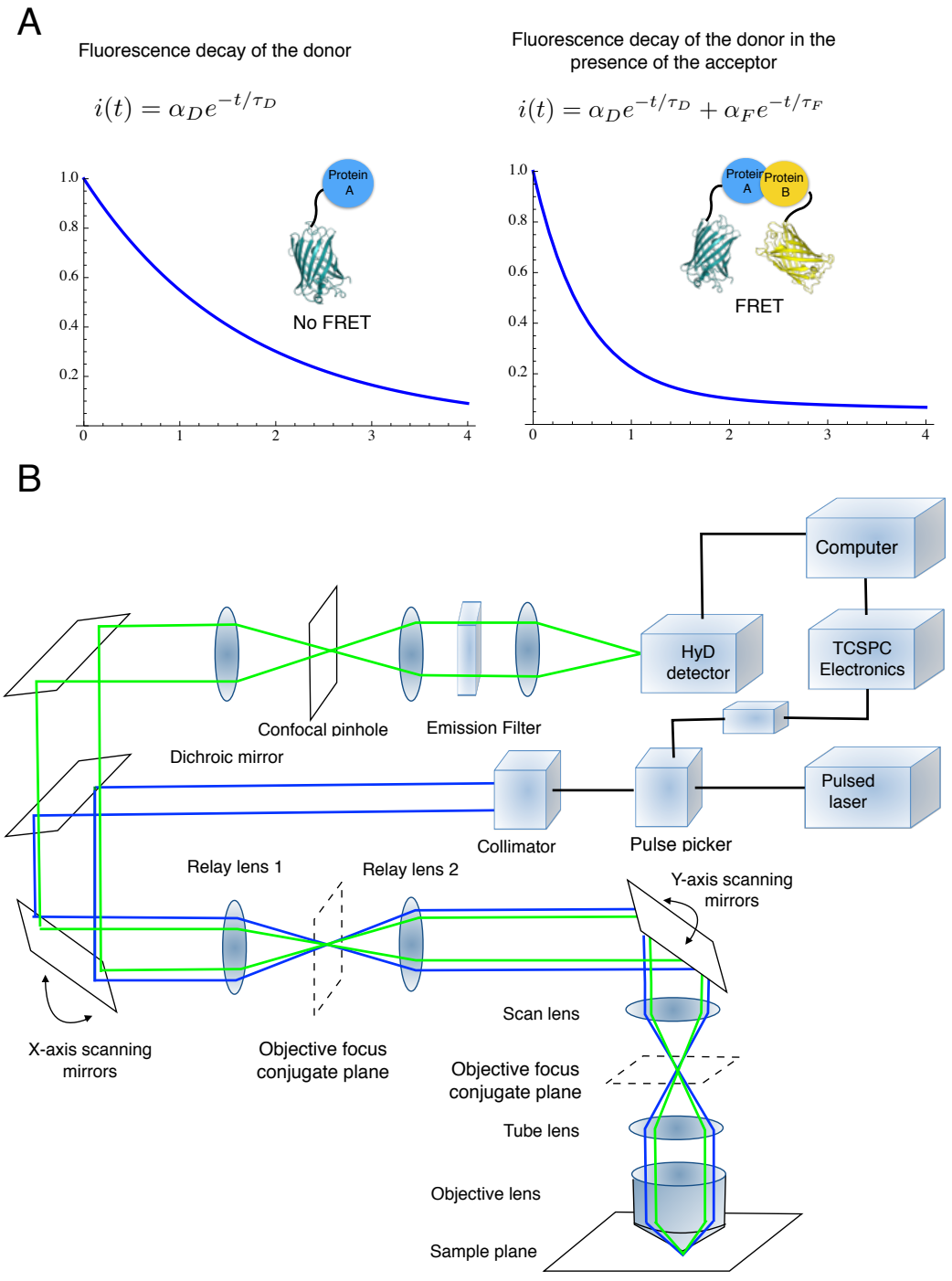


Figure 2.7: The experimental setup of FRET-FLIM. (A) Illustration of the theoretical fluorescence decay profiles of the donor mTFP1 alone (left) and in the presence of the acceptor mVenus (right). (B) Schematic Illustration of the FRET-FLIM setup based on the Leica TCS SP8 microscope system. The excitation and emission lightpaths for mTFP1 are coloured in blue and green, respectively.

2.9 Protein Cluster Analysis Based on Localisation Microscopy

2.9.1 Overview of Localisation Microscopy

In brief, localisation microscopy is a super-resolution microscopy technique which exploits the stochastic properties of photo-activated fluorophores to derive the positions of single fluorescent molecules with high accuracy. In the localisation microscopy experiments discussed here, a laser with high intensity is used to induce all fluorophores into a long-live dark state. Over a course of time, a set of stochastically-returned photons for each fluorescent molecule can be measured, which typically accumulate into centroid distributions. These distributions can be fitted with a two-dimensional Gaussian model, the centers of which represent the positions of single fluorescent molecules at a resolution limit of ~50 nm.

2.9.2 Overview of Protein Cluster Analysis

As the localisation microscopy-based method pinpoints the positions of single protein molecules on the cell membrane, these single molecules can be subjected cluster analysis by analyzing the intermolecular distances using the Ripley's K-function (110):

$$\widehat{K}(h) = \frac{1}{\hat{\lambda} N} \sum_{i=1}^{N} \sum_{j \neq i} I(d_{ij} \leq h) \tag{2.28}$$

where λ is the density of points, r is the distance between points, d_{ij} is the measured distance between point i and point j. I is defined as:

$$I(d_{ij} \leq h) = \left\{ \begin{array}{l} 1 \text{ if } d_{ij} \leq h \\ \\ 0 \text{ otherwise} \end{array} \right\} \tag{2.29}$$

In short, the Ripley's K-function describes the frequency of observing clusters of a certain size for a set of two-dimensionally distributed points. For a more straightforward representation better-suited for protein cluster analysis we employ the Ripley's L-function:

$$L(h) = \sqrt{\frac{K(h)}{\pi}} \, . \tag{2.30}$$

When the Ripley's L-function is plotted over the intermolecular distances r ($L(r) - r$), for a simulated set of points following random (Poisson) distribution we obtain a diagonally ascending line. Any deviation from this line to larger frequencies ($L(r)$) signifies the presence of protein clusters.

2.9.3 Experimental Procedures

COS-7 cells were transiently transfected with PlxnA2 full-length constructs conjugated to mVenus at the C-termini using X-tremeGENE HP DNA Transfection Reagent (Roche) and incubated at 37 °C, 5% CO_2 for 24 hours before fixing with 4% paraformaldehyde (PFA) on ice. The fixed cells were mounted onto the glass coverslip using Prolong Gold^{TM} antifade reagent from Life Technologies^{TM}. A single molecule localisation microscopy technique based on the stochastic switching of standard fluorescent proteins (111), which was similar to the principle of (F)photoactivated localization microscopy (PALM) (112) and Stochastic optical reconstruction microscopy (STORM) (113) which rely on special photo-activatable or photo-switchable fluorophores, was employed. localisation microscopy was performed using an OMX (optical microscope experimental, V2, API) microscopic system modified to enable single molecule localisation microscopy using conventional fluorescent proteins according to the method described previously (111). The intensity of the 488 nm laser was adjusted to ~14 kW/cm^2 in the object plane for localisation microscopy imaging of the mVenus conjugated samples. Single molecule positions were determined using an algorithm based on a maximum-likelihood method optimized for localisation microscopy data (114) and adapted to the OMX hardware configuration. The positions of single molecules were calculated with an estimated mean localisation accuracy of < 21 nm with a maximum likelihood–based algorithm and the protein cluster analysis was performed based on the Ripley's L-function as previously described (25, 115).

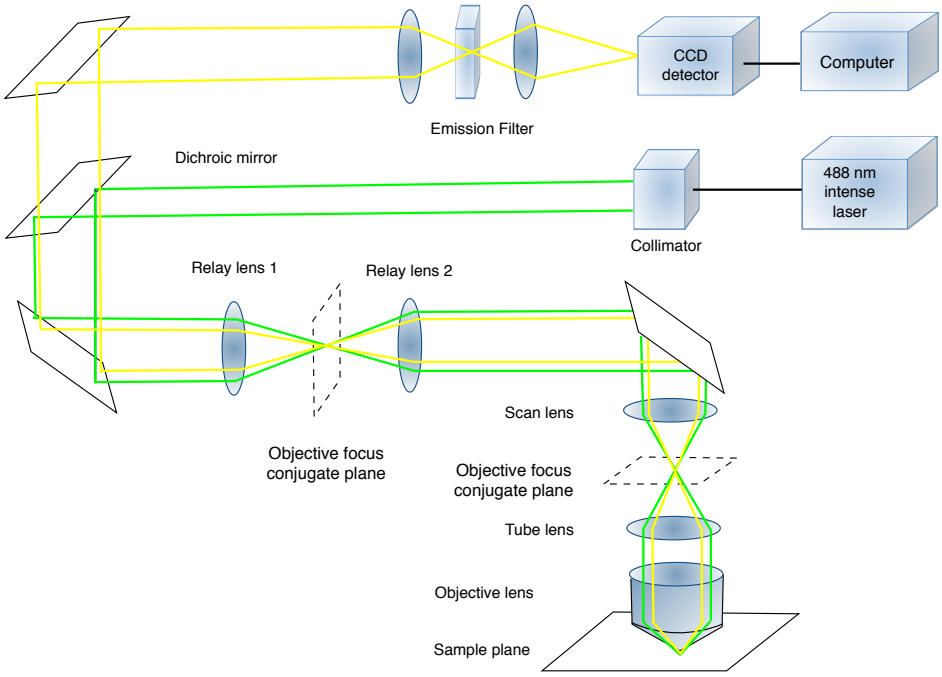


Figure 2.8: Illustration of the localisation microscopy setup. The basic components of the localisation microscopy setup based on the OMX hardware configuration are illustrated as cartoons. The excitation and emission fluorescence light paths for mVenus are coloured in green and yellow, respectively.

2.10 COS-7 Cell Collapse Assay

CV-1 in Origin with SV40 genes-7 (COS-7) cells were cultured in DMEM (high glucose), supplemented with 10% FCS, 1% Penicillin-streptomycin and L-glutamine. Cells were plated on coverslips at a density of $0.5-1\times10^4$ cell/cm^2 in 12-well plate. Following day, the fully confluence COS-7 cells were transfected with 0.5 µg of PlxnA4$_{\mathrm{FL}}$, PlxnA4$_{\mathrm{FL}}$ F689N/E691S, and PlxnA4$_{\Delta\mathrm{ecto}}$ in pcDNA 3.1(+) vector, together with 0.5 µg of eGFP us-

ing transfection reagent XtremeGENE^{TM} 9 (Roche) following manufacture's instruction. After incubated at 37 °C, 5% CO_2 for 24 hours, cells were fixed with 4% PFA (paraformaldehyde) and stained for eGFP. Images were acquired with 40 × objective with AxioScopeA1 (Zeiss) microscope. Cell area was measured using ImageJ (116). COS-7 cell collapse assay was performed in collaborations with Prof. Jeroen Pasterkamp's group in University Medical Center Utrecht.

2.11 Dentate Gyrus (DG) Growth Cone Collapse Assay

DG growth cone collapse assay was also performed in collaborations with Prof. Jeroen Pasterkamp's group in University Medical Center Utrecht. DG granule cell cultures were prepared according to Van Battum et al., 2014 (117). Briefly, hippocampi were dissected from p6-p8 C57b16J mice and cut into 350 µm slices using a tissue chopper (McIllwain). dentate gyrus (DG) was dissected from these slices and collected in 1 × Krebs medium (0.7% sodium chloride, 0.04% potassium chloride, 0.02% monopotassium phosphate, 0.2% sodium bicarbonate, 0.25% glucose and 0.001% phenol red). Cells were dissociated after incubation in 1 × trypsin in Krebs medium for 15 min at 37 °C, the reaction was stopped by adding 20 mg/ml soyabean trypsin inhibitor followed by trituration using a firepolished Pasteur pipette in Krebs medium with 2 mg soyabean trypsin inhibitor and 20 µg/ml DNAseI. Dissociated cells were resuspended in neurobasal medium with B-27® supplement, L-glutamine, penicillin/streptomycin and β-mercaptoethanol and plated on coverslips coated with poly-D-lysine (20 mg/ml) and laminin (40 µg/ml) in 24

well plate and incubated in a humified incubator at 37 °C and 5% CO_2. On DIV1 (day *in vitro* 1) cells were treated with wild type PlxnA1$_{4\text{-}5}$ and mutant PlxnA1$_{4\text{-}5}$ F693N/E695S recombinant proteins (3 mg/ml) for 30 min at 37 °C. Proteins were pre-filtered with 0.45 µm filter before applying onto cells. Cells were fixed in 4% PFA and sucrose for 20 min at room temperature and immunostained with βIII-tubulin and Phalloidin. Images were acquired with 100 × objective with AxioScopeA1 (Zeiss) microscope. Using a scale of 1 to 10 where 1 is healthy growth cone and 10 is a collapsed growth cone basing on the surface area, growth cones were quantified between wild type PlxnA1$_{4\text{-}5}$, mutant PlxnA1$_{4\text{-}5}$ F693N/E695S and untreated condition using ImageJ.

CHAPTER 3

STRUCTURAL STUDIES OF PLXNA EXTRACELLULAR SEGMENTS

3.1 Introduction

The four members of the class A plexins, PlxnA1-A4, particularly important in the neural, cardiovascular, skeletal and immune systems, constitute the best-characterised plexin receptors in vertebrates (39, 41). The PlxnAs can bind to transmembrane class 6 semaphorins (Sema6s) independently or to secreted class 3 semaphroins (Sema3s) via obligatory coreceptor neuropilins 1 and 2 (Nrp-1 and -2) (Fig. 3.1) (17). Recent studies have also uncovered the interactions between transmembrane class 5 semaphorins and PlxnA2, a process potentially mediated by cofactor HSPGs and CSPGs (46). The

downstream binding partners of PlxnAs include several small GTPases such as Rap, Ras and RND1, as well as the microtubule-associated redox enzyme MICAL (118). The PlxnA cytoplasmic domains can also be tyrosine phosphorylated in overexpressing cells, linking the activation of PlxnAs to several potential receptor tyrosine kinases such as MET or non-receptor kinases such as Fes and Fyn (119) although whether tyrosine phosphorylation of PlxnAs is dependent on semaphorin-binding remains elusive (120). Typically, semaphorin signalling through PlxnAs directly changes cytoskeleton dynamics and thus, mediates cell motility. The PlxnAs have also been shown to crosstalk with many other TM1 receptors such as integrins (17, 39, 41).

Despite our current knowledge on the semaphorin–plexin recognition mode and the basis of plexin Rap-GAP activation in the cytoplasmic region, we lack a complete picture on how bivalent semaphorin-binding in the plexin ectodomain is translated into the dimerisation-based activation of its cytoplasmic segments. This is largely due to the missing structural information on the full-length ectodomains of plexins. In particular, recent studies on PlxnAs have also uncovered uncanonical modes of semaphorin-binding, transmembrane class 5 and 6 semaphorins were shown to interact with PlxnAs in *cis* resulting in receptor activation and silencing, respectively (46, 66, 67). However, the current structures of PlxnA ectodomains only cover the N-terminal segments (four out of the ten domains for PlxnA2). When forming the recognition complex with class 3 and 6 semaphorins the C-termini of these partial segments of PlxnA ectodomains diverge. Thus, how the PlxnAs, and indeed all plexins, are activated upon dimerisation remains puzzling.

Understanding the activation and regulation of PlxnA signalling requires structural insights into the full-length PlxnA extracellular segments. Prior to this thesis work a previous member of our group, Dr. Bert J.C. Janssen, obtained a total of thirteen crystals of PlxnA1, A2 and A4 ectodomains which diffracted to 4-10 Å resolutions. Although initial crystal structures could be determined, refinement of these structures proved challenging due to limited resolutions. The major obstacle was the absence of high-resolution structural models for more than half of the PlxnA stalk region (domains 4-10). Therefore, this thesis began with crystallographic studies on the PlxnA ectodomain, during which time high-resolution crystal structures for two segments in the stalk region, PlxnA2$_{4\text{-}5}$ and PlxnA1$_{7\text{-}10}$, were obtained. After the PlxnA full ecotodmain structures were refined by Dr. Janssen aided by these new crystal structures, we found a remarkable ring-like conformations common for all PlxnAs. My subsequent negative stain EM experiments showed that the ring-like conformation is indeed the preferred conformation for PlxnA ectodomains, while also uncovering an alternative, less-frequent, chair-like (twisted-open) conformation. By superposing our PlxnA ectodomain structures onto the existing Sema6A–PlxnA2 and Sema3A–PlxnA2–Nrp1 recognition complexes, we found that the structures of PlxnA ectodomains may provide an elegant solution to the long-standing question regarding PlxnA activation, in which the distinct ring-like conformation of the PlxnA ectodomains translate semaphorin-binding in the extracellular space into activation in the PlxnA cytoplasmic regions. I also hypothesized that the alternative chair-like conformation of PlxnA ectodomains may facilitate *cis*-interactions with transmembrane Sema6s.

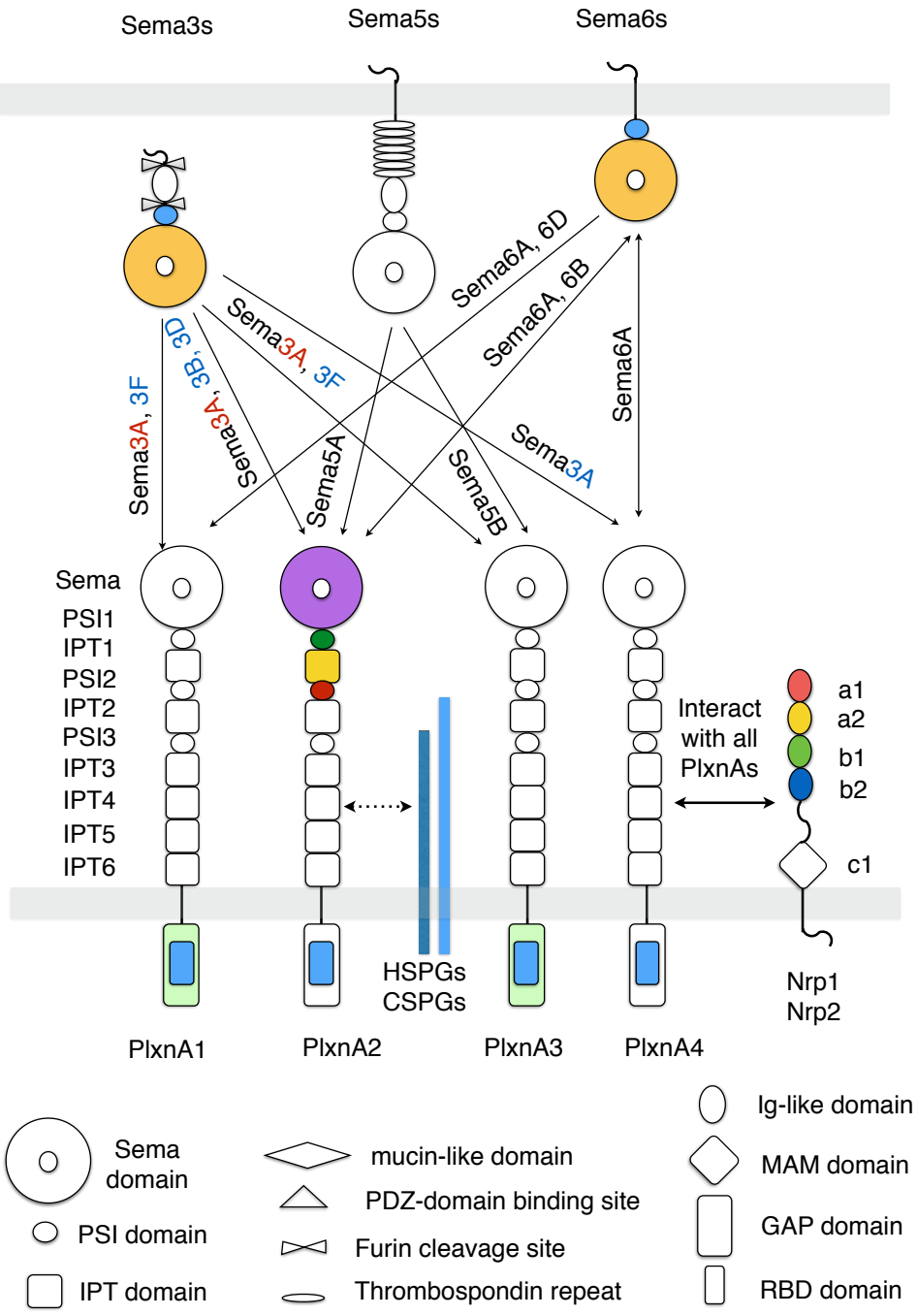


Figure 3.1: The Vertebrate class A plexins. PlxnA1-A4 together with their cognate semaphorin ligands, coreceptor neuropilins and cofactor HSPGs and CSPGs are illustrated. Among the proteins illustrated, only domains with available crystal structures are coloured.

3.2 The 1.36 Å Crystal Structure of PlxnA2$_{4\text{-}5}$

In order to obtain high-resolution crystal structures of the PlxnA stalk region, a wide variety of constructs consisting of different domain combinations in the mouse PlxnA1 and A2 ectodomains were systematically screened. The first successful crystallisation hits yielded a 1.36 Å resolution crystal structure of PlxnA2$_{4\text{-}5}$ (domains PSI2-IPT2) of space group C2. The mouse PlxnA2$_{4\text{-}5}$ was produced in HEK293S cells and purified with IMAC and SEC using standard protocols (as detailed in Chapter 2). After deglyscosylation using endoglycosidase F1 (Endo F1), the purified PlxnA2$_{4\text{-}5}$ was concentrated to 6.5 mg/ml in 10 mM HEPES pH 7.5, 150 mM sodium chloride, and crystallised in 25% (w/v) polyethylene glycol 3,350 and 100 mM BIS-TRIS pH 5.5. Before diffraction data collection crystals were soaked in mother liquor supplemented with 25% (v/v) glycerol. The purification and crystallisation procedures for PlxnA2$_{4\text{-}5}$ are summarised in Fig.3.2.

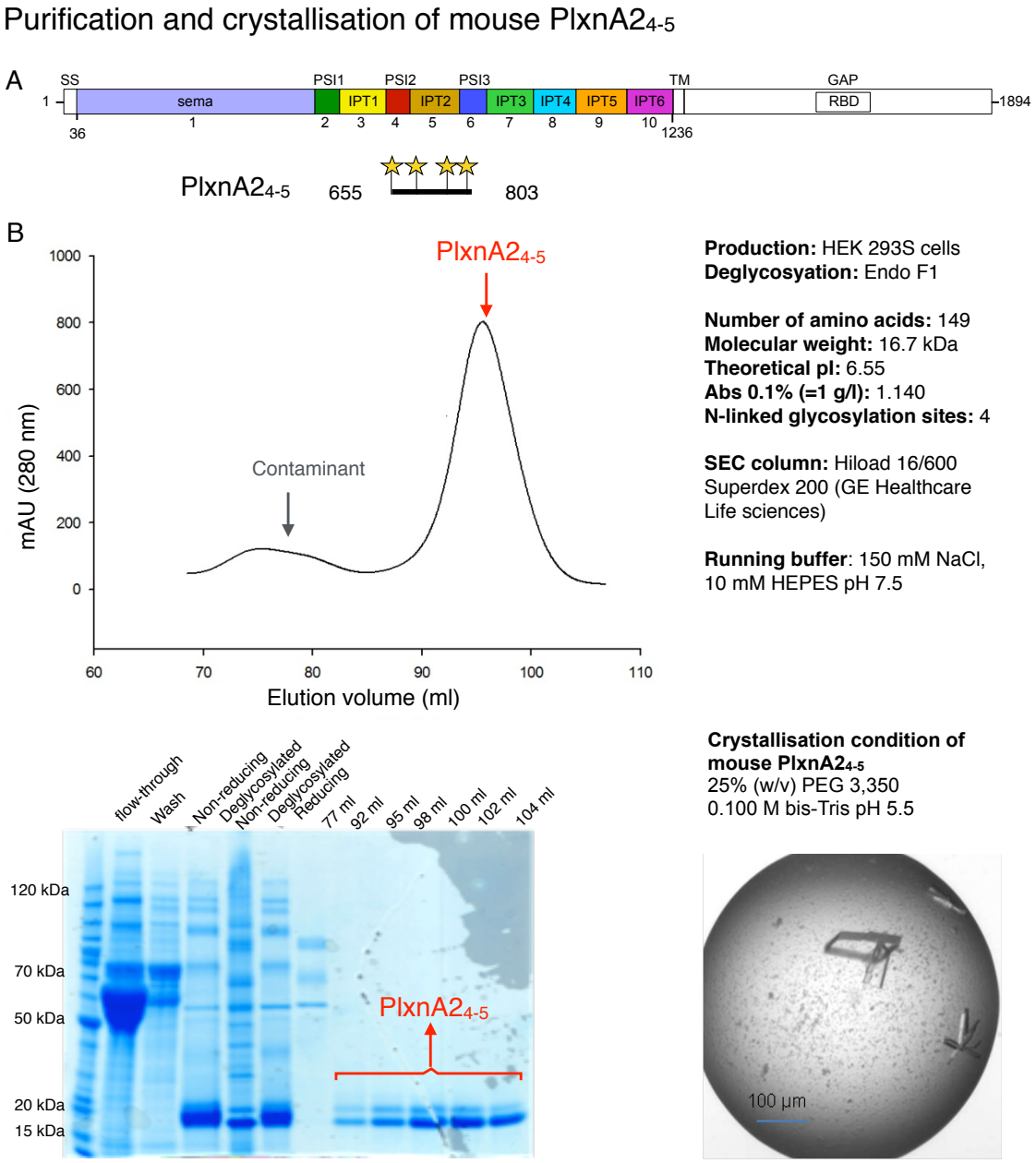
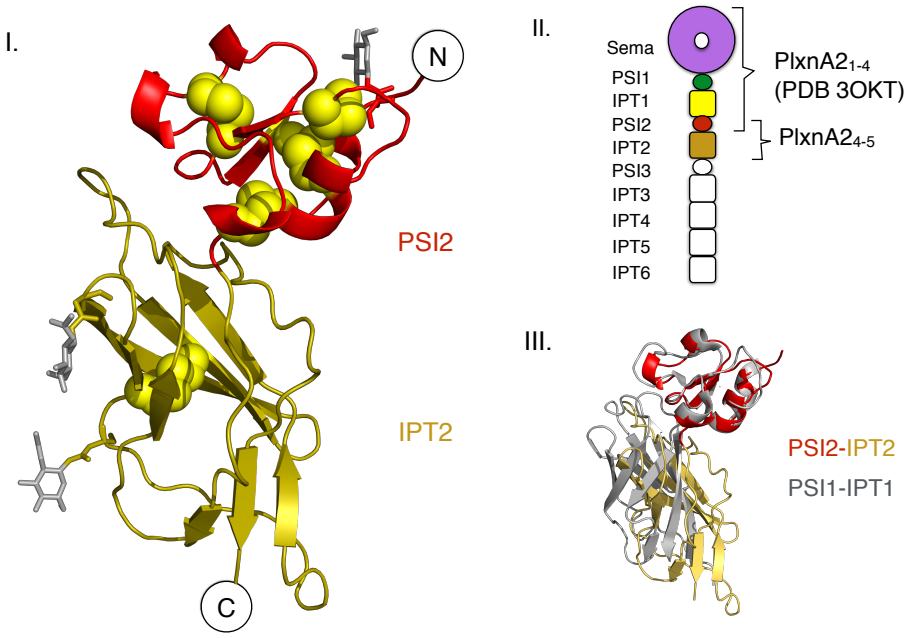


Figure 3.2: Summary of the purification and crystallisation of mouse PlxnA2$_{4-5}$. (A) Schematic illustration of the domain organization of mouse PlxnA2 and the PlxnA2$_{4-5}$ construct. The black line underneath the coloured domains of mouse PlxnA2 represents the segment covered by the current PlxnA2$_{4-5}$ construct (numbers at each end indicated the N-terminal and C-terminal residue, respectively). Yellow stars indicate the positions of predicted N-linked glycans. (B) The elution profile and SDS-PAGE analysis of the SEC experiment conducted for mouse PlxnA2$_{4-5}$. The basic information, production and SEC experimental conditions of PlxnA2$_{4-5}$ are listed at the right of the middle panel. The crystallisation details of PlxnA2$_{4-5}$ are listed at the bottom right.

The 1.36 Å crystal structure of PlxnA2$_{4\text{-}5}$ was solved by molecular replacement (MR) in PHASER (83) using the highly conserved domains 4-5 segment (PSI2-IPT2) of a PlxnA1 ectodomain structure (4 Å resolution, space group $P4_32_12$) previously determined by Dr. Janssen as initial model. The crystallographic and refinement statistics are listed in Fig. 3.3.B. The PlxnA2 PSI2 (domain 4) displays a small cysteine-knot motif typical for the plexin and semaphorin PSI domains, in which 45 residues are locked together by four pairs of cysteine bonds (36). The subsequent PlxnA2 IPT2 (domain 5) has an Ig (immunoglobin)-like fold based on β-sheet-barrel backbones charactersitic for the plexin IPT domains (36, 37). The relative orientation between PlxnA2$_{4\text{-}5}$ (PSI2-IPT2) in the current crystal structure is slightly different from that of the previously characterised PlxnA2$_{2\text{-}3}$ (PSI1-IPT1) (PDB 3OKT) (36) (Fig. 3.3.A).

A The 1.36 Å crystal structure of PlxnA2$_{4-5}$ (space group C2)



B Crystallographic and refinement statistics of PlxnA2$_{4-5}$

Data collection	**PlxnA2$_{4-5}$**
Space group	*C2*
Cell dimensions	
a, *b*, *c* (Å)	107.7, 44.6, 33.0
α, β, γ(°)	90, 104.7, 90
Resolution (Å)	52.1-1.36(1.40-1.36)[a]
R_{merge} (%)	3.8 (47.6)
$I/\sigma I$	17.0 (2.3)
Completeness (%)	85.5 (41.0)
Redundancy	3.7 (3.4)

Refinement	**PlxnA2$_{4-5}$**
Resolution (Å)	52.1-1.36
Number of	27858
R_{work}/R_{free} (%)	14.1/19.6
Number of atoms	1334
Protein	1228
Ligand/ion	3
Water	103
B factors (Å^2)	
Protein	16.3
Ligand/ion	24.9
Water	29.6
R.m.s. deviations	
Bond lengths (Å)	0.020
Bond angles (°)	2.00

Figure 3.3: The 1.36 Å crystal structure of mouse PlxnA2$_{4-5}$. (A) The 1.36 Å crystal structure of PlxnA2$_{4-5}$ (PSI2-IPT2) shown in ribbon representation (red and dark yellow) (I) and superposed onto the previously characterised PSI1-IPT1 (PlxnA2$_{2-3}$) (grey) (PDB 3OKT) based on the two PSI domains (III). Illustration of PlxnA2$_{4-5}$ in relation to the PlxnA2 full ectodomain is shown in (II). The N-linked glycans are represented as grey sticks and cysteines as yellow balls. (B) Crystallographic and refinement statistics of mouse PlxnA2$_{4-5}$ crystal structures. The superscript a (^{a}) indicates that statistics of the highest resolution shell is shown in the parenthesis.

3.3 The 2.2 Å Crystal Structure of PlxnA1$_{7\text{-}10}$

We next obtained a crystal structure of mouse PlxnA1$_{7\text{-}10}$ (domains IPT3-IPT6) at 2.2 Å resolution (space group C222$_1$). This crystal structure has greatly aided the refinement of PlxnA ectodomain crystal structures, as it allowed us to trace the C-terminal domains particularly IPT6 (domain 10), a process otherwise obstructed by missing electron densities. Mouse PlxnA1$_{7\text{-}10}$ was produced in HEK293S cells and purified with IMAC and SEC using standard protocols (as detailed in Chapter 2). As native PlxnA1$_{7\text{-}10}$ did not crytallise despite extensive screening, I performed lysine methylation for PlxnA1$_{7\text{-}10}$ as a routine strategy to reduce the flexibility of surface lysines using dimethylamineborane complex and formaldehyde (121). After deglyscosylation using Endo F1, the lysine-methylated PlxnA1$_{7\text{-}10}$ was concentrated to 7.5 mg/ml in 10 mM TRIS, pH 8.0, 150 mM sodium chloride, and crystallised in 25% (w/v) polyethylene glycol 3,350, 100 mM BIS-TRIS pH 5.5 and 200 mM potassium sodium tartrate. Before diffraction data collection crystals were soaked in mother liquor supplemented with 25% ethylene glycol. The purfication and crystallisation precedures for PlxnA1$_{7\text{-}10}$ are summarised in Fig.3.4.

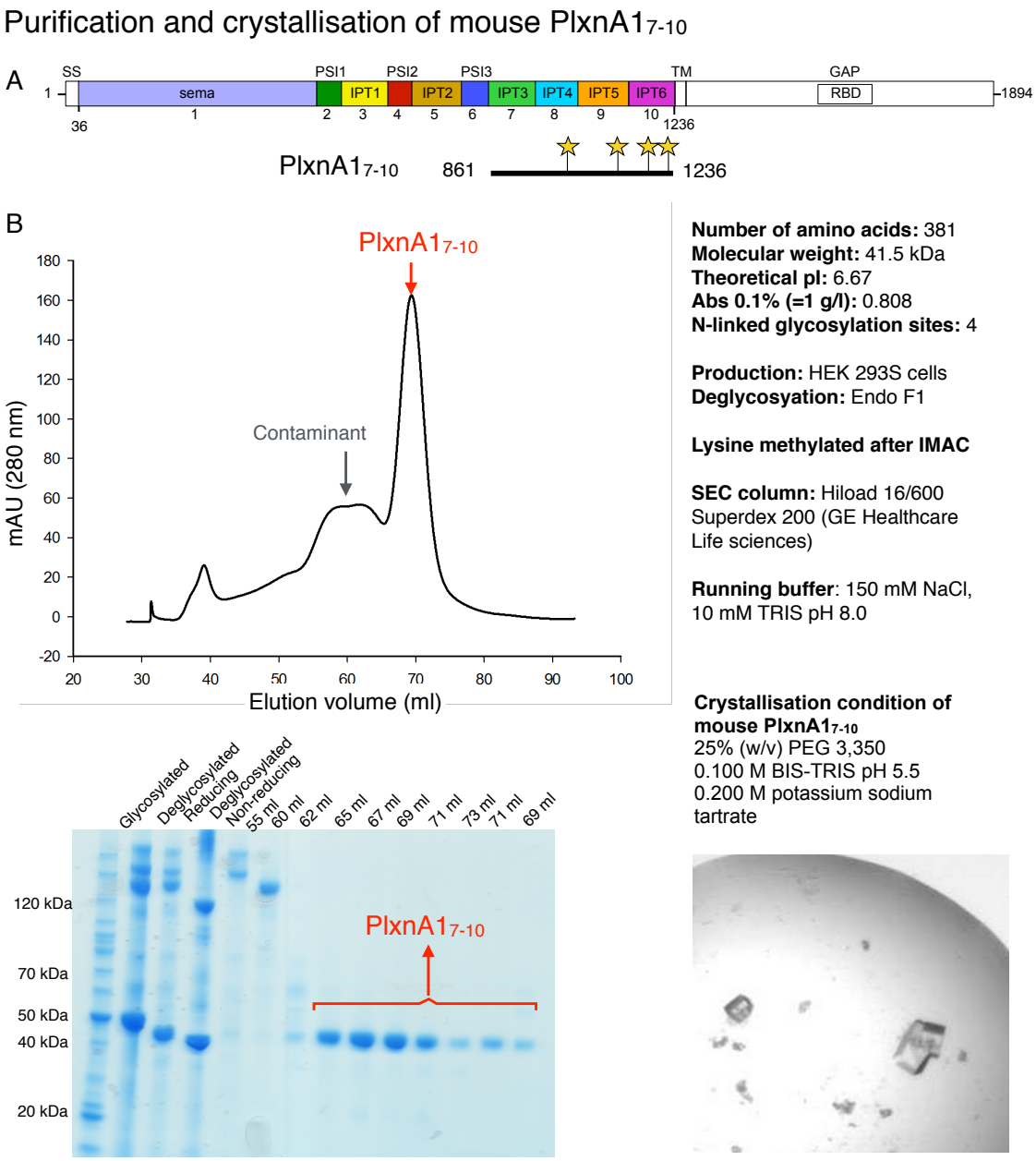
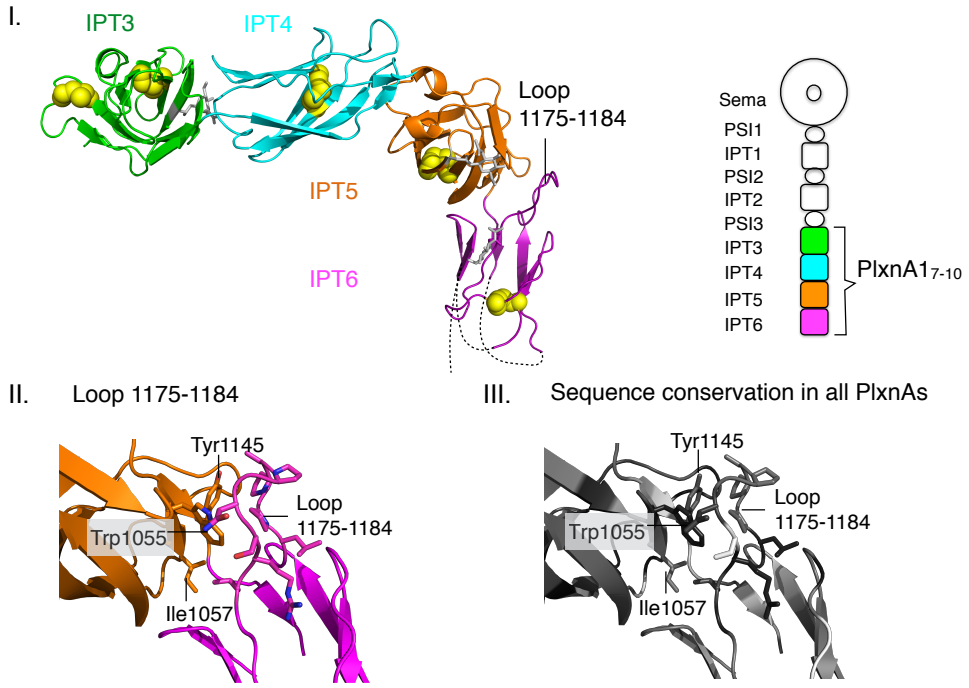


Figure 3.4: Summary of the purification and crystallisation of mouse PlxnA1$_{7\text{-}10}$. (A) Schematic illustration of the domain organization of mouse PlxnA1 and the PlxnA1$_{7\text{-}10}$ construct. The black line underneath the coloured domains of mouse PlxnA1 represents the segment covered by the current PlxnA1$_{7\text{-}10}$ construct (numbers at each end indicated the N-terminal and C-terminal residue, respectively). Yellow stars indicate the positions of predicted N-linked glycans. (B) The elution profile and SDS-PAGE analysis of the SEC experiment conducted for mouse PlxnA1$_{7\text{-}10}$. The basic information, production and SEC experimental conditions of PlxnA1$_{7\text{-}10}$ are listed at the right of the middle panel. The crystallisation details of PlxnA1$_{7\text{-}10}$ are listed at the bottom right.

The 2.2 Å crystal structure of PlxnA1$_{7\text{-}10}$ was solved by molecular replacement (MR) in PHASER (83) using the domains 7-10 (IPT3-IPT6) segment of a PlxnA1 ectodomain structure (4 Å resolution, space group $P4_32_12$) previously determined by Dr. Janssen as initial model. The crystallopgraphic and refinement statistics are listed in Fig. 3.5.B. In the PlxnA1$_{7\text{-}10}$ crystal structure, the four consecutive IPT domains, IPT3-IPT6, display the Ig-like folds characteristic of plexin IPT domains (Fig. 3.5.A). Interestingly, the first three IPT domains, IPT3-IT5, make a gentle curve on the same plane from which the last domain, IPT6, points almost perpendicularly away. In particular, IPT6 (domain 10) stands out from all PlxnA IPT domains for its shortened β-sheet-barrel backbones despite having a similar number of ~100 residues as other IPTs. A distinct, 10-residue loop extending from the second β-sheet of IPT6 directly contacts the outer edge of IPT5 (domain 9) (Fig. 3.5.A.II). This particular loop (1175-1184, LPPAPGNSRL), is most likely to provide hydrophobic interactions to three residues in IPT5, Trp1055, Ile1057 and Tyr1145, which are highly conserved across all vertebrate PlxnAs (Fig. 3.5.A.III). Therefore, this loop may be used to sustain the distinct ~90° turn made by IPT6 from the IPT3-IPT5 curve. One should note that the correct positions of 29 residues at the C-terminal of IPT6 could not be determined due to the absence of electron density, a result from the lack of crystal contacts in this particular region.

A The 2.2 Å crystal structure of PlxnA1$_{7\text{-}10}$ (space group C222$_1$)



B Crystallographic and refinement statistics of PlxnA1$_{7\text{-}10}$

Data collection	**PlxnA1$_{7\text{-}10}$**
Space group	*C222$_1$*
Cell dimensions	
a, *b*, *c* (Å)	75.9, 85.2, 125.6
α, β, γ(°)	90, 90, 90
Resolution (Å)	62.8-2.2 (2.27-2,20)[a]
R_{merge} (%)	12.0 (82.5)
I/σ*I*	9.2 (1.9)
Completeness (%)	100 (99.8)
Redundancy	8.1 (5.0)

Refinement	**PlxnA1$_{7\text{-}10}$**
Resolution (Å)	62.8-2.2
Number of	20219
R_{work}/R_{free} (%)	20.7/25.3
Number of atoms	2841
Protein	2703
Ligand/ion	8
Water	130
B factors (Å^2)	
Protein	58.8
Ligand/ion	78.0
Water	51.0
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.82

Figure 3.5: The 2.2 Å crystal structure of mouse PlxnA1$_{7\text{-}10}$. (A) The 2.2 Å crystal structure of PlxnA1$_{7\text{-}10}$ (IPT3-IPT6) in ribbon representation (I). The missing residues are illustrated as black dotted lines. Illustration of PlxnA1$_{7\text{-}10}$ in relation to the PlxnA1 full ectodomain is shown at the right. Structural details and sequence conservation of loop 1175-1184 (LPPAPGNSRL) are shown in (II) and (III), respectively. The residues in (III) are coloured from nonconserved, white, to conserved, black. The N-linked glycans are represented as grey sticks and cysteines as yellow balls. (B) Crystallographic and refinement statistics of mouse PlxnA1$_{7\text{-}10}$ crystal structures. The superscript a (^{a}) indicates that statistics of the highest resolution shell is shown in the parenthesis.

3.4 Crystal Structures of PlxnA Ectodomains Revealed a Ring-like Conformation

3.4.1 All PlxnA Ectodomain Crystal Structures Have a Ring-like Conformation

With the help of my high-resolution crystal structures of PlxnA2$_{4\text{-}5}$, PlxnA1$_{7\text{-}10}$, and the previously reported PlxnA2$_{1\text{-}4}$ (PDB 3OKT) (36), Dr. Janssen were able to refine the medium- to low-resolution crystal structures of PlxnA extracellular segments. A total of thirteen independent crystal structures of PlxnA1, A2 and A4 ranging in maximum resolution from 4 to 10 Å were derived from eight crystal forms (Fig.3.6). In five out of the thirteen full extracellular structures all ten domains were resolved whereas in the other eight structures for some of the C-terminal domains were not modelled due to fragmentary or absent electron density. This indicates that some flexibility is present in the extracellular segment of plexins, in particular the C-terminal region. In all crystal structures, individual domains in the PlxnA extracellular segments are arranged in a beads-on-a-string-like fashion, forming a distinct, ring-like conformation. In particular, the membrane proximal IPT6 (domain 10) points away from the ring formed by domains 1 to 9 (sema-IPT5). All domains in these PlxnA extracellular segments are joined by short linkers (up to two residues). Each interdomain interface contains a small hydrophobic core and a buried surface area ranging from 716 to 1172 Å^2. These substantial interdo-

main interactions likely contribute to conservation of the overall ring-like conformation for the three PlxnAs across different crystal forms. Indeed, the interdomain orientations between PSI2 and IPT2 (domains 4-5) as well as IPT3-IPT6 (domains 7-10) observed in the PlxnA full ectodomain structure are almost identical to those in my $PlxnA_{4\text{-}5}$ and $PlxnA1_{7\text{-}10}$ crystal structures (Fig. 3.5). Some flexibility in the PlxnA stalk, however, is evident when comparing the thirteen different structures. This flexibility contributes to the ring diameters, which range from 84 to 99 Å. Details of the production of the PlxnA ectodomain crystals as well as the structural solution and refinement of PlxnA ectodomains are described in Appendix 1.1 and 1.2, respectively. Crystallographic and refinement statistics for the PlxnA ectodomain structures are listed in Appendix Figure A.1 and A.2.

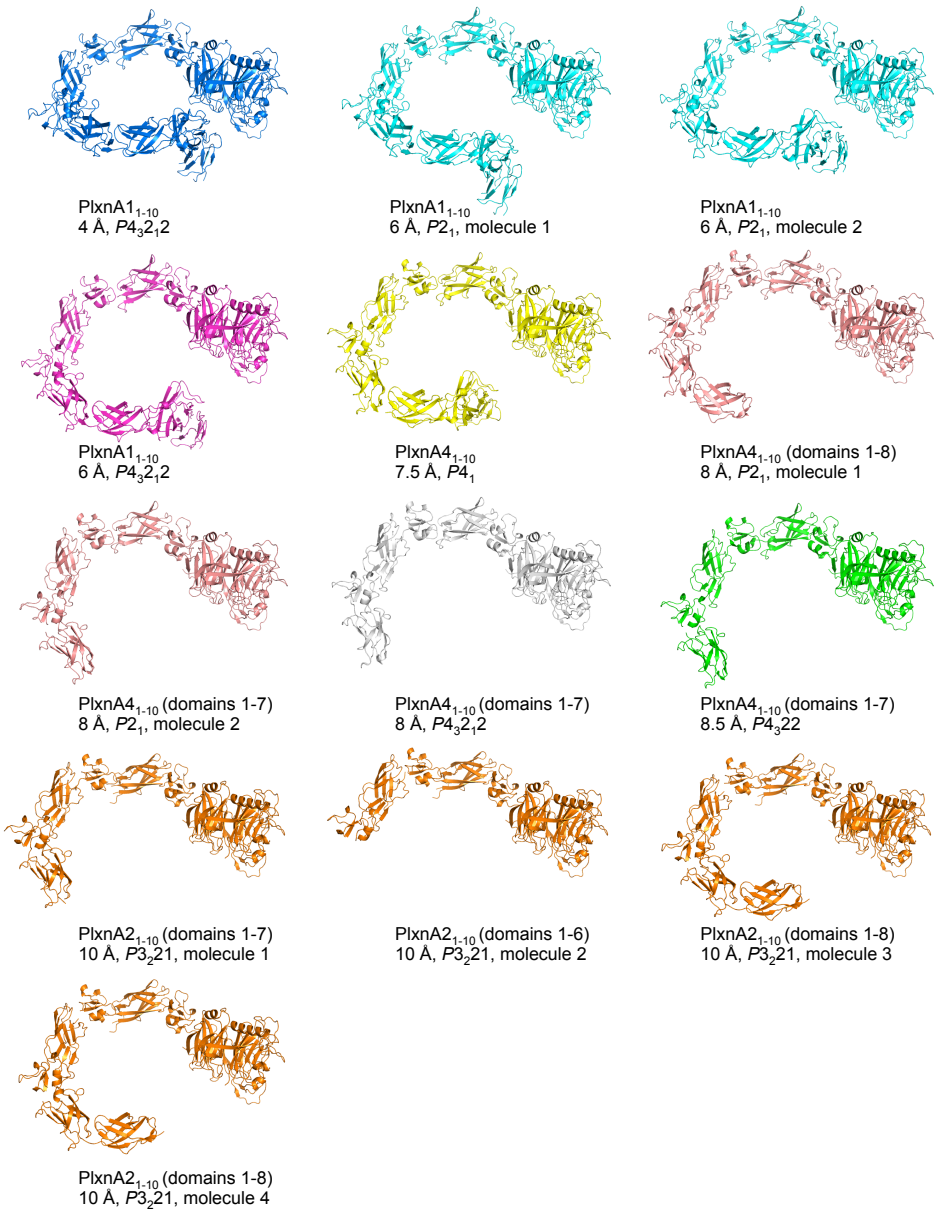


Figure 3.6: Crystal structures of PlxnA1, A2 and A4 extracellular segments. All of the thirteen crystal structures derived from eight different crystal forms for the extracellular segments of PlxnA1, PlxnA2 and PlxnA4 are shown in ribbon representation. Highest resolution and space group of each structure are indicated.

3.4.2 The 4 Å Crystal Structure of PlxnA1$_{1\text{-}10}$

The crystal structure of PlxnA ectodomains at highest resolution we determined is a 4 Å resolution one for PlxnA1$_{1\text{-}10}$ (space group $P4_32_12$) (Fig. 3.7.A). This dataset was obtained from dehydrating PlxnA1$_{1\text{-}10}$ crystals of the same space group which previously diffracted at best to 6 Å resolution. In the PlxnA1$_{1\text{-}10}$ crystal structure, the first nine domains form a 230 Å long ring-like stalk with a diameter of ~84 Å together with the sema head (domain 1). In the present structure we can clearly see that the out-of-the-ring orientation of IPT6 (domain 10), which is consistent with the high-resolution crystal structure for PlxnA1$_{7\text{-}10}$ (as shown in Fig. 3.5). One should note that the distinct orientation of IPT6 is observed in all five PlxnA ectodomain structures for which the electron density is sufficiently well ordered to allow IPT6 to be resolved (as shown in Fig. 3.6).The linker between PlxnA1 IPT6 and the transmembrane helix is predicted to be 6 residues and a similarly short length in all other PlxnAs. Therefore, should the extracellular segments of PlxnAs adopt the ring-like conformation *in vivo*, they are most likely to reside close to the cell surface with the rings orienting parallel to the plasma membrane. We can also hypothesize that should PlxnA1 orient its ring-like stalk relatively parallel to the cell surface to allow full exposure of the ligand-binding sema domain as shown in Fig. 3.7, it is possible that the a small area on the bottom half of the stalk, domains 2-9 (PSI1-IPT5), may interact with the plasma membrane. However, at the bottom half PlxnA1$_{1\text{-}10}$ stalk I cannot find basic patches, i.e. regions of positively charged cell surface residues, which can potentially interact with the negatively charged phospholipid head groups.

Interestingly, in the present $PlxnA_{1\text{-}10}$ structure as well as all other crystals of PlxnA extracellular segments, the interface between PSI3 and IPT3 (domains 6 and 7) differs markedly from those of PSI1-IPT2 (domains 1-2) and PSI2-IPT2 (domains 4-5). These differences allow PSI3 and IPT3 to impose a more acute angle on the relative arrangement of these two domains, and thus reducing the ring circumference (Fig. 3.7.B and C). Although a degree of structural breathing may result in small changes in the PlxnA stalk region, the overall ring-like conformation is likely to be rigid. One should note that only in this $PlxnA_{1\text{-}10}$ crystal structure did we observe a closed ring for PlxnA ectodomains, in which several residues in IPT5 (domain 9) make contacts with the sema domain (domain 1). As this compaction was not observed in other PlxnA ectodomain structures, it can be attributed to the dehydration procedure which induces compression to the overall PlxnA architecture. Thus, contrary to previous speculations on plexin ectodomain structure (22, 36, 122), the PlxnA ectodomain does not make any "head-down" intramolecular interaction in the unliganded state, that is, the sema domain does not make any contacts with the remainder of the plexin other than to the directly adjoining PSI1 domain. Similarly, all other PlxnA extracellular domains associate only via the consecutive domain-domain interactions (as shown in Fig. 3.6).

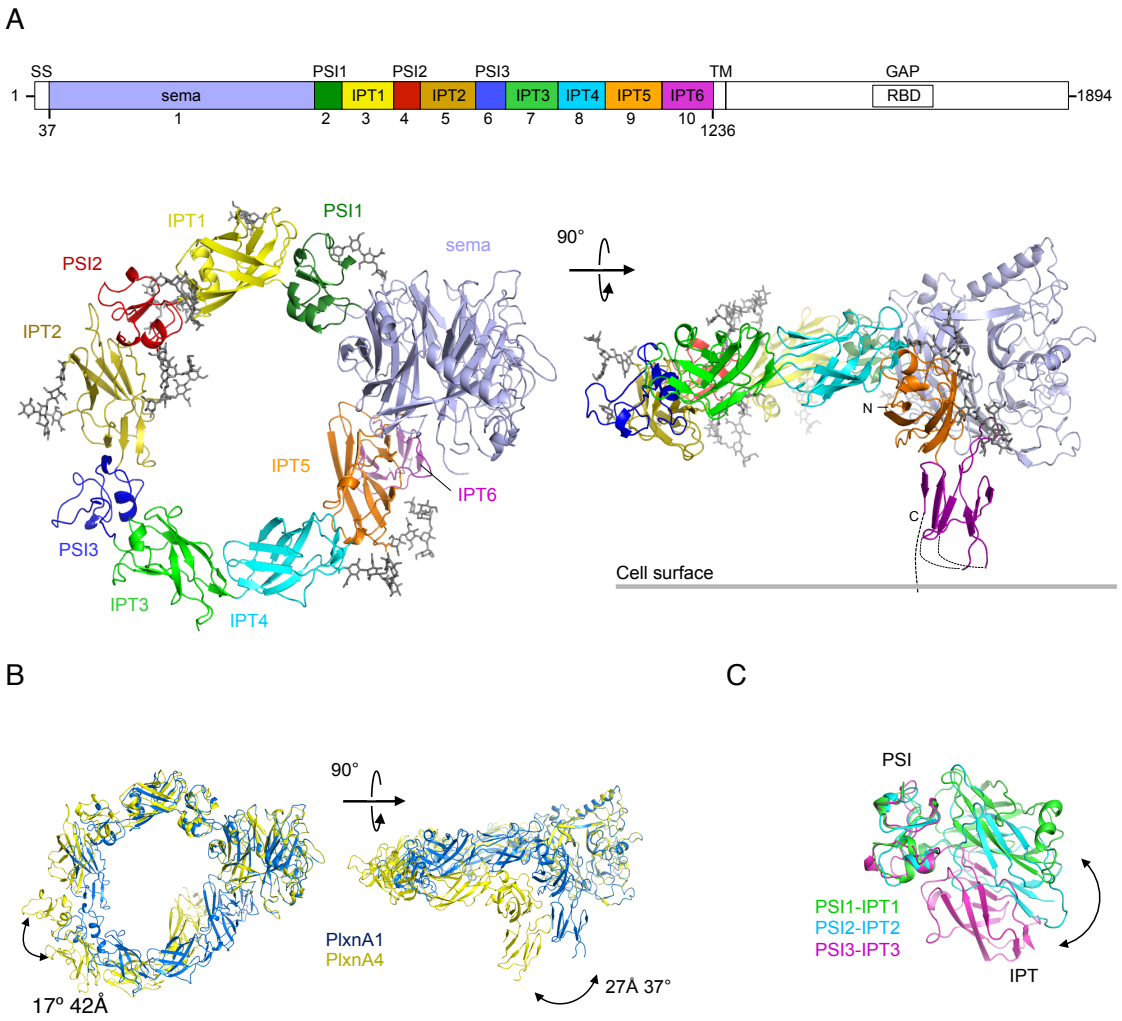


Figure 3.7: The 4Å crystal structure of PlxnA1$_{1\text{-}10}$. (A) Schematic domain organization of PlxnA1 (top) and ribbon representation of the PlxnA1$_{1\text{-}10}$ crystal structure at 4Å resolution (space group $P4_32_12$) in top and side views (left and right, respectively). N-linked glycans are shown in grey. (B) Ribbon representation of the 4 Å PlxnA1$_{1\text{-}10}$ (blue) and the 7.5 Å PlxnA4$_{1\text{-}10}$ (space group $P4_1$) crystal structures superposed on the basis of the sema domain (views as in A). Differences in PSI3 and IPT6 domain orientation (left and right, respectively) are indicated by domain translation and rotation. (C) Ribbon representation of a PSI domain-based superposition of the three PSI-IPT segments from the 4 Å PlxnA1$_{1\text{-}10}$ structure. The PSI3-IPT3 segment has a different interface and a different interdomain angle compared to the PSI1-IPT1 and PSI2-IPT2 segments.

3.5 Negative Stain Electron Microscopy (EM) Revealed Two Alternative Conformations for PlxnA Ectodomains

3.5.1 PlxnA1$_{1\text{-}10}$ Exists in a Major Ring-like and a Minor Chair-like Conformation

To assess the possible conformations of PlxnA ectodomains when released from the constraints of crystal packing, I conducted negative stain electron microscopy (EM) experiments for PlxnA1$_{1\text{-}10}$. Mouse PlxnA1$_{1\text{-}10}$ was produced in HEK293T cells to ensure appropriate glycosylation states. After IMAC and SEC purification using standard protocols, a high-purity sample of PlxnA1$_{1\text{-}10}$ was obtained (as summarised in Fig. 3.8). To minimize the potential artefacts induced by the staining process such as protein aggregation and stain precipitation, freshly gel-filtrated PlxnA1$_{1\text{-}10}$ were stained with newly prepared uranyl formate (pH ~4.5). To achieve image high contrast, electron micrographs were taken at an under-focus value ranging from 1.0 to 1.5 µm and astigmatism was carefully avoided. In the electron micrographs, PlxnA1$_{1\text{-}10}$ is monomeric, as can be expected from the low concentration (7 to 33 nM) of PlxnA1$_{1\text{-}10}$ applied to the EM grid (Fig. 3.9.A). Assessed by eye, the majority of PlxnA1$_{1\text{-}10}$ molecules appeared to orient with the flat face of the ring lying on the grid in a conformation resembling the crystal structures. In contrast, a proportion of the particles adopt a novel, chair-like (twisted-

open) conformation unseen in the crystal structures (Fig. 3.9.A and B).

We next analyzed the conformational variability of PlxnA1$_{1\text{-}10}$ via two-dimensional image classification. A total of 12,645 PlxnA1$_{1\text{-}10}$ particles were manually selected, normailsed, centered and aligned. To best represent the heterogeneity of the sample the aligned particles were categorized into 60 classes, resulting in a total number of 100-408 particles per class (Fig. 3.9.C). In agreement with the initial assessment, the class averages of PlxnA1$_{1\text{-}10}$ can be sorted into two categories, the "ring-like" or the "chair-like (twisted-open)" species. The ring-like species, which includes 66.0% of the particles, essentially matches the PlxnA$_{1\text{-}10}$ crystal structures. A bulky feature at one end of the ring is readily identifiable as the sema domain, which appears in a disk-like or bulb-like projection when orienting flat or side-way on the grid, respectively. From the sema domain onwards the stalk region extends in a range of curved conformations resulting, as in the crystal structures, in some variation in ring circumference. A second species, which includes 30.5% of the particles, can be identified by its distinct chair-like conformation. The N-terminal portion of the stalk region extends from the sema domain as in the ring-like structures, but the C-terminal half appears flipped by 180°, which then curves away rather than completes the ring. For the remaining 3.5% of particles no clear class averages can be generated.

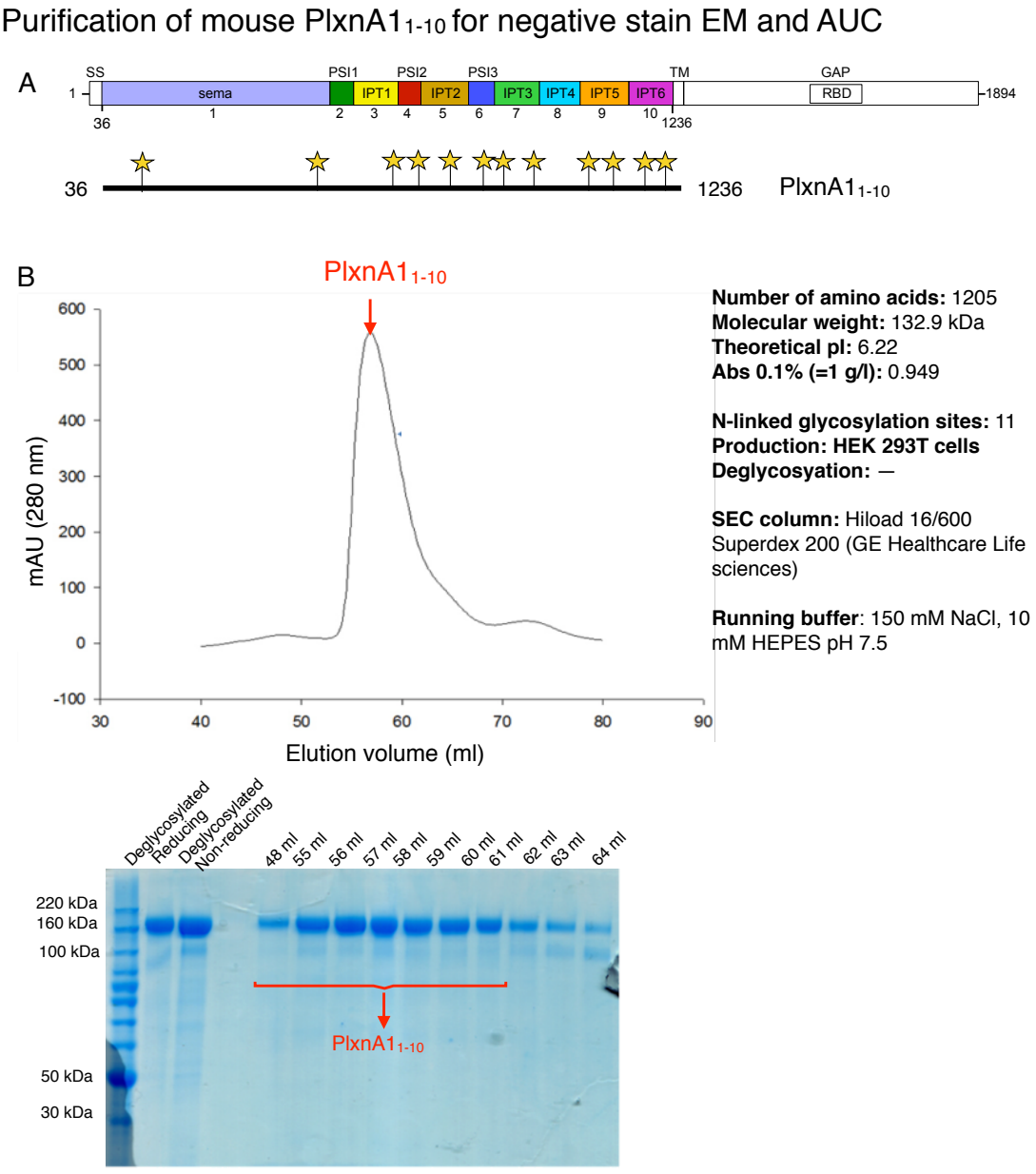


Figure 3.8: Summary of the purification of mouse PlxnA1$_{1\text{-}10}$. (A) Schematic illustration of the domain organization of mouse PlxnA1 and the PlxnA1$_{1\text{-}10}$ construct. The black line underneath the coloured domains of mouse PlxnA1 represents the segment covered by the current PlxnA1$_{1\text{-}10}$ construct (numbers at each end indicated the N-terminal and C-terminal residue, respectively). Yellow stars indicate the positions of predicted N-linked glycans. (B) The elution profile and SDS-PAGE analysis of the SEC experiment conducted for mouse PlxnA1$_{1\text{-}10}$. The basic information, production and SEC experimental conditions of PlxnA1$_{1\text{-}10}$ are listed at the right of the middle panel.

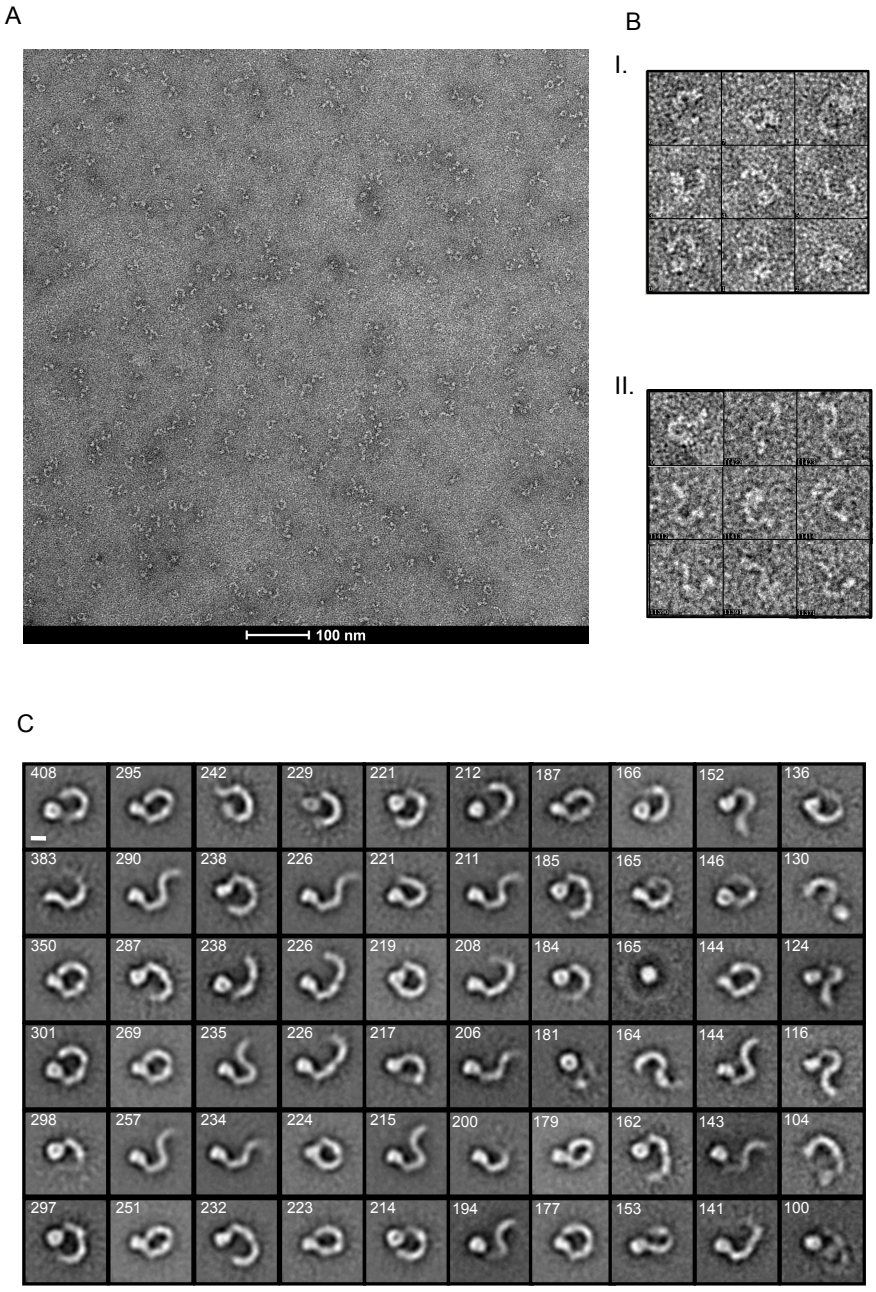


Figure 3.9: Negative stain EM micrograph and class averages of PlxnA1$_{1\text{-}10}$. (A) Representative negative stain EM electron micrograph of PlxnA$_{1\text{-}10}$ particles. Scale bar, 100 nm. (B) Representative raw particles adopting the ring-like (I) and the chair-like (twisted-open) conformations (II). (C) Sixty class averages of PlxnA$_{1\text{-}10}$ particles ranked from top left to bottom right based on the number of particles within each class (listed on the top left corner). The size of each image is 298 Å× 298 Å. Scale bar, 5 nm.

3.5.2 The Two Potential Flexions in PlxnA$_{1\text{-}10}$

By comparing the PlxnA$_{1\text{-}10}$ crystal structures to its negative stain EM class averages, I found that the PlxnA$_{1\text{-}10}$ structural models could be fitted to their class averages by assuming two flexions (Fig. 3.10). To best model the major, ring-like conformations seen in the class averages, the stalk region of a 6 Å PlxnA$_{1\text{-}10}$ crystal structure (space group P2$_1$ as shown in Fig. 3.6) was anchored at the IPT2-PSI3 (domains 5-6) interface and the PSI3-IPT6 (domains 6-10) was flexed as a rigid body closer to or away from the N-terminal half of the stalk region (Fig. 3.10.A.I). For the chair-like (or twisted-open) conformations, domains PSI3-IPT6 (domains 6-10) were first rotated by 180° around the IPT2-PSI3 (domains 5-6) linker region (805Y-806K) and then flexed open and closed to generate the series of relatively closed to opened chair-like conformations (Fig. 3.10.A.II). A total of 10 models (six for the ring-like and four for the chair-like conformation) were generated and filtered to a Gaussian resolution of 10 Å. Two-dimensional projections of the structure models were then produced and cross-correlated with the negative stain EM class averages.

An excellent agreement between the class averages and the projections of seven out of ten PlxnA1$_{1\text{-}10}$ models reinforced the rather surprising conclusion that the ten domain extracellular segments of PlxnAs have relatively limited flexibility (Fig. 3.10.B). Indeed, for the ring-like conformation, variation in the ring circumference can be attributed to flexion between IPT2-PSI3 (domains 5-6). The disc-like projection of the sema domain observed for some ring-like class averages required a rotation of the sema

domain around the sema-PSI1 linker (Fig. 3.10.A.I). As this particular orientation has not been observed in any of the crystal structures, it can be attributed to the fast drying and flatting of molecules in the staining process. For the chair-like conformation, the twist in the stalk region can be attributed to an 180° rotation around the IPT2-PSI3 (domains 5-6) linker region with additional flexion between IPT2-PSI3 (domains 5-6) (Fig. 3.10.A.II). For both conformations, the curvatures of the two halves of the stalk region, as maintained by sema-IPT2 (domains 1-5) and PSI3-IPT6 (domains 6-10), remain rigid. However, using these negatively-stained images the contributions the interface between IPT2-PSI3 (domains 5-6) or even PSI3-IPT3 (domains 6-7) make in flexing the stalk region cannot be resolved.

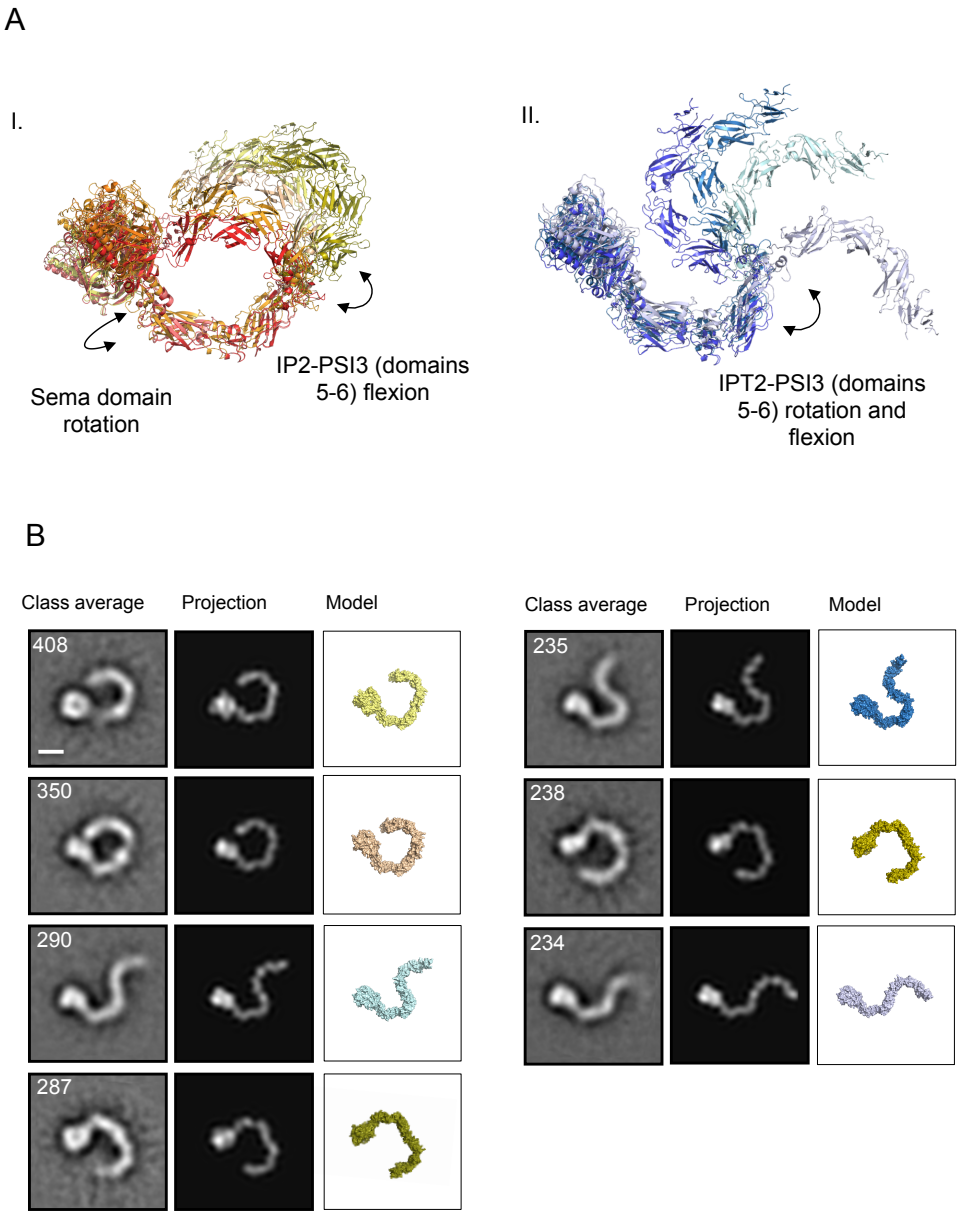


Figure 3.10: Correlation between the negative stain EM class averages and structural models of PlxnA1$_{1\text{-}10}$ revealed two potential flexions. (A) The ring-like (I) and chair-like (II) PlxnA1$_{1\text{-}10}$ structural models manually created based on the 6 Å crystal structure of PlxnA1$_{1\text{-}10}$ (space group, P2$_1$). (B) Seven representative class averages of PlxnA1$_{1\text{-}10}$ shown alongside the correlated projections and corresponding structural models (coloured according to models shown in A). The number of particles within each class is listed on the top left corner. The seven class averages which correlate best with the structure models are shown, which have a correlation coefficient of 0.929, 0.923, 0.918, 0.910, 0.901, 0.890, 0.876, respectively. Scale bar, 5 nm.

3.6 Discussions

3.6.1 The Two Alternative Conformations of PlxnA Ectodomains

Although the structural bases for semaphorin–plexin recognition as well as plexin intracellular Rap-GAP function have been characterised, the activation and regulation mechanisms for plexin signalling remained elusive. Our understanding of the molecular mechanisms that control the semaphorin–plexin signalling system has been severely limited by lack of information on the architecture and interaction modes of plexins, in particular in the plexin extracellular segments. How do plexins, in particular the relatively well-characterised PlxnAs, dimerise and activate their cytoplasmic domain as a result of extracellular semaphorin-binding? In this thesis, I find that the key to this question is the architecture of the plexin extracellular segments as exemplified by the PlxnAs.

Crystallographic studies together with negative stain EM experiments demonstrate that the PlxnA ectodomains exist in two alternative conformations each with limited inter-domain flexibility. These two distinct conformations are first, the preferred, ring-like conformation as shared by all our PlxnA ectodomain crystal structures and recurring in negative stain EM class averages for PlxnA1$_{1\text{-}10}$, and second, the less frequent, twisted-open (chair-like) conformation only observed in negative stain EM experiments. These results argue against the previous speculations on the structures of PlxnA extracellular segments, which suggested an elongated conformation reminiscent of other multidomain TM1 cell-surface receptors (30, 35, 123, 124). Rather, the PlxnA extracellular seg-

ments are most likely to curve into a ring-like form and reside close to the cell membrane. My results also rule out a hypothesized "head-down" conformation based on intramolecular interactions of the PlxnA ectodomains (22, 36), as it would require significant flexibility throughout the stalk region precluded by my analysis. The alternative, chair-like conformation of the PlxnA ectodomains most likely results from a twist-opening midway in the stalk region around potential flexion points between domains 5-6 (IPT2-PSI3). However, as PSI3 (domain 6) remains the only domain without a high-resolution crystal structure despite my extensive crystallisation trials, I could not examine these potential flexion points in detail. Interesting, this suggestd that PSI3 may have a certain degree of inherent structural variability. One should note that despite the lack of extensive interdomain flexibility or multiple transition states between the PlxnA ring-like and chair-like conformations, a certain degree of breathing may still exist within each conformation.

3.6.2 Structural Models of Sema6s–PlxnAs and Sema3s–PlxnAs–Nrps Signalling Complexes

When I superpose the ring-like crystal structures of PlxnA ectodomains ($PlxnA4_{1\text{-}10}$, resolution 7.5 Å, space group $P4_1$) with those in the 2:2 $Sema6A_{1\text{-}2}$–$PlxnA2_{1\text{-}4}$ recognition complex (PDB 3OKT) (36), I arrive at a potential model of Sema6s–PlxnAs activation (Fig. 3.11.A.I). This model suggest that two copies of PlxnA monomer adopting the ring-like conformation may effectively bring together their extracellular C-termini (ranging

from 10 Å to 59 Å for the five PlxnA crystal structures that have all 10 domains resolved), and thus allowing their transmembrane helices as well as cytoplasmic regions to come into proximity. This close distance together with an appropriate orientation of the PlxnA cytoplasmic domains may trigger interactions between the juxtamembrane helices and thus, the subsequent opening of the Rap-binding site in the PlxnA GAP domains (125). Similarly, a model constructed based on the 2:2:2 Sema3A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$–Nrp1$_{1\text{-}4}$ recognition complex (PDB 4GZA) and the same ring-like crystal structures of PlxnA ectodomains may represent the potential Sema3s–PlxnAs–Nrps activation complex (Fig. 3.11.A.II). The large cavity between the two rings of the PlxnA ectodomains is sufficient for two copies of the five-domain Nrp extracellular segments to slot in and buttress the overall heterotrimeric complex. The long and flexible coiled-coil region between domains 1-2 (a1-a2) in Nrps may allow appropriate orientations for the a1 domain to be anchored to the Sema3s–PlxnA1–Nrps binding complex. The rest of the Nrp1 ectodomain apart from the interacting a1 domain, domains 2-5 (a2-c1), may extend downwards to the cell surface oriented by flexible interdomain linkers (45). Thus, the distinctive, ring-like architecture of PlxnA ectodomains may provide a satisfyingly simple solution to a long-standing puzzle regarding the translation of Sema–PlxnA extracellular ligand-binding into cytoplasmic activation.

In a similar modelling process I found that the minor, chair-like (twisted-open) conformation of PlxnA extracellular segments is less likely to be used for binding semaphorins in *trans*. This is because should the PlxnAs adopt the chair-like conformation, the membrane-proximal regions of the PlxnAs would be separated far apart and

point upwards from the cell surface. However, the chair-like PlxnAs fit well in a model where Sema6s bind to PlxnAs in *cis* in a 1:1 mode assuming the Sema6–PlxnA head-to-head interaction destabilises the original Sema6 dimer interface (Fig. 3.11.B). This model was generated by taking a 1:1 Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$ unit from the Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$ recognition complex (PDB 3OKT). It is also possible that the Sema6 dimer coexists with the chair-like PlxnAs to form a 2:1 complex, although a 2:2 Sema6s–PlxnAs complex using the PlxnA chair-like conformation seems highly improbable as an additional PlxnA ectodomain would clash with the cell membrane. The potential Sema6s–PlxnAs *cis*-interaction during which the PlxnA ectodomains adopt the chair-like state may be the basis for the recently discovered mutual silencing of PlxnAs and Sema6s (39, 66, 67). Similarly, the chair-like conformation may also be used in the potential *cis*-interactions between PlxnA2 and Sema5A potentially also incorporating HSPGs and CSPGs (46). However, uncovering what specific role (if any) the chair-like conformation plays will require a significant amount of further experimentation.

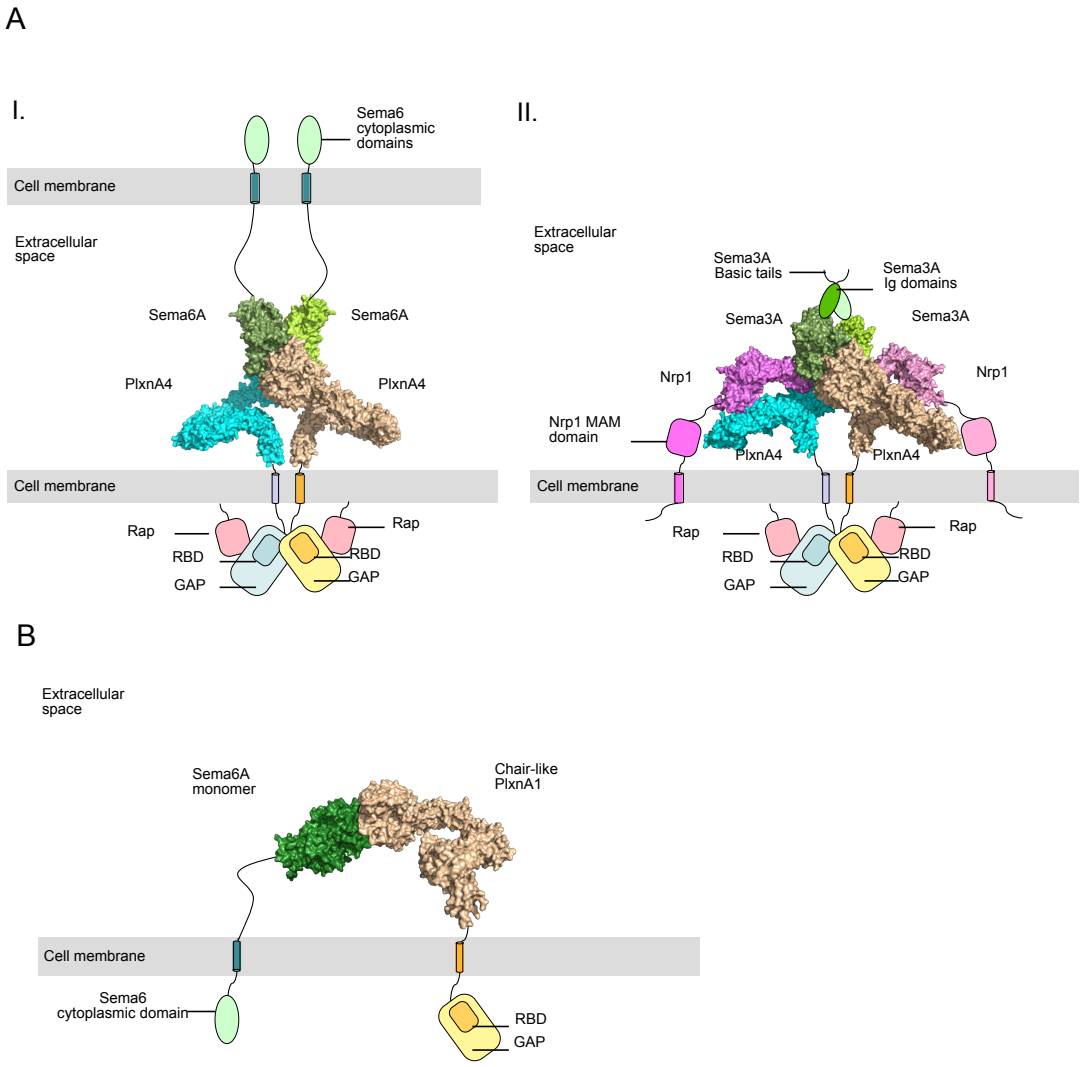


Figure 3.11: Structural models of the Sema–PlxnA signalling complexes. (A) Structural models of the potential Sema6s–PlxnAs (I) and Sema3s–PlxnAs–Nrps (II) signalling complexes based on the ring-like conformation of PlxnA ectodomains (modeled on the 7.5 Å crystal structure of PlxnA4$_{1\text{-}10}$, space group P4$_1$). The structural model of Nrp1 was generated from the crystal structure of Nrp1$_{1\text{-}4}$ (PDB 4GZ9). (B) Structural model of the potential Sema6s–PlxnAs complex based on *cis*-interactions. The chair-like PlxnA extracellular segments were generated from a model of the 6 Å PlxnA1$_{1\text{-}10}$ crystal structure (space group P2$_1$) according to the negative stain EM class averages shown in Fig. 3.10. The transmembrane helices, cytoplasmic region, and the substrate Rap are illustrated as cartoons. Each protein is indicated by its nearby labels.

CHAPTER 4

STRUCTURAL BASIS OF PLXNA AUTOINHIBITION

4.1 Introduction

As detailed in the previous chapter, structural studies uncovered a preferred, ring-like conformation together with an alternative, chair-like conformation for the PlxnA extracellular segments. In particular, the ring-like conformation may be essential for the activation of PlxnAs, as superposing the ring-like PlxnA ectodomains onto the Sema6A–PlxnA2 and the Sema3A–PlxnA2–Nrp1 recognition complexes I arrived at a model in which bivalent semaphorin–plexin binding in the extracellular region is transduced into dimerisation-induced Rap-GAP function in the plexin cytoplasmic region. I further hy-

pothesized that the minor, chair-like conformation of PlxnA ectodomains may potentially be useful for the PlxnA *cis*-interactions with transemembrane class 5 and 6 semaphorins (46, 67).

However, the signalling of plexins is exquisitely balanced between inhibition and activation. A previous study has found that the PlxnAs are autoinhibited pre-signalling and that such autoinhibition is imposed by the PlxnA extracellular segments in particular the sema domains (126). However, the structural basis for this pre-signalling autoinhibitory mechanism remains elusive. It has also been observed that the extracellular segments of PlxnAs may weakly associate in solution and thus potentially cluster on the cell-surface (36). Pre-clustering of cell guidance cue receptors based on their ectodomains may provide advantages for localised and spatial-temporally precise guidance cue signalling, a mechanism exemplified by the Eph receptors as well as other systems (30, 32, 39). These observations have lead to a paradoxical scenario. Should the PlxnAs already pre-associate before semaphorin-binding and dimerisation and/or oligomerisation effectively trigger signalling, then how is ligand-independent constitutive activation of PlxnAs avoided?

In this thesis, I uncover that the key to balancing PlxnA inhibition and activation is pre-signalling receptor *cis*-interactions based on an intermolecular, head-to-stalk (sema–PSI2-IPT2) interface. This interface was first discovered by Dr. Bert J.C. Janssen in all crystal structures of the PlxnA extracellular segments and then validated in solution by me via a palette of structure-guided biophysical assays. Using advanced fluorescence microscopy techniques I was able to assess the PlxnA–PlxnA interactions based on this

particular interface on the surface of live COS-7 cells. Collapse assays based on COS-7 cells and dentate gyrus (DG) growth cones in collaborations with the group of Prof. Jeroen Pasterkamp in University Medical Center Utrecht demonstrated that indeed, the head-to-stalk, *cis*-interactions together with the conformation of the PlxnA ectodomains may impose autoinhibition onto the otherwise constitutively active PlxnAs. This mechanism is particularly important in maintaining a healthy axonal growth cone, which may otherwise collapse prematurely due to a high level of PlxnA basal activities.

4.2 The Sema–PSI2-IPT2 Intermolecular Interface of PlxnAs

4.2.1 Crystal Structures of PlxnA Extracellular Segments Revealed a Common Intermolecular Interface

In all our crystal structures of PlxnA1, A2 and A4 extracellular segments (Chapter 3), we observed a common intermolecular interface between the head, i.e. the sema domain, and domains 4-5 (PSI2-IPT2) in the stalk region for neighbouring PlxnA molecules (Fig. 4.1). Through this sema–PSI2-IPT2 (head-to-stalk) interface, the adjacent PlxnA ectodomains form arrays in the crystal lattices. Although other intermolecular interactions exist in the crystal networks, this sema–PSI2-IPT2 interface is the only common one shared by all our thirteen PlxnA crystal structures across eight different crystal

forms. Previous studies have found that the PlxnA2 extracellular segments have a tendency for weak *cis*-interactions in solution (36). Therefore, I set out to examine whether the PlxnA extracellular segments are engaged in *cis*-interactions via the sema–PSI2-IPT interface as observed in our crystal structures.

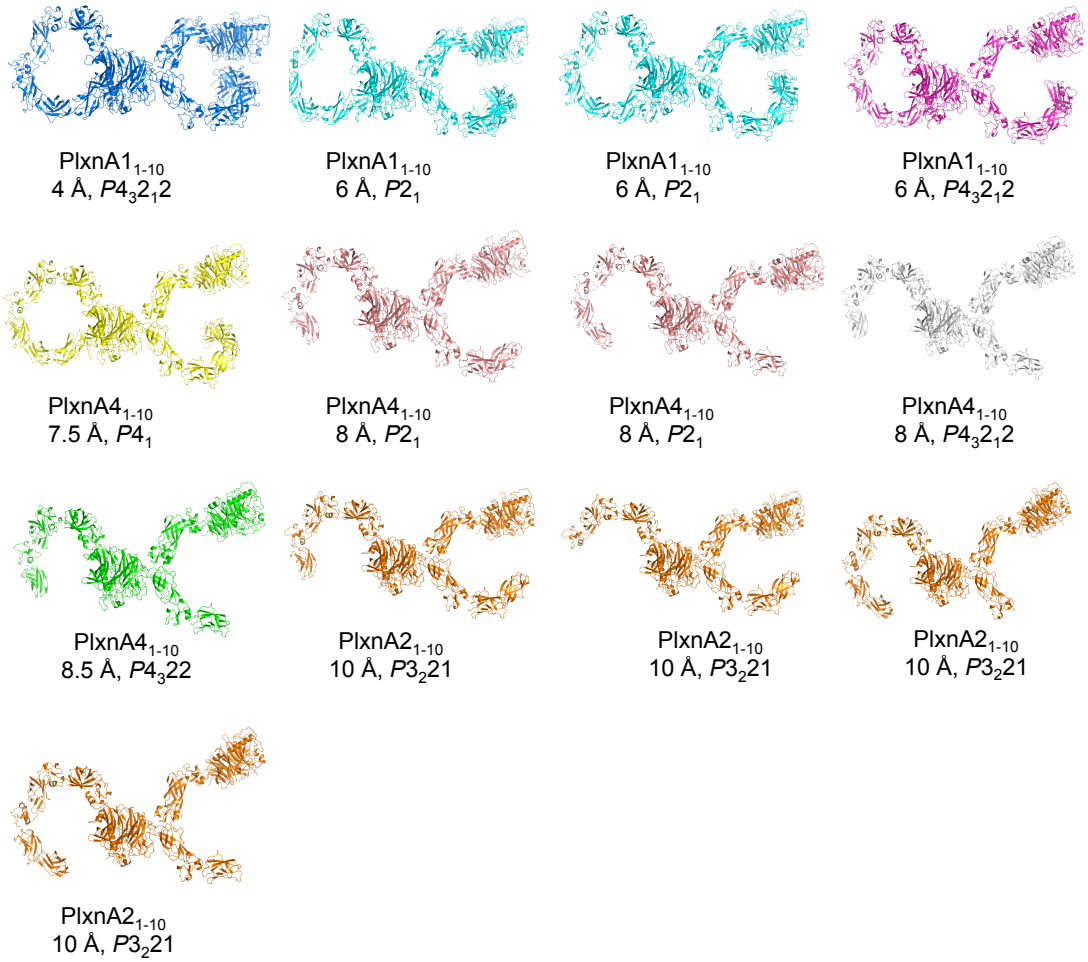


Figure 4.1: A common head-to-stalk interface in all crystal structures of PlxnA extracellular segments. All thirteen crystal structures of PlxnA1, A2 and A4 extracellular segments are shown in ribbon representation in the same sequence as in Fig. 3.6. The highest resolution of the diffraction data and the space group for each crystal structure are indicated.

4.2.2 Structural Details of the PlxnA Sema–PSI2-IPT2 Interface

The PlxnA sema–PSI2-IPT2 (head-to-stalk) interface is formed by blades 1-3 of the sema domain of one PlxnA molecule and the outer side of PSI2-IPT2 (domains 4-5) from the ring of a second PlxnA (Fig. 4.2.A). In the 4 Å PlxnA1$_{1\text{-}10}$ structure the interface has a buried surface area of 1303 Å^2 (355 Å^2 between sema–PSI2 and 1008 Å^2 between sema–IPT2) and a complementary electrostatic charge distribution (Fig. 4.2.B). Sequence conversation analysis revealed that the residues within this interface are conserved among vertebrate PlxnAs but much less conserved among all plexins (Fig. 4.2.C), suggesting that this plexin–plexin interaction may be an important functional feature of the class A plexins. One should note that in Fig. 4.2, the structure of domain IPT6 (domain 10) has several unmodelled loops and hence is revealing its conserved hydrophobic core in these views. Unfortunately, at 4 Å resolution I was not able to determine the precise side chain orientations for residues making the sema–PSI2-IPT2 contacts.

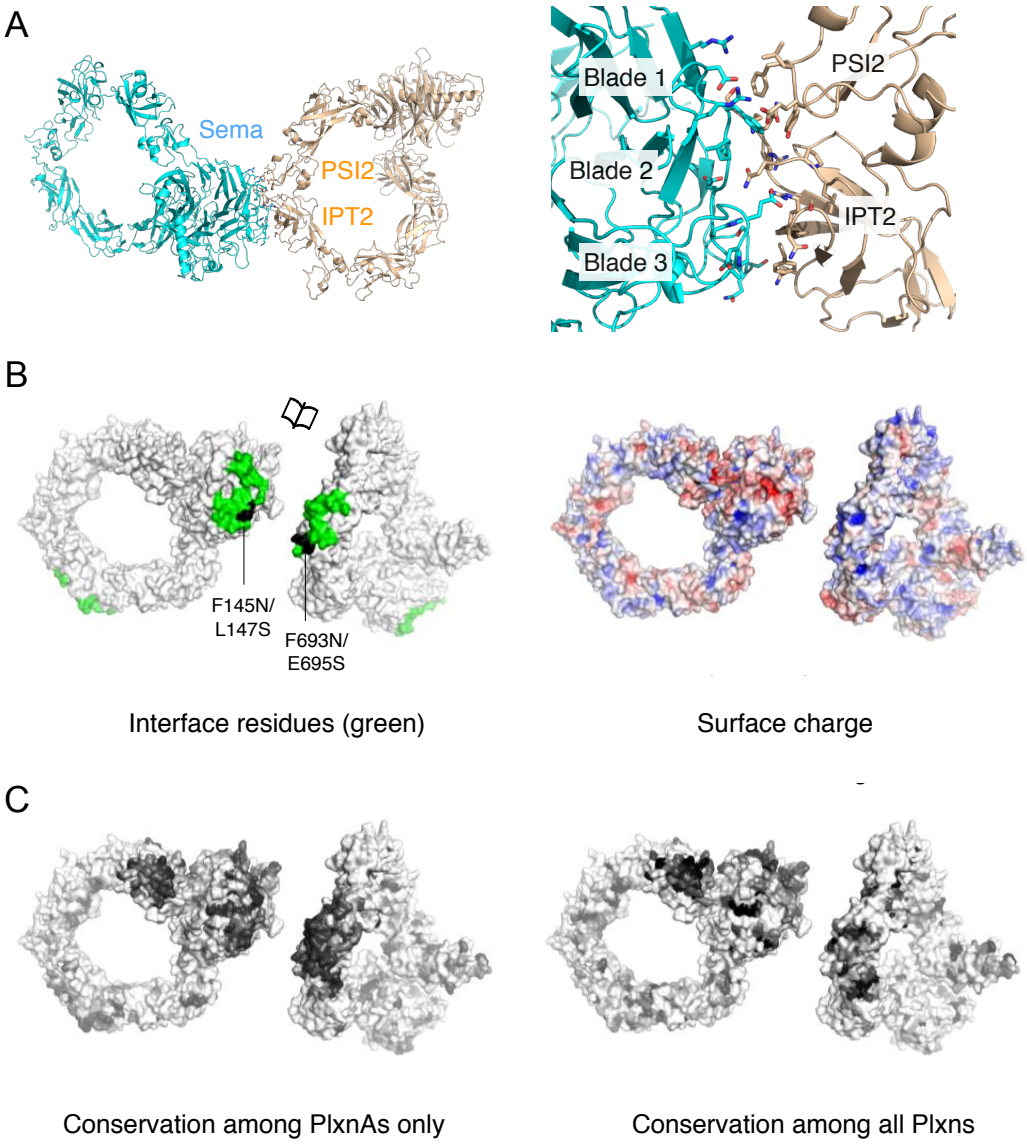


Figure 4.2: Structural details of the sema–PSI2-IPT2 interface of PlxnAs. (A) The sema–PSI2-IPT2 interface in the 4 Å crystal structure of PlxnA1$_{1\text{-}10}$ (space group $P4_32_12$) is shown in ribbon representation. (B) Surface representation of an opened view of two interacting PlxnA1$_{1\text{-}10}$ (4 Å resolution, space group $P4_32_12$) with the residues on the sema–PSI2-IPT2 interface in green and mutation sites (to be discussed in section 4.3) in black (left), and electrostatic potential from red ($-6k_bT/e_c$) to blue ($6k_bT/e_c$) (right). (C) Sequence conservation of surface residues mapped onto two interacting PlxnA1$_{1\text{-}10}$ (same as in B). Residues are colour coded from nonconserved, white, to conserved, black, on the basis of alignments containing vertebrate PlxnA sequences only (left) or vertebrate plexin sequences from all classes (right).

4.2.3 PlxnAs May Form Small Cell-Surface Oligomers Based on the Sema–PSI2-IPT2 Interface

The findings from our crystal structures and sequence analysis raise the possibility that the PlxnAs on the same cell surface can interact in *cis* using the sema–PSI2-IPT2 intermolecular interface to form dimers or small oligomeric clusters. A simple model-building exercise confirmed this possibility whilst arguing against formation of large cell-surface attached plexin arrays through this head-to-stalk interaction mode. For a dimer mediated by the sema–PSI2-IPT2 interface the two plexin rings are tilted relative to each other by 50° rather than abutting in the same plane (Fig. 4.3.A). For this arrangement some flexibility in the plexin stalk region, as revealed by the crystal structures, could well allow formation of small, up to three molecule, cell-surface compatible clusters (Fig. 4.3.A.II). In contrast, it appears that the formation of large arrays through repeated head-to-stalk interactions would inevitably lead to steric clashes with the plasma membrane in the absence of substantial membrane curvature. In a head-to-stalk association the C-termini of plexin ectodomains in ring-like conformations are spaced far apart (99 Å to 121 Å for the five PlxnA crystal structures that have all 10 domains resolved). My modelling further demonstrated that the putative secondary, twisted-open (chair-like) conformation for the PlxnAs as seen in the negative stain EM data could potentially also dimerise via the sema–PSI2-IPT2 intermolecular interface (Fig. 4.3.B). In this case, the architecture would separate two ectodomain C-termini by more than 210 Å. Formation of a PlxnA trimer based on the chair-like conformation or a mixture of

ring-like and chair-like conformations appeared sterically incompatible with cell surface attachment, and is thus not illustrated.

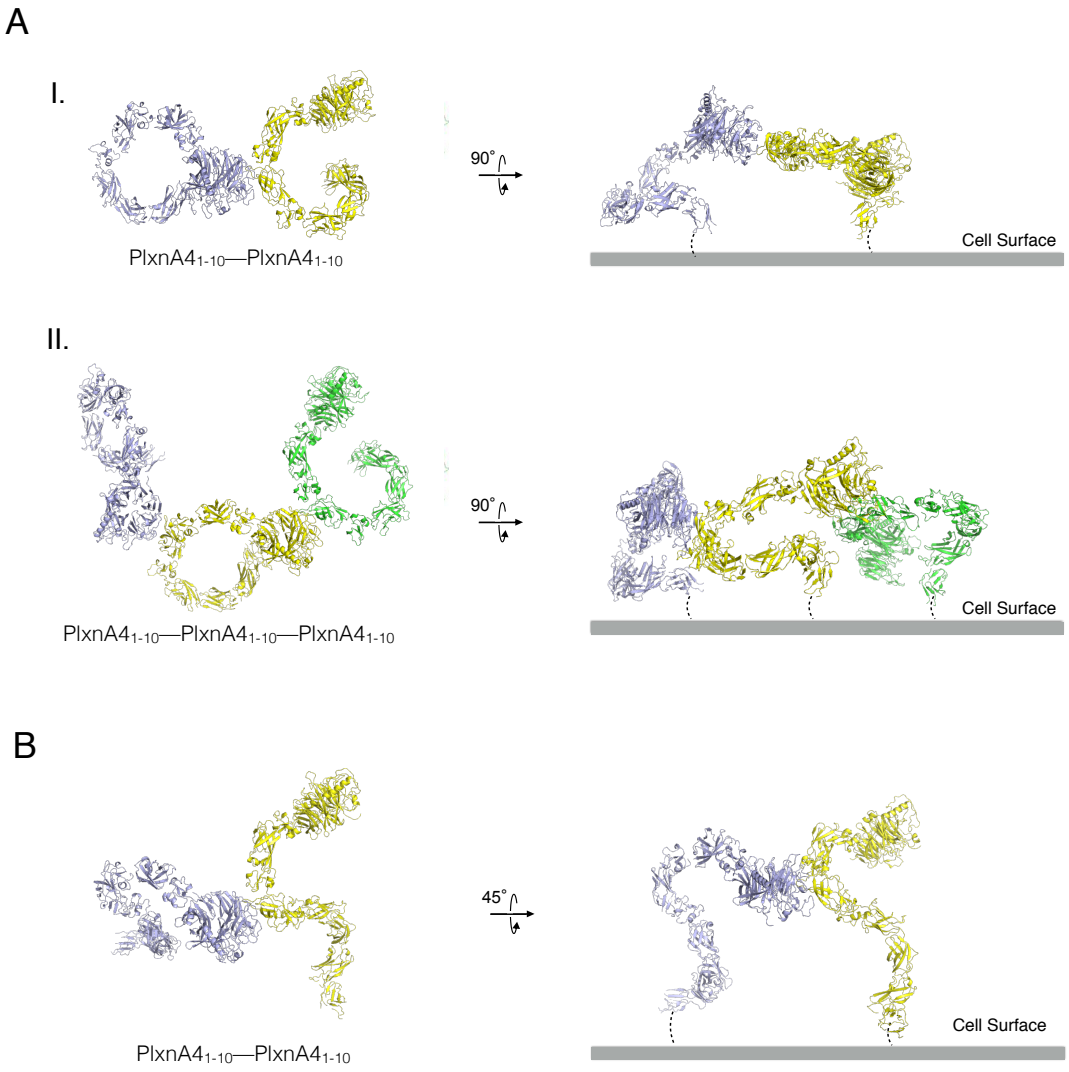


Figure 4.3: Models of potential small PlxnA oligomers based on the sema–PSI2-IPT2 interface. (A) The sema–PSI2-IPT2 interface mediated PlxnA dimers (I) and trimers (II) based on the ring-like conformation and packing of the PlxnA4$_{1-10}$ crystal structure (space group $P4_1$) are illustrated in the ribbon representation. (B) Possible chair-like PlxnA homodimers based on the sema–PSI2-IPT2 interface that may form on a cell surface. The PlxnA4$_{1-10}$ chair-like structural model which correlates best with the most populated negative stain EM class average of PlxnA1$_{1-10}$ (see Chapter 3 Fig. 3.10) is used for this illustration.

4.3 PlxnA Extracellular Segments Interact in Solution via the Sema–PSI2-IPT2 Interface

4.3.1 The Engineering of PlxnA Interface Mutants

In light of the findings detailed in the previous sections I set out to assess whether PlxnAs interact via the sema–PSI2-IPT2 interface in solution. Based on the 4 Å crystal structure of PlxnA1$_{1\text{-}10}$ (space group $P4_32_12$), I engineered a series of mutant constructs of PlxnA1, A2 and A4, in which the sema–PSI2-IPT2 interface is disrupted. I mutated several residues on the sema–PSI2-IPT2 interface of PlxnA1 to a NXS combination, including F145N/L147S (on sema domain) and F693N/E695S (on PSI2) (Fig. 4.4.A). These mutations allow N-linked glycans to be introduced into the interface when produced from mammalian cells. The bulky N-linked glycans should provide ample steric hindrance to the sema–PSI2-IPT2 interactions. Similarly, based on sequence conservation the N-linked glycans were introduced into similar sites in PlxnA2 via mutations F146N/L148S and F690N/E692S, as well as in PlxnA4 via mutations F146N/L148S and F689N/E691S.

4.3.2 MALS Revealed Weak Associations of PlxnA1$_{1\text{-}10}$ via the Sema–PSI2-IPT2 Interface in Solution

To evaluate the potential intermolecular interactions between PlxnA ectodomains I first performed multi-angle light scattering (MALS) experiments coupled to size-exclusion chromatography (SEC) for PlxnA1$_{1\text{-}10}$. When a purified sample of proteins is allowed to interact with the size-exclusion matrices, species of larger sizes (higher molar masses) are separated from the relatively smaller species (lower molar masses) resulting in a shift to lower retention volumes. For a particular protein weakly associating in solution, a reduction in SEC-MALS retention volume, i.e. an increment in measured molecular weight, may occur as a result of increasing concentration. In the SEC-MALS experiments I injected purified PlxnA1$_{1\text{-}10}$ at four concentrations, 28.1 μM (4 mg/ml), 7.0 μM (1 mg/ml), 1.8 μM (0.25 mg/ml), and 0.45 μM (0.0625 mg/ml). One should note that due to the dilution effect of SEC, the concentrations during elution were lower than concentrations at injections. Indeed, PlxnA1$_{1\text{-}10}$ showed a clear shift to lower retention volumes as concentrations increased (Fig 4.4.B.I). In addition, the average molar mass for PlxnA1$_{1\text{-}10}$ determined by MALS increased from 151 kDa (strictly indicative of monomers) to 176 kDa (consistent with a mixture of monomers and dimers) with increments in concentrations (Fig. 4.4.B.I).

To assess whether the sema–PSI2-IPT2 interface is responsible for the PlxnA1$_{1\text{-}10}$ *cis*-interactions in solution, I conducted SEC-MALS experiments for the two interaction-blocking mutants of PlxnA1$_{1\text{-}10}$, PlxnA1$_{1\text{-}10}$ F145N/L147S and PlxnA1$_{1\text{-}10}$ F693N/E695S

under the same conditions. Indeed, these two mutants showed much less propensity for intermolecular interaction. PlxnA1$_{1\text{-}10}$ F145N/L147S had a less pronounced peak shift to lower retention volumes and the weight average mass indicated a predominantly monomeric species (Fig. 4.4.B.II). PlxnA1$_{1\text{-}10}$ F693N/E695S was strictly monomeric at all concentrations tested; there was no peak shift and the weight average mass indicated that only monomers were present (Fig. 4.4.B.III). The mass ranging from 147-156 kDa determined for PlxnA1$_{1\text{-}10}$ F145N/L147S and PlxnA1$_{1\text{-}10}$ F693N/E695S at all concentrations fits well with a calculated mass of 156 kDa for a PlxnA1$_{1\text{-}10}$ monomer. One should note that for PlxnA1$_{1\text{-}10}$ F145N/L147S, despite a peak shift to lower retention volumes at the highest concentration (4 mg/ml) its measured molar mass remained very similar to the PlxnA1$_{1\text{-}10}$ monomer (red curve, molar mass 155 kDa in Fig. 4.4.B.II). This may be because when high concentrations of F145N/L147S travelled through the SEC column it may have initially interacted very weakly, but as it diffused through the column during the run it was gradually diluted and thus the interactions became diminutive. As a result, when the elution fractions were measured by light scattering they yielded a molar mass typical for monomeric PlxnA1$_{1\text{-}10}$.

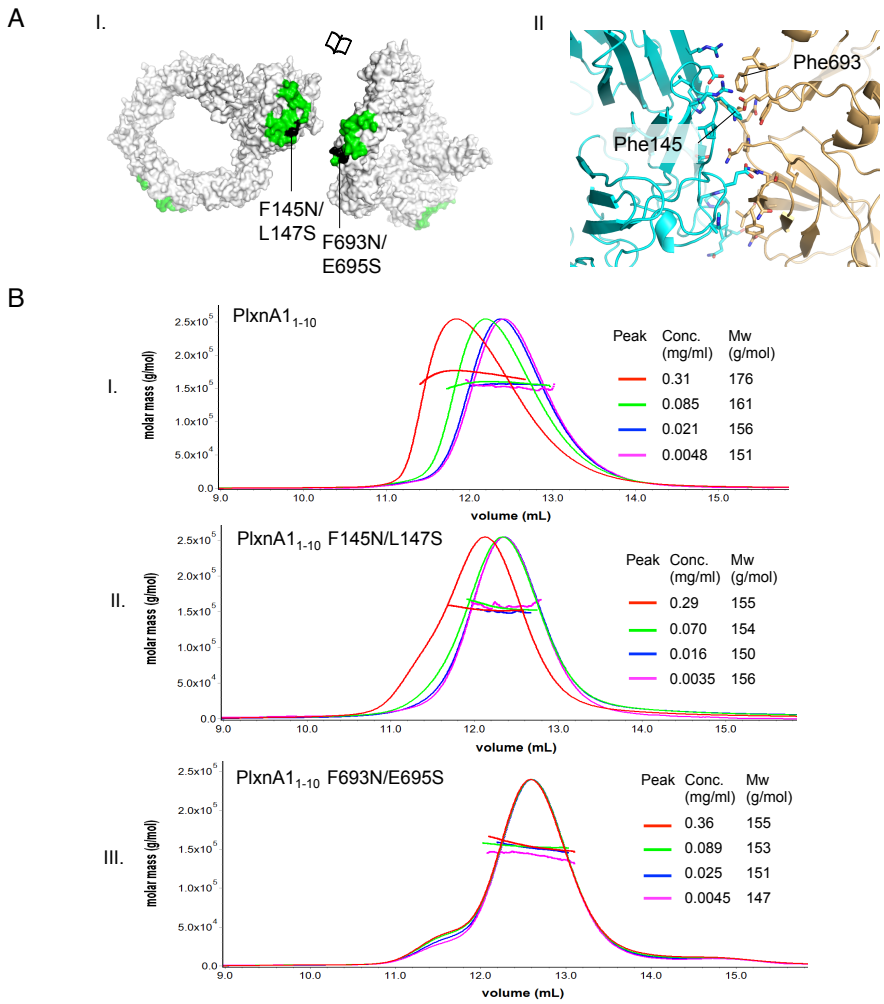


Figure 4.4: SEC-MALS of wild type and interface mutant PlxnA1$_{1\text{-}10}$. (A) Mutations at the sema–PSI2-IPT2 interface of PlxnA1$_{1\text{-}10}$ to yield interface mutants PlxnA1$_{1\text{-}10}$ F145N/L147S and PlxnA1$_{1\text{-}10}$ F693N/E695S. The mutated residues are coloured in black in the surface representation of an opened view of two PlxnA1$_{1\text{-}10}$ molecules (I). The residues mutated into asparagines for N-linked glycan attachment (Phe145 and Phe693) are indicated in II. (B) SEC-MALS elution profiles at four different concentrations of PlxnA1$_{1\text{-}10}$ (I), PlxnA1$_{1\text{-}10}$ F145N/L147S (II) and PlxnA1$_{1\text{-}10}$ F693N/E695S (III). The profiles of samples injected at 4 mg/ml, 1 mg/ml, 0.25 mg/ml and 0.0625 mg/ml are coloured in red, green, blue and magenta, respectively. The concentration of proteins and the molar masses of the peaks, which are indicated by coloured lines intersecting the elution maxima, are listed in the tables.

4.3.3 AUC Revealed Small-Scale Oligomerisation of PlxnA1$_{1\text{-}10}$ via the Sema–PSI2-IPT2 Interface

I next set out to further assess the sema–PSI2-IPT2 interface as well as the oligomerisation states of PlxnA1$_{1\text{-}10}$ in solution. For this purpose I performed AUC (analytical ultracentrifugation) velocity experiments as it allows for the characterisation of protein–protein interactions at high concentrations free from any surfaces or matrices. In agreement with MALS, results of AUC velocity experiments demonstrated that in solution PlxnA1$_{1\text{-}10}$ exists in heterogeneous states encompassing monomers up to tetramers (Fig. 4.5.A). At a high concentration of 28.1 μM (4 mg/ml), PlxnA1$_{1\text{-}10}$ gave four distinct peaks with sedimentation coefficients (s) corresponding to monomer, dimer, trimer and tetramer species, respectively. When the concentration was lowered to 7.0 μM (1 mg/ml), the population of PlxnA1$_{1\text{-}10}$ dimer and higher order oligomers decreased and more of the PlxnA1$_{1\text{-}10}$ remained monomeric. The discrete shape of the sedimentation coefficient distribution peaks is indicative of relatively slow exchange between the monomeric and oligomeric states. To assess whether the measured sedimentation coefficients represent well the oligomerisation states assuming the coexistence of both the ring-like and twisted-open conformations for PlxnA1$_{1\text{-}10}$, we calculated the theoretical values of the oligomerisation state-specific sedimentation coefficients. The predicted sedimentation coefficient for the ring-like monomer was 5.6S and for the chair-like monomer 5.5S; for the ring and chair dimers it was 8.4S and 7.8S respectively, for the ring-like and chair-like trimers it was 10.6S and 9.56S respectively, and for the ring and chair tetramers it

was 12.4S and 11.2S respectively (values calculated as described in section Materials and Methods). Indeed, the sedimentation coefficients predicted from the structural models for the different oligomeric states fall within the spectrum of each peak. It is worth noting that in AUC PlxnA11-10 oligomers consisting of more than four molecules was not observed even at high concentration (4mg/ml). This is consistent with the finding that the relatively weak PlxnA1–PlxnA1 interaction is insufficient to build up large oligomers.

In order to assess directly the contribution of sema—IPT2-PSI2 interface to PlxnA1$_{1\text{-}10}$ oligomerisations the behaviour of PlxnA1$_{1\text{-}10}$ F693N/E695S during AUC velocity experiments was analysed (Fig. 4.5.B). In contrast to wild type PlxnA1$_{1\text{-}10}$, this interaction-blocking mutant showed only one major sedimentation coefficient peak corresponding to the monomer species at both 28.1 μM and 7.0 μM. Small dimer and trimer peaks were observed at 28.1 μM, indicating that although intermolecular interaction was abrogated, it was not completely eliminated. However, as the diminutive dimer and trimer peaks were only observed at high concentration they are likely a result of unspecific interactions from small regions in the PlxnA1 ectodomains. I can therefore conclude that the formation of PlxnA1$_{1\text{-}10}$ oligomers is a direct result of intermolecular sema–IPT2-PSI2 interactions. One should note that the systematic peak broadening and shifts to higher absolute sedimentation coefficients in these experiments is a general observation at decreasing concentrations (127).

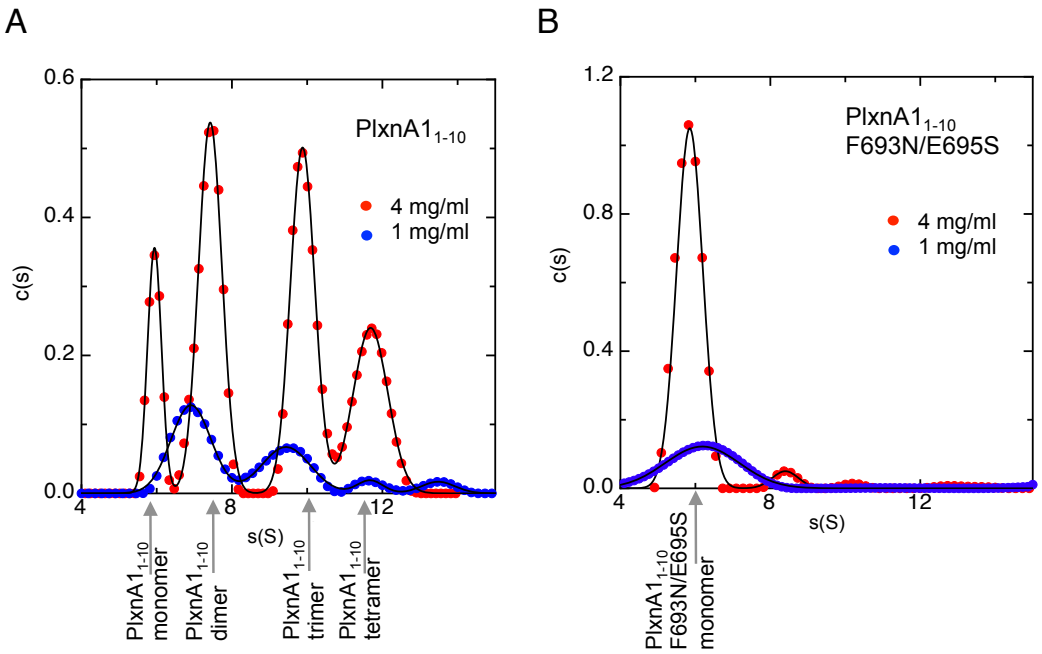


Figure 4.5: Distribution of sedimentation coefficients of wild type and interface mutant PlxnA1$_{1\text{-}10}$ measured by AUC. Distribution of sedimentation coefficients c(s) determined by AUC velocity experiments for PlxnA1$_{1\text{-}10}$ (A) and PlxnA1$_{1\text{-}10}$ F693N/E695S (B) at 4 mg/ml (red) and 1 mg/ml (blue). Calculated from the Lamm equation model, c(s) (in arbitrary units) is plotted against the sedimentation coefficient, s (in Svedbergs). The predicted sedimentation coefficients of different oligomeric states which best correspond to the peak values are indicated by grey arrows.

4.3.4 SPR Revealed the Binding Affinity Between the Sema and PSI2-IPT2 Domains for PlxnA2

To further probe the sema–IPT2-PSI2 interface and determine the affinity of interactions I undertook SPR (surface plasmon resonance) equilibrium experiments (Fig. 4.6). I engineered two separate segments consisting of sema-PSI1 (domains 1-2) and PSI2-

IPT2 (domains 4-5) of PlxnA2; each segment encompassing one side of the putative head-to-stalk intermolecular interface. I used PlxnA2 instead of PlxnA1 for these experiments as the analyte PlxnA2$_{1\text{-}2}$ were much better expressed and more stable in solution that PlxnA1$_{1\text{-}2}$. Similar to the MALS and AUC experiments, I constructed an interface-blocking mutant for PlxnA2$_{4\text{-}5}$, PlxnA2$_{4\text{-}5}$ F690N/E692S, which harbours an introduced N-linked glycosylation site in the PlxnA2 PSI2 domain equivalent to that in PlxnA1$_{1\text{-}10}$ (PlxnA1$_{1\text{-}10}$ F693N/E695S).

In SPR equilibrium experiments PlxnA2$_{4\text{-}5}$ associated to PlxnA2$_{1\text{-}2}$ with an equilibrium dissociation constant (Kd) of $20.9 \pm 0.38 \mu M$ (Fig. 4.6.A). This equilibrium dissociation constant typically indicates relatively weak interactions, which is consistent with MALS and AUC results. In contrast, the interaction-blocking mutant PlxnA2$_{4\text{-}5}$ F690N/E692S completely abolished the interaction with PlxnA2$_{1\text{-}2}$, indicating that the sema–PSI2-IPT2 interaction is solely responsible for PlxnA intermolecular interactions (Fig. 4.6.B). One should note that although the PlxnA head-to-stalk (sema–PSI2-IPT2) interactions appears relatively weak when measured in solution, for *cis*-interactions occurring between plexins tethered to the same plasma membrane there is potential for substantial enhancement in binding affinity (128).

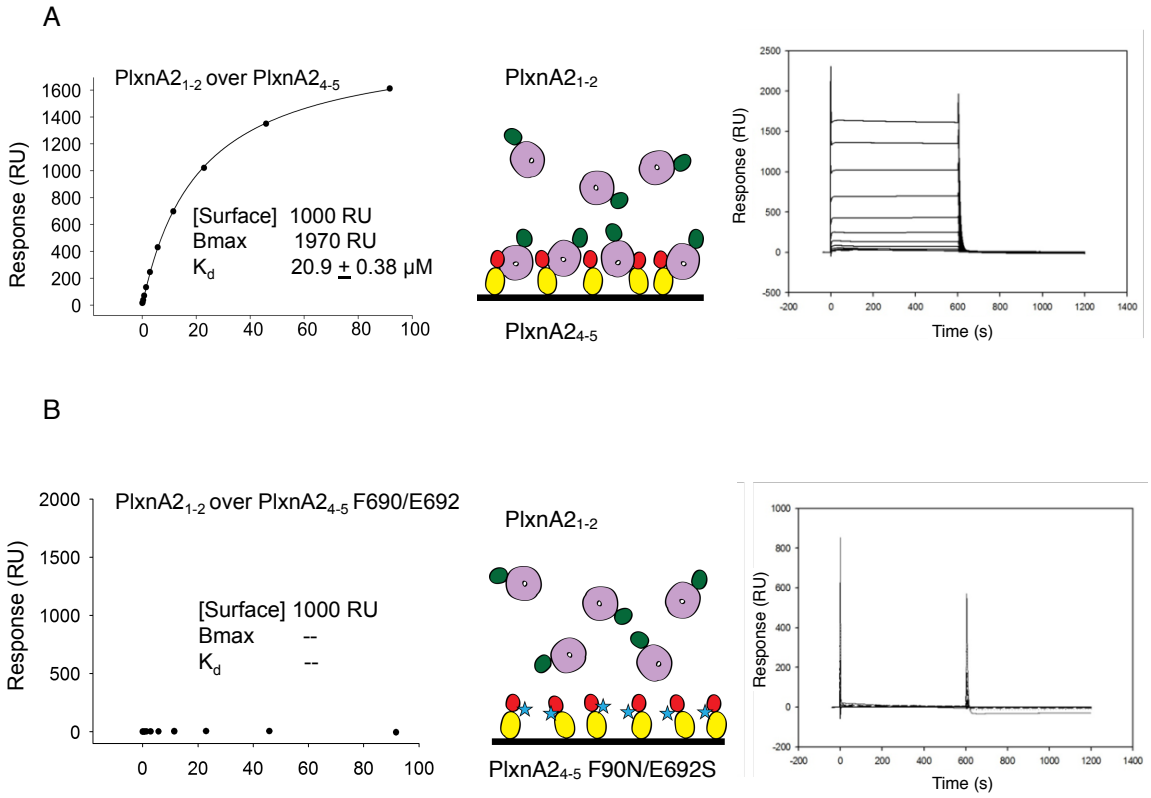


Figure 4.6: SPR equilibrium experiments for PlxnA2$_{1-2}$ binding to wild type PlxnA2$_{4-5}$ and its interaction-blocking mutant. (A) SPR equilibrium experiments for PlxnA2$_{1-2}$ binding to PlxnA2$_{4-5}$ revealed a Kd of 20.9 ± 0.38 μM. (B) SPR equilibrium experiments for PlxnA2$_{1-2}$ and to PlxnA2$_{4-5}$ F690N/E692S revealed no binding. In both (A) and (B), the equilibrium binding plots are shown in the left panels, in which R.U. is response unit. In the illustrations (middel panels), the sema domain, PSI1 domain, PSI2 domain and IPT2 domain are coloured in purple, green, red and yellow, respectively. The N-linked glycans introduced by the F690N/E692S mutation are represented as blue stars. The right panels show the SPR sensorgrams.

4.4 Localisation Microscopy Revealed the Lack of Large PlxnA Clusters on the COS-7 Cell Surface

As I discovered that the PlxnA extracellular segments interact in *cis* in solution, I set out to examine whether the same interactions can be observed on the cell surface. I first conducted localisation microscopy-based protein cluster analysis for PlxnA2 on the COS-7 cell surface. Our localisation microscopy setup, which increases the conventional resolution limit for yellow fluorescent protein (YFP)-tagged receptors (~250 nm) to ~50 nm, allowed determinination of the precise positions of single PlxnA molecules on the cell surface (111). For these experiments, I first expressed full-length, wild type PlxnA2 ($PlxnA2_{FL}$) and its interaction-blocking mutant ($PlxnA2_{FL}$ F690N/E692S) coupled to mVenus (a monomeric YFP variant) in COS-7 cells. To examine the effect of Sema6s binding to the oligomerisation states of PlxnA2, a group of cells expressing PlxnA2s were treated with Fc-tagged Sema6A dimers. All cells were then fixed with PFA (paraformaldehyde) and mounted with special anti-quenching reagents before imaging. After locating the positions of each PlxnA4 molecules in a selected region of interest (ROI) (Fig. 4.7.A), usually at the peripheries of the COS-7 cells, a frequency plot for the distribution of intermolecular distances can be generated (Fig. 4.7.B). The distribution of distances for each PlxnA2-expressing COS-7 cell is then compared with the distribution of distances generated from a set of homogenous, randomly distributed molecules of the same mean density. Therefore, any significant deviations from the simulated random distribution of points will indicate the presence of clustering.

In the frequency plot for the distribution of intermolecular distances, I did not see any significant deviations from the random distribution of points for both $PlxnA2_{FL}$ and $PlxnA2_{FL}$ F690N/E692S. However, at distances smaller than 50 nm a small degree of deviation for all samples treated or untreated with Sema6A was observed. This observation hints at the potential existence of small PlxnA clusters with a diameter of less than 50 nm. Unfortunately, due to the resolution limit of our setup I could not determine the extent of this potential, small-scale clustering of PlxnA2, nor could I compare the wild type and interaction-blocking PlxnA2s. I could only conclude that according to our localisation microscopy-based protein cluster analysis, the PlxnAs are unlikely to form large clusters (larger than 50 nm) on the cell surface. This observation is consistent with my structural models together with the results of the in-solution biophysical assays, which indicate that the formation of large PlxnA2 oligomers is unlikely due to weak affinity between the sema and PSI2-IPT2 domains as well as the steric hindrance created by the ring-like PlxnAs when engaging in head-to-stalk *cis*-interactions.

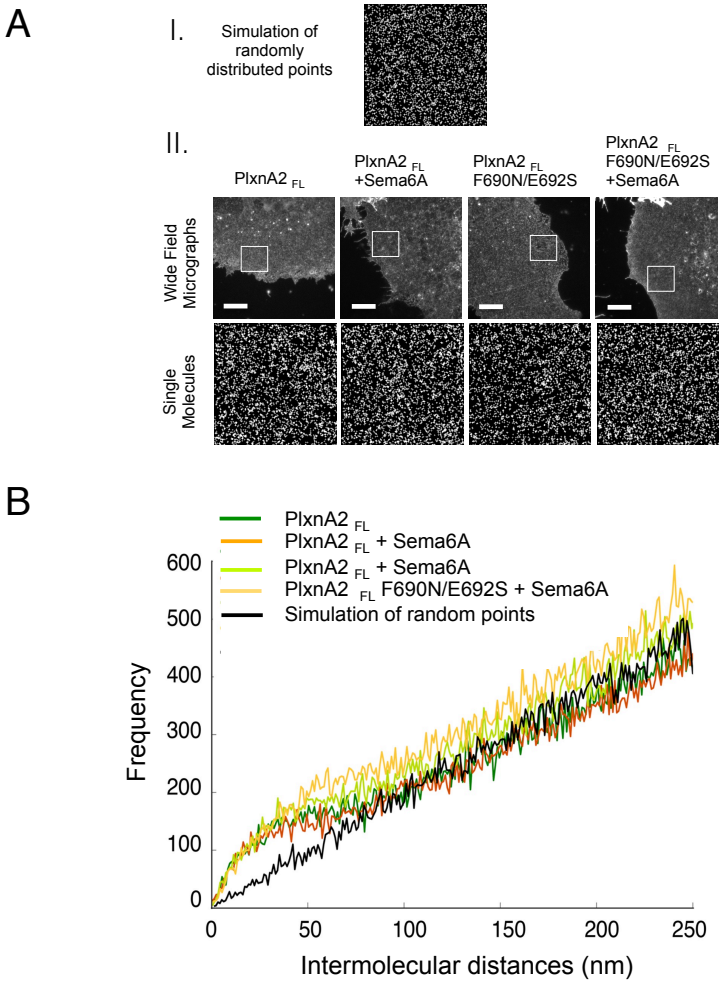


Figure 4.7: localisation microscopy-based cluster analysis for PlxnA2 revealed the lack of large-scale clustering. (A) A simulated image of a set of randomly distributed points (I) together with representative wide-field images (upper panels) and single molecule localisation microscopy images (lower panels) for PlxnA2$_{FL}$ and PlxnA2$_{FL}$ F690N/692S single molecules on COS-7 cells surface (II). The ROIs, indicated by white boxes in the wide field images, are used to generate single molecule localisation microscopy images, in which single PlxnA2 molecules are represented as white crosses. (B) Distribution of distances between single PlxnA2 molecules on the COS-7 cell surface compared to a set of randomly distributed points. Sample size, n = 3.

4.5 Cell Surface PlxnA2 Has Sema–PSI2-IPT2 Mediated *cis*-Interactions

4.5.1 The FRET-FLIM Setup

To probe the potential PlxnA *cis*-interactions on the cell surface as seen in the in-solution studies, I employed the FRET-FLIM (Förster energy transfer by fluorescence lifetime imaging microscopy) technique. FRET-FLIM is ideally suited to measure transient and low-affinity interactions in live cells as may occur for PlxnAs. The FRET donor–acceptor couple mTFP1 (a monomeric CFP variant) and mVenus were cloned onto the intracellular side of PlxnA2 extracellular-transmembrane segments via a RTLEVLFQGP linker. If two fluorescently tagged PlxnA2s interact via the sema–IPT2-PSI2 interface this linker provides sufficient flexibility for the fluorescent proteins to rotate and orient at a distance of less than 10 nm, allowing FRET to occur. To preclude any potential contribution of the cytoplasmic domain to the *cis*-interaction, the PlxnA2 intracellular segment was omitted from these constructs. An interaction-blocking mutant of PlxnA2, PlxnA2 F690N/E692S, was also coupled to the mTFP1–mVenus pair. In a no FRET situation where only mTFP1 tagged PlxnA2 is present, the lifetime of mTFP1 without the influence of the mVenus (τ_D) can be obtained by single exponential fitting of the fluorescence decay as shown in green (Fig. 4.8.A). When both the mTFP1- and mVenus-tagged PlxnA2 are present, should FRET occurs the donor mTFP1 would show accelerated fluorescence decay and a shorter lifetime (Fig. 4.8.B). By fitting the fluorescence

decay curve of mTFP1 in this case with a single (in green) and a double exponential (in orange) model, the fixed donor lifetime (τ_D), the discreet FRET lifetime (τ_F), the fraction of interacting donor (f_D), and the mean fluorescence lifetime ($\langle\tau\rangle$) can be obtained. Moreover, For each FLIM image of a live COS-7 cell, the PlxnA–PlxnA interactions can be quantified by f_D and $\langle\tau\rangle$ in a pixel-by-pixel basis at a spatial resolution of ~2-3 μm^2. A mTFP1-mVenus Tandem protein, which exhibits high level FRET due to the covalently linked mTFP1 and mVenus, was used as positive control.

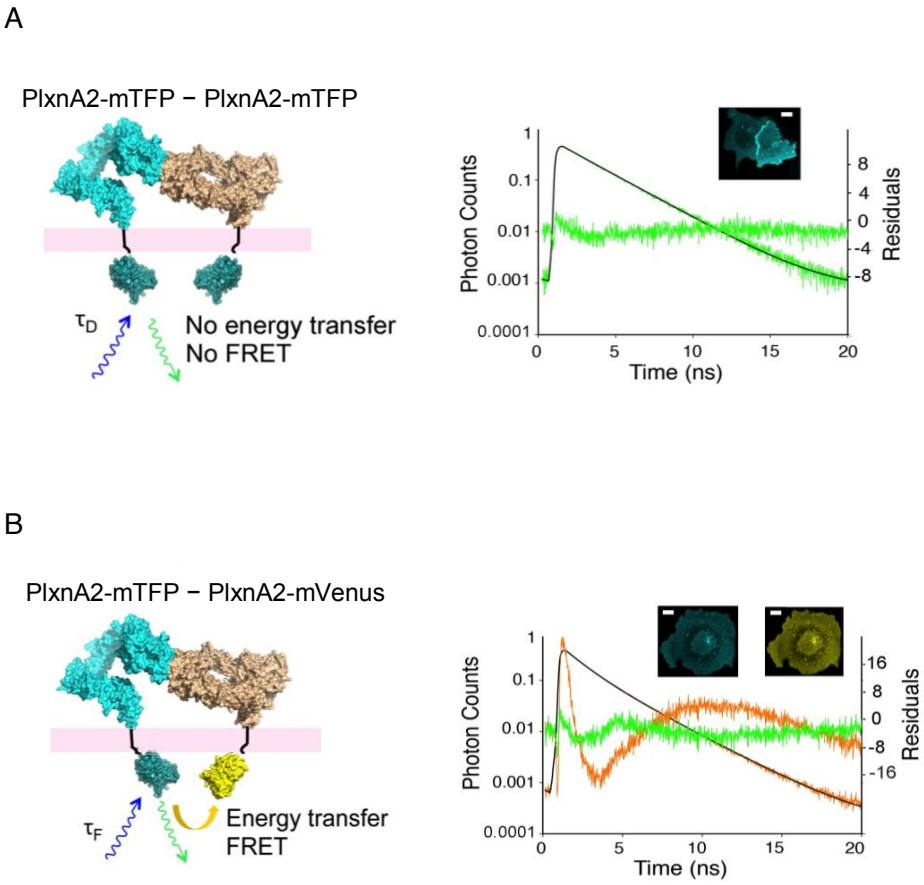


Figure 4.8: The FRET-FLIM setup for measuring PlxnA2–PlxnA2 interactions in live COS-7 cells. (A) In a no-FRET situation, only PlxnA2 labelled with mTPF1 (cyan) would interact (left). This results in a fluorescence decay profile for mTFP1 (green curve), which can be fitted with a single exponential decay model (black curve) (right). (B) During FRET, a mixed population of PlxnA2 labelled with mTPF1 (cyan) and mVenus (yellow) would interact (left). This results in an accelerated fluorescence decay for mTFP1 (orange curve), which can be fitted with a single plus a double exponential decay model (black curve) (right). The residuals of the single and double exponential fit are coloured in green and orange, respectively.

4.5.2 FRET-FLIM Revealed PlxnA–PlxnA Interactions on the Live Cell Membrane

In the FRET-FLIM experiments, most of the fluorophore-tagged PlxnA2 expressed in COS-7 cells were freely diffusing on the membrane. Due to overexpression, a fraction of PlxnA2 were also present in the cytosol or attached to cytoplasmic membranes. The FRET-FLIM results are best represented in Fig. 4.9, in which a representative cell expressing PlxnA2-mTFP1 alone showed a distribution of mean fluorescence lifetime ($\langle\tau\rangle$) from 2.6 ns to 2.8 ns with an average mean lifetime ($\overline{\langle\tau\rangle}$) of 2.73 ns and a distribution for fraction of interacting donors (f_D) with an average fraction of interacting donor $\overline{f_D}$ of 0.05 (Fig. 4.9.I). The $\overline{\langle\tau\rangle}$ in this case represent a no FRET situation as it is consistent with mTFP1 expressed alone as previously reported (106). In contrast, a cell co-expressing PlxnA2-mTFP1 and PlxnA2-mVenus showed a $\overline{\langle\tau\rangle}$ of 2.61 ns and an $\overline{f_D}$ of 0.11 (Fig. 4.9.II). Comparing to wild type PlxnA2, a cell co-expressing PlxnA2 F690N/E692S-mTFP1 and PlxnA2 F690N/E692S-mVenus exhibited lower FRET, with a longer $\overline{\langle\tau\rangle}$ of 2.65 ns and an lower $\overline{f_D}$ of 0.07 (Fig. 4.9.IV). These results indicate that the PlxnAs interact in *cis* on the cell surface and that this interaction is mediated by the sema–PSI2-IPT2 interface.

I next acquired FLIM images after incubating the fluorophore-tagged PlxnA2 on COS-7 cells with pre-dimerised Fc-tagged Sema6A. After Sema6A stimulation, a cell co-expressing PlxnA2-mTFP1 and PlxnA2-mVenus showed a $\overline{\langle\tau\rangle}$ of 2.55 ns and an $\overline{f_D}$ of 0.16 (Fig. 4.9.III). Similarly, a cell co-expressing PlxnA2 F690N/E692S-mTFP1 and PlxnA2 F690N/E692S-mVenus showed a level of FRET as high as the wild type PlxnA2 $\overline{\langle\tau\rangle}$ = 2.54 ns and $\overline{f_D}$ =

0.16) (Fig. 4.9.V). Interestingly, for both wild type and mutant PlxnA2 after Sema6A binding, I observed an increased number of small pixel clusters with high fraction of interacting donors ($f_D = 0.4 - 0.5$) (Fig. 4.9.III and V). These pixel clusters correspond to microdomains exhibiting high FRET with a size of $2 - 3\ \mu m^2$. The homogenous distribution of these microdomains signifies a high level of interaction occurring uniformly across the cell surface. This result points to a trend of increased PlxnA2 interaction upon Sema6A binding.

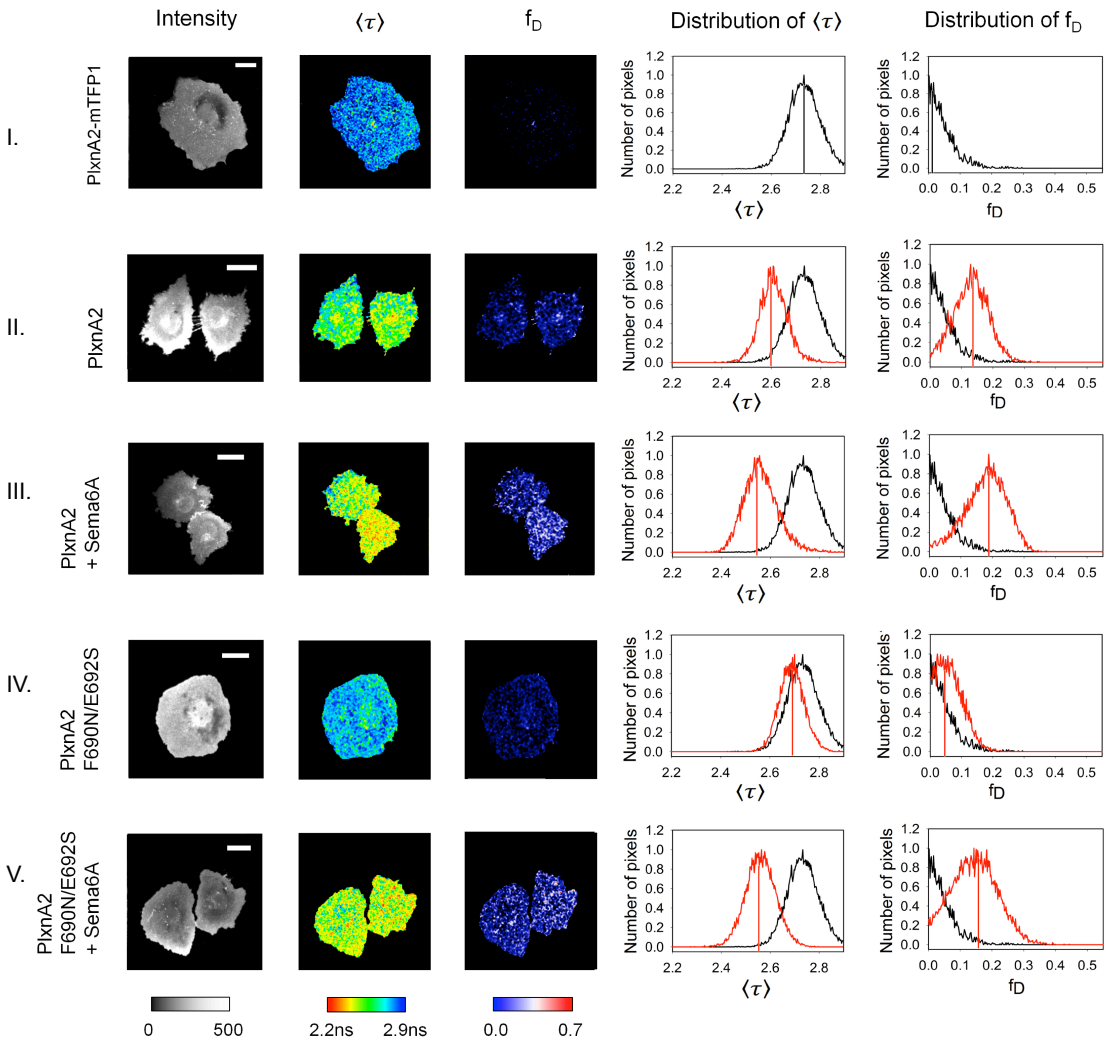


Figure 4.9: FRET-FLIM analysis of wild type PlxnA2 and its interaction-blocking mutants pre- and post-Sema6A binding. Representative cells expressing PlxnA2-mTFP1 alone (I), PlxnA2-mTFP1 and PlxnA2-mVenus before (II) and after (III) Sema6A stimulation, PlxnA2 F690N/E692S-mTFP1 and PlxnA2 F690N/E692S-mVenus before (IV) and after (V) Sema6A stimulation are shown from top to bottom. From left to right, the images represent the fluorescence intensity acquired by FLIM (in number of photons), the mean fluorescence lifetime $\langle\tau\rangle$, and the fraction of interacting donors (f_D) (shown in pseudo-colour scales). The histograms display the distribution of $\langle\tau\rangle$ and f_D of all pixels in each image. The distributions of $\langle\tau\rangle$ and f_D for PlxnA2-mTFP1 expressed alone and for coexpressed PlxnA2-mTFP1 and PlxnA2-mVenus are coloured in black and red, respectively. The lines within the peaks indicate the average mean lifetime ($\overline{\langle\tau\rangle}$) and the average fraction of interacting donor ($\overline{f_D}$). Scale bar, 10 µm.

4.5.3 The Quantification of PlxnA–PlxnA Interactions by FRET-FLIM

Indeed, as I analysed a large population of cells, I found that PlxnA2 F690N/E692S showed significantly less *cis*-interactions than PlxnA2 and that the Sema6A-binding induced increases to the same level of interaction for wild type PlxnA2 and PlxnA2 F690N/E692S (Fig. 4.10.A). In a sample size of more than 50 cells, I established that the $\overline{f_D}$ of PlxnA2-mTFP1 and PlxnA2 F690N/E692S-mTFP1 expressed alone is 0.05 ± 0.02 (Fig. 4.10.A.I), which is close to the zero baseline indicating a no-FRET situation. However, for the mTFP1- and mVenus-tagged PlxnA2 coexpressing on the same COS-7 cells, an $\overline{f_D}$ of 0.11 ± 0.02 was retrieved (Fig. 4.10.A.II). In contrast, significantly less interactions in *cis* were observed for PlxnA2 F690N/E692S, which yielded an $\overline{f_D}$ of 0.07 ± 0.02 (Fig. 4.10.A.IV). After the treatment of Sema6A, the $\overline{f_D}$ increased to the same value of 0.16 ± 0.03 for both PlxnA2 and PlxnA2 F690N/E692S (Fig. 4.10.A.III and V). The FRET-FLIM results demonstrate that indeed, PlxnAs *cis*-interaction is mediated by the sema–PSI2-IPT2 interface on live cell surface. Binding of semaphorins to PlxnAs can induce higher levels of interactions, a process most likely to be independent from the pre-signalling *cis*-interaction of PlxnAs. One should note that when plotting the PlxnA2-mTFP1/PlxnA2-mVenus intensity ratio (D/A) against the $\overline{f_D}$ values for a random selection of cells in each condition showed no correlation (Fig. 4.10.B). Therefore, the extent of FRET and thus, PlxnA–PlxnA interactions, is not related to the expression level of mTFP1- and mVenus-tagged proteins. Only cells expressing more acceptor than donor (PlxnA2-mTFP/PlxnA2-mVenus intensity ratio < 1), which sufficiently allows FRET to occur, were included in my subse-

quent analysis.

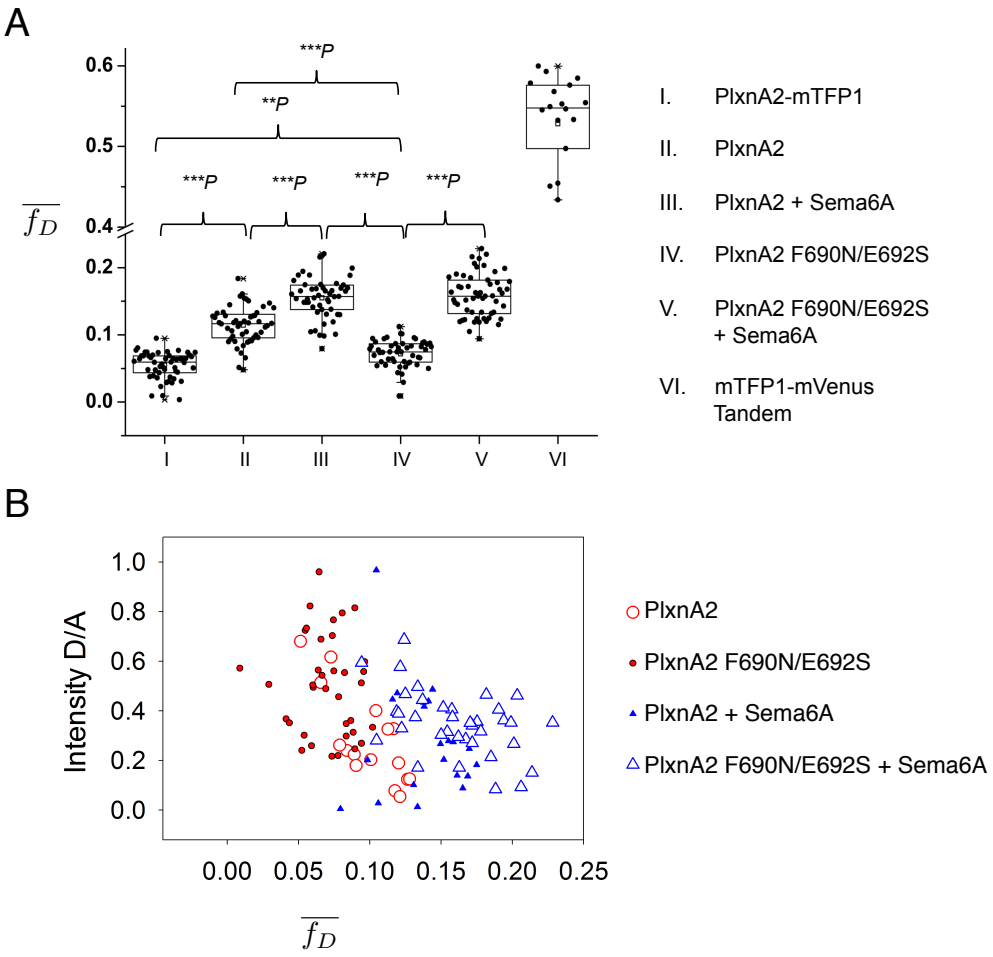


Figure 4.10: The quantification of PlxnA2–PlxnA2 interactions by FRET-FLIM. (A) The distribution of $\overline{f_D}$ for PlxnA2-mTFP1 alone (column I), PlxnA2 (column II), PlxnA2 treated with Sema6A (column III), PlxnA2 F690N/E692S (column IV), PlxnA2 F690N/E692S treated with Sema6A (column V) (mTFP1- and mVenus-tagged proteins in coexpression), and mTFP1-mVenus Tandem (column VI). The upper and lower quartiles of each sample are represented by the upper and lower sides of the boxes, the medians are represented by black lines, and the means by hollow points. The range of the whiskers indicates the statistical outliers with a coefficient of 1.5. The P-values signify the statistical significances between two selected samples determined by paired two-sample t-tests. Sample size, $n > 52$ per condition. $**P$, $P < 0.01$. $***P$, $P < 0.001$. (B) PlxnA2-mTFP1/PlxnA2-mVenus intensity ratio plotted against $\overline{f_D}$ for a set of cells randomly sampled from PlxnA2 and PlxnA2 F690N/E692S-expressing cells untreated or treated with Sema6A.

4.6 PlxnA4 is Autoinhibited by the Sema–PSI2-IPT2 *cis*-Interactions in COS-7 Cells

Having observed the *cis*-interactions based on Sema–PSI2-IPT2 interface for PlxnAs on the cell surface, I asked whether these interactions regulate PlxnA functions. I first hypothesized that the PlxnA–PlxnA interactions may be used for receptor autoinhibition (126). For this purpose I conducted COS-7 cell collapse assay, which is the standard cellular assay for plexin-mediated cytoskeleton collaspe (36, 129), in collaborations with Prof. Jeoren Pasterkamp's lab in University Medical Center Utrecht. We coexpressed full-length, wild type PlxnA4 ($PlxnA4_{FL}$), an interaction-blocking mutant, $PlxnA4_{FL}$ F689N/E691S, and a version of PlxnA4 in which the ectodomain is deleted, $PlxnA4_{\Delta ecto}$, in COS-7 cells. For control we also transfected COS-7 cells with eGFP alone. After 24 hours, the COS-7 cells were fixed and their cell areas were quantified.

We found that COS-7 cells expressing wild type $PlxnA4_{FL}$ showed a healthy and rounded morphology (Fig. 4.11.A.I) identical to the COS-7 cells expressing control protein eGFP (Fig. 4.11.A.II). In contrast, COS-7 cells expressing $PlxnA4_{FL}$ F689N/E691S showed drastic cytoskeleton collapse and a much reduced cell area (Fig. 4.11.A.III). This result is similar to the COS-7 cell-collapse induced by $PlxnA4_{\Delta ecto}$ (Fig. 4.11.A.IV), which was constitutively active consistent with previous findings (67, 126). Indeed, for a large number of COS-7 cells, extensive collapse in the cytoskeleton of cells expressing $PlxnA4_{FL}$ F689N/E691S and $PlxnA4_{\Delta ecto}$ showed significantly smaller cell surface area ($\leq$ 1600

μm^2) than those expressing wild type $PlxnA4_{FL}$ (Fig. 4.11.B). These results indicate that the PlxnAs are indeed autoinhibited by their extracellular segments before semaphorin-binding. Importantly, this autoinhibition of PlxnAs is mediated by the intermolecular interactions via the sema–PSI2-IPT2 interface, rather than the previously hypothesized intramolecular interactions.

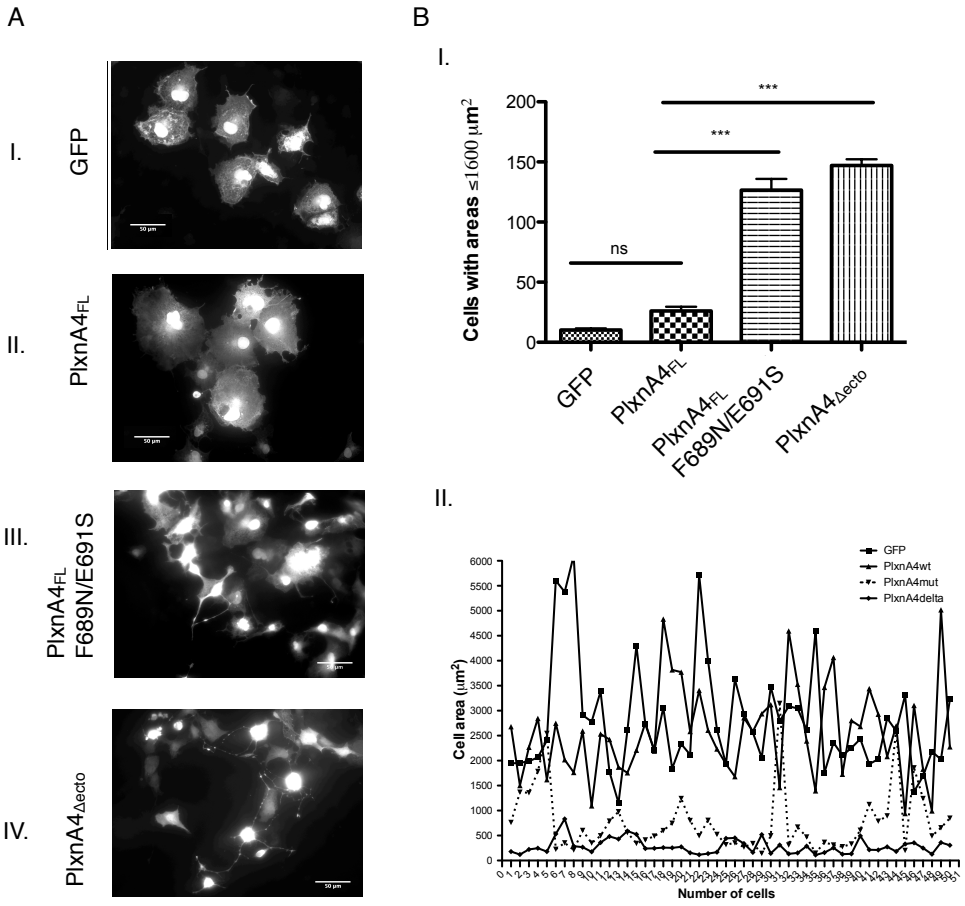


Figure 4.11: The autoinhibition of PlxnA4 based on the sema–PSI2-IPT2 interface in COS-7 cells. (A) Representative images COS-7 cells expressing eGFP (I), PlxnA4$_{\mathrm{FL}}$ (II), PlxnA4$_{\mathrm{FL}}$ F689N/E691S (III) and PlxnA4$_{\Delta\mathrm{ecto}}$ (IV). Scale bar, 50 µm. (B) Quantification of cell surface areas for COS-7 cells expressing eGFP and the PlxnA4 proteins, in which the average number of cells with areas less than 1600 µm^2 is shown in (I) and the cell areas of 50 consecutive cells are listed in (II). One way ANOVA with Bonferroni multiple comparison test was conducted, in which SEM (standard error of means) is shown as bars. Sample size, 350-440 cells per condition. $***P < 0.0001$. Ns, not significant.

4.7 Autoinhibition of PlxnAs Based on the Sema–PSI2-IPT2 Interface Prevents Premature DG Growth Cone Collapse

As PlxnA4 was found to be autoinhibited by their head-to-stalk (sema–PSI2-IPT2) interactions in COS-7 cells, I set out to examine whether this autoinhibitory mechanism is important for PlxnA functions in axon guidance. Again in collaborations with Prof. Jeoren Pasterkamp's group, we conducted functional experiments using dentate gyrus (DG) granule cells. The DG granule cells are characterised by their projecting axons tipped with growth cones endogenously expressing high levels of PlxnA2 and A4 and lower levels of PlxnA1 (130). In our experiments, we extracted DG granule cells on DIV1 (days *in vitro* 1) and incubated them with high concentrations (3 mg/ml) of purified PlxnA1 segments, the wild type PSI2-IT2 ($PlxnA1_{4\text{-}5}$) and its interation-blocking mutant $PlxnA1_{4\text{-}5}$ F693N/E695S. Based on our findings from the in-solution and COS-7 cell-based experiments, we predicted that the high amount of PlxnA1 PSI2-IPT2 proteins in the medium may engage in *trans*-interactions with the sema domains of PlxnAs on the DG growth cones and thus outcompete the original PlxnA *cis*-interactions using the same sema–PSI2-IPT2 interface. Should this outcompetition occur, the PlxnAs on the DG growth cones would be released from autoinhibition and start to constitutively signal, resulting in growth cone collapse. In comparison, as $PlxnA1_{4\text{-}5}$ F693N/E695S is not able to interact with the endogenously expressed PlxnAs in *trans*, it should cause little or no

perturbations to the DG growth cones.

Indeed, we found that when treated with PlxnA1$_{4\text{-}5}$, the majority of DG growth cones showed a retracted or collapsed morphology (Fig. 4.12.A.I). In contrast, incubation with PlxnA1$_{4\text{-}5}$ F693N/E695S caused little changes (Fig. 4.12.A.II), similar to the untreated growth cones (Fig. 4.12.A.III). As the degree of DG growth cone collapse induced by the treatment of PlxnA1 recombinant proteins was heterogenous, for quantification we categorised the extent of collapse using a scale from 1 to 10, where 1 indicates a healthy growth cone and 10 indicates a completely collapsed growth cone basing on the surface area (Fig. 4.12.B). The DG growth cones under the three different conditions (incubated with PlxnA1$_{4\text{-}5}$, PlxnA1$_{4\text{-}5}$ F693N/E695S or untreated) were then classified according to the 10 scales (Fig. 4.12.C). A two-tailed χ^2 test for the contingency of the collapse of DG growth cone and the treatment of PlxnA1 PSI2-IPT2 proteins retrieved a P-value of less than 0.0001, indicating the collapse in DG growth cones is indeed related to the exogenous wild type and mutant PlxnA1 proteins supplied in the medium. Importantly, within the most severely collapsed and the least collasped scales (scale 1-4 and scale 8-10, respectively) the number of untreated DG growth cones and the number of those treated with PlxnA1$_{4\text{-}5}$ F693N/E695S showed no difference according to a pair-wise t-test, while the number of growth cones treated with wild type PlxnA1$_{4\text{-}5}$ is significantly different ($P < 0.05$). These results indicate that the treatment of wild type PlxnA1$_{4\text{-}5}$ but not its mutant (PlxnA1$_{4\text{-}5}$ F693N/E695S) induced a trend of growth cone collapse. All together our results indicate that the sema–PSI2-IPT2 interactions contribute to PlxnA autoinhibition in DG growth cones. However, one should note that the

growth cone retraction observed in our case is less drastic than a full-blown collapse induced by semaphorin-binding (10), suggesting that the basal activity of PlxnAs, despite being constitutive without the autoinhibitory mechanism, is to a lesser extent than the canonical semaphorin-induced PlxnA activation. Nevertheless, suppressing the PlxnA constitutive activity via the sema–PSI2-IPT2 interactions is likely to prevent premature growth cone collapse and thus, is crucial in PlxnA-mediated axon guidance.

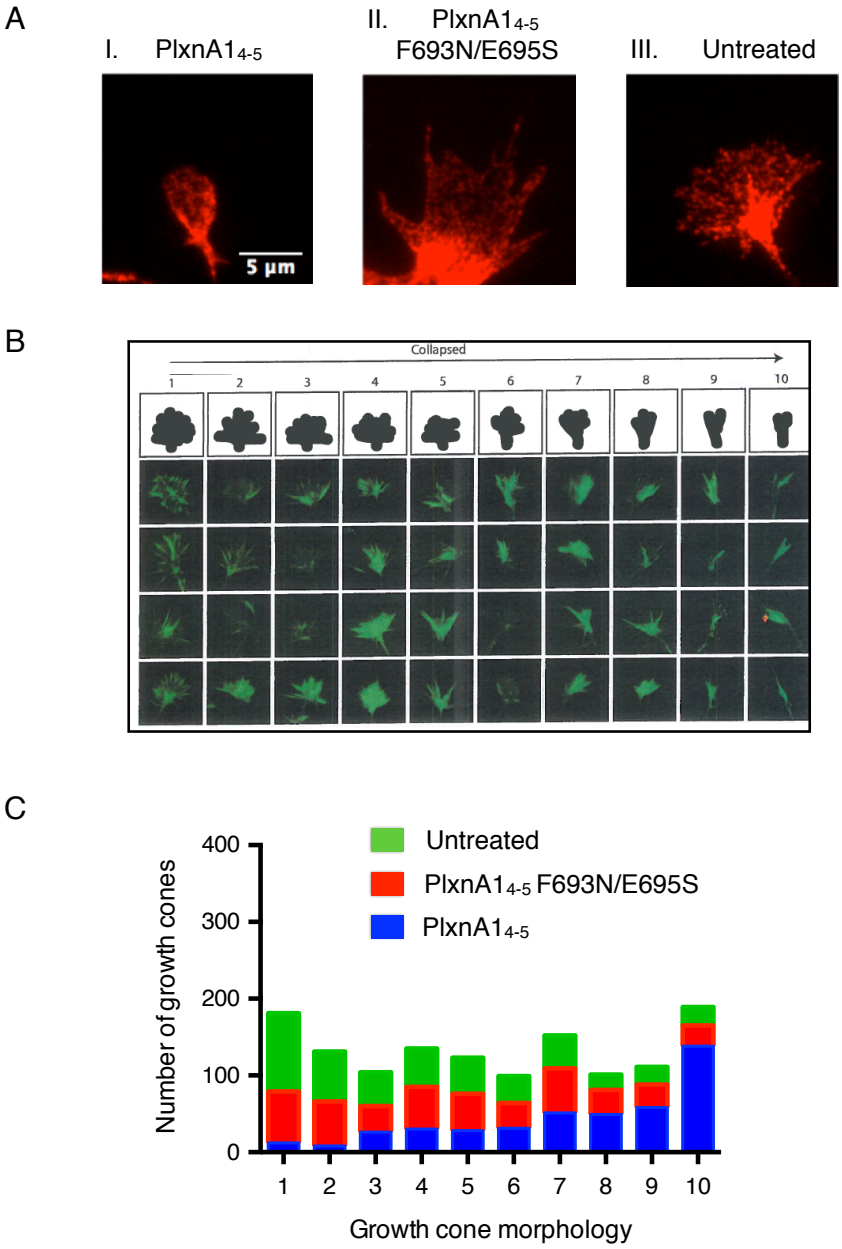


Figure 4.12: PlxnA autoinhibition based on the sema–PSI2-IPT2 interface prevents premature growth cone collapse in DG granule cells. (A) Representative growth cones of DG granule cells when incubated with wild type PlxnA1 PSI2-IPT2 ($PlxnA1_{4\text{-}5}$) (I), mutant PlxnA1 PSI2-IPT2 ($PlxnA1_{4\text{-}5}$ F693N/E695S) (II), and under untreated condition (III). Scale bar, 5 µm. The growth cones are stained for phalloidin (red). (B) Representative DG growth cones (green) categorized by a scale of 1 (healthy) to 10 (fully collapsed) based on their morphology. (C) Quantification of the number of DG growth cones falling into a scale of 1-10 as described in (B) when treated with $PlxnA1_{4\text{-}5}$ (blue), $PlxnA1_{4\text{-}5}$ F693N/E695S (red), and untreated (green).

4.8 Discussion

4.8.1 The Autoinhibition of PlxnAs based on *cis*–Interactions

Semaphorin-plexin signalling is subjected to multi-level checks and balances, in which the large, multidomain plexin extracellular segments present key control points. The regulatory mechanisms of plexins are exemplified by the PlxnAs. Previous studies together with the structures of PlxnA extracellular segments detailed in Chapter 3 pointed to a potential activation mechanism for PlxnAs, in which bivalent semaphorin-binding to PlxnAs in the extracellular region releases them from autoinhibtion and subsequently lead to dimerisation-based activation (22, 36, 37, 58). It also has been reported that the PlxnAs have a tendency to associate in solution and potentially pre-localised on the cell surface. From these observations a paradox emerged: should the PlxnAs have a tendency to pre-cluster on the cell surface and that activation is triggered by receptor dimerisation and potential clustering, then how is oligomerisation-based PlxnA self-activation avoided?

This thesis reveals a pre-signalling autoinhibitory mechanism based on the *cis*-interactions of PlxnA ectodomains, which may provide a simple solution to the aforementioned paradox. This *cis*-interactions based on an intermolecular, sema–PSI2-IPT2 (head-to-stalk) interface of PlxnAs, were first discovered in the crystal structures of PlxnA extracellular segments and later verified in solution as well as on the live cell surface. Functional experiments in COS-7 cells as well as neuronal growth cones further validated

my hypothesis, that is, before signalling the PlxnAs are autoinhibited by their head-to-stalk *cis*-interactions until the binding of semaphorins release them from this autoinhibited state. Together with the distinct ring-like conformation of the PlxnA ectodomains, this head-to-stalk interface may segregate the cytoplasmic region of PlxnAs sufficiently apart, while maintaining small-scale pre-signalling oligomerisation to facilitate bivalent semaphorin-binding.

4.8.2 Structural Models of PlxnA Activation and Regulation

Based on our studies on the class A plexins detailed in the present chapter and Chapter 3, I arrive at a structural model resolving the long-standing puzzle regarding PlxnA activation and regulation (Fig. 4.13). Before semaphorin-binding, the PlxnAs may pre-pack into small oligomers on the cell surface based on the intermolecular head-to-stalk interactions. Adopting the rigid, ring-like conformations, the PlxnA cytoplasmic domains are separated more than 100 Å apart and thus, despite the formation of small clusters the PlxnAs are inhibited from activating their Rap-GAP function (Fig. 4.13.A). This outside-together-inside-apart mechanism is perhaps the most effective way to preclude any accidental dimerisation and signalling of the plexin cytoplasmic regions. Moreover, the pre-associations of PlxnAs render them ready to engage in 2:2 interactions with semaphorins in *trans*. When transmembrane Sema6s come along, the Sema6s–PlxnAs interactions in *trans* outcompete the PlxnA–PlxnA *cis*-interactions, switching autoinhibition to activation by bringing together the C-termini of two PlxnA ectodomains as well as the subsequent cytoplasmic regions. Interestingly, based on our localisation microscopy re-

sults, I did not observe large scale clustering of PlxnAs even after Sema6A binding. This result is not surprising as once the sema domains are released from the pre-signalling head-to-stalk interactions, the PlxnA ectodomains lose their principal structural elements for the formation of *cis*–interacting, small-scale clusters. However, it is possible that the Sema3s–PlxnAs–Nrps complexes would further cluster due to the inherent oligomerisation properties of Nrps.

Similarly, I arrived at a potential model for the mutual silencing of PlxnAs and Sema6s based on *cis*-interactions. Should the PlxnA extracellular segments adopt a chair-like conformation as seen in the negative stain EM class averages (Chapter 3), the cytoplasmic domains are spaced even further apart (~ 130Å) pre-signalling and thus unable to self-activate (Fig. 4.13.B). Once the Sema6s are transported to the same cell surface as PlxnAs, they may again outcompete the original PlxnA–PlxnA *cis*-interactions. This swapping of inhibitory partners may lock the PlxnAs in a monomeric form, rendering them completely unresponsive to any semaphorins in *trans*. One should note that the existence of any of the aforementioned complexes on the cell surface requires further experimentations.

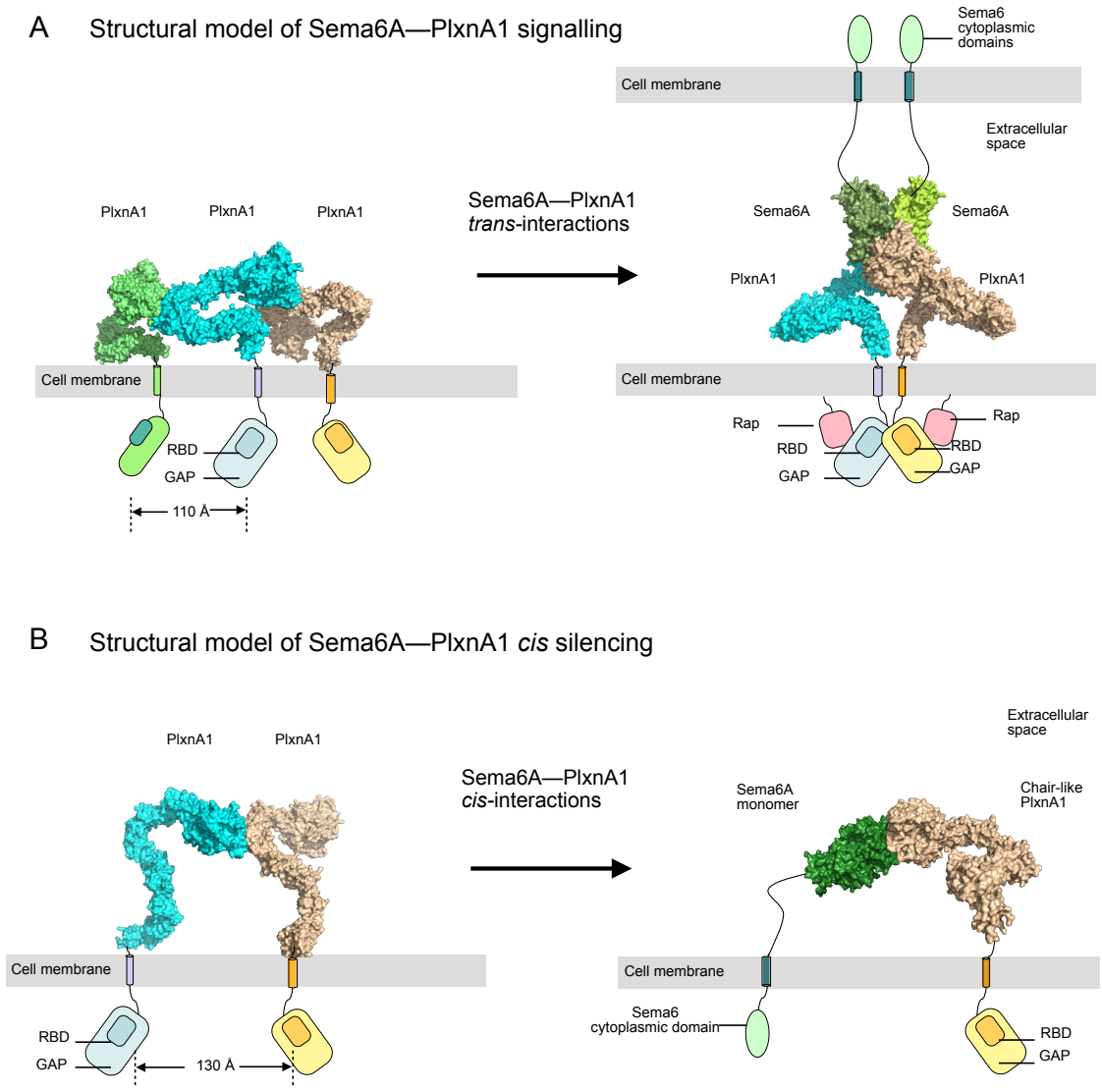


Figure 4.13: Structural models on the regulation of PlxnA signalling. (A) A model for Sema6s–PlxnAs signalling based on the ring-like conformation of the PlxnA ectodomains. Before signalling, PlxnAs may form potential dimers or trimers as modelled using the 6 Å PlxnA$_{1\text{-}10}$ crystal structure (space group P2$_1$). Sema6s-binding (modelled on the Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$ complex, PDB 3OKY) may allow the two ring-like PlxnAs to come into close vicinity and interact in the cytoplasmic region. (B) A model of the potential Sema6s–PlxnAs *cis*-silencing dimers based on the chair-like (twisted-open) conformation of PlxnAs (modelled as detailed in the negative stain EM section in Chapter 3). The formation of a trimer or larger oligomers based on this conformation is unlikely as discussed earlier. The 1:1 Sema6–PlxnA *cis*-complex is generated from the 6 Å crystal structure and negative stain EM class averages of PlxnA$_{1\text{-}10}$ detailed in Fig. 4.3. The C-terminal residues of PlxnA ectodomain, transmembrane, and cytoplasmic regions are represented by cartoons.

CHAPTER 5

STRUCTURAL STUDIES OF PLXND1 EXTRACELLULAR SEGMENTS

5.1 Introduction

5.1.1 Overview on PlxnD1 Signalling

The sole member of the class D plexins, PlxnD1, plays essential roles in the development of cardiovascular and neural systems but little is known regarding its signalling mechanism (41, 62, 131). Up to today, PlxnD1 remains the only class of vertebrate plexins without any published crystal structures. Recent studies have also highlighted the implications of PlxnD1 in cancer, revealing it as a particularly important therapeutic target

(62, 76). The cognate ligands of PlxnD1 include members of the secreted class 3 semaphorins, Sema3A, 3C, 3E, as well as the cell-attached Sema4A (41, 132, 133) (Fig. 5.1). Similar to the secreted class 3 semaphorins signalling through PlxnAs, Sema3A and 3C signal through PlxnD1 in the presence of coreceptor Nrp1 and Nrp2, respectively (41, 62). Transmembrane Sema4A has been shown to interact with PlxnD1 potentially also through Nrp1 (74). It is generally thought that these semaphorins activate PlxnD1 in an canonical manner, inducing discrete changes in the cytoskeleton.

The most interesting aspect of PlxnD1 signalling is its interactions with a particular member of the secreted class 3 semaphorins, Sema3E. It has been discovered a decade ago that Sema3E can uniquely signal through PlxnD1 without any neuropilins (72), while Sema3E signalling through PlxnA1 follows the canonical model where Nrp1 is required (62). The coreceptor-independent Sema3E–PlxnD1 signalling is thus known as an exception to the rule, which has puzzled researchers ever since (131, 134). The Sema3E–PlxnD1 system is further complicated by the discovery that coreceptor Nrp1 can gate signalling outcome, switching Sema3E–PlxnD1 induced axon repulsion to attraction (68). This bi-functionality of Sema3E gated by Nrp1 when signalling through PlxnD1 is reminiscent of the netrin–DCC system, in which coreceptor UNC5 switches the canonical netrin–DCC induced attraction to repulsion (31, 44). Sema3E has also been shown to compete with Sema4A for PlxnD1 binding, a process potentially involving Nrp1 (74). The Sema3E–PlxnD1 signalling axis was later found to cross-interact with other signalling systems such as VEGF–VEFGR2 and EGF-EGFR (HER2), although the formation of a hetero-multimeric complex involving components from two systems remains elu-

sive (55, 73).

What structural features make PlxnD1 unique and versatile? How does Sema3E independently activate PlxnD1 while other semaphorins require Nrp1? What is the structural basis of Nrp1-gated Sema3E bi-functionality? To answer this questions we undertook the structural characterisation of PlxnD1 in particular its extracellular segment.

5.1.2 The Extracellular Segment of PlxnD1

Based on sequence conservation analysis, the domain organization of PlxnD1 extracellular segment is identical to that of PlxnAs, in which an N-terminal sema domain is followed by three consecutive pairs of alternating PSI and IPT domains (PSI1, IPT1, PSI2, IPT2, PSI3, IPT3) and three additional IPT domains (IPT4-IPT6) (Fig. 5.1). For the PlxnD1 sema domain, its structural core should be highly conserved, adopting the typical seven β-propeller fold as seen in all other sema domains of PlxnAs, B1 and C1, as sequence analysis retrieved a ~30% identity between PlxnD1 and PlxnA-C sema domains. However, one may expect substantial differences in the semaphorin- and neuropilin-binding regions to account for the specificity of PlxnD1 in recognizing its ligands in particular Sema3E. Despite a overall sequence identity of ~27.3% (similar to the case for comparisons between other PlxnAs) and an identical domain organization with the PlxnA ectodomains, we cannot predict the conformation of PlxnD1 ectodomain. From the studies of PlxnA ectodomains we know that specific interdomain interfaces may be essential in determining the architecture of PlxnD1 extracellular segment.

Although much is still to be done, in the time span of this thesis I first obtained high-resolution crystal structures for the N-terminal segment of human PlxnD1 ectodomain, PlxnD1$_{1\text{-}2}$ (sema-PSI1). Distinct insertions on the semaphorin-binding interface of the PlxnD1 sema domain hint at determinants of structural specificity. When I subjected the complete PlxnD1 ectodomain (PlxnD1$_{1\text{-}10}$) to SEC, I uncovered surprisingly two major species. AUC for these two PlxnD1$_{1\text{-}10}$ species revealed that PlxnD1$_{1\text{-}10}$ may have a weak oligomerisation tendency and also multiple conformations. Most recently, I determined a 2.35 Å crystal structure for mouse PlxnD1$_{1\text{-}10}$, in which I observed a remarkable closed-ring conformation, although the refinement of this structure has not yet be completed due to the writing of this thesis. The current closed-ring architecture of PlxnD1 ectodomain, although appearing similar to the ring-like conformation of PlxnA ectodomains, showed a few marked differences from the later. These results provide tantalising evidence that the distinct conformation as well as potential structural variability of PlxnD1 ectodomain may be closely linked to its unique signalling modes.

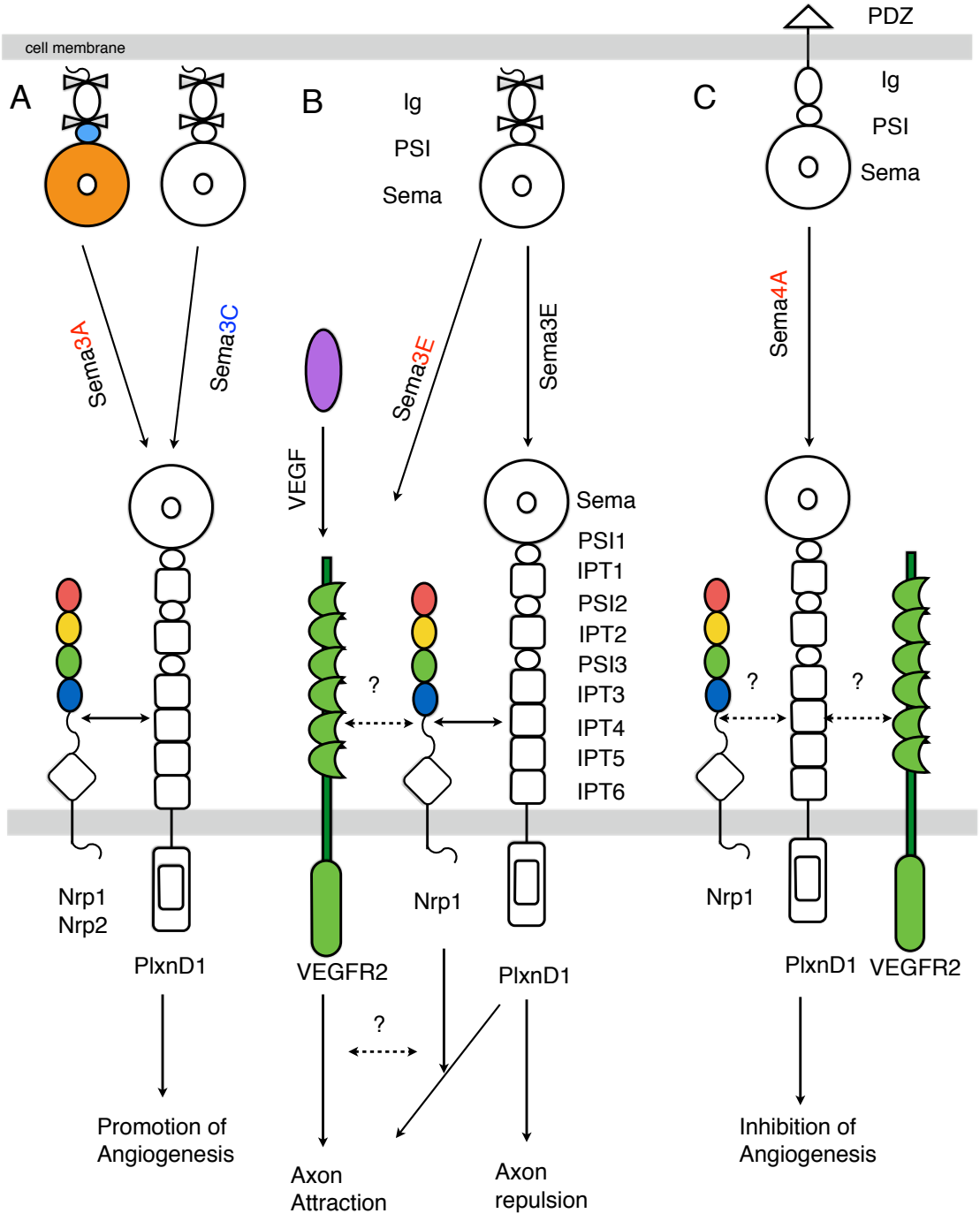


Figure 5.1: A summary of PlxnD1 signalling. Components of the Sema3A/3C–PlxnD1–Nrps (A), Sema3E–PlxnD1 (B), and Sema4A–PlxnD1 (C) signalling systems including the potential cross-interacting VEGF and VEGFR2 are illustrated as cartoons. Only the coloured domains have published crystal structures. The coloured labeling of semaphorins indicate the requirement for Nrp1 (red) and Nrp2 (blue) in signalling. The black solid arrows indicate established interactions while the dotted ones indicate potential interactions.

5.2 The Crystal Structures of Human PlxnD1$_{1\text{-}2}$ at 1.70 Å and 2.59 Å Resolutions

To obtain structural information on the PlxnD1 extracellular segment, a series of constructs containing different domain combinations of human, mouse and frog PlxnD1 were systematically screened for protein crystallisation. From these experiments I first obtained two forms of human PlxnD1$_{1\text{-}2}$ (sema-PSI1) crystals of space group C222$_1$ and P2$_1$2$_1$2, which diffracted to 1.70 Å and 2.59 Å resolutions, respectively. After the production from HEK293S cells the human PlxnD1$_{1\text{-}2}$ recombinant protein was purified using IMAC and SEC via standard protocols (Chapter 2, Section 2.2.2). After deglycosylation using endoglycosidase F1 (Endo F1), purified samples of human PlxnD1$_{1\text{-}2}$ was concentrated to 6.0 mg/ml in 10 mM HEPES pH 7.5, 150 mM sodium chloride and crystallised in space group C222$_1$ in 10 mM BICINE, pH 8.5, 20% (w/v) PEG 10,000. Another form of crystals, which later determined to be space group P2$_1$2$_1$2, crystallised in 20 mM disodium phosphate, 20% (w/v) PEG 3,350. Before data collection the PlxnD1$_{1\text{-}2}$ crystals were soaked in mother liquor supplemented with 25% (v/v) glycerol. The purification and crystallisation procedures for PlxnD1$_{1\text{-}2}$ are summarised in Fig 5.2.

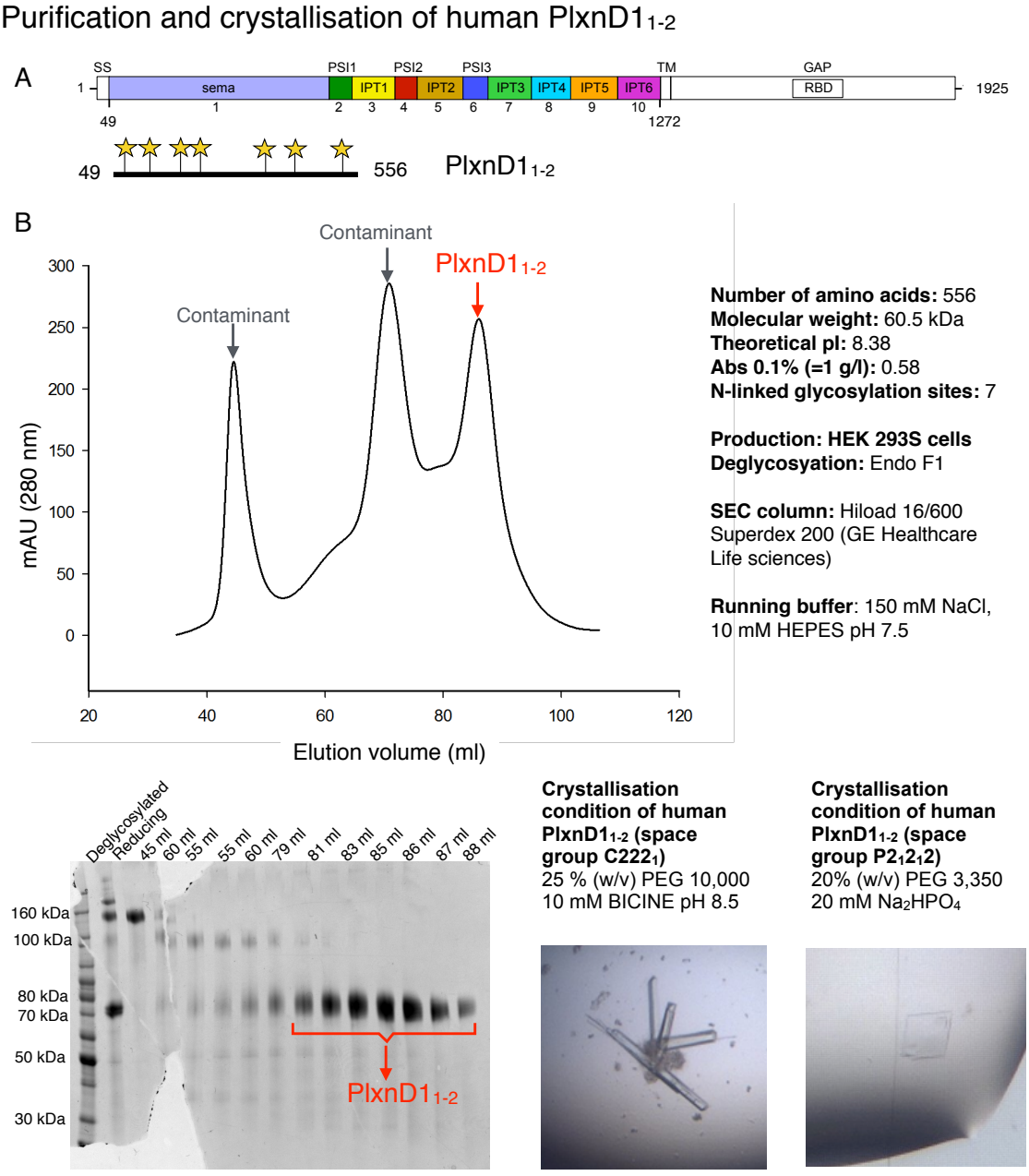


Figure 5.2: Summary of the purification and crystallisation of human PlxnD1$_{1\text{-}2}$. (A) Schematic illustration of the domain organization of human PlxnD1 and the PlxnD1$_{1\text{-}2}$ construct. The black line underneath the coloured domains of human PlxnD1 represents the segment covered by the current PlxnD1$_{1\text{-}2}$ construct (numbers at each end indicated the N-terminal and C-terminal residue, respectively). Yellow stars indicate the positions of predicted N-linked glycans. (B) The elution profile and SDS-PAGE analysis of the SEC experiment conducted for human PlxnD1$_{1\text{-}2}$. The basic information, production and SEC experimental conditions of PlxnD1$_{1\text{-}2}$ are listed at the right of the middle panel. The crystallisation conditions and pictures of PlxnD1$_{1\text{-}2}$ crystals are listed in the bottom right panel.

The two crystal structures of human PlxnD1$_{1\text{-}2}$ were solved by MR in PHASER using the sema-PSI1 domains of PlxnA2$_{1\text{-}4}$ (PDB 3OKT) (27.6% sequence identity) as initial models. The structures were then completed by model building in COOT and refinement in PHENIX. The data collection and refinement statistics of the two PlxnD1$_{1\text{-}2}$ crystal structures are listed in Fig. 5.3. In both crystal structures I observed the characteristic sema domain comprised of seven β-sheet blades as well as the immediate adjacent cysteine-rich PSI1 domain (Fig. 5.4.A). Consistent with the previously determined PlxnAs, B1 and C1 structures, the sema domain of PlxnD1 can be distinguished by specific residues in loop β1D-β2A between blade 1-2 and the extrusion between blade 4-5 (36, 37). The PSI1 domain is closely packed to the PlxnD1 sema domain similar to those in PlxnA2, PlxnB1 and PlxnC1. In the 1.70 Å crystal structure (space group C222$_1$), a total of fifteen residues in the loop regions of the sema domain as well as six residues near the C-terminal of PSI1 could not be determined due to missing electron density (represented as dotted curves in Fig. 5.4.A). The two PlxnD1$_{1\text{-}2}$ crystal structures obtained from different crystal forms are almost identical when superposed onto each other, except for the relative orientations for a few residues in the loop regions as well as different numbers of observable N-Acetylglucosamine (NAG)s (five in the 1.70 Å structure and four in the 2.59 Å structure) (Fig. 5.4.B). The root mean square deviation (rmsd) of all Cα atoms in the two superposed PlxnD1$_{1\text{-}2}$ structures is 0.89 Å, as calculated by the SuperPose server (135). Similarly to the 1.70 Å structure, a total of 18 residues in the sema domain and 6 residues in PSI1 domain of the 2.59 Å structure remain missing due to lack of electron density (not shown).

Interestingly, in the sema domain of PlxnD1 I observed a 13-residue loop extending from blade β3B-β3C (217-229, GSSFFPRNRSLED) (Fig. 5.4.A-C). This particularly long loop, which harbours an NAG linked to N224, is highly conserved in all vertebrate PlxnD1s but non-existent in any other plexins (Fig. 5.4.D and Fig. 5.6). When superposing PlxnD1$_{1\text{-}2}$ onto the previously characterized PlxnA2$_{1\text{-}2}$ (PDB 3OKT) (Fig. 5.4.E), I found that this 13-residue loop extends right into the potential semaphorin–plexin binding interface close to the semaphorin-interacting loop β1D-β2A (36). The similar observation was made when superposing the sema domain of PlxnD1 to those of PlxnB1 and PlxnC1 (omitted here). I therefore suspect that this 13-residue, glycosylated loop unique to PlxnD1 sema domain may potentially contribute to the binding specificity between PlxnD1 and its cognate semaphorins in particular Sema3E, although further experimentations are needed to assess this hypothesis.

	PlxnD1$_{1\text{-}2}$	**PlxnD1$_{1\text{-}2}$**
Data collection		
Space group	*C222$_1$*	*P2$_1$2$_1$2*
Cell dimensions		
a, *b*, *c* (Å)	98.3, 114.9, 105.2	59.1, 101.3, 118.8
α, β, γ(°)	90, 90, 90	90, 90, 90
Resolution (Å)	60.9-1.70 (1.74-1.70)[a]	77-2.59 (2.66-2.59)[a]
R_{merge} (%)	7.4 (85.5)	15.3 (94.1)
R_{pim} (%)	3.2 (28.7)	6.0 (30.1)
$CC_{1/2}$	0.999 (0.729)	0.997 (0.909)
I/σ*I*	16.5 (2.4)	12.4 (2.6)
Completeness (%)	99.5 (99.9)	99.8 (99.5)
Redundancy	6.6 (6.4)	6.5 (6.7)
Refinement		
Resolution (Å)	60.92-1.70	72-8.0
Number of reflections	62338	21695
R_{work}/R_{free} (%)	0.18 / 0.21	0.19 / 0.24
Number of atoms	4229	3991
Protein	4096	3742
Ligand/ion	0	0
Water	333	59
B factors (Å^2)		
Protein	29.9	58.2
Ligand/ion	37.7	41.3
Water	29.1	54.6
R.m.s. deviations		
Bond lengths (Å)	0.008	0.008
Bond angles (°)	1.15	1.24

Figure 5.3: Crystallographic and refinement statistics of human PlxnD1$_{1\text{-}2}$ crystal structures. The superscript a (^{a}) indicates that statistics of the highest resolution shell is shown in the parenthesis.

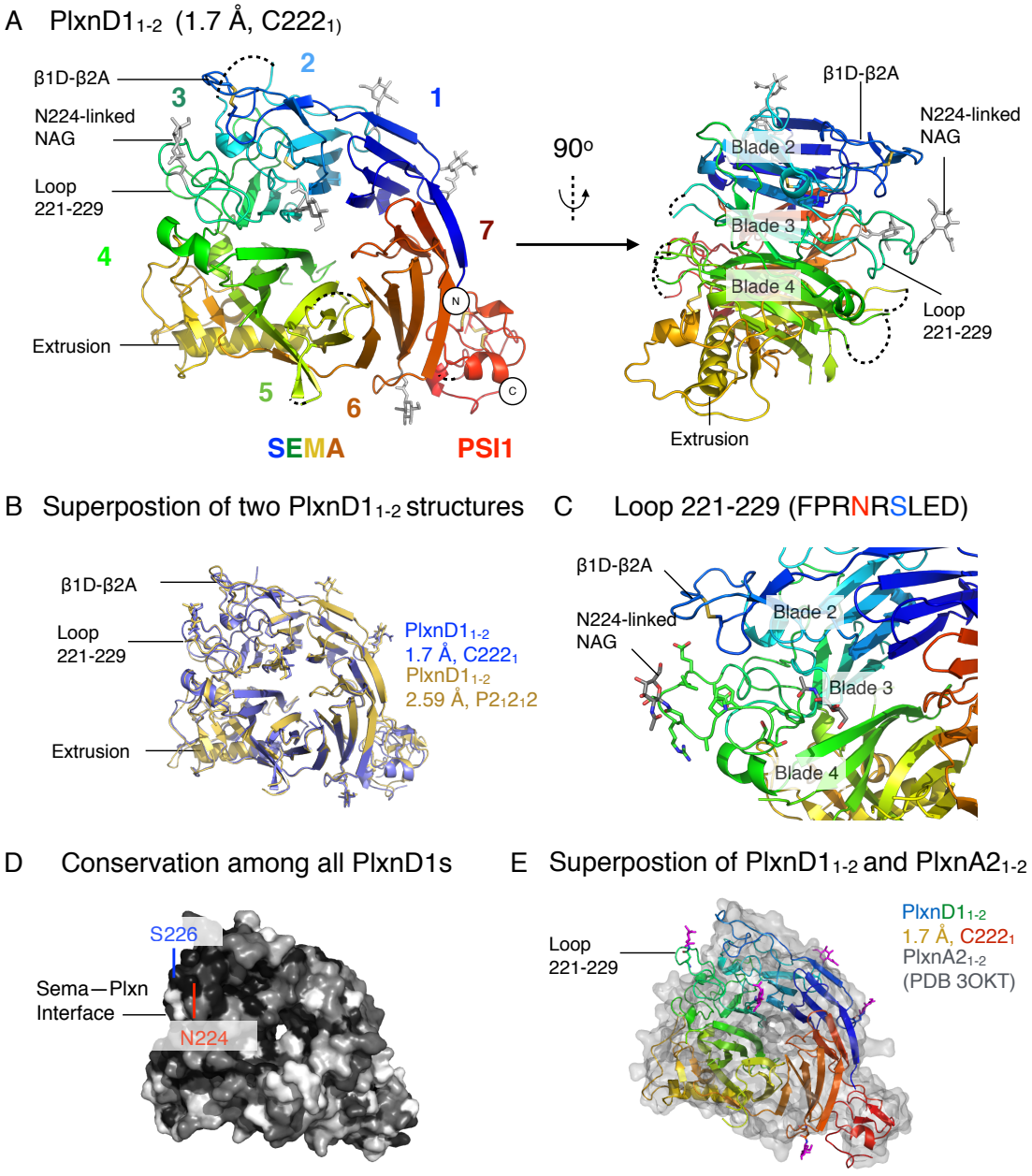


Figure 5.4: The crystal structures of PlxnD1$_{1-2}$. (A) The 1.7 Å crystal structure of PlxnD1$_{1-2}$ (space group C222$_1$) in two views. The missing loops are represented by dotted curves. (B) Superposition of the 1.7 Å (space group C222$_1$) and 2.59 Å (space group P2$_1$2$_1$2) resolution crystal structures of PlxnD1$_{1-2}$. (C) Structural details of the β3B-β3C loop (217-229, GSSFFPRNRSLED) in PlxnD1$_{1-2}$ (shown in the 1.7 Å , space group C222$_1$ structure). (D) Sequence conservation of surface residues in all vertebrate PlxnD1$_{1-2}$ coloured from white, non-conserved to black, conserved. (E) Superposition of PlxnD1$_{1-2}$ (space group P2$_1$2$_1$2) shown in cartoon (rainbow colour) and PlxnA2$_{1-2}$ (PDB 3OKT) shown in surface representation (grey). The N-linked glycans are shown as stick models (grey in A-C and magenta in E).

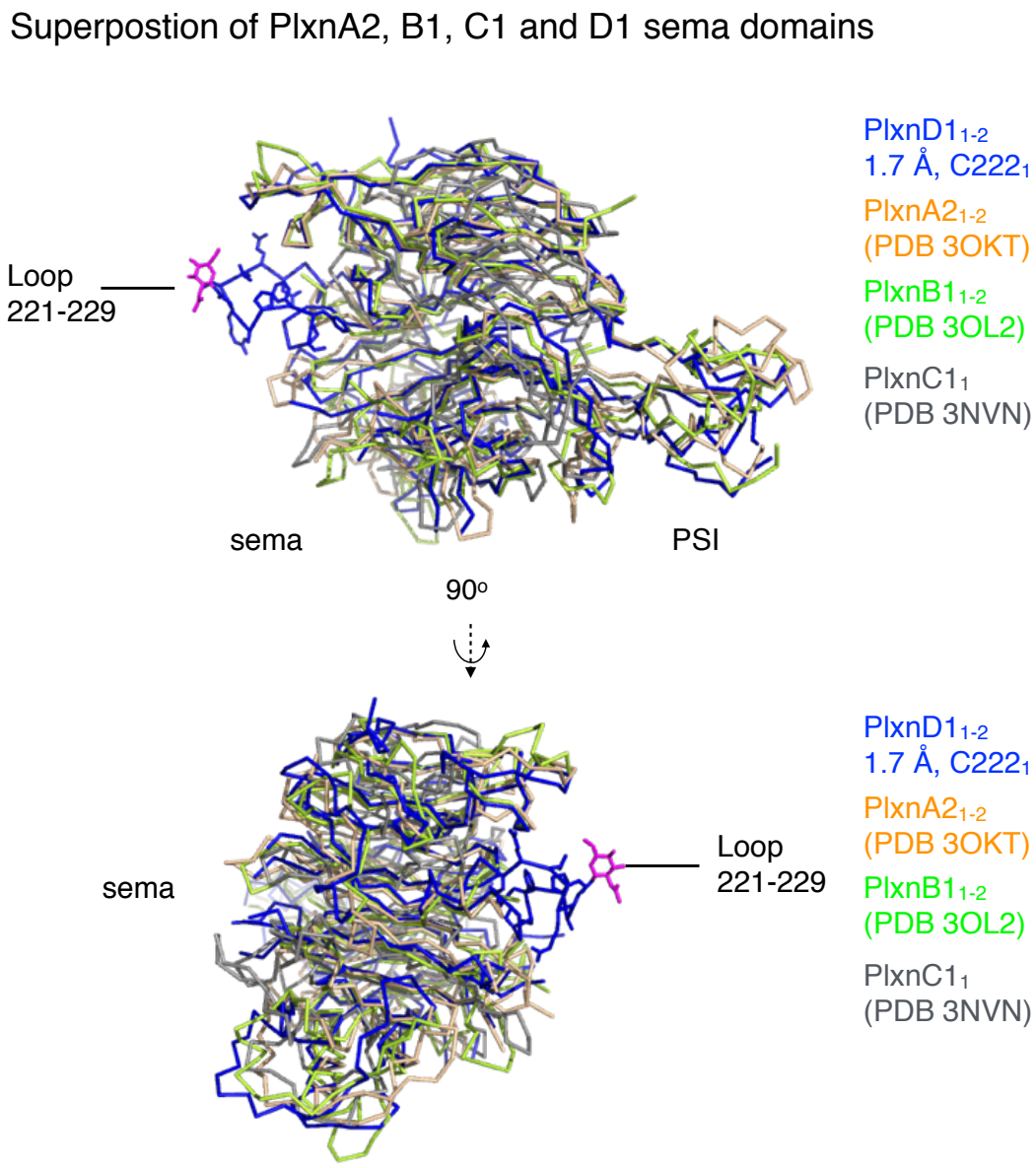


Figure 5.5: Superposition of the sema-PSI domains of class A-D plexins. Side chains on the PlxnD1 $\beta 3B-\beta 3C$ loop (217-229, GSSFFPRNRSLED) of the 1.7 Å PlxnD1$_{1\text{-}2}$ crystal structure (blue, space group C2221) are shown. The N-link Glycan at N224 is coloured in magenta. The sema domains of PlxnA2$_{1\text{-}2}$ (PDB 3OKT), PlxnB1$_{1\text{-}2}$ (PDB 3OL2), and PlxnC1$_1$ (PDB 3NVN) are shown in wheat, light green, and light grey, respectively.

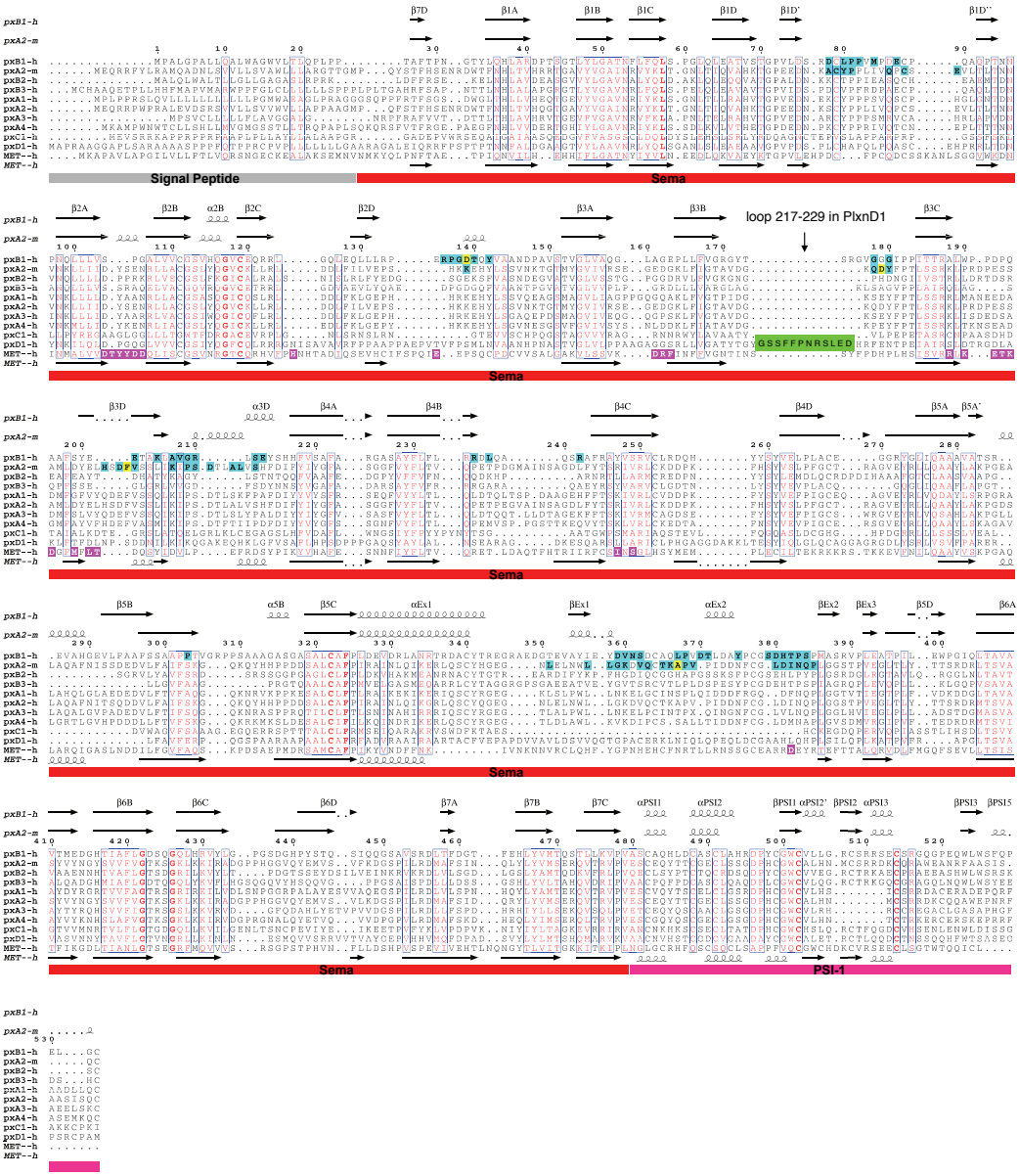


Figure 5.6: Sequence alignment of the sema-PSI1 (domains 1-2) segment of plexin family members and Met receptor. Sequences of the sema-PSI (domains 1-2) of all human plexin family members, mouse PlxnA2 and human Met were aligned based on the structures of the sema domains of human PlxnB1 (PDB 3OL2), mouse PlxnA2 (PDB 3OKT) (36) and human Met19 (PDB 1SHY) (136) using Indonesia (http://xray.bmc.uu.se/dennis) and on sequence similarity with ClustalW (137). Loop 217-229 (GSSFFPRNRSLED) in human PlxnD1 sema domain is highlighted in green. Numbering corresponds to the full length human PlxnB1 (including the secretion signal sequence). Residues on the semaphorin–plexin binding interface in mouse PlxnA2 and human PlxnB1 are coloured in cyan and yellow. Residues on the scatter factor–Met receptor interface are coloured in pink. Domains are indicated by coloured rectangles below the alignment. This figure is modified from a previous publication by our group (36) prepared with ESPript (138).

5.3 The PlxnD1 Ectodomain

5.3.1 Purification of PlxnD1 Full Ectodomain ($PlxnD1_{1\text{-}10}$)

When I subjected the purified and deglycosylated human PlxnD1 full ectodomain ($PlxnD1_{1\text{-}10}$) produced in HEK293S cells to size exclusion chromatography (SEC) using a S200 16/60 Superdex column (GE Healthcares), I found that it associated with one major contaminant. $PlxnD1_{1\text{-}10}$ eluted at retention volume maximum of ~65 ml while the major contaminant eluted at retention volume maximum at ~61 ml (Fig. 5.7.A). SDS-PAGE in a non-reducing environment for the different elution volumes in SEC uncovered two different bands corresponding to the two SEC species, $PlxnD1_{1\text{-}10}$ and major contaminant (Fig. 5.7.A.IV). To examine whether the contaminant could be a different species of $PlxnD1_{1\text{-}10}$ resulted from different complex glycosylations or proteolytic cleavage, I conducted western blot experiments for the two retention volume maximum fractions (Fig. 5.7.A.III). The western blot result shows that only $PlxnD1_{1\text{-}10}$ at elution volume maximum of ~65 ml has the C-terminal His-tag used for purification. Thus I can confirm that species eluted at 55-65 ml are contamination proteins. Further experiments are needed to examine which species of protein the contaminant could be. It would be useful to also test the fully glycosylated versions of $PlxnD1_{1\text{-}10}$ produced from HEK293T cells would behave when also subjected to SEC.

5.3.2 AUC Reveals the Potential Monomer State for PlxnD1$_{1\text{-}10}$

To characterize human PlxnD1$_{1\text{-}10}$ in solution, I first performed AUC velocity experiments. Human PlxnD1$_{1\text{-}10}$ purified as described earlier was separated and subjected to AUC velocity experiments at two concentrations, 1.5 mg/ml (10.5 μM) and 0.75 mg/ml (5.25 μM). My AUC results indicate that PlxnD1$_{1\text{-}10}$ is mostly monomeric in solution, as one dominant peak with sedimentation coefficient as an sedimentation coefficient c(s) value of 6.1S was detected. One should note that several small peaks of large sedimentation coefficients were also seen for PlxnD1$_{1\text{-}10}$. These peaks suggest the presence of weakly associating PlxnD1$_{1\text{-}10}$ oligomers, although further experiments using much high concentrations than 1.5 mg/ml are needed to test this possibility. It is also possible that PlxnD1$_{1\text{-}10}$ adopt diferent distinct conformations which cannot be resolved by my current AUC experiments. Further MALS, protein mass spectrometry, and negative stain EM experiments are needed to examine these possibilities.

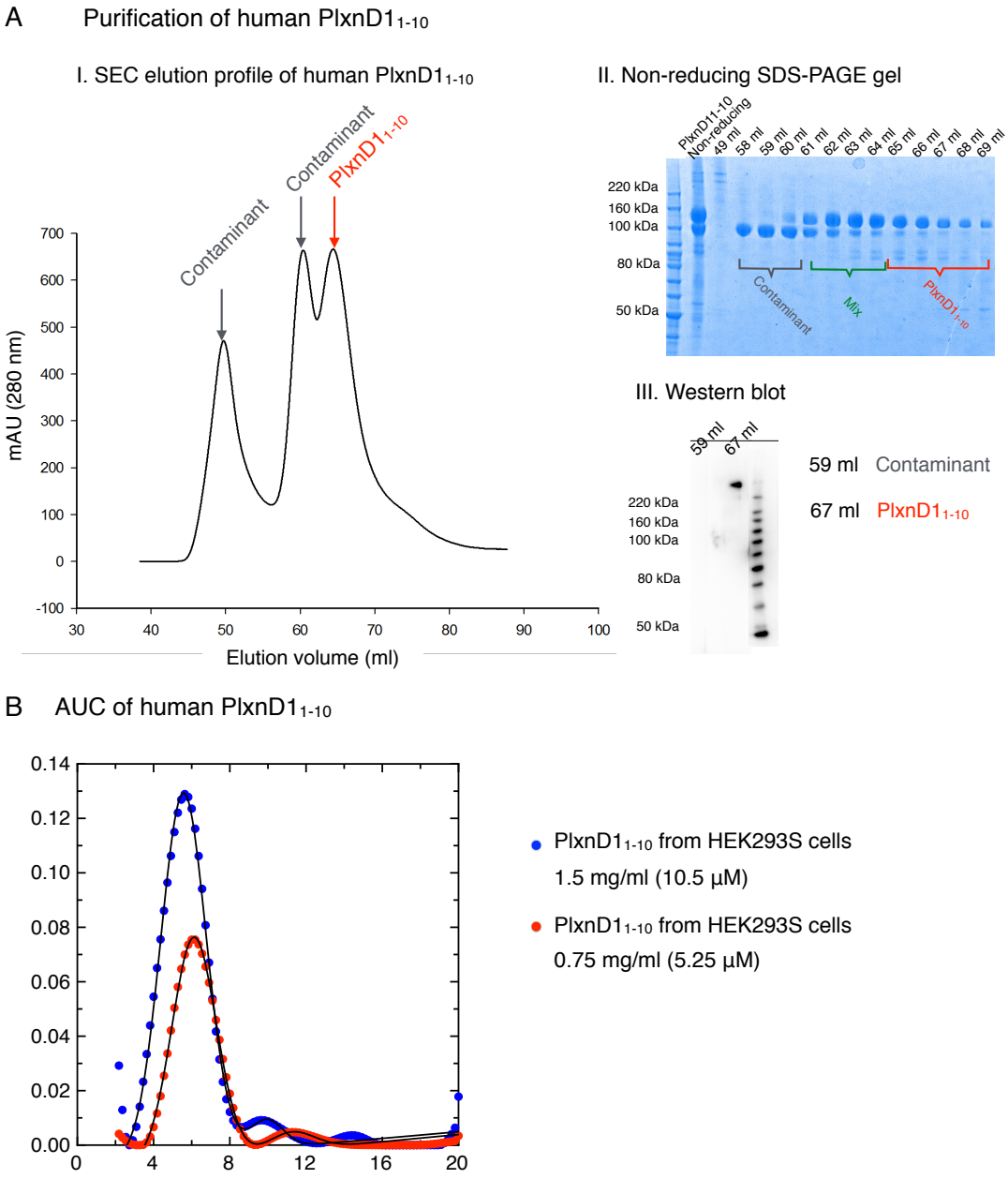


Figure 5.7: Purification and AUC experiments of PlxnD1$_{1\text{-}10}$. (A) The SEC elution profile (I),corresponding SDS-PAGE analysis at reducing conditions (II) and western blot (III) for human PlxnD1$_{1\text{-}10}$. (B) The distribution of sedimentation coefficient (s, in Svedbergs) for PlxnD1$_{1\text{-}10}$ at 1.5 mg/ml (10.5 µM; blue) and 0.75 mg/ml (5.25 µM; red).

5.4 Crystal Structure of Mouse PlxnD1$_{1\text{-}10}$ Revealed a Closed-ring Conformation

A week before the submission of this thesis, I obtained a crystal structure of the full ectodomain of mouse PlxnD1 (PlxnD1$_{1\text{-}10}$) at 2.35 Å resolution (space group C2). Similar to human PlxnD1$_{1\text{-}10}$, mouse PlxnD1$_{1\text{-}10}$ produced from HEK293S cells separated into two species in SEC experiment after IMAC purification and deglycosylation with Endo F1. The two species were isolated and indenpedently concentrated to 6 mg/ml in 10mM HEPES pH 7.5, 150 mM sodium chloride. Within 48 hours, mouse PlxnD1$_{1\text{-}10}$ (species 2) crystallised in space group C2 in 0.19 mM CYMAL®-7, 10 mM HEPES pH 7.5, 40% (v/v) PEG 400. Subsequent optimisation screens based on the same conditions were set up using diluted samples of mouse PlxnD1$_{1\text{-}10}$. From these screens I obtained a large amount of bigger crystals, some of which were made into PlxnD1$_{1\text{-}10}$ platinum derivative crystals by soaking the native PlxnD1$_{1\text{-}10}$ crystals into potassium tetrachloroplatinate solution for 1 hour before freezing. When freezing the PlxnD1$_{1\text{-}10}$ crystals nothing extra was added as precipitants in the crystallisation solution, 40% (v/v) PEG 400, also served as cryo-protectant. The purification and crystallisation procedures are summarised in Fig.5.8.

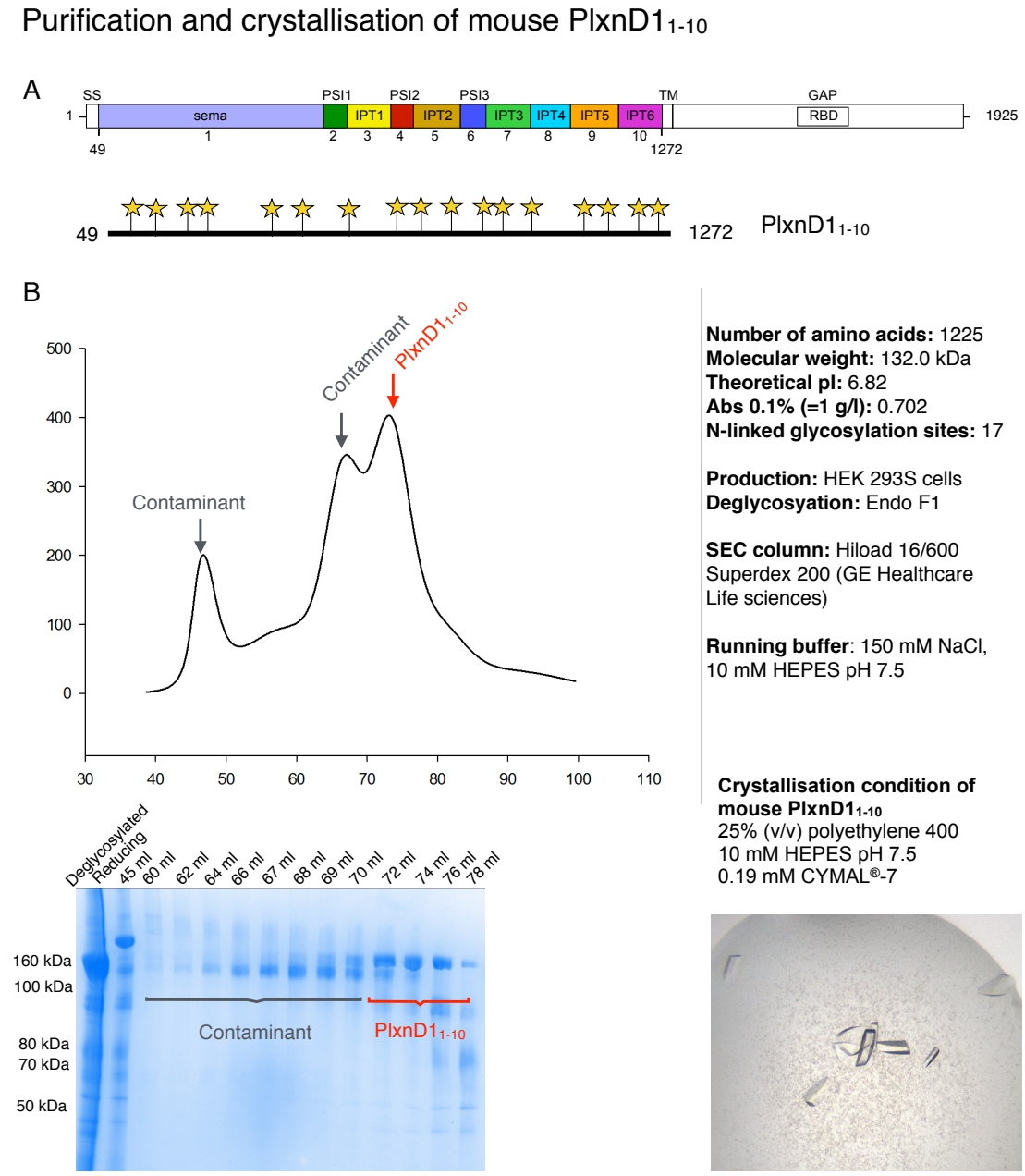


Figure 5.8: Summary of the purification and crystallisation of mouse PlxnD1$_{1\text{-}10}$. (A) Schematic illustration of the domain organization of mouse PlxnD1 and the PlxnD1$_{1\text{-}10}$ construct. The black line underneath the coloured domains of mouse PlxnD1 represents the segment covered by the current PlxnD1$_{1\text{-}10}$ construct (numbers at each end indicated the N-terminal and C-terminal residue, respectively). Yellow stars indicate the positions of predicted N-linked glycans. (B) The elution profile and SDS-PAGE analysis of the SEC experiment conducted for mouse PlxnD1$_{1\text{-}10}$. The basic information, production and SEC experimental conditions of the PlxnD1$_{1\text{-}10}$ recombinant protein are listed at the right of the middle panel. The crystallisation conditions and a picture of the native PlxnD1$_{1\text{-}10}$ crystals are shown in the bottom right panel.

The crystal structure of mouse PlxnD1$_{1\text{-}10}$ was solved using platinum single isomorphous replacement with anomalous scattering (Pt-SIRAS). A total of five datasets collected from five PlxnD1$_{1\text{-}10}$ Pt-derivate crystals were merged using Xia2 3dii (139). The merged dataset showed an excellent anomalous signal up to 4.03 Å resolution. Ten datasets collected from ten PlxnD1$_{1\text{-}10}$ native crystals diffracted to higher than 2.50 Å were also merged to achieve a high redundancy and a final resolution of 2.35 Å. After identifying five platinum sites using HKL2MAP (140), the native and Pt-derivative datasets were used to perform experimental phasing via AutoSol in PHASER (141). The native dataset was also used together with the previously determined human PlxnD1$_{1\text{-}2}$ structure as initial model to generate a partial model of mouse PlxnD1$_{1\text{-}10}$ encompassing the sema–PSI1 segment (PlxnD1$_{1\text{-}2}$) using MR in PHASER. This partial model was used to aid the determination of experimental phases as well as subsequent structure solution and refinement. Experimental phasing in AutoSol yielded an initial electron density map, which after density modification in PHENIX (142), showed excellent density for the sema–PSI1 segment as well as the last four IPT domains (IPT3-IPT6). After placing the homology models of mouse PlxnD1$_{1\text{-}10}$ IPT3-IPT6 generated from the previously determined 4 Å mouse PlxnA1$_{1\text{-}10}$ crystal structure (space group $P4_32_12$) into the corresponding densities, the improved electron density map allowed the placing of the remaining IPT domains (IPT1 and IPT2). The electron density map generated from this model revealed a clear, ring-like form for PlxnD1$_{1\text{-}10}$, in which the last two remaining domains, PSI2 and PSI3, can be placed using a combination of manual and automatic building via AutoBuild in PHENIX (89). Several rounds of manual and automatic re-trained refinement were conducted. Please note that the refinement is incomplete as

this stage due to time constraints. The data collection and current refinement statistics of the mouse $PlxnD1_{1\text{-}10}$ crystal structures are listed in Fig. 5.9.

	$PlxnD1_{1\text{-}10}$ (merged from 10 crystals)	**$PlxnD1_{1\text{-}10}$-Pt (merged from 5 crystals)**
Data collection		
Space group	*C2*	*C2*
Cell dimensions		
a, *b*, *c* (Å)	273.5, 69.6, 88.6	273.3, 69.1, 89.0
α, β, γ(°)	90, 103.0, 90	90, 102.7, 90
Resolution (Å)	133.2-2.35 (2.41-2.35)[a]	133.8-2.84 (2.81-2.84)[a]
R_{merge} (%)	22.5 (240)	20.3 (168)
$I/\sigma I$	28.9 (2.4)	21.6 (1.1)
Completeness (%)	99.5 (99.9)	95.3 (66.0)
Redundancy	127.9 (142.6)	69.5 (19.1)
Refinement		
Resolution (Å)	86.19-2.35	-
Number of reflections	66490	-
R_{work}/R_{free} (%)	21.1/24.6	-
Number of atoms		-
Protein	7452	-
Ligand/ion	0	-
Water	0	-
B factors (Å^2)		
Protein	31.1	-
Ligand/ion	-	-
Water	-	-
R.m.s. deviations		
Bond lengths (Å)	0.003	-
Bond angles (°)	0.67	-

Figure 5.9: Data collection and refinement statistics of mouse $PlxnD1_{1\text{-}10}$ crystal structures. The superscript a (^{a}) indicates that statistics of the highest resolution shell is shown in the parenthesis.

In the crystal structure, PlxnD1$_{1\text{-}10}$ makes a ~230 Å ring similar to the previously characterised PlxnA ectodomains using its first 9 domains (sema-IPT5) (Fig. 5.10.A). Strikingly, the IPT5 (domain 9) stalk region contacts the sema domain, creating a closed-ring structures. The head-to-tail (sema–IP5) interface appears to be maintained by several residues on both sides, although I cannot make out the details at the current stage. In particular, the last domain, IPT6, also has a distinct, out-of-the-ring orientation in PlxnD1 as in PlxnAs. The diameter of the PlxnD1$_{1\text{-}10}$ closed-ring is ~86 Å, which is most similar to the 84 Å of the 4 Å resolution, closed-ring structure of PlxnA1$_{1\text{-}10}$ (space group $P4_32_12$) (Fig. 5.10.B). One should note that while the closed-ring conformation of the 4 Å PlxnA1 ectodomain is mostly likely due to dehydration of the crystals, I do not know at this stage whether the closed-ring conformation of PlxnD1$_{1\text{-}10}$ was also induced by crystal packing or is indeed sustained under physiological conditions. When I superposed the current structural model of PlxnD1$_{1\text{-}10}$ onto the PlxnA full ectodomain structures (two of PlxnA1$_{1\text{-}10}$ and one of PlxnA4$_{1\text{-}10}$ as detailed in Chapter 3, Section 3.3.1) based on their sema domains, I found that the relative interdomain orientations of the PlxnD1 and PlxnA rings are quite different (Fig. 5.10.C).

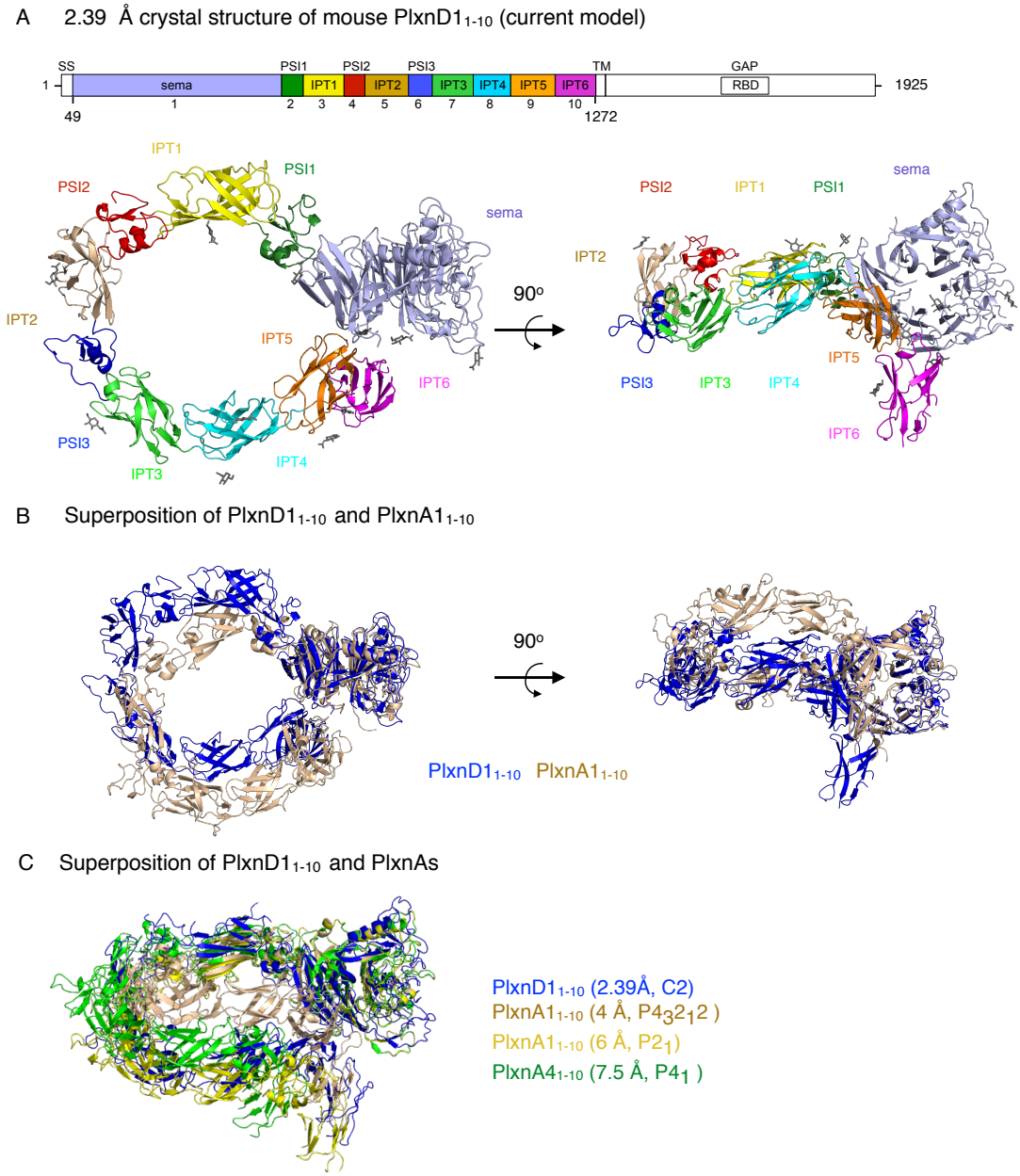
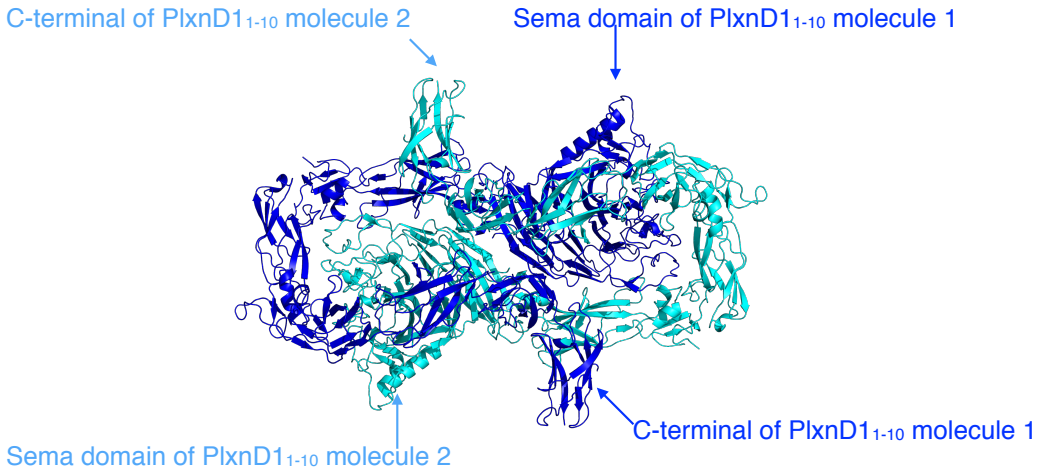


Figure 5.10: Crystal structure of PlxnD1$_{1-10}$ has a closed-ring conformation. (A) Current structural model of PlxnD1$_{1-10}$ obtained from the 2.35 Å merged dataset (space group C2) in two views. (B) Superposition of PlxnD1$_{1-10}$ and PlxnA1$_{1-10}$ (4 Å, space group P4$_3$2$_1$2) closed-ring structures. (C) Superpostion of PlxnD1$_{1-10}$ shown in blue, PlxnA1$_{1-10}$ (4 Å resolution, space group P4$_3$2$_1$2) shown in wheat, PlxnA1$_{1-10}$ (6 Å resolution, space group P4$_3$2$_1$2) shown in magenta, and PlxnA4$_{1-10}$ (7.5 Å resolution, space group P4$_1$) shown in green.

Remarkably, in the tightly packed PlxnD1$_{1\text{-}10}$ crystal structure, two neighboring molecules arrange in an interlocking manner, in which the sema domain of one PlxnD1 slots into the center of the sema-PSI5 ring of another PlxnD1, wrapping its stalk around its partner then closing it by the head-to-tail interactions (Fig. 5.10.A). Although the two PlxnD1$_{1\text{-}10}$ molecules in each pair appear to be related to each other by point symmetry, there is only one molecule in an asymmetric unit. These interlocking pairs then pack together in tight arrays to form extensive crystal contacts (Fig. 5.10.B). Each pair of the interlocking PlxnD1 observed in the crystal is formed by the two-fold symmetry of the C2 space group. However, an arrangement like this would be quite impossible in the context of the cell-surface as the two C-termini point in exact opposite directions. The PlxnD1 ectodomains also make several other crystal contacts. Once structure refinement is completed, it would be interesting to assess whether any of these contacting interfaces are conserved during evolution and result in the weak associations seen in AUC.

A An interlocking pair of mouse PlxnD1$_{1\text{-}10}$ in its crystal (space group C2)



B Arrays formed by interlocking pairs of PlxnD1$_{1\text{-}10}$

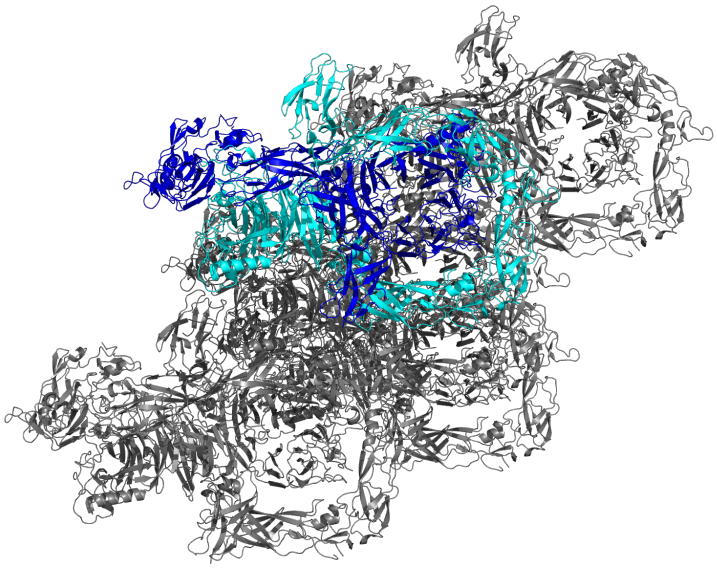


Figure 5.11: Packing of mouse PlxnD1$_{1\text{-}10}$ in its cystal. (A) A pair of PlxnD1$_{1\text{-}10}$ (blue and cyan) packed in an interlocking mode. (B) The interlocking pairs of PlxnD1$_{1\text{-}10}$formed extensive arrays in the crystal. Two PlxnD1$_{1\text{-}10}$ molecules locked into a pair are shown in blue and cyan, while their symmetry mates are shown in grey.

5.5 Discussion

PlxnD1 stands out amongst all plexins for its important roles in cardiovascular development and human cancer (41, 131). However, we know little about its signalling mechanism due to the complete absence of structural data. Uniquely, PlxnD1 can signal with secreted Sema3E without the presence of any coreceptors (72). The Sema3E–PlxnD1 system has also been found to have bi-functionality gated by Nrp1 as well as cross-interactions with several other TM1 receptor systems (68, 131). Our journey to understanding the molecular mechanism of PlxnD1 signalling has just begun.

The high-resolution crystal structures of the PlxnD1 extracellular segment I obtained in thesis have provided first-hand insights into the structural basis of PlxnD1 signalling. The N-terminal segment of human PlxnD1, sema-PSI, has a structural core highly similar to the previously characterized sema-PSI segments of PlxnAs, PlxnB1 and PlxnC1. Interestingly, despite that the structural elements on the semaphorin-binding interface in PlxnD1 are mostly conserved with those of PlxnA2 and PlxnB1 (Appendix 2), I observed a 13-residue, β3B-β3C loop unseen in other plexins. The distinct conformation of the loop as well as an N-linked glycan it harbours hints at potential contribution to the semaphorin-binding specificity of PlxnD1. To test this hypothesis, I would need to perturb this particular loop and assess its effects on PlxnD1 ligand-binding properties via structural-guided in solution assays. So far, all my PlxnD1 crystal structures are in an unliganded form. To the uncover the structural basis of PlxnD1 activation, it is essential to obtain crystal structures of PlxnD1 in complex with its semaphorin ligands. Since

Sema3E–PlxnD1 signalling represents the most interesting aspect of PlxnD1 biology, I have set out to screen for Sema3E–PlxnD1 complex crystals. Unfortunately, my work has met with no success so far. However, I have recently obtained small, needle-like crystals which diffracted very weakly for mouse Sema3E alone. Immediate future experiments will thus focus on obtaining well-diffracting crystals for Sema3E in isolation as well as in complex with PlxnD1.

Should PlxnD1 follows the common, bivalent mode of semaphorin–plexin recognition, which is most likely the case due to high sequence conservation with other plexin receptors, we are again presented with the same puzzle regarding the activation of PlxnD1 signalling. How is semaphorin-binding in the extracellular space translated into dimerisation-based activation of the PlxnD1 cytoplasmic region? From in-depth studies of the PlxnA ectodomains, as documented in Chapter 3-4, we now know that the plexin ectodomains may adopt distinct conformations specially designed to suit this purpose. Remarkably, my new structural model of PlxnD1 full ectodomain has a closed-ring conformation strikingly similar to those of the PlxnAs, in which the first nine domains (sema-IPT5) form a ring from which the tenth domain (IPT6) points perpendicularly away. However, the PlxnD1 and PlxnA architectures are also markedly different in that the PlxD1 ectodomain first, is a closed ring and second, the varied interdomain orientations in the two structures result in a different orientation of ring in relation to the sema domain. However, it is not until I complete the refinement of the PlxnD1$_{1\text{-}10}$ structure that I can analyse the structural elements responsible for the distinct architecture of the PlxnD1 ectodomain. It would also be useful to know whether PlxnD1$_{1\text{-}10}$ sustains the closed-

ring like conformation across different crystal forms. It is possinle that PlxnD1$_{1\text{-}10}$ has conformations other than the close-ring conformation seen in the current crystal structure. To assess the possible structural variability of PlxnD1 ectodomains when released from constraints of crystal packing, it will be important to conduct in-depth AUC, small angle X-ray scattering (SAXS), negative stain or even cryo-EM experiments in the future.

CHAPTER 6

CONCLUSIONS

The development and maintenance of higher organisms depends on the effective communication between cells. One of the key methods of cell–cell communication is ligand-receptor signalling on the cell surface. The TM1 receptors, which play essential roles in the nervous, cardiovascular, and immune systems, constitute the most abundant and diverse type of cell-surface receptors in mammals (4, 20). Signalling by TM1 receptors controls a plethora of cell functions, in particular cell morphogenesis, motility, adhesion and survival. In the past decades a central paradigm for TM1 receptors has emerged, in which signalling is triggerd by receptor dimerisation followed by clustering upon ligand-binding. Although this central paradigm, derived mostly from structural characterisations, can describe the behaviour of many TM1 receptor signalling systems in astonishing detail, it is incomplete. The current theory has little to say about the multifaceted

regulatory mechanisms of TM1 receptors, in particular, the exquisite balance between inhibition and activation of signalling. The regulation of TM1 receptors is examplified by the cell guidance process, in which a strikingly small number of cell guidance cues govern the formation and homeostasis of complex organs, for instance, the wiring of neural circuits. This phenomenon requires that the cell guidance cue signalling systems must diversify their repertoires. In practice, this diversification is often achieved through the multi-level and spatial-temporal regulation of the guidance cue receptors. In this thesis, I characterized two homolog proteins belonging to a classic guidance cue receptor family, the class A and D plexins. My findings point to a model in which the regulation of signalling is achieved by the distinct conformations as well as *cis*-interactions of the receptor ectodomains.

Our understanding on the molecular mechanisms controlling the semaphorin–plexin signalling system has been severely limited by the lack of information on the architecture and interaction modes of full length plexin extracellular segments. My structural analyses in this thesis for the class A plexins now provide the missing links which couple the extracellular interactions of the N-terminal semaphorin binding domain to the transmembrane and cytoplasmic regions for the PlxnAs. Data from multiple crystal structures of PlxnA1, PlxnA2 and PlxnA4 ectodomain revealed that the extracellular segment in class A plexins adopts a distinctive, relatively rigid, ring-like conformation. Negative stain EM experiments demonstrated that this ring-like state is most likely the favored conformation for PlxnA ectodomains. As the precisely arranged dimeric Sema3s and Sema6s interact head-to-head with two diverging PlxnA ectodomain frag-

ments (36, 45), I asked how does a recognition assembly, involving only the N-terminal (sema) of ten 'beads-on-a-string' domains, position the membrane proximal tenth domain (IPT6) to initiate dimerisation, and consequent activation, in the cytoplasmic region? The distinctive PlxnA ectodomain architecture immediately provides a satisfyingly simple solution to this longstanding conundrum. As detailed in Chapter 3, superposition of the crystal structure for a full length PlxnA extracellular segment onto the previously reported Sema6A$_{1\text{-}2}$–PlxnA2$_{1\text{-}4}$ complex (PDB 3OKY) brings the two membrane proximal domains into close proximity. This positioning of the membrane proximal domains from two PlxnAs appears appropriate to promote interaction between the transmembrane regions and, ultimately, dimerisation of the cytoplasmic region. Based on these observations I believe that the distinctive ring-like conformation of the PlxnA extracellular segment provides a mechanism for efficient and rather precisely defined coupling between extracellular and intercellular dimerisation states. Interestingly, negative stain EM also uncovered a less-frequent, twisted-open (chair-like) conformation for PlxnA1 ectodmain most likely generated by a 180° rotation around a hinge midway in the ring-like stalk. The fact that this conformation is the only alternative to the ring-like conformation suggests it may play a role in another aspect of PlxnA signalling, possibly the recently discovered *cis*-interactions between PlxnAs and transmembrane Sema6s (65–67). However, uncovering what specific role (if any) the chair-like conformation plays will require a significant amount of further experimentation.

Another important discovery made in this thesis is the ectodomain-based, intermolecular autoinhibitory mechanism for PlxnAs. Although various models have been hypoth-

esized for autoinhibitory plexin intra- and intermolecular interactions (17, 22, 36, 56, 143), follow-up experiments to verify these hypothesis remain elusive. Based on the structural studies I report here it now seems unlikely that PlxnA ectodomains can make autoinhibitory intramolecular interactions analogous to integrins and other TM1 receptors (5, 144). Conversely, I discovered that the autoinhibition of PlxnAs is imposed by the intermolecular interactions based on a head-to-stalk (sema–PSI2-IPT2) interface in the receptor ectodomains. Results from a combination of structural and functional characterisations for PlxnAs in this thesis favor a novel autoinhibitory mechanism previously not known for TM1 cell-surface receptors. Although speculations have been made based on the elusive pre-signalling dimer structure of the Epo receptor more than a decade ago (145), the mechanism of self-imposed inhibition based on the pre-clustering of receptors has not been revealed until this thesis. In my model, intermolecular, *cis*-interactions between the TM1 ectodomains may serve two important roles. First, together with the distinct ectodomain architectures these intermolecular interactions may maintain effective separation of the receptor transmembrane and cytoplasmic segments and thus, preclude any ligand-independent activation. Second, pre-clustering based on the *cis*-interaction in receptor ectodomains may prepare them for the acute response to bivalent or multivalent ligand-binding and potentially facilitate further clustering.

How do different classes of plexins translate bivalent semaphorin-binding into cytoplasmic activities? Is the ectodomain-based balancing of inhibition and activation seen in PlxnAs a general rule for all plexin classes? My studies on PlxnD1, which represents the last structural frontier of plexins, has provided tantalising insights. Although the

the PlxnD1 ectodomain adopts a closed-ring conformation strikingly similar to the PlxnAs in its crystal structure, via in solution experiments I have discovered potential multiple conformations for PlxnD1. Structural details of the PlxnD1 sema domain alone have also highlighted potential specificity determinants for PlxnD1 in terms of ligand-recognition. These observations suggest that PlxnD1 may be analogous to PlxnAs in some ways but also distinct in its mechanisms of action. Future experiments will focus on further structural characterization of the PlxnD1 ectodomain in particular with its bi-functional ligand Sema3E. More ambitiously, by combining structural and functional analyses we may hopefully link the unique architectures of PlxnD1 to its diverse cytoplasmic interactome.

In sum, my studies on PlxnAs and PlxnD1 uncovered novel mechanisms for the activation and regulation of plexin receptors, prompting fresh consideration of the role of autoinhibitory pre-association and ligand-induced conformational reorganisation in plexin biology. This thesis has open new doors to the reconstruction of a complete picture of semaphorin–plexin signalling. To achieve this ultimate goal, future experiments would need to proceed into two diercetions, first, the structural characterisation of full-length, transmembrane plexins in complex with their ligands and coreceptors incorporating the cell membrane, and second, analyses on the time-evolution and spatial dynamics of the plexin receptor complexes in live cells. Although these future tasks seem daunting, recent advances in membrane protein crystallography and cryoEM has provided us with unique advantages for the former, while development in multiple correlative light microscopy and spectroscopy techniques has made the later,

time-dimensional studies feasible. Indeed, it is an exciting time for the field of interdisciplinary structural studies of receptor biology, as we may expect many questions regarding the mechanistic details of many systems to be addressed in the foreseeable future.

BIBLIOGRAPHY

1. "A time- and cost-efficient system for high-level protein production in mammalian cells," vol. 62, pp. 1243–1250, Sept. 2006.

2. B. P. Ceresa and S. L. Schmid, "Regulation of signal transduction by endocytosis," *Current opinion in cell biology*, vol. 12, pp. 204–210, Apr. 2000.

3. M. A. Lemmon and J. Schlessinger, "Cell Signaling by Receptor Tyrosine Kinases," *Cell*, vol. 141, pp. 1117–1134, June 2010.

4. M. C. Pahl, O. L. Askinazi, C. Hamilton, I. Cheng, K. Cichewicz, J. Kuhn, S. Manohar, and C. D. Deppmann, *Signalling via Single-Pass Transmembrane Proteins*. Chichester, UK: John Wiley & Sons, Ltd, 2013.

5. N. F. Endres, R. Das, A. W. Smith, A. Arkhipov, E. Kovacs, Y. Huang, J. G. Pelton, Y. Shan, D. E. Shaw, D. E. Wemmer, J. T. Groves, and J. Kuriyan, "Conformational Coupling across the Plasma Membrane in Activation of the EGF Receptor," *Cell*, vol. 152, pp. 543–556, Jan. 2013.

6. P. Nurse, "Life, logic and information," *Nature*, vol. 454, pp. 424–426, July 2008.

7. T. W. Yu and C. I. Bargmann, "Dynamic regulation of axon guidance," *Nature neuroscience*, vol. 4, pp. 1169–1176, Nov. 2001.

8. M. A. Lemmon and J. Schlessinger, "Cell Signaling by Receptor Tyrosine Kinases," *Cell*, vol. 141, pp. 1117–1134, June 2010.

9. I. Moraga, J. Spangler, J. L. Mendoza, and K. C. Garcia, "Multifarious determinants of cytokine receptor signaling specificity.," *Advances in immunology*, vol. 121, pp. 1–39, 2014.

10. I. Grunwald, "Axon guidance: receptor complexes and signaling mechanisms," *Current opinion in neurobiology*, vol. 12, pp. 250–259, June 2002.

11. J. B. Spangler, I. Moraga, J. L. Mendoza, and K. C. Garcia, "Insights into cytokine-receptor interactions from cytokine engineering.," *Annual review of immunology*, vol. 33, pp. 139–167, 2015.

12. N. F. Endres, R. Das, A. W. Smith, A. Arkhipov, E. Kovacs, Y. Huang, J. G. Pelton, Y. Shan, D. E. Shaw, D. E. Wemmer, J. T. Groves, and J. Kuriyan, "Conformational Coupling across the Plasma Membrane in Activation of the EGF Receptor," *Cell*, vol. 152, pp. 543–556, Jan. 2013.

13. R. O. Hynes, "Integrins," *Cell*, vol. 110, pp. 673–687, Sept. 2002.

14. M. Sanchez-Lockhart, M. Kim, and J. Miller, "Cutting edge: A role for inside-out signaling in TCR regulation of CD28

ligand binding.," *Journal of immunology (Baltimore, Md. : 1950)*, vol. 187, pp. 5515–5519, Dec. 2011.

15. P. J. Brennan, T. Kumogai, A. Berezov, R. Murali, and M. I. Greene, "HER2/neu: mechanisms of dimerization/oligomerization.," *Oncogene*, vol. 19, no. 53, pp. 6093–6101, 2000.

16. M. Atanasova and A. Whitty, "Understanding cytokine and growth factor receptor activation mechanisms," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 47, pp. 502–530, Dec. 2012.

17. R. P. Kruger, J. Aurandt, and K.-L. Guan, "Semaphorins command cells to move.," *Nature reviews. Molecular cell biology*, vol. 6, pp. 789–800, Oct. 2005.

18. J. Schlessinger, "Cell Signaling by Receptor Tyrosine Kinases," *Cell*, vol. 103, pp. 211–225, Oct. 2000.

19. I. Dudanova and R. Klein, "The axon's balancing act: cis- and trans-interactions between Ephs and ephrins.," *Neuron*, vol. 71, pp. 1–3, July 2011.

20. L. C. Cantley, T. Hunter, R. Sever, and J. Thorner, "Signal Transduction," May 2014.

21. G. J. Bashaw and R. Klein, "Signaling from axon guidance receptors.," *Cold Spring Harbor perspectives in biology*, vol. 2, p. a001941, May 2010.

22. T. Takahashi and S. M. Strittmatter, "PlexinA1 Autoinhibition by the Plexin Sema Domain," *Neuron*, vol. 29, pp. 429–439, Feb. 2001.

23. N. F. Endres, K. Engel, R. Das, E. Kovacs, and J. Kuriyan, "Regulation of the catalytic activity of the EGF receptor.," *Current Opinion in Structural Biology*, vol. 21, pp. 777–784, Dec. 2011.

24. C.-H. Heldin and M. Purton, *Signal Transduction*. Garland Science, 1998.

25. E. Seiradake, A. Schaupp, D. del Toro Ruiz, R. Kaufmann, N. Mitakidis, K. Harlos, A. R. Aricescu, R. Klein, and E. Y. Jones, "Structurally encoded intraclass differences in EphA clusters drive distinct cell responses," *Nature Structural & Molecular Biology*, vol. 20, pp. 958–964, Aug. 2013.

26. E. G. Healey, B. Bishop, J. Elegheert, C. H. Bell, S. Padilla-Parra, and C. Siebold, "Repulsive guidance molecule is a structural bridge between neogenin and bone morphogenetic protein," *Nature Structural & Molecular Biology*, vol. 22, pp. 458–465, May 2015.

27. C. H. Coles, N. Mitakidis, P. Zhang, J. Elegheert, W. Lu, A. W. Stoker, T. Nakagawa, A. M. Craig, E. Y. Jones, and A. R. Aricescu, "Structural basis for extracellular cis and trans RPTPσ signal competition in synaptogenesis.," *Nature communications*, vol. 5, p. 5209, 2014.

28. M. J. Katz and R. J. Lasek, "Invited review: Guidance cue patterns and cell migration in muiticeliuiar organisms," *Cell Motility*, vol. 1, pp. 141–157, Jan. 1980.

29. F. Charron and M. Tessier-Lavigne, "Novel brain wiring functions for classical morphogens: a role as graded positional cues in axon guidance.," *Development*, vol. 132, pp. 2251–2262, May 2005.

30. A. B. Huber, A. L. Kolodkin, D. D. Ginty, and J.-F. Cloutier, "Signaling at the growth cone: ligand-receptor complexes and the control of axon growth and guidance.," *Annual review of neuroscience*, vol. 26, pp. 509–563, 2003.

31. L. Finci, Y. Zhang, R. Meijers, and J.-H. Wang, "Signaling mechanism of the

netrin-1 receptor DCC in axon guidance.," *Progress in biophysics and molecular biology*, vol. 118, pp. 153–160, Sept. 2015.

32. R. Klein, "Eph/ephrin signalling during development," *Development*, vol. 139, pp. 4105–4109, Nov. 2012.

33. K. Kullander and R. Klein, "Mechanisms and functions of eph and ephrin signalling," *Nature reviews. Molecular cell biology*, vol. 3, pp. 475–486, July 2002.

34. E. Seiradake, K. Harlos, G. Sutton, A. R. Aricescu, and E. Y. Jones, "An extracellular steric seeding mechanism for Eph-ephrin signaling platform assembly," *Nature Structural & Molecular Biology*, vol. 17, pp. 398–402, Apr. 2010.

35. K. Xu, Z. Wu, N. Renier, A. Antipenko, D. Tzvetkova-Robev, Y. Xu, M. Minchenko, V. Nardi-Dei, K. R. Rajashankar, J. Himanen, M. Tessier-Lavigne, and D. B. Nikolov, "Neural migration. Structures of netrin-1 bound to two receptors provide insight into its axon guidance mechanism.," *Science (New York, N.Y.)*, vol. 344, pp. 1275–1279, June 2014.

36. B. J. C. Janssen, R. A. Robinson, F. Pérez-Brangulí, C. H. Bell, K. J. Mitchell, C. Siebold, and E. Y. Jones, "Structural basis of semaphorin-plexin signalling.," *Nature*, vol. 467, pp. 1118–1122, Oct. 2010.

37. C. Siebold and E. Y. Jones, "Structural insights into semaphorins and their receptors.," *Seminars in cell & developmental biology*, vol. 24, pp. 139–145, Mar. 2013.

38. E. Hohenester, "Structural insight into Slit-Robo signalling.," *Biochemical Society transactions*, vol. 36, pp. 251–256, Apr. 2008.

39. "Getting neural circuits into shape with semaphorins.," vol. 13, pp. 605–618, Sept. 2012.

40. B. J. Dickson, "Molecular mechanisms of axon guidance.," *Science (New York, N.Y.)*, vol. 298, pp. 1959–1964, Dec. 2002.

41. B. C. Jongbloets and R. J. Pasterkamp, "Semaphorin signalling during development.," *Development*, vol. 141, pp. 3292–3297, Sept. 2014.

42. A. Yaron and D. Sprinzak, "The cis side of juxtacrine signaling: a new role in the development of the nervous system.," *Trends in neurosciences*, vol. 35, pp. 230–239, Apr. 2012.

43. R. Klein, "Bidirectional modulation of synaptic functions by Eph/ephrin signaling," *Nature neuroscience*, vol. 12, pp. 15–20, Jan. 2009.

44. L. I. Finci, N. Krüger, X. Sun, J. Zhang, M. Chegkazi, Y. Wu, G. Schenk, H. D. T. Mertens, D. I. Svergun, Y. Zhang, J.-H. Wang, and R. Meijers, "The Crystal Structure of Netrin-1 in Complex with DCC Reveals the Bifunctionality of Netrin-1 As a Guidance Cue," *Neuron*, vol. 83, pp. 839–849, Aug. 2014.

45. B. J. C. Janssen, T. Malinauskas, G. A. Weir, M. Z. Cader, C. Siebold, and E. Y. Jones, "Neuropilins lock secreted semaphorins onto plexins in a ternary signaling complex," *Nature Structural & Molecular Biology*, vol. 19, pp. 1293–1299, Dec. 2012.

46. Y. Duan, S.-H. Wang, J. Song, Y. Mironova, G.-l. Ming, A. L. Kolodkin, R. J. Giger, and F. Miller, "Semaphorin 5A inhibits synaptogenesis in early postnatal- and adult-born hippocampal dentate granule cells," *eLife*, vol. 3, p. e04390, Nov. 2014.

47. E. Seiradake, K. Harlos, G. Sutton, A. R. Aricescu, and E. Y. Jones, "An extracellular steric seeding mechanism for Eph-ephrin signaling platform assembly," *Nature Structural & Molecular Biology*, vol. 17, pp. 398–402, Apr. 2010.

48. E. Seiradake, A. Schaupp, D. del Toro Ruiz, R. Kaufmann, N. Mitakidis, K. Harlos, A. R. Aricescu, R. Klein, and E. Y. Jones, "Structurally encoded intraclass differences in EphA clusters drive distinct cell responses," *Nature Structural & Molecular Biology*, vol. 20, pp. 958–964, Aug. 2013.

49. H. Liu, Z. S. Juo, A. H.-R. Shim, P. J. Focia, X. Chen, K. C. Garcia, and X. He, "Structural basis of semaphorin-plexin recognition and viral mimicry from Sema7A and A39R complexes with PlexinC1.," *Cell*, vol. 142, pp. 749–761, Sept. 2010.

50. Y. Wang, H. He, N. Srivastava, S. Vikarunnessa, Y.-b. Chen, J. Jiang, C. W. Cowan, and X. Zhang, "Plexins are GTPase-activating proteins for Rap and are activated by induced dimerization.," *Science signaling*, vol. 5, pp. ra6–ra6, Jan. 2012.

51. Y. Wang, H. G. Pascoe, C. A. Brautigam, H. He, and X. Zhang, "Structural basis for activation and non-canonical catalysis of the Rap GTPase activating protein domain of plexin.," *eLife*, vol. 2, p. e01279, 2013.

52. I. Grunwald, "Axon guidance: receptor complexes and signaling mechanisms," *Current opinion in neurobiology*, vol. 12, pp. 250–259, June 2002.

53. R. Barton, A. Driscoll, S. Flores, D. Mudbhari, T. Collins, M. K. Iovine, and B. W. Berger, "Cysteines in the neuropilin-2 MAM domain modulate receptor homooligomerization and signal transduction.," *Biopolymers*, Feb. 2015.

54. G. Neufeld, T. Cohen, N. Shraga, T. Lange, O. Kessler, and Y. Herzog, "The neuropilins: multifunctional semaphorin and VEGF receptors that modulate axon guidance and angiogenesis.," *Trends in cardiovascular medicine*, vol. 12, pp. 13–19, Jan. 2002.

55. W.-J. Oh and C. Gu, "The role and mechanism-of-action of Sema3E and Plexin-D1 in vascular and neural development.," *Seminars in cell & developmental biology*, vol. 24, pp. 156–162, Mar. 2013.

56. T. Nogi, N. Yasui, E. Mihara, Y. Matsunaga, M. Noda, N. Yamashita, T. Toyofuku, S. Uchiyama, Y. Goshima, A. Kumanogoh, and J. Takagi, "Structural basis for semaphorin signalling through the plexin receptor.," *Nature*, vol. 467, pp. 1123–1127, Oct. 2010.

57. E. Gherardi, C. A. Love, R. M. Esnouf, and E. Y. Jones, "The sema domain," *Current Opinion in Structural Biology*, vol. 14, pp. 669–678, Dec. 2004.

58. H. G. Pascoe, Y. Wang, and X. Zhang, "Structural mechanisms of plexin signaling.," *Progress in biophysics and molecular biology*, Mar. 2015.

59. C. H. Bell, A. R. Aricescu, E. Y. Jones, and C. Siebold, "A dual binding mode for RhoGTPases in plexin signalling.," *PLoS biology*, vol. 9, p. e1001134, Aug. 2011.

60. H. He, T. Yang, J. R. Terman, and X. Zhang, "Crystal structure of the plexin A3 intracellular region reveals an autoinhibited conformation through active site sequestration.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 15610–15615, Sept. 2009.

61. S. Kang and A. Kumanogoh, "Semaphorins in bone development, homeostasis, and disease," *Seminars in cell & developmental biology*, vol. 24, pp. 163–171, Mar. 2013.

62. T. Worzfeld and S. Offermanns, "Semaphorins and plexins as therapeutic targets.," *Nature reviews. Drug discovery*, vol. 13, pp. 603–621, Aug. 2014.

63. E. Y. Jones, "Understanding cell signalling systems: paving the way for new therapies.," *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, vol. 373, pp. 20130155–20130155, Mar. 2015.

64. J. Luchino, M. Hocine, M.-C. Amoureux, B. Gibert, A. Bernet, A. Royet, I. Treilleux, P. Lécine, J.-P. Borg, P. Mehlen, S. Chauvet, and F. Mann, "Semaphorin 3E suppresses tumor cell death triggered by the plexin D1 dependence receptor in metastatic breast cancers.," *Cancer cell*, vol. 24, pp. 673–685, Nov. 2013.

65. I. Andermatt, N. H. Wilson, T. Bergmann, O. Mauti, M. Gesemann, S. Sockanathan, and E. T. Stoeckli, "Semaphorin 6B acts as a receptor in post-crossing commissural axon guidance.," *Development*, p. dev.112185, Sept. 2014.

66. L. O. Sun, Z. Jiang, M. Rivlin-Etzion, R. Hand, C. M. Brady, R. L. Matsuoka, K.-W. Yau, M. B. Feller, and A. L. Kolodkin, "On and off retinal circuit assembly by divergent molecular mechanisms.," *Science (New York, N.Y.)*, vol. 342, pp. 1241974–1241974, Nov. 2013.

67. L. Haklai-Topper, G. Mlechkovich, D. Savariego, I. Gokhman, and A. Yaron, "Cis interaction between Semaphorin6A and Plexin-A4 modulates the repulsive response to Sema6A.," *The EMBO journal*, vol. 29, pp. 2635–2645, Aug. 2010.

68. S. Chauvet, S. Cohen, Y. Yoshida, L. Fekrane, J. Livet, O. Gayet, L. Segu, M.-C. Buhot, T. M. Jessell, C. E. Henderson, and F. Mann, "Gating of Sema3E/PlexinD1 signaling by neuropilin-1 switches axonal repulsion to attraction during brain development.," *Neuron*, vol. 56, pp. 807–822, Dec. 2007.

69. A. Bellon, J. Luchino, K. Haigh, G. Rougon, J. Haigh, S. Chauvet, and F. Mann, "VEGFR2 (KDR/Flk1) Signaling Mediates Axon Growth in Response to Semaphorin 3E in the Developing Brain," *Neuron*, vol. 66, pp. 205–219, Apr. 2010.

70. T. Toyofuku, H. Zhang, A. Kumanogoh, N. Takegahara, M. Yabuki, K. Harada, M. Hori, and H. Kikutani, "Guidance of myocardial patterning in cardiac development by Sema6D reverse signalling.," *Nature cell biology*, vol. 6, pp. 1204–1211, Dec. 2004.

71. T. Toyofuku, J. Yoshida, T. Sugimoto, M. Yamamoto, N. Makino, H. Takamatsu, N. Takegahara, F. Suto, M. Hori, H. Fujisawa, A. Kumanogoh, and H. Kikutani, "Repulsive and attractive semaphorins cooperate to direct the navigation of cardiac neural crest cells," *Developmental biology*, vol. 321, pp. 251–262, Sept. 2008.

72. C. Gu, Y. Yoshida, J. Livet, D. V. Reimert, F. Mann, J. Merte, C. E. Henderson, T. M. Jessell, A. L. Kolodkin, and D. D. Ginty, "Semaphorin 3E and plexin-D1 control vascular pattern independently of neuropilins.," *Science (New York, N.Y.)*, vol. 307, pp. 265–268, Jan. 2005.

73. J. Kim, W.-J. Oh, N. Gaiano, Y. Yoshida, and C. Gu, "Semaphorin 3E-Plexin-D1 signaling regulates VEGF function in developmental angiogenesis via a feedback mechanism.," *Genes & development*, vol. 25, pp. 1399–1411, July 2011.

74. T. Toyofuku, M. Yabuki, J. Kamei, M. Kamei, N. Makino, A. Kumanogoh, and M. Hori, "Semaphorin-4A, an activator for T-cell-mediated immunity, suppresses angiogenesis via Plexin-D1," *The EMBO journal*, vol. 26, pp. 1373–1384, Mar. 2007.

75. H.-F. Guo, X. Li, M. W. Parker, J. Waltenberger, P. M. Becker, and C. W. Vander Kooi, "Mechanistic basis for the potent anti-angiogenic activity of semapho-

rin 3F.," *Biochemistry*, vol. 52, pp. 7551–7558, Oct. 2013.

76. L. Tamagnone and M. Rehman, "To die or not to die: Sema3E rules the game.," *Cancer cell*, vol. 24, pp. 564–566, Nov. 2013.

77. A. Casazza, V. Finisguerra, L. Capparuccia, A. Camperi, J. M. Swiercz, S. Rizzolio, C. Rolny, C. Christensen, A. Bertotti, I. Sarotto, M. Risio, L. Trusolino, J. Weitz, M. Schneider, M. Mazzone, M. Mazzone, P. M. Comoglio, and L. Tamagnone, "Sema3E-Plexin D1 signaling drives human cancer cell invasiveness and metastatic spreading in mice.," *The Journal of clinical investigation*, vol. 120, pp. 2684–2698, Aug. 2010.

78. C. H. Coles, "Clustering and σ Proteoglycan-Specific Molecular Switch for RPTP," *Nat Struct Biol*, 1996.

79. T. Schwede, J. Kopp, N. Guex, and M. C. Peitsch, "SWISS-MODEL: An automated protein homology-modeling server.," *Nucleic acids research*, vol. 31, pp. 3381–3385, July 2003.

80. J. Söding, A. Biegert, and A. N. Lupas, "The HHpred interactive server for protein homology detection and structure prediction.," *Nucleic acids research*, vol. 33, pp. W244–8, July 2005.

81. Z. R. Yang, R. Thomson, P. McNeil, and R. M. Esnouf, "RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins.," *Bioinformatics (Oxford, England)*, vol. 21, pp. 3369–3376, Aug. 2005.

82. G. Rhodes, *Crystallography Made Crystal Clear*. A Guide for Users of Macromolecular Models, Academic Press, Dec. 2012.

83. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, "Phaser crystallographic software.," *Journal of applied crystallography*, vol. 40, pp. 658–674, Aug. 2007.

84. T. S. Walter, 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, K. Harlos, and IUCr, "A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization," *Acta crystallographica. Section D, Biological crystallography*, vol. 61, pp. 651–657, June 2005.

85. T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, and A. G. W. Leslie, "iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.," *Acta crystallographica. Section D, Biological crystallography*, vol. 67, pp. 271–281, Apr. 2011.

86. P. R. Evans, "An introduction to data reduction: space-group determination, scaling and intensity statistics," *Acta crystallographica. Section D, Biological crystallography*, vol. 67, pp. 282–292, Mar. 2011.

87. G. Winter, C. M. C. Lobley, and S. M. Prince, "Decision making in xia2.," *Acta crystallographica. Section D, Biological crystallography*, vol. 69, pp. 1260–1273, July 2013.

88. P. Emsley and K. Cowtan, "Coot: model-building tools for molecular graphics," *Acta crystallographica. Section D, Biological crystallography*, vol. 60, pp. 2126–2132, Dec. 2004.

89. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W.

Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, and P. H. Zwart, "PHENIX: a comprehensive Python-based system for macromolecular structure solution.," *Acta crystallographica. Section D, Biological crystallography*, vol. 66, pp. 213–221, Feb. 2010.

90. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, and N. A. Baker, "PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations.," *Nucleic acids research*, vol. 32, pp. W665–7, July 2004.

91. S. Unni, Y. Huang, R. M. Hanson, M. Tobias, S. Krishnan, W. W. Li, J. E. Nielsen, and N. A. Baker, "Web servers and services for electrostatics calculations with APBS and PDB2PQR," *Journal of Computational Chemistry*, vol. 32, pp. 1488–1491, May 2011.

92. F. Glaser, T. Pupko, I. Paz, R. E. Bell, D. Bechor-Shental, E. Martz, and N. Ben-Tal, "ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information," *Bioinformatics (Oxford, England)*, vol. 19, pp. 163–164, Jan. 2003.

93. E. Krissinel and K. Henrick, "Inference of Macromolecular Assemblies from Crystalline State," *Journal of molecular biology*, vol. 372, pp. 774–797, Sept. 2007.

94. D. S. Booth, A. Avila-Sakar, and Y. Cheng, "Visualizing Proteins and Macromolecular Complexes by Negative Stain EM: from Grid Preparation to Image Acquisition," *Journal of Visualized Experiments*, no. 58, pp. e3227–e3227, 2011.

95. M. Ohi, Y. Li, Y. Cheng, and T. Walz, "Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy.," *Biological procedures online*, vol. 6, pp. 23–34, 2004.

96. G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke, "EMAN2: An extensible image processing suite for electron microscopy," *Journal of Structural Biology*, vol. 157, pp. 38–46, Jan. 2007.

97. M. van Heel and W. Keegstra, "IMAGIC: A fast, flexible and friendly image analysis software system," *Ultramicroscopy*, vol. 7, pp. 113–129, Jan. 1981.

98. J. Frank, M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith, "SPIDER and WEB: Processing and Visualization of Images in 3D Electron Microscopy and Related Fields," *Journal of Structural Biology*, vol. 116, pp. 190–199, Jan. 1996.

99. J. F. Flanagan IV, O. Namy, I. Brierley, and R. J. C. Gilbert, "Direct Observation of Distinct A/P Hybrid-State tRNAs in Translocating Ribosomes," *Structure (London, England : 1993)*, vol. 18, pp. 257–264, Feb. 2010.

100. J. L. Cole, J. W. Lary, T. P Moody, and T. M. Laue, "Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium.," *Methods in cell biology*, vol. 84, pp. 143–179, 2008.

101. P. Schuck, "Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling," *Biophysical Journal*, vol. 78, pp. 1606–1619, Mar. 2000.

102. E. Brookes, W. Cao, and B. Demeler, "A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape," *European Biophysics Journal*, vol. 39, pp. 405–414, Feb. 2009.

103. A. van der Merwe, "Surface Plasmon Resonance," pp. 1–50, Feb. 2003.

104. I. Langmuir, "THE ADSORPTION OF GASES ON PLANE SURFACES OF GLASS,

MICA AND PLATINUM.," *Journal of the American Chemical Society*, vol. 40, pp. 1361–1403, Sept. 1918.

105. A. Leray, S. Padilla-Parra, J. Roul, L. Héliot, and M. Tramier, "Spatio-Temporal Quantification of FRET in living cells by fast time-domain FLIM: a comparative study of non-fitting methods [corrected].," *PloS one*, vol. 8, no. 7, p. e69335, 2013.

106. S. Padilla-Parra, N. Audugé, M. Coppey-Moisan, and M. Tramier, "Quantitative FRET Analysis by Fast Acquisition Time Domain FLIM at High Spatial Resolution in Living Cells," *Biophysical Journal*, vol. 95, pp. 2976–2988, Sept. 2008.

107. S. Padilla-Parra, N. Audugé, H. Lalucque, J.-C. Mevel, M. Coppey-Moisan, and M. Tramier, "Quantitative Comparison of Different Fluorescent Protein Couples for Fast FRET-FLIM Acquisition," *Biophysical Journal*, vol. 97, pp. 2368–2376, Oct. 2009.

108. D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, pp. 431–441, June 1963.

109. Y. Zhao, J. Ren, S. Padilla-Parra, E. E. Fry, and D. I. Stuart, "Lysosome sorting of β-glucocerebrosidase by LIMP-2 is targeted by the mannose 6-phosphate receptor.," *Nature communications*, vol. 5, p. 4321, 2014.

110. P. M. Dixon, *Ripley's K Function*. Chichester, UK: John Wiley & Sons, Ltd, 2002.

111. P. Lemmer, M. Gunkel, D. Baddeley, R. Kaufmann, A. Urich, Y. Weiland, J. Reymann, P. Müller, M. Hausmann, and C. Cremer, "SPDM: light microscopy with single-molecule resolution at the nanoscale," *Applied Physics B*, vol. 93, no. 1, pp. 1–12, 2008.

112. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution.," *Science (New York, N.Y.)*, vol. 313, pp. 1642–1645, Sept. 2006.

113. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature methods*, vol. 3, pp. 793–796, Oct. 2006.

114. F. Grüll, M. Kirchgessner, R. Kaufmann, M. Hausmann, and U. Kebschull, *Accelerating Image Analysis for Localization Microscopy with FPGAs*. IEEE, 2011.

115. C. Cremer, R. Kaufmann, M. Gunkel, S. Pres, Y. Weiland, P. Müller, T. Ruckelshausen, P. Lemmer, F. Geiger, S. Degenhard, C. Wege, N. A. W. Lemmermann, R. Holtappels, H. Strickfaden, and M. Hausmann, "Superresolution imaging of biological nanostructures by spectral precision distance microscopy," *Biotechnology Journal*, vol. 6, pp. 1037–1051, Sept. 2011.

116. C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis.," *Nature methods*, vol. 9, pp. 671–675, July 2012.

117. E. Y. Van Battum, R.-A. F. Gunput, S. Lemstra, E. J. N. Groen, K. L. Yu, Y. Adolfs, Y. Zhou, C. C. Hoogenraad, Y. Yoshida, M. Schachner, A. Akhmanova, and R. J. Pasterkamp, "The intracellular redox protein MICAL-1 regulates the development of hippocampal mossy fibre connections.," *Nature communications*, vol. 5, p. 4317, 2014.

118. T. Weide, J. Teuber, M. Bayer, and A. Barnekow, "MICAL-1 isoforms, novel rab1 interacting proteins," *Biochemical and Biophysical Research Communications*, vol. 306, pp. 79–86, June 2003.

119. L. Tamagnone, S. Artigiani, H. Chen, Z. He, G. I. Ming, H. Song, A. Chédotal, M. L. Winberg, C. S. Goodman, M. Poo, M. Tessier-Lavigne, and P. M. Comoglio, "Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates.," *Cell*, vol. 99, pp. 71–80, Oct. 1999.

120. L. Tamagnone and P. M. Comoglio, "Signalling by semaphorin receptors: cell guidance and beyond," *Trends in cell biology*, vol. 10, pp. 377–383, Sept. 2000.

121. T. S. Walter, C. Meier, R. Assenberg, K.-F. Au, J. Ren, A. Verma, J. E. Nettleship, R. J. Owens, D. I. Stuart, and J. M. Grimes, "Lysine Methylation as a Routine Rescue Strategy for Protein Crystallization," *Structure (London, England : 1993)*, vol. 14, pp. 1617–1622, Nov. 2006.

122. A. Antipenko, J.-P. Himanen, K. van Leyen, V. Nardi-Dei, J. Lesniak, W. A. Barton, K. R. Rajashankar, M. Lu, C. Hoemme, and A. W. Püschel, "Structure of the Semaphorin-3A Receptor Binding Module," *Neuron*, vol. 39, pp. 589–598, Aug. 2003.

123. C. H. Coles, N. Mitakidis, P. Zhang, J. Elegheert, W. Lu, A. W. Stoker, T. Nakagawa, A. M. Craig, E. Y. Jones, and A. R. Aricescu, "Structural basis for extracellular cis and trans RPTPσ signal competition in synaptogenesis," *Nature communications*, vol. 5, p. 5209, Nov. 2014.

124. E. Seiradake, A. Schaupp, D. del Toro Ruiz, R. Kaufmann, N. Mitakidis, K. Harlos, A. R. Aricescu, R. Klein, and E. Y. Jones, "Structurally encoded intraclass differences in EphA clusters drive distinct cell responses," *Nature Structural & Molecular Biology*, vol. 20, pp. 958–964, Aug. 2013.

125. Y. Wang, H. G. Pascoe, C. A. Brautigam, H. He, and X. Zhang, "Structural basis for activation and non-canonical catalysis of the Rap GTPase activating protein domain of plexin," *eLife*, vol. 2, pp. e01279–e01279, Jan. 2013.

126. T. Takahashi and S. M. Strittmatter, "Plexina1 autoinhibition by the plexin sema domain.," *Neuron*, vol. 29, pp. 429–439, Feb. 2001.

127. P. H. Brown, A. Balbo, and P. Schuck, *Characterizing Protein-Protein Interactions by Sedimentation Velocity Analytical Ultracentrifugation*. Hoboken, NJ, USA: John Wiley & Sons, Inc., May 2001.

128. Y. Wu, J. Vendome, L. Shapiro, A. Ben-Shaul, and B. Honig, "Transforming binding affinities from three dimensions to two with application to cadherin clustering," *Nature*, vol. 475, pp. 510–513, Apr. 2012.

129. L. J. Turner and A. Hall, "Plexin-induced collapse assay in COS cells.," *Methods in enzymology*, vol. 406, pp. 665–676, 2006.

130. Y. Murakami, F. Suto, M. Shimizu, T. Shinoda, T. Kameyama, and H. Fujisawa, "Differential expression of plexin-A subfamily members in the mouse nervous system.," *Developmental dynamics : an official publication of the American Association of Anatomists*, vol. 220, pp. 246–258, Mar. 2001.

131. W.-J. Oh and C. Gu, "The role and mechanism-of-action of Sema3E and Plexin-D1 in vascular and neural development," *Seminars in cell & developmental biology*, vol. 24, pp. 156–162, Mar. 2013.

132. Y. Zhou, R.-A. F. Gunput, and R. J. Pasterkamp, "Semaphorin signaling: progress made and promises ahead.," *Trends in biochemical sciences*, vol. 33, pp. 161–170, Apr. 2008.

133. J. A. Epstein, H. Aghajanian, and M. K. Singh, "Semaphorin Signaling in Cardiovascular Development," *Cell metabolism*, vol. 21, pp. 163–173, Feb. 2015.

134. M. Klagsbrun and A. Shimizu, "Semaphorin 3E, an exception to the rule.," *The Journal of clinical investigation*, vol. 120, pp. 2658–2660, Aug. 2010.

135. R. Maiti, G. H. Van Domselaar, H. Zhang, and D. S. Wishart, "SuperPose: a simple server for sophisticated structural superposition.," *Nucleic acids research*, vol. 32, pp. W590–4, July 2004.

136. H. H. Niemann, V. Jäger, P. J. G. Butler, J. van den Heuvel, S. Schmidt, D. Ferraris, E. Gherardi, and D. W. Heinz, "Structure of the Human Receptor Tyrosine Kinase Met in Complex with the Listeria Invasion Protein InlB," *Cell*, vol. 130, pp. 235–246, July 2007.

137. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, and D. G. Higgins, "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega," *Molecular Systems Biology*, vol. 7, pp. 539–539, Jan. 2011.

138. P. Gouet, E. Courcelle, D. I. Stuart, and F. Métoz, "ESPript: analysis of multiple sequence alignments in PostScript.," *Bioinformatics (Oxford, England)*, vol. 15, pp. 305–308, Apr. 1999.

139. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson, "Overview of the CCP4 suite and current developments.," *Acta crystallographica. Section D, Biological crystallography*, vol. 67, pp. 235–242, Apr. 2011.

140. T. Pape, T. R. Schneider, and IUCr, "HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs," *Journal of Applied Crystallography*, vol. 37, pp. 843–844, Oct. 2004.

141. T. C. Terwilliger, P. D. Adams, R. J. Read, A. J. McCoy, N. W. Moriarty, R. W. Grosse-Kunstleve, P. V. Afonine, P. H. Zwart, L. W. Hung, and IUCr, "Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard," *Acta crystallographica. Section D, Biological crystallography*, vol. 65, pp. 582–601, June 2009.

142. J. J. Headd, N. Echols, P. V. Afonine, R. W. Grosse-Kunstleve, V. B. Chen, N. W. Moriarty, D. C. Richardson, J. S. Richardson, P. D. Adams, and IUCr, "Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution," *Acta crystallographica. Section D, Biological crystallography*, vol. 68, pp. 381–390, Apr. 2012.

143. A. Antipenko, J.-P. Himanen, K. van Leyen, V. Nardi-Dei, J. Lesniak, W. A. Barton, K. R. Rajashankar, M. Lu, C. Hoemme, A. W. Püschel, and D. B. Nikolov, "Structure of the Semaphorin-3A Receptor Binding Module," *Neuron*, vol. 39, pp. 589–598, Aug. 2003.

144. Y. I. Choi, J. S. Duke-Cohan, W. Chen, B. Liu, J. Rossy, T. Tabarin, L. Ju, J. Gui, K. Gaus, C. Zhu, and E. L. Reinherz, "Dynamic control of β1 integrin adhesion by the plexinD1-sema3E axis.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, pp. 379–384, Jan. 2014.

145. O. Livnah, E. A. Stura, S. A. Middleton, D. L. Johnson, L. K. Jolliffe, and I. A. Wilson, "Crystallographic Evidence for Preformed Dimers of Erythropoietin Receptor Before Ligand Activation," *Science*

(New York, N.Y.), vol. 283, pp. 987–990, Feb. 1999.

146. A. Fiser and A. Šali, "Modeller: Generation and Refinement of Homology-Based Protein Structure Models," in *Macromolecular Crystallography, Part D*, pp. 461–491, Elsevier, 2003.

APPENDIX

A.1 Production of the PlxnA Ectodomain Crystals

Mouse PlxnA1$_{1\text{-}10}$ was concentrated to 8.2 mg/ml in 20 mM HEPES, pH 7.5 and 150 mM sodium chloride and subsequently crystallized in three different crystal forms. One crystal form, with space group $P4_32_12$ grown in 6% (w/v) PEG 4k and 5 mM tricine, pH 8.5 (or 5 mM TRIS, pH 8.5), diffracted to 6.0 Å resolution. Dehydration of these crystals in the plate by increasing the PEG 4k concentration in the mother liquor in three steps to 25% over a time span of 48 hours produced a second, diffracting to 4.0 Å resolution, crystal form with a 53 Å shorter c-axis. A third crystal form with space group $P2_1$ diffracting to 6.0 Å resolution, was grown in 15% (v/v) propanol, 20 mM magnesium chloride and 50 mM 2-(N-morpholino)ethanesulfonic acid (MES), pH 6.0. For PlxnA2$_{1\text{-}10}$ we ob-

tained one crystal form. PlxnA2$_{1\text{-}10}$ was concentrated to 6.6 mg/ml in 20 mM HEPES, pH 7.5 and 150 mM sodium chloride and subsequently deglycosylated with endoglycosidase F1. Crystals of PlxnA2$_{1\text{-}10}$ with space group P3221 were produced in 0.68 M potassium/sodium tartrate and 85 mM HEPES, pH 7.5 and diffracted to 10 Å resolution. PlxnA4$_{1\text{-}10}$ was crystallized in four different crystal forms. PlxnA4$_{1\text{-}10}$ concentrated to 3.8 mg/ml in 20 mM HEPES, pH 7.5, 150 mM sodium chloride, 100 mM dimethylbenzylammonium propane sulfonate, NDSB (non-detergent sulfobetaine)-256, crystallized in space group P4$_1$ in 29% (v/v) glycerol, 4.3% (w/v) PEG 8k and 53 mM TRIS hydrochloride, pH 8.5 and diffracted to 7.5 Å resolution. PlxnA4$_{1\text{-}10}$ concentrated to 4.2 mg/ml in 10 mM citrate, pH 5.1 yielded a second and third crystal form. Crystals with space group P2$_1$ were obtained in 0.5 M sodium succinate, 100 mM 2-bisamino]-2–1,3-propanediol (BIS)-trisaminomethane (TRIS) propane, pH 7.0 and diffracted to 8.0 Å resolution and crystals with space group P4$_3$2$_1$2 were grown in 0.8 M ammonium sulfate, 0.4 M Dimethyl (2-hydroxyethyl) ammonium propane sulfonate, non-detergent sulfobetaines (NDSB)-211, 100 mM TRIS, pH 8.0 and diffracted to 8 Å resolution. PlxnA4$_{1\text{-}10}$ concentrated to 5.8 mg/ml in 20 mM HEPES, pH 7.5, 150 mM sodium chloride and 100 mM NDSB-256 crystallized in a fourth space group, P4$_3$2$_1$2, in 1.5 lithium sulfate, 100 mM TRIS, pH 8.5 and diffracted to 8.5 Å resolution.

Before diffraction data collection crystals were soaked in mother liquor supplemented with 12.5% (v/v) 2,3-butanediol for mouse PlxnA1$_{1\text{-}10}$ crystals, 25% (v/v) glycerol for PlxnA2$_{1\text{-}10}$ and PlxnA4$_{1\text{-}10}$ crystals. In particular, mouse PlxnA4$_{1\text{-}10}$ (space group P4$_1$) and PlxnD1$_{1\text{-}10}$ crystals nothing extra was added as precipitants in the crystallization

solution (29% v/v glycerol) also served as cryo-protectant.

A.2 Structure Solution and Refinement of PlxnA Ectodomains

The structure of PlxnA1$_{1\text{-}10}$ at 4.0 Å resolution was initially solved using the structure of PlxnA2$_{1\text{-}4}$ (PDB 3OKT) (54% sequence identity with PlxnA1$_{1\text{-}4}$) as MR intitial model. After solvent flattening in PARROT (88) this solution revealed clear electron density for PlxnA1 domains IPT2, PSI3, IPT3, IPT4 and IPT5 that were not included in the search model. To further complete the structure of PlxnA1$_{1\text{-}10}$ we generated homology models of these domains with MODELLER (146) using the following template PDB–PlxnA1 domain combinations; 3OKY (36) for IPT2, 3NVN (49) for PSI3, 2UZY (136) for IPT3 and IPT4, and 2CXK (unpublished results) for IPT5, with 16%, 29%, 29%, 29% and 20% sequence identity, respectively. These models were placed by MR in PHASER and the resulting structure was further optimized by manual rebuilding in COOT and refinement in REFMAC using jelly-body restraints (88, 139). High resolution structures of PlxnA2$_{4\text{-}5}$ and PlxnA1$_{7\text{-}10}$ were then placed into the PlxnA1$_{1\text{-}10}$ using PHASER and subsequently refined. We then used the refined structures of domains, IPT3, IPT4, IPT5 and segments of domain IPT6 to replace the original ones in the the 4.0 Å PlxnA1$_{1\text{-}10}$ structure which was then completed by model building in COOT and refinement in REFMAC using jelly-body restraints.

All other PlxnA full extracellular segment structures were solved using the refined 4.0 Å resolution PlxnA1$_{1\text{-}10}$ structure and, for PlxnA2$_{1\text{-}10}$, also the PlxnA2$_{1\text{-}4}$ structure (PDB 3OKT). Homology models for PlxnA2 and PlxnA4 were calculated with MODELLER using the PlxnA1$_{1\text{-}10}$ structure as template for PlxnA2 domains 5 to 10 (IPT2, PSI3, IPT3, IPT4, IPT5 and IPT6) (53% sequence identity) and for all ten PlxnA4 domains (54% sequence identity). The other PlxnA1$_{1\text{-}10}$, PlxnA2$_{1\text{-}10}$ and PlxnA4$_{1\text{-}10}$ crystal forms were solved by MR in PHASER using individual domains and homology models of the PlxnA1$_{1\text{-}10}$, PlxnA2$_{1\text{-}4}$, and PlxnA2$_{5\text{-}10}$ structures search models. Domains were omitted from the models in cases where the electron density was too weak for interpretation due to flexibility. All solutions were refined by rigid-body refinement in PHENIX with each domain as a rigid group and a single B-factor per domain, thus preventing any overfitting.

	PlxnA1$_{1\text{-}10}$	PlxnA1$_{1\text{-}10}$	PlxnA1$_{1\text{-}10}$	PlxnA2$_{1\text{-}10}$	PlxnA4$_{1\text{-}10}$
Data collection					
Space group	$P4_32_12$	$P2_1$	$P4_32_12$	$P3_221$	$P4_1$
Cell dimensions					
a, *b*, *c* (Å)	195, 195, 176	140, 196, 145	198, 198, 229	238, 238, 642	221, 221, 66
α, β, γ (°)	90, 90, 90	90, 94.6, 90	90, 90, 90	90, 90, 120	90, 90, 90
Resolution (Å)	87-4.0 (4.24-4.00)[a]	58-6.0 (6.57-6.00)[a]	112-6.0 (6.71-6.00)[a]	63-10 (10.5-10.0)[a]	62-7.5 (8.39-7.50)[a]
R_{merge} (%)	24.1 (104)	7.0 (72.9)	7.1 (124)	18.7 (88.2)	20.7 (121)
$I/\sigma I$	5.2 (1.5)	6.8 (1.5)	9.7 (1.6)	5.1 (1.5)	12.2 (2.8)
Completeness (%)	99.1 (99.8)	99.7 (99.9)	98.8 (98.9)	93.9 (96.8)	99.8 (100)
Redundancy	5.1 (5.1)	3.2 (3.3)	6.6 (6.9)	3.3 (3.6)	6.9 (7.1)
Refinement					
Resolution (Å)	87-4.0	58-6.0	112-6.0	63-10.0	55-7.5
Number of reflections	28954	19391	11641	11079	4253
R_{work}/R_{free} (%)	25.7 / 30.8	27.4 / 30.0	30.5 / 31.6	33.5 / 37.0	35.3 / 37.1
Number of atoms	9692	19191	9546	28787	9134
Protein	9692	19191	9546	28787	9134
Ligand/ion	0	0	0	0	0
Water	0	0	0	0	0
B factors (Å^2)					
Protein	134	273	283	237	250
Ligand/ion	–	–	–	–	–
Water	–	–	–	–	–
R.m.s. deviations					
Bond lengths (Å)	0.006	0.006	0.006	0.019	0.019
Bond angles (°)	1.2	1.3	1.2	1.7	2.4

[a]Highest resolution shell is shown in parenthesis.

Figure A.1: Crystallographic and refinement statistics of PlxnA ectodomains (1).

	PlxnA4$_{1\text{-}10}$	PlxnA4$_{1\text{-}10}$	PlxnA4$_{1\text{-}10}$	PlxnA2$_{4\text{-}5}$	PlxnA1$_{7\text{-}10}$
Data collection					
Space group	$P2_1$	$P4_32_12$	$P4_322$	$C2$	$C222_1$
Cell dimensions					
a, *b*, *c* (Å)	142, 241, 144	272, 272, 251	189, 189, 253	107.7, 44.6, 33.0	75.9, 82.5, 125.6
α, β, γ (°)	90, 99.8, 90	90, 90, 90	90, 90, 90	90, 104.7, 90	90, 90, 90
Resolution (Å)	48-8.0 (8.95-8.00)[a]	72-8.0 (8.94-8.00)[a]	71-8.5 (9.50-8.50)[a]	52.1-1.36 (1.40-1.36)[a]	62.8-2.2 (2.27-2.20)[a]
R_{merge} (%)	15.8 (76.4)	13.0 (89.4)	33.0 (117)	3.8 (47.6)	12.0 (82.5)
$I/\sigma I$	9.0 (2.9)	14.8 (1.3)	4.2 (1.7)	17.0 (2.3)	9.2 (1.9)
Completeness (%)	99.5 (100)	99.5 (100)	99.7 (100)	85.6 (41.0)	100 (99.8)
Redundancy	4.1 (4.3)	3.7 (3.9)	5.8 (6.1)	3.7 (3.4)	8.1 (5.0)
Refinement					
Resolution (Å)	48-8.0	72-8.0	70-8.5	52.1-1.36	62.8-2.2
Number of reflections	10061	10298	4374	27858	20219
R_{work}/R_{free} (%)	34.9 / 34.9	37.3 / 39.5	40.1 / 41.2	14.1 / 19.6	20.7 / 25.3
Number of atoms	15030	7189	7189	1334	2841
Protein	15030	7189	7189	1228	2703
Ligand/ion	0	0	0	3	8
Water	0	0	0	103	130
B factors (Å^2)					
Protein	264	251	163	16.3	58.8
Ligand/ion	–	–	–	24.9	78.0
Water	–	–	–	29.6	51.0
R.m.s. deviations					
Bond lengths (Å)	0.019	0.019	0.019	0.020	0.003
Bond angles (°)	2.5	2.5	2.5	2.00	0.82

[a]Highest resolution shell is shown in parenthesis.

Figure A.2: Crystallographic and refinement statistics of PlxnA ectodomains (2).

Name of construct	**Residue range (mutation site)**	**Description**	**Vector (C-ter tag)**	**Experimental usage**
PlxnA1$_{1\text{-}10}$	36-1236	PlxnA1 full ectodomain (sema-IPT6)	pHLsec (6 xhistadine)	Crystallisation Negative stain EM AUC MALS
PlxnA1$_{1\text{-}10}$ F145N/ L147S	36-1236 (F145N/L147S)	PlxnA1 full ectodomain (sema-IPT6) interface-blocking mutant 1 (F145N/L147S)	pHLsec (6 x histadine)	MALS AUC
PlxnA1$_{1\text{-}10}$ F693N/ E695S	36-1236 (F693N/E695S)	PlxnA1 full ectodomain (sema-IPT6) interface-blocking mutant 2 (F690N/E695S)	pHLsec (6 x histadine)	MALS
PlxnA1$_{7\text{-}10}$	861-1236	PlxnA1 domains 7-10 (IPT3-IPT6) interface-blocking mutant 2 (F690N/E695S)	pHLsec (6 x histadine)	Crystallisation
PlxnA1$_{4\text{-}5}$	655-804	PlxnA1 domains 4-5 (PSI2-IPT2)	pHLsec (6 x histadine)	Stimulants in DG growth cone assay
PlxnA1$_{4\text{-}5}$ F693N/ E695S	655-803 (F693N/E695S)	PlxnA1 domains 4-5 (PSI2-IPT2) interface-blocking mutant (F690N/E692S)	pHLsec (6 x histadine)	Stimulants in DG growth cone assay
PlxnA2$_{1\text{-}10}$	36-1236	PlxnA2 full ectodomain (sema-IPT6)	pHLsec (6 x histadine)	Crystallisation
PlxnA2$_{1\text{-}2}$	35-560	PlxnA2 domains 1-2 (sema-PIS1)	pHLsec (6 x histadine)	SPR (analyte)
PlxnA2$_{4\text{-}5}$	655-803	PlxnA2 domains 4-5 (PSI2-IPT2)	pHLsec (Avitag)	SPR (ligand)
PlxnA2$_{4\text{-}5}$ F690N/ E692S	655-803 (F690N/E692S)	PlxnA2 domains 4-5 (PSI2-IPT2) interface-blocking mutant (F690N/E692S)	pHLsec (Avitag)	SPR (ligand)
PlxnA2$_{FL}$-mVenus	36-1894	Full length PlxnA2 linked with C-terminal mVenus by RTLEVLFQGP	pHLsec (mVenus)	localisation microscopy
PlxnA2-mVenus	36-1260	PlxnA2 ectodomain and transmembrane helix linked with C-terminal mVenus by RTLEVLFQGP	pHLsec (mVenus)	FRET-FLIM

Figure A.3: List of PlxnA and PlxnD1 constructs (1).

Name of construct	**Residue range (mutation Site)**	**Description**	**Vector (C-terminal tag)**	**Experimental usage**
PlxnA2-mTFP1	36-1260	PlxnA2 ectodomain and transmembrane helix linked with C-terminal mTFP1 by RTLEVLFQGP	pHLsec (mTFP1)	FRET-FLIM
$PlxnA2_{FL}$-mVenus F690N/E692S	36-1894 (F690N/E692S)	Full length PlxnA2 linked with C-terminal mVenus by RTLEVLFQGP, interaction-blocking mutant	pHLsec (mVenus)	localisation microscopy
PlxnA2-mVenus F690N/E692S	36-1260 (F690N/E692S)	PlxnA2 ectodomain and transmembrane helix linked with C-terminal mVenus by RTLEVLFQGP, interaction-blocking mutant	pHLsec (mVenus)	FRET-FLIM
PlxnA2-mTFP1 F690N/E692S	36-1260 (F690N/E692S)	PlxnA2 ectodomain and transmembrane helix linked with C-terminal mTFP1 by RTLEVLFQGP, interaction-blocking mutant	pHLsec (mTFP1)	FRET-FLIM
$Sema6A_{1-2}$-Fc	19-571	Fc-tagged Sema6A domains 1-2 (sema-PSI1_ dimer	pHLsec (Fc-tag, 6 x histadine)	Stimulants in FRET-FLIM and localisation microscopy
$PlxnA4_{1-10}$	24-1236	PlxnA4 full ectodomain (sema-IPT6)	pHLsec (6 x histadine)	Crystallisation
$PlxnA4_{FL}$	24-1893	Full length PlxnA4	pcDNA 3.1 (+)	Expression in COS-7 collapse assay
$PlxnA4_{FL}$ F689N/E691S	24-1893	Full length PlxnA4 interaction-blocking mutant	pcDNA 3.1 (+)	Expression in COS-7 collapse assay
$PlxnA4_{\Delta ecto}$	1236-1893	PlxnA4 ectodomain-deleted construct	pcDNA 3.1 (+)	Expression in COS-7 collapse assay
$PlxnD1_{1-2}$	49-556	PlxnD1 domains 1-2 (sema-PIS1)	pHLsec (6 xhistadine)	Crystallisation AUC
$PlxnD1_{1-10}$	49-1272	PlxnD1 full ectodomain (sema-IPT6)	pHLsec (6 xhistadine)	Crystallisation

Figure A.4: List of PlxnA and PlxnD1 constructs (2).

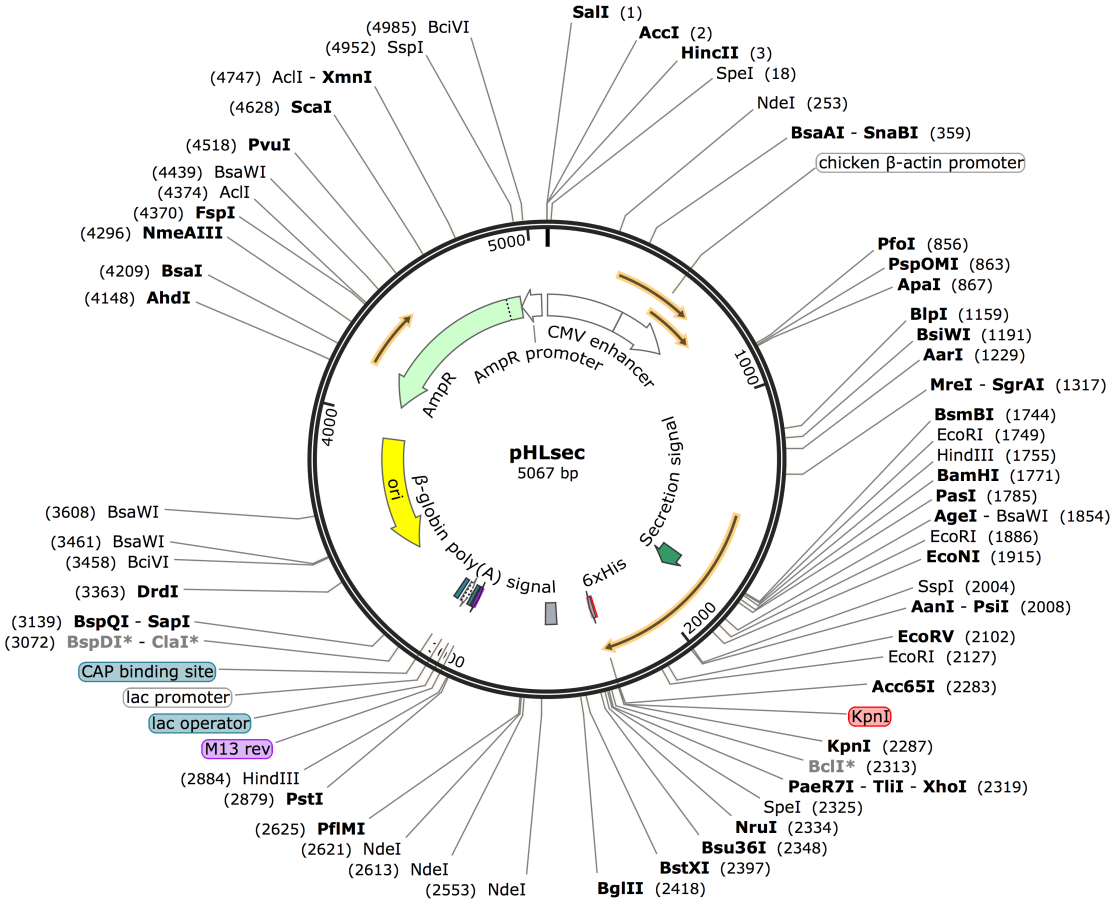


Figure A.5: Map of the pHLsec vector(1).

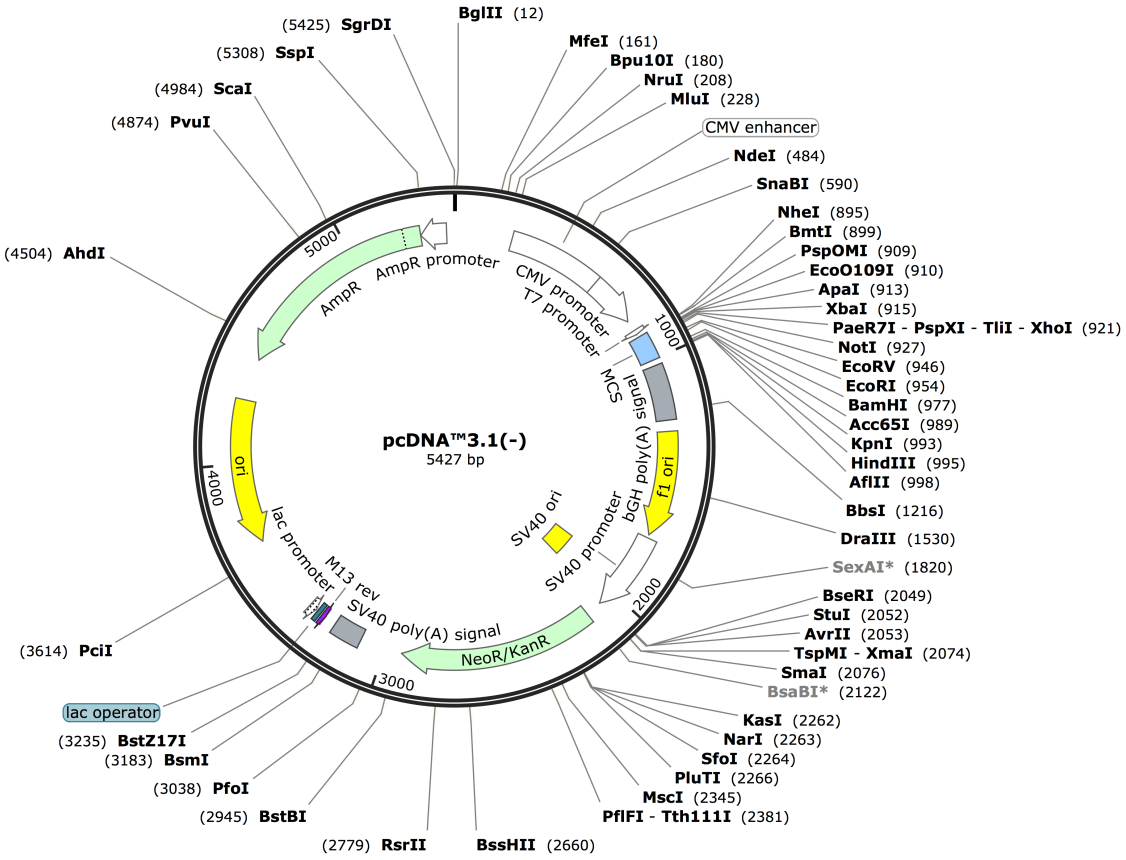


Figure A.6: Map of the pcDNA3.1(+) vector.