TU Delft

Delft University of Technology

Protein nanoarchitecture of neuronal wiring

Gogou, C.

DOI
10.4233/uuid:9da47c94-a3b7-49ea-a6bc-4c99973dc9db

Publication date
2025

Document Version
Final published version

Citation (APA)
Gogou, C. (2025). *Protein nanoarchitecture of neuronal wiring*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:9da47c94-a3b7-49ea-a6bc-4c99973dc9db

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

PROTEIN NANOARCHITECTURE OF NEURONAL WIRING

PROTEIN NANOARCHITECTURE OF NEURONAL WIRING

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen,
Chair of the Board for Doctorates,
to be defended publicly on
Friday 7th February 2025 at 10:00 o'clock.

by

Christos GOGOU

Master of Science in Nanobiology,
Delft University of Technology, The Netherlands,
born in Gorinchem, the Netherlands.

The dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Prof. dr. A.M. Dogterom	Delft University of Technology, promotor
Dr. D.H. Meijer	Delft University of Technology, copromotor
Independent members:	
Prof. dr. C. Joo	Delft University of Technology
Dr. M.E. Aubin-Tam	Delft University of Technology
Dr. P.E. Boukany	Delft University of Technology
Dr. J. Elegheert	CNRS, France
Other member:	
Prof. dr. ir. B.J.C. Janssen	Utrecht University
Reserve member:	
Prof. dr. G.H. Koenderink	Delft University of Technology



Keywords: Structural biology; Cryo-EM; Cell adhesion; Synaptic adhesion; Neurobiology; Protein-protein interactions; Single-molecule biophysics; Membrane proteins; Voltage-sensing proteins.

Cover: Design by the author and Gerard Kruithof.

Printed by: Gildeprint.

Copyright © 2024 by C. Gogou.

Casimir PhD Series.

ISBN 978-94-6496-312-0.

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`.

CONTENTS

Summary vii

Samenvatting ix

1 Introduction 1
- 1.1 Neuronal pathfinding in network formation . . . 2
- 1.2 Cell adhesion and transmembrane signalling . . . 2
- 1.3 SynCAM diversity for synaptic specificity . . . 5
- 1.4 Alternative splicing in neuronal genes . . . 7
- 1.5 Splicing-dependent synaptic complex formation . . . 7
- 1.6 Thesis overview . . . 9
- References . . . 11

2 Alternative splicing controls teneurin-3 compact dimer formation 21
- 2.1 Introduction . . . 22
- 2.2 Results . . . 24
  - 2.2.1 A_1 variants compact dimerise through an extended β-sheet . . . 24
  - 2.2.2 Splice insert A as a potential spacer to dislodge EGF6 . . . 27
  - 2.2.3 All isoforms are compact in a calcium-dependent manner . . . 27
  - 2.2.4 A model for splicing-dependent *trans* interactions . . . 31
  - 2.2.5 Insert B attracts teneurin-3-expressing outgrowing axons . . . 33
- 2.3 Discussion . . . 34
- 2.4 Methods . . . 37
- 2.5 Supplementary Information . . . 45
- References . . . 56

3 Mechanistic insights into microphthalmia via a teneurin-3 mutation 63
- 3.1 Introduction . . . 64
- 3.2 Results . . . 65
  - 3.2.1 The high-pathogenicity A_1 contact harbours the MO mutation . . . 65
  - 3.2.2 A_1 cell clustering is specifically affected by MO mutation . . . 67
  - 3.2.3 MO mutation may obstruct the EGF8-ABD contact . . . 70
- 3.3 Discussion . . . 72
- 3.4 Methods . . . 75
- 3.5 Supplementary Information . . . 79
- References . . . 82

4 Structural insights into the contactin-1 – neurofascin-155 adhesion 87
- 4.1 Introduction . . . 88
- 4.2 Results . . . 90
  - 4.2.1 Cntn1-Nf155 is mediated by conserved Ig1-Ig2 interactions . . . 90
  - 4.2.2 Mannose-rich glycans stabilise the Cntn1-Nf155 contact . . . 90
  - 4.2.3 Homodimer and Cntn1 interfaces on Nf155 overlap . . . 91
  - 4.2.4 Cell clustering assay supports the structural insights . . . 93

4.3 Discussion . . . 94
4.4 Methods . . . 98
4.5 Supplementary Information . . . 102
References . . . 106

5 Setting up a workflow for structural studies of protein voltage-sensitivity 115
5.1 Introduction . . . 116
5.2 Results . . . 118
5.2.1 ASAP2s and Ace2N-mNeon are functionally expressed and localise to the membrane of HEK-T and HEK-E cells . . . 118
5.2.2 ASAP2s but not Ace2N-mNeon is stably purified using whole-cell solubilisation and affinity purification . . . 120
5.2.3 ASAP2s exhibits three melting transitions and displays oligomeric states during mass photometry. . . . 122
5.2.4 ASAP2s is reconstituted into liposomes . . . 124
5.3 Discussion . . . 127
5.4 Methods . . . 129
5.5 Supplementary Information . . . 136
References . . . 140

6 Review: Enabling cryo-EM reconstruction of low-SNR proteins 145
6.1 Introduction . . . 146
6.2 Target binding strategies . . . 146
6.3 Target fusion strategies . . . 149
6.4 Symmetric scaffolding strategies . . . 150
6.5 Summary . . . 151
References . . . 153

7 Conclusion and Outlook 157
References . . . 162

Acknowledgements (Dankwoord) 167

Curriculum vitæ 173

List of publications 175

About the author 177

SUMMARY

Membrane proteins are of vital importance for human life. They reside in the plasma membrane, where they provide the adhesive forces needed for cells to spatially organise themselves relative to one another. Additionally, given their unique localisation - being in contact with both sides of the membrane - membrane proteins enable various modes of sensing and transmembrane signalling. For instance, channel proteins regulate the selective membrane permeation to ions and nutrients. Various other membrane proteins transduce information across the membrane by structural changes in response to direct binding interactions on either the intra- or extracellular side.

In the **first chapter**, we briefly introduce the role of a particular type of membrane proteins, namely synaptic cell adhesion molecules (synCAM) and their importance in organising neuronal networks. We discuss the role of synCAMS in recognizing extracellular cues which constitutes the mechanistic basis for network formation, and focus on the structural basis of synCAM binding specificity.

In the **second chapter**, teneurin-3 (Ten3) is investigated as a model synCAM for splicing-dependent binding specificity. We resolve the structures of various isoforms to reveal how small splice inserts completely reorganise intramolecular domain configuration. These structural changes are shown to underlie *trans*-cellular clustering and the guidance of axon outgrowth.

In the **third chapter**, we investigate the molecular basis of a neurological disease called microphthalmia ("small eye disease"). A unique patient-derived Ten3 missense mutation resides at a intramolecular domain-domain interface specific to two of the splicing isoforms. We show how the point mutation disrupts *trans*-cellular binding which may explain the eye and brain malformations, and thereby the cognitive impairments typical in microphthalmia patients.

In the **fourth chapter**, we look into the structural basis of glycosylation-dependent binding specificity between synCAMs contactin 1 and neurofascin 155. Only specific glycan chains at the interface allow formation of the complex, and subsequently, enable the *trans*-cellular contactin 1 - neurofascin 155 interactions.

In the **fifth chapter**, we set up a workflow for the reconstitution of purified neuronal membrane proteins in the lipid bilayers of vesicles over which transmembrane voltages can be induced. This will allow structural studies of neuronal membrane proteins in an environment more representative of the neuronal membrane, which is subjected to the dynamical voltages of the action potential.

In the **sixth chapter**, we provide an overview of techniques that have enabled the reconstruction of small, low signal-to-noise proteins using electron cryo-microscopy (cryo-EM). These approaches typically involve appending additional protein domains to the target, thereby facilitating particle picking and alignment needed for high-resolution cryo-EM single-particle analysis. This methodological review is aimed to select an approach for reconstructing the membrane-embedded proteins of chapter 5.

SAMENVATTING

Membraaneiwitten zijn van vitaal belang voor het menselijk leven. Ze bevinden zich in het plasmamembraan, waar ze de kracht door adhesie leveren die nodig is voor cellen om zich ruimtelijk ten opzichte van elkaar te organiseren. Bovendien, gezien hun unieke lokalisatie - in contact met beide zijden van het membraan - maken membraaneiwitten verschillende vormen van waarneming en transmembraan signaaloverdracht mogelijk. Bijvoorbeeld, kanaaleiwitten reguleren de selectieve membraanpermeabiliteit voor ionen en voedingsstoffen. Verschillende andere membraaneiwitten dragen informatie door het membraan over door structurele veranderingen als reactie op directe bindinginteracties aan de intra- of extracellulaire zijde.

In het **eerste hoofdstuk** introduceren we kort de rol van een bepaald type membraaneiwit, namelijk synaptische celadhesiemoleculen (synCAM) en hun belang bij het organiseren van neuronale netwerken. We bespreken de rol van synCAMs bij het herkennen van extracellulaire input, wat de mechanistische basis vormt voor netwerkvorming, en richten ons op de structurele basis van synCAM bindingsspecificiteit.

In het **tweede hoofdstuk** wordt teneurin-3 (Ten3) onderzocht als een modelsynCAM voor splicing-afhankelijke bindingsspecificiteit. We lossen de structuren op van verschillende isoformen om te onthullen hoe kleine splicevariaties de intramoleculaire domeinconfiguratie volledig reorganiseren. Deze structurele veranderingen blijken ten grondslag te liggen aan *trans*-cellulaire clustering en de wegbewijzering van axonale uitgroei.

In het **derde hoofdstuk** onderzoeken we de moleculaire basis van een neurologische ziekte genaamd microphthalmie ("klein-oog ziekte"). Een unieke, patiënt-afgeleide Ten3 missense-mutatie bevindt zich op een intramoleculair domein-domein contact die specifiek is voor twee van de splicing isoformen. We tonen aan hoe de puntmutatie *trans*-cellulaire binding verstoort, wat mogelijk de oog- en hersenafwijkingen verklaart, en daarmee de cognitieve beperkingen die typisch zijn voor microphthalmiepatiënten.

In het **vierde hoofdstuk** onderzoeken we de structurele basis van glycosyleringafhankelijke bindingsspecificiteit tussen synCAMs contactin-1 en neurofascin-155. Alleen specifieke glycanenketens op het contactpunt maken de vorming van het complex mogelijk, en stellen vervolgens de *trans*-cellulaire interacties tussen contactin-1 en neurofascin-155 in staat.

In het **vijfde hoofdstuk** stellen we een workflow op voor de reconstructie van gezuiverde neuronale membraaneiwitten in de lipidedubbellagen van blaasjes waarover transmembraansvoltages kunnen worden geïnduceerd. Dit zal structurele studies van neuronale membraaneiwitten mogelijk maken in een omgeving die representatiever is voor het neuronale membraan, welke onderhevig is aan de dynamische voltages van het actiepotentiaal.

In het **zesde hoofdstuk** geven we een overzicht van technieken die de reconstructie van kleine, lage signaal-tot-ruis-verhouding eiwitten met elektronen cryo-microscopie (cryo-EM) hebben mogelijk gemaakt. Deze benaderingen omvatten doorgaans het toevoegen van extra eiwitdomeinen aan het doelwit, wat de deeltjesselectie en uitlijning vergemakkelijkt die nodig zijn voor cryo-EM reconstructies tot hoge resolutie. Dit me-

thodologische naslagwerk is gericht op het selecteren van een benadering voor het reconstrueren van de in membraan geplaatste eiwitten van hoofdstuk 5.

1

INTRODUCTION

"De orde en het verband der ideeën zijn dezelfde als de orde en het verband der dingen."
- Benedictus de Spinoza

1

1.1. NEURONAL PATHFINDING IN NETWORK FORMATION

In his influential work *Ethica* (1677), Spinoza stated that the order and connection of our conceptions are the same as the order and connection of things[1]. Far ahead of his time, he may have been intuitively grasping the structure of the brain - an ordered connection of neurons enabling us to understand the physical world of interrelated objects and events.

The adult human brain comprises nearly 100 billion neurons[2,3], which form a network through an estimated total of 0.15 quadrillion neuron-neuron contacts in the neocortex alone[4,5]. These neuronal junctions, known as synapses*, play a crucial role in brain function. During embryonic development, neurons proliferate to nearly their final number in a process called neurogenesis[6,7]. After their birth, the neonatal neurons then migrate to specific regions to organise the brain's ultimate anatomical substructures. Migratory defects in the central nervous system (CNS) are implicated in various neurological diseases[8]. Prior to forming actual synaptic contacts, neurons prepare for network formation by extending neurites (axons and dendrites) toward their neuronal binding partners[9–11]. Some neurons in the spinal cord can extend axons up to 1 metre long, which is 200,000 times the size of the neuronal cell body. By adolescence, the brain has formed the explosive amount of synapses[6,7], which are subsequently refined through pruning - eliminating roughly 40% of these contacts[12]. The final neuronal network configuration is crucial for neurological health, where increasing evidence points at faulty wiring at the basis of many severe neurological disease, such as autism spectrum disorder[13–16], fragile X chromosome[17], tuberous sclerosis complex[18–20], and bipolar disorder and schizophrenia[16,21].

A fundamental role of the brain network is to ensure efficient communication between the connected neurons. This is the basis of the brain's ability to process all sorts of inputs into behaviour. Information is transferred by two main modes. Within the neuron, an action potential is a spike in transmembrane voltage that propagates intracellularly (Box **1.1**). It typically runs from the neuronal soma along the axon to the axon terminal, which is synaptically connected to the next neuron's dendrite. At the synapse, the incoming action potential then induces neurotransmitter release to effectively transmit the action potential to the next neuron (Box **1.1**). Synapses are highly dynamic entities, undergoing continuous assembly and elimination[11,22]. Nonetheless, to faithfully enable synaptic transmission, the pre- and postsynaptic membranes, including the neurotransmitter release and receptor machinery, are tightly organised[23].

1.2. CELL ADHESION AND TRANSMEMBRANE SIGNALLING

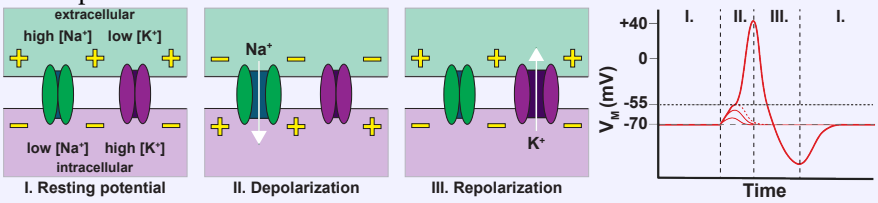
How do neurons robustly achieve the astronomical task of migration and precise network formation? How do neurons 'know' which other neurons to form synapses with, and how are these junctions physically held together? Cellular responses to external cues are prompted by tactile input from the extracellular environment – direct binding interactions with the extracellular domains (ECDs) of proteinaceous cell adhesion molecules

*From the ancient Greek σύν for 'together' and ἅπτειν for 'to fasten', and where σύναψις is a conjunction. Although more micro-organismic contacts are regarded synapses, here we exclusively refer to the neuronal kind.

Box 1.1: Neuronal communication

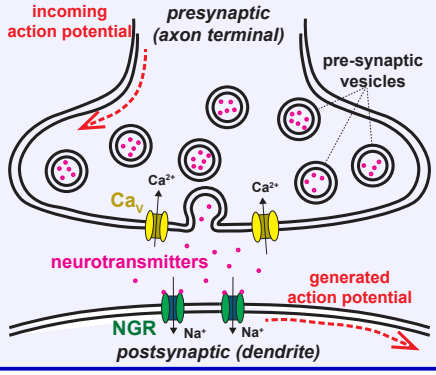
Intracellular: Action potential propagation

During inactivity, a neuron experiences a transmembrane voltage (V_M) of -70 mV, called the resting potential. This results from an electrochemical gradient instated by a Na^+/K^+-ATPase (not shown) that pumps K^+ into and Na^+ out of the cell. Upon stimulation of a membrane patch above a threshold (-55 mV), all voltage-gated Na^+ channels (Na_V) locally open to allow Na^+ influx and depolarisation (I.). K_V then open and Na_V close, causing K^+ outflux and regeneration of the negative V_M (II.). The feedback loop ends with all channels closing and restoration of the resting potential (III.). Upon local membrane depolarisation, nearby Na_V also open, causing a chain reaction where the spike is propagated from initial stimulation along the neuronal membrane to the axon terminal - an action potential.



Intercellular: Synaptic transmission

At the presynaptic axon terminal, the incoming action potential opens voltage-gated calcium channels (Ca_V). In-flowing Ca^{2+} instigates fusion of presynaptic vesicles with the plasma membrane, and their neurotransmitter content is released into the synaptic cleft. Neurotransmitter-gated receptors (NGC) open, and the influx of Na^+ induces an action potential at the postsynaptic dendrite.



(CAMs) that are anchored to the neuronal plasma membrane[11,22,24–30]. These cues often depend on direct cell-cell contact[27,31], but also longer-range diffusive protein cues like ephrins, netrins, semaphorins, and SLITs[24]. The well-studied N-cadherin family of CAMs, for instance, are classically known to guide neuronal migration[32]. Early studies already revealed that N-cadherins use astrocytes as navigation landmarks *in vitro*[33]. Protocadherins are another example of well-studied CAMs that are involved in self-recognition and subsequent self-avoidance of certain types on neurons[34].

External input can be interpreted as attractive (i.e. in the case of N-cadherins) or repulsive (i.e. self-avoiding protocadherins) by migrating cells. A seminal study showed that both stimuli play a role in proper wiring of the mouse hippocampus, involving the CAM families of teneurins and latrophilins[35]. Teneurin-expressing hippocampal axons

grow in the right direction by repulsive interactions with locally expressed latrophilins, and by attractive interactions with teneurins expressed in a more distal hippocampal region[35,36]. Similarly, another study captured teneurin-latrophilin repulsion *in vitro*[37]. Cortical neurons that exressed teneurin and FLRT were slowed down in their migration and axon outgrowth when placed onto surfaces coated with purified latrophilin adhesion domains.

CAMs do not only contribute to overall neuronal migration or neurite guidance. The subset of synaptic CAMs (synCAMs) is well-documented to play specialised roles in synapse architecture and synapse identity[11,23,27,38]. Teneurin-expressing cells can, for example, induce postsynaptic specialisations in neurons that express latrophilin[39]. The presynaptic neurexin[†] synCAM can similarly induce various postsynaptic specialisations in collaboration with neuroligins[40].

The best-characterised synCAMs can be categorised into a few main groups on basis of their ECD composition[22,25,27,28]: The immunoglobulin (Ig) superfamily[41], the cadherin superfamily[32], leucine-rich repeat-containing (LRR) superfamily[42,43], and less-canonical groups such as neurexins[43,44], neuroligins, teneurins and the latrophilin[39] adhesion G protein-coupled receptors (aGPCRs) (Fig. 1.1). Ig superfamily members, commonly contain a variable number of Ig-like domains and a few fibronectin III (FNIII) domains. LRR members are typified by a variable amount of leucine-rich domain repeats[42,45,46]. Teneurins also contain repeated regions, but with a wider variety of domain types. Teneurins are reported to also enable transmembrane signalling in multiple ways. A recent study identified intracellular binding partners via which teneurin regulates the actin cytoskeleton at the tips of growing axons[47,48], while others report that teneurins also affect microtubules further up the axons[49]. Earlier work has even hinted at nuclear translocation of the cleaved teneurin intracellular domain, where it would modulate gene expression[50]. The latrophilin adhesion G protein-coupled receptor is another striking example of how synCAMs participate in transmembrane signalling upon extracellular events. The force of teneurin binding to the latrophilin ECD leads to cleavage of the ECD, and the subsequent severed extracellular loop auto-activates its GPCR domain, thereby triggering intracellular G protein signalling cascades[51–54].

Beyond binding extracellular cues, SynCAMs also contribute to bi-directional synaptic specialisation[11,39,40,44], and participate in transmembrane signalling[11,23,25,26,43,45,46,55]. These functions add to the more intuitive repertoire of synCAM functionality of providing the selectivity in cell-cell recognition, and the adhesive force to physically align the pre- and postsynaptic machinery. Nobel laureate Thomas Südhof even proposed that adhesion can be considered a secondary role of synCAMs - after signalling[29]. It is, however, unambiguous that adhesion is a necessary step to prompt synCAM-mediated signal transduction. In the next section, we will discuss the binding diversity by which synCAMs establish *trans*-cellular binding specificity.

[†]Neurexins are expressed in three flavours: the full-length α-neurexin and the β- and γ-neurexins with largely truncated ECDs (Fig. 1.1). In the rest of this chapter, we will exclusively discuss α-neurexin, which will be referred to as 'neurexin'.

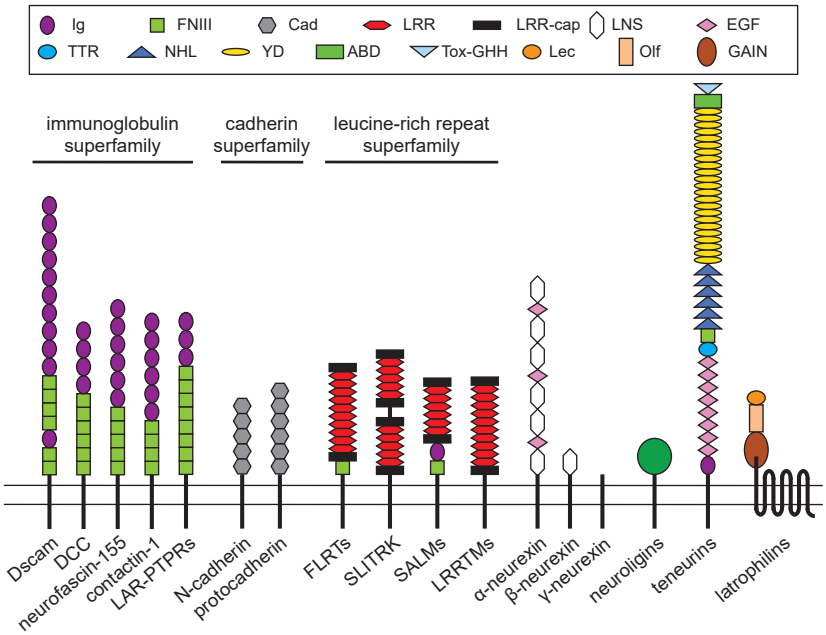


Figure 1.1: Extracellular domain compositions of well-studied synaptic cell adhesion molecule (synCAM) categories. All synCAMs are represented as attached to the membrane (two vertical black lines) by a transmembrane domain. Domain abbreviations: Ig, immunoglobulin; FNIII, fibronectin III; Cad, cadherin; LRR, leucine-rich repeat; LNS, lamin neurexin sex hormone-binding globulin; EGF, endothelial growth factor; TTR, transthyretin; NHL, ncl-1 HT2A lin-41; YD, tyrosine and aspartic acid-rich; ABD, antibiotic-binding domain; Tox-GHH, toxin glycine-histidine-histidine; Lec, lectin; Olf, olfactomedin; GAIN, GPCR autoproteolysis-inducing. Schematics were inspired by Moreland *et al.*[27].

1.3. SYNCAM DIVERSITY FOR SYNAPTIC SPECIFICITY

The first step towards synapse formation is the contact between two neurons displaying compatible synCAMs, informing the outgrowing neurites that their target is reached. The human genome is estimated to contain approximately 470 CAM genes[56], with an even smaller subset of synCAMs. How do neurons then achieve the specificity to form precise networks with more than 100 trillion contacts from less than 500 synCAM genes?

SynCAMs bind to each other on the same side of the synapse, or they can form a complex that bridges the synaptic cleft (Fig. 1.2A). These binding interactions are called *cis* and *trans*, respectively. Secreted proteins that diffuse through the synaptic cleft also participate in either such complexes[57–60]. Furthermore, complexes are formed though either homophilic binding between proteins of the same species, or synCAMs interact heterophilically with other synCAMs. Additional to the existence of different synCAM species, gene duplication throughout evolution[61] also gave rise to synCAMs that have multiple family members to diversify the barcode, such as latrophilins (3 family members), teneurins (4 family members), neurexins (3 family members), and FLRTs (3 family members) (Fig. 1.2B). Different synCAM variants can be expressed simultaneously in variable relative abundance to combinatorically expand the barcode repertoire. The relative abundance of interacting synCAMs on either side of the synapse may have to accurately match, since, for instance, up- or down-regulation of teneurins on either side

leads to mismatching between neurons that would otherwise be connected[48].

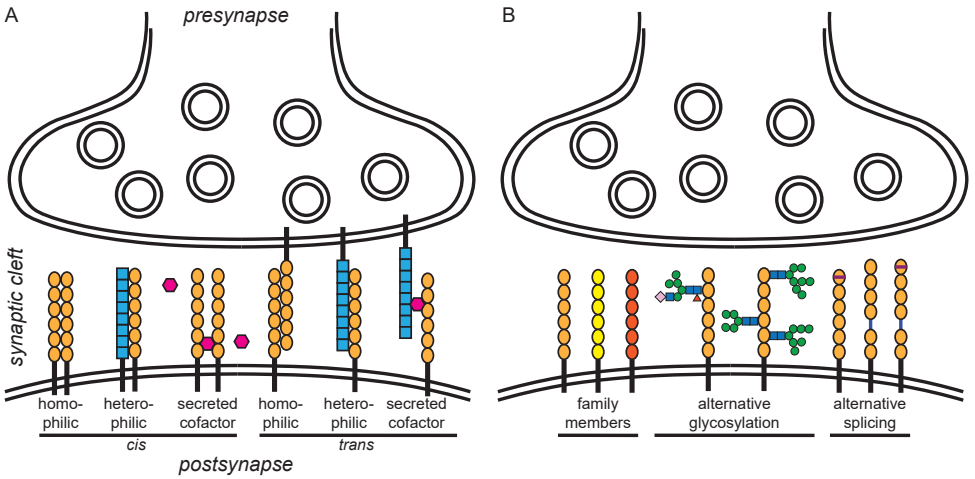


Figure 1.2: Modalities and diversity in synaptic cell adhesion molecule (synCAM) recognition. **A)** SynCAMs can bind each other on the same membrane (*cis*) or on the opposite synaptic membrane (*trans*). SynCAMs can bind their own species (homophilic) or bind other synCAMs (heterophilic). Non-membrane anchored secreted factor can also facilitate these interactions. **B)** SynCAM species typically comprise multiple family members. Glycosylation is a posttranslational modification where polysaccharide chains of variable sugar composition are chemically linked to the synCAM at various sites. Alternative splicing of mRNA results in different translated proteins from the same gene. All modes of variability can happen at either side of the synaptic membranes.

SynCAMs are also subject to posttranslational modification (PTM) (Fig. 1.2B). Glycosylation is a PTM in which polysaccharide chains are chemically linked to specific residues in the protein amino acid sequence[62–66]. The main types are N-linked and O-linked glycans, which differ in glycosylated residues as well as the involved sugars and their concatenation. Glycosylation is the most diverse PTM due to the many possibilities by which sugar groups can be incorporated into the long and often branched glycan chains. Glycans on ECDs are long understood to also affect cellular adhesion[62,67], and structural data is starting to accumulate in which these normally flexible glycan chains stably participate in synCAM binding interfaces[68–70].

Protein variants also emerge from variable posttranscriptional mRNA modifications, including alternative splicing (AS) (Fig. 1.2B). Protein-coding exons can be differentially included or omitted from the ultimately translated protein molecules[71]. RNA-sequencing data suggest, for instance, that AS of neurexins brings forth more than 1300 isoforms[72]. Even more extremely, the splicing sites on Down syndrome cell adhesion molecule (Dscam) in *Drosophila* code for ~38,000 theoretical isoforms - nearly twice the number of isoforms from one fly gene than there are human genes in total. In the next paragraphs, we will discuss alternative splicing in more detail and how high-resolution structural data reveal that added or removed splice inserts affect protein complex formation of synCAMs.

1.4. ALTERNATIVE SPLICING IN NEURONAL GENES

AS occurs in nearly 95% of the human multi-exonic genes[73,74]. On EnSEMBL , the existence of at least 70,000 isoforms is confirmed from the ~20,000-membered human genome[75,76]. In health, the relative expression of different isoforms is adapted to various conditions such as cell and tissue type, but also to more transient conditions like temperature and microbiome composition[77,78]. Although AS thus appears sensitive to environmental factors, the relative abundance of isoforms requires tight regulation, as misregulation has been related to, among others, neurodegenerative disease[79,80]. For the nervous system, RNA-seq data has recently revealed an enrichment in AS of so-called 'micro-exons' (1-9 amino acids), specifically, in neuronal-regulated genes[81]. Although AS micro-exons are associated with many neuronal functions[82], three of the five most significant gene ontology groups in this study[81] were proteins at the synapse; genes involved in neuron- and axonogenesis; and plasma membrane proteins. This makes it tempting to assume that the brain harnesses increased synCAM binding diversity through AS micro-exons to realise the synaptic specificity. In line with that notion, nearly 80% of the micro-exons were located at the protein surface, where they are more likely to affect protein-protein interactions. Kjer-Hansen and Weatheritt provide an extensive overview of the modes by which AS can affect protein function in general[83]. For synCAMs, however, not all these modes seem relevant, and others are even uniquely found in synCAM complex formation. We will discuss the structural basis of splicing-dependent synCAM-synCAM interactions in the next section.

1.5. SPLICING-DEPENDENT SYNAPTIC COMPLEX FORMATION

The scientific literature extensively describes splice inserts affecting synaptic complexes, and the shifting of synCAM binding preferences within the complexity of the synapse. Splice inserts occasionally serve as simple all-or-nothing switches, but the occurrence of multiple ligands in *cis* or *trans* for some synCAMs - with sometimes overlapping binding sites - is better described as binding promiscuity where neuronal phenotypes are ultimately more subtly modulated by alternative splicing.

The most abundantly described way splice inserts modulate protein-protein interactions is through their direct placement at binding interfaces (Fig. 1.3A). Here we will discuss some of the many examples. Neurexins and LAR-PTPRs (leukocyte common antigen-related receptor-type protein tyrosine phosphatases) synCAMs are good examples of an AS-broadened *trans*-synaptic binding palette[11,43,44,84,85]. Neurexin has a total of six splice sites dubbed SS1 through SS6, of which SS2 and SS4 are particularly well-described in affecting complex formation[43,44,58]. SS2 is found very close to the neurexin's binding site with secreted factor neurexophilin[86], strongly affecting their affinity for each other. For binding of LRRTM, SS4 was also found directly at the contact with neurexin[87]. Neurexin also binds in *trans* a synCAM known as neuroligin, which itself is spliced at two sites – ssA and ssB. Neurexin and neuroligin isoforms are differentially expressed throughout mouse brain, and expression of neurexins with constitutive SS4 insertion results in reduced postsynaptic AMPA receptor recruitment *in vivo*[88]. All neurexin and neuroligin isoforms and family members bind to each other with varying affinity[89], but the absence versus the presence of SS4 respectively shifts neurexin's concomitant stimu-

lation of the LLRTM-mediated signalling pathways to that of the GluD1/D2 metabotropic glutamate receptors[44,88]. Various structural studies show that neuroligin's ssB is found directly at the neurexin-neuroligin interface, while ssA and SS4 on neurexin flank the contact, making it less evident how these inserts can regulate affinity[90–93]. A protein known as MDGA binds neuroligin in *cis* – in direct competition with neurexin – and both complexes are affected by ssB[94,95]. Intriguingly, neuroligin's ssB contains an N-linked glycan, which is postulated to preferentially reduce MDGA-neuroligin binding compared to neurexin-neuroligin affinity (~7-fold vs. ~2-fold, respectively), partly due to the glycan's more proximal localisation to the MDGA-neuroligin binding site. ssA does not affect binding of neuroligin to either neurexin or MDGA1/2, despite its proximity to the overlapping binding sites of neurexin and MDGA on neuroligin[95]. Direct manipulation of the binding interface is also found for Dscam homophilic *cis*- and *trans* interactions[96,97]. The enormous variability in Dscam AS isoforms[98], is required for neurite self-avoidance[99], preventing neurons to form synapses with themselves. The variable presence in LAR-PTPRs of two micro-exon splice site (meA and meB) in immunoglobulin repeat 2 (Ig2) and in between Ig2-Ig3, respectively, determine their *trans*-cellular binding partners[43,85,100], and subsequently, pre- and postsynaptic specialisation[84] (Fig. 1.1). The most intuitive effect of splice inserts is found in the binding of family member PTPδ to IL1RAPL1, and to different SLITs and SALMs of the LRR superfamily[101–106]. The presence of MeB, for instance, increases the PTPδ-SLITRK2 affinity by nearly 10-fold[102]. MeB is required for the clustering of cells expressing either of the three LAR-PTPR family members with cells expressing SALM, whereas meA only modulates this effect[100].

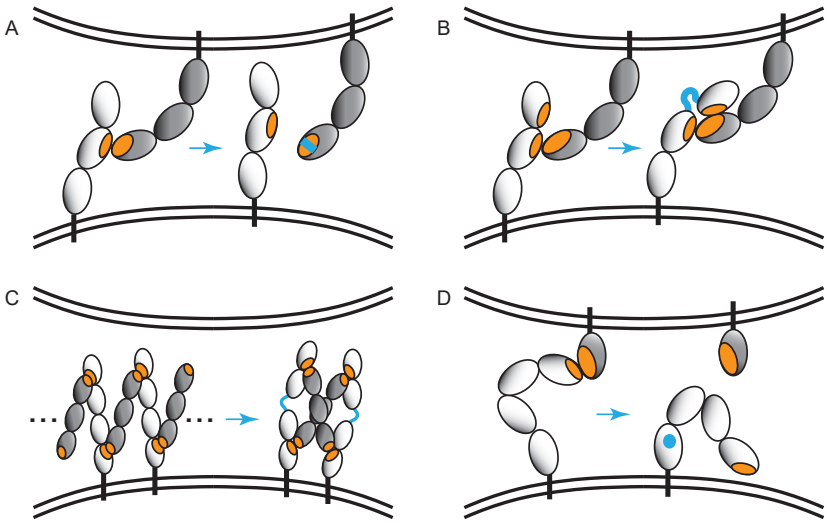


Figure 1.3: Structural modalities of binding modulation by splice inserts. **A)** Affinity between synaptic binding partners can be directly altered by insertion of an alternatively spliced exon directly at the binding interface. **B)** Splice inserts can alter the relative spatial freedom of domains, thereby regulating the participation of flanking domains in an existent binding interface. **C)** The splicing-dependent relative reorganisation of domains can lead to an altered stoichiometry whilst utilising the same binding sites. **D)** Alternative splicing of regions distal from the binding interfaces can cause intramolecular reorganisation of a synCAM, thereby changing exposure of the binding interface for interactions.

MeB very differently affects PTPδ binding to a member of the LRR superfamily called

IL-1RAcP, and to neuroligin-3. In this case, the meB insertion creates a few-amino acid spacer between Ig2-Ig3[101]. The resulting flexibility allows Ig3 to also engage in the pre-existing interface of IL-1RAcP with Ig2, which in turn is directly affected by meA. Conversely, PTPδ exclusively forms a *trans*-synaptic complex with neuroligin-3 in the absence of meB[107]. These examples demonstrate how splice inserts can alter the relative organisation of domains within a molecule to modify binding interfaces and modulate specificity (Fig. 1.3B).

Alternative splicing of DCC (deleted in colorectal cancer) synCAM results in a long and a short isoform – DCC_L and DCC_S, respectively. DCC_L is reported to promote axon outgrowth compared to its shorter isoform[60,108]. Both isoforms bind to netrin-1 with similar affinity[109], but combined structural studies revealed that the netrin-DCC configuration differs significantly. DCC_L-netrin bind in a 2:2 stoichiometry. The first netrin's LN domain links fibronectin repeat 4 (FN4) of one DCC to FN5 of the next DCC via that netrin's LE3 domain. The same domains bind the second netrin-1 to the complex but in opposite direction, thereby closing the complex[109]. DCC_S utilises the same binding sites to dock netrin-1, but the second netrin molecule now recruits a third DCC, resulting in a concatenation of an indefinite number of DCC on the membrane[60,109,110]. This difference is best explained by the increased FN4-FN5 distance in DCC_L, where the resulting increase in flexibility prevents steric hindrance between two netrins that link the same DCC monomers in the 2:2 complex. Here, the AS-dependent flexibility between domains can thus result in altered stoichiometry and severely affected size of the ultimate complex (Fig. 1.3C). It should be noted, however, that, in contrast DCC_S, both the long and short neogenin isoforms are capable of forming the 2:2 complex with netrin-1[109,111].

A unique mechanism is proposed for the teneurin-latrophilin *trans*-synaptic adhesion complex[68,83]. Two teneurin-2 isoforms - differing in the presence of 7-residue splice insert B in the NHL domain – exhibited altered binding to latrophilin[68]. Cells exogenously expressing teneurin-B_0 (insert B absent) do cluster with cells expressing latrophilin, whereas teneurin-B_1-expressing cells do not. Co-culturing of these teneurin-expressing cells leads to different postsynaptic specialisations on latrophilin-positive neurons in a splicing-dependent manner. Interestingly, binding between purified teneurins and latrophilins is observed regardless of teneurin isoform. A shapeshifting mechanism was therefore proposed where the membrane shields the binding site on the teneurin-2 surface – either by a novel membrane-teneurin interaction, or by intramolecular reorganisation of teneurin (Fig. 1.3D). In their report, Kjer-Hansen and colleague mention teneurin as the only example for this splicing-dependent alternative motif-hiding/exposure mechanism[83].

1.6. THESIS OVERVIEW

This thesis is centered around the structural biology of neuronal integration, with a focus on synCAMs in establishing neuronal networks.

In first experimental chapter, the unique splicing-dependent mechanism of *trans*-cellular binding specificity will be investigated for mouse teneurin-3. A combination of structural techniques, such as small-angle X-ray scattering, and negative stain- and cryo-electron microscopy (cryo-EM) reveal how the extracellular domains reorganise upon introduction of the two splice inserts. We thereby confirm the proposed splicing-

dependent motif-hiding/exposing mechanism in teneurins. Cellular assays such as cell clustering and axon guidance *in vitro* shine light on the effect of these intramolecular reorganisations on *trans*-cellular recognition.

Teneurin-3 is expressed in the central nervous system, mainly the hippocampus and the visual system, where it is known to participate in proper neuronal wiring. Ample patient data relate genetic mutations in the *TENM3* gene to microphthalmia, a disease typified by miswired visual system and cognitive impairments. In this thesis, we show that the only known single-residue missense mutation – in contrast to various reported nonsense mutations – affects teneurin-3 *trans*-cellular binding *in vitro*. More specifically, *trans*-cellular complex formation is disrupted for teneurin-3 isoforms that contain the missense mutation at an intramolecular domain-domain contact. We deploy molecular dynamics simulations to delineate the effect of this point mutation on teneurin-3 molecular behaviour for these isoforms.

We next investigate *trans*-cellular recognition between two members of the Ig superfamily of synCAMs: contactin-1 and neurofascin-155 (Fig. 1.1). An X-ray crystallographic structure reveals the involvement of a glycan in the heterophilic *trans* interaction. Cell clustering data in which glycosylation of these synCAMs is altered indeed shows the involvement of this structural glycan in binding specificity.

All membrane-anchored synCAMs are theoretically subject to the dynamic transmembrane potential that neurons use to communicate (Box **1.1**). Many synaptic membrane proteins, such as various ion channels, and two groups of GPCRs – muscarinic- and metabotropic glutamate receptors – are known to be voltage-sensitive in various aspects of their function. Transmembrane proteins have thus evolved to harness the energy from electrical fields for the regulation of protein activity. This demonstrates that representative and complete structural characterisation of neuronal membrane proteins in some cases requires embedding them in membranes with controllable membrane voltages. To this end, we assemble a workflow for a cryo-EM-compatible tool, in which the structure of versicle-embedded voltage-sensitive membrane proteins can be manipulated via chemically-induced transmembrane potentials.

In conclusion, this thesis aims to resolve the structural basis of synCAM binding specificity, focusing on how splicing and posttranslational modifications influence their interactions. It is unknown whether synCAM functionality is voltage-sensitive, a factor that could theoretically impact their synaptic complex formation and transmembrane signalling. To address this, in the final chapters, we sought to establish a cryo-EM method to investigate the structural voltage-dependence of membrane-embedded neuronal proteins. We thereby aim to characterise the structural factors that are key to the formation of neuronal networks in the brain. Perhaps human kind will thereby one day fully grasp how this complex interplay of biomolecules constitutes a working model of the infinitely complex world we inhabit.

REFERENCES

[1] B. de Spinoza, *Ethica* (Rieuwertsz, Amsterdam, 1677) part II, Proposition 7.

[2] F. A. Azevedo, L. R. Carvalho, L. T. Grinberg, J. M. Farfel, R. E. Ferretti, R. E. Leite, W. J. Filho, R. Lent, and S. Herculano-Houzel, *Equal numbers of neuronal and non-neuronal cells make the human brain an isometrically scaled-up primate brain,* Journal of Comparative Neurology **513**, 532 (2009).

[3] S. Herculano-Houzel, J. H. Kaas, and R. de Oliveira-Souza, *Corticalization of motor control in humans is a consequence of brain scaling in primate evolution,* Journal of Comparative Neurology **524**, 448 (2016).

[4] B. Pakkenberg, D. Pelvig, L. Marner, M. J. Bundgaard, H. Jørgen, G. Gundersen, J. R. Nyengaard, and L. Regeur, *Aging and the human neocortex,* Experimental Gerontology **38**, 95 (2003).

[5] Y. Tang, J. R. Nyengaard, D. M. D. Groot, and H. J. G. Gundersen, *Total regional and global number of synapses in the human brain neocortex,* Synapse **41**, 258 (2001).

[6] G. Z. Tau and B. S. Peterson, *Normal development of brain circuits,* Neuropsychopharmacology **35**, 147 (2010).

[7] J. C. Silbereis, S. Pochareddy, Y. Zhu, M. Li, and N. Sestan, *The cellular and molecular landscapes of the developing human central nervous system,* Neuron **89**, 248 (2016).

[8] I. Y. Buchsbaum and S. Cappello, *Neuronal migration in the cns during development and disease: Insights from in vivo and in vitro models,* Development (Cambridge) **146** (2019), 10.1242/dev.163766.

[9] R. Ayala, T. Shu, and L. H. Tsai, *Trekking across the brain: The journey of neuronal migration,* Cell **128**, 29 (2007).

[10] M. Rahimi-Balaei, H. Bergen, J. Kong, and H. Marzban, *Neuronal migration during development of the cerebellum,* Frontiers in Cellular Neuroscience **12** (2018), 10.3389/fncel.2018.00484.

[11] T. C. Südhof, *The cell biology of synapse formation,* Journal of Cell Biology **220** (2021), 10.1083/jcb.202103052.

[12] J. Bourgeois and P. Rakic, *Changes of synaptic density in the primary visual cortex of the macaque monkey from fetal to adult stage,* The Journal of Neuroscience **73**, 2801 (1993).

[13] K. McFadden and N. J. Minshew, *Evidence for dysregulation of axonal growth and guidance in the etiology of asd,* Frontiers in Human Neuroscience (2013), 10.3389/fnhum.2013.00671.

[14] A. D. Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, *et al.*, *The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism,* Molecular Psychiatry **19**, 659 (2014).

[15] R. M. Avital Hahamy, Marlene Behrmann, *The idiosyncratic brain: distortion of spontaneous connectivity patterns in autism spectrum disorder,* Nature Neuroscience **18**, 302 (2015).

[16] Y. Du, Z. Fu, Y. Xing, D. Lin, G. Pearlson, P. Kochunov, L. E. Hong, S. Qi, M. Salman, A. Abrol, and V. D. Calhoun, *Evidence of shared and distinct functional and structural brain signatures in schizophrenia and autism spectrum disorder,* Communications Biology **4** (2021), 10.1038/s42003-021-02592-2.

[17] M. R. Swanson, J. J. Wolff, M. D. Shen, M. Styner, A. Estes, G. Gerig, R. C. McKinstry, K. N. Botteron, J. Piven, and H. C. Hazlett, *Development of white matter circuitry in infants with fragile x syndrome,* JAMA Psychiatry **75**, 505 (2018).

[18] E. Widjaja, G. Simao, S. Z. Mahmoodabadi, A. Ochi, O. C. Snead, J. Rutka, and H. Otsubo, *Diffusion tensor imaging identifies changes in normal-appearing white matter within the epileptogenic zone in tuberous sclerosis complex,* Epilepsy Research **89**, 246 (2010).

[19] F. M. Baumer, J. W. Song, P. D. Mitchell, R. Pienaar, M. Sahin, P. E. Grant, and E. Takahashi, *Longitudinal changes in diffusion properties in white matter pathways of children with tuberous sclerosis complex,* Pediatric Neurology **52**, 615 (2015).

[20] K. Im, B. Ahtam, D. Haehn, J. M. Peters, S. K. Warfield, M. Sahin, and P. E. Grant, *Altered structural brain networks in tuberous sclerosis complex,* Cerebral Cortex **26**, 2046 (2016).

[21] C. H. Xia, Z. Ma, R. Ciric, S. Gu, R. F. Betzel, A. N. Kaczkurkin, M. E. Calkins, P. A. Cook, A. G. de la Garza, S. N. Vandekar, Z. Cui, T. M. Moore, D. R. Roalf, K. Ruparel, D. H. Wolf, C. Davatzikos, R. C. Gur, R. E. Gur, R. T. Shinohara, D. S. Bassett, and T. D. Satterthwaite, *Linked dimensions of psychopathology and connectivity in functional brain networks,* Nature Communications **9** (2018), 10.1038/s41467-018-05317-y.

[22] G. Rudenko, *Dynamic control of synaptic adhesion and organizing molecules in synaptic plasticity,* Neural Plasticity **2017** (2017), 10.1155/2017/6526151.

[23] T. Biederer, P. S. Kaeser, and T. A. Blanpied, *Transcellular nanoalignment of synaptic function,* Neuron **96**, 680 (2017).

[24] B. J. Dickson, *Molecular mechanisms of axon guidance,* Science **298**, 1959 (2002).

[25] S. Yamada and W. J. Nelson, *Synapses: Sites of cell recognition, adhesion, and functional specification,* Annual Review of Biochemistry **76**, 267 (2007).

[26] M. B. Dalva, A. C. McClelland, and M. S. Kayser, *Cell adhesion molecules: Signalling functions at the synapse,* Nature Reviews Neuroscience **8**, 206 (2007).

[27] T. Moreland and F. E. Poulain, *To stick or not to stick: The multiple roles of cell adhesion molecules in neural circuit assembly,* Frontiers in Neuroscience **16** (2022), 10.3389/fnins.2022.889155.

[28] J. R. Sanes and S. L. Zipursky, *Synaptic specificity, recognition molecules, and assembly of neural circuits,* Cell **181**, 536 (2020).

[29] T. C. Südhof, *Towards an understanding of synapse formation,* Neuron **100**, 276 (2018).

[30] D. Chowdhury, K. Watters, and T. Biederer, *Synaptic recognition molecules in development and disease,* (Academic Press Inc., 2021) pp. 319–370.

[31] C. S. G. Marc Tessier-Lavigne, *The molecular biology of axon guidance,* Science **274**, 1123 (1996).

[32] I. Martinez-Garay, *Molecular mechanisms of cadherin function during cortical migration,* Frontiers in Cell and Developmental Biology **8** (2020), 10.3389/fcell.2020.588152.

[33] J. L. Bixby and P. Jhabvala, *N-cadherin, ncam, and integrins promote retinal neurite outgrowth on astrocytes in vitro,* Journal of Cell Biology **107**, 1177 (1988).

[34] J. L. Lefebvre, Y. Zhang, M. Meister, X. Wang, and J. R. Sanes, *γ-protocadherins regulate neuronal survival but are dispensable for circuit formation in retina,* Development **135**, 4141 (2008).

[35] D. T. Pederick, J. H. Lui, E. C. Gingrich, C. Xu, M. J. Wagner, Y. Liu, Z. He, S. R. Quake, and L. Luo, *Reciprocal repulsions instruct the precise assembly of parallel hippocampal networks,* SCIENCE **372**, 1068+ (2021).

[36] D. S. Berns, L. A. DeNardo, D. T. Pederick, and L. Luo, *Teneurin-3 controls topographic circuit assembly in the hippocampus,* Nature **554**, 328 (2018).

[37] D. del Toro, M. A. Carrasquero-Ordaz, A. Chu, T. Ruff, M. Shahin, V. A. Jackson, M. Chavent, M. Berbeira-Santana, G. Seyit-Bremer, S. Brignani, R. Kaufmann, E. Lowe, R. Klein, and E. Seiradake, *Structural basis of teneurin-latrophilin interaction in repulsive guidance of migrating neurons,* Cell **180**, 323 (2020).

[38] K. Gerrow and A. El-Husseini, *Cell adhesion molecules at the synapse,* Frontiers in Bioscience **11**, 2400 (2006).

[39] J. Li, M. Shalev-Benami, R. Sando, X. Jiang, A. Kibrom, J. Wang, K. Leon, C. Katanski, O. Nazarko, Y. C. Lu, T. C. Südhof, G. Skiniotis, and D. Araç, *Structural basis for teneurin function in circuit-wiring: A toxin motif at the synapse,* Cell **173**, 735 (2018).

[40] E. R. Graf, X. Zhang, S. X. Jin, M. W. Linhoff, and A. M. Craig, *Neurexins induce differentiation of gaba and glutamate postsynaptic specializations via neuroligins,* Cell **119**, 1013 (2004).

[41] K. Zinn and E. Özkan, *Neural immunoglobulin superfamily interaction networks,* Current Opinion in Neurobiology **45**, 99 (2017).

[42] A. Schroeder and J. D. Wit, *Leucine-rich repeat-containing synaptic adhesion molecules as organizers of synaptic specificity and diversity,* Experimental and Molecular Medicine **50** (2018), 10.1038/s12276-017-0023-8.

[43] H. Y. Kim, J. W. Um, and J. Ko, *Proper synaptic adhesion signaling in the control of neural circuit architecture and brain function,* Progress in Neurobiology **200** (2021), 10.1016/j.pneurobio.2020.101983.

[44] T. C. Südhof, *Synaptic neurexin complexes: A molecular code for the logic of neural circuits,* Cell **171**, 745 (2017).

[45] J. Ko, *The leucine-rich repeat superfamily of synaptic adhesion molecules: Lrrtms and slitrks,* Molecules and Cells **34**, 335 (2012).

[46] J. de Wit and A. Ghosh, *Control of neural circuit formation by leucine-rich repeat proteins,* Trends in Neurosciences **37**, 539 (2014).

[47] J. Beckmann, R. Schubert, R. Chiquet-Ehrismann, and D. J. Müller, *Deciphering teneurin domains that facilitate cellular recognition, cell-cell adhesion, and neurite outgrowth using atomic force microscopy-based single-cell force spectroscopy,* Nano Letters **13**, 2937 (2013).

[48] C. Xu, Z. Li, C. Lyu, Y. Hu, C. N. McLaughlin, K. K. L. Wong, Q. Xie, D. J. Luginbuhl, H. Li, N. D. Udeshi, T. Svinkina, D. R. Mani, S. Han, T. Li, Y. Li, R. Guajardo, A. Y. Ting, S. A. Carr, J. Li, and L. Luo, *Molecular and cellular mechanisms of teneurin signaling in synaptic partner matching,* Cell (2024), 10.1016/j.cell.2024.06.022.

[49] S. Ichinose, Y. Susuki, N. Hosoi, R. Kaneko, M. Ebihara, H. Hirai, and H. Iwasaki, *Interaction between teneurin-2 and microtubules via eb proteins provides a platform for gabaa receptor exocytosis,* eLife **12** (2023), 10.7554/eLife.83276.

[50] G. Peppino, R. Ruiu, M. Arigoni, F. Riccardo, A. Iacoviello, G. Barutello, and E. Quaglino, *Teneurins: Role in cancer and potential role as diagnostic biomarkers and targets for therapy,* International Journal of Molecular Sciences **22**, 1 (2021).

[51] X. Barros-Álvarez, R. M. Nwokonko, A. Vizurraga, D. Matzov, F. He, M. M. Papasergi-Scott, M. J. Robertson, O. Panova, E. H. Yardeni, A. B. Seven, F. E. Kwarcinski, H. Su, M. C. Peroto, J. G. Meyerowitz, M. Shalev-Benami, G. G. Tall, and G. Skiniotis, *The tethered peptide activation mechanism of adhesion gpcrs,* Nature **604**, 757 (2022).

[52] B. L. Zhong, C. E. Lee, V. T. Vachharajani, M. S. Bauer, T. C. Südhof, and A. R. Dunn, *Piconewton forces mediate gain domain dissociation of the latrophilin-3 adhesion gpcr,* Nano Letters **23**, 9187 (2023).

[53] Y. Qian, Z. Ma, C. Liu, X. Li, X. Zhu, N. Wang, Z. Xu, R. Xia, J. Liang, Y. Duan, H. Yin, Y. Xiong, A. Zhang, C. Guo, Z. Chen, Z. Huang, and Y. He, *Structural insights into adhesion gpcr adgrl3 activation and gq, gs, gi, and g12 coupling,* Molecular Cell **82**, 4340 (2022).

[54] X. Zhu, Y. Qian, X. Li, Z. Xu, R. Xia, N. Wang, J. Liang, H. Yin, A. Zhang, C. Guo, G. Wang, and Y. He, *Structural basis of adhesion gpcr gpr110 activation by stalk peptide and g-proteins coupling,* Nature Communications **13**, 1 (2022).

[55] K. A. Han, J. W. Um, and J. Ko, *Intracellular protein complexes involved in synapse assembly in presynaptic neurons,* (Academic Press Inc., 2019) pp. 347–373.

[56] X. Zhong, J. Drgonova, C. Y. Li, and G. R. Uhl, *Human cell adhesion molecules: Annotated functional subtypes and overrepresentation of addiction-associated genes,* Annals of the New York Academy of Sciences **1349**, 83 (2015).

[57] S. Cheng, A. B. Seven, J. Wang, G. Skiniotis, and E. Özkan, *Conformational plasticity in the transsynaptic neurexin-cerebellin-glutamate receptor adhesion complex,* Structure **24**, 2163 (2016).

[58] T. C. Südhof, *Cerebellin–neurexin complexes instructing synapse properties,* Current Opinion in Neurobiology **81** (2023), 10.1016/j.conb.2023.102727.

[59] S. Fan, S. P. Gangwar, M. Machius, and G. Rudenko, *Interplay between hevin, sparc, and mdgas: Modulators of neurexin-neuroligin transsynaptic bridges,* Structure **29**, 664 (2021).

[60] B. Dailey-Krempel, A. L. Martin, H. N. Jo, H. J. Junge, and Z. Chen, *A tug of war between dcc and robo1 signaling during commissural axon guidance,* Cell Reports **42** (2023), 10.1016/j.celrep.2023.112455.

[61] A. L. Hughes, *Gene duplication and the origin of novel proteins,* Proceedings of the National Academy of Sciences **102**, 8791 (2005), https://www.pnas.org/doi/pdf/10.1073/pnas.0503922102 .

[62] T. N. Wight, M. G. Kinsella, and E. E. Qwarnstrgm, *The role of proteoglycans in cell adhesion, migration and proliferation,* Current Opinion in Cell Biology **4**, 793 (1992).

[63] V. P. Bhavanandan and K. Furukawa, *Biochemistry and oncology of sialoglycoproteins,* Biology of the Sialic Acids , 145 (1995).

[64] D. Medina-Cano, E. Ucuncu, L. S. Nguyen, M. Nicouleau, J. Lipecka, J.-C. Bizot, C. Thiel, F. O. Foulquier, N. Lefort, C. Faivre-Sarrailh, L. Colleaux, I. C. Guerrera, and V. Cantagrel, *High n-glycan multiplicity is critical for neuronal adhesion and sensitizes the developing cerebellum to n-glycosylation defect,* eLife (2018), 10.7554/eLife.38309.001.

[65] M. Rahman, N. J. Ramirez-Suarez, C. A. Diaz-Balzac, and H. E. Bülow, *Specific n - glycans regulate an extracellular adhesion complex during somatosensory dendrite patterning,* EMBO reports **23** (2022), 10.15252/embr.202154163.

[66] K. Scott, T. Gadomski, T. Kozicz, and E. Morava, *Congenital disorders of glycosylation: New defects and still counting,* Journal of Inherited Metabolic Disease **37**, 609 (2014).

[67] A. I. Fogel, Y. Li, J. Giza, Q. Wang, T. K. T. Lam, Y. Modis, and T. Biederer, *N-glycosylation at the syncam (synaptic cell adhesion molecule) immunoglobulin interface modulates synaptic adhesion,* Journal of Biological Chemistry **285**, 34864 (2010).

[68] J. Li, Y. Xie, S. Cornelius, X. Jiang, R. Sando, S. P. Kordon, M. Pan, K. Leon, T. C. Südhof, M. Zhao, and D. Araç, *Alternative splicing controls teneurin-latrophilin interaction and synapse specificity by a shape-shifting mechanism,* Nature Communications **11** (2020), 10.1038/s41467-020-16029-7.

[69] J. Wang, Y. Miao, R. Wicklein, Z. Sun, J. Wang, K. M. Jude, R. A. Fernandes, S. A. Merrill, M. Wernig, K. C. Garcia, and T. C. Südhof, *Rtn4/nogo-receptor binding to bai adhesion-gpcrs regulates neuronal development,* Cell **184**, 5869 (2021).

[70] O. Akkermans, C. Delloye-Bourgeois, C. Peregrina, M. Carrasquero-Ordaz, M. Kokolaki, M. Berbeira-Santana, M. Chavent, F. Reynaud, R. Raj, J. Agirre, M. Aksu, E. S. White, E. Lowe, D. B. Amar, S. Zaballa, J. Huo, I. Pakos, P. T. McCubbin, D. Comoletti, R. J. Owens, C. V. Robinson, V. Castellani, D. del Toro, and E. Seiradake, *Gpc3-unc5 receptor complex structure and role in cell migration,* Cell **185**, 3931 (2022).

[71] C. J. Wright, C. W. Smith, and C. D. Jiggins, *Alternative splicing as a source of phenotypic diversity,* Nature Reviews Genetics **23**, 697 (2022).

[72] D. Schreiner, T. M. Nguyen, G. Russo, S. Heber, A. Patrignani, E. Ahrné, and P. Scheiffele, *Targeted combinatorial alternative splicing generates brain region-specific repertoires of neurexins,* Neuron **84**, 386 (2014).

[73] Q. Pan, O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe, *Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing,* Nature Genetics **40**, 1413 (2008).

[74] E. T. Wang, R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S. F. Kingsmore, G. P. Schroth, and C. B. Burge, *Alternative isoform regulation in human tissue transcriptomes,* Nature **456**, 470 (2008).

[75] P. W. Harrison, M. R. Amode, O. Austine-Orimoloye, A. G. Azov, M. Barba, I. Barnes, A. Becker, R. Bennett, A. Berry, J. Bhai, S. K. Bhurji, S. Boddu, P. R. Lins, L. Brooks, S. B. Ramaraju, L. I. Campbell, M. C. Martinez, M. Charkhchi, K. Chougule, A. Cockburn, C. Davidson, N. H. D. Silva, K. Dodiya, S. Donaldson, B. E. Houdaigui, T. E. Naboulsi, R. Fatima, C. G. Giron, T. Genez, D. Grigoriadis, G. S. Ghattaoraya,

J. G. Martinez, T. A. Gurbich, M. Hardy, Z. Hollis, T. Hourlier, T. Hunt, M. Kay, V. Kaykala, T. Le, D. Lemos, D. Lodha, D. Marques-Coelho, G. Maslen, G. A. Merino, L. P. Mirabueno, A. Mushtaq, S. N. Hossain, D. N. Ogeh, M. P. Sakthivel, A. Parker, M. Perry, I. Piližota, D. Poppleton, I. Prosovetskaia, S. Raj, J. G. Pérez-Silva, A. I. A. Salam, S. Saraf, N. Saraiva-Agostinho, D. Sheppard, S. Sinha, B. Sipos, V. Sitnik, W. Stark, E. Steed, M. M. Suner, L. Surapaneni, K. Sutinen, F. F. Tricomi, D. Urbina-Gómez, A. Veidenberg, T. A. Walsh, D. Ware, E. Wass, N. L. Willhoft, J. Allen, J. Alvarez-Jarreta, M. Chakiachvili, B. Flint, S. Giorgetti, L. Haggerty, G. R. Ilsley, J. Keatley, J. E. Loveland, B. Moore, J. M. Mudge, G. Naamati, J. Tate, S. J. Trevanion, A. Winterbottom, A. Frankish, S. E. Hunt, F. Cunningham, S. Dyer, R. D. Finn, F. J. Martin, and A. D. Yates, *Ensembl 2024,* Nucleic Acids Research **52**, D891 (2024).

[76] R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, C. R. Bertozzi, E. S. Boja, C. E. Costello, B. F. Cravatt, C. Fenselau, B. A. Garcia, Y. Ge, J. Gunawardena, R. C. Hendrickson, P. J. Hergenrother, C. G. Huber, A. R. Ivanov, O. N. Jensen, M. C. Jewett, N. L. Kelleher, L. L. Kiessling, N. J. Krogan, M. R. Larsen, J. A. Loo, R. R. O. Loo, E. Lundberg, M. J. Maccoss, P. Mallick, V. K. Mootha, M. Mrksich, T. W. Muir, S. M. Patrie, J. J. Pesavento, S. J. Pitteri, H. Rodriguez, A. Saghatelian, W. Sandoval, H. Schlüter, S. Sechi, S. A. Slavoff, L. M. Smith, M. P. Snyder, P. M. Thomas, M. Uhlén, J. E. V. Eyk, M. Vidal, D. R. Walt, F. M. White, E. R. Williams, T. Wohlschlager, V. H. Wysocki, N. A. Yates, N. L. Young, and B. Zhang, *How many human proteoforms are there?* Nature Chemical Biology **14**, 206 (2018).

[77] S. Suzuki, N. Ayukawa, C. Okada, M. Tanaka, S. Takekoshi, Y. Iijima, and T. Iijima, *Spatio-temporal and dynamic regulation of neurofascin alternative splicing in mouse cerebellar neurons,* Scientific Reports **7**, 1 (2017).

[78] L. E. Marasco and A. R. Kornblihtt, *The physiology of alternative splicing,* Nature Reviews Molecular Cell Biology **24**, 242 (2023).

[79] D. Nikom and S. Zheng, *Alternative splicing in neurodegenerative disease and the promise of rna therapies,* Nature Reviews Neuroscience **24**, 457 (2023).

[80] Q. Liu, L. Fang, and C. Wu, *Alternative splicing and isoforms: From mechanisms to diseases,* Genes **13** (2022), 10.3390/genes13030401.

[81] M. Irimia, R. J. Weatheritt, J. D. Ellis, N. N. Parikshak, T. Gonatopoulos-Pournatzis, M. Babor, M. Quesnel-Vallières, J. Tapial, B. Raj, D. O'Hanlon, M. Barrios-Rodiles, M. J. Sternberg, S. P. Cordes, F. P. Roth, J. L. Wrana, D. H. Geschwind, and B. J. Blencowe, *A highly conserved program of neuronal microexons is misregulated in autistic brains,* Cell **159**, 1511 (2014).

[82] E. Furlanis and P. Scheiffele, *Regulation of neuronal differentiation, function, and plasticity by alternative splicing,* Annual Review of Cell and Developmental Biology Annu. Rev. Cell Dev. Biol **34** (2018), 10.1146/annurev-cellbio-100617.

[83] P. Kjer-Hansen and R. J. Weatheritt, *The function of alternative splicing in the proteome: rewiring protein interactomes to put old functions into new contexts,* Nature Structural and Molecular Biology **30**, 1844 (2023).

[84] H. Takahashi and A. M. Craig, *Protein tyrosine phosphatases ptpδ, ptp, and lar: Presynaptic hubs for synapse organization,* Trends in Neurosciences **36**, 522 (2013).

[85] K. Sakamoto, T. Ozaki, Y. Suzuki, and K. Kadomatsu, *Type iia rptps and glycans: Roles in axon regeneration and synaptogenesis,* International Journal of Molecular Sciences **22** (2021), 10.3390/ijms22115524.

[86] S. C. Wilson, K. I. White, Q. Zhou, R. A. Pfuetzner, U. B. Choi, T. C. Südhof, and A. T. Brunger, *Structures of neurexophilin–neurexin complexes reveal a regulatory mechanism of alternative splicing,* The EMBO Journal **38** (2019), 10.15252/embj.2019101603.

[87] A. Yamagata, S. Goto-Ito, Y. Sato, T. Shiroshima, A. Maeda, M. Watanabe, T. Saitoh, K. Maenaka, T. Terada, T. Yoshida, T. Uemura, and S. Fukai, *Structural insights into modulation and selectivity of transsynaptic neurexin–lrrtm interaction,* Nature Communications **9**, 1 (2018).

[88] J. Aoto, D. C. Martinelli, R. C. Malenka, K. Tabuchi, and T. C. Südhof, *Xpresynaptic neurexin-3 alternative splicing trans-synaptically controls postsynaptic ampa receptor trafficking,* Cell **154**, 75 (2013).

[89] J. Koehnke, P. S. Katsamba, G. Ahlsen, F. Bahna, J. Vendome, B. Honig, L. Shapiro, and X. Jin, *Splice form dependence of β-neurexin/neuroligin binding interactions,* Neuron **67**, 61 (2010).

[90] D. Araç, A. A. Boucard, E. Özkan, P. Strop, E. Newell, T. C. Südhof, and A. T. Brunger, *Structures of neuroligin-1 and the neuroligin-1/neurexin-1β complex reveal specific protein-protein and protein-ca2+ interactions,* Neuron **56**, 992 (2007).

[91] X. Chen, H. Liu, A. H. Shim, P. J. Focia, and X. He, *Structural basis for synaptic adhesion mediated by neuroligin-neurexin interactions,* Nature Structural and Molecular Biology **15**, 50 (2008).

[92] P. Leone, D. Comoletti, G. Ferracci, S. Conrod, S. U. Garcia, P. Taylor, Y. Bourne, and P. Marchot, *Structural insights into the exquisite selectivity of neurexin/neuroligin synaptic interactions,* EMBO Journal **29**, 2461 (2010).

[93] H. Tanaka, N. Miyazaki, K. Matoba, T. Nogi, K. Iwasaki, and J. Takagi, *Higher-order architecture of cell adhesion mediated by polymorphic synaptic adhesion molecules neurexin and neuroligin,* Cell Reports **2**, 101 (2012).

[94] J. A. Kim, D. Kim, S. Y. Won, K. A. Han, D. Park, E. Cho, N. Yun, H. J. An, J. W. Um, E. Kim, J. O. Lee, J. Ko, and H. M. Kim, *Structural insights into modulation of neurexin-neuroligin trans-synaptic adhesion by mdga1/neuroligin-2 complex,* Neuron **94**, 1121 (2017).

[95] J. Elegheert, V. Cvetkovska, A. J. Clayton, C. Heroven, K. M. Vennekens, S. N. Smukowski, M. C. Regan, W. Jia, A. C. Smith, H. Furukawa, J. N. Savas, J. de Wit, J. Begbie, A. M. Craig, and A. R. Aricescu, *Structural mechanism for modulation of synaptic neuroligin-neurexin signaling by mdga proteins,* Neuron **95**, 896 (2017).

[96] R. Meijers, R. Puettmann-Holgado, G. Skiniotis, J. H. Liu, T. Walz, J. H. Wang, and D. Schmucker, *Structural basis of dscam isoform specificity,* Nature **449**, 487 (2007).

[97] J. Cheng, Y. Yu, X. Wang, X. Zheng, T. Liu, D. Hu, Y. Jin, Y. Lai, T. M. Fu, and Q. Chen, *Structural basis for the self-recognition of sdscam in chelicerata,* Nature Communications **14** (2023), 10.1038/s41467-023-38205-1.

[98] D. Schmucker, J. C. Clemens, H. Shu, C. A. Worby, J. Xiao, M. Muda, and J. E. Dixon, *Drosophila dscam is an axon guidance receptor exhibiting extraordinary molecular diversity modified by neuronal activity (reviewed by albright et al., 2000). extraordinary progress has been made in identifying,* Cell **101**, 671 (2000).

[99] B. J. Matthews, M. E. Kim, J. J. Flanagan, D. Hattori, J. C. Clemens, S. L. Zipursky, and W. B. Grueber, *Dendrite self-avoidance is controlled by dscam,* Cell **129**, 593 (2007).

[100] Y. Li, P. Zhang, T. Y. Choi, S. K. Park, H. Park, E. J. Lee, D. Lee, J. D. Roh, W. Mah, R. Kim, Y. Kim, H. Kwon, Y. C. Bae, S. Y. Choi, A. M. Craig, and E. Kim, *Splicing-dependent trans-synaptic salm3-lar-rptp interactions regulate excitatory synapse development and locomotion,* Cell Reports **12**, 1618 (2015).

[101] A. Yamagata, T. Yoshida, Y. Sato, S. Goto-Ito, T. Uemura, A. Maeda, T. Shiroshima, S. Iwasawa-Okamoto, H. Mori, M. Mishina, and S. Fukai, *Mechanisms of splicing-dependent trans-synaptic adhesion by ptpδ -il1rapl1/il-1racp for synaptic differentiation,* Nature Communications **6** (2015), 10.1038/ncomms7926.

[102] A. Yamagata, Y. Sato, S. Goto-Ito, T. Uemura, A. Maeda, T. Shiroshima, T. Yoshida, and S. Fukai, *Structure of slitrk2-ptp complex reveals mechanisms for splicing-dependent trans-synaptic adhesion,* Scientific Reports **5** (2015), 10.1038/srep09686.

[103] J. W. Um, K. H. Kim, B. S. Park, Y. Choi, D. Kim, C. Y. Kim, S. J. Kim, M. Kim, J. S. Ko, S. G. Lee, G. Choii, J. Nam, W. D. Heo, E. Kim, J. O. Lee, J. Ko, and H. M. Kim, *Structural basis for lar-rptp/slitrk complex-mediated synaptic adhesion,* Nature Communications **5** (2014), 10.1038/ncomms6423.

[104] Z. Lin, J. Liu, H. Ding, F. Xu, and H. Liu, *Structural basis of salm5-induced ptpδ dimerization for synaptic differentiation,* Nature Communications **9** (2018), 10.1038/s41467-017-02414-2.

[105] S. Goto-Ito, A. Yamagata, Y. Sato, T. Uemura, T. Shiroshima, A. Maeda, A. Imai, H. Mori, T. Yoshida, and S. Fukai, *Structural basis of trans-synaptic interactions between ptpδ and salms for inducing synapse formation,* Nature Communications **9** (2018), 10.1038/s41467-017-02417-z.

[106] S. Fukai and T. Yoshida, *Roles of type iia receptor protein tyrosine phosphatases as synaptic organizers,* FEBS Journal **288**, 6913 (2021).

[107] L.-Y. Li, A. Imai, H. Izumi, R. Inoue, Y. Koshidaka, K. Takao, H. Mori, and T. Yoshida, *Differential contribution of canonical and noncanonical nlgn3 pathways to early social development and memory performance,* Molecular Brain **17**, 16 (2024).

[108] J. C. Leggere, Y. Saito, R. B. Darnell, M. Tessier-Lavigne, H. J. Junge, and Z. Chen, *Nova regulates dcc alternative splicing during neuronal migration and axon guidance in the spinal cord,* eLife **5** (2016), 10.7554/eLife.14264.001.

[109] K. Xu, Z. Wu, N. Renier, A. Antipenko, D. Tzvetkova-robev, Y. Xu, M. Minchenko, V. Nardi-dei, K. R. Rajashankar, and J. Himanen, *Guidance mechanism,* Sciemce **344**, 1275 (2014).

[110] L. I. Finci, N. Krüger, X. Sun, J. Zhang, M. Chegkazi, Y. Wu, G. Schenk, H. D. Mertens, D. I. Svergun, Y. Zhang, J. huai 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 **83**, 839 (2014).

[111] R. A. Robinson, S. C. Griffiths, L. L. van de Haar, T. Malinauskas, E. Y. van Battum, P. Zelina, R. A. Schwab, D. Karia, L. Malinauskaite, S. Brignani, M. H. van den Munkhof, O. Düdükcü, A. A. De Ruiter, D. M. Van den Heuvel, B. Bishop, J. Elegheert, A. R. Aricescu, R. J. Pasterkamp, and C. Siebold, *Simultaneous binding of guidance cues net1 and rgm blocks extracellular neo1 signaling,* Cell **184**, 2103 (2021).

2

ALTERNATIVE SPLICING CONTROLS TENEURIN-3 COMPACT DIMER FORMATION

Neuronal network formation is facilitated by recognition between synaptic cell adhesion molecules at the cell surface. Alternative splicing of cell adhesion molecules provides additional specificity in forming neuronal connections. For the teneurin family of cell adhesion molecules, alternative splicing of the EGF-repeats and NHL domain controls synaptic protein-protein interactions. Here we present cryo-EM structures of the compact dimeric ectodomain of two teneurin-3 isoforms that harbour the splice insert in the EGF-repeats. This dimer is stabilised by an EGF8-ABD contact between subunits. Cryo-EM reconstructions of all four splice variants, together with SAXS and negative stain EM, reveal compacted dimers for each isoform, with variant-specific dimeric arrangements. This results in specific trans-cellular interactions, as tested in cell clustering and stripe assays. The compact conformations provide a structural basis for teneurin homo- and heterophilic interactions. Altogether, our findings demonstrate how alternative splicing results in rearrangements of the dimeric subunits, influencing neuronal recognition and likely circuit wiring.

This chapter is published as *Christos Gogou*, J. Wouter Beugelink, Cátia P. Frias, Leanid Kresik, Natalia Jaroszynska, Uwe Drescher, Bert J.C. Janssen, Robert Hindges, and Dimphna H. Meijer, "Alternative splicing controls teneurin-3 compact dimer formation for neuronal recognition", *Nature Communications (2024)* [1].

2.1. INTRODUCTION

During the formation of the central nervous system, specific guidance and recognition signals are required to orchestrate the billions of outgrowing neurons as they form complex circuits. Important cues are especially provided by combinations of cell adhesion molecules expressed on the surface of neurons. Among these, teneurins, a conserved family of type II transmembrane proteins, play a crucial role in circuit wiring across various species[2]. However, the molecular mechanisms of how teneurins facilitate the formation of functional neuronal circuits remain poorly understood.

In mammals, there are four family members of teneurins, each expressed in distinct and often interconnected areas of the developing nervous system[3,4]. They have the capability to form homophilic interactions with other teneurins[5–9], as well as heterophilic interactions with other synaptic proteins, enabling a diverse range of functions in circuit development and maintenance[10]. For instance, the interaction between mouse teneurin-3 (Ten3) and latrophilin-2 (Lphn2), an adhesion G-protein coupled receptor (GPCR), orchestrates proper circuit wiring in the hippocampus[8]. This interaction involves a Ten3-Lphn2 repulsive mechanism, where axons expressing Ten3 avoid connecting with Lphn2-expressing regions. Instead, Ten3-expressing axons grow toward other areas enriched in Ten3, pointing toward homophilic *trans*-synaptic interactions[8,11]. Additional players such as FLRT proteins and glypican-3 may contribute to ternary complex formation with teneurins to control diverse functions such as neuronal cell migration, synapse formation, and pruning[9,12,13]. Finally, genetic abnormalities of all members of the teneurin family have been associated with diverse neurological disorders[14]. For example, teneurin-1 is linked to autism[15], teneurin-2 to major depressive disorder[16], teneurin-3 to microphthalmia[17], and teneurin-4 to essential tremor[18] and schizophrenia[19].

The teneurin intracellular domain (ICD), approximately 45 kDa in size, exhibits less conservation across different homologues compared to the extracellular domain (ECD)[20]. However, common features and domains can be identified, such as an EF-hand-like Ca^{2+}-binding site, potential phosphorylation sites, and polyproline-rich regions known to bind SH3-containing adaptor proteins[21]. The ICD may be released by a mechanism known as regulated intramembrane proteolysis and result in translocation of the ICD to the nucleus where it can coordinate transcription[21–23]. Structural studies of the teneurin ectodomains have revealed a multi-domain ~1850-residue core superfold, prominently featuring a tyrosine-aspartate rich barrel-shaped structure termed YD-shell[24,25]. This YD-shell is adorned with a calcium-binding C-rich domain, a transthyretin-like (TTR) domain and a fibronectin-like region (FN-plug), as well as with a 6-bladed β-propeller NHL domain (NCL-1, HT2A, Lin-41), all located at its N-terminal end[24,25]. At the C-terminal side of the ectodomain, structurally adjacent to the YD-shell, lies an antibiotic-binding domain-like (ABD) fold and a domain akin to the Tox-GGH class of DNases[26]. Connecting the transmembrane helix with the core superfold are a predicted immunoglobulin (Ig)fold and eight predicted EGF domains (EGF1-8). Notably, EGF2 and EGF5 contain free cysteines that directly contribute to covalent *cis* dimers formation in all teneurins[27,28]. A homophilic compact *cis*-dimer was resolved for human teneurin-4, which demonstrated a key role for the ABD domain in the dimer interface[7]. In this compact

conformation, the ABD domain contacts the C-rich domain, the YD-shell, and the ABD domain of the other subunit (subunit refers to one of the two chains in the covalent homodimer). Further X-ray crystallography analysis of this teneurin-4 C-rich domain has revealed that three calcium ions are coordinated by eight acidic residues in the C-rich domain. These residues are conserved in mouse and human teneurin paralogues. SAXS analysis showed that calcium binding stabilises the compact dimer conformation of teneurin-4. More recently, Li and coworkers reported a *Drosophila* teneurin (Ten-m) homodimer connected through a TTR-NHL interface[29]. This asymmetric complex was shown to be capable of forming zipper-like higher-order oligomers, most probably as linear superstructures. Two subsequent studies detailed the structural basis of the teneurin-latrophilin interaction. Both works revealed a binding site for the Lec domain of latrophilin on the YD-shell of teneurin-2, directly opposite to the ABD and Tox-GHH domains[12,30]. Together, these structural studies unveiled the intricate architecture of teneurin ectodomains, showcasing their multi-domain core superfold and revealing critical insights into their dimeric conformations and interactions, both homo- and heterophilic.

Often, alternative splicing of CAMs plays a crucial role in diversifying their functions and fine-tune cellular interactions. For instance, several synaptic CAMs, such as Dscam, the LAR-RPTP family, neurexin, and neuroligin, harbour few-residue splicing variations that change *trans*-cellular binding sites directly[31–40]. The ectodomains of teneurin proteins also contain two alternatively spliced regions that may combinatorically contribute to the code for molecular recognition at the synapse. The Ten3 splice inserts reside between EGF7 and EGF8 (AHYLDKIVK in mouse Ten3), referred to as splice insert A, and between the first and the second blade of the NHL β-propeller (RNKDFRH in mouse Ten3) referred to as splice insert B. Berns and colleagues performed cDNA sequencing of postnatal mouse brain to demonstrate that all Ten3 A and B splice variants are expressed in mouse subiculum, and all except A_0B_0 are found in the CA1 region of the hippocampus[8]. Functionally, the alternative splice variants of Ten3 were found to affect binding partner specificity[9,25,30]. Namely, mouse Ten3-A_0B_0, lacking both splice inserts, is not capable of homophilic cell clustering in K562 cells, whereas the presence of either or both splice inserts (A_1B_0, A_0B_1, A_1B_1) was reported to induce clustering in homophilic fashion[11]. Interestingly, Ten3-A_0B_0 (as well as Ten3-A_1B_1, other variants not tested) is capable of inducing heterophilic clusters with latrophilin-3[11]. Alternative splicing of teneurin-2 at similar sites has also been functionally characterised. Teneurin-2 without splice insert B (Ten2-A_0B_0) was shown to induce cluster formation with latrophilin-3-expressing HEK cells, whereas the presence of splice insert B in teneurin-2 (Ten2-A_0B_1) obstructed cluster formation with latrophilin-3-expressing cells[25,30]. Strikingly, this difference in latrophilin binding was only observed with full-length, membrane-tethered, teneurin-2 but not when teneurin-latrophilin binding was assessed using soluble ectodomain variants. This suggests a shapeshifting mechanism in which the latrophilin binding site becomes inaccessible in the presence of splice insert B due to conformational restraints, rather than slice insert B changing the latrophilin binding site directly. In addition, HEK-neuron co-culturing experiments revealed that overexpression of Ten2-A_0B_0 induces excitatory postsynaptic specialisations in neurons, whereas overexpression of Ten2-A_0B_1 induces inhibitory synapse formation[30].

It is therefore evident that the role of alternative splicing in teneurins is crucial for controlling trans-cellular interactions, yet the structural basis of splicing-dependent assembly formation remains unclear. Here, we present a structural comparison of four splice variants of mouse Ten3 full ectodomains. Using single particle cryo-electron microscopy (cryo-EM), in combination with small-angle X-ray scattering (SAXS) and negative stain EM, we reveal a homo-dimeric interface for Ten3 A_1B_0 and A_1B_1 (A_1 variants), resulting in a compact dimer conformation. This interface is stabilised by an extended β-sheet interaction between EGF8 and the ABD domain of interacting subunits. Also, the A_0B_0 and A_0B_1 variants (A_0 variants) are compact dimers in solution but they are distinct from the compact conformation for the A_1 isoforms. For the A_0 variants, EGF6 and EGF7 are additionally present in the reconstructed cryo-EM density map, where for A_0B_0 a contact site is observed between the EGF6 and the YD/ABD domains. Notably, each A_1 dimer displays a different arrangement of the two core superfolds. Finally, cell clustering and neuronal stripe assays highlight the capability of the A and B splice insert to directly mitigate *trans*-cellular interactions *in vitro*. These findings support a model wherein the subunits of the Ten3 dimers adopt different orientations due to the presence of relatively small splice inserts. These structural rearrangements, in turn, expose different binding interfaces that play a crucial role in orchestrating cellular interactions essential for specifying neuronal circuitry formation.

2.2. RESULTS

2.2.1. A_1 VARIANTS COMPACT DIMERISE THROUGH AN EXTENDED β-SHEET

To examine the structural basis of the alternative splice variant A_1B_1 of teneurin-3 (Ten3-A_1B_1), we expressed the complete extracellular domain (ECD, residues 342-2715) of N-terminal hexahistidine-tagged mouse Ten3-A_1B_1 in HEK293-E cells (Fig. 2.1A). We purified the proteins by Ni-NTA affinity chromatography followed by size-exclusion chromatography in the presence of calcium (Supplementary Fig. 2.8A). The purified ECD forms a constitutive dimer, linked by two disulfide bonds in EGF2 and -5 (Cys549 and Cys648, respectively, Fig. 2.1A). The dimer structure was resolved to 3.1 Å using single particle cryo-electron microscopy (cryo-EM), revealing a compact conformation (Fig. 2.1B, Supplementary Fig. 2.8B-D, and Supplementary Table 2.1). The core of the compact dimer could be further resolved to 3.0 Å (Supplementary Fig. 2.8D). The ECD (residues 342-2708, Fig. 2.1A) of the A_1B_0 isoform was additionally purified and reconstructed by cryo-EM, resulting in a density map with 3.9 Å overall resolution (Supplementary Fig. 2.9). In the compact conformation, the NHL domains are facing away from each other (Fig. 2.1B and Supplementary Fig. 2.9E, middle) and the two subunits are in-plane when viewed from the NHL domain (Fig. 2.1B and Supplementary Fig. 2.9E, right). The total dimensions of the Ten3-A_1B_1 compact in-plane dimer are 122 Å by 141 Å by 70 Å. The interface is mainly composed of hydrophilic and hydrophobic surface (Supplementary Fig. 2.11A).

An atomic model was built into the 3.1-Å resolution A_1B_1 dimer map. The main compact dimer interface is formed through an extended edge-to-edge β-sheet between residues 761-764 in EGF8 and residues 2594-2599 in the ABD of the two respective subunits (interface I, Fig. 2.2A-E and Supplementary Fig. 2.10A). This interface is only per-

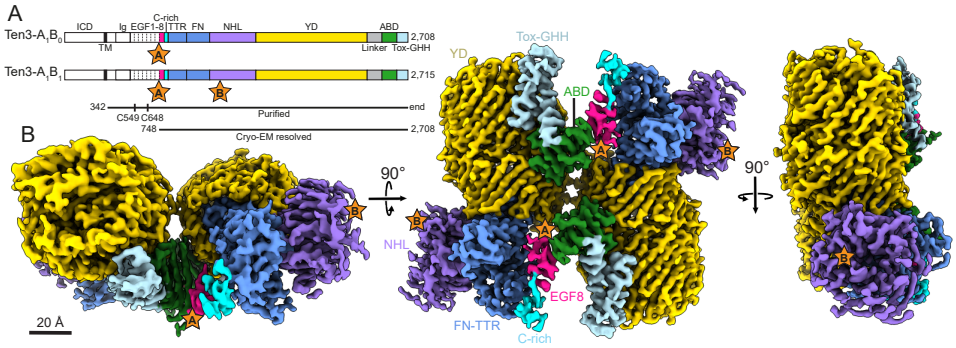


Figure 2.1: Teneurin-3 A_1B_1 core superfolds form a compact dimer. **A)** Linear representations of the full-length teneurin-3 A_1B_0 and A_1B_1 domain compositions. Splice inserts are indicated as stars: splice insert A in between EGF7 and EGF8; splice insert B in the NHL domain. The expressed and purified ectodomain for both isoforms are indicated with the black line below. Free cystines in EGF2 (C549) and EGF5 (C648) that enable constitutive dimerization are indicated. The bottom black line indicates domains (EGF8 through Tox-GHH) resolved for the A_1B_1 isoform using cryo-EM. ICD, intracellular domain; Ig, immunoglobulin fold; EGF, epidermal growth factor repeat domains 1 through 8; C-rich, cysteine-rich region; TTR, transthyretin-related; FN, fibronectin plug; NHL, NCL, HT2A and Lin-41; YD, tyrosine-aspartate; ABD, antibiotic-binding domain; Tox-GHH, toxin-glycine-histidine-histidine **B)** Three different views of the density map of the teneurin-3 A_1B_1 compact dimer at 3.1 Å resolution (EMD-18889). Domain annotations are coloured corresponding to the linear representation in A. Positions of the splice inserts are indicated with stars. Scale bar is 20 Å.

mitted upon displacement of a β-strand on the ABD side, which remains present in a non-compacted subunit reconstruction (Fig. 2.2C, and Supplementary Fig. 2.11B-D). Interestingly, a mutation associated with microphthalmia in humans, namely R2563W in Ten3-A_0B_0 [17] – corresponding to R2579W in A_1B_1 –, is located directly at the border of the missing β-strand (Fig. 2.2C). To further stabilise the extended β-sheet, S758-T2600 side chains undergo hydrogen bonding, and R761 forms a salt bridge with D2611 and a hydrogen bond with S2598 (Fig. 2.2D). The EGF8-ABD interface is also supported by hydrophobic interactions between W769 and L764 on EGF8, and L2575, L2578, and V2595 on ABD (Fig. 2.2D-E). Splice insert A residues V748-K749 are resolved at the N-terminal side of EGF8 (Fig. 2.2D), and focussed classification of all dimer particles results in extra connected density that folds back onto the ABD domain, which presumably represents the remainder of the insert (Supplementary Fig. 2.8D). The compact dimer is further stabilised by YD-ABD (interface II, Fig. 2.2A and F, and Supplementary Fig. 2.10B) and YD-YD (interface III, Fig. 2.2A-B, G, and Supplementary Fig. 2.10C) contacts. Interface II contains T1607-N2589 and K1612-N2534 hydrogen bonding (Fig. 2.2F). Interface III contains a T1604-G1618 hydrogen bond (Fig. 2.2G). Individually, the interface-contributing residues in EGF8 span a total interface area (IA) of 420 $Å^2$, those in the YD-shell 476 $Å^2$, and 657 $Å^2$ in the ABD domain (Fig. 2.2B). A separate subset of dimer particles harbours no YD-YD interactions, resulting in relatively flexible YD domains hinging at the C-rich/EGF8 region ("open conformation"). The A_1B_0 (Supplementary Fig. 2.9E) compact dimer is found in this open conformation – lacking interface III – as opposed to the A_1B_1 closed compact dimer (Fig. 2.1B) that includes this additional YD-YD contact.

What are the differences between the mouse teneurin-3 compact dimer presented here and the human teneurin-4 (Ten4) compact dimer published previously [7]? The main

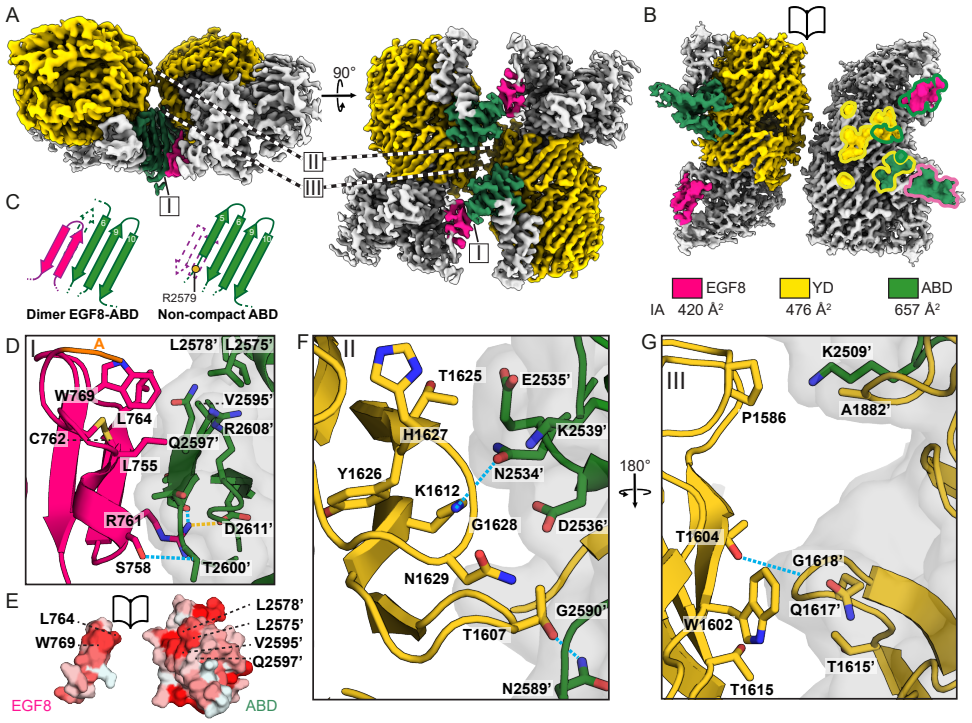


Figure 2.2: The teneurin-3 A_1B_1 compact dimer interface comprises three distinct interaction sites. **A)** Density map of the Ten3-A_1B_1 compact dimer with domains that participate in the interfaces shown in colour. Colour code is shown below panels B, along with the interface area (IA) per domain. **B)** Open book representation of the density map. The left subunit shows the participating domain in colour, and the right subunit displays the specific interface residues. Fill colours indicate the domains in which the residues are located, and outline colours indicate the domains in which the contacting residues reside. **C)** Schematic representation of EGF8-ABD interface in the compact dimer versus the ABD domain of a Ten3 A_1B_1 non-compact subunit reconstruction. Residues 2579-2588 are displaced or disordered in the compact dimer form and residues 2582-2587 form β-strand 5 in ABD of the non-compact subunit. **D)** Close-up of interface I, the EGF8-ABD extended β-sheet (PDB:8R50). Resolved residues from insert A (labelled A) are coloured orange **E)** Open book surface hydrophobicity representation of the EGF8-ABD interface. **F)** Close-up of interface II, comprising YD-ABD contacts. **G)** Close-up of interface III, the YD-YD and YD-ABD contacts. In the close-ups, ionic bonds are indicated in dashed yellow lines and hydrogen bonds in dashed blue lines, and transparent surface representation is used to distinguish between subunits within the compact dimer. Only residues with >5% per-residue buried surface area are represented. For clarity, L2530, S2596, and R2624 are not displayed.

difference between the Ten3 and Ten4 structures is the orientation of the two subunits with respect to each other (Fig. 2.3A). The subunits in the human Ten4-A_1B_1 dimer are each turned outwards at a 30° angle, thereby creating a 'transverse dimer' (Fig. 2.3A). The subunits in the compact dimer configuration of mouse Ten3-A_1B_1, however, align in the same plane (Fig. 2.1B). In both structures, the ABD domain is central in the interface, but different interactions are formed. In the human Ten4 transverse dimer, the ABD domain contacts the C-rich domain, the YD shell and the ABD domain of the other subunit. In the dimer of mouse Ten3, the ABD domain contacts EGF8 and the YD shell (Fig. 2.3B). The difference in dimer conformation results in a molecule with smaller dimensions for mouse Ten3 compared to human Ten4 (122 Å by 141 Å by 70 Å versus 133 Å by 166 Å by 112 Å, respectively). A detailed structural comparison reveals that simi-

lar residues on the ABD domain and the YD shells are involved in both the transverse versus and in-plane dimer interactions (Fig. 2.3C). Interestingly, a centrally positioned ABD loop in Ten4 interface – could not be resolved in the in-plane Ten3 compact dimer structure (missing N2531-A2533, Fig. 2.2F). Altogether, our cryo-EM density map reveals a compact conformation of covalently dimerised mouse Ten3, which is different from previously published dimeric conformations of teneurin homologues [7,24,29].

2.2.2. SPLICE INSERT A AS A POTENTIAL SPACER TO DISLODGE EGF6

To investigate the effect of splice insert A on the architecture of the EGF-repeat domains, we expressed and purified the full ectodomains and acquired cryo-EM data of all splice variant combinations of mouse Ten3: Ten3-A_0B_0, A_0B_1, A_1B_0 and A_1B_1 (Fig. 2.4A, Supplementary Fig. 2.12A-C, and Supplementary Table 2.1). Full ectodomain structures of compact dimers could not be resolved for the A_0 variants, presumable due to air-water interface issues during cryo-EM grid preparation. Similar issues were also observed previously for the Ten4 dimer [7]. Instead, we resolved non-compacted subunits containing partial EGF-repeat domains for each of the proteins (Fig. 2.4B and Supplementary Fig. 2.12D). Comparison of the four structures reveals that the subunits adopt similar conformations with largest differences in the orientation of EGF6 and EGF7. In absence of splice insert A (A_0B_0 and A_0B_1 isoforms), EGF-repeats 6-8 run across the FN domain towards the YD shell, whereby EGF6 is located in a groove between the YD shell and ABD domain of the same subunit (Fig. 2.4C). EGF6 directly contacts several negatively charged residues on the YD β-barrel an2506 in the ABD domain. However, in the presence of splice insert A (A_1B_0 and A_1B_1 isoforms) – predicted to be a 9-residue helix inserted in the linker between EGF7-8 -, EGF6 and EGF7 cannot be resolved anymore, suggesting increased flexibility of EGF6 with respect to the YD shell. Possibly this splice insert acts as a physical spacer to dislodge EGF6 out of its binding groove, hereby preventing the EGF6-YD/ABD interface to form. Low-resolution EM density of the potential splice insert A can indeed be observed hovering over the ABD domain (Supplementary Fig. 2.8D), thus dislodging EGF6 from its YD/ABD groove. The increased rigidity of EGF6 and EGF7 with respect to the YD shell in the absence of splice insert A appears to hinder EGF8 from interacting with the ABD domain on the other subunit (Fig. 2.4D-E) – an interaction necessary for the compact dimerisation observed in A_1B_1 (Fig. 2.2C-E). Thus, depending on splice insert A, the EGF-repeat region reorganises relative to the core superfolds, with potential implications for the orientation of the two subunits themselves, and further homo- or heterodimeric interactions.

2.2.3. ALL ISOFORMS ARE COMPACT IN A CALCIUM-DEPENDENT MANNER

To further investigate the splicing-dependent effects on the conformation of the Ten3 dimers in solution, we performed small-angle X-ray scattering (SAXS, Fig. 2.5A-B and Supplementary Fig. 2.13A-D). In presence of calcium, all isoforms are compact in solution (Fig. 2.5A-B, and Supplementary Fig. 2.13A-D), with average radii of gyration (Rg) 7.3, 8.0, 7.7, and 7.9 nm, for Ten3-A_0B_0, A_0B_1, A_1B_0, and A_1B_1, respectively (Fig. 2.5A). This relation in compactness is additionally reflected in the D_{max} values per isoform (Fig. 2.5B). The absence of calcium, as well as the addition of calcium chelator EDTA, result in increased Rg and D_{max} for all isoforms, as well as a loss of the distinct compact-protein

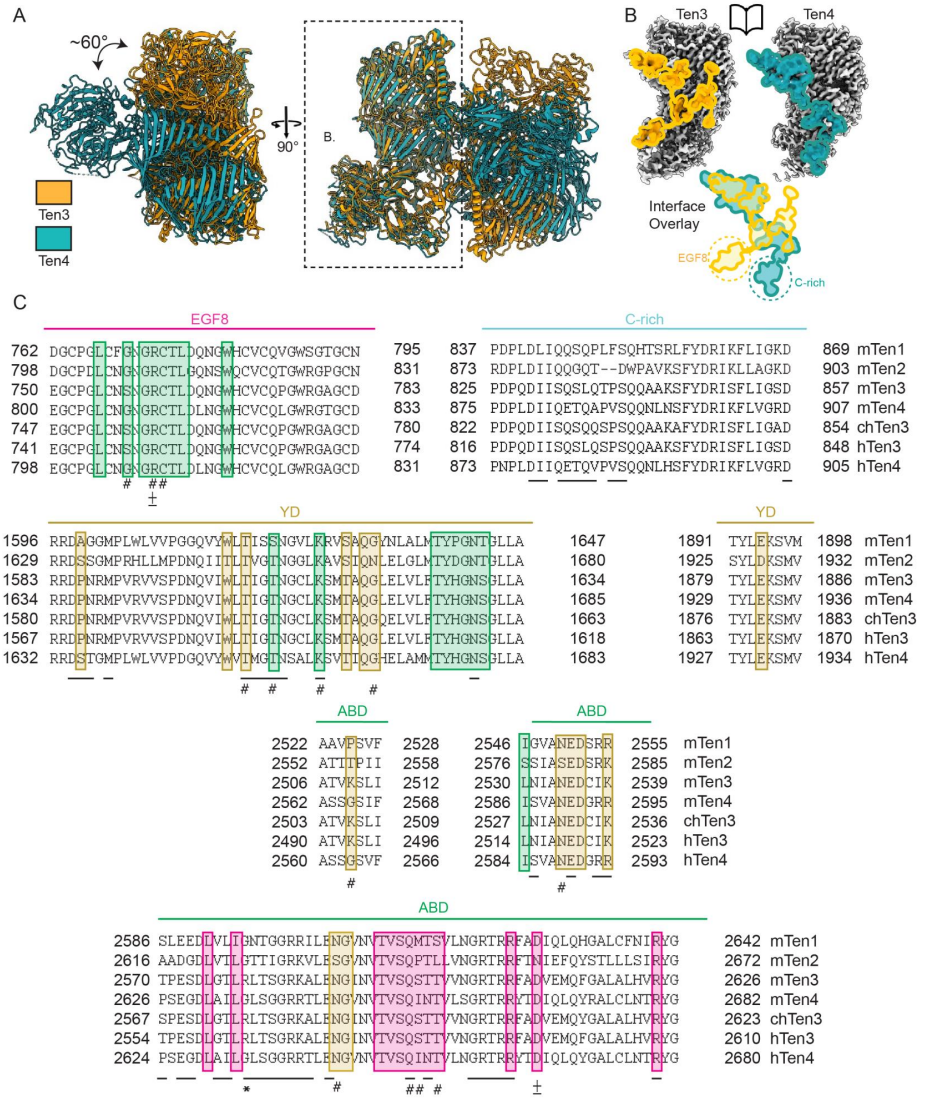


Figure 2.3: Teneurin-3 and teneurin-4 compact dimer interfaces are different. **A)** Alignment of one subunit from the teneurin-3 (Ten3, orange) compact dimer conformation with one subunit from the teneurin-4 (Ten4, PDB:7BAM, blue) dimer. The remaining subunits from each dimer orient at a relative angle of approximately 60°. Dashed outline indicates the subunit represented in panel B. **B)** Comparison of the mouse Ten3 and human Ten4 compact dimer interface projections on open-book subunits. Below: overlays of the interface projections. Dashed circles indicate the domains that exclusively participate in one or the other compact dimer. **C)** Teneurin alignment across orthologues of the interface residues for the EGF8, C-rich, YD and ABD domains. Boxes around the residues indicate the contacting domains in mouse teneurin-3. Domain colouring corresponds to that in Fig. 1. Only residues with >5% per-residue buried surface area are represented. ±, ionic-bonding residues in Ten3; #, hydrogen-bonding residues in Ten3; _, interface residues in Ten4; *, R2579, the A_1B_1 equivalent of microphthalmia-associated R2563 on human Ten3-A_0B_0.

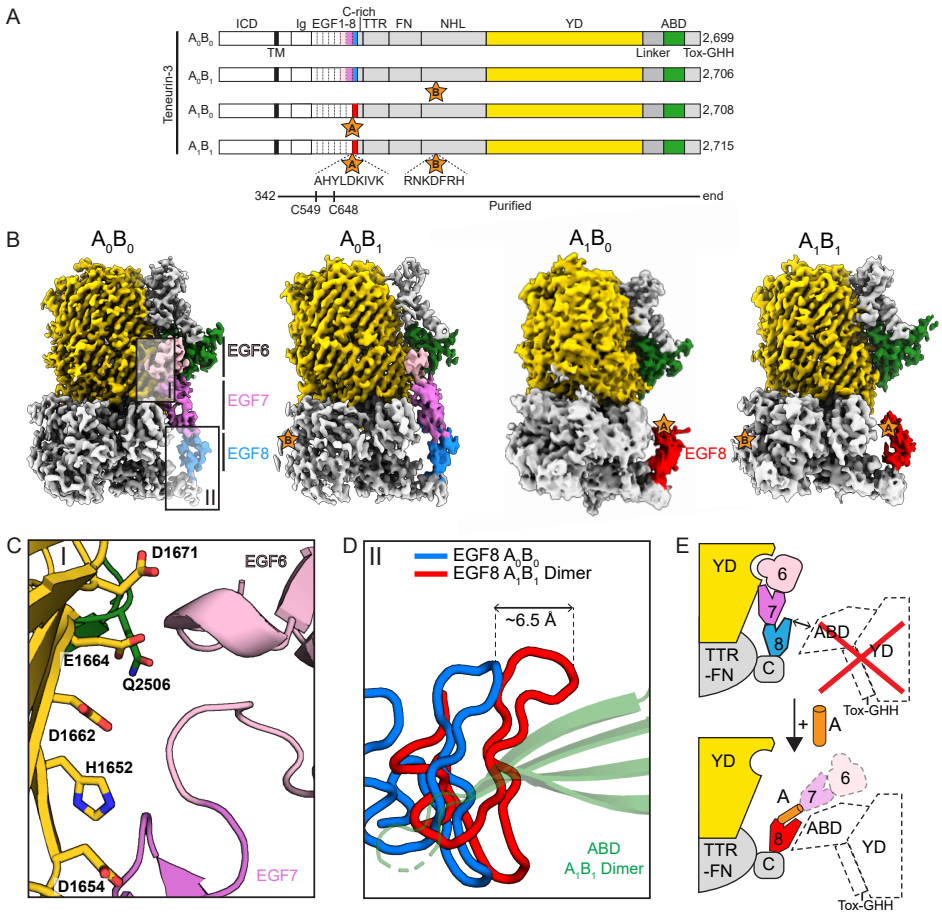


Figure 2.4: Subunits of A_0 isoforms display a contact between EGF6 and the YD/ABD domains. **A)** Linear representations of the four full-length Ten3 isoforms with domain composition. Presence of splice inserts is indicated with stars. Domains are labelled according to Fig. 1. The expressed and purified ectodomain for each isoform is indicated with the black line below. **B)** Cryo-EM reconstructions of isoform subunits (left to right: EMD-18891, EMD-18900, EMD-18902, EMD-18890). Dashed boxes I-II indicate close-ups shown in panels C-D, respectively. **C)** EGF6-YD/ABD interface (PDB:8R54). Interacting residues on the YD- and ABD-side of the interface are displayed and annotated. **D)** Overlay of EGF8 domains of Ten3-A_0B_0 and A_1B_1. The relative displacement of EGF8 is indicated. **E)** Schematic representation of dislodging by splice insert A. Top: Absence of A is accompanied by an additional interface between EGF6-YD/ABD. This stabilisation impedes EGF8 β-sheet extension with ABD for compact dimerisation. Below: Splice insert A creates distance between EGF7 and EGF8, in turn dislodging the EGF6-YD interface. Loss of this interface and the presence of splice insert A increases the flexibility of the EGF region and EGF8 is now free to interact with the opposite ABD domain.

profile in the dimensionless Kratky plots and pair-distance distributions (Supplementary Table 2.2, Supplementary Fig. 2.13A-D). These effects are similar to the transition of Ten4 from a compact conformation into a more elongated conformation in the absence of calcium[7]. Notably, in the presence of calcium, out of the four splice variants A_0B_0 is most compact, whereas A_0B_1 deviates most from a compact form (Fig. 2.5A-B). The increased compactness of A_0B_0 could possibly be explained by a predicted inter-

action between any of the flexible Ig or EGF domains and the splice site B within the NHL domain (Supplementary Fig. 2.14), that would condense the A_0B_0 compact dimer even further. The data further show that Ten3-A_1B_0 and A_1B_1 have highly similar SAXS profiles. Thus, it appears that splice insert B affects compact dimerisation only in absence of splice insert A. The general compactness in presence of calcium is supported by increased protein stability observed as a thermal shift in the protein melting curves of all splice variants (Fig. 2.5C and Supplementary Fig. 2.13E). Note that calcium-binding residues of the C-rich domain of human Ten4 are well-conserved in mouse Ten3 (Fig. 2.3C), suggesting that the C-rich domain of Ten3 also selectively binds calcium ions.

To corroborate the SAXS profiles indicating compact conformations for all splice variants, we collected negative stain transmission electron microscopy (nsTEM) datasets for all variants. The 2D classifications of all Ten3 splice variants indeed revealed compact conformations for each variant (Fig. 2.5D and Supplementary Fig. 2.15A). Whereas Ten3-A_1B_0 and A_1B_1 have the same conformations as their cryo-EM structures (Fig. 2.1 and Supplementary Fig. 2.9E), A_0B_0 and A_0B_1 revealed alternative architectures. While the low resolution of the data precludes detailed analysis, a potential YD-YD interface is observed for A_0B_0 – also observed at low abundance in cryo-EM 2D classes (Supplementary Fig. 2.15B) -, and the A_0B_1 compact dimer could be explained by an FN-FN interaction. 3D reconstructions and low-resolution fitting confirmed these potential dimer interfaces (Fig. 2.5E). Map-model correlations were calculated for all map-model combinations, and highest correlations were found between model-derived densities and the corresponding experimental densities into which the model was built (Supplementary Fig. 2.15C). In addition, the model for A_1B_1 correlates well with the map of A_1B_0 and vice versa, which is in line with their similarities in cryo-EM and SAXS analysis. Splice site B is not directly involved in any of the observed compact dimer interfaces, hinting at an unresolved allosteric mechanism that could explain the shifts between A_0B_0 and A_0B_1 conformations. Notably, an additional dimer interface was observed in the nsTEM dataset of the A_0B_0 variant, in which the NHL domains are connected (Supplementary Fig. 2.15D). This conformation would, however, not be compatible with a compact dimer due to restraints imposed by the EGF2-2/5-5 disulfide bonds and may represent an additional higher-order *cis* contact.

We next used the SAXS profiles and compact dimer interfaces from nsTEM to model the structures of the full ectodomain of all the splice variants (Supplementary Fig. 2.15E). For SAXS modelling, the dimer interfaces obtained from nsTEM were restrained, whereas the domains in the EGF-Ig stalk were treated as independent connected rigid bodies, with EGF2-4 restrained to enforce the intermolecular disulfide bonds. All modelled splice variants, except for the A_0B_1 dimer, explain their SAXS profiles well (Supplementary Fig. 2.15E). Possibly, the SAXS data of A_0B_1 cannot be explained by the nsTEM-based model, because A_0B_1 in solution does not exhibit the same level of compactness or adopts multiple conformations. Nonetheless, the combination of the nsTEM and SAXS data indicates at least three different dimer conformations, with A_1B_0 and A_1B_1 sharing the same EGF8-ABD mediated dimer, also observed in their cryo-EM structures, and splice variants A_0B_0 and A_0B_1 each adopting a distinct dimer conformation.

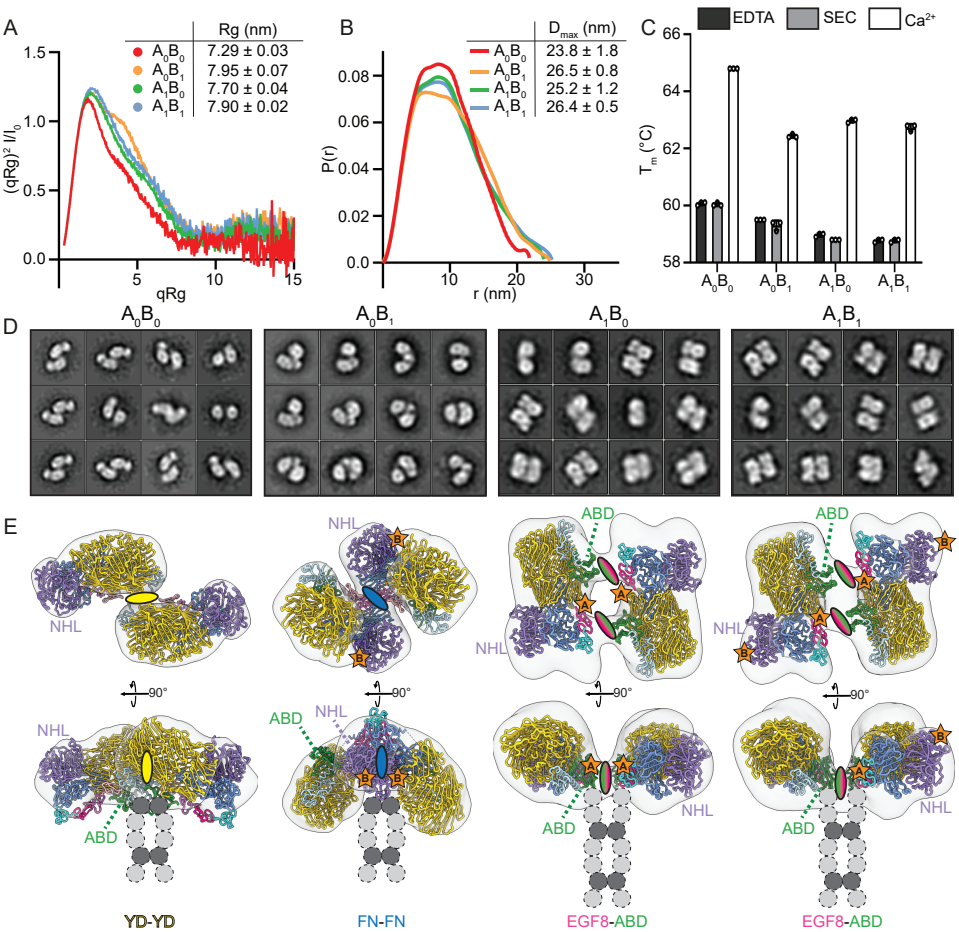


Figure 2.5: All teneurin-3 isoforms are compact in solution with isoform-specific conformations. **A)** Dimensionless Kratky-plot of SAXS data for all Ten3 isoforms at 0.5 mg/mL in the presence of 2 mM Ca^{2+}. Inset, radii of gyration (Rg) derived from SAXS data for each Ten3 isoform averaged from three protein concentrations (mean ± SEM). **B)** Pair-distance distribution for all Ten3 isoforms at 0.5 mg/mL in the presence of 2 mM Ca^{2+}. Inset, maximum diameter (D_{max}) derived from the pair-distance distribution for each Ten3 isoform averaged from three protein concentrations (mean ± SEM). **C)** Melting temperatures (T_m) for each isoform in the presence (Ca^{2+}) and absence (SEC) of 2 mM calcium, as well as under addition of 5 mM EDTA (EDTA). Data are presented as mean ± SD; n=3 melting temperatures per N=1 independent experiment. **D)** 2D classes of negative stain TEM for all Ten3 isoforms in the presence of calcium. **E)** 3D reconstructions for all splice variants, shown at two different angles. Domains are coloured according to colouring in Fig. 1. Ovals indicate location of the compact dimer interface, and the colour denotes the involved domain(s). Attachment of the EGF repeats that are not included in the map or model are added schematically as dashed circles. Disulfide-linked EGFs 2 and 5 shown in dark grey.

2.2.4. A MODEL FOR SPLICING-DEPENDENT *trans* INTERACTIONS

The alternative conformations as shown in Fig. 2.5 may affect the capabilities of Ten3 splice variants for *trans*-cellular interactions. To test this, we electroporated GFP-tagged full-length transmembrane variants of all four Ten3 splice variants in K562 hematopoietic cells. *Trans*-cellular interactions were quantified by assessing cluster formation of

the Ten3 electroporated K562 cells (Fig. 2.6A-B and Supplementary Fig. 2.16A). Similar to previously shown[11], all isoforms except A_0B_0 form homophilic clusters. In addition, we observe a quantitative difference between clustering of Ten3-A_0B_1, A_1B_0 and A_1B_1, whereby A_1B_0 and A_1B_1 have only limited clustering capabilities, while electroporation with A_0B_1 results in the highest clustering index and cluster size averages (Fig. 2.6B). As such, similar to the structural data (Fig. 2.5), splice insert B has a large effect, but exclusively when splice insert A is absent.

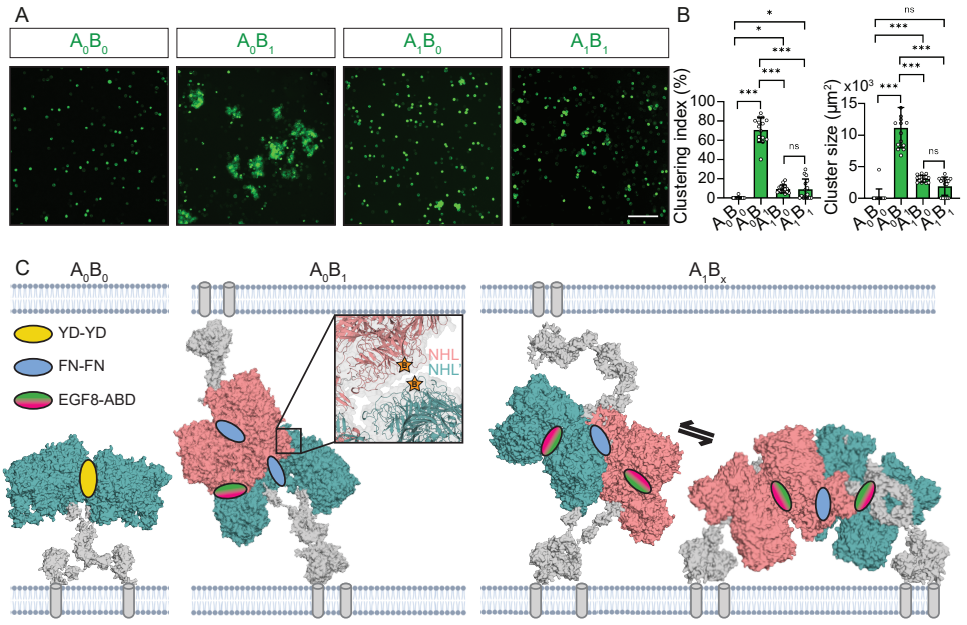


Figure 2.6: Models for teneurin-3 homophilic *trans*-cellular complex formation. **A)** Clustering assay of K562 hematopoietic cells electroporated with the mouse Ten3 isoforms. Scale bar is 100 µm. **B)** Analysis of the clustering indices and mean cluster area of the conditions in A. Data points for cluster size correspond to the per-image averaged values (n=5 images per N=3 independent experiments, 15 images total). Data are presented as mean ± SEM; ns: not significant, *: $p < 0.05$, ***: $p < 0.001$; P-values for clustering index comparisons from top-down are p=0.0312, p=0.0253, p<0.0001, p<0.0001, p<0.0001 and for cluster size comparisons from top-down are p=0.0983, p=0.0008, p<0.0001, p<0.0001, p<0.0001. One-Way ANOVA, Tukey's multiple comparisons test. **C)** A non-clustering Ten3-A_0B_0 compact dimer, a *trans*-cellular dimer-of-dimer of Ten3-A_0B_1, and a *trans*-cellular dimer-of-dimer of Ten3-A_1B_x. The A_0B_x dimer-of-dimers is also compatible with a cis configuration due to the EGF organisation. Close-up of the A_0B_1 *trans* dimer-of-dimers display an additional contact between the NHL domains of subunits not involved in the EGF8-ABD contact. Dimers located on opposing membranes are distinctly coloured salmon and dark teal. The rigid-body models of each of the EGF-Ig stalks were calculated using the SAXS data of each isoform (Fig. 5 and Supplementary Fig. 8E). Transmembrane helices are indicated by grey tubes.

No *trans* interaction is observed in the cell clustering assay for Ten3-A_0B_0 (Fig. 2.6C). Possibly, the A_0B_0 may not engage in *trans*-cellular interactions because it forms homophilic complexes in *cis*, through the above-mentioned NHL-NHL contact (Supplementary Fig. 2.15D). To model higher-order oligomers that may explain the observed *trans*-cellular interactions for the other variants, we analysed dimer-of-dimer formation of the SAXS- and nsTEM-based full ectodomain models. For compact dimer Ten3-A_0B_1, harbouring a potential FN-FN interface, the ABD domain would be accessible to

form a *trans* dimer-of-dimer interaction with the EGF8 domain of a dimer at the opposing membrane (Fig. 2.6C). This interface could not be employed for dimer-of-dimers in *cis* due to increased constraints induced by the EGF6-ABD/YD interaction. Remarkably, the A_0B_1 *trans* dimer-of-dimer displays an additional NHL-NHL interface between them (close-up, Fig. 2.6C), with splice inserts B located at this interface. This interface is reminiscent of the teneurin-2 crystal contact reported previously[24]. Compact dimers A_1B_0 and A_1B_1 could form *trans*-cellular interactions by means of the FN-FN interface (Fig. 2.6C). This same interface could also be employed for an additional *cis*-interaction (Fig. 2.6C). A presumed equilibrium between *cis*- and *trans* dimers of dimers would explain the limited cell-clustering capability of the A_1B_0 and A_1B_1 dimers, as *cis* dimers-of-dimers would harbour no potential for *trans* interactions. The distinct conformations of compact dimers may also have implications for their ability to bind *trans*-cellularly with latrophilin[12,30], and with FLRT in *cis*[41]. As illustrated in Supplementary Fig. 2.16B, A_0B_1 may not readily engage in *trans* interaction with latrophilin, given that the latrophilin-interaction surface appears to be oriented towards the membrane, obstructing access to the Olfactomedin-Lectin domain. The other three splice variants seem to be more conducive to latrophilin binding. This observation aligns with earlier findings that indicated the incapability of a teneurin-2 A_0B_1 version to bind latrophilin, and thus provides a structural basis for the proposed shapeshifting mechanism[8,11,30]. None of the membrane-tethered compact dimer conformations occlude the latrophilin and FLRT binding sites (Supplementary Fig. 2.16B). However, FLRT may not engage in binding A_0B_1 if the teneurin-latrophilin *trans* interaction is sterically obstructed by the intermembrane setting[41].

2.2.5. INSERT B ATTRACTS TENEURIN-3-EXPRESSING OUTGROWING AXONS

We hypothesised that the emergent NHL-NHL contact in Ten3-A_0B_1 (close-up Fig. 2.6C) could provide additional *trans*-cellular stabilisation in a B-dependent manner. To address this hypothesis and whether this contact by itself provides sufficient adhesive strength, we set up chick retinal stripe assays[42,43]. First, we characterised Ten3 expression patterns in the chick retina, using RNA fluorescent *in situ* hybridisation (FISH). We observed that Ten3 mRNA is highly expressed in the ventral, but not dorsal, retina. (Fig. 2.7A). Then, to address whether Ten3 acts as an attractant or repellent for neurons from either retinal regions, we plated retinal chick explants from both regions and grew them on Ten3 stripe assays. To directly evaluate the impact of splice insert B on *trans*-cellular interactions, we utilised a monomeric variant of mouse Ten3 lacking the Ig and EGF domains (lacking residues 1-845). This specific monomer is incapable of establishing EGF8-ABD interactions, shown in this work to restrict the influence of splice insert B. Furthermore, the designed Ten3 monomeric variant cannot participate in the A_0B_1 EGF8-ABD-based dimer-of-dimer *trans* interaction, thereby isolating the contribution of B in direct *trans*-cellular binding. At 2-3 days after explant growth, it was observed that ventral neurons localised preferentially to the Ten3-B_1-covered regions with their axons growing parallel to the stripes, whereas the dorsal neurons covered the entire surface with axons growing in random directions (Fig. 2.7B-D). No such difference was observed when culturing ventral or dorsal neurons on Ten3-B_0 stripes (Fig. 2.7E-F). Notably, when only labelled and unlabelled control antibody was immobilised in the stripe assay, nei-

ther the ventral nor the dorsal neurons showed any preferential growth over particular stripes. (Fig. 2.7C-F). The K562 and retinal explant studies together show that, while the presence of splice insert A may allow for *trans*-cellular clustering in a manner unaffected by B, the presence of splice insert B itself can support and induce robust *trans*-cellular interactions and serves as an attractant to chick retinal neurons from the ventral region.

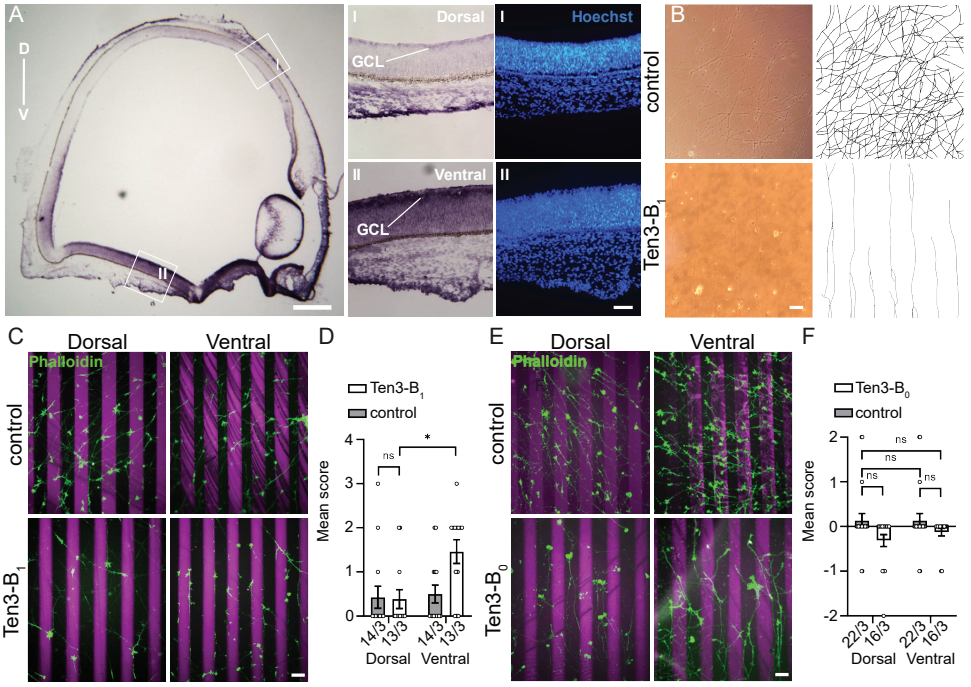


Figure 2.7: Splice insert B directly guides axons. **A)** *In situ* hybridisation on E6 chick retinal cryosections showing chick Ten3 mRNA expression in a gradient along dorso-ventral (D-V) axis. GCL, ganglion cell layer. Scale bar is 400 µm. I. Close up of dorsal (D) region of the retina showing little to no expression of Ten3, and Hoechst staining of nuclei (blue) in D retina. II. Close up of ventral (V) region of the retina showing high expression of Ten3, and Hoechst staining of nuclei (blue) in V retina. Scale bar is 50 µm. Three retinae were imaged with similar results. **B)** Representative light microscope image of V RGC axons cultured on control versus control (top) or Ten3-B_1 versus control (bottom) stripes. Scale bar is 50 µm. Right: Tracing of axons in the light microscopy images. Statistics identical to panel D. **C)** Control stripe assays with cultured retinal explants from D and V retinas. Scale bar is 100 µm. Bottom: Ten3 stripe assays with 10 µg/mL Ten3-B_1 (magenta), with retinal explants from D and V retina. **D)** Quantification of RGC axon guidance from V vs D retina in panel C. Guidance decisions of panels D and F were scored between -3, -2, -1, 0, 1, 2 or 3, reflecting varied degrees or repulsion or attraction; data of panels D and F are presented as mean ± SEM; numbers under each bar of panel D and F represent n/N, with n scored images per N independent experiments; ns: p=0.9992 , *: p=0.013. Two-Way ANOVA, Tukey's multiple comparisons test. **E)** Control stripe assays with cultured retinal explants from D and V retinas. Stripes were labelled with goat anti-Rabbit Alexa-568 secondary antibody and RGC axons labelled with Phalloidin-Alexa Fluor 488. Scale bar is 100 µm. Ten3 stripe assays with 10 µg/mL Ten3-B_0 (magenta), with retinal explants from D and V retina. **F)** Quantification of RGC axon guidance from V vs D retina in panel E. P-values from top-down left-right are p=0.5794, p>0.999, p=0.1025, p=0.5794; ns: not significant. Two-Way ANOVA, Tukey's multiple comparisons test.

2.3. DISCUSSION

Cell adhesion molecules (CAMs) play a prominent role in the process of neuronal wiring by establishing a molecular recognition code at the synapse through their combina-

torial expression pattern. Teneurins are a family of synaptic CAMs that exhibit additional *trans*-synaptic specificity through alternative splicing of two sites, referred to as splice site A and splice site B. Here, we reveal a compact dimer conformation of the ectodomains of mouse Ten3-A_1B_0 and A_1B_1. The compact dimer interface is formed through an extended β-sheet interaction between the EGF8 and ABD domains on interacting monomers. Additionally, the YD-shell and ABD domain participate in further stabilisation. We resolved the cryo-EM subunits of each splice variant, and these data demonstrate an additional interaction between EGF6 and the ABD/YD domain, only when splice insert A is absent. We further provide small-angle X-ray scattering (SAXS)-based evidence, complemented with negative stain EM, that all splicing variants are compact in presence of calcium. This data corroborate that A_1B_0 adopts the same conformation as A_1B_1, while A_0B_0 and A_0B_1 display different compact conformations. These large rearrangements of the two subunits have practical implications for *trans*-cellular interactions, as demonstrated in the hematopoietic cell line K562 and chick retinal neurons. Our data provide a model where relatively short splice inserts of ~7 amino acids result in large structural rearrangements, hereby impacting *trans*-cellular interactions and neural circuit wiring.

Trans-synaptic adhesion is commonly regulated by alternatively spliced variants, often directly affecting the *trans*-cellular contacts[11,44–47]. For instance, the neuroligin-neurexin complex is regulated by the presence of two neuroligin splice inserts (A, 30 residues; B, 8 residues) that directly occlude neurexin binding[31]. Wilson and colleagues further revealed that the presence of a neurexin splice insert (SS2A, 8 residues) elongates a β-sheet near the interface with neurexophilin, directly stabilising this complex[38]. Similarly, two so-called micro-exons A and B (meA, 9 residues; meB, 4 residues) strongly determine *trans*-synaptic binding for PTPδ, PTPσ, and LAR, three members of the LAR-RPTP family[46]. The inserts were shown to either directly remodel the *trans*-synaptic interface[31–39], or - in the case of meB as a linker between Ig2 and Ig3 - by increasing the flexibility of Ig3 relative to Ig1/Ig2[40], thereby extending the interface. These inserts then determine selective binding of PTPδ to IL1RAPL1 and IL-1RAcP[40], SLITRK1/2[37,39], and SALM5[34], and the interactions of PTPσ with SALM3[48] and TrkC[49]. Similarly, cell-specific expression of Dscam splice variants – with a total of 38,016 possible Dscam isoforms - underlies dendritic self-avoidance[50], and various structural works have shown how Dscam splicing directly remodels the homophilic binding interface to orchestrate different *cis* and *trans* interactions[32,33,35,36]. These findings are all examples in which the *trans*-cellular binding interface is directly modulated by small splice variations.

In contrast, our data illustrate how alternative splicing of teneurin-3 results in large rearrangements of the individual subunits, referred to as shapeshifting by Li et al[30], forming different compacted dimers, and thereby exposing new binding sites. Specifically, the presence of splice insert A increases the flexibility of the EGF repeat region, thereby facilitating EGF8-ABD compact dimers to form. Additionally, the presence of splice insert B facilitates an FN-FN compact dimer, but only in absence of splice insert A. Our data do not explain why Ten3-A_0B_0 and Ten3-A_0B_1 are structurally different. Although splice insert B locally only modifies the loop in which it is located (Supplementary Fig. 2.14A), we speculate that the NHL domain (including splice site B) may be engaged in a direct contact with any of the Ig or EGF domains (Supplementary Fig.

2.14B-D). This would impose or remove additional restraints resulting in FN-FN compact dimers for the Ten3-A_0B_1 variant. Importantly, these different compact dimer conformations expose different surfaces available for dimer-of-dimer formations in *trans*. For instance, the EGF8-ABD compacted dimers are exposing FN domains for FN-FN *trans* interactions. Conversely, the FN-FN compact dimers are exposing the EGF8 and ABD domains for *trans* and *cis* dimers-of-dimers.

In recent years, it has become evident that the pre- and postsynaptic membranes are tightly organised[51]. For instance, postsynaptic receptors are aligned with presynaptic vesicle release sites, hereby forming 'nanocolumns', and potentially enhancing the efficiency of signal transfer. Cell adhesion molecules may also play a significant role in this process by directly interacting with receptors to orchestrate the molecular organisation at the synaptic membrane. Indeed, few studies have now shown that cell adhesion molecules are also present in nanocolumns, exemplified by neurexins[52,53]. More recently, using STORM super-resolution microscopy, researchers discovered that teneurin-3 is also organised in nano-size clusters, of approximately 80 nm in size, localised at the presynaptic side[54]. It would be interesting to explore further what is required to form such 80 nm clusters, as our *cis*- and *trans* interactions were only compatible with linear arrays (data not shown). This outcome resembles what is predicted from the asymmetric dimeric contact observed in fly teneurin[29]. Presumably, formation of such clusters may require the presence of additional cell adhesion molecules, membrane-embedded proteins, lipids, or interacting receptors.

Taken together, an underlying principle is emerging regarding the relationship between alternative splicing and neural connectivity. Teneurin's splicing-dependent conformational reorganisation itself presents a unique example of large rearrangements causing indirect remodelling of *trans*-synaptic interactions. These splicing-dependent conformation acrobatics shed light on how small changes in protein sequence realise large conformational shifts at the molecular level, and orchestrate neuronal circuitry formation up to the scale of entire nervous systems. Future studies will have to address the cell-specific expression of such splice variants and the consequences for neuronal wiring specificity.

2.4. METHODS

PLASMIDS

The ectodomain and full-length version of Ten3-A_0B_0 (MGC premier cDNA clone BC145284, Biocat GmbH, BamH1 sites removed with silent mutations) were cloned into the pUPE106.03 and pUPE3820 vectors (U-protein Express), containing a N-terminal cystatin secretion signal followed by a hexa-His tag or a N-terminal GFP tag, respectively, by using BamHI and NotI restriction sites. Ectodomains and full-length splice variants of Ten3-A_0B_1, Ten3-A_1B_0, Ten3-A_1B_1 were constructed from full-length Ten3-A_0B_0 using PmlI-AgeI, EcoRI-PmlI and EcoRI-AgeI restriction ligation cloning. Monomeric mouse Ten3 (residues 846-2715) for stripe assays as in Jackson *et al.* 2018[24].

PROTEIN EXPRESSION AND PURIFICATION

Epstein–Barr virus nuclear antigen I-expressing HEK293 cells (HEK-E; U-Protein Express) were used to express all secreted Ten3 ECD isoforms. Cells were cultured in FreeStyle293 expression medium with GlutaMAX (FreeStyle; Gibco) supplemented with 0.2% fetal bovine serum (FBS, Gibco) and 0.1% Geneticin (G418 Sulfate; Gibco), in a shaking incubator at 37°C and 5%. Prior to transfection, HEK-E cells were seeded at 0.25-0.32·10^6 cells/mL onto 1 L Erlenmeyer cell culture flasks in FreeStyle medium without supplements. 24 h later, cells were transfected with a total of 125 µg DNA encoding the full ECD per 250 mL of cells using polyethylenimine (PEI, 1:3 DNA:PEI ratio; Polysciences), according to the manufacturer's protocol, and treated with 5.5% Primatone in Freestyle medium 6-24h post-transfection. Six days after transfection, the medium was collected by pelleting the cells through centrifugation at 300 g for 10 min. The medium was once again centrifuged at 4000 g for 10 min to remove any remaining cells from the medium. Proteins were purified by Ni-NTA affinity chromatography using an elution buffer containing 25 mM HEPES (pH 7.8), 500 mM NaCl, and 500 mM Imidazole, followed by size-exclusion chromatography (SEC) using a Superose6 Increase 10/300 GL column (Cytiva) into a final buffer composition (SEC-buffer) of 20 mM HEPES (pH 7.8), 150 mM NaCl. After each Ni-NTA affinity and SEC purification step, the proteins were concentrated using a 15 mL Amicon ultra centrifugal filter at 100 kDa cut-off. The final SEC-buffer was either supplemented with 2 mM $CaCl_2$ or concentrated proteins were dialysed overnight at 4°C using 3.5K MWCO Slide-A-Lyzer MINI Dialysis Devices (Thermo Scientific) in 100 mL SEC buffer containing 2 mM $CaCl_2$. Purified proteins were diluted to 20 µg/mL, boiled for 5 min at 98 °C in the presence or absence of β-mercaptoethanol and loaded onto a TGX mini protean precast gel (Biorad) for 50 min at 200V.

SEC-MALS

The oligomeric and conformational states of the different purified isoforms were probed and directly compared by using analytical size exclusion chromatography with subsequent multi-angle light scattering (SEC-MALS). The Ten3 isoforms, along with a bovine serum albumin (BSA) reference, were diluted to a final concentration of 0.5 mg/mL in SEC buffer (20 mM HEPES pH 7.8, 150 mM NaCl.) supplemented with 2 mM $CaCl_2$ and loaded onto a Superose6 Increase 10/300 GL column (Cytiva) integrated on a high-performance liquid chromatography (HPLC) unit (1260 Infinity II, Agilent) with an on-

line UV detector (1260 Infinity II VWD, Agilent), an 8-angle static light scattering detector (DAWN HELEOS 8+, Wyatt Technology), and a refractometer (Optilab T-rEX, Wyatt Technology) in series. On the basis of the measured Rayleigh scattering at different angles and the established differential refractive index increment of value of 0.185 ml/g for proteins in solution with respect to the change in protein concentration (dn/dc), weight-averaged molar masses for each species were calculated using ASTRA software (Wyatt Technology; v.7.3.1).

CRYO-ELECTRON MICROSCOPY AND DATA PROCESSING

All cryo-EM imaging of Ten3 ECDs was performed on purified protein samples diluted in SEC-buffer supplemented with 2 mM $CaCl_2$ to a final concentration of 0.5 mg/mL. 3.0 µL of diluted sample was deposited onto Quantifoil R1.2/1.3 Cu 300 mesh grids that were glow-discharged for 1 min at 15 mA. Excess sample volume was blotted away for 2.5-4.0 s with blot force of -3.0 at 22°C and 100% humidity before plunging into liquid ethane using a Vitrobot Mark IV (Thermo Scientific). Movies were acquired using a K3 detector (Gatan) in counting super-resolution mode with a 20 keV (Gatan) slit width at 105,000x nominal magnification, a corresponding super-resolution pixel size of 0.418 Å, and a total dose of 50 e-/$Å^2$ per movie. A total of 10,927; 3,888; 4,412 and 8,447 movies were acquired for the A_0B_0, A_0B_1, A_1B_0, and A_1B_1 isoforms, respectively, at a defocus range of -0.8 to -2.0 µm (-0.8 to -2.3 µm for the A_0B_0 isoform). All datasets were processed using RELION3.1(Scheres, 2012) unless otherwise stated. Beam-induced motion and drift correction was performed using MotionCorr2[55] with a pixel binning factor of 2 (0.836 Å physical pixel size), and Ctf estimation by using Gctf[56]. Particles were initially picked from the corrected images using RELION's own Laplacian-of-Gaussian method, upon which a manual selection from subsequent 2D classification was used for reference-based particle re-picking. 2D classification was used to clean data of non-particle objects like foil hole edges and ice crystals from the extracted particle sets. Random subsets of particles were further 2D classified to select class averages for *ab initio* reconstruction of a reference for 3D classification on the full set of particles that were cleaned through 2D classification. Automated particle picking of A_1B_0 compact dimers was performed in RELION using the A_1B_1 compact dimer refined map as a reference after 20 Å low-pass filtering. For all datasets, 3D map reconstructions were performed according to the regular workflows of Relion3.1, and as reported in the corresponding Supplementary figures. C2-to-C1 symmetry expansions for the A_1B_1 compact dimer were performed on refined particle datasets. Particle subtraction was performed without automated centering of the resulting maps. For particle subtraction in symmetry-expanded data, the C2-to-C1 particles were first refined with a full-map reference in C1 before masking regions of interest on a single subunit for subtraction outside the masked region. To select particles within datasets that contained good-quality signal for the EGF domains and splice insert A, focussed 3D classification was performed with increased regularisation parameters T, while masking the region of interest. To visualise the compact dimer movement from open to closed conformation, 3D variability analysis in CryoSPARC[57] was used. CryoSPARC's local refinement was used for reconstruction of the dimer particles containing splice insert A density.

MODEL BUILDING AND REFINEMENT

Monomeric Ten3-A_0B_0 (PDB:6FAY) was combined with models predicted by ColabFold[58] for EGF6-8 (A_0), C-rich, ABD, and Tox-GHH domains. Non-compact subunit models for the other isoforms were created by predicting structures for the EGF6-8 containing splice insert A, and NHL containing B, and substituting them in the initial A_0B_0 model. The Ten3 'stalk' was predicted by ColabFold multimer for a dimeric amino acid input spanning the residues of the ECD through EGF5, which exclusively resulted in stalks with EGF2-2/EGF5-5 configuration. To model the compact dimer, two A_1B_1 subunit models were refined in the compact dimer closed conformation full map by iterative cycles of real space refinement in phenix.real_refine of the Phenix package[59] and manual real-space refinement in Coot[60]. EGF6-7 and A that were outside the map were removed from the model. Residues and side chains lacking sufficient density were removed from the model to enhance the refinement process. At the final stages, the model was refined in the full map of teneurin-3 A_1B_1 compact dimer using C2 symmetry constraints in Phenix (Supplementary Table 2.1), and remaining Ramachandran outliers were solved manually in Coot. A single subunit A_0B_0 model containing EGF6 through the C-terminal Tox-GHH domain was refined in the map in which EGF6 and 7 were additionally resolved. Before refinement, this model's C-rich domain was substituted for the C-rich domain from the final refined A_1B_1 compact dimer model. The stereochemistry of the models was checked with MolProbity[61], the interfaces were analysed with the Protein Interfaces, Surfaces and Assemblies (PISA) web server[62]. The structural figures of maps and models were generated with Chimera[63] and PyMOL (Schrödinger Molecular Graphics System, DeLano Scientific, San Carlos, CA). Protein sequence alignments were generated using Clustal Omega[64].

SMALL-ANGLE X-RAY SCATTERING

SAXS was performed at the European Synchrotron Radiation Facility BM29 beamline, at 12.5 keV operating energy equipped with a Pilatus3 2M detector. Samples at approximately 1 mg/mL were diluted two-fold twice in series, and then 50 microliters of each was injected, i.e., at ~1, ~0.5 and ~0.25 mg/mL for each of the four Ten3 splice variants and at three buffer conditions. The three conditions, all containing SEC buffer (150 mM NaCl, 20 mM HEPES, pH 7.8), were either without additional calcium, with 2 mM $CaCl_2$ or with 5 mM EDTA added. Data were collected at 20°C over 10 frames, one frame per second, with a scattering vector range of 0.0456–5.15 nm^{-1}. Images were radially averaged and normalised using the BM29 EDNA-based pipeline[65]. Radiation damage was monitored and data frames were selected manually in the PRIMUS GUI[66], which was also used for frame averaging, buffer subtraction, Guinier modelling, and determining the pair-distance distribution P(r). Kratky analyses were performed according to Durand *et al.* (2010)[67]. In preparation for rigid-body fitting, complex glycan trees were built manually on glycosylation sites of the compact models in Coot(Emsley & Cowtan, 2004). A model for the N-terminal domains that could not be resolved by cryo-electron microscopy was generated in ColabFold[58]. These were then taken together, and short flexible linkers were introduced between domains, conserving the interfaces of the compact dimers, as well as the EGF2/EGF5 covalent dimerisation interface, and fit to the datasets of each splice insert was optimised using CORAL[68]. The fit of each resulting

model to each dataset was then calculated in CRYSOL[69], using 50 as "number of spherical harmonics" and allowing for constant subtraction. All programs were used as implemented in ATSAS 3.1.3[70]. SAXS measurements of protein dilutions for which insufficient coherent frames were collected to reliably calculate an averaged scattering curve were excluded, with a maximum of one exclusion per protein sample.

THERMAL STABILITY ASSAY

Thermal shift assay (TSA) assays were performed using purified human Ten3 isoforms at a concentration of 0.5 mg/mL. Prior to the experiments, the samples were buffer exchanged into SEC buffer (150-mM NaCl, 20 mM HEPES, pH 7.8) without additional calcium denoted "SEC" condition, into SEC buffer supplemented with 2 mM $CaCl_2$ (physiological calcium) denoted "Ca^{2+}" condition, or into SEC buffer supplemented with 5 mM EDTA (no calcium) denoted "EDTA" condition. SYPRO Orange Dye (Invitrogen) was diluted to a concentration of 5X concentrated solution and filtered using a 0.22 µm membrane. Final protein concentrations were 60 µg/mL protein, 1× dye and final buffer concentrations 132 mM NaCl, 17.6 mM HEPES, pH 7.8, supplemented with 1.76 mM $CaCl_2$ or EDTA. A temperature ramp from 5°C to 95°C was set up at a speed of 0.02°C/s on a QuantStudio 3 Real-Time PCR system (Thermo Fisher Scientific). All measurements were performed in triplicates. The melting temperatures T_m were determined as the intersection of the x-axis with the smoothened second derivative of the melting curve between the valley and peak shown in the raw melting curves.

NEGATIVE STAIN IMAGING AND ANALYSIS

Purified protein samples were diluted to 20 µg/mL before depositing 3.5 µL of the resulting sample volume onto glow-discharged carbon-coated copper grids for 2 min incubation. Excess sample volume was then blotted away, followed by three washing cycles of depositing 10 µL of SEC buffer supplemented with 2 mM $CaCl_2$ with subsequent blotting. The grid was then incubated with 3.5 µL 2% uranyl acetate or 1% uranyl formate (for A_0B_0) for another 60 sec before blotting the excess away. Grids were left to dry at RT for up to 2 min before imaging or storage. At least 20 micrographs were acquired for each sample using a JEOL JEM-1400Plus transmission electron microscope operating at 120 kV with a standard magnification of 30,000x, 1000 ms exposure, and a defocus value of -2.4 µm (or, Tecnai12 microscope operated at 120 kV for A_0B_0). Micrographs were processed using RELION3.1(Scheres, 2012) before following standard reconstruction workflow. In short, particles were first picked using RELION's own Laplacian-of-Gaussian implementation, followed by reference-based particle picking after 2D classification. An initial 3D model was generated for multiple rounds of 3D classification of all particles. The final selection was refined and postprocessed in RELION. To determine model-map correlations, two non-compact subunits of the isoform corresponding to the nsTEM electron densities (with EGF6-8 for A_0 and without for A_1 isoforms) were fit using "Fit in map" tool in Chimera. A calculated low-resolution filtered map (molmap) of 20 Å resolution was created from each dimeric model and fit to all experimental densities using "Fit in map" to acquire correlation values.

CELL CLUSTERING ASSAY

Cell electroporation for the clustering assays was performed as previously reported[71]. K562 cells (catalog nr. ACC 10, Leibniz Institute DSMZ) were cultured in RPMI-1640 medium (Gibco), supplemented with 10% FBS (Gibco) and 1% Penicillin/Streptomycin (Gibco), and grown in a shaking incubator at 37°C and 5% CO_2. Prior to electroporation, K562 cells were centrifuged for 5 min at 300 g and washed in 1x PBS (Gibco). Cells were once again centrifuged for 5 min at 300 g and resuspended in buffer R (Gibco). Per condition, 2×10^6 cells were incubated with a total amount of 15 µg of DNA (Ten3:empty vector ratio was 1:5) for 15 min at room temperature. After the incubation, K562 cells were electroporated with the Neon Transfection System (Thermo Fisher Scientific), using the following parameters: 1450 V, 10 ms pulse length, and 3 pulses[8]. After electroporation, cells were directly seeded randomly onto 5 mL of pre-warmed RPMI-1640 medium with 10% FBS in 6-well plates. Cells were placed in a shaking incubator at 37°C and 5% CO_2 for approximately 20h. Cells were imaged blindly on an EVOS M5000 microscope with a 10x objective (0.25 NA; EVOS, Thermo Fisher Scientific), using the EVOS LED GFP cube (Thermo Fisher Scientific). The researcher was blind to which condition was being imaged to prevent bias for cluster formation. In the GFP channel, regions of interest (ROIs) larger than 100 pixels were identified using Fiji[72] Analyze Particles after rolling ball background subtraction with 50-pixel radius was performed. The area of each ROI was then measured. A cell cluster was defined as an object three times larger than the mean large single cell size (800 pixels). The clustering index was determined as the summed cluster area divided by summed area of all ROIs (clusters + non-clusters) times 100%. The cluster size and clustering index were averaged per image, and data from three independent experiments (5 images per experiment, 15 images total per condition). Statistical significance was determined by performing a one-way ANOVA followed by a Tukey's multiple comparison test using GraphPad Prism (GraphPad Software, San Diego, California USA). All analysed data are represented as mean ± SEM.

ANIMALS FOR NEURONAL ASSAYS

Fertilised chicken eggs (*Gallus gallus*, white leghorn) were obtained from a commercial dealer and incubated in a forced draft incubator at 37°C and 65% humidity until the desired embryonic stage.

In situ HYBRIDISATION (ISH)

Eyes from chick embryos (*Gallus gallus*, White Leghorn strain, obtained from Henry Steward Ltd.) at developmental day (E) 6 were fixed in 4% paraformaldehyde (PFA) solution for 2 days, followed by perfusion with 4% sucrose/0.5% PFA solution, until samples sunk to the bottom of the tube. Eyes were embedded in O.C.T. cryo-embedding matrix prior to cryo-sectioning. 20-µm thick coronal sections were cut on the cryostat to obtain dorsoventral sections of the chick retina and stored at -80°C. Sections were rehydrated for 5 min in PBS-T (1x PBS + 0.1% Tween-20), and permeabilised by digestion with 5µg/mL Proteinase K for 5 min at room temperature. Digestion was inhibited by 2mg/mL (2%) glycine in PBS-T, followed by 2 times 5 min washes with PBS-T and fixation with 4% PFA for 20 min. Residual PFA was removed by washing twice with PBS-T for 5 min, and sections were prehybridised in hybridisation mix (50% formamide, 5x

SSC pH 4.5, 50 µg/mL tRNA, 1% SDS and 50 µg/mL Heparin) for 1 h at 65°C, prior to hybridisation of probe overnight at 65°C. Sections were washed 3 times for 15 min with a solution containing 50% formamide, 5 X SSC pH 4.5 and 1% SDS and 3 times for 15 min in solution comprising 50% formamide, 2 X SSC pH 4.5, at 65°C and 60°C, respectively. After blocking for 1 h at room temperature with 10% sheep serum in TBS-T (1x TBS + 0.1% Tween-20), sections were incubated with polyclonal Alkaline Phosphatase (AP)-conjugated anti-DIG antibody (Roche, 11093274910) diluted 1:2000 in 10% sheep serum in TBS-T overnight at 4°C. Samples were extensively washed to remove excess antibody. Sections were then washed thrice for 10 min in NTMT buffer containing 100 mM NaCl, 100mM Tris-HCl pH 9.5, 100mM $MgCl_2$ and 1% Tween-20. The colour staining was developed at room temperature and in the dark using NBT/BCIP stock solution (75 mg/mL and 50 mg/mL, respectively) (Roche) diluted 1:200 in NTMT buffer. The reaction was terminated by 2 washes for 10 min in NTMT, followed by a wash in PBS-T pH 5.5 for 10 min and 2 washed for 10 min in 1x PBS. Samples were postfixed in 4% PFA in PBS for 30 min at room temperature and stained with Hoechst diluted 1:10,000 in 1x TBS before a final wash in PBS and mounting in 90% glycerol.

STRIPE ASSAYS

Stripe assays were performed according to a previously established protocol[42]. Prior to the experiment, specially designed silicone matrices were boiled in autoclaved water, glass coverslips were coated in 10 µg/mL poly-L-lysine (PLL) for 1 h and both were left to dry overnight in a laminar flow hood. On the day of the experiment, dry matrices were fixed onto glass coverslips. Solutions were applied using a Hamilton syringe which was rinsed before use with ethanol and Hank's Basic Salt Solution (HBSS). Monomeric Ten3-B_0/B_1 proteins were diluted in HBSS (ThermoFisher) and used at a final concentration of 10 µg/mL. Before application, monomeric Ten3-B_0/B_1 was clustered using monoclonal anti-6xHis Tag antibodies (GeneTex, GTX44514) and fluorescently labelled by adding a small amount of polyclonal goat anti-rabbit AlexaFluor 568 secondary antibody (ThermoFisher, A-11011) enabling subsequent identification of 1st and 2nd applied stripes. The anti-6xHis Tag antibody (10 µg/mL) alone was used for generating control stripes (2nd stripe). Laminin and merosin solutions were prepared to final concentrations of 5 µg/µL and 1µg/mL, respectively. Neurobasal (NB) medium was supplemented with 2% methylcellulose, 1% B27, 1% glutamine, 1% Penicillin/Streptomycin (Pen/Strep) and 0.01% forskolin. To generate the first set of stripes, clustered purified protein solutions were injected into the matrices using a Hamilton syringe and culture dishes were incubated at 37°C/5% CO_2 for 45 min. Following the removal of matrices from the coverslips and HBSS wash, unlabelled second stripe solution, containing the anti-6xHis Tag antibody in HBSS, was pipetted on top of the stripes. Culture dishes were incubated for 45 min at 37°C and coverslips subsequently coated with laminin/merosin solution for 1 h. Freshly dissected chick retinas at embryonic stage E6 were flat-mounted on nitrocellulose filter paper and cut into ~2 mm wide strips along the dorsoventral axis using a tissue chopper. Explant strips were placed perpendicular on top of protein stripes and incubated at 37°C/ 5% CO_2 for up to three days in neurobasal media. Explant strips were fixed in 4% PFA/0.33% sucrose for 15-20 min at room temperature and permeabilised by washing twice for 5 min with 1x PBS/0.1%Triton and subsequently stained by incubation

with 1:200 phalloidin conjugated with AlexaFluor 488 for 20 min at room temperature, in a wet chamber and under aluminium foil. After a series of washes to remove excess staining solution, coverslips were mounted onto glass slides using Mowiol. Stripe assays were imaged using a Fluorescence microscope (Zeiss) at 20x magnification. Three independent rounds of stripe assays were performed, each round consisting of 4-6 individual assays per condition (Control vs monomeric Ten3-B_0/B_1), using explants from at least 2 animals. For individual stripe assay, two images were taken from the ventral and dorsal regions of each retinal explant strip. Axon guidance decisions in all images were scored blindly by 3 people, in accordance with a previously established scoring system, with values between -3, 0 and +3, representing strong repulsion, no guidance and strong attraction, respectively[42,43]. Mean scores were calculated and plotted using GraphPad (mean ± SEM), and statistical significance was determined with Two-way ANOVA with Tukey's multiple comparisons tests.

ACKNOWLEDGEMENTS

Structural data from Protein Data Bank (PDB) were used under the CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license. Cryo-EM data were collected at The Netherlands Centre for Electron Nanoscopy (NeCEN) with assistance from Willem Noteborn. We acknowledge the European Synchrotron Radiation Facility (ESRF) for provision of synchrotron radiation facilities, and we would like to thank Petra Pernot for assistance and support in using beamline BM29. UD is supported by a BBSRC grant (BB/T013753/1). BJCJ is supported by an NWO grant (OCENW.KLEIN.026). DHM is supported by an NWO computing grant (2021.058) and NWO Veni grant (722.016.004).

2.5. SUPPLEMENTARY INFORMATION

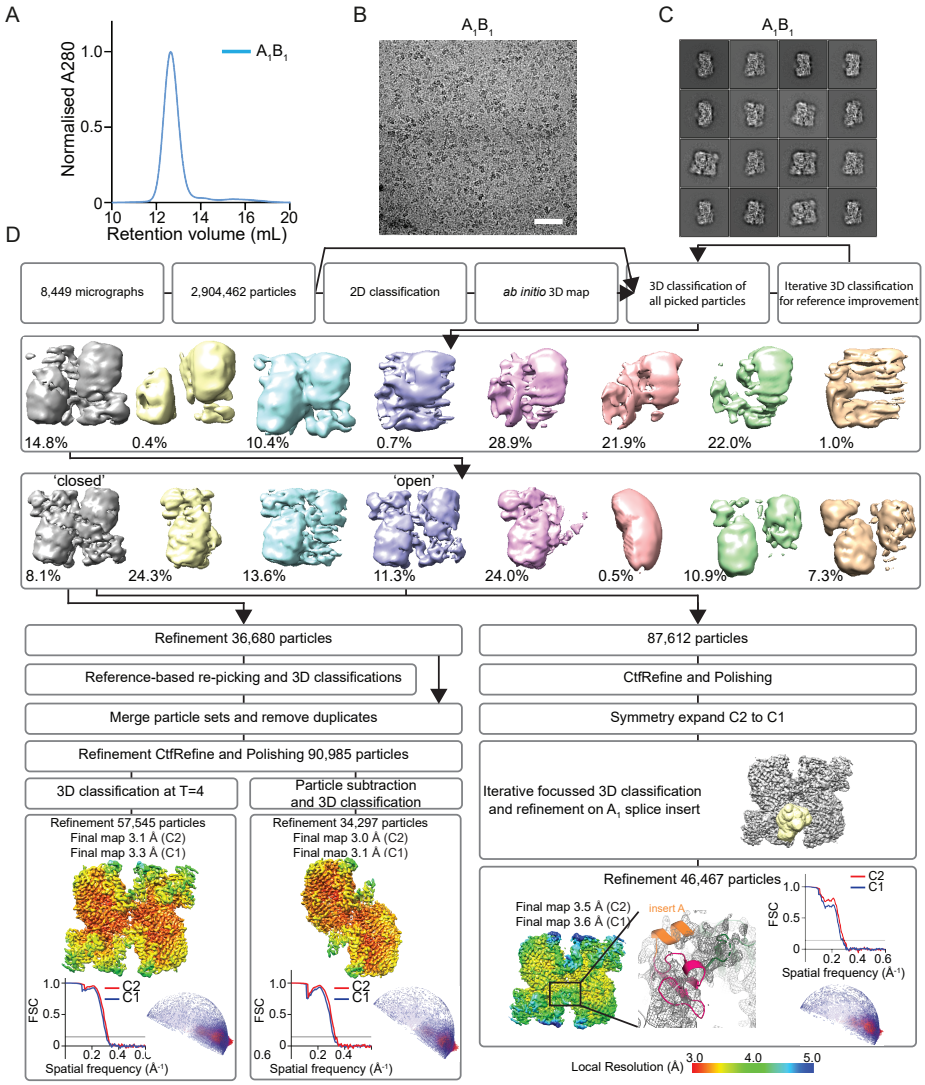


Figure 2.8: Purification, cryo-EM data collection and reconstruction of teneurin-3 A_1B_1 compact dimer. **A)** Size exclusion chromatography (SEC) trace of the Ten3-A_1B_1 ectodomain after nickel affinity purification. **B)** Representative cryo-EM micrograph of vitrified Ten3-A_1B_1 ectodomain. Scale bar is 50 nm. **C)** Single-particle analysis (SPA) 2D classes of compact dimeric Ten3-A_1B_1 particles in 'closed' conformation. **D)** SPA workflow for the reconstruction of cryo-EM electron density of compact dimeric Ten3-A_1B_1. Left: reconstruction workflow of the full dimeric map. Second to left: dimeric map reconstruction after particles subtraction around interface domain. Third to left: C2-to-C1 symmetry-expanded reconstruction of subunit with density outside YD through Tox-GHH subtracted. Right: focussed classification and refinement of splice insert A (orange helix) with dimers from the closed conformation and more flexible open conformation. Corresponding local resolution values, Fourier shell correlation (FSC) graphs, and angular distribution of particles in final refinement are shown for each map. FSC cut-off is 0.143.

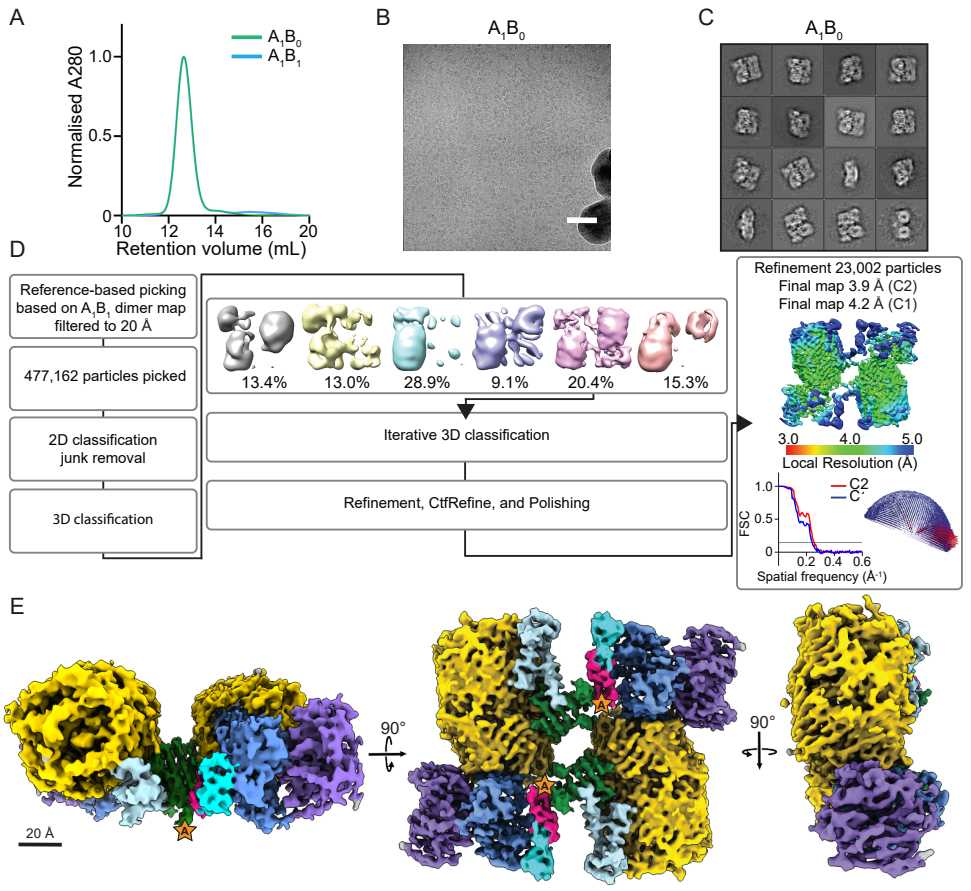


Figure 2.9: Purification, cryo-EM data collection and reconstruction of teneurin-3 A_1B_0 compact dimer. **A)** Size exclusion chromatography (SEC) trace of the Ten3-A_1B_0 versus A_1B_1 ectodomains after nickel affinity purification. Traces essentially overlap. The A_1B_1 trace is equal to that in Supplementary Fig. 2.8A. **B)** Representative cryo-EM micrograph of vitrified Ten3-A_1B_0 ectodomain. Scale bar is 50 nm. **C)** Single-particle analysis (SPA) 2D classes of compact dimeric Ten3-A_1B_0 particles in 'open' conformation (no YD-YD contact). **D)** SPA workflow for the reconstruction of the cryo-EM density map of compact dimeric Ten3-A_1B_0. Corresponding local resolution values, Fourier shell correlation (FSC) graphs, and angular distribution of particles in final refinement are shown. FSC cut-off is 0.143. **E)** Three different views of the density map of the teneurin-3 A_1B_0 compact dimer at 3.9 Å resolution (EMD-19409). Domains are coloured corresponding to the linear representation in Fig. 1A. Positions of the splice insert A are indicated with stars. Scale bar is 20 Å.

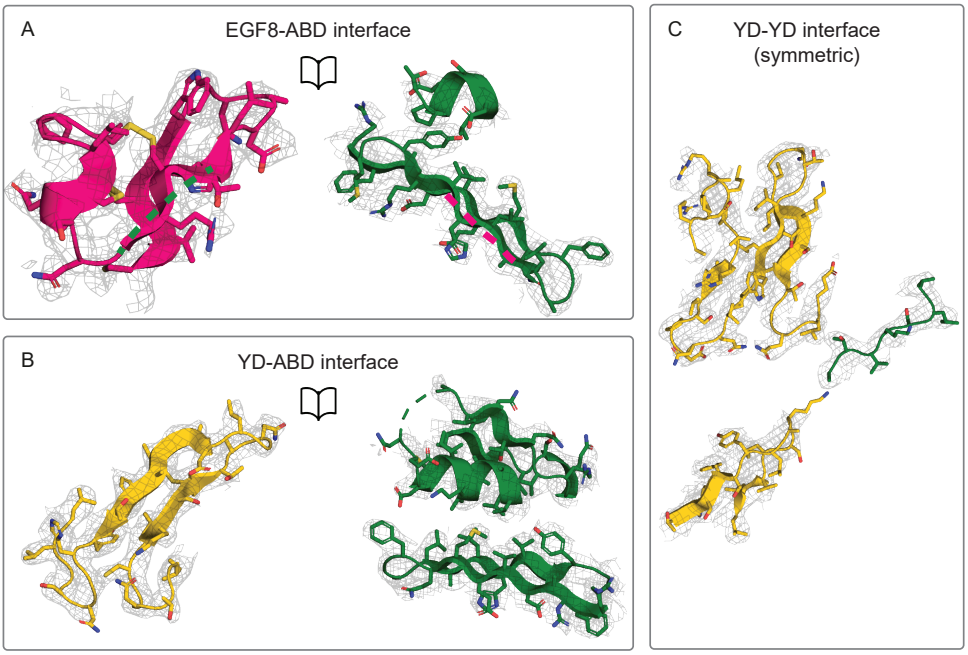


Figure 2.10: Cryo-EM densities for secondary structure elements in the teneurin-3 compact homo-dimeric interface of the A_1 isoforms. **A)** Open book representation of the secondary structure elements within the EGF8-ABD interface I. Dashed lines in the colour of the complementary domain indicate the extended β-sheet contact. **B)** Open book representation of the secondary structure elements within the YD-ABD interface II. **C)** Secondary structure elements within the YD-YD interface III on the side of one non-compact subunit. All density is visualised at a contour level of 4.2 RMSD. Domains are coloured according to colouring in Fig. 1.

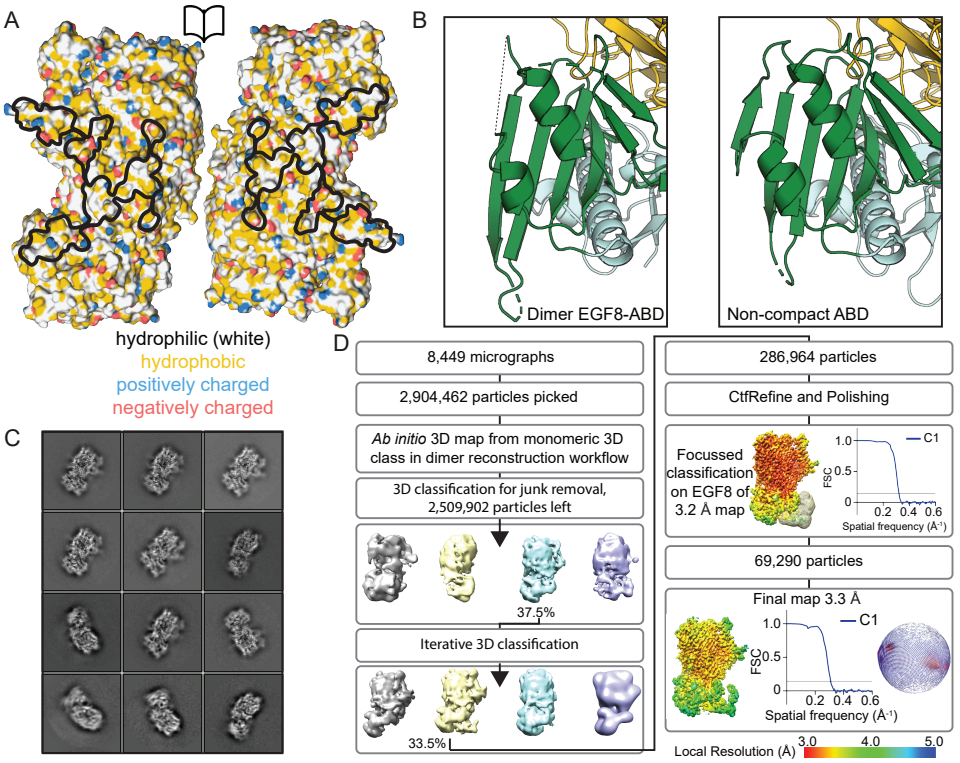


Figure 2.11: Teneurin-3 A_1B_1 compact interface analysis and non-compact subunit reconstruction. **A)** Open book electrostatics and hydrophobicity surface representation of the Ten3 compact dimer. Black lines indicate the outline of compact dimer interface. **B)** Structural comparison of the Ten3- A_1B_1 non-compact subunit (PDB:8R51) ABD versus the compact dimeric (PDB:8R50) EGF8-contacting ABD. The non-compact ABD displays the presence of an additional β-strand at solvent-facing edge of the ABD β-sheet. Missing residues in the non-compact subunit are indicated with dashed black line. **C)** Single-particle analysis (SPA) 2D classes of Ten3-A_1B_1 non-compact subunit particles after focussed classification on EGF8. **D)** SPA workflow for the reconstruction of the Ten3-A_1B_1 non-compact subunit, including subsequent focussed classification and refinement of the EGF8-containing particles. Corresponding local resolution values, Fourier shell correlation (FSC) graphs, and angular distribution of particles in final refinement are shown for each map. FSC cut-off is 0.143.

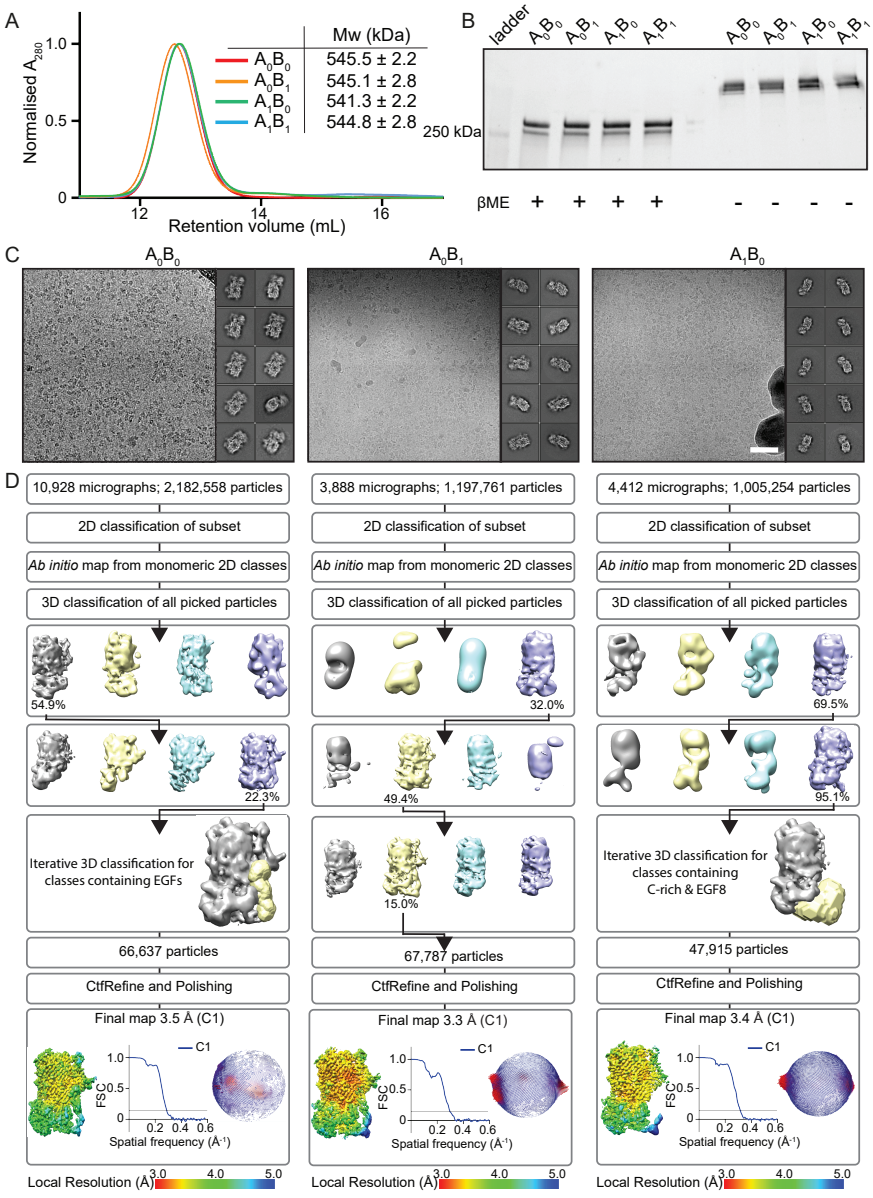


Figure 2.12: Purification, cryo-EM data collection and reconstruction of teneurin-3 A_0B_0, A_0B_1, A_1B_0, non-compact subunits. **A)** Size exclusion chromatography (SEC) trace of all Ten3 ectodomains after nickel affinity purification. Traces essentially overlap. Inset displays the weight-averaged molar mass determined from multi-angle light scattering (MALS) data (mean ± SEM) of all Ten3 isoform at a protein concentration of 1.0, 0.5, and 0.25 mg/mL in presence of 2 mM Ca^{2+} directly after exiting the SEC column. The A_1B_0 and A_1B_1 traces are equal to those in Supplementary Fig. 2A. **B)** PAGE gel of all purified Ten3 isoform ectodomain in the presence and absence of reducing agent β-mercaptoethanol (βME). This gel was repeated twice with similar results. Full uncropped image of the gel is available. **C)** Representative cryo-EM micrographs of vitrified Ten3 ectodomain and single-particle analysis (SPA) 2D classes for the non-compact A_0B_0, A_0B_1, and A_1B_0 isoform subunits. Scale bars are 50 nm. Per isoform 2D classes of the final reconstructions from single-particle (SPA) analysis in D are shown. The representative micrograph for the A_1B_0 isoform is identical to Supplementary Fig. 2B. **D)** SPA workflow for the reconstruction of the cryo-EM density map of Ten3 isoform subunits. Workflows include subsequent focussed classification and refinement of the EGF-containing particles. Corresponding local resolution values, Fourier shell correlation (FSC) graphs, and angular distribution of particles in final refinement are shown for each map. FSC cut-offs are 0.143.

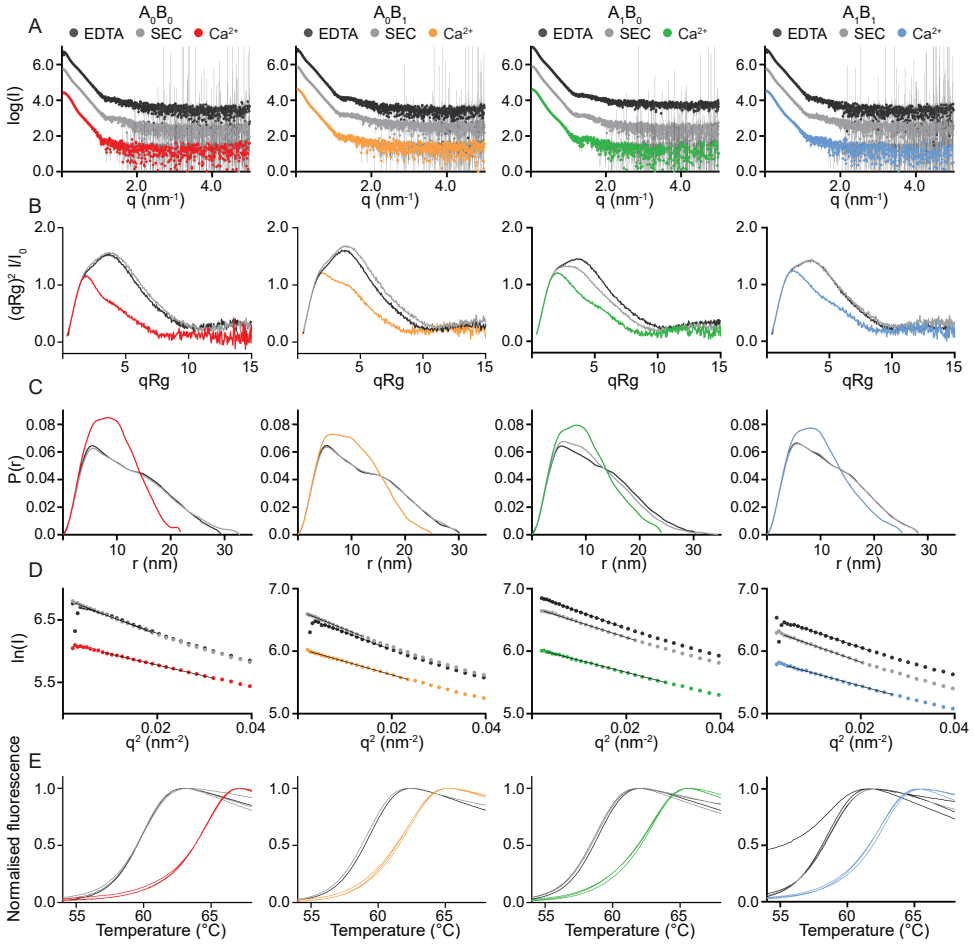
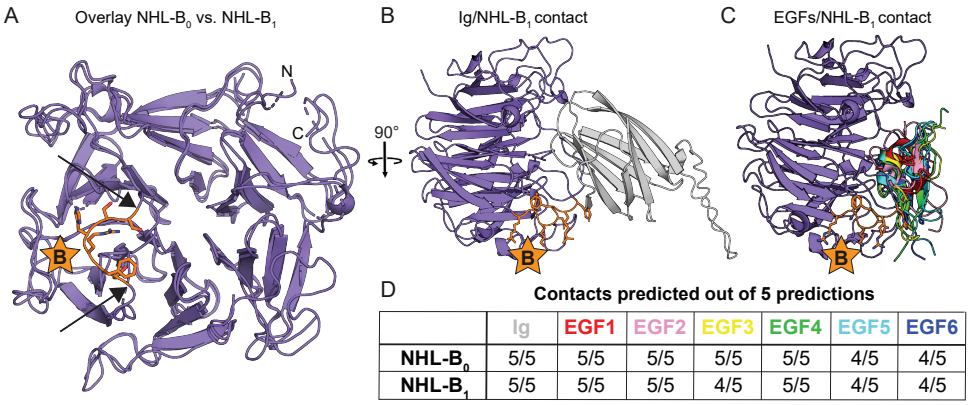


Figure 2.13: The compactness of all four Ten3 isoforms is influenced by the presence of calcium ions. **A)** Small-angle X-ray scattering profile for all purified teneurin-3 (Ten3) isoforms. Scattering curves (left) are shifted by an arbitrary offset along the y-axis for clarity. **B)** Dimensionless Kratky plot for all purified Ten3 isoforms. **C)** Pair-distance distribution, normalised to the area under the distribution, for all purified Ten3 isoforms. **D)** Guinier plot for all purified Ten3 isoforms. Only datapoints in the q^2 range indicated by the Guinier model linear fit (black line) were included in the linear fit. All R^2 values for linear fits are greater than 0.997 (see Supplementary Supplementary Table 2.2), supporting the SAXS data quality. **E)** Melting curves for all purified Ten3 isoforms. Each graph compares the conditions for a Ten3 isoform in the presence (Ca^{2+}) and absence (SEC) of 2 mM calcium, as well as under addition of 5 mM EDTA (EDTA).

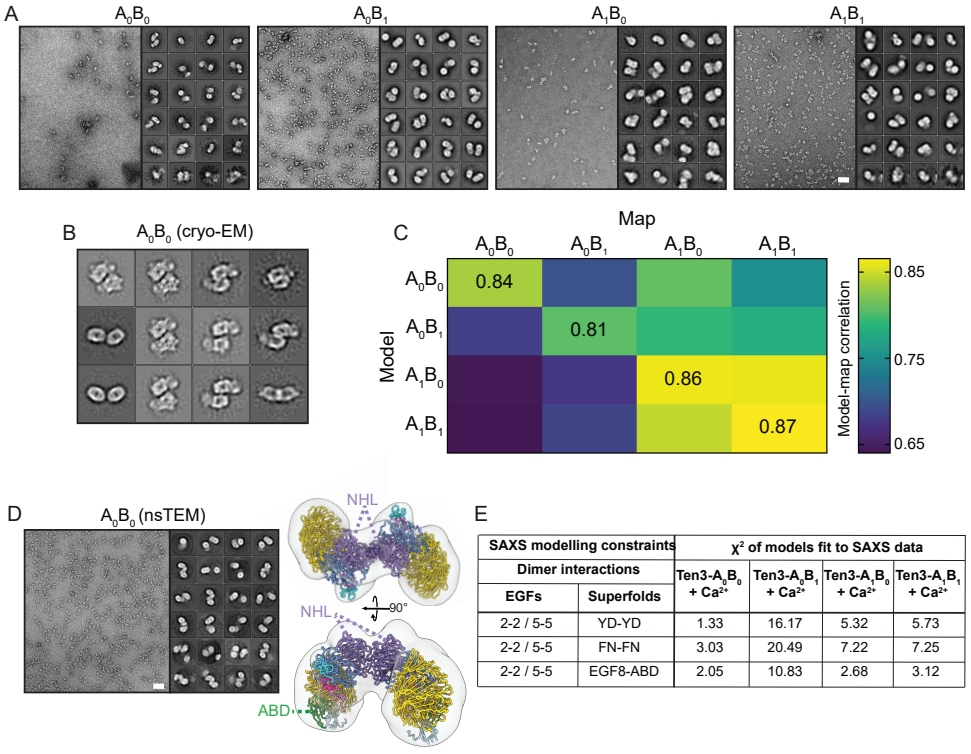


D

Contacts predicted out of 5 predictions

	Ig	EGF1	EGF2	EGF3	EGF4	EGF5	EGF6
NHL-B_0	5/5	5/5	5/5	5/5	5/5	4/5	4/5
NHL-B_1	5/5	5/5	5/5	4/5	5/5	4/5	4/5

Figure 2.14: Effect of splice insert B on local NHL domain structure and on interactions with EGF and Ig domains. **A)** Structural comparison of NHL domains lacking (B_0) and containing (B_1) splice insert B. B_0 and B_1 models were built into Ten3-A_0B_0 non-compact subunit and A_1B_1 compact dimer subunit densities, respectively. Splice insert B is shown as orange stick representation. Boundaries of splice insert B are indicated with arrows. N- (N) and C-termini (C) of the NHL domain are indicated. **B)** Colabfold predictions of the NHL-B_1 contact with Ig domain (grey). Splice insert B is shown as orange stick representation, and indicated with a star. **C)** Colabfold predictions of the NHL-B_1 contact with EGF1 through EGF6. Different EGFs predictions have different rainbow colouring, and splice insert B is shown as orange stick representation, and indicated with a star. **D)** Table showing the proportion of predictions containing a direct contact between the EGF or Ig domains and the NHL loop harbouring splice site B - out of 5 total predictions per NHL-EGF or NHL-Ig pair.



E

SAXS modelling constraints		χ^2 of models fit to SAXS data			
Dimer interactions		Ten3-A_0B_0 + Ca^{2+}	Ten3-A_0B_1 + Ca^{2+}	Ten3-A_1B_0 + Ca^{2+}	Ten3-A_1B_1 + Ca^{2+}
EGFs	Superfolds				
2-2 / 5-5	YD-YD	1.33	16.17	5.32	5.73
2-2 / 5-5	FN-FN	3.03	20.49	7.22	7.25
2-2 / 5-5	EGF8-ABD	2.05	10.83	2.68	3.12

Figure 2.15: Teneurin-3 isoform-specific compactness and assessment using negative stain electron microscopy. **A)** Representative negative stain EM micrographs of Ten3 ectodomain and single-particle analysis (SPA) 2D classes of all picked particles for each isoform indicated above. Scale bars is 50 nm. **B)** Cryo-EM SPA 2D classes of Ten3 A_0B_0 in a YD-YD-connected compact conformation. **C)** Quantification of model-to-map fitting in panel D by correlation of the model-derived calculated low-resolution filtered (20-Å) maps with the experimental nsTEM density maps. Heat map representing correlation values for each model to each map, and numbers (on diagonal) in the heat map indicate the highest correlation values per experimental map. **D)** Negative stain TEM micrograph with SPA 2D classes and 3D refinement of the A_0B_0 NHL-NHL connected compact dimer. Scale bars is 50 nm. **E)** Fits of rigid-body models to SAXS data. Rigid-body models were calculated by restraining the superfold interfaces described in Fig. 2.5E and the covalent disulfide bonds in the EGF stalk and using the SAXS data corresponding to the isoform for which the superfold interfaces are found, i.e. Ten3-A_0B_0 for YD-YD, Ten3-A_0B_1 for FN-FN and Ten3-A_1B_1 for EGF8-ABD. χ^2 values are calculated by comparing the calculated models against the SAXS data of all isoforms.

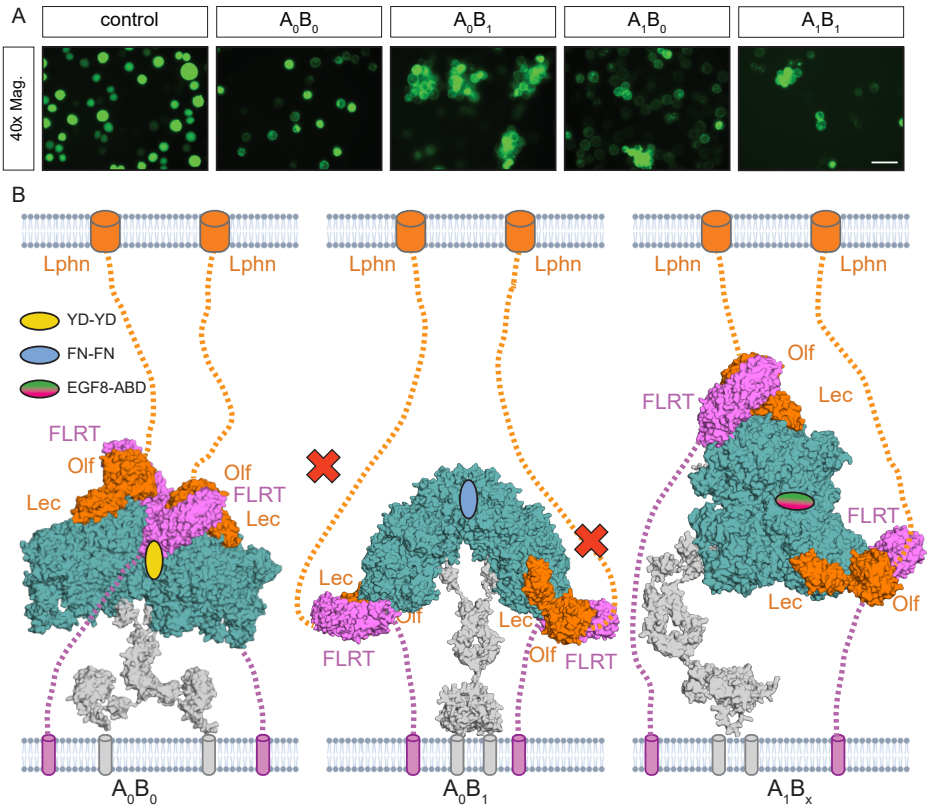


Figure 2.16: Model for isoform-specific teneurin-latrophilin trans-cellular complex formation. **A)** Clustering assay of K562 hematopoietic cells electroporated with the mouse Ten3 isoforms at 40x magnification. For all isoforms, GFP signal localises to the membrane in contrast to the soluble GFP that occupies the entire cell area. Scale bar is 100 µm. **B)** Exposure of known teneurin-latrophilin (Olf-Lec, PDB:6SKA, in orange) [12,30] and latrophilin-FLRT (ankyrin-repeat domain, PDB:5FTU, in magenta) [41] binding sites for each splicing-dependent compact dimer. In all panels, coloured ellipses indicate the type of homophilic compact dimeric contacts formed, according to the interface colour codes in Fig. 2.5E. The rigid-body models of each of the EGF-Ig stalks were calculated using the SAXS data of each isoform (Fig. 2.5 and Supplementary Fig. 2.15E). Transmembrane domains are indicated by tubes. The crosses indicate the unfavourable orientation of latrophilin in the teneurin A_0B_1-latrophilin trans model.

Table 2.1: Cryo-EM data collection and refinement statistics.

Data collection and processing	**Teneurin-3 A_1B_1 Compact dimer**	**Teneurin-3 A_1B_0 Compact dimer**	**Teneurin-3 A_1B_1 Subunit**	**Teneurin-3 A_0B_0 Subunit**	**Teneurin-3 A_0B_1 Subunit (EGFs focused)**	**Teneurin-3 A_1B_0 Subunit (EGFs focused)**
EMD ID	EMD-18889	EMD-19409	EMD-18890	EMD-18891	EMD-18900	EMD-18902
PDB ID	8R50	-	8R51	8R54	-	-
Magnification	105,000x	105,000x	105,000x	105,000x	105,000x	105,000x
Electron exposure (e-/Å^2)	50	50	50	50	50	50
Voltage (kV)	300	300	300	300	300	300
Defocus range (μm)	-0.8 to -2.0	-0.8 to -2.0	-0.8 to -2.0	-0.8 to -2.3	-0.8 to -2.0	-0.8 to -2.0
Pixel size (Å)	0.836	0.836	0.836	0.836	0.836	0.836
Symmetry imposed	C2	C2	C1	C1	C1	C1
Initial particle images (n)	736,451	477,162	2,904,462	2,182,558	1,197,761	1,005,254
Final particle images (n)	57,545	23,002	286,964	66,637	67,787	47,915
Map resolution (Å)	3.1	3.9	3.2	3.4	3.2	3.4
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	3.1-4.5	3.8-7.1	3.1-4.5	3.4-4.8	3.2-5.4	3.2-6.2
Refinement						
Initial model used (PDB code)	6FAY		6FAY	6FAY		
Model resolution (Å)	3.4		3.2	3.6		
FSC threshold	0.5		0.5	0.5		
Model composition	EGF8 through Tox-GHH		EGF8 through Tox-GHH	EGF6 through Tox-GHH		
Non-hydrogen atoms	29,752		13,982	15,900		
Protein residues	3,778		1,793	1,982		
Ligands (glycans)	32		13	24		
B factors (Å^2)						
Protein	69.29		45.3	87.56		
Ligand	78.95		49.9	100.24		
R.m.s.d						
Bond lengths (Å)	0.002		0.003	0.005		
Bond angles (°)	0.485		0.469	0.609		
Validation						
MolProbity score	2		1.75	2		
Clash score	9.64		6.96	9.06		
Poor rotamers (%)	0.16		0.14	0.35		
Ramachandran						
Favoured (%)	91.95		94.6	91.16		
Allowed (%)	7.99		5.34	8.79		
Disallowed (%)	0.05		0.06	0.05		

Table 2.2: SAXS data collection parameters and analysis.

	Teneurin-3 A_0B_0 (EDTA, SEC, Ca^{2+})	**Teneurin-3 A_0B_1 (EDTA, SEC, Ca^{2+})**	**Teneurin-3 A_1B_0 (EDTA, SEC, Ca^{2+})**	**Teneurin-3 A_1B_1 (EDTA, SEC, Ca^{2+})**
	Data collection parameters			
Instrument	ESRF BM29			
Beam geometry	100 µm x 100 µm			
Wavelength (Å)	0.9918			
q-range (nm^{-1})	0.0456–5.15			
Exposure time (sec)	1 per frame			
Concentration (mg/mL)	1.03,	1.23,	1.68,	1.30,
	0.87,	1.29,	1.38	0.97,
	0.46	0.72	0.74	0.62
Temperature (K)	293			
Flux (photons/sec)	~$8.17 \cdot 10^{11}$			
	Structural parameters			
Expected mass (kDa)	536			
Molar mass (kDa) [from I_0]	547,	585,	594,	549,
	659,	604,	595,	592,
	606	590	589	577
I_0 [From Guinier, autoRg]	563.6 ± 2.3,	719.3 ± 3.2,	998.1 ± 2.5,	713.9 ± 2.6,
	573.4 ± 1.8,	779.6 ± 1.3,	820.6 ± 2.1,	574.5 ± 1.4,
	279.2 ± 0.5	422.0 ± 0.7	423.7 ± 1.0	348.2 ± 1.1
Rg (nm) [from Guinier, autoRg]	9.04 ± 0.05,	9.12 ± 0.06,	9.06 ± 0.03,	8.79 ± 0.04,
	9.43 ± 0.04,	9.61 ± 0.02,	8.60 ± 0.03,	8.89 ± 0.03,
	7.26 ± 0.02	7.97 ± 0.02	7.67 ± 0.03	7.87 ± 0.04
$q_{min}Rg - q_{max}Rg$ used for Guinier	0.64 – 1.29,	0.56 – 1.26,	0.41 – 1.25,	0.67 – 1.30,
	0.48 – 1.30,	0.37 – 1.27,	0.44 – 1.27,	0.54 – 1.27,
	0.67 – 1.30	0.40 – 1.22	0.43 – 1.29	0.52 – 1.28
Correlation coefficient R^2 of Guinier fit	0.9983,	0.9973,	0.9991,	0.9989,
	0.9987,	0.9995,	0.9990,	0.9992,
	0.9997	0.9996	0.999	0.9985
I_0 [from Guinier, GNOM]	570.6,	730.5,	1000.2,	716.0,
	577.7,	774.5,	829.9,	572.8,
	276.4	419.1	423	346.7
I_0 [from P(r), GNOM]	570.7,	730.6,	1000.2,	716.1,
	577.8,	774.5,	830.0,	572.9,
	276.5	419.1	423	346.7
Rg (nm) [from Guinier, GNOM]	9.40,	9.55,	9.30,	9.00,
	9.78,	9.67,	8.97,	9.02,
	7.19	7.95	7.75	7.93
Rg (nm) [from P(r), GNOM]	9.45,	9.59,	9.32,	9.04,
	9.81,	9.69,	9.00,	9.05,
	7.2	7.96	7.76	7.95
Dmax (nm) [from P(r), GNOM]	29.4,	30.3,	30.7,	28.4,
	32.8,	27.8,	34.7,	28.4,
	21.2	25.1	24.2	25.4
Porod volume V_p (nm^3) [from p(r), GNOM]	1026,	1015,	1056,	1000,
	1087,	1057,	1070,	1010,
	1025	1020	1065	1089
Molar mass (kDa) [from V_p]	604,	597,	621,	588,
	639,	622,	629,	594,
	603	600	626	641
GNOM goodness of fit	0.7817,	0.7073,	0.7630,	0.7984,
	0.7081,	0.7529,	0.6542,	0.7975,
	0.8469	0.8101	0.8563	0.8396
	Software employed			
Primary data reduction	ESRF BM29 data analysis pipeline			
1D data processing	Primus55, autoRg, Gnom from ATSAS 3.1.3			
Rigid-body modelling	Coral from ATSAS 3.1.3			
Model-data fits	Crysol from ATSAS 3.1.3			
	SASBDB identifiers			
	SASDT83,	SASDTD3,	SASDTJ3,	SASDT43,
	SASDT53,	SASDT93,	SASDTE3,	SASDTY2,
	SASDT63	SASDTB3	SASDTG3	SASDT24

REFERENCES

[1] C. Gogou, J. W. Beugelink, C. P. Frias, L. Kresik, N. Jaroszynska, U. Drescher, B. J. Janssen, R. Hindges, and D. H. Meijer, *Alternative splicing controls teneurin-3 compact dimer formation for neuronal recognition,* Nature Communications **15** (2024), 10.1038/s41467-024-47763-x.

[2] R. P. Tucker, J. Beckmann, N. T. Leachman, J. Schöler, and R. Chiquet-Ehrismann, *Phylogenetic analysis of the teneurins: Conserved features and premetazoan ancestry,* Molecular Biology and Evolution **29**, 1019 (2012).

[3] A. Cheung, K. E. Trevers, M. Reyes-Corral, P. Antinucci, and R. Hindges, *Expression and roles of teneurins in zebrafish,* Frontiers in Neuroscience **13** (2019), 10.3389/fnins.2019.00158.

[4] X. H. Zhou, O. Brandau, K. Feng, T. Oohashi, Y. Ninomiya, U. Rauch, and R. Fässler, *The murine ten-m/odz genes show distinct but overlapping expression patterns during development and in adult brain,* Gene Expression Patterns **3**, 397 (2003).

[5] J. Beckmann, R. Schubert, R. Chiquet-Ehrismann, and D. J. Müller, *Deciphering teneurin domains that facilitate cellular recognition, cell-cell adhesion, and neurite outgrowth using atomic force microscopy-based single-cell force spectroscopy,* Nano Letters **13**, 2937 (2013).

[6] A. Cheung, G. Schachermayer, A. Biehler, A. Wallis, M. Missaire, and R. Hindges, *Teneurin paralogues are able to localise synaptic sites driven by the intracellular domain and have the potential to form cis-heterodimers,* Frontiers in Neuroscience **16** (2022), 10.3389/fnins.2022.915149.

[7] D. H. Meijer, C. P. Frias, J. W. Beugelink, Y. N. Deurloo, and B. J. C. Janssen, *Teneurin4 dimer structures reveal a calcium-stabilized compact conformation supporting homomeric trans-interactions,* The EMBO Journal **41** (2022), 10.15252/embj.2020107505.

[8] D. T. Pederick, J. H. Lui, E. C. Gingrich, C. Xu, M. J. Wagner, Y. Liu, Z. He, S. R. Quake, and L. Luo, *Reciprocal repulsions instruct the precise assembly of parallel hippocampal networks,* SCIENCE **372**, 1068+ (2021).

[9] R. Sando, X. Jiang, and T. C. Südhof, *Latrophilin gpcrs direct synapse specificity by coincident binding of flrts and teneurins,* Science **363** (2019), 10.1126/science.aav7969.

[10] A. A. Boucard, S. Maxeiner, and T. C. Südhof, *Latrophilins function as heterophilic cell-adhesion molecules by binding to teneurins: Regulation by alternative splicing,* Journal of Biological Chemistry **289**, 387 (2014).

[11] D. S. Berns, L. A. DeNardo, D. T. Pederick, and L. Luo, *Teneurin-3 controls topographic circuit assembly in the hippocampus,* Nature **554**, 328 (2018).

[12] D. del Toro, M. A. Carrasquero-Ordaz, A. Chu, T. Ruff, M. Shahin, V. A. Jackson, M. Chavent, M. Berbeira-Santana, G. Seyit-Bremer, S. Brignani, R. Kaufmann, E. Lowe, R. Klein, and E. Seiradake, *Structural basis of teneurin-latrophilin interaction in repulsive guidance of migrating neurons,* Cell **180**, 323 (2020).

[13] O. Spead, T. Moreland, C. J. Weaver, I. D. Costa, B. Hegarty, K. L. Kramer, and F. E. Poulain, *Teneurin trans-axonal signaling prunes topographically missorted axons,* Cell Reports **42** (2023), 10.1016/j.celrep.2023.112192.

[14] T. J. Mosca, *On the teneurin track: A new synaptic organization molecule emerges,* Frontiers in Cellular Neuroscience **9** (2015), 10.3389/fncel.2015.00204.

[15] R. K. Yuen, B. Thiruvahindrapuram, D. Merico, S. Walker, K. Tammimies, N. Hoang, C. Chrysler, T. Nalpathamkalam, G. Pellecchia, Y. Liu, M. J. Gazzellone, L. D'Abate, E. Deneault, J. L. Howe, R. S. Liu, A. Thompson, M. Zarrei, M. Uddin, C. R. Marshall, R. H. Ring, L. Zwaigenbaum, P. N. Ray, R. Weksberg, M. T. Carter, B. A. Fernandez, W. Roberts, P. Szatmari, and S. W. Scherer, *Whole-genome sequencing of quartet families with autism spectrum disorder,* Nature Medicine **21**, 185 (2015).

[16] N. R. Wray, S. Ripke, M. Mattheisen, M. Trzaskowski, E. M. Byrne, A. Abdellaoui, M. J. Adams, E. Agerbo, T. M. Air, T. M. Andlauer, S. A. Bacanu, M. Bækvad-Hansen, A. F. Beekman, T. B. Bigdeli, E. B. Binder, D. R. Blackwood, J. Bryois, H. N. Buttenschøn, J. Bybjerg-Grauholm, N. Cai, E. Castelao, J. H. Christensen, T. K. Clarke, J. I. Coleman, L. Colodro-Conde, B. Couvy-Duchesne, N. Craddock, G. E. Crawford, C. A. Crowley, H. S. Dashti, G. Davies, I. J. Deary, F. Degenhardt, E. M. Derks, N. DIrek, C. V. Dolan, E. C. Dunn, T. C. Eley, N. Eriksson, V. Escott-Price, F. H. F. Kiadeh, H. K. Finucane, A. J. Forstner, J. Frank, H. A. Gaspar, M. Gill, P. Giusti-Rodríguez, F. S. Goes, S. D. Gordon, J. Grove, L. S. Hall, E. Hannon, C. S. Hansen, T. F. Hansen, S. Herms, I. B. Hickie, P. Hoffmann, G. Homuth, C. Horn, J. J. Hottenga, D. M. Hougaard, M. Hu, C. L. Hyde, M. Ising, R. Jansen, F. Jin, E. Jorgenson, J. A. Knowles, I. S. Kohane, J. Kraft, W. W. Kretzschmar, J. Krogh, Z. Kutalik, J. M. Lane, Y. Li, Y. Li, P. A. Lind, X. Liu, L. Lu, D. J. MacIntyre, D. F. MacKinnon, R. M. Maier, W. Maier, J. Marchini, H. Mbarek, P. McGrath, P. McGuffin, S. E. Medland, D. Mehta, C. M. Middeldorp, E. Mihailov, Y. Milaneschi, L. Milani, J. Mill, F. M. Mondimore, G. W. Montgomery, S. Mostafavi, N. Mullins, M. Nauck, B. Ng, M. G. Nivard, D. R. Nyholt, P. F. O'Reilly, H. Oskarsson, M. J. Owen, J. N. Painter, C. B. Pedersen, M. G. Pedersen, R. E. Peterson, E. Pettersson, W. J. Peyrot, G. Pistis, D. Posthuma, S. M. Purcell, J. A. Quiroz, P. Qvist, J. P. Rice, B. P. Riley, M. Rivera, S. S. Mirza, R. Saxena, R. Schoevers, E. C. Schulte, L. Shen, J. Shi, S. I. Shyn, E. Sigurdsson, G. B. Sinnamon, J. H. Smit, D. J. Smith, H. Stefansson, S. Steinberg, C. A. Stockmeier, F. Streit, J. Strohmaier, K. E. Tansey, H. Teismann, A. Teumer, W. Thompson, P. A. Thomson, T. E. Thorgeirsson, C. Tian, M. Traylor, J. Treutlein, V. Trubetskoy, A. G. Uitterlinden, D. Umbricht, S. V. D. Auwera, A. M. V. Hemert, A. Viktorin, P. M. Visscher, Y. Wang, B. T. Webb, S. M. Weinsheimer, J. Wellmann, G. Willemsen, S. H. Witt, Y. Wu, H. S. Xi, J. Yang, F. Zhang, V. Arolt, B. T. Baune, K. Berger, D. I. Boomsma, S. Cichon, U. Dannlowski, E. C. D. Geus, J. R. Depaulo, E. Domenici, K. Domschke, T. Esko, H. J. Grabe, S. P. Hamilton, C. Hayward, A. C. Heath, D. A. Hinds, K. S. Kendler, S. Kloiber, G. Lewis, Q. S. Li,

S. Lucae, P. F. Madden, P. K. Magnusson, N. G. Martin, A. M. McIntosh, A. Metspalu, O. Mors, P. B. Mortensen, B. Müller-Myhsok, M. Nordentoft, M. M. Nöthen, M. C. O'Donovan, S. A. Paciga, N. L. Pedersen, B. W. Penninx, R. H. Perlis, D. J. Porteous, J. B. Potash, M. Preisig, M. Rietschel, C. Schaefer, T. G. Schulze, J. W. Smoller, K. Stefansson, H. Tiemeier, R. Uher, H. Völzke, M. M. Weissman, T. Werge, A. R. Winslow, C. M. Lewis, D. F. Levinson, G. Breen, A. D. Børglum, and P. F. Sullivan, *Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression,* Nature Genetics **50**, 668 (2018).

[17] B. Singh, P. Srivastava, and S. R. Phadke, *Sequence variations in tenm3 gene causing eye anomalies with intellectual disability: Expanding the phenotypic spectrum,* European Journal of Medical Genetics **62**, 61 (2019).

[18] P. P. Burbach and D. H. Meijer, *Latrophilin's social protein network,* Frontiers in Neuroscience **13** (2019), 10.3389/fnins.2019.00643.

[19] X. Yi, M. Li, G. He, H. Du, X. Li, D. Cao, L. Wang, X. Wu, F. Yang, X. Chen, L. He, Y. Ping, and D. Zhou, *Genetic and functional analysis reveals tenm4 contributes to schizophrenia,* iScience **24** (2021), 10.1016/j.isci.2021.103063.

[20] A. D. Minet, B. P. Rubin, R. P. Tucker, S. Baumgartner, and R. Chiquet-Ehrismann, *Teneurin-1, a vertebrate homologue of the drosophila pair-rule gene ten-m, is a neuronal protein with a novel type of heparin-binding domain,* Journal of Cell Science **112**, 2019 (1999).

[21] S. M. Nunes, J. Ferralli, K. Choi, M. Brown-Luedi, A. D. Minet, and R. Chiquet-Ehrismann, *The intracellular domain of teneurin-1 interacts with mbd1 and cap/ponsin resulting in subcellular codistribution and translocation to the nuclear matrix,* Experimental Cell Research **305**, 122 (2005).

[22] C. Bagutti, G. Forro, J. Ferralli, B. Rubin, and R. Chiquet-Ehrismann, *The intracellular domain of teneurin-2 has a nuclear function and represses zic-1-mediated transcription,* Journal of Cell Science **116**, 2957 (2003).

[23] D. Kenzelmann, R. Chiquet-Ehrismann, N. T. Leachman, and R. P. Tucker, *Teneurin-1 is expressed in interconnected regions of the developing brain and is processed in vivo,* BMC Developmental Biology **8**, 1 (2008).

[24] V. A. Jackson, D. H. Meijer, M. Carrasquero, L. S. V. Bezouwen, E. D. Lowe, C. Kleanthous, B. J. Janssen, and E. Seiradake, *Structures of teneurin adhesion receptors reveal an ancient fold for cell-cell interaction,* Nature Communications **9** (2018), 10.1038/s41467-018-03460-0.

[25] J. Li, M. Shalev-Benami, R. Sando, X. Jiang, A. Kibrom, J. Wang, K. Leon, C. Katanski, O. Nazarko, Y. C. Lu, T. C. Südhof, G. Skiniotis, and D. Araç, *Structural basis for teneurin function in circuit-wiring: A toxin motif at the synapse,* Cell **173**, 735 (2018).

[26] T. L. Dodsworth and D. A. Lovejoy, *Role of teneurin c-terminal associated peptides (tcap) on intercellular adhesion and communication,* Frontiers in Neuroscience **16** (2022), 10.3389/fnins.2022.868541.

[27] K. Feng, X. H. Zhou, T. Oohashi, M. Mörgelin, A. Lustig, S. Hirakawa, Y. Ninomiya, J. Engel, U. Rauch, and R. Fässler, *All four members of the ten-m/odz family of transmembrane proteins form dimers,* Journal of Biological Chemistry **277**, 26128 (2002).

[28] T. Oohashi, X.-H. Zhou, K. Feng, B. Richter, M. Mörgelin, M. T. Perez, W.-D. Su, R. Chiquet-Ehrismann, U. Rauch, and R. Fässler, *Mouse ten-m/odz is a new family of dimeric type ii transmembrane proteins expressed in many tissues,* The Journal of Cell Biology **145**, 563 (1999).

[29] J. Li, S. J. Bandekar, and D. Araç, *The structure of fly teneurin-m reveals an asymmetric self-assembly that allows expansion into zippers,* EMBO reports **24** (2023), 10.15252/embr.202256728.

[30] J. Li, Y. Xie, S. Cornelius, X. Jiang, R. Sando, S. P. Kordon, M. Pan, K. Leon, T. C. Südhof, M. Zhao, and D. Araç, *Alternative splicing controls teneurin-latrophilin interaction and synapse specificity by a shape-shifting mechanism,* Nature Communications **11** (2020), 10.1038/s41467-020-16029-7.

[31] D. Araç, A. A. Boucard, E. Özkan, P. Strop, E. Newell, T. C. Südhof, and A. T. Brunger, *Structures of neuroligin-1 and the neuroligin-1/neurexin-1β complex reveal specific protein-protein and protein-ca2+ interactions,* Neuron **56**, 992 (2007).

[32] J. Cheng, Y. Yu, X. Wang, X. Zheng, T. Liu, D. Hu, Y. Jin, Y. Lai, T. M. Fu, and Q. Chen, *Structural basis for the self-recognition of sdscam in chelicerata,* Nature Communications **14** (2023), 10.1038/s41467-023-38205-1.

[33] S. A. Li, L. Cheng, Y. Yu, J. H. Wang, and Q. Chen, *Structural basis of dscam1 homodimerization: Insights into context constraint for protein recognition,* Science Advances **2** (2016), 10.1126/sciadv.1501118.

[34] Z. Lin, J. Liu, H. Ding, F. Xu, and H. Liu, *Structural basis of salm5-induced ptpδ dimerization for synaptic differentiation,* Nature Communications **9** (2018), 10.1038/s41467-017-02414-2.

[35] R. Meijers, R. Puettmann-Holgado, G. Skiniotis, J. H. Liu, T. Walz, J. H. Wang, and D. Schmucker, *Structural basis of dscam isoform specificity,* Nature **449**, 487 (2007).

[36] M. R. Sawaya, W. M. Wojtowicz, I. Andre, B. Qian, W. Wu, D. Baker, D. Eisenberg, and S. L. Zipursky, *A double s shape provides the structural basis for the extraordinary binding specificity of dscam isoforms,* Cell **134**, 1007 (2008).

[37] J. W. Um, K. H. Kim, B. S. Park, Y. Choi, D. Kim, C. Y. Kim, S. J. Kim, M. Kim, J. S. Ko, S. G. Lee, G. Choii, J. Nam, W. D. Heo, E. Kim, J. O. Lee, J. Ko, and H. M. Kim, *Structural basis for lar-rptp/slitrk complex-mediated synaptic adhesion,* Nature Communications **5** (2014), 10.1038/ncomms6423.

[38] S. C. Wilson, K. I. White, Q. Zhou, R. A. Pfuetzner, U. B. Choi, T. C. Südhof, and A. T. Brunger, *Structures of neurexophilin–neurexin complexes reveal a regulatory mechanism of alternative splicing,* The EMBO Journal **38** (2019), 10.15252/embj.2019101603.

[39] A. Yamagata, Y. Sato, S. Goto-Ito, T. Uemura, A. Maeda, T. Shiroshima, T. Yoshida, and S. Fukai, *Structure of slitrk2-ptp complex reveals mechanisms for splicing-dependent trans-synaptic adhesion,* Scientific Reports **5** (2015), 10.1038/srep09686.

[40] A. Yamagata, T. Yoshida, Y. Sato, S. Goto-Ito, T. Uemura, A. Maeda, T. Shiroshima, S. Iwasawa-Okamoto, H. Mori, M. Mishina, and S. Fukai, *Mechanisms of splicing-dependent trans-synaptic adhesion by ptpδ -il1rapl1/il-1racp for synaptic differentiation,* Nature Communications **6** (2015), 10.1038/ncomms7926.

[41] V. A. Jackson, S. Mehmood, M. Chavent, P. Roversi, M. Carrasquero, D. del Toro, G. Seyit-Bremer, F. M. Ranaivoson, D. Comoletti, M. S. Sansom, C. V. Robinson, R. Klein, and E. Seiradake, *Super-complexes of adhesion gpcrs and neural guidance receptors,* Nature Communications **11184**, 1 (2016).

[42] B. Knöll, C. Weinl, A. Nordheim, and F. Bonhoeffer, *Stripe assay to examine axonal guidance and cell migration,* Nature Protocols **2**, 1216 (2007).

[43] J. Walter, B. Kern-Veits, J. Huf, B. Stolze, and F. Bonhoeffer, *Recognition of position-specific properties of tectai cell membranes by retinal axons in vitro,* Development **101**, 685 (1987).

[44] B. Chih, L. Gollan, and P. Scheiffele, *Alternative splicing controls selective trans-synaptic interactions of the neuroligin-neurexin complex,* Neuron **51**, 171 (2006).

[45] E. Furlanis and P. Scheiffele, *Regulation of neuronal differentiation, function, and plasticity by alternative splicing,* Annual Review of Cell and Developmental Biology Annu. Rev. Cell Dev. Biol **34** (2018), 10.1146/annurev-cellbio-100617.

[46] Y. Li, P. Zhang, T. Y. Choi, S. K. Park, H. Park, E. J. Lee, D. Lee, J. D. Roh, W. Mah, R. Kim, Y. Kim, H. Kwon, Y. C. Bae, S. Y. Choi, A. M. Craig, and E. Kim, *Splicing-dependent trans-synaptic salm3-lar-rptp interactions regulate excitatory synapse development and locomotion,* Cell Reports **12**, 1618 (2015).

[47] S. Suzuki, N. Ayukawa, C. Okada, M. Tanaka, S. Takekoshi, Y. Iijima, and T. Iijima, *Spatio-temporal and dynamic regulation of neurofascin alternative splicing in mouse cerebellar neurons,* Scientific Reports **7**, 1 (2017).

[48] S. Karki, A. V. Shkumatov, S. Bae, H. Kim, J. Ko, and T. Kajander, *Structural basis of salm3 dimerization and synaptic adhesion complex formation with ptpσ,* Scientific Reports **10** (2020), 10.1038/s41598-020-68502-4.

[49] 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 **5** (2014), 10.1038/ncomms6209.

[50] W. M. Wojtowicz, W. Wu, I. Andre, B. Qian, D. Baker, and S. L. Zipursky, *A vast repertoire of dscam binding specificities arises from modular interactions of variable ig domains,* Cell **130**, 1134 (2007).

[51] T. Biederer, P. S. Kaeser, and T. A. Blanpied, *Transcellular nanoalignment of synaptic function,* Neuron **96**, 680 (2017).

[52] B. A. Lloyd, Y. Han, R. Roth, B. Zhang, and J. Aoto, *Neurexin-3 subsynaptic densities are spatially distinct from neurexin-1 and essential for excitatory synapse nanoscale organization in the hippocampus,* Nature Communications **14** (2023), 10.1038/s41467-023-40419-2.

[53] J. H. Trotter, J. Hao, S. Maxeiner, T. Tsetsenis, Z. Liu, X. Zhuang, and T. C. Südhof, *Synaptic neurexin-1 assembles into dynamically regulated active zone nanoclusters,* Journal of Cell Biology **218**, 2677 (2019).

[54] X. Zhang, P. Y. Lin, K. Liakath-Ali, and T. C. Südhof, *Teneurins assemble into presynaptic nanoclusters that promote synapse formation via postsynaptic non-teneurin ligands,* Nature Communications **13** (2022), 10.1038/s41467-022-29751-1.

[55] S. Q. Zheng, E. Palovcak, J. P. Armache, K. A. Verba, Y. Cheng, and D. A. Agard, *Motioncor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy,* Nature Methods **14**, 331 (2017).

[56] K. Zhang, *Gctf: Real-time ctf determination and correction,* Journal of Structural Biology **193**, 1 (2016).

[57] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, *Cryosparc: Algorithms for rapid unsupervised cryo-em structure determination,* Nature Methods **14**, 290 (2017).

[58] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger, *Colabfold: making protein folding accessible to all,* Nature Methods **19**, 679 (2022).

[59] 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 **66**, 213 (2010).

[60] P. Emsley and K. Cowtan, *Coot: Model-building tools for molecular graphics,* Acta Crystallographica Section D: Biological Crystallography **60**, 2126 (2004).

[61] C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, and D. C. Richardson, *Molprobity: More and better reference data for improved all-atom structure validation,* Protein Science **27**, 293 (2018).

[62] E. Krissinel and K. Henrick, *Inference of macromolecular assemblies from crystalline state,* Journal of Molecular Biology **372**, 774 (2007).

[63] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, and T. E. Ferrin, *Ucsf chimerax: Structure visualization for researchers, educators, and developers,* Protein Science **30**, 70 (2021).

[64] 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 **7** (2011), 10.1038/msb.2011.75.

[65] P. Pernot, A. Round, R. Barrett, A. D. M. Antolinos, A. Gobbo, E. Gordon, J. Huet, J. Kieffer, M. Lentini, M. Mattenet, C. Morawe, C. Mueller-Dieckmann, S. Ohlsson, W. Schmid, J. Surr, P. Theveneau, L. Zerrad, and S. McSweeney, *Upgraded esrf bm29 beamline for saxs on macromolecules in solution,* Journal of Synchrotron Radiation **20**, 660 (2013).

[66] P. V. Konarev, V. V. Volkov, A. V. Sokolova, M. H. Koch, and D. I. Svergun, *Primus: A windows pc-based system for small-angle scattering data analysis,* Journal of Applied Crystallography **36**, 1277 (2003).

[67] D. Durand, C. Vivès, D. Cannella, J. Pérez, E. Pebay-Peyroula, P. Vachette, and F. Fieschi, *Nadph oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers,* Journal of Structural Biology **169**, 45 (2010).

[68] M. V. Petoukhov, D. Franke, A. V. Shkumatov, G. Tria, A. G. Kikhney, M. Gajda, C. Gorba, H. D. Mertens, P. V. Konarev, and D. I. Svergun, *New developments in the atsas program package for small-angle scattering data analysis,* Journal of Applied Crystallography **45**, 342 (2012).

[69] D. Svergun, C. Barberato, and M. H. J. Koch, *Crysol-a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates,* J Appl. Cryst **28**, 768 (1995).

[70] K. Manalastas-Cantos, P. V. Konarev, N. R. Hajizadeh, A. G. Kikhney, M. V. Petoukhov, D. S. Molodenskiy, A. Panjkovich, H. D. Mertens, A. Gruzinov, C. Borges, C. M. Jeffries, D. I. Sverguna, and D. Frankea, *Atsas 3.0: Expanded functionality and new tools for small-angle scattering data analysis,* Journal of Applied Crystallography **54**, 343 (2021).

[71] L. M. Chataigner, C. Gogou, M. A. den Boer, C. P. Frias, D. M. Thies-Weesie, J. C. Granneman, A. J. Heck, D. H. Meijer, and B. J. Janssen, *Structural insights into the contactin 1 – neurofascin 155 adhesion complex,* Nature Communications **13** (2022), 10.1038/s41467-022-34302-9.

[72] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, *Fiji: An open-source platform for biological-image analysis,* Nature Methods **9**, 676 (2012).

3

MECHANISTIC INSIGHTS INTO MICROPHTHALMIA VIA A TENEURIN-3 MUTATION

Neurons use physical interactions between cell adhesion molecules (CAMs) on their cell surface as guiding cues to ensure their proper integration into neuronal networks. Disruption of trans-cellular protein-protein interactions can lead to miswiring of the nervous system with clinical implications. Teneurin-3 is a synaptic CAM important for axon guidance and development of the visual system. Various teneurin-3 mutations are associated with microphthalmia, a disease characterised by severe visual and cognitive impairments. Here, we show that the microphthalmia-associated arginine-to-tryptophan missense mutation in the ABD domain is localised directly at an intramolecular contact that is required for teneurin-3 compact dimerisation in cis. We use cell clustering assays to show that this mutation completely disrupts homophilic trans interactions, likely due to complex destabilisation in cis. Molecular dynamics simulations reveal how the mutation possibly stabilises a β-strand that needs to be displaced for the compact cis-dimer to form. Interestingly, the missense mutation only affects the two particular splice variants that adopt this specific compact dimeric conformation out of the four known teneurin-3 splicing isoforms in total. Based on these insights, we provide a model for how disruption of cis-complex formation may affect binding in trans. Together, these data delineate the pathological impact of a single point mutation on the intricate wiring of the visual system in the brain.

This chapter is in preparation for publication as *Christos Gogou*, Cátia P. Frias, Stephany Laclé, Nikolina Šoštarić, and Dimphna H. Meijer, "Mechanistic insights into microphthalmia from a teneurin-3 point mutation."

3.1. INTRODUCTION

Microphthalmia is a congenital neurological disease that is typified by severely malformed eyes, often in combination with delayed cognitive development. Microphthalmia occurs in up to an estimated 14 per 100,000 newborns as either a stand-alone disease or as part of one of ~200 syndromes[1–3]. Roughly 30% of the syndromic patients exhibit brain malformation, and a similar percentage displays intellectual disabilities[2]. According to the ClinVar database, 22 pathological missense mutations in various genes are associated with this disease[4,5]. One such pathological missense mutation is the R2563W mutation in the *TENM3* gene on chromosome 4 that encodes teneurin-3 (Ten3) in humans. Aside from the Ten3 R2563W missense mutation listed in ClinVar[5], various homozygous nonsense mutations in Ten3 are also reported in multiple microphthalmia case reports[6–10]. These nonsense mutations lead to truncated proteins that are usually non-functional, unstable and are quickly degraded in the cell[9]. The patient that carried the R2563W mutation also possessed an additional mutation in Ten3, A1349G, thereby classifying it as compound heterozygous. Since the patient's brother and father also carried the A1349G mutation, but did not display any disease symptoms, only R2563W is regarded as pathogenic.

Of the four family members of the type II transmembrane CAM family, Ten3 is involved in proper wiring of the hippocampus and visual system of different animal species [11–18]. In zebrafish, Antinucci and colleagues found that reducing Ten3 levels led to improper formation of the inner plexiform layer of the retina, and caused defects in connections between retinal ganglion cells (RGCs) and their target area, the tectal neuropil[13]. As a result, the fish had difficulty responding to certain directional and orientational inputs of visual stimuli. In mice that lack teneurin-3 expression, defects were observed in ipsilateral projections between RGCs and the dorsal lateral geniculate nucleus (dLGN), leading to difficulties in tasks that require binocular vision[18]. More recently, miswiring in teneurin-3 knockout mice was also observed between a light-sensitive subtype of RGCs (the intrinsically photosensitive RGCs or ipRGCs) and the suprachiasmic nucleus (SCN), a structure of the brain that regulates the circadian rhythm, resulting in higher sensitivity to light[16]. At the molecular level, a combination of repulsive interactions with adhesion G protein-coupled receptor (aGPCR) latrophilin (Lphn) together with attractive homophilic interactions act as guidance cues for the proper wiring[15]. For instance in the mouse hippocampus, Ten3-expressing axons from the proximal CA1 project across the Lphn-expressing distal CA1 and proximal subiculum, over to the distal subiculum where Ten3 is expressed[14,15]. *In vitro* studies with retinal neurons have recapitulated these attractive homophilic interactions, showing that Ten3-expressing neurons from chicken retinal explants preferentially grew on substrates coated with purified Ten3[19].

Structural biology studies of various Teneurin family members have revealed that Teneurins are composed of a large globular superfold at the extracellular side comprising multiple domains, starting from an N-terminal calcium-binding C-rich domain, a transthyretin (TTR)-like domain, a fibronectin domain (FN-plug), and followed by an NHL β-propeller[20–23]. The tyrosine- and aspartate-rich YD shell is the central and largest domain around which the other domains are stabilised. An additional linker domain threads the YD shell which upon exit is followed by the antibiotic-binding domain-like (ABD) and Tox-GHH-like domain. This full assembly is attached to a single-pass trans-

membrane helix via a an immunoglobulin (Ig) fold and 8 flexible repeats of an endothelial growth factor (EGF)-like domain. Many of the different domains are found to be involved in intramolecular contacts across teneurin homologues, as well as in intermolecular contacts in *cis* and in *trans* [19,20,22–25]. The YD shell, for instance, harbours a heterophilic binding site for Lphn [22,23].

The extracellular regions of mouse and human Ten3 contain two alternatively spliced regions, called splice sites A and B, in between EGF7-EGF8 and in the NHL domain, respectively. The resulting splice variants are referred to as A_0B_0 (both absent), A_1B_1 (both present), and A_0B_1 and A_1B_0 (insert B or insert A present, respectively). We previously solved the structures of the different isoforms, which provided the mechanistic basis for the splicing-dependent differences in *trans*-cellular clustering [19]. Shortly, in absence of splice insert A, splice insert B instigates a global reorganisation of the superfolds within the ECD, where each isoform forms a different 'compact dimer' contact between the superfolds. Insert A brings about yet another conformational shift that is independent of the presence of B. These splicing-dependent conformations are distinguished by isoform-specific intramolecular contacts.

In this work, we map the human microphthalmia-associated R2563W (A_0B_0 isoform; corresponding to R2570W, R2572W, and R2579W in A_0B_1, A_1B_0, and A_1B_1, respectively) to the EGF8-ABD intramolecular interface that is specific to the splice insert A-containing isoforms. We use *in vitro* cell clustering assays to show that introducing the microphthalmia (MO) mutation into full-length mouse teneurin-3 ($Ten3^{MO}$) results in complete loss of homophilic clustering for A_1 isoforms - but not for A_0 isoforms - and quantitative differences in binding with Lphns. Ten3 thereby appears a unique case in which a missense mutation is present in all isoforms but only affects a specific subset. We then use molecular dynamics simulations to show that the mutation locally interferes with the EGF8-ABD contact, and provide a model for how this influences the compact *cis* dimer and subsequent *trans*-cellular complexes. These structural and functional insights will help to determine the underlying mechanistic basis of microphthalmia, and possibly guide future treatments for the deleterious effects of this mutation.

3.2. RESULTS

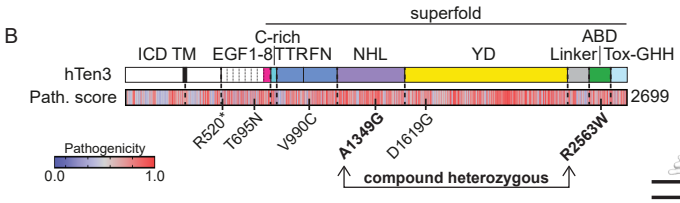
3.2.1. THE HIGH-PATHOGENICITY A_1 CONTACT HARBOURS THE MO MUTATION

In recent years, various teneurin-3 (Ten3) mutations in patients have been linked to the onset of microphthalmia (Fig. 3.1A-B). These mutations, except for R2563W (corresponding to A_0B_0), result in codon frameshifts quickly followed by a nonsense mutation (Fig. 3.1A-B). From the structure, it can be seen that most of the truncations result in a near-complete loss of the superfold that carries all the intra- and intermolecular binding sites identified thus far (Fig. 3.1C and Supplementary Fig. 3.7) [19,22–24].

To address the pathogenicity of any mutation along the Ten3 amino acid sequence, the AlphaMissense database was consulted (Fig. 3.1B) [26]. This algorithm predicts pathogenicity on basis of structural context, residue conservation across species, genetic variation among humans, and clinical data of mutant-disease association. The projection of pathogenicity scores onto the human Ten3 structure highlights the well-characterised

A

Mutation	Mutation type	Prematur stop codon site (aa)	Allelic status	Ref.
R520*	Nonsense	520	Homozygous	[6]
T695N	Frameshift	708	Homozygous	[7]
V990C	Frameshift	1003	Homozygous	[8]
D1695G	Frameshift	1697	Homozygous	[9]
A1349G	**Missense**	-	**Compound**	[10]
R2579W	**Missense**	-	**heterozygous**	



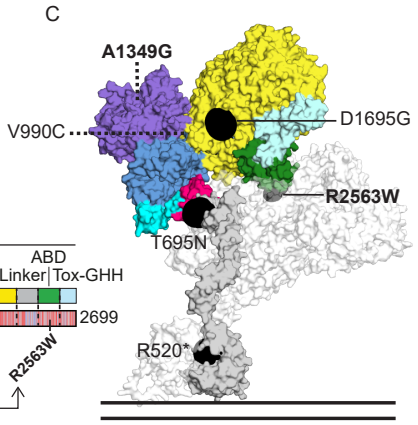


Figure 3.1: Microphthalmia-associated mutations in human teneurin-3 A_0B_0 isoform. **A)** Human teneurin-3 (hTen3) mutations found inpatients with microphthalmia. **B)** Linear representations of hTen3 domains and pathogenicity. The domains participating in the superfold are indicated by a vertical line above. Top: Protein domains in the full-length hTen3. Positions of the dimerising cystine at C549 in EGF2 and C648 in EGF5 are not shown. Domain abbreviations are shown: ICD, intracellular domain; TM, transmembrane helix; Ig, immunoglobulin fold; EGF, epidermal growth factor repeat domains 1 through 8; C-rich, cysteine-rich region; TTR, transthyretin-related; FN, fibronectin plug; NHL, NCL, HT2A and Lin-41; YD, tyrosine-aspartate; ABD, antibiotic-binding domain; Tox-GHH, toxin-glycine-histidine-histidine. Bottom: Average pathogenicity score of any substitution along the human hTen3 amino acid sequence. Pathogenicity scores are colour coded according to legend. Pathogenic mutations are indicated below the sequence. **C)** Surface representation of a model of hTen3 extracellular domain in compact dimer configuration. One covalently linked monomer is coloured corresponding to the linear representation in panel A, and the other is shown as transparent surface. Black spheres represent the mutations. Black dashed lines indicated mutations that are buried in the structure. Vertical black line represents the membrane. The unique compound heterozygous missense mutations are indicated in bold in all panels.

binding site for Lphn[22,23] as highly pathogenic (Supplementary Fig. 3.7A). We further assessed the predicted pathogenicity of splice isoform-specific intramolecular contacts (Supplementary Fig. 3.7B-E). The average pathogenicity scores of residues in these combined interfaces is 0.78 ± 0.009 compared to 0.68 ± 0.005 (mean ± SEM) for superfold residues (EGF through Tox-GHH domains) that are not involved in binding. The compact dimer of A_1-containing structures is connected through EGF8-ABD, YD-ABD, and YD-YD contacts, and the average pathogenicity for these interface residues is 0.72 ± 0.023 (mean ± SEM). With this, any random amino acid substitution is predicted to be most pathogenic in the superfold when compared to any other region – except for the transmembrane helix – of the full-length hTen3 (Supplementary Fig. 3.7E).

A closer look at the interfaces of the compact dimers of the different Ten3 mouse isoforms reveals an interesting difference between the A_0 and the A_1 variants. A solvent-exposed β-strand (β_{rival}) needs to be displaced so that EGF8 is able to form an extended β-sheet with the opposing ABD domain, thereby forming the specific A_1 interface (Fig. 3.2A). In case of A_0 variants, β_{rival} is a stable strand of the ABD domain. Simultaneous binding of EGF8 and β_{rival} would result in a steric clash between them (Supplementary Fig. 3.7C). Importantly, β_{rival} is directly flanked by R2563, one of the microphthalmia-associated amino acids (Fig. 3.1A). Note that although AlphaMissense predicts high pathogenicity for a tryptophan substitution at any position throughout hTen3 domains (Fig. 3.2C and Supplementary Fig. 3.7B), an R2563W substitution is predicted to be am-

biguous in pathogenicity (Fig. 3.2B). This is slightly surprising, as 12/19 possible mutations at this specific position are predicted to be pathogenic, and a solvent-exposed tryptophan may structurally be expected to be quite disruptive.

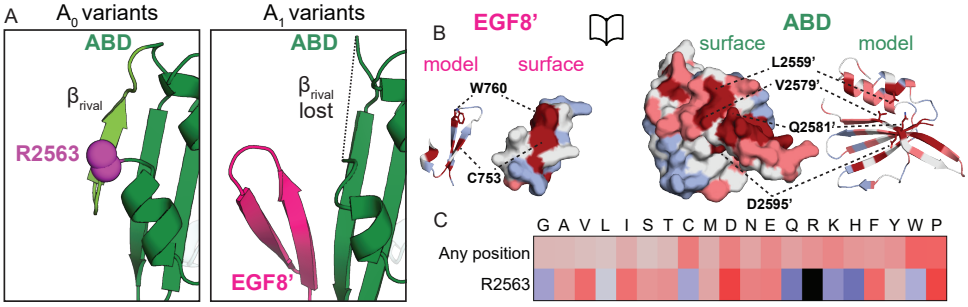


Figure 3.2: Elevated pathogenicity for EGF8-ABD contact but not for R2563, specifically **A)** The difference between the ABD domain in the A_0 and the A_1 compact dimer interfaces. The A_0 ABD (dark green; PDB:8R54) has an extra solvent-exposed β-strand (β_{rival}, light green), which is seen in all structural data of single-subunit reconstructions[19–25]. For the formation of the A_1 compact dimer (PDB:8R50), this strand becomes flexible to allow EGF8 (magenta) β-sheet extension with ABD, and β_{rival} is not resolved anymore. The structural representations are taken from chapter2 Supplementary Fig. 2.11B. **B)** Pathogenicity comparison at any position along the hTen3 amino acid sequence and at position R2563, specifically. A black box denotes a synonymous mutation. **C)** Open book surface representation of predicted pathogenicity at the EGF8-ABD interface. Corresponding models are shown besides the surface representations. Hydrophobic, as well as H-bond- and salt bridge-forming interface residues are labelled and indicated with black dashed lines. Residue numbers correspond to the A_0B_0 positional equivalents of those found in the mouse A_1B_1 compact dimer.Pathogenicity colours in panels B, and C, correspond to the legend in Fig. 3.1B.

Further inspection of the EGF8-ABD interface as a whole does, however, recapitulate high pathogenicity for the residues that were identified as contributing to the interface (Fig. 3.2C). Especially the hydrophobic residues on either side are predicted highly pathogenic, as is the D2595 in ABD that forms a salt bridge with R752 (human homologues of D2611-R761) on EGF8.

Together, these data indicate quite a strong correspondence between AlphaMissense predictions and the known binding sites on the Ten3 superfold. The residues involved in the EGF8-ABD contact are predicted to be sensitive to mutations, while the role of R2563W that directly flanks this contact is computationally not specifically highlighted.

3.2.2. A_1 CELL CLUSTERING IS SPECIFICALLY AFFECTED BY MO MUTATION

We next introduced the microphthalmia (MO) mutation into the full-length mouse Ten3 constructs ($Ten3^{MO}$) encoding the four different splice variants. The MO mutation is located on the β-strand that needs to be dislodged to form the A_1 compact dimer interfaces (Fig. 3.2A). We hypothesised that changing a solvent-exposed positively charged arginine into a large and hydrophobic tryptophan might interfere with the displacement of β_{rival}, thereby disrupting *cis* dimers and affecting the *trans*-cellular complex formation of the A_1 variants.

To test this hypothesis, the full-length versions of the four Ten3 wildtype (WT) isoforms and their respective $Ten3^{MO}$ mutants the were expressed in K562 hematopoietic cancer cells, that by themselves do not express cell adhesion molecules and are therefore natively single cells in suspension. As previously published, ectopic expression of Ten3 A_0B_1, A_1B_0, and A_1B_1 on the membrane prompts cell clustering, whereas cells expressing

A_0B_0 remain single cells in suspension [14,15,19,22] (Fig. 3.3A-B). A_0B_1 homophilic binding appears strongest through the formation of the most and largest clusters, while the A_1 variants behave comparable to each other in forming fewer and smaller clusters. Strikingly, when comparing the four WT Ten3 isoforms to their respective $Ten3^{MO}$ mutants, clustering is almost completely eliminated in A_1 -containing splice variants (Fig. 3.3A-B). The total number of clusters is significantly reduced, and the sizes of the remaining clusters also dropped. Together these numbers show a sharp decrease in number of clusters, cluster sizes and clustering index for cells expressing $Ten3^{MO}$ A_1 variants. A small quantitative effect on $Ten3^{MO}$-A_0B_1 cluster count and size is also observed. No defects in trafficking of $Ten3^{MO}$ to the membrane are observed (Supplementary Fig. 3.8A).

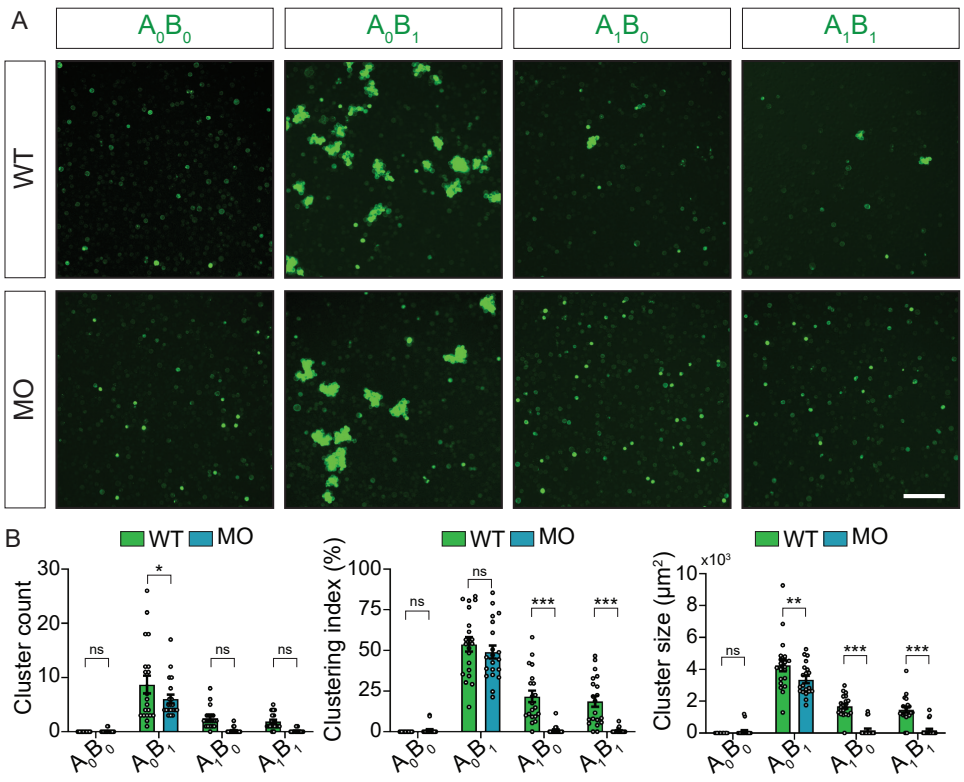


Figure 3.3: The microphthalmia mutation abrogates homophilic *trans*-cellular interactions for the teneurin-3 isoforms containing splice insert A. **A)** Representative homophilic clustering data of K562 cells expressing the wildtype (WT) and microphthalmia mutant (MO) full-length teneurin-3 (Ten3) isoforms indicated above. Scale bar is 100 µm. **B)** Quantification of clustering data in panel A. Bar plots represent the number of clusters found (Cluster count), the percentage these clusters contribute to the total segmented cell area (Clustering index) per image, and the average sizes of the clusters (Cluster size) per image (n=5 images per N=4 independent experiments, 20 images total). Data are presented as mean ± SEM; ns: not significant, *: p > 0.05, **: p > 0.01, ***: p > 0.001. Two-Way ANOVA, Šìdák's multiple comparisons test.

The MO mutation is structurally not located near the Lphn binding site in any of the isoforms' compact dimeric organisations (Supplementary Fig. 3.7A), but the relative reorganisation of the Ten3 superfolds might still affect teneurin-latrophilin *trans*-

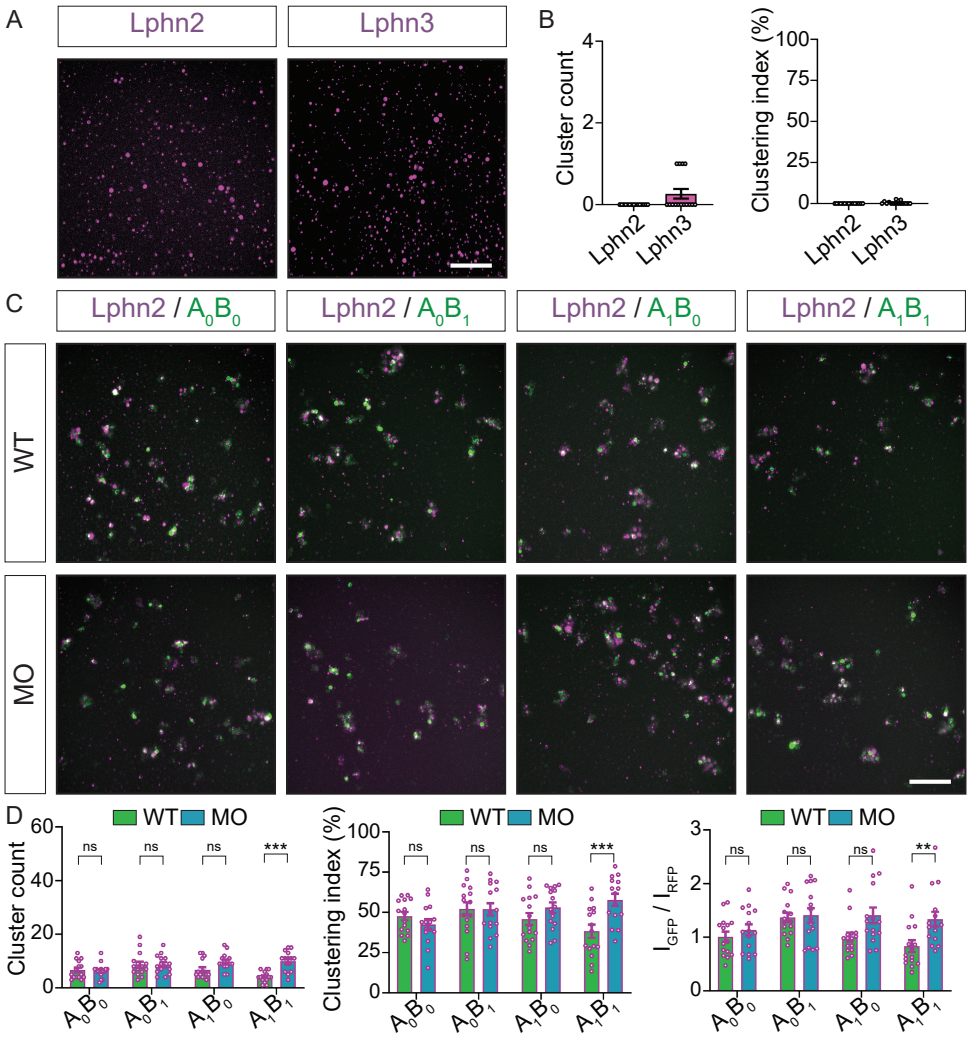


Figure 3.4: The microphthalmia mutation increases latrophilin-2 cluster formation for the teneurin-3 A_1 isoforms. **A)** Representative homophilic clustering data of K562 cells expressing full-length latrophilin-2 (Lphn2) or latrophilin-3 (Lphn3) only. Scale bar is 100 µm. **B)** Quantification of clustering data in panel A as in Figure 3.3B (n=5 images per N=3 independent experiments, 15 images total). **C)** Representative heterophilic clustering data of K562 cells expressing wildtype (WT) or microphthalmia mutant (MO) full-length teneurin-3 (Ten3) isoforms, and cells expressing latrophilin-2 (Lphn2). Scale bar is 100 µm. **D)** Quantification of clustering data in panel C. Bar plots represent the number of clusters found (Cluster count), the percentage these clusters contribute to the total segmented cell area (Clustering index) per image, and the per-image average co-clustering ratio (I_{GFP} / I_{RFP}) denoting the relative abundance of GFP signal per mCherry signal within each cluster (n=5 images per N=3 independent experiments, 15 images total). Data are presented as mean ± SEM; ns: not significant, **: p > 0.01, ***: p > 0.001. Two-Way ANOVA, Šìdák's multiple comparisons test.

cellular binding. Especially the dissociation of the A_1 compact dimers in *cis* might logically disrupt downstream *trans*-cellular Ten-Lphn binding. To test this, we expressed Lphn2 and Lphn3 in K562 as well (Fig. 3.4A-B and Supplementary Fig. 3.8B). As previ-

ously reported, both Lphn family members do not engage in homophilic *trans*-cellular interactions that can instigate cell clustering[14,15,22] (Fig. 3.3C-D and Supplementary Fig. 3.8B). When mixing Lphn2 or -3 with cells expressing either one of the Ten3 isoforms, heterotypic clusters are observed for all conditions (Fig. 3.4C-D and Supplementary Fig. 3.2C-D). Although Lphn displays signs of distorted membrane localisation after 48h experimentation, sufficient membrane expression is observed at earlier time points, and the clustering across conditions also underlines its proper membrane localisation (Fig. 3.4 and Supplementary Fig. 3.8C-D). The MO mutation does not qualitatively seem to alter teneurin-latrophilin heterophilic *trans*-cellular interactions(Fig. 3.4C and Supplementary Fig. 3.8C). Upon quantification, however, some differences are observed (Fig. 3.4D and Supplementary Fig. 3.8D). Lphn2 exhibits increased numbers of clusters, along with a higher clustering index for clustering with Ten3^{MO}-A_1B_0 and -A_1B_1 mutants compared to the respective WTs (Fig. 3.4D), but this difference was only found statistically significant for A_1B_1. Similarly, the relative abundance of GFP signal compared to RFP per cluster (I_{GFP} / I_{RFP}) is significantly elevated for this Ten3^{MO} isoform, and does not seem affected in A_1B_0 or the others. These differences are also observed for binding with Lphn3, but are not significant (Supplementary Fig. 3.8C-D). This difference is not explained by a reduced RFP signal, as the Lphn-expressing cells are shared between all conditions. The non-significant but elevated GFP/RFP intensity ratio for the Lphn3 binding with Ten3^{MO} A_1 is similarly not explained by increased A_1B_0-mutant fluorescence, as this is not changed in binding with Lphn2. Therefore, a relative increase is apparent of participation in heterophilic *trans*-cellular interactions by A_1 isoforms carrying the MO mutation, which results in a higher count and clustering index with Lphn2.

These findings indicate that the arginine-to-tryptophan subsitutation likely influences A_1 compact dimerisation in *cis* so that it prevents homophilic *trans* interactions, while also mildly affecting Lphn binding in *trans*.

3.2.3. MO MUTATION MAY OBSTRUCT THE EGF8-ABD CONTACT

To understand the mechanistic basis of the inhibitory effect on homophilic *trans* interactions, we employed all-atom molecular dynamics (MD) simulations to probe the effect of R2579W on the A_1B_1 compact *cis*-dimer. We carried out the molecular dynamics simulations using a partial ECD, and we focused on the A_1B_1 splice variant, which was shown to be structurally identical to A_1B_0 [19]. Therefore, we specifically refer to the MO mutation as R2579W in this section.

As described above, R2579 is located at the EGF8-ABD intramolecular interface of the A_1B_1 splice variant, (Fig 3.2A). This contact is not formed in the A_0B_0 and A_0B_1 compact dimers. R2579 itself is not resolved in the compact dimer structures of the A_1-containing splice variants, and is most likely part of an unstructured loop. To test the effect of the R2579W mutation, MD simulations were performed where both EGF8 and β_{rival} connected the ABD domain at initiation t = 0 ns (Supplementary Fig. 3.9A-B), where they are expected to hinder each other in competition for ABD-binding (Supplementary Fig. 3.7C). Structural inspection of this contact upon completion of the simulation (t = 550 ns) reveals that the mutated residue results in stabilisation of R2579W-flanking β_{rival} within the ABD domain(Fig. 3.5A-B). Furthermore, EGF8 on the interacting subunit is displaced, resulting in the disruption of the extended EGF8-ABD β-sheet that is char-

acteristic of the A_1-containing compact dimers. The R2579W substitution itself now interacts with H1627 in the YD shell of the other subunit – a hydrophobic contact with π-π stacking potential (Fig. 3.5B). This results in a reduced distance between their Cα's from 14.4 to 9.2 Å. The R2579W-H1627 interaction may stabilise β_{rival} in its competition with EGF8 for binding the ABD domain. The observed difference between WT and mutant in EGF8-ABD interface is quantified in the pair-wise distances between the interface residues (Fig. 3.5C). Conversely, quantifying the proximity of β_{rival} to the next β-strand in the ABD reveals that the WT amino acids near R2579 are located further away from the ABD domain - they are being dislodged (Fig. 3.5D). This also signifies formation of the extended EGF8-ABD β-sheet, whereas the residues in β_{rival} remain lodged in the ABD domain in the mutant. It should be noted that at our current simulation times, EGF8 is displaced from the ABD domain in only one of the two mutant contacts (Fig. 3.5C-D). At the other WT contact, both EGF8 and the competing β-strand are still connected to the ABD, similar to the initial state (Supplementary Fig. 3.9A).

What is the structural effect of the R2579W point mutation on a more global scale of the Ten3 compact dimer? The per-residue root-mean-squared fluctuation (RMSF) for the subunits in the WT versus the mutant reveals that the domains that are not involved in the compact dimer interface (the C-rich, TTR, Tox-GHH, and the C-terminal side of the YD shell) are slightly more flexible in the WT compared to the mutant (Fig. 3.5E). The largest difference in the per-residue binding free energy is found in those regions that do participate in the compact dimer interface, suggesting a global effect from the R2579W mutation on compact dimer stability (Fig. 3.5F). To further analyse differences in global movements between domains in the WT and mutant compact dimer, principal component analysis was performed on the MD simulation data. Modes 1 and 2 accounted for 27% and 17%, respectively, of the movement in the WT, and 30% and 13% in the mutant. For the Ten3 WT, the first mode of movement mainly describes the overall transition from in-plane to an angle between subunits – a 'hugging' movement (Fig. 3.5G and Supplementary Fig. 3.9C). The long axis of each subunit tilt in opposite direction relative to the initial plane. The WT's second mode principally describes a rotation of each subunit around its own long axis, with the result that the YD shells alternate between making contact and not – a 'butterfly' movement (Supplementary Fig. 3.9C). This latter movement best describes mode 1 of the mutant, whereas mode 2 of the mutant exhibits contraction of the domains within each subunit (data not shown). Notably, 3D variability analysis of the cryo-EM dataset from which the experimental compact dimer was reconstructed, reveals similar movements (Supplementary Fig. 3.9C) [19]. More specifically, experimental mode 1 aligns well with the hugging movement of the WTs simulated mode 1, whereas experiment mode 2 is highly comparable to the butterfly movements in mode 2 of the WT and mode 1 of the mutant. The simulations are thereby consistent with experimental data in capturing global intramolecular flexibility.

The relative organisation and flexibility of subunits within compact dimers is possibly of key importance for the formation of the higher-order oligomers needed for *trans*-cellular adhesion. Subunit reorganisation may then alter available binding sites for compact dimerisation in *cis*, with consequences for *trans*-cellular interactions. The simulations thus provide novel insights into how a single missense mutation can locally affect compact dimerisation, which may have strong implications for the formation of higher-

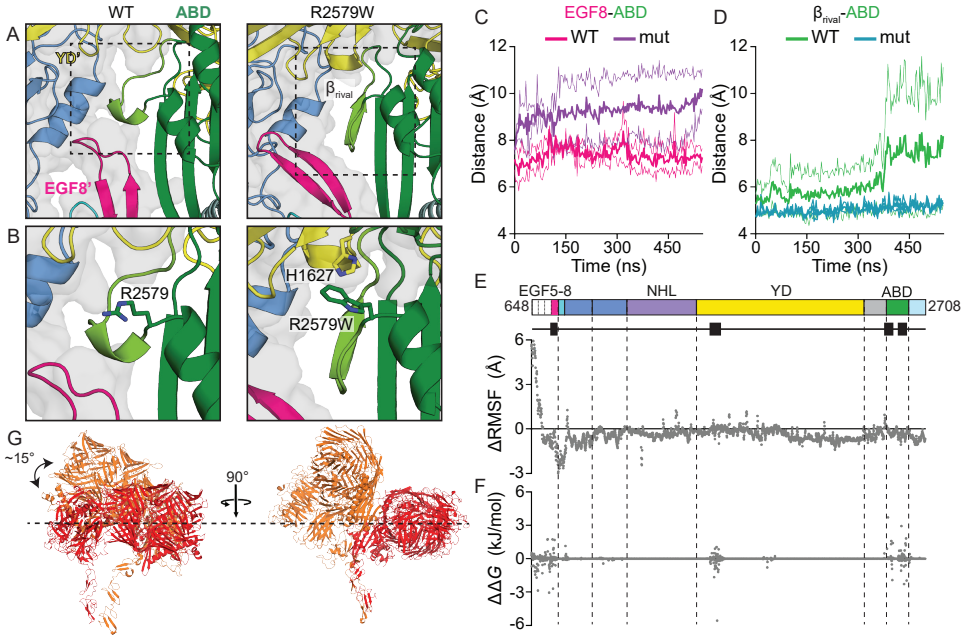


Figure 3.5: Molecular dynamics simulations reveal how EGF8 dislodges β_{rival} from the ABD. **A)** Comparison of the ABD (dark green) competitive binding between EGF8 (magenta) and β textsubscriptrival (light green) after 550 ns of simulated time. In the WT, the competing β_{rival} is dislodged from its ABD domain and transitions in secondary structure, while EGF8 remains ABD-bound. In the R2579 mutant, β_{rival} has remained fixed, while EGF8 has moved away from the ABD domain. Dashed boxes indicate the regions enlarged in panel B. **B)** Insets on dashed boxes in panel A. The native arginine 2579 is shown to be facing the solvent in the WT. Conversely, the tryptophan substitution R2579W at this position has moved inward to interact with histidine 1627 in the YD shell (yellow) of the other subunit. Domains are coloured corresponding to the linear representation in Fig. 3.1 A. Transparent surface representation is used to distinguish between subunits within the compact dimer. **C)** Distance between EGF8 of one subunit and the ABD domain of the other throughout simulated time for the WT vs. the R2579W mutant. These distances are averages of the pairwise distances between Cα's of β-strand-forming residues within T761-Ser764 (EGF8) and L2596-R2599 (ABD). **D)** Distance between the β_{rival} and the ABD domain within the same subunit over simulated time. Here, the averaged pair-wise distance between residues S2582-R2584 (β textsubscriptrival) and S2594-S2596 (ABD). Thin traces in panels C and D represent the two individual interfaces within on WT/mutant molecule, and thick traces in corresponding colours are averages of these two. **E)** Per-residue calculated difference in RMSF (ΔRMSF) between WT and mutant (mutant minus WT). **F)** Per-residue difference ($\Delta\Delta G$) in binding free energy (ΔG) between WT and mutant (mutant minus WT). Dashed lines through panels E and F indicate protein domain borders in the linear representation above. The linear representation only contains the domains that were included in the simulation. **G)** Cartoon representation of the wildtype (WT) 'hugging configuration' after 550 ns of simulated time. The subunits make a 15° angle relative to that between the subunits in the experimental A_1B_1 compact dimer (indicated as 'plane' along long axis of one subunit with dashed line).

order complexes that constitute *trans*-cellular binding.

3.3. DISCUSSION

In this chapter, we characterised a arginine-to-tryptophan substition in the ECD of teneurin-3 that has been associated with the eye disease microphthalmia (MO mutation): R2563W, R2570W, R2572W, and R2579W in A_0B_0, A_0B_1, A_1B_0, and A_1B_1, respectively. First, we demonstrated that this substitution is localised in the ABD domain, precisely where it interacts with EGF8 to form a compact dimer in the case of splice insert A-containing

teneurin-3 isoforms. We then consulted AlphaMissense (AM) to predict the pathogenicity of any mutation in the Ten3 amino acid sequence, which predicts that mutations in the intramolecular binding sites of the different isoforms are increased in pathogenenicity, including at the ABD-EGF8 contact. It should be noted, however, that AM does not provide any information on the mechanism by which such a mutation on the amino acid sequence level is translated into medical phenomena. We next showed that the MO mutants of the Ten3-A_1 isoforms specifically exhibit disrupted homophilic cell clustering. We also observed a small but significant effect on cell clustering of the missense mutation in the A_0B_1 isoform. This could be possibly explained by the postulated involvement of the EGF8-ABD contact in A_0B_1 dimer-of-dimer formation, as predicted in Chapter 2 Figure 2.6C) [19]. To explore why *trans*-cellular binding is affected by R2579W in A_1B_1 (and A_0B_1), we deployed molecular dynamics (MD) simulations. This revealed that the mutation could strongly impact the structure locally, with potential consequences for global subunit organisation, including dimer-of-dimer formation in *trans*.

As mentioned above, AM revealed elevated pathogenicity scores of the different intramolecular contacts, as well as of the latrophilin binding site. This is remarkable, as structural residues like those in the hydrophobic core of proteins - that are not part of the intra-or intermolecular interfaces in teneurin-3 - typically display elevated pathogenicity scores [27]. It should be noted that none of these interfaces were known when the AM neural network was trained [27]. This suggests an application for AM to identify putative interfaces, although other specific software [28,29] is available for this purpose as well. While AM predicted high pathogenicity for EGF8 residues involved in the A_1 EGF8-ABD contact, the output is ambiguous for the MO substitution. This could be explained by the dependency of AM on existing protein structures, where all publicly available teneurin structures have the competing β-strand (β_{rival}) in place that is absent from the ABD in the A_1 compact dimers. This subtle structural alteration may lead to a different AM pathogenicity score, that cannot be reproduced in the algorithm yet.

With our cell clustering data, we show that all Ten3 isoforms cluster with both Lphn2 and Lphn3. This is unlike teneurin-2, where the presence of B splice inserts disrupts heterophilic clustering [21,22]. Recent findings in fly Ten-m indicate that the effective level of teneurin proteins on pre- and postsynapses need to match for the contact to be established [30]. Possibly, the difference in expression levels between teneurin-2 and teneurin-3 may explain the difference in splice insert B sensitivity. Our data also show that interactions between cells expressing Ten3 isoforms and cells expression Lphn2 or Lphn3 are largely unaffected by introducing the MO mutation. A slight increase in clustering with Lphn-expressing cells, however, is observed for the mutant A_1 variants. Notably, this difference is only found statistically significant for A_1B_1 with Lphn2 (Fig. 3.4). Crucially, Pederick and colleagues demonstrated how the loss of either homophilic or heterophilic interactions with outgrowing Ten3-expressing axons leads to global miswiring in mouse hippocampus [15]. These findings suggests that also in the visual circuit, disruption of any of these sensitive balances in partner matching may greatly impact neuronal wiring, and possibly even lead to neurological defects, as observed in microphthalmia patients. Future axon outgrowth assays and animal models are needed to assess the impact of the missense mutation on neuronal wiring, that would likely result from the changes in *trans*-cellular binding as presented in this work.

The MD simulations in this chapter reveal global differences in principal modes of movement between the subunits in the compact dimer when comparing WT and mutant variants. This movement may be necessary for the protein to carry out its *trans*-cellular binding (Fig. 3.6A). Intrinsic freedom of subunits might even be required for WT binding specificity, where the stabilisation of a specific conformation in the mutant disrupts the higher-order complex formation required for homophilic *trans* interactions (Fig. 3.6B). It is also imaginable how the R2579W-induced dislodging of EGF8 from the ABD may eventually completely dissociate the compact dimer (Fig. 3.6B). We want to speculate that these dissociated subunits could reorganise into a conformation more akin to the A_0 variant. A conformational shift to the A_0B_0-specific conformation could also explain the loss of homophilic clustering. The observed increases in ${A_1}^{MO}$ heterophilic clustering with Lphn-expressing cells may be explained by the increased availability of Ten that has ceased forming homophilic interactions in *cis* - although no evidence can be provided for this mechanism at the moment. The duration of our simulations does not allow further insight into these larger conformational changes, however, as the time scale of larger conformational changes is typically in the µs scale[31]. Longer simulations are therefore desirable, and are expected to generate more insights into the global reorganisation of the subunits.

In addition, it should be noted that our simulations were performed without the initial placement of structural calcium ions in the C-rich domain[24], in the absence of glycans, and with incomplete N- and C-termini. Experimental data has shown that both Ten3 and -4 are stabilised by calcium[19,24], and this may be expected to affect the overall dynamics of the complex. Ten3 is also heavily glycosylated at its 17 N-linked glycosylation sites. These posttranslationalnal modifications (PTMs) are typically negatively charged and can add up to 2.5 kDa per glycosylated site, and this PTM is known to affect *trans*-neuronal complex formation[23,32–35]. The A_0B_0 and A_0B_1 isoforms would also be interesting candidates for MD simulation, possibly revealing how the presence of splice insert B in the NHL domain causes the global reorganisation of the superfold. Full-length versions of these extracellular domains could then capture putative NHL-Ig/EGF interactions required for this shift, as mentioned in Chapter 2 Supplementary Figure 2.14 of this thesis[19]. Finally, using MD simulations, the elusive higher-order dimer-of-dimers may be identified that enable the observed *trans* interactions. Besides MD simulations, complementary small-angle X-ray scattering experiments may reveal the degree of intramolecular reorganisation by the microphthalmia-associated mutation in the A_1 variants. Also, performing K562 clustering with Ten3 isoforms that do not form covalent dimers – by mutagenic substitution of cystines 549 and 648 in EGF2 and EGF5, respectively - will show whether Lphn binding is possible without compact dimer formation.

To our best knowledge, no missense mutations have been described thus far that are expressed in all isoforms but only affect the function of some, as is the case for the MO substitution in Ten3. This can be explained by the unique mechanism of teneurin-3 where its splicing isoforms alternatively hide and expose binding motifs[36]. Missense mutations are, however, more commonly described to affect alternative splicing at various levels. In frontotemporal dementia and Alzheimer's disease, various mutations in *MAPT* gene results in an expression imbalance between the native 3R and 4R isoforms of tau protein[37]. Disease mutations have also been reported that result in unnatural

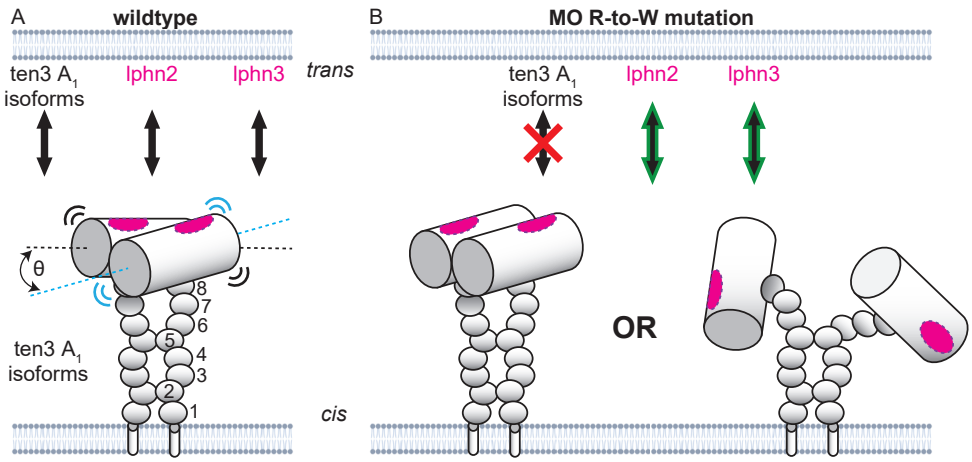


Figure 3.6: Model for the effects of altered compact dimer flexibility on *trans*-cellular interactions. **A)** Schematic representation of wildtype (WT) teneurin-3 A_1 isoform extracellular domain (ECD) attached to a membrane. There is moderate mobility of each superfold (cylinder) within the compact dimer relative to one another. The arrows indicate that the WT interacts homophilically as well as heterophilically with latrophilin-2 (Lphn2) and -3 (Lphn3) in *trans*. EGF repeats that connect the superfolds to the membrane are numbered 1-8. **B)** Schematic representation of the MO mutant ECD on a membrane. On basis of the simulations, there is less relative motion between the superfolds in a compact dimer (left). On longer timescales, the instability of the EGF8-ABD contact might still lead to complect dissociation of the compact dimer interface (right), which, in turn, could be an intermediate step in the transition from one compact dimeric conformation to another. The clustering analysis reveals that the indirect effect of the MO mutation, regardless, is the loss (red cross) of homophilic *trans* interactions, and the slight strengthening (green outlines) of heterophilic binding. Magenta patches on the superfolds represent latrophilin binding sites.

isoforms, for instance, by skipping canonical exons or generating new ones[38,39]. A few known missense mutations are located in exons where they logically only affect isoforms that contain this exon. For instance, two Noonan syndrome-associated (typified by a wide range of bodily disfiguration) missense mutations are described that only affect KRAS isoform B but not A, as they are located on exon 6 that is only included in isoform B[40].

Taken together, our findings provide the first direct evidence for the pathogenicity of microphthalmia-associated arginine-to-tryptophan substitution by interfering with Ten3 *trans*-cellular recognition. This is expected to unsettle the repulsive and attractive cues of teneurin-expressing neurons that are required for proper synaptic matching, and could thereby lead to connectopathic wiring and ocular malformations, as observed in microphthalmia patients.

3.4. METHODS

ALPHAMISSENSE STRUCTURAL DATA

AlphaMissense[27] pathogenicity data for human teneurin-3 (Q9P273) was acquired from the AlphaFold structure Protein Structure Database[41]. Mouse teneurin-3 protein-protein interaction sites were based the known interface residues (ten3 A_1B_0/A_1B_1) or determined on basis of proximity (ten3 A_0B_0/A_0B_1, and ten-Lphn) to the interacting domain using the select zone option within 20 Å and 10 Å, respectively, in the Chimera GUI[42].

For compact dimeric representations with pathogenicity surface projections, hTen3 was overlayed with mouse teneurin-3 in the negative stain compact dimer models (Chapter 2 Fig. 2.5E) using the Chimera mm command. Human teneurin-3 was similarly superimposed with chicken teneurin-2 in complex with latrophilin Lec-Olf domains (PDB:6SKA). Region-specific average pathogenicity scores were calculated by using these binding sites and the domain boundaries of human teneurin-3 UniProt entry. Average pathogenicity scores are the average of pathogenicity scores for each of the 19 possible substitution at a given amino acid position. Structural data visualisation was performed in PyMol (Schrödinger Molecular Graphics System, DeLano Scientific, San Carlos, CA). AlphaMissense annotated PDB hTen3 data was downloaded from the Hege Lab database, and colour values were visualised in PyMol using the Hege Lab coloram.py script.

PLASMIDS

The full-length wildtype versions of the four mouse teneurin-3 isoforms were cloned into the pUPE3820 vector (U-protein Express) containing a N-terminal GFP as described previously[19]. The R2579W mutation was introduced by cloning a custom gBlock Gene Fragment (residues 2497-264; Integrated DNA Technologies) into the full-length teneurin-3 A_0B_0 isoform by KasI-AscI (New England Biolabs) restriction-ligation. The mutation was then subcloned from there into the other isoforms by either EcoRI-KasI or EcoRI-AgeI (New England Biolabs) restriction-ligation. Full-length mouse latrophilin-2 (BC172691, Biocat) and latrophilin-3 (BC088989, Biocat) were cloned into the mCherry2-N1 vector containing a C-terminal mCherry2 tag (gift from Michael Davidson; Addgene plasmid# 54517; RRID: Addgene_54517) using NheI/BamHI and XhoI/BamHI cloning sites, respectively.

CLUSTERING ASSAY

Cell electroporation for clustering assays was conducted following a previously established protocol[15,19]. K562 cells (catalog nr. ACC 10, Leibniz Institute DSMZ) were cultured in RPMI-1640 medium (Gibco), supplemented with 10% FBS (Gibco) and 1% Penicillin/Streptomycin (Gibco), and grown in a shaking incubator at 37°C and 5% CO_2. K562 cells pelleted by for 5 min centrifugation at 300 g and washed in 1x PBS (Gibco) before electroporation. Cells were resuspended in buffer R (Gibco) with 2×10^6 cells per condition and were incubated with 15 µg of total amount of DNA (Ten3:empty vector ratio was 1:5) for 15 min at RT. K562 cells were then electroporated with the Neon Transfection System (Thermo Fisher Scientific) by 3x 10-ms pulses of 1450 V. Cells were then seeded into 5 mL of pre-warmed RPMI-1640 medium with 10% FBS in 6-well plates. After 20 hrs recovery in a shaking incubator, cells were collected and centrifuged for 3 minutes at 200 g. Cell resuspended in assay medium (RPMI-1640 supplemented with 10% FBS) were treated for 15 min at 37°C with DNase I (Invitrogen). Cells were once again centrifuged, resuspended in assay medium and passed through a 40 µm cell strainer. A total of $2{\cdot}10^5$ cells per clustering condition was plated in 1 mL of assay medium per condition in 12-well plates for 24 h of clustering on a shaking incubator at 37°C and 5% CO_2. Clusters were then imaged on an EVOS M5000 microscope with a 10x objective (0.25 NA; EVOS, Thermo Fisher Scientific), using the EVOS LED GFP and RFP cubes (Thermo Fisher Scientific). Clustering was analysed using FIJI (Schindelin. Et al).

GFP and mCherry channels were combined and Triangle thresholding was applied after Guassian blurring. Regions of interest (ROIs) larger than 100 pixels were determined, and rolling ball background subtraction with 50-pixel radius was performed on the individual GFP and mCherry channels before measuring the area and mean intensities of GFP and mCherry channels for each ROI. Clusters were identified on basis of a size four times larger than the mean large single cell size (650 pixels). The clustering index was defined as the total cluster area over the total area of all ROIs within an image (clusters + non-clusters) times 100%. The co-clustering ratio was determined by dividing the background-subtracted GFP by mCherry (I_{GFP} / I_{RFP}) signal per identified cluster. The cluster size and co-cluster ratio were averaged per image, and data from three to four independent experiments as described in the corresponding captions. Statistical significance was determined by performing a two-way ANOVA followed by Šìdák's multiple comparisons test using GraphPad Prism (GraphPad Software, San Diego, California USA). All analysed data are represented as mean ± SEM.

MOLECULAR DYNAMICS SIMULATIONS

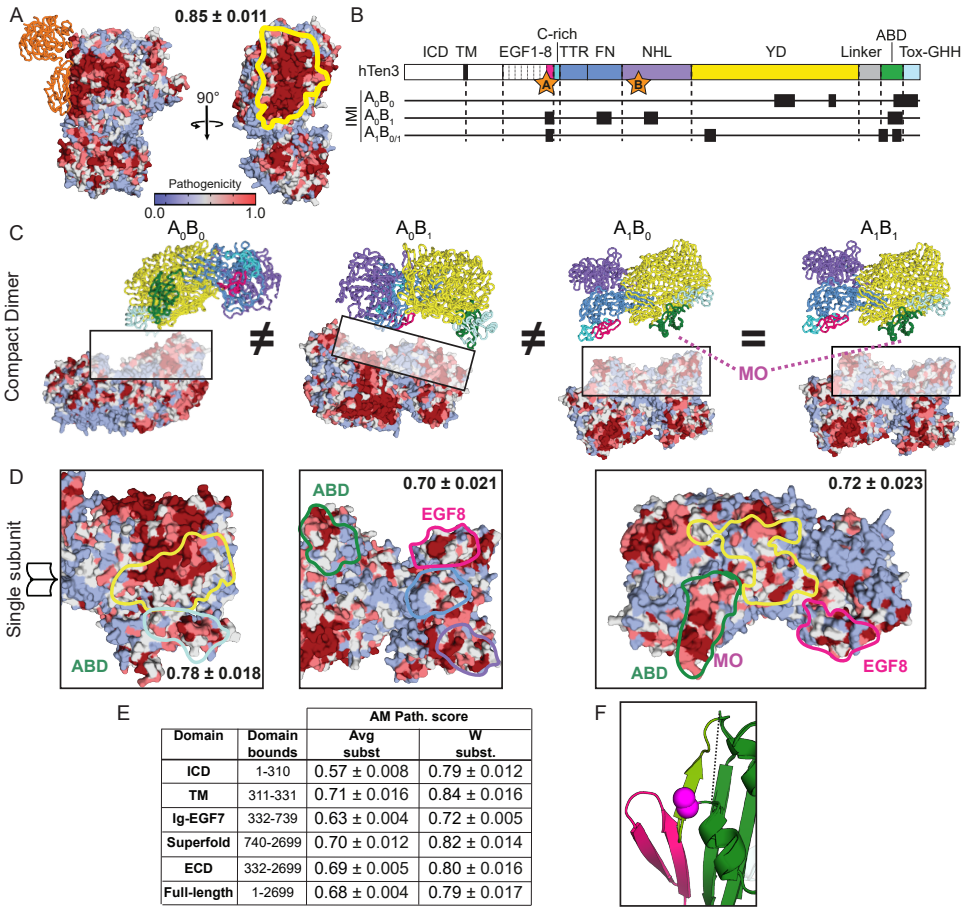
The full extracellular domain of mouse teneurin-3 A_1B_1 compact dimer was used as a starting point, which was modelled in CRYSOL[43]. 1995) to the small-angle X-ray scattering (SAXS) data of that isoform in the presence of calcium (Chapter 2 Fig. 2.6C)[19]. This SAXS-based initial model was derived from cryo-EM data (PDB:8R50) into which all missing EGF repeats were added that were predicted using ColabFold[29,44], and all complex glycans were manually built onto known glycosylation sites using Coot[45] before SAXS modelling. Regions that were unresolved using cryo-EM were taken from whole-domain ColabFold predictions to patch the structure. The competing beta-strand was also derived from ColabFold-predicted ABD domain, and residues that clashed with the EGF8 in the EGF8-ABD compact interface were manually curated in Coot to remove clashes. To minimise simulated box size and computation time, the flexible N-terminal region was removed up to the second dimerising disulfide link at C648 in EGF5, thereby preserving the native covalent linkage. All glycans were stripped from the structure, and unnatural N- and C-termini that were present in the structure were capped with acetyl and methyl groups, respectively. From the resulting wildtype structure, arginine 2579 was substituted by tryptophan to generate the R2579W mutant. Each amino acid was computationally protonated corresponding to a pH of 7.0 using Pdb4amber from AMBER Tools[46], and the structure enclosed in a box with 10 Å padding around the structure using tleap in AMBER Tools[46]. Ions were added to neutralise the protein nett charge and ions were then added to simulate a concentration of 150 mM NaCl and 2 mM $CaCl_2$ was simulated. Finally, the explicit TIP3P water molecules were added. Applied AMBER force fields were TIP3P for water, ff14SB for Protein, ions1lm_126_tip3p for monovalent ions, and ions234lm_1264_tip3p for divalent ions. The AmberTools tleap output was converted to GROMACS-compatible formats using ParmEd, and 4-step molecular simulations - consisting of consecutive energy minimisation step, NVT and NPT ensemble equilibration steps, and production step - were performed in the GROMACS software package[47–51] on DelftBlue supercomputing cluster[52]. Energy minimisation step was performed using steepest-decent minimisation with a force threshold of 103 kJ/mol/nm and 0.001 and 0.0001 step sizes for wildtype and mutant, respectively. NVT ensemble

equilibration was performed for 100 ps with the leap-frog integrator, and the temperature was simultaneously increased to 310 K with randomised initial velocity assignment. Pressure was increased to 1 bar during NPT ensemble equilibration by steps of 100 ps (leap-frog) with constant particles, pressure, and temperature. Root mean square fluctutations (RMSF)was calculated using the GROMACS rmsf command[51]. All these structural parameters were determined relative to structure after NPT ensemble equilibration. Principal component analysis (PCA) was performed in VMD software[53] with the Normal Mode Wizard and Prody plugins[54]. Principal components of movement for the experimental A_1B_1 compact dimer were calculated in CryoSPARC 3D Variability Analysis[55], of which the first and final (19^{th}) components were used here. Visual comparison of principal component volumes was performed using PyMol.

ACKNOWLEDGEMENTS

AlphaMissense structural data was acquired for academic and commercial use under a CC-BY-4.0 licence. Structural data from Protein Data Bank (PDB) were used under the CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license.

3.5. SUPPLEMENTARY INFORMATION



Domain	Domain bounds	AM Path. score Avg subst	W subst.
ICD	1-310	0.57 ± 0.008	0.79 ± 0.012
TM	311-331	0.71 ± 0.016	0.84 ± 0.016
Ig-EGF7	332-739	0.63 ± 0.004	0.72 ± 0.005
Superfold	740-2699	0.70 ± 0.012	0.82 ± 0.014
ECD	332-2699	0.69 ± 0.005	0.80 ± 0.016
Full-length	1-2699	0.68 ± 0.004	0.79 ± 0.017

Figure 3.7: Binding interfaces exhibit elevated predicted pathogenicity. **A)** Surface projection of predicted human teneurin-3 (hTen3) pathogenicity proximal to the latrophilin (Lphn) binding site (PDB:6SKA). Left: Lphn Olf-Lec domains are shown in orange, and human teneurin-3 (hTen3) pathogenicity surface is superimposed with the original chicken teneurin-2 structure. Right: hTen3-Lphn interface outlined in yellow. Value next to yellow interface indicates mean ± SEM pathogenicity score of entire interface. Colour codes for the pathogenicity scores are represented in the legend below. **B)** Linear representations of teneurin-3 domains, binding sites, and pathogenicity. Top: Protein domains in the full-length hTen3. Stars denote the positions of splice sites. Domain abbreviations as in Fig. 3.1 A. Middle: compact dimer intramolecular interfaces (IMI) are indicated per isoform as black boxes. **C)** Compact dimer conformations of the mouse Ten3 isoforms (Chapter 2 Fig. 2.5E). The lower subunit surface of each dimer is coloured by the per-residue average pathogenicity score of any amino acid substitution. Equal (= or ≠) signs denote whether isoforms adopt the same compact dimer conformation or not. Boxes outline the domains involved in the compact dimer interface and the region enlarged in panel D. **D)** Inset on the 'open book' (90° turn) representation of the region denoted by the box in panel C. Compact dimer IMI are outlined in colours of the involved domains on that subunit. Value in the boxes indicate mean ± SEM pathogenicity score that interface. These regions are denoted as black boxes IMI in panel B. Pathogenicity colours in panels C and D correspond to the legend in panel B, and positions of the microphthalmia mutation (MO) are idicated. **E)** Pathogenicity scores (mean ± SEM) for varying hTen3 regions. The extracellular domain (ECD) is defined as the Ig through Tox-GHH domains, and the superfold as EGF8 through Tox-GHH. The amino acid numbers that define the domain boundaries are indicated. **F)** Cartoon representation of A_0 and A_1 compact dimer ABD domains to show the clash upon simultaneous binding of EGF8 (magenta) and β textsubscriptrival (light green) to the rest of the ABD domain (dark green).

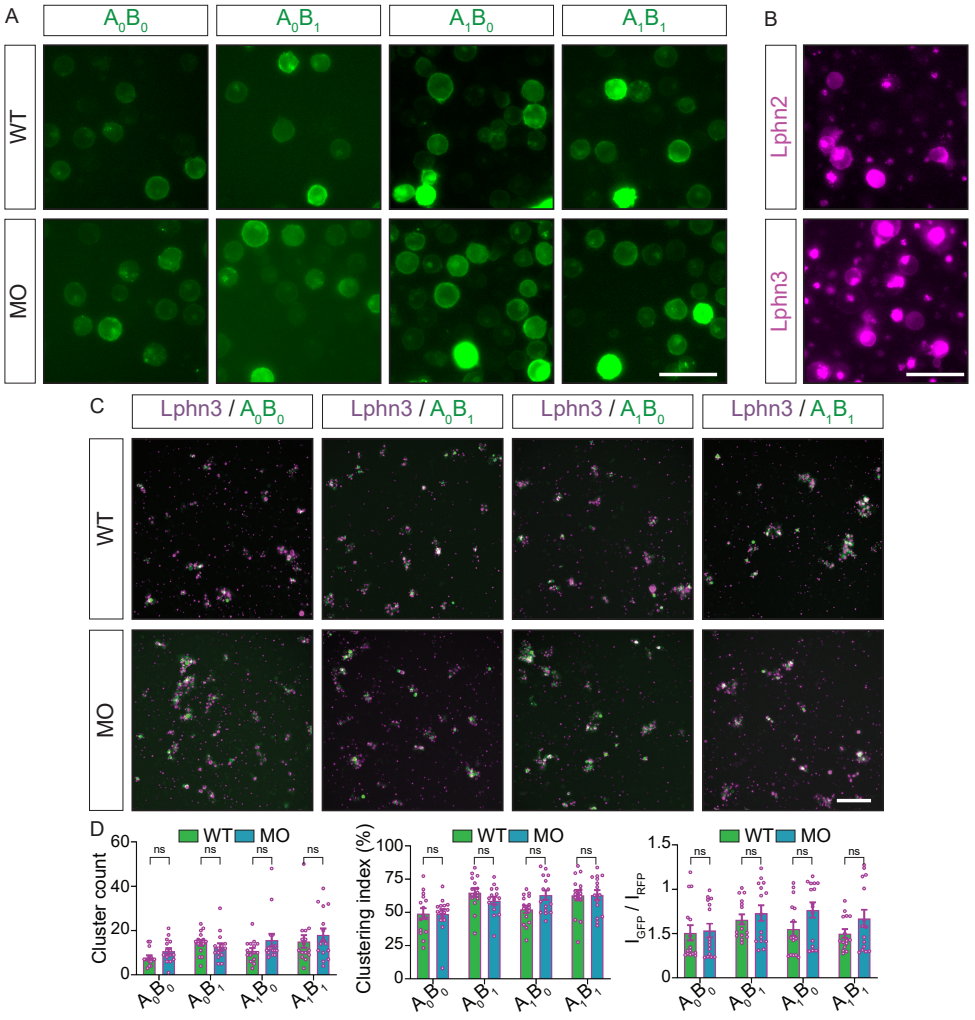


Figure 3.8: All teneurin-3 and latrophilin full-length proteins localise to the membrane and no significant effect of the microphthalmia mutation on teneurin-latrophilin clustering. **A)** Representative fluorescence microscopy images of non-clustering K562 cells at 40x magnification expressing the wildtype (WT) teneurin-3 isoforms and microphthalmia mutants (MO) indicated. Scale bar is 50 µm. **B)** Representative fluorescence microscopy images of non-clustering K562 cells at 40x magnification expressing latrophilin-2 and -3 as indicated on the left. Scale bar is 50 µm. **C)** Representative heterophilic clustering data between K562 cells expressing the full-length Ten3 isoforms indicated above, and cells expressing latrophilin-3 (Lphn3). WT Ten3 versions of the isoforms are compared with the MO mutants, as indicated on the left. Scale bar is 100 µm. **D)** Quantification of clustering data in panel C by metrics described for Fig. 3.4D (n=5 images per N=3 independent experiments, 15 images total). Data are presented as mean ± SEM; ns: not significant, **: p > 0.01, ***: p > 0.001. Two-Way ANOVA, Šìdák's multiple comparisons test.

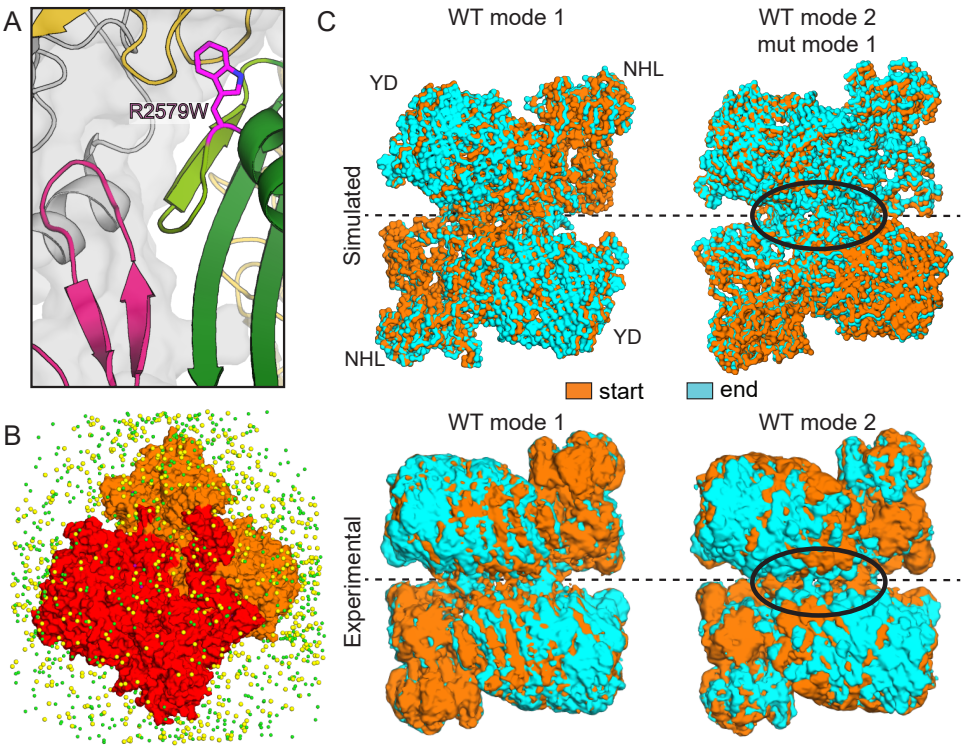


Figure 3.9: Structural data from the teneurin-3 A_1B_1 molecular dynamics simulations. **A)** Cartoon representation of the EGF8-ABD interface including β textsubscriptrival. Both EGF8 (magenta) and β textsubscriptrival (light green) are attached to the rest of the ABD (dark green) without initially clashing. The R2579W substitution is displayed as stick representation in purple. Transparent surface representation is used to distinguish between subunits within the compact dimer. **B)** The simulated solvent around the initial structure. Calcium ions are shown as purple, chloride as yellow, and sodium as green spheres. Water molecules are not shown. The initial structure in panels A and B are after the simulated equilibration steps. **C)** Top: Extreme states of the simulated modes of movement for wildtype (WT) and R2579W mutant (mut) from the principal component analysis. The WT first mode of displays a movement of each YD shell out of the sheet, recognised by predominantly cyan surface of the last timepoint (end, $t = 550$ ns). An opposite movement is observed for the NHL domains into the sheet, recognised by protruding orange surface representing the first time point in the mode of movement – the 'hugging' movement. For the WT second mode and mut first mode of movement, one of the subunits displays movement out of the sheet while the other stays roughly at the same position. This movement is typified by the YD shells of each subunit coming closer together – the so-called 'butterfly' movement. Bottom: Extreme states of the experimental modes of movement for WT from the 3D variability analysis. Experimental mode 1 is most similar to the simulated mode 1 of the WT, and experimental mode 2 is most similar to the WT mode 2 and mut mode 1. The black oval outlines the increased contact between the YD shells in the $t = 550$ ns timepoints (end, cyan) of the corresponding modes. The dashed black line indicates the separate subunits in the compact dimers.

REFERENCES

[1] A. S. Verma and D. R. FitzPatrick, *Anophthalmia and microphthalmia,* Orphanet Journal of Rare Diseases **2**, 47 (2007).

[2] L. Roos, H. Jensen, K. G. nskov, and Z. T. René Holst, *Congenital microphthalmia, anophthalmia and coloboma among live births in denmark,* Ophthalmic Epidemiology **23**, 324 (2016).

[3] A. L. Groot, M. M. Kuijten, J. Remmers, A. Gilani, D. L. Mourits, E. Kraal-Biezen, P. de Graaf, P. J. Zwijnenburg, A. C. Moll, S. Tan, P. Saeed, and D. T. Hartong, *Classification for treatment urgency for the microphthalmia/anophthalmia spectrum using clinical and biometrical characteristics,* Acta Ophthalmologica **98**, 514 (2020), https://onlinelibrary.wiley.com/doi/pdf/10.1111/aos.14364 .

[4] L. M. Reis and E. V. Semina, *Conserved genetic pathways associated with microphthalmia, anophthalmia, and coloboma,* Birth Defects Res C Embryo Today **105**, 96 (2015).

[5] M. J. Landrum, J. M. Lee, G. R. Riley, W. Jang, W. S. Rubinstein, D. M. Church, and D. R. Maglott, *Clinvar: Public archive of relationships among sequence variation and human phenotype,* Nucleic Acids Research **42** (2014), 10.1093/nar/gkt1113.

[6] F. Islam, S. Htun, L. W. Lai, M. Krall, M. Poranki, P. M. Martin, N. Sobreira, E. S. Wohler, J. Yu, A. T. Moore, and A. M. Slavotinek, *Exome sequencing in patients with microphthalmia, anophthalmia, and coloboma (mac) from a consanguineous population,* Clinical Genetics **98**, 499 (2020).

[7] M. A. Aldahmesh, J. Y. Mohammed, S. Al-Hazzaa, and F. S. Alkuraya, *Homozygous null mutation in odz3 causes microphthalmia in humans,* Genetics in Medicine **14**, 900 (2012).

[8] N. Chassaing, N. Ragge, J. Plaisancié, O. Patat, D. Geneviève, F. Rivier, C. Malrieu-Eliaou, C. Hamel, J. Kaplan, and P. Calvas, *Confirmation of tenm3 involvement in autosomal recessive colobomatous microphthalmia,* American Journal of Medical Genetics, Part A **170**, 1895 (2016).

[9] S. G. Yarahmadi, F. Sarlaki, and S. Morovvati, *Novel mutation in tenm3 gene in an iranian patient with colobomatous microphthalmia,* Clinical Case Reports **10** (2022), 10.1002/ccr3.5532.

[10] B. Singh, P. Srivastava, and S. R. Phadke, *Sequence variations in tenm3 gene causing eye anomalies with intellectual disability: Expanding the phenotypic spectrum,* European Journal of Medical Genetics **62**, 61 (2019).

[11] S. Merlin, S. Horng, L. R. Marotte, M. Sur, A. Sawatari, and C. A. Leamey, *Deletion of ten-m3 induces the formation of eye dominance domains in mouse visual cortex,* Cerebral Cortex **23**, 763 (2013).

[12] N. Dharmaratne, K. A. Glendining, T. R. Young, H. Tran, A. Sawatari, and C. A. Leamey, *Ten-m3 is required for the development of topography in the ipsilateral retinocollicular pathway,* PLoS ONE **7** (2012), 10.1371/journal.pone.0043083.

[13] P. Antinucci, N. Nikolaou, M. P. Meyer, and R. Hindges, *Teneurin-3 specifies morphological and functional connectivity of retinal ganglion cells in the vertebrate visual system,* Cell Reports **5**, 582 (2013).

[14] D. S. Berns, L. A. DeNardo, D. T. Pederick, and L. Luo, *Teneurin-3 controls topographic circuit assembly in the hippocampus,* Nature **554**, 328 (2018).

[15] D. T. Pederick, J. H. Lui, E. C. Gingrich, C. Xu, M. J. Wagner, Y. Liu, Z. He, S. R. Quake, and L. Luo, *Reciprocal repulsions instruct the precise assembly of parallel hippocampal networks,* SCIENCE **372**, 1068+ (2021).

[16] J. L. Hunyara, K. M. Daly, K. Torres, M. E. Yurgel, R. Komal, S. Hattar, and A. L. Kolodkin, *Teneurin-3 regulates the generation of nonimage- forming visual circuitry and responsiveness to light in the suprachiasmatic nucleus,* PLoS Biology **21** (2023), 10.1371/journal.pbio.3002412.

[17] K. Liakath-Ali and T. C. Südhof, *The perils of navigating activity-dependent alternative splicing of neurexins,* Frontiers in Molecular Neuroscience **14**, 1 (2021).

[18] C. A. Leamey, S. Merlin, P. Lattouf, A. Sawatari, X. Zhou, N. D. K. A. Glendining, T. Oohashi, M. Sur, and R. Fässler, *Ten_m3 regulates eye-specific patterning in the mammalian visual pathway and is required for binocular vision,* PLOS Biology **5**, e241 (2007).

[19] C. Gogou, J. W. Beugelink, C. P. Frias, L. Kresik, N. Jaroszynska, U. Drescher, B. J. Janssen, R. Hindges, and D. H. Meijer, *Alternative splicing controls teneurin-3 compact dimer formation for neuronal recognition,* Nature Communications **15** (2024), 10.1038/s41467-024-47763-x.

[20] V. A. Jackson, D. H. Meijer, M. Carrasquero, L. S. V. Bezouwen, E. D. Lowe, C. Kleanthous, B. J. Janssen, and E. Seiradake, *Structures of teneurin adhesion receptors reveal an ancient fold for cell-cell interaction,* Nature Communications **9** (2018), 10.1038/s41467-018-03460-0.

[21] J. Li, M. Shalev-Benami, R. Sando, X. Jiang, A. Kibrom, J. Wang, K. Leon, C. Katanski, O. Nazarko, Y. C. Lu, T. C. Südhof, G. Skiniotis, and D. Araç, *Structural basis for teneurin function in circuit-wiring: A toxin motif at the synapse,* Cell **173**, 735 (2018).

[22] J. Li, Y. Xie, S. Cornelius, X. Jiang, R. Sando, S. P. Kordon, M. Pan, K. Leon, T. C. Südhof, M. Zhao, and D. Araç, *Alternative splicing controls teneurin-latrophilin interaction and synapse specificity by a shape-shifting mechanism,* Nature Communications **11** (2020), 10.1038/s41467-020-16029-7.

[23] D. del Toro, M. A. Carrasquero-Ordaz, A. Chu, T. Ruff, M. Shahin, V. A. Jackson, M. Chavent, M. Berbeira-Santana, G. Seyit-Bremer, S. Brignani, R. Kaufmann,

E. Lowe, R. Klein, and E. Seiradake, *Structural basis of teneurin-latrophilin interaction in repulsive guidance of migrating neurons,* Cell **180**, 323 (2020).

[24] D. H. Meijer, C. P. Frias, J. W. Beugelink, Y. N. Deurloo, and B. J. C. Janssen, *Teneurin4 dimer structures reveal a calcium-stabilized compact conformation supporting homomeric trans-interactions,* The EMBO Journal **41** (2022), 10.15252/embj.2020107505.

[25] J. Li, S. J. Bandekar, and D. Araç, *The structure of fly teneurin-m reveals an asymmetric self-assembly that allows expansion into zippers,* EMBO reports **24** (2023), 10.15252/embr.202256728.

[26] J. Cheng, G. Novati, J. Pan, C. Bycroft, A. Žemgulyte, T. Applebaum, A. Pritzel, L. H. Wong, M. Zielinski, T. Sargeant, R. G. Schneider, A. W. Senior, J. Jumper, D. Hassabis, P. Kohli, and Žiga Avsec, *Accurate proteome-wide missense variant effect prediction with alphamissense,* Science **381** (2023), 10.1126/science.adg7492.

[27] J. Cheng, Y. Yu, X. Wang, X. Zheng, T. Liu, D. Hu, Y. Jin, Y. Lai, T. M. Fu, and Q. Chen, *Structural basis for the self-recognition of sdscam in chelicerata,* Nature Communications **14** (2023), 10.1038/s41467-023-38205-1.

[28] S. J. de Vries, M. van Dijk, and A. M. J. J. Bonvin, *The haddock web server for data-driven biomolecular docking,* Nature Protocols **5**, 883 (2010).

[29] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, *Highly accurate protein structure prediction with alphafold,* NATURE **596**, 583+ (2021).

[30] C. Xu, Z. Li, C. Lyu, Y. Hu, C. N. McLaughlin, K. K. L. Wong, Q. Xie, D. J. Luginbuhl, H. Li, N. D. Udeshi, T. Svinkina, D. R. Mani, S. Han, T. Li, Y. Li, R. Guajardo, A. Y. Ting, S. A. Carr, J. Li, and L. Luo, *Molecular and cellular mechanisms of teneurin signaling in synaptic partner matching,* Cell (2024), 10.1016/j.cell.2024.06.022.

[31] S. A. Hollingsworth and R. O. Dror, *Molecular dynamics simulation for all,* Neuron **99**, 1129 (2018).

[32] L. M. Chataigner, C. Gogou, M. A. den Boer, C. P. Frias, D. M. Thies-Weesie, J. C. Granneman, A. J. Heck, D. H. Meijer, and B. J. Janssen, *Structural insights into the contactin 1 – neurofascin 155 adhesion complex,* Nature Communications **13** (2022), 10.1038/s41467-022-34302-9.

[33] J. Wang, Y. Miao, R. Wicklein, Z. Sun, J. Wang, K. M. Jude, R. A. Fernandes, S. A. Merrill, M. Wernig, K. C. Garcia, and T. C. Südhof, *Rtn4/nogo-receptor binding to bai adhesion-gpcrs regulates neuronal development,* Cell **184**, 5869 (2021).

[34] S. H. Wang, T. J. Wu, C. W. Lee, and J. Yu, *Dissecting the conformation of glycans and their interactions with proteins,* Journal of Biomedical Science **27** (2020), 10.1186/s12929-020-00684-5.

[35] O. Akkermans, C. Delloye-Bourgeois, C. Peregrina, M. Carrasquero-Ordaz, M. Kokolaki, M. Berbeira-Santana, M. Chavent, F. Reynaud, R. Raj, J. Agirre, M. Aksu, E. S. White, E. Lowe, D. B. Amar, S. Zaballa, J. Huo, I. Pakos, P. T. McCubbin, D. Comoletti, R. J. Owens, C. V. Robinson, V. Castellani, D. del Toro, and E. Seiradake, *Gpc3-unc5 receptor complex structure and role in cell migration,* Cell **185**, 3931 (2022).

[36] P. Kjer-Hansen and R. J. Weatheritt, *The function of alternative splicing in the proteome: rewiring protein interactomes to put old functions into new contexts,* Nature Structural and Molecular Biology **30**, 1844 (2023).

[37] A. Kar, D. Kuo, R. He, J. Zhou, and J. Y. Wu, *Tau alternative splicing and frontotemporal dementia,* Alzheimer Disease & Associated Disorders **19**, S29 (2005).

[38] A. Anna and G. Monika, *Splicing mutations in human genetic disorders: examples, detection, and confirmation,* Journal of Applied Genetics **59**, 253 (2018).

[39] P. Baeza-Centurion, B. Miñana, J. Valcárcel, and B. Lehner, *Mutations primarily alter the inclusion of alternatively spliced exons,* eLife **9**, 1 (2020).

[40] C. Carta, F. Pantaleoni, G. Bocchinfuso, L. Stella, I. Vasta, A. Sarkozy, C. Digilio, A. Palleschi, A. Pizzuti, P. Grammatico, G. Zampino, B. Dallapiccola, B. D. Gelb, M. Tartaglia, G. Rotondo, and A. Pizzuti, *Report germline missense mutations affecting kras isoform b are associated with a severe noonan syndrome phenotype,* The American Journal of Human Genetics **79**, 129 (2006).

[41] M. Varadi, D. Bertoni, P. Magana, U. Paramval, I. Pidruchna, M. Radhakrishnan, M. Tsenkov, S. Nair, M. Mirdita, J. Yeo, O. Kovalevskiy, K. Tunyasuvunakool, A. Laydon, A. Žídek, H. Tomlinson, D. Hariharan, J. Abrahamson, T. Green, J. Jumper, E. Birney, M. Steinegger, D. Hassabis, and S. Velankar, *Alphafold protein structure database in 2024: providing structure coverage for over 214 million protein sequences,* Nucleic Acids Research **52**, D368 (2024).

[42] T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, and T. E. Ferrin, *Ucsf chimerax: Meeting modern challenges in visualization and analysis,* PROTEIN SCIENCE **27**, 14 (2018).

[43] D. Svergun, C. Barberato, and M. H. J. Koch, *Crysol-a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates,* J Appl. Cryst **28**, 768 (1995).

[44] M. Mirdita, K. Schütze, Y. Moriwaki, L. Heo, S. Ovchinnikov, and M. Steinegger, *Colabfold: making protein folding accessible to all,* Nature Methods **19**, 679 (2022).

[45] P. Emsley and K. Cowtan, *Coot: Model-building tools for molecular graphics,* Acta Crystallographica Section D: Biological Crystallography **60**, 2126 (2004).

[46] D. A. Case, H. M. Aktulga, K. Belfon, D. S. Cerutti, G. A. Cisneros, V. W. D. Cruzeiro, N. Forouzesh, T. J. Giese, A. W. Götz, H. Gohlke, S. Izadi, K. Kasavajhala, M. C. Kaymak, E. King, T. Kurtzman, T. S. Lee, P. Li, J. Liu, T. Luchko, R. Luo, M. Manathunga, M. R. Machado, H. M. Nguyen, K. A. O'Hearn, A. V. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, A. Risheh, S. Schott-Verdugo, A. Shajan, J. Swails, J. Wang, H. Wei, X. Wu, Y. Wu, S. Zhang, S. Zhao, Q. Zhu, T. E. Cheatham, D. R. Roe, A. Roitberg, C. Simmerling, D. M. York, M. C. Nagan, and K. M. Merz, *Ambertools,* Journal of Chemical Information and Modeling **63**, 6183 (2023).

[47] D. V. D. Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, *Gromacs: Fast, flexible, and free,* Journal of Computational Chemistry **26**, 1701 (2005).

[48] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindah, *Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers,* SoftwareX **1-2**, 19 (2015).

[49] B. Hess, C. Kutzner, D. V. D. Spoel, and E. Lindahl, *Grgmacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation,* Journal of Chemical Theory and Computation **4**, 435 (2008).

[50] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. V. D. Spoel, B. Hess, and E. Lindahl, *Gromacs 4.5: A high-throughput and highly parallel open source molecular simulation toolkit,* Bioinformatics **29**, 845 (2013).

[51] E. Lindahl, B. Hess, and D. van der Spoel, *Gromacs 3.0: A package for molecular simulation and trajectory analysis,* Journal of Molecular Modeling **7**, 306 (2001).

[52] Delft High Performance Computing Centre (DHPC), *DelftBlue Supercomputer (Phase 2),* `https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2` (2024).

[53] W. Humphrey, A. Dalke, and K. Schulten, *Vmd: Visual molecular dynamics,* (1996).

[54] A. Bakan, L. M. Meireles, and I. Bahar, *Prody: Protein dynamics inferred from theory and experiments,* Bioinformatics **27**, 1575 (2011).

[55] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, *Cryosparc: Algorithms for rapid unsupervised cryo-em structure determination,* Nature Methods **14**, 290 (2017).

4

STRUCTURAL INSIGHTS INTO THE CONTACTIN-1 – NEUROFASCIN-155 ADHESION

Cell-surface expressed contactin-1 and neurofascin-155 control wiring of the nervous system and interact across cells to form and maintain paranodal myelin-axon junctions. The molecular mechanism of contactin-1 - neurofascin-155 adhesion complex formation is unresolved. Crystallographic structures of complexed and individual neurofascin-155 binding regions presented here, provide a rich picture of how competing interfaces, posttranslational glycosylation, and splice differences enable formation of diverse adhesion sites. Structural, biophysical, and cell clustering analysis reveal how conserved Ig1-2 interfaces form competing heterophilic contactin-1 - neurofascin-155 and homophilic neurofascin-155 complexes. The structures explain how the heterophilic Ig1-Ig4 horseshoes in the contactin-1 - neurofascin-155 complex define the 7.4 nm paranodal spacing and how the remaining six domains enable bridging of distinct intercellular distances.

This chapter is a reduced version of the article published as Lucas M.P. Chataigner, *Christos Gogou*, Maurits A. den Boer, Cátia P. Frias, Dominique M.E. Thies-Weesie, Joke C.M. Granneman, Albert J.R. Heck, Dimphna H. Meijer, Bert J.C. Janssen, Structural insights into the contactin 1-neurofascin 155 adhesion complex, *Nature Communications (2022)* [1]. Selected structural data are presented here that support our observations from the clustering assays we conducted.

4.1. INTRODUCTION

The immunoglobulin-like cell adhesion molecule (IgCAM) family functionally patterns multicellular structures by providing cellular recognition, mechanical support, and contributing to membrane microdomain assembly at adhesion sites between cells[2,3]. Two IgCAM subfamilies, contactin and L1, have roles in the development and function of tissues[4,5], most notably of the nervous system[6,7] by controlling processes of neurite extension, axon guidance, synapse formation, myelination, and axo-glia domain assembly[6–11]. Underlying many of these processes is the formation of oligomers and complexes between and within contactin and L1 family proteins[8,9]. Contactin-1 (Cntn1) and neurofascin-155 (Nf155) exemplify the two subclasses and together form a heterophilic adhesion complex at the myelin-axon paranode[10–15] and the neuronal synapse[16]. Dysfunction of contactin and L1 subfamilies are associated with deficits in learning and memory regulation, wide ranging neuropsychiatric diagnoses, neurodevelopmental disorders, cancers, neurodegenerative diseases, and neuropathies[4–7,17–19]. Cntn1 and neurofascin particularly have well-documented roles in the etiology of various demyelinating neuropathies, neurodegenerative diseases including Alzheimer's and Parkinson's, and cancers[4–7,17,18,20].

Cntn1 (glycoprotein gp135, F11 in chicken, and F3 in mouse) and five other paralogues (contactin-2 to -6) form the vertebrate contactin subfamily of the IgCAM family[21]. Defining subfamily features are six extracellular N-terminal immunoglobulin-like domains (Ig), four fibronectin type III domains (FnIII) and posttranslational modifications, i.e. N-glycosylation at multiple sites and a C-terminal glycophosphatidylinositol (GPI)-membrane anchor. Crystal structures of one to four domain-containing segments of contactins have shown that the first four domains form a characteristic horseshoe supramodule with versatile binding capacity[22–26], that the Ig5-FnIII2 domains are arranged in an extended head-to-tail conformation and that the FnIII2-3 connection is bent[26]. Membrane-anchored or soluble forms of contactins are expressed across tissues with strongest expression in the nervous system[5]. Contactin proteins are critical to neuronal development and homeostasis[19] and function through ubiquitous interactions, which can be homophilic or heterophilic, and *cis* or *trans*[9]. Proteins from the contactin-associated proteins (CASPR), receptor tyrosine phosphatases (PTPR), amyloid beta precursor protein (APP), notch, and L1 signalling protein families interact with contactins[9,19,27]. The Cntn1-Nf155 *trans* interaction, necessary for proper paranode formation and maintenance, is amongst the best documented interactions with robust cellular level data supporting the concomitant role of these molecules in this setting[28–30].

Neurofascin and three other paralogues (L1, CHL1, and NrCam) form the vertebrate L1 subfamily of the IgCAMs[8] that has a prominent role in nervous system development and homeostasis[7]. All four members are type I transmembrane proteins with six extracellular N-terminal Ig like domains, three to five FnIII domains, various predicted N-glycosylation sites, a single transmembrane spanning helix, and a ~120 cytoplasmic region with a highly conserved ankyrin binding sequence[4]. The only structures available for this subfamily are of neurofascin Ig1-Ig4 that show, similar to the contactins, a characteristic horseshoe supramodule with a conserved homodimerisation interface located on Ig1-Ig2[31,32]. Neurofascin consists of over 50 possible splice variants[33], at extracellular and cytosolic sites, with tightly regulated expression patterns intimately associated

with their selective biological functions[34,35]. Two isoforms, Nf155 and neurofascin-186 (Nf186), are predominantly expressed in the mature nervous system. Generally referred to as "glial" and "axonal" isoforms respectively, owing to their cell type-specific expression, they both play key roles in myelination[15,36,37].

Paranodes flank the nodes of Ranvier and are formed by distal uncompacted loops of myelin that contact the axon at tight 7.4 nm junctions[38]. The Cntn1-Nf155 *trans*-complex together with CASPR1 are required for the formation and maintenance of these structures, acting as molecular rulers maintaining the rather narrow intercellular spacing[10–15,28–30,36,37]. This ternary complex forms a larger-order assembly that spirals around the axon acting as a membrane diffusion barrier, as denoted by loss of paranodal attachment and axonal compartmentalisation observed upon loss of either of the three components[10,11]. Recent research highlights the importance of this barrier as it acts as an incomplete conductive seal for correct action potential propagation[39], and remarkably displays plasticity as the boundary it creates is dynamically displaced over time during learning[40]. Interestingly, Cntn1 and Nf155 also interact at the synapse[16] at a much larger intercellular distance of ~20-25 nm[41–43].

While the Cntn1-Nf155 *trans*-complex underlies the attachment of the paranode, many contactin and L1 family members localise and interact in close proximity in nodal, juxtaparanodal, an internodal regions establishing neighouring molecular domains along the myelinated axon[10,11]. How these homologous molecules retain interaction specificities is of great interest in this context, given the high identity between paralogues. It seems part of this inquiry can be answered through the role of isoforms and glycoforms which appear to regulate interaction strength. For Cntn1 – neurofascin interactions both N-linked glycosylation of Cntn1[44–46], and splicing of neurofascin[12] are reported to modulate interaction between the two molecules. In the absence of structural information on the Cntn1-Nf155 complex, its interaction mode, adhesion mechanism, parameters determining specificity and role of isoforms and glycosylation have remained unresolved. Moreover, whilst fragments of contactin[22–26] and L1 subfamily[31,32] proteins have informed on the structural basis of their function, how heterophilic complexes form and their possible stoichiometries has not yet been elucidated.

To address these questions, we structurally characterised the Cntn1 and Nf155 interacting regions, i.e. Ig1-Ig6, individually and in complex. We show that both proteins have a characteristic Ig1-Ig4 horseshoe fold and that Ig5 has distinct connections with Ig4 comparing Cntn1 and Nf155. A conserved surface on Ig1-2 on both molecules plays an important role in complex formation. Several N-linked glycans are involved in intermolecular interactions in the complex and their composition on Nf155 affects interaction affinity. Nf155 homodimer formation and Cntn1-Nf155 heterocomplex formation are mutually exclusive due to overlapping interaction sites. Structure-based insights are verified by mutations that prevent Cntn1–Nf155 mediated cell clustering. Together our data suggest that the interacting Ig1-Ig4 horseshoes in Cntn1-Nf155 determine the paranodal spacing and that the more extended Ig5-FnIII-4 tails could be used to span the larger intercellular spacing encountered in the neuronal synapse.

4.2. RESULTS

4.2.1. CNTN1-NF155 IS MEDIATED BY CONSERVED IG1-IG2 INTERACTIONS

Functional analyses have identified the six Ig-domains of contactin-1 (Cntn1) and neurofascin-155 (Nf155) as the extracellular regions that mediate *trans* interactions[12,46,47]. We determined the structure of glycosylated Cntn1$^{Ig1\text{-}6}$ in complex with Nf155$^{Ig1\text{-}6}$ from a crystal that diffracted anisotropically to 8-4.8 Å (PDB:7OL4) (Fig. 4.1A-C, Fig. 4.2A, and Table 4.1). Structure determination of the low-resolution complex was aided by the structures of the unliganded Cntn1$^{Ig1\text{-}6}$ (PDB:7OL2) and Nf155$^{Ig1\text{-}6}$ (PDB:7OK5) that we determined to a resolution of 3.9 and 3.0 Å, respectively. The Ig1-Ig4 horseshoes of Cntn1$^{Ig1\text{-}6}$ and Nf155$^{Ig1\text{-}6}$ interact in a near orthogonal edge-on orientation each using a large part of the Ig2 domain side for interaction. At the center of the heteromeric complex, an intermolecular "super" β-sheet is formed by antiparallel hydrogen bonding of the outer G β-strands located in the Ig2 GFC β-sheet of both molecules (Fig. 4.1C, Fig. 4.2B and Supplementary Fig. 4.6A). At the "bottom side", the super β-sheet forms a concave half barrel that is closed off by the Ig2 CD loops of both molecules and characterised by extensive hydrophobic contacts in the core (Fig. 4.2B). At the "top side" the intermolecular interactions are less prominent (Supplementary Fig. 4.6A). Here Ig1 AB loops of both molecules contact edges of the opposing Ig2 domains formed by the Ig2 N-terminus and G β-strand (Supplementary Fig. 4.7). At both sides of the super β-sheet intramolecular salt-bridges are formed. Overall, the Cntn1-Nf155 complex buries ~2089 Å^2 of solvent accessible area[48] and the interface is of a mixed hydrophobic hydrophilic nature with complementary electrostatic interactions (Fig. 4.2A and Supplementary Fig. 4.7A).

The Cntn1-Nf155 interaction interface is conserved among vertebrate orthologues indicating that the interface features we identify in the complex structure are likely crucial to Cntn1-Nf155 mediated interactions (Supplementary Fig. 4.7B). In particular, high conservation of β-strands G and C, CD loop residues and specific N-glycosylation sites are observed across orthologues. Plotting conservation of L1-specified paralogues, i.e. L1, CHL1, NrCAM and neurofascin reveals relatively lower conservation for the C β-strand and CD loop residues (Supplementary Fig. 4.7B). Analysis of conservation among Cntn1 paralogues shows conservation of FC strand and CD loop residues and lower conservation of G strand residues (Supplementary Fig. 4.7B). Taken together, the conservation among orthologues and paralogues of residues located in the Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ interface suggests that this site may play a role in the adhesion mechanisms within these families.

4.2.2. MANNOSE-RICH GLYCANS STABILISE THE CNTN1-NF155 CONTACT

N-glycosylation of Cntn1 has been shown to regulate binding to neurofascin[44–46], where Cntn1 containing mannose-rich glycans is expected to be relevant in paranodal adhesion[49]. For structure determination we produced the mannose-rich version of the N-linked glycans. In the Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ complex, mannose-rich glycans attached to conserved Cntn1 glycosylation sites at N208 on Ig2, and N258 on Ig3, buttress the main protein-protein interface and may form intermolecular glycan-glycan and glycan-protein interfaces, respectively (Fig. 4.1C). The glycan at Cntn1 N208 is in near vicinity

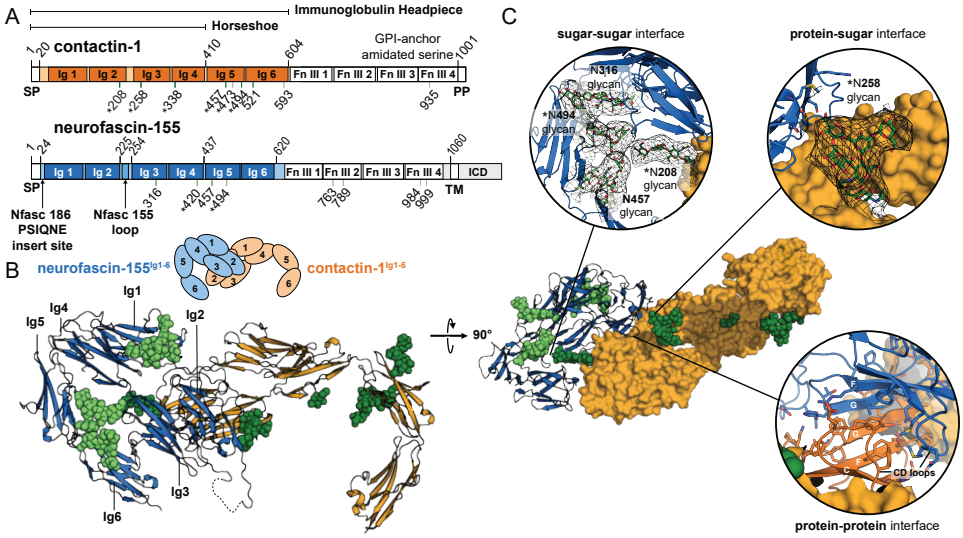


Figure 4.1: Structural and biophysical insights into contactin-1 - neurofascin-155 Ig1-2 mediated complex formation. **A)** Annotated sequence schematic of contactin-1 (Cntn1) and neurofascin-155 (Nf155). Immunoglobulin-like (Ig) and fibronectin type III (FnIII) domains are denoted with domain boundaries (top lines), and crystallised Ig segments coloured in. N-glycosylation sites (bottom lines) confirmed by structures (green), predicted (grey), and highly conserved in orthologue conservation analysis (*). Signal peptide (SP), pro peptide (PP), transmembrane domain (TM), glycophosphatidylinositol (GPI) anchor and intracellular domain (ICD) are also indicated. Neurofascin (Nf). **B)** Cartoon representation of Cntn1^{Ig1-6}–Nf155^{Ig1-6} complex structure, with glycan residues (spheres), missing loop segment (dashed line) and a schematic indicating the composition. The interface is located on domain Ig1 and 2 of both molecules. **C)** Rotated view of the Cntn1^{Ig1-6}–Nf155^{Ig1-6} complex structure with Cntn1^{Ig1-6} in surface representation. Insets show various interfaces with glycan density (2mFobs-DFcalc) at 1σ shown as black mesh.

to Ig2 of Nf155, while the glycan on contactin N258 extends into proximity of glycans attached to N457 and N494 on Ig5 of Nf155. These additional glycan-mediated interfaces may extend or, more importantly, sterically hinder the protein-protein interface, depending on glycan composition (Fig. 4.1C). Possibly the glycans on Ig5 of Nf155, of which N494 is conserved, and that interface with the glycan on N258 of Cntn1, play a role in modulating the affinity.

4.2.3. HOMODIMER AND CNTN1 INTERFACES ON NF155 OVERLAP

The conformation of six Ig domains in unliganded Nf155^{Ig1-6}, determined to a maximum resolution of 3 Å (PDB:7OK5) (Fig. 4.2C and Table 4.1), is similar to that of Nf155^{Ig1-6} bound to Cntn1^{Ig1-6} (Fig. 4.2A). The first four domains of Nf155^{Ig1-6} display a characteristic horseshoe arrangement in which Ig1 contacts Ig4, and Ig2 contacts Ig3 in an antiparallel fashion enabled by a 180° V-turn in the Ig2-Ig3 unit-connecting loop. Another Ig4-Ig5 mediated V-turn is stabilised by extensive hydrophobic interactions, in which a central F466 of Ig5 is surrounded by I25, I27, P28 and L31 of the N-terminus, I358 and L437 of Ig4, and V439, L444 and I470 of Ig5 (Supplementary Fig. 4.6B). In addition, sidechains of D30 with R442, and K355 with Q493 are poised to form two salt bridges sandwiching the interface.

In the unliganded form, Nf155^{Ig1-6} forms a homodimer (Fig. 4.2D-E and Supple-

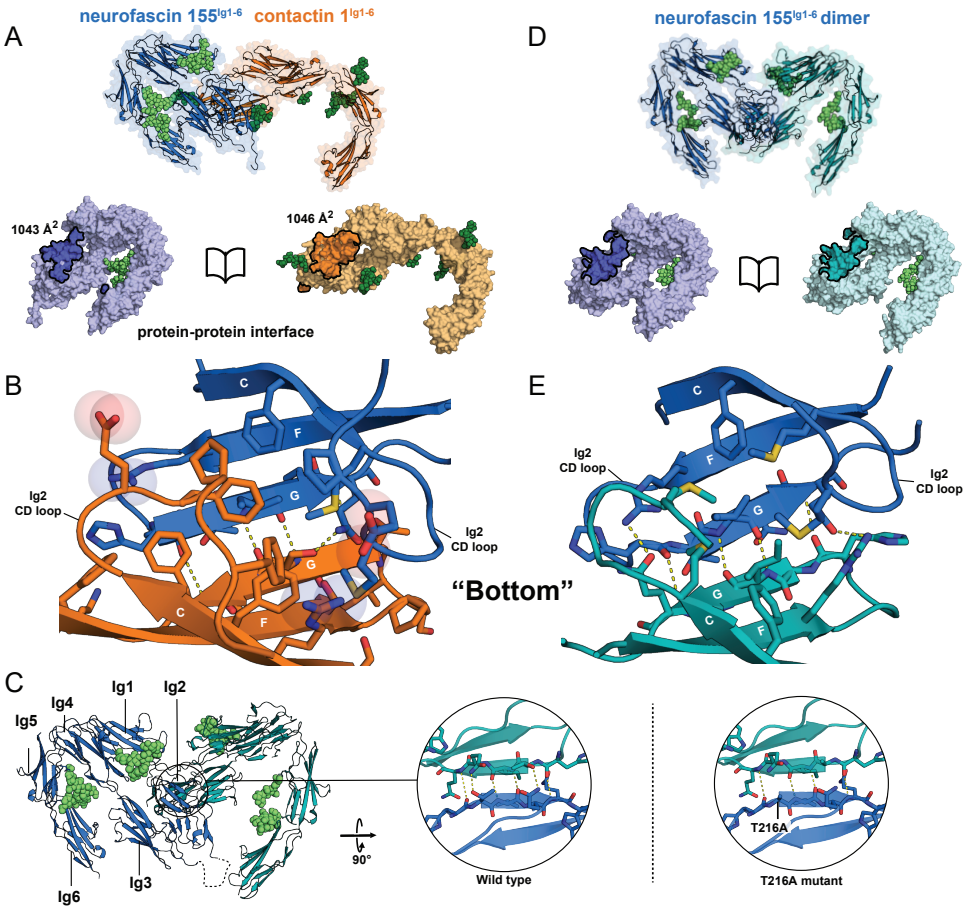


Figure 4.2: Neurofascin-155^{Ig1-6} dimerises using an Ig1-2 interface and the monomer has a distinctive garden hoe architecture. **A)** Cntn1^{Ig1-6}–Nf155^{Ig1-6} complex with open book surface representation below with interaction surfaces outlined. Buried solvent accessible surface areas are indicated per interface. **B)** Residues at the "bottom" of the Cntn1^{Ig1-6}–Nf155^{Ig1-6} GFC super beta sheet involved in interface formation. Nf155^{Ig1-6} dimerises using a similar interface as in the Cntn1^{Ig1-6}–Nf155^{Ig1-6} complex. **C)** Nf155^{Ig1-6} dimerisation site as wildtype and as modelled T279A mutant. Glycans are represented as green spheres. **D)** Nf155^{Ig1-6} dimer with open book surface representation below. **E)** Residues at the "bottom" of the Nf155^{Ig1-6}–Nf155^{Ig1-6} GFC super beta sheet involved in interface formation.

mentary Fig. 4.6C). The dimerisation interface overlaps with the binding site for Cntn1, suggesting that competition may exist between Nf155 dimerisation and Cntn1 binding (Fig. 4.2D). The organisation of the Nf155^{Ig1-6} homodimer resembles that of the Cntn1^{Ig1-6} -Nf155^{Ig1-6} complex except that a Nf155^{Ig1-6} homodimer molecule is replaced by a Cntn1^{Ig1-6} molecule. At the center of the Nf155 homodimerisation interface, established by Ig2 domains with more modest Ig1 contributions, a super β-sheet is formed by joining the G β-strands of both Ig2 domains (Fig. 4.2C,E, and Supplementary Fig. 4.6C). The hydrophobic interface at the "bottom side" of this super β-sheet is smaller in the Nf155^{Ig1-6} homodimer compared to the heterophilic complex, due to a larger

hydrophobic area on Cntn1 (Fig. 4.2E). The homodimerisation interface in Nf155$^{Ig1\text{-}6}$ buries ~1700 Å^2 of solvent accessible area, slightly less than the 2089 Å^2 that is buried in the Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ complex (Fig. 4.2A,D). Nf155$^{Ig1\text{-}6}$ dimerises in solution as assessed by a concentration-dependent peak shift and molecular weight increase in size-exclusion chromatography with multiangle-light scattering; (SEC-MALS) (Supplementary Fig. 4.6D). Consistent with crystallographic observations the type of N-glycosylation does not impact this process (Supplementary Fig. 4.6E). Nf155$^{Ig1\text{-}6}$ dimerisation is substantially reduced by a T216A mutation reported to disrupt Nf186$^{Ig1\text{-}4}$ dimerisation[32] by preventing two intermolecular hydrogen bonds in the G β-strands of Ig2 from forming (Fig. 4.2C). In SEC-MALS analysis the T216A mutant does not undergo a concentration-depended peak shift nor an increase in molecular weight (Supplementary Fig. 4.6D) indicating dimerisation is disrupted. The concentration-dependent dimerisation observed for wild-type Nf155$^{Ig1\text{-}6}$ in SEC-MALS - with mostly monomeric protein below 5 μM - suggests a K_D of dimerisation in the ~5-30 μM range (Supplementary Fig. 4.6D-E).

4.2.4. CELL CLUSTERING ASSAY SUPPORTS THE STRUCTURAL INSIGHTS

Seeking further validation of structural insights into Cntn1 and Nf155 adhesion, we set up a previously reported cell clustering assay to characterise adhesion complex formation[22,50,51]. Previous reports have suggested, that in the cellular context a high-mannose form of Cntn1 is a pre-requisite for Nf155 binding[45]. We confirmed this finding by treating Cntn1-expressing K562 cells with the mannosidase I inhibitor kifunensine, resulting in high-mannose type glycosylation. Cntn1-expressing cells did not cluster by themselves, irrespective of kifunensine treatment (Fig 3, Supplementary Fig. 4.8). This shows that Cntn1 does not engage in homophilic *trans* interactions On the other hand, Nf155-expressing cells engaged in homophilic cell-clustering, supporting previous reports of neurofascin mediated homophilic adhesion[32], and this interaction was not influenced by kifunensine treatment (Fig. 4.3, Supplementary Fig. 4.8). Kifunensine-treated Cntn1-expressing cells co-clustered with Nf155-expressing cells as indicated by the co-clustering ratio that shows the clusters contained both Cntn1 and Nf155-expressing cells (Fig. 4.3B). Untreated cells did not form such heterophilic co-clusters, i.e. the clusters consisted predominantly of Nf155-expressing cells (Fig. 4.3B). Overall, we confirm previous reports[45] that presence of high mannose glycans on Cntn1 are required for heterophilic interactions with Nf155 in the cellular setting. Interestingly, we additionally found that Cntn1-expressing cells not treated with kifunensine and thus containing complex glycans, co-clustered with Nf155-expressing cells as long as those cells were treated with kifunensine (Fig. 4.3 and Supplementary Fig. 4.8). When both cell types were treated with kifunensine, i.e., Cntn1 and Nf155 both have high-mannose glycans, no co-clustering occurred (Fig. 4.3 and Supplementary Fig. 4.8). These findings suggest that specific glycosylation types and patterns are required for interaction of Cntn1-Nf155 in the cellular context.

Based on the Cntn1-Nf155 structure we designed two interface mutant versions in which hydrophobic residues at the "bottom side" of the Ig2 GFC super β-sheet (Fig. 4.2A and Supplementary Fig. 4.6A) are changed to charged residues that we predict would prevent Cntn1-Nf155 heterophilic and Nf155 homophilic *trans* interaction. Surface residues F177, F180 and F212, were mutated to aspartate in Cntn1^{Mut} and surface residues F168, M170, M174 and I217 were mutated to aspartate in Nf155^{Mut} (Fig.

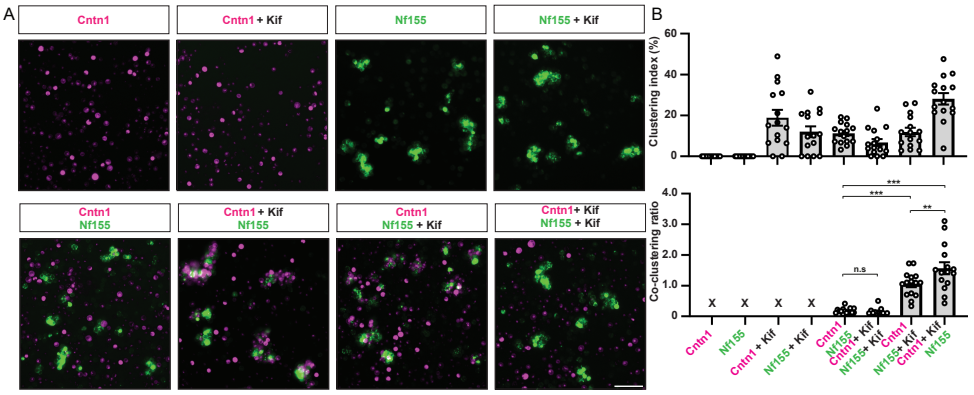


Figure 4.3: Contatin-1 and neurofascin-155 glycosylation regulates cell co-clustering. **A)** Representative cell clustering images of K562 cells expressing contactin-1 (mCherry; magenta; Cntn1) and neurofascin-155 (GFP; green; Nf155). Nf155 but not Cntn1-expressing cells exhibit homophilic clustering. Cntn1-Nf155 co-clusters only form when either Cntn1 or Nf155 is expressed in the presence of kifunensine (Kif; 10 µM), but not when neither or both are expressed in the presence of kifunensine. Scale bar is 100 µm. **B)** Quantification of clustering data in panel A. Bar plots represent the percentage that clusters contribute to the total segmented cell area (clustering index) per image, and the per-image average co-clustering ratio denoting the relative abundance of GFP signal per mCherry signal within each cluster. Per bar n=5 images per N=3 independent experiments, 15 images total. Data are presented as mean ± SEM; ns: not significant, **: $p < 0.01$, ***: $p < 0.001$. Statistical significance was determined by performing a one-way ANOVA followed by a Tukey's multiple comparison test. 'X' denote single-channel conditions in which no fluorescence ratio could be calculated.

4.4A). As expected, cells expressing Cntn1^{Mut} (treated with kifunensine) and cells expressing Nf155^{Mut} did not co-cluster (Fig. 4.4A and Supplementary Fig. 4.9). In addition, Cntn1^{Mut}-expressing cells (treated with kifunensine) did not co-cluster with Nf155^{WT}-expressing cells, nor did Nf155^{Mut}-expressing cells co-cluster with Cntn1^{WT}-expressing cells (treated with kifunensine). These experiments illustrate that the mutant versions independently abolish co-clustering and indicate that the hydrophobic surfaces we identify, both on the Cntn1 and Nf155 Ig2 domains, are required for heterophilic *trans* interaction. Furthermore, as expected, Nf155^{Mut}-expressing cells did not form homophilic cell clusters either (Fig. 4.4A and Supplementary Fig. 4.9), confirming that the heterophilic adhesion site of Cntn1-Nf155 overlaps with the homophilic adhesion site of Nf155.

Taken together, the cell clustering experiments show that Cntn1-Nf155 *trans* interaction is mediated through Ig2 as observed in the complex structure, that Nf155 homophilic *trans* interaction is mediated through the interface we observe in the Nf155$^{Ig1\text{-}6}$ dimer structure and, as expected, that Cntn1 does not engage in homophilic *trans* interactions.

4.3. DISCUSSION

Members of the contactin and L1 families act concomitantly through functionally diverse oligomers and complexes to establish adhesion underlying neuronal tissue wiring and connectivity[8]. The consensus view from structural and biophysical methods has been that in these families the first Ig1-4 domains have a characteristic backfolded horseshoe architecture[22–26,31,32,52,53]. In many cases, immunoglobulin domains and particu-

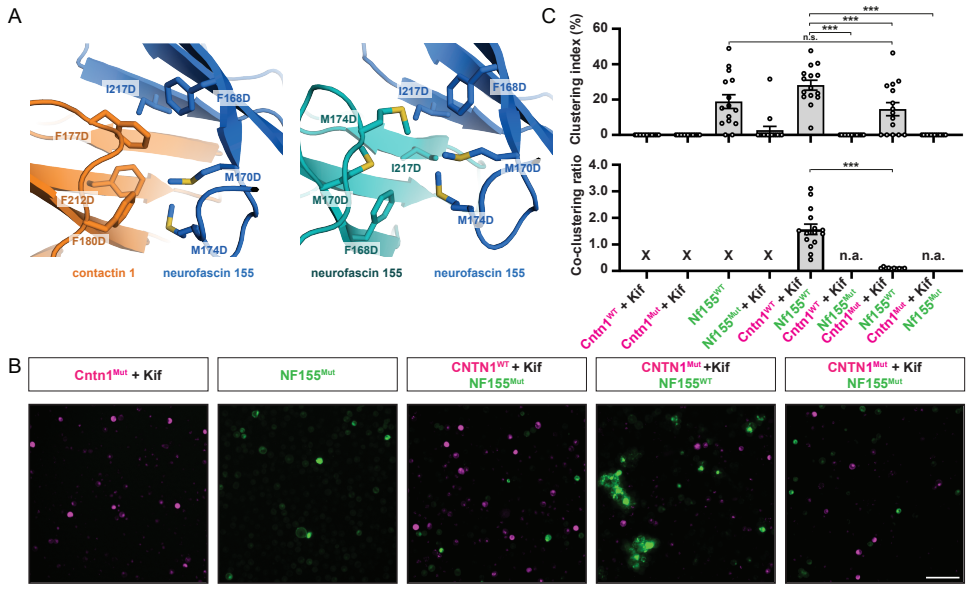


Figure 4.4: The resolved X-ray crystallography-resolved interfaces mediate clustering. **A)** Locations on contactin-1 (Cntn1) and neurofacin-155 (Nf155) contacts where the structure-based mutations are introduced to disrupt homophilic and heterophilic clustering. Cntn1-Nf155 interface residues, shown in stick representation, at the "bottom side" of the Ig2 super β-sheet (left panel) are mutated to aspartates to prevent heterophilic *trans* interactions. The same Nf155 residues are also located in the Nf155 homodimer interface (right panel) and mutating them to aspartates prevents homophilic *trans* interactions. **B)** Representative cell clustering images of K562 cells expressing contactin-1 (mCherry; magenta; Cntn1) and neurofascin-155 (GFP; green; Nf155). Wildtype (WT) conditions are compared to mutants (Mut) with the mutations represented in panel A. As a baseline, the most clustering WT condition was taken as a baseline, in which Cntn1 but not Nf155 is expressed in the presence of kifunensine (Fig. 4.3B). Scale bar is 100 µm. **C)** Quantification of clustering data in panel B. Bar plots represent the clustering index and co-clustering defined in Figure 4.3. Per bar n=5 images per N=3 independent experiments, 15 images total (14 images for $Cntn1^{Mut}$-$Nf155^{Mut}$. Data are presented as mean ± SEM; ns: not significant, ***: $p < 0.001$. Statistical significance was determined by performing a one-way ANOVA followed by a Tukey's multiple comparison test. In absence of clusters for analysis 'n.a' is indicated, and 'X' denote single-channel conditions in which no fluorescence ratio could be calculated.

larly the horseshoe supramodule, have been established as necessary and sufficient for *trans*-adhesion [12,23,31,32,46,47,52,54–58]. Here we show that the mouse $Cntn1^{Ig1\text{-}6}$–$Nf155^{Ig1\text{-}6}$ complex, and the $Nf155^{Ig1\text{-}6}$ dimer are formed via Ig1-2 interactions between the horseshoe modules in a mode similar to that described for the human contactin-2 Ig1-4 [23] and Nf186 Ig1-4 [32] horseshoe homodimers. This common mode of interaction suggests that the heterophilic interaction between Cntn1 and Nf155 may have emerged from duplication of ancestral homodimerising proteins [8,9]. Heterophilic interactions emerging in protein families from ancestral homodimers are proposed to lead to heterodimeric species with varying affinities that may be exploited to finetune spatio-temporal control of cellular responses [59]. This notion is supported on a subcellular level by the similarity in interaction mechanisms for the Cntn1-Nf155 heterophilic *trans*-complex that functions in myelin paranode junction formation [13,15,60] and the contactin-2 homophilic *trans*-complex that is required for myelin juxtaparanode formation [10,11] (Fig. 4.5).

Cntn1 and Nf155, together with CASPR1, control paranodal adhesion [13–15,49] and

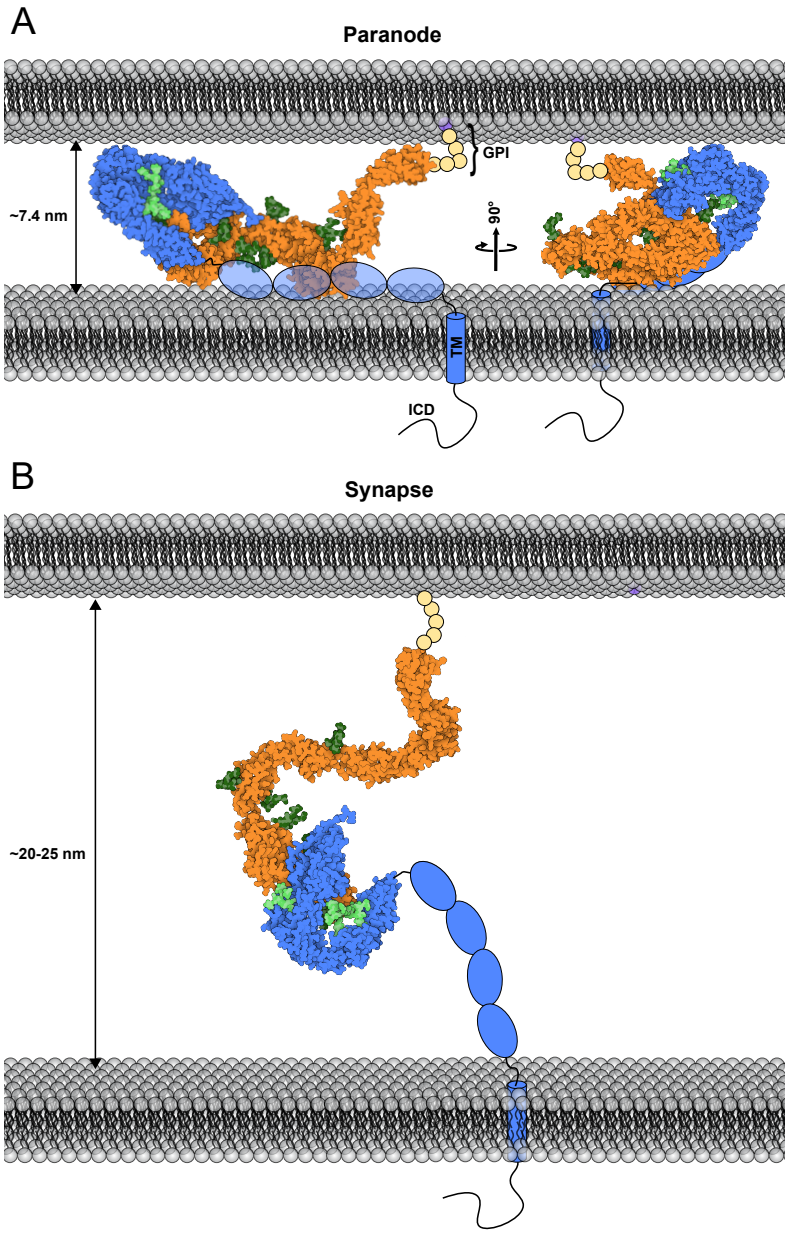


Figure 4.5: Contactin-1 – neurofascin-155 *trans*-adhesion can span distinct intercellular distances. **A)** Model of the Cntn1-Nf155 complex in the ~7.4 nm paranodal intermembrane spacing, with molecules laying flat on the membranes and horseshoes determining minimum intermembrane distance. **B)** Model of the Cntn1–Nf155 complex spanning the ~20-25 nm synaptic cleft, enabled by the length and flexibility of the C-terminal domains. Missing FnIII domains are shown as ovals. Glycophosphatidylinositol (GPI) anchor, intracellular domain (ICD).

maintain the intercellular distance of 7.4 nm ± 0.6 nm[38] important for maturation and homeostasis of myelinated fibers[28–30] ensuring effective saltatory conduction[39,40]. The Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ complex has an edge-on size of ~7 nm (Fig. 4.5A) indicating that both proteins must lay flat between the two cellular surfaces with the horseshoe modules determining the paranodal spacing, similar to what has been proposed for horseshoe-containing sidekick proteins in the retina[61]. Interestingly, Cntn1 and Nf155 are also found[62,63] and interact in the neuronal synapse[16] which has a much larger intercellular distance of ~20-25 nm[41–43]. The flexibility reported by others in the FnIII 1-3 contactin segment[26,64], and flexibility at the membrane attachment sites may allow to span the intercellular synapse space (Fig. 4.5B) and enable these molecules to function at different cell-cell distances. If and how structural plasticity of Cntn1 and Nf155 plays a role in cellular adhesion will require further functional verification. The expression levels of specific isoforms[15,34,35], the localisation in space and time[8,9] and posttranslational modification of Cntn1[27–29] and neurofascin[12,65–67] impact the development and function of the nervous system. Nf155 homodimerisation and Cntn1-Nf155 complex interfaces overlap, indicating these interactions exclude one another. In the biological context, the preference for either of the interactions is likely modulated by factors affecting effective concentration of components, such as expression levels, anchoring and trafficking processes, distances between molecules on apposed membranes, and whether the competing interactions can occur in a *cis* or in a *trans* setting. In addition, neurofascin has many isoforms[35] and their expression levels and location are often distinct[15,34–37,68,69]. The structure of Nf155$^{Ig1\text{-}6}$ is very similar to Nf186 Ig1-4 and together show that the Cntn1 binding site on both molecules is very similar (data not shown). Given isoform splice differences are outside of the Ig1-Ig2 region, all neurofascin isoforms could likely bind Cntn1 using the interface reported here on Ig1-Ig2. In line with our model, Cntn1 would be able to interact with any neurofascin isoform, a finding supported by earlier neurofascin isoform – Cntn1 interaction studies that show unambiguous binding with differing relative efficiencies[12]. The distinctive Nf155$^{Ig1\text{-}6}$ garden hoe-shaped architecture may be regulated by splicing at the N-terminus (Fig. 4.2 and Supplementary Fig 4.6B), possibly affecting the presentation of the Cntn1 binding site and explaining splicing-dependent difference in interaction[12] and absence of interaction in a cellular context upon deletion of the Ig 5-6 domains in neurofascin[47]. Additionally, N-linked glycosylation has been shown to regulate binding of Cntn1 and Nf155 in a cellular context[44–46] and we show in this context that high-mannose glycans, either on Cntn1 or on Nf155 but not on both, enable heterophilic cell-cell interactions. From the Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ structure, conserved glycosylation sites in the immunoglobulin segments appear poised to extend or sterically hinder interaction depending on glycosylation microheterogeneity. Taken together, a balance of pre- and posttranslational modification, localisation and protein concentrations underlies Cntn1 – neurofascin interaction.

Evidence in recent years has emerged implicating both Cntn1 and Nf155 as central players in a host of pathologies. Both proteins are intimately associated with autoimmune neuropathies[70], such as multiple sclerosis[20,71]. Furthermore, contactin may play a role in neurodegenerative diseases with reports implicating it in Parkinson's[72] and Alzheimer's disease[73,74], but also cancers where it impacts cancer progression and metastasis[75,76]. Our work on the Cntn1-Nf155 complex and the individual proteins,

may provide a steppingstone for the development of future therapeutics targeting these molecules.

4.4. METHODS

PLASMIDS

Contactin-1 (Cntn1) based on Image Clone 30099512, and human cell line expression codon optimised Nf155 (Nf155) based on NCBI transcript variant NM_001160316.1 obtained from (GeneArt ThermoFisher), were used as templates to generate Cntn1$^{Ig1\text{-}6}$ (residues 21-604), Nf155$^{Ig1\text{-}6}$ (residues 25-633), and full ectodomains (fe) Cntn1^{fe} (residues 21-996) and Nf155^{fe} (residues 25-1059) constructs by polymerase chain reaction (PCR). Nf155$^{Ig1\text{-}6}$ T216A, Cntn1^{fe} F177D, F180D, F212D (Cntn1^{Mut}) and Nf155^{fe} F168D, M170D, M174D, I217D (Nf155^{Mut}) mutants were created using overlapping primers. Cntn1^{Mut} was subcloned into pUPE107.58 (cystatin secretion signal peptide, C-terminal transmembrane helix–mCherry fluorophore–His6 tag, U-Protein Express) and Nf155^{Mut} that was subcloned into pUPE107.21 (cystatin secretion signal peptide, C-terminal transmembrane helix–eGFP fluorophore–His6 tag, U-Protein Express). Nf155$^{Ig1\text{-}6}$ were subcloned using BamHI/NotI sites into pUPE107.03 (cystatin secretion signal peptide, C-terminal His6 tag, U-Protein Express), and Cntn1$^{Ig1\text{-}6}$ was subcloned into pUPE10762 (cystatin secretion signal peptide, C-terminal biotin acceptor peptide-His6 tag) vector (U-Protein Express). Cntn1^{fe} and Nf155^{fe} were subcloned into pUPE107.58 and pUPE107.21, respectively (called Cntn1^{WT} and Nf155^{WT}).

PROTEIN EXPRESSION AND PURIFICATION

Complex glycan (CG) proteins were produced in suspension preparations of Epstein–Barr virus nuclear antigen I (EBNA1)-expressing HEK293 cells (HEK293-E) (U-Protein Express), while high mannose (HM) proteins were produced in N-acetylglucoaminyltransferase I-deficient (GnTI-) EBNA1-expressing HEK293 cells (HEK293-ES) (U-Protein Express). Medium was harvested six days after transfection and cells were spun down by 10 min of centrifugation at 1000 g. Cellular debris was then spun down from medium for 15 min at 4000 g. Protein was purified using Ni Sepharose excel (GE Healthcare) affinity chromatography followed by size exclusion chromatography (SEC) on either Superdex200 Hiload 16/60 (GE Healthcare) or Superdex200 10/300 (GE Healthcare) columns equilibrated in SEC buffer (25 mM HEPES pH 7.5, 150 mM NaCl). Protein was then concentrated to 5-10 mg/mL and stored at -80°C. Purity was evaluated by SDS-PAGE and Coomassie staining.

CRYSTALLISATION AND X-RAY DATA COLLECTION

Sitting-drop vapour diffusion at 4°C was used for all crystallisation trials, by mixing 150 nL of protein solution with 150 nL of reservoir solution. Crystals of Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ (equimolar ratio final mixture 6 mg/mL) grew from a condition with 2% v/v TacsimateTM pH 5.0, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, and 16% w/v PEG 3,350. Crystals of Nf155 $^{Ig1\text{-}6}$ (8 mg/mL) grew from a condition with 20% (w/v) PEG 8000 and 100 mM HEPES/Sodium hydroxide pH 7.5. Reservoir solution supplemented with 30% of glycerol was added as cryo-protectant to the crystals before plunge freezing them in liquid nitro-

gen. Cntn1$^{Ig1\text{-}6}$ – Nf155 $^{Ig1\text{-}6}$ and Nf155$^{Ig1\text{-}6}$ data sets were collected at 100K at Diamond Light Source beamline I03 at a wavelength of 0.9763 Å.

STRUCTURE DETERMINATION AND REFINEMENT

Integrated data were obtained for Nf155 $^{Ig1\text{-}6}$ from the xia2 dials diamond beamline data auto processing pipeline [77], and further processed in AIMLESS [78]. Resolution limit cut off was determined based on mean intensity correlation coefficient of half-data sets, $CC_{1/2}$. Unmerged and unscaled data were obtained from the xia2 dials diamond beamline data auto processing pipeline [77] for two isomorphic Cntn1$^{Ig1\text{-}6}$ – Nf155 $^{Ig1\text{-}6}$ datasets collected in succession from the same crystal. These data were used to produce a combined unmerged and unscaled dataset using POINTLESS [79]. Anisotropic cut-off, merging and scaling of the combined dataset was performed by the STARANISO [80] webserver. Following recommendations from the STARANISO webserver, unobserved and unobservable reflections that lie outside the diffraction cut-off surface were removed from the merged data file. All structures were solved by molecular replacement using PHASER [81]. Initial search model for unliganded Nf155$^{Ig1\text{-}6}$ was (PDB: 3P3Y) [32] for neurofascin Ig1-4 residues 31-437, and a PHYRE2 [82] model was generated with >90% confidence for Ig5 and Ig6 domains. For Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ complex structure, refined structures from unliganded datasets were used as search models. Structure refinement of unliganded datasets was performed using PHENIX [83] with automatic weighting options. Manual model building was done in COOT [84]. Manual inspection and correction of unliganded structures was performed iteratively in COOT [84] in between automated refinement runs to correct register errors. Readjusting the positioning of residues or correcting the register was done on the following parts: Nf155 Ig 5 residues 445-449 and 482-492, Nf155 Ig 6 residues 564-584, Cntn1 Ig 5 residues 415-422, and Cntn1 Ig 6 residues 549-566. To minimise overfitting of the unliganded Nf155 $^{Ig1\text{-}6}$, secondary structure and NCS restraints were applied. Unliganded Nf155$^{Ig1\text{-}6}$ was additionally refined with TLS parameters given the observed mobility of domains between chains within the dataset. Structure refinement of the Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ dataset was initially performed using REFMAC [85] with manual model building performed in COOT [84]. Given the low resolution and highly anisotropic data for the Cntn1$^{Ig1\text{-}6}$–Nf155$^{Ig1\text{-}6}$ complexed dataset, special care was given in refinement to avoid over parametrisation. MR placed domain positions were first refined using jelly body restraints and conservative weighting (0.0002) between X-ray and geometric restraints. B factors were also further set to a constant value and only TLS group B factor refinement was performed on the structure. To optimise glycan geometry the final refinement step was performed in PHENIX where the same TLS groups were used and a single grouped B factor per chain was refined with conservative X-ray/geometry weighting (wxc of 0.05) and optimisation of X-ray/ADP weighting. Also here starting model, NCS and secondary structure restraints were used. MOLPROBITY [86] was used for structure validation. The composite OMIT map was calculated in PHENIX using the "simple" method [87].

STRUCTURAL ANALYSES

Structural analyses were performed using various relevant programs. Interface properties and buried surface areas were determined for analysis with the jsPISA [48] server. Hy-

drophobic surface representation colouring was obtained using the YRB colouring scheme [88]. Electrostatic surface properties at pH 7.4 were obtained using the PDB2PQR[89] and APBS[90] webservers. Conservation analyses were performed using CONSURF[91] with curated sequence lists retrieved for vertebrate orthologues from UNIPROT database and chosen paralogues from the NCBI database[92]. For paralogue conservation analyses, sequences L1 and contactin paralogues from human, mouse, chicken, xenopus and zebrafish were selected. Figures were generated with PyMol (Schrödinger), and the ILLUSTRATE[93] webserver. Sequence alignment of L1 mouse paralogues highlighting architecture was prepared using ESPript webserver[94].

SIZE EXCLUSION CHROMATOGRAPHY AND MULTI-ANGLE LIGHT SCATTERING

SEC analysis was performed on $Nf155^{Ig1\text{-}6}$ HM and CG wt, and T216A HM samples to characterise monomer dimer exchange through peak shift. Purified samples (1-100 µL) were injected onto a Superdex200 10/300 increase (GE Healthcare) column equilibrated in SEC buffer and separated with a flow rate of 0.75 mL/min. For molecular weight characterisation of $Nf155^{Ig1\text{-}6}$ HM and T216A HM samples, light scattering measurements were performed using a miniDAWN TREOS multi-angle light scattering detector (Wyatt), connected to a differential refractive index monitor (Shimadzu, RID-10A) used for protein concentration quantification. Collected chromatograms were analyzed and processed using ASTRA6 software (Wyatt, using a calculated dn/dc value of 0.182 mL/g, determined from dn/dc of 0.188 and 0.145 for the protein and glycan parts respectively, and 8% glycosylation estimated from crystallographically confirmed glycosylation sites). Instrument calibration was assessed by injection of 5 mg/mL monomeric conalbumin (Sigma-Aldrich), using in this case a dn/dc value of 0.185 mL/g.

CELL CLUSTERING ASSAY, IMAGING AND ANALYSIS

K562 cells (kind gift from Dr. Bas van Steensel, Netherlands Cancer Institute) were cultured in RPMI-1640 medium (Gibco), supplemented with 10% FBS (Gibco) and 1% Penicillin/Streptomycin (Gibco), and grown in a shaking incubator at 37°C and 5% CO_2. Prior to electroporation, K562 cells were collected from culture flasks and centrifuged for 5 min at 300 g. Cells were then washed in 1x PBS (Gibco), centrifuged for 5 min at 300 g and resuspended in buffer R (Gibco). Per condition, $2{\cdot}10^6$ cells were incubated with a total amount of 15 µg of DNA for 15 min at room temperature. The ratio of $Cntn1^{WT}$ or $Nf155^{WT}$ plasmid to empty vector was 10:1, while the ratio of $Cntn1^{Mut}$ or $Nf155^{Mut}$ plasmid to empty vector was 1:1. After the incubation, K562 cells, in presence of DNA mixtures, were electroporated with the Neon Transfection System (Thermo Fisher Scientific), using 100 µL tips and the following settings: 1450 V, 10 ms pulse length, and 3 pulses[95]. After electroporation, cells were directly plated in 6-well plates onto 5 mL of pre-warmed RPMI-1640 medium with 10% FBS and either 0.1% DMSO (Sigma-Aldrich) or 10 µM Kifunensine in 0.1% DMSO (Kif; Sigma-Aldrich). Cells were allowed to recover for approximately 20 h in a shaking incubator at 37°C and 5% CO_2. After recovery, cells were collected and centrifuged for 3 min at 200 g. Cells were then resuspended in assay medium (RPMI-1640 supplemented with 10% FBS) and treated with DNase I (Invitrogen) for 10 min at 37°C. Cells were once again centrifuged, resuspended in assay

medium and passed through a 40 µm cell strainer. Cells were counted using the Countess 3 FL (Thermo Fisher Scientific), and a total of $2{\cdot}10^5$ cells per clustering condition were plated in a 12-well plate in 1 mL of assay medium. Cells were left to cluster for 24 h on a shaking incubator at 37°C and 5% CO_2 and clusters were imaged on an EVOS M5000 microscope with a 10x objective (0.25 NA; EVOS, Thermo Fisher Scientific), using the EVOS LED GFP and RFP cubes (Thermo Fisher Scientific). For the analysis, GFP and mCherry channels were combined and Otsu thresholding was applied. Regions of interest (ROIs) larger than 50 pixels were identified using Analyze Particles[96]. Rolling ball background subtraction with 50-pixel radius was performed on the individual GFP and mCherry channels, before measuring the area and mean intensities of GFP and mCherry channels of each ROI. A cell cluster was defined as an object three times larger than the determined mean large single cell size calculated from the largest 2.5% in a single image consisting of 1,637 single cell ROIs after removal of the largest 10 ROIs. The clustering index was determined as the summed cluster area divided by summed area of all ROIs (clusters + non-clusters) times 100%. The cluster size, co-cluster ratio, and clustering index were averaged per image, and data from three independent experiments (5 images per experiment, 15 images total per condition) were analysed using Python 3[97] and Seaborn[98] statistical data visualisation.

For the comparison of multiple groups, we used a one-way ANOVA followed by a Tukey's multiple comparison test using GraphPad Prism (GraphPad Software, San Diego, California USA). Differences between conditions were considered significant when $p < 0.05$ (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). In all figure legends, N indicates the number of independent experiments, and n indicates the number images analyzed. Data are represented as mean values ± SEM.

ACKNOWLEDGEMENTS

Structural data from Protein Data Bank (PDB) were used under the CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license. We thank the staff of the DLS beamlines I03 and I24 for help with X-ray diffraction data collection. L.M.P.C thanks Nick Pearce, Jitse van der Horn, and Gijs van der Schot, for the instructional conversations regarding crystallography. K562 cells were a kind gift from Dr. Bas van Steensel at Netherlands Cancer Institute. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme with grant agreement No. 677500 (to B.J.C.J.). D.H.M. acknowledges support from Parents in KIND grant, sponsored by the Kavli Institute of Nanoscience, the Department of Bionanoscience in Delft, and the NWO Spinoza Prize. M.A.d.B and A.J.R.H acknowledge support from the Netherlands Organisation for Scientific Research (NWO) funding the Netherlands Proteomics Centre through the X-omics Road Map program (project 184.034.019).

4.5. SUPPLEMENTARY INFORMATION

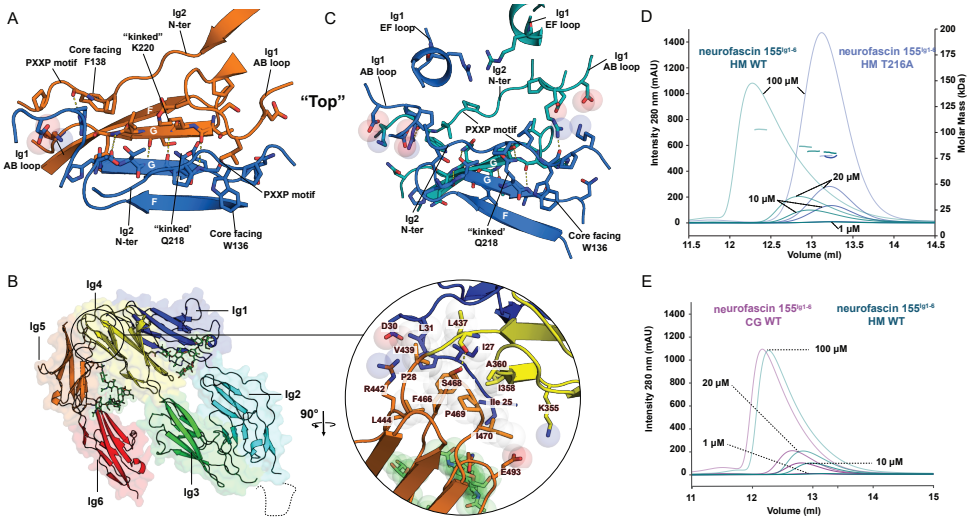


Figure 4.6: Characterisation of neurofascin-155 monomeric architecture, and of homodimer structure and binding affinity. **A)** Residues at the "top" of the Cntn1^{Ig1-6}–Nf155^{Ig1-6} GFC super beta sheet involved in interface formation. **B)** Nf155^{Ig1-6} monomer (domain rainbow coloured) and Nf155^{Ig1-6} dimer (coloured by chain) with glycans shown as spheres (lime). The monomer has a hoe-like shape in which the Ig1-Ig4 horseshoe represents the "blade" and domains Ig5 and Ig6, that tovgether make a V-turn with respect to Ig3-Ig4, represent the "handle". **C)** Residues at the "top" of the Nf155^{Ig1-6}–Nf155^{Ig1-6} GFC super beta sheet involved in interface formation. Nf155^{Ig1-6} dimerises using a similar interface as in the Cntn1^{Ig1-6} – Nf155^{Ig1-6} complex. **D)** Nf155^{Ig1-6} wildtype (WT) and T279A mutant SEC-MALS analysis shows the mutant disrupts dimerisation. Protein concentrations at injection are indicated. **E)** Nf155^{Ig1-6} WT high-mannose (HM) and complex glycan (CG) SEC analysis.

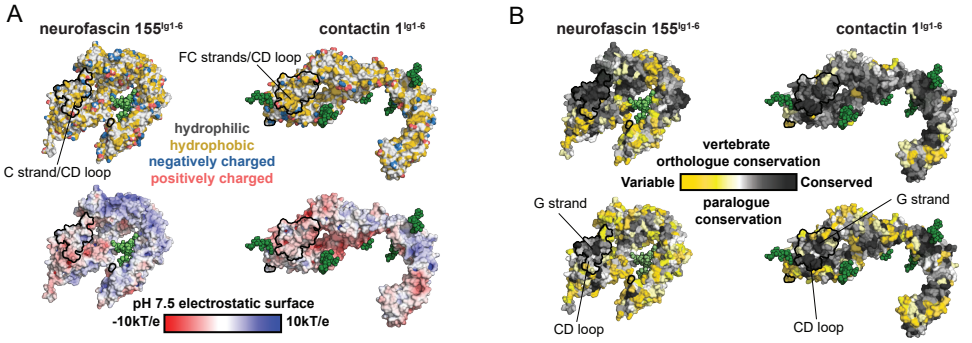


Figure 4.7: Physical nature and conservation of the Cntn1^{Ig1-6}–Nf155^{Ig1-6} interface. **A)** Open book representation of the Cntn1^{Ig1-6}-Nf155^{Ig1-6} complex in surface representation coloured according to various properties and with interaction surfaces outlined. The interface is relatively conserved, suggesting a role in adhesion.

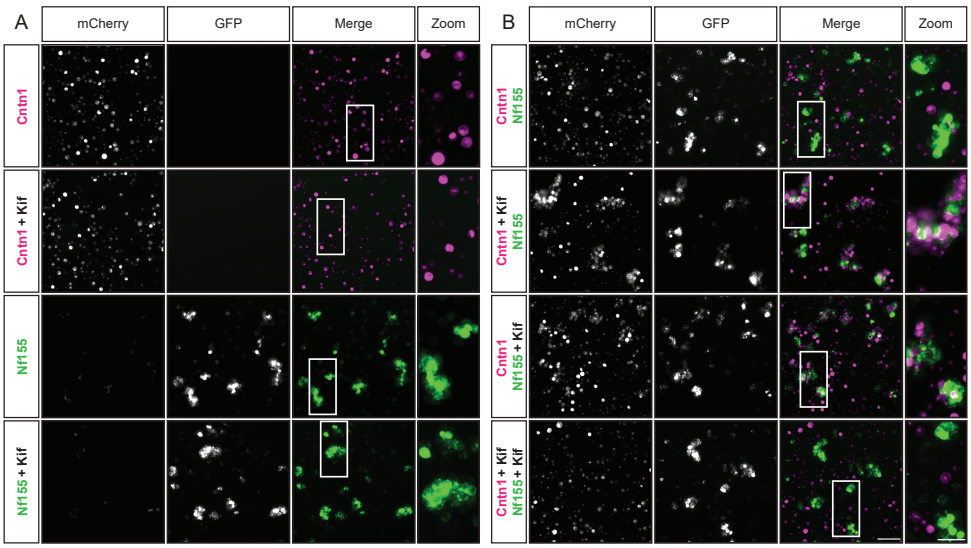


Figure 4.8: Cntn1-Nf155 glycosylation regulates cell co-clustering. **A)** Representative cell clustering images of K562 cells expressing contactin-1 (mCherry; magenta; Cntn1) and neurofascin-155 (GFP; green; Nf155) individual conditions. **B)** Co-clustering between Cntn-1 and Nf155-expressing cells in variable presence of kifunensine (Kif; 10 µM). Scale bar is 100 µm.

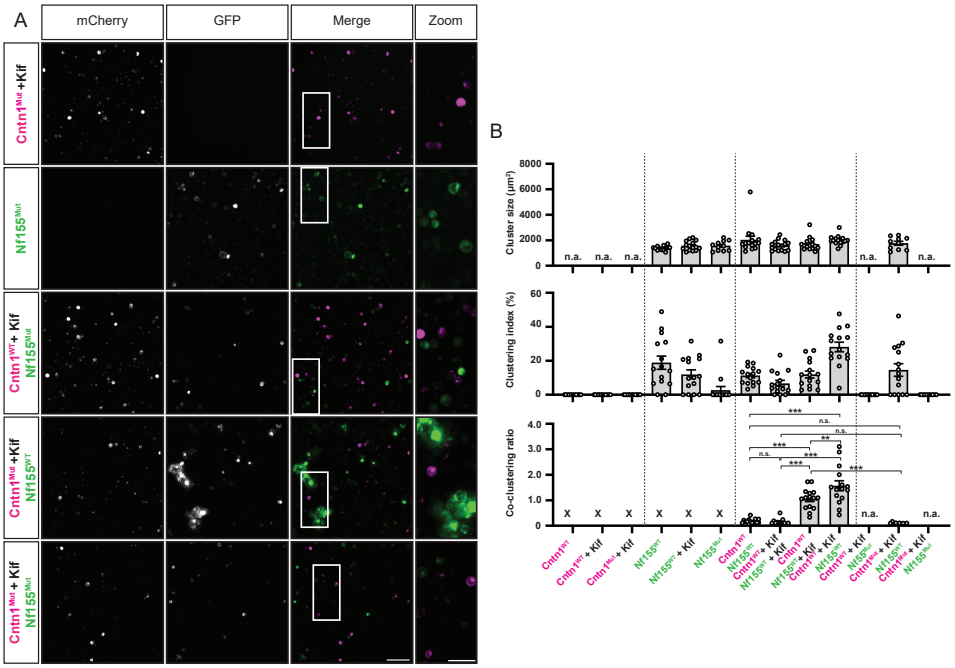


Figure 4.9: The interfaces from the structural data mediate clustering. **A)** Representative cell clustering images of K562 cells expressing contactin-1 (mCherry; magenta; Cntn1) and neurofascin-155 (GFP; green; Nf155). Wildtype (WT) conditions are compared to mutants (Mut) for which the specific mutagenic sites are represented in Figure 4.4A. Scale bar is 100 µm. **B)** Quantification of clustering data in panel A. Bar plots represent the cluster size, the percentage that clusters contribute to the total segmented cell area (clustering index) per image, and the per-image average co-clustering ratio denoting the relative abundance of GFP signal per mCherry signal within each cluster. Per bar n=5 images per N=3 independent experiments, 15 images total. Data are presented as mean ± SEM; ns: not significant, **: $p < 0.01$, ***: $p < 0.001$. Statistical significance was determined by performing a one-way ANOVA followed by a Tukey's multiple comparison test. In absence of clusters for analysis 'n.a' is indicated, and 'X' denote single-channel conditions in which no fluorescence ratio could be calculated.

Table 4.1: X-ray diffraction data collection and refinement statistics.

Crystal	**Cntn1$^{Ig1–6}$-Nf155$^{Ig1–6}$**	**Nf155$^{Ig1–6}$**
Data collection		
Space group	*P2$_1$*	*P2$_1$2$_1$2$_1$*
Cell dimensions		
a, *b*, *c* (Å)	146.8, 151.8, 162.3	77.6, 96.3, 238.6
α, β, γ (°)	90, 111.8, 90	90, 90, 90
Resolution (Å)	75.9 – 4.8 (5.3 – 4.8)a	79.5 – 3.0 (3.02 – 2.97)
R_{merge}	0.312 (1.926)	0.168 (1.938)
$I/\sigma I$	4.5 (1.3)	6.2 (1.0)
Completeness (%) (spherical)	48.2 (8.7)	99.7 (97.4)
Redundancy	7.1 (7.5)	6.8 (6.1)
$CC_{1/2}$	0.993 (0.399)	0.976 (0.399)
Anisotropic processing		
Fitted ellipsoid diffraction limit (Å)		
0.999a - 0.033c	8.01	N/A
b	4.78	N/A
-0.298a + 0.955c	5.35	N/A
Lowest diffraction limit	7.85	N/A
Worst diffraction limit	8.17	N/A
Best diffraction limit	4.8	N/A
Overall anisotropy tensor Eigenvectors and Eigenvalues		
0.999a - 0.049c	16.94	N/A
b	37.54	N/A
-0.287a + 0.958c	37.44	N/A
Refinement		
Resolution (Å)	75.4–4.8	60.4–3.0
No. of reflections	15662	37575
R_{work}/R_{free}	0.29/0.32	0.22/0.27
No. of atoms		
Protein	18,329	9191
Ligand/carb	971	355
Clashscore	6.1	5.6
B factors (Å^2)		
Protein	257	112
Ligand/carb	302	151
TLS groups	24	8
R.m.s.d.		
Bond lengths (Å)	0.004	0.002
Bond angles (°)	0.79	0.55
Ramachandran		
Favored (%)	93.1	97.2
Outliers (%)	0	0.1
Rotamer		
Outliers (%)	3.4	1.5
PDB ID	7OL4	7OK5

REFERENCES

[1] L. M. Chataigner, C. Gogou, M. A. den Boer, C. P. Frias, D. M. Thies-Weesie, J. C. Granneman, A. J. Heck, D. H. Meijer, and B. J. Janssen, *Structural insights into the contactin 1 – neurofascin 155 adhesion complex,* Nature Communications **13** (2022), 10.1038/s41467-022-34302-9.

[2] A. R. Aricescu and E. Y. Jones, *Immunoglobulin superfamily cell adhesion molecules: zippers and signals,* CURRENT OPINION IN CELL BIOLOGY **19**, 543 (2007).

[3] B. Honig and L. Shapiro, *Adhesion protein structure, molecular affinities, and principles of cell-cell recognition,* CELL **181**, 520 (2020).

[4] L. R. Herron, M. Hill, F. Davey, and F. J. Gunn-Moore, *The intracellular interactions of the l1 family of cell adhesion molecules,* BIOCHEMICAL JOURNAL **419**, 519 (2009).

[5] J. Morris, S. S. Y. Leung, M. E. Bailey, B. Cullen, A. Ferguson, N. Graham, K. J. A. Johnston, D. M. Lyall, L. M. Lyall, J. Ward, D. J. Smith, and R. J. Strawbridge, *Exploring the role of contactins across psychological, psychiatric and cardiometabolic traits within uk biobank,* Genes **11** (2020), 10.3390/genes11111326.

[6] A. N. Mohebiany, S. Harroch, and S. Bouyain, *New insights into the roles of the contactin cell adhesion molecules in neural development.* Advances in neurobiology **8**, 165 (2014).

[7] V. Sytnyk, I. Leshchyns'ka, and M. Schachner, *Neural cell adhesion molecules of the immunoglobulin superfamily regulate synapse formation, maintenance, and function,* TRENDS IN NEUROSCIENCES **40**, 295 (2017).

[8] M. Hortsch, *Structural and functional evolution of the l1 family: Are four adhesion molecules better than one?* MOLECULAR AND CELLULAR NEUROSCIENCE **15**, 1 (2000).

[9] Y. Shimoda and K. Watanabe, *Contactins emerging key roles in the development and function of the nervous system,* CELL ADHESION & MIGRATION **3**, 64 (2009).

[10] C. Lubetzki, N. Sol-Foulon, and A. Desmazieres, *Nodes of ranvier during development and repair in the cns,* NATURE REVIEWS NEUROLOGY **16**, 426 (2020).

[11] M. N. Rasband and E. Peles, *Mechanisms of node of ranvier assembly,* Nature Reviews Neuroscience **22**, 7 (2021).

[12] H. Volkmer, U. Zacharias, U. Nörenberg, and F. Rathjen, *Dissection of complex molecular interactions of neurofascin with axonin-1, f11, and tenascin-r, which promote attachment and neurite formation of tectal cells,* JOURNAL OF CELL BIOLOGY **142**, 1083 (1998).

[13] M. Boyle, E. Berglund, K. Murai, L. Weber, E. Peles, and B. Ranscht, *Contactin orchestrates assembly of the septate-like junctions at the paranode in myelinated peripheral nerve,* NEURON **30**, 385 (2001).

[14] P. Charles, S. Tait, C. Faivre-Sarrailh, G. Barbin, F. Gunn-Moore, N. Denisenko-Nehrbass, A. Guennoc, J. Girault, P. Brophy, and C. Lubetzki, *Neurofascin is a glial receptor for the paranodin/caspr-contactin axonal complex at the axoglial junction,* CURRENT BIOLOGY **12**, 217 (2002).

[15] D. Sherman, S. Tait, S. Melrose, R. Johnson, B. Zonta, F. Court, W. Macklin, S. Meek, A. Smith, D. Cottrell, and P. Brophy, *Neurofascins are required to establish axonal domains for saltatory conduction,* NEURON **48**, 737 (2005).

[16] M. A. Gonzalez-Lozano, F. Koopmans, P. F. Sullivan, J. Protze, G. Krause, M. Verhage, K. W. Li, F. Liu, and A. B. Smit, *Stitching the synapse: Cross-linking mass spectrometry into resolving synaptic protein interactions,* SCIENCE ADVANCES **6** (2020), 10.1126/sciadv.aax5783.

[17] P. F. Maness and M. Schachner, *Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration,* NATURE NEUROSCIENCE **10**, 19 (2007).

[18] F. Colombo and J. Meldolesi, *L1-cam and n-cam: From adhesion proteins to pharmacological targets,* TRENDS IN PHARMACOLOGICAL SCIENCES **36**, 769 (2015).

[19] M. Chatterjee, D. Schild, and C. E. Teunissen, *Contactins in the central nervous system: role in health and disease,* NEURAL REGENERATION RESEARCH **14**, 206 (2019).

[20] M. Chatterjee, M. J. Koel-Simmelink, I. M. Verberk, J. Killestein, H. Vrenken, C. Enzinger, S. Ropele, F. Fazekas, M. Khalil, and C. E. Teunissen, *Contactin-1 and contactin-2 in cerebrospinal fluid as potential biomarkers for axonal domain dysfunction in multiple sclerosis.* Multiple sclerosis journal - experimental, translational and clinical **4**, 2055217318819535 (2018).

[21] A. Zuko, S. Bouyain, B. van der Zwaag, and J. P. H. Burbach, *Contactins: Structural aspects in relation to developmental functions in brain disease,* in *ADVANCES IN PROTEIN CHEMISTRY AND STRUCTURAL BIOLOGY, VOL 84*, Advances in Protein Chemistry and Structural Biology, Vol. 84, edited by R. Donev (2011) pp. 143–180.

[22] J. Freigang, K. Proba, L. Leder, K. Diederichs, P. Sonderegger, and W. Welte, *The crystal structure of the ligand binding module of axonin-1/tag-1 suggests a zipper mechanism for neural cell adhesion,* CELL **101**, 425 (2000).

[23] M. Moertl, P. Sonderegger, K. Diederichs, and W. Welte, *The crystal structure of the ligand-binding module of human tag-1 suggests a new mode of homophilic interaction,* PROTEIN SCIENCE **16**, 2174 (2007).

[24] S. Bouyain and D. J. Watkins, *The protein tyrosine phosphatases ptprz and ptprg bind to distinct members of the contactin family of neural recognition molecules,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **107**, 2443 (2010).

[25] S. Lamprianou, E. Chatzopoulou, J.-L. Thomas, S. Bouyain, and S. Harroch, *A complex between contactin-1 and the protein tyrosine phosphatase ptprz controls the development of oligodendrocyte precursor cells,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **108**, 17498 (2011).

[26] R. M. Nikolaienko, M. Hammel, V. Dubreuil, R. Zalmai, D. R. Hall, N. Mehzabeen, S. J. Karuppan, S. Harroch, S. L. Stella, and S. Bouyain, *Structural basis for interactions between contactin family members and protein-tyrosine phosphatase receptor type g in neural tissues,* JOURNAL OF BIOLOGICAL CHEMISTRY **291**, 21335 (2016).

[27] S. J. Karuppan, A. Vogt, Z. Fischer, A. Ladutska, J. Swiastyn, H. F. McGraw, and S. Bouyain, *Members of the vertebrate contactin and amyloid precursor protein families interact through a conserved interface,* JOURNAL OF BIOLOGICAL CHEMISTRY **298** (2022), 10.1016/j.jbc.2021.101541.

[28] M. Djannatian, S. Timmler, M. Arends, M. Luckner, M.-T. Weil, I. Alexopoulos, N. Snaidero, B. Schmid, T. Misgeld, W. Moebius, M. Schifferer, E. Peles, and M. Simons, *Two adhesive systems cooperatively regulate axon ensheathment and myelin growth in the cns,* NATURE COMMUNICATIONS **10** (2019), 10.1038/s41467-019-12789-z.

[29] N. Elazar, A. Vainshtein, K. Rechav, M. Tsoory, Y. Eshed-Eisenbach, and E. Peles, *Coordinated internodal and paranodal adhesion controls accurate myelination by oligodendrocytes,* JOURNAL OF CELL BIOLOGY **218**, 2887 (2019).

[30] A. Klingseisen, A.-M. Ristoiu, L. Kegel, D. L. Sherman, M. Rubio-Brotons, R. G. Almeida, S. Koudelka, S. K. Benito-Kwiecinski, R. J. Poole, P. J. Brophy, and D. A. Lyons, *Oligodendrocyte neurofascin independently regulates both myelin targeting and sheath growth in the cns,* DEVELOPMENTAL CELL **51**, 730+ (2019).

[31] Y. He, G. J. Jensen, and P. J. Bjorkman, *Cryo-electron tomography of homophilic adhesion mediated by the neural cell adhesion molecule l1,* STRUCTURE **17**, 460 (2009).

[32] H. Liu, P. J. Focia, and X. He, *Homophilic adhesion mechanism of neurofascin, a member of the l1 family of neural cell adhesion molecules,* JOURNAL OF BIOLOGICAL CHEMISTRY **286**, 797 (2011).

[33] B. Hassel, F. Rathjen, and H. Volkmer, *Organization of the neurofascin gene and analysis of developmentally regulated alternative splicing,* JOURNAL OF BIOLOGICAL CHEMISTRY **272**, 28742 (1997).

[34] T. Pruss, E. Kranz, M. Niere, and H. Volkmer, *A regulated switch of chick neurofascin isoforms modulates ligand recognition and neurite extension,* MOLECULAR AND CELLULAR NEUROSCIENCE **31**, 354 (2006).

[35] M. Kriebel, J. Wuchter, S. Trinks, and H. Volkmer, *Neurofascin: A switch between neuronal plasticity and stability,* International Journal of Biochemistry and Cell Biology **44**, 694 (2012).

[36] B. Zonta, S. Tait, S. Melrose, H. Anderson, S. Harroch, J. Higginson, D. L. Sherman, and P. J. Brophy, *Glial and neuronal isoforms of neurofascin have distinct roles in the assembly of nodes of ranvier in the central nervous system,* JOURNAL OF CELL BIOLOGY **181**, 1169 (2008).

[37] A. M. Pillai, C. Thaxton, A. L. Pribisko, J.-G. Cheng, J. L. Dupree, and M. A. Bhat, *Spatiotemporal ablation of myelinating glia-specific neurofascin (nfasc^{nf155}) in mice reveals gradual loss of paranodal axoglial junctions and concomitant disorganization of axonal domains,* JOURNAL OF NEUROSCIENCE RESEARCH **87**, 1773 (2009).

[38] A. Nans, S. Einheber, J. L. Salzer, and D. L. Stokes, *Electron tomography of paranodal septate-like junctions and the associated axonal and glial cytoskeletons in the central nervous system,* JOURNAL OF NEUROSCIENCE RESEARCH **89**, 310 (2011).

[39] C. C. H. Cohen, M. A. Popovic, J. Klooster, M.-T. Weil, W. Moebius, K.-A. Nave, and M. H. P. Kole, *Saltatory conduction along myelinated axons involves a periaxonal nanocircuit,* CELL **180**, 311+ (2020).

[40] C. L. Cullen, R. E. Pepper, M. T. Clutterbuck, K. A. Pitman, V. Oorschot, L. Auderset, A. D. Tang, G. Ramm, B. Emery, J. Rodger, R. B. Jolivet, and K. M. Young, *Periaxonal and nodal plasticities modulate action potential conduction in the adult mouse brain,* CELL REPORTS **34** (2021), 10.1016/j.celrep.2020.108641.

[41] V. Lucic, T. Yang, G. Schweikert, F. Förster, and W. Baumeister, *Morphological characterization of molecular complexes present in the synaptic cleft,* STRUCTURE **13**, 423 (2005).

[42] B. Zuber, I. Nikonenko, P. Klauser, D. Muller, and J. Dubochet, *The mammalian central nervous synaptic cleft contains a high density of periodically organized complexes,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **102**, 19192 (2005).

[43] C.-L. Tao, Y.-T. Liu, R. Sun, B. Zhang, L. Qi, S. Shivakoti, C.-L. Tian, P. Zhang, P.-M. Lau, Z. H. Zhou, and G.-Q. Bi, *Differentiation and characterization of excitatory and inhibitory synapses by cryo-electron tomography and correlative microscopy,* JOURNAL OF NEUROSCIENCE **38**, 1493 (2018).

[44] L. Gollan, D. Salomon, J. Salzer, and E. Peles, *Caspr regulates the processing of contactin and inhibits its binding to neurofascin,* JOURNAL OF CELL BIOLOGY **163**, 1213 (2003).

[45] C. Bonnon, C. Bel, L. Goutebroze, B. Maigret, J.-A. Girault, and C. Faivre-Sarrailh, *Pgy repeats and n-glycans govern the trafficking of paranodin and its selective association with contactin and neurofascin-155,* MOLECULAR BIOLOGY OF THE CELL **18**, 229 (2007).

[46] M. Labasque, B. Hivert, G. Nogales-Gadea, L. Querol, I. Illa, and C. Faivre-Sarrailh, *Specific contactin n-glycans are implicated in neurofascin binding and autoimmune*

targeting in peripheral neuropathies, JOURNAL OF BIOLOGICAL CHEMISTRY **289**, 7907 (2014).

[47] C. Thaxton, A. M. Pillai, A. L. Pribisko, M. Labasque, J. L. Dupree, C. Faivre-Sarrailh, and M. A. Bhat, *In vivo deletion of immunoglobulin domains 5 and 6 in neurofascin (nfasc) reveals domain-specific requirements in myelinated axons,* JOURNAL OF NEUROSCIENCE **30**, 4868 (2010).

[48] E. Krissinel, *Stock-based detection of protein oligomeric states in jspisa,* NUCLEIC ACIDS RESEARCH **43**, W314 (2015).

[49] J. Rios, C. Melandez-Vasquez, S. Einheber, M. Lustig, M. Grumet, J. Hemperly, E. Peles, and J. Salzer, *Contactin-associated protein (caspr) and contactin form a complex that is targeted to the paranodal junctions during myelination,* JOURNAL OF NEUROSCIENCE **20**, 8354 (2000).

[50] M. Yamagata and J. R. Sanes, *Expanding the ig superfamily code for laminar specificity in retina: Expression and role of contactins,* JOURNAL OF NEUROSCIENCE **32**, 14402 (2012).

[51] Y.-R. Peng, N. M. Tran, A. Krishnaswamy, D. Kostadinov, E. M. Martersteck, and J. R. Sanes, *Satb1 regulates contactin 5 to pattern dendrites of a mammalian retinal ganglion cell,* NEURON **95**, 869+ (2017).

[52] C. Rader, B. Kunz, R. Lierheimer, R. Giger, P. Berger, P. Tittmann, H. Gross, and P. Sonderegger, *Implications for the domain arrangement of axonin-1 derived from the mapping of its ngcam binding site,* EMBO JOURNAL **15**, 2056 (1996).

[53] K. Mikulska-Ruminska, A. J. Kulik, C. Benadiba, I. Bahar, G. Dietler, and W. Nowak, *Nanomechanics of multidomain neuronal cell adhesion protein contactin revealed by single molecule afm and smd,* SCIENTIFIC REPORTS **7** (2017), 10.1038/s41598-017-09482-w.

[54] J. Haspel, D. Friedlander, N. Ivgy-May, S. Chickramane, C. Roonprapunt, S. Chen, M. Schachner, and M. Grumet, *Critical and optimal ig domains for promotion of neurite outgrowth by l1/ng-cam,* JOURNAL OF NEUROBIOLOGY **42**, 287 (2000).

[55] E. De Angelis, J. MacFarlane, J. Du, G. Yeo, R. Hicks, F. Rathjen, S. Kenwrick, and T. Brümmendorf, *Pathological missense mutations of neural cell adhesion molecule l1 affect hemophilic and heterophilic binding activities,* EMBO JOURNAL **18**, 4744 (1999).

[56] B. Kunz, R. Lierheimer, C. Rader, M. Spirig, U. Ziegler, and P. Sonderegger, *Axonin-1/tag-1 mediates cell-cell adhesion by a cis-assisted trans-interaction,* JOURNAL OF BIOLOGICAL CHEMISTRY **277**, 4551 (2002).

[57] O. Pavlou, K. Theodorakis, J. Falk, M. Kutsche, M. Schachner, C. Faivre-Sarrailh, and D. Karagogeos, *Analysis of interactions of the adhesion molecule tag-1 and its domains with other immunoglobulin superfamily members,* MOLECULAR AND CELLULAR NEUROSCIENCE **20**, 367 (2002).

[58] R. M. Gouveia, C. M. Gomes, M. Sousa, P. M. Alves, and J. Costa, *Kinetic analysis of l1 homophilic interaction - role of the first four immunoglobulin domains and implications on binding mechanism,* JOURNAL OF BIOLOGICAL CHEMISTRY **283**, 28038 (2008).

[59] I. Ispolatov, A. Yuryev, I. Mazo, and S. Maslov, *Binding properties and evolution of homodimers in protein-protein interaction networks,* NUCLEIC ACIDS RESEARCH **33**, 3629 (2005).

[60] S. Chamberland, H. H. Yang, M. M. Pan, S. W. Evans, S. Guan, M. Chavarha, Y. Yang, C. Salesse, H. Wu, J. C. Wu, T. R. Clandinin, K. Toth, M. Z. Lin, and F. St-Pierre, *Fast two-photon imaging of subcellular voltage dynamics in neuronal tissue with genetically encoded indicators,* eLife **6**, e25690 (2017).

[61] H. Tang, H. Chang, Y. Dong, L. Guo, X. Shi, Y. Wu, Y. Huang, and Y. He, *Architecture of cell-cell adhesion mediated by sidekicks,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **115**, 9246 (2018).

[62] C. FAIVRESARRAILH, G. GENNARINI, C. GORIDIS, and G. ROUGON, *F3/f11 cell-surface molecule expression in the developing mouse cerebellum is polarized at synaptic sites and within granule cells,* JOURNAL OF NEUROSCIENCE **12**, 257 (1992).

[63] S. Pourhoseini, D. Goswami-Sewell, and E. Zuniga-Sanchez, *Neurofascin is a novel component of rod photoreceptor synapses in the outer retina,* FRONTIERS IN NEURAL CIRCUITS **15** (2021), 10.3389/fncir.2021.635849.

[64] L. M. Chataigner, L. Thärichen, J. W. Beugelink, J. C. Granneman, N. J. Mokiem, J. Snijder, F. Förster, and B. J. Janssen, *Contactin 2 homophilic adhesion structure and conformational plasticity,* Structure **32**, 60 (2024).

[65] S. Tuvia, T. Garver, and V. Bennett, *The phosphorylation state of the figqy tyrosine of neurofascin determines ankyrin-binding activity and patterns of cell segregation,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **94**, 12957 (1997).

[66] Q. Ren and V. Bennett, *Palmitoylation of neurofascin at a site in the membrane-spanning domain highly conserved among the l1 family of cell adhesion molecules,* JOURNAL OF NEUROCHEMISTRY **70**, 1839 (1998).

[67] S. Pacharra, F.-G. Hanisch, and I. Breloy, *Neurofascin 186 is o-mannosylated within and outside of the mucin domain,* JOURNAL OF PROTEOME RESEARCH **11**, 3955 (2012).

[68] A. Zhang, A. Desmazieres, B. Zonta, S. Melrose, G. Campbell, D. Mahad, Q. Li, D. L. Sherman, R. Reynolds, and P. J. Brophy, *Neurofascin 140 is an embryonic neuronal neurofascin isoform that promotes the assembly of the node of ranvier,* JOURNAL OF NEUROSCIENCE **35**, 2246 (2015).

[69] A. Freal, D. Rai, R. P. Tas, X. Pan, E. A. Katrukha, D. van de Willige, R. Stucchi, A. Aher, C. Yang, A. F. M. Altelaar, K. Vocking, J. A. Post, M. Harterink, L. C. Kapitein, A. Akhmanova, and C. C. Hoogenraad, *Feedback-driven assembly of the axon initial segment,* NEURON **104**, 305+ (2019).

[70] P. Stathopoulos, H. Alexopoulos, and M. C. Dalakas, *Autoimmune antigenic targets at the node of ranvier in demyelinating disorders,* NATURE REVIEWS NEUROLOGY **11** (2015), 10.1038/nrneurol.2014.260.

[71] O. W. Howell, A. Palser, A. Polito, S. Melrose, B. Zonta, C. Scheiermann, A. J. Vora, P. J. Brophy, and R. Reynolds, *Disruption of neurofascin localization reveals early changes preceding demyelination and remyelination in multiple sclerosis,* BRAIN **129**, 3173 (2006).

[72] M. Chatterjee, I. van Steenoven, E. Huisman, L. Oosterveld, H. Berendse, W. M. van der Flier, M. Del Campo, A. W. Lemstra, W. D. J. van de Berg, and C. E. Teunissen, *Contactin-1 is reduced in cerebrospinal fluid of parkinson's disease patients and is present within lewy bodies,* BIOMOLECULES **10** (2020), 10.3390/biom10081177.

[73] R. A. Bamford, J. Widagdo, N. Takamura, M. Eve, V. Anggono, and A. Oguro-Ando, *The interaction between contactin and amyloid precursor protein and its role in alzheimer's disease,* NEUROSCIENCE **424**, 184 (2020).

[74] M. Chatterjee and C. Teunissen, *Contactins & alzheimer's disease: Synaptic proteins, contactins may contribute to the pathology of alzheimer's disease,* NEUROSCIENCE **424**, 182 (2020).

[75] Y. Gu, T. Li, A. Kapoor, P. Major, and D. Tang, *Contactin 1: An important and emerging oncogenic protein promoting cancer progression and metastasis,* GENES **11** (2020), 10.3390/genes11080874.

[76] Y. Liang, C. Ma, F. Li, G. Nie, and H. Zhang, *The role of contactin 1 in cancers: What we know so far,* FRONTIERS IN ONCOLOGY **10** (2020), 10.3389/fonc.2020.574208.

[77] G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, and G. Evans, *Dials: implementation and evaluation of a new integration package,* ACTA CRYSTALLOGRAPHICA SECTION D-STRUCTURAL BIOLOGY **74**, 85 (2018).

[78] P. R. Evans and G. N. Murshudov, *How good are my data and what is the resolution?* ACTA CRYSTALLOGRAPHICA SECTION D-BIOLOGICAL CRYSTALLOGRAPHY **69**, 1204 (2013).

[79] P. R. Evans, *An introduction to data reduction: space-group determination, scaling and intensity statistics,* ACTA CRYSTALLOGRAPHICA SECTION D-BIOLOGICAL CRYSTALLOGRAPHY **67**, 282 (2011).

[80] I. J. Tickle, C. Flensburg, P. Keller, W. Paciorek, A. Sharff, C. Vonrhein, and G. Bricogne, *Staraniso,* `http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi` (2018).

[81] 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 **40**, 658 (2007).

[82] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, and M. J. E. Sternberg, *The phyre2 web portal for protein modeling, prediction and analysis,* NATURE PROTOCOLS **10**, 845 (2015).

[83] P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, and P. D. Adams, *Towards automated crystallographic structure refinement with phenix.refine,* ACTA CRYSTALLOGRAPHICA SECTION D-STRUCTURAL BIOLOGY **68**, 352 (2012).

[84] P. Emsley and K. Cowtan, *Coot: Model-building tools for molecular graphics,* Acta Crystallographica Section D: Biological Crystallography **60**, 2126 (2004).

[85] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, and A. A. Vagin, *Refmac5 for the refinement of macromolecular crystal structures,* ACTA CRYSTALLOGRAPHICA SECTION D-STRUCTURAL BIOLOGY **67**, 355 (2011).

[86] C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, and D. C. Richardson, *Molprobity: More and better reference data for improved all-atom structure validation,* Protein Science **27**, 293 (2018).

[87] P. V. Afonine, N. W. Moriarty, M. Mustyakimov, O. V. Sobolev, T. C. Terwilliger, D. Turk, A. Urzhumtsev, and P. D. Adams, *Fem: feature-enhanced map,* ACTA CRYSTALLOGRAPHICA SECTION D-STRUCTURAL BIOLOGY **71**, 646 (2015).

[88] D. Hagemans, I. A. E. M. van Belzen, T. Moran Luengo, and S. G. D. Rudiger, *A script to highlight hydrophobicity and charge on protein surfaces,* FRONTIERS IN MOLECULAR BIOSCIENCES **2** (2015), 10.3389/fmolb.2015.00056.

[89] T. Dolinsky, J. Nielsen, J. McCammon, and N. Baker, *Pdb2pqr: an automated pipeline for the setup of poisson-boltzmann electrostatics calculations,* NUCLEIC ACIDS RESEARCH **32**, W665 (2004).

[90] N. Baker, D. Sept, S. Joseph, M. Holst, and J. McCammon, *Electrostatics of nanosystems: Application to microtubules and the ribosome,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **98**, 10037 (2001).

[91] H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, and N. Ben-Tal, *Consurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules,* NUCLEIC ACIDS RESEARCH **44**, W344 (2016).

[92] N. R. Coordinators, *Database resources of the National Center for Biotechnology Information,* Nucleic Acids Research **46**, D8 (2017), https://academic.oup.com/nar/article-pdf/46/D1/D8/23162308/gkx1095.pdf .

[93] D. S. Goodsell, L. Autin, and A. J. Olson, *Illustrate: Software for biomolecular illustration,* STRUCTURE **27**, 1716+ (2019).

[94] X. Robert and P. Gouet, *Deciphering key features in protein structures with the new endscript server,* NUCLEIC ACIDS RESEARCH **42**, W320 (2014).

[95] D. T. Pederick, J. H. Lui, E. C. Gingrich, C. Xu, M. J. Wagner, Y. Liu, Z. He, S. R. Quake, and L. Luo, *Reciprocal repulsions instruct the precise assembly of parallel hippocampal networks,* SCIENCE **372**, 1068+ (2021).

[96] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, *Fiji: An open-source platform for biological-image analysis,* Nature Methods **9**, 676 (2012).

[97] G. Van Rossum and F. L. Drake Jr, *Python reference manual* (Centrum voor Wiskunde en Informatica Amsterdam, 1995).

[98] M. L. Waskom, *seaborn: statistical data visualization,* Journal of Open Source Software **6**, 3021 (2021).

5

SETTING UP A WORKFLOW FOR STRUCTURAL STUDIES OF PROTEIN VOLTAGE-SENSITIVITY

The ability of protein molecules to sense transmembrane voltages is of central importance in, for instance, the opening of vital ion channels. Resolving voltage-dependent conformational changes has recently been added to the repertoire of cryo-EM structural biology. It works by embedding voltage-sensitive membrane proteins (VSMPs) in liposomal membranes with a chemically induced transmembrane potential during cryo-EM sample vitrification. So far, this was only applied to the best-characterised voltage-sensing mechanism, while other such mechanisms remain poorly understood despite their medical relevance – putative voltage-sensing in GPCRs, for instance. To enable such future structural investigation, the first steps are made in this work to also assemble this liposome-based VSMP voltage-clamp for cryo-EM – here referred to as CryoClamp. As a proof-of-principle the ASAP2s genetically-encode voltage indicator (GEVI) is purified for reconstitution into CryoClamp. ASAP2s directly reports changes in voltages as variable fluorescence. ASAP2s is isolated at high yield and purity using whole-cell solubilisation of human embryonic kidney cells. A generalisable workflow for membrane protein purification is thereby established. Stabilising detergents and buffer conditions are identified using thermostability assays. A combination of thermostability, gel electrophoresis, and mass photometry provide clues towards ASAP2s higher-order oligomeric states. ASAP2s is successfully inserted into liposomal membranes over which transmembrane voltages were previously measured. Once voltages are successfully induced in ASAP2s-containing vesicles, CryoClamp provides an in vitro approach for the cell-free and structure-based optimisation of GEVI functionality. Ultimately, CryoClamp enables the reconstitution of theoretically any purified VSMP for fundamental understanding of yet unresolved voltage-sensing mechanism, and for the structure-based development of treatments against disease-related voltage-sensor dysfunctionality.

Christos Gogou, Koen Wentinck, Phuong Nguyen, Nynke Meesters, Cátia P. Frias, Srividya Ganapathy, Daan Brinks, and Dimphna H. Meijer

5.1. INTRODUCTION

The universality of biophysical processes that depend on transmembrane ion gradients has fostered the idea that electrochemical potentials are life's primordial energy source[1–4]. Voltages in the range of, for instance, the 70-mV neuronal resting potential across 5-nanometre membranes generate an electrical field of roughly 14 million Volts per metre – 1000 times greater than required for thunder bolts.

Voltage-sensitive membrane proteins (VSMPs) are exposed to these electrical fields within biological membranes. They have evolved ways to convert this source of energy into mechanical work in the form of conformational changes. Vital cellular process such as neuronal action potential propagation[5–7], mitochondrial ATP-synthesis[8,9], and cell cycle regulation[10] depend on VSMPs in various ways. This functional versatility arises from the modularity of dedicated voltage-sensing domains (VSDs), being physically coupled to varying functional domains such as ion pores, proton pumps, or even enzymes. This VSD-to-functional domain coupling tunes VSMP activity to the ambient transmembrane potential[11].

The pathological implications of VSD dysregulation by mutations are varied and debilitating, mostly for ion channels[12]. For instance, detailed structural insights have elucidated how mutations in the VSD and its coupling mechanism in sodium channels ($Na_v1.7$) can result in severe pain syndrome[13]. Even so-called 'hotspots' have been mapped for various other sodium channels where mutations can be directly linked to severe channelopathies[14,15]. Such structural insights have also enabled computational identification of antiarrhythmic drugs targeting the VSD domain of cardiac potassium channels[16]. These examples showcase the utility of detailed structural data in understanding and countering disease.

Technological advantage is taken of VSD modularity in the design of genetically-encoded voltage indicators (GEVIs), which can be expressed at the plasma membrane of live neurons for the real-time visualisation of action potentials[17–19]. Most GEVIs visualise spiking action potentials as spikes in fluorescence, by coupling to a fluorescent protein (FP) domain. In some cases, the VSD itself is also fluorescent[17]. GEVIs are typically voltage-sensitive in their direct brightness or in their response and decay times upon changing voltage. After first design, optimisation of the fusion protein often takes place to improve features such as expression, brightness, signal-to-noise, and response times[20].

A limited number of molecular voltage-sensitive mechanisms is currently known, of which certain mechanisms have been studied more in-depth. The VSD - pervasive among voltage-gated ion channels - has been known the longest and is the best-studied mechanism[21,22]. It comprises four transmembrane α-helices (4TM domain) that are inserted perpendicular to the plane of the lipid bilayer[6,23]. These four helices are dubbed S1 through S4[22,23]. Helix S4 typically contains a motif of positively charged residues at every third position in the helix, which unidirectionally points the residues inward to the so-called gating transfer centre of the 4TM domain's core[22]. When the strength or direction of the transmembrane potential changes, the electrostatic force moves the entire S4 helix along this electrical field across the membrane[24–27]. In the case of ion channels, this movement is coupled to a pore domain – typically α-helices S5 and S6 – to modulate the membrane's ion permeability. In the case of the GEVI ASAP2s, a circularly permuted

GFP (cpGFP) is inserted between helices S3-S4 to convert S4 movement into altered fluorescence[28,29].

Different voltage-sensing mechanisms are now starting to be characterised among some of the G protein-coupled receptors (GPCRs). GPCRs have seven transmembrane helices (7TM) and this superfamily of integral membrane proteins spans almost 25% of the human genome[30] and contains ~60% of all drug targets, making them medically highly relevant[31,32]. The M1 and M2 muscarinic receptor (M1R and M2R) are voltage-dependent in their binding to acetylcholine by a yet incompletely specified mechanism, although three structural features appear to be involved[32–36]. Firstly, molecular dynamics (MD) simulations hint at a mechanism where a motile cation inhabits the M2R water-filled central pocket[32,33]. In response to changing transmembrane voltage, it translocates within the pocket and alters the receptor's function. Second, electrophysiology assays highlighted the importance of three inward-pointing tyrosines covering the orthosteric binding site[37]. The polar residues of this 'tyrosine lid' are postulated to alter conformation within a changing membrane potential. Third, by constructing MR1-MR2 chimeras, a charged KDKKE motif in TM3 was also found to contribute, but the importance of this motif is still under debate[38]. G protein association of mGluR1 and mlGluR3 metabotropic glutamate receptors GPCRs is also found to be voltage-dependent[39]. Here, low-resolution insights also ascribe roles to intracellular loops 2 and 3. Despite experimental evidence of GPCR voltage-sensitivity and the structural clues discussed here, no clear-cut domain is identifiable as is the case for the canonical 4TM VSD commonly found in ion channels.

Much effort has been invested in exploring the structural basis of protein voltage-sensitivity. MD simulations provide computational insight into the voltage-sensing dynamics of the M2R[32,33] and the 4TM-containing hyperpolarisation-activated cyclic nucleotide-gated (HCN) channel[26]. Site-specific mutations in combination with heavy metal cross-linking was used to physically lock this HCN VSD in its hyperpolarised conformation for electron cryo-microscopy (cryo-EM) reconstruction[25]. Similarly, a single amino acid was mutated in the 4TM VSD of a *Ciona intestinalis* voltage-sensing phosphatase (VSP) to shift its activation potential before structure determination by X-ray crystallography[27]. This resulted in a synthetic 'up' state for direct comparison with the wildtype in 'down' state. These approaches resulted all in important mechanistic insights. However, the experimental set-ups are far from physiological as mutations and toxic chemicals are required, while membrane leaflets and transmembrane potentials are absent. Direct atomic structure determination of membrane-embedded VSMPs under the influence of variable transmembrane potentials has only recently been realised[40,41]. In the works by Mandala and colleagues, the Eag and KCNQ1 voltage-sensitive potassium channels were reconstituted into the bi-layers of sub-100-nm liposomes with high potassium concentration inside and low K^+ outside. Both contain multiple copies of the 4TM VSDs. Transmembrane potentials were induced over the vesicles by specific permeation to potassium using valinomycin; the VSMP conformational state is thus controlled via the resulting K^+ concentration-dependent membrane voltage. Finally, the different conformational states were reconstructed using cryo-EM single-particle analysis (SPA) of these proteoliposomes within the thin vitreous ice layer. This approach revealed that the KCNQ1 channel is voltage-dependent in its binding to

the PIP_2 lipid, which is required to couple VSD movements to the opening of its pore domain.

In this chapter, we will set up and optimise a similar workflow to reconstitute VSMPs in liposomes for high resolution cryo-EM structure determination, a set-up here referred to as CryoClamp. The workflow for CryoClamp consists of high-purity and stable VSMP production, of VSMP reconstitution into liposomal membranes, and the induction of electrochemical potentials over these membranes for the use in future cryo-EM studies. Two GEVIs with diverging voltage-sensing mechanisms, ASAP2s and Ace2N-mNeon, are selected as proof-of-principle for the workflow. Upon liposomal reconstitution, both GEVIs are expected to directly report voltage-sensor movements through changes in fluorescence. Their fluorescent protein domains that point outwards of the membrane provides control over the orientation of liposomal insertion, and are expected to facilitate particle picking and -alignment during cryo-EM SPA. We demonstrate how affinity-tagged ASAP2s and Ace2N-mNeon are efficiently produced in mammalian suspension cells and how ASAP2s is successfully extracted from the membrane using whole-cell solubilisation. After affinity purification and size-exclusion chromatography, we performed thermostability assays to screen detergents and buffer conditions for further stabilisation of soluble ASAP2s. Gel electrophoretic bands at integer multiples of ASAP2s molecular weight, together with a two-step thermostability melting curve and single-molecule mass photometry provide preliminary evidence for ASAP2s oligomeric states. Purified ASAP2s is inserted into vesicles via detergent-mediated reconstitution with subsequent detergent removal via gel filtration. Finally, their voltage-sensing functionality in the vesicles is assessed using valinomycin-induced transmembrane potentials. Follow-up structural characterisation by cryo-EM *in vitro* would offer a cell-free approach for GEVI characterisation and optimisation for their applications *in vivo*. The workflow of CryoClamp may also enable future high-resolution structural investigation of virtually any purified VSMP, such as the elusive voltage-sensing mechanisms that was started to be uncovered in GPCRs.

5.2. RESULTS

5.2.1. ASAP2S AND ACE2N-MNEON ARE FUNCTIONALLY EXPRESSED AND LOCALISE TO THE MEMBRANE OF HEK-T AND HEK-E CELLS

In order to set up the workflow for cryoClamp, we selected the genetically-encoded voltage indicators (GEVIs) ASAP2s and Ace2N-mNeon. For both proteins, human embryonic kidney (HEK) cells were transfected with the corresponding plasmids. This mammalian production approach will ensure incorporation of any functional posttranslational modifications. Both expression plasmids were engineered to contain a C-terminal hexahistidine purification tag (his-tag). To assess correct folding and plasma membrane trafficking in human embryonic kidney cells, both GEVIs were first transiently transfected into adherent human embryonic kidney 293-T (HEK293-T) cells (Fig. 5.1). The GFP signal of both GEVIs colocalise with outer edges of the phalloidin stain, which demarcates the cell surface (Fig. 5.1A). Quantification of GFP signal distribution relative to the cell radius confirms that ASAP2s and Ace2N-mNeon display enriched localisation toward the cell periphery (Fig. 5.1B). Contrastingly, cells transfected with a cytoplasmic GFP control

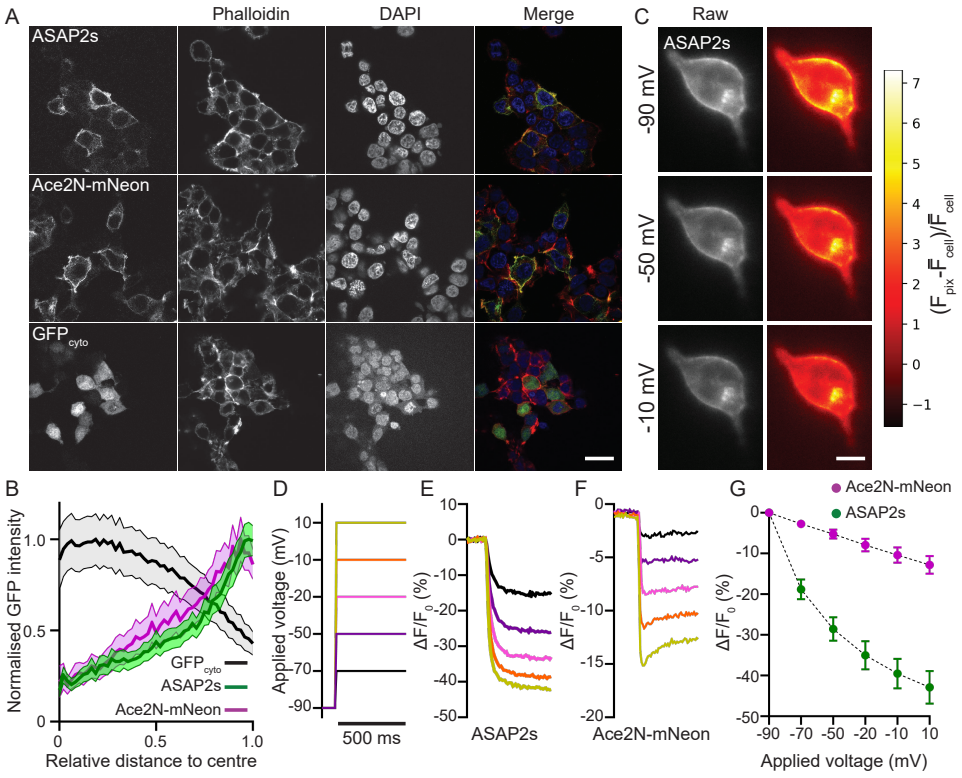


Figure 5.1: ASAP2s and Ace2N-mNeon localise functionally to the membrane of adherent HEK cells. **A)** Representative fluorescence confocal images of ASAP2s and Ace2N-mNeon expression in adherent HEK293-T cells. GFP signal shows the localisation of transiently transfected ASAP2s, Ace2N-mNeon, and cytosolic GFP (GFPcyto). Phalloidin stains filamentous actin in the cytoskeleton to visualise the entire cell surface, and DAPI visualises the cell nucleus. Scale bar is 25 µm. **B)** Quantification of relative GFP signal distribution between the cell's centre of mass (0) and the membrane at the cell's edge (1.0). Data represent mean ± SEM from normalised-to-maximum distributions from 71, 72, and 59 cells for ASAP2s, Ace2N-mNeon, and GFP, respectively; $N = 1$ experiment. **C)** Representative GFP confocal images of ASAP2s-expressing adherent HEK-E in whole-cell patch-clamp experiments at different holding voltages. Left: Raw fluorescence at indicated voltages. Right: Corrected pixel intensity (intensity at pixel minus the cell-wide average pixel intensity) relative to the cell-wide average pixel intensity. Scale bar is 10 µm. **D)** The applied 20-mV incremental step voltages in the patch-clamp electrophysiology. **E)** ASAP2s and Ace2N-mNeon GFP intensity response ($\Delta F/ F_0$) to the step voltages on the patched cells. ΔF is the measured fluorescence minus the fluorescence at -90 mV ($\Delta F = F - F_0$; $F_0 = F$ at -90 mV). The colours of the response traces correspond to the applied voltage steps in panel D. Response traces are averages of traces from the 10% most anti-correlated pixels (Supplementary Fig. 5.5C) from 4 cells for ASAP2s and Ace2N-mNeon. **F)** Fluorescent responses plotted as function of applied voltage for ASAP2s and Ace2N-mNeon. Data represent the fluorescence response 250 ms after applying the step voltage and are mean ± SEM of this metric, 4 cells for ASAP2s and Ace2N-mNeon each.

show decreasing signal from the geometric cell centre outward, as well as colocalisation with the nucleus. In conclusion, both GEVIs are fluorescent and trafficked to the plasma membrane - both signs of proper folding and functionality.

Next, the GEVI constructs were expressed in Epstein–Barr virus nuclear antigen I-expressing HEK293 cells (HEK-E) cells - human embryonic kidney cells in suspension that are optimised for increased protein production[42]. While in a fraction of the suspension cells bright spots were observed - possibly due to improper folding and trafficking

as the result of high levels of overexpressed proteins - in most cells, both GEVIs were localised to the membrane (Supplementary Fig. 5.5A). After allowing the HEK-E cells to adhere, whole-cell patch-clamp electrophysiology experiments were carried out to probe GEVI functionality and to quantify their fluorescent response in this cell type (Supplementary Fig. 5.5B). As with HEK293-T, the GEVIs are enriched at the periphery of HEK-E cells (Fig. 5.1C and Supplementary Fig. 5.5C). Starting from a -90 mV baseline voltage, the transmembrane potential was depolarised in 500-ms discrete steps of 20 mV (Fig. 5.1D). The change in fluorescence was monitored for the pixels with highest correlation to the changing voltage, which are the pixels at the cell edges (Supplementary Fig. 5.5C). For the his-tagged Ace2N-mNeon plasmid, no response could be observed – likely due to misfolding upon overexpression or by detrimental effects from the his-tag itself (data not shown). We therefore reverted to expressing Ace2N-mNeon from the original plasmid without his-tag to assess its voltage-sensitivity in HEK-E. Ace2N-mNeon's GFP response was then approximately linear to the voltage at roughly -10%/100 mV (Fig. 5.1F) compared to approximately -40%/100 mV for ASAP2s (Fig. 5.1E), which responds more strongly to changing transmembrane potentials in the more negative regime (Fig. 5.1G). These values correspond well with values reported in literature[29]. Moreover, these data show that Ace2N-mNeon and ASAP2s are both functionally expressed at the membrane of HEK-E cells.

5.2.2. ASAP2S BUT NOT ACE2N-MNEON IS STABLY PURIFIED USING WHOLE-CELL SOLUBILISATION AND AFFINITY PURIFICATION

In order to purify GEVIs, transfected HEK-E cells were solubilised with 1% detergent to free the integral membrane proteins before two-step purification (Fig. 5.2A). Before large-scale purification of both GEVIs, we probed at a smaller scale whether the proteins could be released from the membrane by using detergents DDM or digitonin, and whether they could be subsequently purified using the his-tag. ASAP2s appeared as a discrete band at the expected molecular weight (48.67 kDa) on western blot using both detergents (Supplementary Fig. 5.6A). The brighter band in the DDM-solubilised fraction of the small-scale affinity purification indicates that DDM may be more suitable to extract ASAP2s from the membrane. The low amount of protein in the flow-through indicates proper binding of the C-terminal his-tag to the Ni-NTA beads. The band in the digitonin-solubilised fraction was dimmer and no distinct differences between the input, flow-though, and eluted fractions could be seen (Supplementary Fig. 5.6A). Digitonin therefore appears less effective in solubilising ASAP2s, and the affinity tag is less accessible for binding during affinity purification. However, neither DDM nor digitonin were suitable to detect his-tagged Ace2N-mNeon on western blot, suggesting that it could not be produced in sufficient quantities or be stably solubilised for further purification. Because of this and of the absent voltage-sensitivity for the his-tagged Ace2N-mNeon construct, it was decided to only continue large-scale his-tag purification for ASAP2s using DDM for solubilisation.

For large-scale purification, ASAP2s-transfected HEK-E cells were solubilised with 1% DDM and incubated with Ni-NTA beads for affinity purification. After overnight binding to the nickel affinity beads, the protein was eluted with a step-wise incremental imidazole concentration (Supplementary Fig. 5.6B). Fractions eluted at 40% and

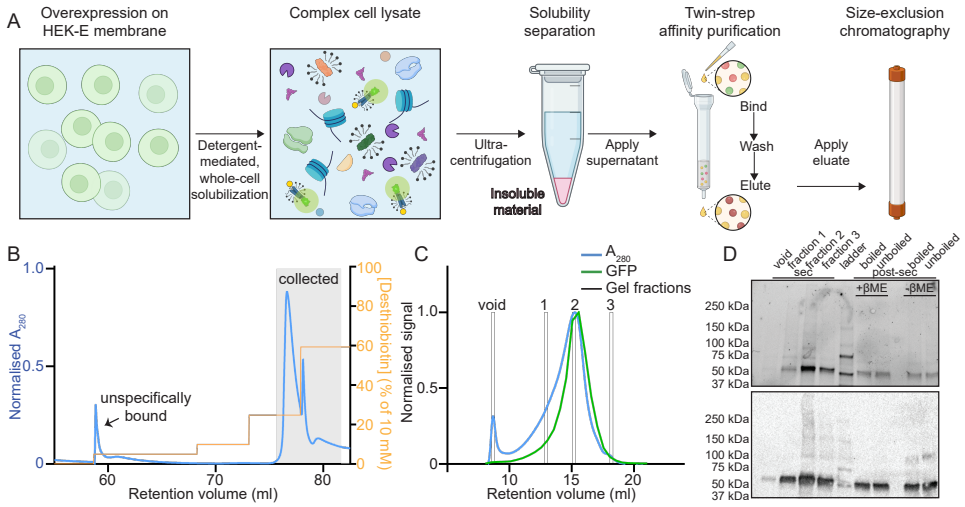


Figure 5.2: ASAP2s is purified from detergent-solubilised whole-cell lysate using twin-strep affinity purification. **A)** Schematic overview of the ASAP2s purification workflow. Transiently transfected cells are solubilised to free ASAP2s from the plasma membrane, resulting in a complex whole-cell lysate. Insoluble material is removed from the lysate by ultracentrifugation before binding to a strep-tactin affinity column. The column is washed with low concentrations of desthiobiotin before full elution of twin-strep-tagged ASAP2s. Size-exclusion chromatography (SEC) is performed as a subsequent round of purification. **B)** Twin-strep affinity purification trace of tagged ASAP2s. Two distinct A_{280} peaks are observed: A peak at 0.5 mM (5%) desthiobiotin that represents protein that is un-specifically bound being washed off the strep-tactin column; A larger peak appears upon elution at 2.5 mM (25%) desthiobiotin that represents ASAP2s elution. The grey area denotes the collected fractions for concentration and subsequent SEC. **C)** SEC trace of ASAP2s. A_{280} was measured during the run, while overlayed GFP was measured separately from the fractions collected during SEC. All fractions were collected for concentration and storage. Bars denote the fractions collected for gel electrophoresis and are labelled according to the numbering in panel D. **D)** Coomassie gel (top) and subsequent anti-twin-strep western blot (bottom) of fractions collected from SEC in panel C, and of entire pooled SEC trace (post-sec) in the presence/absence of reducing agent β-mercaptoethanol (βME) with/without boiling before gel electrophoresis.

100% (200 and 500 mM, respectively) of imidazole were collected for further purification through size-exclusion chromatography (SEC), while both GFP fluorescence and A_{280} were monitored (Supplementary Fig. 5.6C). The fluorescence trace exhibits a single peak – representing monomeric ASAP2s - with a shoulder that elutes earlier, which hints at the presence of higher-order oligomers or detergent micelles containing multiple ASAP2s molecules. Although the A_{280} peak elutes simultaneous with the fluorescence peak, A_{280} is also strongly elevated before and after the monomeric peak. Gel electrophoresis in combination with anti-his-tag western blot confirm the high relative abundance of contaminant proteins across SEC elution fractions (Supplementary Fig. 5.6D).

As the presence of impurities may negatively affect ASAP2s function in subsequent experiments, we substituted the C-terminal his-tag by a twin-strep-tag. The higher affinity of this purification tag for its bait could improve protein purity. A TEV cleavage site is additionally included between the ASAP2s-coding sequence and the tag to enable enzymatic tag removal.

Similar to the his-tagged constructs, we tested purification of the strep-tagged proteins at a small scale before moving on to larger scale protein production. Small scale pu-

rification demonstrated stable extraction from the membrane using five different detergents (Supplementary Fig. 5.6E). Upon overnight incubation with TEV-protease at varying temperatures, those conditions were negative for anti-twin-strep western blot, indicating effective removal of the tag (Supplementary Fig. 5.6F). We therefore moved on to large-scale purification. In contrast to his-tag purification, the A_{280} signal drops to baseline before elution of twin-strep-tagged ASAP2s from the strep-tactin column (Fig. 5.2B), indicating effective clearing of unspecific protein. ASAP2s elutes at 2.5 mM desthiobiotin (25% DTB). These fractions were pooled and concentrated for SEC (Fig. 5.2C). GFP and A_{280} SEC traces display higher mutual similarity and a single peak during gel filtration, further demonstrating a significant reduction of impurities (Fig. 5.2C). Different SEC fractions were collected for gel electrophoresis, followed by anti-strep western blot (Fig. 5.2D and Supplementary Fig. 5.6G). Although the bands in each well are brightest at the expected monomeric molecular weight of 50 kDa, the SEC fraction containing the peak values additionally reveals protein bands at roughly 100 and 150 kDa – double and triple the size of ASaP2s, possibly corresponding to dimers and trimers, respectively. In an attempt to preserve these oligomeric states on gel, purified ASAP2s was compared on gel with and without prior boiling at 95°C, but no relative increase in the presence of higher bands is observable (Fig. 5.2D and Supplementary Fig. 5.6G). On western blot, however, omitting reducing agent results in a more intense 100-kDa band. ASAP2s possesses only one accessible free cystine - C45. It is located on the N-terminal part of helix S1 that is predicted by AlphaFold to stick out of the membrane and bends back onto the surface of the membrane in a parallel fashion (Supplementary Fig. 5.6H). If this residue would constitute disulfide linkage of ASAP2s homodimers, however, no monomeric bands should be expected on western blot in the absence of reducing agent. AlphaFold also predicts dimers, trimers, and tetramers, but the monomers in these predicted complexes take on widely variable orientations with low-confidence contacts (data not shown). The orientations are also incompatible with a membrane. The experimental data shown in Fig. 5.2D do, however, support the existence of ASAP2s oligomers which may in part be covalently stabilised by disulfide bonds.

5.2.3. ASAP2S EXHIBITS THREE MELTING TRANSITIONS AND DISPLAYS OLIGOMERIC STATES DURING MASS PHOTOMETRY

A thermostability assay (TSA) was performed on purified twin-strep-tagged ASAP2s using nano differential scanning fluorimetry (nanoDSF). With nanoDSF, the change in exposure of aromatic tryptophan and tyrosine residues upon unfolding is measured as a shift in intrinsic fluorescence. This can be quantified by calculating the ratio in fluorescence at 330 and 350 nm (F_{330nm}/F_{350nm}). Melting transitions are recognised by a peak in this ratio's first derivative to temperature: $d(F_{330nm}/F_{350nm})/dT$ (Fig. 5.3A). Protein aggregation is simultaneously monitored through light scattering.

The stability of ASAP2s was assessed in the presence of different detergents, at a range of pH values, and at varying concentrations of sodium and potassium while maintaining a constant total cation concentration of 180 mM (Fig. 5.3B-D). At starting conditions (i.e. in DDM detergent, at pH 8.0, and in presence of 180 mM Na), ASAP2s displays three transitions in the F_{330nm}/F_{350nm} graph: two sharp transitions at approximately 40°C and 55°C ($T_{m,1}$ and $T_{m,2}$, respectively), and a broader transition near 80°C ($T_{m,3}$)

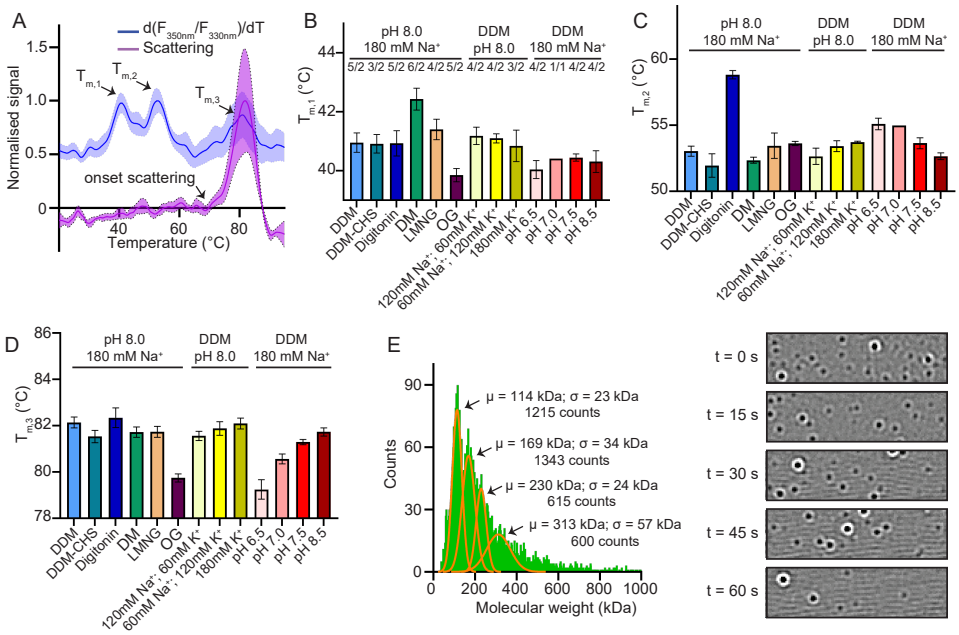


Figure 5.3: ASAP2s has displays three melting transitions and occupies higher-order oligomeric states. **A)** representative ASAP2s melting curve using nanoDSF. The curves are normalised mean ± SEM of *N* measurements from *n* experiments in control conditions with DDM detergent, pH 8.0, and 180 mM Na^{+}. B-D) Quantification of nanoDSF-derived ASAP2s melting temperatures in different buffer compositions. Data represent mean ± SEM from *N* technical repeats from *n* independent experiments, where N/n is presented above each condition in panel B. **E)** Mass photometry (MP) histogram displaying the molecular weight distribution of particles in the purified DDM-ASAP2s sample. Raw frames at various timepoints throughout data acquisition are shown on the right to show the non-aggregated sample state.

which is accompanied by a peak in scattering (Fig. 5.3A). Note that $T_{m,1}$ corresponds well with the his-tagged ASAP2s Tm calculated during on-column TSA (Supplementary Fig. 5.7A-B). The onset of the nanoDSF scattering peak at around 70°C, is also observed during on-column TSA as diminishing GFP signal (Supplementary Fig. 5.7B). This transition therefore likely represents protein aggregation upon GFP unfolding. The first two sharp transitions then probably represent structural changes within the transmembrane domain itself. When testing the different conditions, several changes in the three transition peaks are observed (Fig. 5.3B-D). For $T_{m,1}$, detergents DM and LMNG have a stabilising effect compared to DDM. Increasing potassium or decreasing sodium is slightly destabilising for this transition peak, and the acidity appears optimal at pH 7.0 and 7.5, although the peak shifts are minimal (Fig. 5.3B). For $T_{m,2}$, Digitonin appears most stabilising compared to DDM, and an opposite effect is seen in relative cation concentrations compared to $T_{m,1}$ (Fig. 5.3C). The pH is optimal at 6.5 and 7.0 for this transition peak. The stability of GFP ($T_{m,3}$) is most positively affected by higher relative K^{+} concentrations, and at more alkaline pH. From these findings together, it could be concluded that from the detergents tested, Digitonin has the most beneficial effect on ASAP2s stability. However, it shielded the C-terminal his-tag from binding to the stationary phase during small-scale affinity purification (Supplementary Fig. 5.6A), and will thus not be

considered as a suitable detergent for this workflow. While DM is found to be most stabilising to $T_{m,1}$ on-column and with nanoDSF, the detergent DDM performs well across parameters and TSA approaches - without any apparent destabilisation - and therefore remains a suitable candidate detergent for ASAP2s storage and experimentation.The different cation ratio's that were tested resulted in opposite effects for $T_{m,1}$ and $T_{m,2}$, and we therefore conclude that no optimised cation ratio has been found. Finally, when comparing the different pH conditions, it can be concluded that ASAP2s transmembrane domain is most stable at pH 7.0 – with the highest $T_{m,1}$ and $T_{m,2}$.

Where $T_{m,1}$ probably represents dissociation of the ASAP2s oligomers, the $T_{m,2}$ likely reflects unfolding of the monomeric transmembrane domain (Fig. 5.2A) Mass photometry (MP) was performed to further assess the oligomerisation state of ASAP2s (Fig. 5.3E). We performed on-stage dilution below DDM critical micelle concentration (CMC) to minimise background signal from empty detergent micelles. Individual peaks are discernible for ASAP2s particles of ~114 and 169 kDa, which largely correspond to dimers and trimers, respectively. A large tail is also observed comprising particles of higher molecular weight. In that tail, the higher-molecular weight peaks are less reliably resolved. However, peaks can still be recognised at multiples of ~60 kDa upon gaussian curve fitting. Together with the high-molecular weight bands observed on the gel and western blot images (Fig. 5.2D and Supplementary Fig. 5.6G), this indicates the presence of higher-order oligomers.

5

5.2.4. ASAP2S IS RECONSTITUTED INTO LIPOSOMES

To generate vesicles over which a transmembrane potential can be applied, we produced unilamellar lipid vesicles of 3:1 POPE:POPG lipid composition in different potassium (K^+) buffers. The vesicles were first prepared through buffer resuspension of a desiccated lipid film and subsequent freeze-thawing cycles. Extrusion through a membrane with 400 nm pore size resulted in vesicles of approximately 400 nm as observed with negative stain electron microscopy (Fig. 5.4A). These lipid-only vesicles were produced in a low-potassium chloride buffer, to encapsulate this buffer inside. Transfer of the liposomes into a high-K^+ buffer creates a potassium gradient across the membrane. To circumvent osmotic pressure, total ionic strength inside and outside are equalised using complementary sodium chloride concentrations (150 mM total cation concentration). Addition of the K^+-specific ionophore valinomycin allows for accumulation of positive charge through potassium influx down the chemical gradient, thereby creating an electrical field across the membrane (Fig. 5.4B). The strength of the transmembrane potential at equilibrium is calculated from the initial inner and outer potassium concentrations using the Nernst equation:

$$V_M = \frac{RT}{zF} \ln \frac{[K^+]_{out}}{[K^+]_{in}} \tag{5.1}$$

Where V_M is the transmembrane voltage, R is the universal gas constant at 8.31 $J \cdot K^{-1} \cdot mol^{-1}$, T is the temperature, z is the integer ion valence (+1 for K^+), and F is Faraday's constant at $9.65 \cdot 10^4$ $C \cdot mol^{-1}$, and $[K^+]_{in}$ and $[K^+]_{out}$ are the potassium concentrations inside and outside the liposomes, respectively. The transmembrane voltage for two different inner potassium concentrations is visualised in Supplementary Figure 5.8A. To

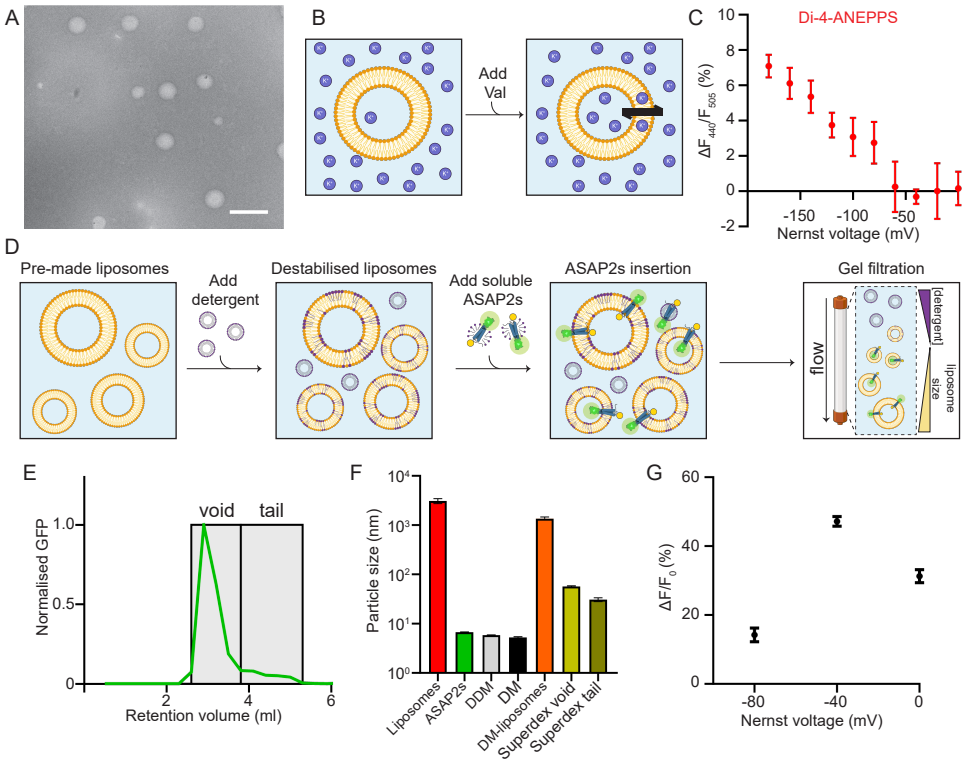


Figure 5.4: Nernst voltages are successfully induced, and ASAP2s is reconstituted into liposomes. **A)** Negative stain electron micrograph of 400-nm extruded unilamellar vesicles. Scale bar is 1 µm. **B)** Schematic representation of valinomycin-induced transmembrane voltages. Left: Vesicles contain a high potassium (K^+) concentration inside, and the extraliposomal buffer contains low potassium. Right: Upon addition of valinomycin (Val), K^+ will selectively flow inward down the gradient. This will result in a nett positive charge inside, thereby inducing a transmembrane electrical field. The Nernst voltage is the electrical potential at which K^+-outflux due to the positive charge inside equilibrates with the influx due to the chemical gradient. **C)** Ratiometric analysis of Di-4-ANEPPS fluorescence at varying valinomycin-induced Nernst potentials. Fluorescence values at 440 nm and 505 nm emission are measured and the change in their ratio correlates ($\Delta F_{440nm}/F_{505nm}$) with the Nerst potential. Data represent mean ± std from three independent experiments, in which the value was averaged for the first 9 time points (8 min) of each experiment. **D)** Schematic overview of the ASAP2s reconstitution workflow. Liposomes are pre-made through dried lipid film resuspension followed by freeze-thawing without extrusion. Detergent is added to destabilise the liposomes membrane through passive insertion of detergent molecules from ambient micelles. DDM-solubilised, pure solubilised ASAP2s is added for passive insertion into the destabilised vesicles. The detergent is removed by gel filtration and the proteoliposomes are collected. **E)** Sephadex G50 gel filtration GFP trace of applied ASAP2s in destabilised vesicles. Grey boxes with black outlines indicate the collected fractions. **F)** Dynamic light scattering (DLS) particle size measurements at different stages of the reconstitution workflow in panel D. Values are Number-derived mean ± SEM of 3 sizes in 1 experiment. **G)** Quantification of ASAP2s fluorescence response ($\Delta F/F_0$) as function of the theoretical Nernst potential on basis of the inner and outer potassium concentration. F_0 is the time-averaged intensity measured before addition of valinomycin, and ΔF is the measured fluorescence upon valinomycin addition minus F_0 ($F\text{-}F_0$). Full time series are presented in Supplementary Fig. 5.8H-I.

demonstrate that different K^+ concentrations inside and outside of liposomes can result in a set membrane potential, we used the lipophilic dye Di-4-ANEPPS. Di-4-ANEPPS localises to the membrane and reports changes in transmembrane potential through shifts in fluorescence excitation/emission spectra[43–46]. Indeed, changes in ratiometric fluo-

rescence emissions after 440 nm and 505 nm exitation ($\Delta F_{440nm}/F_{505nm}$) to a sensitivity of ~4%/100mV demonstrated successful induction of membrane voltage in the 400-nm liposomes that we produced (Fig. 5.4C).

To reconstitute purified ASAP2s into the liposomal membranes, the vesicles were first destabilised by adding 0.5% DM detergent. Vesicles were not extruded to a size of 400 nm. Instead, DDM-solubilised ASAP2s was added for a passive insertion of the proteins into the destabilised membrane. The mixture was subsequently applied to a 5 ml superdex G50 gel for size separation under gravity flow to remove the detergent (Fig. 5.4D). Large ASAP2s proteoliposomes do not interact with the gel and elute in the void volume, observed as a peak in ASAP2s GFP fluorescence (Fig. 5.4E). Previously, reconstitution of his-tagged ASAP2s on an 18-ml superdex column demonstrated the colocalisation of the GFP peak with that of lissamine rhodamine (LissRhod)-tagged lipids in the liposomes, as well as with absorbance at 280 nm (Supplementary Fig. 5.8B). This underpins the simultaneous elution of proteins and vesicles. Smaller proteoliposomes appear to interact more strongly with the gel, recognised as a tail at larger retention times (Fig. 5.4E and Supplementary Fig. 5.8B). Much smaller detergent molecules and micelles should migrate even slower through the gel, thereby clearing it from the proteoliposomal assembly. To minimise the probability that smaller vesicles would still contain traces of detergent that may cause leakiness, only the vesicles from the void volume fractions were collected (Fig. 5.4E).

Using dynamic light scattering (DLS), the average particle size at different reconstitution stages was assessed (Fig. 5.4F and Supplementary Fig. 5.8C). The initial vesicles without extrusion display a broadly distributed size at around 3 μm (Fig. 5.4F). Pure DDM-embedded ASAP2s adopts a size of 6.7 nm, while DDM as well as DM micelles have smaller sizes at 5.8 and 5.2 nm, respectively, indicating that empty micelles are smaller than ASAP2s in DDM micelles. Addition of 0.5% DM shrinks the liposomes to approximately 1300 nm, compared to 3 μm without detergent, indicating that the vesicles are indeed destabilised but not fully dissolved. The two collected fractions from the G50 gel filtration differ in size, with the vesicles in the void volume adopting an average size of ~60 nm, almost double the vesicle size in the tail, displaying a slightly larger size than 30 nm on average. The tail fraction is also more polydisperse, seen from the multiple shoulders in the autocorrelation curve (Supplementary Fig. 5.8C). This makes sense as the gel-interacting tail contains vesicles that were size-separated before being pooled back together. Despite the lack of a liposome extrusion step, the liposomes and destabilised versicles appear monodisperse (Supplementary Fig. 5.8C). Although DDM alone is quite polydisperse with three populations in the autocorrelation curve, the presence of ASAP2s results in relative monodispersity, probably owing to micelle formation of discrete size around a stable protein core rather than randomly aggregating micelles.

Theoretically, detergent removal by gel filtration - as performed here - may induce aggregation of the exposed ASAP2s transmembrane domains, which could possibly also end up in the void volume. To verify that the GFP peak in the void volume indeed represents membrane reconstitution instead of aggregation, we immobilised the vesicles on a 2D surface to asses protein-vesicle colocalisation (Supplementary Fig. 5.8D-E). For this approach, LissRhod-tagged lipids were added to the liposome composition. ASAP2s aggregates would logically be recognised as GFP-only foci, whereas proteoliposomes are

recognised by colocalising GFP and LissRhod. Indeed, the large majority of the ASAP2s foci display such colocalisation (Supplementary Fig. 5.8E-F). Specifically, around 80% of the ASAP2s particles represent proteoliposomes, and the remainder 20% are probably protein aggregates. Conversely, ~85% of the vesicles do not contain ASAP2s in their membrane, whereas 15% are proteopliposomes – vesicles that contain ASAP2s.

Next, we assessed the functionality of the reconstituted ASAP2s by measuring its voltage-coupled fluorescence. A dilution series was measured to establish a proper proteoliposome concentration for the valinomycin experiments (Supplementary Fig. 5.8G). The log-log plot of the titration curve is still linear in the 100- to 200-fold diluted regime. A 20-fold dilution was therefore selected, as this would still allow for sensitivity at -90% fluorescence response to applied voltage. Proteoliposomes from the concentrated (1.33 mg/ml protein) void fraction (Fig. 5.4E) were diluted in buffer varying in lower K^+ concentration relative to the 300 mM inside the vesicles (Supplementary Fig. 5.8A). ASAP2s fluorescence does not exhibit a change in fluorescence that correlates with the valinomycin-induced Nernst voltage (Fig. 5.4G and Supplementary Fig. 5.8H-I). Directly upon addition of valinomycin, however, ASAP2s fluorescence increases by up to ~80%. Surprisingly, it also rises by nearly 60% with equal K^+ concentrations inside as outside the vesicles, which should correspond to no transmembrane voltage at all. We conclude from these data that reconstituted ASAP2s was not able to report on the changes in membrane potential as set up in our system.

5.3. DISCUSSION

Voltage-sensitive membrane proteins (VSMP) modulate their function along a changing transmembrane potential. Over the course of time, this feature has been integrated into various vital cellular processes such as action potential propagation[6,7], cell cycle regulation, and regulation of development and regeneration[47]. It is known that proteins are able to undergo conformational changes in specialised domains as a response to this electrical transmembrane potential. To fully explore the molecular mechanisms of VSMPs, it is therefore crucial to incorporate controllable transmembrane voltage into the repertoire of structural biology approaches. This has recently been demonstrated by a detailed cryo-EM analysis of the 4TM voltage-sensor domain in the presence of changing transmembrane voltages[40,41] – an approach that is here referred to as CryoClamp. The next step is to investigate the VSMPs that possess other, less well-understood voltage-sensing mechanisms, such as those starting to be identified in GPCRs. In this chapter, our goal was to set up CryoClamp in our own lab to enable such further investigation. To this end, we successfully achieve VSMP purification and stabilisation, as well as reconstitution into liposomal membranes over which voltages were induced in the absence of protein. As proof-of-principle, we set up the workflow for a known 4TM VSMP with direct voltage-sensitive fluorescence read-out - the ASAP2s genetically-encoded voltage indicator (GEVI). Unfortunately, we were not able to demonstrate the functionality of our CryoClamp approach. This requires further optimisation of the established protocols.

While ASAP2s has traditionally been regarded as a monomeric protein, our SEC data, mass photometry, and nanoDSF thermostability data together suggest that ASAP2s may exist as homodimers, or even oligomers of higher order. This may have functional implications to ASAP2s usage, since the function of numerous proteins, including mem-

brane proteins, relies on the formation of oligomeric assemblies[48,49]. Recent advances in rational protein design harness this dependency by making so-called protein grafts – synthetic oligomerisation modules to enhance various functionalities[50]. In attempt to optimise ASAP2s sensitivity or response speed, its ability to form homomeric complexes may be further explored. Given the observation of multimers of varying stoichiometry, exploring their individual properties may be fruitful. Other GEVIs, like opsin-based ones with typically much faster response times, are also expected to rely on functional oligomerisation, as is for instance the case in certain rhodopsin proton pumps[51].

Several future steps can be taken to improve our CryoClamp set up. First, to minimise the probability of liposome leakiness due to residual detergent, the sephadex gel filtration could be performed on an 18-ml column to improve separation of vesicles from detergent. An adsorption step can be added to remove trace detergent[40,41]. Second, to improve the strength of the Nernst voltage, the external buffer can be exchanged using ion exchange columns instead of the direct resuspension approach in this chapter. The latter results in residual potassium in the outer buffer as high as 15 mM (300 mM in reconstitution buffer diluted 20x), which has a strong effect on the achievable valinomycin-induced Nernst voltage (Supplementary Fig. 5.8A). Reverting to low-K^+ inside is additionally advisable, as control over stronger Nerst potentials via high-K^+ outside is then less limited. Third, a uniform orientation of ASAP2s in the proteoliposomes is crucial to avoid cancelling responses in fluorescence between two populations of opposing membrane insertion (Fig. 5.1G). This orientations was not assessed[52] in this work, although it is expected to be biased towards one with an extra-liposomal GFP given that the liposomes were not completely dissolved before ASAP2s insertion[53]. The twin-strep affinity can be used to enrich the sample in proteoliposomes with a single orientation: ASAP2s with GFP inside the lumen have an outward-facing twin-strep tag, which binds strep-tactin. The unbound fraction will be enriched in outward-facing GFP. A uniform population will not suffer from the opposite fluorescence responses to applied voltage that cancel each other out when measured in bulk.

After optimizing the applied voltage, the proteoliposomes are ready for electron cryomicroscopy (cryo-EM) analysis, to reconstruct ASAP2s in different voltage-dependent conformational states. The ~60 nm vesicles are compatible with the vitreous ice layer. Cryo-EM single-particle analysis (SPA), however, relies on good contrast between particle and background for particle picking and -alignment steps required for 3D reconstruction[40,41,54–58]. In the next chapter 6, we discuss recent developments to improve the cryo-EM signal-to-noise (SNR) for the reconstruction of challengingly small proteins (<100 kDa). Such approaches typically attach extra domains to a protein of interest to effectively increase the molecular weight and therefore the SNR. Attaining good contrast is particularly challenging in the context of membrane-embedded particles, as phospholipids possess high electron density compared to transmembrane domains. ASAP2s was selected, in part, on basis of its GFP domain protruding from the membrane, which should be more clearly discernible against the aqueous background. Especially the presence of multiple GFPs in a ASAP2s oligomer could facilitate SPA. Otherwise the GFP domains provide a convenient epitope for recognition by nanobody- and/or Fab-based size-augmentations, as many GFP-targeting nanobodies are available[59,60].

Our cell-free, vesicle-based approach to investigate GEVIs offers novel opportunities

as well. Unlike neuron-based studies, this method allows accurate control over ion compositions on either side of the membrane. Real-time monitoring of ion concentrations is still problematic in live cells, and requires additional indicators[61]. Exact control over buffer compositions in our vesicles thus allows for more accurate calibration of GEVI responses to different ion concentrations and the emergent transmembrane potential. Although it is often reported for various GEVIs that their expression does not disturb membrane organisation, it is unknown what fraction of the plasma membrane they occupy during overexpression in neurons, and in turn, how this could affect the membrane potentials that are being measured. Other than manipulation of membrane compositions, the GEVI occupation of liposomal membranes could also be tightly controlled to investigate this.

Furthermore, ultrafast spectroscopy (UFS) is becoming more prominent in the exploration of fast conformational dynamics in various light-sensitive opsins[62–64]. A large proportion of the GEVIs, including Ace2N-mNeon are opsin-based. UFS is typically applied to purified opsins in detergent micelles or nanodiscs[63,64]. Especially in the case of opsin-based GEVIs, the presence of intact liposome membranes and voltage will enable more native exploration and optimisation of their functionality throughout their sub-microsecond photocycles.

Ultimately, any VSMPs of fundamental or medical interest can be inserted in CryoClamp and this would enable characterisation of novel mechanisms or even contribute to development of treatments against malfunctioning voltage-sensing domains in neurological and cardiac disorders.

5.4. METHODS

PLASMIDS

Plasmids for mammalian expression of ASAP2s (SVG20) and Ace2N-mNeon (SVG17) were generously provided by the Daan Brinks lab (Department of Imaging Physics, Delft University of Technology). To enable affinity purification, the protein-coding sequences were PCR-amplified using primers incorporating N-terminal BamHI and C-terminal NotI restriction sites. The PCR products were cloned into the pUPE3913 vector (U-protein express) - with hexahistidine tag directly downstream the NotI site - using BamHI-NotI (New England Biolabs) restriction-ligation, resulting in constructs pCG005 for Ace2N-mNeon and pCG006 for ASAP2s. To generate the twin-strep-tagged ASAP2s plasmid, a gBlock gene fragment (Integrated DNA Technologies) was designed and cloned between unique NotI-XbaI sites, replacing the N-terminal his-tag with a twin-strep tag, yielding construct pNM2 for ASAP2s.

HEK293-T CONFOCAL FLUORESCENCE MICROSCOPY AND MEMBRANE LOCALISATION ANALYSIS

Human embryonic kidney 293-T (HEK293-T) adherent cells were cultured in DMEM/F12 medium (Gibco) supplemented with 0.2% fetal bovine serum (FBS) and 1% Penicillin/Streptomycin (PS, Gibco) on coverslips coated with 0.2% gelatin. Following 24 h of incubation period at 37°C and 5% CO_2, transfection mediated by polyethylenimine (PEI) was performed using 0.02 µg of protein-encoding plasmid (GFP, ASAP2s, or Ace2N-mNeon)

and 1.98 µg of dummy plasmid. After an additional 24 h incubation under the same conditions, the cells were fixed with 4% paraformaldehyde (PFA), stained with Alexa568 Phalloidin (1:1000), and mounted in DAPI mounting medium (Vectorshield). Imaging was conducted using a Nikon Eclipse TI inverted microscope equipped with a 100x/1.4 Apo TIRF oil immersion objective lens (Nikon) and an ANDOR iXON EMCCD camera with a numerical aperture of 1.49 and a pinhole diameter of 67 µm. DAPI was excited at 401 nm (93 HV), GFP/ASAP2s/Ace2N-mNeon at 471 nm (100 HV), and Alexa568 Phalloidin at 561 nm (32 HV). Z-stacks were acquired with a step size of 1 µm.

From each z-stack, one slice containing the most expressing cell volumes in the focal plane was selected. The Cellpose Python module was employed to segment cells based on the Phalloidin channel. After background subtraction, segmented cells with a mean intensity from the GFP channel above 50 were used for further analysis. The centre of mass (CoM) of each cell and the distance of each pixel within the cell to the CoM were calculated using the "ndimage" module from Scipy. For each cell, distances to the CoM were binned into 50 equal-width bins, and the average pixel intensity per bin was calculated. Subsequently, the average pixel intensity with respect to distance to the CoM was averaged over all cells for each bin.

HEK-E PATCH-CLAMP ELECTROPHYSIOLOGY

Patch-clamp experiments were performed in collaboration with the Daan Brinks lab as previously described. Epstein–Barr virus nuclear antigen I-expressing HEK293 cells (HEK-E; U-Protein Express) were cultured in Freestyle medium supplemented with 10% FBS and 1% Penicillin/Streptomycin (PS, Gibco), and transfected with 0.02 µg ASAP2s/Ace2N-mNeon plasmid and 2.48 µg dummy plasmid using PEI. After 4-6 h at 37°C and 5% CO_2, cells were treated with 5.5% Primatone and 50,000 cells were seeded onto fibronectin-coated coverslips to attach overnight at 37°C and 5% CO_2. A custom microscope set-up was used for patch-clamp recordings with a multiphoton objective (Olympus XLPLN25XWMP2, NA 1.05). Widefield fluorescence imaging was done using 488-530 nm excitation-emission. After manual sealing of the patch-clamp pipette, cells were voltage clamped on whole cell mode. Voltage step depolarisations were applied from holding potential at -90 mV to a range of -70 to +10 mV in steps of 20 mV. Time at each step was 500 ms.

Patch-clamp data analysis was done in Python with background subtraction and bleach correction for each time frame. Manual and Otsu thresholding were used to segment patch-clamped cells. Pearson product-moment correlation was calculated for each pixel, and the top 10% anticorrelating pixels were averaged. The his-tagged Ace2N-mNeon (pCG005) did not functionally express in the pUPE vector and therefore expression was done with the original CaMKII (FCK) plasmid containing a CMV promotor (SVG17).

SMALL-SCALE EXPRESSION OF HIS-TAGGED ACE2N-MNEON AND ASAP2S

His-tagged ASAP2s and Ace2N-mNeon plasmids were expressed in HEK-E suspension cells by electroporation. Cells, grown in FreeStyle293 expression medium with GlutaMAX supplemented with 10% FBS and 1% PS, were harvested by centrifugation, resuspended to $2{\cdot}10^6$ cells per 100 µL in buffer R, and electroporated with 1.36 µg plasmid us-

ing the Neon Transfection System (Thermo Fisher Scientific) with a single 25-ms pulse of 1100 V before transfer into 5 mL of warm Freestyle medium without FBS and PS. After 24 h at 37°C and 5% CO_2, cells were imaged using an EVOS M5000 microscope with 10x objective (0.25 NA; EVOS, Thermo Fisher Scientific) using the EVOS LED GFP cubes, and stored at -80°C after harvesting by centrifugation and freezing in liquid nitrogen.

LARGE-SCALE ASAP2S PURIFICATION

HEK-E cells were cultured in FreeStyle293 expression medium with GlutaMAX supplemented with 10% FBS and 1% PS, in a shaking incubator at 37°C and 5% CO_2. Cells were seeded at 0.25-0.32·10^6 cells/mL in 0.5 L Erlenmeyer flasks and transfected with 5 µg DNA encoding his-tagged or twin-strep-tagged ASAP2s per 250 mL using PEI, followed by treatment with 5.5% Primatone within 4-6 h. Cells were collected 24 h post-transfection by centrifugation, after which cells were washed with PBS, pelleted, and frozen for storage at -80°C.

For large-scale purification of his-tagged ASAP2s, frozen cell pellets from 50 mL HEK-E culture volume was thawed on ice. Pellets were resuspended in 7 mL solubilisation buffer (25 mM HEPES, 25 mM NaCl, 1% n-Dodecyl-β-D-maltoside (DDM, Anatrace), supplemented with 0.1 µL of benzonase, and EDTA-free protease inhibitor) and solubilised while rotating at 4°C overnight. Remaining insoluble material was removed through ultracentrifugation at 100,000 g for 1 h in the SW 60 Ti rotor (Beckman Coulter). Supernatant containing the solubilised membrane proteins was incubated for affinity purification with Ni Sepharose excel beads (Cytiva) overnight while rotating at 4°C. ASAP2s eluted with buffer A (25 mM HEPES, 500 mM NaCl, 0.03% DDM, filtered at 0.45 µm) plus 200 mM imidazole and peak absorption at 280 nm (A_{280}) fractions were concentrated using Amicon Ultra centrifugal filter (15 mL, 30 kDa cutoff). Subsequently, the sample was applied to size-exclusion chromatography (SEC) (Superose 6 Increase 10/300 GL, Cytiva), pre-equilibrated in SEC buffer (20 mM HEPES, 150 mM NaCl, 0.1% FC12 detergent, filtered at 0.45 µm). 15 µL of each fraction was taken for fluorescence measurements using the QuantStudio 5 Real-Time PCR System (ThermoFisher) at 470-520 nm excitation-emission. Fluorescence peak fractions were concentrated using Amicon Ultra centrifugal filter (4mL, 30 kDa cutoff). ASAP2s was aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C.

For large-scale purification of twin-strep-tagged ASAP2s 450 mL HEK-E cell culture volume was thawed on ice and resuspended in 50 mL of solubilisation buffer (25mM HEPES, 25mM NaCl, 1mM $MgCl^2$, 5% glycerol, 1% DDM, supplemented with 0.1 µL of benzonase, and EDTA-free protease inhibitor, pH 8.0). The cells are solubilised for 1 h at 4°C on a rotator. Insolubilised material was removed by ultracentrifugation at 31,300 g for 1 h at 4∘C in the 45 Ti rotor (Beckman Coulter). Supernatant containing the solubilised membrane proteins was applied to a 1-mL StrepTrap HP affinity chromatography column (Cytiva). ASAP2s was eluted at more than 2.5 mM desthiobiotin (IBA) was collected and concentrated using an Amicon Ultra centrifugal filter (15 mL, 50 kDa cutoff). Subsequently, the sample was applied to SEC, pre-equilibrated in SEC buffer (20 mM HEPES, 150 mM NaCl, 0.06% DDM, pH 8.0). Samples from each fraction were taking for GFP fluorescence measurement using the Clariostar Plus microplate reader (BMT Labtech) with excitation-emission at 485±8 - 520±12 nm at room temperature. All

fractions after the void volume were collected and concentrated using an Amicon Ultra centrifugal filter (15 mL, 50 kDa cutoff). ASAP2s was stored at -80°C.

GELS AND WESTERN BLOTTING

Samples were diluted 1.5-fold in sample buffer (30% (w/v) glycerol, 6% SDS, 0.03% bromophenol blue, 187.5 mM Tris, 0.06% beta-mercaptoethanol) and boiled at 98°C for 5 min. Mini-Protean TGX Precast Gels (Bio Rad) were loaded with 10-12 µL boiled sample, and run at 200 V for 30 min. After gel imaging (Gel Doc EZ imager, Bio-Rad), gels were transferred to a membrane using the Trans-Blot Turbo Transfer System (Bio-Rad).

For anti-his antibody-based western blotting, the membrane was blocked in 15 mL coffee creamer dissolved in TBS-T buffer (20 mM Tris, 150 mM NaCl, 0.1% Tween 20) for 1 h at room temperature or overnight at 4°C. Membrane was washed on shaker for 10 min in TBS-T buffer before incubating for 1h with primary antibodies (0.5 µL penta-his, 0.5 µL poly-his diluted in 1 mL TBS-T buffer) at room temperature. Three washes of 15 min in 10-15 mL TBS-T buffer were done before incubating with secondary antibodies (0.2 µL HPR anti-mouse diluted in 1 mL TBS-T buffer) at room temperature. Another three washes of 15 min in 10-15 mL TBS-T buffer were done before 5 min activation using two-component activation (Luminol Enhancer Solution, Thermo Scientific). Imaging of Chemiluminescence was done using the Gel Doc XR+ imaging system (Bio-Rad).

For anti-twin-step western blotting, The membrane was blocked with 3% BSA in PBS buffer for 1 h at room temperature or overnight at 4°C. After being washed 3 times with TBS-T buffer (10 mM Tris, 150 mM NaCl, 0.1% Tween 20) for 5 min with a rotator, the membrane was incubated with strep-tactin HRP conjugate 1:100, 10 µL in 10 mL (1:100,000 final) of TBS-T buffer for 1 h on a rotator. Incubated membrane was then washed 2 times with TBS-T buffer for 2.5 min and 2 times with PBS buffer for 2.5 min. Two-component HRP activation was then performed for 2 min before chemiluminescence imaging of the membrane with ChemiDoc imaging system (Bio-Rad).

NANODSF THERMAL STABILITY ASSAY

NanoDSF thermal stability assays were performed using a Prometheus NT.48 device (NanoTemper Technologies). Purified ASAP2s in DDM was diluted to a final concentration of 0.2 mg/mL in each sample. Samples with various detergents contained a final detergent concentration equivalent to three times the critical micelle concentration (CMC) of the respective detergent in water. The samples were subjected to a temperature gradient from 20°C to 95°C at a rate of 1°C per min. Fluorescence emission was measured at 330 nm and 350 nm, and the melting temperatures (Tm) were determined by the peak in the ratio of fluorescence emission at 350 nm to 330 nm. Raw scattering data was analysed to identify peaks in scattering.

ON-COLUMN THERMAL STABILITY ASSAY

Affinity-purified his-tagged ASAP2s samples of 50 µL (1 mg/mL) were heated at various temperatures for 10 min. After centrifugation at 16,000 g for 10 min at 4°C to remove aggregates, samples were applied to a SEC column (Superose 6 Increase 10/300 GL, GE Healthcare) connected to a SEC-MALS system (Wyatt), pre-equilibrated with the buffer (20 mM HEPES, 150 mM NaCl, 0.03% DDM, filtered at 0.1 µm). Fluorescence was mea-

sured with 470 nm excitation and 530 nm emission. For detergent screening, 5 µL ASAP2s was diluted 5x in 45 µL total volume of various detergents at 2-3x their respective CMC. After 2 h incubation at 4°C, samples were heated to 50°C for 10 min and subsequently centrifuged at 16,000 g for 10 min at 4°C. 40 µL supernatant was applied to SEC. Baseline subtraction on A_{280} and fluorescence spectra was done using linear curve fitting on 0-7 mL and 22-25 mL, and for the fluorescence curves from 0.5-5 mL and 22-25 mL. Right and left peak positions were based on the fluorescence peak at >10mL and at <10 mL retention volume of the run, respectively. Ratios were normalised to the max values and sigmoid curves were fit using GraphPad Prism10 (Dotmatics) statistical analysis. The integrated fluorescence intensity was calculated by summing all the fluorescence values along the curve at constant flow.

ASAP2S STRUCTURE DETERMINATION

The monomeric structure of ASAP2s was predicted using the AlphaFold server and visualised using PyMol (Schrödinger Molecular Graphics System, DeLano Scientific, San Carlos, CA).

MASS PHOTOMETRY

Mass photometry data were acquired using a Refeyn OneMP instrument. Purified ASAP2s in DDM was diluted to 130 µg/mL in SEC buffer containing 0.03% detergent, and then diluted 40x on the stage to significantly reduce the DDM concentration below CMC directly before the measurement. Data acquisition and analysis were performed using DiscoverMP software. The molecular weight calibration curve was determined from two measurements of NativeMark protein standard (ThermoFisher) 1:100 dilution, with a max mass error of 9.7 % (R^2=0.992).

LIPOSOME PREPARATION

Liposomes were prepared by mixing POPE and POPG lipid in chloroform solutions (Avanti) in a 3:1 molar ratio, followed by evaporation under a gentle argon flow and rehydration in liposome buffer (20 mM HEPES, 0.1 mM KCl, 149.9 mM NaCl, 10% glycerol, pH 8.0) to a final lipid concentration of 8 mg/mL. Unilamellar vesicles of monodisperse size were formed by 5 freeze-thaw cycles and extrusion by 11-19 passages through a 0.4 µm membrane. Liposomes were stored at -80°C after freezing in liquid nitrogen. For the reconstitution of twin-strep-tagged ASAP2s, vesicles were produced in liposome buffer containing 300 mM KCl inside instead of 0.1 mM, and no extrusion step was performed.

NEGATIVE STAIN ELECTRON MICROSCOPY

Liposome were diluted 6,250-fold in liposome buffer and 3 µL was applied to glow-discharged carbon-coated grids. Excess sample was blotted away, followed by 3 washing and blotting cycles with liposome buffer. The grid was stained with 3 µL 2% uranyl acetate, followed by blotting and drying, and was then imaged using a 120kV Jeol JEM-1400plus transmission electron microscope at 15,000x magnification.

ASAP2S RECONSTITUTION

Sephadex G50 powder (Sigma) were swelled overnight in liposome buffer (12 mL per gram sephadex powder) at 4°C and was packed into a tricorn column to a final bed volume of ~18 mL. For reconstitution of his-tagged ASAP2s, liposomes were solubilised by incubation with 1% n-Decyl-β-D-Maltopyranoside (DM, Anatrace) for 30 min at 4°C. Purified ASAP2s was added in 1:1000 protein-to-lipid molar ratio and incubated for another 2 h at 4°C. The sample was applied to the Sephadex G50 column at 4°C by gravity flow. Eluted fractions of 250-500 µL were manually collected and protein content was measured with A_{280} and fluorescence on Quantstudio. A_{280} and fluorescence values were corrected by subtraction of the first value before the void volume fraction and all negative values were set to zero.

For reconstitution of twin-strep-tagged ASAP2s, 170 µL liposomes were destabilised by incubation with 0.5 % DM for 45 min at 4°C in a final volume of 357 µL (300 mM KCl, 20 mM HEPES, 10% glycerol, 0.5% DM, pH 8.0). Swelled sephadex G50 was packed to a final bed volume of 5 mL. 50 µL of 1.33 mg/mL purified ASAP2s was added and incubated for another 2h at 4°C. The sample was applied to the Sephadex G50 column at 4°C by gravity flow. Eluted fractions of ~500 µL were manually collected and protein fluorescence was measured using the Clariostar plate reader. Fluorescence values were corrected by subtraction of the first value before the void volume fraction and all negative values were set to zero.

BULK LIPOSOME ELECTROPHYSIOLOGY

For the empty liposomes, Di-4-ANEPPS was diluted to 0.5 µM in buffers with varying KCl concentrations higher than 0.1 mM KCl inside (and 149.9 NaCl mM) – in all other buffers the concentration KCl plus NaCl also add up to 150 mM, while the rest of the buffer is identical to liposome buffer. Liposomes were added to a final lipid concentration of 18 µg/mL, and valinomycin was added to 22 µM. For each Nernst voltage condition, a control without valinomycin was simultaneously measured. Emission values at 650 nm were measured for excitation at 440 nm and 505 nm after addition of valinomycin using Clariostar Plus microplate reader (BMT Labtech). The difference in the ratio of these emission values between was calculated between the conditions with and without was calculated ($\Delta F_{440nm}/F_{505nm}$).

For Twin-strep-tagged ASAP2s proteoliposomes, concentrated (0.52 mg/mL protein) vesicles from the Sephadex gel filtration void fraction were diluted 1:20 in buffers of varying low-potassium buffer. Fluorescence was measured using BioSPX Synergy H1 plate reader (Agilent) every 15 s for 3 min. Then, valinomycin was added to a final concentration of 1.2 µM and measurements were continued every 15 s for another 30 min. $\Delta F/F_0$ values were calculated by dividing each time point by the average of all time points before addition of valinomycin in the respective condition.

DYNAMIC LIGHT SCATTERING

Dynamic light scattering measurements were performed using a Zetasizer Nano ZS (Malvern) in 173° backscattering mode. Three repeats were measured with automatically determined number of runs. Autocorrelation curves were reported raw, and numbers-distribution size values were analysed when numbers-volume-intensity size distri-

butions were acceptable in comparability among repeats.

PROTEOLIPOSOME IMMOBILISATION

Liposomes were produced with DOPE:DOPG:PE-LissRhod:DOPG-biotin in 749:249:1:1 molar ratios. Reconstitution of his-tagged ASAP2s in liposomes was performed as described above. Proteoliposomes were immobilised on a cover slip with chambers that were washed with MQ and incubated with 10 µL BSA:BSA-biotin (1 mg/mL) for 10 min. After two additional washes with MQ, surfaces were incubated with 10 µL neutravidin (1 mg/mL) for 10 min. Chambers were washed three times with liposomes buffer. 15 µL of the Sephadex void fraction was applied to the chambers in 1000x dilution. Confocal imaging was done on a Nikon Eclipse TI inverted microscope as described above, operated with a pinhole size of 56 µm. ASAP2s was excited at 488 nm (12 HV) and LissRhod at 561nm (7 HV).

For image analysis, particles were picked by thresholding (0.2 for ASAP2s and Otsu thresholding for liposomes) on the normalised intensity distributions. ASAP2s particles with an area of more than 2 pixels or mean intensity less than 50 pixels were discarded as background particles or aggregates, respectively. Colocalisation was defined as particles in separate channels with their centroid within 3 pixel radius. Colocalisation parameters diverging two times the standard deviation from the mean were seen as outliers and excluded from analysis.

ACKNOWLEDGEMENTS

NanoDSF thermal stability experiments were performed in collaboration with Alex Fish in the Anastassis Perrakis group at the Netherlands Cancer Institute (NKI), supported by the Instruct ERIC tech-sci grant (PID:30554). DHM is supported by an NWO grant (OCENW.XS4.059).

5.5. SUPPLEMENTARY INFORMATION

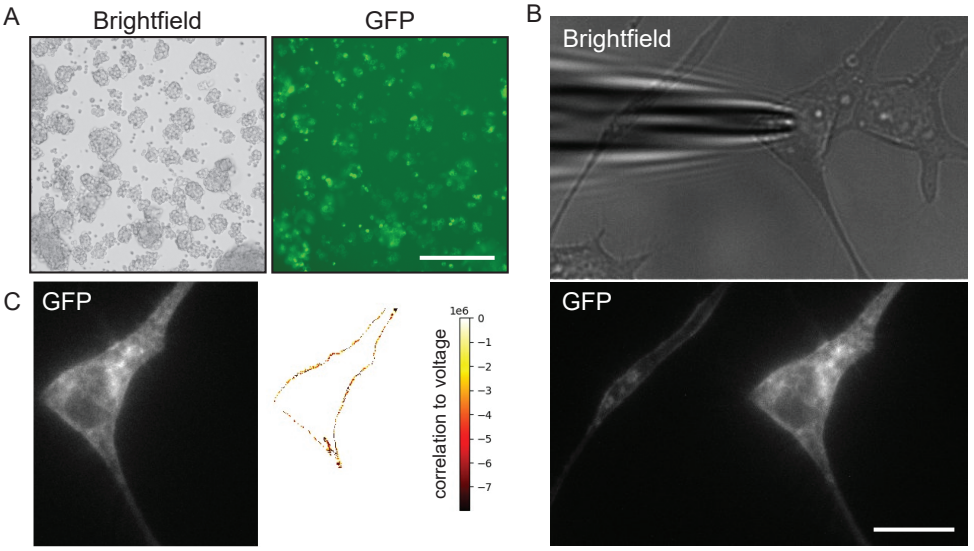


Figure 5.5: ASAP2s functionally localises to the membrane. **A)** Representative brightfield and GFP images of HEK-E cells 24 h after transient transfection. Scale bar is 300 µm. **B)** Representative brightfield and GFP images of a patch-clamped HEK-E cell expressing ASAP2s. The patch-clamp pipette is visible as a dark object attaching to the cells from the right in the brightfield image. Scale bar is 20 µm. **C)** Representative segmentation of top 10% pixels that anti-correlated most strongly with applied voltage.

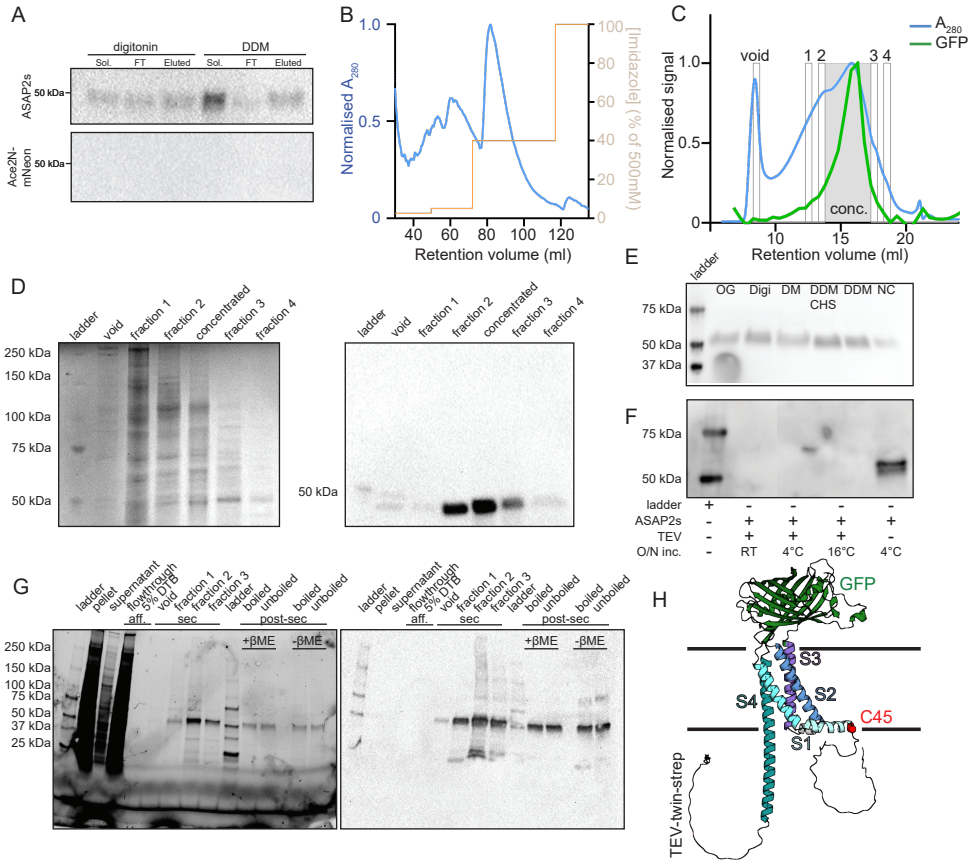


Figure 5.6: His-tag affinity purification results in low-purity ASAP2s. **A)** Anti-his western blot of small-scale purification of his-tagged ASAP2s and Ace2N-mNeon using DDM vs. digitonin detergents. Input solubilised (Sol) fractions are compared with protein in the unbound flowthrough (FT) and purified protine in the eluted fraction. **B)** Large-scale his-tag affinity purification trace of tagged ASAP2s. The strongest A_{280} peak is observed that represents ASAP2s during elution with 200 mM (40%) imidazole. The preceding peak at lower imidazole concentrations has not fallen back to baseline before elution of ASAP2s, which indicates un-specifically bound proteins still are eluted together with ASAP2s. **C)** Fluorescent SEC (fSEC) traces of eluent from affinity purification. The A_{280} represents all protein and the GFP trace visualises ASAP2s specifically. The grey area denotes the collected fractions for concentration and subsequent SEC, and bars denote the fractions collected for gel electrophoresis and are labelled according to the numbering in panel D. **D)** Coomassie gel (left) and subsequent anti-his western blot (right) of fractions collected from SEC in panel C. **E)** Anti-twin-strep western blot of small-scale expression of twin-strep-tagged ASAP2s using five different whole-cell solubilisation detergents vs. the negative control (NC). **F)** Western blot of twin-strep-ASAP2s after overnight incubation with TEV-protease a different temperatures. In all cases, the purification tag is digested and not observed on gel. The table below indicates the presence of ASAP2s, protease, and incubation temperature. **G)** Uncropped gel and western blot of Fig. 2D. **H)** AlphaFold3[65] predicted structure for ASAP2s-TEV-twin-strep. Vertical bars represent the approximate position of the lipid bi-layer surface. Cystine 45 (C45) is indicated as red spheres, helices S1 through S4 in the transmembrane region are indicated in shades of blue, circularly permuted GFP in green, and the C-terminal TEV-twin-strep-tag in black.

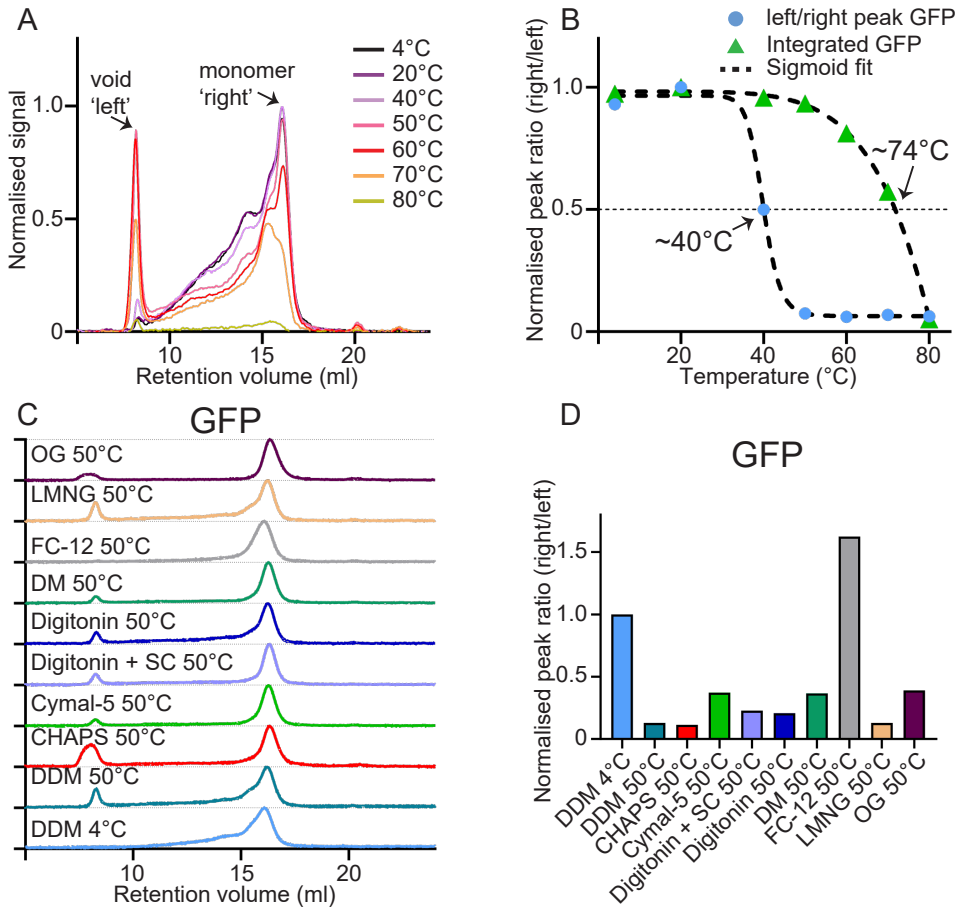


Figure 5.7: On-column ASAP2s thermal stability assay using fSEC. **A)** fSEC GFP traces of affinity-purified ASAP2s after 10 min incubation at varying temperatures. Left peak: Void volume fraction that contains proteins and aggregates that do not interact with the gel. Right peak: Stable (monomeric) ASAP2s. **B)** GFP-based melting curve of ASAP2s in DDM detergent. The normalised ratio of right/left peak values in panel A represents the relative abundance of stable (right peak) vs. aggregated (left peak, void) ASAP2s as function of incubation temperature. To visualise diminishing fluorescence by unfolding of the ASAP2s GFP, the integrated GFP represents the sum of all GFP values along the fSEC trace in panel A for a specific incubation temperature. The half-max of the fitted sigmoid to each curve is taken as the melting temperatures of ASAP2s in DDM. **C)** GFP fSEC traces of purified ASAP2s incubated for 2 h at 4°C in varying detergents before heating to 50°C for 10 min. Normalised traces are off-set by integer multiples of 1 for visualisation purposes. **D)** Right/left GFP peak ratios for the corresponding traces in panel C, normalised to the control ratio value for ASAP2s in DDM at 4°C.

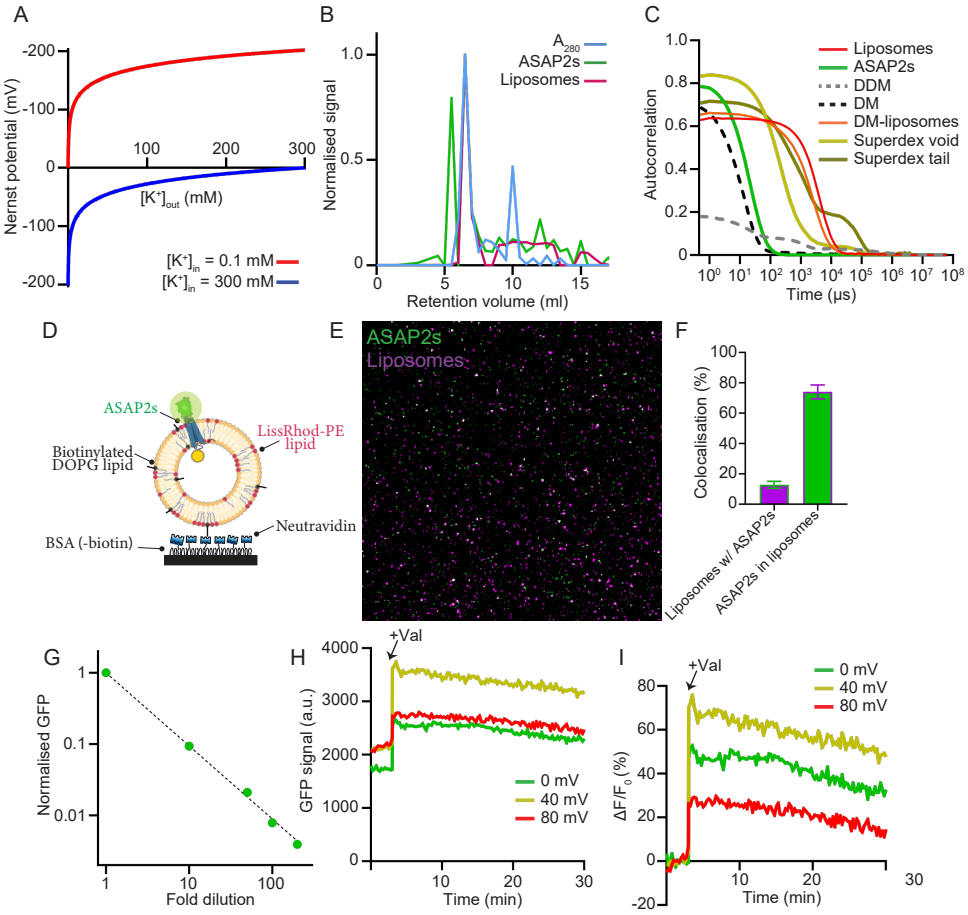


Figure 5.8: ASAP2s is successfully reconstituted in liposomes. **A)** Nernst transmembrane voltages at room temperature (293K) as a function of potassium concentration in the outer buffer ([K textsuperscript+]$_{out}$) for given intraliposomal. potassium concentrations ([K textsuperscript+]$_{in}$) of 0.1 (red) and 300 mM (blue). **B)** Sephadex G50 gel filtration traces of applied his-tagged ASAP2s in destabilised vesicles. LissRhod signal represents vesicles containing rhodamine-tagged lipids. A_{280} coincides with GFP signal, which both represent ASAP2s migration through the gel. **C)** DLS autocorrelation as function of time for samples from different stages of ASAP2s reconstitution. **D)** Schematic representation of immobilised his-tagged ASAP2s proteoliposomes on a 2D glass surface. The passivated glass surface is coated with neutravidin to which the vesicles are bound that contain biotinylated lipid. Lissamin rhodamine B sulfonyl-tagged (LissRhod) lipids are included in proteoliposome production to visualise the vesicles. **E)** Representative fluorescence confocal image of immobilised ASAP2s proteoliposomes. The imaging orientation is from the top relative to the schematic in panel D. Green foci are ASAP2s and magenta foci represent LissRhod-containing vesicles. **F)** Quantification of the co-localisation between the ASAP2s and LissRhod foci. The bar on the left denotes the fraction of all magenta foci that also colocalise with green signal. The right bar represents the fraction of all green foci that colocalise with magenta. **G)** Titration of ASAP2s fluorescence at different dilutions. **H)** Raw GFP time traces of bulk proteoliposomes electrophysiology experiments at different Nernst voltages. **I)** GFP response time traces derived from the traces in panel H. Quantification of ASAP2s fluorescence response ($\Delta F/F_0$) as function of the theoretical Nernst potential on basis of the inner and outer potassium concentration. F_0 is the time-averaged intensity measured before addition of valinomycin, and ΔF is the measured fluorescence upon valinomycin addition minus F_0 (F- F_0). The time at which valinomycin is added (+Val) is indicated by an arrow in panels H and I.

REFERENCES

[1] P. Mitchell, *Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism,* Nature (1961).

[2] M. Saraste, *Oxidative phosphorylation at the fin de siècle,* Science **283**, 1488 (1999).

[3] N. Lane, J. F. Allen, and W. Martin, *How did luca make a living? chemiosmosis in the origin of life,* BioEssays **32**, 271 (2010).

[4] N. Lane, *Proton gradients at the origin of life,* BioEssays **39** (2017), 10.1002/bies.201600217.

[5] B. Alberts, A. Johnson, J. Lewis, *et al.*, *Molecular Biology of the Cell*, 4th ed. (Garland Science, New York, 2002) Chap. Ion Channels and the Electrical Properties of Membranes.

[6] F. Bezanilla, *How membrane proteins sense voltage,* Nature Reviews Molecular Cell Biology **9**, 323 (2008).

[7] M. H. Grider, R. Jessu, and R. Kabir, *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2023).

[8] P. Dimroth, G. Kaim, and U. Matthey, *Crucial role of the membrane potential for atp synthesis by f_1f_o atp synthases,* Journal of Experimental Biology **203**, 51 (2000).

[9] A. P. Wescott, J. P. Kao, W. J. Lederer, and L. Boyman, *Voltage-energized calcium-sensitive atp production by mitochondria,* Nature Metabolism **1**, 975 (2019).

[10] C. D. McCaig, B. Song, and A. M. Rajnicek, *Electrical dimensions in cell science,* Journal of Cell Science **122**, 4267 (2009).

[11] J. Zhang, X. Chen, Y. Xue, N. Gamper, and X. Zhang, *Beyond voltage-gated ion channels: Voltage-operated membrane proteins and cellular processes,* Journal of Cellular Physiology **233**, 6377 (2018).

[12] F. Miceli, M. V. Soldovieri, P. Ambrosino, M. de Maria, L. Manocchio, A. Medoro, and M. Taglialatela, *Molecular pathophysiology and pharmacology of the voltage-sensing module of neuronal ion channels,* Frontiers in Cellular Neuroscience **9** (2015), 10.3389/fncel.2015.00259.

[13] G. Wisedchaisri, T. M. G. El-Din, N. Zheng, and W. A. Catterall, *Structural basis for severe pain caused by mutations in the s4-s5 linkers of voltage-gated sodium channel nav1.7,* Proceedings of the National Academy of Sciences of the United States of America **120** (2023), 10.1073/pnas.2219624120.

[14] X. Pan, Z. Li, X. Jin, Y. Zhao, G. Huang, X. Huang, Z. Shen, Y. Cao, M. Dong, J. Lei, and N. Yan, *Comparative structural analysis of human na v 1.1 and na v 1.5 reveals mutational hotspots for sodium channelopathies,* Proceedings of the National Academy of Sciences **118** (2021), 10.1073/pnas.2100066118.

[15] Z. Li, X. Jin, T. Wu, X. Zhao, W. Wang, J. Lei, X. Pan, and N. Yan, *Structure of human nav1.5 reveals the fast inactivation-related segments as a mutational hotspot for the long qt syndrome,* Proceedings of the National Academy of Sciences **118**, e2100069118 (2021).

[16] Y. Lin, S. Z. Grinter, Z. Lu, X. Xu, H. Z. Wang, H. Liang, P. Hou, J. Gao, C. Clausen, J. Shi, J. Cui, X. Zou, I. S. Cohen, G. Zhang, P. W. Kang, L. Zhao, K. M. White, Y. Liu, Z. Ma, and W. Zhao, *Modulating the voltage sensor of a cardiac potassium channel shows antiarrhythmic effects,* PNAS **118** (2021), 10.1073/pnas.2024215118/-/DCSupplemental.

[17] T. Knöpfel and C. Song, *Optical voltage imaging in neurons: moving from technology development to practical tool,* Nature Reviews Neuroscience **20**, 719 (2019).

[18] Y. Bando, M. Sakamoto, S. Kim, I. Ayzenshtat, and R. Yuste, *Comparative evaluation of genetically encoded voltage indicators,* Cell Reports **26**, 802 (2019).

[19] H. H. Yang and F. St-Pierre, *Genetically encoded voltage indicators: Opportunities and challenges,* Journal of Neuroscience **36**, 9977 (2016).

[20] M. Kannan, G. Vasan, and V. A. Pieribone, *Optimizing strategies for developing genetically encoded voltage indicators,* Frontiers in Cellular Neuroscience **13**, 1 (2019).

[21] F. H. Yu, V. Yarov-Yarovoy, G. A. Gutman, and W. A. Catterall, *Overview of molecular relationships in the voltage-gated ion channel superfamily,* Pharmacological Reviews **57**, 387 (2005).

[22] J. R. Groome and L. Bayless-Edwards, *Roles for countercharge in the voltage sensor domain of ion channels,* Frontiers in Pharmacology **11** (2020), 10.3389/fphar.2020.00160.

[23] K. J. Swartz, *Sensing voltage across lipid membranes,* Nature **456**, 891 (2008).

[24] T. M. G. El-Din, T. Scheuer, and W. A. Catterall, *Tracking s4 movement by gating pore currents in the bacterial sodium channel nachbac,* Journal of General Physiology **144**, 147 (2014).

[25] C. H. Lee and R. MacKinnon, *Voltage sensor movements during hyperpolarization in the hcn channel,* Cell **179**, 1582 (2019).

[26] M. A. Kasimova, D. Tewari, J. Cowgill, W. Carrasquel-Ursulaez, J. L. Lin, L. Delemotte, and B. Chanda, *Helix breaking transition in the s4 of hcn channel is critical for hyperpolarization-dependent gating,* eLife **8** (2019), 10.7554/eLife.53400.

[27] Q. Li, S. Wanderling, M. Paduch, D. Medovoy, A. Singharoy, R. Mcgreevy, C. A. Villalba-Galea, R. E. Hulse, B. Roux, K. Schulten, A. Kossiakoff, and E. Perozo, *Structural mechanism of voltage-dependent gating in an isolated voltage-sensing domain,* Nature Structural and Molecular Biology **21**, 244 (2014).

[28] F. St-Pierre, J. D. Marshall, Y. Yang, Y. Gong, M. J. Schnitzer, and M. Z. Lin, *High-fidelity optical reporting of neuronal electrical activity with an ultrafast fluorescent voltage sensor,* Nature Neuroscience **17**, 884 (2014).

[29] S. Chamberland, H. H. Yang, M. M. Pan, S. W. Evans, S. Guan, M. Chavarha, Y. Yang, C. Salesse, H. Wu, J. C. Wu, T. R. Clandinin, K. Toth, M. Z. Lin, and F. St-Pierre, *Fast two-photon imaging of subcellular voltage dynamics in neuronal tissue with genetically encoded indicators,* eLife **6**, e25690 (2017).

[30] L. Fagerberg, K. Jonasson, G. V. Heijne, M. Uhlén, and L. Berglund, *Prediction of the human membrane proteome,* Proteomics **10**, 1141 (2010).

[31] H. B. Schiöth and M. C. Lagerström, *Structural diversity of g proteincoupled receptors and significance for drug discovery,* Nature Reviews Drug Discovery **7**, 339 (2008).

[32] O. N. Vickery, J. P. Machtens, and U. Zachariae, *Membrane potentials regulating gpcrs: insights from experiments and molecular dynamics simulations,* Current Opinion in Pharmacology **30**, 44 (2016).

[33] O. N. Vickery, J. P. Machtens, G. Tamburrino, D. Seeliger, and U. Zachariae, *Structural mechanisms of voltage sensing in g protein-coupled receptors,* Structure **24**, 997 (2016).

[34] Y. Ben-Chaim, B. Chanda, N. Dascal, F. Bezanilla, I. Parnas, and H. Parnas, *Movement of 'gating charge' is coupled to ligand binding in a g-protein-coupled receptor,* Nature **444**, 106 (2006).

[35] Y. B. Chaim, S. Bochnik, I. Parnas, and H. Parnas, *Voltage affects the dissociation rate constant of the m2 muscarinic receptor,* PLoS ONE **8** (2013), 10.1371/journal.pone.0074354.

[36] Y. Ben-Chaim, C. Broide, and H. Parnas, *The coupling of the m2 muscarinic receptor to its g protein is voltage dependent,* PLoS ONE **14** (2019), 10.1371/journal.pone.0224367.

[37] O. Barchad-Avitzur, M. F. Priest, N. Dekel, F. Bezanilla, H. Parnas, and Y. Ben-Chaim, *A novel voltage sensor in the orthosteric binding site of the m2 muscarinic receptor,* Biophysical Journal **111**, 1396 (2016).

[38] M. Tauber and Y. Ben-Chaim, *Voltage sensors embedded in g protein-coupled receptors,* International Journal of Molecular Sciences **25** (2024), 10.3390/ijms25105295.

[39] L. Ohana, O. Barchad, I. Parnas, and H. Parnas, *The metabotropic glutamate g-protein-coupled receptors mglur3 and mglur1a are voltage-sensitive,* Journal of Biological Chemistry **281**, 24204 (2006).

[40] V. S. Mandala and R. Mackinnon, *Voltage-sensor movements in the eag kv channel under an applied electric field,* Proceedings of the National Academy of Sciences of the United States of America **46** (2022), 10.1073/pnas.

[41] V. S. Mandala and R. MacKinnon, *The membrane electric field regulates the pip2-binding site to gate the kcnq1 channel,* Proceedings of the National Academy of Sciences of the United States of America **120** (2023), 10.1073/pnas.2301985120.

[42] R. Tom, L. Bisson, and Y. Durocher, *Culture of HEK293-EBNA1 Cells for Production of Recombinant Proteins* (Expression Systems, 2007) Chap. 12.

[43] L. M. Loew, L. B. Cohenw, J. Dix, E. N. Fluhler:, V. Montanat, G. Salamaa, and W. Jian-Youngw, *Membrane biology a naphthyl analog of the aminostyryl pyridinium class of potentiometric membrane dyes shows consistent sensitivity in a variety of tissue, cell, and model membrane preparations,* J. Membrane Biol **130**, 1 (1992).

[44] A. P. Demchenko, G. Duportail, S. Oncul, A. S. Klymchenko, and Y. Mély, *Introduction to fluorescence probing of biological membranes,* Methods in molecular biology (Clifton, N.J.) **1232**, 19 (2015).

[45] V. Montana, D. L. Farkas, and L. M. Loew, *Dual-wavelength ratiometric fluorescence measurements of membrane potential*,* **28**, 4536 (1989).

[46] W. Y. Kao, C. E. Davis, Y. I. Kim, and J. M. Beach, *Fluorescence emission spectral shift measurements of membrane potential in single cells,* Biophysical Journal **81**, 1163 (2001).

[47] M. P. Harris, *Bioelectric signaling as a unique regulator of development and regeneration,* Development (Cambridge) **148** (2021), 10.1242/DEV.180794.

[48] D. T. Moore, B. W. Berger, and W. F. DeGrado, *Protein-protein interactions in the membrane: Sequence, structural, and biological motifs,* Structure **16**, 991 (2008).

[49] A. Fink, N. Sal-Man, D. Gerber, and Y. Shai, *Transmembrane domains interactions within the membrane milieu: Principles, advances and challenges,* Biochimica et Biophysica Acta - Biomembranes **1818**, 974 (2012).

[50] C. Liu and J. Luo, *Protein oligomer engineering: A new frontier for studying protein structure, function, and toxicity,* Angewandte Chemie - International Edition **62** (2023), 10.1002/anie.202216480.

[51] A. Iizuka, K. Kajimoto, T. Fujisawa, T. Tsukamoto, T. Aizawa, N. Kamo, K. H. Jung, M. Unno, M. Demura, and T. Kikukawa, *Functional importance of the oligomer formation of the cyanobacterial h+ pump gloeobacter rhodopsin,* Scientific Reports **9** (2019), 10.1038/s41598-019-47178-5.

[52] S. Veit, L. C. Paweletz, and T. G. Pomorski, *Determination of membrane protein orientation upon liposomal reconstitution down to the single vesicle level,* Biological Chemistry **404**, 647 (2023).

[53] N. Ritzmann, J. Thoma, S. Hirschi, D. Kalbermatter, D. Fotiadis, and D. J. Müller, *Fusion domains guide the oriented insertion of light-driven proton pumps into liposomes,* Biophysical Journal **113**, 1181 (2017).

[54] L. Wang and F. J. Sigworth, *Structure of the bk potassium channel in a lipid membrane from electron cryomicroscopy,* Nature **461**, 292 (2009).

[55] L. Tonggu and L. Wang, *Structure of the human bk ion channel in lipid environment,* Membranes **12** (2022), 10.3390/membranes12080758.

[56] X. Yao, X. Fan, and N. Yan, *Cryo-em analysis of a membrane protein embedded in the liposome,* Proceedings of the National Academy of Sciences **117**, 18497 (2020).

[57] Z. Melville, K. Kim, O. B. Clarke, and A. R. Marks, *High-resolution structure of the membrane-embedded skeletal muscle ryanodine receptor,* Structure **30**, 172 (2022).

[58] N. Sengupta, A. K. Mondal, S. Mishra, K. Chattopadhyay, and S. Dutta, *Single-particle cryo-em reveals conformational variability of the oligomeric vcc β-barrel pore in a lipid bilayer,* Journal of Cell Biology **220** (2021), 10.1083/JCB.202102035.

[59] R. Nygaard, J. Kim, and F. Mancia, *Cryo-electron microscopy analysis of small membrane proteins,* CURRENT OPINION IN STRUCTURAL BIOLOGY **64**, 26 (2020).

[60] K. Wentinck, C. Gogou, and D. H. Meijer, *Putting on molecular weight: Enabling cryo-em structure determination of sub-100-kda proteins,* Current Research in Structural Biology **4**, 332 (2022).

[61] G. Rong, E. H. Kim, K. E. Poskanzer, and H. A. Clark, *A method for estimating intracellular ion concentration using optical nanosensors and ratiometric imaging,* Scientific Reports **7** (2017), 10.1038/s41598-017-11162-8.

[62] I. H. van Stokkum, Y. Hontani, J. Vierock, B. S. Krause, P. Hegemann, and J. T. Kennis, *Reaction dynamics in the chrimson channelrhodopsin: Observation of product-state evolution and slow diffusive protein motions,* Journal of Physical Chemistry Letters **14**, 1485 (2023).

[63] A. Silapetere, S. Hwang, Y. Hontani, R. G. F. Lahore, J. Balke, F. V. Escobar, M. Tros, P. E. Konold, R. Matis, R. Croce, P. J. Walla, P. Hildebrandt, U. Alexiev, J. T. Kennis, H. Sun, T. Utesch, and P. Hegemann, *Quasar odyssey: the origin of fluorescence and its voltage sensitivity in microbial rhodopsins,* Nature Communications **13**, 1 (2022).

[64] S. Kaziannis, M. Broser, I. H. M. van Stokkum, J. Dostal, W. Busse, A. Munhoven, C. Bernardo, M. Kloz, P. Hegemann, and J. T. M. Kennis, *Multiple retinal isomerizations during the early phase of the bestrhodopsin photoreaction,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **121** (2024), 10.1073/pnas.2318996121.

[65] J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, S. W. Bodenstein, D. A. Evans, C.-C. Hung, M. O'Neill, D. Reiman, K. Tunyasuvunakool, Z. Wu, A. Žemgulytė, E. Arvaniti, C. Beattie, *et al.*, *Accurate structure prediction of biomolecular interactions with alphafold 3,* Nature **630**, 493 (2024).

6

REVIEW: ENABLING CRYO-EM RECONSTRUCTION OF LOW-SNR PROTEINS

Significant advances in the past decade have enabled high-resolution structure determination of typically large proteins by cryogenic electron microscopy single particle analysis. Despite improved sample preparation, next-generation imaging hardware, and advanced single particle analysis algorithms, sub-100 kDa proteins remain elusive for reconstruction due to low signal-to-noise and the lack of distinctive structural features. Multiple efforts have therefore been directed at the development of size-increase techniques for small proteins. Here we review the latest methods for increasing effective molecular weight of proteins <100 kDa through target protein binding or target protein fusion - specifically by using nanobody-based assemblies, fusion tags, and symmetric scaffolds. Finally, we summarise these state-of-the-art techniques into a decision-tree to facilitate the design of tailored future approaches, and thus for further exploration of ever-smaller proteins that make up the largest part of the human genome.

This chapter is published as Koen Wentinck*, *Christos Gogou**, and Dimphna H. Meijer, "Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins", *Current Research in Structural Biology)*[1]. *Contributed equally to this work.

6.1. INTRODUCTION

In the last decade, cryogenic electron microscopy (cryo-EM) single particle analysis (SPA) has become a prominent method in structural biology for protein structure determination[2,3]. Compared to other structural methods such as X-ray crystallography and NMR, cryoEM SPA has significant advantages: It is compatible with large, heterogenous macromolecular protein complexes in intact membrane environments, and it also typically requires much lower sample concentrations[4,5]. However, high-resolution structure determination of small proteins (<100 kDa) remains difficult using cryo-EM SPA. The current structures of human proteins on EMDB below 100 kDa and better than 4.0 Å resolution covers less than 3% of the total number of human-protein entries. In contrast, nearly 75% of the known human proteome members are smaller than 50 kDa[6] (Fig. 6.1A). High-resolution reconstruction of small proteins from cryo-electron micrographs is problematic for two main reasons (Fig. 6.1B): (i) Fewer electrons are scattered by particles with a lower molecular weight (MW), which reduces the signal-to-noise ratio (SNR) of small proteins compared to larger ones. Therefore, smaller proteins are more difficult to extract from noisy EM data. (ii) Small proteins intrinsically have fewer distinctive morphological features, which complicates particle alignment needed for protein reconstructions[7]. The inability to resolve small protein structures is currently a serious limitation for the general applicability of cryo-EM.

Despite great improvements in particle reconstruction software and electron microscopy hardware – such as direct detectors, energy filters, and phase plates – low SNR and lack of distinctive morphological features remain an impairment for faithful reconstruction of small protein targets[8]. An elegant solution for enabling structure determination of small proteins with cryo-EM would then be to increase the MW of the target protein (Fig. 6.1C). Here we review recent endeavors to overcome these size-dependent limitations by effectively increasing target protein size through tightly attached proteinaceous augmentations. This tight attachment is crucial as intrinsic flexibility reduces resolution, observed as blurring of molecule density during alignment and averaging in the SPA workflow. We focus on techniques that are adaptable to a wide range of small proteins. We discuss target protein binding and target protein fusion methods in which the resulting oligomers or chimeras, respectively, directly increase target protein size, while the added MW also serves as a fiducial marker for particle localisation and alignment. Ideally, such fiducials add asymmetric features to reduce ambiguity in the alignment of differently oriented particles. Additional examples will be addressed in which augmentations form higher-order symmetrical scaffolds, further facilitating high-resolution reconstructions of ever smaller proteins. Finally, we offer an overview of requirements, considerations, and suggestions to assist the design of future small-protein studies toward generalised applicability of size-increase techniques for SPA.

6.2. TARGET BINDING STRATEGIES

To append MW to a target protein without direct modification of the target protein itself, several approaches have been developed. Non-fusion size-increase for cryo-EM SPA has been realised using scaffolding DNA[9], thermostabilised antagonists[10], ubiquitin[11], antigen-binding fragments (Fabs)[8,12,13], nanobodies[14], and design ankyrin repeat

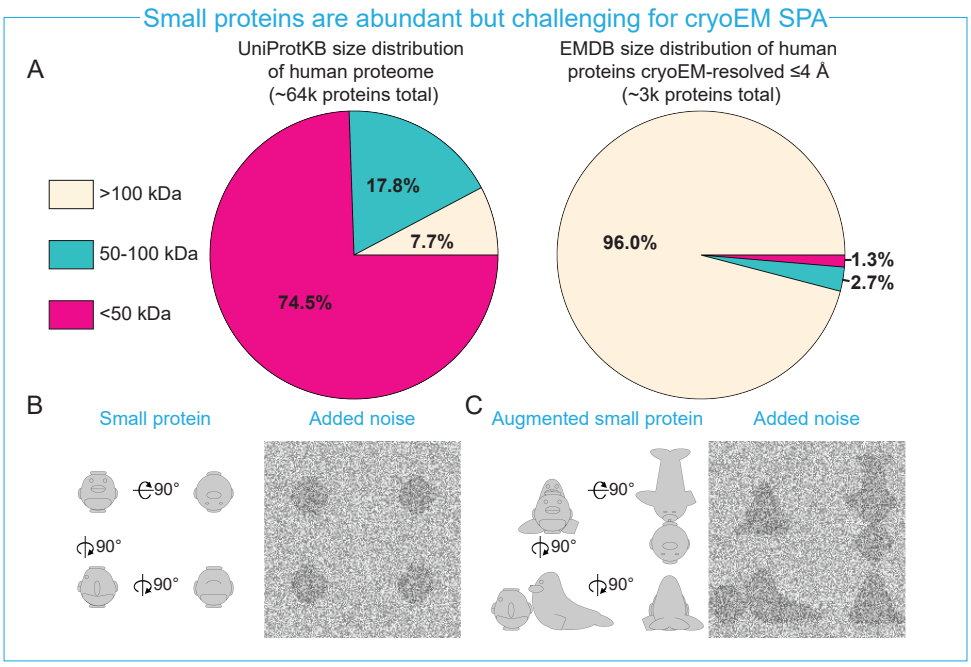


Figure 6.1: **A)** Size distribution of the UniProtBK human proteome compared to size distribution of EMDB entries of human proteins resolved below 4 Å resolution. Small proteins (<100 kDa) are under-represented in the SPA electron density database relative to their known abundance in the human proteome **B)** Low signal-to-noise ratio and lack of morphological features of small proteins (depicted as penguins) in electron micrographs complicates proper alignment of particles. **C)** Augmenting small proteins by appending extra molecular weight (depicted as seals) increases the signal-to-noise ratio and adds distinctive features to the assembly. The appended body allows different orientations of the assembly to be distinguishable for unambiguous alignment.

proteins (DARPins) [15–18]. The three latter options are generally applicable since they are compatible with in vitro display technologies [19] to obtain high affinity binders against a wide range of target proteins and will be discussed in more detail below.

As early as 2012, Fabs were proposed to serve as fiducial markers for alignment [13]. Fabs are ~50 kDa proteins and have a very characteristic "elbow" shape. In cryo-EM, Fabs have been proven to be very successful, especially for resolving small membrane proteins that often lack a characteristic feature when embedded in a detergent micelle or nanodisc. This topic has been excellently covered by Nygaard and colleagues [8]. Unfortunately, a stable and rigid Fab is not commercially available for every protein, and a single Fab might still add insufficient MW and features for structure determination.

A nanobody (Nb) is smaller (~12 kDa) but internally more rigid than a Fab, and typically locks a target protein in a specific conformation [14]. Because of their stabilising properties but limited addition of MW, Nbs are often deployed in complex with other proteins. For instance, Bloch and coworkers directed a Fab fragment against a conserved surface in Nbs, dubbing it 'NabFab' [20]. Subsequently, they increased the rigidity of the Fab by directing a Fab-targeting Nb. This complex formed a stable fiducial marker and aided in structure determination of two 50-kDa transmembrane proteins, a bacte-

rial MATE transporter (VcNorM) at 3.5 Å, and a bacterial divalent metal ion transporter (ScaDMT) at 3.8 Å resolution (Fig. 6.2A).

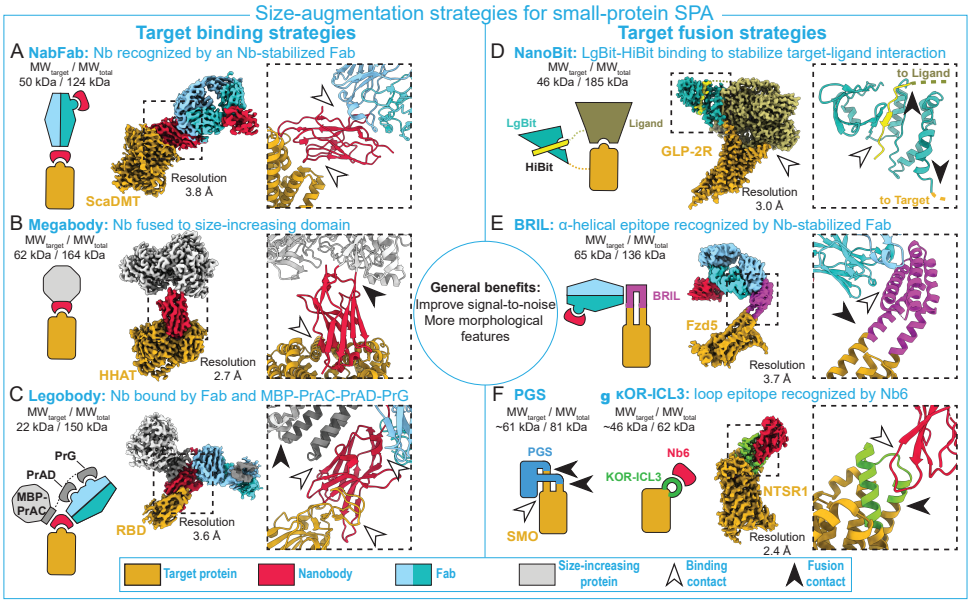


Figure 6.2: **A)** NabFab enabled 3.8 Å-resolution reconstruction of 50-kDa target protein ScaDMT (EMDB:13438, PDB:7PIJ). The target is rigidly bound by a Nb, which in turn is bound by a Nb-specific Fab. An additional Nb is bound to minimise the Fab internal flexibility. **B)** Example of a Megabody enabling a 2.7 Å-resolution reconstruction of 62-kDa target protein HHAT (EMDB: 13764, PDB:7Q1U). A Mb, consisting of a target-specific Nb fused with a large protein, is stably bound to the target protein. **C)** Use of a Legobody enabled 3.6 Å-resolution reconstruction of 22-kDa target protein SARS-CoV-2 RBD (EMDB:24729, PDB:7RXD). The target is rigidly bound by a Nb, which in turn is bound by a Nb-specific Fab. This assembly is stabilised by a MBP-PrAC fused to Fab-binding PrAD-PrG which binds Nb as well as Fab. **D)** Nanobit enabled high-resolution reconstruction of 46-kDa target protein GLP-2R (EMDB:30590, PDB:7D68). LgBit-fused GPCR is bound to its HiBit-fused natural G protein ligand. The LgBit-HiBit interaction stabilises the physiological interaction between GPCR and G protein. The originally unresolved LgBit and HiBit structures were predicted using AlphaFold[21]. NanoBit schematic not to scale. **E)** A so-called BRIL insert enabled 3.7 Å-resolution reconstruction of 65-kDa target protein Fzd5 (EMDB:21927, PDB:6WW2). BRIL is a two-point insertion between two anti-parallel α-helices of the target protein. BRIL serves as binding site for a Nb-stabilised Fab. **F)** Schematic of a PGS-insertion strategy that enabled the 3.7 Å-resolution reconstruction of ~61-kDa target protein SMO. PGS is fused to the target protein via a flexible two-point α-helical insertion, whereby the structural rigidity of the target-PGS fusion is ensured by a hydrophobic interaction between the target and PGS. **G)** Use of a κOR-ICL3 insertion with anti-ICL3 Nb6 enabled the 2.4 Å-resolution reconstruction of ~46 kDa NTSR1 (EMDB:26589, PDB:7UL2). κOR-ICL2 is a two-point insertion between two anti-parallel α-helices of the target protein, and is bound by a Nb (Nb6). Left of each panel: Schematics of assemblies. Middle: Colored electron density maps of full assemblies. Right: Zoom-ins of area indicated with black dashed line in middle panels. Colored dashed lines in panels c-d denote unresolved fusion connections. The molecular weight of the target protein (MW_{target}) and the total assembly (MW_{total}) are indicated above the schematics. Electron densities and structural models were produces using UCSF ChimeraX[22].

Similarly, a conserved β-turn within a nanobody was fused to the adhesin domain of *Helicobacter pylori* (HopQ) and a *Escherichia coli* K12 glucosidase (YgjK), to produce 50-100 kDa 'Megabodies' (Mb)[23]. Mbs are modular as the fused Nb can be exchanged. Moreover, Mbs can be directly matured with yeast and phage display technology. Recently, a Mb was used as fiducial to allow structure determination of the hedgehog acetyltransferase transmembrane enzyme (HHAT, 62 kDa) to 2.7 Å resolution (Fig. 6.2B)[24], and of the Na^+–taurocholate co-transporting polypeptide (NTCP, 38 kDa) to 3.7 and 3.3

Å resolution using two different Mbs[25]. Aside from this success with fiducial-assisted cryo-EM on a small protein, Mbs have also been effective in solving preferred orientation issues. In addition, Mbs were developed that target nanodisc-forming membrane scaffold proteins (MSP) for general membrane protein-targeting[23,25].

Even more complex assemblies were realised in which a Nb is stably bound by an anti-Nb Fab on one side, and by domain C of protein A of *Staphylococcus aureus* (PrAC) fused with a maltose binding protein (MBP) on the other side. This PrAC-MBP is extended with the Fab-binding protein AD (PrAD) and protein G (PrG) domains to stabilise the assembly, together referred to as 'Legobodies'. Legobodies facilitated structure determination of the 22-kDa SARS-CoV-2 receptor binding domain (RBD) (Fig. 6.2C) at 3.6 Å, and the small membrane KDEL receptor (23 kDa) at 3.2 Å[26]. Like NabFabs and Mbs, Legobodies are also very promising as a generally applicable method since they only require a Nb directed against the target protein.

6.3. TARGET FUSION STRATEGIES

To increase the MW of small proteins, several fusion strategies have also been developed. These approaches utilise, for instance, large multimerising proteins or small epitope motifs to which then high-affinity Fabs are available. For obtaining high resolution, the fusion with the target protein should be very rigid. A common tactic to achieve this rigidity is to extend C- or N-terminal α-helices. For cryo-EM SPA, however, a single connection through an extended α-helix is often still too flexible for high-resolution reconstruction[10,27]. Therefore, additional interactions between the fusion segment and target protein are typically required. We now discuss examples in which the target protein is covalently fused by two or three points of attachment.

Duan and co-authors have realised a fusion strategy that harnesses innate protein-protein interactions to render the size and morphology of the resulting complex compatible with SPA[28]. They stabilised a target-ligand complex by fusing a so-called LgBit fragment of a split luciferase to the target protein and a LgBit-complementary HiBit peptide fragment to the ligand, creating a 'NanoBit'. This two-point strategy has been particularly useful in - but probably not limited to – resolving G protein-coupled receptors (GPCRs) in complex with their G protein, such as the 3.0-Å reconstruction of the 46-kDa glucagon-like peptide 2 receptor (GLP-2R) (Fig. 6.2D)[28–30].

Another broadly applicable approach is to fuse epitopes to the target protein, thereby enabling subsequent binding of a Fab or Nb for fiducial-assisted SPA. Epitope tags as short as 14 amino acids have been tried on <50 kDa proteins with varying success[31,32]. An engineered thermostable variant of apocytochrome b562 (BRIL) has been most effective so far through its two-point insertion between two α-helices of target proteins. The BRIL element is recognised by an anti-BRIL Fab which can be used for fiducial-assisted SPA[27]. This strategy – with BRIL extending two α-helices by substituting the intracellular loop that initially connects them – facilitated structure determination of a 65 kDa frizzled-5 receptor (Fzd5)[33] (Fig. 6.2E), the GPCR adenosine A_{2A} receptor (45 kDa) at <4 Å resolution[10], and an inactive Epstein-Barr virus-induced GPCR 2 (EBI2) at 3.0 Å[34]. BRIL fusions, through an extension of single C- or N-terminal α-helices, were shown too flexible for fiducial-assisted SPA[10,27], but were effective in overcoming aggregation problems in one instance[35].

Similar to the two-point BRIL insertion strategy, the 87-kDa smoothened receptor (SMO) GPCR was resolved at 3.7 Å resolution by fusing a thermostable glycogen synthase domain from *Pyrococcus abyssi* (PGS) via a two-point α-helical extension [10] (Fig. 6.2F). Although the extended α-helices were found to be nicked, the required rigidity of the PGS domain was rescued due to an additional hydrophobic interaction between PGS and SMO.

Also, the 39-kDa AmpC beta-lactamase was utilised to replace an intracellular loop of the inactive β1 adrenoreceptor (45 kDa), and facilitated structure determination at 3.6 Å [36 36]. Notably, AmpC was selected as a rigid insertion candidate by searching the Protein Data Bank for structures with anti-parallel α-helices 11 Å apart, compatible with the distance between GPCR helices. A 4.2 Å-resolution structure of an inactive β2 adrenoreceptor GPCR was acquired through a three-point fusion with calcineurin, specifically by an insertion between transmembrane helices TM5-TM6, and by extension of TM7 [37].

Most recently, Robertson and colleagues developed a universal fiducial for inactive state GPCRs by substituting the conserved intracellular loop 3 (ICL3) for the κ-opioid receptor (κOR) ICL3 that is recognised by a specific Nb (Nb6), as well as by a Nb6-derived Mb (Mb6) [38]. This resulted in the high-resolution reconstruction of three sub-70-kDa inactive GPCRs, namely neurotensin 1 (NTSR1) (Fig. 6.2G), μ-opioid receptor, and the somatostatin receptor 2.

Fusion strategies have shown to be effective in determining the structure of small proteins if one or multiple α-helices can be extended, making it broadly applicable. However, it should be noted that even with increasingly accurate methods for protein structure prediction such as AlphaFold [21] and RoseTTAFold [39], and molecular dynamics simulations [38], screening over multiple variants is typically required for obtaining a rigid connection [27,33,34].

6.4. SYMMETRIC SCAFFOLDING STRATEGIES

So far, we have discussed instances where small protein target binding and target fusion led to a direct increase in MW of the targets. Here we discuss examples in which the appended units organise into symmetric scaffolds achieving at least 20-fold MW-increase, with regularly spaced target proteins docked at the scaffold's periphery (Fig. 6.3A).

Liu and co-authors fused an anti-GFP DARPin to one of the subunits of a symmetric protein cage [15,40]. DARPins are small (14 kDa) single-domain adapter proteins that consist of short α-helices separated by ankyrin repeat loops. The resulting ~588 kDa symmetric cage-like assembly was able to bind 12 copies of GFP (26 kDa) simultaneously (Fig. 6.3B). This scaffold was easily distinguished from the background and its 12-copy binding additionally overcame the cryo-EM SPA-specific problem of preferred orientation, allowing structure determination at 3.8 Å resolution for GFP. Designing DARPin-based scaffolds is not trivial, however. Flexibilities in the DARPin-to-target binding prevented Yao and colleagues from reaching sub-4 Å resolution for GFP with another scaffold design, even though their scaffolding complex itself could be resolved at 3 Å [17]. Subsequently, the DARPin motif was dimerised to increase stability of the complex [16]. In this case issues with preferred orientation limited the resolution to 4.8 Å for GFP. Dimerisation did improve stability of the complex, since a 4 Å-resolution structure for a 66 kDa human serum albumin (HSA) could be obtained.

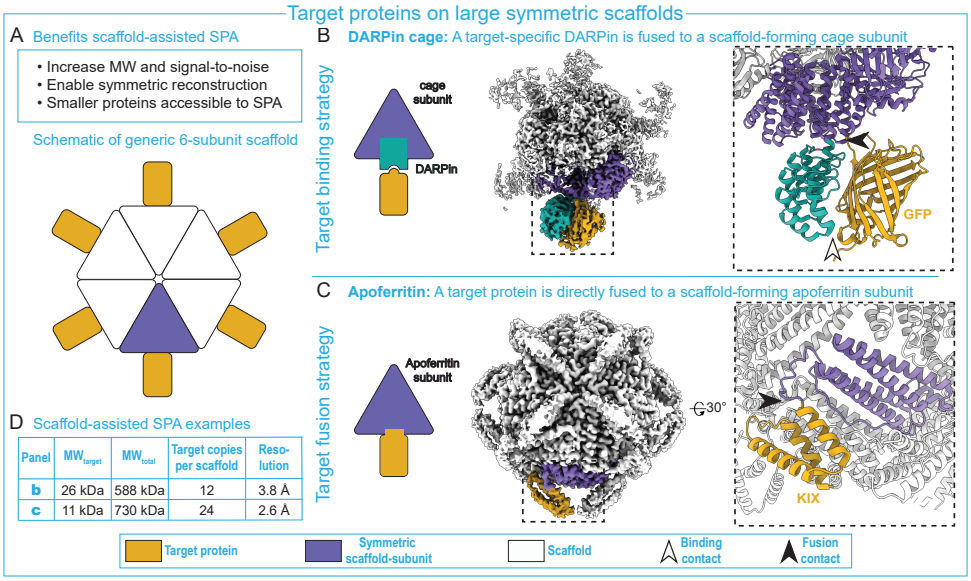


Figure 6.3: **A)** Benefits of using a symmetric scaffold for SPA as compared to non-scaffolding size augmentations. A schematic is shown of a generic symmetric scaffold with a single in-plane 6-fold symmetry axis and with target proteins docked at the periphery. **B)** Use of a target-specific DARPin fused with a cage subunits enabled the 3.8 Å-resolution reconstruction of 26-kDa GFP upon self-assembly of the DARPin-cage symmetric scaffold (EMDB:9373&9374, PDB:6NHV&6NHT). **C)** Direct fusion with apoferritin enabled the 2.6 Å-resolution reconstruction of 11-kDa KIX domain upon self-assembly of the apoferritin symmetric scaffold (EMDB:25791, PDB:7TB3. Left of panels b-c: Schematics of assemblies. Middle: Colored electron density maps of full assemblies. Right: Zoom-ins of area indicated with black dashed line in middle panels. **D)** Scaffold parameters corresponding to the examples given in panels B and C, including the molecular weight of the target protein (MW_{target}) and the total assembly (MW_{total}). Electron densities and structural models were produces using UCSF ChimeraX (Goddard et al., 2018).

A symmetric scaffolding approach relying on a single fusion was used to resolve the 11-kDa KIX domain of the CREB-binding protein (Fig. 6.3C)[41]. Zhang and colleagues fused the KIX domain, through a single-point α-helical extension, rigidly to apoferritin – a protein that self-assembles into a 24-subunit symmetric shell, allowing KIX structure determination at 3-4 Å resolution. It should be noted that two cysteine mutations and a deletion of a flexible segment of the KIX motif were needed to achieve the required stability for high-resolution structure determination.

Due to the additional benefits of increased size and symmetric features, on top of their modular design, symmetric scaffolds are currently being explored as a universal method for small-protein cryo-EM SPA[18,40].

6.5. SUMMARY

In summary, various techniques have become available in recent years to extend structure determination by cryo-EM SPA into the realm of small proteins (<100 kDa), which make up the vast majority of the human proteome. Size-dependent resolution limits can be overcome by clever size augmentation approaches, such as Nb-based assemblies, fused Fab epitopes, and symmetric scaffolding. The effectiveness of all above-

mentioned approaches mainly depends on the rigidity of the resulting complex[10,27,33], for which predictive computational approaches are recently emerging[21,38,39]. In Figure 6.4, we offer an overview of the discussed approaches, including considerations and suggestions for employing target protein binding versus target protein fusion strategies.

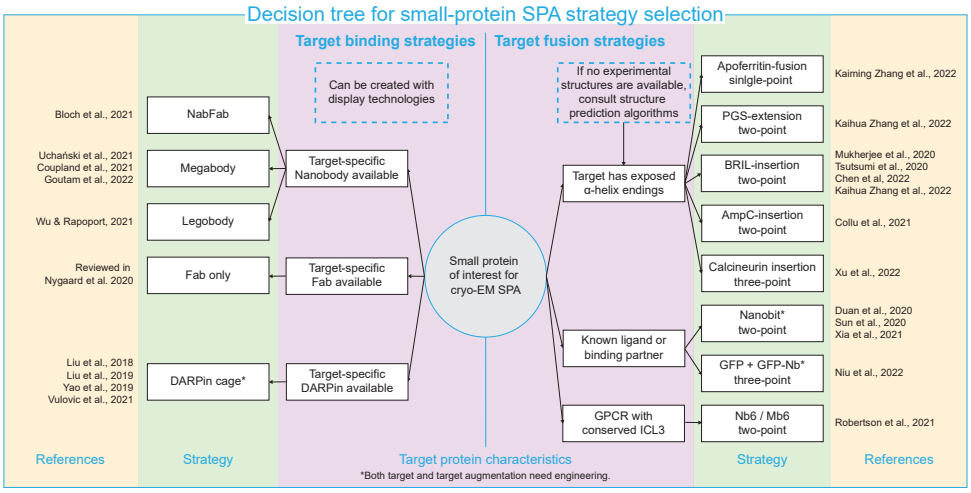


Figure 6.4: To enable cryo-EM SPA of small target proteins of interest (center vignette), the characteristics of the protein can guide choice of strategy for protein size-augmentation. A target binding (left) or a target fusion (right) strategy can be employed, each having multiple variants that together cover a wide range of small protein targets. Boxes in dashed blue lines are additional suggestions.

ACKNOWLEDGEMENTS

Structural data from Protein Data Bank (PDB) were used under the CC0 1.0 Universal (CC0 1.0) Public Domain Dedication license. DHM is supported by an NWO grant (OCENW.XS4.059).

REFERENCES

[1] K. Wentinck, C. Gogou, and D. H. Meijer, *Putting on molecular weight: Enabling cryo-em structure determination of sub-100-kda proteins,* Current Research in Structural Biology **4**, 332 (2022).

[2] J.-P. Renaud, A. Chari, C. Ciferri, W.-t. Liu, H.-W. Remigy, H. Stark, and C. Wiesmann, *Cryo-em in drug discovery: achievements, limitations and prospects,* NATURE REVIEWS DRUG DISCOVERY **17**, 471 (2018).

[3] C. Wigge, A. Stefanovic, and M. Radjainia, *The rapidly evolving role of cryo-em in drug design.* Drug discovery today. Technologies **38**, 91 (2020).

[4] J. Garcia-Nafria and C. G. Tate, *Cryo-electron microscopy: Moving beyond x-ray crystal structures for drug receptors and drug development,* in *ANNUAL REVIEW OF PHARMACOLOGY AND TOXICOLOGY, VOL 60*, Annual Review of Pharmacology and Toxicology, Vol. 60, edited by P. Insel (2020) pp. 51–71.

[5] J. P. Overington, B. Al-Lazikani, and A. L. Hopkins, *Opinion - how many drug targets are there?* NATURE REVIEWS DRUG DISCOVERY **5**, 993 (2006).

[6] L. Brocchieri and S. Karlin, *Protein length in eukaryotic and prokaryotic proteomes,* NUCLEIC ACIDS RESEARCH **33**, 3390 (2005).

[7] G. C. Lander and R. M. Glaeser, *Conquer by cryo-em without physically dividing,* BIOCHEMICAL SOCIETY TRANSACTIONS **49**, 2287 (2021).

[8] R. Nygaard, J. Kim, and F. Mancia, *Cryo-electron microscopy analysis of small membrane proteins,* CURRENT OPINION IN STRUCTURAL BIOLOGY **64**, 26 (2020).

[9] N. Aissaoui, J. Lai-Kee-Him, A. Mills, N. Declerck, Z. Morichaud, K. Brodolin, S. Baconnais, E. Le Cam, J. B. Charbonnier, R. Sounier, S. Granier, V. Ropars, P. Bron, and G. Bellot, *Modular imaging scaffold for single-particle electron microscopy,* ACS NANO **15**, 4186 (2021).

[10] K. Zhang, H. Wu, N. Hoppe, A. Manglik, and Y. Cheng, *Fusion protein strategies for cryo-em study of g protein-coupled receptors,* NATURE COMMUNICATIONS **13** (2022), 10.1038/s41467-022-32125-2.

[11] Y.-H. Chiu, K.-T. Ko, T.-J. Yang, K.-P. Wu, M.-R. Ho, P. Draczkowski, and S.-T. D. Hsu, *Direct visualization of a 26 kda protein by cryo-electron microscopy aided by a small scaffold protein,* BIOCHEMISTRY **60**, 1075 (2021).

[12] J. A. Coleman, V. Navratna, D. Antermite, D. Yang, J. A. Bull, and E. Gouaux, *Chemical and structural investigation of the paroxetine-human serotonin transporter complex,* ELIFE **9** (2020), 10.7554/eLife.56427.

[13] S. Wu, A. Avila-Sakar, J. Kim, D. S. Booth, C. H. Greenberg, A. Rossi, M. Liao, X. Li, A. Alian, S. L. Griner, N. Juge, Y. Yu, C. M. Mergel, J. Chaparro-Riggers, P. Strop, R. Tampe, R. H. Edwards, R. M. Stroud, C. S. Craik, and Y. Cheng, *Fabs enable single particle cryoem studies of small proteins,* STRUCTURE **20**, 582 (2012).

[14] T. Uchanski, E. Pardon, and J. Steyaert, *Nanobodies to study protein conformational states,* CURRENT OPINION IN STRUCTURAL BIOLOGY **60**, 117 (2020).

[15] Y. Liu, S. Gonen, T. Gonen, and T. O. Yeates, *Near-atomic cryo-em imaging of a small protein displayed on a designed scaffolding system,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **115**, 3362 (2018).

[16] I. Vulovic, Q. Yao, Y.-J. Park, A. Courbet, A. Norris, F. Busch, A. Sahasrabuddhe, H. Merten, D. D. Sahtoe, G. Ueda, J. A. Fallas, S. J. Weaver, Y. Hsia, R. A. Langan, A. Plueckthun, V. H. Wysocki, D. Veesler, G. J. Jensen, and D. Baker, *Generation of ordered protein assemblies using rigid three-body fusion,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **118** (2021), 10.1073/pnas.2015037118.

[17] Q. Yao, S. J. Weaver, J.-Y. Mock, and G. J. Jensen, *Fusion of darpin to aldolase enables visualization of small protein by cryo-em,* STRUCTURE **27**, 1148+ (2019).

[18] T. O. Yeates, M. P. Agdanowski, and Y. Liu, *Development of imaging scaffolds for cryo-electron microscopy,* CURRENT OPINION IN STRUCTURAL BIOLOGY **60**, 142 (2020).

[19] J. Helma, M. C. Cardoso, S. Muyldermans, and H. Leonhardt, *Nanobodies and recombinant binders in cell biology,* JOURNAL OF CELL BIOLOGY **209**, 633 (2015).

[20] J. S. Bloch, S. Mukherjee, J. Kowal, E. Filippova, V, M. Niederer, E. Pardon, J. Steyaert, A. A. Kossiakoff, and K. P. Locher, *Development of a universal nanobody-binding fab module for fiducial-assisted cryo-em studies of membrane,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **118** (2021), 10.1073/pnas.2115435118.

[21] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, and D. Hassabis, *Highly accurate protein structure prediction with alphafold,* NATURE **596**, 583+ (2021).

[22] T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, and T. E. Ferrin, *Ucsf chimerax: Meeting modern challenges in visualization and analysis,* PROTEIN SCIENCE **27**, 14 (2018).

[23] T. Uchanski, S. Masiulis, B. Fischer, V. Kalichuk, U. Lopez-Sanchez, E. Zarkadas, M. Weckener, A. Sente, P. Ward, A. Wohlkonig, T. Zogg, H. Remaut, J. H. Naismith, H. Nury, W. Vranken, A. R. Aricescu, E. Pardon, and J. Steyaert, *Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-em,* NATURE METHODS **18**, 60+ (2021).

[24] C. E. Coupland, S. A. Andrei, T. B. Ansell, L. Carrique, P. Kumar, L. Sefer, R. A. Schwab, E. F. X. Byrne, E. Pardon, J. Steyaert, A. Magee, I, T. Lanyon-Hogg, M. S. P. Sansom, E. W. Tate, and C. Siebold, *Structure, mechanism, and inhibition of hedgehog acyltransferase,* MOLECULAR CELL **81**, 5025+ (2021).

[25] K. Goutam, F. S. Ielasi, E. Pardon, J. Steyaert, and N. Reyes, *Structural basis of sodium-dependent bile salt uptake into the liver,* NATURE **606**, 1015+ (2022).

[26] X. Wu and T. A. Rapoport, *Cryo-em structure determination of small proteins by nanobody-binding scaffolds (legobodies),* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **118** (2021), 10.1073/pnas.2115001118.

[27] S. Mukherjee, S. K. Erramilli, M. Ammirati, F. J. D. Alvarez, K. F. Fennell, M. D. Purdy, B. M. Skrobek, K. Radziwon, J. Coukos, Y. Kang, P. Dutka, X. Gao, X. Qiu, M. Yeager, H. E. Xu, S. Han, and A. A. Kossiakoff, *Synthetic antibodies against bril as universal fiducial marks for single-particle cryoem structure determination of membrane proteins,* NATURE COMMUNICATIONS **11** (2020), 10.1038/s41467-020-15363-0.

[28] J. Duan, D.-d. Shen, X. E. Zhou, P. Bi, Q.-f. Liu, Y.-x. Tan, Y.-w. Zhuang, H.-b. Zhang, P.-y. Xu, S.-J. Huang, S.-s. Ma, X.-h. He, K. Melcher, Y. Zhang, H. E. Xu, and Y. Jiang, *Cryo-em structure of an activated vip1 receptor-g protein complex revealed by a nanobit tethering strategy,* NATURE COMMUNICATIONS **11** (2020), 10.1038/s41467-020-17933-8.

[29] W. Sun, L.-N. Chen, Q. Zhou, L.-H. Zhao, D. Yang, H. Zhang, Z. Cong, D.-D. Shen, F. Zhao, F. Zhou, X. Cai, Y. Chen, Y. Zhou, S. Gadgaard, W. J. C. van der Velden, S. Zhao, Y. Jiang, M. M. Rosenkilde, H. E. Xu, Y. Zhang, and M.-W. Wang, *A unique hormonal recognition feature of the human glucagon-like peptide-2 receptor,* CELL RESEARCH **30**, 1098 (2020).

[30] R. Xia, N. Wang, Z. Xu, Y. Lu, J. Song, A. Zhang, C. Guo, and Y. He, *Cryo-em structure of the human histamine h₁ receptor/g_q complex,* NATURE COMMUNICATIONS **12** (2021), 10.1038/s41467-021-22427-2.

[31] B. C. McIlwain, A. L. Erwin, A. R. Davis, B. Ben Koff, L. Chang, T. Bylund, G.-Y. Chuang, P. D. Kwong, M. D. Ohi, Y.-T. Lai, and R. B. Stockbridge, *N-terminal transmembrane-helix epitope tag for x-ray crystallography and electron microscopy of small membrane proteins,* JOURNAL OF MOLECULAR BIOLOGY **433** (2021), 10.1016/j.jmb.2021.166909.

[32] R. Tamura-Sakaguchi, R. Aruga, M. Hirose, T. Ekimoto, T. Miyake, Y. Hizukuri, R. Oi, M. K. Kaneko, Y. Kato, Y. Akiyama, M. Ikeguchi, K. Iwasaki, and T. Nogi, *Moving toward generalizable nz-1 labeling for 3d structure determination with optimized epitope-tag insertion,* ACTA CRYSTALLOGRAPHICA SECTION D-STRUCTURAL BIOLOGY **77**, 645 (2021).

[33] N. Tsutsumi, S. Mukherjee, D. Waghray, C. Y. Janda, K. M. Jude, Y. Miao, J. S. Burg, N. G. Aduri, A. A. Kossiakoff, C. Gati, and K. C. Garcia, *Structure of human frizzled5*

by fiducial-assisted cryo-em supports a heterodimeric mechanism of canonical wnt signaling, ELIFE **9** (2020), 10.7554/eLife.58464.

[34] H. Chen, W. Huang, and X. Li, *Structures of oxysterol sensor ebi2/gpr183, a key regulator of the immune response,* STRUCTURE **30**, 1016+ (2022).

[35] Z. Deng, Y. Zhao, J. Feng, J. Zhang, H. Zhao, M. J. Rau, J. A. J. Fitzpatrick, H. Hu, and P. Yuan, *Cryo-em structure of a proton-activated chloride channel tmem206,* SCIENCE ADVANCES **7** (2021), 10.1126/sciadv.abe5983.

[36] G. Collu, I. Mohammed, A. Lafita, T. Bierig, E. Poghosyan, S. Bliven, and R. M. Benoit, *Cryo-em structure of a single-chain β1-adrenoceptor – ampc β-lactamase fusion protein,* BioRxiv (2021), 10.1101/2021.09.25.461805.

[37] J. Xu, G. Chen, H. Wang, S. Cao, J. Heng, X. Deupi, Y. Du, and B. K. Kobilka, *βcalcineurin-fusion facilitates cryo-em structure determination of a family a gpcr,* bioRxiv (2022), 10.1101/2022.03.27.485993.

[38] M. J. Robertson, M. M. Papasergi-Scott, F. He, *et al.*, *Structure determination of inactive-state gpcrs with a universal nanobody,* Nature Structural & Molecular Biology **29**, 1188 (2022).

[39] M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, C. Millan, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. van Dijk, A. C. Ebrecht, D. J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C. K. Yip, J. E. Burke, K. C. Garcia, N. V. Grishin, P. D. Adams, R. J. Read, and D. Baker, *Accurate prediction of protein structures and interactions using a three-track neural network,* SCIENCE **373**, 871+ (2021).

[40] Y. Liu, D. T. Huynh, and T. O. Yeates, *A 3.8 Å resolution cryo-em structure of a small protein bound to an imaging scaffold,* NATURE COMMUNICATIONS **10** (2019), 10.1038/s41467-019-09836-0.

[41] K. Zhang, N. Horikoshi, S. Li, A. S. Powers, M. A. Hameedi, G. D. Pintilie, H.-D. Chae, Y. A. Khan, C.-M. Suomivuori, R. O. Dror, K. M. Sakamoto, W. Chiu, and S. Wakatsuki, *Cryo-em, protein engineering, and simulation enable the development of peptide therapeutics against acute myeloid leukemia,* ACS CENTRAL SCIENCE **8**, 214 (2022).

7

Conclusion and Outlook

In this chapter, future directions are proposed that build forward upon the progress made in this thesis.

This thesis covers various fundamental principles that govern the functionality of membrane proteins at the neuronal synapse, where we provide novel insights into diversification of synCAM binding specificity. X-ray crystallography- and cryo-EM-derived atomic structures are used to, respectively, showcase how glycans can modulate protein-protein binding interfaces directly, and how alternative splicing globally changes protein conformation to modify binding. Additionally, we made significant steps towards generalised cryo-EM reconstruction of membrane proteins within a transmembrane potential to more closely approximate their native neuronal context.

HIGH-RESOLUTION RECONSTRUCTION OF A_0 ISOFORMS

In chapter 2, we provided insight into how inserts of a few amino acids can completely alter the conformation of the nearly 3,000-amino acid teneurin-3 protein, and how these structural changes differentially guide the outgrowing axons of neurons. The single-residue resolution of the cryo-EM reconstructions that was achieved for only two of the four isoforms already led to important insights into how a microphthalmia-associated mutation may lead to improper neuronal wiring - as discussed in chapter 3.

Resolving the other, A_0, isoforms of teneurin-3 (Ten3) at high resolution in the future may provide similar insights. It may, for instance, reveal how splice insert B can result in the conformation difference observed between the A_0B_0 and A_0B_1 isoforms. Various approaches might ensure the stability of these compact dimer conformations during cryo-EM sample preparation. Mild cross-linking conditions using various chemical agents are commonly used to capture unstable complexes in the vitreous ice on the sample grid [1–4]. Similarly, complexes are occasionally stabilised by appending tags with high mutual affinity to the protomers of interest [5]. We postulated that the aggresive hydrophobic nature of air-water interface (AWI) on the sample grid causes our A_0 compact dimers to dissociate. The complexes could therefore also be protected by shielding the AWI, for instance by using sub-critical micelle concetration (CMC) detergent concentrations as protective layers [6]. Recently, small proteins were developed that localise to the AWI, thereby also shielding the AWI, which has led to the successful reconstruction of proteins that were prone to aggregate on-grid [7].

TENEURIN-3 MOLECULAR DYNAMICS *in vitro*

The structural techniques used in chapter 2 intrinsically provide snap-shot insights on the molecular behaviour of Ten3. Small-angle X-ray scattering (SAXS) data of the different isoforms indicated that A_0B_1, for instance, may adopt multiple conformations (Chapter 2 Fig. 2.5A-B). It is conceivable that each isoform does not exclusively inhabit either of the resolved structures, but that the splice inserts change the relative occupation of these states per isoform. Future experimentation into the dynamical behaviour will be necessary to address such questions.

Single-molecule Förster resonance energy transfer (FRET, a technique to probe the dynamical behaviour of proteins) experiments of immobilised Ten3 on a 2D surface may prove feasible to asses the intramolecular reorganisations of the two superfolds [8,9].

Ten3 specifically possesses a few surface-accessible cystines that may be chemically labelled with Cy3/Cy5 FRET-pairs, for instance. The difference in distance between these pairs within each compact dimer could provide conformation-specific FRET signal, and thereby allow monitoring the conformational state occupation over time. Immobilisation on a 2D surface additionally offers a set-up more representative of the synaptic membrane than extracellular domains in solution. The intermolecular dynamics could similarly also be assessed, and is theoretically less challenging. In this case, a batch of purified Ten3 could be tagged with Cy3, for instance, and immobilised on the surface, while another batch could be labeled with Cy5 and stays in solution. The binding kinetics can then be monitored over time, provide K_D values, but also time-resolved binding and unbinding events of single molecules, which are not possible to obtain with surface plasmon resonance, for instance. High-speed atomic force micropscopy (AFM) also offers a way to assess the spatiotemporal dynamics of single Ten3 molecules [10,11]. The large size of Ten3 dimers (~600kDa) probably provides the resolution needed to resolve the adopted compact dimer conformations, and allows monitoring these conformations over time. Moreover, this technique is compatible with planar membranes to which teneurins can be tethered, either via purification tags or via included transmembrane domains [10,11]. Such attachment to a membrane substrate permits free diffusion of Ten3 along the plane of the membrane and allows the formations of higher-order *trans*-complexes such as the pre-synaptic Ten3 nano-size clusters reported by others [12].

MICROPHTHALMIA IN NEURONAL RECOGNITION

In chapter 3, we reveal that a single-residue mutation in the teneurin-3 ECD abolishes homophilic clustering for the A_1 isoforms. Heterophilic interactions with latrophilin were also affected, albeit less prominently. Although the molecular dynamics simulations provided insight into the local structural effects, open questions remain about how this mutation changes the overall structure of the ECD. Moreover, the effect of altered *trans*-cellular binding on axon guidance and neuronal wiring remain to be investigated.

The techniques we used to delineate wildtype (WT) Ten3 structure-function relations in chapter 2 offer a great way to start tackling these questions. Cryo-EM, SAXS data and thermal stability assays of the purified ECDs will reveal how the mutation affects the conformation. Axon outgrowth stripe assays of all isoforms versus their mutated versions could provide the first insight into how the structural changes affect neuronal wiring.

TENEURIN-3 MOLECULAR DYNAMICS *in silico*

The molecular dynamics (MD) simulations performed in chapter 3 provided informative insight into the local disruption of the EGF8-ABD interface specific to the splice insert A-containing isoforms of mouse Ten3. It will be crucial to also perform these MD simulations for human Ten3.

MD simulations are, of course, not limited to the assessment of mutagenic effects. It will be interesting to obtain more insights into the wildtype versions of teneurin-3 by incorporating additional native structural features and applying constraints that Ten3 experiences on the synaptic membrane. Firstly, the full ECD can be simulated that is

tethered to a membrane. This may recapitulate putative interactions with the membrane[13] or the NHL-Ig/EGF interactions we propose in chapter 2 to explain the difference between the A_0 isoforms. Secondly, the 17 glycans of heavily-glycosylated Ten3 ECD should also be included, as other groups[13–15] and we[16] demonstrated that glycans play important structural roles in synCAM complex formation. Finally, the yet unresolved dimer-of-dimer interactions[17,18] may be captured in MD simulations with multiple dimer molecules. More course-grained MD simulations may be considered to attain simulated time scales that can capture large conformational changes of more and larger molecules.

MEMBRANE-INSERTED ASAP2S CRYO-EM RECONSTRUCTION

In chapter 5, one of the few remaining steps in the workflow is the cryo-EM reconstruction of ASAP2s. In chapter 6, we review that cryo-EM reconstruction of <100-kDa proteins is generally challenging. Especially the high electron density of phospholipids in the lipoosomal membrane would obscure ASAP2s signal.

Earlier works have, nonetheless, realised the high-resolution reconstruction of various membrane protein in vesicles using cryo-EM[19–25]. These examples mostly include membrane protein complexes much larger than 100 kDa, with more easily detectable domains sticking out of the membrane. ASAP2s is roughly 50 kDa, and reconstruction of the monomeric protein without size augmentation would prove challenging, if not impossible. Our finding that ASAP2s forms homo-oligomers may render them suitable for particle picking and alignment. For instance, a homo-trimer of transmembrane protein "AcrB" of ~350 kDa in size was also recently realised[21]. If this intrinsic increase of molecular weight does not suffice for ASAP2s reconstruction from vesicles, we can resort to one of the methods described in chapter 6. ASAP2s was selected, in part, as it also has a GFP protruding from the membrane. This GFP offers a basis to dock one or multiple NabFabs[26,27]. Various constructs are commercially available for the production of nanobodies that bind GFP at non-overlapping sites[28–30]. Appending nanobodies to the GFP domain - especially in the case of multiple GFPs in ASAP2s oligomers - may already aid reconstruction. A NabFab could be attached to each of the GFP-bound nanobodies to further facilitate cryo-EM single-particle analysis.

RECONSTRUCTING BIOLOGICALLY RELEVANT VOLTAGE-SENSITIVE MEMBRANE PROTEINS

In chapter 5, we aim to enable cryo-EM sample preparation of membrane-embedded voltage-sensitive membrane proteins (VSMP) in the presence of actual membrane voltage[24,25]. To this end, we purified the ASAP2s GEVI at high yield and purity, and inserted it into vesicles of sizes that are compatible with the vitreous ice layer.

While the voltage-sensitive function of ASAP2s still needs to be recovered for this workflow, the purification of other membrane proteins can be conducted simultaneously. As discussed in more detail in chapter 5, two G protein-coupled receptor (GPCR) families are particularly interesting candidates of fundamental and medical relevance - members of the metabotropic glutamate receptors [31] and of the muscarinic recep-

tors [32–39]. They both possess yet unspecified voltage-sensing mechanisms, and existing data indicate that these mechanisms could be resolved using cryo-EM reconstructions. For instance, the translocation of sodium [37,38], or the concerted movement of three tyrosines[32,35] in the tyrosine lid. The ability of GPCRs to bind large G protein complexes outside the membrane makes them excellent candidates for cryo-EM reconstruction as ample recent endeavours have enabled faithful reconstruction of GPCRs in various states[40,41].

FINAL REMARKS

In this thesis we addressed - from the structural perspective of neuronal membrane proteins - two fundamental characteristics of neurons: Neuronal integration into networks and their electrochemical signalling within those networks. Although the task of mapping the biological basis of the experiential world is vast, given, for instance, the added diversity of alternative splicing to network formation alone, we should not be discouraged. We are not only endeavouring to understand how we perceive the world of things, and how we can find our way through it; we are the conscious brain becoming increasingly aware of itself.

REFERENCES

[1] J. D. Strauss and T. Wagenknecht, *Structure of glutaraldehyde cross-linked ryanodine receptor,* Journal of Structural Biology **181**, 300 (2013).

[2] K. Adamus, S. N. Le, H. Elmlund, M. Boudes, and D. Elmlund, *Agarfix: Simple and accessible stabilization of challenging single-particle cryo-em specimens through crosslinking in a matrix of agar,* Journal of Structural Biology **207**, 327 (2019).

[3] A. K. Shukla, G. H. Westfield, K. Xiao, R. I. Reis, L. Y. Huang, P. Tripathi-Shukla, J. Qian, S. Li, A. Blanc, A. N. Oleskie, A. M. Dosey, M. Su, C. R. Liang, L. L. Gu, J. M. Shan, X. Chen, R. Hanna, M. Choi, X. J. Yao, B. U. Klink, A. W. Kahsai, S. S. Sidhu, S. Koide, P. A. Penczek, A. A. Kossiakoff, V. L. Woods, B. K. Kobilka, G. Skiniotis, and R. J. Lefkowitz, *Visualization of arrestin recruitment by a g-protein-coupled receptor,* Nature **512**, 218 (2014).

[4] B. Kastner, N. Fischer, M. M. Golas, B. Sander, P. Dube, D. Boehringer, K. Hartmuth, J. Deckert, F. Hauer, E. Wolf, H. Uchtenhagen, H. Urlaub, F. Herzog, J. M. Peters, D. Poerschke, R. Lührmann, and H. Stark, *Grafix: Sample preparation for single-particle electron cryomicroscopy,* Nature Methods **5**, 53 (2008).

[5] J. Duan, D.-d. Shen, X. E. Zhou, P. Bi, Q.-f. Liu, Y.-x. Tan, Y.-w. Zhuang, H.-b. Zhang, P.-y. Xu, S.-J. Huang, S.-s. Ma, X.-h. He, K. Melcher, Y. Zhang, H. E. Xu, and Y. Jiang, *Cryo-em structure of an activated vip1 receptor-g protein complex revealed by a nanobit tethering strategy,* NATURE COMMUNICATIONS **11** (2020), 10.1038/s41467-020-17933-8.

[6] S. Chen, J. Li, K. R. Vinothkumar, and R. Henderson, *Interaction of human erythrocyte catalase with air–water interface in cryoEM,* Microscopy **71**, i51 (2022), https://academic.oup.com/jmicro/article-pdf/71/Supplement_1/i51/42548217/dfab037.pdf .

[7] K. M. Abe, G. Li, Q. He, T. Grant, and C. J. Lim, *Small lea proteins mitigate air-water interface damage to fragile cryo-em samples during plunge freezing,* Nature Communications **15**, 7705 (2024).

[8] H. Mazal and G. Haran, *Single-molecule fret methods to study the dynamics of proteins at work,* Current Opinion in Biomedical Engineering **12**, 8 (2019).

[9] M. Filius, T. J. Cui, A. N. Ananth, M. W. Docter, J. W. Hegge, J. V. D. Oost, and C. Joo, *High-speed super-resolution imaging using protein-assisted dna-paint,* Nano Letters **20**, 2264 (2020).

[10] A. Karner, B. Nimmervoll, B. Plochberger, E. Klotzsch, A. Horner, D. G. Knyazev, R. Kuttner, K. Winkler, L. Winter, C. Siligan, N. Ollinger, P. Pohl, and J. Preiner, *Tuning membrane protein mobility by confinement into nanodomains,* Nature Nanotechnology **12**, 260 (2017).

[11] S. Maity, G. Trinco, P. Buzón, Z. R. Anshari, N. Kodera, K. X. Ngo, T. Ando, D. J. Slotboom, and W. H. Roos, *High-speed atomic force microscopy reveals a three-state elevator mechanism in the citrate transporter cits,* Proceedings of the National Academy of Sciences **119**, e2113927119 (2022), https://www.pnas.org/doi/pdf/10.1073/pnas.2113927119 .

[12] X. Zhang, P. Y. Lin, K. Liakath-Ali, and T. C. Südhof, *Teneurins assemble into presynaptic nanoclusters that promote synapse formation via postsynaptic non-teneurin ligands,* Nature Communications **13** (2022), 10.1038/s41467-022-29751-1.

[13] J. Li, Y. Xie, S. Cornelius, X. Jiang, R. Sando, S. P. Kordon, M. Pan, K. Leon, T. C. Südhof, M. Zhao, and D. Araç, *Alternative splicing controls teneurin-latrophilin interaction and synapse specificity by a shape-shifting mechanism,* Nature Communications **11** (2020), 10.1038/s41467-020-16029-7.

[14] J. Wang, Y. Miao, R. Wicklein, Z. Sun, J. Wang, K. M. Jude, R. A. Fernandes, S. A. Merrill, M. Wernig, K. C. Garcia, and T. C. Südhof, *Rtn4/nogo-receptor binding to bai adhesion-gpcrs regulates neuronal development,* Cell **184**, 5869 (2021).

[15] O. Akkermans, C. Delloye-Bourgeois, C. Peregrina, M. Carrasquero-Ordaz, M. Kokolaki, M. Berbeira-Santana, M. Chavent, F. Reynaud, R. Raj, J. Agirre, M. Aksu, E. S. White, E. Lowe, D. B. Amar, S. Zaballa, J. Huo, I. Pakos, P. T. McCubbin, D. Comoletti, R. J. Owens, C. V. Robinson, V. Castellani, D. del Toro, and E. Seiradake, *Gpc3-unc5 receptor complex structure and role in cell migration,* Cell **185**, 3931 (2022).

[16] L. M. Chataigner, C. Gogou, M. A. den Boer, C. P. Frias, D. M. Thies-Weesie, J. C. Granneman, A. J. Heck, D. H. Meijer, and B. J. Janssen, *Structural insights into the contactin 1 – neurofascin 155 adhesion complex,* Nature Communications **13** (2022), 10.1038/s41467-022-34302-9.

[17] D. H. Meijer, C. P. Frias, J. W. Beugelink, Y. N. Deurloo, and B. J. C. Janssen, *Teneurin4 dimer structures reveal a calcium-stabilized compact conformation supporting homomeric trans-interactions,* The EMBO Journal **41** (2022), 10.15252/embj.2020107505.

[18] C. Gogou, J. W. Beugelink, C. P. Frias, L. Kresik, N. Jaroszynska, U. Drescher, B. J. Janssen, R. Hindges, and D. H. Meijer, *Alternative splicing controls teneurin-3 compact dimer formation for neuronal recognition,* Nature Communications **15** (2024), 10.1038/s41467-024-47763-x.

[19] L. Wang and F. J. Sigworth, *Structure of the bk potassium channel in a lipid membrane from electron cryomicroscopy,* Nature **461**, 292 (2009).

[20] L. Tonggu and L. Wang, *Structure of the human bk ion channel in lipid environment,* Membranes **12** (2022), 10.3390/membranes12080758.

[21] X. Yao, X. Fan, and N. Yan, *Cryo-em analysis of a membrane protein embedded in the liposome,* Proceedings of the National Academy of Sciences **117**, 18497 (2020).

[22] Z. Melville, K. Kim, O. B. Clarke, and A. R. Marks, *High-resolution structure of the membrane-embedded skeletal muscle ryanodine receptor,* Structure **30**, 172 (2022).

[23] N. Sengupta, A. K. Mondal, S. Mishra, K. Chattopadhyay, and S. Dutta, *Single-particle cryo-em reveals conformational variability of the oligomeric vcc β-barrel pore in a lipid bilayer,* Journal of Cell Biology **220** (2021), 10.1083/JCB.202102035.

[24] V. S. Mandala and R. Mackinnon, *Voltage-sensor movements in the eag kv channel under an applied electric field,* Proceedings of the National Academy of Sciences of the United States of America **46** (2022), 10.1073/pnas.

[25] V. S. Mandala and R. MacKinnon, *The membrane electric field regulates the pip2-binding site to gate the kcnq1 channel,* Proceedings of the National Academy of Sciences of the United States of America **120** (2023), 10.1073/pnas.2301985120.

[26] R. Nygaard, J. Kim, and F. Mancia, *Cryo-electron microscopy analysis of small membrane proteins,* CURRENT OPINION IN STRUCTURAL BIOLOGY **64**, 26 (2020).

[27] J. S. Bloch, S. Mukherjee, J. Kowal, E. Filippova, V, M. Niederer, E. Pardon, J. Steyaert, A. A. Kossiakoff, and K. P. Locher, *Development of a universal nanobody-binding fab module for fiducial-assisted cryo-em studies of membrane,* PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **118** (2021), 10.1073/pnas.2115435118.

[28] A. Kirchhofer, J. Helma, K. Schmidthals, C. Frauer, S. Cui, A. Karcher, M. Pellis, S. Muyldermans, C. S. Casas-Delucchi, M. C. Cardoso, *et al.*, *Modulation of protein properties in living cells using nanobodies,* Nature Structural & Molecular Biology **17**, 133 (2010).

[29] H. Cai, H. Yao, T. Li, C. A. Hutter, Y. Li, Y. Tang, M. A. Seeger, and D. Li, *An improved fluorescent tag and its nanobodies for membrane protein expression, stability assay, and purification,* Communications Biology **3**, 1 (2020).

[30] A. T. Q. Cong, T. L. Witter, and M. J. Schellenberg, *High-efficiency recombinant protein purification using mcherry and yfp nanobody affinity matrices,* Protein Science **31**, e4383 (2022), https://onlinelibrary.wiley.com/doi/pdf/10.1002/pro.4383 .

[31] L. Ohana, O. Barchad, I. Parnas, and H. Parnas, *The metabotropic glutamate g-protein-coupled receptors mglur3 and mglur1a are voltage-sensitive,* Journal of Biological Chemistry **281**, 24204 (2006).

[32] O. Barchad-Avitzur, M. F. Priest, N. Dekel, F. Bezanilla, H. Parnas, and Y. Ben-Chaim, *A novel voltage sensor in the orthosteric binding site of the m2 muscarinic receptor,* Biophysical Journal **111**, 1396 (2016).

[33] Y. Ben-Chaim, B. Chanda, N. Dascal, F. Bezanilla, I. Parnas, and H. Parnas, *Movement of 'gating charge' is coupled to ligand binding in a g-protein-coupled receptor,* Nature **444**, 106 (2006).

[34] Y. B. Chaim, S. Bochnik, I. Parnas, and H. Parnas, *Voltage affects the dissociation rate constant of the m2 muscarinic receptor,* PLoS ONE **8** (2013), 10.1371/journal.pone.0074354.

[35] Y. Ben-Chaim, C. Broide, and H. Parnas, *The coupling of the m2 muscarinic receptor to its g protein is voltage dependent,* PLoS ONE **14** (2019), 10.1371/journal.pone.0224367.

[36] A. Hoppe, M. Marti-Solano, M. Drabek, M. Bünemann, P. Kolb, and A. Rinne, *The allosteric site regulates the voltage sensitivity of muscarinic receptors,* Cellular Signalling **42**, 114 (2018).

[37] O. N. Vickery, J. P. Machtens, and U. Zachariae, *Membrane potentials regulating gpcrs: insights from experiments and molecular dynamics simulations,* Current Opinion in Pharmacology **30**, 44 (2016).

[38] O. N. Vickery, J. P. Machtens, G. Tamburrino, D. Seeliger, and U. Zachariae, *Structural mechanisms of voltage sensing in g protein-coupled receptors,* Structure **24**, 997 (2016).

[39] M. Tauber and Y. Ben-Chaim, *Voltage sensors embedded in g protein-coupled receptors,* International Journal of Molecular Sciences **25** (2024), 10.3390/ijms25105295.

[40] K. Wentinck, C. Gogou, and D. H. Meijer, *Putting on molecular weight: Enabling cryo-em structure determination of sub-100-kda proteins,* Current Research in Structural Biology **4**, 332 (2022).

[41] W. Shihoya, A. Iwama, F. K. Sano, and O. Nureki, *Cryo-EM advances in GPCR structure determination,* The Journal of Biochemistry **176**, 1 (2024), https://academic.oup.com/jb/article-pdf/176/1/1/58387362/mvae029.pdf .

ACKNOWLEDGEMENTS

Nothing presented in this dissertation would have been of much value to me if it were not for the meaning added to it by many others during the journey. I would therefore like to dedicate some words to thanking everyone who contributed to this work by shaping me and keeping me indisputably happy.

Dimphna Thank you for your reliable knowledgeability and decisiveness, and for you trust in mine. Both made me feel completely safe to try and err, which has taught me so much. It was exactly what I needed and wished for when joining your then still maturing lab. It has been beautiful to see the group flourish and to have been a part of that. I hope with all my heart that you will always keep valuing the individuality and creative source in all your students, as you have mine. I think we made a great team! **Cátia** You have been a warm, patient, and reliable safe haven to so many of us over the years. It was incredibly reassuring to always have someone around who is knowledgeable on so many levels, and who actively made me feel welcome to knock on the door for those matters. Thank you for all the support, guidance, and most of all, your friendship in all these great and sometimes tough times. **Agathe** A true companion in discussing almost anything together without judgment or offense, and mainly with a lot of fun and laughter. Thank you for making me feel like a valuable colleague during what seemed like such a short time together. **Koen** A great example of how efficient work and great fun go well together. Thanks for being such a great student that I could feel like a great teacher. **Angelica** You are a wonderful combination of innocence and competence – contagiously overflowing with emotion which eradiates trustworthiness. Thank you for lighting up the room. **Leo** Probably without knowing it yourself, you so clearly enact that you know what you are doing, in work and in life. Thanks for the derailed humour we could share together, but also the scientific help and sparring sessions. **Maja** In the short time we worked together we have already proven ourselves an effective an fun team. Thank you for all your thorough expertise for the experiments, and for being good company during the trip. **Lars** Your demeanour spreads calmness, and you possess a great deal of curiosity and good humour. Thanks for making me feel like some of the many things I blurt out might someday be of use to a young scientist. Thanks to **Martijn, Nynke, and Phuong** for facing so many scientific puzzles and challenges together, and your diligent contributions to the wonderful work we performed. **Joyce, Yanthi, Selina, Alin, Amke, Deep, Yasha, Jeroen, Lukas, Jurriaan, Elisa, Marlena, Dominique, Myta** Thank you for making the DM lab a dynamic, spirited, enthusiastic, safe, fun, and warm place.

Mama en Papa Dankjulliewel dat ik geen moment in mijn leven heb hoeven twijfelen dat mijn ouders van mij houden, in mij geloven, en onvoorwaardelijk klaar staan. Bedankt dat jullie iemand van mij hebben gemaakt die in ieder geval de intentie heeft om een goed mens te zijn. **Jorgo** Bedankt voor het banen van de weg, en voor het laten zien dat dingen des te mooier zijn wanneer ze niet eenvoudig verkregen zijn, en dat we ondanks onze duidelijke verschillen zo goed kunnen aanvoelen wat onvoorwaardelijke broederschap is. **Mignon** Dankjewel voor het laten zien hoe onbetwijfelbare liefde voelt en hoe essentieel dat is voor een volwaardig bestaan. **Poppie** Bedankt dat je er altijd bij

bent geweest, vooral voor onze lieve grootmoeder. **Christina** Bedankt dat je mijn eigenaardigheid accepteert en welkom bij de familie.

Sam (van Beljouw) A man with infinitely many seats in his heart for other people and ideas, achieved by being so much in touch with yourself. I was sure that regaining childlike wisdom was only reserved for a much later stage of life. Thanks for bringing that experience forward many decades in our friendship. **Kasper** For some reason, everything you say has instant validity in my experience, which clearly emerges from my endless respect for you. Thanks for the infinite fun and conversation, and for giving me the feeling that I just might be as respectable. **Hairy** I didn't know that brotherly feelings for another person were truly possible, but they are. Thanks for showing that there is always something you can do, even when that something is only talking. **Sandro** My first tutor and my bigger brother in the scientific practice. You have taught me a lot that I still use every day, foremost by good example - not only scientifically and professionally but also socially and energetically. Thanks for exemplifying the value of being yourself.

Throughout the years I have lived in many homes – places where I would return every day and would truly feel at home. This has been in so many different formats, but the common denominator is the familial connection I had with all the people in them. **Marc** Someone who has been my bar buddy, travel companion, squash rival, roommate, and basically my (little) brother through already so many years. Thanks for all the beautiful and annoying times together, and let's continue that trend. **Dennis** The embodiment of Dutch sobriety, whom I had the honour of enjoying our quiet and less quiet times at home together with. Thanks for being such a reliable and relatable factor the last few years. **Morris** Thanks for the many conversations and times of kind relaxation. I love all the times that we were in sync with each other, but the memories of where we couldn't possibly agree on something force me to smile the most. **Maurits** The imperturbable giant with an infinite stock of good humour (not as in 'good jokes', they are all terrible). Thanks for all the laughs, and of course, also all the more serious discussions. **Toga** The extent to which you know what you want and do what you like is humbling – especially when it comes to gadgets and fidgeting. Thanks for all the unpredictable and exciting situations as roommates. **Laurent** You are one of the most brilliant people when it comes to understanding the human experiential world. We have been great brothers in arms during the many in-house group conflicts. **Mauryyyn** It is remarkable how the short time that you lived with us makes my memories feel like you were there all these years. Thanks for you kindness an innocence. Apparently, those traits are a lot of fun, too. **Twips** I will never forget your 'je bent hier thuis', which has so much enhanced that feeling throughout my 5 years at the PC. Thanks for being a big-brotherly housemate. **Art** There is so little distance between what we see and what we get, and I admire that. Thanks for savouring Ludovico Einaudi with me at random times in the middle of the night. **Knoch** Geology truly did fit you, for you have always been like a rock: either brightly fluid or immovably rigid depending on the temperature of your mood. Thanks for your keeping me young, and for all the mismatched philosophical discussions. **Nadège** I have never seen anyone handle such difficulty in such a healthy way, without diminishing the complexity by looking away. Our theatre man has shown some great character development. Thanks for all the loud noises and weird stuff, which were so much fun. **Breezer** A man of so many faces throughout the years, yet ever-so-

little facial expression; a man of true, active personal development. Thanks for all the fun ridiculousness and good conversations. **Dibo** You are my first example in life that some people just stay 'your little baby' forever. I cannot deny however, that my little baby has grown into a competent young man capable of choosing his way and overcoming great adversity. Thanks for the hilarious times. Many loving thanks to **Skyp, Pamper, Fluf, Bruno, Lulu**.

My academic career started with some less-academic activities together with a great new group of friends. A group of such different people, still going on great trips to the other side of the world, as well as the comfy occasional get-together. It is heart-warming how effortlessly and uniquely we manage all of this still. **Yannick** I sometimes brag about how clearly you always see what the right thing to do is, and how reliably you enact this. Thanks for the great connection we have always had and the unlimited pool of social discussion material to fathom mankind. **Steven** Good-spiritedness manifest with a giant brain. Despite your deep understanding of philosophy, nature, and humanity, there exists no sign of negativity in you, and that gives me hope. Thanks for all the discussions deep into the night, and for always being a most reliable companion in all our adventures. **Matthijs** It is great to see how you do everything with the fullest embodiment: eating, cursing, working, and even being quiet. Thanks for all the fun, parties, workouts, adventures, and effortless mutual understanding. **René** What a massive transformation I have seen from close by and far away - from a sometimes too gentle giant into someone who firmly decides and claim what he wants, taking you to far edges of the world and even to happy fatherhood. Thanks for your friendship and our broad ranges of adventures. **Bor** You are the proof that a mental state can be reached where negativity ceases to exist. So far, I have only seen you express love, happiness, and 'yes'. Thanks for being an example in that, and for being an awesome buddy. **Thomas** The steam engine of calm determination. It is other-worldly how calmly you break through obstacles and how you devour set goals. Thanks for always being so enthusiastically there. **Sam (Sawyer)** It has been so cool to see how you acquired confident calmness without losing any of your excitement. Thanks for all the fun, craziness, and kindness. **Bas** It is almost boring how your invariable sociability and your salient competence has brought you exactly where I expected: There where great things grow from those qualities. **Rogier** I love your idiosyncrasy and how you wear it. **Mick** Being too kind might in the end not be negative after all. Thanks for your sincerity and listening ear. **Mathijs** I admire how you have really mastered doing things your own way in spite of opposing loud noises from all around. **Kenneth** I often still struggle to accept how you have become such a knowledgeable and responsible figure - the jester mounting the throne. I would also like to thank **Geert, Niels, Sjors, Bernice, Lisette** for forming an amazing team with which we learned such incredibly valuable lessons during this less-academic time in managing long-term difficult projects.

One group of friends has been, and still is, crucial in my playful development to tackle any kind of situation in this complex life. It goes without saying how influential that has been in managing the many facets of a PhD. **Sebas** I am sure that my words are not needed to express how much we've been – despite our obvious differences - one and the same person during the years that were so crucial in our social development. Thank you for that unquestionable connection and already eternal friendship. **Frans** You have in

my experience always been incredibly intelligent, and it was good to see how you also grew beautifully sensitive throughout the years. Thanks for your understanding. **Bob** It was awesome to playfully learn with (and mostly from) you the subtle art of professional teasing, but of course also the many years of unrestricted and open friendship. I would have been a completely different person without you. **Mats** By far the most impressive example of goal-directed and successful personal development. Thanks for all the good conversations, the great fun, and the proactivity in pulling the group towards adventure. **Daan** There is a lot of strength and sincerity in your no-nonsense attitude. It really made all our awesome nonsensical adventures in nature and culture all the more enjoyable. **Rik** So open with infinite room, kindness and understanding for all kinds of ideas an people. Thank you for being an example of the value of broadened acceptance, and of course, your warm friendship and all the random thing we've encountered. **Ivan** I like what I think to recognise as the development of self-satisfaction into self-peace. Thanks for the great hospitality all those years. It really improved our friendship, mutual understanding and memories. **Paul** As long as I can remember, you always happily dominated the social scene and parties without lifting a finger. I'm incredibly happy that that is also where you've found to still be at place, and where your energy is meaningfully spent. **Tom** The Dutchest example of sober solution-driven thinking and acting – often where no solutions were even needed... You have always been an immensely reliable, resourceful and fun friend to experience anything with. **Stijn** Another one of the stable and stabilising factors in our group, even through some incredible hardships that no one should have to face at younger age. And, or course, thanks for all the social events you held in all these different places. **Jim** You always seemed to clearly have your game together with your calmness, competence, diligence, and sometimes your destructive ice-cold criticality. You lately have also beautifully demonstrated how strong your drive is to be a good person in selflessly helping others who can use it. Thanks for inspiring me. **Christiaan** A kind and competent guy with always clearly the best of intentions.

I cannot deny that my university education was extremely challenging. I was lucky enough to have a wonderful group of people with whom we could take on this challenge in a positive-spirited way. **Célina** You may very well be the strongest person I know. You always kept moving forward despite quite some adversity, and often despite the possibility of receiving true understanding from those nearest to you. Thanks you for your direct and caring friendship from the very first moment we met. **Mindy** Without a doubt the most competent factor of our group from the very start; A go-getter mentality that deserves all the confidence attainable. Thanks for all the great times and all the explorations of how to keep oneself sane. **Kirsten** Without perhaps knowing, you exemplified how people meaningfully gravitate toward others with personality traits that they themselves unknowingly lack and should develop. It made me aware of the calmness that I still want to acquire in my life. **David** Throughout our academic adventure, I noticed immense growth in your passion for whatever you are doing, and I believe it has already brought you many great things and development. Thanks for years of fun and battling a shared educational foe. **Fiona** Although we have only gotten to know each other later, from a distance you were already an incredibly competent and hard-working individual, with a justified amount of calmness in that. Coming closer, that has only added the clear features of being fun and caring. **Valérie** I can still remember almost hearing the actual

'clicking' when we first really interacted, and how truthful this instantaneous moment of a clicking sensation between two people can be. Thanks for the almost inseparable friendship we have had, and for the cheeky exploration of topics that probably not many get to discuss. **Rochelle** It is great to see how your infinite chillness has been met with great professional success. I even noticed lately that I've been very proud to brag about your achievements. Thanks for all the fun interactions on random moments, and our nice walks in the park.

I never doubted about extending my stay at BN for four more years because of the mentality that was so obviously present with many of its members. I always loved, and wondered at, the willingness of everyone to so selflessly help out, the enthusiasm to spar about any (scientific) problem, and of course, just the occasional chats about anything in the hallway. It remains a personal goal to achieve the calmness and control from which it is able to stably provide it to others. My sincerest gratitude for this family-like environment to **Esi, Ilja, Anke, Theo, Jan, Tracey, Heleen**, and my special thanks to **Eli** who so often calmly demonstrated to me that brilliance is achieved by taking time to think. **Wiel** for sharing his incredible wealth of technical knowledge and never losing his patience with students that ask so much. **Cecilia** for being so compassionate regarding my experiments and taking time and responsibility to ensure correct understanding of what I am doing. **Ramon** for always having time to answer any experimental question, 'borrowing' me so many things, and for unhinged horsing around at any time of the day. **Jaco** for helping me out on many occasions throughout the years and our fun sparring about the academic institution. **Jacob** for those random conversations that necessarily always went quite a bit deeper and which were only received with bad humour during severe mutual misunderstanding.

I would also like to specifically thank the people that have generated and propagated the scientific core of BN, and have personally helped me in my scientific endeavours. **Cees** for showing trust in my abilities from a very early point and the many opportunities and support to further develop them. **Marileen** for providing critical guidance in managing my PhD project. **Gijsje** for the opportunities and trust to design and teach education at university level, and of course, for our fun chats in the hallway. **Daan (Brinks)** for his seemingly unlimited willingness to help by finding time for me on so many occasions to just spar or do some experiments together. **Stan** for his mentorship, both the more formal occasions and the more informal lessons between the lines of just chatting, sharing enthusiasm about our biological systems, and plenty of joking. **Nina** for her incredible enthusiasm for teaching students anything, which appears to be only limited by the fixed number of hours that fit into a day. **Arjen** for his substitute supervision, his enthusiasm to spar about structural biological puzzless, and the fun random chats at the coffee machine. I also acknowledge the influence that **Martin (Depken), Timon, Kristin, Pouyan, Marie-Eve, Christophe, Liedewij** have on what happens at and to BN, but mainly thank them for their incredibly relatable and comforting humanity.

I want to thank **Nils, Lucia, Alejandro, Eva, Hidde, Fede, Beatriz, Stefan (Loonen)**, for the wonderful added dimension of (organising) social activity at BN. Despite our unavoidable individual dedication to our projects, each one of us showed clear passion for bringing colleagues together and really thinking about how to do that.

BN has proved to be such a warm and fun place of all too human people that just

love to chat, greet, spar, think, complain, drink a beer, or help each other out. Many thanks to my fellow BN'ers over the years **Alicia, Aswin, Enzo, Leila, Marieke, Nynke (Hettema), Frank, Michal, Anouk, SaFyre, Sonam, Hande, Irene, Jeff, James, Marcos, Charu, Ana, Celine, Marijn, Moritz, Ran, Nicole, Miyase, Ivy, Tanja, Stefan (Huber), Clémence, Lennart, Alok, Reza, Yash, Daniel (Remirez Montero), Ali, Sumit, Filip, Louis, Humberto, Roland, Friedrich, JJ, Arash, Kijun, Carlos, Ivo, Jaïrus, Allard, Wayne, Sergii, Daniel (Verschueren), Alberto, Sonja, Sabina, Federico (Fanalista), Paola, Anthony, Alessio, JK, Yoones, Brian, Tišma, Martin (Holub), Roman, Rafael, Richard, Ashmiani**.

Raman The many sparring sessions for writing papers, and designing experiments and analysis have been much fun and very fruitful. Thanks for validating with me that successful science is playful exploration at its base. **Mike** Your 'mechanic-like' dry attitude in science is truly artistic. It tells me that whatever is not robust beyond an experimenter's precision is not worth the optimisation. **Ruben (Boot)** I loved our always-too-short coffee breaks, just hissing at 'the system'. Thanks for making me feel like my untested speculations were of any motivational value, and for reminding me that the PhD defense most preferably is an entertaining occasion for performer, judge and audience. **Stefan (Loonen)** Only a short time of working together was needed for me to recognise your incredible competence. What is even more reassuring that you will have a successful PhD track is the ease and happiness with which you carry this competence around.

Only thanks to our collaborators throughout the years, we were able to produce the multi-disciplinary findings presented in this thesis. Thanks to **Lucas, Nina, Stefan (Loonen), Stephany, Robert, Natalia** for such smooth and exciting collaboration, which I dare call successful and something we can all be truly proud of. A special word of gratitude in this regard to **Bert** and **Wouter** who have been so welcoming in allowing me how to become a more complete structural biologist, but even more for the fun and enthusiastic way of doing science of high quality.

My final words of thanks go to the members of the defense committee that were not already mentioned. Many thanks to **Chirlmin** for so enthusiastically joining my committee, and to **Jonathan** for using his deep expertise in our field to provide such valuable feedback on my thesis, and coming all the way to the Netherlands to join the committee in person.

CURRICULUM VITÆ

Christos GOGOU

04-11-1993 Born in Gorinchem, the Netherlands.

EDUCATION

2018 – 2020 *Master of Science* in Nanobiology, *Cum Laude*.
Delft University of Technology, the Netherlands.
Thesis: Constructing a cryo-EM assay for molecular voltage-sensitivity of liposome-reconstituted membrane proteins.
Supervisor: Dr. D.H. Meijer.

2013 – 2018 *Bachelor of Science* in Nanobiology.
Delft University of Technology, the Netherlands.

WORK EXPERIENCE

2024 – present *Postdoctoral researcher*, Dimphna Meijer lab at the Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, the Netherlands.

2020 – 2024 *Doctor of Philosophy*, Dimphna Meijer lab at the Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, the Netherlands.
Dissertation: Protein nanoarchitecture of neuronal wiring.
Promotors: Prof. dr. A.M. Dogterom.
Dr. D.H. Meijer.

2019 *Master intern*, Peng Yin lab at the Wyss Insitute, Harvard Medical School, Boston, USA.

2018 *Student research assistant*, Cees Dekker lab at the Kavli Institute of Nanoscience, Department of Bionanoscience, Delft University of Technology, the Netherlands.

GRANTS AND AWARDS

2024 *Poster prize*, EMBO Molneuro Workshop, Fodele, Crete.

2020 *Best master thesis award* runner-up, BEI bioengineering institute, Delft University of Technology.

2018 *Justus & Louise van Effen excellence scholarship*, Delft University of Technology.

OTHER ACHIEVEMENTS

2024 *Music bingo champion*, Quo Vadis Bionanoscience, Delft University of Technology.

2023 *Rotterdam marathon*, finished in 4:09:28.

2023 *The Hague half marathon*, finished in 1:57:07.

LIST OF PUBLICATIONS

THIS DISSERTATION

3. **C. Gogou**, J.W. Beugelink, C.P. Frias, L. Kresik, N. Jaroszynska, U. Drescher, B.J.C. Janssen, R. Hindges, and D.H. Meijer "Alternative splicing controls teneurin-3 compact dimer formation for neuronal recognition.", *Nature Communications*, 2024.

2. L.M.P. Chataigner, **C. Gogou**, M.A. den Boer, C.P. Frias, D.M.E. Thies-Weesie, J.C.M. Granneman, A.J.R. Heck, D.H. Meijer, B.J.C. Janssen, "Structural insights into the contactin 1 – neurofascin 155 adhesion complex", *Nature Communications*, 2022.

1. K. Wentinck*, **C. Gogou***, D.H. Meijer, "Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins", *Current Research in Structural Biology*, 2022. *Contributed equally to this work.

OTHER JOURNAL PUBLICATIONS

4. R.C. Boot, A. van der Net, **C. Gogou**, P. Mehta, D.H. Meijer, G.H. Koenderink, P.E. Boukany, "Cell spheroid viscoelasticity is deformation-dependent", *Scientific Reports*, 2024.

3. A. Japaridze, R. van Wee, **C. Gogou**, J.W.J. Kerssemakers, D.F. van den Berg, C. Dekker, "MukBEF-dependent chromosomal organization in widened Escherichia coli", *Frontiers in Microbiology*, 2023.

2. **C. Gogou***, A. Japaridze*, C. Dekker, "Mechanisms for chromosome segregation in bacteria", *Frontiers in Microbiology*, 2021. *Contributed equally to this work.

1. A. Japaridze, **C. Gogou**, J.W.J. Kerssemakers, H. M. Nguyen, C. Dekker, "Direct observation of independently moving replisomes in *Escherichia coli*", *Nature Communications*, 2020.

CONFERENCE PRESENTATIONS

9. NWO CHAINS, Veldhoven, the Netherlands, *Oral Presentation*, 2024.

8. Dutch Biophysics, Veldhoven, the Netherlands, *Poster Presentation (Poster Pitch)*, 2024.

7. EMBO Molneuro Workshop, Fodele, Crete, *Poster Presentation (Poster Prize)*, 2024.

6. Dutch Biophysics, Veldhoven, the Netherlands, *Oral Presentation*, 2023.

5. EMBO Molneuro Workshop, Fodele, Crete, *Poster Presentation*, 2022.

4. Dutch Biophysics, Veldhoven, the Netherlands, *Poster Presentation*, 2023.

3. Dutch Neuroscience, Tiel, the Netherlands, *Poster Presentation*, 2022.

2. Dutch Biophysics, Veldhoven, the Netherlands, *Poster Presentation*, 2022.

1. Molecular Mechanisms of Neuronal Connectivity (online), CSHL, *Poster Presentation*, 2020.

About the author



Before there was earth or sea or sky or structural insights on teneurin isoforms, Nature was what we call Chaos, discordant atoms nothing but unstable form and unretained shape, and it was not until 1993 AD, or 1 CG, that this conflict was ended. A relationship between a mortal and a God in the lands of Gorinchem created out of the raw confused mass and obscurity a creature now of mythical proportions, named **Christos Gogou**, a Greek that desired to be Roman.

From a young age, **Christos Gogou** wielded his polished pretoogjes and his crystal-clear intellect swiftly through the educational tribulations that would prepare him for the master trial in Academia. The physicality of mankind is in constant turnover, with each component of the body renewed within seven years, making the eternal pretoogjes of **Christos Gogou** a remarkable exception. It is with those shining brightly that **Christos Gogou** arrives at work, each morning fearlessly and 9 AM sharp, man of the Chronos that he is.

On his first day, in 26 CG, **Christos Gogou** pauses for a second at the doorsteps of notorious Building 58, the shadowy place where the mentally ill are generated, to stoically observe the encouraging incantation above the door, 'Whatever doesn't kill you makes you tired.' A nostalgic smile appears on his face, as **Christos Gogou** went to the same preliminary school as the originator of this phrase, Friedrich Nietzsche, with whom **Christos Gogou** already conversed in aphorisms before learning Simia, Nota, Oogst. Standing there in the presence of the grey-white building of his future and the thoughts of his past, **Christos Gogou** takes another minute to envision that, one day, his own legacy will be written in the stones of Building 58. A strong candidate for such purpose is his famous phrase 'Dying must be annoying, but being dead must be fantástic,' by which **Christos Gogou**, little did he know, prophetically summarized the average moment of his coming four years.

This occasion of silent self-absorption gives us the chance to observe what would become the daily phenotype of **Christos Gogou**. Over his shoulder his ever-present sleek brown briefcase full of pretentious Latin quotes, which **Christos Gogou** will recite from memory if only the smallest of opportunities arises, likely in order to compensate for his non-Spartan appearance. Around his neck his lucky Roman coin, which presents an interesting paradox within the phenomenon **Christos Gogou**, because **Christos Gogou** doesn't believe in the Gods and their interventions. If anything, the Gods believe in him. **Christos Gogou** believes in the structures of mind and brain, **Christos Gogou** is faithful to Logos, which will be the name of his first son once **Christos Gogou** finds his way through the labyrinth of the woman's heart, and, thus through an unresolved bit of mystery, **Christos Gogou** believes in his lucky Roman coin. Noteworthy is the colour of the rope of this necklace, namely brown, just like the colour of the jacket **Christos Gogou** is always wearing, which is a few shades lighter brown than the brown of his hair and beard, but perfectly matched with the brown of his Timberlands, always the brown Tim-

berlands, always and everything brown, can we see in his Timberlands and the colour brown traces of the eternal as obviously present as in the pretoogjes of **Christos Gogou**, whose pupils are framed with irises that are, indeed, brown.

Echoing characteristics of the final faith befallen to his youth friend Nietzsche, it slowly came to the awareness of **Christos Gogou** that parallel to his journey through the nanocosmos, a voyage of peaks and valleys through his own psyche manifested, one that **Christos Gogou** was luckily enough — perhaps through the providence of the piece of metal around his neck — well-adapted to travel, being well-versed in the maps of Carl Jung's archetypes, Jordan Peterson's phenomena, Marcus Aurelius' life advises, Dostoyevsky's suffering, and the psychological struggles of the whole lineage of Roman emperors, as well as being a seasoned expert by experience when it comes to the radical phenomenology of the crisis.

Then, is it here that the plot of this odyssey is starting to weave together? Was Building 58 the destiny of **Christos Gogou** in order to, once and for all, through the scientific metá-hodós of elucidating infinite detail after detail after detail to conveniently bypass the daunting mystery of the whole by death staring of its fragments, enforce synthetic Order upon the unbearable Chaos of the psyche? And was there a better place for **Christos Gogou** to fulfil this faith than in the structural nuts and bolts of the neuron, the cold but controllable substrate which the priests of science have hubristically dubbed the dwelling place of the uncontrollable internal?

Indeed, it turned out that **Christos Gogou** was destined to disentangle the elements of the nanoworld, the cosmos at her smallest; the mythologists largely grope in the dark to whether this faith holds any causal relation to the minuteness of the physical dimensions of **Christos Gogou** himself. While laboriously sculpting a brain at the nanoscale, the malevolent and highly contagious laughter of **Christos Gogou**, when his little structures behave according to the spitting image of his own brain physiology, indicated the fall in love with his creations. It goes well with **Christos Gogou** when it goes well with his research, **Christos Gogou** refines when his structures refine, and vice versa. On refined days, **Christos Gogou** runs material and mental marathons, and is the texture of his dreams determined by the conformations of his structural darlings when **Christos Gogou** moves from the waking to the sleeping realm where, as the mythologists feel comfortable speculating, **Christos Gogou** enacts the dress rehearsals for his alleged Frankensteinian ambitions to design the nanoarchitecture of the first Homo cogniticus. On coarse days, one finds **Christos Gogou** nights and weekends—'if you are more than God, you don't have to rest on Sundays,' another favourite in the aphoristic repertoire of **Christos Gogou**—sailing over the misty waters of his internal Styx into cloning hell when his plasmids decided, once again - days in a row, weeks in row, months in a row, not a thousand Gibson reactions could lighten up to misery of **Christos Gogou** - to speak Babylonian languages.

Prometheus and Dimphna, staring and sighing, both as impressed as concerned, wondering why they were ever convinced to give the TU Delftian fire to **Christos Gogou**, the semantic wonderchild, his tongue and pen verbal sledgehammers of which many ask what it exactly is the Gogouan prose strikes at. Chiselling with chirurgical precision and clinically logical grammatical structures — because poetry is strictly forbidden, **Christos Gogou** doesn't believe in poetry, although this self-proclaimed detail has been recently

challenged by new archaeological findings, having revealed the occurrence of the suspiciously metaphorical ‘conformation acrobatics’ (1x) and ‘orchestrate’ (5x) in Gogou et al., 2024 —, **Christos Gogou** bears his lengthy tomes that are destined to end in Nature Miscommunications or related journals.

Today, in the year 31 CG, the academic exodus of **Christos Gogou** comes to an end, begging insights into the morals of his story. Some mythologists have suggested that it was no coincidence that **Christos Gogou** aimed to reproduce a brain, to replicate the throne of the psyche, to control in vitro what is uncontrollable in the flesh. It remains to be seen what the myth of **Christos Gogou** will reveal in future blooming, but based on his warm personality, nurturing presence, professional achievements and depth of character, one can expect a lot. May his pretoogjes shine on thou and eternity.

(This statement has been humbly crafted by Sam van Beljouw, based on careful observation between the years 26 CG – 31 CG of a rewarding friendship with the legend **Christos Gogou**.)