Biochemical and structural studies of the voltage-gated sodium channels hNav1.8 and NavMs

Biochemical and Structural Studies of the Voltage-gated Sodium Channels hNav1.8 and NavMs

Thesis submitted by Callum Haste to Birkbeck, University of London, for the degree of Doctor of Philosophy (PhD)

May 2023

Institute of Structural and Molecular Biology,
Department of Biological Sciences

Birkbeck, University of London

Malet Street, London, WC1E 7HX

The candidate confirms that the work submitted is their own. Where work has been contributed to or data provided by others, the candidate confirms that appropriate credit has been given within the thesis with reference made to the work of others.

Acknowledgements and Dedications

The BBSRC LIDo Doctoral Training Programme funded this PhD position.

Firstly, I would like to thank Professor Bonnie Wallace for her unwavering support throughout my PhD. You have been understanding when I have had difficult times and been encouraging when I needed a push, and still managed to supervise us all with your full attention even when it could only be through a screen for so long. Thank you to Dr Altin Sula, who was an excellent teacher when I was still new to the sodium channel world and has become a friend since. I thank all the members of the Wallace group, with whom it has been a joy to work with these last few years and with whom I will endeavour to stay in touch after I am finished at Birkbeck. Thank you to Dr Claire Bagnéris, Dr Natasha Lukoyanova, and Dr Shu Chen, who have all been invaluable help with my experimental work, and whom I have learned so much from. Thank you to the LIDo programme, through which I have made so many friends that have made doing this PhD even more worthwhile. I am grateful to my thesis committee members Dr Anthony Roberts and Dr Amandine Maréchal for their helpful advice and encouraging words in our meetings.

Thank you to my parents, who have always been supportive of my path in life, and without whom I would not have been able to get a second chance at higher education and becoming a scientist like I had always wanted. Finally, thank you to Harriet; for making every day a joy, and giving me something to look forward to even when I'm stuck in the lab blocking western blots at 9 pm.

Abstract

Voltage-gated sodium channels (VGSCs or Navs), responsible for the action potential in excitable tissues, are drug targets for a variety of serious diseases linked to Navs, termed channelopathies; however, many of these drugs cause significant side effects due to their non-specificity between the nine distinct human isoforms (hNavs). The treatment of these diseases is complicated by the multiple states cycled through during the action potential, with open, closed, and inactivated channels adopting different conformations, and the fact that drugs can bind differentially to each state. Recent structural studies of eukaryotic and prokaryotic sodium channels (which share high sequence identities at key residues, plus similar structural features and drug binding affinities to eukaryotic Navs) have revealed a wealth of information about their function and interaction with various animal toxins and small molecule drugs, but hNavs remain challenging targets for structural study. The aim of this research was to use a novel expression system, *Tetrahymena thermophila*, to aid in expressing and purifying the peripheral nervous system sodium channel hNav1.8, associated with pain disorders, in order to study it by cryo-transmission electron microscopy (cryo-TEM) and gain insights into the hNav1.8 mechanism. This research marked the first time our lab was able to visualise hNav1.8 particles on a cryo-TEM grid without graphene oxide support, however, it was deemed unsuitable for obtaining high-resolution structural data due to the high level of degradation of the protein. The next aim of the research was to develop a method for the cryo-TEM analysis of the prokaryotic VGSC NavMs so that this useful model for hNavs could be studied closer to physiological conditions and ideally in unstudied conformations, which with the use of membrane scaffold protein (MSP) nanodiscs was ultimately a success, leading to a 3.07 Å reconstruction and a model with features distinct from previously solved NavMs structures.

Table of Contents

List of Figures

List of Abbreviations

AG	Activation gate
BacNav	Bacterial voltage-gated sodium channel
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
CCD	Charge-coupled device
CHAPS	(3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate)
CHS	Cholesteryl hemisuccinate
CNS	Central nervous system
Cryo-TEM	Cryo-transmission electron microscopy
CTD	C-terminal domain
DDM	N-dodecyl-β-D-maltopyranoside
DNA	Deoxyribonucleic acid
DRG	Dorsal root ganglion
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EM	Electon microscopy
EMDB	Electron microscopy data bank
FPLC	Fast protein liquid chromatography
GDN	Glyco-diosgenin
GMT	Gating modifier toxin
GSFSC	Gold-standard Fourier shell correlation
HEGA-10	Decanoyl-N-hydroxyethylglucamide
hNav	Human voltage-gated sodium channel
HWTX-IV	Huwentoxin-IV
IC_{50}	Half-maximal inhibitory concentration
IG	Inactivation gate
IPTG	Isopropyl-β-D-thiogalactopyranoside
LB	Lysogeny broth
LC-MS	Liquid chromatography-mass spectrometry
LDS	Lauryl dodecyl sulfate
MALDI-TOF	Matrix-assisted laser desorption/ionisation-time-of-flight
MOPS	3-((3-cholamidopropyl)dimethylammonio)-1-propanesulfonate
MSP	Membrane scaffold protein
Nav	Voltage-gated sodium channel
NCC	Normalised cross-correlation
PBS	Phosphate-buffered saline
PD	Pore domain
PDB	Protein data bank
pI	Isoelectric point
P-loop	Pore loop
PMSF	Phenylmethylsulfonyl fluoride

PNS	Peripheral nervous system
POPC	1-palmitoyl-2-oleoyl-glycero-3-phosphocholine
PVDF	Polyvinylidene difluoride
RMSD	Root mean square deviation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SF	Selectivity filter
SNR	Signal-to-noise ratio
SOC	Super optimal broth with catabolite repression
STX	Saxitoxin
TEM	Transmission electron microscopy
TEV	Tobacco etch virus
TTX	Tetrodotoxin
UA	Uranyl acetate
UV	Ultraviolet
VGSC	Voltage-gated sodium channel
VSD	Voltage-sensor domain
WT	Wild type

List of Tables

Chapter 1

Literature Review

1.1. Introduction to voltage-gated sodium channels

1.1.1. Overview of voltage-gated sodium channels

Voltage-gated sodium channels, commonly abbreviated as VGSCs or Navs, are integral membrane proteins present in all animal life, as well as some unicellular eukaryotes and prokaryotes (Moran *et al.*, 2015; Helliwell *et al.*, 2020). Part of the voltage-gated ion channel superfamily along with voltage-gated potassium and calcium channels (Yu *et al.*, 2005), Navs form channels in cell membranes that allow sodium ion flux in response to changes in voltage across the membrane, and the sequential opening of these channels is responsible for forming the action potential that propagates signals along excitable tissues such as cardiac muscle or neurons (Hodgkin and Huxley, 1952; Payandeh *et al.*, 2011). Many processes fundamental to animal life rely on this process, such as avoiding injury. As an example, the action potential is required to send a signal of a pain stimulus from the peripheral nervous system (PNS) caused by touching something hot, after which action potentials are needed to cause skeletal muscle contraction to move away from the source of pain. Action potentials in cardiac muscle are necessary to coordinate heartbeats and send oxygenated blood to the muscles; all the while, action potentials are firing in the central nervous system (CNS), processing this sensory information (de Lera Ruiz and Kraus, 2015). Due to the importance of sodium channel function in animal life, it is perhaps unsurprising that many species have evolved toxin molecules which target Navs, both for defence and predation, whilst genetic mutations in Nav channels are known to be responsible for a range of serious diseases, termed channelopathies (Graves and Hanna, 2005; Kalia *et al.*, 2015).

1.1.2. General structure and function of Navs

Nine Nav isoforms have been identified in mammals, which in humans are termed hNav1.1-1.9. These isoforms are differentially expressed around the body: hNav1.1, hNav1.2, and hNav1.3 are predominantly expressed in CNS neurons; hNav1.4 in skeletal muscle; hNav1.5 in cardiac and smooth muscle; hNav1.6 in both the CNS and peripheral nervous system (PNS); and hNav1.7, hNav1.8, and hNav1.9 in the dorsal root ganglia

(DRG) neurons of the PNS (Catterall, Goldin and Waxman, 2005), all of which have been identified as linked with channelopathy diseases (see Table 1.1).

Table 1.1. Table of human voltage-gated sodium channel isoforms. Human voltage-gated sodium channel (Nav) protein and gene nomenclature, expression localisation, channelopathies, and classification by tetrodotoxin resistance. All localisation and disease data taken from Catterall, Goldin and Waxman, 2005 unless otherwise specified. Tetrodotoxin (TTX) resistance information taken from Bennett *et al*., 2019. TTX-sensitive indicates half-maximal inhibitory concentration (IC_{50}) <30 nM.

Isoform	Gene	Localisation	Associated diseases (Channelopathies)	TTX sensitive
hNav1.1	SCN1A	CNS neurons, cardiac muscle	Dravet syndrome, febrile epilepsy, GEFS+ (Catterall, Goldin and Waxman, 2005)	Yes
hNav1.2	SCN2A	CNS neurons	Inherited febrile seizures, epilepsy	Yes
hNav1.3	SCN3A	CNS neurons	Epilepsy, pain disorders, brain development disorders (Smith et al., 2018)	Yes
hNav1.4	SCN4A	Skeletal muscle	Hyperkalemic periodic paralysis, paramyotonia congenita, potassium-aggravated myotonias	Yes
hNav1.5	SCN5A	Cardiac muscle, smooth muscle cells	Long QT syndrome, various arrythmias (Chockalingam and Wilde, 2012)	No
hNav1.6	SCN8A	CNS neurons, dorsal root ganglia, glial cells	Epilepsy, ataxia	Yes
hNav1.7	SCN9A	Dorsal root ganglia, Schwann cells, sympathetic neurons	Inherited erythromelalgia, paroxysmal extreme pain disorder (Eberhardt et al., 2015)	Yes
hNav1.8	SCN10A	Dorsal root ganglia	Small fibre neuropathy (Bennett et al., 2019)	No
hNav1.9	SCN11A	Dorsal root ganglia	Pain disorders (Bennett et al., 2019)	No

Each of these isoforms is a single pore-forming 220-260 kDa polypeptide pseudotetramer with 24 transmembrane helices, split into 4 highly homologous but non-identical domains (termed domains I-IV from N-to C-terminus) connected by linker regions, termed the α subunit. Each α subunit is associated with one or two 30-40 kDa β subunits termed β1-β4, of which β1 and β3 associated non-covalently and β2 and β4 associate covalently. These β subunits have a range of suggested functions, including modulation of channel gating, cell adhesion, interaction with the cytoskeleton and more (Hull and Isom, 2018), however the α subunit themselves are capable of functioning as voltage sensing, ion conducting channels without their respective β subunits.

Both eukaryotic and prokaryotic Navs are dynamic proteins which cycle through an open conducting state, a non-conducting inactivated state, and a closed resting state, to regulate the flow of Na^+ ions across a membrane (see Figure 1.1), however in eukaryotes two types

of inactivation are seen, slow and fast, whereas in prokaryotes only slow inactivation occurs (Bagnéris *et al.*, 2015). The cycling of states is triggered by changes in the potential difference across the membrane; at hyperpolarised membrane potentials, such as those of neuron cells, Navs are most likely to be in their resting state, but when the membrane depolarises, this initiates conformational changes in the α subunit that open or 'activate' the channel, allowing sodium ions to conduct through the channel pore down a concentration gradient and enter cells. In eukaryotic Navs fast inactivation may now occur, which involves the rapid (millisecond-scale) inactivation of the channel to prevent repeated opening of channels, entering a temporary refractory period. Slow inactivation occurs after extended depolarisation of the membrane, and occurs over a longer, second-scale time span (Vilin and Ruben, 2001). The membrane potential repolarises as potassium channels open and allow K^+ to pass out of the cell, before ATP-driven Na^+/K^+ pumps restore Na^+/K^+ gradient of the resting state (Glitsch, 2001).

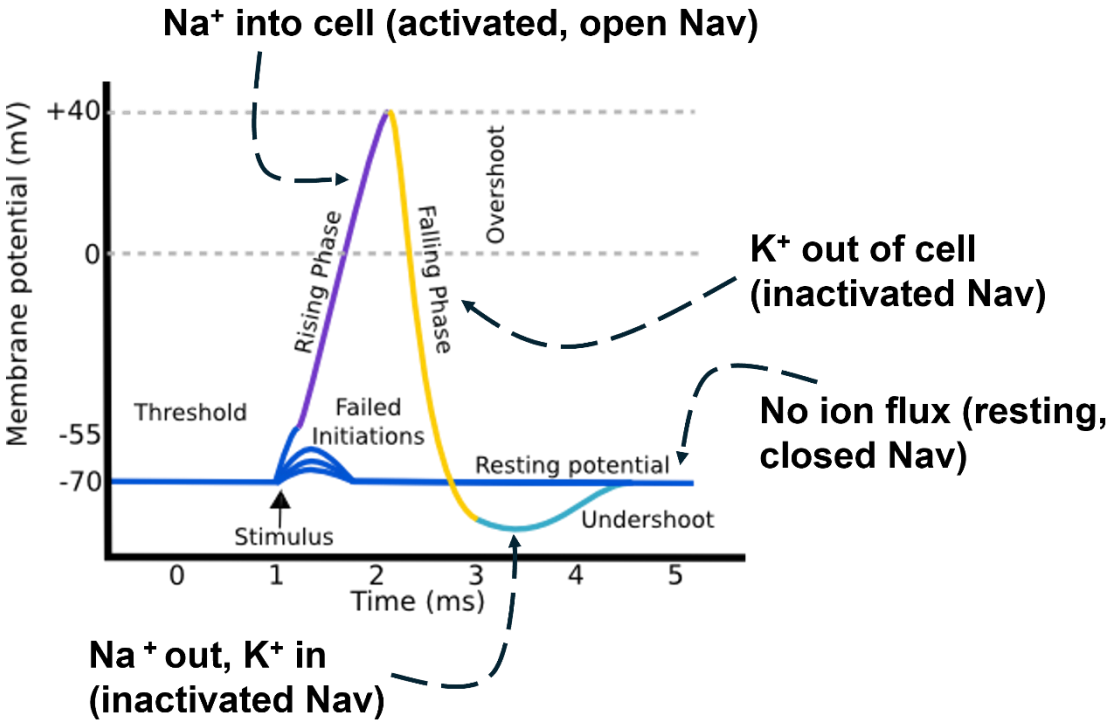

Figure 1.1. Graphical representation of an action potential. Diagram of an action potential, with changes in membrane potential and associated movement of ions across the membrane indicated. Nav = voltage-gated sodium channel.

The action potential and the cycling between states is made possible by the structure of the α subunit. The four highly homologous domains, termed DI to DIV, each contain 6 of the total 24 transmembrane helices, which are termed S1-S6 of each of their respective domains (Ahern *et al.*, 2016) (see Figure 1.2), a general architecture which is shared with

prokaryotic Navs, with the distinction that prokaryotic Navs are made up of four identical ~30 kDa tetramers rather than a single polypeptide, and therefore lack inter-domain loops. There are several distinct subdomains and structural motifs within each domain, which can be seen in Figures 1.2 and 1.3, and are as follows; the voltage-sensor domain (VSD), the pore domain (PD), the pore-loops (P-loops) which house the selectivity filter (SF), the intracellular gate (IG) which forms the tightest constriction of the pore, blocking ion conduction, the Ile-Phe-Met inactivation motif (IFM) on the intracellular DIII-DIV linker, extracellular loops above the pore domain, and a regulatory C-terminal domain (CTD). The VSD is formed of helices S1-S4, and, as the name suggests, is responsible for detecting membrane depolarisation, hence modulating the 'activation', or opening, of the channel. The PD, which is connected to the VSD by a membrane-parallel S4-S5 linker, is formed of S5 and S6 of each of the 4 domains and is responsible for forming the channel through which sodium ions can flow in the open state, with residues towards the intracellular, C-terminal end of S6 forming the IG. Both prokaryotic and eukaryotic Navs, along with other channels in the voltage-gated ion channel superfamily, show a 'domain-swapped' assembly, meaning that the VSD of one repeat/monomer forms an interface with the PD of another repeat/monomer (see Figure 1.2). A final structural feature in both prokaryotic and eukaryotic channels to note is the 'fenestrations', which are formed at interdomain interfaces; these were first theorised as an explanation for how hydrophobic molecules could block a sodium channel without having to enter through the aqueous environment of the extracellular pore (Hille, 1977). Structural study of Navs has shown that in certain conformational states there are up to 4 openings (1 between each domain) from the transmembrane region of a Nav channel through to the pore lumen (see Figure 1.3) and that drugs, such as cannabidiol, can indeed enter and block prokaryotic and eukaryotic pores by this pathway (Sait *et al.*, 2020; Huang *et al.*, 2023).

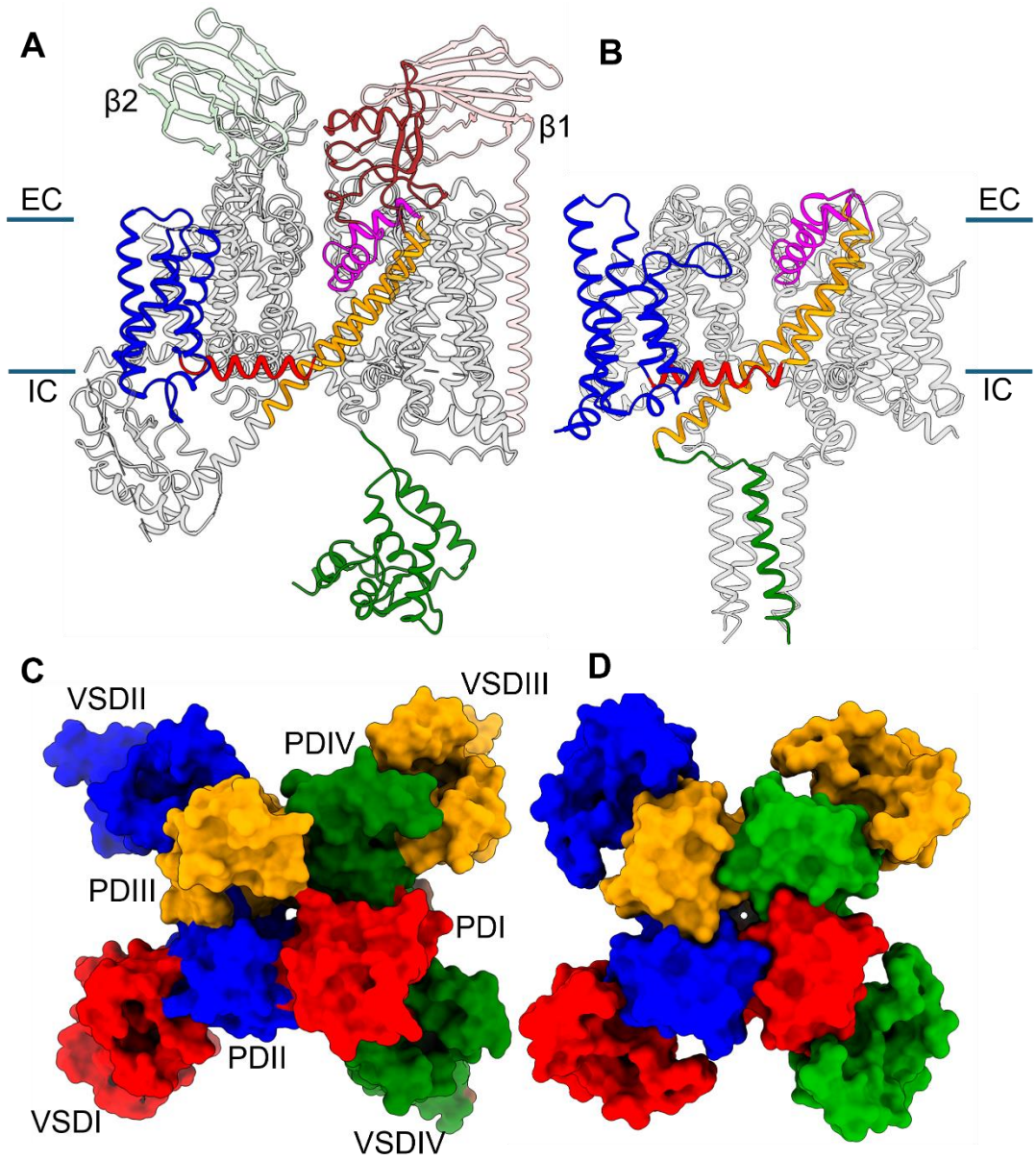

Figure 1.2. Voltage-gated sodium channel domain architecture. (A) shows the architecture of a eukaryotic voltage-gated sodium channel (Nav) domain (hNav1.7, PDB 7W9K (G. Huang *et al.*, 2022)), with subunits β1 and β2 labelled. **(B)** shows equivalent regions, if present, in a prokaryotic Nav (NavMs, PDB 6SX5, Sula *et al.,* 2021). Blue = voltage sensor domain (VSD) helices S1-S4, red = S4-S5 linker helix, orange = pore domain (PD) helices S5 and S6, magenta = selectivity filter (SF), brown = extracellular loops, green = C-terminal domain (CTD). Extracellular (EC) and intracellular (IC) regions marked with blue lines and labelled. **(C)** and **(D)** show extracellular views of hNav1.7 and NavMs respectively, with each repeat of hNav1.7 (DI-DIV) and monomer of NavMs coloured to show the domain-swapped, with the VSD of one domain associated with the PD of its neighbour. Figure made in ChimeraX (Pettersen *et al.,* 2021).

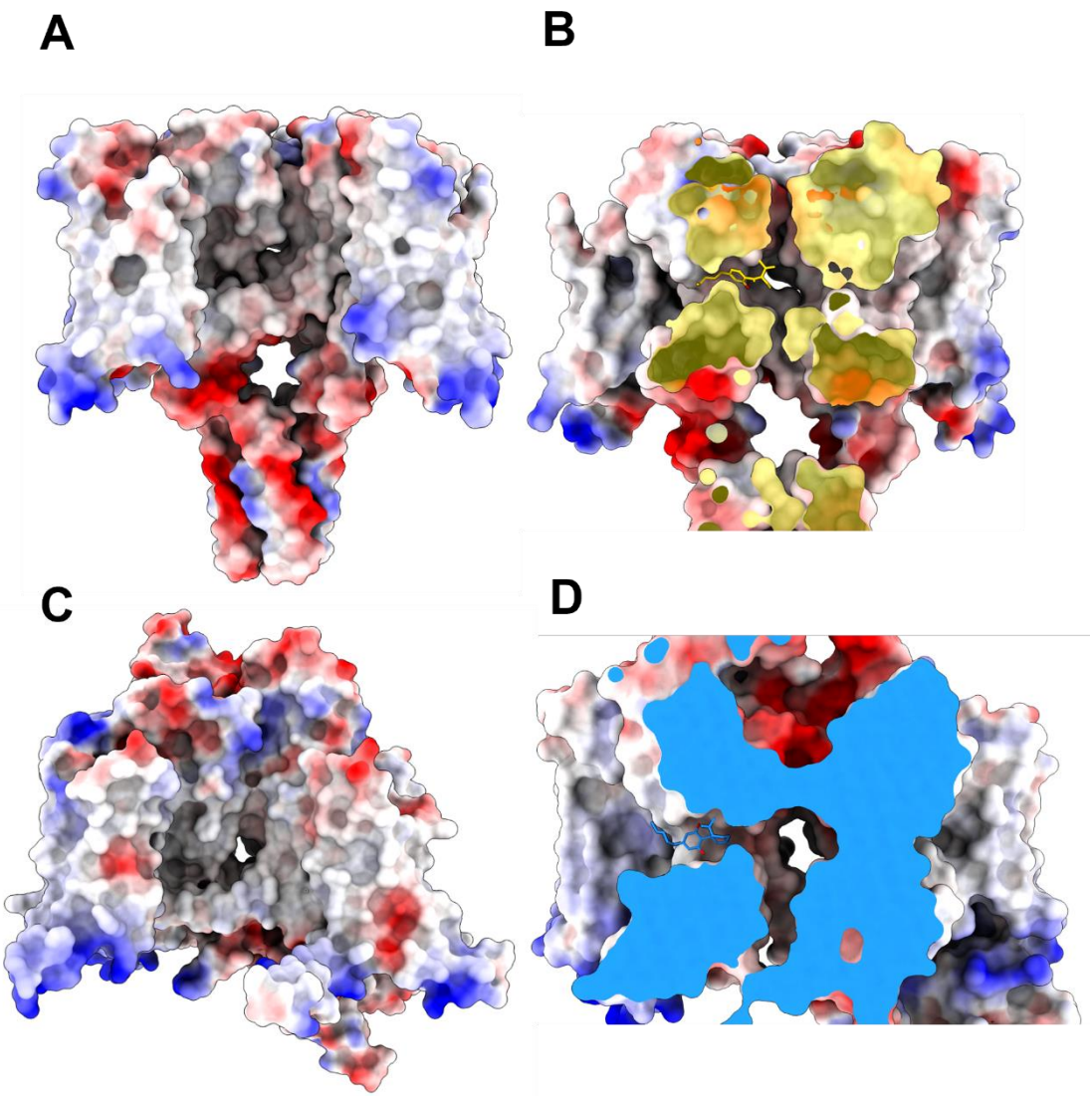

Figure 1.3. Fenestrations as entry points for hydrophobic molecules. (A) and **(B)** show transmembrane views of NavMs (PDB 6YZ0, Sait *et al.,* 2020), with surfaces coloured by electrostatics, white being neutral, blue positively charged, and red negatively charged. **(B)** shows a sliced view, showing the hydrophobic molecule cannabidiol in the fenestration. **(C)** and **(D)** show the same views and colourings of hNav1.7, also with cannabidiol in the fenestration (PDB 8G1A, Huang *et al.,* 2023). Figure made in ChimeraX (Pettersen *et al.,* 2021).

1.2. The structural basis for Nav function

The first high-resolution structural data of full-length Navs began with X-ray crystal structures of the prokaryotic channel NavAb of *Arcobacter butzleri* (Payandeh *et al.*, 2011), and was followed by others such as NavRh of *Rickettsiales sp. HIMB114* (PDB 4DXW, (Zhang *et al.*, 2012)) and NavMs of *Magnetococcus marinus* (Sula *et al.*, 2017). This was later followed by cryo-transmission electron microscopy structures of non-human Navs, the first being Nav1.4 from the electric eel *Electrophorus electricus* (Yan *et al.*, 2017) then NavPas, a sodium channel from the cockroach *Periplaneta americana* (Shen *et al.*, 2018). The first high-resolution structure of a human Nav was released soon after, of hNav1.4 in complex with the β1 subunit (Pan *et al.*, 2018). This structural data gave insight into the conformational changes which accompany the different states of sodium channels, along with the structural basis for important aspects of Nav function, such as selectivity for sodium ions at the SF, interactions in the VSD that support the 'sliding helix' model of channel activation (DeCaen *et al.*, 2008, 2009), the formation of the pore constriction at the IG, and, in eukaryotic Navs, the process of fast inactivation by the IFM motif.

1.2.1 The selectivity filter

The SF forms a constriction at the extracellular end of the pore which confers selectivity for sodium ions as larger monovalent cations, such as potassium, are too large to fit through, and the higher charge density of divalent cations such as magnesium makes their passage energetically unfavourable (Naylor *et al.*, 2016). In NavMs, a symmetric SF is formed of residues Thr176, Leu177, Glu178, and Ser179 of each monomer, are able to coordinate and dehydrate a hydrated sodium ion for passage through the pore. These residues were shown to form three sites for sodium ions in a pore-only construct of NavMs (Naylor *et al.*, 2016), which was later corroborated by the full-length crystal structure (Sula *et al.*, 2017), with Glu178 forming the tightest constriction. This symmetric SF and 'EEEE' motif is shared amongst prokaryotic Navs (Dudev and Lim, 2014). Mutational studies had suggested that an equivalent motif, the 'DEKA' motif (made up of an aspartate of DI, a glutamate of DII, a lysine of DIII and an alanine of DIV), conserved amongst eukaryotic Navs, was responsible for conferring selectivity in eukaryotes (Favre, Moczydlowski and Schild, 1996). The cryo-TEM structure of NavPas (Shen *et al.*, 2018) shed further light on this, showing a sodium ion coordinated

asymmetrically by a 'DEE' motif (see Figure 1.4); the ion permeation path was predicted in the same paper by molecular dynamics studies, suggesting that the sodium ion travels through an asymmetric path, with the lysine of DEKA likely making the passage of larger or more positively charged ions energetically unfavourable.

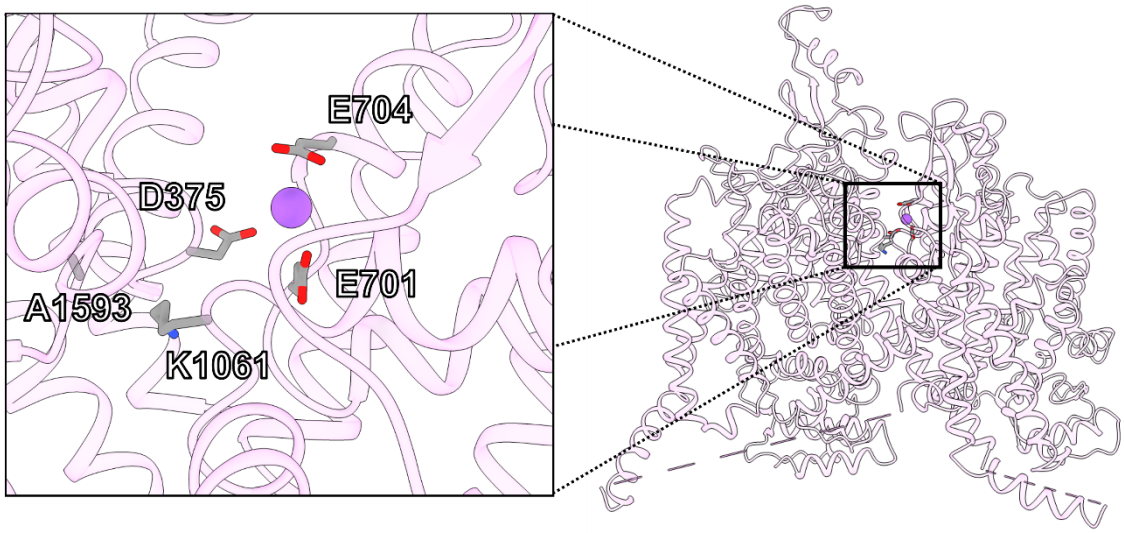

Figure 1.4. The selectivity filter of a eukaryotic voltage-gated sodium channel. The left figure shows the selectivity filter of the eukaryotic voltage-gated sodium channel NavPas (PDB 6A90, Shen *et al.*, 2018), with the 'DEKA' and 'DEE' motifs labelled, and the coordinated sodium ion shown as a purple sphere. Right shows its position at the extracellular end of the channel pore. Figure made in ChimeraX (Pettersen *et al.,* 2021).

1.2.2. The sliding helix and the intracellular gate

To allow sodium ions across a membrane during an action potential, Nav channels must undergo conformational changes to form a conductive pore; the source of this conformational change was theorised as the 'sliding helix' model. Studies involving the mutation of charged residues in the VSD to cystines so that they became crosslinked showed sequential ion pair formation between arginines of the S4 helix and negatively charged residues on S2 and S3, aiding the movement of S4 through the membrane (DeCaen *et al.*, 2008, 2009) and catalysing channel activation. This was again confirmed by structural studies of sodium channels; the Catterall group used disulfide crosslinking in the VSD to hold a NavAb channel in the activated state and obtain a crystal structure (see Figure 1.5). Alongside this, they mutated multiple residues in another construct to modify the voltage of activation and disulfide-trap NavAb in a resting state, using a maltose-binding protein fusion tag to enlarge NavAb and make it more amenable to

analysis by cryo-TEM, as the extensive mutations in this construct were judged to cause it to be unable to crystallise (Wisedchaisri *et al.*, 2019).

Comparing the two structures, in the activated state the gating charges R1-R3 of S4 are above a conserved occluding hydrophobic constriction site in the VSD (Phe56 in NavAb) and R4 is below, but in the resting state S4 has moved downward through the membrane and R3 is now beneath Phe56. This is coupled with a tightening of the S4-S5 linkers, which in turn constricts the pore helices, leading to a narrower intracellular gate (a hydrophobic constriction in the pore of Nav channels, which in an inactivated or closed state is too narrow for sodium ions to pass through (>2.3 Å radius) (Oelstrom *et al.*, 2014; Sula *et al.*, 2017)), formed by hydrophobic residues on S6. The same constriction at the intracellular gate can be seen in hNavs, where it is a conserved DI leucine, D2 phenylalanine, DIII isoleucine (valine in hNav1.8) and DIV tyrosine (G. Huang *et al.*, 2022), shown in Figure 1.6. The IG is a hydrophobic constriction in the pore of Nav channels, which in an inactivated or closed state is too narrow for sodium ions to pass through (>2.3 Å radius) (Oelstrom *et al.*, 2014; Sula *et al.*, 2017).

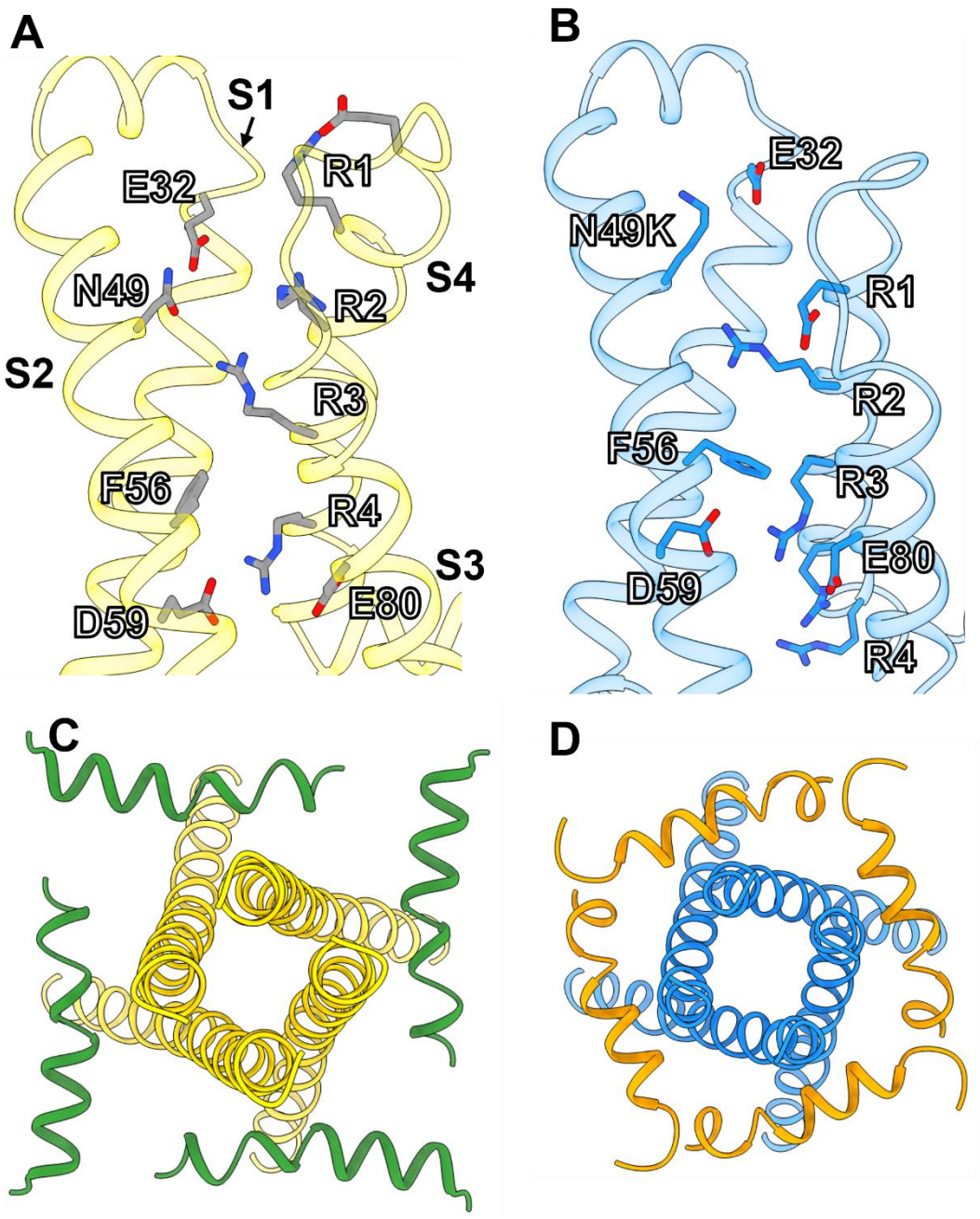

Figure 1.5. The sliding helix model of channel activation. (A) and **(B)** show the voltage sensors of NavAb mutated to be disulfide-locked in the activated and resting states, respectively. Conserved gating charge residues on S4 are labelled R1-R4, and the occluding hydrophobic constriction residue Phe56 is highlighted. **(C)** and **(D)** the effect on the pore of the position of the gating charges in the activated and resting conformations, respectively, from an intracellular view. S4-S5 helices are shown in green in **(C)** and orange in **(D)**, whilst S6 helices are shown in yellow in **(C)** and blue in **(D)**. Figure made in ChimeraX (Pettersen *et al.,* 2021).

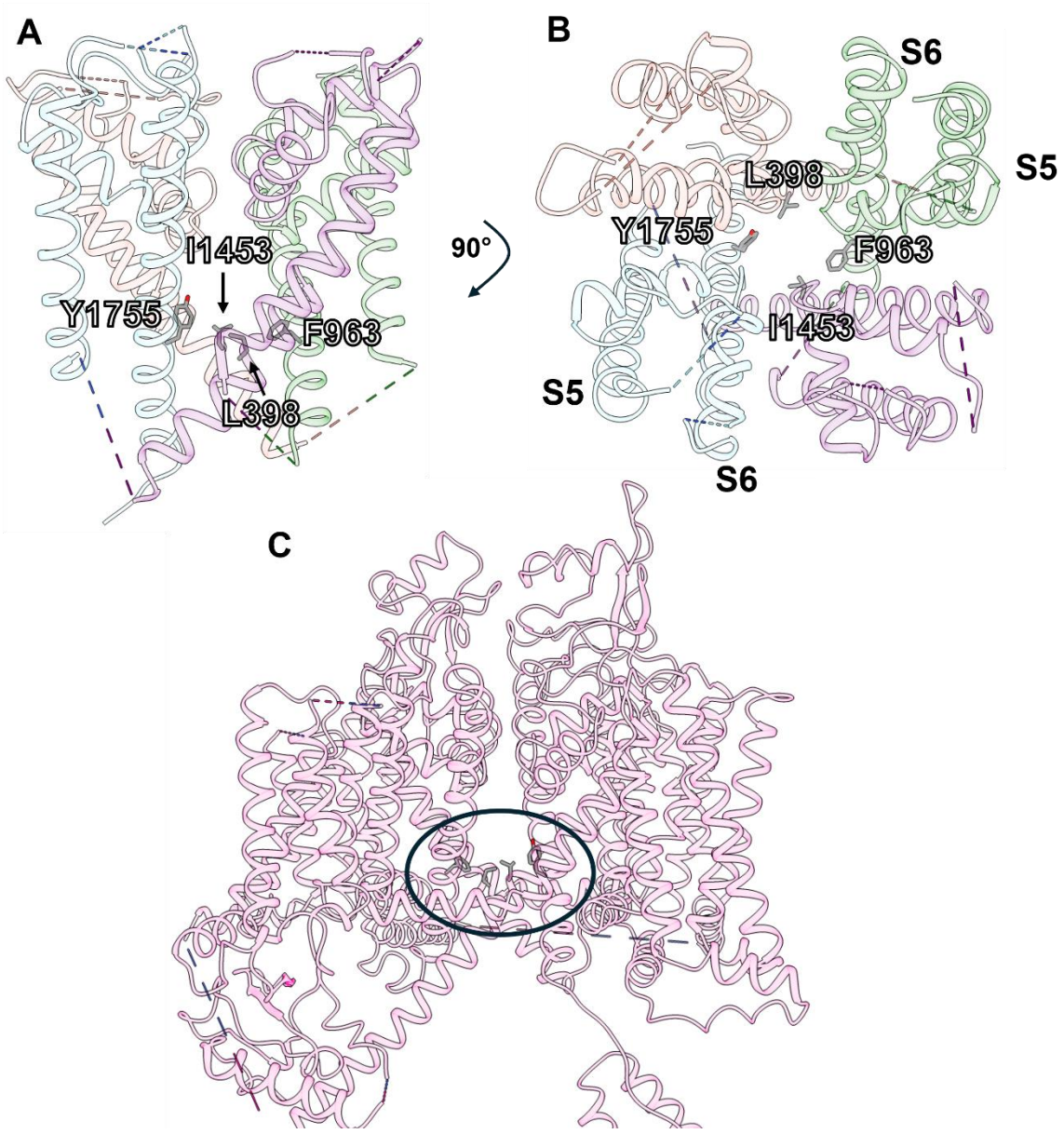

Figure 1.6. The intracellular gate of the human voltage-gated sodium channel hNav1.7. (A) and **(B)** show cross-membrane and extracellular views respectively of the intracellular gate (IG) of hNav1.7 (PDB 7W9K, G. Huang *et al*., 2022), with domain DI coloured salmon, DII coloured green, DIII coloured purple and DIV coloured blue. The IG is a hydrophobic constriction site at the extracellular end of the pore S6 helices which, in inactivated and resting channels, blocks the conduction of sodium ions through the channel. The location of the IG in the channel is shown in **(C)**. Figure made in ChimeraX (Pettersen *et al.,* 2021).

1.2.3. Fast inactivation by the IFM motif

As mentioned previously, eukaryotic Navs (but not prokaryotic) are capable of undergoing 'fast inactivation'. Within a few milliseconds after the channel has opened and the resulting sodium ion flux has depolarised the membrane potential, the channel becomes non-conductive, but without immediately returning to the resting state, as the membrane potential has not yet been repolarised to resting potential. Mutagenesis studies indicated that residues on the DIII-IV linker were responsible (McPhee *et al.*, 1994), which was confirmed when cryo-TEM structures of eukaryotic Navs became available (Pan *et al.*, 2018). This showed that three residues on the DIII-IV linker, the Ile-Phe-Met (IFM) motif, were responsible. The voltage sensor remains in its activated position, but the IFM motif inserts into a hydrophobic cavity formed by the S4-S5 linker and S6 helices of DIII and DIV, moving S6 so that the pore is constricted (see Figure 1.7).

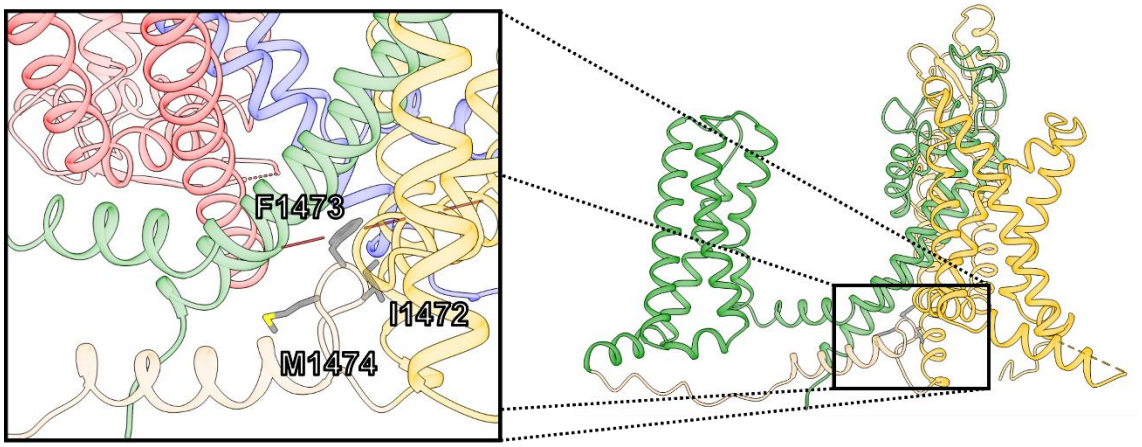

Figure 1.7. The IFM motif responsible for fast inactivation. Left: showing residues Ile1472, Phe1473, and Met1474, which form the IFM motif of the DIII-IV linker of hNavs. DI shown in red, DII in blue, DIII in yellow, DIII-IV linker in beige and DIV in green. The insertion of the IFM motif into a hydrophobic cavity between S4-5 and S6 of DIII and DIV, allosterically constricting the channel. **Right:** the position of the IFM, with only DIII, DIV, and the DIII-IV linker shown for clarity. Figure made in ChimeraX (Pettersen *et al.,* 2021).

The transmembrane regions of prokaryotic channels share approximately 30% sequence identity with hNavs as well as very similar folding and conserved amino acids at key functional sites (Sula and Wallace, 2017). In fact, NavMs, a voltage-gated sodium channel found in the extremophile bacterium *Magnetococcus marinus*, is structurally close enough to hNavs to show similar drug binding, selectivity of ion translocation, and conductance (Ulmschneider *et al.*, 2013; Bagnéris *et al.*, 2014), leading to its use as a functional model for human channels.

1.3. Navs as Drug Targets

As seen in Table 1.1, each of the 9 hNavs has a range of associated diseases. These diseases, or channelopathies, occur due to a wide variety of mutations which affect the function of the channel, for example gain-of-function mutations in Navs involved in nociception (hNav1.7-1.9) leading to chronic pain disorders, or loss-of-function meaning the afflicted person cannot feel pain (Huang *et al.*, 2017). Unsurprisingly from the function of Navs and their links to such a large number of serious and debilitating diseases, they have long been explored as drug targets for epilepsy, cardiac arrythmias, chronic pain, and as a suggested mechanism for anaesthesia (Catterall, 1987; Bagal *et al.*, 2015). Unfortunately, the search for effective therapeutics is hampered by the fact that all 9 subtypes of Navs share similar sequences (see Table 1.2) and structures, meaning that drugs must be highly selective to avoid harmful side effects by action on channels expressed in non-target tissues. This is complicated further by differential drug binding depending on the state of the channel; for example, tamoxifen, a commonly prescribed oestrogen receptor modulator and anti-cancer drug, has been shown to bind to the inactivated NavMs channel and block current by delaying recovery (Sula *et al.*, 2021). These two problems mean that high-resolution structures of each Nav in multiple

conformations are required to design subtype-specific drugs for the greatest medical benefit and understanding of their mechanism of action.

Table 1.2. Percentage identity matrix of hNavs, NavMs, and NavAb. This table shows a percentage identity matrix of human voltage-gated sodium channel isoforms and prokaryotic NavMs and NavAb. The high level of sequence identity between isoforms means that non-selective drugs have a high risk of off-target side effects. Prokaryotic Navs, e.g., NavMs, are similar enough to human Navs to replicate drug binding in transmembrane regions. Table created using Clustal Omega (Sievers *et al*., 2011).

Nav	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	Ms	Ab
1.1	-	89	86	78	71	80	79	65	61	36	39
1.2	89	-	89	79	72	80	81	66	62	37	39
1.3	86	88	-	77	72	79	79	66	62	36	38
1.4	78	79	77	-	74	76	76	68	62	36	38
1.5	71	72	72	74	-	71	70	72	63	37	38
1.6	80	80	79	76	71	-	75	66	61	36	38
1.7	79	81	79	76	70	75	-	65	62	37	38
1.8	65	66	66	68	72	66	65	-	63	34	37
1.9	61	62	62	62	63	61	62	63	-	33	35
Ms	36	37	36	36	37	36	37	34	33	-	53
Ab	39	39	38	38	38	38	38	37	35	53	-

Small molecules that act on sodium channels, such as lidocaine, have been used for the better part of a century (Cummins, 2007) but only in recent decades has the pharmacological profile of these drugs and their mode of action been empirically explored. The main classes of medicinal drugs which act on sodium channels include anaesthetics such as the aforementioned lidocaine, antiarrhythmics such as flecainide, and anticonvulsants and antiepileptics such as phenytoin and carbamazepine (see Figure 1.8); these so called ‘first-generation’ sodium channel modulators share similar chemical properties, and mutagenesis studies on Nav1.2 showed that they share a common binding site in the highly conserved S6 transmembrane helix (Ragsdale *et al.*, 1996; Bagal *et al.*, 2015), generally named the local anaesthetic binding site. As expected these drugs are subtype nonselective and lead to a range of side effects, however subtype-unselective drugs such as these can still be therapeutically useful due to the concept of ‘use-dependent’ inhibition (Starmer, Grant and Strauss, 1984). This essentially means that as the molecule is more able to enter the channel in an open state and then forms its tightest binding site in the inactivated state, thus preventing the channel from opening again (Quan, Mok and Wang, 1996). The small molecule will preferentially bind to channels

which are open more frequently (those which are misfiring and producing unwanted action potentials) over more frequently closed, normally functioning channels.

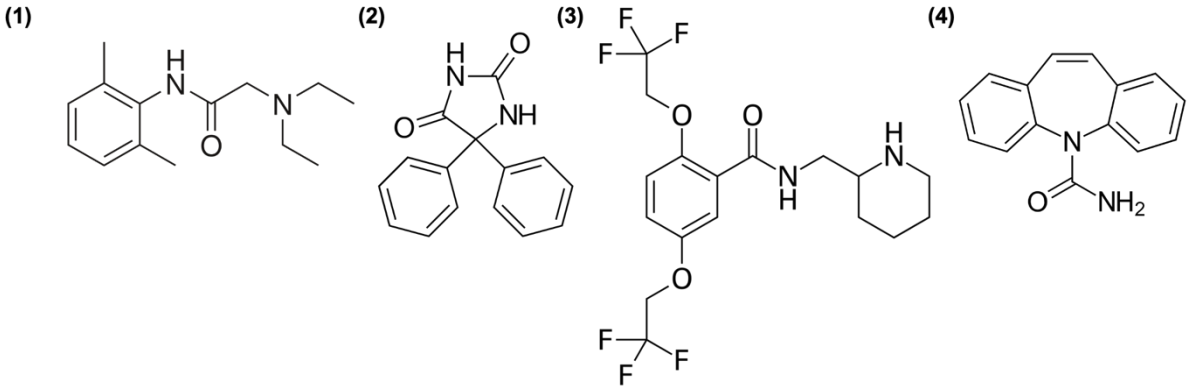

Figure 1.8. First-generation sodium channel modulators. A selection of first-generation sodium channel modulators. (1) Local anaesthetic lidocaine (2) anticonvulsant phenytoin (3) antiarrhythmic flecainide (4) anticonvulsant carbamazepine. Each drug shares common functional groups e.g., amides and arenes, and the same binding site in hNav1.2 shown by site-directed mutagenesis (Ragsdale *et al*., 1996).

What these first-generation drugs also have in common is that they were all in medical use for many years before being characterised and having their pharmacology explored in terms of sodium channel modulation; research has since moved on to attempts to discover molecules which are expected to have subtype-specific activity, or 'second-generation' sodium channel modulators. Due to the high prevalence of chronic pain in developed nations (Breivik *et al.*, 2006), and known mutations associated with loss or gain of pain (Goldberg *et al.*, 2007; Xiao *et al.*, 2019) hNav1.7 and 1.8 in particular have been consistent targets for drug companies looking to produce a subtype-selective molecule with reduced CNS and cardiac side-effects. This has led to trials of more complex molecules such as the aryl sulfonamide PF-05089771 being developed, with additional moieties intended to confer subtype selectivity, however only modest efficacy has been shown and second-generation sodium channel modulators have had difficulty moving past phase 2 clinical trials (McDonnell *et al.*, 2018). Despite the difficulty, the quest for subtype selective drugs is being aided by the wealth of structural information now available of hNavs, many structures available in apo state and with drugs or toxins bound, for example; hNav 1.1 (Pan *et al.*, 2021), 1.2 (Pan *et al.*, 2019), 1.3 (Li *et al.*, 2022), 1.4 (Pan *et al.*, 2018), 1.5 (Li, Jin, Wu, Huang, *et al.*, 2021; Li, Jin, Wu, Zhao, *et*

al., 2021), 1.6 (Fan *et al.*, 2023; Li *et al.*, 2023), 1.7 (Shen *et al.*, 2019; G. Huang *et al.*, 2022; Zhang *et al.*, 2022), and 1.8 (X. Huang *et al.*, 2022). This has allowed structural mapping of drug and toxin binding sites to well-studied hNavs such as hNav1.7, which can now act as blueprint for future drug development (Wu *et al.*, 2023) (see Figure 1.). As well as structures of hNavs, prokaryotic Navs are still used as a useful model for hNavs, as they are easier to express, but still have many features in common; the transmembrane regions of prokaryotic channels share approximately 30% sequence identity with hNavs as well as very similar folding and conserved amino acids at key functional sites (Sula and Wallace, 2017). In fact, NavMs is structurally close enough to hNavs to show similar drug binding, selectivity of ion translocation, and conductance (Ulmschneider *et al.*, 2013; Bagnéris *et al.*, 2014). As well as drug molecules binding to sites already present on prokaryotic Navs (Gamal El-Din *et al.*, 2018; Sait *et al.*, 2020; Sula *et al.*, 2021), it is also possible to add humanisng mutations to prokaryotic Navs, for example in the VSD, to allow the study of hNav-specific drugs and toxins (Ahuja *et al.*, 2015; Gao *et al.*, 2020).

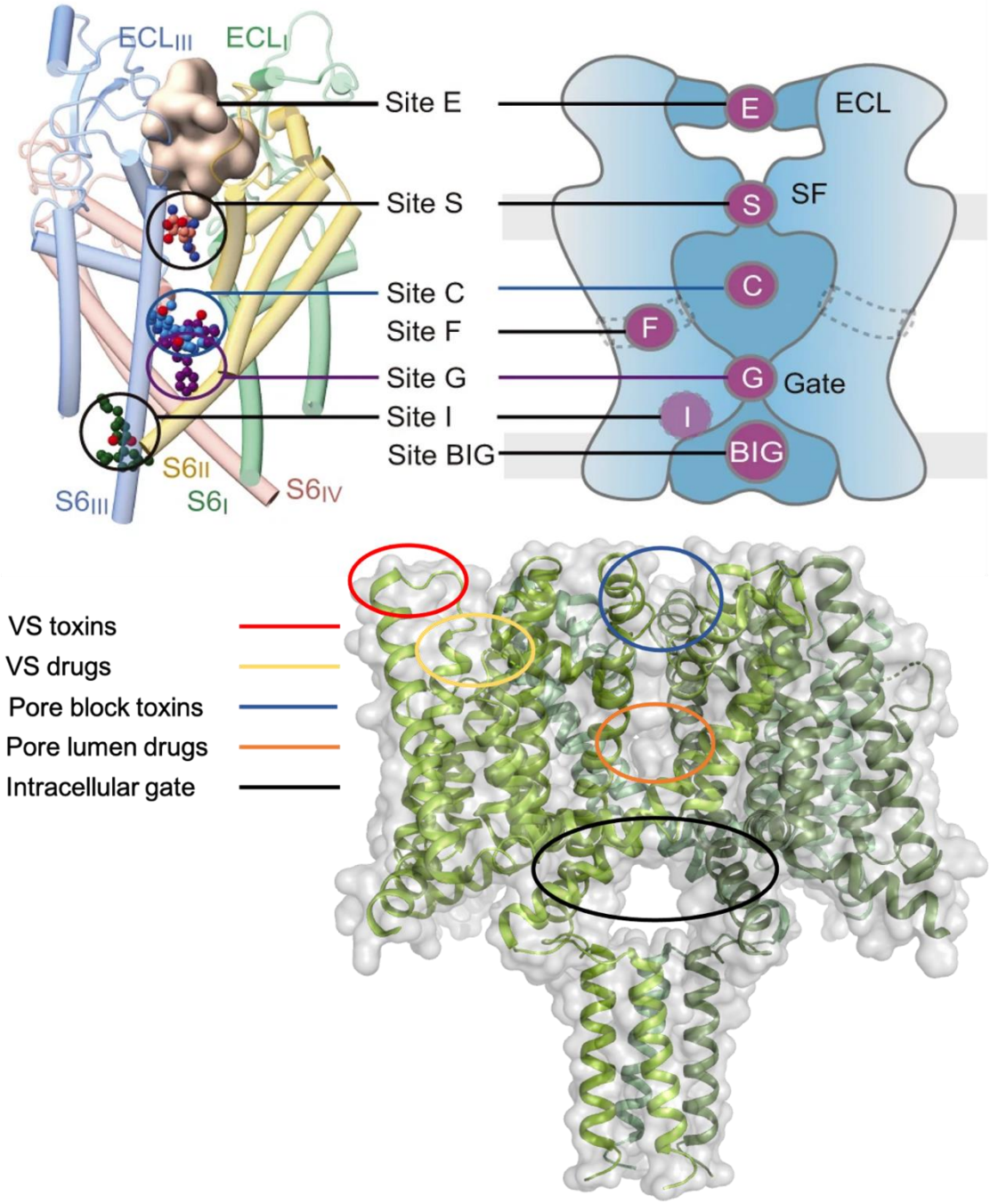

Figure 1.9. Mapping of drug and toxin sites to Nav channels. Top: Figure showing drug sites mapped to the hNav1.7 pore domain (PD), adapted from Wu *et al.,* 2023. **Top left** shows μ-conotoxin KIIIA (pale pink surface) in the extracellular loop (E) site, tetrodotoxin (TTX) in the selectivity filter (SF, Site S). Quinidine is in the pore lumen cavity (site C), propafenone in the intracellular gate (site G) and cannabidiol in the inactivation site (site I) **Right:** schematic of drug sites in the PD, with the fenestration, where molecules such as cannabidiol can also enter and bind, labelled as site F. **Bottom:** Crystal structure of NavMs (PDB 5HVX, Sula *et al.,* 2019) showing the sites of voltage sensor drug and toxin binding on prokaryotic-human Nav chimeras (Ahuja *et al.,* 2015; Gao *et al.,* 2020), pore blocking toxins on eukaryotic Navs (Pan *et al.,* 2019), pore blocking drugs (Gamal El-Din *et al.,* 2018) and tamoxifen and its metabolites below the intracellular gate (Sula *et al.,* 2021). Bottom figure made in Pymol (version 2.0; Schrödinger, LLC).

1.4. Insights from the structure of hNav1.8 in apo state and bound to A-803467

At the start of this project, hNav1.8 did not yet have a cryo-TEM structure, however this was published in 2022 in the apo state and in complex with the subtype-specific channel blocker A-803467, shown previously to reduce pain in rats (Jarvis *et al.*, 2007; X. Huang *et al.*, 2022) by binding to the resting and inactivated states of the channel, with preference for the inactivated state. As hNav1.8 is the area of interest of much of this thesis, this research will be discussed here in detail. hNav1.8 is distinguished from other hNav isoforms by its resistance to tetrodotoxin (along with hNavs 1.5 and 1.9); its higher voltage (i.e., less negative, more depolarised) of activation; and its greater persistent current, associated with slower inactivation kinetics than other channels, allowing hyperexcitability in dorsal root ganglion neurons (involved in pain signalling). This hyperexcitability in relation to pain has led to hNav1.8 being investigated as a target for non-opioid, and by extension non-addictive pain medication; without structural data however it was not possible to determine what enables these specific characteristics and therefore how to intelligently design a drug to ameliorate pain symptoms.

Table 1.3. Table of published hNav1.8 structures.

Structure	PDB code	Resolution (Å)	Residues modelled
Apo	7WFW	3.1	998 (51%)
A-803467 Class I	7WE4	2.7	1122 (57%)
A-803467 Class II	7WEL	3.2	1102 (56%)
A-803467 Class III	7WFR	3.0	1100 (56%)

The resultant apo state structure was resolved to 3.1 Å as a global average with 998 out of 1956 residues modelled (51%), whilst the A-803467 bound structures were split into 3 distinct classes (classes I, II and III, resolved to 2.7, 3.2 and 3.0 Å respectively and with 57%, 56%, and 56% of residues modelled respectively) due to differences in the conformation of voltage sensor domain I (VSDI). The modelled residues are mostly

confined to the transmembrane region of the protein, as is typical for hNav cryo-TEM structures published thus far (see Figure 1.10).

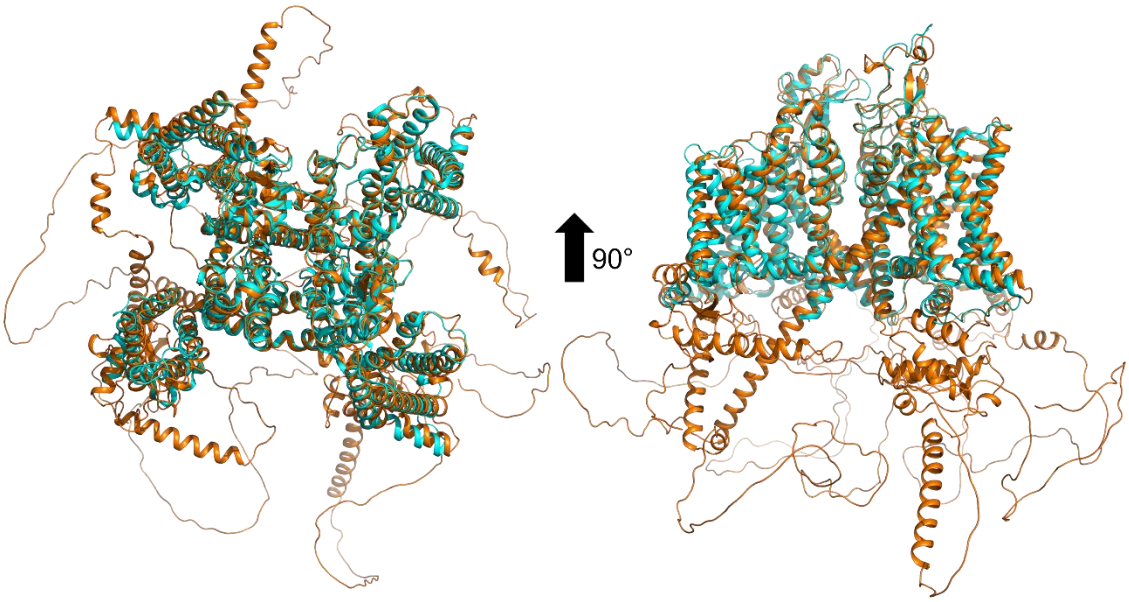

Figure 1.10. Representation of the extent of modelled residues in the published hNav1.8 structure. Figure showing a structural alignment of class I of hNav1.8 bound to A-803467 (cyan, PDB 7WE4) (X. Huang *et al*., 2022) with the Alphafold structural prediction of hNav1.8 (orange), with top (left) and side (right) views. The structure of class I has 1122 out of 1956 residues modelled (57%), which, as can be seen in the above figure are limited to the transmembrane region of hNav1.8, typical of the hNav cryo-TEM structures published to date. The cytosolic loop regions shown in the Alphafold model are designated as low or very low confidence and should not be interpreted as the true structure of hNav1.8. Figure made using PyMol molecular graphics software (version 2.0; Schrödinger, LLC).

As the models for the apo and class III structures did not have a high enough resolution in the variable VSDI region to confidently assign side chain positions, further analysis of the structures focused on only class I and II of the A-803467 bound structures. For each structure, VSDs II-IV are in identical 'up' (depolarised, activated) positions, with only VSDI differing between the structures; whilst the rest of the structure appears essentially identical, VSDI of class I and II vary by a root mean square deviation (RMSD) of 2.8 Å over 88 Cα atoms (see Figure 1.11, top). VSDI undergoes a 12° rotation around the central pore, pivoting S4 in the process. However, it appears that the whole voltage sensor does not move as a single rigid body, with each of the four helices moving out of concert. The gating charge residues R2 to R4 are above the occluding Tyr164 of S2 in both classes, which indicates an activated voltage sensor, and in class I the intradomain binding of VSDI is very similar to that of other published hNav structures, but the binding between residues on S2 and S4 of the voltage sensor differs in class II. Of particular note is that the occluding Tyr164 residue of S2 in class II is forming a π-cation interaction with gating charge residue R4, causing Tyr164 to swing downwards compared to class I, a conformation not seen in other hNav structures. Charge interactions are also reduced in class II, for example the anionic residues E157 and E167 of S2 are no longer in reach of R2 and K5 of S4 (see Figure 1.11, bottom). This shift of VSDI also reduces the interactions between the interfaces of the VSD and pore domain (PD) II; in class I, 7 residues of S4I interact with 8 on S5II of the PD, whereas in class II a kink in chain S4I means that only 4 residues on S4I are interacting with 5 on S5II. At the same time, S1I of class I interacts more closely with the PD, as the cytosolic end of S1I in class II swings outwards away from S5 and P1 of PDII; the publication suggests that this extracellular contact between S1I and the PDII is the pivot of where VSDI shifts during activation.

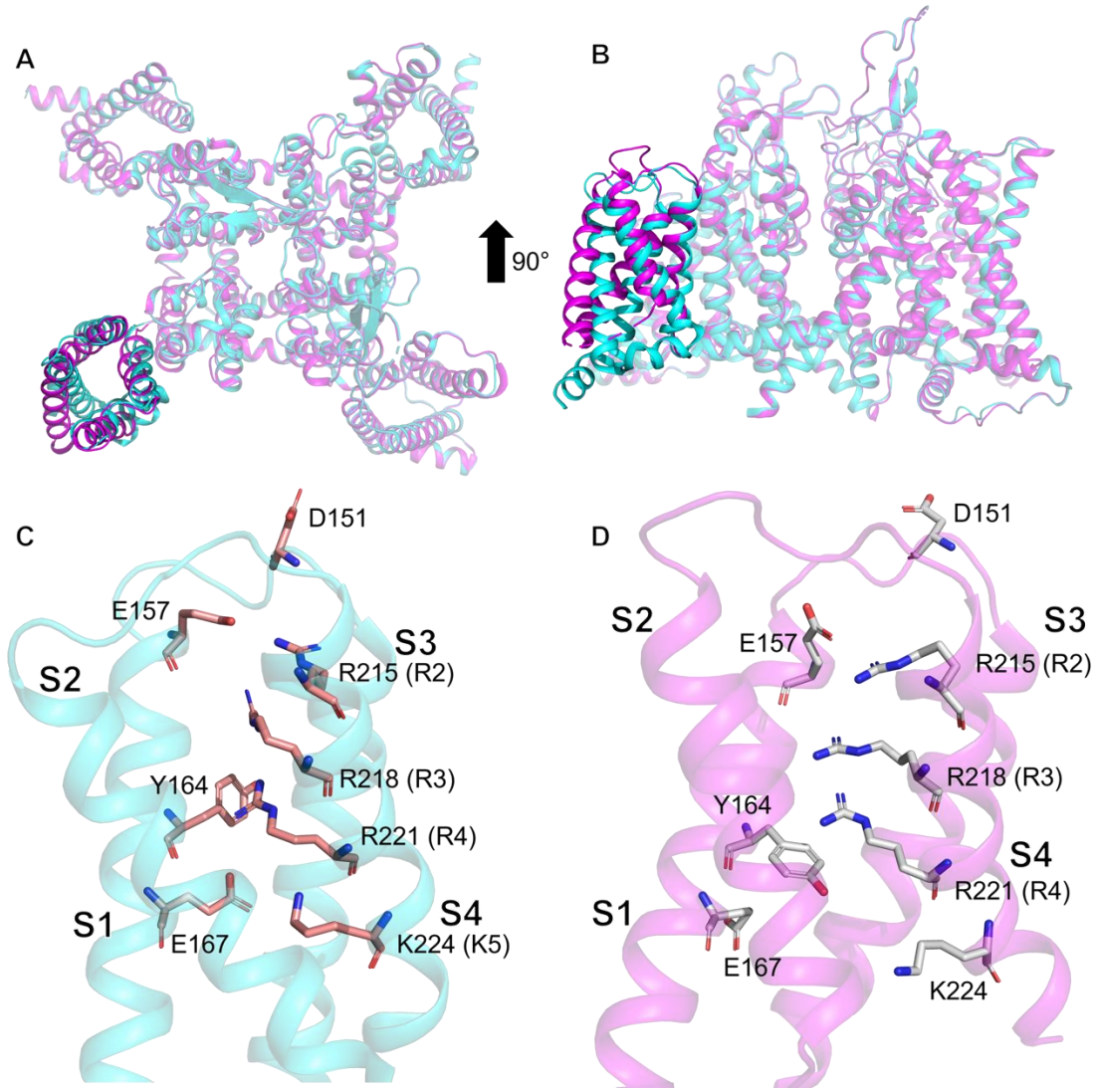

Figure 1.11. Differences in the voltage-sensor domain between classes I and II. Figure highlighting the differences in voltage-sensor domain I (VSDI) between class I (cyan, PDB 7WE4) and class II (pink, PDB 7WEL) of hNav1.8 (X. Huang *et al.*, 2022). (**A**) and (**B**) show top and side views respectively of a structural alignment of class I and class II, performed in PyMol. VSDI in class II has undergone a 12 ° rotation with respect to the central axis of the channel compared to class I, while the rest of the structure remains closely aligned. (**C**) and (**D**) show key gating charge residue positions in class I and II respectively, with both images made from the same viewpoint. Grey = C, blue = N, red = O. The occluding Y164 residue of S2 has swung downwards in class II and is forming a π-cation interacting with R221 of S4, and the anionic residues E157 and E167 are out of reach of gating charge residues of S4, showing that the four helices have not moved as a rigid body. Figure made using PyMol molecular graphics software (version 2.0; Schrödinger, LLC).

When comparing the structures of class I and II of hNav1.8 with that of hNav1.7 (G. Huang *et al.*, 2022), the bulk of the three structures are very similar (RMSD of 0.66 Å over 803 Cα atoms), with the only region of difference being VSDI; X. Huang *et al* therefore investigated this region to determine whether it is the source of hNav1.8's depolarised activation potential compared to other hNavs. Comparing the sequences of each hNav, they found that out of all the interfaces between VSDs and PDs in each, only the interface of VSDI and S5II has residues unique to hNav1.8, where 6 substitutions are present (from hNav1.7 to hNav1.8; I226V on S4I, and V793I, I800V, V803L, M806K, and F809L on S5II). Since movements of VSDs I-III caused by depolarisation are known to be associated with channel activation (Gilchrist and Bosmans, 2018), they suspected that this might be why hNav1.8 activates at a more depolarised voltage, and tested this hypothesis by site-directed mutagenesis. The substituted residues on hNav1.8 were mutated to those of the rest of the hNavs and the effect on activation measured by electrophysiology; of these, V800I had a negligible effect, and I226V and V793I were not tested as these interact directly with each other. Mutation of L803V actually moved the activation slightly towards depolarisation, but mutations K806M or L809F did shift the activation towards that of hNav1.7, and the two of these combined shifted it still further, with the equivalent mutations of hNav1.7 to hNav1.8 residues having the opposite effect, suggesting that these residues are indeed involved in the differences in activation between hNav1.8 and other hNavs.

A-803467, a subtype-specific blocker of hNav1.8, is resolved in the pore of the non-apo structures. It sits with the dimethoxybenzene group of the molecule in the fenestration between domains I and IV, clenching S6IV as it does so (see Figure 1.12), and likely impeding conductance by forming a hydrophobic barrier underneath the selectivity filter (SF). Looking at Figure 1.12, the molecule does not significantly cover the central axis of the pore, which may explain why the half-maximal inhibitory concentration IC_{50} of A-803467 is nearing the micromolar range, measured by X. Huang *et al* at 0.73 ± 0.08 μM. Comparing the pore domains of the apo and bound structures does not show a difference between backbone carbons, other than M1713 of S6IV has swung around to avoid steric clash with the blocker. This fits with A-803467 being suspected to preferentially bind the inactivated state, which the published structures have been interpreted as. To determine how A-803467 is subtype-specific, which is very important for the design of future

subtype specific drugs, the group mutated residues in the binding site to see what effect, if any, this would have on the potency of the drug. Surprisingly, only one residue in the binding site, M1713, is specific to hNav1.8, and when this was mutated to valine to match other hNavs the IC_{50} was actually reduced to 0.47 ± 0.07 μM. As the drug preferentially binds the inactivated state over resting, the group considered that the specificity may be related to conformational changes in the intracellular gate or fenestrations that occur between resting and inactivation; limiting their search to hNav1.8 specific residues in these areas, the group found that a mutation of G1406 of S6III to serine (present in all other isoforms), 11 Å away from the binding site decreased the potency of A-803467 roughly 7-fold, to 5.12 ± 0.86 μM, and mutation of T397, a helical turn below the intracellular gate and very distant from the binding site, to alanine reduced the IC_{50} to 2.23 ± 0.51 μM. Together, these effects identified A-803467 as inhibiting hNav1.8 allosterically, a mechanism which would be difficult to predict when designing subtype-specific drugs.

Overall, the published structures of hNav1.8 go some way to explaining the differential characteristics of this isoform, and the knowledge that subtype specificity can be achieved allosterically may be able to be exploited for the safe treatment of channelopathies. Despite this success however, and the fact that almost all hNavs now have published structures, there is still more to investigate; structures in resting and activated conformations, rather than only inactivated, are also needed, as it is not possible to determine why drugs such as A-803467 preferentially target certain states without this information, and gaps are left in our knowledge of the mechanism of opening, inactivation, and closing without them. Structures of hNavs published so far also are limited to only modelling 50-60% of the protein, with the loop regions (the site of much of the sequence heterogeneity between isoforms, and therefore a source of interest for subtype-specific drug development) remaining unresolved.

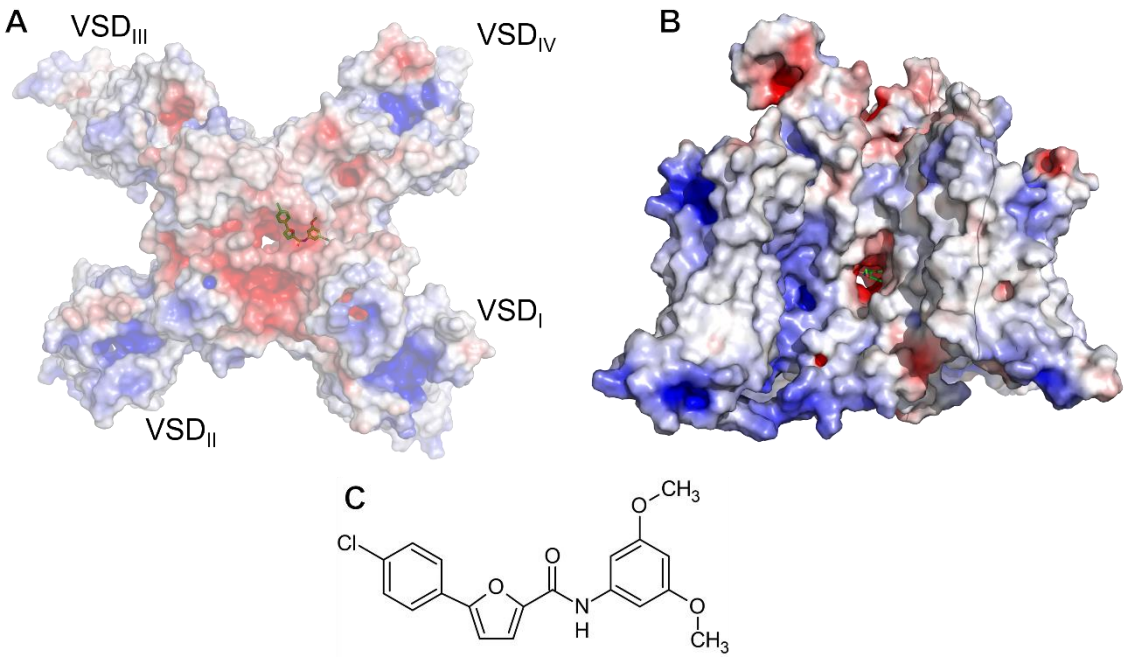

Figure 1.12. The binding site of A-803467 in hNav1.8. Figure highlighting the binding site of A-803467 in hNav1.8, using a space-filling representation of class II (PDB 7WEL) with volume transparency set to 0.6 in PyMol. (**A**) shows a top view of the protein with voltage-sensor domains (VSDs) labelled; the drug is bound with the dimethoxybenzene head lodged in the fenestration between domains I and IV. (**B**) shows A-803467 visible through this fenestration, where it is bound just below the selectivity filter, clenching S6 of domain IV and possibly creating a hydrophobic barrier to hydrated sodium ion conductance. (**C**) shows the molecular structure of A-803467. Figure made using PyMol molecular graphics software (version 2.0; Schrödinger, LLC).

1.5. Aims and objectives

1.5.1. Chapter 5: Purification and Electron Microscopy Analysis of hNav1.8 Recombinantly Expressed in *Tetrahymena thermophila*

The overall aim of this project was to optimise the purification of hNav1.8 expressed in *Tetrahymena thermophila*, followed by optimisation of cryo-TEM sample preparation with the aim of obtaining a high-resolution structure. This structure would then be used to probe the mechanism of this medically important channel, and ideally identify hNav1.8's role in pain disorders. This would be achieved by trialling different purification methods inspired by other membrane protein targets and taking advantage of *T. thermophila's* high cell membrane surface area as an abundant source of expressed recombinant hNav1.8. Success in this would provide the first structure of hNav1.8, and could give information on the poorly understood mechanism of voltage-gated sodium channel conformational changes.

1.5.2. Chapter 6: Purification and High-resolution Structural Determination of NavMs in Membrane Scaffold Protein Nanodiscs

The overall aim of this project was to develop a method for the structural study of NavMs by cryo-TEM and obtain a high-resolution structure. X-ray crystallography structures are already available, however, without the binding of drug ligands it has only been possible to view it in the open conformation. Ideally, the protein could be studied in conditions closer to that of a physiological membrane than can be produced using X-ray crystallography, and allow the elucidation of other conformational states, as well as opening up the target to further study of binding of drugs to voltage-gated sodium channels which are difficult to assess by X-ray crystallography. As previous attempts by the Wallace group to study NavMs by cryo-TEM with the protein solubilised in detergent or amphipol had been unsuccessful, it was decided to develop a protocol for the reconstitution of NavMs in membrane scaffold protein nanodiscs, which if successful could also be adapted to study other difficult membrane protein targets.

Chapter 2

Introduction to Techniques

2.1. Chapter Summary

The following chapter gives background information on the primary techniques described in this thesis, with specific reference to the expression and purification of integral membrane proteins, including non-detergent methods of membrane protein solubilisation and details of electron microscopy theory and workflow.

2.2. Overview of expression and purification of membrane proteins

At the time this project began, only one of the nine hNavs isoforms, hNav1.4, in complex with β1 (Pan *et al.*, 2018), had yet been successfully expressed, purified, and analysed by structural methods; this is because these proteins are particularly challenging targets. Membrane proteins, whether prokaryotic or eukaryotic, are generally more difficult to express and purify than cytosolic proteins. Firstly, overexpression of membrane proteins such as ion channels is more toxic to the host cell, meaning not as much can be produced (Gubellini *et al.,* 2011). If expression is successful, then the hydrophobic nature of the transmembrane regions of membrane proteins means that they must be solubilised using amphipathic molecules such as detergents (see Figure 2.1) to avoid aggregating and stay correctly folded; this requires investigation into a suitable detergent or other methods of solubilisation, as not all will work with any target, or may be useful for solubilisation but unsuitable for downstream analysis. Even with solubilisation in materials such as detergent or amphipols, this remains a very different environment to the cell membrane, and therefore the protein may not keep the same functional activity. To overcome this difficulty expressing membrane proteins in prokaryotes, specific strains have been developed that are more able to withstand toxicity, such as C41(DE3) (Gubellini *et al.*, 2011); however, expression of hNavs in *E. coli* is not viable, as prokaryotic expression systems, lacking the requisite cofactors and chaperones, are not able to express and correctly fold these complex proteins, and are not able to post-translationally modify the

proteins as eukaryotes can. Mammalian expression systems can overcome this, and some, such as the human embryonic kidney (HEK) 293F cell line (Portolano *et al.*, 2014), have been developed for faster growth which is more suitable for toxic proteins, but generally, these are still more costly and do not produce eukaryotic membrane proteins at the same scale as prokaryotic proteins can be expressed.

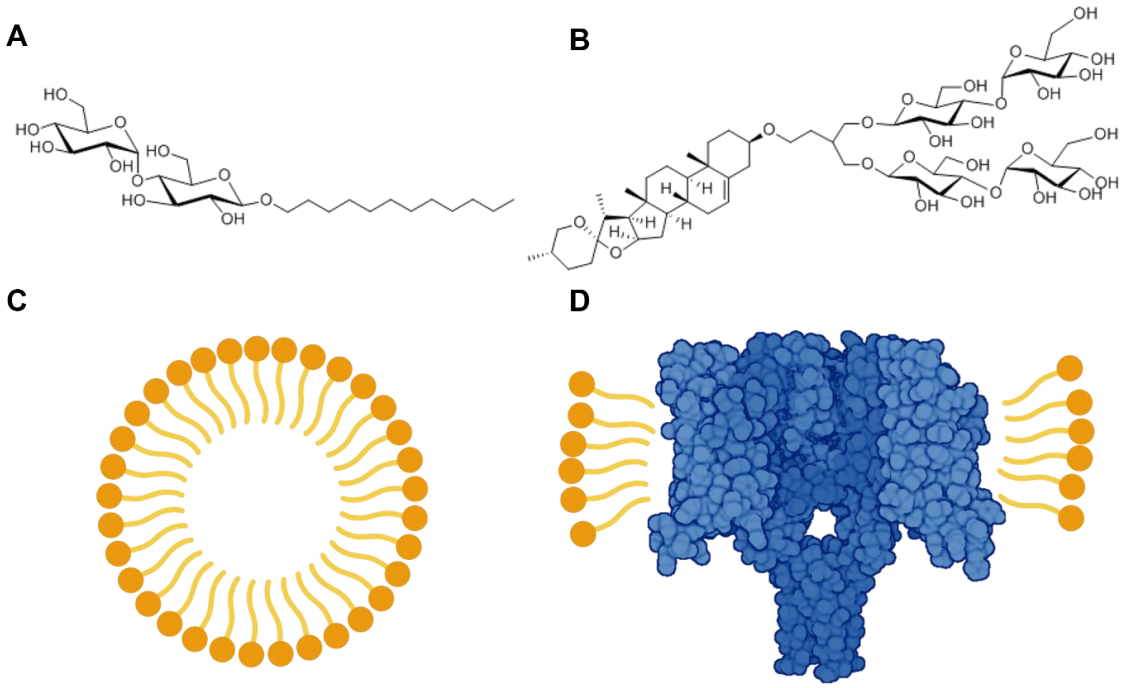

Figure 2.1. Detergent examples and schematic of solubilised membrane proteins in micelles. Figure showing examples of detergents and how they solubilise integral membrane proteins. (**A**) and (**B**) Molecular structures of the non-ionic detergents N-dodecyl-β-D-maltopyranoside (DDM) and glyco-diosgenin (GDN), respectively. Note the polar 'head' regions with many alcohol groups which form hydrogen bonds with water and non-polar 'tail' regions which bind to and solubilise the integral membrane regions of a protein. (**C**) shows a cross-section of a detergent micelle in aqueous conditions; the polar head groups point outwards to the surrounding water whilst the non-polar tails point inwards, forming a sphere with no water in the core. (**D**) shows a representation of how detergent molecules solubilise the transmembrane region of an integral membrane protein, in this case NavMs (PDB code 5HVX). (**C**) and (**D**) created with BioRender.com.

When membrane proteins can be successfully overexpressed, purification follows a general protocol; cells are lysed and centrifuged to clear unwanted cell debris, and the soluble fraction left over is then ultracentrifuged at speeds of up to 200000 x *g*, separating the membrane fraction and causing it to sediment as a pellet. This membrane pellet can then be solubilised by the aforementioned detergents or lipids, and purification continues as with soluble proteins, for example, by affinity chromatography and size exclusion chromatography (SEC). Different detergents are often used for the solubilisation and purification steps, and can also be exchanged before further analysis; N-dodecyl-β-D-maltopyranoside (DDM) is often used for solubilisation as it is non-polar and has a low critical micelle concentration (CMC) of 0.0087% in water, but this same low CMC means that many free micelles are produced, which can be difficult to distinguish from micelles containing protein in cryo-transmission electron microscopy (cryo-TEM), and can interfere with the growth of crystals when attempting X-ray crystallography, meaning that after solubilisation samples are often exchanged into other detergents such as decanoyl-N-hydroxyethylglucamide (HEGA-10) or glyco-diosgenin (GDN). Newer methods have been developed to improve this process, such as by giving the protein greater biochemical stability via amphipathic polymers like A8-35 amphipol (Zoonens and Popot, 2014), or in the case of membrane protein nanodiscs (see section 2.3 for more detail) more closely mimicking or even extracting the native membrane environment, meaning the sample can be more physiologically relevant and suitable for analysis by methods like cryo-TEM.

2.3. Detergent-free membrane protein purification

As discussed in section 2.2, recombinantly expressed membrane proteins require solubilisation before they can be stable for purification and structural or functional analysis in aqueous environments. Detergent molecules which form a micelle around the hydrophobic regions of the protein have been the method of choice for the majority of membrane protein structures produced, but limitations do exist. For example, detergent in buffers can often cause produce artifacts in negative stain electron microscopy, leading to difficulties in screening samples; protein-free micelles can be difficult to tell apart from those containing protein (Gewering *et al.*, 2018), and free detergent can create background noise even below the detergent's stated critical micelle concentration (CMC). It has also been suggested that detergents can interfere with the biochemical stability of

membrane proteins by interaction with the extramembrane soluble domains, altering thermodynamic properties or conformation (Yang *et al.*, 2014), and detergents do not accurately mimic the phospholipid environment of a cell membrane. Finally, if the detergent used for solubilisation is too weak, it will not effectively solubilise the protein, but if too strong, it can interfere with the interactions between proteins in a complex, breaking apart the very structures they are being used to study. Due to these difficulties and the desire to reconstitute membrane proteins closer to their native lipid environment, a range of methods have been developed to replace detergents in membrane protein solubilisation and purification.

2.3.1. Amphipols

Amphipols, sometimes referred to as APols, are short, synthetic carbon-chain polymers with both hydrophobic and hydrophilic side chains (Tribet, Audebert and Popot, 1996). This allows them to wrap around the transmembrane region of a protein, hydrophobic side chains pointing inwards and hydrophilic out to the aqueous environment, thereby solubilising the protein. The detergent can then be safely removed by diluting out or adsorption to Bio-Beads (Rigaud *et al.*, 1997). This has advantages over solubilisation by detergent as once the protein is solubilised, the amphipol remains tightly bound to the hydrophobic regions, conferring greater biochemical stability than detergents (Le Bon *et al.*, 2018). The amphipol A8-35 (Pocanschi *et al.*, 2006) in particular has been successfully used for solubilisation and structural or functional analysis of proteins, for example the 3.45 Å cryo-TEM structure of the receptor tyrosine kinase KIT (Krimmer *et al.*, 2023) or purification of a functionally active respiratory supercomplex (Shinzawa-Itoh *et al.*, 2016). They can also be isotopically or otherwise labelled, giving them extra functionality in non-structural studies (Le Bon, Popot and Giusti, 2014). Beyond these advantages though, amphipols are not any closer to a native lipid environment than detergents, unlike the methods described in sections 2.3.2 and 2.3.3.

2.3.2. Styrene-maleic acid lipid particles (SMALPS)

Styrene-maleic acid (SMA) copolymers are capable of extracting proteins directly from the cell membrane into nanometre-scale discs, taking with them the native membrane environment, with no detergents used even for the initial solubilisation (Lee *et al.*, 2016; Esmaili and Overduin, 2018). As some proteins are not capable of being solubilised by detergents, and large complexes of proteins may not form in a detergent environment, the

extraction of membrane proteins alongside their native lipids into SMALP nanoparticles has potential for more accurate examination of difficult samples. As with any extraction and purification method, SMALP extraction benefits from optimisation of parameters such as the ratio of SMA to protein, buffer pH conditions, and accessibility of affinity tag sites on the target protein (Lee *et al.,* 2016). SMALP discs are dynamic and structurally flexible, with lipid molecules exchanging between discs, which can be unhelpful for cryo-TEM as it complicates alignment and averaging between particles. This method has still led to a few successful structural elucidations, including with interesting data that would be difficult to discern with other methods, such as protein-protein interactions within a larger complex or the position of native lipids in protein cavities or binding sites (Unger *et al.*, 2021).

2.3.3. Membrane scaffold protein (MSP) nanodiscs

Membrane scaffold protein (MSP) nanodiscs were engineered by truncation of human apolipoprotein AI (ApoA-I), involved in lipid and cholesterol transport in the liver (Denisov *et al.*, 2004). These amphipathic proteins self-assemble with lipids as a disc-shaped dimer containing a lipid bilayer (the nanodisc) and can stabilise a membrane protein within that lipid bilayer (see Figure 2.2). They can assemble around membrane proteins in their native environment, or the protein can be purified in detergent before transfer into nanodiscs. This is a relatively simple procedure, as the MSP, lipid, and protein can be incubated together with Bio-Beads which gradually adsorb the detergent and promote integration of the protein into the lipid bilayer. A range of MSPs are now available of different lengths, which then produce nanodiscs of different diameters to accommodate larger or smaller transmembrane domains, from 7-8 nm diameter MSP1D1 to 17 nm diameter MSP2N2 (Denisov and Sligar, 2016, 2017), and individual phospholipids or mixtures to mimic a native environment are readily commercially available. Drawbacks to MSP nanodiscs are that a specific MSP:lipid:protein ratio is required for stability, along with choosing the correct lipids and MSP length, which if similar proteins have not already been structurally elucidated in MSP nanodiscs can require significant time and resources to determine by trial and error. Multiple proteins have had high-resolution structures determined using MSP nanodiscs (Luo *et al.*, 2021; Nakanishi *et al.*, 2022), including a 3.1 Å cryo-TEM structure of the prokaryotic voltage-gated sodium channel NaChBac (Gao *et al.*, 2020).

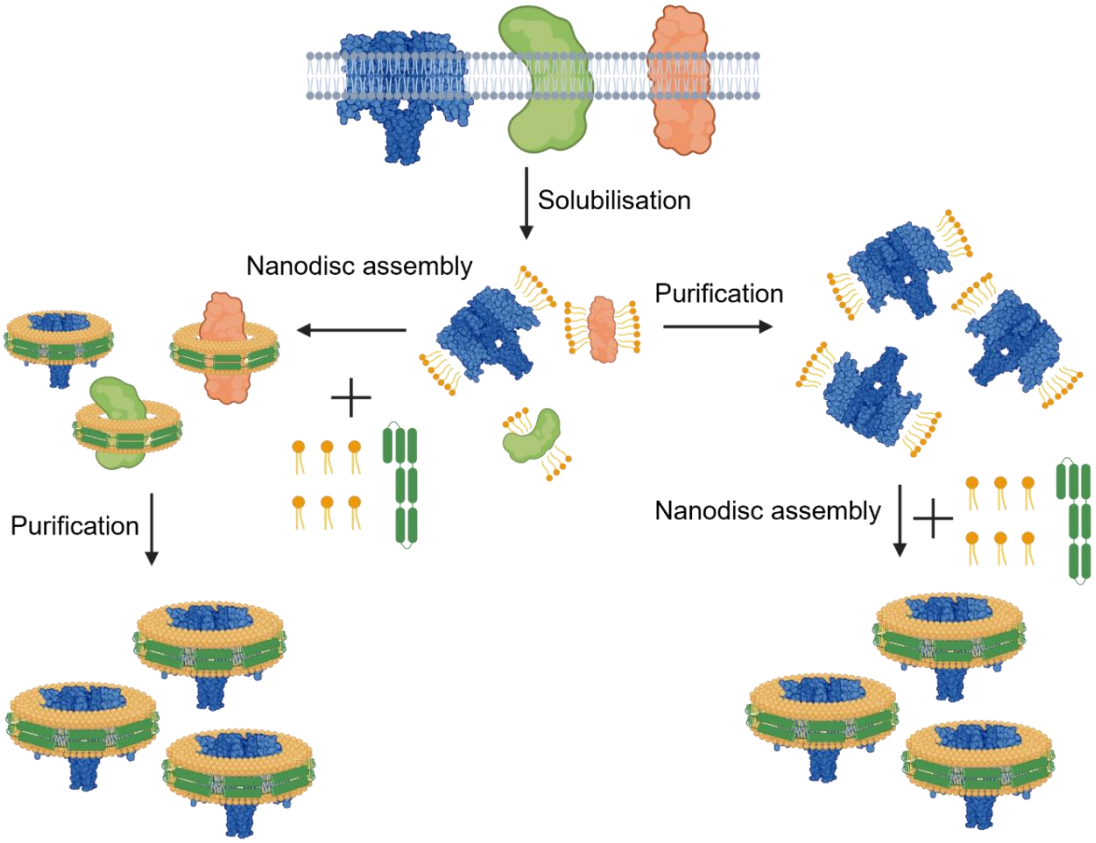

Figure 2.2. Workflow of membrane protein reconstitution in MSP nanodiscs. Figure showing the workflow of reconstitution of membrane proteins into membrane scaffold protein (MSP) nanodiscs. Cell membranes containing target protein are solubilised with detergent, and then two routes are possible: on the left, we see the addition of phospholipids (yellow) and MSP (green) to solubilised proteins, forming nanodiscs around each, followed by purification to obtain only nanodiscs with the target protein. On the right, we see the purification of target protein from solubilised proteins whilst still in detergent, followed by addition of phospholipids and MSP to form nanodiscs around target protein (here shown as NavMs, PDB code 5HVX). Created with BioRender.com.

2.4. Electron Microscopy

2.4.1. The history of electron microscopy

Microscopy has been used to study nature on a scale smaller than the human eye can discern since the 17th century when Robert Hooke published his *Micrographia* (Hooke, 1665), using a single (objective) lens optical microscope to produce approximately 20x magnified illustrations of insects and plants, coining the term 'cell' in the process, and setting off a revolution in biology. Anton van Leeuwenhoek made improvements to the simple microscopes available at the time and was the first to physically observe biological material at up to 300x magnification, describing what we now know as bacteria as 'animalcules' (van Leeuwenhoek, 1677), although the difficulty in producing microscopes at the time meant that his work was not thoroughly reproduced or explored further for some time. In the 1800s innovations by scientists such as Ernst Abbe pushed the capabilities of microscopes still further with the development of double-lens microscopes consisting of an objective and a condenser lens, and the apochromatic lens system which allowed accurate representation of colour. However, these advances also led to an understanding of the limitations of optical microscopes, when Abbe demonstrated that two objects could only be resolved from each other if their distance was greater than half the wavelength of the light source, or:

$$d = \frac{\lambda}{2NA} \tag{2.1}$$

Where d is the possible resolution, λ is the wavelength of illumination and NA is the numerical aperture, defined by the refractive index of the imaging medium (e.g., air). When imaging with visible light, such as with a light microscope, this means that the best resolution possible is approximately 300 nm, whereas a resolution of a thousand times smaller than this is needed to reliably image atoms.

Clearly, a different source was needed if fundamental questions in biology were to be answered; over the next century, X-rays and electrons became the answer. X-rays have a wavelength of 0.1 to 10 nm (1 to 100 Å), and can take advantage of the periodicity of a crystal to reach resolutions beyond the limit that would be suggested by equation 2.1. A great leap forward was made when the first X-ray crystal structures of macromolecules

were elucidated, leading to Nobel Prizes in Chemistry for John Kendrew and Max Perutz for determination of the structures of myoglobin and haemoglobin (Kendrew *et al.*, 1958; Perutz *et al.*, 1960), and in Medicine for the model of DNA after work by Rosalind Franklin, Francis Crick and James Watson (Watson and Crick, 1953). X-ray crystallography (and to a lesser extent nuclear magnetic resonance (NMR) spectroscopy), has been the go-to method to obtain high-resolution structural data of proteins for decades, reflected in it being the experimental source of roughly 85% of the depositions in the PDB. X-ray crystallography is still the source of many protein structures, but it has its own limitations: the sample must itself be able to form crystals, which can become difficult with membrane proteins and proteins of high molecular weight; it requires a high concentration of homogeneous, pure source material, such as 10 mg/mL; a crystal structure is one static form of the target protein which may cycle through many orientations, and also may not be representative of its native, physiological state.

When electrons were first discovered by J.J. Thomson in 1897, they were thought of only as particles, and therefore not capable of diffraction like photons; this changed with the advent of quantum mechanics and Louis de Broglie's suggestion of wave-particle duality, whereby electrons possess a relativistic wavelength proportional to their momentum via the de Broglie relationship, which is as follows:

$$\lambda = \frac{h}{mv} \tag{2.2}$$

Where λ represents wavelength, h is the Planck constant, m is the mass of the particle and v is its velocity. This means that if an electron is accelerated with a voltage of 300 kV, as is possible with an FEI Titan Krios transmission electron microscope, it has a relativistic wavelength of approximately 1.97 pm (0.0197 Å), far smaller than is needed to resolve between two atoms. Electron microscopes, using electromagnets rather than glass lenses to focus an electron beam, (see Figure 2.3 in next section for the schematic of a modern electron microscope) were first developed as a prototype in 1931 by Max Knoll, an electrical engineer, and Ernst Ruska, a physicist (Franken *et al.*, 2020), with electron microscopes available for commercial use and obtaining better resolutions than were available with light microscopes by 1939 (Ruska was jointly awarded the Nobel Prize in Physics for this work in 1986). Despite this, there were many limitations that meant it would be decades before the first atomic-resolution protein structures were found by modern electron microscopy, for example that aqueous samples could not be looked at in

the vacuum conditions of an electron microscope, and that biological samples would be too damaged by the electron beam to give useful data. The method of negative stain electron microscopy (Brenner and Horne, 1959) went some way to dealing with these issues: a biological sample is 'stained' in a heavy metal substrate such as uranyl acetate or uranyl formate, protecting it from radiation damage and removing the issue of evaporation of sample in vacuum conditions. This, along with the development of methods to find the orientation of 2D projections of particles relative to each other in Fourier space, made the three dimensional reconstruction of the tail of bacteriophage T4 possible (De Rosier and Klug, 1968); whilst only to a resolution of 30-35 Å, this gave far more detail than would be possible with a light microscope, and of a structure that would not be suitable for crystallisation. Iterative improvements were made to this process, such as using known symmetry of structures to aid in alignment of projections, and reducing the electron dose required so that biological samples would not be as damaged (Frank, 1975), but as the resolution of negative stain electron microscopy is limited by the size of the grains of stain, rarely reaching past 20 Å, and the highly acidic environment means that proteins are unlikely to be fully representative of their native state, stainless, low-dose methods would have to be implemented to obtain higher resolution data.

This led to the development of cryo-transmission electron microscopy (cryo-TEM), first performed by freezing samples on a grid in liquid nitrogen (Taylor and Glaeser, 1976). Freezing the sample meant that the protein, in this case catalase, was more protected from radiation damage than a wet sample, and a resolution of 11.5 Å was achieved. The limiting factor then became the crystalline nature of ice producing a large amount of background noise, which if looking at a low-contrast biological sample means that it can be difficult to pick out true signal: to remedy this, Jacques Dubochet developed a rapid freezing system to produced amorphous, non-crystalline (vitrified) ice. Computational methods of 3D reconstruction were developed, and when combined with vitrification allowed the reconstruction at 35 Å of the Semliki Forest virus envelope in an aqueous environment (Vogel *et al.*, 1986), and within several years of this Richard Henderson demonstrated the true potential of cryo-TEM by publishing a structure of bacteriorhodopsin with details at 3.5 Å resolution (Henderson *et al.*, 1990). Frank, Dubochet, and Henderson would be awarded the Nobel Prize in Chemistry in 2017 for the impact of this work.

Despite this huge leap in the technical capabilities of cryo-TEM, high-resolution structures like that of bacteriorhodopsin did not immediately become the norm, and

electron microscopy could still not be used to elucidate biomolecule structures with as much detail or as readily as via NMR or X-ray crystallography. This finally changed with the replacement of charge coupled device (CCD) detectors, which after excitation by the electron beam release a subsequent photon which is then then relayed as a voltage by fibre optic cable (Roberts, Chapman and MacLeod, 1982), with direct electron detectors. CCD detectors are much less sensitive; they have a lower 'detection quantum efficiency' (DQE), meaning that they only accurately detect about 10% of the incoming information (Henderson and Glaeser, 1985). The greater efficiency of electron detection meant that lower doses were needed to gather the same data, aiding with the issue of radiation damage, and the increased rate of detection by digital cameras meant that images could be taken much more rapidly, allowing for 'movies' to be made of multiple frames over the timeframe of several seconds. This revealed another method of obtaining a better signal-to-noise (SNR) ratio, and therefore resolution, as computational alignment of frames was used to correct for motion of particles caused by the energy of the electron beam itself (Brilot *et al.*, 2012). Another limiting factor in cryo-TEM reconstructions is the Nyquist limit, which corresponds to double the pixel size of the detector. For example, if the magnification setting of the data collection is such that 1 Å of the sample spans a single pixel on the detector, then the pixel size is 1 Å, the Nyquist limit is 2 Å, and spatial frequencies surpassing this limit will suffer from artifacts. Direct electron detectors have made it possible to surpass this limit with the advent of super-resolution mode, as a single electron can be detected in more than one pixel, centring the event to a quadrant of a single pixel (Li *et al.*, 2013; Feathers, Spoth and Fromme, 2021).

Technological advancements, such as new and improved models of direct electron detectors for more accurate and faster data collection, and Selectris image filters, which improve contrast of TEM images, along with the availability of plunge-freezing devices such as the Vitrobot Mark IV, and advances in computational capabilities (both hardware and software) have allowed what is termed a 'resolution revolution' in cryo-TEM (Kühlbrandt, 2014), with greater numbers of PDB depositions each year. While X-ray crystallography still produces more PDB depositions than cryo-TEM per year and is more suitable in some circumstances (for example, cryo-TEM becomes more challenging the smaller the protein due to lower SNR), electron microscopy has become invaluable in finding atomic-resolution structures of macromolecular complexes, such as an entire 5.8 MDa respiratory supercomplex (Mühleip *et al.*, 2023), or a 3.72 Å photosystem I structure

from cyanobacterium (Semchonok *et al.*, 2022), which would be extremely challenging or impossible by X-ray crystallography. The inherent flexibility and hydrophobicity of membrane proteins also makes them a more suitable target for cryo-TEM, in particular proteins which cycle through multiple states (such as voltage-gated sodium channels), as these different states can be captured and separated out to multiple classes and give greater detail of how a protein is performing its function.

2.4.2. Optics in an electron microscope

Before beginning the extensive process of sample preparation, screening, data collection and image analysis that leads to a high-resolution protein structure it is helpful to first understand how images are formed in an electron microscope. This begins with the layout of the electron microscope itself, starting with providing a source of electrons. Older electron microscopes, or less advanced screening microscopes often use a tungsten or lanthanum hexaboride filament which is heated to high enough temperatures that electrons are emitted thermionically, while newer, more advanced microscopes will commonly use a field emission gun (FEG), which can be made of a single crystal of tungsten sharpened to a radius at the tip of approximately 100 nm. This tip is held at a negative potential versus a positive electrode at the other end of the vacuum-held microscope column (for example at a 300 kV potential difference in a Titan Krios microscope), causing electrons to be drawn out of the crystal by the applied electric field. This has the advantage of creating a narrower, more coherent, and brighter electron beam than a filament, albeit being more costly to produce. The electron beam then passes through a series of water-cooled electromagnetic lenses (electromagnetic coils surrounding the column which can be strengthened by increasing the current) and apertures which focus and refine the electron beam (see Figure 2.3) beginning with a condenser lens or system of lenses which condense and parallelise the electron beam (Franken *et al.*, 2020). This is followed by a condenser aperture which blocks the path of electrons travelling wider than a certain angle away from the optical axis of the microscope column before interacting with the sample, where electrons can pass through the specimen without interacting, be absorbed, or be inelastically (transferring energy), elastically (no energy transfer), or back scattered (reflected backwards) by charges in the sample. This includes both the positively charged nucleus and the negatively charged electron cloud of an atom, in contrast to X-rays which are diffracted by electrons only (Wang and Moore, 2017). The electrons then pass through the objective lens and aperture

system, which focuses the electrons towards a back focal plane where the objective aperture is situated and a diffraction pattern is formed, before producing a two-dimensional image at the first image plane. Intermediate and projector lenses can then be used to refocus and magnify the image further, flipping the image 180° each time until the electrons meet the detector at the final image plane.

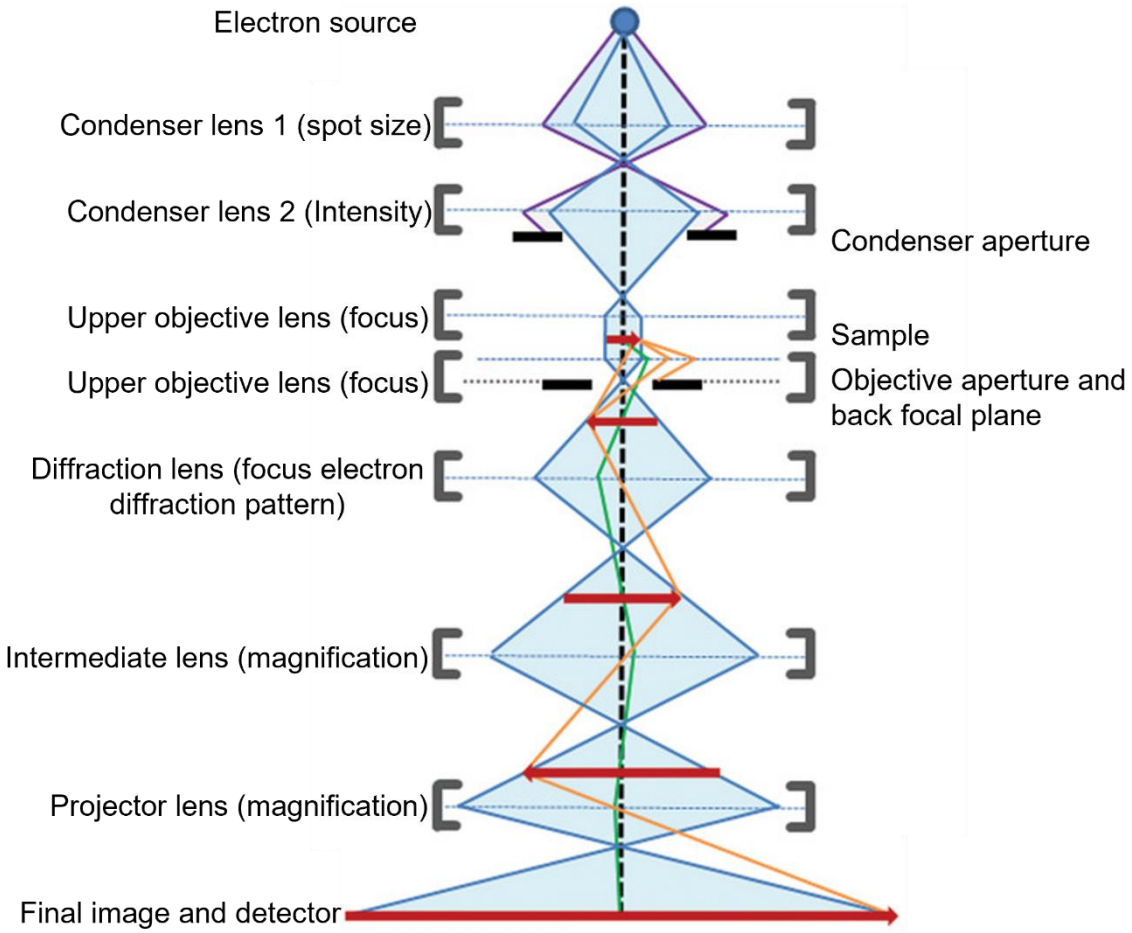

Figure 2.3. General electron microscope structure. A schematic of an electron microscope showing lens and aperture systems, electron pathways, scattering, and image formation. The blue circle at the top of the schematic represents the electron source, grey brackets represent lenses (labelled with their functions), black rectangles represent apertures, black dotted line represents the incident electron beam, blue and purple lines represent the pathways of the electron beam, yellow and green lines represent the pathways of scattered electrons, and red arrows represent the sample and its image planes, with the image flipping with each change of direction of the arrow. Adapted from Franken *et al*., 2020.

2.4.3. Scattering and image formation

The image produced by the microscope is formed by varying contrast between different areas of the image, which is in turn decided by how the electrons interact with the specimen. As mentioned earlier, this interaction is made up of several types and extents of scattering, which, thanks to the wave-particle duality of electrons postulated by de Broglie, allow the electrons to be detected as possessing differing amplitudes and phases, both of which contribute to image contrast. We will first discuss amplitude contrast, made possible by an electron's particle-like nature. Electrons which are absorbed (not common with protein samples as they are made up primarily of light atoms), backscattered, or scattered at such a high angle that they are blocked by the subsequent aperture, will not reach the detector, and therefore produce areas of low amplitude signal. Inelastically scattered electrons are undesirable as the lack of knowledge of how much energy has been lost to the sample means that their detection does not provide conclusive information, and because electrons with lower energy (and therefore longer wavelength) are focused differentially to elastically scattered electrons, causing chromatic aberration (Orlova and Saibil, 2011); therefore, they are removed before reaching the detector by an energy filter, essentially another form of aperture which only allows the passage of electrons possessing more than a certain energy. This leaves only electrons which have been scattered and refocused towards the image plane by the objective lens, and non-interacting electrons (also known as the carrier beam) being detected, forming a contrast in amplitude caused by higher and lower areas of adjacent charge density, or 'Coulomb potential' (Marques, Purdy and Yeager, 2019) in the specimen.

Phase contrast is made possible by the wave-like properties of an electron; elastically scattered electrons take a longer journey to the detector than those which travel in a straight line undeterred, meaning that they will exhibit a later phase than electrons in the carrier beam- this difference in phase is the phase contrast, and, like amplitude contrast, highlights differing areas of charge density in the sample. As mentioned earlier in this section however, biological samples mostly consist of light atoms such as C, N, O, and H, and scattering from these atoms does not cause strong shifts in phase due to their relatively low charge density, leading to weak contrast signal. Whilst only non-interacting or elastically scattered electrons are involved in phase contrast, other electrons will affect the resultant image due to their interaction with the sample. As energy from the electron beam is transferred to the specimen, the protein can be ionised, bonds can be broken and

reformed, or X-rays and radicals can be produced which damage and further alter the sample (Orlova and Saibil, 2011). This limits the exposure time possible during data collection, in turn limiting the SNR and possible resolution of a resultant structure as the maximum electron dose possible before damage is low.

As referred to in the previous section, once the electron beam has passed through the sample, the electrons which have been scattered are focused and magnified by the objective and projector lenses, forming an image at the image plane. According to quantum mechanics, these electrons can be modelled as a 'wavefunction,' oscillating with space and time, with the squared magnitude of the wavefunction at a given space and time giving the likelihood of finding the electron there. When an electron interacts with the sample by elastic scattering, this wavefunction is also scattered at multiple angles. Not all these elastic scattering events are equivalent; signal scattered at higher angles carries higher resolution data, as the scattering angle is proportional to the spatial frequency spectrum in Fourier space. In terms of what is going on practically, an electron which has been scattered at a high angle will travel longer before reaching the detector than those scattered at low angles, and as it travels it will be continuing to change phase. As the phase of an electron acting as a wave is cyclical, the different elements of the wavefunction can be in phase with each other, detected as positive contrast, as the magnitude of the wave is increased due to constructive interference; out of phase, detected as negative contrast where the amplitude of the wave is decreased due to destructive interference; or can have no effect at all, depending on what angle in or out of phase the wave is when it meets the electron detector. Because of this interference caused by different elements of the spectrum of spatial frequencies, certain frequencies will be detected as higher contrast than others, or negative contrast; to obtain high-resolution data, the influence of all of these varying spatial frequencies is modelled by the 'contrast transfer function' (CTF), which must be estimated and corrected for when processing cryo-TEM data (Singer and Sigworth, 2020). The mathematical model of CTF used by cryoSPARC (Punjani *et al.*, 2017), the software used for data processing in this thesis, is given in equation 2.3:

$$CTF = -\cos\left(\pi \Delta z \lambda_e f^2 - \frac{\pi}{2} C_s \lambda_e^3 f^4 + \Phi\right) \tag{2.3}$$

Where Δz is the defocus of the microscope (the distance away from the detector that the image is focused to), λ_e is the wavelength of the incident electrons, C_s is the spherical

aberration (a measure of the lens focusing rays differently depending on where relative to the middle of the lens they are being focused from, which is specific to each individual microscope), f is the spatial frequency of each Fourier component, and Φ is a phase shift factor, which is relevant when a phase plate is in use (a device used to increase the phase contrast). The CTF can be modelled as a wave oscillating above and below zero, which if using a grid with a carbon foil (see section 2.4.6) can be overlain on a Fourier transform of an EM image so that the points where the wave crosses zero match with dark rings, known as Thon rings (see Figure 2.4). What the CTF provides us with is an estimate of how the contrast information has been modulated for each spatial frequency detected, allowing a correction of contrast and the retrieval of high spatial frequency data that would otherwise be lost. This only works to correct for interference produced by an individual spatial frequency, meaning that it cannot correct for phase interference between one wave and another scattered (or unscattered) at a different spatial frequency, leading to another source of lost information.

Loss of information due to phase interference is not the only source of degradation of high-resolution data. As in any real-world application, the microscope is not perfect; the aforementioned spherical aberration combines with other factors such as astigmatic aberration (stronger focusing in different directions) and chromatic aberration (unequal focusing of electrons of different wavelengths), as well as disparity in the energies of electrons in the beam and small differences in the angle of said beam as it meets the sample, which respectively give rise to temporal and spatial incoherence and lead to a dampening of higher spatial frequency information. This combination of losses of high spatial frequency information can be plotted against spatial frequency as the 'envelope function', a measure of the coherence of scattered waves, which tends towards zero (meaning more information lost) as the spatial frequency increases. In real space, this envelope function can be combined with the CTF to produce the 'point spread function' (PSF), which is essentially a description of how much an object (or point) in the sample is spread, i.e., blurred, in the final image, and is given by equation 2.4:

$$F\{PSF(\vec{r})\} = CTF(\vec{R}) \cdot E(\vec{R}) \tag{2.4}$$

Where $F\{PSF\,\vec{r}\}$ is the Fourier transform of the PSF in a vector $\vec{r}$ in the image plane, and $\vec{R}$ is spatial frequency in Fourier space (Orlova and Saibil, 2011).

Fortunately, some of this lost high spatial frequency information can be retained due to the fact that the CTF is a function of the defocus level of the microscope. Low defocus means that high spatial frequency data can be picked up whilst low spatial frequencies are lost, and high defocus leads to the opposite; in terms of how this affects the microscope user, at low defocus there will be low contrast for larger objects, meaning that particles are harder to pick out against the vitreous ice background, but high-resolution data is present, and conversely for high defocus. During data collection, images are taken at a range of defocus values, so that both high and low spatial frequency data can be obtained.

Combining the use of a range of defocus values, with estimation of the CTF, and correcting for known or estimable values in the envelope function such as spherical aberration and astigmatism can retrieve useful data lost in a degraded EM image. Despite this, biological samples still present a challenge as even with CTF correction phase contrast is weak. To counter this, data collection will often involve taking thousands of micrographs, each with hundreds of particles per image, meaning that a dataset can contain millions of particles to average together data from. Using this many particles also means that these particle images, each a projection of the full three-dimensional specimen, will ideally show all the high-resolution detail of the outside and inside of a protein from every possible angle, which can then be used to reconstruct a three-dimensional Coulomb potential map, from which the structure of the protein of interest can be modelled. To obtain a Coulomb potential map that accurately represents the protein and is of high enough resolution to glean biologically important information requires optimisation of every step of the cryo-TEM process, from protein purification, to grid preparation, screening, data collection and data processing; the workflow of a typical single-particle cryo-TEM investigation will be discussed in the following sections.

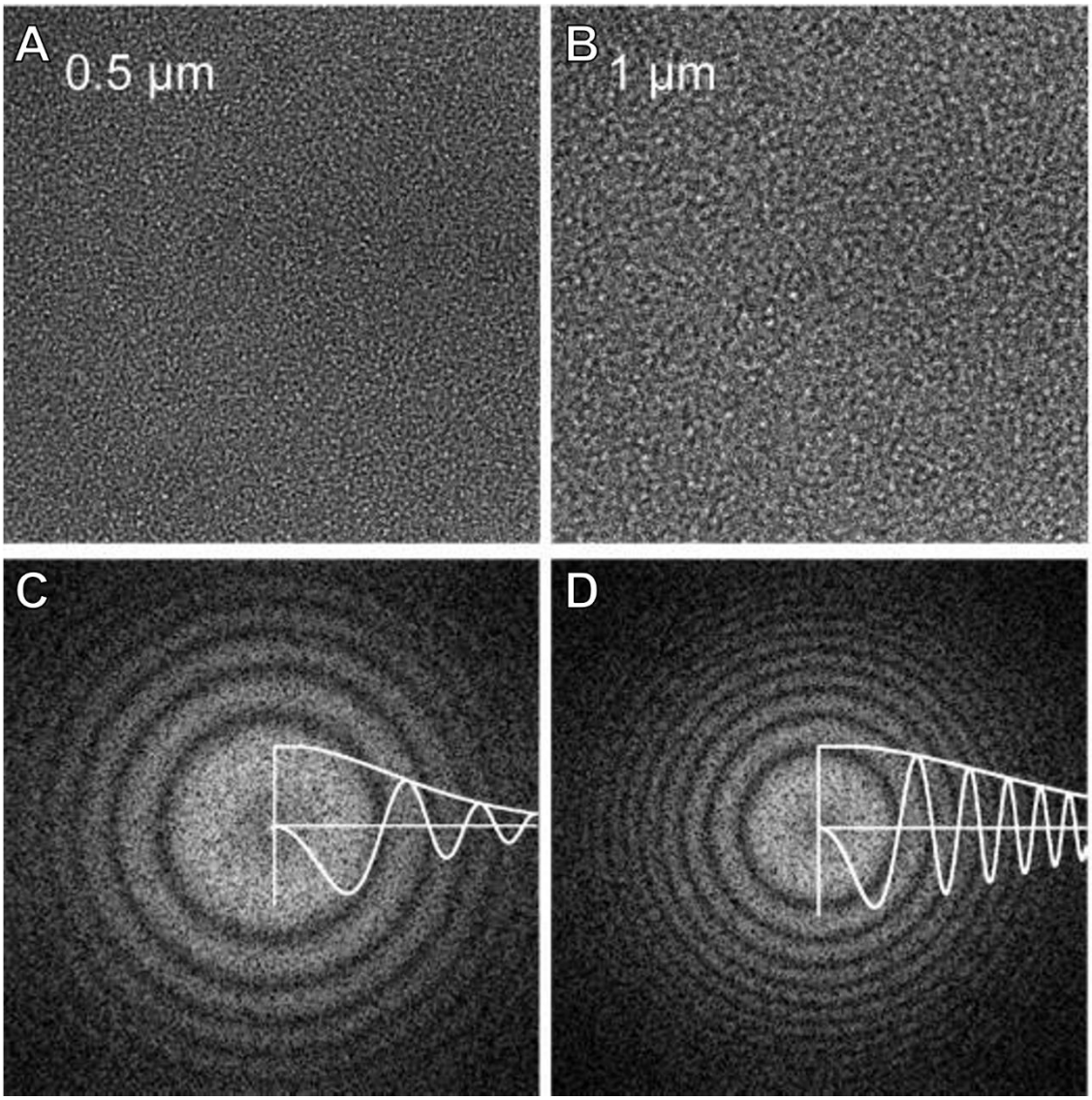

Figure 2.4. Example of the effect of defocus on the contrast transfer function. Figure showing electron microscopy images of carbon films, taken at 0.5 µm (**A**) and 1 µm (**B**) defocus, with their corresponding diffraction patterns (seen as Thon rings) shown underneath (**C** and **D**). The contrast transfer function (CTF) curves are overlaid on the Thon rings, showing how the CTF oscillates between positive and negative contrast, with dark rings where the contrast reaches zero. Increasing the defocus causes the CTF curve to oscillate more rapidly, meaning that more Thon rings are visible in the higher defocus image and the first zero value is closer to the origin. Adapted from Orlova and Saibil, 2011.

2.4.4. Sample conditions

While sample purity and concentration do not have to be quite as high with cryo-TEM as with X-ray crystallography, sample quality is still integral to being able to produce a high-resolution structure, and often multiple rounds of purification, for example affinity chromatography followed by size exclusion chromatography will be required. The presence of some known impurities or inhomogeneity can be accounted for during particle picking and 2D classification, but the greater the presence of these the more difficult and computationally demanding the data processing will be, especially if the impurities bear some resemblance to the target protein. Even with a sample containing only the target protein, if the protein is present in multiple conformations, then this can complicate data processing further as mobile regions will have to be separated into separate classes or appear as a low-resolution blur, although if multiple conformations are possible to resolve well then this can be beneficial. Techniques such as circular dichroism (CD) or functional assays can be performed prior to microscopy to judge whether the protein is correctly folded and functional, as otherwise, the resulting structure may not be physiologically relevant. Finally, buffer conditions are very important, especially when dealing with membrane proteins, as some buffer components such as glycerol or detergents can increase background noise, and empty micelles can be difficult to tell apart from those containing protein particles (Gewering *et al.*, 2018), although addition of a small amount of detergent, below the CMC, has been shown to sometimes improve samples by shielding the protein from denaturation at the air-water interface (Chen *et al.*, 2019). Another recent advance in sample preparation in order to allow the cryo-TEM analysis of very small (sub 100 kDa) proteins is by rigidly binding the protein to larger macromolecules, such as Fab antibody fragments or maltose binding protein, which effectively increase the size of the entire structure, thereby increasing the SNR, although this involves producing antibody fragments specific to the protein and is therefore a significant investment of resources (Wu and Rapoport, 2021).

2.4.5. Negative stain electron microscopy

Negative stain electron microscopy is an invaluable method of determining the quality of a sample for electron microscopy without the added difficulty and variables of producing a vitrified specimen. As briefly mentioned in section 2.4.1, the purified protein sample is

added to a carbon film-covered grid, usually made of copper, which is then coated in a heavy metal stain such as uranyl acetate or formate, which coats the grid and stains the areas where protein is absent more than it stains the protein itself, creating contrasting lighter regions where the protein is, hence 'negative' stain. The heavy metal atoms scatter electrons much more strongly than those in a protein, meaning that particles are easily visualised compared to protein in vitreous ice. This is useful for early screening as the general shape and size of particles can be judged by eye to see if the purification has produced particles resembling the expected appearance of the target protein. Issues such as aggregation can be identified, and modified samples quickly analysed again to see if problems have been overcome. Also, as little as a tenth of the sample concentration is required compared to cryo-TEM, which is especially useful for challenging purifications when samples are very precious and early on in the optimisation of a protocol. The vastly increased contrast between background and particles means that fewer particles are needed to produce a dataset, and said dataset can be processed faster, while data can be collected on less advanced screening microscopes with lower voltages.

Negative stain datasets are useful for judging the identity of proteins in a sample, as well as whether it is forming oligomers or is in multiple conformational states; however, due to the size of the grains of stain coating the outside of the protein the resolution is generally limited to greater than 20 Å, and the highly acidic, non-aqueous environment means that negative stain datasets are further from the native conditions of the protein than a cryo-TEM sample. Even though negative stain is not suitable for producing high-resolution protein structures, it can be used to great effect to improve sample conditions prior to more costly and complex cryo-TEM.

2.4.6. Cryo-TEM grid choice and conditions

A vast variety of grids are now available for cryo-TEM, and each protein will have different properties that make a particular grid more suitable for sample preparation. These grids are approximately 3 mm diameter discs made up of a mesh base and a foil, with the mesh typically made of copper, nickel, or gold, as these materials are able to efficiently conduct electrons away from the sample to reduce radiation damage, dissipate heat well, and are mechanically stable to reduce damage when handling the grid (Karuppasamy *et al.*, 2011), with copper being the most common due to its balance of low cost and high conductivity. The density of the mesh can be varied, generally between

200 to 400 grid bars per square inch, with lower density giving more free area to visualise protein but also less mechanical support. The foil is layered on top of these metal grids and has holes in which the aqueous sample will be suspended before freezing; this foil can vary in both material and the pattern of holes. Most grids have foil made of amorphous carbon, which can be useful as it allows for the visualisation of Thon rings in Fourier space, from which microscopes can be easily aligned by the user (McMullan, Vinothkumar and Henderson, 2015). The layer of carbon can be continuous for even ice layer formation (although this has the drawback of inelastic scattering from the film, adding background noise, and reducing contrast between the protein and background), 'holey' carbon films with holes of a range of sizes and shapes, and 'lacey', which has larger, very irregularly sized and shaped holes, providing a large amount of open area, useful for long protein complexes such as microtubules (Consolati *et al.*, 2020). The drawback of holey and lacey grids is that the irregular shape and position of the holes mean that data collection cannot be automated, which severely limits the length of data collection possible. This challenge was overcome by the production of Quantifoil carbon support films, which have a regular, predefined arrangement of identically sized and shaped holes, meaning that data collection can be automated and run for extended periods with less operator input. The pattern of holes is categorised by the spacing between each hole and the diameter of the hole itself, for example, the R1.2/1.3 grids used in this study have a hole radius of 1.2 μm and a hole spacing of 1.3 μm; larger holes mean that more images can be taken per hole without having to move to another hole and refocus, so data collection is faster, but beam-induced particle motion can be more severe, leading to lower resolution data, while larger distances between holes means less imageable area but a stronger grid.

A more recent advance in grid construction is the UltrAuFoil grid, which instead of having a metal mesh and carbon foil uses gold for both elements of the grid. One of the advantages of these is that the gold foil is more adept than carbon at conducting electrons and dissipating heat, reducing radiation damage to the protein and giving added protection from ice melting (Russo and Passmore, 2016). More importantly, though, using a mesh and foil of differently conducting materials means that the copper or gold mesh and carbon foil contract differentially when frozen, leading to buckling or wrinkling of the grid surface and increased beam-induced motion. This problem is not present when using an entirely gold grid, with the downside of the loss of Thon rings used for quick

microscope alignment and significantly greater expense than carbon film grids. Gold grids have been optimised even further with the invention of HexAuFoil® grids, which have been developed to further reduce buckling of the suspended ice layer by modifying the pattern of holes, leading to the reduction of particle movement to less than 1 Å (Naydenova, Jia and Russo, 2020), although these are not yet available for general purchase.

Whatever the grid used and whether performing negative stain or cryogenic electron microscopy, before deposition of the sample the grid must be treated by plasma cleaning. This is partly to remove any contaminants present on the grid, but also because the untreated surface of the foil can be hydrophobic, meaning that an aqueous sample will not adhere to it and particles will not enter the holes. Plasma cleaning involves placing the grids in a low-pressure vacuum and subjecting them to ionised plasma, which in the case of ambient air gives the grid an overall negative charge due to the deposition of ionised oxygen and water molecules, making the grid inherently hydrophilic and therefore wettable with a water-soluble protein or solubilised membrane protein. Different gases can be used to vary the effect of the treatment depending on the needs of the target protein, such as amylamine to deposit a positive charge, and the time and strength of the treatment can also be optimised to different grid types.

Unfortunately, even with this treatment protein particles can be attracted towards the grid surface rather than the holes leading a reduction in visible particles, a particular issue when dealing with naturally hydrophobic membrane proteins, or particles can be drawn towards the air-water interface leading to denaturation. Grid treatment methods have been developed to combat this, generally either by chemically encouraging the particles towards the holes of the grid or by depositing a continuous film onto the grid which the particles more readily adhere to. As well as the addition of low concentrations of detergents mentioned in the previous section, another example of the first method is ‘PEGylation’ (Meyerson *et al.*, 2014), where a layer of PEG-thiol is deposited onto the grid, which is able to self-assemble into a monolayer. This has a similar effect to plasma cleaning in that it increased the wettability of the grid via the polyethylene glycol (PEG) group, which is highly hydrophilic, directing the protein into the holes of the grid while reducing interactions between the protein and the mesh or foil of the grid itself. A more common method of improving the number of particles visible in the holes is the second strategy, deposition of a continuous film over the grid, which can also be helpful in

directing particles towards different preferred orientations to compensate for preferred orientation of another view in a complementary dataset, or for limiting air denaturation (Noble *et al.*, 2018). Commonly this is a layer of amorphous carbon, which ideally will be as thin as possible (10 to 100 Å thickness), which the protein will adhere to and therefore be more evenly spread across the entire surface of the grid, including the holes. This is also helpful if only small quantities of protein can be produced as it has the effect of increasing the proportion of particles which end up on the grid, although has drawbacks as the film increases background noise, reducing the SNR; this is especially disadvantageous when studying small proteins that already have a low SNR. An improvement over amorphous carbon films is the deposition of a graphene support; this forms a monolayer a single atom thick over the grid which produces negligible background noise compared to an amorphous carbon layer, and has the added bonus of being an extremely good conductor of electrons to reduce radiation damage (Heersche *et al.*, 2007). This can be very helpful with small, low concentration or asymmetric proteins; recently, a graphene monolayer was used to promote the formation of uniformly thick ice, leading to a 2.2 Å structure of the 52 kDa protein streptavidin, along with a 2.6 Å structure of the asymmetric 67 kDa α-fetoprotein (Zheng *et al.*, 2023). Unfortunately this also has its disadvantages as the process of graphene deposition is more complex than amorphous carbon layers (Passmore and Russo, 2016; Naydenova, Peet and Russo, 2019), and the graphene is itself hydrophobic, meaning further treatment is required after deposition. An alternative is using a film of graphene oxide (Palovcak *et al.*, 2018), which is more easily handled and has inherent hydrophilicity due to the oxide functionalisation, although there can be issues with reproducibility of monolayers, meaning higher background noise than graphene. Finally, films can be functionalised to bind to the target protein, whether by using an affinity tag which has already been used for purification, such with Ni-NTA groups on the film, or even with antibodies specific to the protein (Llaguno *et al.*, 2014; Yu, Li and Jiang, 2016).

2.4.7. Vitrification

One of the major hindrances to high-resolution structural data of proteins is not the quality of the sample or the data processing, but the production of vitrified grids. Despite all the advances in electron microscopy over the last decade, the process of vitrification of the sample is surprisingly similar to that originally developed by Jacques Dubochet in the 1980s; after glow-discharging or other grid treatments, the protein sample is applied to

the grid via pipette, saturating the holes, before blotting of excess liquid with filter paper to make a thin layer (as thin as 50 nm can be required for a 2 μm diameter hole (Rice *et al.*, 2018)) and rapid plunging into liquid nitrogen-cooled ethane (or ethane/propane mix) so that the ice freezes so quickly that crystals are not formed. Reproducibility has been improved by the availability of plunge freezing machines such as the FEI Vitrobot or Leica EM GP2, however even with the ability to set reproducible parameters such as temperature and humidity control to reduce evaporation or specific blotting time and force, there are still a huge number of variables which are much more difficult to control for, such as how long the sample is on the grid before the operator can start the plunging sequence, uneven application of the sample to the grid, or the specific biochemical attributes of a protein leading to different rates of adsorption or denaturation (Glaeser *et al.*, 2016). The fact that so many iterations of this stage may be needed before a suitable sample is produced, and that excess sample is needed to saturate the holes in the grid (typically 3-5 μL is applied) means that the decreased concentrations required versus X-ray crystallography can be negated by how much protein is used up in vitrification trials.

There have been some recent innovations to improve the quality of grid production in standard machinery, such as using multiple rounds of sample application and blotting for samples with low particle density (Snijder *et al.*, 2017) as well as the ability to blot from only one side or from the opposite side of the grid to where the sample has been applied, useful for structures such as microtubules or filaments (Jakobi *et al.*, 2020), however methods which remove the blotting step altogether appear to be showing great promise in improving reproducibility and reducing the amount of sample needed per grid, and even allowing time-resolved microscopy, where rapid mixing and freezing means that states that only exist for a matter of milliseconds can be viewed. Some examples of these include: Spotiton (Jain *et al.*, 2012; Dandey *et al.*, 2018), which sprays picolitre to nanolitre volumes of sample through a device similar to an inkjet, and has been used to study transient molecular states by spraying subsequent droplets of sample onto a grid with highly specific mixing times before freezing; the use of acoustic waves to generate an ultrasonic spray of femtolitre sized droplets (Ashtiani *et al.*, 2018); and microfluidic chips which are able to spray sample at a consistent thickness using varying gas pressure (Feng *et al.*, 2017; Kaledhonkar *et al.*, 2018). These methods often use specialised self-wicking grids, which have $Cu(OH)_2$ nanowires on coating the grid bars that act as a sponge, thinning the layer of aqueous sample by capillary action (Wei *et al.*, 2017). This

creates a more uniform layer of ice, as each hole is undergoing the same wicking action, versus traditional filter paper blotting which typically creates a gradient of ice thickness across the grid. The fact that the level of blotting is proportional to the wicking time (the time between application of sample and freezing) coupled with the millisecond timescale of these techniques means that ice thickness can also be tuned far more specifically than with filter paper blotting. As well as improving reproducibility and reducing sample volumes, the incredible speed of these techniques drastically decreases the amount of time that the protein can be interacting with the air-water interface, thereby reducing the rate of denaturation events and improving the quality of the sample along with permitting the capture of millisecond scale molecular events (Darrow *et al.*, 2019; Dandey *et al.*, 2020).

Plunge freezing itself presents difficulties as it is very easy to damage or contaminate the frozen sample after plunging, for example by touching it to the sides of a grid box when dropping the grid in for storage, or by taking too long to deposit the grid leading to melting and re-freezing of the ice- even the users breath is a possible issue, as moisture in the air can contaminate the grid with ice crystals. A new technique called jet vitrification seeks to remedy this; in this process, samples are deposited on a pre-clipped 'autogrid', which is a grid clipped into the jet freezing device by a copper ring. A sub-nanolitre volume of sample is deposited on the grid by a method called pin printing, where a pin is dipped in a solution of protein sample and then deposited as a thin layer by moving it at close range over the surface of the grid. Jets of cryogenic fluid, rather than a reservoir of ethane as in plunge freezing, are then sprayed onto either side of the grid, evenly and rapidly freezing the grid while still held in the autogrid holder. Along with the rapidity and uniformity of this method, the entire process can be performed automatically, including loading into microscopes equipped with autoloaders, removing the risk of the grid melting before data collection (Ravelli *et al.*, 2020).

2.4.8. Cryo-TEM screening

Due to demand for data collection time on high powered microscopes such as the 300 keV Titan Krios, the range of grid and vitrification conditions will generally first be screened on a less powerful, higher throughput microscope, such as an Tecnai T12 (120 keV) or Glacios 2 (200 keV). This can be another bottleneck, as significant trial and error is involved in trying different purification, grid, and vitrification conditions before visualising each sample in the screening microscope, then going back to the sample

preparation steps for optimisation based on parameters such as the appearance of particles (whether they are the appropriate size and shape), optimal ice thickness, and particle density. Small, preliminary datasets can also be obtained at this point to perform early data processing and determine particle quality. Reproducibility of conditions becomes very important here, especially if the screening microscope available does not have an autoloader (making grids difficult to reuse after screening), as once a sample has been produced that looks suitable for data collection on a more powerful microscope, an identical grid will need to be supplied. While this remains a complex and lengthy procedure, high-throughput grid screening through automated data collection and machine learning methods, such as those used by Smart Leginon (Cheng *et al.*, 2023) are aiming to make it more convenient through more rapid assessment of grids and reduced operator time.

2.4.9. Data collection

As discussed in section 2.4.1, hardware such as direct electron detectors and more powerful microscopes, along with advances in dedicated data processing software such as RELION (Scheres, 2012) and cryoSPARC (Punjani *et al.*, 2017) have significantly increased the availability of high-resolution data from cryo-TEM. Data collection, typically performed on a 300 keV Titan Krios microscope, involves thousands of movies being taken of the sample over hundreds or thousands of holes, each lasting a few seconds and consisting of 40 or so frames, so that a balance is struck between each frame having a low electron dose (limiting sample damage) and enough data being gathered. Images are taken at a range of defocus values to account for the loss of spatial frequency data at each defocus, as discussed in section 2.4.3. Due to the low electron dose applied, usually about 10 e^-/A^2, each of the frames of these movies has a low SNR for each particle against the ice background.

2.4.10 Pre-processing and particle picking

Once the data has been collected, the first step is to correct for motion over the movie, which can be caused by drifting of the sample stage or by energisation of the sample by the electron beam (anisotropic deformation). This is performed using software packages such as MotionCor2 (Zheng *et al.*, 2017), or the patch motion correction job within cryoSPARC, which split each micrograph up into small patches to account for variations in ice thickness and track the movement of each particle over the movie, without requiring

an input of particle locations, allowing the movie stacks to be aligned into single micrographs. The next step is patch CTF estimation, where automated software such as CTFFIND4 (Rohou and Grigorieff, 2015) or Gctf (Zhang, 2016) transforms multiple areas of each micrograph into Fourier space before estimation of the CTF and correction for negative phase contrast and loss of spatial frequency linked to the envelope function. CryoSPARC produces plots of cross-correlation of the computed CTF versus the observed power spectrum of each micrograph, showing the frequency at which these no longer fit each other well (judged as a cross-correlation threshold of 0.3), which can be used to judge the quality of micrographs. This can be combined with other diagnostic information such as estimated ice thickness (Olek *et al.*, 2022) and the extent of particle motion in a micrograph to remove those which are not expected to contain valuable data i.e., a low resolution CTF fit or very high motion of particles. Motion correction and CTF estimation are together known as pre-processing and are a necessary first step for every data processing workflow.

Once unsuitable micrographs have been removed from the dataset, the next step is particle picking. There are multiple methods to go about this, along with multiple software packages in which to perform it, such as EMAN2 (Tang *et al.*, 2007), RELION, or within cryoSPARC. Initially particle picking will be performed without inputting a reference, as this can introduce bias towards a model, for example by using the blob picker in cryoSPARC, where the user inputs an estimated diameter range for the target particle and every blob of the appropriate size will be picked. Particles can also be selected by eye (although this can be difficult with membrane protein particles that do not look particularly different to junk, such as empty micelles or empty nanodiscs) and then used to train machine learning software such as TOPAZ (Bepler *et al.*, 2019) to continue picking based on this initial input. The picked particles can then be inspected to remove those that appear to be junk en masse, before ‘extraction’ along with their CTF data for further processing. High-resolution CTF data, which is spread out to high spatial frequencies, can require extraction from a box size more than twice the diameter of the particle itself, however for initial processing where high spatial frequency data is not as important, the box size can be smaller, or the particles can be binned to reduce the computational load, as especially for a small membrane protein many millions of particles can be picked.

2.4.11. Classification and refinement

The next step is 2D classification, where each extracted particle projection is correlated with other projections through rotation and translation and aligned with similar-looking particles to produce 2D class averages representing different views or conformations, which due to the accumulation of signal can have more visible features than individual projections. Particles representing low quality classes, such as obvious ice contamination, background noise, or aggregates can be removed at this step and new iterations of 2D classification performed on the leftover 'good' particles. At this step in cryoSPARC, classes representing top and side views can be used as templates to run a 'template picker' job, where the programme will go back to the micrographs and only pick particles that match these chosen templates, leading to better quality picks, and further rounds of 2D classification.

Once satisfied with the 2D classes, these 'good' classes can be combined and used to produce a 3D reconstruction through different methods. One is through 'back-projection'; as mentioned earlier, the particles seen in cryo-TEM data are not 2D images of each particle, but a projection through the entire protein; ideally, all orientations of the protein will have been projected onto the image plane, and therefore all of these projections can be aligned and combined in Fourier space to reconstruct the 3D detail of the protein. This is performed on a per-particle basis through a process called projection matching (Estrozi and Navaza, 2008), where the Fourier transform of each projection is matched up against a Fourier transform of a 3D reference volume, usually of a closely related protein of which structural data is already available. When the transformed 2D projection matches the reference, it is added to the 3D Fourier transform, iteratively adding more data and producing a more detailed reconstruction. The 3D volume produced is then transformed back into real space to be visualised as a 3D Coulomb potential map. The whole processes is automated according to a probabilistic distribution which is dependent on the level of noise in the sample, known as maximum-likelihood (Sigworth, 1998). As this method involves a user-provided reference, it is unfortunately susceptible to bias towards the reference structure. To circumvent this, the cryoSPARC programme contains an algorithm for unsupervised '*ab initio*' 3D reconstruction (Punjani *et al.*, 2017). This allows the generation of 3D volumes, including of multiple conformational states, without prior knowledge of what the protein is expected to look like or providing an initial structure, and without making assumptions about the relative resemblance of each of these states.

This is done using an algorithm called stochastic gradient descent (SGD), which, in simple terms, uses an arbitrary, computer-generated random initialisation map which is incrementally improved over many 'noisy' steps. This random initialisation will be judged against randomly selected subsets of particle images, with each iteration producing a new map that is more likely to explain the particle images fed into it, and if told to produce multiple classes, can build multiple maps of different conformations based on image data. This produces a relatively low-resolution map (i.e., >5 Å, not enough to assign the positions of amino acids), which, if certain orientations are underrepresented, will show as smearing or blurred volume, which can be improved by further data collection to fill in the gaps, such as using a tilted sample during acquisition or by grid conditions that favour certain orientations. Beyond avoiding bias to a user-inputted reference, it also has the benefit of being less computationally intensive than other cryo-TEM software packages as the programme only deals with small groups of images for each iteration, reducing the cost of computer hardware needed and the time taken to analyse EM data.

Before high resolution can be reached, this structure needs to go through a multi-step process of refinement. The next step in cryoSPARC is often heterogeneous refinement; this is where particles from an initial volume (i.e., the volume produced in *ab-initio* reconstruction) are simultaneously separated into different classes as the same time as refining these classes to higher resolution, allowing for high-resolution differences between particles to be found; this is particularly useful with proteins that have mobile regions or cycle through multiple states, but can also be used to clear out junk particles, such as by differentiating between filled and empty MSP nanodiscs. Once the user is satisfied that all conformational states or junk particles have been separated out, the resultant volumes can be put through homogeneous refinement, which uses another algorithm called a 'branch-and-bound' approach. This aims to find an optimal pose for a 3D map to match with the supplied images, iteratively eliminating poses that exceed a certain level of error until the best candidate pose, in other words the most accurate Coulomb potential map, is produced. Version 3 of cryoSPARC introduced 'non-uniform refinement' (Punjani, Zhang and Fleet, 2020), a method of refinement specifically tailored to small, low SNR or membrane proteins, which can account for regions of disordered or

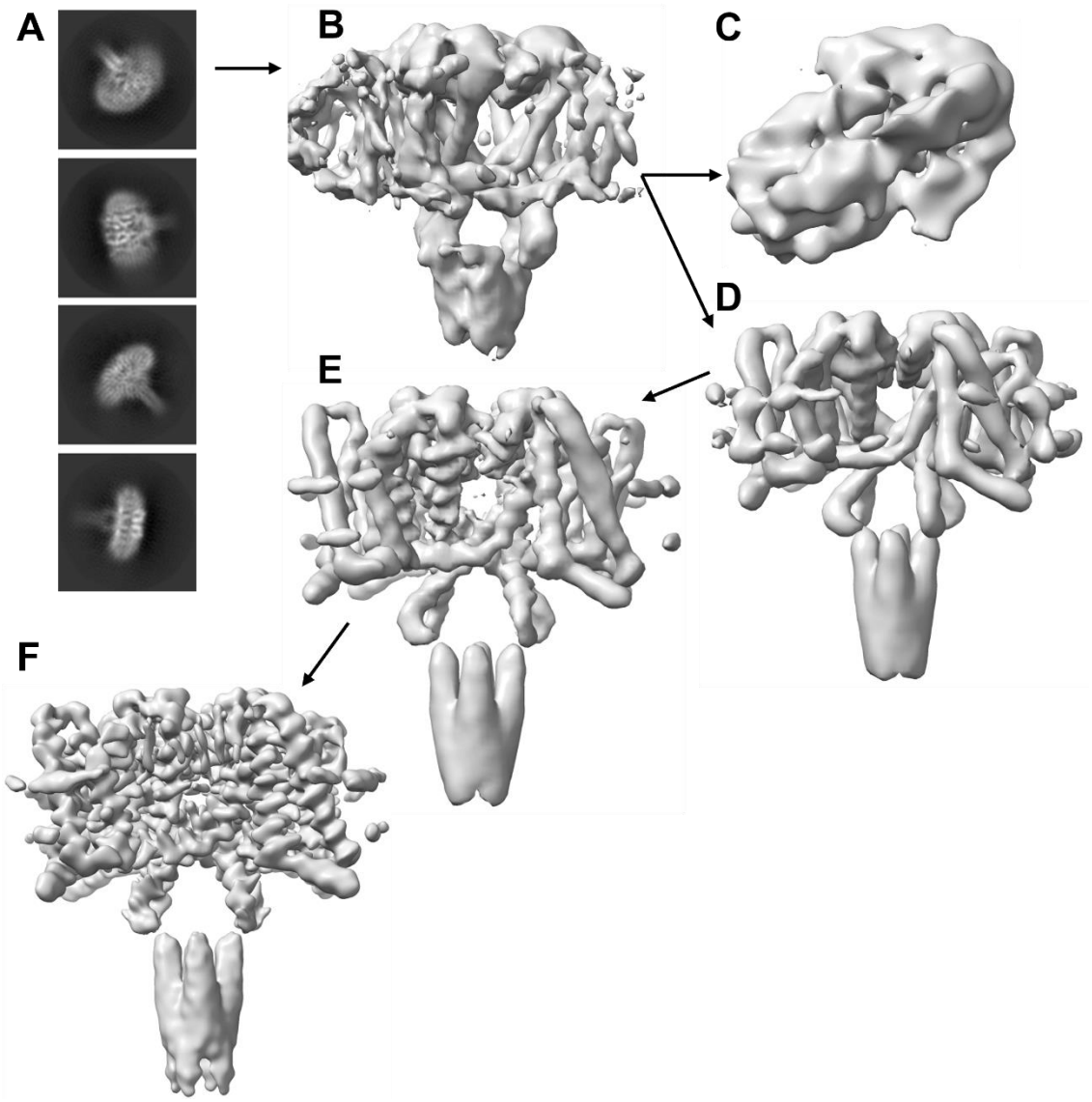

Figure 2.5. Simplified example of a classification and refinement pipeline. (A) shows 2D selected to produce the *ab initio* 3D reconstruction **(B)**. This undergoes heterogeneous refinement so separate out junk particles, which give volume **(C)**, and good particles, which produce volume **(D)**. Multiple rounds of non-uniform refinement of **(D)** give a more detailed (i.e., higher resolution) reconstruction **(E)**. Particles are extracted again with a larger box size to obtain higher spatial frequency data, and a new *ab initio* reconstruction undergoes multiple rounds of non-uniform refinement to give the final reconstruction **(F)**.

flexible density, such as micelles or lipids in nanodiscs, by removing unhelpful noisy signal from disordered areas of the protein whilst keeping signal which can be used for aligning particle images.

Another refinement step which can be used is local refinement; if one domain or region of a protein is flexible of mobile with reference to another, this can interfere with refining the entire structure as one body. A 'mask' can be produced which covers a specific sub-

volume of the protein, and that masked area alone can then be refined without reference to the area outside the mask (or vice versa) without interference from the relative motion of each region. This can be performed separately on each sub-volume and then the subsequent maps combined to create a map with improved global resolution, although it is generally not viable for proteins less than 100-150 kDa as the sub-volumes will be too small to align effectively.

Once the map is refined to a suitable resolution, macromolecular modelling software such as Coot (Emsley *et al.*, 2010) or Phenix (Liebschner *et al.*, 2019) can be used to produce a model by building the sequence of the target protein into the map.

2.4.12. Validation

To determine whether a 3D reconstruction of a protein actually contains high enough resolution data to make assignments of details such as the position of amino acid residues, the model needs to be validated. The need for stringent validation of models produced from alignment and averaging of many particles is best demonstrated by the "Einstein from noise" problem, highlighted in a 2013 paper from Richard Henderson (Henderson, 2013). The issue is that if provided with enough images, a template-based computational particle picking programme will be able to find areas of image that it considers matches for the template (in this case a picture of Albert Einstein), amongst white noise, with no particles present. This problem is made worse with high background noise or with smaller proteins, as there is inherently a lower SNR, meaning a computer will be less able to discern between particle and noise. This highlighted the need for scrutiny of electron microscopy data when automated, template-led processes are involved, and encouraged structural biologists not to take their data at face-value.

The current method for determining the resolution of an Coulomb potential map produced by cryo-TEM has been developed to be comparable with structures determined by X-ray crystallography. This is called the gold-standard Fourier shell correlation (FSC) and involves taking the particles used to create a reconstruction and randomly splitting them into two equally sized datasets, or 'half-maps'. These half-maps are cross-correlated with each other in Fourier space at increasingly higher spatial frequencies, until they reach a frequency where the correlation drops a value of 0.143 (although some prefer a more stringent value of 0.5) (Scheres and Chen, 2012); any data at higher spatial frequencies than this is considered indistinguishable from noise. Though widely adopted this method

does have its limitations, for example a global resolution is usually assigned to the protein as a whole, even though there will often be a range of local resolutions over the structure, being higher around an axis of symmetry or lower in flexible regions- unscrupulous researchers can use this to boost claims about structural details in areas of the protein that the data may not actually show. It is therefore paramount to be critical of all EM data and not take claims of resolution or structural detail at face value.

Chapter 3

Materials and Methods: Purification and Electron Microscopy Analysis of hNav1.8 Recombinantly Expressed in *Tetrahymena thermophila*

3.1. Chapter summary

The following chapter describes the materials and methods used in the expression, purification, and analysis of hNav1.8 purified from the protozoan *Tetrahymena thermophila*, along with descriptions of some of the techniques used.

3.2. Bioinformatics

Uniprot accession codes used were as follows: NavMs (A0L5S6), NavAb (A8EVM5), NaChBac (Q9KCR8), hNav1.1 (P35498), hNav1.2 (Q99250), hNav1.3 (Q9NY46), hNav1.4 (P35499), hNav1.5 (Q14524), hNav1.6 (Q9UQD0), hNav1.7 (Q15858), hNav1.8 (Q9Y5Y9), hNav1.9 (Q9UI33) (The Uniprot Consortium, 2023).

Data such as the molecular mass of proteins, their theoretical isoelectric point (pI) and their extinction coefficients were calculated from amino acid sequences using the ExPASy ProtParam tool (Gasteiger *et al.*, 2005). The Clustal Omega tool was used to perform multiple sequence alignments, with default settings including mBed-like clustering enabled for more rapid guide tree generation (Sievers *et al.*, 2011). The EMBOSS Needle tool (Rice, Longden and Bleasby, 2000) was used to perform pairwise sequence alignments, with the BLOSUM62 scoring matrix and default gap open and gap extend penalty values of 10 and 0.5 respectively. The ExPASy translate tool (Gasteiger, 2003) was used to translate sequenced plasmid samples to amino acid residues, before comparison with target protein via EMBOSS Needle.

Protein Data Bank (PDB) accession codes (Berman, 2000) of published voltage-gated sodium channel structures used in the analysis of results, alongside stated resolution,

sequence coverage, and where relevant Electron Microscopy Data Bank (EMDB) accession codes (Lawson *et al.*, 2016) are included in the following table (Table 3.1).

Table 3.1. Table of proteins used for analysis. Table shows the identities and relevant details of protein structure models used in the analysis of results in this thesis.

Protein	Drugs or toxins	PDB code	EMDB code	Reso-lution (Å)	Cove-rage	Reference
NavMs	Apo	5HVX	Crystal	2.45	92%	Sula *et al.*, 2017
NavMs F208L	Apo	6SX5	Crystal	2.20	96%	Sula *et al.*, 2021
NavMs F208L	4-hydroxytamo-xifen	6SXC	Crystal	2.50	98%	"
NaChBac in lipid nanodisc	Apo	6VWX	21425	3.10	82%	Gao *et al.*, 2020
hNav1.2	μ-conotoxin KIIIA, β2	6J8E	9780	3.00	56%	Pan *et al.*, 2019
hNav1.4	β1	6AGF	9617	3.20	61%	Pan *et al.*, 2018
hNav1.8	Apo	7WFW	32476	3.10	51%	Huang *et al.*, 2022
hNav1.8 class I	A-803467	7WE4	32439	2.70	57%	"
hNav1.8 class II	A-803467	7WEL	32451	3.20	56%	"
hNav1.8 class III	A-803467	7WFR	32475	3.00	56%	"

The AlphaFold Protein Structure Database was used to download a predicted structure for hNav1.8 made by the machine learning programme Alphafold2 (Jumper *et al.*, 2021).

Structural alignments and figures were made in PyMol molecular graphics software (version 2.0; Schrödinger, LLC) and ChimeraX (Pettersen *et al.*, 2021) (versions 1.5-1.7; Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco).

3.3. Materials and equipment

Below is a table of materials and equipment used in this thesis alongside suppliers.

Table 3.2. List of materials and equipment. Table showing the materials and equipment used in this thesis and from where they were obtained.

Item	Supplier
1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC)	Avanti Polar Lipids

20x 3-(N-morpholino)propanesulfonic acid (MOPS) running buffer	Invitrogen
3-((3-cholamidopropyl)dimethylammonio)-1-propanesulfonate (CHAPS)	Anatrace
4-15% Mini-PROTEAN TGX precast protein gels	Bio-Rad
4x Laemmli sample buffer	Bio-Rad
5 M sodium chloride (NaCl) solution	Invitrogen
Agarose gel powder	Sigma-Aldrich
AKTA purifier fast protein liquid chromatography (FPLC) system	Cytiva (formerly GE Healthcare Life Sciences)
Amicon Ultra-15 centrifugal filters (100 kDa cut-off)	Sigma-Aldrich
Amicon Ultra-4 centrifugal filters (100 kDa cut-off)	Sigma-Aldrich
Amphipol A8-25	Anatrace
Ampicillin	Melford
Anti-FLAG M2 affinity gel	Merck Millipore
Anti-hNav1.7 primary antibody	Donated by J.Wood lab
BenchMark protein ladder	Thermo Fisher Scientific
Bio-Beads SM-2 resin	Bio-Rad
CF300-Cu TEM support grids	Electron Microscopy Sciences
Chemically competent C41 *E. coli* cells	New England Biolabs
Chemically competent DH5α *E. coli* cells	New England Biolabs
Chemidoc MP imaging system	Bio-Rad
Cholesteryl hemisuccinate (CHS)	Anatrace
Choline chloride	Sigma-Aldrich
Clarity enhanced chemiluminescence (ECL) western blotting substrate	Bio-Rad
cOmplete mini EDTA-free protease inhibitor cocktail	Roche
Decanoyl-N-hydroxyethylglucamide (HEGA-10)	Anatrace
Econo-Pac chromatography columns	Bio-Rad
Emulsiflex-C3 high-pressure homogeniser	Avestin
FLAG peptide	Merck Millipore
Glycerol	Sigma-Aldrich
Glycine	Sigma-Aldrich
Glyco-diosgenin (GDN)	Anatrace
Goat anti-rabbit IgG horseradish peroxidase-conjugated secondary antibody	Invitrogen
Harrick plasma cleaner with PlasmaFlo gas mixer	Harrick Plasma
HisTrap HP 1 mL column	Cytiva (formerly GE Healthcare Life Sciences)
iBlot 2 dry blotting system	Thermo Fisher Scientific
Imidazole	Thermo Fisher Scientific

InstantBlue protein stain	Expedeon
Isopropyl-β-D-1-thiogalactopyranoside (IPTG)	Generon
Lidocaine	Sigma-Aldrich
Lysogeny broth (LB)	Formedium
Magnesium sulfate ($MgSO_4$)	Thermo Fisher Scientific
Membrane scaffold protein 2N2 (MSP2N2)	Sigma-Aldrich
NanoDrop ND-1000 spectrophotometer	Thermo Fisher Scientific
N-cyclohexyl-3-aminopropanesulfonic acid (CAPS)	Sigma-Aldrich
N-dodecyl-β-D-maltopyranoside (DDM)	Anatrace
Ni sepharose 6 fast flow gel	Cytiva (formerly GE Healthcare Life Sciences)
NuPAGE 4-12% Bis-Tris Gel	Invitrogen
NuPAGE LDS 4x sample buffer	Invitrogen
PageRuler Plus pre-stained protein ladder	Thermo Fisher Scientific
PELCO easiGLOW glow discharge cleaning system	Ted Pella Inc
Peristaltic pump	Cytiva (formerly GE Healthcare Life Sciences)
Phenylmethylsulfonyl fluoride (PMSF)	VWR International
Phosphate-buffered saline (PBS) tablets	Sigma-Aldrich
QuantiFoil Cu R1.2/1.3 300 mesh holey-carbon grids	Electron Microscopy Sciences
SIGMAFAST 3,3'-diaminobenzidine (DAB)	Sigma-Aldrich
Skimmed milk powder	Thermo Fisher Scientific
Sodium chloride (NaCl)	Thermo Fisher Scientific
Super optimal broth with catabolite repression (SOC) outgrowth medium	New England Biolabs
Superdex 200 Increase 10/300 GL column	Cytiva (formerly GE Healthcare Life Sciences)
Superose 6 Increase 10/300 GL column	Cytiva (formerly GE Healthcare Life Sciences)
Syringe filters	Sartorius Biotech
T. thermophila cells expressing hNav1.8	Tetragenetics Inc
Tecnai T12 electron microscope	Thermo Fisher Scientific
Thrombin (restriction grade) and 10x cleavage buffer	Merck Millipore
Titan Krios microscope (K3 electron detector)	Thermo Fisher Scientific
Tobacco etch virus (TEV) protease	Expressed and purified by Dr D. Hollingworth (Wallace lab)
Tris(hydroxymethyl)aminomethane (tris base)	Thermo Fisher Scientific
TURBO DNAse I	Sigma-Aldrich
Tween-20	VWR International
UltrAuFoil R1.2/1.3 300 mesh holey-gold grids	Electron Microscopy Sciences
Uranyl acetate	Polysciences, Inc

Vitrobot Mark IV	Thermo Fisher Scientific
Wizard Plus SV Miniprep kit	Promega
α-poly His horseradish peroxidase-conjugated antibody	Invitrogen
β-mercaptoethanol	Sigma-Aldrich

3.4. Constructs and cloning

The full-length, wild type (WT) hNav1.8 construct, first identified and sequenced in the lab of Professor J. Wood (Akopian, Sivilotti and Wood, 1996), was modified with a C-terminal 10x histidine-tag (His-tag) and FLAG tag (sequence DYKDDDDK) and cloned into *T. thermophila* by our collaborator Tetragenetics Inc.

3.5. Protein expression and purification

3.5.1. Expression of hNav1.8

In this project, the single-celled protozoan *T. thermophila* was selected as the expression system as it is eukaryotic, so would be more able than *E. coli* to fold the protein correctly, but also because of its large membrane surface area, increased by the incorporation of 18-21 rows of cilia into upwards of 750 'membranelle' structures (Bayless, Galati and Pearson, 2015; Bayless, Navarro and Winey, 2019), as seen in Figure 3.1.

Preliminary research by our collaborator Tetragenetics Inc showed that membrane proteins expressed in the system could be detected by western blot in the cilia fraction of cells; ideally, this increased membrane area could mean that the cell has a greater capacity for the expression of membrane proteins. As a single-celled aquatic organism, *T. thermophila* can also be conveniently cultured in liquid medium, similarly to *E. coli*.

The expression itself was performed by Tetragenetics Inc: *T. thermophila* was genetically modified to express recombinant hNav1.8, expressed in the presence of 1 mM lidocaine, and pelleted and flash frozen in 1x phosphate buffered saline (PBS), then kept at -80 °C upon receipt until ready for purification.

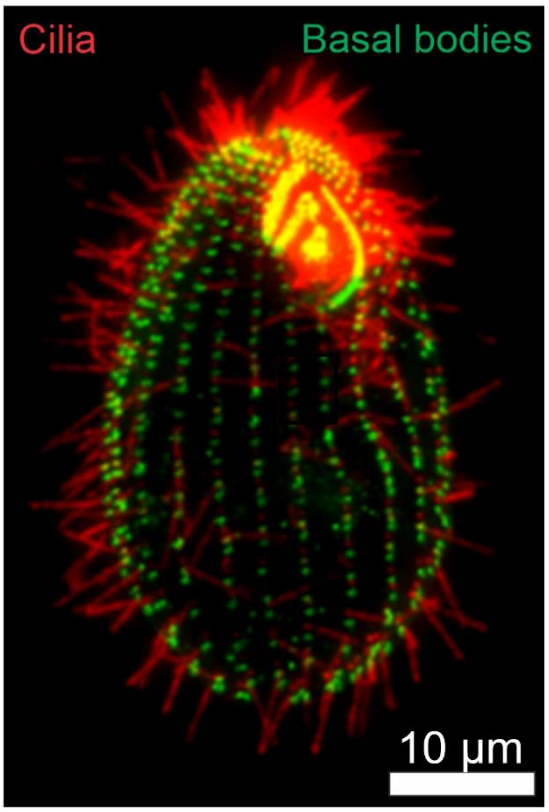

Figure 3.1. Immunofluorescence image of *Tetrahymena thermophila*. *T. thermophila* is stained for cilia (red) and basal bodies (green). *T. thermophila* possesses 18-21 rows of cilia anchored in basal bodies, visible over the surface of the cell, which are encompassed by the cell membrane. Figure adapted from Bayless, Navarro and Winey, 2019.

3.5.2. Purification of hNav1.8

Preliminary work by Dr A. Sula of the Wallace group established a general protocol for purification of recombinant hNav1.8 from *T. thermophila* (see section 5.2). Different combinations of FLAG, His, and strep affinity tags were tested, with strep being ruled out due to issues with solubility. N-dodecyl-β-D-maltopyranoside (DDM) was chosen for solubilisation, and glyco-diosgenin (GDN), a synthetic analogue of digitonin, was used during purification due to its success in purifying hNav1.4 (Pan *et al.*, 2018). Purification by FLAG affinity, then Ni^{2+} affinity, and finally size-exclusion chromatography (SEC) was shown to reliably produce hNav1.8 (confirmed by western blot and solution mass spectrometry) although in small amounts, and cryo-transmission electron microscopy (cryo-TEM) of the sample did not yield suitable results for high-resolution structural analysis. This protocol, hereafter referred to as the 'standard protocol,' was as follows.

The entire purification was performed at 4 °C. Cell pellets from 6 L of growth medium were thawed and resuspended in 300 mL lysis buffer (1x PBS, 50 mM choline chloride, 10% glycerol, 1 mM lidocaine, 0.75 µL/mL TURBO™ DNase, and cOmplete™ mini EDTA-free Protease Inhibitor cocktail tablets at pH 7.4) and lysed by Dounce homogeniser. Lysed cells were centrifuged for 40 min at 20000 x *g*, giving a pellet (containing membrane fraction), supernatant, and viscous white material settled on top. The white material and supernatant were removed by pipette, and the pellet resuspended in 100 mL extraction buffer (lysis buffer with 1% DDM (115x CMC of 0.0087%), minus DNase), before incubation at 4 °C for 2 h and centrifugation for 1 h at 235000 x *g*. After ultracentrifugation, more of the viscous white material was removed by pipette. The supernatant (minus this material) was filtered with a 1.2 µm syringe filter, incubated with 4 mL anti-FLAG M2 affinity resin overnight at 4 °C, then loaded into an Econo-Pac gravity chromatography column. Incubated gel beads were washed with 50 column volumes of FLAG buffer (extraction buffer with 0.075% GDN instead of DDM, 35x CMC of 0.0021%). Target protein was eluted in a two-step process; in the first step, 350 µL of 5 mg/mL FLAG peptide stock was added to 4 mL of FLAG buffer, and M2 gel beads were incubated in this for 60 minutes with occasional mixing before elution. In the second step, M2 gel beads were mixed with 16 mL FLAG buffer containing 150 µL FLAG peptide stock and incubated and eluted in the same manner. The two eluents were combined and incubated with 0.5 mL Ni^{2+} Sepharose 6 Fast Flow resin overnight at 4 °C in Ni^{2+} binding buffer (FLAG buffer plus 300 mM NaCl and 50 mM imidazole) before

loading onto a new gravity chromatography column, where gel beads were washed with 10 column volumes of Ni^{2+} binding buffer and 3 column volumes of Ni^{2+} wash buffer (Ni^{2+} binding buffer with 100 mM imidazole) before elution with 5 column volumes of Ni^{2+} elution buffer (Ni^{2+} binding buffer with 500 mM imidazole). Fractions containing target protein were pooled and concentrated using 100 kDa cut-off Amicon centrifugal concentrators at 2500 x *g* (faster speeds caused precipitation). Precipitation was also reduced by raising the NaCl concentration by addition of 5 M NaCl solution up to a final concentration of 0.5 M NaCl. Samples were further purified by SEC using a Superose 6 Increase 10/300 column in SEC buffer (1x PBS, 1 mM lidocaine, 0.075% GDN, 300 mM NaCl, pH 7.4). Modifications to this protocol made with the aim of increasing sample quality are shown in results section 5.3.

3.5.3. Determination of protein concentration

The protein concentration after purification was measured using a Nanodrop UV spectrophotometer, measuring absorbance of 280 nm light (A_{280}) by the amino acids tryptophan, tyrosine, and cysteine. Concentration is calculated from absorbency by the Beer-Lambert Law:

$$A = \varepsilon c l \tag{3.1}$$

Where A = absorbance of sample, ε is the molar extinction coefficient of the protein in M^{-1} cm^{-1}, c is the molar concentration of the protein, and l is the pathlength of the light in cm. The extinction coefficient of a protein is calculated from the weighted sum of the molar extinction coefficients of each the tyrosines (Y), tryptophans (W), and cysteines (C) in its sequence, described by:

$$\varepsilon = (nW\ x\ 5500) + (nY\ x\ 1490) + (nC\ x\ 125) \tag{3.2}$$

Where n is the number of each residue present, and the values are their individual extinction coefficients.

3.5.4. Amphipol exchange

One method used to attempt to improve the sample of hNav1.8 for cryo-TEM over the standard protocol was to exchange it from GDN into A8-35 amphipol. For this purification, membranes were solubilised from *T. thermophila* by the same method as the standard protocol and purified by FLAG affinity as normal. The sample was then

concentrated by 100 kDa cut-off Amicon centrifugal filter to approximately 250 µL and incubated with A8-35 at a 1:5 protein to amphipol ratio overnight at 4 °C. The complexes were isolated by SEC using a Superose 6 Increase 10/300 GL column in buffer containing 1x PBS, 1 mM lidocaine, 300 mM NaCl, pH 7.4, removing excess detergent in the process.

3.6. Biochemical analysis of hNav1.8

3.6.1. SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Weber and Osborn, 1969) is a technique commonly used to separate proteins by molecular mass, which can then be directly stained in the gel to visualise the size of proteins present in the sample, or transferred for western blotting (see section 3.6.2). Protein samples are mixed with running buffer containing sodium dodecyl sulphate (SDS) and reducing agent, which denatures the protein so that it becomes linear and is uniformly negatively charged. Often samples will also be heated to e.g., 95 °C at this point to ensure they are fully denatured, although this is not possible for some membrane proteins as it can cause the sample to aggregate. After loading into a polyacrylamide gel, an electric potential is applied so that the proteins travel through the gel towards a positive electrode, and as they have been linearised they then migrate through the polyacrylamide lattice at distances proportional to their molecular mass (higher molecular mass, shorter distance travelled). A mixture of proteins of known mass are also run through alongside as a 'ladder' from which approximate molecular weight can be inferred.

For SDS-PAGE analysis of purified hNav1.8 samples were mixed with 4x sample buffer, 12 µL protein to 4 µL Laemmli sample buffer (made up in 1 mL batches of 900 µL 4x Laemmli sample buffer and 100 µL β-mercaptoethanol as reducing agent). These were not cooked for denaturation as the sample could not then run through the gel due to aggregation. Samples were then loaded into 10-well Bio-Rad Mini-Protean TGX 4-15% precast gels alongside Pageruler Plus prestained protein ladder, and gels were placed in a tank containing tris-glycine-SDS running buffer (25 mM tris base, 192 mM glycine, 0.1% SDS, pH 8.6). Gels were run for 45 minutes at 180 V. Gels were rinsed with distilled water and either stained with InstantBlue protein stain for 1 h before destaining with distilled water until bands were visible (all with rocking motion) or used for western blotting.

3.6.2. Western blotting

Western blotting (Burnette, 1981) is a technique where proteins separated by SDS-PAGE are transferred (blotted) onto a membrane, where they can subsequently be incubated with an antibody to identify whether a particular target protein is present. The SDS-PAGE gel is placed in a stack with a nitrocellulose or polyvinylidene difluoride (PVDF) membrane and a current is run across to draw the protein from the gel to the membrane. The current and blotting time must be adjusted for the sample for optimal transfer of protein.

In the case of hNav1.8, following SDS-PAGE the gels were blotted onto a PVDF membrane via iBlot 2 dry blotting system. The membrane must then be 'blocked' from non-specific binding of the primary antibody with the membrane itself by incubation (gentle shaking at 4 °C) in 1x phosphate-buffered saline (PBS) buffer with 3% w/v skimmed milk powder for 30 minutes. The membrane was then washed 3 times for 5 minutes each in 1x PBS before incubation overnight at 4 °C with 7 µL (1:2000 dilution in 1x PBS 1% skimmed milk powder) of a rabbit source antibody produced by and donated from the J. Wood lab at UCL, with specificity to hNav1.7 (54% sequence identity by pairwise sequence alignment in EMBOSS Needle). The solution was poured off and 7 µL of a secondary anti-rabbit antibody conjugated to horseradish peroxidase at 1:2000 dilution in 1x PBS was added and incubated at 4 °C for 1h. After secondary antibody incubation the membrane was washed 5 times for 5 minutes each with a solution of 1x PBS and 0.05% v/v Tween-20 detergent and developed with Bio-Rad Clarity ECL (enhanced chemiluminescence) Western Blotting Substrate, before imaging by ChemiDoc MP Imaging System.

3.6.3. Liquid chromatography-mass spectrometry

Whilst western blotting is useful for confirming the presence of a known protein for which an antibody is available, it cannot identify unknown impurities which are seen on an SDS-PAGE gel. Liquid chromatography-mass spectrometry (LC-MS) (Karpievitch *et al.*, 2010) is used to attempt to identify all proteins in a sample, by fragmentation into peptides and identification of the source of these peptides by comparison to a database of protein sequences. Samples can be analysed in solution or after separation by SDS-PAGE. In the case of SDS-PAGE samples, bands are cut out of the gel and subjected to digestion into peptides at specific residue sites by trypsin or chymotrypsin protease, before separation by liquid chromatography. The resultant peptides are ionised by a method called matrix-

assisted laser desorption/ionisation (MALDI) and accelerated; the distance the molecules travel is detected by a time-of-flight (TOF) mass spectrometer, and this time-of-flight used to calculate their mass/charge ratio (shorter time-of-flight meaning higher mass compared to charge). The specific site cleavage means that detection of different mass/charge ratios act as a 'fingerprint' for the protein, which can then be run against a database of protein sequences using the MASCOT search engine. Two important parameters that are given as results are a probabilistic score that the spectra of hits matches particular proteins (Perkins *et al.*, 1999), and the exponentially modified protein abundance index (emPAI) (Ishihama *et al.*, 2005), which estimates the abundance of each identified protein in the sample by a ratio of the number of peptides that would be expected to be detected (essentially proportional to the size of the protein) against the number experimentally found. This is used to judge the major species present in the sample, although this can be less accurate for membrane proteins as the hydrophobic, membrane-integral regions of membrane proteins sometimes are less likely to be detected by mass spectrometry, leading to an underestimate of the protein's abundance.

In this study, hNav1.8 sample purified by the standard protocol was separated by SDS-PAGE and stained (see section 3.6.1) before selected bands were cut out of the gel using a sterile scalpel and sent for trypsin digest and LC-MS analysis at the BSRC Mass Spectrometry & Proteomics Facility, University of St Andrews. Resulting spectra were searched against the MASCOT server to match with proteins in human-only and all-species protein sequence databases.

3.7. Electron microscopy screening of hNav1.8

3.7.1. Negative stain screening

Negative stain electron microscopy was used to screen the effects of different purification conditions and the quality of different SEC fractions of target protein. 3 μL of sample at approximately 0.1 mg/mL was added to glow discharged (PELCO easiGlow™, 30 second discharge) standard thickness (5-6 nm) carbon film 300 mesh copper TEM support grids and left to adhere for 30 seconds. The grid was submerged in two drops of Millipore filtered water to rinse off excess detergent, before quick submersion in one drop of 2% uranyl acetate (UA), blotting with ashless filter paper, 30 seconds of submersion in a second drop of UA, and a final blot. Grids were left to air dry before storage. Micrographs

of negative stain grids were screened using a Tecnai 12 microscope (Thermo Fisher Scientific) operating at 120 kV with a 4k x 4k charge-coupled device (CCD) camera.

3.7.2. Cryo-TEM screening

Cryo-TEM samples were concentrated to 1-2 mg/mL and 3.5 µL of sample added to glow-discharged (PELCO easiGlow™, 30 second discharge) Quantifoil Cu R1.2/R1.3 300 mesh holey-carbon grids and UltrAuFoil R1.2/1.3 300 mesh holey gold grids. Grids were flash-frozen in liquid ethane using a Vitrobot Mark IV at 4 °C and 100% humidity, with 0 seconds wait time, -10 blotting force and between 3-5 seconds of blotting time. Micrographs of cryogenically frozen grids were screened using a Tecnai 12 microscope operating at 120 kV with a 4k x 4k CCD camera, in low dose mode.

Chapter 4

Materials and Methods: Cryo-transmission electron microscopy analysis of prokaryotic voltage-gated sodium channels

4.1. Chapter summary

The following chapter describes the materials and methods used in the expression, purification, and electron microscopy analysis of the prokaryotic voltage-gated sodium channel NavMs, along with descriptions of some of the techniques used.

4.2. Bioinformatics

Please refer to section 3.2 for details of bioinformatics resources used in this thesis.

4.3 Materials and equipment

Please see section 3.3 for a table of all materials and equipment used in this thesis alongside suppliers.

4.4. Constructs and cloning

The NavMs gene was isolated from the marine bacterium *Magnetococcus marinus* (formerly known as *Magnetococcus spirulina*) strain MC-1, cloned as a C-terminally truncated pore-only construct (McCusker *et al.*, 2011) and donated to the Wallace lab by the group of Dr D.E. Clapham. The C-terminal domain (CTD) was modified to a codon-optimised synthetic gene from residue H237 and subsequently cloned as full-length wild type (WT) NavMs into a pET15b plasmid vector with an N-terminal 6x His-tag and thrombin cleavage site by Dr C. Bagnéris (McCusker *et al.*, 2012; Sula *et al.*, 2017).

4.5. NavMs plasmid propagation and sequencing

4.5.1. Transformation

The NavMs pET15b construct was amplified to increase availability of plasmid material. 2 µL of 1000 ng/µL plasmid DNA was added to 50 µL of chemically competent DH5α *E.*

coli, which was then incubated on ice for 30 minutes. It was then heat-shocked at 42 °C for 45 seconds, then returned to ice for a further 2 minutes. 250 µL of super optimal broth with catabolite repression (SOC) medium was added before 30 minutes of incubation at 37 °C. 100 µL of sample was spread on a lysogeny broth (LB) agar plate with 0.1 mg/mL ampicillin and incubated at 37 °C overnight. Single colonies were picked and used to inoculate 10 mL of LB medium with 0.1 mg/mL ampicillin. Finally, inoculated broth samples were incubated at 37 °C overnight with 200 rpm shaking.

4.5.2. Plasmid extraction and purification

Plasmids were extracted and purified via Wizard Plus SV Miniprep kit. Cells were lysed with cell lysis solution and alkaline protease solution provided with kit, and lysate loaded onto spin column for deoxyribonucleic acid (DNA) binding. The column was washed, and DNA eluted in nuclease-free water. Plasmid concentration was measured using a Nanodrop spectrophotometer by absorbance at 260 nm.

4.5.3. Sequencing

To check that no mutations or errors were introduced during the plasmid propagation process, 20 µL of 100 ng/µL sample were sent for Sanger DNA sequencing at Source Bioscience. The returned sequence was translated into amino acids by ExPASy translate and compared to the wild type NavMs sequence by EMBOSS Needle (Needleman and Wunsch, 1970).

4.6. Expression and purification of NavMs

4.6.1. Expression of NavMs

Full-length wild type NavMs plasmid was transformed into C41(DE3) *E. coli* cells as described in section 4.5.1. Single colonies were picked and used to inoculate 800 mL LB media with 0.1 mg/mL ampicillin in 2 L conical flasks (8 flasks, 6.4 L total), which were then incubated at 37 °C with shaking at 220 rpm. The cell density was checked after 2 h by measuring the OD_{600} of 1 mL fractions using a UV-Vis spectrophotometer. Once an OD of 0.8 was reached, expression was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG), with cultures then incubated for a further 3 h under the same conditions. Cultures were harvested by centrifugation at 6000 x *g* for 20 minutes and subsequent pellets stored at -20 °C until ready for purification.

4.6.2. Purification of NavMs

The protocol for purification of NavMs for structural study was developed by previous members of the Wallace lab (Sula *et al.*, 2017). The entire purification was performed at 4 °C. Cell pellets were resuspended in lysis buffer containing 150mM NaCl, 20mM tris base, ethylenediaminetetraacetic acid-free (EDTA-free) protease inhibitor tablets, 0.2 mM phenylmethanesulfonyl fluoride (PMSF), 2 mg/mL DNase, and 2.5 mM $MgSO_4$ at pH 7.5, and lysed by 3 passes through an Avestin Emulsiflex C3 at 15000 psi. The cell debris fraction was removed by centrifugation of the lysate at 12000 x *g* for 30 min, and cell membranes isolated by further centrifugation at 235000 x *g* for 2 h. Membrane pellets were resuspended in 100 mL solubilisation buffer (1.5% N-dodecyl-β-D-maltopyranoside (DDM), (172x CMC), 150 mM NaCl and 20 mM tris base pH 7.5) using a Dounce homogeniser before 2 h incubation at 4 °C. This was diluted with 50 mL dilution buffer (20 mM tris base pH 7.5, 600 mM NaCl, 90 mM imidazole) to make 1% DDM (115x CMC), 20 mM tris base pH 7.5, 300 mM NaCl and 30 mM imidazole (to reduce non-specific binding) then syringe filtered (0.45 μm) and loaded onto a 1 mL Histrap™ HP column by peristaltic pump. The column was attached to an AKTA purifier fast protein liquid chromatography (FPLC) system and washed with buffer containing 0.52% decanoyl-N-hydroxyethylglucamide (HEGA-10) (2x CMC of 0.26%), 300 mM NaCl, and increasing concentrations of imidazole (30 mM, 150 mM, 300 mM, 10 column volumes each) before elution with 1 M imidazole. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and western blotting (see sections 4.8.1 and 4.8.2) were used to identify fractions containing target protein, which were pooled and concentrated with 100 kDa cut-off Amicon centrifugal concentrators to approximately 250 μL and further purified by size-exclusion chromatography (SEC) using a Superdex 200 Increase 10/300 GL column in SEC buffer (150 mM NaCl, 20 mM tris base pH 7.5, 0.52% HEGA-10).

4.6.3. Determination of protein concentration

The concentration of purified NavMs, membrane scaffold protein 2N2 (MSP2N2), and NavMs in MSP lipid nanodiscs was determined as in section 3.5.3.

4.7. Membrane scaffold protein nanodisc reconstitution and purification

The protocol developed here for reconstitution of NavMs in MSP nanodiscs was informed by the publication of cryo-TEM structures of another prokaryotic Nav in nanodiscs, NaChBac (Gao *et al.*, 2020). Two datasets were collected of NavMs in MSP2N2 nanodiscs (Grinkova, Denisov and Sligar, 2010), hereby named dataset 1 and dataset 2. Samples for dataset 2 underwent an extra step of purification to remove empty nanodiscs, detailed in sections 4.7.1 and 4.7.2.

4.7.1. Nanodisc reconstitution

Lipids for nanodisc reconstitution were prepared by transferring 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) in chloroform into a round bottomed flask and drying out in a fume hood by rotation under argon gas flow. The lipid film in the flask was resuspended in lipid resuspension buffer (150 mM NaCl, 20 mM tris base, pH 7.5), with 30 seconds of 30% amplitude sonication to ensure homogeneous suspension. The suspension was then aliquoted into Eppendorf tubes before topping off with argon gas and freezing at -20 °C for later use. On the day of nanodisc reconstitution, aliquots were defrosted and 10% 3-((3-cholamidopropyl)dimethylammonio)-1-propanesulfonate (CHAPS) detergent was added in 4 equal stages, with a total of 2 minutes of sonication (10 seconds on, 10 seconds off, 10% amplitude with microtip probe) after each addition of CHAPS to solubilise the lipid, which was judged by the mixture turning clear. This made a 10 mM solution of solubilised POPC.

Bio-Beads SM-2 resin was prepared by washing with inversion in 4x the Bio-Bead volume of 100% methanol, followed by the same in 100% ethanol, then 3 times in distilled water, and finally two washes with Bio-Bead buffer (150 mM NaCl, 20 mM tris base, pH 7.5) before storage at 4 °C until use.

Membrane scaffold protein 2N2 (MSP2N2) was defrosted on ice and centrifuged at 23000 *x g* for 15 minutes at 4 °C to remove aggregated MSP2N2. The concentration of the supernatant and of purified NavMs in 0.52% HEGA-10 was measured to determine what volume of each was needed to produce a molar ratio of NavMs tetramer:MSP2N2 monomer:POPC of 1:12:520 (final concentrations approximately 1.8 μM NavMs tetramer, 22.0 μM MSP2N2 monomer, and 951.5 μM POPC for dataset 1, and 3.7 μM NavMs tetramer, 43.91 μM MSP2N2 monomer and 1.9 mM POPC for dataset 2). The

molar ratio of 1:12:520 was chosen for initial trials due to its successful implementation in producing the cryo-TEM structure of another prokaryotic Nav, NaChBac (Gao *et al.*, 2020). The components were added in this sequence: MSP2N2, then POPC, then 5.2% stock HEGA-10 solution (to give a final concentration of 0.52%), then NavMs last, before 1 h of incubation with gentle rotation at 4 °C. For dataset 2 only, tobacco etch virus (TEV) protease (produced and provided by Dr David Hollingworth of the Wallace group) was added at the start of incubation to cleave the poly His-tag from MSP2N2, with a final molar ratio of 1:10 TEV to MSP2N2.

After 1 h of incubation, prepared Bio-Beads were added to the mixture at a w/v ratio of 0.3g/mL to begin the adsorption of detergent and formation of nanodiscs. This was left to incubate overnight at 4 °C with gentle rotation.

4.7.2. Purification of NavMs-MSP2N2 lipid nanodiscs

The resultant mixture was centrifuged at 23000 *x g* for 15 minutes at 4 °C to remove aggregate. For dataset 2, the supernatant was incubated with 1 mL of Ni Sepharose 6 Fast Flow resin for 4 h in nanodisc Ni^{2+} binding buffer (30 mM imidazole, 300 mM NaCl, 20 mM tris base, pH 7.5) before loading onto a gravity chromatography column. Gel beads were washed with 10 column volumes of nanodisc Ni^{2+} binding buffer and 3 column volumes of nanodisc Ni^{2+} wash buffer (nanodisc Ni^{2+} binding buffer with 100 mM imidazole) before elution with 5 column volumes of nanodisc Ni^{2+} elution buffer (Ni^{2+} binding buffer with 500 mM imidazole). Fractions containing target protein were identified by SDS-PAGE (see section 4.8.1) and pooled. Samples for datasets 1 and 2 were both concentrated using 100 kDa cut-off Amicon centrifugal concentrators and further purified by SEC using a Superose 6 Increase 10/300 GL column in nanodisc SEC buffer (300 mM NaCl, 20 mM tris base, pH 7.5).

4.8. Biochemical analysis of NavMs

4.8.1. SDS-PAGE

For SDS-PAGE analysis of purified NavMs and NavMs in nanodiscs, samples were mixed with 4x NuPAGE lauryl dodecyl sulfate (LDS) sample buffer. These were not cooked for denaturation as the sample could not then run through the gel due to aggregation. Samples were then loaded into 10 or 15-well Invitrogen NuPAGE 4-12% Bis-Tris precast gels alongside Benchmark protein ladder, and gels were placed in a tank

containing 3-(*N*-morpholino)propanesulfonic acid (MOPS) running buffer. Gels were run for 45 minutes at 200 V. Gels were rinsed with distilled water and either stained with InstantBlue protein stain for 1 h before destaining with distilled water until bands were visible (all with rocking motion) or used for western blotting.

4.8.2. Western blotting

Western blotting of NavMs was performed as follows: following SDS-PAGE, the gels were blotted onto a polyvinylidene difluoride (PVDF) membrane via iBlot 2 dry blotting system. The membrane was blocked from non-specific binding by incubation (gentle shaking at 4 °C) in 1x phosphate-buffered saline (PBS) buffer with 3% w/v skimmed milk powder for 30 minutes. The membrane was then washed 3 times for 5 minutes each in 1x PBS before incubation overnight at 4 °C with anti-polyHis horseradish peroxidase-conjugated antibody (1:2000 dilution in 1x PBS 1% skimmed milk powder). The solution was poured off and the membrane was washed 5 times for 5 minutes each with a solution of 1x PBS and 0.05% v/v Tween-20 detergent, then washed with SIGMAFAST 3,3'-diamonibenzidine (DAB) tablets in 5 mL distilled water. After 5 minutes water was added to quench the reaction before pouring off and the blot analysed by eye.

4.9. Electron microscopy screening of NavMs

4.9.1. Negative stain screening

3 µL of sample at approximately 0.1 mg/mL was added to glow discharged (Harrick Plasma Cleaner with PlasmaFlo gas mixer, low power, 45 seconds) 300 mesh copper transmission electron microscopy (TEM) support grids and left to adhere for 30 seconds. The grid was submerged in two drops of Millipore filtered water to rinse off excess detergent, before quick submersion in one drop of 2% uranyl acetate (UA), blotting with ashless filter paper, 30 seconds of submersion in a second drop of UA, and a final blot. Grids were left to air dry before storage. Micrographs of negative stain grids were screened using a Tecnai 12 microscope (Thermo Fisher Scientific) operating at 120 kV with a 4k x 4k charge-coupled device (CCD) camera.

4.9.2. Cryo-TEM screening

Cryo-transmission electron microscopy (cryo-TEM) samples were concentrated to 1-2 mg/mL and 3.5 µL of sample added to glow-discharged (Harrick Plasma Cleaner with PlasmaFlo gas mixer, high power, 1 minute) Quantifoil Cu R1.2/R1.3 300 mesh holey-

carbon grids and UltrAuFoil R1.2/1.3 300 mesh holey-gold grids. Grids were flash-frozen in liquid ethane using a Vitrobot Mark IV at 4 °C and 100% humidity, with 0 seconds wait time, -10 blotting force and between 3-5 seconds of blotting time. Micrographs of cryogenically frozen grids were screened using a Tecnai 12 microscope operating at 120 kV with a 4k x 4k CCD camera, in low dose mode.

4.10. Cryo-TEM data collection and processing of NavMs in nanodiscs

4.10.1. Data collection

Both datasets 1 and 2 were collected by Dr Natalya Lukoyanova on a Titan Krios microscope equipped with Gatan post-GIF K3 direct electron detector in super resolution mode, binned by 2. Specific details of each dataset are shown in Table 4.1 below.

Table 4.1. Cryo-TEM data collection parameters.

Parameter	Dataset 1	Dataset 2
Microscope	Titan Krios	Titan Krios
Voltage (kV)	300	300
Detector	K3	K3
Magnification	105 kx	130 kx
Pixel size (Å / pix)	0.83	0.65
Exposure time (s)	1.6	1.55
Dose per $Å^2$ per second (e- / $Å^2$ / s)	31.5	32.66
Total dose (e^- / $Å^2$)	50.4	50.6
Defocus range (μm)	-1.2 to -2.7	-1.2 to -2.4
Nr of frames	20,620	20,207
Tilt (degrees)	0	0

4.10.2. Data processing

All cryo-TEM data processing was carried out in cryoSPARC v3.3.1 through the cryoSPARC interface (Punjani *et al.*, 2017). Specific details of data processing steps are given in Chapter 6. Overall, images were imported into cryoSPARC and subjected to pre-processing steps of patch motion correction (multi) and patch contrast transfer function (CTF) estimation. Exposures were curated by setting cut-offs for CTF fit resolution and total full-frame motion distance to filter exposures with issues during data collection. Blob picking was performed on accepted exposures, and particle picks curated by normalised cross-correlation (NCC) score and power threshold to remove obviously incorrect picks such as ice contamination, empty space, and hole edges. Selected particle

picks were subjected to 2D classification and representative top and side views used as references for template-based particle picking. Particles curated from this selection were then again subjected to multiple rounds of 2D classification to remove junk particles such as ice or aggregate before *ab-initio* reconstruction. The resultant Coulomb potential map was improved by multiple rounds of heterogeneous refinement (to further clean out bad particles), before multiple rounds of non-uniform refinement to improve the resolution of the final map (Punjani, Zhang and Fleet, 2020). 3D classification and local refinement were also used in an attempt to improve resolution and detect subpopulations that could exist due to flexibility of the protein, with masks made in ChimeraX (Pettersen *et al.*, 2021). Global resolution was calculated by cryoSPARC with each stage of refinement, using the gold-standard Fourier shell correlation (GSFSC) with a correlation coefficient of 0.143 (Scheres and Chen, 2012), and local resolution calculated by the local resolution estimation job using the GSFSC with a correlation coefficient of 0.5. Coulomb potential maps were visualised and analysed in ChimeraX (Pettersen *et al.*, 2021).

4.10.3. Model building and analysis

Final Couloumb potential maps were locally sharpened using the autosharpen tool in Phenix (Liebschner *et al.,* 2019). A predicted model for NavMs tetramer was produced by Alphafold Multimer using Colabfold (Evans *et al.*, 2022; Mirdita *et al.*, 2022) and used as a starting model. The model was docked into the sharpened map and refined using the dock in map and real-space refinement tools in Phenix with default settings (secondary structure, Ramachandran, and rotamer restraints on). Ramachandran, geometry, and rotamer outliers were adjusted manually in Coot version 0.9.8 (Emsley *et al.*, 2010) before subsequent runs of real-space refinement. Remaining errors in side chain placement were manually fixed in ISOLDE (Croll, 2018). Potential ligand sites were identified by searching for unmodelled blobs in Coot, with two POPC ligands (code POV) placed per monomer and a single sodium ion placed in the selectivity filter of the tetramer. POPC ligands were manually refined in Coot with Phenix eLBOW used to generate ligand restraints, and the final structure validated for map-model FSC and model quality using the cryo-TEM validation tool in Phenix and MolProbity (Williams *et al.*, 2018). Figures of models were produced in UCSF ChimeraX (Pettersen *et al.*, 2021) and PyMol molecular graphics software. The pore radius of the channel was analysed and visualised using HOLE2 and Visual Molecular Dynamics (Humphrey, Dalke and Schulten, 1996; Smart *et al.*, 1996) by Fran Thomas of the Wallace group.

Chapter 5

Results: Purification and Electron Microscopy Analysis of hNav1.8 Recombinantly Expressed in *Tetrahymena thermophila*

5.1. Chapter summary

Accurate structural models of each human voltage-gated sodium channel (hNav) isoform are important for understanding differences in their mechanisms of action, such as why different isoforms are be activated at depolarisation voltages. Furthermore, considerable research effort is underway to design effective and subtype-specific drugs to treat channelopathy conditions associated with each hNav isoform, which would benefit from high-resolution structural data of each isoform in open, closed, and inactivated conformations. As hNavs are difficult to express and purify in large quantities and their hydrophobic and highly mobile nature makes crystallisation difficult, cryo-transmission electron microscopy (cryo-TEM) rather than X-ray crystallography is currently the method of choice for analysing these proteins.

At the time this investigation began, there was only one published hNav structure available (hNav1.4 in complex with β1 at 3.2 Å (Pan *et al.*, 2018)), with no published structure or established purification protocol available for hNav1.8. This isoform is involved in pain signalling (Bennett *et al.*, 2019) and the subject of drug trials for analgesic medication, such as the compound VX-548 developed by Vertex for treatment of acute pain, (Mullard, 2022). As referred to in section 2.2, expression of hNavs in *E. coli* is not viable, so a collaboration was formed with Tetragenetics Inc. (since acquired by AbCellera Biologics Inc.), who, with funding from the National Institute of Health (NIH) and under principal investigator Paul Colussi had success using the eukaryotic single-celled organism *Tetrahymena thermophila* to recombinantly express hNavs with the original purpose of developing therapeutic antibodies for treatment of channelopathies. Previous work by Dr Altin Sula of the Wallace group had established a general method for purification of hNav1.8 from recombinant *T. thermophila* cell pellets,

expressed and provided by Tetragenetics Inc (see section 5.2 for details of previous work). However, this produced a very low yield with suspected impurities, meaning that up to the beginning of this project it had only been possible to effectively visualise hNav1.8 by negative stain electron microscopy (EM) (unsuitable for high-resolution structural analysis) or on cryo-TEM grids with a graphene oxide support, which led to poor 2D classes and reconstructions (see Figure 5.2). The aim of this project therefore was to improve the purity and yield of samples of the human voltage-gated sodium channel hNav1.8, to the point where they could then be used to obtain high-resolution cryo-TEM data and explore its mechanism of action, elucidating structural details that differentiate it from other hNavs.

A range of purification conditions were trialled, some of which were informed by published purification and cryo-TEM analysis of other hNavs published during the project, detailed in section 5.3. As these did not significantly improve the result of the purification, multiple preparations were pooled together, at which point visualisation on cryo-TEM micrographs was successful, however this needed further optimisation due to the presence of impurities and low occupancy per micrograph. Identification of these impurities by mass spectrometry (see section 5.3.2) suggested that they were due to degradation of hNav1.8 by the host organism; attempts to reduce this degradation by modifications to the purification protocol were unsuccessful (see section 5.3.3). Soon after this, high-resolution cryo-TEM structures of hNav1.8 in apo form and in complex with the subtype-selective blocker A-803467 were published by the group of Nieng Yan (X. Huang *et al.*, 2022) (discussed in section 7.1.2), leading to conclusion of the project.

5.2. Preliminary research

As mentioned in the previous section, preliminary work for the project was undertaken by Dr Altin Sula, who alongside Tetragenetics Inc. established a general protocol for purification of hNav1.8 from *T. thermophila* cells, involving a construct of full-length wild type hNav1.8 with a 10x poly-His tag and FLAG tag undergoing two rounds of affinity chromatography prior to size exclusion chromatography (SEC) (see Methods section 3.5.2). A twin streptavidin tag was introduced at the N-terminus, but the construct did not solubilise well and so work continued with the FLAG and poly-His construct. This purification leads to a characteristic SEC trace with four peaks, hereby named P1, P2, P3, and P4 (see Figure 5.1). Samples from each of these peaks were taken directly

from SEC fractions and analysed by liquid chromatography-mass spectrometry (LC-MS), giving positive results for hNav1.8 in P1, P2, and P3. P1 and P2 samples had sequence coverage from domains I-IV, whereas P3 only produced peptide hits from the N-and C-termini, suggesting that intact hNav1.8 is present in P1 and P2 and only fragments in P3.

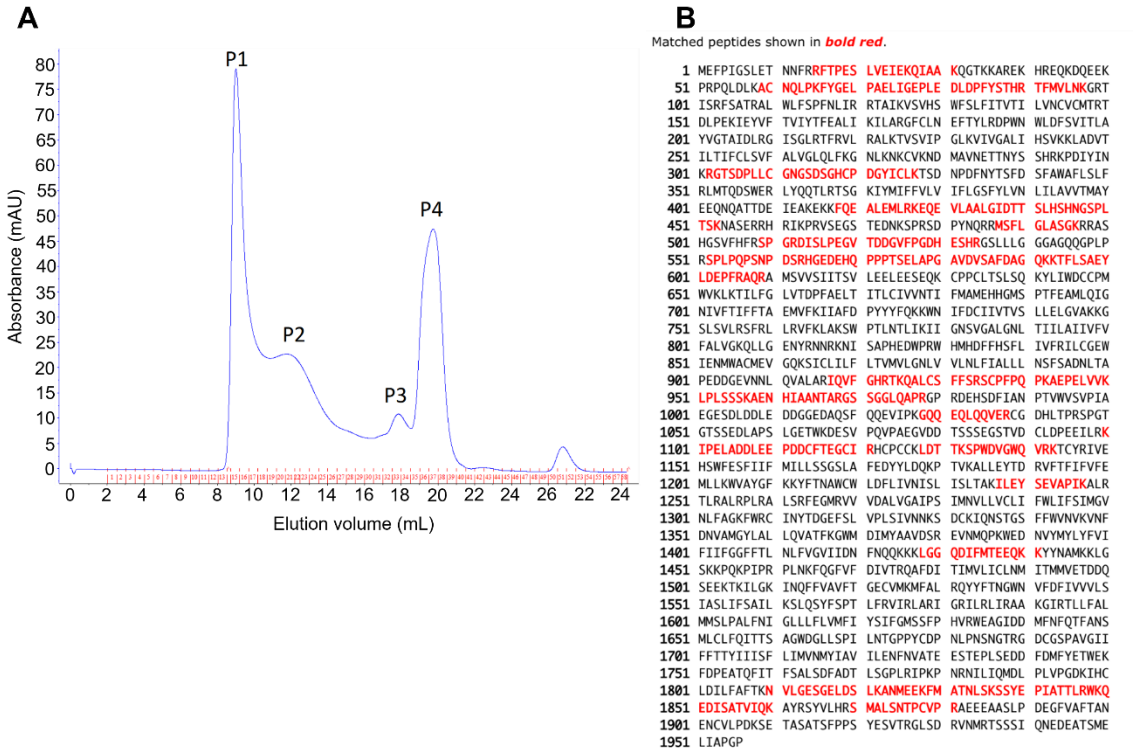

Figure 5.1. Preliminary hNav1.8 purification results. Figure showing preliminary results of hNav1.8 purification and mass spectrometry analysis. (**A**) shows a size exclusion chromatography (SEC) trace after following the standard protocol (see Methods section 3.5.2). Peaks are labelled P1-4 according to elution volume. (**B**) shows solution mass spectrometry peptide matches of peak P2 with wild type hNav1.8, with matches shown in red. Figure adapted from results provided by Dr Altin Sula.

Negative stain electron microscopy of P1 (see Figure 5.2), coupled with the fact that the peak elutes at approximately 8 mL (the void volume of the column) suggested that P1 was formed of aggregate material, whereas negative stain micrographs of P2 gave monodispersed particles of the expected size for hNav1.8 (approximately 10-12 nm across), identifying this as the target peak. Cryo-TEM micrographs of grids prepared from P2 did not show hNav1.8 particles; in an attempt to remedy this, grids were prepared with a support of graphene oxide, which adheres protein particles allowing for samples of lower concentration to be visualised (Palovcak *et al.*, 2018), however data processing of approximately 45,000 particles from these samples was only able to produce a 12 Å ab-initio Coulomb potential map. A homology model based on the structure of electric eel Nav1.4 (Yan *et al.*, 2017) was fitted into this map showing that they were similar sizes, although at 12 Å resolution no detail of secondary structure was identifiable. It was therefore decided that improvements to the purification process were needed to increase the yield and purity of hNav1.8 before further data collection.

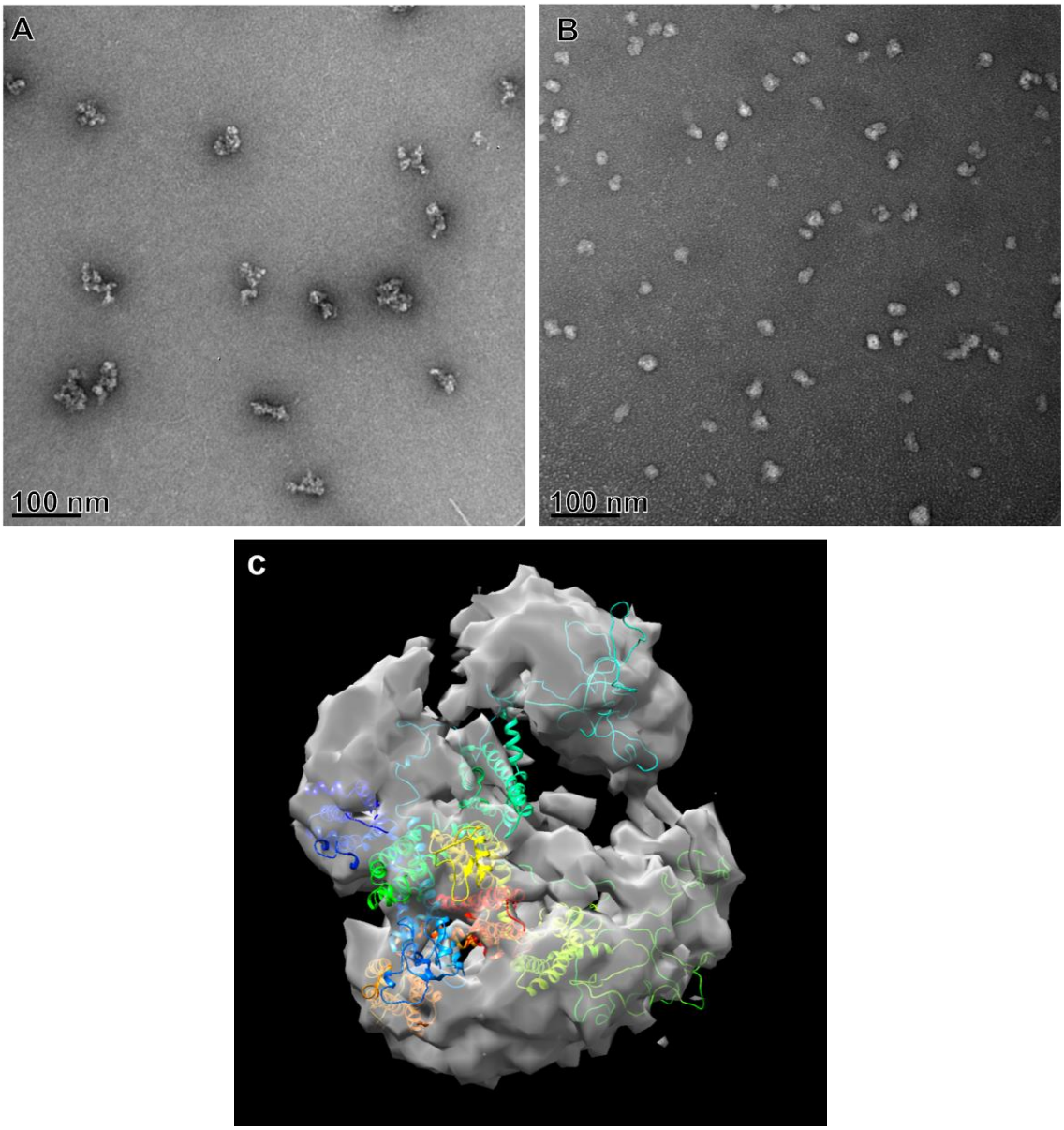

Figure 5.2. Preliminary electron microscopy screening of hNav1.8. Negative stain electron micrographs of (**A**) aggregated particles from a sample of peak 1 and (**B**) dispersed particles from a sample of peak 2 of the final step of purification of hNav1.8 following the standard protocol. **(C)** shows a homology model of hNav1.8 based on the structure of electric eel Nav1.4 (Yan *et al*., 2017) fitted into a 12 Å *ab initio* reconstruction from cryo-TEM screening of peak 2. Figure adapted from results provided by Dr Altin Sula.

5.3. Purification of hNav1.8 expressed in recombinant *T. thermophila*

5.3.1. Purification trial conditions

A range of purification condition modifications were performed with the aim of improving the yield and purity of hNav1.8 in comparison to what will be referred to as the 'standard protocol' (developed by Dr Altin Sula), which, in short, consists of solubilisation of membrane pellets in 1% N-dodecyl-β-D-maltopyranoside (DDM) and a 3-step purification in glyco-diosgenin (GDN) using FLAG and Ni^{2+} affinity chromatography followed by size exclusion chromatography (SEC) (see Methods section 3.5.2 for full details). The following purification methods were undertaken, some of which followed published protocols for purifying other eukaryotic Navs:

- Expression and purification in the presence and absence of lidocaine
- Use of cell pellets produced from 6 L or 12 L growths
- Purification in DDM throughout (i.e., no exchange into GDN)
- Addition of 0.2% cholesteryl hemisuccinate Tris salt (CHS) to extraction buffer (after Jiang *et al.*, 2020)
- Purification at pH 6 with 25 mM imidazole (after Pan *et al.*, 2019)
- Addition of 0.001% sodium dodecyl sulfate (SDS) (after Pan *et al.*, 2019)

The effects of these modifications versus the standard protocol were analysed by western blot and negative stain EM to judge whether they had improved the purification. In some cases, the modifications led to a reduction in yield of the target protein, such as when lidocaine was not included in expression or purification, seen as a smaller target peak at approximately 12 mL elution volume and larger aggregate peak at the 8 mL void volume in the resulting SEC trace (see Figure 5.3). Western blots of the FLAG affinity and Ni^{2+} affinity purification steps of this preparation show only faint signal in the region expected for the hNav1.8 construct (220 kDa) and stronger bands at lower molecular weights, suggesting that the protein had degraded leaving a 55-70 kDa fragment with the antigen present (see Figure 5.4). The standard protocol in comparison also shows bands at low molecular weights but displays more intense signal around the expected molecular weight also, suggesting better stability.

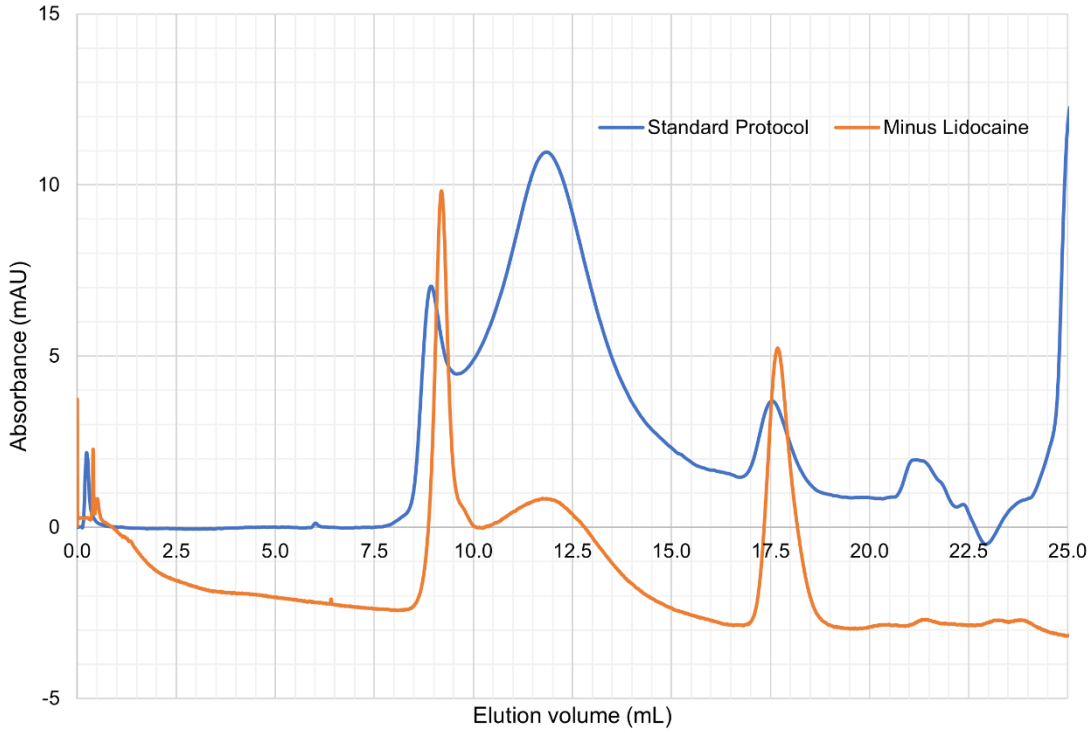

Figure 5.3. Expression and purification of hNav1.8 with and without lidocaine. Size-exclusion chromatography (SEC) traces of purification of hNav1.8 by the standard protocol (blue) versus expression and purification without lidocaine (orange).

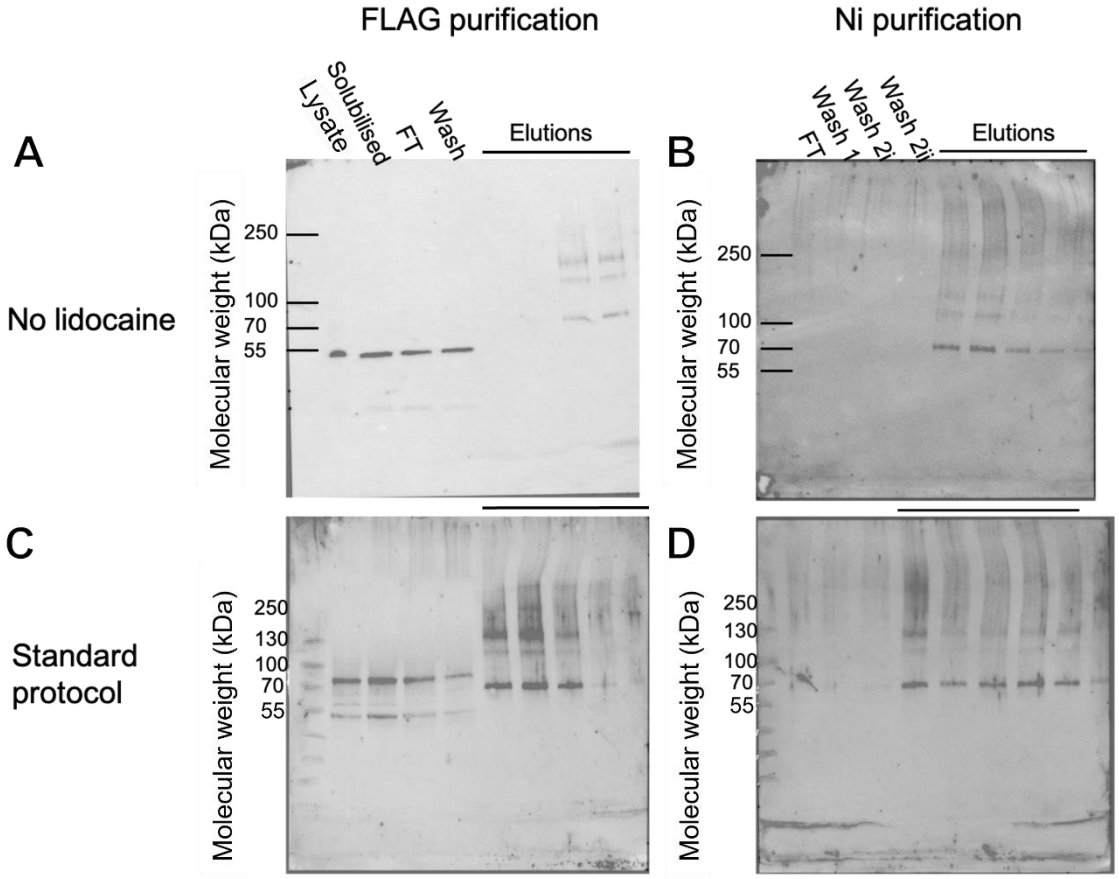

Figure 5.4. **Western blots of hNav1.8 purification steps.** Analysis of purification of hNav1.8 by standard protocol versus purification in the absence of lidocaine. (**A**) and (**B**) show western blots of FLAG purification and Ni affinity purification steps (respectively) in the absence of 1mM lidocaine, versus (**C**) and (**D**), showing the same purification steps in the presence of 1 mM lidocaine, i.e., the standard protocol (FT = flow through).

Adding 0.001% SDS produced no discernible difference to the standard protocol. Using pellets from 12 L of growth medium rather than the standard 6 L produced a slightly greater magnitude target peak, though not enough to warrant double the consumption of precious *T. thermophila* material (see Figure 5.5). Analysis by negative stain electron microscopy was used to determine whether the sample was suitable for cryo-TEM; this was judged by particle aggregation, the level of dispersion of particles, or heterogeneity in particle size or shape. The negative stain micrographs in Figure 5.6 show that in samples without lidocaine, samples purified in DDM instead of GDN, or purified at pH 6 with 25 mM imidazole show increased aggregation compared to the standard protocol, suggesting that these modifications would not be helpful in producing cryo-TEM grids for high-resolution structural data.

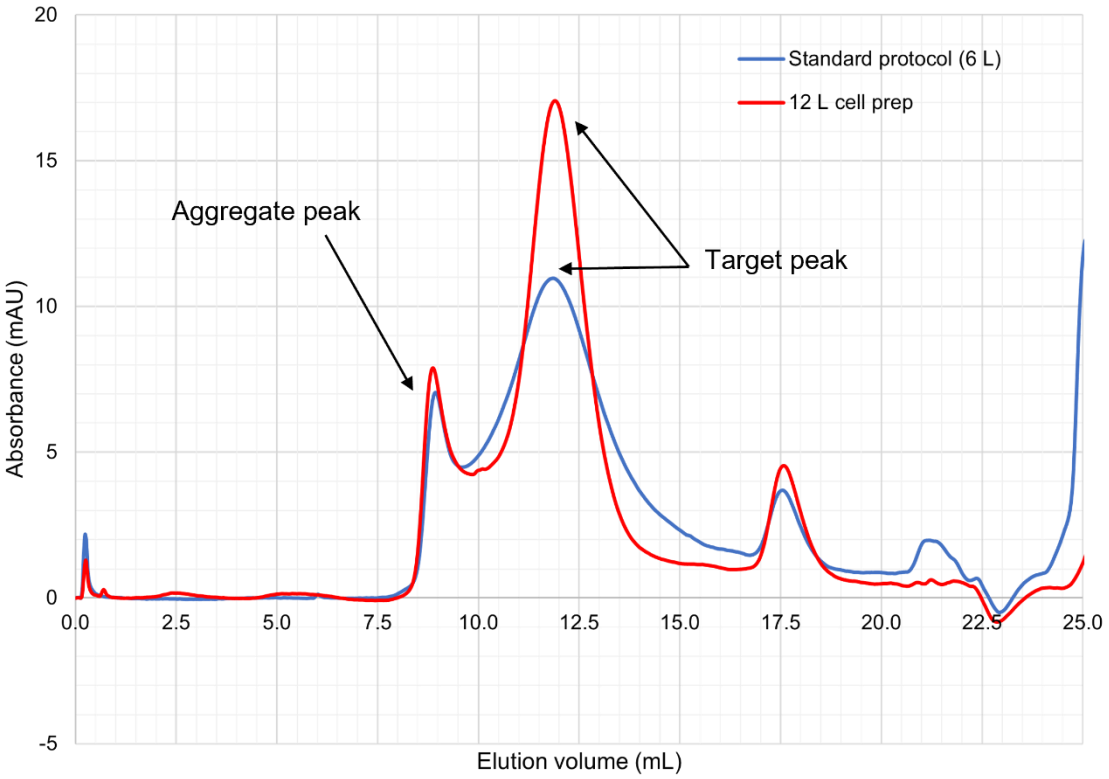

Figure 5.5. Comparison of standard protocol and 12 L growth purifications. Size-exclusion chromatography (SEC) traces of standard protocol purification (using 6 L of cell growth medium, blue) vs 12 L of cell growth medium (red).

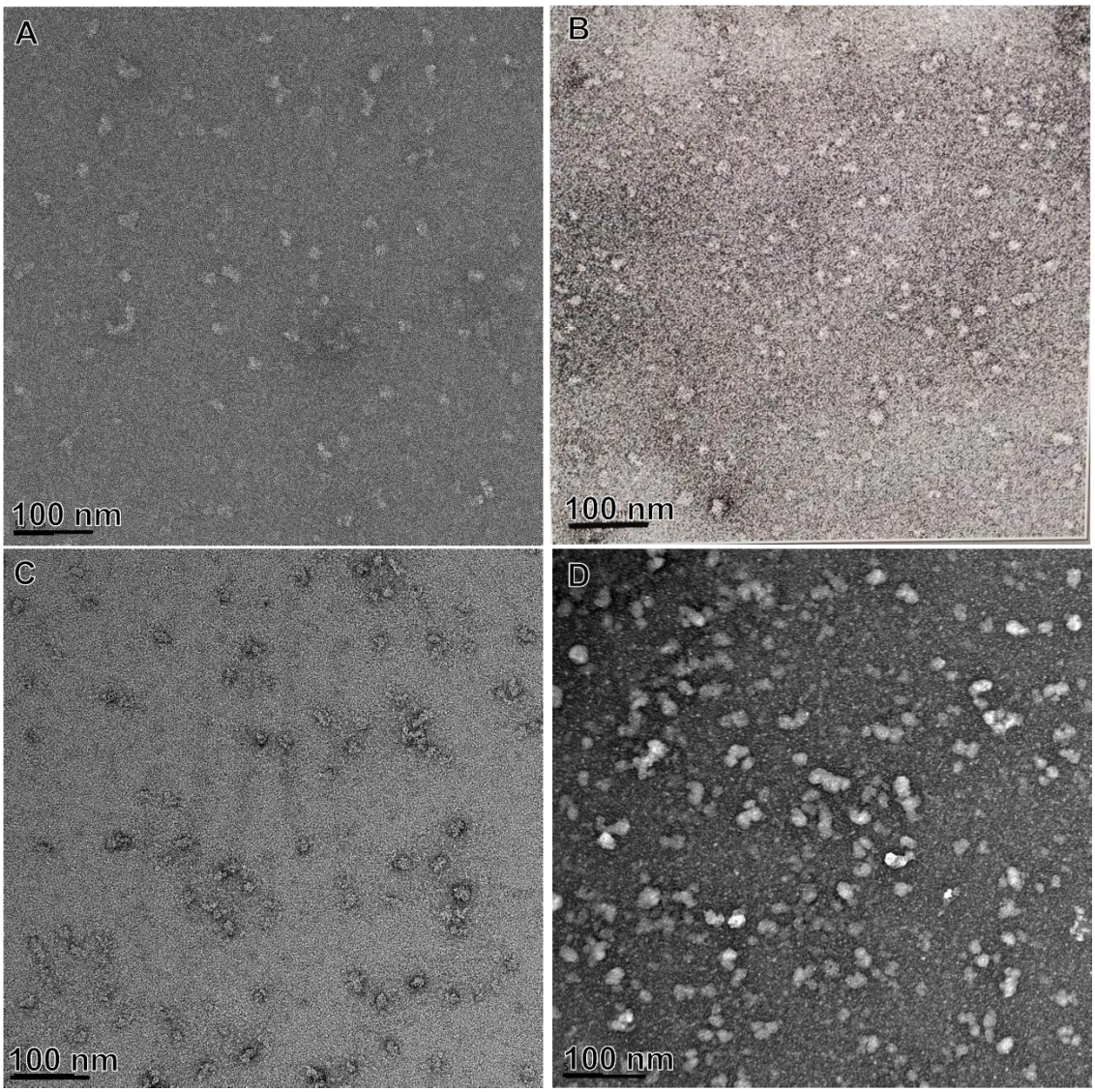

Figure 5.6. Negative stain screening of hNav1.8 purification conditions. Negative stain micrographs of hNav1.8 after various purification conditions. (**A)** Expressed and purified in the absence of lidocaine; (**B)** Standard protocol (final buffer 1x PBS, 300 mM NaCl, 0.075% GDN, 1 mM lidocaine, pH 7.4); (**C)** Purified in 1% DDM throughout instead of GDN; (**D**) Purified by standard protocol plus 25 mM imidazole at pH 6. Particles of hNav1.8 are expected to be 10-12 nm across based on previous hNav protein structures (Pan *et al,* 2019).

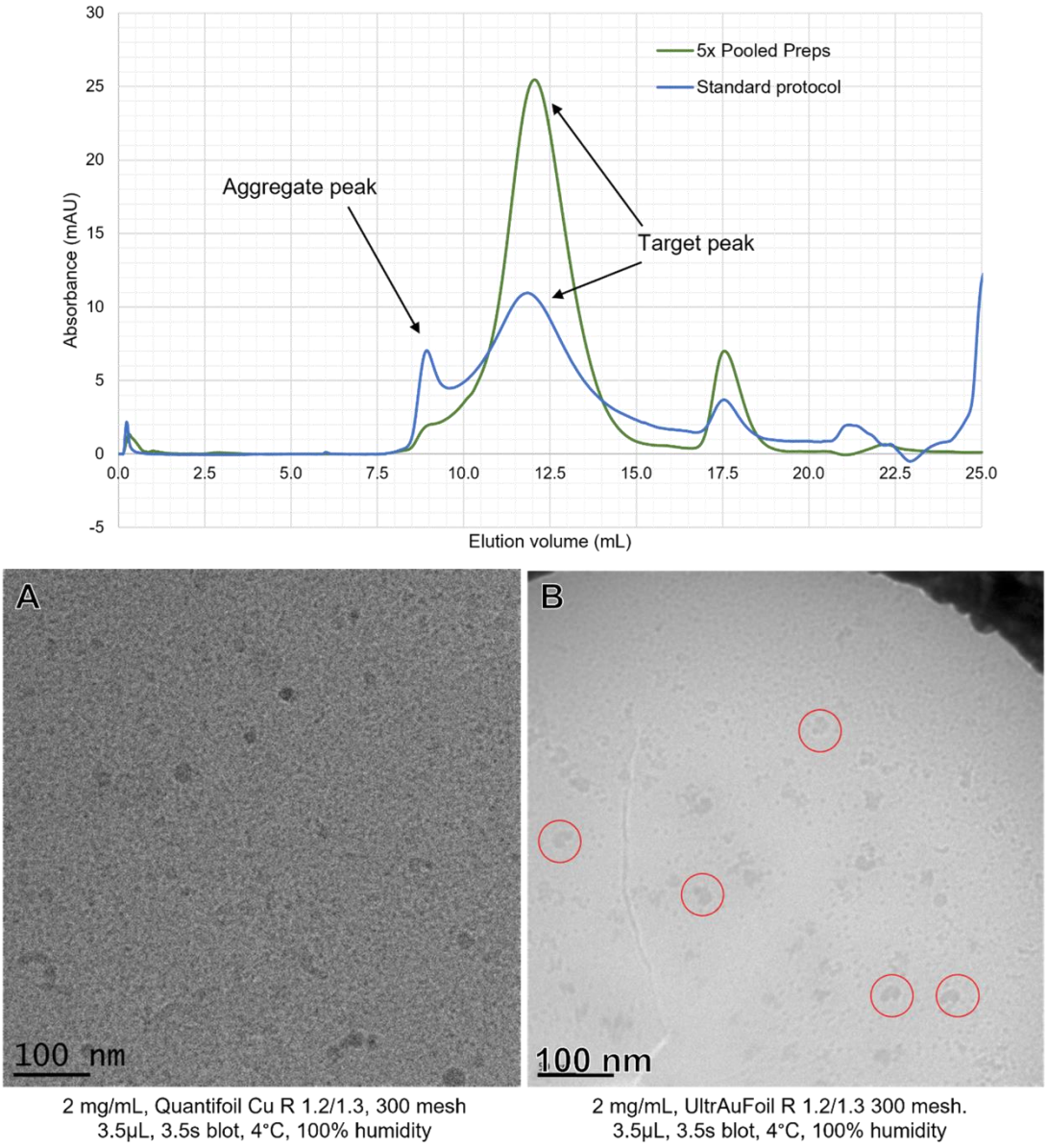

Figure 5.7. The effect of pooling purifications on hNav1.8 SEC traces and cryo-TEM screening. Top: Comparisons of a standard protocol hNav1.8 purification size-exclusion chromatography (SEC) trace versus 5 purifications pooled together and subjected to an additional size-exclusion chromatography (SEC) step. **Bottom:** Comparisons of cryo-electron micrographs of (**A**) hNav1.8 from a single standard purification protocol versus (**B**) 5 purifications pooled together, with vitrification conditions below each micrograph. Suspected hNav1.8 particles are circled in red.

At this point, a range of conditions had been attempted without significant improvement in yield of quality of negative stain micrographs over the standard protocol; as, however, the standard protocol could be reliably reproduced, it was decided to perform multiple purifications, snap-freezing the resultant protein, and eventually pooling the target peaks together. This pooled sample was then further purified by an added SEC step to remove any aggregate or degraded protein produced by the freeze-thaw process. In total, 5

samples were pooled from successful purifications, leading to an increase in peak magnitude of approximately 2.5x over a standard purification (see Figure 5.7, top), whilst the peak judged to represent aggregates was reduced in comparison to a single purification. Excitingly, when 3.5 µL of this pooled sample at 2 mg/mL was applied to an UltraAuFoil, 300 mesh, R1.2/1.3 cryo-TEM grid, well-dispersed particles expected to be hNav1.8 (judged by comparison with micrographs from previously published hNav structures (Pan *et al.*, 2018, 2019)) were visualised for the first time without a graphene oxide support (see Figure 5.7, bottom). Considering that samples from the standard protocol were also able to be concentrated to 2 mg/mL (according to their absorbance at 280 nm) but did not show particles recognisable as hNav1.8, it may be that in a heterogeneous sample the pooled purifications contained a larger proportion of non-degraded target protein versus fragments of hNav1.8 or small impurities. Some background impurities are still visible in the micrographs of the pooled sample and the occupancy of hNav1.8 per micrograph is low, so further optimisation was deemed necessary at this point. Overall, however, visualising the protein without a support was an important milestone.

To quantitatively compare the effect of different purification protocols, the area under the curve of each target peak was calculated (see Table 5.1). This showed that purification in the presence of 0.2% CHS and 25 mM imidazole slightly reduced the yield relative to the standard protocol (89%), and expression and purification without 1 mM lidocaine present drastically reduced it (45%). The yield was increased to 128% of the standard protocol by doubling the cell culture volume purified from, indicating that using larger cell culture volumes was not an efficient use of resources. Finally, pooling 5 standard protocol

Table 5.1. Quantitative comparison of purification method yields. Yields in mAU*mL are calculated by measuring the area under the target peak of size-exclusion chromatography (SEC) traces. CHS = 0.2% cholesteryl hemisuccinate, IMZ = 25 mM imidazole.

Method	Yield (mAU*mL)	Relative to Standard (%)
Standard	23	100
CHS+IMZ	20	89
12 L culture	29	128
No lidocaine	10	45
5 pooled purifications	53	231

purifications together and performing another gel filtration gave a target peak 231% the size of an individual standard protocol, which combined with a visibly smaller aggregate peak suggested that it may have improved the sample quality despite losing almost half of the sample.

5.3.2. Identification of impurities

To remove the impurity particles seen on the cryo-TEM micrographs, it was decided to first identify them (along with the presence of target protein) by performing liquid chromatography-mass spectrometry (LC-MS) on SDS-PAGE bands (see Figure 5.8 for analysed gel bands). A standard purification of hNav1.8 did not produce enough material to both produce cryo-TEM grids and visualise by SDS-PAGE, so this also provided an opportunity to examine how heterogeneous the sample was after three steps of purification. The gel itself exhibits many bands at low molecular weights, although these may be degradation products of a range of molecular weights rather than multiple impurity species. It could also be caused by heterogeneous glycosylation leading to different apparent molecular weights from the same protein, which can also explain the 'smearing' seen at the upper range of molecular weights; hNavs are known to be modulated by glycosylation, suggested to be important for cellular trafficking and protein-protein interactions, and gating modulation (Swanwick, Pristerá and Okuse, 2010; Laedermann, Abriel and Decosterd, 2015).

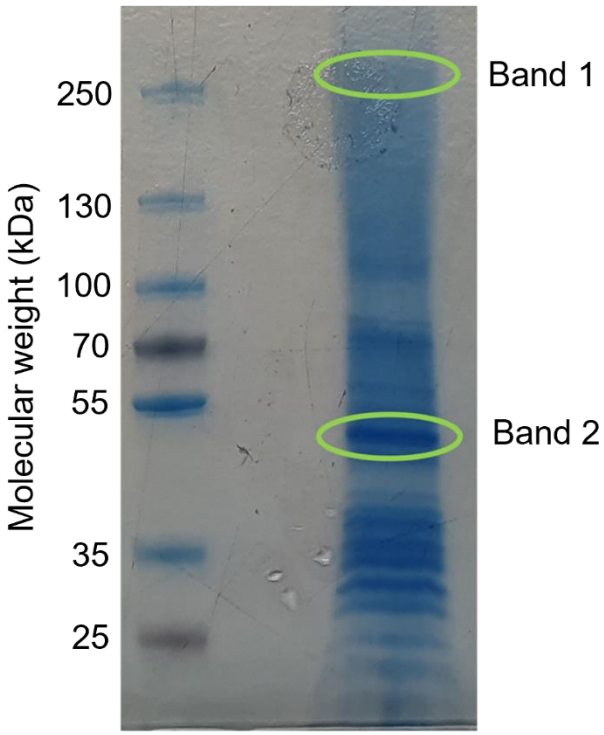

Figure 5.8. SDS-PAGE gel of hNav1.8 for LC-MS. SDS-PAGE gel of hNav1.8 sample after purification by standard protocol and concentration to 2 mg/mL. Bands sent off for mass spectrometry are circled in green.

Two bands were analysed by LC-MS, highlighted in Figure 5.8. Band 1 was chosen from the top of the gel (approximately 250 kDa) to ensure that full-length hNav1.8 was present in the sample, whilst Band 2 (approximately 50 kDa) was chosen as it appears to be the highest intensity band in the gel and does not match with a western blot band from the later stages of purification (as seen in Figure 5.4), and could therefore be a major source of impurity.

The peptide spectrum results of the LC-MS were run against a database of protein sequences from all species using the MASCOT search engine (Perkins *et al.*, 1999), giving relative probabilistic scores that the spectra match each protein. This indicated that hNav1.8 (score 4865), human keratin (score 1318, commonly present in gel band mass spectrometry from user handling of the gel (Hodge *et al.*, 2013)), and a 15 kDa *Tetrahymena* ubiquitin (score 906) were present in Band 1 (see Figures 5.9 and 5.11). The molecular weight of the sample sent and the fact that the peptide matches are spread consistently from the N- to C- terminus of the sequence (44% coverage total) indicate that full-length hNav1.8 is present; speculatively, as hNav1.8 itself is approximately 220 kDa,

it is possible that this result suggests the band consists of the target protein having been tagged with two 15 kDa ubiquitins by the expression organism for subsequent degradation (Ciechanover, 2005).

Band 2 at approximately 50 kDa gave the highest probability score as hNav1.8 (score 1113), but with only 24% coverage, although the peptide hits were present from N- to C-termini (see Figures 5.10 and 5.11). The next highest probability match was β-tubulin from the expression organism (score 242), a 50 kDa cytoskeletal protein. This result could indicate that the band consists of approximately 50 kDa proteolytic fragments of hNav1.8 which together produce hits over the length of the protein; looking back to the western blots in Figure 5.4, the band at 50 kDa is no longer present after Ni affinity purification, which could support it being β-tubulin, although this does not rule out degradation of hNav1.8, as fragments of hNav1.8 could simply not contain the epitope targeted by the primary antibody during western blotting. Consequently, it was decided to modify the purification to reduce degradation.

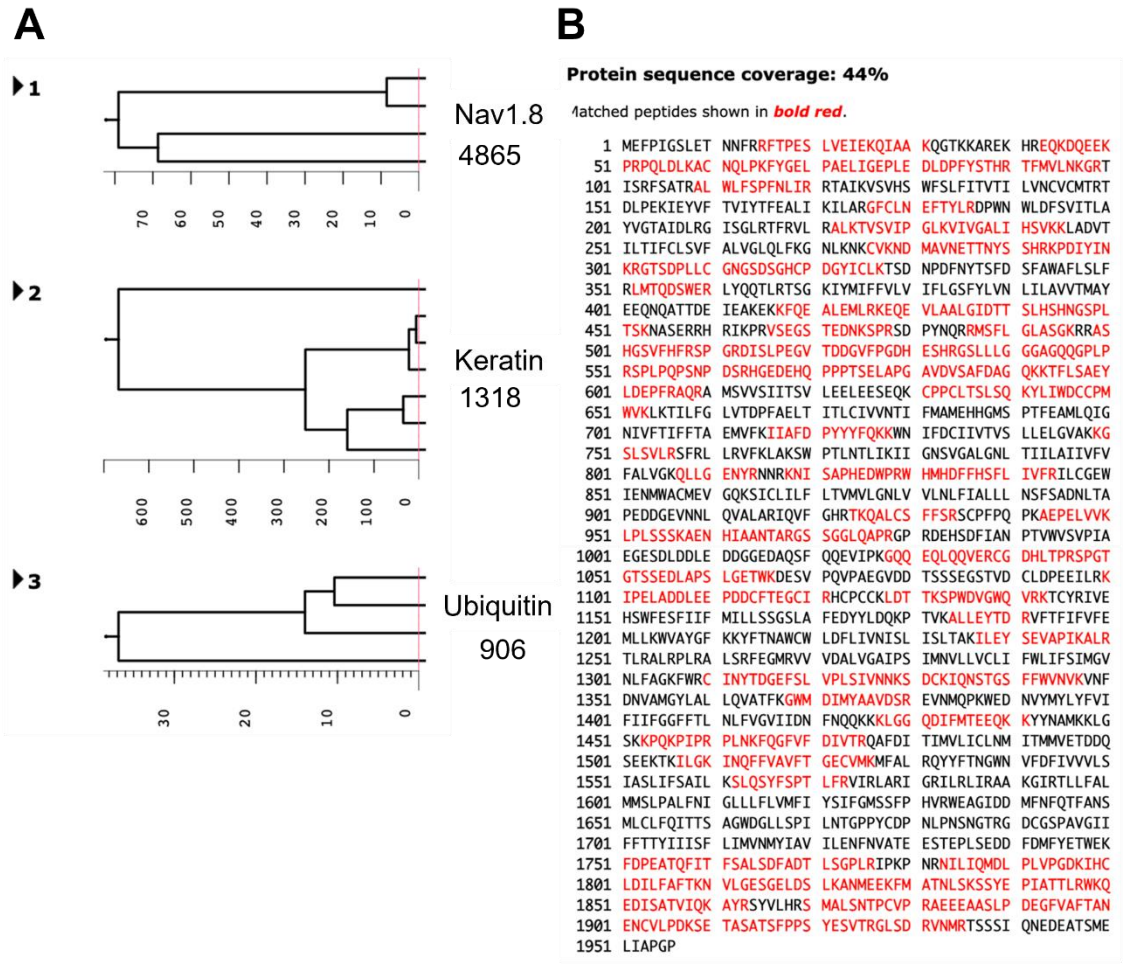

Figure 5.9. LC-MS results of band 1. Liquid chromatography-mass spectrometry (LC-MS) results of Band 1. (**A**) shows trees produced by the MASCOT server of top protein matches with scores of relative probabilities for each protein. The top result is Nav1.8 of various species, with hNav1.8 getting the highest score at tree 1; tree 2 is human keratin; tree 3 is ubiquitin, including from *Tetrahymena.* (**B**) shows sequence coverage of hNav1.8 by matching detected peptides in red, with 44% total coverage over the length of the protein.

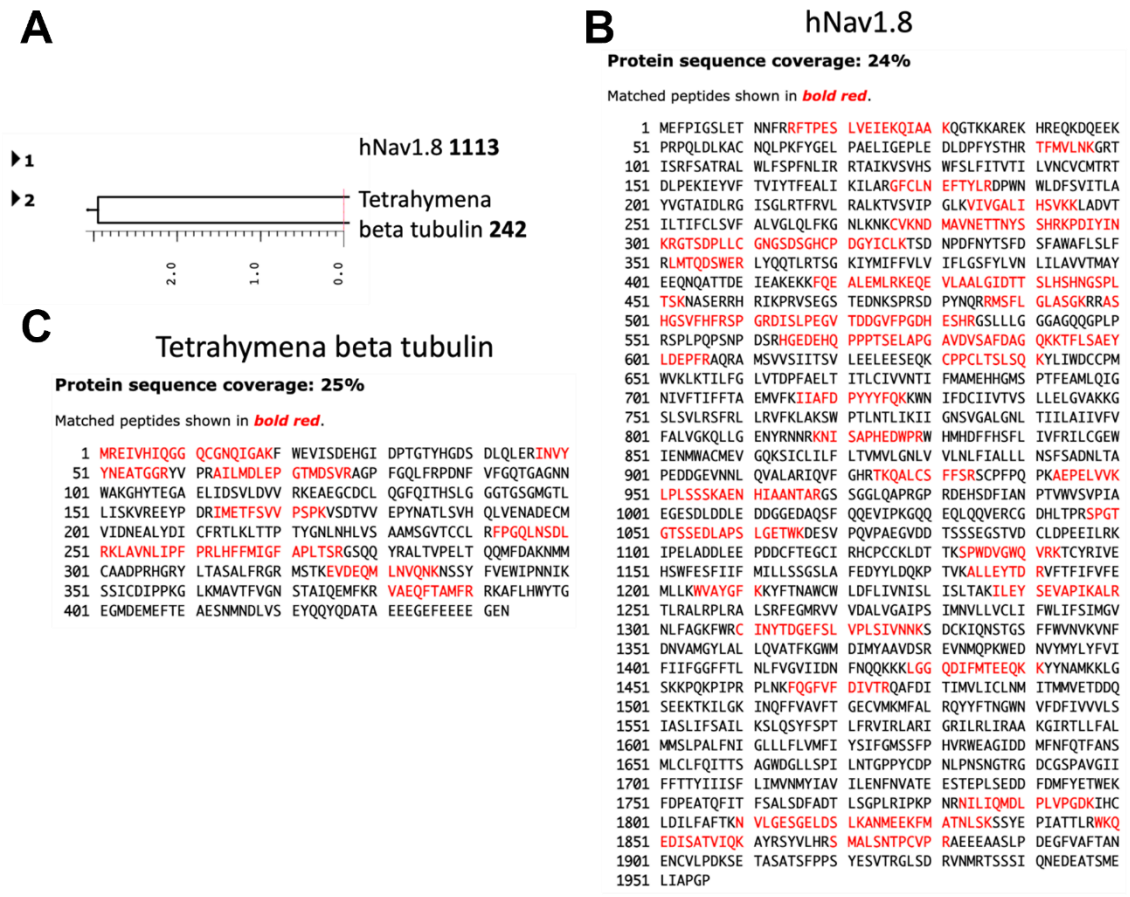

Figure 5.10. LC-MS results of band 2. Liquid chromatography-mass spectrometry (LC-MS) result of Band 2. (**A**) shows trees produced by the MASCOT server of top protein matches with scores of relative probabilities for each protein. The top result is hNav1.8 and second result is *Tetrahymena* β-tubulin, a cytoskeletal protein. (**B**) shows sequence coverage of hNav1.8 by matching detected peptides in red, with 24% total coverage consistent over the length of the protein. (**C**) shows sequence coverage of *Tetrahymena* β-tubulin (approximately 50 kDa), as in (**B**).

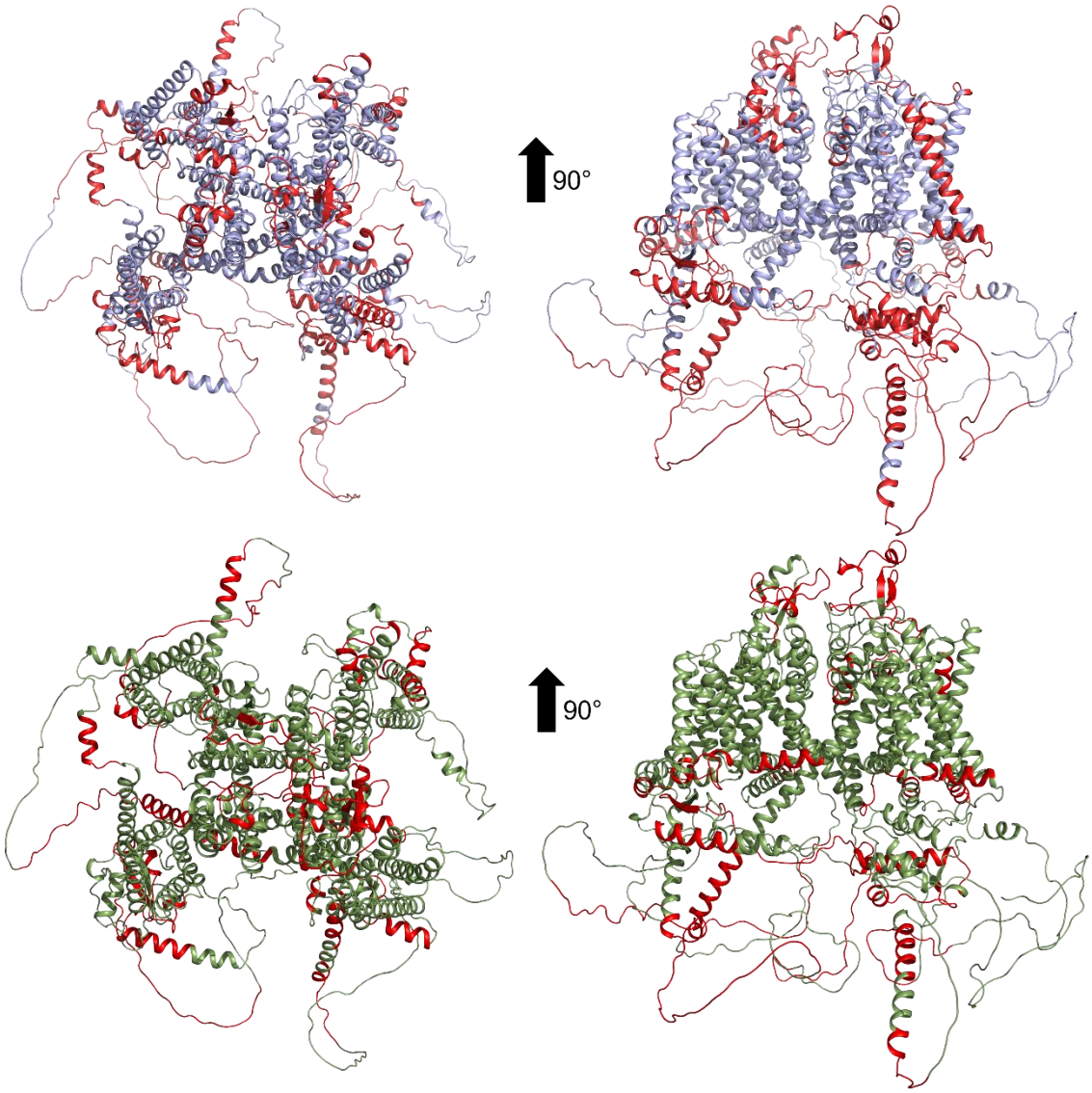

Figure 5.11. Representation of peptide hit coverage on the AlphaFold2 structure prediction of hNav1.8. Comparison of sequence coverage of hNav1.8 from liquid chromatography-mass spectrometry (LC-MS) results of Band 1 (top, blue) and Band 2 (bottom, green), with peptide matches highlighted in red on a predicted structure of hNav1.8 produced by AlphaFold2 (Jumper *et al*., 2021). Band 1 gives peptide matches over a greater proportion of the protein (44% versus 24%), whilst both give matches over each of the four domains of the protein. Figure made using PyMol molecular graphics software (version 2.0; Schrödinger, LLC).

5.3.3. Attempts to reduce degradation

The first attempt to reduce degradation involved reducing the time taken for the purification from 3 days to 2, to decrease the time available for degradation to occur. Usually in this project, the sample has overnight incubations with both the anti-FLAG M2 affinity gel and then, after elution the next morning, another overnight incubation with Ni^{2+} Sepharose 6 Fast Flow resin, before elution and SEC the following day. In this case, the anti-FLAG incubation was reduced from overnight to 3 h and then the elution loaded onto the Ni resin so that only one overnight incubation was needed. As seen in Table 5.2 as well as Figures 5.12 and 5.13, this resulted in a slight reduction in yield versus the standard protocol (78% compared to the standard protocol), but degradation was still present, with multiple bands appearing in western blots of purified sample.

Table 5.2. Quantitative comparison of attempts to reduce degradation. Yields in mAU*mL are calculated by measuring the area under the target peak of size-exclusion chromatography (SEC) traces. Standard and pooled purification yields are included for comparison.

Method	Yield (mAU*mL)	Relative to Standard (%)
Standard	23	100
5 pooled purifications	53	231
2 day purification	18	78
A8-35 exchange	26	113

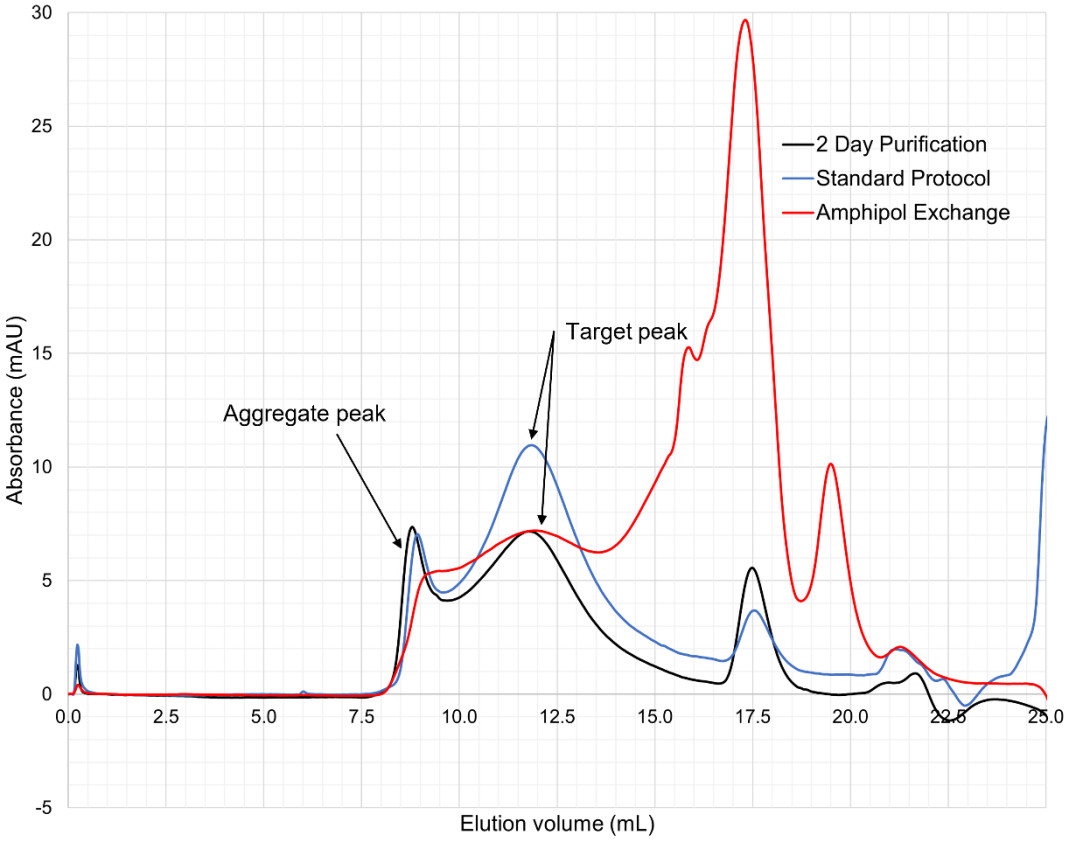

Figure 5.12. UV absorbance traces comparing attempts to reduce degradation. Comparison of size-exclusion chromatography (SEC) traces of the purification of hNav1.8 by standard protocol (blue), 2-day purification (black) or by A8-35 amphipol exchange after FLAG affinity purification (red).

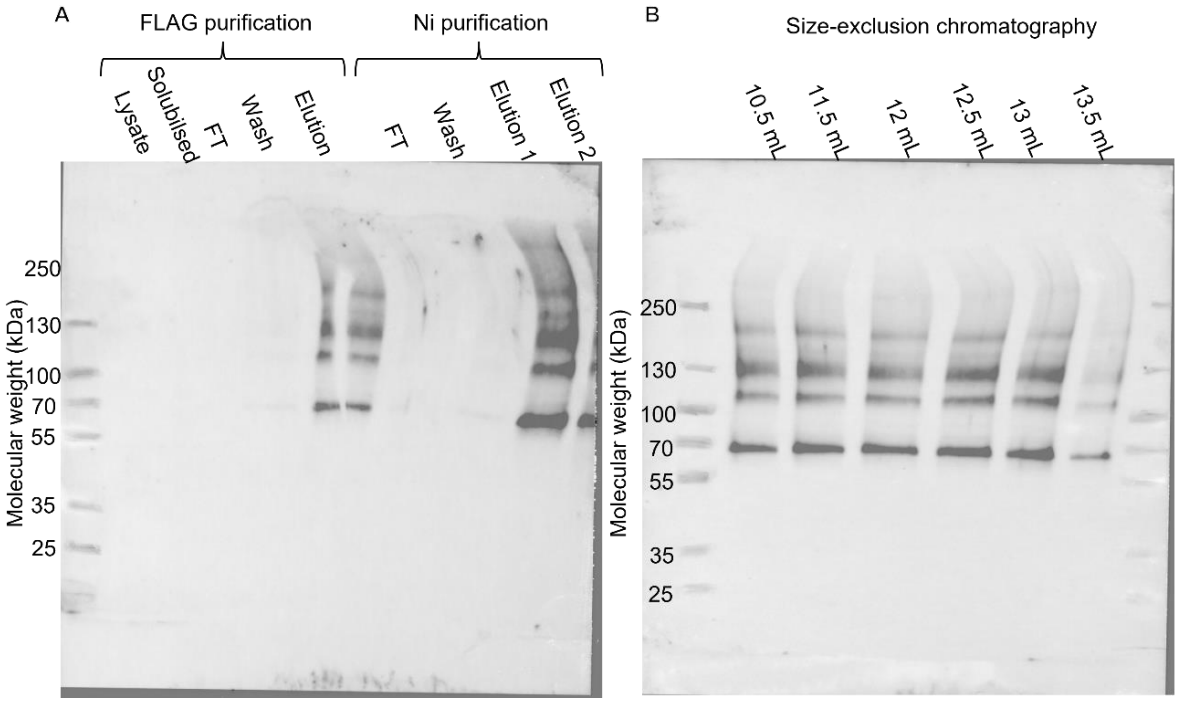

Figure 5.13. Western blots assessing degradation reduction. Western blots of the 2-day purification of hNav1.8. FT = flow through. (**A**) shows a western blot of FLAG and Ni affinity purification steps and (**B**) shows fractions over the target peak from size-exclusion chromatography (SEC).

As with the western blots shown in Figure 5.4, the approximately 50 kDa band of the SDS-PAGE seen in Figure 5.8 sent for LC-MS analysis does not appear on the western blots, which could mean that it is in fact β-tubulin from the expression organism, although it could also be fragment of hNav1.8 that does not contain the antigen targeted by the antibody. As membrane proteins can be more stable when solubilised in amphipol compared to detergent (see section 2.3.1), another method that was attempted was to exchange the sample into A8-35 amphipol by overnight incubation after elution from the anti-FLAG resin. The Ni^{2+} resin step was excluded from this preparation to reduce the complexity of knowing whether the protein had exchanged into the amphipol or not. Figure 5.12 shows that this led to a broader but not higher magnitude magnitude target peak, with larger impurity peaks which eluted later, and due to the further protein loss that would likely occur performing the full purification it was decided not to continue. A quantification of the area under the curve of the target peak is shown in Table 5.2, which suggests an increase of 13% over the standard protocol, although the lack of resolution between the target peak and previous and subsequent peaks makes it difficult to definitively discern from other species. Negative stain micrographs of the target peak fractions from amphipol and 2 day preps showed particles were well-dispersed but a range

of sizes and shapes, suggesting a non-homogenous sample (see Figure 5.14). At this point, due to the wide variety of purification conditions attempted, as well as the expense and effort required to produce enough protein for visualisation by cryo-TEM, it was decided that recombinant *T. thermophila* cells were not a suitable expression system for production of purified hNav1.8 and this aspect of the project was concluded.

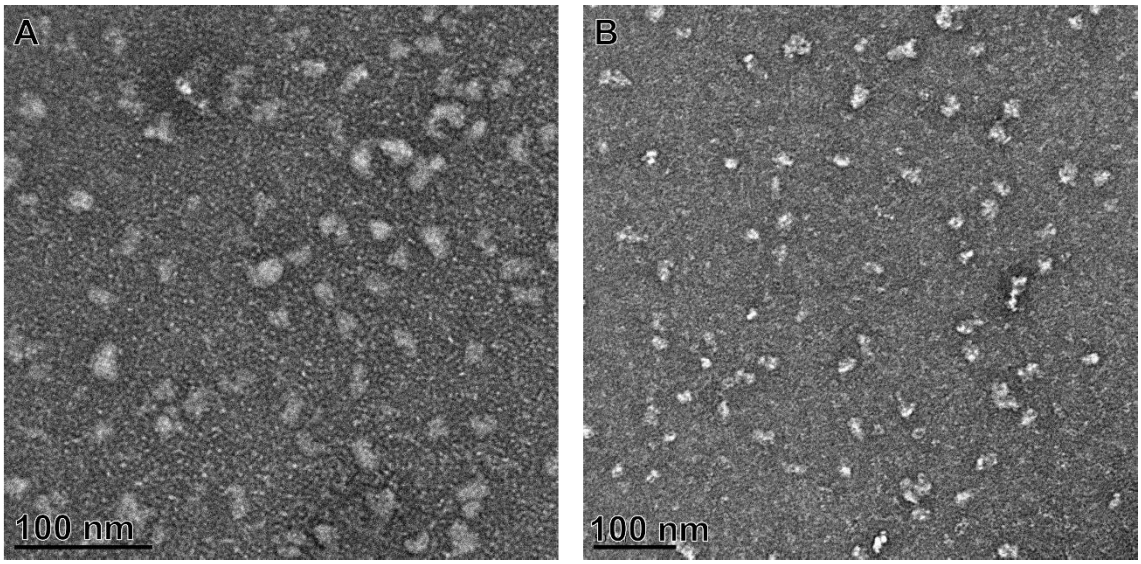

Figure 5.14. Negative stain screening assessing degradation reduction attempts. Negative stain micrographs of target peak fractions from (**A**) 2-day purification and (**B**) exchange into A8-35 after FLAG affinity purification.

5.4. Conclusion: *T. thermophila* is not suitable as an expression system for hNavs

The aim of this project was to obtain high-resolution structural data of hNav1.8, and in the process show whether *T. thermophila* was an effective system for expressing human membrane proteins; extensive trials of purification conditions unfortunately did not lead to an appropriate sample, despite the range of conditions applied. Western blotting showed that multiple degradation products are likely present in the sample, and LC-MS suggested that hNav1.8 may have been bound to ubiqutin, known as a tag for degradation by proteasome complexes (Ciechanover, Orian and Schwartz, 2000). Examination of the genome of *T. thermophila* has shown that it has undergone lineage-specific gene family expansions in genes related to detecting and responding to changes and stress response to factors in its local environment (Ruehle, Orias and Pearson, 2016), in particular proteases such as the ubiquitin carboxyl-terminal hydrolase family (Eisen *et al.*, 2006). This could help explain the high level of proteoylsis of hNav1.8, counteracting any supposed improvement in recombinant membrane protein expression.

If *T. thermophila* had been successful in expressing hNav1.8, it could have had potential to take the place of mammalian cell culture in some conditions, as media for *T. thermophila* cell growth is significantly cheaper than mammalian cell growth media. Two which have been shown to be effective, Neff medium and SSP medium, can be bought for between £15-25 per litre, whilst medium for HEK293 can cost upwards of £100 per litre. This would require issues with proteolysis to be overcome, for example with protease inhibitors specific to *T. thermophila*; this is an active area of research, with E-64 cysteine protease inhibitor added during cultivation shown to increase the yield of extracellularly secreted target proteins (Dündar Orhan *et al.*, 2024).

Chapter 6

Results: Purification and High-resolution Structural Determination of NavMs in Membrane Scaffold Protein Nanodiscs

6.1. Chapter summary

Since the first discovery of a prokaryotic voltage-gated sodium channel (VGSC, Nav), NaChBac from the Gram-positive bacterium *Bacillus halodurans,* by the D.E. Clapham group in 2001 (Ren *et al.*, 2001), their high sequence and structural homology with human Navs (hNavs) has led to their effective use as structural, mechanistic, and pharmacological models for hNavs (Nurani *et al.*, 2008; Payandeh *et al.*, 2011; McCusker *et al.*, 2012; Bagnéris *et al.*, 2014; Boiteux *et al.*, 2014; Zanatta *et al.*, 2019; Sait *et al.*, 2020; Sula *et al.*, 2021; Wisedchaisri *et al.*, 2023). Coupled with their suitability as models for hNavs, they are also more amenable to recombinant expression, meaning that prokaryotic Navs can be utilised as resource for higher-throughput examination of drug and toxin binding, the effects of disease mutations, and attempts to explore the mechanism of opening, closing, and inactivation of sodium channels. The majority of prokaryotic Nav structures released to date have been obtained by X-ray crystallography, which proved more suitable a technique than cryo-transmission electron microscopy (cryo-TEM) prior to the 'resolution revolution' (Kühlbrandt, 2014); the low molecular weight of prokaryotic Navs meant that the first cryo-TEM structures of these were only possible when the size of the particle was artificially increased by methods such as binding to Fab antibody fragments (Xu *et al.*, 2019) or linking maltose-binding protein to the N-terminus of each monomer (Wisedchaisri *et al.*, 2019).

Structures of full-length NavMs have previously been elucidated, but only by X-ray crystallography (Sula *et al.*, 2017, 2021; Sait *et al.*, 2020). Cryo-TEM is a useful complementary technique as it allows for structural analysis in conditions closer to that of the native state of the protein in question and can capture multiple conformations or states within a sample. Specifically in the case of using prokaryotic Navs to investigate drug binding in hNavs, drugs which bind to the outer surface of the protein (such as gating

modifier toxins (GMTs) which bind to the voltage-sensor domain (VSD)) may be excluded due to crystal packing of each individual protein in the lattice; this is not a concern with dispersed particles in a cryo-TEM sample. A previous attempt to obtain a high-resolution cryo-TEM structure of NavMs (of *Magnetococcus marinus*) by Dr Jennifer Booker of the Wallace lab was ultimately unsuccessful (described in more detail in section 6.2.), but the advancements in cryo-TEM equipment and processing in the intervening years, coupled with the publication of cryo-TEM structures in detergent and membrane scaffold protein (MSP) nanodiscs of another prokaryotic Nav, NaChBac (Gao *et al.*, 2020), suggested that cryo-TEM analysis of NavMs was worth pursuing further. With this in mind, the aim of this project was to show that cryo-TEM could be used to obtain atomic-resolution structural data of NavMs, opening up a new method for the analysis of different conformational states and drug binding to an important model for human voltage-gated sodium channels.

The investigation involved expressing and purifying NavMs in detergent as in Sula *et al.*, 2017, so that it could be reconstituted into MSP nanodiscs and analysed by cryo-TEM. Successful trials of reconstituting NavMs into nanodiscs with varying molar ratios of NavMs:MSP:phospholipid were performed based around the molar ratio used in Gao *et al.*, 2020, with the most promising ratio, judged by size-exclusion chromatography (SEC) trace, used for scaled-up reconstitutions. The presence of NavMs in MSP2N2 nanodiscs was confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and western blot, and negative stain electron microscopy showed particles of the expected size and shape for MSP2N2 nanodiscs. Grid conditions for cryo-TEM were optimised for particle visibility and ice thickness until judged appropriate for data collection; the resultant data processing in cryoSPARC (Punjani *et al.*, 2017) suggested that reconstituting NavMs into nanodiscs had been achieved, however there were many empty nanodiscs in the sample which were dominating the reconstruction. A step to remove empty nanodiscs was introduced, and data processing of the resultant dataset led to the successful reconstruction of a 3.07 Å global resolution Coulomb potential map with characteristics distinct from previously published X-ray crystallography structures of NavMs.

6.2. Preliminary research

Previous work in the Wallace lab had established a protocol for the expression and purification of full-length NavMs for structural study by X-ray crystallography (Sula *et al.*, 2017), involving expression in *E. coli,* solubilisation in N-dodecyl-β-D-maltopyranoside (DDM) before exchange into decanoyl-N-hydroxyethylglucamide (HEGA-10) during Ni^{2+} affinity chromatography, followed by a final SEC purification. As mentioned in the previous section, Dr Jennifer Booker of the Wallace lab began an investigation to adapt this protocol for cryo-TEM analysis of NavMs (Booker, 2019); negative stain and cryo-TEM screening of NavMs in HEGA-10 showed heterogeneous detergent micelles, and empty micelles could not be discerned from those containing NavMs. NavMs was therefore exchanged into amphipol A8-35 and screened on holey-carbon grids, at which point the expected 'mushroom' shape of NavMs was seen, and the sample was judged suitable for data collection on a Titan Krios microscope.

Processing of the resultant dataset in RELION 2.0 (Scheres, 2012) gave 2D classes showing internal features of NavMs, however with significant streaking and blurring, with the flexible C-terminal domain (CTD) found in a range of positions relative to the transmembrane region of the protein, and limited top or bottom (i.e., extracellular, or intracellular) views available (see Figure 6.1). When compared to the previously published structures of 'open' NavMs (PDB 5HVX) (Sula *et al.*, 2017) and 'closed' NavAb (PDB 5VB2) (Lenaeus *et al.*, 2017), the position of the CTD suggested that NavMs in this dataset was more similar to the closed conformation (see Figure 6.1).

A second dataset was collected with a 30° tilt in an attempt to retrieve more top views, but 2D classes did not appear significantly different to the 0° tilt dataset; it was suggested that top views were not being aligned accurately due to the flexibility of the VSD and CTD. 3D reconstruction was attempted using an initial model of a 30 Å low-pass filtered volume of the 'closed' NavAb crystal structure; from initial particle picks of 287,608 from dataset 1 and 312,734 of dataset 2, 48,655 and 61,486 'good' particles were selected for each reconstruction respectively. This resulted in low resolution reconstructions which were asymmetric (NavMs should show C4 symmetry), not matching the known channel structure other than possessing a general transmembrane region and CTD. To improve the reconstruction, datasets 1 and 2 were merged and 3D refinement attempted again, both with and without the CTD masked out due to its flexibility with respect to the

transmembrane region. The resultant structures showed slight improvement as some helices became visible, as seen in Figure 6.2, but were still not well defined and were judged to be highly anisotropic. The project was paused at this point, but interest in developing a strategy for cryo-TEM study of prokaryotic Navs remained, and the project was taken up once again, due to the release of the aforementioned NaChBac structures and the improvements developed in cryo-TEM grids, direct electron detectors, and data processing hardware and software in the intervening years.

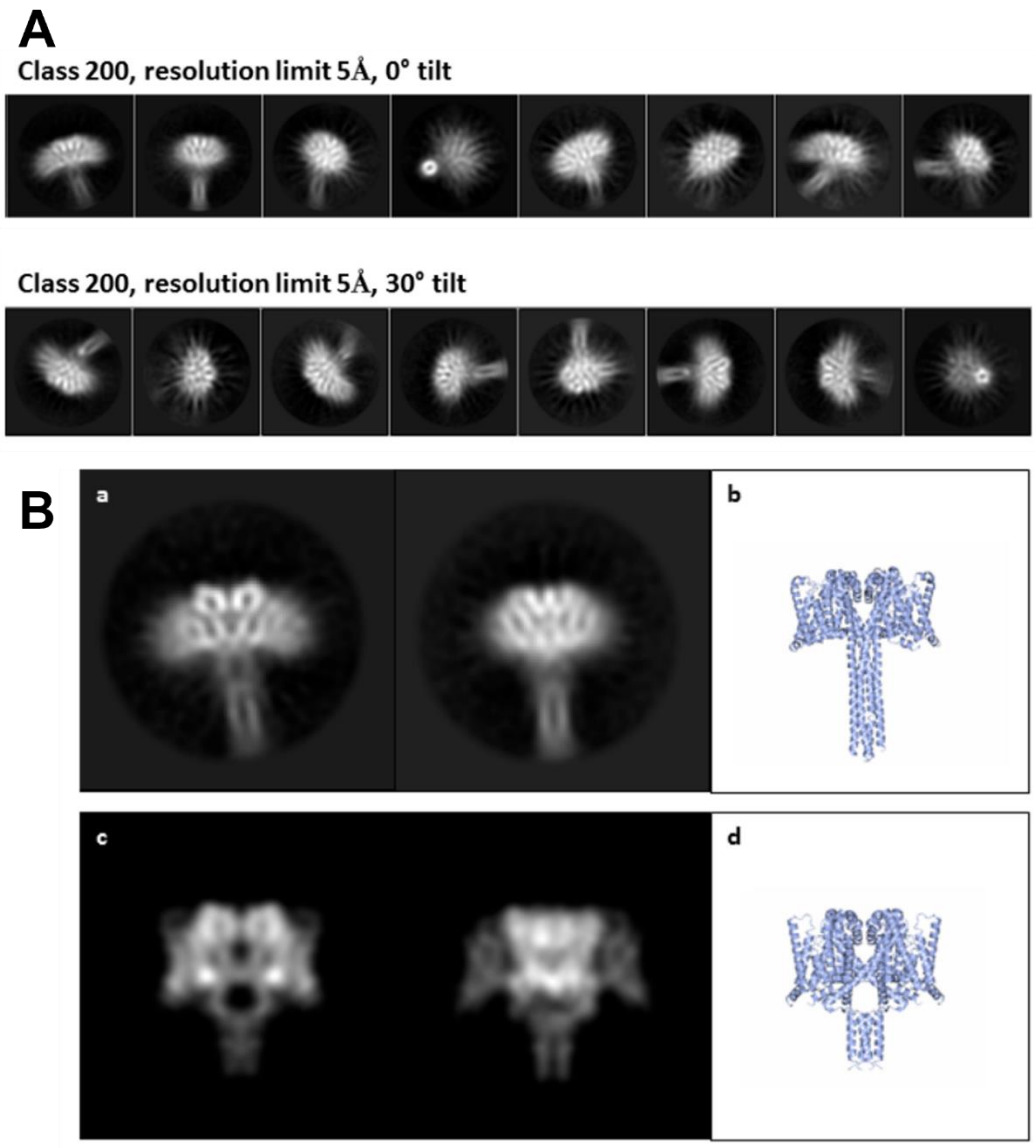

Figure 6.1. 2D classification of NavMs in A8-35 amphipol. Figure showing results of 2D classification of NavMs in A8-35 amphipol by Dr Jennifer Booker of the Wallace lab prior to this investigation. (**A**) shows representative 2D classes, with transmembrane 'head' and mobile C-terminal domain (CTD) 'tail' with internal structural detail. (**B**) (**a**) shows representative 2D classes for comparison with (**b**)**,** the 'closed' crystal structure of NavAb (PDB 5VB2, Lenaeus *et al*., 2017). Beneath this in (**c**) are matched reprojections of the 'open' NavMs crystal structure seen in (**d**) (PDB 5HVX, Sula *et al*., 2017). Figure adapted from Booker, 2019.

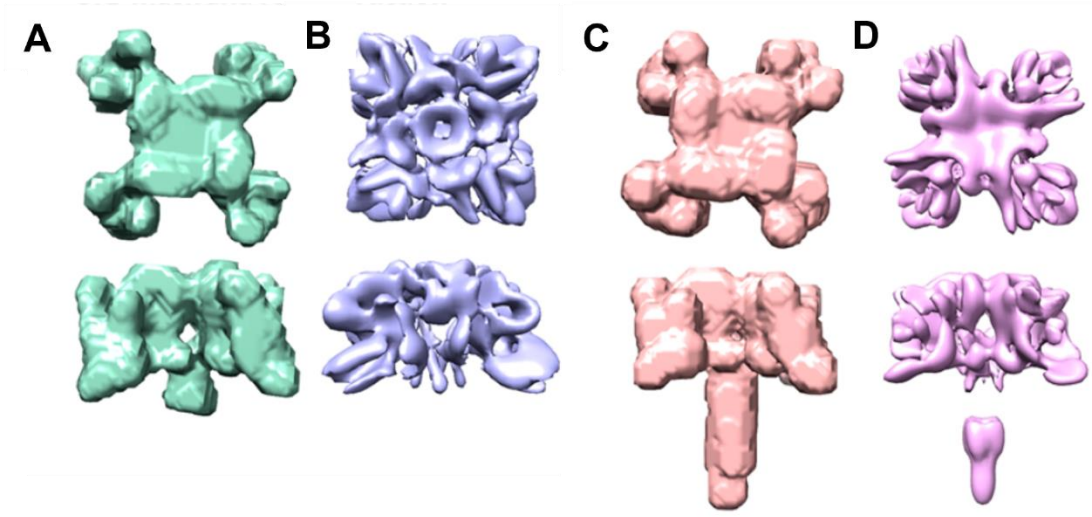

Figure 6.2. 3D reconstructions of NavMs in A8-35 amphipol. Figure showing 3D reconstructions of NavMs in A8-35 amphipol by Dr Jennifer Booker of the Wallace lab prior to this investigation, with and without masking of the C-terminal domain (CTD) to aid in alignment. From left to right we see top (top) and side (bottom) views of (**A**) ΔCTD mask (**B**) ΔCTD reconstruction (**C**) Full-length mask (**D**). Figure adapted from Booker, 2019.

6.3. Expression, purification, and reconstitution of NavMs into membrane scaffold protein nanodiscs

6.3.1. Expression and purification of NavMs in detergent

Please refer to Methods section 4.6 for full details of the expression and purification of full-length, wild type (WT) NavMs.

Ni^{2+} affinity chromatography: After recombinant expression in *E. coli*, lysing of cells, isolation of cell membrane fractions by ultracentrifugation, and solubilisation the of cell membrane fraction in 1.5% DDM, NavMs was isolated from the cell membrane fraction by HisTrap Ni^{2+} affinity chromatography (using a 6x poly-His tag at the N-terminus), with the presence of protein eluting from the column shown by ultraviolet (UV) absorbance trace (see Figure 6.3). This trace shows peaks with each wash (15% and 30% Buffer B containing 1 M imidazole) as histidine-rich proteins native to the expression system which have non-specifically bound to the column are removed by competition with imidazole. A larger peak after elution with 100% Buffer B (i.e., 1 M imidazole) can be seen, which is known from previous work in the Wallace group to contain purified NavMs, confirmed here in Figure 6.4 by SDS-PAGE (see Methods section 4.8 for details). Bands at approximately 25 kDa, 50 kDa, 75 kDa and 100 kDa are known to represent NavMs as the monomer, dimer, trimer, and tetramer respectively, as the association between each chain is not completely disrupted by SDS denaturation. These molecular weights are slightly below that of multiples of NavMs monomers and oligomers (a NavMs monomer is approximately 31 kDa), however the migration rate of membrane proteins on SDS-PAGE gels is known to be shifted by the altered interaction with SDS compared to cytosolic proteins (Rath *et al.*, 2009). Streaking in each lane of the SDS-PAGE gel is visible, but is also a known phenomenon in membrane proteins (Zhou *et al.*, 2000), and has been observed for pure NavMs in previous work by the Wallace group. Fractions from the final elution peak were collected and concentrated to approximately 250 μL in preparation for purification by SEC.

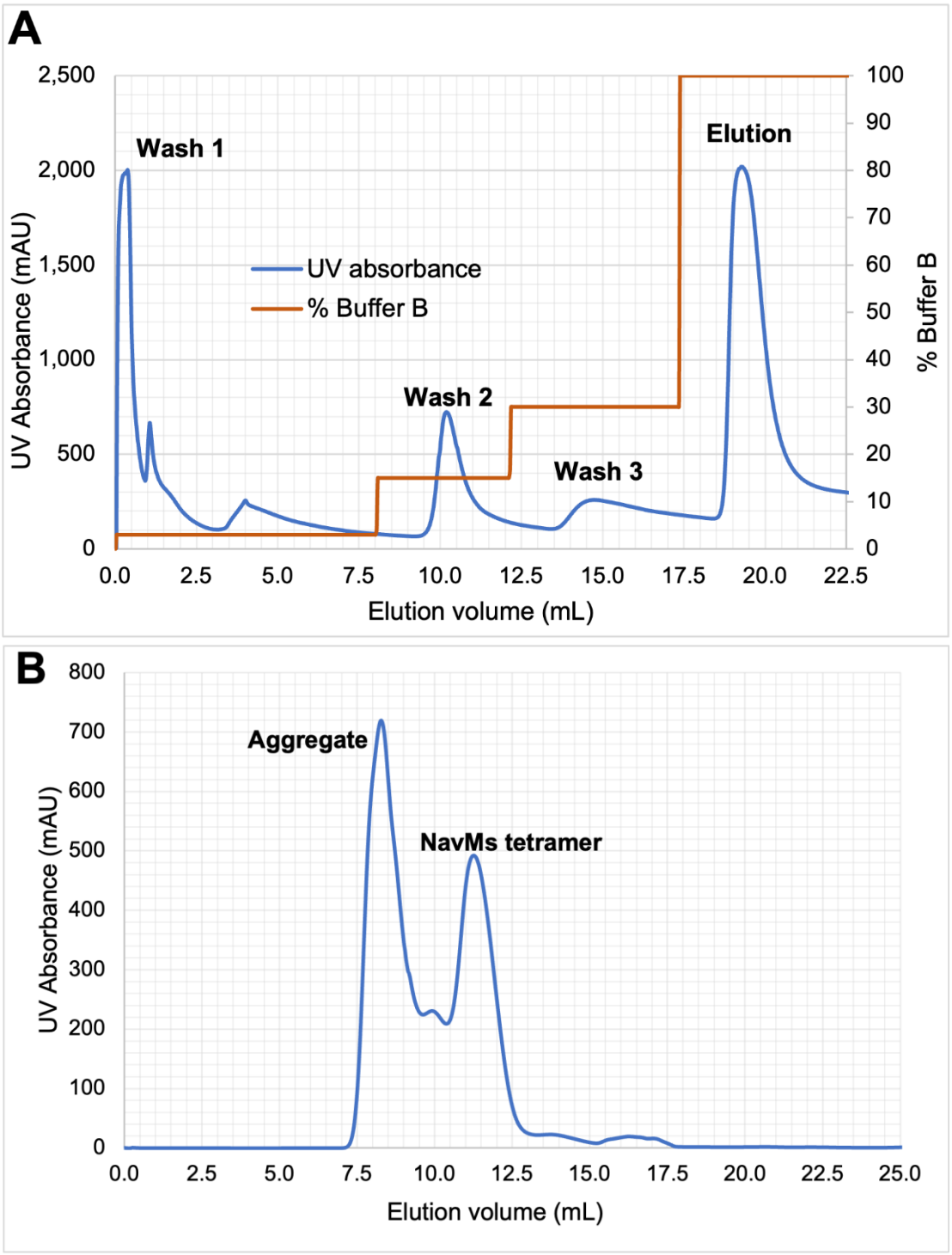

Figure 6.3. UV traces of purification of NavMs in detergent. Figure showing UV absorbance traces of (**A**) Histrap Ni^{2+} affinity chromatography and (**B**) size-exclusion chromatography (SEC) purification steps of NavMs in HEGA-10 detergent on a Superdex 200 Increase 10/300 GL column with 8 mL void volume.

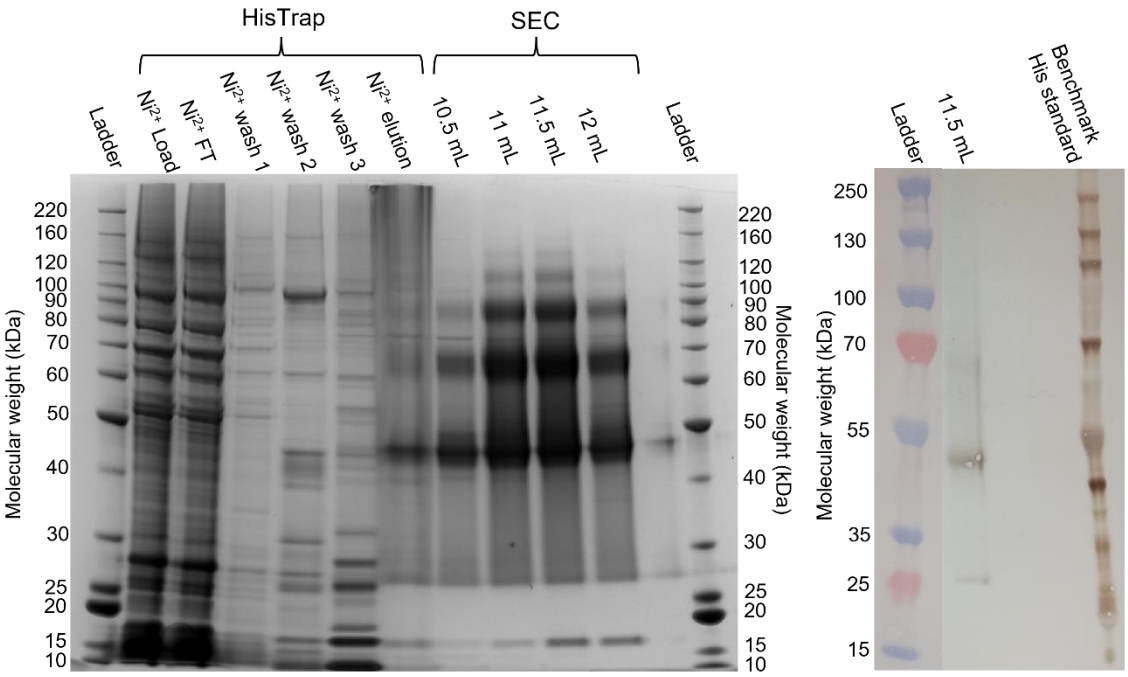

Figure 6.4. SDS-PAGE and western blot of Histrap and SEC purification fractions. Figure showing an SDS-PAGE gel (left) and western blot (right) of fractions from Histrap Ni^{2+} affinity chromatography and size-exclusion chromatography (SEC) purification steps of NavMs in HEGA-10 detergent. Bands representing NavMs monomer, dimer, trimer, and tetramer are expected to be seen at 25, 50, 75, and 100 kDa respectively. The western blot uses an anti-polyHis horseradish peroxidase-conjugated antibody.

Size-exclusion chromatography: Ni^{2+} affinity chromatography elution fractions were loaded onto a Superdex 200 Increase 10/300 GL column to separate out proteins in the sample by molecular weight. The UV trace of this purification step shows peaks at elution volumes of approximately 8 mL (the void volume of this column), 10 mL and 11.5 mL (see Figure 6.3); this is typical of previous study of NavMs by the Wallace group. Unpublished circular dichroism (CD) analysis by the group suggested that the 11.5 mL peak contains a greater proportion of secondary structure than the previous peaks and is therefore likely to be more folded, and therefore most useful for structural analysis; further confirming this, fractions from this peak were used for successful crystallographic studies of NavMs (Sula *et al.*, 2017). In this case, SDS-PAGE and western blotting were used to confirm the presence and purity of NavMs in the 11.5 mL peak fraction (see Figure 6.4), which was then used for reconstitution in MSP2N2 nanodiscs.

6.3.2. Reconstitution of MSP2N2 nanodiscs

In general, nanodiscs are nanometre-scale discs formed of membrane phospholipids encircled and stabilised by amphipathic protein, with hydrophobic side chains pointing inward towards the fatty tails of the phospholipids and hydrophilic side chains pointing out towards the aqueous environment (for full details, see Technical Introduction section 2.3). Variations in the application of nanodiscs include whether to use them to encapsulate protein directly from solubilised membranes, versus purifying the target in detergent and then reconstituting into nanodiscs, removing detergent with Bio-Beads; the choice of MSP, as this determines the size of the disc and therefore the size of the protein that can fit inside it; the choice of which phospholipid is most appropriate for the target protein; and the molar ratio of target protein, to MSP, to phospholipid, which is related to the size of the transmembrane region of the target protein, and the size of the head group of the phospholipid (head group occupancy). This ratio is important and must be optimised to avoid losing target protein to aggregation, formation of excess empty nanodiscs (which can be difficult to tell apart from full discs with a largely transmembrane protein like NavMs), or aggregation of the MSP itself. As NavMs and NaChBac share approximately 57% sequence similarity according to EMBOSS Needle (Rice, Longden and Bleasby, 2000) and share the same 24 transmembrane helix architecture, it was decided to use the same MSP and phospholipid as was successfully employed for NaChBac, these being MSP2N2 (Grinkova, Denisov and Sligar, 2010) and 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC). MSP2N2 has a diameter of between 15-17 nm, whilst the

transmembrane region of NavMs (and NaChBac) is approximately 10 nm across, meaning that it should fit comfortably into the disc. Using the molar ratio of 1:12:520 of target protein (tetramer), to MSP2N2 (monomer), to POPC employed for NaChBac as a starting point, trials were conducted to ensure that this ratio was suitable. For full details of the reconstitution process, see Methods section 4.7.

Reconstitution of NavMs into nanodiscs was trialled at 1/3 full scale at molar ratios of 1:12:470, 1:12:520, and 1:12:570. Concentrations of freshly purified NavMs taken from the 11.5 mL peak fraction and MSP2N2 were measured by Nanodrop spectrophotometer, with the required volume of POPC calculated from a known concentration of 10 mM. After 1 h of incubation at 4 °C, the mixture was left incubating for 1 h with 0.3 g/mL Bio-Beads to form nanodiscs. After this 1 h incubation, the mixture was removed from the Bio-Beads and added to fresh Bio-Beads for overnight incubation, and Bio-Beads refreshed again for another 1 h incubation before centrifugation to remove aggregate and SEC (using a Superose 6 Increase 10/300 GL column) to separate out resultant species. The ensuing SEC UV traces are seen in Figure 6.5, alongside the SEC trace provided in the supplementary information of Gao *et al.*, for comparison. Each of the molar ratios show the same two-peak trace, with the first eluting at approximately 14 mL and second just past 16 mL. These peaks match the positions of NaChBac in nanodiscs and empty nanodiscs seen with NaChBac, suggesting that reconstitution was successful; however, the molar ratios of 1:12:470 and 1:12:570 each have ‘shoulders’ before the peak, suggesting another larger species was also present. Negative stain EM was used to judge whether the resultant nanodiscs had the expected appearance, which can be seen in comparison to a sample cryo-TEM micrograph of NaChBac in nanodiscs in Figure 6.5. The discs appear visually similar to NaChBac in nanodiscs in size and shape and appear as discrete discs rather than aggregate; it was therefore decided to use 1:12:520 as the molar ratio for a full-scale reconstitution and purification.

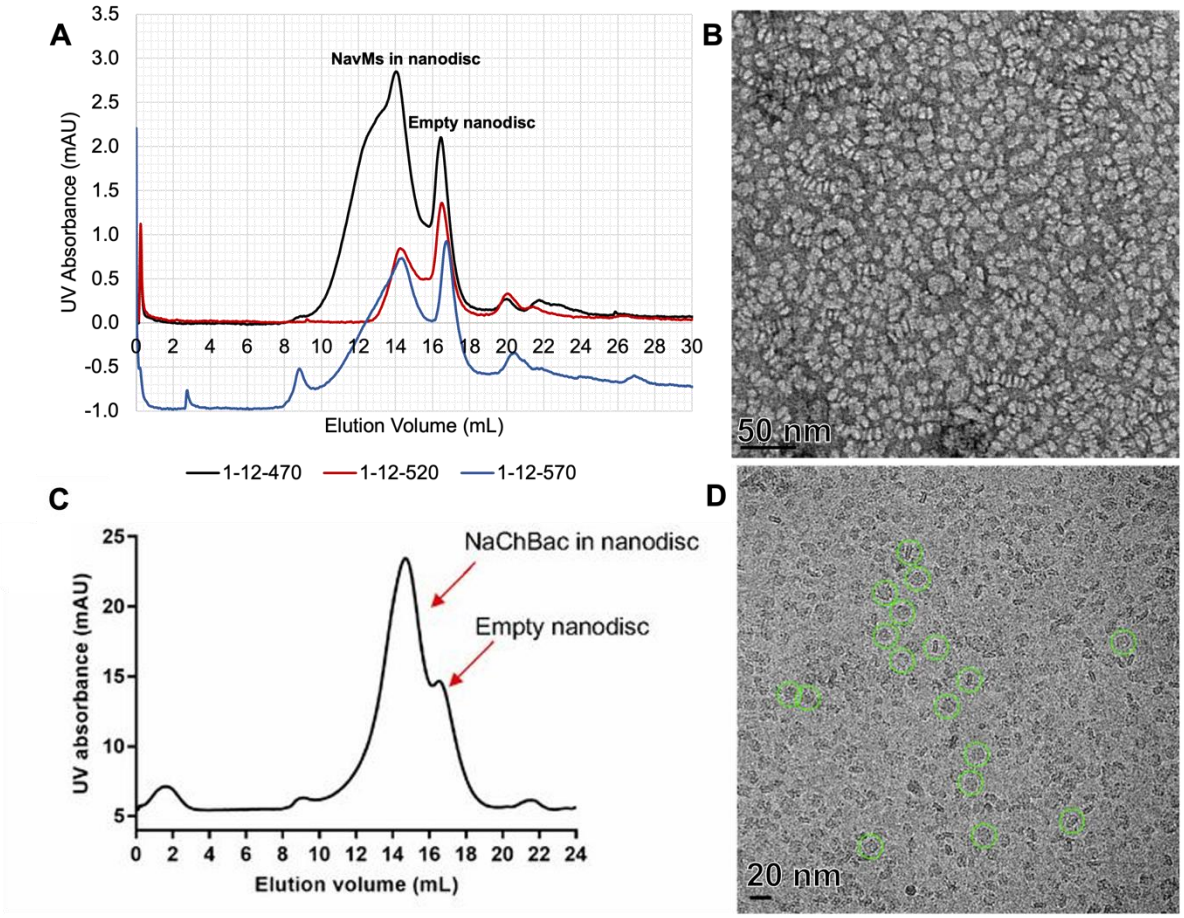

Figure 6.5. Molar ratio trials for reconstitution and purification of NavMs in nanodiscs. Figure showing the results of trials of varying molar ratios of NavMs:MSP2N2:lipid for membrane scaffold protein (MSP) nanodisc reconstitution and purification. (**A**) shows UV absorbance traces of size-exclusion chromatography (SEC) performed after overnight incubation of nanodisc components with Bio-Beads. Equivalent UV trace of NaChBac in nanodiscs published in Gao *et al.*, 2020, shown underneath in (**C**) for comparison. (**B**) shows a negative stain electron micrograph of the peak fraction of the 1:12:520 ratio, with (**D**), a cryo-transmission electron micrograph published by Gao *et al.,* 2020 showing NaChBac in nanodiscs underneath for comparison. (**C**) and (**D**) were adapted from Gao, *et al.,* 2020.

The SEC UV trace of the full-scale purification with the same multi-step Bio-Bead incubation as the trials showed the same two-peak pattern, as seen in Figure 6.6, although the NavMs-nanodiscs peak appeared smaller in relation to the empty nanodiscs peak in this purification. Fractions were collected from the suspected NavMs-nanodisc peak and screened by negative stain EM directly from the SEC fraction as well as concentrated to approximately 1.5 mg/mL for cryo-TEM screening, using Quantifoil Cu 300 mesh 1.2/1.3 grids for cryo-TEM in this case (see Methods section 4.9); blotting conditions were 100% humidity, 4 °C, 5 s blot, 0 s wait time, and a blotting force of -10, with 3.5 µL of sample used for both negative stain and cryo-TEM (see Figure 6.6 for cryo-TEM screening micrographs). Again, the sample appeared suitable by negative stain, however blotting for 5 s as described in the Methods of Gao *et al.*, appeared to make the ice too thin, and only vacuum was visible on subsequent cryo-TEM micrographs. The second reconstitution was altered slightly, with Bio-Beads being added once for the entire incubation rather than being refreshed after the first hour of incubation and again for an hour of incubation the following morning; a greater yield of the NavMs-nanodisc peak was seen under these conditions, which, speculatively, could be due to detergent adsorbing to fresh Bio-Beads at a faster rate than Bio-Beads which already have some detergent adsorbed, meaning that the components have less time to form nanodiscs without aggregating. Blotting conditions for cryo-TEM screening of this reconstitution used a 3 s blot rather than 5 s and UltrAufoil Au 300 mesh 1.2/1.3 grids rather than Quantifoil, and showed faint particles, as seen in Figure 6.6, although the ice was too thick to view particles for much of the grids. A third reconstitution, again using a single Bio-Bead addition, was performed using NavMs that had been snap-frozen and defrosted (rather than freshly purified NavMs), and a 4 s blot proved successful, showing well-dispersed particles without noticeable aggregation or contamination. A grid from the same preparation and produced in the same blotting conditions was used for data collection on a Titan Krios microscope. This dataset will be referred to as Dataset 1 from hereon.

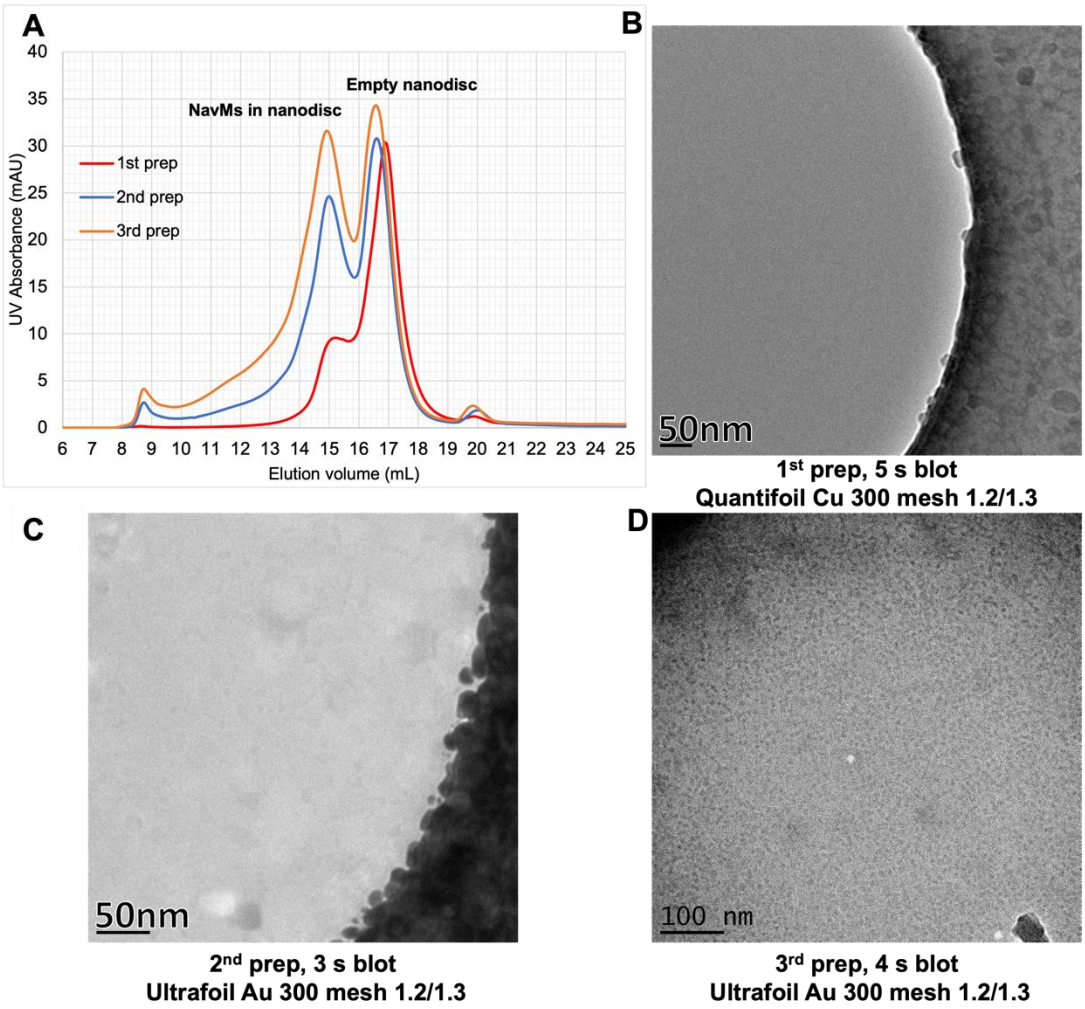

Figure 6.6. UV traces and cryo-TEM screening of full-scale nanodisc preparations. Figure showing the results of full-scale preparations of NavMs reconstituted in membrane scaffold protein (MSP) nanodiscs. (**A**) shows UV absorbance traces of size-exclusion chromatography (SEC) of reconstituted nanodiscs. The 1st preparation (red, multi-stage Bio-Bead addition) involved multiple additions of fresh Bio-Beads at 0.3 mg/mL during nanodisc reconstitution, whilst the 2nd and 3rd used only a single addition of 0.3 mg/mL Bio-Beads. (**B**), (**C**), and (**D**) show cryo-TEM screening micrographs from the target peak fractions (NavMs in nanodiscs) of the 1st, 2nd, and 3rd preparation respectively, with blotting conditions shown underneath.

6.4. Data collection and processing (Dataset 1)

6.4.1. Data collection

A total of 20,620 frames were collected using 3 shots per hole, with collection in super resolution mode and recorded binned by 2 to match a physical pixel size of 0.828 Å (for full details of data collection see Methods section 4.10.1). Movies were imported into cryoSPARC, and all subsequent data processing performed in cryoSPARC version v3.3.1.

6.4.2. Data processing

Unless otherwise specified, jobs were built using the default settings in cryoSPARC. A graphical depiction of the processing pipeline performed on Dataset 1 resulting in 2D classes identifiable as NavMs reconstituted into nanodiscs can be seen in Figure 6.7.

Pre-processing: After importation, Patch motion correction (multi) was performed on all 20,620 frames and patch contrast transfer function (CTF) estimation (multi) performed on motion-corrected micrographs. The manually curate exposures job was used to exclude micrographs with poor characteristics such as high levels of motion or low resolution CTF fit resolution. In this case, micrographs were excluded with a CTF fit resolution of > 20 Å and total full-frame motion distance of > 40 pixels, leading to 20,100 micrographs being accepted.

Particle picking and 2D classification: The blob picker job was then used with a minimum and maximum particle diameter of 100 Å and 200 Å respectively, which picked 13,940,135 particles. These picks were manually inspected and those with a normalised cross correlation (NCC) score of below 0.350, indicating poor fit to the parameters, excluded, as well as picks in empty ice space or of high-contrast ice contamination or hole edges, curating the particle selection down to 6,865,695. Particle locations were extracted with a box size of 410 Å (larger than the size of the particle in order to retain high-resolution CTF data which is spread to a larger spatial frequency), and a 50 class 2D classification job performed with force max over poses/shifts turned off, number of online-EM iterations set to 40, batchsize per class set to 200 (each of these are recommended in cryoSPARC documentation for dealing with small membrane proteins) and a circular mask diameter of 200 Å. None of these classes definitively showed NavMs in nanodiscs, but possible side and top views (representing 458,503 of the 5,893,985 particles classified) of nanodiscs were picked from the 50 resultant classes and used as

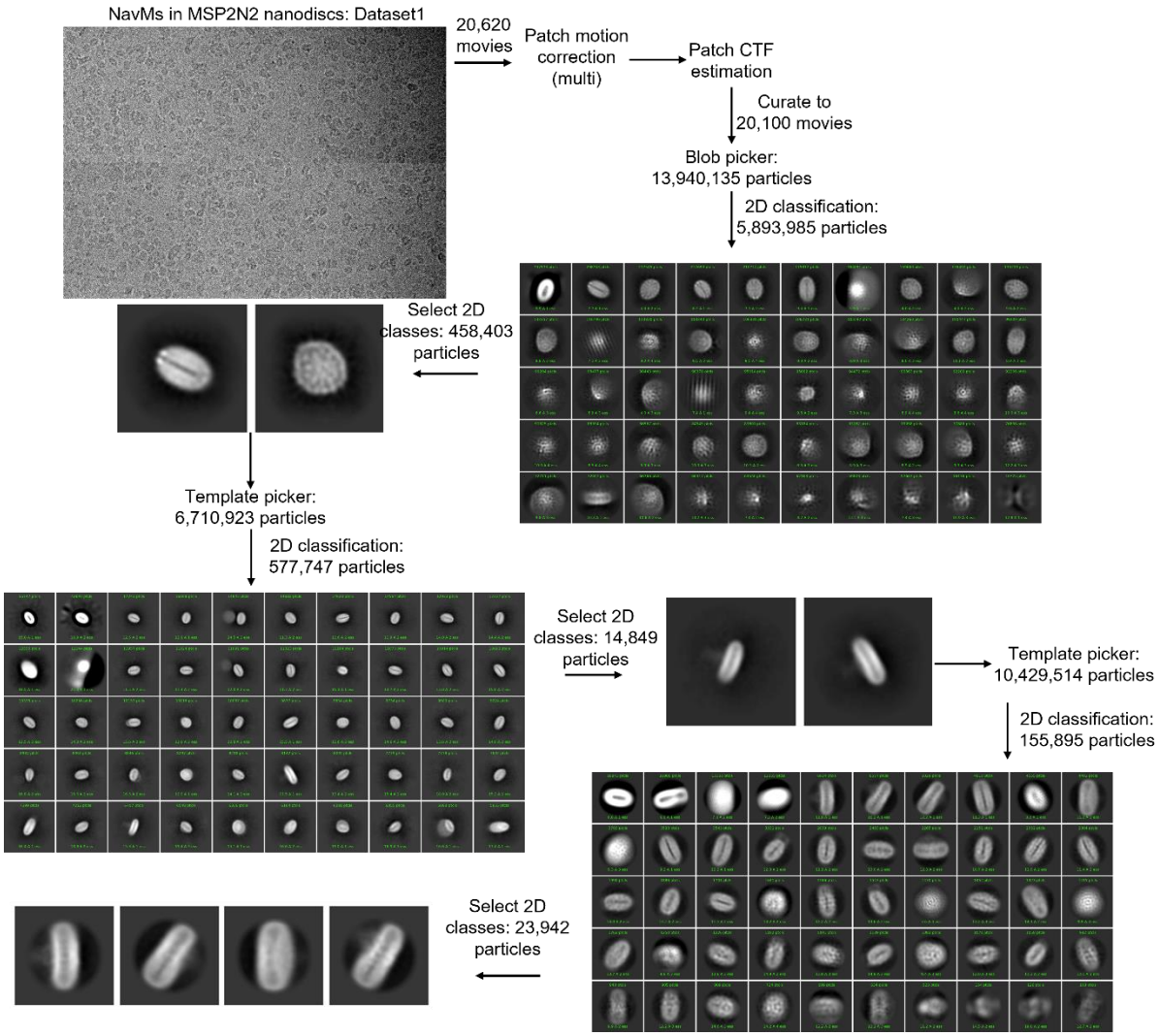

Figure 6.7. Cryo-TEM data processing pipeline of Dataset 1 of NavMs in MSP2N2 nanodiscs. Figure showing the general pipeline of the processing of Dataset 1. All processing performed in cryoSPARC version v3.3.1. 20,620 movies were imported into cryoSPARC for pre-processing by patch motion correction and patch contrast transfer function estimation. Resultant micrographs were curated to remove unsuitable frames and subjected to blob picking and 2D classification. Classes judged to possibly represent NavMs in nanodiscs were used as templates for two rounds of template picking and further 2D classification, which gave 4 classes consisting of 23,942 particles with visible internal secondary structure or possible NavMs C-terminal domains. The small number of particles identified as NavMs in nanodiscs versus likely empty nanodiscs indicated that steps should be taken to remove empty nanodiscs.

references for the template picker job with the particle diameter setting at 310 Å. 6,710,923 particles were picked, with curation of these picks bringing the number down to 699,892. Particle locations were extracted and 2D classification performed with the same parameters as before, and 2 out of the 50 classes (representing 14,849 of 577,747 classified particles) produced showed a disc shape with a protruding tail (see Figure 6.8**A**), which, encouragingly, appeared similar to example 2D classes shown in Gao *et al*., 2020; this suggested that, although it was not present in abundance, NavMs had been successfully incorporated into the nanodiscs. The other 48 classes appeared to contain ice or what may be empty nanodiscs. These 2 classes were used as templates for further picking to try to find more NavMs-nanodisc complex particles. This time, with a particle diameter setting of 250 Å, 10,429,514 particles were picked; this was curated down to 167,500 particles, extracted with a box size of 410 pixels, and submitted for 2D classification with the same parameters as previously except for a reduced box size of 170 Å. Most of the classes again appeared to be junk particles, showing the putative empty nanodiscs or ice, with only 4 classes representing 23,942 particles out of 155,895 classified likely representing the target; however, within these classes, some internal helices could be made out, as well as a CTD (see Figure 6.8**B**).

At this point another nanodisc reconstitution was ready to be performed, and as most of the particles and classes from Dataset 1 appeared to show empty nanodiscs (empty nanodiscs are known to be difficult to fully remove even with good peak separation in SEC), it was decided that for the next reconstitution an extra step should be added to the purification to remove said empty nanodiscs before preparing grids.

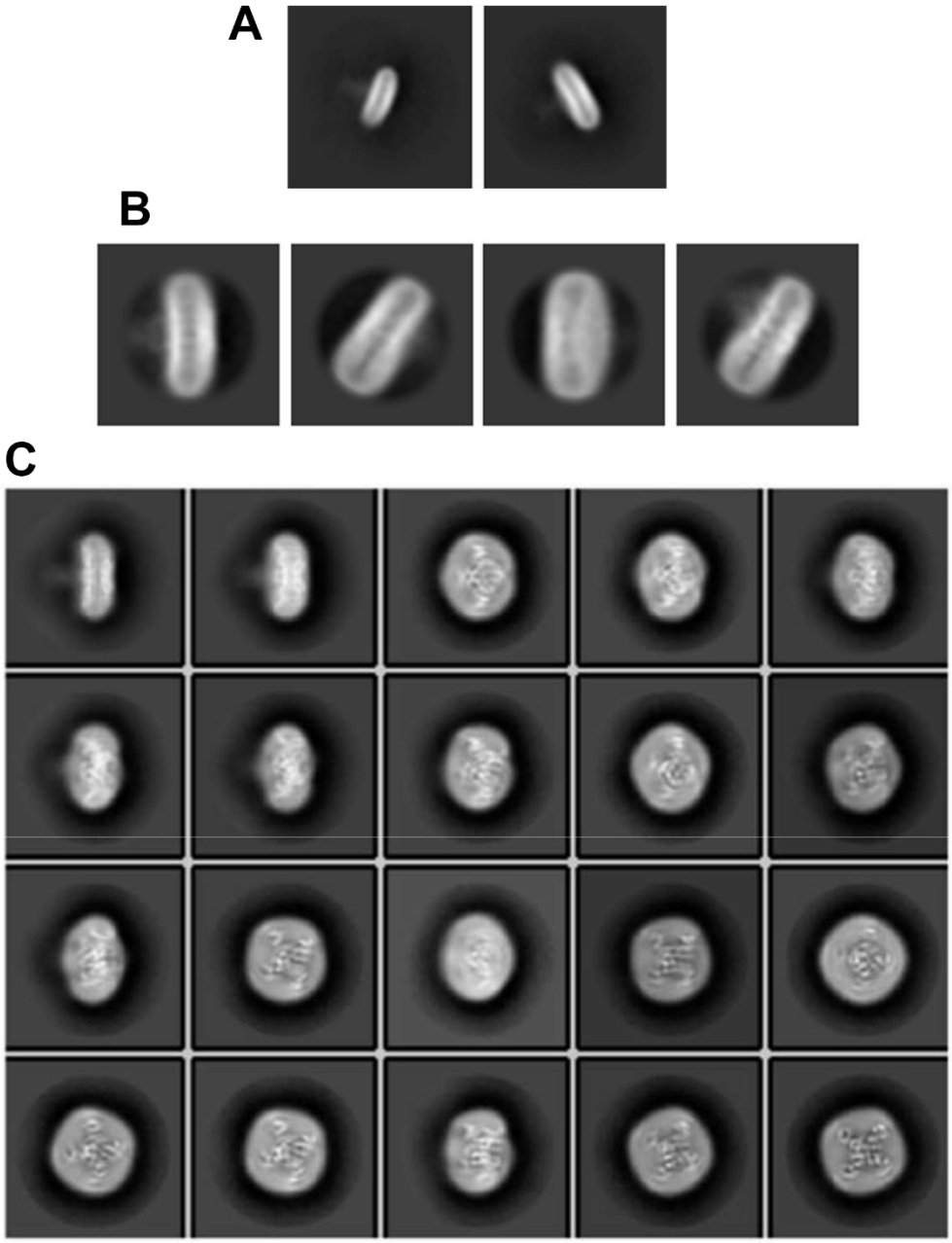

Figure 6.8. 2D classes obtained from Dataset 1 showing NavMs C-terminal domains and internal secondary structure. Figure showing (**A**) 2D classes obtained after blob picking, 2D classification, template picking and a second round of 2D classification, with classes showing transmembrane 'head' regions and faint C-terminal domain 'tails'. (**B**) shows 2D classes obtained using the classes from (**A**) as templates for further picking before another round of 2D classification, again showing 'head' regions and C-terminal domains. (**C**) Representative 2D classes of NaChBac in lipid nanodiscs, adapted from Gao *et al.,* 2020, for comparison. For scale: (**B**) circular mask diameter is 140 Å, (**C**) circular mask diameter is 220 Å.

6.5. Optimisation to remove empty MSP nanodiscs

A method to remove excess empty nanodiscs is made possible by the fact that both the target protein and the MSP2N2 have an N-terminal poly-His tag for purification by Ni^{2+} affinity chromatography, but the MSP2N2 has a tobacco etch virus (TEV) protease cleavage site after the tag, whereas the recombinant NavMs used here has a thrombin cleavage site. This was utilised by adding TEV protease (expressed, purified, and provided by Dr David Hollingworth of the Wallace group) to the reconstitution mixture prior to incubation, removing the His-tag from the MSP prior to an added Ni^{2+} affinity chromatography step; if the His-tag has been successfully removed from the MSP, then only nanodiscs containing NavMs (i.e., with an uncleaved His-tag) will bind to the column (please refer to Methods section 4.7 for full details). After this step, the sample is concentrated, and SEC is performed as before. This reconstitution was performed with double the concentration of each nanodisc constituent (NavMs, MSP2N2, and POPC) versus earlier attempts, to safeguard against loss of target protein during the extra purification step.

As can be seen in Figure 6.9, the resultant UV absorbance trace is markedly different from the UV trace of uncleaved MSP. The target peak is reduced in amplitude compared to the uncleaved sample, peaking at approximately 13 mAU versus 34 mAU. Alongside this, the empty nanodisc peak is almost entirely removed, suggesting that the added purification step was a success. While the target peak is reduced in comparison to the uncleaved sample, comparing each by SDS-PAGE (see Figure 6.9) shows that the band representing NavMs is a similar intensity in each, whilst the cleaved MSP band is noticeably reduced in intensity (and molecular weight) compared to the uncleaved MSP. This indicates that the amplitude of the target peak is reduced due to removal of empty nanodiscs rather than deleterious loss of NavMs-nanodisc complexes. Bands representing possible impurities can be seen at 90-100 kDa in the 15 mL uncleaved fraction, which, as they are also present in the MSP2N stock lane, seem likely to be from the nanodisc stock rather than caused by contamination during the purification of NavMs. The peak fractions were concentrated to approximately 1.5 mg/mL as before (although knowing that the previous sample contained many empty nanodiscs suggests that the true concentration of NavMs-nanodiscs complexes was previously lower than this) and cryo-TEM grids were produced with the same conditions as Dataset 1 (100% humidity, 4 °C, 4 s blot, 0 wait

time, -10 blotting force, 3.5 µL sample). Data was then collected on a Titan Krios microscope, producing Dataset 2.

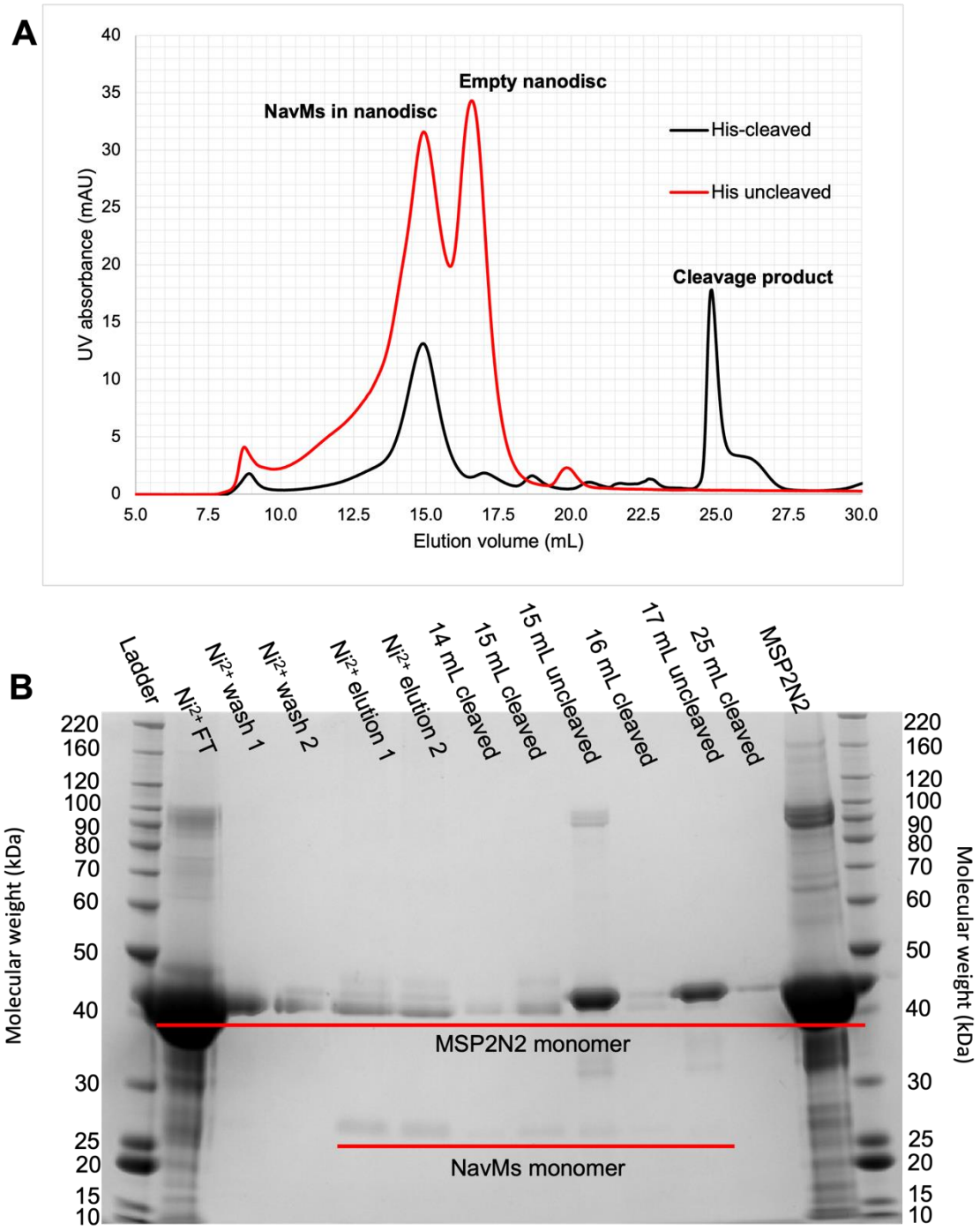

Figure 6.9. UV absorbance traces and SDS-PAGE gel comparing purification of nanodiscs with and without removal of empty nanodiscs. (**A**) UV absorbance traces of final size-exclusion chromatography (SEC) steps of nanodisc reconstitution and purification. The poly-His tag of the 'His-cleaved' sample was cleaved by tobacco etch virus and Ni^{2+} affinity chromatography performed prior to SEC, with the aim of only nanodiscs containing NavMs being retained via the N-terminal His-tag of NavMs. (**B**) shows an SDS-PAGE gel of Ni^{2+} affinity purification of cleaved sample and SEC steps of both cleaved and uncleaved samples.

6.6. Data collection and processing (Dataset 2)

6.6.1. Data collection

A total of 20,207 frames were collected using 3 shots per hole, with collection in super resolution mode and recorded binned by 2 to match a physical pixel size of 0.65 Å (for full details of data collection see Methods section 4.10.1). Movies were imported into cryoSPARC, and all subsequent data processing performed in cryoSPARC version v3.3.1.

6.6.2. Data processing and 3.07 Å reconstruction

Unless otherwise specified, jobs were built using the default settings in cryoSPARC. A graphical depiction of the processing pipeline performed on Dataset 2 resulting in a 3.07 Å global resolution reconstruction is shown in Figures 6.10 and 6.11.

Pre-processing: After importation, Patch motion correction (multi) was performed on all 20,207 frames and patch contrast transfer function (CTF) estimation (multi) performed on motion-corrected micrographs. The manually curate exposures job was used to exclude micrographs with poor characteristics such as high levels of motion or low resolution CTF fit resolution. In this case, micrographs were excluded with a CTF fit resolution of > 8 Å and total full-frame motion distance of > 20 pixels, leading to 18,663 micrographs being accepted.

Particle picking and 2D classification: The blob picker job was then used with a minimum and maximum particle diameter of 100 Å and 250 Å respectively, which picked 8,310,254 particles. These picks were manually inspected and curated as with Dataset 1, bringing the particle selection down to 3,626,392. Particle locations were extracted with a box size of 520 pixels, and a 50 class 2D classification job performed with force max over poses/shifts turned off, number of online-EM iterations set to 40, batchsize per class set to 200, and a circular mask diameter of 220 Å. Two 2D classes (representing 232,691 particles out of 2,891,300 classified) appearing to show a side view of a NavMs-nanodisc complex with CTD and a top/bottom view of a nanodisc with possible internal structural detail (see Figure 6.12) were selected and used as templates for the template picker job with a particle diameter of 180 Å, resulting in 10,966,186 particle picks. Curation of these picks reduced the particles down to 2,402,952, and these particle coordinates were extracted with a reduced extraction box size of 320 pixels Fourier cropped to 160 pixels to reduce the computational load, as high spatial frequency CTF data is not as important

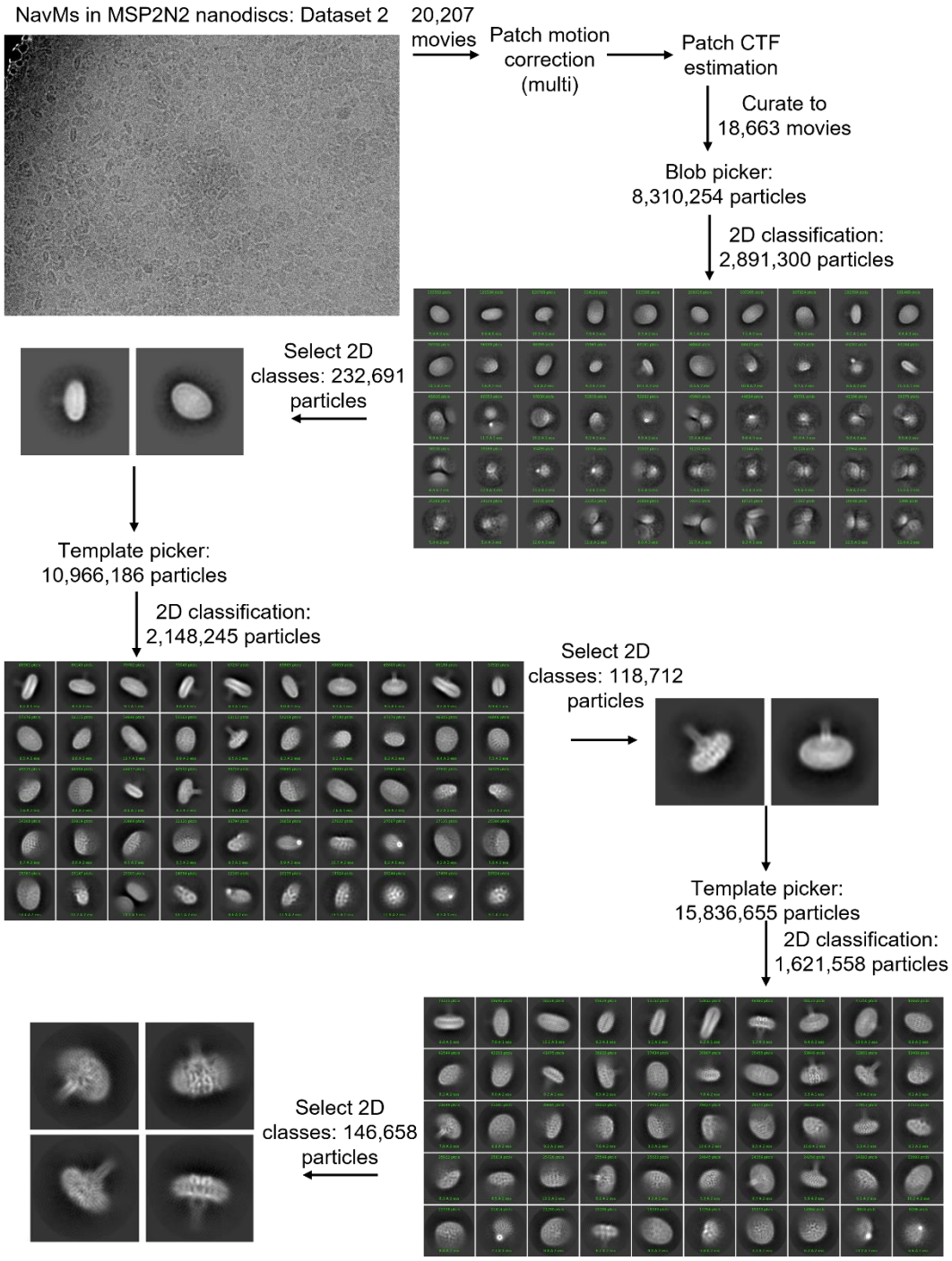

Figure 6.10. Cryo-TEM data processing pipeline of Dataset 2 of NavMs in MSP2N2 nanodiscs, part 1. Figure showing the general pipeline of the processing of Dataset 2 from pre-processing to obtaining good 2D classes. All processing performed in cryoSPARC version v3.3.1. 20,207 movies were imported into cryoSPARC for pre-processing by patch motion correction and patch contrast transfer function estimation. Resultant micrographs were curated subjected to blob picking and 2D classification. Classes resembling top and side views of NavMs in nanodiscs were identified and used for two rounds of template picking and 2D classification. The 4 highest resolution classes with internal details and C-terminal domains representing 146,658 particles were chosen for further processing.

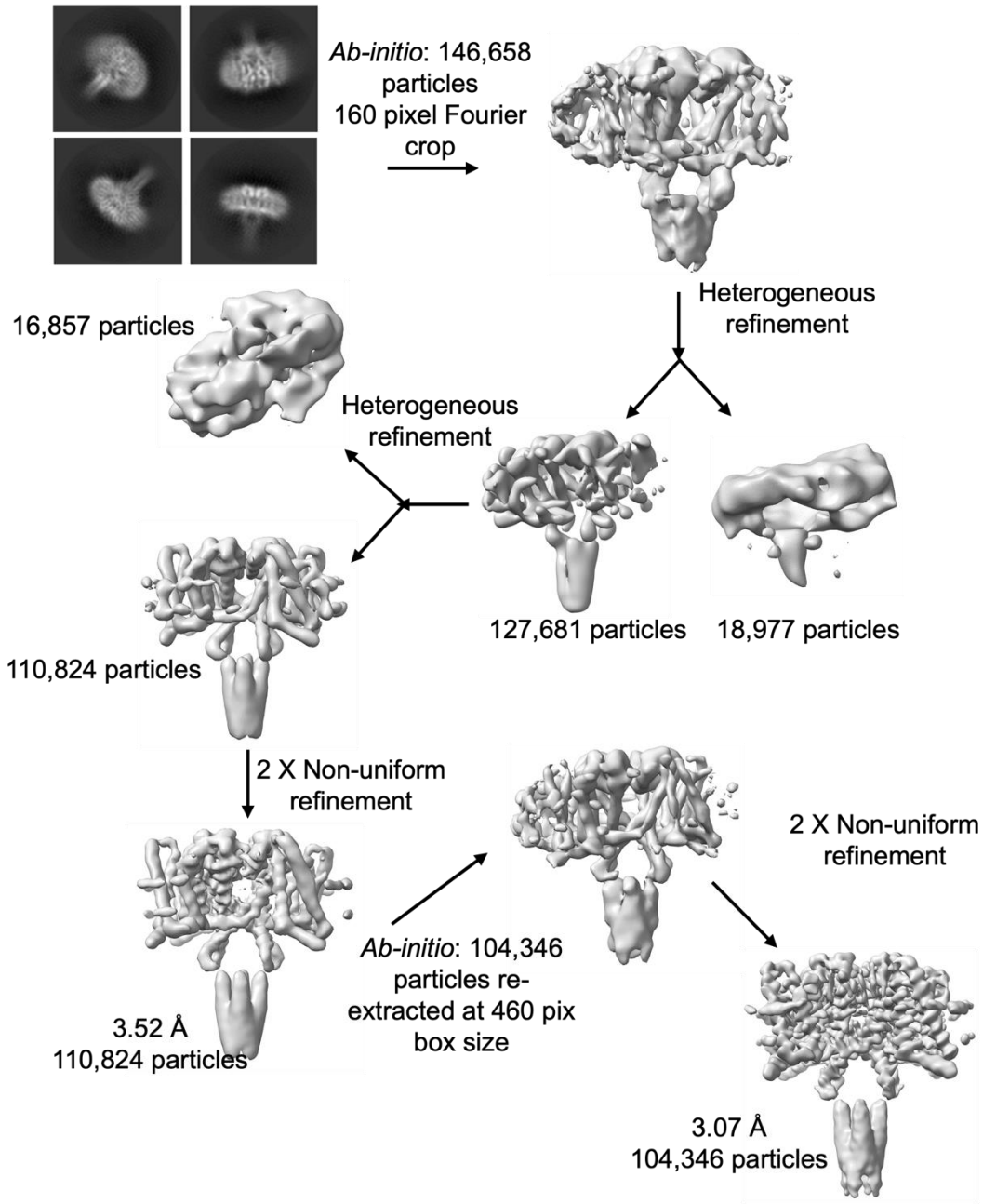

Figure 6.11. Cryo-TEM data processing pipeline of Dataset 2 of NavMs in MSP2N2 nanodiscs, part 2. Figure showing the general pipeline of the processing of Dataset 2 from *ab-initio* reconstruction to a 3.07 Å Coulomb potential map. All processing performed in cryoSPARC version v3.3.1. The 4 highest-resolution classes showing clear secondary structure were used to produce an *ab-initio* reconstruction. This was put through 2 rounds of heterogeneous refinement with a junk decoy to remove bad particles, and 2 rounds of non-uniform refinement gave a 3.52 Å volume. Particles were re-extracted with a 460-pixel box size to account for signal displacement caused by the contrast transfer function and a new *ab-initio* reconstruction produced. 2 rounds of non-uniform refinement pushed this to 3.07 Å resolution with 104,346 particles, at which point further refinement reduced the resolution. Map figures made in ChimeraX (Pettersen *et al*., 2021).

at this stage of processing. 2D classification was performed with an initial classification uncertainty factor of 1 (a value of 1 leads the algorithm to quickly decide on classes and assignments, leading to greater identification of 'junk', helpful in early processing), a circular mask diameter of 200 Å, force max over poses/shifts off, online-EM iterations set to 40 and a batchsize per class of 400. This produced 2D classes which were more clearly identifiable as NavMs-nanodisc complexes than the previous 2D classification step, appearing at this stage by looking at the 'interaction domain' or 'motif' where helix S3, the S4-S5 domain linker, and the CTD interact with each other (Sula and Wallace, 2017; Sula *et al.*, 2017), to resemble an 'open' conformation (see Figure 6.12), alongside junk particles. The 2D classes appear to favour side views at this stage. Two of these (representing 118,712 particles out of 2,148,245 classified), showing a side and partial bottom view with internal structural detail and clear CTDs, were selected and used for another round of template picking with a particle diameter of 170 Å to try to improve the particle selection. 15,836,655 particles were picked, which was curated down to 1,753,290 picks and extracted as previous. This resultant 1,621,558 particles were subject to 2D classification with the same parameters as previous, again producing 2D classes with internal structural detail and visible CTDs, although again favouring side views. The 4 highest resolution class averages which showed internal structural detail and CTD, representing 146,658 particles, were used to build a reference-free *ab-initio* model (Punjani *et al.*, 2017).

***Ab-initio* model and refinement:** The *ab-initio* reconstruction job was performed with 1 *ab-initio* class, maximum and initial resolutions set to 5 and 12 Å respectively (recommended for small membrane proteins), and no symmetry applied, the result of which can be seen in Figure 6.13. As expected from the fact that the 2D classes favoured side views, the angular viewing distribution heatmap provided by cryoSPARC does not show a continuous spread of views, however, comparing the reconstruction with a crystal structure of NavMs (PDB 6SX5) (Sula *et al.*, 2021) it still appears to be a promising reconstruction. Viewing the map in ChimeraX (Pettersen *et al.*, 2021), this reconstruction shows the expected 'mushroom' shape of NavMs in a nanodisc, with pore domain (PD) helices already visible at this stage and a resolved CTD.

Refinement: Heterogeneous refinement was performed to see if classes could be separated out from the *ab-initio* reconstruction, with force hard classification set to on.

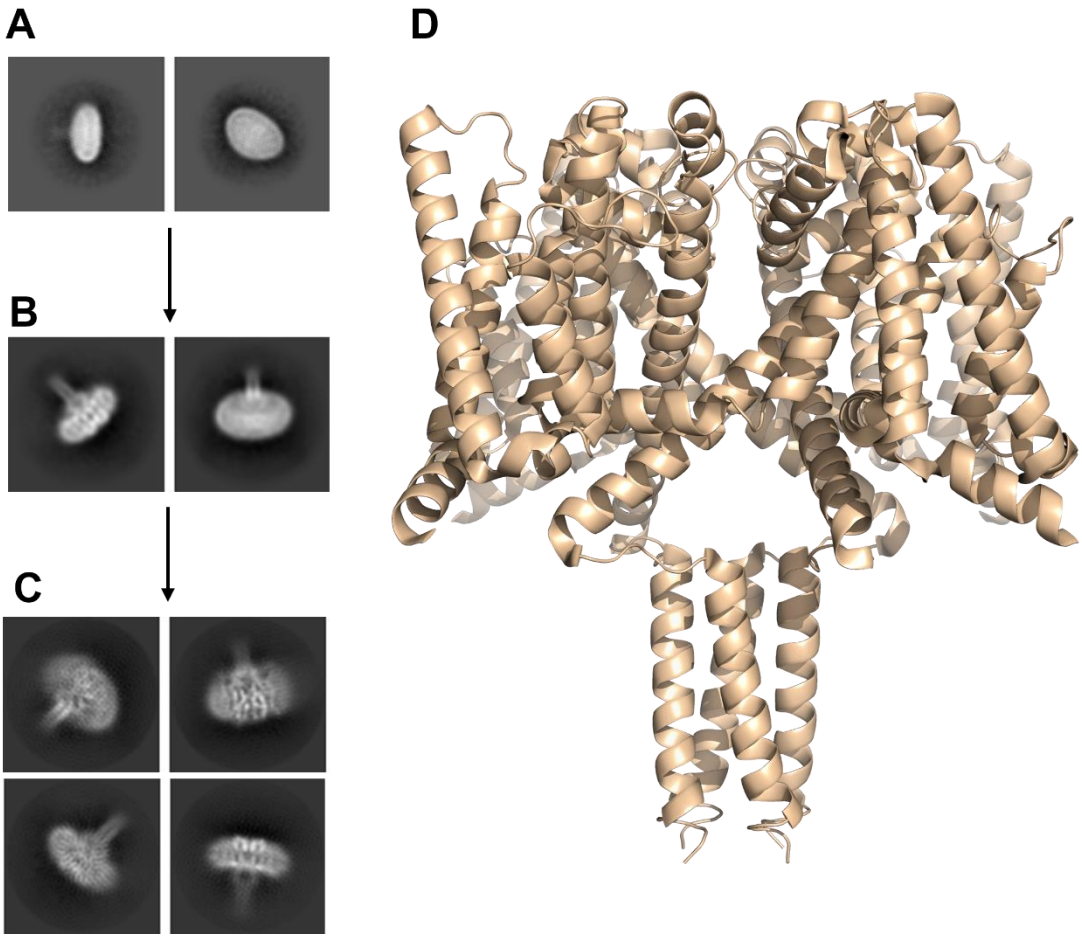

Figure 6.12. 2D classes of Dataset 2 with NavMs crystal structure for comparison. Classes in (**A**) were produced by 2D classification of blob picking and were used as references for template picking. 2D classification of these picks produced (**B**), and the same process again produced the 2D classes shown in (**C**). C-terminal domains, and pore and voltage-sensor helices can be seen. (**D**) shows the 2.20 Å crystal structure of 'open' NavMs F208L (PDB 6SX5) published by Sula *et al.,* 2021 for comparison. Figure made using PyMol molecular graphics software (version 2.0; Schrödinger, LLC).

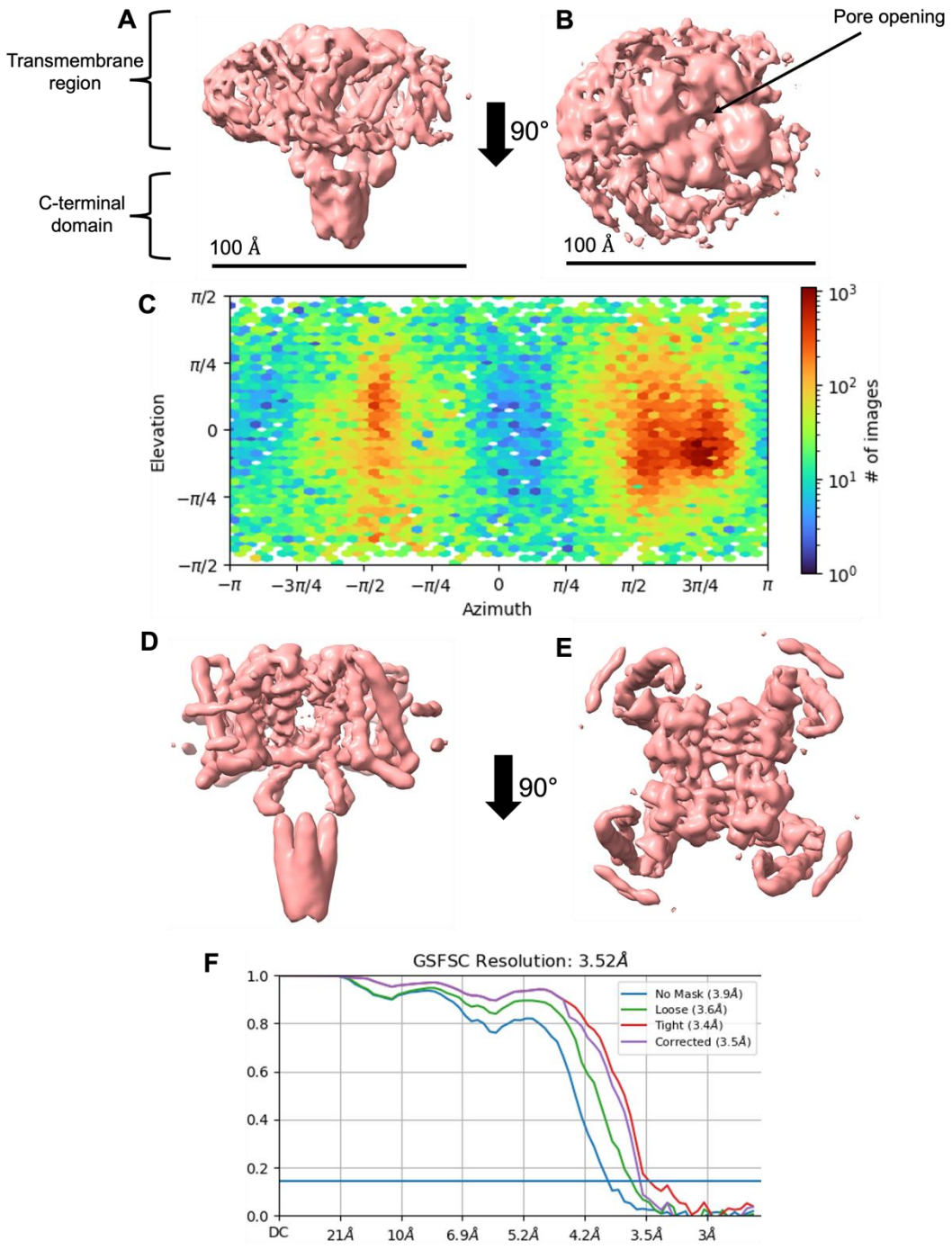

Figure 6.13. 3D reconstructions of Dataset 2 reaching 3.52 Å resolution. (**A**) and (**B**) show side and top (extracellular) views of the first *ab-initio* reconstruction. 100 Å scale bar shown. (**C**) shows a heatmap of viewing angles produced by cryoSPARC, with red indicating preferred views. (**D**) and (**E**) show equivalent side and top views of the 3.52 Å reconstruction achieved after 2 rounds of heterogeneous refinement and 2 rounds of non-uniform. (**F**) shows gold-standard Fourier shell correlation estimating global resolution at 3.52 Å. Volume figures made in ChimeraX (Pettersen *et al.*, 2021).

This causes cryoSPARC to assign each particle to only one class at each iteration instead of partially assigning particles to each class, which is recommended for improving results with low signal-to-noise ratio particles such as small membrane proteins. The previously described *ab-initio* volume and a second 'junk' *ab-initio* volume made from 2D classes without an obvious CTD were used as input volumes. This produced a 7.10 Å class which again resembled a NavMs-nanodisc complex, produced from 127,681 'good' particles, and a junk class of 18,977 particles. To further remove junk particles, the 7.10 Å volume was subjected to non-uniform refinement (developed specifically for refinement of membrane proteins (Punjani, Zhang and Fleet, 2020) with C4 symmetry imposed (NavMs being a homotetramer with C4 symmetry), and the resultant volume used alongside the junk class from the previous job as input volumes for a second round of heterogeneous refinement, resulting in a 16,857-particle junk class and a 110,824-particle NavMs-nanodisc class at 4.23 Å resolution. Non-uniform refinement of the NavMs-nanodisc class with C4 symmetry produced a 3.52 Å global resolution Coulomb potential map, with well-defined pore helices and identifiable VSDs, interaction domains and CTD helices (see Figure 6.13). The MSP2N2 dimer is also visible encircling the transmembrane domain, as can also be seen in the structure of NaChBac in nanodiscs published by Gao *et al*., 2020. As further refinement steps did not improve resolution, it was decided that these 110,824 particles should be re-extracted with a larger box size to catch higher spatial frequency CTF information and subject to refinement again.

Particles were re-extracted with a box size of 460 pixels, calculated using the particle diameter; 100 Å for NavMs itself, with a pixel size of 0.65 Å, multiplied by 3 to account for information spread. 104,346 particles were extracted, and an *ab-initio* reconstruction performed, this time with maximum resolution and initial resolution set to 7 and 9 Å respectively. The resultant reconstruction again showed visible internal transmembrane helices, interaction motif, and CTD, shown in Figure 6.14. A single step of non-uniform refinement with C4 symmetry applied produced a 3.22 Å reconstruction, and a second round with initial lowpass resolution set to 15 Å improved this further to a global resolution of 3.07 Å (see Figure 6.14). Further refinement steps worsened the resolution, indicating that further refinement was leading to overfitting and therefore the resolution would not be increased in this manner.

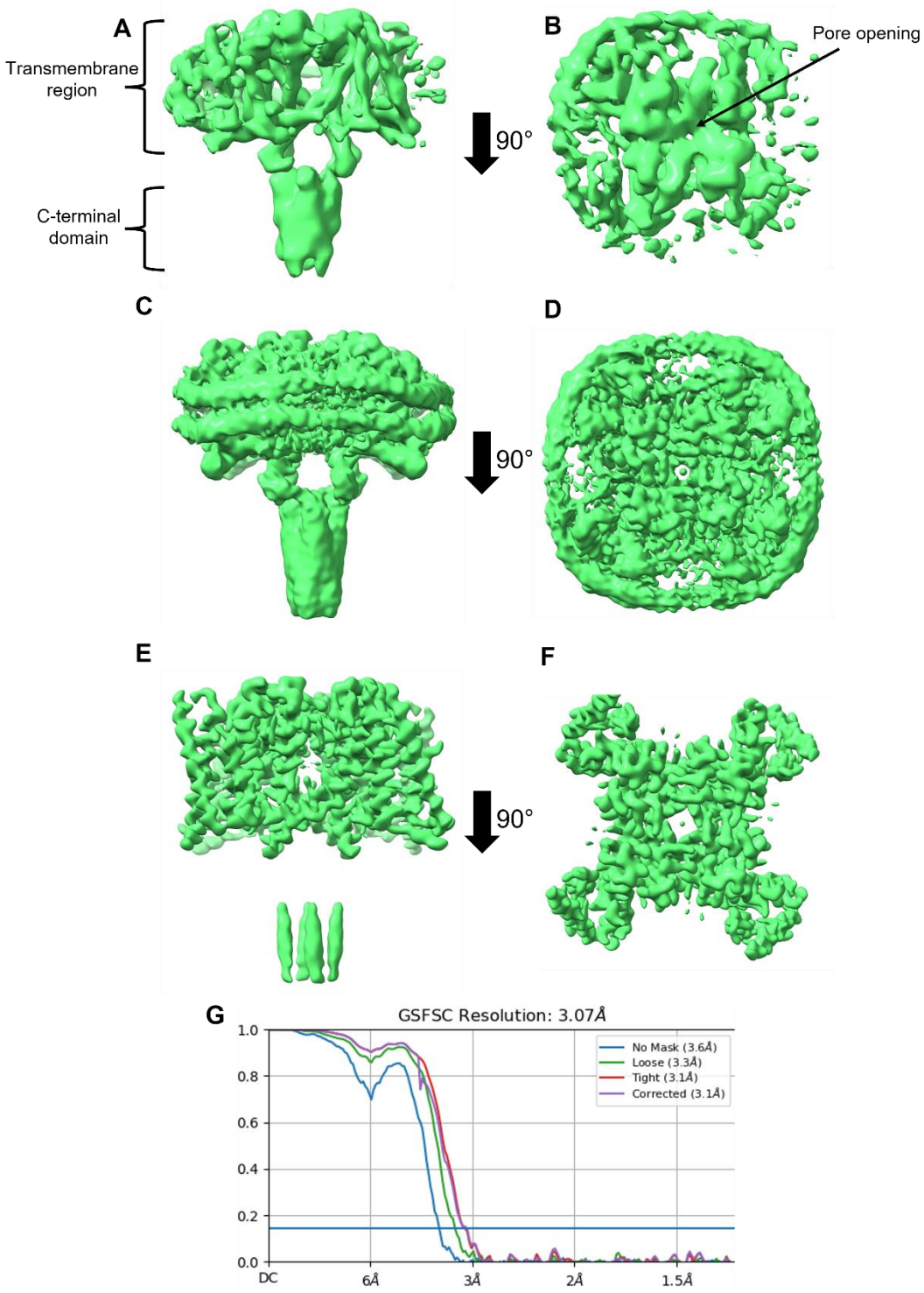

Figure 6.14. 3D reconstructions of Dataset 2 reaching 3.07 Å resolution. (**A**) and (**B**) show side and top (extracellular) views of the *ab-initio* reconstruction after re-extraction of particles with a 460-pixel box to capture higher spatial frequency information. (**C**) and (**D**) show the 3.07 Å volume after 2 rounds of non-uniform refinement at volume level 0.036 in ChimeraX, with (**E**) and (**F**) showing the same reconstruction at volume level 0.083. (**G**) shows gold-standard Fourier shell correlation estimating global resolution at 3.07 Å. Density volume figures made in ChimeraX (Pettersen *et al.*, 2021).

6.6.3. Attempts to improve resolution beyond 3.07 Å

Local refinement: Looking at the local resolution of the reconstruction (see Figure 6.15), it is clear that the transmembrane region is better resolved than the CTD, with pore helices reaching approximately 2.8 Å (identified by values given when mapping local resolution to the volume in ChimeraX), the VSD regions reaching 3 to 3.5 Å (4 Å at the membrane surface), and the CTD at approximately 5-5.5 Å resolution. As the CTD is flexible and able to move independently of the rest of the protein, a mask was made in ChimeraX covering only the transmembrane region and local refinement reformed with the CTD masked out to see if this would improve resolution. As seen in Figure 6.16, this had little effect, with the global resolution worsening from 3.07 Å to 3.35 Å. This could be due to the fact that NavMs is already a small target for cryo-TEM, with removing the CTD making the target even smaller and more difficult for cryoSPARC to align, or because with only the transmembrane region there is little to indicate which way 'up' the volume is during refinement.

3D classification: 3D classification was also performed on the 104,346 particles used to produce the 3.07 Å reconstruction to see if more than one conformation of the mobile CTD was present, as it is flexibly linked to the rest of the channel and mobile. This resulted in two classes split almost evenly, with 53,446 particles in one and 50,900 in the other; Looking at the two volumes in ChimeraX (see Figure 6.16), they appear essentially identical even in the CTD, suggesting that varying CTD conformation is not driving a reduction in resolution in this case.

Using the ResLog analysis job in cryoSPARC gives a plot showing the change in resolution of a structure as particles are added into the reconstruction (Stagg *et al.*, 2014); if the resolution is continuing to rise as more particles are added, this means that collecting more data and adding more particles into the reconstruction should continue to improve the resolution. Performing this job on the particles used for the 3.07 Å reconstruction (see Figure 6.16) shows that the spatial frequency in 1/Å is still steadily increasing up to the full number of particles, suggesting that collecting another dataset would continue to push the resolution higher. As more refinement did not appear to be capable of improving the reconstruction, further analysis was performed using the 3.07 Å Coulomb potential map.

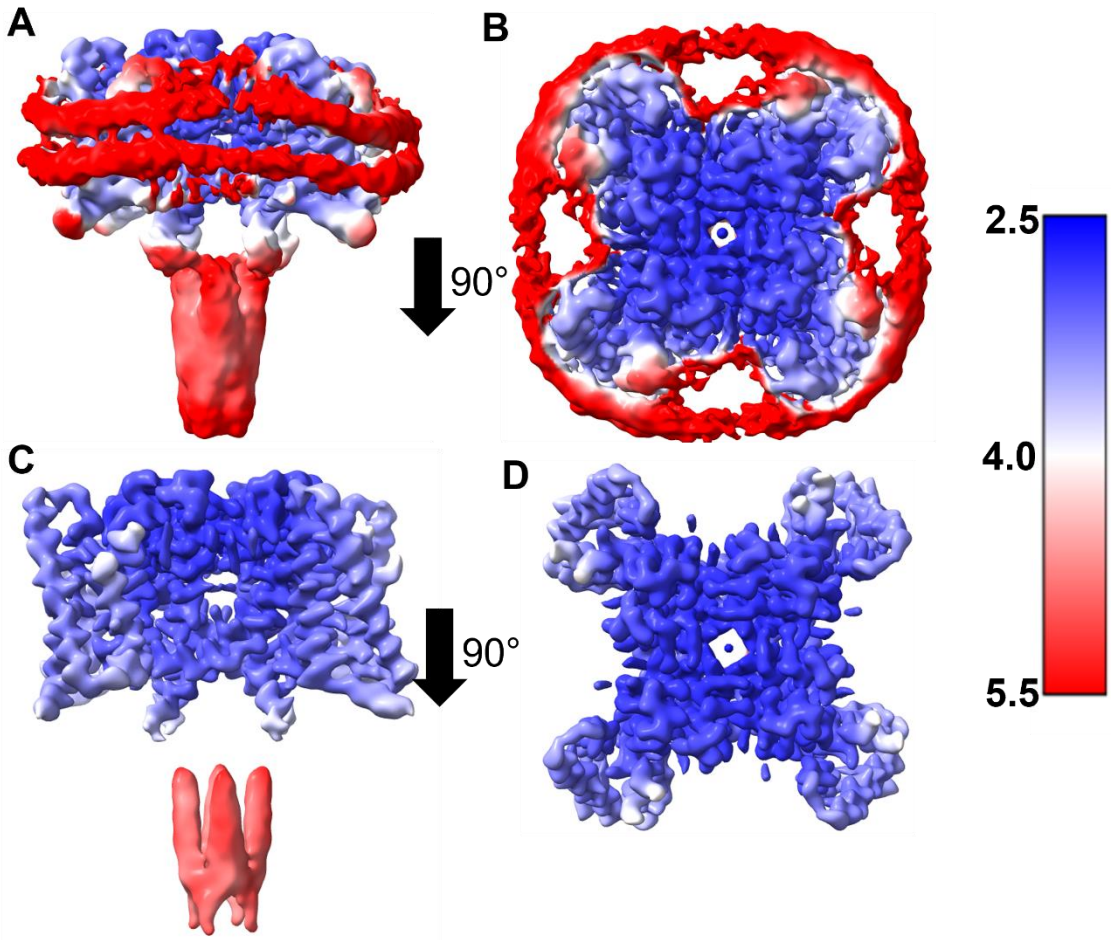

Figure 6.15. Local resolution maps of the NavMs-nanodisc complex at varying sigma levels. (**A**) and (**B**) show side and top views of the 3.07 Å NavMs-nanodisc reconstruction with local resolution mapped to the density volume (scale bar to right) at volume level 0.037. (**C**) and (**D**) show the same volumes and local resolution maps at a higher sigma value (0.07). Worse resolution in the C-terminal domain (CTD) of the reconstruction identified masking out of the mobile CTD followed by local refinement as a possible route to increasing the resolution of the transmembrane regions. Density volume figures made in ChimeraX (Pettersen *et al.*, 2021).

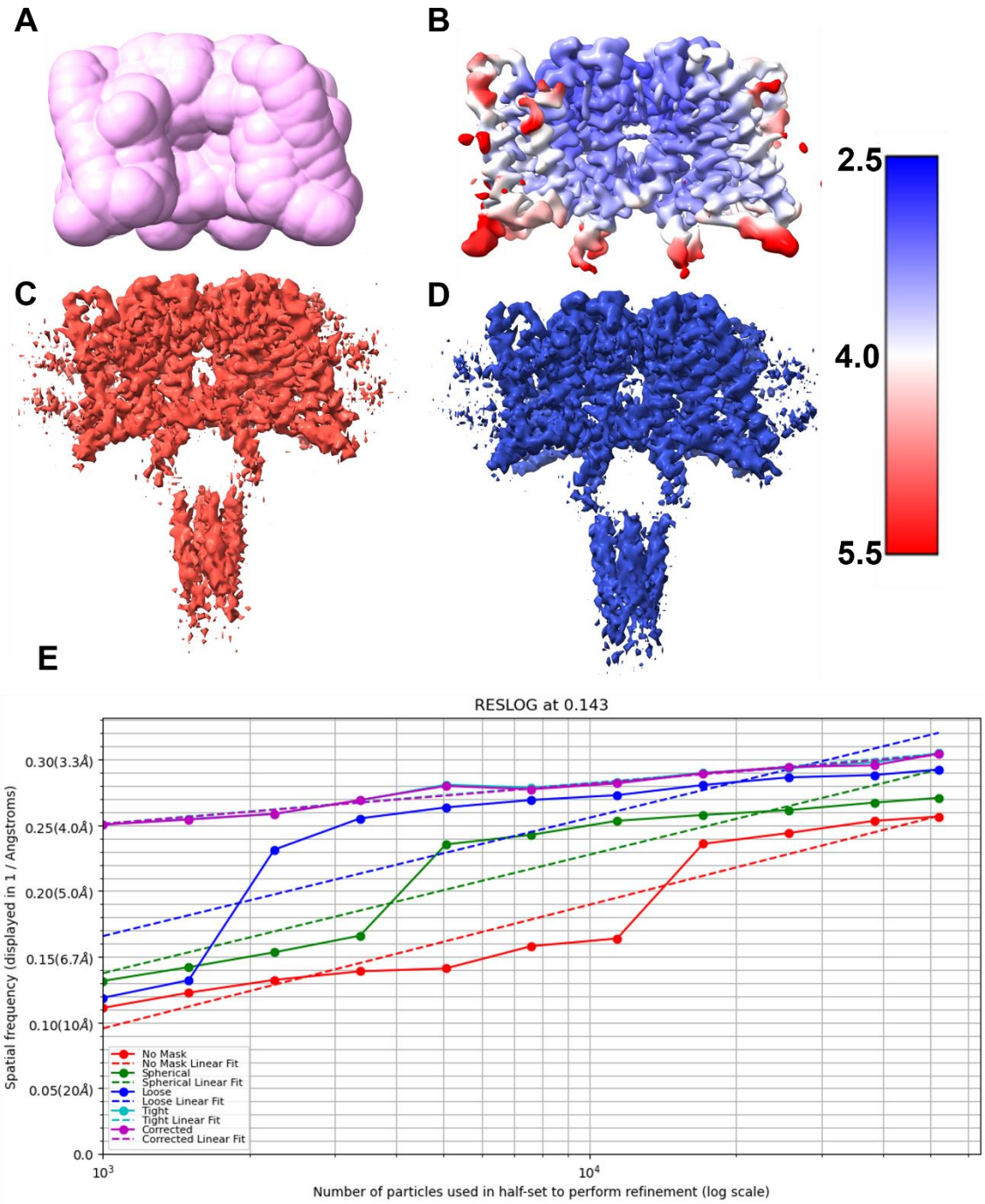

Figure 6.16. Attempts to improve the reconstruction by local refinement and 3D classification. (**A**) and (**B**) show the mask used to mask out the C-terminal domain and the resulting local refinement respectively, with local resolution shown (scale bar to right). (**C**) and (**D**) show the two volumes produced by 3D classification. The classes are almost equally populated (53,446 and 50,900 particles respectively). (**E**) shows ResLog analysis (Stagg *et al.,* 2014) of the effect on resolution (displayed as 1/Å) as particles are added to the reconstruction. Density volume figures made in ChimeraX (Pettersen *et al.*, 2021).

6.7. Discussion

A 3.07 Å resolution Coulomb potential map should be suitable for the confident assignment of backbone carbons, and side chains in areas of increased local resolution, suggesting that the aim of the project, to develop a method for the electron microscopy study of NavMs, has been achieved. As was seen in Figures 6.14 and 6.15, the final NavMs-nanodisc reconstruction shows defined helices throughout the transmembrane domain, and density within which side chains could likely be placed, especially within the pore region where local resolution surpasses 3 Å, whilst the four helices of the CTD coiled-coil bundle are visible although not well-defined enough to visualise side chain positions, as would be expected with a local resolution of 5 Å. Finally, the linkers between the pore and CTD are also less well resolved, as would be expected for such a mobile region (these linkers are unstructured in published NavMs models), with a similar resolution to the MSP2N2 protein itself. Unfortunately, due to time constraints, it was not possible to use macromolecular modelling software to build the NavMs sequence into the map produced in cryoSPARC. However, comparisons can be made between previously published NavMs structures and the NavMs-nanodiscs map to see whether any novel features can be discerned.

6.7.1. The conformation of the NavMs-nanodiscs complex differs from previous apo NavMs structures at the interaction motif

The conformation of a Nav structure (i.e., whether is it in its resting, activated, or inactivated state) can be judged by the conformation of specific regions of the protein (Catterall, Wisedchaisri and Zheng, 2020). It is considered open or activated based on the position of the VSD helix S4 relative to the rest of the VSD, and if the intracellular gate at the intracellular end of the pore is open wide enough for sodium ion conductance (Sula *et al.*, 2017, 2021; Sait *et al.*, 2020). To identify the conformation of the NavMs-nanodisc complex, the 2.20 Å model of apo NavMs F208L published in Sula *et al.,* 2021 (PDB 6SX5) was fitted into the NavMs-nanodisc Coulomb potential map in ChimeraX, shown

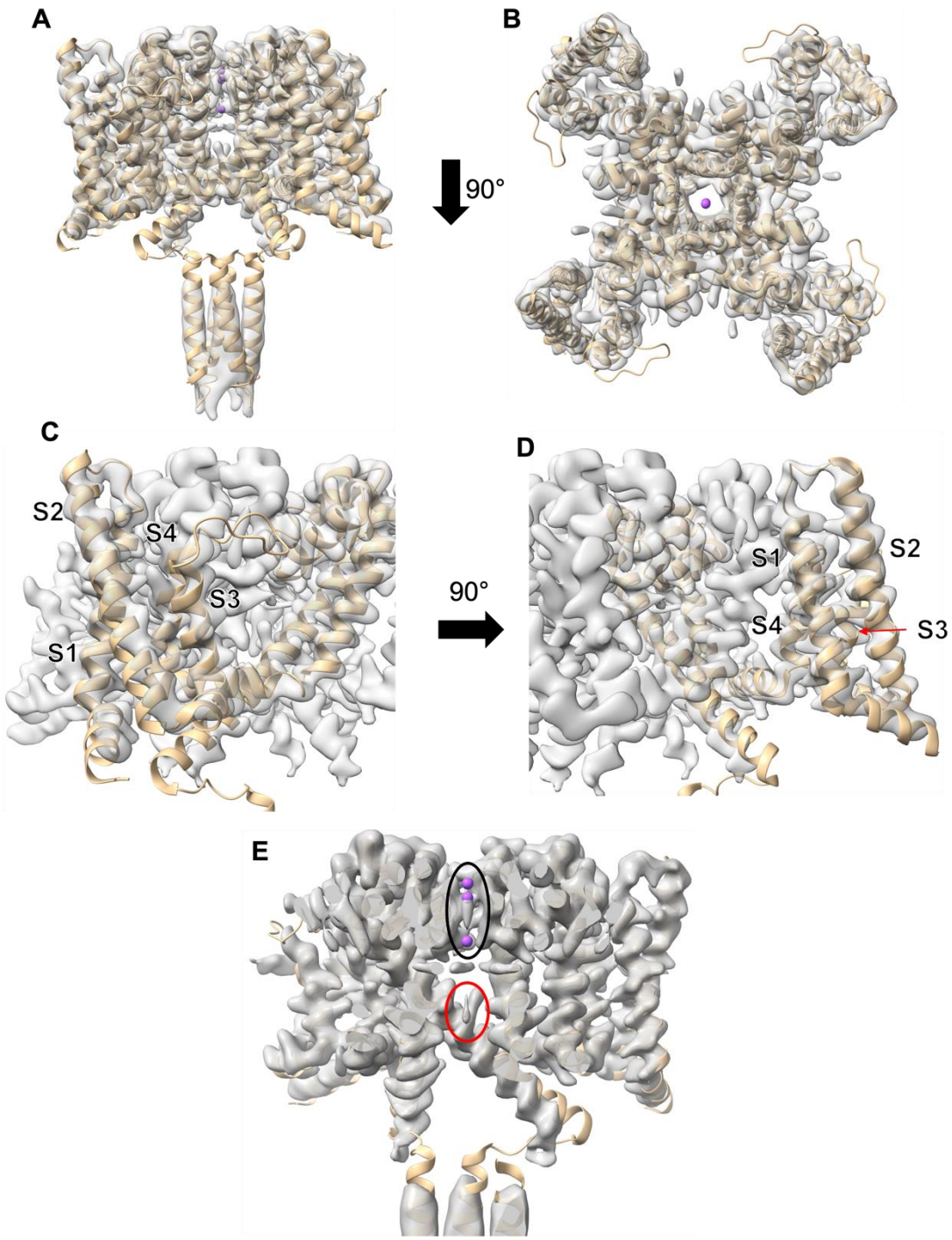

Figure 6.17. Comparison of 'open' NavMs F208L and the NavMs-nanodisc reconstruction. (**A**) and (**B**) show top and side views of 'open' NavMs F208L (PDB 6SX5) fitted into the NavMs-nanodisc Coulomb potential map. (**C**) and (**D**) show two views of one of the voltage-sensor domains (other monomers hidden for clarity). (**E**) shows density in the pore of the reconstruction, with Na^{2+} ions modelled in 6SX5 shown as purple spheres and highlighted in black. Additional pore density in the reconstruction is highlighted in red. Density volume figures made in ChimeraX (Pettersen *et al.*, 2021).

in Figure 6.17. The VSD helices of the NavMs F208L model are in broadly similar positions to equivalent density in the NavMs-nanodisc volume, suggesting that the NavMs-nanodisc density may also have its VSDs in the activated conformation; this

would be expected given the conditions of the sample, as there is no resting membrane potential applied. There is also matching density in the region where Na^+ ions are modelled in the NavMs F208L structure. Additional density not modelled in crystal structures of NavMs is visible in the intracellular gate region of the pore lumen, which has been speculated to represent glycol-diosgenin (GDN) detergent in cryo-TEM reconstructions of hNavs (G. Huang *et al*., 2022) in inactivated conformations.

Where the models also appear to differ is at the intracellular ends of the PD S6 helices, part of the aforementioned interaction motif where helix S3, the S4-S5 domain linker, and the CTD interact with each other. The NavMs-nanodisc density appears to show helix S6 tilted further away from the transmembrane region. In particular, the E229 residue of S6, which forms a salt bridge with R119 of the S4-S5 linker in the open channel, does not appear to be close to the electron density of the NavMs-nanodisc reconstruction, shown in Figure 6.18. This reduced interaction between S6 and the S4-S5 linker could suggest that the intracellular gate is forming a tighter constriction in the NavMs-nanodisc reconstruction; when compared to the model of 4-hydroxytamoxifen-bound NavMs (PDB 6SXG), held in an inactivated state by interaction with the ligand, the NavMs-nanodisc reconstruction density appears to fit the position of S6 more closely (see Figure 6.18), however the sequence should be modelled into the density properly before any conclusions are made. Additional density is also visible in each of the fenestrations of the NavMs-nanodisc reconstruction near the pore lumen; this position in NavMs F208L is filled by HEGA-10 detergent, which should have been removed from the sample by incubation with Bio-Beads during nanodisc reconstitution and from two steps of purification with no detergent present; the density could perhaps be indicative of POPC lipid tails entering the fenestration, however this is purely speculative.

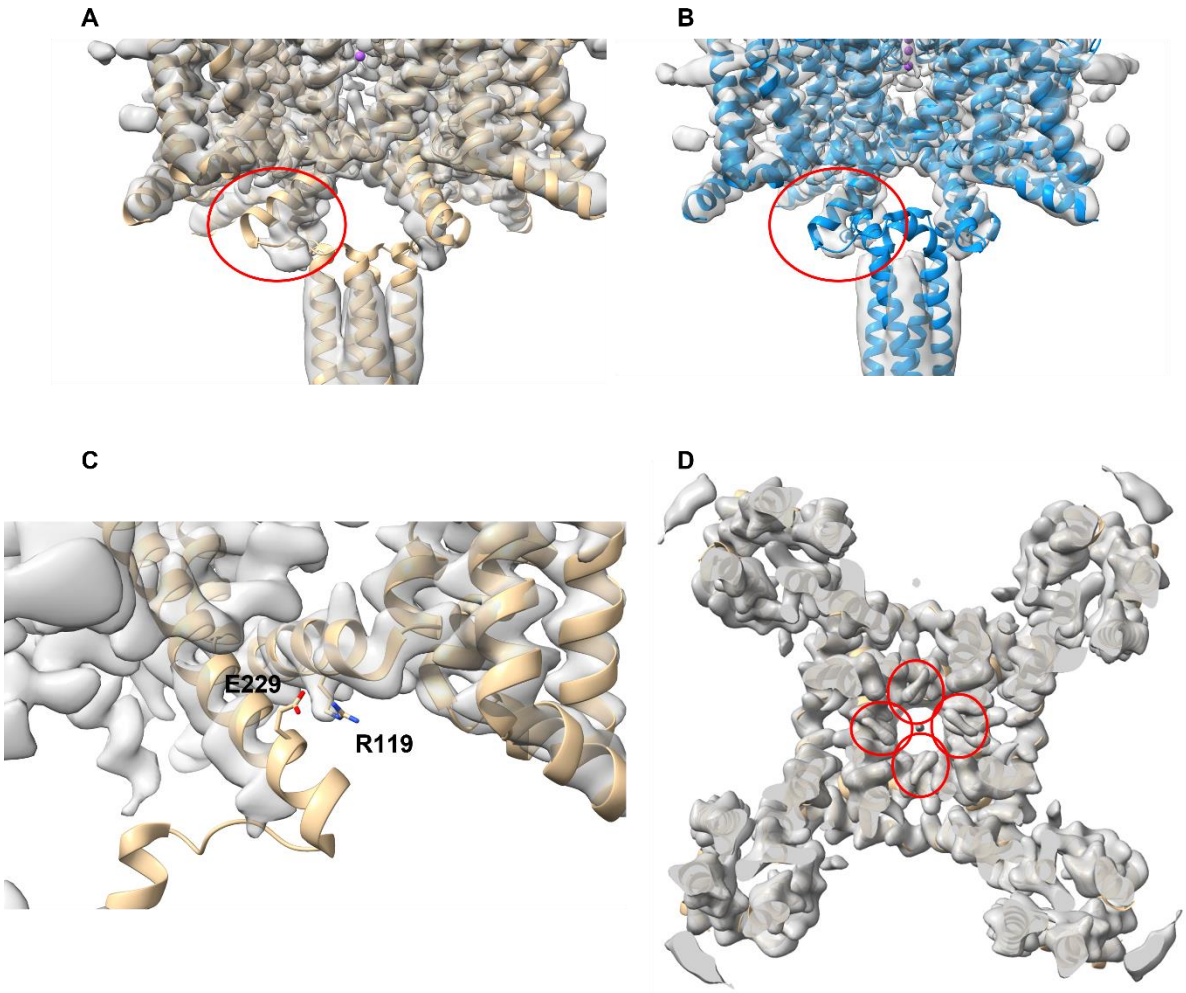

Figure 6.18. Differences between 'open' NavMs F208L and the NavMs-nanodisc reconstruction. (**A**) and (**B**) show views of 'open' NavMs F208L (PDB 6SX5) and 'inactivated' NavMs bound to 4-hydroxytamoxifen (PDB 6SXG) respectively (Sula *et al.,* 2021), fitted into the NavMs-nanodisc Coulomb potential map. The intracellular end of the S6 helix of 6SX5, which forms interactions with S3 and the S4-S5 linker to stabilise the open channel, is highlighted in red. (**C**) shows R119 of the S4-S5 linker and E229 of S6 of model 6SX5, residues in the interaction motif which form a salt bridge to help stabilise the open channel. (**D**) shows density in the fenestrations of the NavMs-nanodisc complex reconstruction using a sliced view to hide density between the viewer and the fenestrations. Density volume figures made in ChimeraX (Pettersen *et al.*, 2021).

6.8. Addendum: Atomic model of NavMs in nanodiscs

After original submission of this thesis, it was possible to use the previously refined density map to produce an atomic model of residues 1-266 (per monomer, out of 274) of NavMs in nanodiscs with 8 associated POPC lipid molecules and a sodium ion ligand bound by Glu178 of the selectivity filter (see Figures 6.19 and 6.20). The density map was sharpened using the autosharpen function in Phenix (Liebschner *et al.*, 2019), and a tetrameric model of NavMs, produced using Alphafold multimer (Evans *et al.*, 2022), was docked into the resulting map in ChimeraX (Pettersen *et al.*, 2021) before refinement using Phenix real-space refine. Errors in residue placement which were not fixed by Phenix real-space refinement were fixed manually in ISOLDE (Croll, 2018) before addition of ligands, including a single sodium ion bound in the selectivity filter and two POPC lipid molecules per monomer; these were placed and refined in Coot (Emsley *et al.*, 2010) using restraints generated by Phenix eLBOW. Representative regions of the model in the sharpened density map can be seen in Figure 6.20, and model refinement and validation statistics can be seen in Appendix 1.

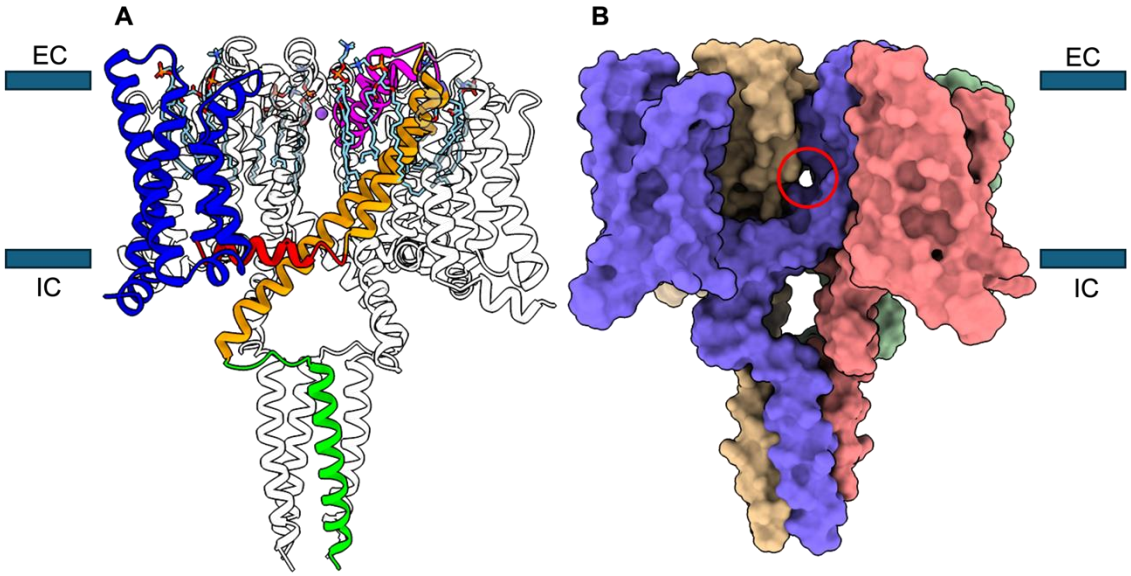

Figure 6.19. Atomic model of NavMs in nanodiscs showing NavMs structural features. (A) shows highlighted domains of the NavMs-nanodisc model. Blue = voltage sensor module (helices S1-S4), Red = voltage sensor/pore linker, Orange = pore module (S5-6), Magenta = selectivity filter, Green = C-terminal domain. Sodium ion bound in selectivity filter shown as a purple sphere. POPC ligands shown as cyan sticks. Chains 2-4 are transparent for clarity. **(B)** shows a surface model of NavMs-nanodiscs with monomers in different colours and fenestration highlighted in a red circle. Extracellular (EC) and intracellular (IC) boundaries shown by dark blue rectangles. Figure made in ChimeraX (Pettersen *et al.*, 2021).

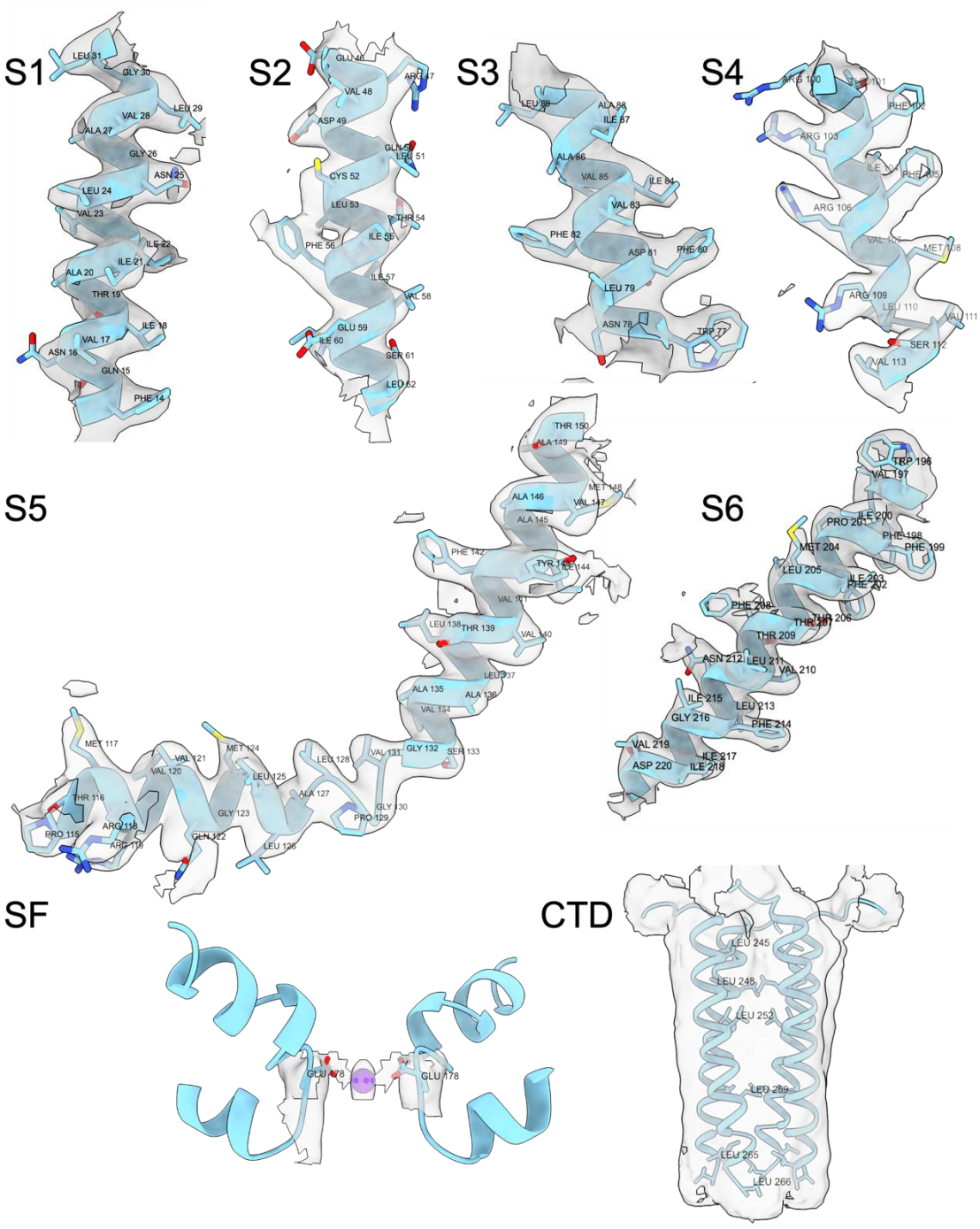

Figure 6.20. Cryo-electron microscopy map and model of representative regions of NavMs in nanodiscs. Helices S1-S6 are shown as well as a side view of the selectivity filter (SF) and the C-terminal domain (CTD). S1-S6 helices and SF are contoured to level 10, and to level 6 in the CTD, where local resolution is lower ChimeraX (Pettersen *et al.,* 2021). Only 2 monomers are shown of the SF for clarity. Purple sphere indicates a sodium ion. Leucine residues are labelled in the CTD figure to show appropriate positioning of helices for a coiled-coil motif. Figure made in ChimeraX (Pettersen *et al.*, 2021).

6.8.1. The NavMs-nanodsics model represents an inactivated channel

Previously obtained apo-state crystal structures of NavMs (Sula *et al.*, 2017, 2021; Sait *et al.*, 2020) show NavMs in an open, activated state, with a channel radius wide enough for hydrated sodium ions to conduct through (>2.3 Å) and S4 of the voltage sensor domain in an 'up', activated position. The NavMs-nanodisc model shows that whilst the voltage sensor of NavMs-nanodiscs is in the 'up' position, with S4 in a gating residues R1-3 above the occluding Phe56 residue and R4 below (see Figure 6.21**A**), the hydrophobic constriction of the lumen (formed by Ile215 at its narrowest radius) of S6 of each monomer is too narrow for conduction, reducing to a radius of 1.8 Å at Ile215 (see Figure 6.21**B** and **C**). This is consistent with an inactivated channel, as seen in the structure of NaChBac in nanodiscs (Gao *et al.*, 2020) and of the crystal structure of NavMs bound to 4-hydroxytamoxifen (Sula *et al.*, 2021).

6.8.2. Comparison of NavMs-nanodiscs and previous NavMs structures

As can be seen by rendering the NavMs-nanodisc model by its root mean square deviation (RMSD) against both open, apo state (for example, PDB 6SX5) and inactivated, 4-hydroxytamoxifen bound (PDB 6SXG) crystal structure conformations, helices S1 and S2 of the VSD, the S4-S5 linker, S5 of the pore domain, and the selectivity filter are near-identical in each model (see Figure 6.22). Interestingly however, the NavMs-nanodisc model shows features distinct from each; firstly, in the VSD, helix S3 is tilted slightly away from S2 when compared to the crystal structures (see Figures 6.22 and 6.23), with the interaction of Glu81 of S3 and R4 seen in crystal structure of NavMs missing (see Figure 6.21). A more drastic difference is that S4 of NavMs-nanodiscs is more structured, with two added helical turns of 3-10 helix at the N-terminal end of S4, putting the first gating charge R100 in a structured region of the helix rather than an S3-S4 loop, as seen in apo and ligand-bound NavMs (see Figures 6.22 and 6.23). As this extended S4 helix is also seen in cryo-TEM structures of eukaryotic Navs and crystal and cryo-TEM structures of prokaryotic Navs other than NavMs, it could be that this helix is disrupted by crystal contacts in prior NavMs structures. When quantitatively comparing these structures using the 2StrucCompare server and DSSP criteria (Kabsch and Sander, 1983; Drew and Janes, 2019), NavMs-nanodiscs has a total helical content of 81.6%, versus 76.5% for apo NavMs and 79.1% for 4-hydroxytamoxifen bound NavMs..

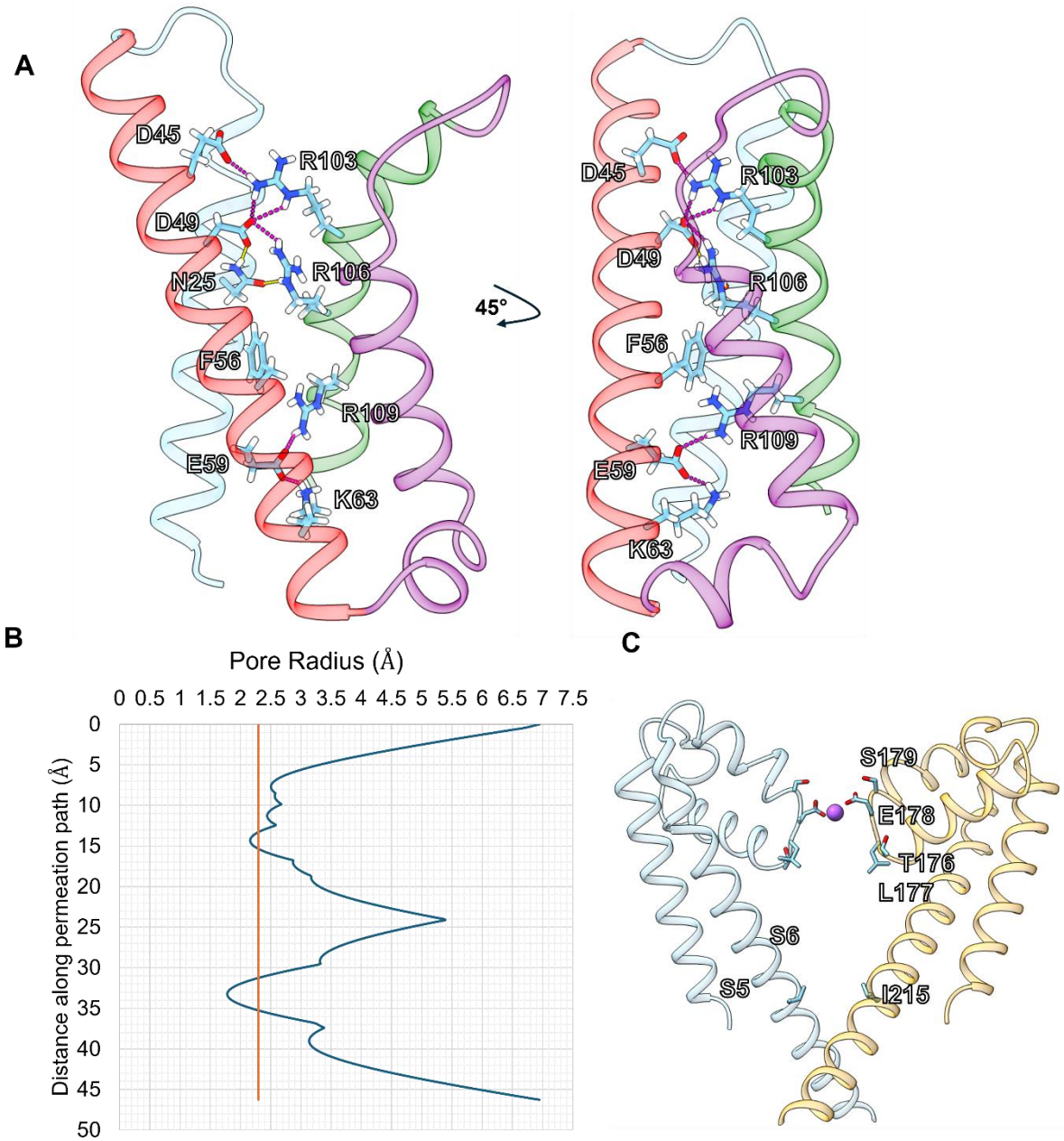

Figure 6.21. Features of the NavMs-nanodisc model indicating an inactivated channel. (A) shows the voltage sensor domain (VSD) of the NavMs-nanodisc model, with hydrogen bond (yellow) and salt bridge (magenta) interactions between arginine gating charges and their interaction partners shown. The occluding Phe56 residue is shown for comparison with arginine positions. Blue ribbon = helix S1, red = S2, purple = S3, green = S4. Pink lines indicate salt bridges, yellow indicate hydrogen bonds. **(B)** shows the plot of distance along the ion permeation path of the model versus the pore radius, produced using the HOLE2 software (Smart *et al.*, 1996). Radius of the channel lumen in Å is shown in blue and the 2.3 Å cutoff for conduction of sodium ions in orange. **(C)** shows helices S5 and S6 and the selectivitiy filter of adjacent monomers indicating the position of constrictions shown in **(B)**. Sodium ion coordinated by Glu178 of the selectivity (purple sphere). Figure made in ChimeraX (Pettersen *et al.*, 2021).

Crystal contacts may also explain why the CTD of NavMs-nanodiscs is less well-resolved than in crystal structures (approximately 5-5.5 Å, see Figure 6.15), as in a nanodisc the CTD would have a wider range of mobility than a crystal lattice; in the case of NavMs-nanodiscs, only the positions of the helices of the CTD were definitively placed (see Figure 6.20). Despite the lower resolution, there is a clear difference in the position of the N-terminal end of the coiled-coil helices of the CTD between the inactivated crystal structure and NavMs-nanodiscs. In the crystal structure, the N-terminal of each CTD helix is shifted towards the membrane region and bent towards the C-terminal end of S6 so that it is positioned approximately 8 Å away from equivalent residues in NavMs-nanodiscs, whilst the same region of NavMs-nanodiscs is placed within only 3-4 Å of the equivalent residues in the open, apo NavMs model (see Figure 6.22). This could suggest that this CTD orientation in the 4-hydroxytamoxifen bound structure is present only due to interactions with the ligand bound below the intracellular gate, rather than exemplifying an inactivated NavMs conformation.

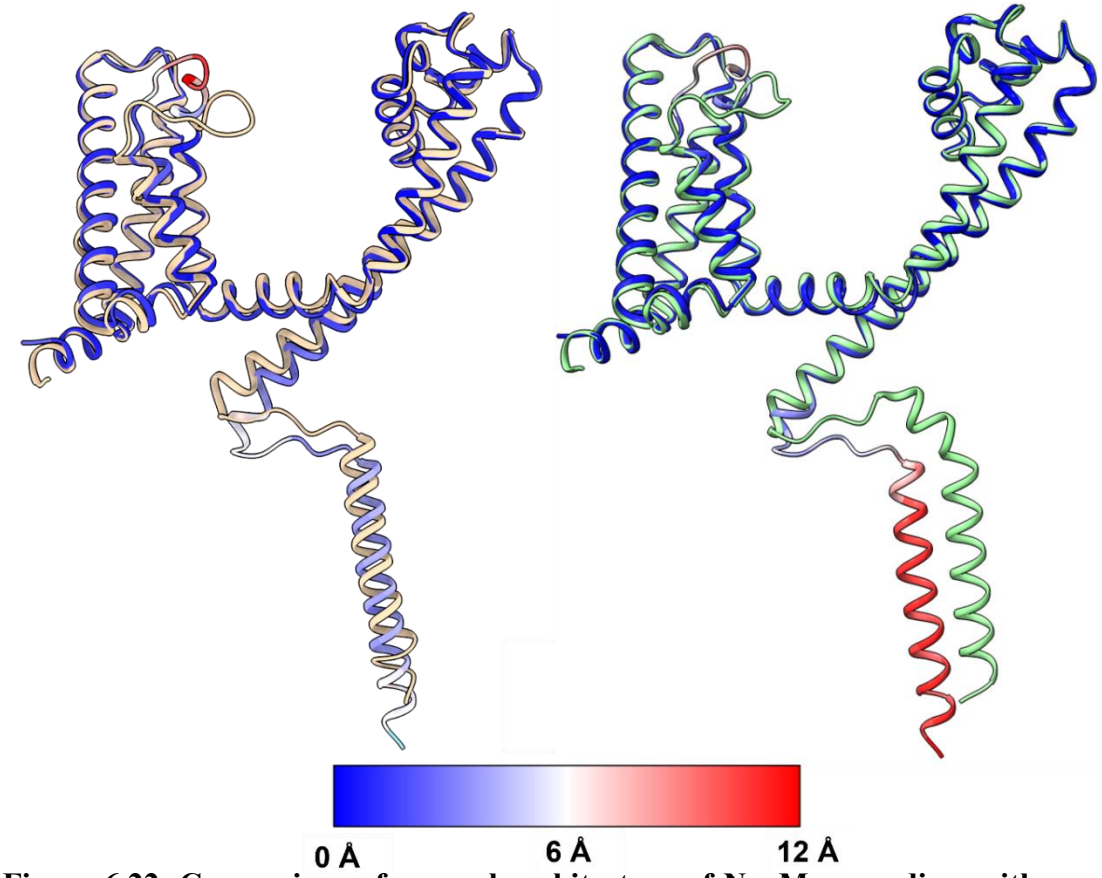

Figure 6.22. Comparison of general architecture of NavMs-nanodiscs with open apo state and inactivated 4-hydroxytamoxifen-bound NavMs crystal structures. Figure shows the NavMs-nanodiscs model aligned to open (left, wheat colour, PDB 6SX5) and inactivated (right, green colour, PDB 6SXG) NavMs crystal structures (Sula *et al.*, 2021). NavMs-nanodiscs has been coloured by its root mean squared deviation (RMSD) to each crystal structure. Figure made in ChimeraX (Pettersen *et al.*, 2021).

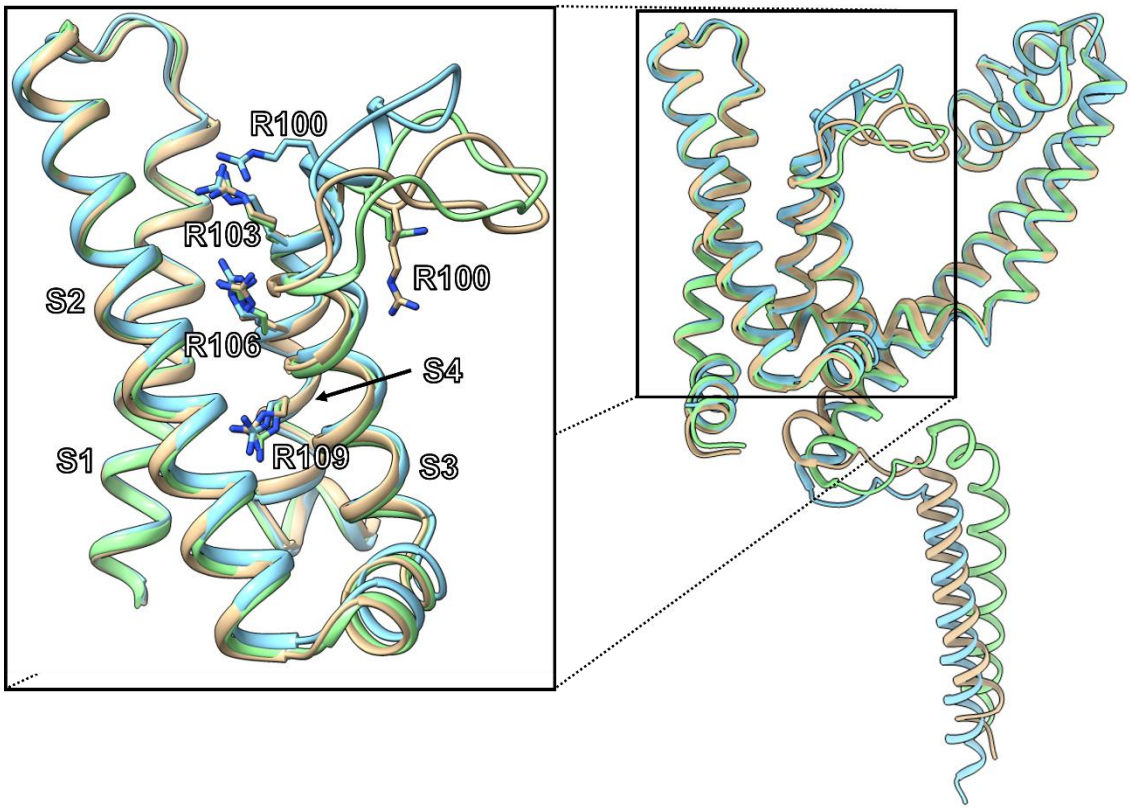

Figure 6.23. Comparison of voltage sensor domains of NavMs-nanodiscs with open apo state and inactivated 4-hydroxytamoxifen-bound NavMs crystal structures. Figure shows helices S1-S4 of the voltage sensor domain (VSD) of the NavMs-nanodiscs model aligned to open (wheat colour, PDB 6SX5) and inactivated (green colour, PDB 6SXG) NavMs crystal structures (Sula *et al.*, 2021). Arginine gating charge residies R100, R103, R106, and R109 (R1-R4, respectively) are shown for comparison. Figure made in ChimeraX (Pettersen *et al.*, 2021)

Building of the model also allowed the speculated shift of the C-terminal end of S6 away from the membrane (see section 6.7.1) and reduced interaction between the S4-S5 linker and S6 to be analysed. The NavMs-nanodisc model does indeed show this shift (see Figure 6.24), along with reduced interactions in the 'interaction motif' which links Trp77 of S3 with the C-terminal end of S6 via the S4-S5 linker helix to stabilise the open pore (Sula *et al.*, 2017). As can be seen in Figure 6.24, the crystal structure of the open channel shows a hydrogen bond interaction between Trp77 of S3 and Gln122 of the S4-S5 linker, which also forms a hydrogen bond with Gln228 of the C-terminal end of S6. A third interaction is formed between Arg119 of the S4-S5 linker and Glu229 of S6. In contrast, the NavMs-nanodisc model only shows the Trp77-Gln122 and R119-E229 interactions, with Gln228 too far from the S4-S5 linker to be forming a hydrogen bond. This interaction is reduced further in the 4-hydroxytamoxifen model, where only Gln122 and Gln228 are interacting. This agrees with the narrower radius of the intracellular gate, measured in the 4-hydroxytamoxifen model to be constricted to ~1.4 Å (Sula *et al.*, 2021), versus 1.8 Å

in NavMs-nanodiscs, reinforcing the role of this interaction motif in affecting the radius, and therefore conductance, of the channel.

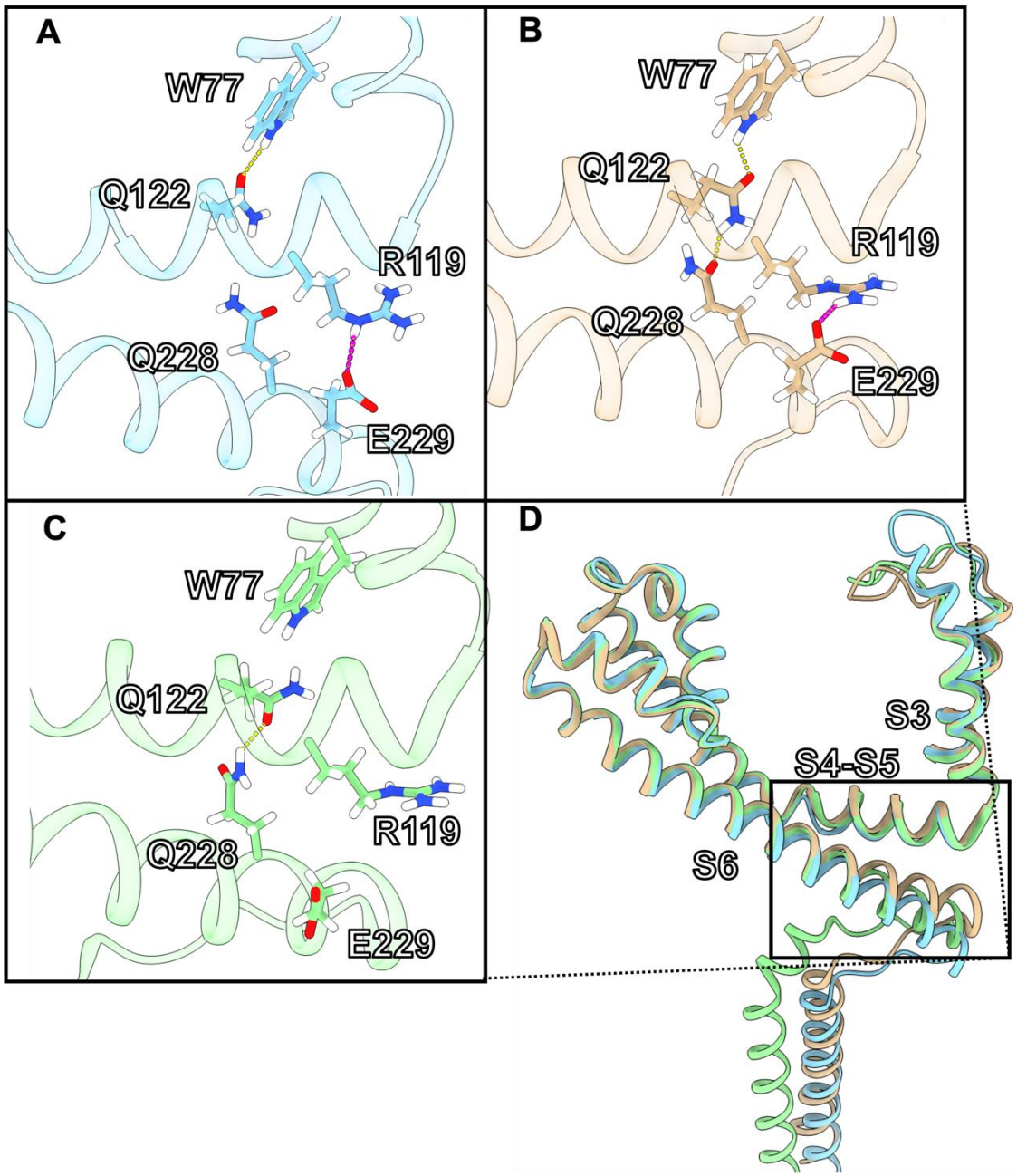

Figure 6.24. Comparison of interaction motif regions of NavMs-nanodiscs with open apo state and inactivated 4-hydroxytamoxifen-bound NavMs crystal structures. Figure shows inter-residue interactions in the interaction motif of S3, the S4-S5 linker, and S6. **(A)** shows the NavMs-nanodiscs model, **(B)** shows apo open NavMs (PDB 6SX5), **(C)** shows 4-hydroxytamoxifen-bound inactivated NavMs (PDB 6SXG) (Sula *et al.*, 2021). Yellow dashes indicate hydrogen bonds, magenta dashes indicate salt bridges. **(D)** shows an alignment of the three structures with the region of interest highlighted. Figure made in ChimeraX (Pettersen *et al.*, 2021).

6.9. Addendum: Discussion of NavMs-nanodiscs atomic model

The atomic model of NavMs in nanodiscs built after initial submission of this thesis allows for direct comparison of NavMs in a membrane mimetic, lipid environment with that of NavMs solubilised in detergent and visualised by X-ray crystallography. This has revealed that the structure of the N-terminal end of the S4 helix, which hosts the R1-R4 gating charges, is likely to be more structured than seen by crystallography, closer to that seen in the cryo-TEM structures of hNavs such as hNav1.8 (X. Huang *et al.*, 2022). It also suggested that the structure of the CTD seen in the inactivated NavMs bound to 4-hydroxytamoxifen may be an effect of the ligand, rather than the conformation of the channel itself. In addition, a reduced chain of interactions in the interaction motif was seen alongside a missing salt bridge between Glu81 of S3 and the gating charge R4, which could indeed be linked as changes in the conformation of the VSD drive conformational changes in the rest of the protein. A fully resolved VSD is important for understanding the function and conformational changes of the channel, but beyond this could also be useful for studying the effect of VSD-binding drugs and toxins to hNavs via humanising mutations, as seen in the case of NaChBac in nanodiscs (Gao *et al.*, 2020). It would also be interesting to see the effect that incorporation of NavMs into nanodiscs has on known drug binding, such as with cannabidiol (Sait *et al.*, 2020). Another interesting piece of research that could be performed now that the NavMs-nanodisc model has been successfully built would be to analyse it by molecular dynamics simulations, which have previously only been performed on crystal structures of NavMs (Ke *et al.*, 2018; Choudhury *et al.*, 2022). If structures could also be found of NavMs in nanodiscs in open and resting states, for example through modification of the lipid environment, or by advances in cryo-electron tomography allowing for visual analysis of NavMs in a membrane-like environment with a resting potential applied at higher resolutions than currently possible (Chang *et al.*, 2023), this could elucidate a voltage-gated sodium channel's full range of conformations in a near-native environment, providing valuable information that may be applied to the studies into hNavs.

Chapter 7

Discussion and Final Conclusions

7.1. Purification and electron microscopy analysis of hNav1.8 recombinantly expressed in *Tetrahymena thermophila*

7.1.1. Comparison of method to published hNav structure protocols

Over the course of this project structures were published of almost all of the human sodium channel isoforms, often also with modulating drugs or toxins bound; to date, this includes hNavs 1.1 (Pan *et al.*, 2021), 1.2 (Pan *et al.*, 2019), 1.3 (Li *et al.*, 2022), 1.4 (Pan *et al.*, 2018), 1.5 (Li, Jin, Wu, Huang, *et al.*, 2021; Li, Jin, Wu, Zhao, *et al.*, 2021), 1.6 (Fan *et al.*, 2023; Li *et al.*, 2023), 1.7 (Shen *et al.*, 2019; G. Huang *et al.*, 2022; Zhang *et al.*, 2022), and 1.8 (X. Huang *et al.*, 2022), with only the structure of hNav1.9 yet to be experimentally determined. Several conclusions can be drawn from the methods of these publications.

Firstly, human embryonic kidney (HEK) cells are clearly the most successful system, specifically the 293F strain developed for faster growth expression; typically when expressing cytotoxic recombinant proteins, the faster an expression can be completed, the better. An expression system based on human cells will have all the chaperones and other apparatus needed to correctly fold human sodium channels, and will also have the appropriate regime of post-translational modification to produce a functional and stable human protein, which may be crucial for expressing hNavs as glycosylation has been shown to be important in hNav function; for example, by affecting gating by modulating the electric field around the voltage sensor domains (VSDs) or for protein-protein interaction and trafficking (Swanwick, Pristerá and Okuse, 2010; Laedermann, Abriel and Decosterd, 2015). Cell lines derived from other sources, e.g., the Sf9 insect cell line derived from *Spodoptera frugiperda*, have been shown to differ in post-translational modification compared to human-derived cell lines (Jenkins, Parekh and James, 1996); as a eukaryote, *T. thermophila* is also capable of post-translational modification, but may not implement it in the same way, i.e., different levels of glycosylation or other post-translational modification could be applied to the protein.

Secondly, our construct differed from those used in these published studies by containing a FLAG-tag and 10x His-tag rather than a FLAG-tag and twin Strep-tag, due to solubility

issues when a FLAG-tag and Strep-tag was attempted. Twin Strep-tag based affinity chromatography involves binding of a 16 residue amino acid sequence (two copies of the eponymous Strep-tag II sequence, Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) to the biotin binding pocket of a bioengineered form of streptavidin (Schmidt and Skerra, 2007), Strep-Tactin, which is conjugated to the surface of sepharose beads. His-tag affinity chromatography in mammalian systems requires vigorous optimisation, as eukaryotic host cell proteins are present which bind non-specifically to Ni^{2+}, reducing the effectiveness of the purification step and necessitating optimised washes where target protein can be lost. Conversely, non-specific binding to Strep-Tactin is much less common, leading to improved purity in a single step of binding to beads and elution with desthiobiotin. It may be that continued optimisation of solubilisation conditions of the hNav1.8 FLAG and twin Strep-tag construct would have enabled purification in this manner.

Another significant difference in purification protocols is that in each case of published hNav structures, the target protein has been transiently coexpressed with at least one of its associated β subunits, whereas the α subunit of hNav1.8 was expressed alone in our protocol. Apart from hNav1.5 and hNav1.8 the β subunits were resolved in the final structure, albeit generally at a lower resolution than the α subunits, but even when unresolved and shown by mass spectrometry to be barely present in the final sample (likely due to weaker associations between some isoforms and their β subunits than others meaning they did not survive the entire purification) it was noted that cotransfection alongside β subunits increased the final yield of target protein.

These issues may have been possible to overcome if a successful route had been found to reduce degradation. As mentioned in Chapter 5, we speculated that the presence of *T. thermophila* ubiquitin, indicated by mass spectrometry, coupled with knowledge of high levels of proteases in the host system (Eisen *et al.*, 2006; Ruehle, Orias and Pearson, 2016) could mean that hNav1.8 was being targeted for proteolytic degradation. It may be that the protease inhibitors used during purification were not effective against *T. thermophila* proteases, and specialised protease inhibitors would make *T. thermophila* a cost-effective recombinant expression system, with up to 5 x cheaper cultivation media and even 10 x cheaper cell line stocks than mammalian expression systems such as HEK293 cells.

In summary, it appears that the key to successfully expressing and purifying hNavs and producing cryo-TEM samples with enough good quality particles to obtain high-

resolution structural data is expression in HEK293F cells alongisde the relevant β subunit of that isoform, followed by two stages of affinity chromatography by twin Strep-tag and FLAG tag before size exclusion chromatography. However, hNavs still remain challenging targets for structural studies. They require expensive detergents and other reagents such as sepharose beads for Strep-tag and FLAG-tag affinity purification, which are significantly higher cost than Ni^{2+} sepharose, and high volumes of mammalian tissue expression are needed; the highest resolution hNav structure published so far, the 2.2 Å model of the hNav1.7-β1-β2 complex (G. Huang *et al.*, 2022), used 40 L of transfected HEK293F cells.

7.2. Purification and high-resolution structural determination of NavMs in membrane scaffold protein nanodiscs

7.2.1. MSP nanodiscs are an effective method for studying NavMs by cryo-TEM

The aim of this project was to obtain high-resolution structural data of NavMs by cryo-TEM, and in the process develop a protocol for the study of this useful mechanistic and pharmacological model for medically important hNavs. MSP nanodiscs provide an environment closer to physiological conditions than detergent, being a mimetic of a cell membrane, and therefore can impart structural information that may be lost in X-ray crystallography conditions, and ideally would allow the study of the binding of drugs or toxins which are not suited to study by X-ray crystallography due to crystal packing at the surface of the protein. Reconstruction of the NavMs-nanodisc complex at a global resolution of 3.07 Å, enough to resolve the position of side chains in the transmembrane region of NavMs, shows that this has been a success. Of particular interest is the extension of 3-10 helix at the N-terminal end of S4, which was modelled as an unstructured loop or left unmodelled in previous NavMs crystal structures (Sula *et al.*, 2017, 2021; Sait *et al.*, 2020), and visualisation of the inactivated conformation without modulating ligands, which was previously only visible by binding to tamoxifen and its metabolites (Sula *et al.*, 2021). Now that an effective purification, reconstitution, grid preparation and data processing pipeline has been developed, this could be exploited for high-throughput screening of drug interactions, as well as the structural and functional effect of mutations linked to channelopathies in hNavs. Difficulties in crystallisation of WT NavMs mean that an F208L mutant is used, as the Wallace group has found that it is more readily crystallisable; while this mutation is not expected to have an effect on channel function,

it is still useful to have a method of reliably visualising NavMs without mutations and in the apo state.

All hNav structures so far published have been solubilised in detergent, so it would be interesting to isolate hNavs in MSP nanodiscs for structural study and see whether there are noticeable differences in structure, however isolation in MSP nanodiscs does require more protein sample than in detergents due to loss at the reconstitution stage, which may be challenging with notoriously difficult hNav targets. The fact that the NavMs-nanodisc complex appears to differ at the interaction motif (where S3, the S4-5 linker and the C-terminal end of S6 interact) to previously published crystal structures of apo NavMs (Sula *et al.*, 2017, 2021) suggests that the change in conditions has influenced the conformation of the protein and warrants further investigation.

7.2.2. Further research and future directions

Model building: As mentioned in the previous section, time constraints meant that building the NavMs sequence into a molecular model was not possible. The next step, given time, would be to build this model, for example in Coot (Emsley *et al.*, 2010) and Phenix (Liebschner *et al.*, 2019) so that it can be more accurately and reliably compared with other molecular models, rather than through qualitative visual judgment. The hypothesised reduced binding in the interaction motif, coordination of Na^+ near the intracellular end of the pore lumen, and identification of lipid interactions all require the building of a molecular model.

Improving the NavMs-nanodisc reconstruction: Whilst 3.07 Å is a suitable resolution for the assignment of side chains, these placements would be more reliable at higher resolution. Pushing the resolution higher would also possibly allow us to answer the question of whether the transition of a segment of S6 from α-helix to π-helix in order to hydrate the pore, simulated by molecular dynamics (Choudhury *et al.*, 2022) is a real phenomenon in NavMs, as has been reported for hNav1.7 (G. Huang *et al.*, 2022). As indicated by the ResLog analysis described in section 6.6.3, this could be achieved simply by collecting more data and adding a greater number of particles into the reconstruction. Another option to improve the resolution could be to make use of neural network particle picking methods such as Topaz (Bepler *et al.*, 2019), which may be more effective at picking NavMs-nanodisc complexes over empty nanodiscs. Finally, as only one type of phospholipid was used in this investigation (POPC), other phospholipid mixtures such as

brain polar lipid, soybean polar lipid extract, or *E. coli* polar lipid extract could be trialled, in case this has an effect on resolution via greater stabilisation or leads to different conformations of NavMs.

Drug binding to NavMs: As mentioned earlier, this method was developed partially as a method to analyse a wider range of drug binding to NavMs than is available by crystallographic methods. Gating modifier toxins are interesting in terms of drug development as they often show subtype specificity for particular hNav isoforms, and binding of these has been shown to be amenable to study by cryo-TEM previously (Gao *et al.*, 2020; Wisedchaisri *et al.*, 2021; G. Huang *et al.*, 2022), and humanising mutations in NavMs could be used to the same effect.

Computational analysis: Previous published apo NavMs structures are indicated to be on the open, conducting conformation (Sula *et al.*, 2017, 2021), however, molecular dynamics simulations of the 2017 structure (PDB 5HVX) indicated that there may be an energetic barrier to the conductance of hydrated Na^+ ions, with a discontinuous 'water wire' through the channel (Ke *et al.*, 2018; Choudhury *et al.*, 2022). Performing the same simulations on the NavMs-nanodisc complex, as well as measuring the dimensions of the pore lumen using the HOLE2 software (Smart *et al.*, 1996) would be important in determining whether the channel is in an activated, open, and conducting state.

All hNav structures so far published have been solubilised in detergent, so it would be interesting to isolate hNavs in MSP nanodiscs for structural study and see whether there are noticeable differences in structure, however isolation in MSP nanodiscs does require more protein sample than in detergents due to loss at the reconstitution stage, which may be challenging with notoriously difficult hNav targets. The fact that the NavMs-nanodisc complex appears to differ at the interaction motif to previously published crystal structures of apo NavMs suggests that the change in conditions has influenced the structure of the protein and warrants further investigation.

7.3. Final remarks

Over the course of this PhD, the revolutions in cryo-TEM microscope direct electron detectors, the huge range of cryo-TEM grids and blotting devices now available, mimetics of physiological cell membranes for integral membrane proteins, massive software advances such as the rapid processing power of cryoSPARC (Punjani *et al.*, 2017) and

the advent of deep learning neural networks in cryo-TEM data processing such as Topaz and CryoDRGN (Bepler *et al.*, 2019; Zhong *et al.*, 2021), mean that more and more challenging targets are becoming amenable to study by cryo-TEM. Indicative of this is the rapid release of all but one hNav isoform (hNav1.9) in the time it took to complete the PhD. The release of these structures does not however mean that all the work is done; the lack of structures of each isoform in different conformational states still clouds their precise mechanism, and the wide range of diseases associated with hNavs means that there is still much to learn from studying the binding of drugs and toxins to these enigmatic proteins, including using prokaryotic Navs as models, without which a novel binding site and interaction with the commonly prescribed cancer prophylactic drug tamoxifen would not have been identified (Sula *et al.*, 2021). It would certainly also be worthwhile to study hNavs stabilised in membrane mimetics such as membrane scaffold protein nanodiscs, as this may give insights into the physiological function of channels which are not possible when solubilised in detergent, as all published hNav structures have been so far. New advances in nanodiscs themselves are also occurring; the development of covalently circularised nanodiscs has increased the size limit of nanodiscs from 15-17 nm up to as much as 90 nm diameter (Nasr and Wagner, 2018; Padmanabha Das *et al.*, 2020), meaning that large protein complexes rather than single membrane proteins can be isolated, and nanodiscs may be able to more closely mimic real cell membranes, for example by introducing membrane curvature.

It may be that the next great advances in cryo-TEM will not be higher-powered (and more expensive) microscopes and another resolution revolution, but a revolution in accessibility to cryo-TEM for researchers without the resources to host their own Titan Krios microscope. It has been demonstrated that even a 100 keV microscope is capable of elucidating structural details given the right circumstances (Naydenova *et al.*, 2019). The ability to perform structural study without high-demand 300 keV microscopes would be extremely useful for investigating drug binding on possible medical targets, as more targets could be probed without waiting for the most powerful microscopes. Any discussion of the future of structural biology must also mention AlphaFold2, (Jumper *et al.*, 2021) which has itself revolutionised the prediction of protein structures; as computational hardware and artificial intelligence grow in capability, algorithms such as AlphaFold2 will only grow in power alongside them.

Bibliography

Ahern, C.A., Payandeh, J., Bosmans, F. and Chanda, B. (2016) 'The hitchhiker's guide to the voltage-gated sodium channel galaxy', *Journal of General Physiology*, 147(1), pp. 1–24

Ahuja, S., Mukund, S., Deng, L., Khakh, K., Chang, E., Ho, H., Shriver, S., Young, C., Lin, S., Johnson, J.P., Wu, P., Li, J., Coons, M., Tam, C., Brillantes, B., Sampang, H., Mortara, K., Bowman, K.K., Clark, K.R., Estevez, A., Xie, Z., Verschoof, H., Grimwood, M., Dehnhardt, C., Andrez, J.-C., Focken, T., Sutherlin, D.P., Safina, B.S., Starovasnik, M.A., Ortwine, D.F., Franke, Y., Cohen, C.J., Hackos, D.H., Koth, C.M. and Payandeh, J. (2015) 'Structural basis of Nav1.7 inhibition by an isoform-selective small-molecule antagonist', *Science*, 350(6267)

Akopian, A.N., Sivilotti, L. and Wood, J.N. (1996) 'A tetrodotoxin-resistant voltage-gated sodium channel expressed by sensory neurons', *Nature*, 379(6562), pp. 257–262

Ashtiani, D., Venugopal, H., Belousoff, M., Spicer, B., Mak, J., Neild, A. and de Marco, A. (2018) 'Delivery of femtolitre droplets using surface acoustic wave based atomisation for cryo-EM grid preparation', *Journal of Structural Biology*, 203(2), pp. 94–101

Bagal, S.K., Marron, B.E., Owen, R.M., Storer, R.I. and Swain, N.A. (2015) 'Voltage gated sodium channels as drug discovery targets', *Channels*, 9(6), pp. 360–366

Bagnéris, C., DeCaen, P.G., Naylor, C.E., Pryde, D.C., Nobeli, I., Clapham, D.E. and Wallace, B.A. (2014) 'Prokaryotic NavMs channel as a structural and functional model for eukaryotic sodium channel antagonism', *Proceedings of the National Academy of Sciences*, 111(23), pp. 8428–8433

Bagnéris, C., Naylor, C.E., McCusker, E.C. and Wallace, B.A. (2015) 'Structural model of the open–closed–inactivated cycle of prokaryotic voltage-gated sodium channels', *The Journal of General Physiology*, 145(1), pp. 5–16

Bayless, B.A., Galati, D.F. and Pearson, C.G. (2015) 'Tetrahymena basal bodies', *Cilia*, 5(1), p. 1

Bayless, B.A., Navarro, F.M. and Winey, M. (2019) 'Motile Cilia: Innovation and Insight From Ciliate Model Organisms', *Frontiers in Cell and Developmental Biology*, 7

Bennett, D.L., Clark, A.J., Huang, J., Waxman, S.G. and Dib-Hajj, S.D. (2019) 'The Role of Voltage-Gated Sodium Channels in Pain Signaling', *Physiological Reviews*, 99(2), pp. 1079–1151

Bepler, T., Morin, A., Rapp, M., Brasch, J., Shapiro, L., Noble, A.J. and Berger, B. (2019) 'Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs', *Nature Methods*, 16(11), pp. 1153–1160

Berman, H.M. (2000) 'The Protein Data Bank', *Nucleic Acids Research*, 28(1), pp. 235–242

Boiteux, C., Vorobyov, I., French, R.J., French, C., Yarov-Yarovoy, V. and Allen, T.W. (2014) 'Local anesthetic and antiepileptic drug access and binding to a bacterial voltage-gated sodium channel', *Proceedings of the National Academy of Sciences*, 111(36), pp. 13057–13062

Le Bon, C., Marconnet, A., Masscheleyn, S., Popot, J.-L. and Zoonens, M. (2018) 'Folding and stabilizing membrane proteins in amphipol A8-35', *Methods*, 147, pp. 95–105

Le Bon, C., Popot, J.-L. and Giusti, F. (2014) 'Labeling and Functionalizing Amphipols for Biological Applications', *The Journal of Membrane Biology*, 247(9–10), pp. 797–814

Booker, J. (2019) *Structural Basis of Voltage-gated Sodium Channel Function*. Birkbeck College, University of London

Breivik, H., Collett, B., Ventafridda, V., Cohen, R. and Gallacher, D. (2006) 'Survey of chronic pain in Europe: Prevalence, impact on daily life, and treatment', *European Journal of Pain*, 10(4), pp. 287–287

Brenner, S. and Horne, R.W. (1959) 'A negative staining method for high resolution electron microscopy of viruses', *Biochimica et Biophysica Acta*, 34, pp. 103–110

Brilot, A.F., Chen, J.Z., Cheng, A., Pan, J., Harrison, S.C., Potter, C.S., Carragher, B., Henderson, R. and Grigorieff, N. (2012) 'Beam-induced motion of vitrified specimen on holey carbon film', *Journal of Structural Biology*, 177(3), pp. 630–637

Burnette, W.N. (1981) '"Western Blotting": Electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A', *Analytical*

Biochemistry, 112(2), pp. 195–203

Catterall, W.A. (1987) 'Common modes of drug action on Na+ channels: local anesthetics, antiarrhythmics and anticonvulsants', *Trends in Pharmacological Sciences*, 8(2), pp. 57–65

Catterall, W.A., Goldin, A.L. and Waxman, S.G. (2005) 'Nomenclature and Structure-Function Relationships of Voltage-Gated Sodium Channels', *Pharmacological Reviews*, 57(4), pp. 397–409

Catterall, W.A., Wisedchaisri, G. and Zheng, N. (2020) 'The conformational cycle of a prototypical voltage-gated sodium channel', *Nature Chemical Biology*, 16(12), pp. 1314–1320

Chang, S.-Y.S., Dijkman, P.M., Wiessing, S.A. and Kudryashev, M. (2023) 'Determining the structure of the bacterial voltage-gated sodium channel NaChBac embedded in liposomes by cryo electron tomography and subtomogram averaging', *Scientific Reports*, 13(1), p. 11523

Chen, J., Noble, A.J., Kang, J.Y. and Darst, S.A. (2019) 'Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO', *Journal of Structural Biology: X*, 1, p. 100005

Cheng, A., Kim, P.T., Kuang, H., Mendez, J.H., Chua, E.Y.D., Maruthi, K., Wei, H., Sawh, A., Aragon, M.F., Serbynovskyi, V., Neselu, K., Eng, E.T., Potter, C.S., Carragher, B., Bepler, T. and Noble, A.J. (2023) 'Fully automated multi-grid cryoEM screening using Smart Leginon', *IUCrJ*, 10(1), pp. 77–89

Choudhury, K., Kasimova, M.A., McComas, S., Howard, R.J. and Delemotte, L. (2022) 'An open state of a voltage-gated sodium channel involving a π-helix and conserved pore-facing asparagine', *Biophysical Journal*, 121(1), pp. 11–22

Ciechanover, A. (2005) 'Proteolysis: from the lysosome to ubiquitin and the proteosome', *Molecular Cell Biology*, 6(January), p. 9

Ciechanover, A., Orian, A. and Schwartz, A.L. (2000) 'Ubiquitin-mediated proteolysis: biological regulation via destruction', *BioEssays*, 22(5), pp. 442–451

Consolati, T., Locke, J., Roostalu, J., Chen, Z.A., Gannon, J., Asthana, J., Lim, W.M., Martino, F., Cvetkovic, M.A., Rappsilber, J., Costa, A. and Surrey, T. (2020)

‘Microtubule Nucleation Properties of Single Human γTuRCs Explained by Their Cryo-EM Structure’, *Developmental Cell*, 53(5), pp. 603-617.e8

Consortium, T.U. (2023) ‘UniProt : the Universal Protein Knowledgebase in 2023’, *Nucleic Acids Research*, 51(D1), pp. D523–D531

Croll, T.I. (2018) ‘ISOLDE : a physically realistic environment for model building into low-resolution electron-density maps’, *Acta Crystallographica Section D Structural Biology*, 74(6), pp. 519–530

Cummins, T.R. (2007) ‘Setting up for the block: the mechanism underlying lidocaine’s use-dependent inhibition of sodium channels’, *The Journal of Physiology*, 582(1), pp. 11–11

Dandey, V.P., Budell, W.C., Wei, H., Bobe, D., Maruthi, K., Kopylov, M., Eng, E.T., Kahn, P.A., Hinshaw, J.E., Kundu, N., Nimigean, C.M., Fan, C., Sukomon, N., Darst, S.A., Saecker, R.M., Chen, J., Malone, B., Potter, C.S. and Carragher, B. (2020) ‘Time-resolved cryo-EM using Spotiton’, *Nature Methods*, 17(9), pp. 897–900

Dandey, V.P., Wei, H., Zhang, Z., Tan, Y.Z., Acharya, P., Eng, E.T., Rice, W.J., Kahn, P.A., Potter, C.S. and Carragher, B. (2018) ‘Spotiton: New features and applications’, *Journal of Structural Biology*, 202(2), pp. 161–169

Darrow, M.C., Moore, J.P., Walker, R.J., Doering, K. and King, R.S. (2019) ‘Chameleon: Next Generation Sample Preparation for CryoEM based on Spotiton’, *Microscopy and Microanalysis*, 25(S2), pp. 994–995

DeCaen, P.G., Yarov-Yarovoy, V., Sharp, E.M., Scheuer, T. and Catterall, W.A. (2009) ‘Sequential formation of ion pairs during activation of a sodium channel voltage sensor’, *Proceedings of the National Academy of Sciences*, 106(52), pp. 22498–22503

DeCaen, P.G., Yarov-Yarovoy, V., Zhao, Y., Scheuer, T. and Catterall, W.A. (2008) ‘Disulfide locking a sodium channel voltage sensor reveals ion pair formation during activation’, *Proceedings of the National Academy of Sciences*, 105(39), pp. 15142–15147

Denisov, I.G., Grinkova, Y. V., Lazarides, A.A. and Sligar, S.G. (2004) ‘Directed Self-Assembly of Monodisperse Phospholipid Bilayer Nanodiscs with Controlled Size’, *Journal of the American Chemical Society*, 126(11), pp. 3477–3487

Denisov, I.G. and Sligar, S.G. (2016) ‘Nanodiscs for structural and functional studies of

membrane proteins', *Nature Structural & Molecular Biology*, 23(6), pp. 481–486

Denisov, I.G. and Sligar, S.G. (2017) 'Nanodiscs in Membrane Biochemistry and Biophysics', *Chemical Reviews*, 117(6), pp. 4669–4713

Drew, E.D. and Janes, R.W. (2019) '2StrucCompare: a webserver for visualizing small but noteworthy differences between protein tertiary structures through interrogation of the secondary structure content', *Nucleic Acids Research*, 47(W1), pp. W477–W481

Dudev, T. and Lim, C. (2014) 'Ion Selectivity Strategies of Sodium Channel Selectivity Filters', *Accounts of Chemical Research*, 47(12), pp. 3580–3587

Dündar Orhan, Y., Üstüntanır Dede, A.F., Duran, Ş. and Arslanyolu, M. (2024) 'Use of E-64 cysteine protease inhibitor for the recombinant protein production in Tetrahymena thermophila', *European Journal of Protistology*, 94, p. 126085

Eisen, J.A., Coyne, R.S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J.R., Badger, J.H., Ren, Q., Amedeo, P., Jones, K.M., Tallon, L.J., Delcher, A.L., Salzberg, S.L., Silva, J.C., Haas, B.J., Majoros, W.H., Farzad, M., Carlton, J.M., Smith, R.K., Garg, J., Orias, E., *et al.* (2006) 'Macronuclear Genome Sequence of the Ciliate Tetrahymena thermophila, a Model Eukaryote', *PLoS Biology*. Edited by M. Gelfand, 4(9), p. e286

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) 'Features and development of Coot', *Acta Crystallographica Section D Biological Crystallography*, 66(4), pp. 486–501

Esmaili, M. and Overduin, M. (2018) 'Membrane biology visualized in nanometer-sized discs formed by styrene maleic acid polymers', *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1860(2), pp. 257–263

Estrozi, L.F. and Navaza, J. (2008) 'Fast projection matching for cryo-electron microscopy image reconstruction', *Journal of Structural Biology*, 162(2), pp. 324–334

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J. and Hassabis, D. (2022) 'Protein complex prediction with AlphaFold-Multimer', *bioRxiv*, p. 2021.10.04.463034

Fan, X., Huang, J., Jin, X. and Yan, N. (2023) 'Cryo-EM structure of human voltage-

gated sodium channel Na v 1.6', *Proceedings of the National Academy of Sciences*, 120(5), p. e2220578120

Favre, I., Moczydlowski, E. and Schild, L. (1996) 'On the structural basis for ionic selectivity among Na+, K+, and Ca2+ in the voltage-gated sodium channel', *Biophysical Journal*, 71(6), pp. 3110–3125

Feathers, J.R., Spoth, K.A. and Fromme, J.C. (2021) 'Experimental evaluation of super-resolution imaging and magnification choice in single-particle cryo-EM', *Journal of Structural Biology: X*, 5, p. 100047

Feng, X., Fu, Z., Kaledhonkar, S., Jia, Y., Shah, B., Jin, A., Liu, Z., Sun, M., Chen, B., Grassucci, R.A., Ren, Y., Jiang, H., Frank, J. and Lin, Q. (2017) 'A Fast and Effective Microfluidic Spraying-Plunging Method for High-Resolution Single-Particle Cryo-EM', *Structure*, 25(4), pp. 663-670.e3

Frank, J. (1975) 'Averaging of low exposure electron micrographs of non-periodic objects', *Ultramicroscopy*, 1(2), pp. 159–162

Franken, L.E., Grünewald, K., Boekema, E.J. and Stuart, M.C.A. (2020) 'A Technical Introduction to Transmission Electron Microscopy for Soft-Matter: Imaging, Possibilities, Choices, and Technical Developments', *Small*, 16(14), p. 1906198

Gamal El-Din, T.M., Lenaeus, M.J., Zheng, N. and Catterall, W.A. (2018) 'Fenestrations control resting-state block of a voltage-gated sodium channel', *Proceedings of the National Academy of Sciences*, 115(51), pp. 13111–13116

Gao, S., Valinsky, W.C., On, N.C., Houlihan, P.R., Qu, Q., Liu, L., Pan, X., Clapham, D.E. and Yan, N. (2020) 'Employing NaChBac for cryo-EM analysis of toxin action on voltage-gated Na + channels in nanodisc', *Proceedings of the National Academy of Sciences*, 117(25), pp. 14187–14193

Gasteiger, E. (2003) 'ExPASy: the proteomics server for in-depth protein knowledge and analysis', *Nucleic Acids Research*, 31(13), pp. 3784–3788

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D. and Bairoch, A. (2005) 'The Proteomics Protocols Handbook', *The Proteomics Protocols Handbook*, pp. 571–608

Gewering, T., Januliene, D., Ries, A.B. and Moeller, A. (2018) 'Know your detergents:

A case study on detergent background in negative stain electron microscopy', *Journal of Structural Biology*, 203(3), pp. 242–246

Gilchrist, J. and Bosmans, F. (2018) 'Using voltage-sensor toxins and their molecular targets to investigate Na V 1.8 gating', *The Journal of Physiology*, 596(10), pp. 1863–1872

Glaeser, R.M., Han, B.-G., Csencsits, R., Killilea, A., Pulk, A. and Cate, J.H.D. (2016) 'Factors that Influence the Formation and Stability of Thin, Cryo-EM Specimens', *Biophysical Journal*, 110(4), pp. 749–755

Glitsch, H.G. (2001) 'Electrophysiology of the Sodium-Potassium-ATPase in Cardiac Cells', *Physiological Reviews*, 81(4), pp. 1791–1826

Goldberg, Y., MacFarlane, J., MacDonald, M., Thompson, J., Dube, M.-P., Mattice, M., Fraser, R., Young, C., Hossain, S., Pape, T., Payne, B., Radomski, C., Donaldson, G., Ives, E., Cox, J., Younghusband, H., Green, R., Duff, A., Boltshauser, E., Grinspan, G., Dimon, J., Sibley, B., Andria, G., Toscano, E., Kerdraon, J., Bowsher, D., Pimstone, S., Samuels, M., Sherrington, R. and Hayden, M. (2007) 'Loss-of-function mutations in the Nav1.7 gene underlie congenital indifference to pain in multiple human populations', *Clinical Genetics*, 71(4), pp. 311–319

Graves, T.D. and Hanna, M.G. (2005) 'Neurological channelopathies', *Postgraduate Medical Journal*, 81(951), pp. 20–32

Grinkova, Y. V., Denisov, I.G. and Sligar, S.G. (2010) 'Engineering extended membrane scaffold proteins for self-assembly of soluble nanoscale lipid bilayers', *Protein Engineering Design and Selection*, 23(11), pp. 843–848

Gubellini, F., Verdon, G., Karpowich, N.K., Luff, J.D., Boël, G., Gauthier, N., Handelman, S.K., Ades, S.E. and Hunt, J.F. (2011) 'Physiological Response to Membrane Protein Overexpression in E. coli', *Molecular & Cellular Proteomics*, 10(10), p. M111.007930

Heersche, H.B., Jarillo-Herrero, P., Oostinga, J.B., Vandersypen, L.M.K. and Morpurgo, A.F. (2007) 'Bipolar supercurrent in graphene', *Nature*, 446(7131), pp. 56–59

Helliwell, K.E., Chrachri, A., Koester, J.A., Wharam, S., Taylor, A.R., Wheeler, G.L. and Brownlee, C. (2020) 'A novel single-domain Na1-selective voltage-gated channel in

photosynthetic eukaryotes', *Plant Physiology*, 184(4), pp. 1674–1683

Henderson, R. (2013) 'Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise', *Proceedings of the National Academy of Sciences*, 110(45), pp. 18037–18041

Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H. (1990) 'Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy', *Journal of Molecular Biology*, 213(4), pp. 899–929

Henderson, R. and Glaeser, R.M. (1985) 'Quantitative analysis of image contrast in electron micrographs of beam-sensitive crystals', *Ultramicroscopy*, 16(2), pp. 139–150

Hille, B. (1977) 'Local anesthetics: hydrophilic and hydrophobic pathways for the drug-receptor reaction.', *Journal of General Physiology*, 69(4), pp. 497–515

Hodge, K., Have, S. Ten, Hutton, L. and Lamond, A.I. (2013) 'Cleaning up the masses: Exclusion lists to reduce contamination with HPLC-MS/MS', *Journal of Proteomics*, 88, pp. 92–103

Hodgkin, A.L. and Huxley, A.F. (1952) 'A quantitative description of membrane current and its application to conduction and excitation in nerve', *The Journal of Physiology*, 117(4), pp. 500–544

Hooke, R. (1665) *Micrographia: or, Some physiological descriptions of minute bodies made by magnifying glasses. With observations and inquiries thereupon*. London: J. Martyn and J. Allestry

Huang, G., Liu, D., Wang, W., Wu, Q., Chen, J., Pan, X., Shen, H. and Yan, N. (2022) 'High-resolution structures of human Nav1.7 reveal gating modulation through α-π helical transition of S6IV', *Cell Reports*, 39(4), p. 110735

Huang, J., Fan, X., Jin, X., Jo, S., Zhang, H.B., Fujita, A., Bean, B.P. and Yan, N. (2023) 'Cannabidiol inhibits Nav channels through two distinct binding sites', *Nature Communications*, 14(1), p. 3613

Huang, W., Liu, M., Yan, S.F. and Yan, N. (2017) 'Structure-based assessment of disease-related mutations in human voltage-gated sodium channels', *Protein & Cell*, 8(6), pp. 401–438

Huang, X., Jin, X., Huang, G., Huang, J., Wu, T., Li, Z., Chen, J., Kong, F., Pan, X. and Yan, N. (2022) 'Structural basis for high-voltage activation and subtype-specific inhibition of human Na v 1.8', *Proceedings of the National Academy of Sciences*, 119(30), p. e2208211119

Hull, J.M. and Isom, L.L. (2018) 'Voltage-gated sodium channel β subunits: The power outside the pore in brain development and disease', *Neuropharmacology*, 132, pp. 43–57

Humphrey, W., Dalke, A. and Schulten, K. (1996) 'VMD: Visual molecular dynamics', *Journal of Molecular Graphics*, 14(1), pp. 33–38

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J. and Mann, M. (2005) 'Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein', *Molecular & Cellular Proteomics*, 4(9), pp. 1265–1272

Jain, T., Sheehan, P., Crum, J., Carragher, B. and Potter, C.S. (2012) 'Spotiton: A prototype for an integrated inkjet dispense and vitrification system for cryo-TEM', *Journal of Structural Biology*, 179(1), pp. 68–75

Jakobi, A.J., Huber, S.T., Mortensen, S.A., Schultz, S.W., Palara, A., Kuhm, T., Shrestha, B.K., Lamark, T., Hagen, W.J.H., Wilmanns, M., Johansen, T., Brech, A. and Sachse, C. (2020) 'Structural basis of p62/SQSTM1 helical filaments and their role in cellular cargo uptake', *Nature Communications*, 11(1), p. 440

Jarvis, M.F., Honore, P., Shieh, C.-C., Chapman, M., Joshi, S., Zhang, X.-F., Kort, M., Carroll, W., Marron, B., Atkinson, R., Thomas, J., Liu, D., Krambis, M., Liu, Y., McGaraughty, S., Chu, K., Roeloffs, R., Zhong, C., Mikusa, J.P., Hernandez, G., Gauvin, D., Wade, C., Zhu, C., Pai, M., Scanio, M., Shi, L., Drizin, I., Gregg, R., Matulenko, M., Hakeem, A., Gross, M., Johnson, M., Marsh, K., Wagoner, P.K., Sullivan, J.P., Faltynek, C.R. and Krafte, D.S. (2007) 'A-803467, a potent and selective Na v 1.8 sodium channel blocker, attenuates neuropathic and inflammatory pain in the rat', *Proceedings of the National Academy of Sciences*, 104(20), pp. 8520–8525

Jenkins, N., Parekh, R.B. and James, D.C. (1996) 'Getting the glycosylation right: Implications for the biotechnology industry', *Nature Biotechnology*, 14(8), pp. 975–981

Jiang, D., Shi, H., Tonggu, L., Gamal El-Din, T.M., Lenaeus, M.J., Zhao, Y., Yoshioka,

C., Zheng, N. and Catterall, W.A. (2020) ‘Structure of the Cardiac Sodium Channel’, *Cell*, 180(1), pp. 122-134.e10

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P. and Hassabis, D. (2021) ‘Highly accurate protein structure prediction with AlphaFold’, *Nature*, 596(7873), pp. 583–589

Kabsch, W. and Sander, C. (1983) ‘Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features’, *Biopolymers*, 22(12), pp. 2577–2637

Kaledhonkar, S., Fu, Z., White, H. and Frank, J. (2018) ‘Time-Resolved Cryo-electron Microscopy Using a Microfluidic Chip’, in J. Marsh (ed.) *Protein Complex Assembly*. New York, NY: Humana Press, pp. 59–71

Kalia, J., Milescu, M., Salvatierra, J., Wagner, J., Klint, J.K., King, G.F., Olivera, B.M. and Bosmans, F. (2015) ‘From Foe to Friend: Using Animal Toxins to Investigate Ion Channel Function’, *Journal of Molecular Biology*, 427(1), pp. 158–175

Karpievitch, Y. V., Polpitiya, A.D., Anderson, G.A., Smith, R.D. and Dabney, A.R. (2010) ‘Liquid chromatography mass spectrometry-based proteomics: Biological and technological aspects’, *The Annals of Applied Statistics*, 4(4)

Karuppasamy, M., Karimi Nejadasl, F., Vulovic, M., Koster, A.J. and Ravelli, R.B.G. (2011) ‘Radiation damage in single-particle cryo-electron microscopy: effects of dose and dose rate’, *Journal of Synchrotron Radiation*, 18(3), pp. 398–412

Ke, S., Ulmschneider, M.B., Wallace, B.A. and Ulmschneider, J.P. (2018) ‘Role of the Interaction Motif in Maintaining the Open Gate of an Open Sodium Channel’, *Biophysical Journal*, 115(10), pp. 1920–1930

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H. and Phillips, D.C. (1958) ‘A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis’, *Nature*, 181(4610), pp. 662–666

Krimmer, S.G., Bertoletti, N., Suzuki, Y., Katic, L., Mohanty, J., Shu, S., Lee, S., Lax, I., Mi, W. and Schlessinger, J. (2023) 'Cryo-EM analyses of KIT and oncogenic mutants reveal structural oncogenic plasticity and a target for therapeutic intervention', *Proceedings of the National Academy of Sciences*, 120(13), p. e2300054120

Kühlbrandt, W. (2014) 'The Resolution Revolution', *Science*, 343(6178), pp. 1443–1444

Laedermann, C.J., Abriel, H. and Decosterd, I. (2015) 'Post-translational modifications of voltage-gated sodium channels in chronic pain syndromes', *Frontiers in Pharmacology*, 6

Lawson, C.L., Patwardhan, A., Baker, M.L., Hryc, C., Garcia, E.S., Hudson, B.P., Lagerstedt, I., Ludtke, S.J., Pintilie, G., Sala, R., Westbrook, J.D., Berman, H.M., Kleywegt, G.J. and Chiu, W. (2016) 'EMDataBank unified data resource for 3DEM', *Nucleic Acids Research*, 44(D1), pp. D396–D403

Lee, S.C., Knowles, T.J., Postis, V.L.G., Jamshad, M., Parslow, R.A., Lin, Y., Goldman, A., Sridhar, P., Overduin, M., Muench, S.P. and Dafforn, T.R. (2016) 'A method for detergent-free isolation of membrane proteins in their local lipid environment', *Nature Protocols*, 11(7), pp. 1149–1162

van Leeuwenhoek, A. (1677) 'Observations, communicated to the publisher by Mr. Antony van Leewenhoeck, in a dutch letter of the 9th Octob. 1676. here English'd: concerning little animals by him observed in rain-well-sea- and snow water; as also in water wherein pepper had lain infus', *Philosophical Transactions of the Royal Society of London*, 12(133), pp. 821–831

Lenaeus, M.J., Gamal El-Din, T.M., Ing, C., Ramanadane, K., Pomès, R., Zheng, N. and Catterall, W.A. (2017) 'Structures of closed and open states of a voltage-gated sodium channel', *Proceedings of the National Academy of Sciences*, 114(15), pp. E3051–E3060

de Lera Ruiz, M. and Kraus, R.L. (2015) 'Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications', *Journal of Medicinal Chemistry*, 58(18), pp. 7093–7118

Li, X., Xu, F., Xu, H., Zhang, S., Gao, Y., Zhang, H., Dong, Y., Zheng, Y., Yang, B., Sun, J., Zhang, X.C., Zhao, Y. and Jiang, D. (2022) 'Structural basis for modulation of human NaV1.3 by clinical drug and selective antagonist', *Nature Communications*, 13(1),

pp. 1–10

Li, X., Zheng, S.Q., Egami, K., Agard, D.A. and Cheng, Y. (2013) 'Influence of electron dose rate on electron counting images recorded with the K2 camera', *Journal of Structural Biology*, 184(2), pp. 251–260

Li, Y., Yuan, T., Huang, B., Zhou, F., Peng, C., Li, X., Qiu, Y., Yang, B., Zhao, Y., Huang, Z. and Jiang, D. (2023) 'Structure of human NaV1.6 channel reveals Na+ selectivity and pore blockade by 4,9-anhydro-tetrodotoxin', *Nature Communications*, 14(1), p. 1030

Li, Z., Jin, X., Wu, T., Huang, G., Wu, K., Lei, J., Pan, X. and Yan, N. (2021) 'Structural Basis for Pore Blockade of the Human Cardiac Sodium Channel Na v 1.5 by the Antiarrhythmic Drug Quinidine', *Angewandte Chemie International Edition*, 60(20), pp. 11474–11480

Li, Z., Jin, X., Wu, T., Zhao, X., Wang, W., Lei, J., Pan, X. and Yan, N. (2021) 'Structure of human Na v 1.5 reveals the fast inactivation-related segments as a mutational hotspot for the long QT syndrome', *Proceedings of the National Academy of Sciences*, 118(11)

Liebschner, D., Afonine, P. V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Poon, B.K., Prisant, M.G., Read, R.J., Richardson, J.S., Richardson, D.C., Sammito, M.D., Sobolev, O. V., Stockwell, D.H., Terwilliger, T.C., Urzhumtsev, A.G., Videau, L.L., Williams, C.J. and Adams, P.D. (2019) 'Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix', *Acta Crystallographica Section D Structural Biology*, 75(10), pp. 861–877

Llaguno, M.C., Xu, H., Shi, L., Huang, N., Zhang, H., Liu, Q. and Jiang, Q.-X. (2014) 'Chemically functionalized carbon films for single molecule imaging', *Journal of Structural Biology*, 185(3), pp. 405–417

Luo, Y., Wan, G., Zhang, X., Zhou, X., Wang, Q., Fan, J., Cai, H., Ma, L., Wu, H., Qu, Q., Cong, Y., Zhao, Y. and Li, D. (2021) 'Cryo-EM study of patched in lipid nanodisc suggests a structural basis for its clustering in caveolae', *Structure*, 29(11), pp. 1286-1294.e6

Marques, M.A., Purdy, M.D. and Yeager, M. (2019) 'CryoEM maps are full of potential',

Current Opinion in Structural Biology, 58, pp. 214–223

McCusker, E.C., Bagnéris, C., Naylor, C.E., Cole, A.R., D'Avanzo, N., Nichols, C.G. and Wallace, B.A. (2012) 'Structure of a bacterial voltage-gated sodium channel pore reveals mechanisms of opening and closing', *Nature Communications*, 3(1), p. 1102

McCusker, E.C., D'Avanzo, N., Nichols, C.G. and Wallace, B.A. (2011) 'Simplified Bacterial "Pore" Channel Provides Insight into the Assembly, Stability, and Structure of Sodium Channels', *Journal of Biological Chemistry*, 286(18), pp. 16386–16391

McDonnell, A., Collins, S., Ali, Z., Iavarone, L., Surujbally, R., Kirby, S. and Butt, R.P. (2018) 'Efficacy of the Nav1.7 blocker PF-05089771 in a randomised, placebo-controlled, double-blind clinical study in subjects with painful diabetic peripheral neuropathy', *Pain*, 159(8), pp. 1465–1476

McMullan, G., Vinothkumar, K.R. and Henderson, R. (2015) 'Thon rings from amorphous ice and implications of beam-induced Brownian motion in single particle electron cryo-microscopy', *Ultramicroscopy*, 158, pp. 26–32

McPhee, J.C., Ragsdale, D.S., Scheuer, T. and Catterall, W.A. (1994) 'A mutation in segment IVS6 disrupts fast inactivation of sodium channels.', *Proceedings of the National Academy of Sciences*, 91(25), pp. 12346–12350

Meyerson, J.R., Rao, P., Kumar, J., Chittori, S., Banerjee, S., Pierson, J., Mayer, M.L. and Subramaniam, S. (2014) 'Self-assembled monolayers improve protein distribution on holey carbon cryo-EM supports', *Scientific Reports*, 4(1), p. 7084

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. and Steinegger, M. (2022) 'ColabFold: making protein folding accessible to all', *Nature Methods*, 19(6), pp. 679–682

Moran, Y., Barzilai, M.G., Liebeskind, B.J. and Zakon, H.H. (2015) 'Evolution of voltage-gated ion channels at the emergence of Metazoa', *The Journal of Experimental Biology*, 218(4), pp. 515–525

Mühleip, A., Flygaard, R.K., Baradaran, R., Haapanen, O., Gruhl, T., Tobiasson, V., Maréchal, A., Sharma, V. and Amunts, A. (2023) 'Structural basis of mitochondrial membrane bending by the I–II–III2–IV2 supercomplex', *Nature*, 615(7954), pp. 934–938

Mullard, A. (2022) 'Vertex's NaV1.8 inhibitor passes phase II pain point', *Nature*

Reviews Drug Discovery, 21(5), pp. 327–327

Nakanishi, H., Hayashida, K., Nishizawa, T., Oshima, A. and Abe, K. (2022) 'Cryo-EM of the ATP11C flippase reconstituted in Nanodiscs shows a distended phospholipid bilayer inner membrane around transmembrane helix 2', *Journal of Biological Chemistry*, 298(1), p. 101498

Nasr, M.L. and Wagner, G. (2018) 'Covalently circularized nanodiscs; challenges and applications', *Current Opinion in Structural Biology*, 51, pp. 129–134

Naydenova, K., Jia, P. and Russo, C.J. (2020) 'Cryo-EM with sub–1 Å specimen movement', *Science*, 370(6513), pp. 223–226

Naydenova, K., McMullan, G., Peet, M.J., Lee, Y., Edwards, P.C., Chen, S., Leahy, E., Scotcher, S., Henderson, R. and Russo, C.J. (2019) 'CryoEM at 100 keV: a demonstration and prospects', *IUCrJ*, 6(6), pp. 1086–1098

Naydenova, K., Peet, M.J. and Russo, C.J. (2019) 'Multifunctional graphene supports for electron cryomicroscopy', *Proceedings of the National Academy of Sciences*, 116(24), pp. 11718–11724

Naylor, C.E., Bagnéris, C., DeCaen, P.G., Sula, A., Scaglione, A., Clapham, D.E. and Wallace, B. (2016) 'Molecular basis of ion permeability in a voltage-gated sodium channel', *The EMBO Journal*, 35(8), pp. 820–830

Needleman, S.B. and Wunsch, C.D. (1970) 'A general method applicable to the search for similarities in the amino acid sequence of two proteins', *Journal of Molecular Biology*, 48(3), pp. 443–453

Noble, A.J., Wei, H., Dandey, V.P., Zhang, Z., Tan, Y.Z., Potter, C.S. and Carragher, B. (2018) 'Reducing effects of particle adsorption to the air–water interface in cryo-EM', *Nature Methods*, 15(10), pp. 793–795

Nurani, G., Radford, M., Charalambous, K., O'Reilly, A.O., Cronin, N.B., Haque, S. and Wallace, B.A. (2008) 'Tetrameric Bacterial Sodium Channels: Characterization of Structure, Stability, and Drug Binding', *Biochemistry*, 47(31), pp. 8114–8121

Oelstrom, K., Goldschen-Ohm, M.P., Holmgren, M. and Chanda, B. (2014) 'Evolutionarily conserved intracellular gate of voltage-dependent sodium channels', *Nature Communications*, 5(1), p. 3420

Olek, M., Cowtan, K., Webb, D., Chaban, Y. and Zhang, P. (2022) 'IceBreaker: Software for high-resolution single-particle cryo-EM with non-uniform ice', *Structure*, 30(4), pp. 522-531.e4

Orlova, E. V. and Saibil, H.R. (2011) 'Structural Analysis of Macromolecular Assemblies by Electron Microscopy', *Chemical Reviews*, 111(12), pp. 7710–7748

Padmanabha Das, K.M., Shih, W.M., Wagner, G. and Nasr, M.L. (2020) 'Large Nanodiscs: A Potential Game Changer in Structural Biology of Membrane Protein Complexes and Virus Entry', *Frontiers in Bioengineering and Biotechnology*, 8

Palovcak, E., Wang, F., Zheng, S.Q., Yu, Z., Li, S., Betegon, M., Bulkley, D., Agard, D.A. and Cheng, Y. (2018) 'A simple and robust procedure for preparing graphene-oxide cryo-EM grids', *Journal of Structural Biology*, 204(1), pp. 80–84

Pan, X., Li, Z., Huang, X., Huang, G., Gao, S., Shen, H., Liu, L., Lei, J. and Yan, N. (2019) 'Molecular basis for pore blockade of human Na + channel Na v 1.2 by the μ-conotoxin KIIIA', *Science*, 363(6433), pp. 1309–1313

Pan, X., Li, Z., Jin, X., Zhao, Y., Huang, G., Huang, X., Shen, Z., Cao, Y., Dong, M., Lei, J. and Yan, N. (2021) '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(11)

Pan, X., Li, Z., Zhou, Q., Shen, H., Wu, K., Huang, X., Chen, J., Zhang, J., Zhu, X., Lei, J., Xiong, W., Gong, H., Xiao, B. and Yan, N. (2018) 'Structure of the human voltage-gated sodium channel Na v 1.4 in complex with β1', *Science*, 362(6412)

Passmore, L.A. and Russo, C.J. (2016) 'Specimen Preparation for High-Resolution Cryo-EM', in R.A. Crowther (ed.) *Methods in Enzymology*. Academic Press, pp. 51–86

Payandeh, J., Scheuer, T., Zheng, N. and Catterall, W.A. (2011) 'The crystal structure of a voltage-gated sodium channel', *Nature*, 475(7356), pp. 353–358

Perkins, D.N., Pappin, D.J.C., Creasy, D.M. and Cottrell, J.S. (1999) 'Probability-based protein identification by searching sequence databases using mass spectrometry data', *Electrophoresis*, 20(18), pp. 3551–3567

Perutz, M.F., Rossmann, M.G., Cullis, A.F., Muirhead, H., Will, G. and North, A.C.T. (1960) 'Structure of Hæmoglobin: A three-dimensional fourier synthesis at 5.5-.

resolution, obtained by X-ray analysis', *Nature*, 185(4711), pp. 416–422

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. (2021) 'UCSF ChimeraX: Structure visualization for researchers, educators, and developers', *Protein Science*, 30(1), pp. 70–82

Pocanschi, C.L., Dahmane, T., Gohon, Y., Rappaport, F., Apell, H.-J., Kleinschmidt, J.H. and Popot, J.-L. (2006) 'Amphipathic Polymers: Tools To Fold Integral Membrane Proteins to Their Active Form', *Biochemistry*, 45(47), pp. 13954–13961

Portolano, N., Watson, P.J., Fairall, L., Millard, C.J., Milano, C.P., Song, Y., Cowley, S.M. and Schwabe, J.W.R. (2014) 'Recombinant Protein Expression for Structural Biology in HEK 293F Suspension Cells: A Novel and Accessible Approach', *Journal of Visualized Experiments*, (92), p. e51897

Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. (2017) 'cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination', *Nature Methods*, 14(3), pp. 290–296

Punjani, A., Zhang, H. and Fleet, D.J. (2020) 'Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction', *Nature Methods*, 17(12), pp. 1214–1221

Quan, C., Mok, W.M. and Wang, G.K. (1996) 'Use-dependent inhibition of Na+ currents by benzocaine homologs', *Biophysical Journal*, 70(1), pp. 194–201

Ragsdale, D.S., McPhee, J.C., Scheuer, T. and Catterall, W.A. (1996) 'Common molecular determinants of local anesthetic, antiarrhythmic, and anticonvulsant block of voltage-gated Na+ channels.', *Proceedings of the National Academy of Sciences*, 93(17), pp. 9270–9275

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G. and Deber, C.M. (2009) 'Detergent binding explains anomalous SDS-PAGE migration of membrane proteins', *Proceedings of the National Academy of Sciences of the United States of America*, 106(6), pp. 1760–1765

Ravelli, R.B.G., Nijpels, F.J.T., Henderikx, R.J.M., Weissenberger, G., Thewessem, S., Gijsbers, A., Beulen, B.W.A.M.M., López-Iglesias, C. and Peters, P.J. (2020) 'Cryo-EM structures from sub-nl volumes using pin-printing and jet vitrification', *Nature*

Communications, 11(1), p. 2563

Ren, D., Navarro, B., Xu, H., Yue, L., Shi, Q. and Clapham, D.E. (2001) 'A Prokaryotic Voltage-Gated Sodium Channel', *Science*, 294(5550), pp. 2372–2375

Rice, P., Longden, L. and Bleasby, A. (2000) 'EMBOSS: The European Molecular Biology Open Software Suite', *Trends in Genetics*, 16(6), pp. 276–277

Rice, W.J., Cheng, A., Noble, A.J., Eng, E.T., Kim, L.Y., Carragher, B. and Potter, C.S. (2018) 'Routine determination of ice thickness for cryo-EM grids', *Journal of Structural Biology*, 204(1), pp. 38–44

Rigaud, J.-L., Mosser, G., Lacapere, J.-J., Olofsson, A., Levy, D. and Ranck, J.-L. (1997) 'Bio-Beads: An Efficient Strategy for Two-Dimensional Crystallization of Membrane Proteins', *Journal of Structural Biology*, 118(3), pp. 226–235

Roberts, P.T.E., Chapman, J.N. and MacLeod, A.M. (1982) 'A CCD-based image recording system for the CTEM', *Ultramicroscopy*, 8(4), pp. 385–396

Rohou, A. and Grigorieff, N. (2015) 'CTFFIND4: Fast and accurate defocus estimation from electron micrographs', *Journal of Structural Biology*, 192(2), pp. 216–221

De Rosier, D.J. and Klug, A. (1968) 'Reconstruction of three dimensional structures from electron micrographs', *Nature*, 217(5124), pp. 130–134

Ruehle, M.D., Orias, E. and Pearson, C.G. (2016) 'Tetrahymena as a Unicellular Model Eukaryote: Genetic and Genomic Tools', *Genetics*, 203(2), pp. 649–665

Russo, C.J. and Passmore, L.A. (2016) 'Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens', *Journal of Structural Biology*, 193(1), pp. 33–44

Sait, L.G., Sula, A., Ghovanloo, M.R., Hollingworth, D., Ruben, P.C. and Wallace, B.A. (2020) 'Cannabidiol interactions with voltage-gated sodium channels', *eLife*, 9, p. e58593

Scheres, S.H.W. (2012) 'RELION: Implementation of a Bayesian approach to cryo-EM structure determination', *Journal of Structural Biology*, 180(3), pp. 519–530

Scheres, S.H.W. and Chen, S. (2012) 'Prevention of overfitting in cryo-EM structure determination', *Nature Methods*, 9(9), pp. 853–854

Schmidt, T.G. and Skerra, A. (2007) 'The Strep-tag system for one-step purification and

high-affinity detection or capturing of proteins', *Nature Protocols*, 2(6), pp. 1528–1535

Semchonok, D.A., Mondal, J., Cooper, C.J., Schlum, K., Li, M., Amin, M., Sorzano, C.O.S., Ramírez-Aportela, E., Kastritis, P.L., Boekema, E.J., Guskov, A. and Bruce, B.D. (2022) 'Cryo-EM structure of a tetrameric photosystem I from Chroococcidiopsis TS-821, a thermophilic, unicellular, non-heterocyst-forming cyanobacterium', *Plant Communications*, 3(1), p. 100248

Shen, H., Li, Z., Jiang, Y., Pan, X., Wu, J., Cristofori-Armstrong, B., Smith, J.J., Chin, Y.K.Y., Lei, J., Zhou, Q., King, G.F. and Yan, N. (2018) 'Structural basis for the modulation of voltage-gated sodium channels by animal toxins', *Science*, 362(6412)

Shen, H., Liu, D., Wu, K., Lei, J. and Yan, N. (2019) 'Structures of human Na v 1.7 channel in complex with auxiliary subunits and animal toxins', *Science*, 363(6433), pp. 1303–1308

Shinzawa-Itoh, K., Shimomura, H., Yanagisawa, S., Shimada, S., Takahashi, R., Oosaki, M., Ogura, T. and Tsukihara, T. (2016) 'Purification of Active Respiratory Supercomplex from Bovine Heart Mitochondria Enables Functional Studies', *Journal of Biological Chemistry*, 291(8), pp. 4178–4184

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. and Higgins, D.G. (2011) 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Molecular Systems Biology*, 7(1), p. 539

Sigworth, F.J. (1998) 'A Maximum-Likelihood Approach to Single-Particle Image Refinement', *Journal of Structural Biology*, 122(3), pp. 328–339

Singer, A. and Sigworth, F.J. (2020) 'Computational Methods for Single-Particle Electron Cryomicroscopy', *Annual Review of Biomedical Data Science*, 3(1), pp. 163–190

Smart, O.S., Neduvelil, J.G., Wang, X., Wallace, B.A. and Sansom, M.S.P. (1996) 'HOLE: A program for the analysis of the pore dimensions of ion channel structural models', *Journal of Molecular Graphics*, 14(6), pp. 354–360

Snijder, J., Borst, A.J., Dosey, A., Walls, A.C., Burrell, A., Reddy, V.S., Kollman, J.M. and Veesler, D. (2017) 'Vitrification after multiple rounds of sample application and

blotting improves particle density on cryo-electron microscopy grids', *Journal of Structural Biology*, 198(1), pp. 38–42

Stagg, S.M., Noble, A.J., Spilman, M. and Chapman, M.S. (2014) 'ResLog plots as an empirical metric of the quality of cryo-EM reconstructions', *Journal of Structural Biology*, 185(3), pp. 418–426

Starmer, C.F., Grant, A.O. and Strauss, H.C. (1984) 'Mechanisms of use-dependent block of sodium channels in excitable membranes by local anesthetics', *Biophysical Journal*, 46(1), pp. 15–27

Sula, A., Booker, J., Ng, L.C.T., Naylor, C.E., DeCaen, P.G. and Wallace, B.A. (2017) 'The complete structure of an activated open sodium channel', *Nature Communications*, 8, p. 14205

Sula, A., Hollingworth, D., Ng, L.C.T., Larmore, M., DeCaen, P.G. and Wallace, B.A. (2021) 'A tamoxifen receptor within a voltage-gated sodium channel', *Molecular Cell*, 81(6), pp. 1160-1169.e5

Sula, A. and Wallace, B.A. (2017) 'Interpreting the functional role of a novel interaction motif in prokaryotic sodium channels', *Journal of General Physiology*, 149(6), pp. 613–622

Swanwick, R.S., Pristerá, A. and Okuse, K. (2010) 'The trafficking of NaV1.8', *Neuroscience Letters*, 486(2), pp. 78–83

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I. and Ludtke, S.J. (2007) 'EMAN2: An extensible image processing suite for electron microscopy', *Journal of Structural Biology*, 157(1), pp. 38–46

Taylor, K.A. and Glaeser, R.M. (1976) 'Electron microscopy of frozen hydrated biological specimens', *Journal of Ultrastructure Research*, 55(3), pp. 448–456

Tribet, C., Audebert, R. and Popot, J.-L. (1996) 'Amphipols: Polymers that keep membrane proteins soluble in aqueous solutions', *Proceedings of the National Academy of Sciences*, 93(26), pp. 15047–15050

Ulmschneider, M.B., Bagnéris, C., McCusker, E.C., DeCaen, P.G., Delling, M., Clapham, D.E., Ulmschneider, J.P. and Wallace, B.A. (2013) 'Molecular dynamics of ion transport through the open conformation of a bacterial voltage-gated sodium channel',

Proceedings of the National Academy of Sciences, 110(16), pp. 6364–6369

Unger, L., Ronco-Campaña, A., Kitchen, P., Bill, R.M. and Rothnie, A.J. (2021) 'Biological insights from SMA-extracted proteins', *Biochemical Society Transactions*, 49(3), pp. 1349–1359

Vilin, Y.Y. and Ruben, P.C. (2001) 'Slow Inactivation in Voltage-Gated Sodium Channels: Molecular Substrates and Contributions to Channelopathies', *Cell Biochemistry and Biophysics*, 35(2), pp. 171–190

Vogel, R.H., Provencher, S.W., von Bonsdorff, C.-H., Adrian, M. and Dubochet, J. (1986) 'Envelope structure of Semliki Forest virus reconstructed from cryo-electron micrographs', *Nature*, 320(6062), pp. 533–535

Wang, J. and Moore, P.B. (2017) 'On the interpretation of electron microscopic maps of biological macromolecules', *Protein Science*, 26(1), pp. 122–129

Watson, J.D. and Crick, F.H.C. (1953) 'Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid', *Nature*, 171(4356), pp. 737–738

Weber, K. and Osborn, M. (1969) 'The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis.', *The Journal of Biological Chemistry*, 244(16), pp. 4406–12

Wei, H., Dandey, V., Zhang, Z., Raczkowski, A., Carragher, B. and Potter, C.S. (2017) 'Self-Blotting Nanowire Grids for Cryo-EM Sample Preparation', *Microscopy and Microanalysis*, 23(S1), pp. 848–849

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., Jain, S., Lewis, S.M., Arendall, W.B., Snoeyink, J., Adams, P.D., Lovell, S.C., Richardson, J.S. and Richardson, D.C. (2018) 'MolProbity: More and better reference data for improved all-atom structure validation', *Protein Science*, 27(1), pp. 293–315

Wisedchaisri, G., Gamal El-Din, T.M., Zheng, N. and Catterall, W.A. (2023) 'Structural basis for severe pain caused by mutations in the S4-S5 linkers of voltage-gated sodium channel Na V 1.7', *Proceedings of the National Academy of Sciences*, 120(14), p. e2219624120

Wisedchaisri, G., Tonggu, L., Gamal El-Din, T.M., McCord, E., Zheng, N. and Catterall,

W.A. (2021) ‘Structural Basis for High-Affinity Trapping of the NaV1.7 Channel in Its Resting State by Tarantula Toxin’, *Molecular Cell*, 81(1), pp. 38-48.e4

Wisedchaisri, G., Tonggu, L., McCord, E., Gamal El-Din, T.M., Wang, L., Zheng, N. and Catterall, W.A. (2019) ‘Resting-State Structure and Gating Mechanism of a Voltage-Gated Sodium Channel’, *Cell*, 178(4), pp. 993-1003.e12

Wu, Q., Huang, J., Fan, X., Wang, K., Jin, X., Huang, G., Li, J., Pan, X. and Yan, N. (2023) ‘Structural mapping of Nav1.7 antagonists’, *Nature Communications*, 14(1), p. 3224

Wu, X. and Rapoport, T.A. (2021) ‘Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies)’, *Proceedings of the National Academy of Sciences*, 118(41), p. e2115001118

Xiao, Y., Barbosa, C., Pei, Z., Xie, W., Strong, J.A., Zhang, J.M. and Cummins, T.R. (2019) ‘Increased resurgent sodium currents in nav1.8 contribute to nociceptive sensory neuron hyperexcitability associated with peripheral neuropathies’, *Journal of Neuroscience*, 39(8), pp. 0468–18

Xu, H., Li, T., Rohou, A., Arthur, C.P., Tzakoniati, F., Wong, E., Estevez, A., Kugel, C., Franke, Y., Chen, J., Ciferri, C., Hackos, D.H., Koth, C.M. and Payandeh, J. (2019) ‘Structural Basis of Nav1.7 Inhibition by a Gating-Modifier Spider Toxin’, *Cell*, 176(4), pp. 702-715.e14

Yan, Z., Zhou, Q., Wang, L., Wu, J., Zhao, Y., Huang, G., Peng, W., Shen, H., Lei, J. and Yan, N. (2017) ‘Structure of the Nav1.4-β1 Complex from Electric Eel’, *Cell*, 170(3), pp. 470-482.e11

Yang, Z., Wang, C., Zhou, Q., An, J., Hildebrandt, E., Aleksandrov, L.A., Kappes, J.C., DeLucas, L.J., Riordan, J.R., Urbatsch, I.L., Hunt, J.F. and Brouillette, C.G. (2014) ‘Membrane protein stability can be compromised by detergent interactions with the extramembranous soluble domains’, *Protein Science*, 23(6), pp. 769–789

Yu, F.H., Yarov-Yarovoy, V., Gutman, G.A. and Catterall, W.A. (2005) ‘Overview of Molecular Relationships in the Voltage-Gated Ion Channel Superfamily’, *Pharmacological Reviews*, 57(4), pp. 387–395

Yu, G., Li, K. and Jiang, W. (2016) ‘Antibody-based affinity cryo-EM grid’, *Methods*,

100, pp. 16–24

Zanatta, G., Sula, A., Miles, A.J., Ng, L.C.T., Torella, R., Pryde, D.C., DeCaen, P.G. and Wallace, B.A. (2019) 'Valproic acid interactions with the NavMs voltage-gated sodium channel', *Proceedings of the National Academy of Sciences*, 116(52), pp. 26549–26554

Zhang, J., Shi, Y., Huang, Z., Li, Y., Yang, B., Gong, J. and Jiang, D. (2022) 'Structural basis for NaV1.7 inhibition by pore blockers', *Nature Structural and Molecular Biology*, 29(12), pp. 1208–1216

Zhang, K. (2016) 'Gctf: Real-time CTF determination and correction', *Journal of Structural Biology*, 193(1), pp. 1–12

Zhang, X., Ren, W., DeCaen, P., Yan, C., Tao, X., Tang, L., Wang, Jingjing, Hasegawa, K., Kumasaka, T., He, J., Wang, Jiawei, Clapham, D.E. and Yan, N. (2012) 'Crystal structure of an orthologue of the NaChBac voltage-gated sodium channel', *Nature*, 486(7401), pp. 130–134

Zheng, L., Liu, N., Gao, Xiaoyin, Zhu, W., Liu, K., Wu, C., Yan, R., Zhang, J., Gao, Xin, Yao, Y., Deng, B., Xu, J., Lu, Y., Liu, Z., Li, M., Wei, X., Wang, H.-W. and Peng, H. (2023) 'Uniform thin ice on ultraflat graphene for high-resolution cryo-EM', *Nature Methods*, 20(1), pp. 123–130

Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y. and Agard, D.A. (2017) 'MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy', *Nature Methods*, 14(4), pp. 331–332

Zhong, E.D., Bepler, T., Berger, B. and Davis, J.H. (2021) 'CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks', *Nature Methods*, 18(2), pp. 176–185

Zhou, F.X., Cocco, M.J., Russ, W.P., Brunger, A.T. and Engelman, D.M. (2000) 'Interhelical hydrogen bonding drives strong interactions in membrane proteins', *Nature Structural & Molecular Biology*, 7, pp. 154–160

Zoonens, M. and Popot, J.-L. (2014) 'Amphipols for Each Season', *The Journal of Membrane Biology*, 247(9–10), pp. 759–796

Appendix

Processing data	
Symmetry imposed	C4
Particles in final reconstruction	104346
Map resolution (Å)	3.07
FSC threshold	0.143
Map sharpening *B* factor (Å^2)	-190
Model composition	
Non-hydrogen atoms	8829
Protein residues	1064
Ligands	9
B factors (Å^2)	
Protein	150.06
Ligands	144.63
R.M.S deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.9
Validation	
MolProbity score	1.67
Clashscore	14.68
Poor rotamers (%)	0
Ramachandran plot	
Favoured (%)	98.48
Allowed (%)	1.52
Disallowed (%)	0
Model	
CC_mask	0.81
CC_volume	0.81
Mean CC for ligands	0.45
Map-model FSC at CC 0.5 (Å)	3.4

Appendix 1. Cryo-electron microscopy map and refined model statistics. Statistics of the map used to produce the atomic model of NavMs in nanodiscs, and the resultant model. Model statistics taken from Phenix validation (Liebschner *et al.*, 2019).