Structural and functional studies of cannabidiol interactions with voltage-gated sodium channels

https://eprints.bbk.ac.uk/id/eprint/48458/

Version: Full Version

Citation: Goodyer Sait, Lily Jane (2022) Structural and functional studies of cannabidiol interactions with voltage-gated sodium channels. [Thesis] (Unpublished)

Structural and Functional Studies of Cannabidiol Interactions with Voltage-gated Sodium Channels

Thesis submitted by Lily Goodyer Sait

to Birkbeck University

for the degree of Doctor of Philosophy

September 2021

Institute of Molecular and Structural Biology, Department of Biological Sciences,
Birkbeck, University of London, Malet Street, London, WC1 7HX

Acknowledgments and Dedications

The UCL-Birkbeck MRC Doctoral Training Programme provided funding for this project.

I would like to first thank Prof. Bonnie Wallace for her continual support throughout this time, especially since she continued to be a superb supervisor throughout the uncertain and frustrating times that arose from the Covid-19 pandemic. I would also like to extend my gratitude to Dr. Altin Sula who has offered his unwavering tutelage and support throughout my time at Birkbeck. In addition he acted as co-first author on (Sait *et al.*, 2020). I would also like to thank the other members of Prof. Bonnie Wallace's lab, who were always there both to offer scientific advice and to have a nice chat when things were challenging. I would also like to thank Dr. Claire Bagnéris who was the best lab manager throughout this entire time, despite having to contend with a global pandemic.

I would like to extend my thanks to Prof. Peter Ruben and Dr. Mohammad Reza Ghovanloo, who collaborated with us and carried out electrophysiology experiments, which complimented the structural work done, with Dr. Mohammad Reza acting as a co-first author on (Sait *et al.*, 2020).

I would like to first dedicate this work to my parents, Ian Goodyer and Georgina Sait. You have always believed in me and supported whatever I wanted to aim for; for that I am forever grateful. Second, I would like to dedicate this work to my grandmother, Jane Sait. I have missed you so much over the last year in which we could not meet due to Covid. However, may the submission of this thesis mark a time where we can finally see each other again and go on our adventures. Finally, I also dedicate this work to my partner Edward Wise, who offered constant support, and always had faith in my ability to do this research. Love you all.

Abstract

Cannabidiol (CBD), is a major non-psychoactive compound isolated from the cannabis plant and has been associated with the treatment of a range of conditions which are often related to voltage-gated sodium ion channels (VGSCs). The aim of this research was to use X-ray crystallography to provide a detailed insight into the interactions which occur between CBD and the prokaryotic VGSC NavMs. CBD was found to bind at a novel site deep within the fenestration of NavMs, near the central hydrophobic cavity (Sait *et al*., 2020). Binding at this site would block sodium ion translocation, thus providing a mechanistic explanation for CBD's channel inhibitory effects, which were validated via electrophysiology experiments performed on designed mutants in collaborative studies with the Ruben lab (Simon Fraser University). In addition, modelling studies conducted suggested why the closely related psychoactive compound tetrahydrocannabinol (THC) may have different binding interactions with VGSCs. Comparisons were also made between the proposed Transient Receptor Potential Cation Subfamily V member 2 (TRPV2) channel CBD binding site and the NavMs binding site. Finally, thermal melt circular dichroism spectroscopic experiments were carried out to explore CBD interactions with NavMs, which showed CBD does not affect NavMs stability during interaction. In summary, this study provides, for the first time, an insight into the possible mechanism underlying CBD interactions with sodium channels.

Table of Contents

List of Figures

List of Chapter Tables

Abbreviations

Word	Abbreviation
2-arachidonoyl glycerol	2-AG
3-(N-morpholino)propanesulfonic acid	MOPS
3,3′-Diaminobenzidine	DAB
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid	HEPES
Activation gate	AG
Brief Psychiatric Rating Scale	BPRS
Caesium chloride	CsCl
Caesium hydroxide	CsOH
Caesium fluoride	CsF
Calcium chloride	$CaCl_2$
Cannabidiol	CBD
Cannabinoid Receptor 1	CB1
Cannabinoid Receptor 2	CB2
Charged-coupled device	CCD
Circular dichroism	CD
Complementary DNA	cDNA
Cryogenic electron microscopy	Cryo-EM
Dalton	Da
Decanoyl-N-Hydroxyethylglucamide	HEGA-10
Deoxyribonucleic acid	DNA
Deoxyribonucleotide triphosphate	dNTPs
Dimethyl sulfoxide	DMSO
Dorsal root ganglion	DRG
Enhanced green fluorescent protein	eGFP
Epilepsy Therapy Screening Program	ETSP
Escherichia coli	E. coli
Ethylenediaminetetraacetic acid	EDTA
Fast protein liquid chromatography	FPLC
Gamma-Aminobutyric acid	GABA
Generalised epilepsy with febrile seizures plus	GEFS+
Glyco diosgenin	GDN
Gravitational force	g
High-tension voltage	HT
Inactivation gate	IG
Isopropyl β-D-1-thiogalactopyranoside	IPTG
Kilo Dalton	kDa
Kilohertz	kHz
Lysogeny broth	LB
Magnesium chloride	$MgCl_2$

Word	**Abbreviation**
Magnesium sulfate	$MgSO_4$
Mean residue ellipticity	MRE
Mean residue weight	MRW
Megaohm	MΩ
Messenger ribonucleic acid	mRNA
Milli-absorbance unit	mAU
Milli volts	mV
Movement Disorder Society-Unified Parkinson’s Disease Rating scale	MDS-UPDRS
Multiple sequence alignments	MSA
N-arachidonoylethanolamine	Anandamide
n-Dodecyl β-D-maltopyranoside	DDM
N-terminal hexa-histidine tag	His-tag
National Institute of Neurological Disorders and Stroke	NINDS
Normalised root-mean-square deviation	NRMSD
Nuclear magnetic resonance	NMR
Optical density	OD
Pairwise sequence alignments	PSA
Phenylmethylsulfonyl fluoride	PMSF
Phosphate-buffered saline	PBS
Polyethylene glycol	PEG
Polysorbate-20	Tween-20
Polyvinylidene difluoride	PVDF
Pore domain	PD
Positive and Negative Syndrome Scale	PANSS
Pounds per square inch	psi
Protein circular dichroism data bank	PCDDB
Protein data bank	PDB
Revolutions per minute	RPM
Root mean square deviation	RMSD
Selectivity filter	SF
Site-directed Ligase-independent Mutagenesis Polymerase Chain Reaction	SLIM PCR
Size Exclusion Chromatography	SEC
Sodium chloride	NaCl
Sodium dodecyl sulfate	SDS
Sodium dodecyl sulfate and polyacrylamide gel	SDS-PAGE
Sodium fluoride	NaF
Standard deviation	SD
Super optimal broth with catabolite repression	SOC
Synchrotron radiation circular dichroism	SRCD
Tetrahydrocannabinol /Δ-9-tetrahydrocannabinol	THC
Tetrahydrocannabinolic	THCA
Tris(hydroxymethyl)aminomethane	Tris base

Word	**Abbreviation**
Tris(hydroxymethyl)aminomethane-acetate-ethylenediaminetetraacetic acid	TAE
Ultraviolet	UV
Volts	V
Voltage-gated sodium ion channel	VGSC
Voltage-sensing domain	VSD
Wild type	WT

Table of word abbreviations.

Chapter 1: Literature Review

Despite the wealth of electrophysiological and functional studies which supply evidence and insight into the interactions between VGSCs and CBD, interactions have not yet been studied or identified on a structural level. This lack of structural work means that interactions between CBD and VGSCs are not well understood on a structural level, resulting in a lack of knowledge of the mechanism/s underlying CBD's apparent action as a sodium ion channel blocker.

1.1 Voltage-gated sodium channels:

1.1.1 Overview

VGSCs are a subset of voltage-gated ions channels, which are transmembrane proteins involved in the electrical signalling of cells. VGSCs specifically allow for the passage of sodium ions across cell membranes, leading to currents which underpin action potential initiation and propagation within nerves and muscle. These action potentials have a range of outcomes; action potentials within cardiac cells lead to contractions of the heart muscle, which enables circulation of blood around the body. In contrast, when an axon of a neuron is sufficiently stimulated by an external heat stimulus, action potentials transmit information across the neuron and through the subsequent nervous system to the brain, where the stimulus can be perceived as harmful. This perception can lead to information being sent out to the site of the heat stimulus to defend the body from the heat stimulus (de Lera Ruiz and Kraus, 2015; Ahern *et al.*, 2016).

There are nine homologous mammalian VGSC subtypes named Nav 1.1-Nav 1.9. Nav 1.1-1.9 have differential expression patterns. Nav 1.1, Nav 1.2 and Nav 1.3 are predominantly expressed within the central nervous system. Nav 1.7, Nav 1.8 and Nav 1.9 are expressed within the peripheral nervous system. In contrast, Nav 1.6 is expressed in both the peripheral and central nervous system, whilst Nav 1.4 and Nav 1.5 are expressed within muscle (de Lera Ruiz and Kraus, 2015).

Looking at mammalian sodium VGSC structures, these channels form a large pseudotetrameric pore α subunit (260 kilo Daltons (kDa)), which can interact with one or two β subunits (30-40 kDa) (de Lera Ruiz and Kraus 2015). The α subunit is made up of four segments which are connected by intracellular linkers. These four segments are known as domains 1-4 (D1-4). D1-4 are homologous in structure, and each are made up of six transmembrane segments (S1-S6) (Fig 1.1) (Ahern *et al.*, 2016). S1-S6 are involved in the formation of the voltage-sensing domain (VSD), pore domain (PD), inactivation gate (IG) and activation gate (AG), which will each be discussed in more detail below.

VSD:
This structure is involved in regulating the opening of the channel upon the detection of membrane depolarisation. S1-S4 make up the VSD in each domain, with the detection of membrane potential being achieved via positively charged amino acid residues along the S4 helix. In response to the detection of membrane depolarisation, the VSDs of each domain move, leading to the opening of the PD, which is discussed below.

PD:
The pore domain is made up of the S5 and S6 segments from each of the four domains, along with the connecting extracellular pore loops. The pore facilitates the movement of sodium ions across the membrane. Within this segment is a structure known as the selectivity filter (SF). The SF in mammalian VGSCs is made up aspartate, glutamate, lysine, and alanine. The SF is the narrowest part of the pore and confers the selectivity of the VGSC for hydrated sodium ions.

IG:
The inactivation gate is an intracellular loop connecting D3 to D4. Due to its positioning, it acts as a ‘hinged lid’ that can act to cover the intracellular side of the pore during inactivation.

AG:
The activation gate is involved in regulating access to the channel’s extracellular pore. This structure is made up of four hydrophobic amino acid residues, with each of these coming from the end of the S6 segments of each of the four domains. These residues form a small intracellular cavity structure.

(de Lera Ruiz and Kraus, 2015)

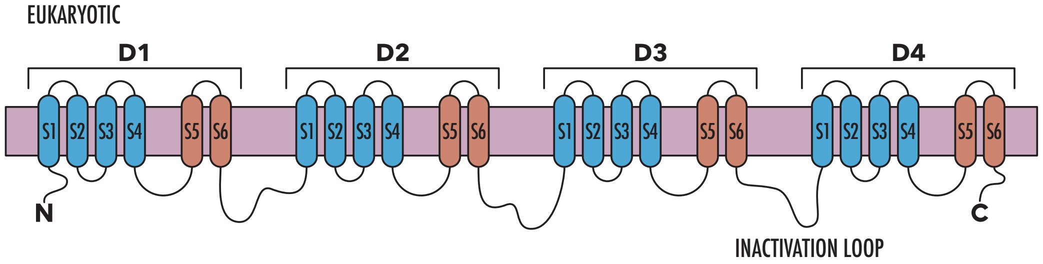


Figure 1.1: Schematic depicting eukaryotic VGSC. *Figure indicates D1-4 of the eukaryotic sodium channel. Each domain consists of six membrane spanning segments. S1-4 form the voltage-sensing domains, whilst the S5-S6 segments form the pore domain. Produced using Affinity Designer (Serif Europe, Nottingham, UK).*

VGSCs are dynamic in structure, occupying three main states: an open activated state, an inactivated state, and a closed resting state. These states are governed by the membrane potential, making the membrane potential vital to VGSC function. At hyperpolarised resting potentials, VGSCs have little chance of existing in the open active state and instead usually occupy their resting state. As the membrane depolarises, this allows for conformational changes to occur within the α subunit, leading to the channel occupying an open activated state, which is facilitated by movements of the voltage-sensing domains. Whilst in an open activated state, the pore is opened, allowing for an inward movement of sodium ions across the membrane. In fast inactivation, open channels will transition from being open and active into an inactivated state within the space of milliseconds, shunting movement of sodium ions across the membrane (Fig 1.2). During slow inactivation, membranes are depolarised for extended periods of time, which can be underpinned by there being incomplete fast inactivation of VGSCs, which leads to persistent currents. Nav 1.4, Nav 1.6 and Nav 1.9 have been shown to give rise to persistent currents (Raman and Bean, 1997; Raman *et al.*, 1997; Baker *et al.*, 2003). In addition, mutations have been found within the genes for Nav 1.1, 1.2, 1.3 and 1.6 which cause defects in inactivation gating, increasing persistent currents, leading to ataxia and epilepsy (Scalmani *et al.*, 2006; Estacion *et al.*, 2010; Volkers *et al.*, 2011; O’Brien and Meisler, 2013).

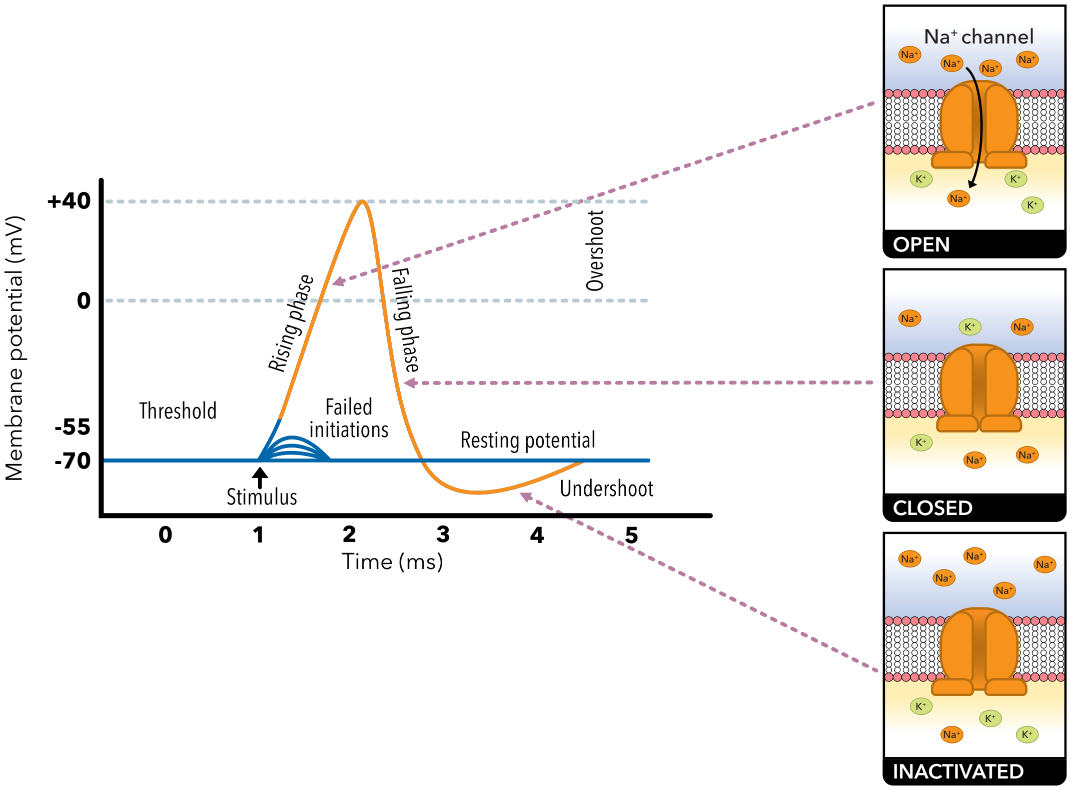


Figure 1.2: Action potential. *Diagram of an action potential, in correlation with the activation, closing and inactivation of VGSCs. Produced using Affinity Designer (Serif Europe, Nottingham, UK).*

1.1.2 Channelopathies involving VGSCs

Because of the extensive expression and presence of VGSCs throughout the human body, mutations within VGSCs have been related to a number of disorders, including epilepsy, pain disorders, neuromuscular disorders and cardiovascular disorders. The association of VGSCs with multiple functions within the body, and consequent association with a variety of disorders, has resulted in VGSCs being a major drug target within medicine.

Mutations within Nav 1.1, Nav 1.2, Nav 1.3 and Nav 1.6 have been associated with epileptic disorders. Nav 1.1 is preferentially expressed within gamma-aminobutyric acid (GABA)ergic interneurons within the brain. These interneurons produce GABA, the major inhibitory neurotransmitter. Missense mutations within Nav 1.1 lead to impairment of GABAergic interneuron excitement, thus impairing GABA production. This causes hyperexcitability within the brain, leading to epilepsy. Mutations within

Nav 1.1 have been associated with the autosomal dominant epilepsy disorders termed generalised epilepsy with febrile seizures plus (GEFS+) (Escayg *et al.*, 2000), and Dravet syndrome (Fukuma *et al.*, 2004), a severe myoclonic epilepsy associated with infancy. Nav 1.2. mutations have been associated with benign familial neonatal-infantile seizures (Heron *et al.*, 2002; Striano *et al.*, 2006). Nav 1.2 is preferentially expressed within excitatory pyramidal neurons, and mutations within Nav 1.2 tend to cause an increase in sodium ion current, thus increasing neuronal hyperexcitability (Scalmani *et al.*, 2006). A mutation within Nav 1.3, which appears to increase persistent current and hyperexcitability within hippocampal neurons, has been associated with cryptogenic partial epilepsy (Estacion *et al.*, 2010). Nav 1.6 is associated with both persistent and resurgent currents. Mutations within Nav 1.6 have been associated with epileptic encephalopathies. These mutations have been characterised as gain-of-function, leading to increased resurgent and/or persistent currents, and causing hyperexcitability within the brain (de Kovel *et al.*, 2014; Patel *et al.*, 2016).

Nav 1.4 is vital within muscle contraction due to its preferential expression within skeletal muscle cells. Mutations with this channel can lead to increased channel activity, which has been associated with hereditary muscular movement disorders, such as hyper- and hypokalemic periodic paralysis, paramyotonia congenital, and congenital myasthenic syndrome (Jurkat-Rott *et al.*, 2010; de Lera Ruiz and Kraus, 2015).

Nav 1.5 is preferentially expressed within cardiac cells. Because of this, mutations within Nav 1.5 have been associated with cardiac conditions like long QT syndrome. Long QT syndrome is defined by mutations which impair Nav 1.5 fast inactivation, leading to the production of persistent currents and prolonging of ventricular action potentials. These prolonged ventricular action potentials cause ventricular arrythmias, which can be fatal (Antzelevitch *et al.*, 2014).

Nav 1.7 has been identified as key within the detection of pain stimuli by sensory neurons. Genetic studies have associated gain-of-function Nav 1.7 mutations to hereditary pain disorders like erythromelalgia, paroxysmal extreme pain disorder and small fibre neuralgia. These gain-of-function mutations appear to cause hyperexcitability within dorsal root ganglion (DRG) neurons (de Lera Ruiz and Kraus, 2015). Loss-of-function mutations within Nav 1.7 have been associated with hereditary insensitivities to pain (Cox *et al.*, 2006).

Nav 1.8 is expressed within the dorsal root ganglia and peripheral nerve axons. Gain-of-function mutations within this channel have been associated with small-fibre neuropathy, a disorder which affects the small nerves in the skin, causing pain or numbness. Functional studies showed that these gain-of-function mutations increased the channel's response to depolarisation, leading to hyperexcitability of the dorsal root ganglia (Faber *et al.*, 2012).

Nav 1.9 is also associated with pain disorders due to its preferential expression within the peripheral nervous system. In one study looking at familial loss of pain disorders, it was found that a gain-of-function mutation of Nav1.9 caused excessive activity at resting voltages, leading to sustained depolarisation of pain receptors, along with impaired action potential generation and abnormal synaptic transmission. These findings suggested that this mutation contributes to the observed loss of pain perception within the study patients (Leipold *et al.*, 2013). In another study, which focused on familial episodic pain, other gain-of-function mutations within Nav 1.9 were identified (Zhang *et al.*, 2013). In both (Leipold *et al.*, 2013; Zhang *et al.*, 2013) it is posited that these gain-of function mutations cause overactivity of Nav 1.9. However, within (Leipold *et al.*, 2013) which looked at loss of pain perception, the excess sodium ion influx at rest caused by Nav 1.9 overactivity may cause other ion channels, such as Nav1.7, Nav1.8 and voltage-gated calcium ion channels in DRG neurons, to progressively inactivate, resulting in an overall conduction block. In comparison, it is posited that the gain-of-function mutation discerned by (Zhang *et al.*, 2013) leads to an overall increase in dorsal root ganglion excitability, causing manifestation of familial episodic pain.

1.2 Evidence for cannabidiol as a treatment within voltage-gated sodium channel related conditions:

1.2.1 Cannabis and phytocannabinoids overview

Cannabis is a genus of flowering plant belonging to the family *Canabaceae,* which also includes the genus *Humulus* (hops), which are mostly known for their use within beer making. Cannabis has an extensive history of use across the globe for medical purposes. The proposed molecules responsible for cannabis' effects within the body are cannabinoids.

Cannabinoids derived from cannabis plants are termed phytocannabinoids, and there are currently 113 phytocannabinoids which have been identified and derived from the cannabis plant (Aizpurua-Olaizola *et al.*, 2016). The two most prevalent phytocannabinoids produced by the cannabis plant are cannabidiol (CBD) and tetrahydrocannabinol (specifically Δ-9-tetrahydrocannabinol) (THC). The structures for both CBD and THC were elucidated within Raphael Mechoulam's lab in 1963 and 1964 respectively (Mechoulam and Shvo, 1963; Gaoni and Mechoulam, 1964). The structures only differ by a single hydroxyl group (Fig 1.3), and yet display very different effects within the body. THC is a psychoactive constituent of the cannabis plant, and is responsible for the effects normally associated with cannabis use, such as altered perception of time, changes in mood and decreases in anxiety levels (Green *et al.*, 2003). THC imparts its psychoactive effects via its action as an agonist to both CB1 and CB2. In comparison, CBD has little affinity for both CB1 and CB2 receptors but has been recorded to act as a weak negative allosteric modulator of these receptors. Because of CBD's actions as a negative allosteric modulator, the potency and efficacy of present orthosteric ligands, such as THC, may be attenuated (Freeman *et al.*, 2019; Hayakawa *et al.,* 2008; Laprairie *et al.*, 2015). This would explain findings which suggest that CBD does indeed counter some of the actions and consequent effects of THC (McPartland and Russo, 2015, pp 280-297).

THC has been shown to have beneficial therapeutic effects against a number of conditions including epilepsy, nausea, vomiting, low appetite, and pain (Carlini, 2004). Unfortunately, THC has been associated with a lot of side effects due to being psychoactive. Side effects include acute impairment of cognitive functions, and the occurrence of unpleasant experiences such as: disconnected thoughts, panic reactions, feelings of depersonalisation, disturbing changes in perception, delusions and hallucinatory experiences (Campbell *et al.*, 2001; Carlini, 2004). CBD has also been evidenced to have beneficial therapeutic effects for conditions such as epilepsy (Stockings *et al.*, 2018), pain disorders (Campbell *et al.*, 2001; Manzanares *et al.*, 2006) and muscle spasticity (Leehey, 2019; Wong and Wilens, 2017), but since CBD has the added benefit of not being psychoactive, this makes it an overall more attractive candidate for medical research and development (Rosenberg, Patra, and Whalley 2017).

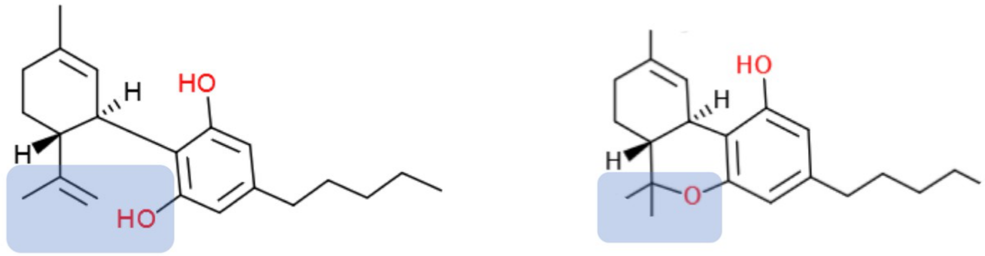


Figure 1.3: CBD and THC (cited from (Sait et al., 2020)). *Molecular structure of CBD (left) and THC (right). Differences in structure highlighted in Blue. Produced using ChemDoodle (Todsen, 2014).*

1.2.2 Preclinical and clinical evidence for medicinal CBD usage

CBD's use within medicine is fuelled by the growing body of preclinical and clinical evidence for its effective use within several conditions, which are often VGSC related. Major examples include epilepsy (Kaplan *et al*., 2016), Parkinson's (Wang *et al*., 2019), schizophrenia (Rees *et al*., 2019) and pain disorders (Dib-Hajj *et al*., 2017). Preclinical and clinical research into these conditions and the use of CBD as a treatment are discussed below.

Pre-clinical evidence: The Epilepsy Therapy Screening Program (ETSP), funded by the National Institute of Neurological Disorders and Stroke (NINDS), has played a vital part in identifying possible new epilepsy treatments. Due to about a third of epilepsy patients still suffering from pharmacoresistant seizures, the screening program was refined to focus on pre-clinical models characterised by pharmacoresistant seizures so as to help identify promising drugs in this field. This re-focused screening program was used to analyse CBD to see what the pre-clinical models indicated about its effectiveness in seizure control (Huizenga *et al*. 2018). Overall, the large body of preclinical work studied in this screening process suggested CBD does possess anti-seizure properties but may be more efficacious against generalised tonic-clonic seizures compared to clonic seizures. In addition, CBD did seem to exert an effect across most models, but its potency did differ across the models. This may be explained by CBD acting at distinct targets within the different models (Klein *et al*. 2017).

Pre-clinical rodent studies have been used to research the possible effectiveness of CBD as a treatment within neurological, neurodegenerative, and psychiatric disorders, such as chronic pain, Parkinson's, and schizophrenia respectively. Several studies have recorded promising effects from CBD when treating pain. In (Ward *et al*., 2014) it was asked whether CBD could help with chemotherapy-induced neuropathic pain, using mouse models exposed to the chemotherapy drug Paclitaxel. This study found that CBD is protective against pain caused by Paclitaxel consumption. In addition, CBD did not attenuate the anti-cancer properties of Paclitaxel, suggesting that CBD may be a safe and effective tool when preventing or reducing chemotherapy-induced pain. The effect of CBD on both neuropathic and inflammatory pain was also studied with the use of rat models. Administration of CBD was shown in both models to reduce hyperalgesia to thermal and mechanical stimuli (Costa *et al*., 2007). In another study, it was examined whether cannabinoids, including CBD, were effective against the toxicity caused by 6-hydroxydopamine, *in vivo* and *in vitro*, which is of relevance to Parkinson's disease. It was discerned that CBD, among other cannabinoids tested, did appear to reduce neurodegeneration caused by 6-hydroxydopamine and restore dopamine levels. The data collected during this study indicated that these neuroprotective effects could be due to the antioxidant properties of CBD (Lastres-Becker *et al*., 2005). In addition, CBD has been evidenced to modulate glial function, which would also add to its apparent neuroprotective effects (Esposito *et al*., 2007). In a study focusing on schizophrenia, rat models displaying schizophrenic-like behavioural alterations, and which fit the neurodevelopmental hypothesis of schizophrenia via prenatal exposure to the antimitotic agent methylazoxymethanol acetate at gestational day 17, underwent peripubertal treatment with CBD (30 mg/kg/day). This treatment reversed the schizophrenic-like phenotype, suggesting that early peripubertal treatment with CBD may prevent the appearance of schizophrenic-like deficits in adulthood in those children identified as at risk of developing schizophrenia (Stark *et al*., 2019).

Clinical trial evidence: In addition to pre-clinical screening, human clinical trials have been coordinated and run to examine the use of CBD as an anti-epileptic. These trials looked at patients with Dravet or Lennox-Gastaut syndrome, both of which are severe genetic forms of epilepsy which become apparent during infancy/childhood. Both aforementioned syndromes are rare early onset epilepsies that often possess a genetic component. Dravet patients often have mutations within the human sodium channel gene SCN1A, which encodes Nav 1.1, whilst those with Lennox-Gastaut syndrome have been recorded with mutations within several genes, including genes encoding human sodium channels (Klein *et al*. 2017, Huizenga *et al*. 2018, Wong and Wilens 2017, Stockings *et al*. 2018). Across all these trials, it was discerned that those using CBD saw a significant decrease in seizure frequency and intensity, with these results also being significantly different to those recorded in the control groups. However, side effects attributed to CBD use were recorded across all of the trials, including somnolence, nausea and temporarily increased aminotransferase enzyme levels within the liver (Cross *et al*., 2017; Devinsky *et al*., 2016, 2017, 2018; Thiele *et al*., 2018). These recordings of negative side-effects were backed up by further trials which focused on the long-term safety of taking CBD as an anti-epileptic (Devinsky *et al*., 2019; Halford *et al*., 2018). Even with identified side-effects, CBD is still favourable when compared to other antiepileptic drugs when treating Dravet and Lennox-Gastaut syndromes, since these disorders are often pharmacoresistant, and other antiepileptics carry more serious side-effects, such as severe mucocutaneous reactions, alopecia, and pancreatitis (Perucca and Gilliam, 2012). In accordance with these clinical trial results, one formulation of CBD (GW Research Ltd., UK) has been recently approved by the European Medicines Agency and the Federal Drug Administration for use in children for the treatment-resistant epilepsies Dravet Syndrome and Lennox-Gastaut Syndrome (Nahar, Onder and Sarker, 2019).

There have also been follow on clinical trials studying the use of CBD within Parkinson's. These trials looked at the effect of CBD on motor, cognitive and psychological symptoms associated with Parkinson's. One recent Phase 1 trial study conducted by Dr. Maureen Leehey looked at the safety and tolerability of CBD when administered to Parkinson's patients. This trial recorded cognitive, anxiety, depression, movement, emotional and sleep assessments, along with observing any changes in the Movement Disorder Society-Unified Parkinson's Disease Rating scale (MDS-UPDRS) tremor score. This study reported decreased scores in the: MDS-UPDRS tremor score, depression short form, REM sleep behaviour disorder screening questionnaire, emotional and behavioural control short form and in the scales for outcomes in Parkinson's disease-sleep-night-time sleep score. The same study found increases in the scores in the: changes in Montreal cognitive assessment and the anxiety short form. These all indicate positive effects of CBD on a range of symptoms associated with Parkinson's (Leehey, 2019; Pauli *et al*., 2020). Another clinical trial studied the effects of CBD on psychosis related to Parkinson's within a small number of patients. This study gives an indicator that CBD appears to significantly decrease psychosis symptoms, whilst not worsening motor symptoms or contributing any significant side effects (Zuardi *et al*., 2009; Pauli *et al*., 2020).

Clinical trials have started with the intent of scrutinizing CBD's use as a treatment within psychiatric disorders (Khoury *et al*., 2019). Of particular interest is the use of CBD as a treatment for schizophrenia. One study involved the performance of a double-blind randomised control trial, to assess whether CBD was non-inferior to amisulpride in the treatment of acute schizophrenia. CBD or amisulpride were administered to participants at 200 mg/day, which was titrated up to 800 mg/day. Symptoms were simultaneously assessed by both the Positive and Negative Syndrome Scale (PANSS) and the Brief Psychiatric Rating Scale (BPRS) during the trial. At the end of this trial there was a comparable reduction of psychotic symptoms between those who were given either CBD or amisulpride, with the CBD treated patients also experiencing less severe adverse effects. In addition, levels of anandamide were increased in the group treated by CBD, which the study goes on to suggest may contribute to the observed antipsychotic effects of CBD (Khoury *et al*., 2019; Leweke *et al*., 2012). In a second double-blind randomised control trial focusing on schizophrenia, CBD was compared to treatment with a placebo. Patients were started on either CBD 600 mg or placebo and half-way through the trial were switched to the other treatment. Using the PANSS scale to assess psychotic symptoms, the study reported significant improvement in positive symptoms when compared to baseline in those treated with CBD (Khoury *et al*., 2019; M. Leweke, 2013). Although these results are promising, it must be noted that in both clinical trials there were only small sample sizes and short durations, meaning that for increased confidence in CBD's use within schizophrenia treatment, larger scale trials would have to be undertaken.

CBD is often reported to be used for pain management (Corroon and Phillips, 2018), but even with this anecdotal evidence, clinical trials into its use as a tool for managing pain are few in number. This being said, there are several trials which have investigated the use of a pharmaceutical product called Nabixmol/Sativex and the use of CBD-rich oils and gels, all of which report these as being promising treatments for the treatment of a range of pain disorders (Blake *et al*., 2006; Cuñetti *et al*., 2018; Hunter *et al*., 2018; Johnson *et al*., 2013; Nurmikko *et al*., 2007; Urits *et al*., 2020). The issue with these trials, however, is that many of these treatments (including Sativex) also contain THC, meaning that the results from the trials cannot be solely attributed to CBD. In addition, many of these other trials that did not use Sativex were small in size, open label, and included treatments containing other cannabinoids other than CBD.

1.3 Research into cannibidiol’s interactions with voltage-gated sodium channels:

As discussed, thanks to preclinical and clinical trial work, CBD appears to be of medical significance within a range of conditions associated, at least in part, with VGSCs.

Within the literature, there is research which indicates that VGSCs may be a direct CBD target. One study showed that CBD at 10 µM significantly reduced action potentials within rat CA1 hippocampal neurons, as well as reducing Nav current density within human blastoma cells and mouse cortical neurons (Hill *et al.*, 2014). In addition, CBD has been shown to bind and inhibit channel activities of human Nav 1.1-1.7, and the prokaryotic Nav homologue NachBac, at physiologically relevant IC50s of 1.9-3.8 µM (Ghovanloo *et al.*, 2018). Functional studies involving human sodium channels and homologous prokaryotic channels have suggested that CBD inhibits VGSC activity via stabilisation of the inactivated state (Patel *et al.*, 2016; Ghovanloo *et al.*, 2018). In addition, CBD has been shown to inhibit both resurgent and persistent sodium currents of human Nav 1.2 at concentrations of 1 µM (Mason and Cummins, 2020). CBD has also been evidenced to inhibit persistent and resurgent currents produced by the human Nav 1.6 channel (Patel *et al.*, 2016). This is significant, since these forms of abnormal current have been related to various epileptic conditions, and shown to be controlled by CBD (de Kovel *et al*., 2014; Mason and Cummins, 2020; Patel *et al*., 2016).

(Le Marois *et al*., 2020) and (Fouda *et al*., 2020) used *in vitro* functional and electrophysiological techniques to investigate CBD and its inhibition of Nav 1.5, which is predominantly expressed within cardiac cells. (Le Marois *et al*., 2020) discerned during their experiments that CBD inhibited Nav 1.5 at an IC50 of 7.4 µM, leading to the conclusion that caution must be exercised when prescribing CBD for treating conditions such as epilepsy, cardiac channelopathies, or to those already taking other drugs known to affect heart rhythm or contractility. In comparison, (Fouda *et al*., 2020) suggest that CBD could be used to protect against high-glucose induced oxidative stress and cytotoxicity within cardiac cells, which is a major cause of mortality within diabetes patients. This conclusion was drawn because the research found that CBD’s inhibition of Nav 1.5 activities confers protection on Nav1.5 against high glucose-elicited hyperexcitability and cytotoxicity. These papers highlight that interaction between VGSCs (Nav 1.5 in this instance) and CBD can be beneficial or possibly harmful depending on the overall situation of the patient, and the condition you are trying to treat.

In a recent follow up paper to (Ghovanloo *et al.*, 2018), the modularity mechanisms of CBD's effects were explored using human Nav 1.4 (Ghovanloo *et al.*, 2021). In this paper they tested the two hypotheses about how CBD exerts its inhibitory effects, which came from (Ghovanloo *et al.*, 2018). The first hypothesis is that CBD acts as a modulator via effects on membrane elasticity. The second hypothesis is that CBD molecule acts as a physical block of the sodium channel pore. First, they performed molecular dynamics simulations which visualised CBD within the membrane, which was then confirmed via nuclear magnetic resonance (NMR) experiments. Next, Gramicidin-based fluorescence assays were carried out which highlighted that CBD does appear to alter and increase membrane elasticity. Afterwards they then performed whole cell patch-clamp experiments with Nav 1.4 F1586A mutant (F1586 is a local anaesthetic binding site), which indicated that removing this site reduced the blockage of the sodium ion current. Computational mutagenesis of four residues to tryptophan D1-F432, D2-V787, D3-1280, and D4-I1583), which would lead to the occlusion of fenestrations, was carried out. Results from this indicate that 10 µM of CBD would no longer be able to block the sodium channel, suggesting that CBD may be entering via the sodium channel fenestrations.

1.4 Structural studies of voltage-gated sodium ion channels and cannabidiol:

Despite the wealth of clinical trials and direct research indicating that CBD acts as an inhibitor of VGSCs, interactions between VGSCs and CBD have not yet been studied on a structural level.

1.4.1 Structural studies of TRPV2 in complex with cannabidiol

The only confirmed structure of CBD bound to a protein target before this thesis' research is that of CBD in complex with the Transient Receptor Potential Cation Channel Subfamily V Member 2 (TRPV2) channel. TRPV2 is an homotetramic ion channel that facilitates the non-specific movement of both sodium and calcium ions through plasma membranes (Qin *et al.*, 2008). TRPV2, like VGSCs, have four domains, with each domain containing six transmembrane segments (S1-S6). S1-S4 forms the VSD whilst S5-S6 go towards forming the pore. TRPV2 differ from VGSCs in structure, however, by the presence of an intracellular ankyrin repeat domain (Fig 1.4). TRPV2 has been associated with epilepsy (Morelli *et al.*, 2013), though less direct than VGSCs' link to epilepsy. TRPV2 has also been associated with pain due to TRPV2 being expressed within medium- to large-diameter Aδ mechanosensitive and thermosensitive neurons in the rat dorsal root ganglion (Caterina *et al.*, 1999), the trigeminal ganglion (Ichikawa and Sugimoto, 2000), and the spinal cord (Lewinter *et al.*, 2004). In addition, TRPV2 has been recorded as being upregulated in rat DRG after nerve injury (Frederick *et al.*, 2007) and inflammation (Shimosato *et al.*, 2005). However, TRPV2 does have an overall wide distribution pattern, which may suggest a broad range of physiological functions (Qin *et al.*, 2008). CBD has been suggested to be an activator of TRPV2 (Qin *et al.*, 2008; Nabissi *et al.*, 2013), and according to electrophysiology studies, CBD activates rat TRPV2 with an EC50 of 3.7 µM (Qin *et al*, 2008). The already mentioned TRPV2-CBD structure was recently elucidated via the use of cryo-EM to a nominal resolution of 3.2 Å (Pumroy *et al*, 2019). The structure indicates the presence of a CBD-binding pocket, which is located between the S5 and S6 helices of adjacent TRPV2 monomers (Pumroy *et al*, 2019) (Figure). However, the resolution of the structure does not allow detailed analysis of its binding site, although it does still indicate that the CBD binding site is lined with mostly hydrophobic and aromatic residues (Fig 1.5), and that binding requires a partial refolding of the adjacent region of the protein polypeptide. Within the TRPV2 structure CBD was oriented with the headgroup on the interior of the pocket and the tail extending into the lipid bilayer since the electron density for the ligand within the interior of the pocket fitted a linear acyl tail poorly (Pumroy *et al*, 2019).

Since TRPV2 is a tetramic ion channel, much like VGSCs, it was interesting to have this structure ready for comparison with VGSC:CBD structures, to observe if there are similarities between the two in how CBD interacts, for instance, does the CBD interact at a similar binding pocket within VGSCs? These possible comparisons were even more interesting since CBD, as already discussed, had been evidenced by previous research to activate TRPV2 but inhibit VGSCs.

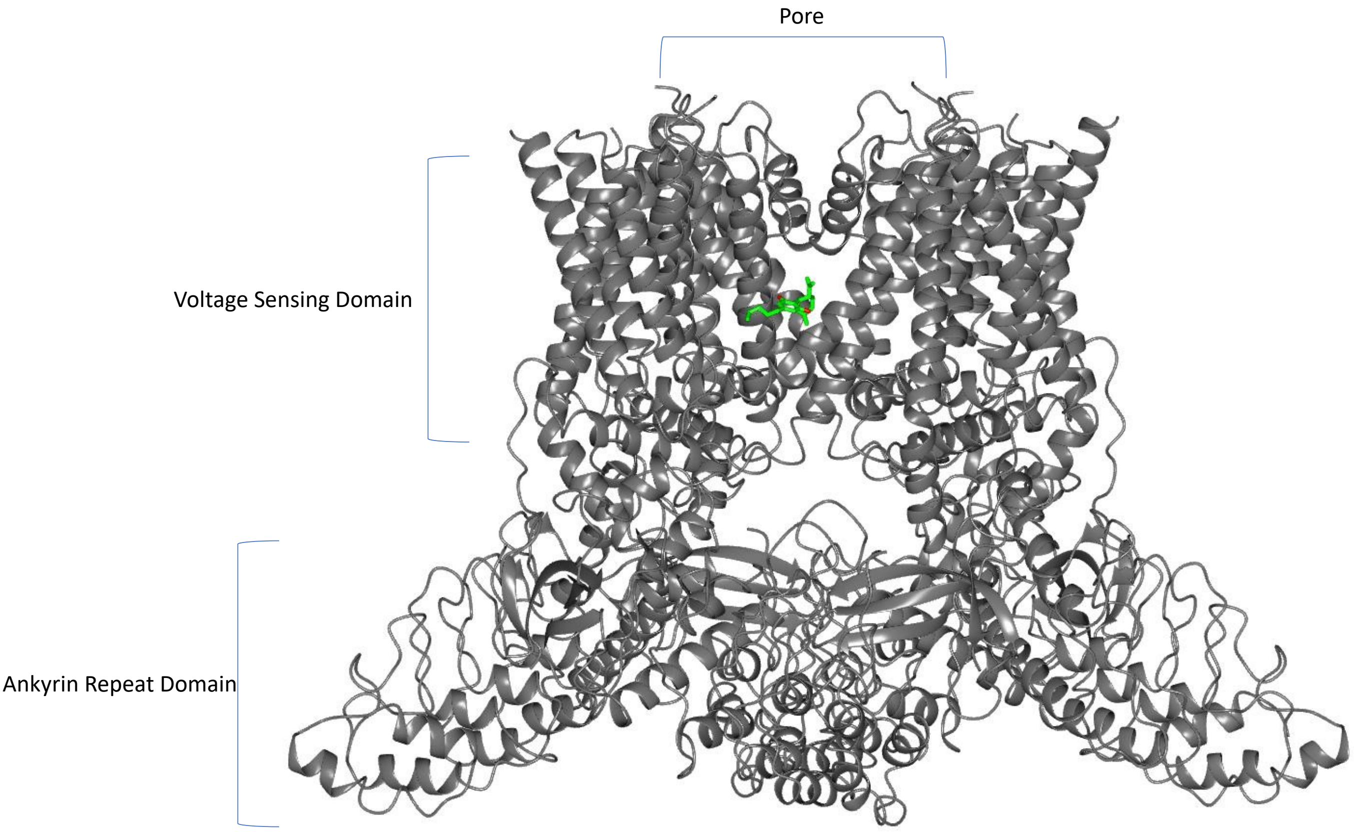


Figure 1.4: TRPV2 with CBD bound structure. *Structure of TRPV2 in its tetramic state with CBD bound (PDB ID: 6U88). The pore, VSD and ankyrin repeat domain are indicated. Produced using CCP4mg (McNicholas et al., 2011).*

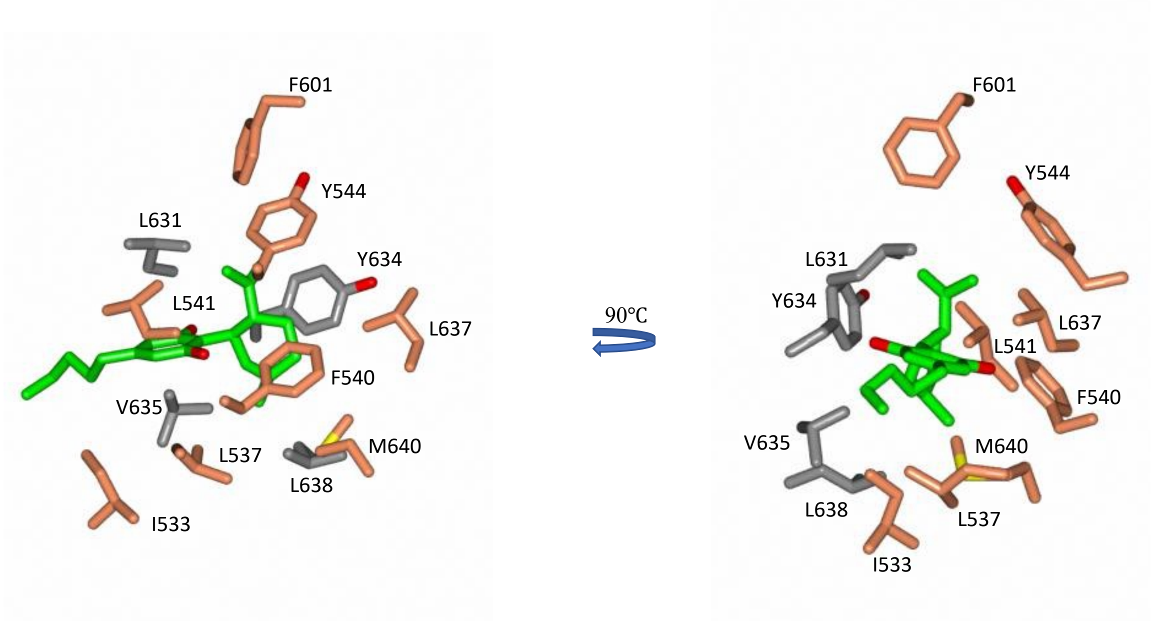


Figure 1.5 : CBD binding residues in TRPV2 binding pocket. *A view of residues which lie within 4 Å of the CBD (green stick representation, with red representing oxygen atoms) location within the TRPV2 with CBD bound structure (PDB ID: 6U88). The binding site residues are shown in two orientations, with the second orientation representing a 90° rotation from the first orientation. Residues are coloured by domain (orange and grey for the contributing sub domains). Residue atoms are coloured by atom type (carbon depicted as grey or orange depending on subdomain, oxygen as red, nitrogen as blue, and sulphur as yellow). Figure produced using CCP4mg (McNicholas et al., 2011).*

1.4.2 Challenges of membrane protein structural work

To study the interactions between VGSCs and CBD on a structural level, VGSC protein has to be expressed and purified in readiness for setting up experiments. However, VGSCs are membrane-residing and expressing membrane proteins for structural studies comes with a set of unique problems, thanks to membrane protein expression being toxic to cells, and membrane proteins having to be maintained within lipid or amphipathic environments. However, the ability to over-express and purify membrane proteins has progressed significantly (Schlegel *et al.*, 2010). The major difference in membrane protein expression techniques, compared to soluble protein protocols, is the isolation of membranes. Once cells have been lysed and centrifuged to remove cell debris, the soluble fraction is taken and spun in an ultracentrifuge (up to 40000 X gravitational force (g)) for two or more hours. This leads to the membrane fraction sedimenting out and forming a pellet, which is then collected and resolubilised in lipids or detergents, and then used within the next protein purification steps to attain the membrane protein (Schlegel *et al.*, 2010). .

1.4.3 Use of prokaryotic voltage-gated sodium ion channels within structural research

Structural research into the interactions between VGSCs and ligands/drugs has often utilised prokaryotic models (Bagnéris et al., 2014; Boiteux *et al.*, 2014; Nurani *et al.*, 2008; Sula *et al.*, 2021; Zanatta *et al.*, 2019). Using a prokaryotic VGSC as a model carries practical benefits, and as demonstrated by previous research, discussed below, prokaryotic channels offer fundamental insights into the structure and function of sodium ion channels, as well as drug/ligand interactions, which can be translated over to human VGSCs.

The practical benefits of expressing prokaryotic VGSCs instead of human VGSCs stem from the ability to express prokaryotic VGSCs within *E. coli. E. coli* is the cheapest host available for protein expression, and there are already a range of commercial genetic vectors, and a number of *E. coli* strains tailored for membrane protein over-expression (Gubellini *et al.*, 2011). Consequently, *E. coli* expression makes over-expressing several different membrane protein constructs cheap and significantly easier than trying to achieve the equivalent with human VGSCs.

Prokaryotic VGSCs hold structural similarity with mammalian VGSCs in that they also contain voltage-sensitive domains and pore domains (Payandeh *et al.*, 2011; Sula *et al.*, 2017; Zhang *et al.*, 2012). In addition they have been crystallised in several different conformational states (open, closed and inactivated) (Bagnéris *et al.*, 2013; Lenaeus *et al.*, 2017; McCusker *et al.*, 2012; Payandeh *et al.*, 2011, 2012; Sula *et al.*, 2017; Zhang et al., 2012). This has allowed for research into sodium ion channel functioning mechanisms (Bagnéris *et al.*, 2015; Ke *et al.*, 2018), which appear to be applicable to human VGSCs thanks to follow up investigations (Bagnéris *et al.*, 2014; Oelstrom et al., 2014). In addition, structural work with prokaryotic channels have offered an insight into the structural basis behind state-dependent binding of sodium channel-blocking drugs (Lenaeus *et al.*, 2017). It must be noted, however, that prokaryotic VGSCs are not perfect models for human VGSCs, since structurally there are differences; prokaryotic VGSCs are true tetramers rather than pseudotetramers, and prokaryotic VGSCs do not have and inactivation gate, thus only go through slow inactivation. Nethertheless, prokaryotic channels offer a model of core structural features which are translatable across VGSCs, and offer knowledge and a model of key structure-function relationships such as ion translocation (Ulmschneider *et al*., 2013) and ligand/drug interactions (Bagnéris et al., 2014; Boiteux *et al*., 2014; Nurani *et al*., 2008; Sula *et al*., 2021; Zanatta *et al*., 2019).

1.4.4 NavMs as a prokaryotic voltage-gated sodium ion channel model within structural studies of cannabidiol interactions

NavMs (isolated originally from *Magnetococcus marinus*) is a prokaryotic channel that has extensive previous research that has demonstrated that this channel holds sequence, structural and functional similarities with human VGSCs, as well as evidencing similar binding affinities for sodium channel blockers (Bagnéris *et al*., 2013, 2014; Ke *et al*., 2018; McCusker *et al*., 2012; Sula *et al*., 2017, 2021; Ulmschneider *et al*., 2013; Zanatta *et al*., 2019). In addition, NavMs has the added benefit of having a crystal structure of the entire channel to 2.45 Å resolution, making in depth investigations of structure and binding sites possible (Sula *et al*., 2017).

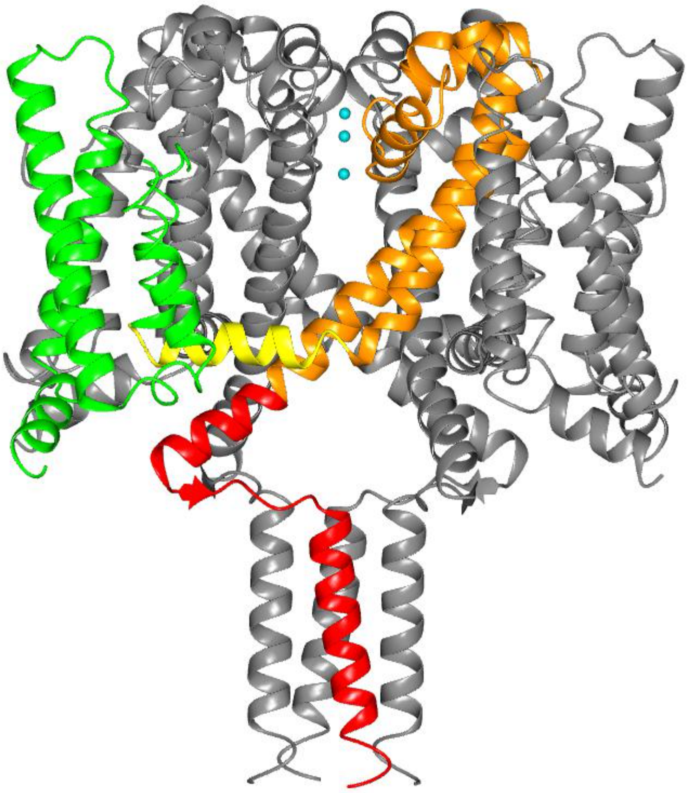


Figure 1.6: Diagram of the full-length structure of the activated sodium channel from Magnetococcus marinus (NavMs) (after (Sula et al., 2017)). *This full-length structure (PDB ID: 5HVX) is to 2.45Å and indicates the structural similarities between NavMs and human VGSCs (see section 1.1.1 for overview of structural features). The VSD is in green, the pore domain in orange, with both the IG and AG being located within the pore formed by the 4 domains. The S4-S5 linker (which acts to physically link the VSD with the PD, allowing for electromagnetic coupling between the two structures) is in yellow, and the C-terminal domain is in red. Also, the three turquoise spheres represent hydrated sodium ions within the pore. Figure produced using CCP4mg software (McNicholas et al., 2011).*

CBD is a hydrophobic molecule that has been evidenced via electrophysiology work to act as an inhibitor of VGSCs. In this sense, CBD is similar to a large number of sodium channel blockers, such as PI1, lamotrigine and lidocaine, which are also hydrophobic molecules and are known to act as sodium ion channel blockers. As mentioned, NavMs has been used to successfully investigate the binding of said sodium channel blockers (Bagnéris *et al*., 2014). During these investigations it was first ascertained that NavMs had remarkably similar binding affinities for these sodium channel blockers when compared to human VGSCs. In addition, consequent crystallographic and docking studies suggested PI1 binds at the top of the pore cavity, near a structure known as the fenestration, a hydrophobic pocket structure which lies within the transmembrane region of the channel, and has been previously proposed to be the entry site to the interior of VGSCs for hydrophobic drugs (Hille, 1977). This site was validated by follow up electrophysiology and mutational work, and the identified binding site in NavMs was shown to share equivalent residues that have been previously identified as important for drug binding in human channels. Consequently, NavMs has already proven to be an excellent and valuable model when wanting to investigate the binding of hydrophobic sodium channel blockers, thus making it a great candidate model for the study of VGSC:CBD interactions, since it can be hypothesised that CBD may be acting in a similar fashion to these sodium channel blockers in its inhibition of VGSCs.

In addition, because of previous work involving NavMs within the lab, the NavMs gene has already been successfully cloned into a Pet15b plasmid vector, along with a N-terminal hexa-histidine tag (His-tag) for nickel column purification purposes. Consequent expression and purification has also already been characterised by previous research (Sula *et al*., 2017), making NavMs an accessible VGSC model for rapidly producing high quality protein at sufficient volumes for subsequent studies, meaning that this research could focus on investigating the interactions between VGSCs and CBD.

All the above factors made NavMs the perfect channel to use when wanting to carry out novel investigations into sodium channel interactions with CBD at the structural level.

It must be noted that cryogenic electron microscopy (cryo-EM) structures for human VGSCs (Nav 1.1, Nav 1.2, Nav 1.4, Nav 1.5 and Nav 1.7) have recently been deposited with the protein database (Li *et al*., 2021; Pan *et al.*, 2018, 2019, 2021; Shen *et al.*, 2019). All the eukaryotic channels reported are bound to β subunits and/or toxins, and are to a resolution of 3.3-3 Å. These structures allow for important insights into human VGSC structures and possible functioning; however, the lower resolution means that binding sites for ligands or drugs cannot be studied in detail. In addition, the expression systems for these human VGSCs yield small amounts of protein when compared to prokaryotic VGSC *E. coli* expression systems (Charalambous and Wallace, 2011; Gubellini *et al*., 2011), making it more challenging and time consuming to yield the required amount of purified protein for structural studies.

1.5 X-ray crystallography as a structural method for researching NavMs:cannabidiol interactions:

X-ray crystallography was utilised within this research to collect high quality diffraction data of the NavMs protein in the presence of CBD, granting structural insight into the inhibitory effects of CBD, which have been observed via electrophysiological experiments (Fouda *et al*., 2020; Ghovanloo *et al*., 2021; Ghovanloo *et al*., 2018; Patel *et al*., 2016; Sait *et al*., 2020).

X-ray crystallography is a technique utilised to obtain three-dimensional molecular structures from crystals of a protein of interest. Crystals are exposed to an X-ray beam, and data on the diffracted X-ray beam is collected. Diffraction data is then consequently processed, yielding information about the structure of the protein forming the diffracting crystal. X-ray crystallography has been fundamental within the history of studying macromolecular structures (Shi, 2014). In particular, X-ray crystallography has allowed for a deepened understanding of drug/ligand interactions with macromolecules, enabling structural insight into the mechanistic underpinnings of how drugs/ligands elicit their effects (Aitipamula and Vangala, 2017). X-ray crystallography has already been used to provide insights into the structure and function of the NavMs sodium channel, and has highlighted how NavMs can be utilised as a model for human sodium channels (Bagnéris *et al*., 2013, 2014; Ke *et al*., 2018; McCusker *et al*., 2012; Sula *et al*., 2017, 2021; Sula and Wallace, 2017; Ulmschneider *et al*., 2013). Since X-ray crystallography has already been successfully used to attain structural data from NavMs protein crystals formed in the presence and absence of ligands, X-ray crystallography was a perfect technique to include within this research to aid investigations into the structural underpinnings of CBD interactions with VGSCs.

1.6 Circular dichroism as a biophysical method for researching NavMs:cannabidiol interactions:

Membrane proteins can be challenging and temperamental to express, purify, and coax into crystalline forms (due to their amphipathic environment), making it difficult to gain structural data from classic X-ray crystallographic methods (Miles and Wallace, 2016). Consequently, circular dichroism (CD) is valuable because it offers another means of gaining insight into a membrane protein’s stability, structure, function, and/or interactions with ligands.

The aim of this research was to utilise CD, specifically thermal melt CD, to analyse NavMs thermal stability with and without the presence of CBD, giving a general indication of protein quality as well as giving an indication of binding interactions.

CD spectroscopy is a technique that utilises circularly polarised light to gain information about the structural characteristics of optically active molecules. CD spectroscopy is used to investigate the structures of optically active biological molecules, including proteins. Consequently, CD is valuable because it offers another means of gaining insight into a membrane protein’s structure, function, and/or interactions with ligands. CD can be used for studying protein secondary structures (Greenfield, 2007; Whitmore and Wallace, 2008), stability (Ireland *et al*., 2018; Lai *et al*., 2018), dynamics (Huck *et al*., 2020; Linares *et al*., 2019), folding pathways (Alamry *et al*., 2020; Lamba *et al*., 2009; Salamanca and Chang, 2005), and interactions, including protein-protein interactions, protein-ligand interactions, and interactions between proteins and metal ions (Hosseinzadeh *et al*., 2021; Jin *et al*., 2021; Manieri *et al*., 2021; Najar *et al*., 2021; Zanatta *et al*., 2019). CD has been used successfully in the study of membrane-bound proteins such as histidine kinases (Keegan *et al*., 2010; Kim *et al*., 2010) and membrane transport proteins (Psakis *et al*., 2009). In addition, CD has also been successfully utilised within ligand-binding studies involving membrane proteins (Gaspar *et al*., 2011; Patching *et al*., 2012; Siligardi *et al*., 2014).

Thermal melt CD (recording changes in CD as a function of temperature) is used to determine the effects of mutations, ligands, and protein environments on protein stability. These measurements can give general insight into protein stability, as well as offer intelligence on protein-ligand interactions and optimal protein environments. Thermal melt CD has already been implemented within studies of proteins, including NavMs (Di Lella *et al*., 2010; Fan *et al*., 2007; Güler *et al*., 2020; Huus *et al*., 2006; Ireland *et al*., 2018; Marada *et al*., 2016; Metrick *et al*., 2013; Szenczi *et al*., 2006).

1.7 Whole cell patch clamp studies as an electrophysiology technique for researching NavMs:cannabidiol interactions:

Electrophysiology broadly pertains to the study of electrical properties of biological cells and tissues. Whole cell patch clamp is a form of electrophysiological experiment, and involves recording currents through multiple channels simultaneously over a large span of cell membrane. This technique has already been successfully utilised to investigate the inhibitory effects of CBD on VGSCs (Fouda *et al.*, 2020; Ghovanloo *et al.*, 2018, 2021; Hill *et al.*, 2014; Mason and Cummins 2020; Patel *et al.*, 2016)

Specific to this research electrophysiological research allows for insight into the real-time functioning of VGSCs, and the effects on VGSC-related currents when drug/ligand is present. Consequently, inclusion of electrophysiological studies, specifically whole cell patch clamp studies, allowed for functional analysis of CBD's effects on NavMs. In addition, studies with specifically mutated NavMs acted as a validation tool for any proposed binding sites that arose from structural studies.

1.8 Conclusion:

Regarding CBD, there have been extensive preclinical and clinical trials looking at CBD's effectiveness as a treatment within a range of conditions that are often associated with VGSCs, including epilepsy, pain disorders, Parkinson's, and schizophrenia. CBD has recently been approved by the European Medicines Agency and the Federal Drug Administration for use in treating Dravet syndrome and Lennox-Gastaut Syndrome due to the recent and extensive clinical trials into its effective use as an anti-convulsant within these epileptic conditions. The trials into its use within other conditions, such as pain, Parkinson's and schizophrenia are currently smaller in size than the epilepsy clinical trials, but still indicate promising results.

Interactions between VGSCs and CBD are supported by a multitude of functional and electrophysiology studies. These studies indicate that CBD acts as an inhibitor of VGSC activity at physiologically relevant conditions, which could explain its beneficial effect within the aforementioned channelopathies and the conditions associated with VGSCs.

However, there is currently no structural research into CBD's interactions with VGSCs, with the only deposited structural complex being with TRPV2 at 3.2 Å resolution. This structure is interesting as TRPV2 is also a tetramic ion channel, thus comparisons could be made between TRPV2 and future VGSC:CBD structures, possibly giving insight into why CBD acts as an inhibitor of VGSCs but an activator of TRPV2. Further structural research would help bolster our understanding of CBD's interactions with VGSCs, since this work would serve as a complement to the functional work that has already been carried out, allowing for insight into the structural mechanisms underlying CBD's interaction, thus aiding our continued use and understanding of CBD within medicinal uses.

Mammalian VGSCs are large pseudotetramers made up of four segments (D1-D4), connected by intracellular linkers. In comparison, prokaryotic structures are similar in structure but are true tetramers. In addition, several of these prokaryotic channels have been shown to have similarities with mammalian VGSCs. NavMs (from *Magnetococcus marinus)* has sequence, structural and functional similarities to human VGSCs. NavMs also has a full-length crystal structure at 2.45 Å, which is of the highest resolution full-length VGSC structures to-date. In addition, NavMs has already been extensively used within the Wallace lab, thus its expression and purification has been already optimised, allowing for the focus to be on investigating key VGSC:CBD interactions, rather than on trying to express and purify protein. These several factors made NavMs a great candidate for modelling human VGSCs and their interaction with CBD. There are recent cryo-EM structures for human VGSCs channels. These structures are important; however, the lower resolution means that binding sites for ligands or drugs cannot be studied in detail. Also,

Plans for this doctoral research project involved using X-ray crystallography, electrophysiology, and CD spectroscopy, on purified sodium channels to examine the interactions between CBD and VGSCs on a structural and functional level. NavMs, the sodium channel from *Magnetococcus marinus*, will be used as a model for human sodium channels, since it has been previously characterised, and has suitable binding and affinity kinetics for human sodium channel blockers (this includes common local anaesthetics and anti-epileptics, such as lidocaine and lamotrigine) (Sula *et al*. 2016, Bagnéris *et al*. 2014), making it an ideal candidate model for studying CBD interactions, since previous functional evidence suggests that CBD acts as an inhibitor of VGSC activity, thus could be hypothesised to be interacting with NavMs in a similar fashion as previously researched sodium channel blockers.

Chapter 2: Materials, Tools and Methods

Methods in sections 2.2-2.5 had been previously optimised by research within the Wallace lab (Sula *et al.*, 2017).

Methods in section 2.6.2-2.6.3 were adapted from previous methods set out by (Ireland *et al.*, 2018).

Methods in section 2.7.1 used initial NavMs protein concentrations of 10mg/ml based on previous NavMs X-ray crystallography research having crystallisation success at this concentration (Sula *et al.*, 2017).

Item	**Supplier**
2X KAPA HiFi HotStart ReadyMix	Kapa Biosystems
2-log DNA ladder	New England Biolabs
6X Loading buffer	New England Biolabs
α-poly His horseradish peroxidase-conjugated antibody	Invitrogen
Acetic acid	Fisher Scientific
Agarose gel powder	Sigma Aldrich
AKTA purifier fast protein liquid chromatography (FPLC) system	Cytiva (formerly GE Healthcare life sciences)
Amicon Ultra-4 centrifugal filters (100 kDa cut-off)	Sigma-Aldrich
Amicon Ultra-15 centrifugal filters (100 kDa cut-off)	Sigma-Aldrich
Ampicillin	Melford
Aviv 430 circular dichroism instrument	Aviv Biomedical
C41 *E. coli* cells	New England Biolabs
Cannabidiol (CBD)	Sigma-Aldrich (structural work) / Toronto Research Chemicals (Electrophysiology work)
Chinese hamster ovary-K1 cells (CHO-K1)	Cedarline Laboratories
CutSmart buffer	New England Biolabs
DH5α *E. coli* cells	Thermo Fisher
Dimethyl sulfoxide (DMSO)	Fisher Scientific
DNAse I	Sigma-Aldrich
DpnI enzyme	New England Biolabs

Item	Supplier
Ethylenediaminetetraacetic acid (EDTA)	Thermo Fisher
EDTA-free protease inhibitor cocktail	Roche
EmulsiFleX-C3 high-pressure homogeniser	Avestin
EPC-9 patch clamp amplifier	HEKA Elektrotronik
Glyco diosgenin (GDN)	Anatrace
Decanoyl-N-Hydroxyethylglucamide (HEGA-10)	Anatrace
Histrap-HP column	Cytiva (formerly GE-Healthcare)
iBlot 2 Dry blotting system	Invitrogen
iBlot 2 transfer stacks	Invitrogen
Imidazole	Thermo Fisher
InstantBlue Protein Stain	Expedeon
Isopropyl β-d-1-thiogalactopyranoside (IPTG)	Generon
ITC-16 interface	Instrutech
Lysogeny broth (LB)	Formedium
LB broth with agar	Formedium
Magnesium sulphate ($MgSO_4$)	Thermo Fisher
MemMeso	Molecular Dimensions
MemChannel	Molecular Dimensions
MemGold	Molecular Dimensions
MemGold 2	Molecular Dimensions
MemStart/MemSys	Molecular Dimensions
3-(N-morpholino)propanesulfonic acid (MOPS)	Sigma-Aldrich
Morpheous	Molecular Dimensions
mosquito LCP liquid handling robot	TTP Labtech
MRC 2 Well Crystallisation Plate	Swissci
n-dodecyl-β-D-maltopyranoside (DDM)	Anatrace
NanoDrop ND-1000 spectrophotometer	Thermo Fisher
NuPAGE MOPS SDS running buffer	Invitrogen
NuPAGE Bis-Tris 4-12 % protein gel	Invitrogen
PageRuler Plus pre-stained protein ladder	Thermo Fisher
Phosphate-buffered saline (PBS) tablets	Sigma-Aldrich
pET15b plasmid encoding NavMs	PMID:28205548
Peristaltic pump	GE Healthcare
Phenylmethylsulfonyl fluoride (PMSF)	VWR International
Polyethylene glycol-400 (PEG-400)	Sigma-Aldrich
Polysorbate-20 (Tween-20)	Thermo Fisher
Primers	Sigma-Aldrich

Item	Supplier
pTracer-CMV2, IRES GFP plasmid encoding NavMs	https://www.addgene.org/100004/
Quartz suprasil "bottle cell" pathlength 0.01cm	Hellma Ltd.
SIGMAFAST^{TM} 3,3′-diaminobenzidine (DAB) tablets with metal enhancer	Sigma-Aldrich
Skimmed milk powder	Thermo Fisher
Super optimal broth with catabolite repression (SOC) outgrowth medium	New England Biolabs
Sodium chloride (NaCl)	Thermo Fisher
Sodium fluoride (NaF)	Thermo Fisher
Superdex 200 Increase - 10/300 GL	GE Healthcare
SybrSafe DNA gel stain	Thermo Fisher
Syringe filters	Sartorious Biotech
Thrombin (restriction grade) and 10Xx cleavage buffer	Merck Millipore
Tris(hydroxymethyl)aminomethane (tris base)	Thermo Fisher
Wizard Plus SV Miniprep kit	Promega

Table 2.1: List of materials and equipment used.

Software/Online tool name	**Source or PubMed ID (PMID) or digital object identifier (doi)**
Affinity Designer	SERIF Europe
Aimless	PMID: 23793146
BUSTER	PMID: 22505257
CCP4	PMID: 15299374
CCP4mg	PMID: 21460457
CDToolX	PMID: 30168221
ChemDoodle	PMID: **25072815**
Clustal Omega	PMID: 28884485
Coot	PMID: 20383002
Dichroweb	PMID: 15215473
EMBOSS NEEDLE	PMID: 30976793
ExPASy Translate	PMID: 12824418
FitMaster	HEKA Elektronik
HOLE	PMID: 9195488
IGOR Pro	Wavemetrics
Microsoft Excel	Microsoft Corporation
Microsoft PowerPoint	Microsoft Corporation
Molprobity	PMID: 29067766
Oligo Calc	PMID: 17452344
PatchMaster	HEKA Elektronik
PROCHECK	doi:10.1107/s0021889892009944
RCSB PDB	PMID: 10592235
REFMAC	PMID: 21460454
SnapGene	snapgene.com
Uniprot	PMID: 30395287
XDS	PMID: 20124692

Table 2.2: List of software and online tools used.

2.1 Bioinformatics:

The following Uniprot (Bateman, 2019) code were used: NavMs (A0L5S6), Nav 1.1 (P35498), Nav 1.2 (Q99250), Nav 1.3 (Q9NY46), Nav 1.4 (P35499), Nav 1.5 (Q14524), Nav 1.6 (Q9UQD0), Nav 1.7 (Q15858), Nav 1.8 (Q9Y5Y9), Nav 1.9 (Q9UI33).

Multiple sequence alignments were achieved using Clustal Omega (Sievers and Higgins, 2018). Clustal Omega parameters were set to their default settings (refer to https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/Clustal+Omega+Help+and+Documentation for details of parameters and their default settings). Pair-wise alignments were created using EMBOSS NEEDLE with settings set to default (refer to https://www.ebi.ac.uk/seqdb/confluence/display/JDSAT/EMBOSS+Needle+Help+and+Documentation for details of parameters and their default settings) (Madeira *et al*., 2019).

DNA sequence translation was carried out using the ExPASy translate tool with the genetic code set to standard (Gasteiger *et al*., 2003).

2.2 NavMs plasmid construct cloning:

Molecular cloning is vital to amplifying up and producing desired DNA sequences. Molecular cloning as a method involves taking a target section of DNA sequence and inserting it into some form of genetic vector for propagation. The copies of the recombinant genetic vector may then be used to express the resulting protein encoded by the inserted DNA sequence, allowing for study of the protein's structure and function.

Often during molecular cloning, the genetic vector is a plasmid (a self-replicating section of DNA), which can then be used to transform bacteria, or any relevant host, allowing for propagation to occur thanks to host multiplication. A commonly used host is *Escherichia coli* (*E. coli)* because it is easily genetically manipulate, cheap to maintain, and easy to grow up.

The first instance of the NavMs gene being isolated from *Magnetococcus marinus* occurred within D.E. Clapham's lab, (McCusker *et al*., 2012) and was donated to the Wallace lab. Dr. Claire Bagnéris altered the C-terminal domain of the gene by replacing this section with a codon-optimised synthetic gene starting at H237 (McCusker *et al.*, 2012). Dr. Claire Bagnéris is also responsible for successfully cloning this altered NavMs gene into a Pet15b plasmid vector, along with a N-terminal hexahistidine tag (His-tag) (McCusker *et al.*, 2012) (PMID:28205548).

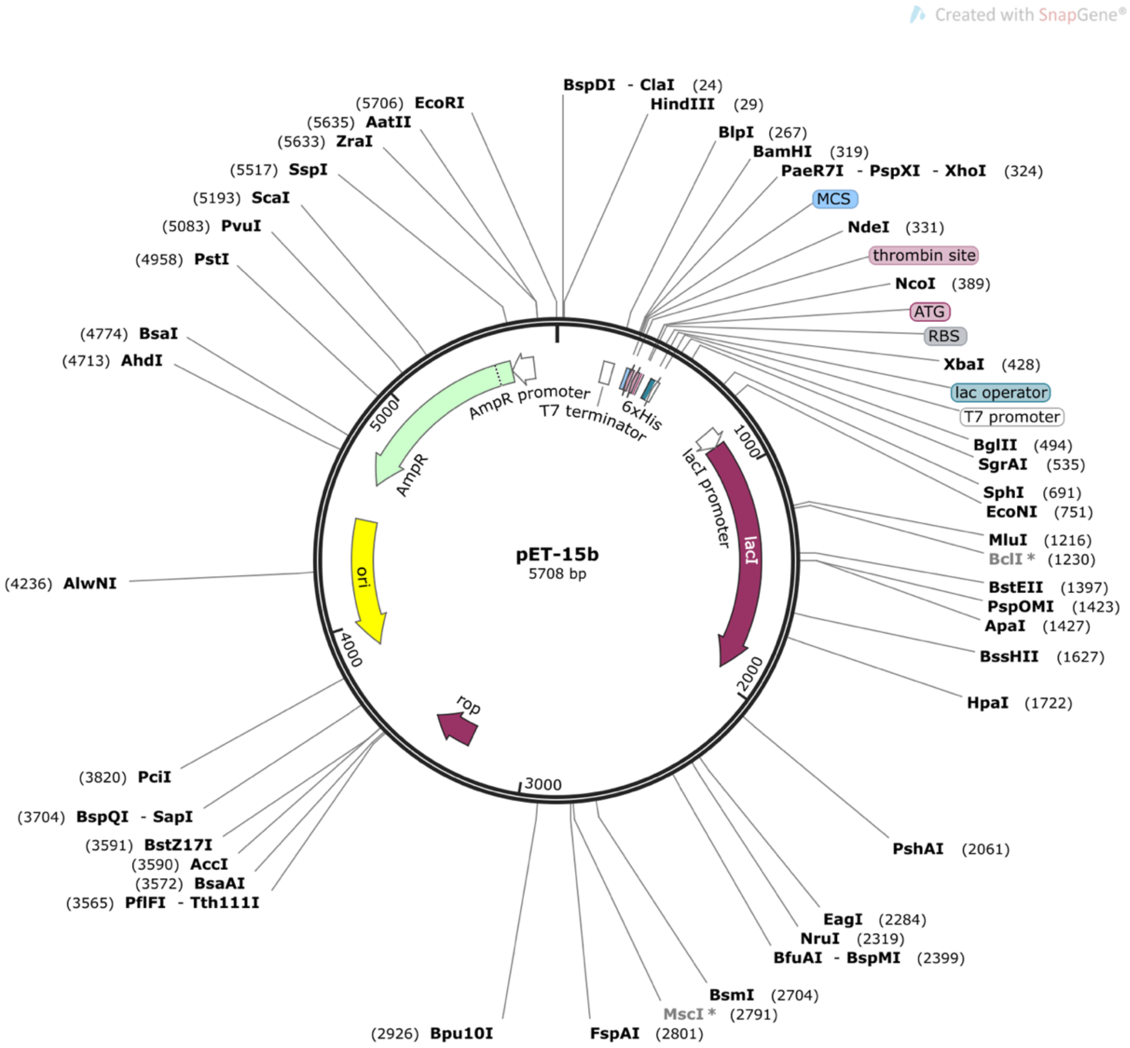


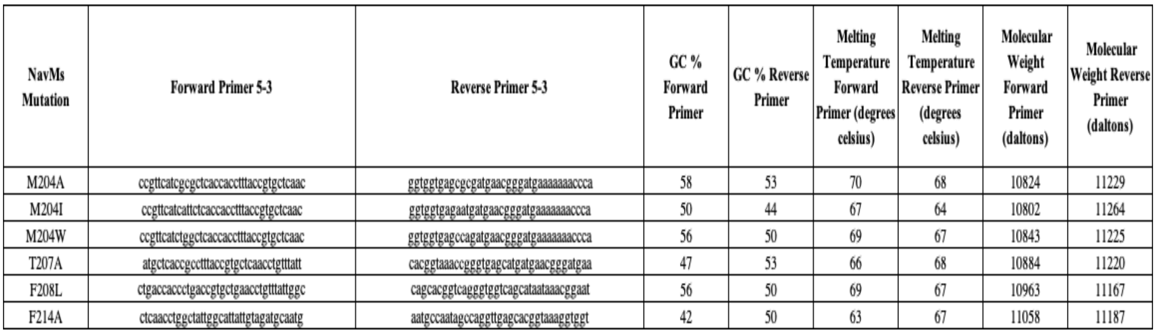
Figure 2.1 : ***Plasmid map of pET-15b plasmid.*** *NavMs was previously successfully cloned into this plasmid, allowing for expression of N-terminally His-tagged NavMs protein. Produced using SnapGene software (from Insightful Science; available at snapgene.com).*

2.3 Site-specific point mutation cloning into NavMs plasmid constructs:

To introduce site-specific point mutations into a gene construct, Site-directed Ligase-independent Mutagenesis (SLIM) can be used (Chiu *et al.*, 2004). This method works by ligase-independent cloning, which is a polymerase chain reaction (PCR) method that involves the use of primers (containing the point mutation of interest) with complementary overhangs, allowing for ligation of the PCR-plasmid, and the consequent creation of cloneable plasmids with the desired mutation.

2.3.1 Primer design for SLIM PCR cloning

Primers were designed by hand and synthesised by Sigma-Aldrich. Primers were checked using OligoCalc (Kibbe, 2007) to minimise the risk of hairpin formation and dimer formation and also check primers had similar melting temperatures. ExPASy Translate (Gasteiger *et al.*, 2003) was used to check that the primers correctly encoded the target sequence.



NavMs Mutation	Forward Primer 5-3	Reverse Primer 5-3	GC % Forward Primer	GC % Reverse Primer	Melting Temperature Forward Primer (degrees celsius)	Melting Temperature Reverse Primer (degrees celsius)	Molecular Weight Forward Primer (daltons)	Molecular Weight Reverse Primer (daltons)
M204A	ccgttcatcgcgctcaccacctttaccgtgctcaac	ggtggtgagcgcgatgaacgggatgaaaaaaaccca	58	53	70	68	10824	11229
M204I	ccgttcatcattctcaccacctttaccgtgctcaac	ggtggtgagaatgatgaacgggatgaaaaaaaccca	50	44	67	64	10802	11264
M204W	ccgttcatctggctcaccacctttaccgtgctcaac	ggtggtgagccagatgaacgggatgaaaaaaaccca	56	50	69	67	10843	11225
T207A	atgctcaccgcctttaccgtgctcaacctgtttatt	cacggtaaaccgggtgagcatgatgaacgggatgaa	47	53	66	68	10884	11220
F208L	ctgaccaccctgaccgtgctgaacctgtttattggc	cagcacggtcagggtggtcagcataataaacggaat	56	50	69	67	10963	11167
F214A	ctcaacctggctattggcattattgtagatgcaatg	aatgccaatagccaggttgagcacggtaaaggtggt	42	50	63	67	11058	11187

Table 2.3: Primers designed for SLIM PCR cloning to introduce point mutations. *The above primers were designed to introduce point mutations into the NavMs protein. Primers for M204I, M204W and F208L were designed and previously deployed by Dr. Altin Sula to produce mutants via SLIM PCR, but are displayed here for full transparency. Primers were designed by hand and then checked using ExPASy Translate (Gasteiger et al., 2003) to see that the sequence encoded was correct, and OligoCalc (Kibbe, 2007) was used to check for and minimise the risk of hairpin and dimer formation.*

2.3.2 SLIM PCR

KAPA HiFi HotStart ReadyMix was used for PCR amplification of the plasmids. The ready mix contains KAPA HiFi HotStart DNA Polymerase (0.5 mM per 25 µL reaction), deoxyribonucleotide triphosphates (dNTPs) (0.3 mM of each dNTP at 1X), $MgCl_2$ (2.5 mM at 1X) and stabilisers.

Components	Volume (µL)
2X KAPA HiFi HotStart ReadyMix	12.5
Template (NavMs WT) DNA (100 µM)	1.0
Forward primer (10 µM)	1.0
Reverse primer (10 µM)	1.0
DMSO	1.0
MilliQ filtered water	8.5
Total volume	25.0

Table 2.4: List of components used in SLIM PCR reactions.

Stage	Temperature (degrees Celsius)	Time (Minutes, Seconds)	Cycle number
Initial	95	3.00	1
Denaturation	98	0.20	35
Annealing	65	0.25	
Extension	72	0.25	
Final extension	72	5.00	1
Cooling/Storage	4	Hold	1

Table 2.5: Depiction of the heat cycling programme used within the PCR reactions.

Gel electrophoresis to check mutant plasmid constructs: Gels consisted of 1 % agarose gel, 1 % tris(hydroxymethyl)aminomethane-acetate-ethylenediaminetetraacetic acid (TAE) buffer (40 mM tris base, 20 mM acetic acid, 1 mM EDTA, pH 8.0), and SybrSafe DNA gel stain. Gels were set in casting tray along with a comb. Gels were placed inside a gel tank filled with 1 % TAE buffer. 1 µL of each PCR sample was mixed with 6X loading buffer and loaded into the gel wells. A few µL of 2-log DNA ladder was added to the first well to act as a molecular weight marker. 120 V was applied for forty minutes. DNA bands were visualised and imaged using an ultraviolet illuminator machine.

Restriction enzyme digestion: After the PCR products had been visualised, they were consequently digested with DpnI (which cleaves methylated DNA) to remove any template DNA. 1 µL of Dpn1enzyme was added per sample, along with 2 µL of CutSmart buffer. The sample was then incubated at 37 °C for thirty minutes.

2.4 NavMs plasmid propagation and sequencing:

2.4.1 Transformations for plasmid construct propagation

Chemically competent *E. coli* (DH5α) was used. 2 µL (at 1000 ng/µL) of NavMs Pet15b plasmid construct was added to 50 µL of cells, incubated on ice for twenty minutes, then heat-shocked (incubated at 45°C for thirty seconds). Samples were then placed on ice for two minutes. 250 µL of super optimal broth with catabolite repression (SOC) outgrowth medium was added, and samples were then incubated at 37°C for thirty minutes. 100 µL of the sample was spread on to a lysogeny broth (LB) agar plate containing 0.1mg/mL ampicillin and incubated at 37°C overnight. Single colonies were picked and used to inoculate 10 mL of LB broth containing 0.1 mg/mL ampicillin. Inoculated broth incubated overnight at 37°C, being shaken at 200 revolutions per minute (rpm).

Inoculated broth was used for plasmid harvesting and purification process (described below). The same process of transformation was used to transform C41 *E. coli* cells, which were used for protein expression.

2.4.2 Plasmid harvesting and purification

Plasmids were extracted and purified using the miniprep purification kit ‘Wizard Plus SV Miniprep kit’. Final concentrations were determined using a NanoDrop spectrophotometer set to measure absorbance at 260 nm.

2.4.3 Sequencing of plasmid constructs

20 µL of purified DNA samples at 100 ng/µL were sent to Source Bioscience and sequenced using their in-house T7 promoter primers. Sequences were then compared to the NavMs WT protein sequence using ExPASy translate (Gasteiger *et al.*, 2003) and EMBOSS NEEDLE (Madeira *et al.*, 2019).

2.5 Protein expression and purification:

2.5.1 Expression system

For the expression of NavMs, transformations were performed using C41 *E. coli* cells.

Cells were cultured at 37°C at 220 rpm within twelve 2 litre flasks containing 0.8 litres of LB with 0.1 mg/mL ampicillin. 1 mL samples were systematically removed, and cell density tested using a UV-Vis spectrophotometer, set to optical density (OD)-600. Once OD-600 reached ~0.8 expression was induced by adding 0.5 mg/mL isopropyl β-D-1-thiogalactopyranoside (IPTG) per flask. Flasks further incubated at 37°C and 220 rpm for three hours. Cells were harvested by centrifuging at 6000 g for twenty minutes, then were split between two 50 mL conical flasks and stored at -20°C.

2.5.2 Protein purification

8 litres worth of pelleted cell sample (two 50 mL conical flasks) were removed from -20°C storage, defrosted on ice, and kept at 4°C throughout the purification process. 50 mL of cell lysis buffer (Table 2.6) was added to the defrosted cells per conical flask, with two ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor tablets, 0.2 mM phenylmethylsulfonyl fluoride (PMSF), 2 mg/mL DNase I and 2.5 mM magnesium sulphate ($MgSO_4$). Cells were re-suspended with a hand-held homogeniser, then lysed using the EmulsiFlex C3 homogeniser, performing a minimum three passes at a minimum pressure of 15000 pounds per square inch (psi). The lysate was collected and centrifuged at 10000 g for thirty minutes to pellet out cell debris. Membrane-containing supernatant from the centrifugation was collected and centrifuged at 41000 g for two hours to pellet out membranes. Pelleted membranes were collected and resolubilised in 100 mL of solubilisation buffer (Table 2.6), then rotated at 4°C for two hours. The sample was filtered using a 0.45 µm filter and had 50 mL of dilution buffer (Table 2.6) added afterwards to adjust the buffer to 20 mM tris(hydroxymethyl)aminomethane (tris base), 300 mM NaCl, 30 mM imidazole, 1 % n-dodecyl-β-D-maltopyranoside (DDM), pH 7.5. This helped avoid non-specific interactions with the HisTrap column later on.

A peristaltic pump was washed with 10 mL of distilled water, and then had a nickel affinity HisTrap column attached. The HisTrap column was then washed with ~15 mL distilled water and then with ~15 mL 2:1 solubilisation buffer:dilution buffer (20 mM tris base, 300 mM NaCl, 30 mM imidazole, 1 % DDM, pH 7.5). The filtered sample was loaded on to the HisTrap column using the peristaltic pump. The loaded HisTrap was then attached to an AKTA purifier fast protein liquid chromatography (FPLC) purification system. The detergent was exchanged using buffer A (Table 2.6), and the protein eluted using a step process. The step process involved increasing the percentage amount of buffer B (Table 2.6) starting at 15 %, then 23 %, then finishing with the imidazole concentration at 500 mM (100 % buffer B) to fully wash out protein (the step process was optimised during previous work purifying NavMs (Sula *et al*., 2017)). Fractions were analysed using SDS-PAGE, and consequently identified peak protein fractions from the HisTrap affinity chromatography step were pooled and concentrated using two Amicon ultra-15 centrifugal filters (100 kDa cut-off) to a total of 2 mL (1 mL per centrifugal filter). 200 μL of 10X thrombin cleavage buffer was added to the sample, along with 10 μL of thrombin protease to cleave the His-tag, and incubated overnight at 4°C.

The 2 mL sample containing the now cleaved protein was further concentrated to ~500 μL using an amicon ultra-4 centrifugal filter (100 kDa cut-off). The concentrated sample was then further purified via SEC using buffer A and a Superdex 200 10/300 GL column attached to an AKTA purifier FPLC purification system. SEC Fractions were analysed using SDS-PAGE and immunoblotting to check for cleavage and the presence of purified NavMs protein. Peak fractions from the SEC were consequently collected and pooled, and protein concentration was determined using a Nanodrop spectrophotometer. Protein was concentrated using an Amicon ultra-4 centrifugal filter (100 kDa cut-off) to 10 mg/mL and stored at -80°C in 50 μL aliquots.

Buffer	**NaCl (mM)**	**NaF (mM)**	**Tris base (mM)**	**Detergent (%w/v)**	**imidazole (mM)**	**pH**
Cell lysis buffer	100	N/A	20	N/A	N/A	7.5
Solubilisation buffer	150	N/A	20	1.5 % DDM	N/A	7.5
Dilution buffer	600	N/A	20	N/A	90	7.5
Buffer A	300	N/A	20	0.52 % HEGA-10	N/A	7.5
Buffer B	300	N/A	20	0.52 % HEGA-10	500	7.5

Table 2.6: List of buffer components.

2.5.3 SDS-PAGE for protein sample analysis

Samples were mixed with 6X loading buffer then loaded into fifteen well NuPAGE Bis-Tris gels within a tank filled with 1X 3-(N-morpholino)propanesulfonic acid (MOPS) SDS running buffer. The gels were run at 200 volts (V) for forty minutes. Gels were removed and stained with Instant Blue stain for one hour. The gels were removed from the Instant Blue stain and kept within distilled water to stop the staining process and maintain clear banding.

2.5.4 Immunoblotting

Following an SDS-PAGE gel, the sample-containing gel was blotted onto a polyvinylidene difluoride (PVDF) membrane using an iBlot 7-minute blotting system. The membrane was blocked using 1X (phosphate-buffered saline) PBS with 3 % w/v skimmed milk powder and shaken for thirty minutes. The membrane was washed in PBS three times for five minutes, then incubated with 7 μL of α-poly His horseradish peroxidase-conjugated antibody (a His-tag targeting antibody) (1:2000 dilution) in 5 mL PBS with 1 % w/v skimmed milk for two hours at room temperature. The membrane then went through three rounds of washing, in PBS with 0.05 % v/v polysorbate-20 (Tween-20) for fifteen minutes each time. The membrane was then washed with 3, 3'-diaminobenzidine and urea substrates (SIGMAFAST^{TM} 3, 3'-diaminobenzidine (DAB) tablets with metal enhancer) and dissolved in 5 mL distilled water for five minutes, after which the membrane was left to dry. The horseradish peroxidase associated with the antibody oxidised the diaminobenzidine, causing a brown precipitate to form, indicating His-tagged protein locations on the membrane.

2.6 Thermal melt circular dichroism spectroscopy:

2.6.1 NavMs Glyco diosgenin (GDN) buffer transfer via size exclusion chromatography for CD experiments

50 μL aliquots attained from NavMs protein purification were loaded on to a Superdex 200 10/300 column (GE Healthcare) (attached to an AKTA purifier FPLC purification system (Cytiva)) equilibrated with GDN transfer buffer (20 mM tris base, 100 mM sodium fluoride (NaF), 0.1 % GDN and pH 7.5). GDN concentration was set to 0.1 % so it was above its critical micelle concentration (0.0021 % w/v) (Chae *et al.*, 2012). Peak detergent-exchanged fractions were collected, pooled, and concentrated using an Amicon (Sigma Aldrich) ultra-15 centrifugal filter (100 kDa cut off) to ~1 mg/mL (as

determined using an A280 nanodrop with a calculated extinction coefficient at 29450 M^{-1} cm^{-1}) and stored at -80°C in 100 μL aliquots. GDN was used as the detergent because decanoyl-N-Hydroxyethylglucamide (HEGA-10) absorbs in the far UV range and has a distinct CD signal, which would complicate CD measurements.

Each NavMs (in GDN transfer buffer) protein sample aliquot used in an experiment was removed from -80°C storage, thawed on ice, and diluted with GDN transfer buffer to produce a protein concentration of ~0.5 mg/mL. This sample was then used to obtain CD measurements. In the case of samples with CBD added (to investigate the effects of CBD binding on thermal stability) the sample was thawed, kept on ice, and diluted to ~0.5 mg/mL as above, but then had 1μL stock CBD solution (32 mM in ethanol) added directly per 100 μL of sample, producing a final solution containing ~0.5 mg/mL NavMs WT protein, 0.32 mM CBD, and 1 % v/v ethanol. Ethanol does not absorb within the far UV range, and thus could be added to the solution. DMSO does absorb in this region, so CBD dissolved within DMSO (as in the crystallisation screens) could not be used in the CD experiments.

2.6.2 CD spectroscopy

Optimisation of protein concentration was carried out using ten-scan runs over wavelengths 260-190 nm at 1 nm steps for protein at concentrations of ~0.9 mg/mL, ~0.7 mg/mL, ~0.6 mg/mL, and ~0.5 mg/mL. The temperature was held constant at 20°C during the set of ten scans. These scans enabled analysis of the high-tension voltage (HT) value, and whether a sample would be usable at that given concentration for thermal melt spectral runs. All protein and baseline samples contained 20 mM tris base, 100 mM NaF, 0.1 % GDN, and pH 7.5. NaF was used instead of NaCl in the CD buffer to allow for the collection of data at lower wavelengths as NaCl exhibits absorbance in the far UV range (Miles and Wallace, 2016).

Data was collected using an Aviv 430 circular dichroism instrument and a quartz Suprasil (Hellma Ltd.) “bottle” cell [part number: 121-0.10-40] with a pathlength of 0.01 cm.

Once protein concentration was optimised (~0.5 mg/mL), a spectral run protocol was put into place for collecting data from a sample. First, a ten-scan run over wavelengths 260-190 nm in 1 nm steps was undertaken. These scans allowed for analysis of the HT, and whether the sample was appropriate for a subsequent thermal melt spectral run.

After the ten-scan run, a thermal spectral run was undertaken. The protocol for this was three repeat spectra collected at each temperature, over the wavelength range from 260 to 190 nm (again in 1 nm steps). The temperature range was from 20°C to 80°C in 5°C increments. A three minute equilibration period was enacted between measurements after each temperature increment. After reaching 80°C, the sample was cooled to 20°C, and a subsequent recording of three repeat spectra was done. This final cooling step allowed for comparison with the initial 20°C spectrum for the sample.

For the GDN transfer buffer baseline, three repeat spectra over the wavelength range from 260-190 nm (in 1 nm steps), were also obtained at each temperature. Buffer baselines were also collected, which consisted of measurements of the GDN transfer buffer with just ethanol added (1 % ethanol v/v), and with both CBD and ethanol added (0.32 mM CBD and 1 % v/v ethanol).

At least two repeats of the above spectral run protocol (each repeat was run using a fresh sample) were undertaken for NavMs WT protein samples under all conditions: without CBD, with CBD (32 mM CBD, 1 % v/v ethanol), and with ethanol (1 % v/v ethanol) (control for ethanol effects).

2.6.3 Data processing and data analysis

Data processing was carried out using CDToolX software (Miles and Wallace, 2018). Replicate scans for each sample were averaged, baselines subtracted, and spectra zeroed between 255 nm and 260 nm and converted to delta epsilon values using a mean residue weight of 113 kDa.

Secondary structures were calculated for samples using the DICHROWEB server (Whitmore and Wallace, 2004) at http://dichroweb.cryst.bbk.ac.uk/html/home.shtml. The CONTIN algorithm was used (van Stokkum *et al.*, 1990), along with the SMP180 reference dataset (Abdul-Gader *et al.*, 2011), since SMP180 was created specifically for use with membrane protein CD data.

The secondary structure for NavMs WT was derived from its crystal structure (PDBID: 5HVX) using the DSSP algorithm (Kabsch and Sander, 1983).

Single value decomposition (SVD) analysis for each sample type was achieved using CDToolX software (Miles and Wallace, 2018). The two key components of SVD corresponded to spectra for folded and unfolded structures as a function of temperature.

The magnitude of these components provided an overall indication of the thermal stability of the protein in this environment, allowing for comparisons in thermal stability between samples with CBD added and those without CBD added.

2.7 X-ray crystallography:

2.7.1 Crystal growth and screens

'Sitting Drop' experiments were set up to produce protein crystals. NavMs WT and NavMs F208L protein at 10 mg/mL (measured using a Nanodrop spectrophotometer) were removed from -80°C storage and thawed on ice. Crystal screens were set up, and protein samples were mixed in a 2:1 protein:mother liquor fashion, using the mosquito LCP liquid handling robot. The commercial crystallisation screens used were MemMeso, MemChannel, MemGold 1, MemGold 2, Morpheous and MemStart/MemSys (Gorrec, 2009; Kato *et al*., 2012; Newstead *et al*., 2008; Parker and Newstead, 2012, 2016). All were supplied by Molecular Dimensions. Trays were kept at 4°C and checked once a week for crystal "hits".

Co-crystallisation screens with CBD were set up in the same fashion, with 1uL of CBD (stock 100 mM solution in dimethyl sulfoxide (DMSO)) added directly to freshly thawed protein to form a final protein concentration of ~10 mg/mL, 2 mM CBD and 2 % v/v DMSO. This mixture was used to set up crystallisation screens as described above.

Crystals were fished out of their respective drops using small nylon loops. Crystals were then stored within liquid nitrogen during transport to the designated synchrotron for screening and data collection.

2.7.2 Model building

Data integration, scaling, and merging: Diffraction images collected from synchrotron visits were fed into the CCP4 program suite (Winn *et al*., 2011), which provides an integrated set of programmes for structural determination, allowing for ease of pipeline movement from raw data to structural model. Diffraction images were first integrated and scaled using the XDS software package (Kabsch *et al*., 2010). XDS also allowed for the resolution cutoff for the data to be adjusted according to statistics such as the CC ½ value and completeness. The CC ½ value, which is the correlation coefficient between two random half data sets associated with the target protein. This value acts as an assessor of model resolution and has a range between one and zero. As values approach one, correlation approaches 100 %. At the lower relative resolution values for a given data set, correlation is generally close to one, but this falls sharply in value at higher resolutions for a data set, where intensities become weaker (Evans and Murshudov, 2013). A conservative resolution cutoff would be to where the CC ½ falls to 0.7. Completeness is a percentage which describes the theoretically observable reflections we have measured in our experiment. Generally, the aim is to achieve a completeness percentage >96%. Integrated and scaled diffraction data was then automatically fed into Aimless for merging (Evans and Murshudov, 2013).

Molecular replacement: The NavMs WT structure is already known (PDBID: 5HVX) (Sula *et al*., 2017) and was used as the phasing model. Phasing was done using a program called PHASER, which accepts phase information from a model structure, in this case the NavMs structure elucidated by (Sula *et al.*, 2017), and combines this with the diffraction data to form initial phase estimates (McCoy *et al*., 2007), and consequently and initial model of structure.

Model building and refinement: Electron density maps and atomic model viewing and manual model building (adapting the model to fit the calculated electron density better) were carried out within Coot (Winn *et al*., 2011). Preliminary refinement where the adapted atomic model from coot, with its improved phase estimates, were fitted to the diffraction data in an iterative fashion, was carried out using REFMAC (Murshudov *et al*., 2011). Refinement was finished off using BUSTER because BUSTER is very good at modelling final parts of the structure which an atomic model is not yet available, such as water molecules (Bricogne *et al*., 2011; Smart *et al*., 2012). Refinement iterations were finished once R-factors (R_{work} and R_{free}) (produced and displayed by both REFMAC and BUSTER) no longer decreased in value, indicating that the model could not be further improved. In general, R-factors with a value between 0.15 and 0.3 are considered good for a final model. Structure quality (checking if residues were in allowed conformations) was checked using PROCHECK (Laskowski *et al*., 1993) and MolProbity (Chen *et al*., 2010; Williams *et al*., 2018).

Structural coordinates and any associated metadata were deposited into the Protein Data Bank (PDB) (rcsb.org) (Berman *et al*., 2000).

Figures and structural alignments: Structural figures were created using CCP4mg (McNicholas *et al*., 2011). CCP4mg was used to create structural alignments using the General Efficient Structural Alignment of Macromolecular Targets (GESAMT) algorithm. HOLE (Smart *et al*., 1996) was a software used to create visual model outputs of NavMs pore dimensions, and works via the simple input of a protein database (pdb) file of the protein structure in question by the user.

2.8 Electrophysiology

Electrophysiology experiments were carried out by Prof. Paul Ruben and Dr. Mohammad-Reza Ghovanloo and followed the below methodology, which is also covered within (Sait *et al*., 2020):

CBD was dissolved in 100 % DMSO and used to make extracellular solutions with a total DMSO concentration of less than 0.5 %. Chinese hamster ovary (CHOK1) cells were transiently co-transfected with complementary DNA (cDNA) which encoded enhanced green fluorescent protein (eGFP), the β1-subunit and the NavMs α-subunit (https://www.addgene.org/100004/) (Fig 2.2). The PolyFect protocol was followed for transfections, and after each transfection, cells were incubated for at least eight hours before being plated on to sterile coverslips. Cells were then incubated at 37°C at 5 % carbon dioxide.

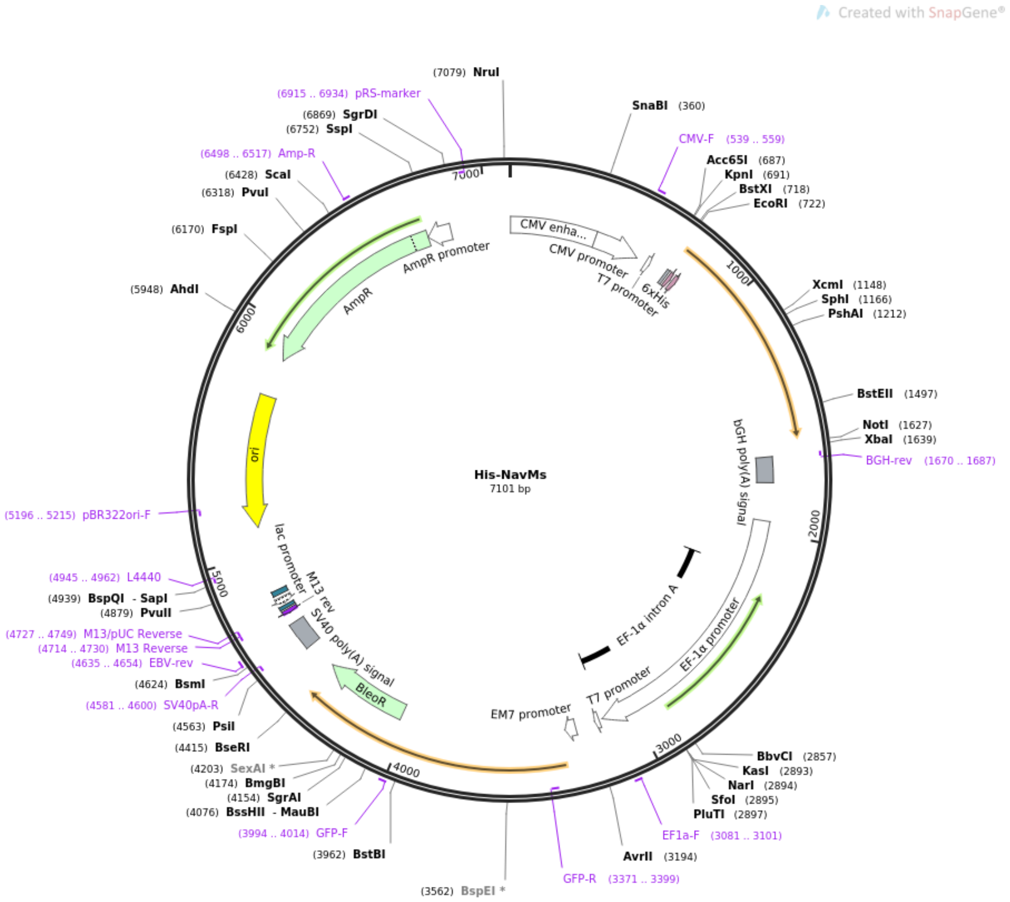


Figure 2.2 : Plasmid map of AddGene plasmid 100004. *Map of commercially available plasmid construct containing His-tagged NavMs used by Prof. Paul Ruben and Dr. Mohammad-Reza Ghovanloo within transfection of CHOK1 cells for electrophysiology experiments. Produced using SnapGene software (from Insightful Science; available at snapgene.com).*

Whole-cell patch-clamp recordings were then carried out in an extracellular solution consisting of: 140 mM sodium chloride (NaCl), 2 mM calcium chloride ($CaCl_2$), 1 mM magnesium chloride ($MgCl_2$), and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.4). Solutions were adjusted to meet a pH of 7.4 via the use of caesium hydroxide (CsOH). Pipettes were filled with intracellular solution consisting of 120 mM caesium fluoride (CsF), 20 mM caesium chloride (CsCl), 10 mM NaCl, and 10 mM HEPES. Recordings were done with an EPC-9-patch-clamp amplifier (HEKA Elektronik, Lambrecht, Germany) and digitised at 20 kilohertz (kHz) with an ITC-16 interface (Instrutech, Great Neck, New York, USA). Data acquisition and voltage clamping were carried out via the use of PatchMaster software (HEKA Elektronik, Lambrecht, Germany), which was run on an Apple iMac. Current got low-pass-filtered at 10 kHz, and leak subtraction was automatically carried out by software using the P/N procedure following the test pulse. Gigaohm seals were stabilised in the on-cell configuration for one minute prior to whole-cell configuration. Series resistance was consistently less than 5 megaohm (MΩ) for all recordings. Series resistance compensation reaching up to 80 % was used when necessary. All data recordings were acquired at least one minute after establishing whole-cell configuration. Prior to each run of the protocol, the cell membrane was hyperpolarised to -180 milli volts (mV), ensuring the complete removal of NavMs channel inactivation. Experiments were conducted at 22°C +/- 2°C.

Analysis and graphing of data was performed using FitMaster software (HEKA Elektronik, Lambrecht, Germany) and Igor Pro (Wavemetrics, Lake Oswego, Oregon, USA), which were also run on an Apple iMac. Continuous variables are represented as means with +/- standard error values, and had a normal distribution. T-tests were used to compare responses, with a significance level $\alpha = 0.05$ being used in all overall tests, and effects with a p-value of less than 0.05 being considered statistically significant.

Chapter 3: Expression and Purification of NavMs Wild Type (WT), NavMs F208L and hypothesised binding site mutants

3.1 Introduction:

The main aim of this research was to undertake studies of CBD interactions with VGSCs at the structural level, using X-ray crystallographic methods, so as to expand understanding beyond the previous electrophysiological work that had been carried out in other scientific groups. To carry out X-ray crystallographic work sufficient amounts of high quality VGSC protein had to be initially expressed and purified.

The prokaryotic NavMs channel was selected for this research as the protein that would be expressed, purified and used in studies for several practical reasons; firstly, NavMs is cheap to produce due to being expressed in an *E. coli* vector. Secondly, NavMs was previously successfully cloned into the Pet15B plasmid, along with a N-terminal His-tag for nickel affinity chromatography purification, thus ready for use within expression methods. Thirdly, consequent expression and purification has also already been characterised and optimised by previous research (Sula *et al.*, 2017) (see methods section 2.5), making NavMs an accessible VGSC model for producing sufficient volumes of protein for structural and biophysical studies into the interactions between VGSCs and CBD.

In addition to the practical benefits expression and purification of NavMs allowed, NavMs was a good candidate model since NavMs has extensive previous research associated with it that demonstrated this channel's sequence, structural and functional similarities with human VGSCs, as well as evidencing NavMs' similar binding affinities for sodium channel blockers, which like CBD, are lipophilic molecules which appear to inhibit sodium ion channel functioning (Bagnéris *et al.*, 2013, 2014; Ke *et al.*, 2018; McCusker *et al.*, 2012; Sula *et al.*, 2017, 2021; Ulmschneider *et al.*, 2013; Zanatta *et al.*, 2019). In addition, NavMs already has a crystal structure of the entire channel to 2.45 Å resolution(Sula *et al.*, 2017), suggesting high-resolution investigations of binding site interactions were feasible with NavMs.

Both NavMs wildtype (WT) and the NavMs mutant F208L were initially expressed and purified. The mutant NavMs F208L plasmid had been successfully produced in the lab previously via SLIM PCR (see methods section 2.3), and had been produced because of an interest in investigating how this mutation may affect structure, since in human VGSCs half of the homologous positions (positions in two of the domains) are phenylalanine (F) whilst in the other half the homologous positions are leucine (L). NavMs F208L was specifically included for initial expression, purification and consequent structural work for two reasons. The first being that including it would allow for investigations into this mutant's structure and how this compared to NavMs WT, thus giving insight into the effect of having the F208 position mutated to a L. Second, previous crystallisation trials with the F208L mutant (unpublished) indicated that the NavMs F208L mutant crystalises more readily than NavMs WT. Including NavMs F208L not only allowed an opportunity for structural investigations of the mutant itself, but also increased chances of crystallisation events with CBD, thus increased chances of obtaining high quality diffracting crystals which would offer first ever structural data insights into VGSC:CBD interaction. Initial samples of Pet15b NavMs WT and Pet15b NavMs F208L plasmids were donated by Dr. Altin Sula, ready for use in expression and purification.

In addition to NavMs WT and F208L expression and purification, mutants at position M204, T207 and F214 were cloned, expressed and purified. These positions were chosen because T207 and F214 correspond to positions in human Navs that are known to be involved in the binding of channel-blocking compounds (Desaphy *et al*., 2010; Hanck *et al*., 2009; Liu *et al*., 2003), and have themselves been evidenced to be involved in the binding of sodium channel blocking drugs within NavMs via structural and electrophysiological work which utilised knockout alanine mutants at these positions (Bagnéris *et al*., 2014). Since CBD has been evidenced to act as a sodium ion channel inhibitor, it was hypothesised that these positions may be involved CBD binding, thus mutations at these positions could affect interactions. M204 was selected because this position in NavMs, and the equivalent position in NavAb, has been associated with influencing the entry dimensions of the fenestration structure (Gamal El-Din *et al*., 2018; Montini *et al*., 2018). The fenestration is a hydrophobic pocket structure created between the NavMs subunits, and runs perpendicular to the channel (Montini *et al*., 2018). This fenestration structure occurs four-fold within the NavMs structure due to NavMs being a tetramic structure. Fenestrations lie within the transmembrane region of the protein, and have been proposed to be the entry sites for hydrophobic drugs, which can then gain entry to the interior of sodium ion channels

(Hille, 1977). Since CBD is a hydrophobic drug, the hypothesis was that mutations at this position could alter CBD's ability to enter NavMs and cause inhibition.

In terms of mutation, it was decided to create alanine knockout mutations of both T207 and F214, since this would shorten the residue, making the distance between it any possible CBD molecule exceed 4 Å, thus removing it as a residue capable of forming hydrophobic interactions with the CBD molecule. An alanine mutant of M204 was also decided upon, since this would increase the size of the fenestration opening (as demonstrated in previous work (Montini *et al.*, 2018)). M204 was also mutated to either isoleucine (I) or tryptophan (W). Isoleucine would result in a slightly larger fenestration due to the residue being less bulky than a methionine. The tryptophan side chain is larger than the methionine side chain, and would result in a restriction of the fenestration, thus possibly inhibiting CBD entry. Initial samples of Pet15b NavMs M204I and NavMs M204W plasmids were donated by Dr. Altin Sula, ready for use in expression and purification. NavMs M204A, T207A and F214A had to be cloned prior to expression and purification.

This chapter will first cover the results from expressing and purifying NavMs WT, then will cover the results for expressing and purifying the NavMs F208L mutant, finishing with results from cloning expressing and purifying other mutants.

Please refer to methods sections 2.2-2.4 for details of methods used for plasmid cloning, propagation and sequencing, and please refer to methods section 2.5 for protein expression and purification methods. These methods had been previously elucidated and optimised for NavMs cloning, expression and purification within the Wallace lab.

3.2 Results:

3.2.1 Sequencing of NavMs WT DNA

Plasmid sample for NavMs WT was donated by Dr. Altin Sula, and after efforts to propagate this plasmid, a plasmid sample for NavMs WT was sent for sequencing to confirm that there were no mutations (see methods section 2.4) (Fig 3.1). Sequencing results were translated into an amino acid sequence of the NavMs protein then aligned with the NavMs WT sequence (Uniprot code: A0L5S6), to confirm that no mutations had been introduced. DNA for NavMs WT plasmid was confirmed to be successfully produced, and ready to be used in the transfection of *E. coli* C41 cells, thus beginning the protein expression process.

```
WT|1-50      MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ
             ||||||||||||||||||||||||||||||||||||||||||||||||||
WT-S|1-50    MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ

WT|51-100    LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR
             ||||||||||||||||||||||||||||||||||||||||||||||||||
WT-S|51-100  LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR

WT|101-150   TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT
             ||||||||||||||||||||||||||||||||||||||||||||||||||
WT-S|101-150 TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT

WT|151-200   NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI
             ||||||||||||||||||||||||||||||||||||||||||||||||||
WT-S|151-200 NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI

WT|201-250   PFIMLTTFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD
             ||||||||||||||||||||||||||||||||||||||||||||||||||
WT-S|201-250 PFIMLTTFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD

WT|251-274   RLDRIEKQLAQNNELLQRQQPQKK
             ||||||||||||||||||||||||
WT-S|251-274 RLDRIEKQLAQNNELLQRQQPQKK
```

Figure 3.1: Sequencing of the pet15b NavMs WT sample. *ExPASy translate (Gasteiger et al., 2003) used to translate sequencing returned from Source Bioscience. Translated sequences were then aligned with the NavMs WT sequence from Uniprot (Bateman, 2019) (Uniprot code: A0L5S6) using EMBOSS NEEDLE (Madeira et al., 2019).'WT' denotes the Uniprot reference sequence, and 'WT-S' denotes the sample sequence.*

3.2.2 Expression and purification of NavMs WT

Refer to methods section 2.5 for details of the expression and purification process.

HisTrap affinity chromatography: After expressing NavMs WT, and collecting membranes from this expression, purification steps were run according to the protocol laid out in the methods section 2.5. A UV trace was collected from the AKTA during the initial HisTrap nickel affinity chromatography stage (which aimed to purify out NavMs due to the expressed protein still having an N-terminal His-tag) to indicate protein coming off of the column during the step process elution with imidazole.

In (Fig 3.2A) there are three distinct UV trace peaks apparent, thus three distinct protein elution events. Each peak correlates to a gradient increase in buffer B (see methods section 2.3.3), thus an increase in the imidazole, which competes with bound protein in the HisTrap nickel affinity column (15 % buffer B for the first peak, 23 % buffer B for the second peak and 100 % buffer B for the third and largest peak). Thanks to previous work in Wallace's lab to optimise NavMs protein purification, it is known that UV traces for the HisTrap purification process indicate the presence of purified NavMs protein within the third peak. Fractions from all three peaks were collected and run on an SDS-PAGE to study the protein content of these fractions and confirm the presence of NavMs protein within the fractions taken from the third peak.

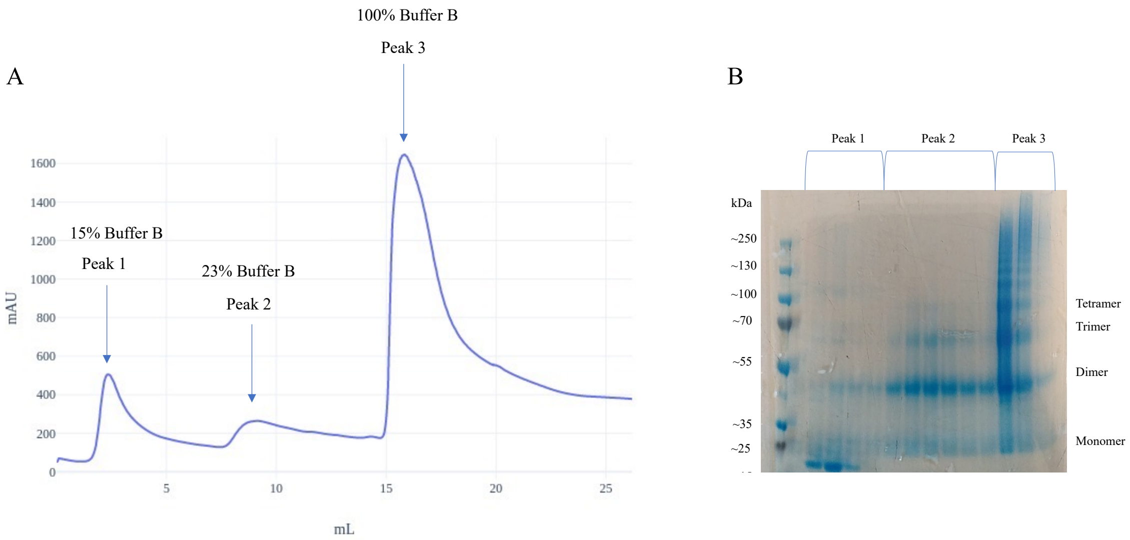


Figure 3.2: UV trace from NavMs WT HisTrap affinity chromatography and SDS-PAGE of collected HisTrap fractions. *A) This graph indicates the three UV trace peaks which occurred during the HisTrap affinity chromatography step, which correlate with increases in buffer B percentage, and thus imizadole concentration. The third peak is known to contain the NavMs protein from previous work within Wallace's lab. Fractions were taken from each peak and run on an SDS-PAGE. Figure produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA). B) Image of the SDS-PAGE gel containing fractions from the first, second and third peak.*

SDS-PAGE gels were run, with the fractions collected from each of the three peaks for NavMs WT (Fig 3.2B). For NavMs WT, all fractions from the first to third peaks can be seen to contain bands at positions ~25 kDa, ~50 kDa, ~70 kDa and ~100 kDa, which are known to indicate NavMs protein and its multimers (monomer, dimer, trimer and tetramer respectively). However, fractions from the first and second peak were observed to contain extra protein bands that are uncharacteristic of NavMs protein, as well as the banding characteristic of NavMs appearing to be much weaker in these fractions when compared to fractions from the third peak. The extra uncharacteristic banding can be attributed to proteins which non-specifically bind to the column due to being histidine-rich, but are washed out by the initial steps in imidazole gradient. The presence of multimeric state bands indicates that NavMs exists in very stable multimeric states which cannot be easily disrupted by the denaturing action of the SDS- detergent. These bands do not directly correlate with the molecular weight of NavMs and its oligomers (molecular weight of NavMs is ~31 kDa), but this phenomenon has already been observed for membrane proteins, including NavMs. This mismatch between the bands and the true molecular weight of NavMs can be explained by increased association with SDS detergent molecules (Rath *et al*., 2009). In addition, all the fractions had streaking occur within the SDS-PAGE lanes. Again, this has been previously observed for pure NavMs samples, and has also been observed as a characteristic of other pure samples of membrane proteins (Zhou *et al*., 2000).

Size exclusion chromatography: Once the HisTrap fractions containing the NavMs WT protein had been collected, concentrated, and put through His-tag cleavage, to remove the N-terminal His-tag, the samples were loaded on to a Superdex 200 10/300 GL column to carry out size exclusion chromatography (SEC) as a second purification step. UV traces were collected from the AKTA during SEC for NavMs WT (Fig 3.3A) to discern fractions containing protein. The UV trace indicated two peaks within the trace, with the first peak starting at around 7 mL and the second peak starting at around 9 mL. This shape of trace, and the volumes at which protein began to fraction off supports previous characterisation of SEC traces collected during the NavMs purification process. It has been evidenced from previous work within the Wallace group that fractions within the first peak are a different species of the NavMs protein, hence elution at a different volume. Previous circular dichroism work (unpublished) within the Wallace lab indicated that protein from the second peak from the SEC contained more secondary structure than first peak protein, indicating more folded protein, thus suggesting protein from this peak is more suitable for crystallographic studies since the protein maintains more of its folded structure. Protein from the first peak has been trialled in crystallisation screens previously but was found to only form aggregates during the sitting drop process, whereas protein from the second peak has gone on to form crystals. These points made protein fractions from the second peak the choice for collection.

SDS-PAGE gels were run with fractions from both the first and second peak, allowing for investigation of the fractions and confirmation of NavMs protein presence after removal of the His-tag (Fig 3.3B). Fractions from the first and second peaks appeared to have the distinctive banding pattern associated with NavMs, with the fractions from the second peak having stronger banding, indicating a higher presence of the stable, folded NavMs protein.

To confirm that protein fractions from SEC had actually undergone successful His-tag cleavage, and that no histidine-rich protein contaminants were present in the final SEC fractions, immunoblots were run with fraction samples collected from both the HisTrap affinity chromatography and SEC purification steps. Samples of the NavMs protein before it was loaded on to the HisTrap column, the flow through after loading the sample, and the sample of the first wash applied to the HisTrap column, were included as checks for contaminants and to observe if any of the still His-tagged NavMs protein was lost at any stage during the purification process prior to His-tag cleavage and SEC. Successful cleavage of the His-tag would cause any fraction samples from the SEC to not show up on the immunoblot since there would not be any His-tag present for the α-poly His horseradish peroxidase-conjugated antibody to recognise. As seen in (Fig 3.4) HisTrap fractions had NavMs bands present, and no contaminant bands, indicating that the His-tagged NavMs sample had undergone major purification during HisTrap purification, and the NavMs protein had not been lost. SEC fractions did not have any visible bands. This coupled with the distinct NavMs banding seen in the SDS-PAGE for SEC fractions indicated that there was purified protein present in the SEC fractions, and this protein had successfully gone through His-tag cleavage, thus was suitable for pooling and concentrating to form a final sample of purified NavMs WT protein at 10 mg/mL. Fractions from the load and flow through were full of histidine rich protein, as indicated by the multiple bands in each. This indicates both that the sample is not pure prior to HisTrap chromatography, and that the HisTrap nickel affinity column did not bind all of the protein in the sample, thus there is a loss of some protein at the stage of loading the HisTrap column, but a lot of this protein was probably contaminants that needed to be removed so as to move towards a pure NavMs WT sample. The wash fraction contained no discernible protein bands, which indicated that no bound protein was lost whilst washing the column.

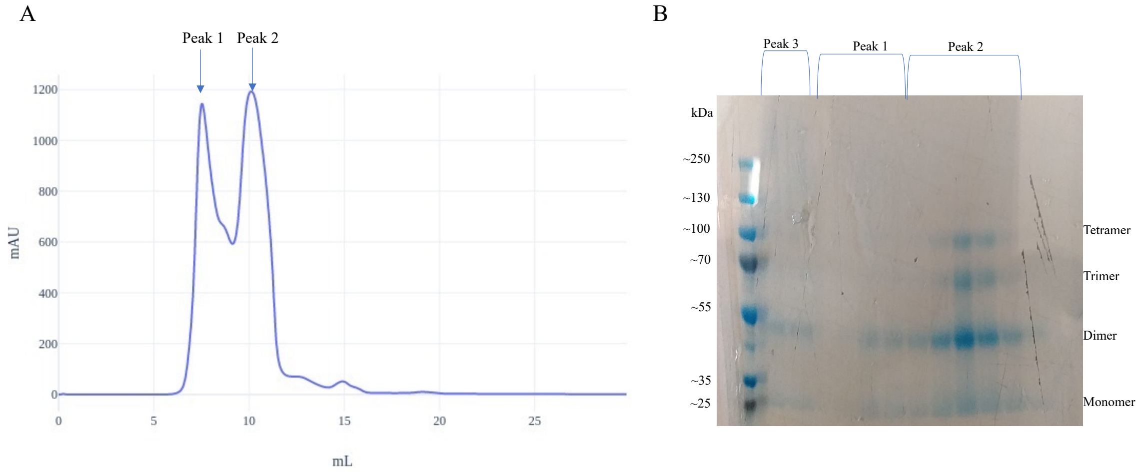


Figure 3.3: UV trace from NavMs WT SEC chromatography and SDS-PAGE of collected SEC fractions. *A) The UV trace from the SEC step of purification indicated two distinct peaks in protein elution, which is characteristic of NavMs purification. Protein from fractions pertaining to the first peak are a different species, contain less secondary structure, and do not crystallise when compared to NavMs protein fractions collected from the second peak. Figure produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA). B) The initial couple of fractions belong to the third peak of the HisTrap affinity chromatography step (see Fig B). Fractions from the first and second peak of the SEC are labelled and showed the characteristic NavMs banding pattern, with banding being stronger in the second peak.*

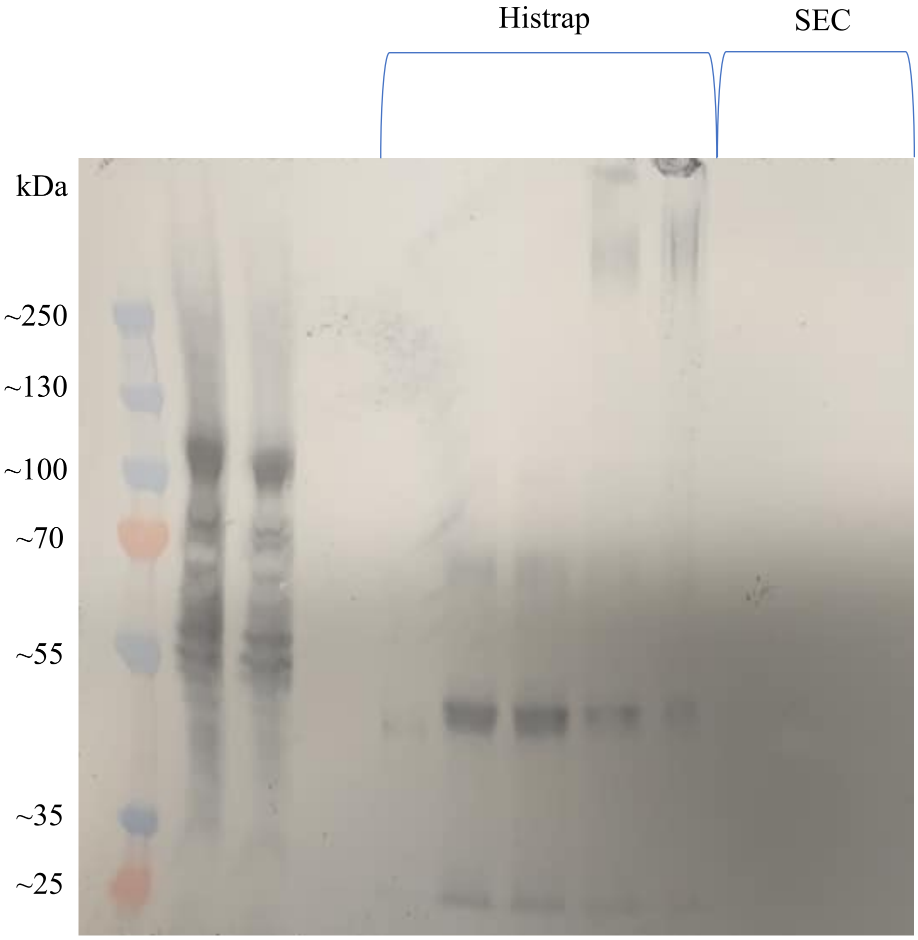


Figure 3.4: NavMs WT immunoblot. *The first three lanes in this gel contained: the load sample before it was loaded on to the HisTrap column, the sample of the flow through after loading the sample, and the sample of the first wash applied to the HisTrap column. The next labelled lanes contained fraction samples from the HisTrap purification step, and then fractions from the SEC purification step.*

3.2.3 Sequencing of NavMs F208L DNA

Plasmid sample for NavMs F208L was also donated by Dr. Altin Sula. This plasmid sample went through plasmid propagation and a consequent plasmid sample was sent for sequencing to confirm the presence of the single F208L mutation (see methods section 2.4) (Fig 3.5). Sequencing results were translated and aligned with the NavMs WT sequence (Uniprot code: A0L5S6), which confirmed the presence of the single F208L mutation. NavMs F208L plasmid was consequently ready for transfection of *E. coli* C41 cells.

```
WT|1-50        MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ
               ||||||||||||||||||||||||||||||||||||||||||||||||||
F208L|1-50     MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ

WT|51-100      LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR
               ||||||||||||||||||||||||||||||||||||||||||||||||||
F208L|51-100   LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR

WT|101-150     TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT
               ||||||||||||||||||||||||||||||||||||||||||||||||||
F208L|101-150  TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT

WT|151-200     NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI
               ||||||||||||||||||||||||||||||||||||||||||||||||||
F208L|151-200  NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI

WT|201-250     PFIMLTTFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD
               |||||||.||||||||||||||||||||||||||||||||||||||||||
F208L|201-250  PFIMLTTLTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD

WT|251-274     RLDRIEKQLAQNNELLQRQQPQKK
               ||||||||||||||||||||||||
F208L|251-274  RLDRIEKQLAQNNELLQRQQPQKK
```

Figure 3.5: Sequencing of the pet15b NavMs F208L sample. *ExPASy translate (Gasteiger et al., 2003) used to translate sequencing returned from Source Bioscience. Translated sequences were then aligned with the NavMs WT sequence from Uniprot (Bateman, 2019) (Uniprot code: A0L5S6) using EMBOSS NEEDLE (Madeira et al., 2019). 'WT' denotes the Uniprot reference sequence, and 'F208L' denotes the sample sequence. Highlighted in red is the position of the point mutation from phenylalanine (F) to leucine (L).*

3.2.4 Expression and purification of NavMs F208L

Refer to methods section 2.5 for details of the expression and purification process.

HisTrap affinity chromatography: As with NavMs WT, a UV trace was collected from the AKTA during the initial HisTrap nickel affinity chromatography stage.

In (Fig 3.6A) there are three distinct UV trace peaks apparent, thus three distinct protein elution events, mirroring that observed for NavMs WT. The peaks are of comparable size to those observed for NavMs WT, which indicated NavMs F208L behaves in a similar fashion to NavMs WT, thus the mutation does not appear to affect expression and initial His-tag purification. Fractions from all three peaks were collected and run on an SDS-PAGE to study the protein content of these fractions. These SDS-PAGE results indicated that banding patterns from all three peaks matched that observed for NavMs WT, confirming similar purification behaviour for the NavMs F208L mutant when compared to NavMs WT (Fig 3.6B).

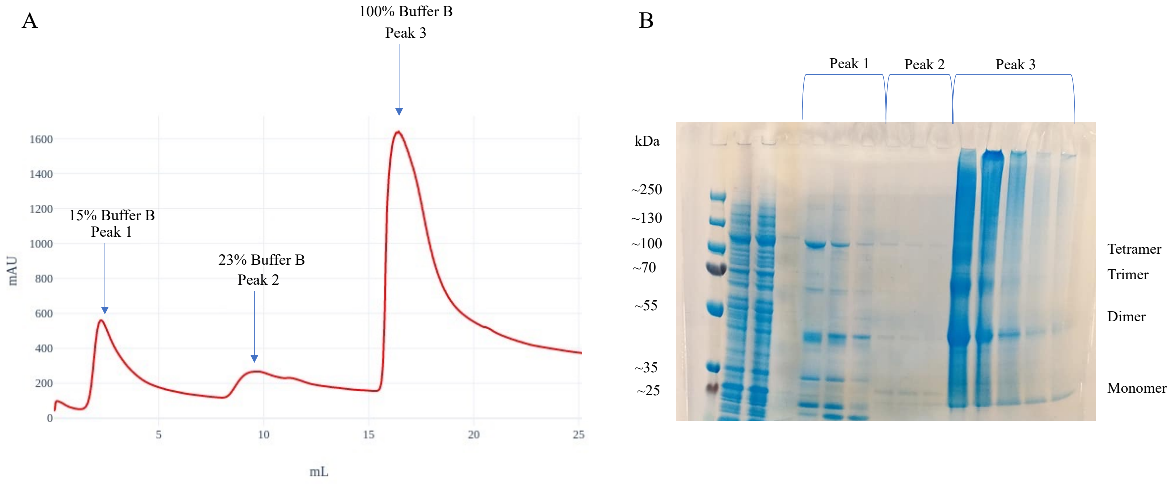


Figure 3.6: UV trace from NavMs F208L HisTrap affinity chromatography and SDS-PAGE of collected HisTrap fractions. *A) UV trace graph indicates the three trace peaks correlated with increases in buffer B percentage. Fractions were taken from each peak and run on an SDS-PAGE. Figure produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA). B) Image of the SDS-PAGE gel containing fractions from the first, second and third peak. The first three lanes contain samples of the protein prior to loading on to the HisTrap column, the flow through from loading the column, and the first wash of the HisTrap column.*

Size exclusion chromatography: The HisTrap fractions from the third peak containing the NavMs F208L were collected, concentrated, put through His-tag cleavage, then loaded on to a Superdex 200 10/300 GL column to carry out SEC. UV traces were collected from the AKTA during SEC for NavMs F208L (Fig 3.7A) to discern fractions containing protein. The UV trace indicated two peaks within the trace. These peaks mirrored peaks observed for NavMs WT, indicating that NavMs F208L behaves in the same manner as NavMs WT during SEC purification.

SDS-PAGE of fractions from the two peaks, along with samples of the protein sample prior to HisTrap column loading, HisTrap column flow through, and HisTrap wash, indicated the same results as those observed for NavMs WT, thus confirming that NavMs F208L purification was successful and was not disrupted by the presence of the F208L mutation, which indicates that this mutation does not significantly affect NavMs folding, since characteristic NavMs banding was maintained (Fig 3.7B). A follow-up immunoblot (Fig 3.8) confirmed that SEC protein fractions had undergone successful

His-tag cleavage, and were not contaminated with random proteins, like equivalent fractions during NavMs WT purification. Purified NavMs F208L fractions from the second SEC peak were pooled, concentrated to 10mg/ml then stored for experiments.

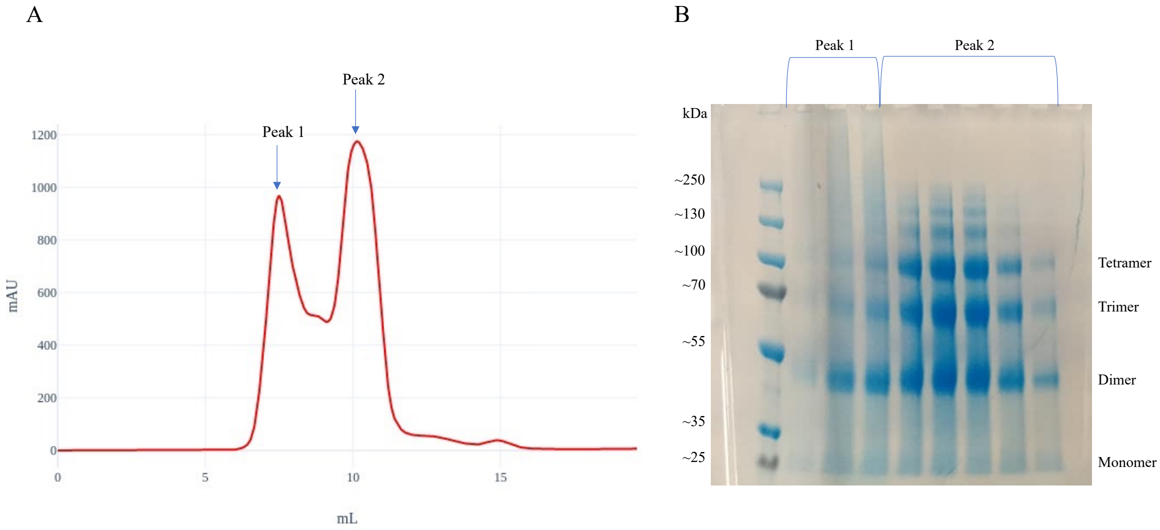


Figure 3.7: UV trace from NavMs F208L SEC chromatography and SDS-PAGE of collected SEC fractions. *A) The UV trace from the SEC step of purification indicated two distinct peaks in protein elution, characteristic of NavMs purification. Figure produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA). B) Fractions from the first and second peak of the SEC are labelled and showed the characteristic NavMs banding pattern, with banding being stronger in the second peak.*

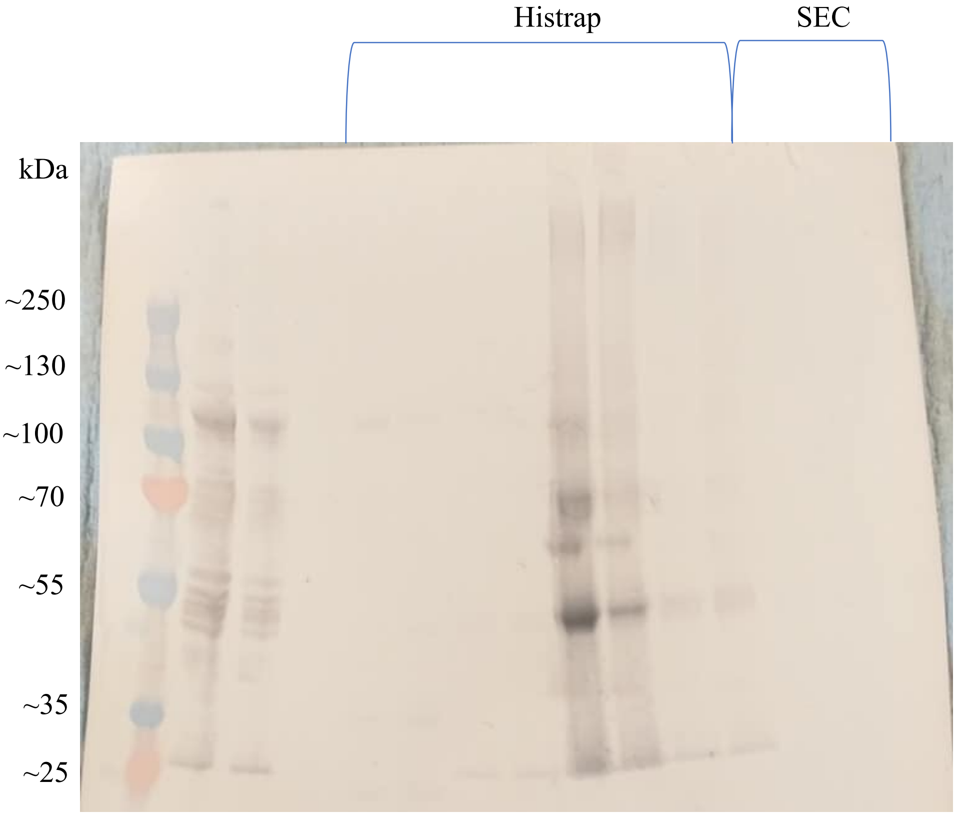


Figure 3.8: NavMs F208L immunoblot. *The first three lanes in this gel contained: the load sample prior to HisTrap column loading, the flow through sample after loading, and the sample of the first wash applied to the HisTrap column. The next labelled lanes contained fraction samples from the HisTrap purification step, and then fractions from the SEC purification step.*

3.2.5 Production of NavMs M204A, T207A and F214A mutant plasmid constructs

Plasmid constructs for M204A, T207A and F214A had to be successfully cloned to allow for subsequent expression and purification. M204I and M204W mutant plasmid were donated by Dr. Altin Sula, thus cloning was not required for these mutants.

SLIM PCR: Primers that would introduce the point mutations necessary to create M204A, T207A, and F214A protein mutants via SLIM PCR were designed and used within the SLIM PCR protocol (see methods section 2.3). Samples of the PCRs were subsequently loaded on to agarose gels and imaged to observe whether there was DNA product. For all three mutants, there was a clear band, indicating PCR product (Fig 3.9). Consequently, these PCRs were digested with Dpn1 to remove parent methylated DNA and ensure only the newly synthesised DNA product remained. These DNA samples were then ready for use within *E. coli* transformations, and the subsequent DNA miniprepping and sequencing, to confirm whether the single point mutations were present.

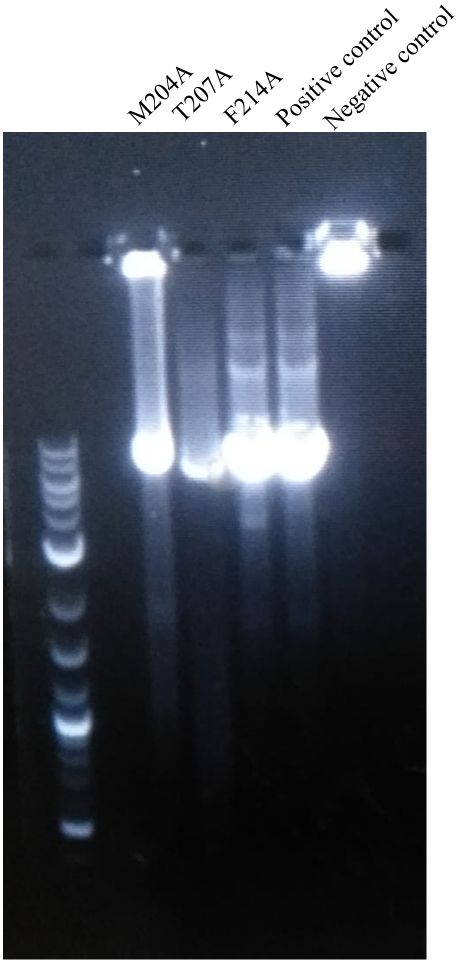


Figure 3.9: SLIM PCR gel electrophoresis. *SLIM PCR samples for M204A, T207A and F214A, along with positive and negative controls. All SLIM PCRS had a definite strong band, indicating plasmid amplification.*

Sequencing NavMs M204A, T207A and F214A DNA samples: Miniprepped DNA samples from successful *E. coli* transformations that used the SLIM PCR products for M204A, T207A, and F214A were sent for sequencing to confirm that the DNA samples only contained the desired point mutation. It was confirmed via sequencing that the NavMs M204A, T207A, and F214A mutant plasmids all only contained the point mutation of interest (Fig 3.10-3.12), and thus could be stored and used within transformation, expression, and purification trials (along with the NavMs M204I and M204W plasmids which were already created and stored by the Wallace lab).

```
WT|1-50          MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
M204A|1-50       MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ

WT|51-100        LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
M204A|51-100     LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR

WT|101-150       TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
M204A|101-150    TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT

WT|151-200       NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
M204A|151-200    NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI

WT|201-250       PFIMLTTFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD
                 |||.||||||||||||||||||||||||||||||||||||||||||||||
M204A|201-250    PFIALTTFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD

WT|251-274       RLDRIEKQLAQNNELLQRQQPQKK
                 ||||||||||||||||||||||||
M204A|251-274    RLDRIEKQLAQNNELLQRQQPQKK
```

Figure 3.10: Sequencing of Pet15bNavMs M204A sample. *ExPASy translate (Gasteiger et al., 2003) was used to translate the sequencing returned from Source Bioscience. Translated sequences were then aligned with the NavMs WT sequence from Uniprot (Bateman, 2019) (Uniprot code: A0L5S6) using EMBOSS NEEDLE (Madeira et al., 2019). 'WT' denotes the Uniprot reference sequence, and 'M204A' denotes the sample sequence. Letters in red highlight the M204A point mutation.*

```
WT|1-50          MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
T207A|1-50       MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ

WT|51-100        LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
T207A|51-100     LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR

WT|101-150       TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
T207A|101-150    TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT

WT|151-200       NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
T207A|151-200    NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI

WT|201-250       PFIMLTTFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD
                 ||||||.|||||||||||||||||||||||||||||||||||||||||||
T207A|201-250    PFIMLTAFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD

WT|251-274       RLDRIEKQLAQNNELLQRQQPQKK
                 ||||||||||||||||||||||||
T207A|251-274    RLDRIEKQLAQNNELLQRQQPQKK
```

Figure 3.11: Sequencing of Pet15bNavMs T207A sample. *ExPASy translate (Gasteiger et al., 2003) was used to translate the sequencing returned from Source Bioscience. Translated sequences were then aligned with the NavMs WT sequence from Uniprot (Bateman, 2019) (Uniprot code: A0L5S6) using EMBOSS NEEDLE (Madeira et al., 2019). 'WT' denotes the Uniprot reference sequence, and 'T207A' denotes the sample sequence. Letters in red highlight the T207A point mutation.*

```
WT|1-50          MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
F214A|1-50       MSRKIRDLIESKRFQNVITAIIVLNGAVLGLLTDTTLSASSQNLLERVDQ

WT|51-100        LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
F214A|51-100     LCLTIFIVEISLKIYAYGVRGFFRSGWNLFDFVIVAIALMPAQGSLSVLR

WT|101-150       TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
F214A|101-150    TFRIFRVMRLVSVIPTMRRVVQGMLLALPGVGSVAALLTVVFYIAAVMAT

WT|151-200       NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
F214A|151-200    NLYGATFPEWFGDLSKSLYTLFQVMTLESWSMGIVRPVMNVHPNAWVFFI

WT|201-250       PFIMLTTFTVLNLFIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD
                 |||||||||||||.||||||||||||||||||||||||||||||||||||
F214A|201-250    PFIMLTTFTVLNLAIGIIVDAMAITKEQEEEAKTGHHQEPISQTLLHLGD

WT|251-274       RLDRIEKQLAQNNELLQRQQPQKK
                 ||||||||||||||||||||||||
F214A|251-274    RLDRIEKQLAQNNELLQRQQPQKK
```

Figure 3.12: Sequencing of Pet15bNavMs F214A sample. *ExPASy translate (Gasteiger et al., 2003) was used to translate the sequencing returned from Source Bioscience. Translated sequences were then aligned with the NavMs WT sequence from Uniprot (Bateman, 2019) (Uniprot code: A0L5S6) using EMBOSS NEEDLE (Madeira et al., 2019). 'WT' denotes the Uniprot reference sequence, and 'F214A' denotes the sample sequence. Letters in red highlight the F214A point mutation.*

3.2.6 Expression and purification of NavMs M204, T207 and F214 mutants

HisTrap affinity chromatography of the mutant protein: After expressing NavMs M204A, M204I, M204W, T207A, and F214A, then collecting membranes from these expressions, purification steps were run according to the protocol laid out in methods section 2.5. UV traces were collected from the AKTA during the HisTrap affinity chromatography for all of the mutants. As with NavMs WT and NavMs F208L, there were three distinct UV trace peaks apparent per mutant, with each peak correlating to a gradient increase in buffer B, and thus an increase in imidazole (15 % buffer B for the first peak, 23 % buffer B for the second peak and 100 % buffer B for the third and largest peak). From previous work within the Wallace group, it has been identified that UV traces for the HisTrap purification process indicate the presence of NavMs protein within the third peak. Fractions from all three peaks for each mutant were collected and run on an SDS-PAGE to study protein content and identify NavMs protein within the fractions taken from the third peak. When comparing all of the mutant traces, along with the NavMs WT and NavMs F208L UV traces (Fig 3.13), there is a variation in sizes for the first peak and second peak, indicating varying levels of contaminants caused by non-specific binding to the HisTrap column. The third peak across all of the samples shows slight variation in magnitude (ranging from ~1500 – 1800 milli-absorbance units (mAU)), which could suggest slight differences in protein produced via expression, but not enough of a difference that one can say, at least at this stage, that any of the point mutations appeared to affect protein expression, and the ability to undergo HisTrap purification. The peaks in (Fig 3.13) are not completely aligned due to the step ups in buffer B gradient being applied at slightly different points during the HisTrap purification step for each mutant.

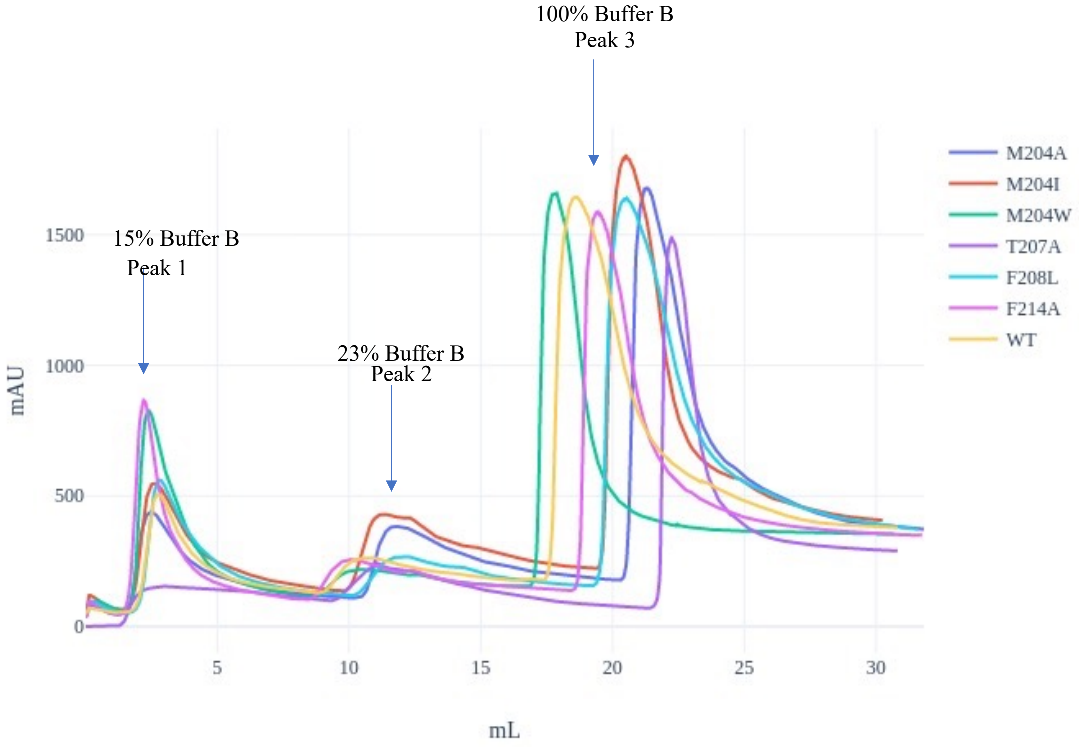


Figure 3.13: UV trace from NavMs HisTrap affinity chromatography. *This graph indicates the three UV trace peaks per mutant which occurred during the HisTrap affinity chromatography step,as well as the peaks for NavMs WT and NavMs F208L, which are included for reference. Figure produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).*

SDS-PAGE gels were run, with the fractions collected from each of the three peaks for all the mutants (Fig 3.14-3.15). For all the mutants, banding patterns were characteristic of NavMs purification, mirroring that observed for NavMs WT and NavMs F208L. Fractions from the third peak for all mutants (Fig 8) contained bands at positions ~25 kDa, ~50 kDa, ~70 kDa and ~100 kDa, which are known to indicate NavMs protein and its multimers (monomer, dimer, trimer, and tetramer respectively). This banding indicates that NavMs protein stability and the ability to form multimers was not affected by any of the point mutations.

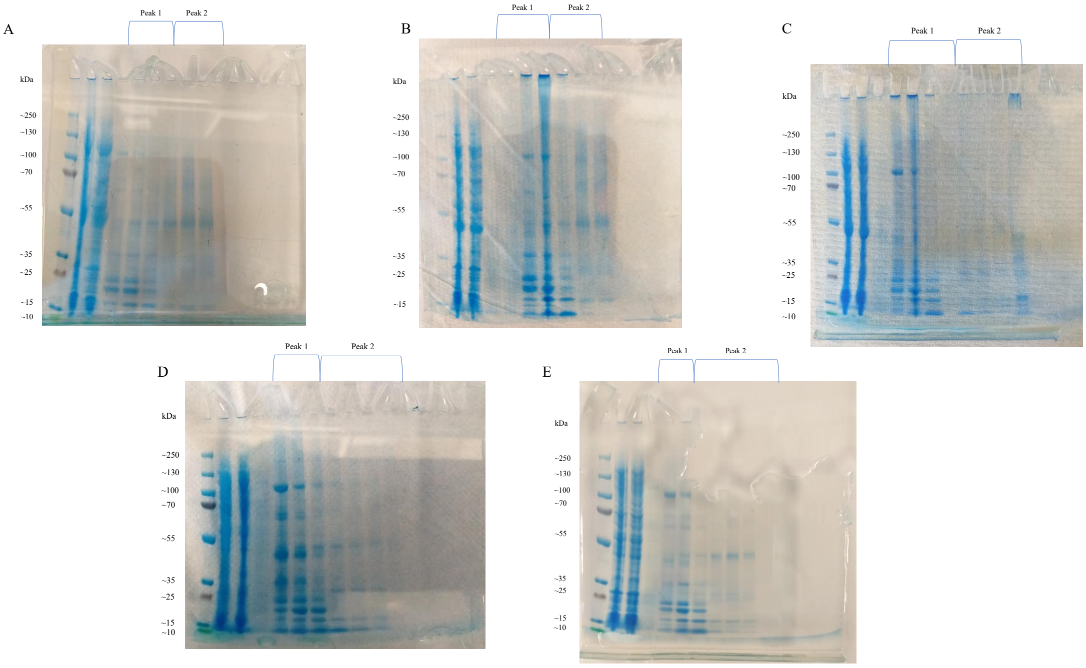


Figure 3.14: Fractions from the first and second peak relating to the HisTrap affinity chromatography. *For all of the images, the first three lanes correspond to (from left to right) the load sample before it was loaded on to the HisTrap column, a sample of the flow through after loading the sample, and a sample of the first wash applied to the HisTrap column. These first three samples acted to check if protein was lost at any steps prior to the HisTrap affinity chromatography. In addition, for all images, fractions from the first peak and second peak are indicated. For all of the images, and thus mutants, these first and second peak fractions appear to mostly consist of a range of different protein contaminants of varying size. A) Gel for NavMs M204A. B) Gel for NavMs M204I. C) Gel for NavMs M204W. D) Gel for NavMs T207A. E) Gel for NavMs F214A.*

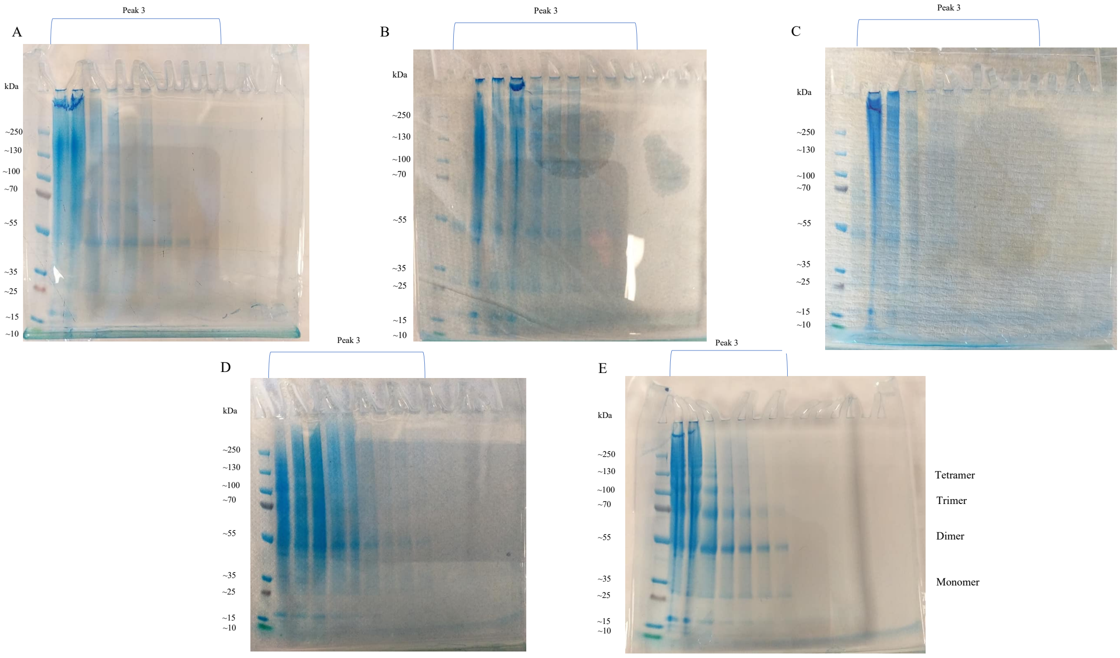


Figure 3.15: Fractions from the third peak relating to the HisTrap affinity chromatography. *For all of the NavMs mutant SDS-PAGE gels shown, the third peak fractions contained the characteristic banding pattern of NavMs protein, demonstrating that all of the NavMs mutants maintained the distinctive banding pattern associated with NavMs WT. The presence of banding signified that the ability of NavMs to form multimeric states is not disrupted by any of the point mutations. Streaking and mismatch between band location and true protein size was observed but this is a known property of running membrane proteins (Rath et al., 2009; Zhou et al., 2000). A) Gel for NavMs M204A. B) Gel for NavMs M204I. C) Gel for NavMs M204W. D) Gel for NavMs T207A. E) Gel for NavMs F214A.*

Size exclusion chromatography (SEC) of mutant protein: HisTrap fractions from the third peak for all mutants were pooled, concentrated, underwent His-tag cleavage, and were then separately loaded on to a Superdex 200 10/300 GL column to carry out SEC. UV traces were collected from the AKTA during SEC (Fig 3.16). All SEC UV traces for the mutants had two peaks within the trace. The first peak started at ~7 mL, and the second peak started at ~9 mL. Both the appearance of a double peak, and the volumes at which peaks started to appear, align with what is observed for NavMs WT both within this research and within previous characterisation carried out within the Wallace lab. This indicated that the general behaviour of NavMs during the SEC process had gone unaffected by the point mutations. As with the His-tag affinity chromatography involving the mutants, no significant difference was observed between the traces for the mutants, or when compared to NavMs WT and NavMs F208L traces, indicating that these point mutations cause no significant effect on expression or purification. One difference was observed between the traces was that of the varying ratios between the first and second peaks for the different mutants, with some mutants seeing higher relative first peaks than second peaks, such as in the case of NavMs F214A. An increased relative magnitude in the first peak would reduce the end yield of usable purified protein, since a larger amount of the NavMs protein ended up in these first peak fractions, which, as already stated, are not as suitable for crystallographic trials. With this said, all mutants did produce purified protein fractions which aligned with the second peak. This protein was pooled, concentrated, and then stored.

SDS-PAGE gels were run for all mutants, with fractions from both the first and second peaks allowing for investigation of the fractions and the confirmation of purified NavMs protein presence (Fig 3.17). As with NavMs WT (see Fig 3.6-3.7), the first and second peaks contained the distinctive banding pattern associated with NavMs.

Finally, to confirm that protein for all of the mutants from the end of the SEC process had successfully gone through His-tag cleavage, immunoblots were carried out with fraction samples collected from the HisTrap affinity chromatography and SEC purification steps (Fig 3.18). For all point mutants, HisTrap fractions had NavMs bands present, but SEC fractions did not have any visible bands, confirming that protein from the SEC fractions had successfully gone through His-tag cleavage.

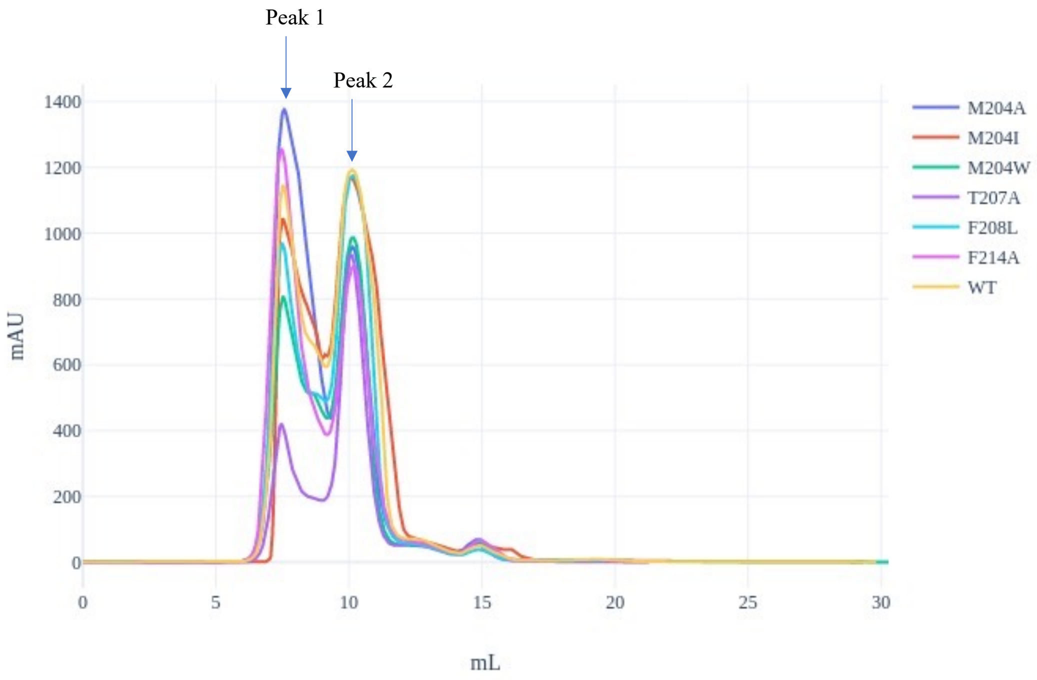


Figure 3.16: SEC UV trace for NavMs mutants. *The UV traces from SEC purification of all mutants, along with the SEC traces for NavMs WT and NavMs F208L for reference. SEC traces indicated two distinct peaks in protein elution for all mutants, which is characteristic of NavMs purification. Figure produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).*

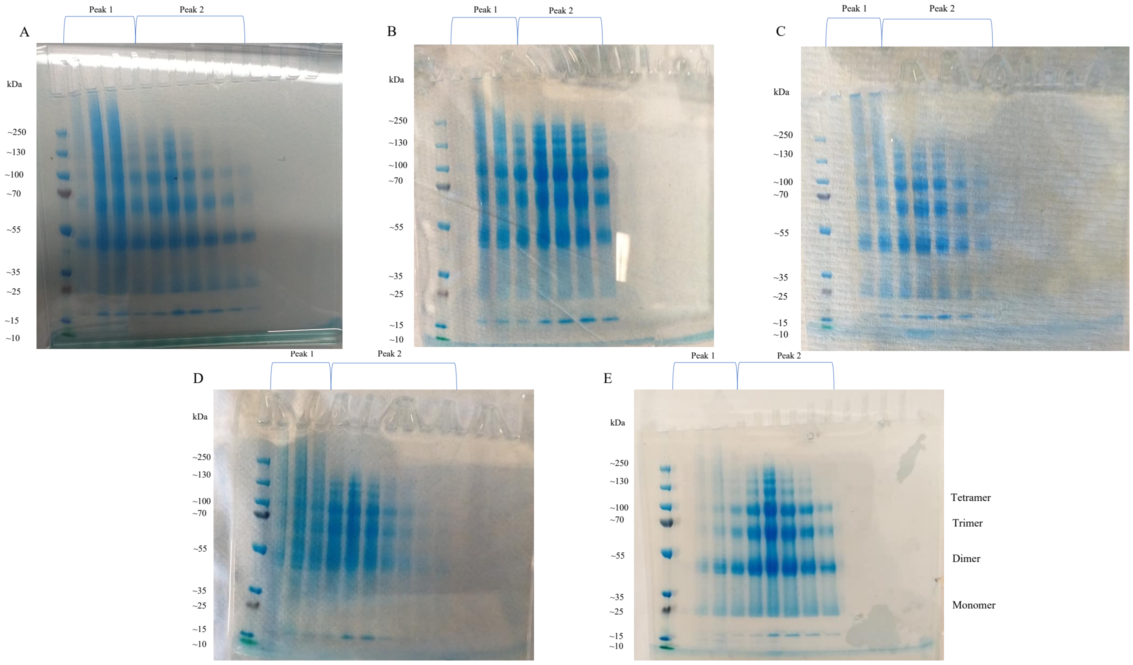


Figure 3.17: SDS-PAGE of NavMs mutant SEC purifications. *Fractions from the first and second peak of the SEC are labelled in all gel images. Fractions from the first and second peak for all of the mutants showed the characteristic NavMs banding pattern, signifying the presence of NavMs and that none of the mutants affected the characteristics of NavMs banding, and thus the folding and formation of multimers. Fractions from the second peak for all mutants were pooled and concentrated to form the final purified NavMs mutant protein. A) Gel for NavMs M204A. B) Gel for NavMs M204I. C) Gel for NavMs M204W. D) Gel for NavMs T207A. E) Gel for NavMs F214A.*

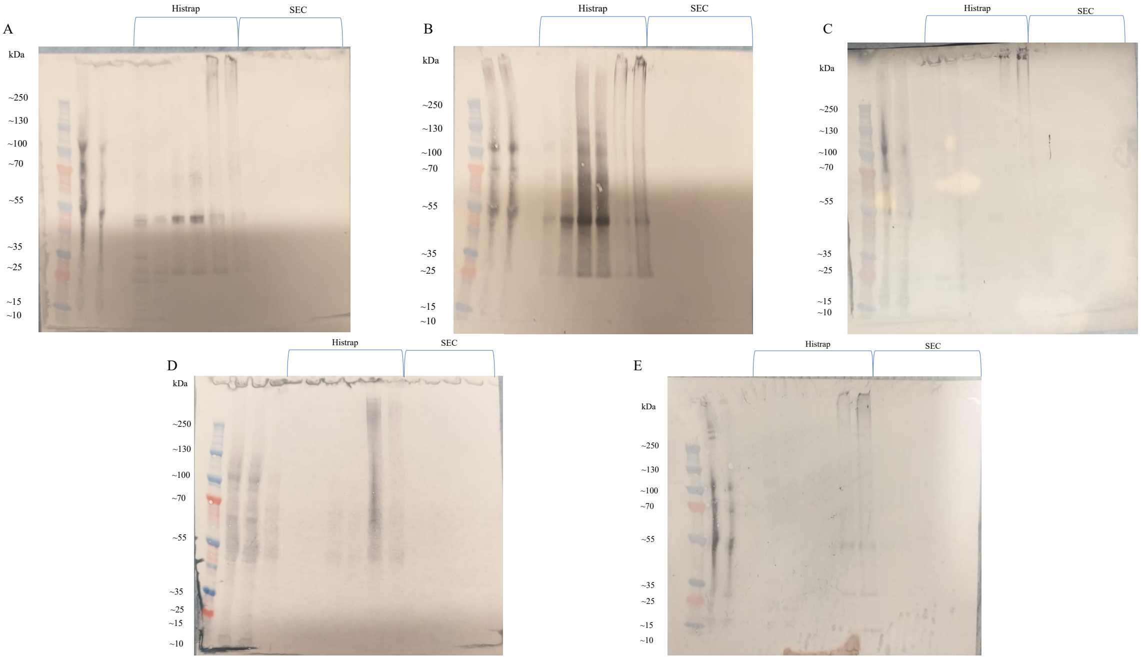


Figure 3.18: NavMs mutant immunoblots. *For all of the images, the first three lanes correspond to (from left to right) the load sample before it was loaded on to the HisTrap column, a sample of the flow through after loading the sample, and a sample of the first wash applied to the HisTrap column. The wash sample was included, along with the load and flow through samples, to act as a check to see if protein was being lost at any stage. The next labelled lanes contained fraction samples from the HisTrap purification step, and then fraction from the SEC purification step. Banding was not visible for SEC fractions, but banding was clear for the load sample, the flow through and for the HisTrap fraction samples. A) Gel for NavMs M204A. B) Gel for NavMs M204I. C) Gel for NavMs M204W. D) Gel for NavMs T207A. E) Gel for NavMs F214A.*

2.5 Discussion:

The first aim was to express and purify NavMs WT and NavMs F208L protein in order to have correctly folded protein in sufficient levels, so as to be able to carry out crystallisation experiments with CBD. NavMs WT and NavMs F208L protein were successfully expressed and purified using a protocol which had previously been optimised within the Wallace lab, and had allowed for earlier research involving the NavMs protein (Ireland *et al*., 2018; Montini *et al*., 2018; Sula *et al*., 2017; Sula and Wallace, 2017; Zanatta *et al*., 2019). Looking at the AKTA UV traces, SDS-PAGE gels, and immunoblots, it was confirmed that purified NavMs protein was produced and purified, since at each step, the traces, SDS-PAGE gels, and immunoblots were characteristic of NavMs protein. In addition, the immunoblots allowed for confirmation that the final purified protein had successfully undergone His-tag cleavage, meaning that the protein was ready for use within further experiments. Comparing the AKTA UV traces and subsequent SDS-PAGE gels for the NavMs WT and NavMs F208L protein highlighted that the F208L mutation did not affect how the protein behaved. This is shown by the fact that: 1: the same elution patterns are observed for the His-tag affinity chromatography and the SEC for NavMs WT and NavMs F208L, 2: The characteristic SDS-PAGE banding pattern associated with NavMs protein and its multimeric states was unaltered for NavMs F208L, indicating that the protein was still able to fold and form multimers like NavMs WT. In addition, the F208L mutation did not appear to affect the protein's expression levels. This suggested that F208L mutation does not appear to affect *E. coli* expression of the NavMs protein, meaning that appropriate levels of this mutated protein can be produced via the previously optimised expression and purification protocol, and could be used as a protein within further crystallisation trials with CBD.

The second aim was to produce mutant NavMs protein which could be used to investigate the hypothesis of CBD binding and entry into the sodium ion channel being similar to that previously observed for other sodium ion channel blockers. All mutants were successfully produced using a protocol which had been previously optimised within the Wallace lab, and looking at the AKTA UV traces, SDS-PAGE gels, and immunoblots for all mutants in comparison to NavMs WT (Fig 3.13-3.18), it was confirmed that the ability to produce purified NavMs protein via the previously characterised protocol was unaffected by the presence of any of the point mutations. Comparing the AKTA UV traces and subsequent SDS-PAGE gels for the NavMs mutants with those observed for NavMs WT highlighted that all of the mutations did not affect NavMs protein characteristics during purification, since elution patterns were the same across both HisTrap and SEC, and NavMs banding was observed across all gels. Overall, no mutations appeared to affect the protein's initial expression levels. However, there are differences in ratio between the first and second peak during the SEC stage, which does lead to a greater loss of end usable protein for some of the mutants (such as NavMs F214A) (Fig 3.16), since it has been previously suggested that protein from the second peak is more suitable for use within experiments, such as crystallographic studies and circular dichroism. With this in mind, purified NavMs protein was still collected for each mutant in levels that would allow for future experimentation, proving that mutants do not appear to affect the overall stability or characteristics of NavMs during the expression and purification stages, and that the production of these mutants does not require any optimisation of the expression or purification process.

Chapter 4: Circular Dichroism Spectroscopy of NavMs Proteins

4.1 Introduction:

Circular Dichroism spectroscopy uses circularly polarised light absorbance by optically active molecules to gain information about structural characteristics. As mentioned in (Chapter 1: Literature review) CD can be used for studying protein secondary structures (Greenfield, 2007; Whitmore and Wallace, 2008), stability (Ireland *et al*., 2018; Lai *et al*., 2018), dynamics (Huck *et al*., 2020; Linares *et al*., 2019), folding pathways (Alamry *et al*., 2020; Lamba *et al*., 2009; Salamanca and Chang, 2005), and interactions, including protein-protein interactions, protein-ligand interactions, and interactions between proteins and metal ions (Hosseinzadeh *et al*., 2021; Jin *et al*., 2021; Manieri *et al*., 2021; Najar *et al*., 2021; Zanatta *et al*., 2019).

Thermal melt CD (recording changes in CD as a function of temperature) is used to study effects of mutations, ligands, and protein environments on protein stability. Results from thermal melt CD can give insight into protein stability, offer insights on protein-ligand interactions and optimal protein environments. Thermal melt CD has been implemented within studies of proteins, including NavMs (Di Lella *et al*., 2010; Fan *et al*., 2007; Güler *et al*., 2020; Huus *et al*., 2006; Ireland *et al*., 2018; Marada *et al*., 2016; Metrick *et al*., 2013; Szenczi *et al*., 2006).

The aim of utilising CD, specifically thermal melt CD, was to analyse NavMs WT protein stability with and without CBD. These observations indicated if the NavMs WT protein was generally stable and folded with and without CBD present, thus signifying if NavMs was still a feasible model for crystallographic studies. These studies were also used to investigate whether the addition of CBD had any impact on the stability of the NavMs protein when compared to the protein without CBD, thus giving an indicator of binding interactions.

Please refer to methods section 2.6 for details surrounding thermal melt CD experiments, and experimental data processing.

4.2 Results:

4.2.1 GDN buffer transfer of NavMs WT protein

Initial trials to transfer NavMs WT protein from HEGA-10 into GDN buffer were set up, since HEGA-10 cannot be used during CD experiments because HEGA-10 causes its own unique absorption pattern within the wavelength range of interest, making results difficult to interpret. Trials began with 50 µL aliquots of NavMs WT at 10 mg/mL being loaded on to a Superdex 200 10/300 column equilibrated with GDN transfer buffer (Fig 4.1). Overall, these initial trials indicated that this method of transferring NavMs WT was neither reproducible nor effective, since most of the protein was lost in the column, possibly due to aggregation.

Next, a 200 µL aliquot of NavMs WT at 10 mg/mL was used to see if the loss of protein was a fixed amount, or whether the loss of protein was proportionate with the amount of initial NavMs WT protein used (Fig 4.2). A peak eluting at around 10 mL, indicative of purified NavMs protein, reaching ~90 mAU, was observed. Fractions from this peak were collected and concentrated to 1 mg/mL and used in subsequent CD experiments.

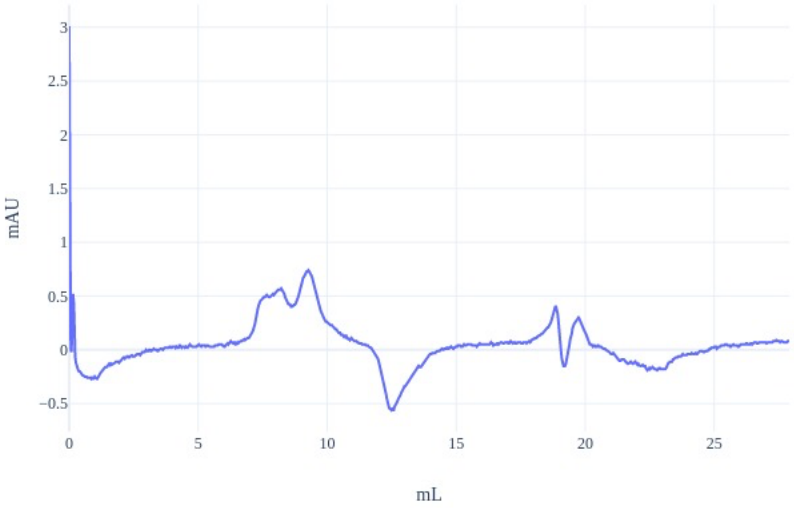


Figure 4.1: AKTA trace from GDN buffer exchange trials using 50 µL NavMs WT aliquot at 10 mg/mL. *GDN buffer exchange where the entirety of the NavMs WT protein sample appeared to be lost in the column, indicating loss due to aggregation. Produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).*

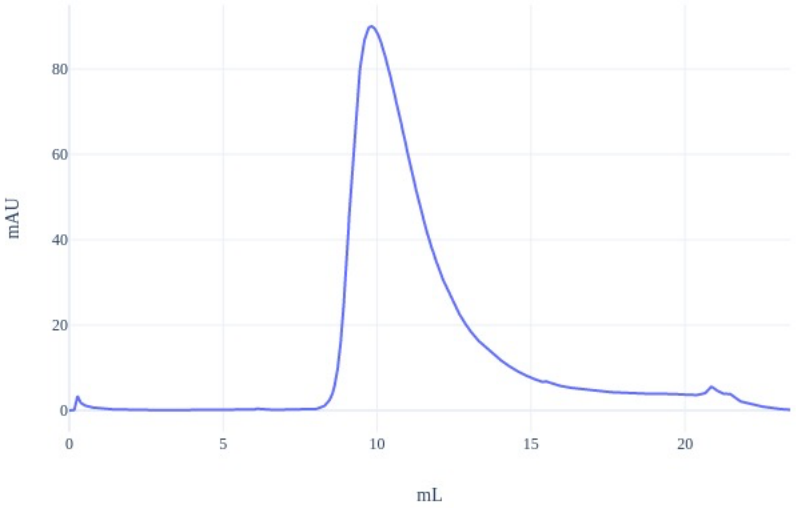


Figure 4.2: AKTA trace from GDN buffer exchange using 200 µL of NavMs WT at 10 mg/mL. *A peak was observed at around 10 mL, which is indicative of correctly folded NavMs protein. This trial with a larger volume of NavMs WT indicated that during the GDN buffer transfer process a fixed amount of protein is lost, possibly due to aggregation. This fixed amount of protein lost makes up a smaller percentage of the larger initial NavMs WT sample, and thus an appreciable volume of successfully transferred protein could be collected and concentrated to 1 mg/mL for CD experiments. Produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).*

4.2.2 Optimisation of NavMs protein concentration for CD measurements

When light passes through a sample and hits the detector in a CD instrument, a current is induced. However, the CD instrument maintains a constant current by altering the voltage, otherwise known as the high-tension voltage (HT) value. At lower wavelengths, the sample's total absorbance will increase, causing the level of light reaching the detector to decrease and the HT value to increase (Greenfield, 2007). It has been suggested that for CD experiments involving membrane proteins any wavelength data where the HT value reaches over 500 mV should be discounted, since at this point the data are often too noisy and unreliable. Consequently, the next step was to optimise NavMs concentration for CD experiments, since protein concentration can affect the HT value during CD experiments (Greenfield, 2007; Miles and Wallace, 2016).

Optimisation was carried out using ten-scan runs over wavelengths 260-190 nm in 1 nm steps, and at concentrations of ~0.9 mg/mL, ~0.7 mg/mL, ~0.6 mg/mL, and ~0.5 mg/mL. All concentrations led to CD spectra with similar magnitude peaks at ~194 nm, ~208 nm and ~223 nm (Fig 4.3). Peaks at these wavelengths suggest that the protein is folded and mostly α-helical, which aligns with the NavMs crystal structure (PDB ID: 5HVX) (Sula *et al.*, 2017). The NavMs protein concentrations correlated with their respective HT values (Fig 4.3). 0.9 mg/mL had the greatest rate of HT increase, with the HT value reaching 500 mV at ~199 nm. For 0.7 mg/mL, 500 mV was reached at ~198 nm. At 0.6 mg/mL, the HT reached 500 mV at ~194 nm, and at 0.5 mg/mL it reached 500 mV at ~192 nm. Consequently, a concentration of 0.5 mg/mL was used for further CD experiments, since at this concentration the maximum range wavelength data could be included in later data analysis.

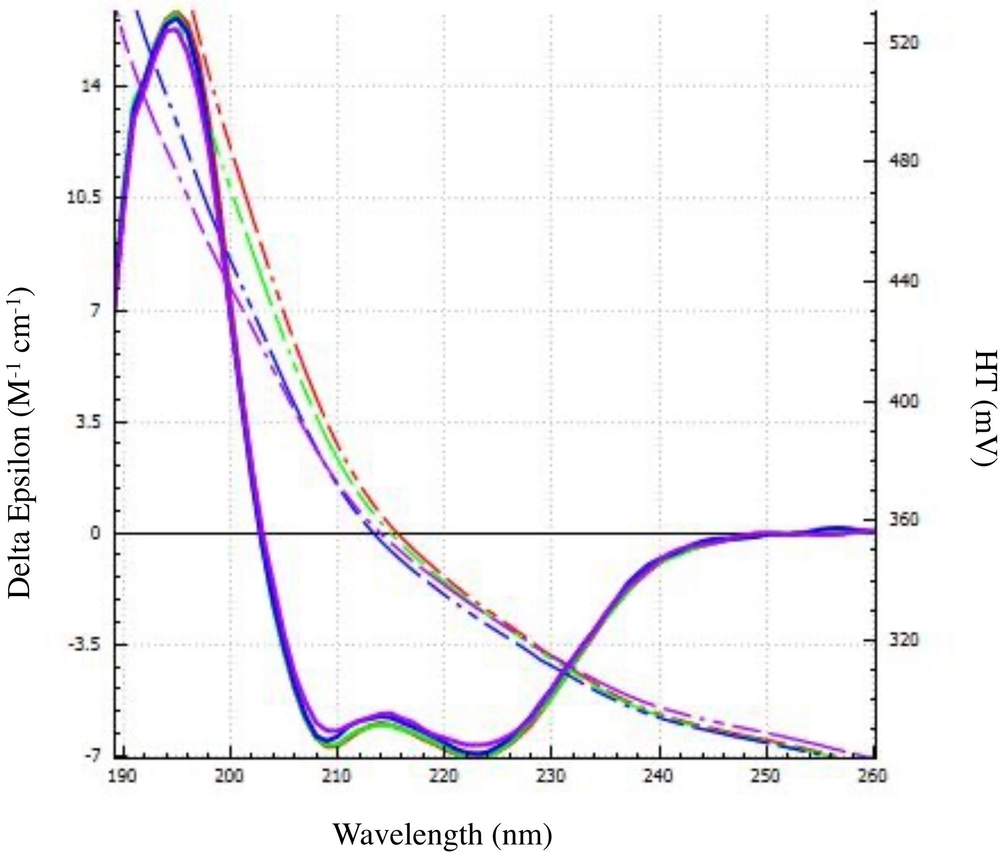


Figure 4.3: CD spectra of NavMs WT samples at differing concentrations. *Processed spectra for NavMs WT protein samples in GDN buffer at 0.9 mg/mL (red), 0.7 mg/mL (green), 0.6 mg/mL (blue), and 0. 5 mg/mL (purple). All spectra exhibited the same peaks at ~194 nm, ~208 nm and ~223 nm, indicating folded, mostly α-helical structures, which corresponds with the NavMs crystal structure (PDB ID: 5HVX). Produced using CDToolX (Miles and Wallace, 2018).*

4.2.3 Comparison of NavMs WT calculated secondary structures under all conditions

At 20°C for NavMs WT, all conditions (without CBD, with CBD and ethanol carrier control) produced peaks at ~194 nm, ~208 nm, and ~223 nm, indicating that the NavMs WT protein maintained its α-helical and folded structure in the presence of CBD and in the ethanol control (Fig 4.4). According to calculated α-helix content, all conditions at 20°C had average α-helix percentages that are in close accordance with that calculated for the NavMs WT crystal structure (~67 %) (Table 4.1) (Fig 4.5). At increasing temperatures, the calculated α-helical content of all the samples decreased at almost identical rates (Fig 4.5). These results suggest ethanol does not have an overall significant effect on NavMs WT secondary structure, thus can be discounted from causing any effects observed with the CBD condition. However, looking at (Fig 4.4), it is apparent that the peak at ~223 nm was slightly but significantly larger for the ethanol sample compared to the other samples. This could suggest some form of effect by ethanol, however, when one considers that peaks occurred at the same wavelengths, and that calculated α-helix content was comparable across all conditions, this difference in peak magnitude at ~223 nm is most likely due to general experimental error.

In previous studies it had been observed that sodium channel binding drugs appear to significantly stabilise or destabilise the structure (Charalambous *et al*., 2009; Nurani *et al*., 2008; O'Reilly *et al*., 2008; Zanatta *et al*., 2019). This can be indicated via significant differences in α-helix content (calculated from CD scan data) at intermediate thermal melt temperatures for samples which contain the drug versus samples which do not (Zanatta *et al*., 2019). The addition of CBD did not cause any significant differences in α-helix content at any temperature points in the thermal melt experiments, which indicates that CBD interactions did not significantly stabilise or destabilise the NavMs WT protein as a function of temperature. Taking the values from (Table 4.1) (assuming maintained standard deviation (SD) values for each respective sample, and going to one decimal place during calculations), for there to have been a significant difference in α-helix content for NavMs with CBD added (according to +/- 1 SD errors), when compared to NavMs without CBD and with ethanol samples at 60°C, the average α-helix percentage would have to have been either ~5.4 % higher in value (68.3 % +/- 2.1 % versus 62.9 % +/- 2.1 %) or ~5.4 % lower (57.5 % +/- 2.1 % versus 62.9 % +/- 2.1 %). At 70°C, the average α-helix content for the NavMs WT sample with CBD added would have needed to be either ~4.5 % higher in value (57.4 % +/- 3.0 % versus 52.9 % +/- 3.0 %) or ~3.4 % lower (49.5 % +/- 3.0 % versus 52.9 % +/- 3.0 %) in value.

It is notable that even at the highest temperature (80°C) all samples still had high α-helical contents, an indicator of the highly stable nature of the NavMs protein structure. This NavMs stability has been observed previously (Ireland *et al*., 2018; Zanatta *et al*., 2019). The back-calculated spectrum derived from the calculated secondary structures at all temperatures had normalised root-mean-square deviation (NRMSD) values < 0.1, indicating that there was close correspondence between the measured data and the back-calculated spectra derived from the calculated secondary structures.

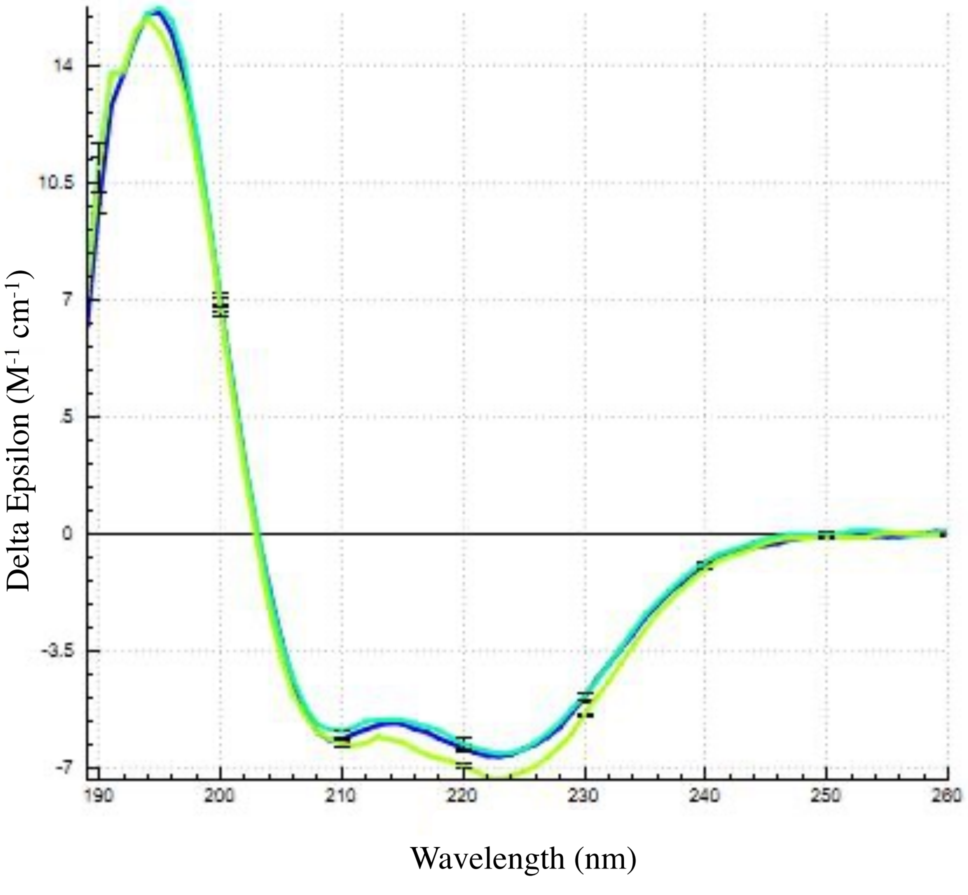


Figure 4.4: Processed CD spectra of NavMs with CBD, without CBD and with ethanol. *Processed spectra for NavMs WT without CBD (dark blue), NavMs WT with CBD (turquoise) and with NavMs WT ethanol (green). All samples have peaks at ~ 194 nm, ~208 nm and ~223 nm, indicating a folded and mostly α-helical secondary structure. Error bars represent 1 SD. Produced using CDToolX (Miles and Wallace, 2018).*

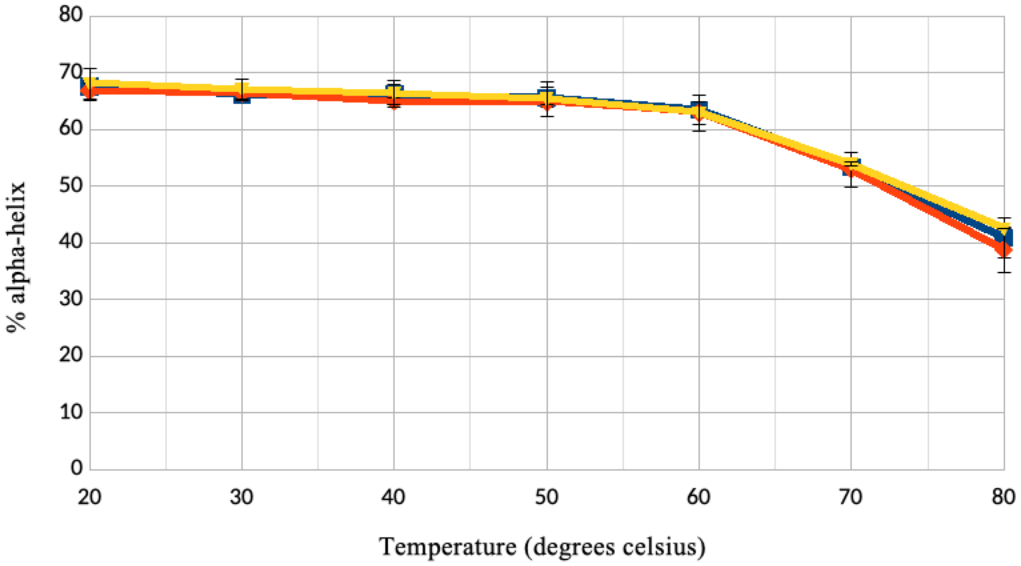


Figure 4.5: Calculated α-helix content of NavMs WT in respect to temperature without CBD, with CBD and with ethanol. *Calculated average α-helix percentage values for NavMs WT samples without CBD (blue), with CBD (red) and with ethanol (yellow). All samples show comparable calculated α-helix fraction values. Plots are shown with error bars corresponding to 1 SD. Produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).*

Temperature (Degrees Celsius)	Without CBD (%)	With CBD (%)	With ethanol (%)
20	67.5 +/- 0.7	66.8 +/- 1.8	68.1 +/- 2.8
60	63.5 +/- 0.4	62.9 +/- 2.1	62.9 +/- 3.2
70	53.4 +/- 0.8	52.9 +/- 3.0	53.9 +/-0.4
80	40.9 +/- 3.5	38.7 +/- 4.0	42.5 +/- 0.1

Table 4.1: Average calculated α-helix percentages for NavMs WT samples. *Table with the calculated average α-helix content percentages for samples without CBD, with CBD and with ethanol (1 decimal place). All samples had average α-helix percentages at 20°C comparable to the α-helix percentage calculated for the crystal structure of NavMs (~67 %) (PDB ID: 5HVX). There were no evident differences in α-helix percentages at intermediate temperatures (60°C and 70°C) when CBD was added, indicating that CBD did not have a stabilising or destabilising effect on the structure. The α-helix percentages at 80°C are comparable and still appreciably high, indicating the stability of the NavMs protein structure. α-helix content percentage values are shown with +/- 1 SD for each value. All values are shown to 1 decimal place.*

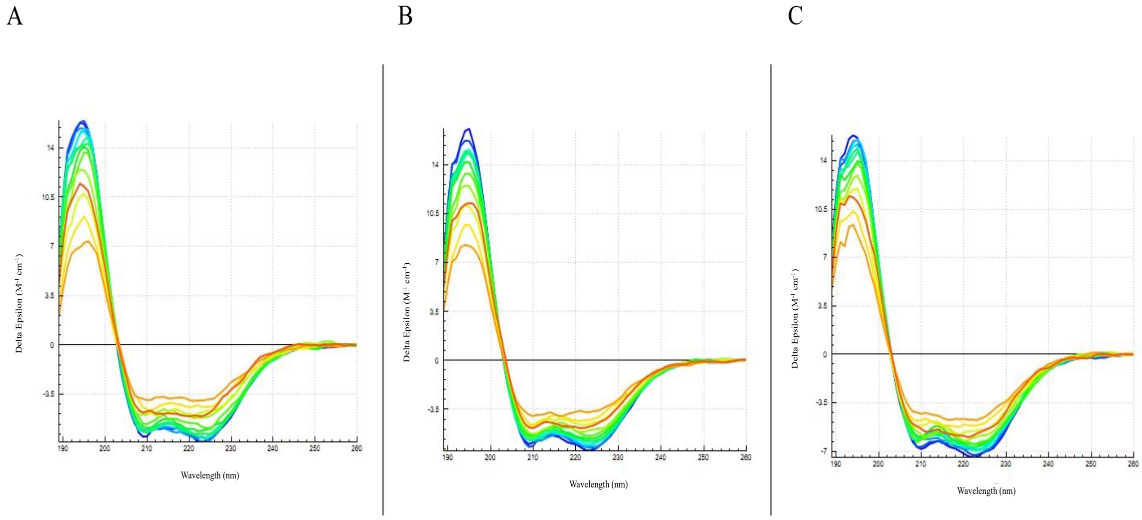
4.2.4 Thermal melt experiments for the study of NavMs WT stability with and without the presence of CBD

Thermal melt experiments were carried out to examine the relative stability of NavMs WT under all conditions (Fig 4.6). Differences in protein stability during the thermal melt could indicate CBD binding and can be recorded via the measurement of the relative peak magnitude at 223 nm, and via SVD analysis.

The peak at 223 nm is due almost entirely to the presence of α-helical structures. In addition, the signal-to-noise ratio at 223 nm tends to be higher than for peaks at 208 nm and 194 nm, since the HT at this point is lower than it is for either the 208 nm peak or the 194 nm peak. Consequently, analysis of 223 nm peak magnitude changes as a function of temperature were included; this analysis gives a strong indication of whether the NavMs WT protein (which is mostly α-helical in nature) is being stabilised or destabilised structurally.

SVD analysis of spectra gives two principal spectral components, with the first component corresponding to the contribution of folded protein to the spectra, and the second to the contribution of unfolded protein to the spectra. Inclusion of SVD analysis of the spectra collected across different temperatures, especially analysis of changes in the contribution of basis component one to the spectra, gives an indicator of whether a protein structure is being stabilised/destabilised (unfolding slower/faster), hence why it was included within the analysis.

All NavMs WT conditions appeared to result no significant difference in the relative change in 223 nm peak magnitude (Fig 4.7). In addition, there was no significant difference in the values for all of the conditions when looking at basis component one of the SVD analysis (Fig 4.8). Overall, these results suggest that when CBD was added to NavMs WT it did not have a significant effect upon the protein's thermal stability. This observation was corroborated by the observed lack of significant difference in average α-helix content at intermediate temperatures between all of the conditions. This could be explained by the hypothesis that CBD slots into the fenestration, thus not significantly altering NavMs structure, or holding any specific residues in a particular position, but this hypothesis would have to be later explored via crystallographic studies of CBD binding.



Figures 4.6: Thermal melt CD spectra for all NavMs WT conditions. *Thermal melt spectra obtained over 260-190 nm in 1 nm steps for NavMs WT without CBD (A), NavMs WT with CBD (B), and NavMs WT with ethanol (C). Thermal melt data were processed using CDToolX (Miles and Wallace, 2018), and used within SVD analysis also using CDToolX (Miles and Wallace, 2018). In all conditions, as temperature was increased, the magnitude of the peaks at 194 nm, 208 nm and 223 nm decreased, and there was a relative decrease in the 194 nm peak when compared to the other peaks. This overall decrease in signal magnitude and change in shape are due to protein unfolding and the formation of aggregates. Colour scheme (blue to red gradient) for the thermal melt spectra correlates with the spectral run protocol temperature order. Produced using CDToolX (Miles and Wallace, 2018).*

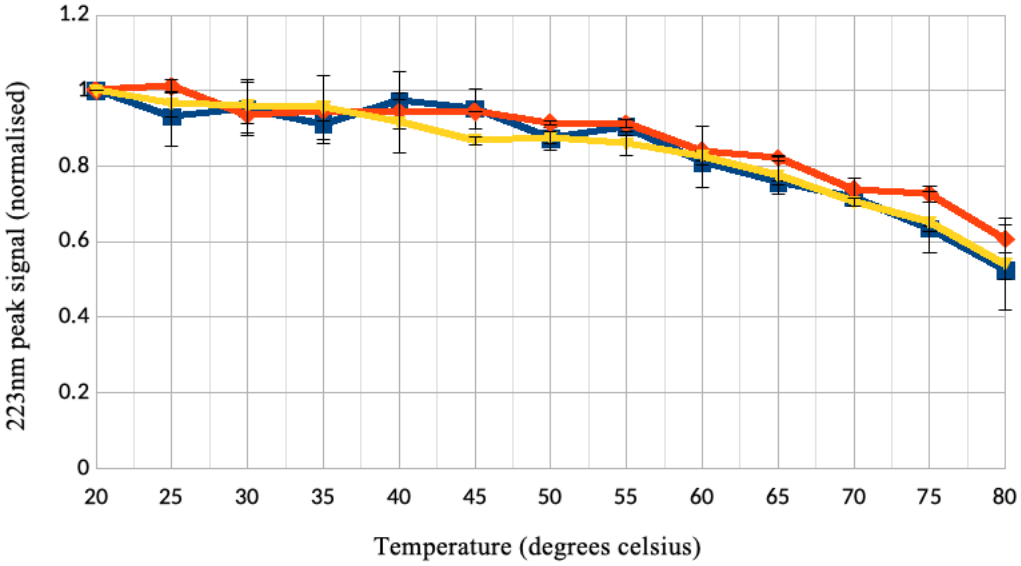


Figure 4.7: Normalised comparisons of thermal denaturation experiments at 223 nm peak. *Graph recording the relative normalised changes in 223 nm peak magnitude as a function of temperature for NavMs WT without CBD (blue), NavMs WT with CBD (red), and NavMs WT with ethanol (yellow). There was no significant difference in peak magnitudes between the conditions, which indicates that neither CBD nor ethanol had a significant effect on the relative helical content at each given temperature. Produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).*

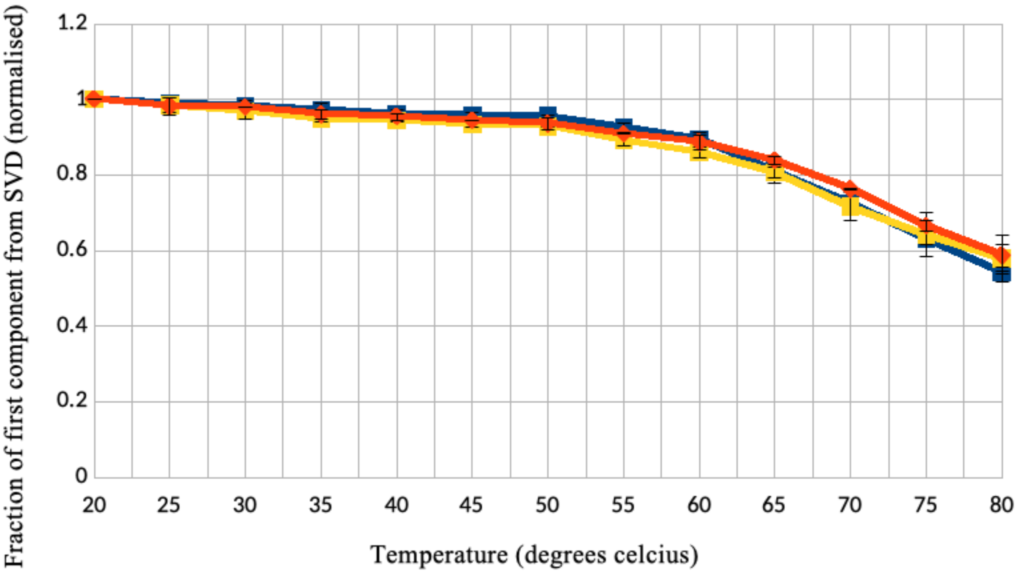


Figure 4.8: SVD analysis as a function of temperature. *Graphing of the first SVD component, which corresponds to folded NavMs WT, for NavMs WT without CBD (blue), NavMs WT with CBD (red), and NavMs WT with ethanol (yellow). NavMs WT with CBD showed no significant difference in the relative change in contribution of the first SVD component across all temperatures. Since the first SVD component corresponds to the contribution of the folded protein to the spectra, this suggests that NavMs WT was unfolding at a similar rate within all the conditions, thus neither CBD nor ethanol appeared to have a significant effect on NavMs WT stability. Produced using Microsoft Excel (Microsoft Corporation, Albuquerque, New Mexico, USA).*

4.2.5 NavMs WT refolding attempts

In none of the conditions, after heating to 80°C and then cooling back to 20°C, was it observed that the final CD spectra resembled the initial CD spectra at 20°C (Fig 4.9). This indicates that the unfolding process is not fully reversible, which is often observed for membrane proteins (Miles and Wallace, 2016). In addition, the spectra shape had also changed, with the peak at 223 nm being slightly shifted towards smaller wavelengths, and the relative peak magnitude at 194 nm being depressed with respect to the other peaks at 208 nm and 223 nm. This change in shape was also combined with an increase in HT value. Overall, both the change in spectra shape and the increase in HT value suggest that this decrease in CD signal is due to both unfolding and the formation of aggregates. These observations of NavMs have been recorded previously in studies of NavMs (Ireland *et al.*, 2018).

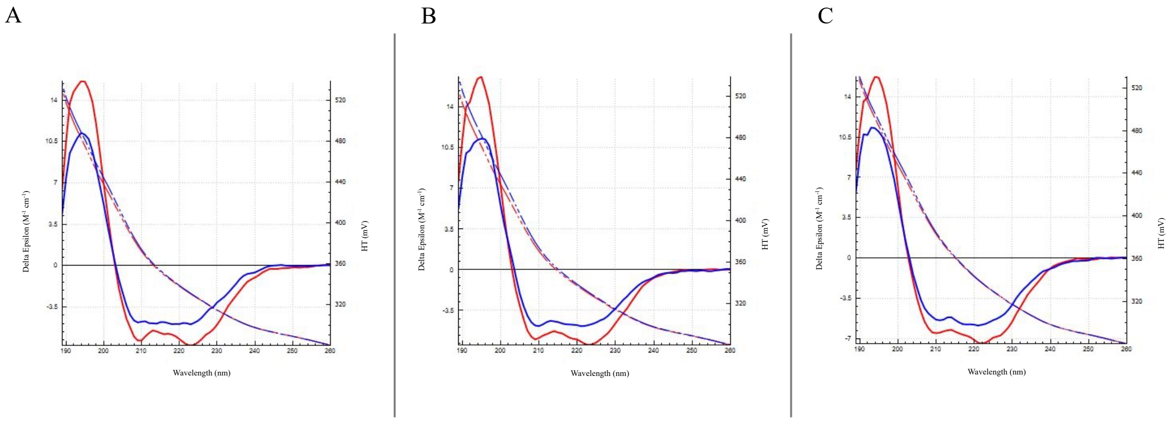


Figure 4.9: Spectra at 20°C prior to heating and at 20°C after heating to 80°C for NavMs WT samples. *Graphs of the relative signals at 20°C prior to heating (red) and at 20°C after heating to* 80°C *(blue), for NavMs WT without CBD (A), NavMs WT with CBD (B), and NavMs WT with ethanol (C). These indicate that for all conditions, unfolding was not completely reversible, since the peak did not return to its original magnitude or shape. Decreases in magnitude and shape can be attributed to unfolding and the formation of aggregates, as previously observed (Ireland et al., 2018). Produced using CDToolX (Miles and Wallace, 2018).*

4.3 Discussion:

4.3.1 Protein GDN transfer

From the trials of transferring NavMs protein into GDN buffer, it quickly became apparent that a certain amount of protein is lost in the process, most likely due to aggregation. This amount of protein lost appears to be relatively fixed, rather than positively correlated with the amount of NavMs initially provided to the GDN buffer transfer process. By increasing the initial total volume of NavMs protein, the set amount of protein lost to aggregation made up a much smaller percentage of the total, leading to measurable peaks at the 10 mL mark (Fig 4.2). GDN transferred fractions were thus able to be successfully collected and concentrated.

4.3.2 Concentration optimisation

At all concentrations of NavMs WT there appeared to be no effect on secondary structure, with all spectra having peaks at ~194 nm ~208 nm and ~223 nm. These peaks have previously been observed for NavMs WT (Ireland *et al*., 2018), indicating that concentration was not a factor that needed to be considered when studying CD spectra shape.

The wavelength point at which the HT value reached 500 mV positively correlated with concentration values. 0.5 mg/mL was the lowest concentration that could be used with the quartz Suprasil (Hellma Ltd.) "bottle" cell, and also had the lowest observed rate of HT increase, with HT reaching 500 mV at ~192 nm. Consequently, 0.5 mg/mL as a concentration allows for the collection of reliable spectral data over the largest range of wavelengths, making it the optimal concentration to use when collecting CD data for NavMs stored in GDN.

4.4.3 CD spectra of NavMs WT without CBD, with CBD, and with ethanol

According to the CD spectra collected for NavMs WT without CBD, with CBD, and with ethanol, the α-helix fractional content for each of these respectively at 20°C is ~67.5 % +/- 0.7 %, ~66.8 % +/- 1.8 % and ~68.1 % +/- 2.8 %. According to present and previous calculations using the DSSP algorithm (Ireland *et al.*, 2018; Kabsch and Sander, 1983) the α-helix content of NavMs WT based on its crystal structure is ~67 %. All of the calculated α-helix content values are in close accordance with this value. In addition, at the majority of the temperatures at which spectra were calculated, there were no significant differences in α-helix content between the conditions (Table 4.1). Also, when looking at the change in α-helix content for all of the conditions (Fig 4.5), the rate of decrease in α-helix content was almost identical. Taking all of these observations together, it was concluded that NavMs is stable in the presence and absence of CBD, thus was suitable for crystallographic studies.

It has been observed that in the case of other sodium channel binding drugs, there is a recorded stabilisation or destabilisation of sodium channels by the given ligand during a thermal melt, which is captured as significant differences in α-helix folded content between the condition with the ligand versus conditions without the ligand at intermediate temperatures (Charalambous *et al.*, 2009; Nurani *et al.*, 2008; Zanatta *et al.*, 2019). SVD analysis of thermal melt data indicated that NavMs WT with CBD did not show any significant difference in the relative amounts of basis component one (which corresponds with folded protein) at all temperatures. Analysis of the relative peak signal at ~223 nm (a peak that can be mostly attributed to α-helical content) also indicated no significant difference in relative ~223 nm peak signal for NavMs WT with

CBD added compared to NavMs WT with nothing added and NavMs WT with ethanol added. Overall, these results suggested that, unlike with some of the other sodium channel binding drugs investigated, CBD does not have any significant effects upon channel stability, and thus cannot be registered via CD measurements.

This apparent lack of effect on stability could be speculated to be the consequence of the CBD molecule not significantly altering NavMs structure or holding multiple specific residues in a particular positional conformation, but this hypothesis needed investigation via follow-on crystallographic research.

4.3.5 Overall

Overall, the data indicated that NavMs was stable with and without CBD present, and that CBD does not have an effect upon NavMs thermal stability when compared to NavMs without CBD. CBD has been evidenced to be an inhibitor of eukaryotic and prokaryotic sodium channels (Fouda *et al.*, 2020; Ghovanloo *et al.*, 2021; Ghovanloo *et al.*, 2018; Sait *et al.*, 2020), thus the lack of apparent effect on stability is an interesting observation, since in previous research known sodium channel targeting drugs have been shown to affect NavMs stability as a function of temperature (Charalambous *et al.*, 2009; Nurani *et al.*, 2008; O'Reilly *et al.*, 2008; Zanatta *et al.*, 2019). To explore whether this observation was due to the nature of CBD:VGSC interactions, crystallographic studies needed to be carried out to investigate CBD binding interactions.

Chapter 5: NavMs WT and NavMs F208L Crystallisation Trials and X-ray Crystallography

5.1 Introduction:

X-ray crystallography has been fundamental within the history of studying macromolecular structures (Shi, 2014). In particular, X-ray crystallography has allowed for a deepened understanding of drug/ligand interactions with macromolecules, enabling structural insight into the mechanistic underpinnings of how drugs/ligands elicit their effects (Aitipamula and Vangala, 2017). X-ray crystallography has already been oused to provide insights into the structure and function of the NavMs sodium channel, and has highlighted how NavMs can be utilised as a model for human sodium channels (Bagnéris *et al.*, 2013, 2014; Ke *et al.*, 2018; McCusker *et al*., 2012; Sula *et al*., 2017, 2021; Sula and Wallace, 2017; Ulmschneider *et al*., 2013). Since X-ray crystallography had already been successfully used to attain structural data from NavMs protein crystals formed in the presence and absence of ligands, X-ray crystallography was a perfect technique to include within this research to aid investigations into the structural underpinnings of CBD interactions with VGSCs.

The aim of including X-ray crystallography using the NavMs protein was to collect diffraction data of the NavMs WT and NavMs F208L protein with and without the presence of CBD. Collecting this data allowed for several observations. First, this data allowed for structural comparisons between NavMs WT and NavMs F208L, allowing for increased understanding of NavMs structure and the significance of changing a phenylalanine to leucine at that position. Second, NavMs F208L is already known from previous unpublished trials within the Wallace group to crystallise more readily than NavMs WT, making it a favourable candidate to include when wanting to increase the chance of forming CBD co-crystals. However, an understanding of the NavMs F208L structure, and whether it was comparable to the already known NavMs WT structure, allowed for evaluation of the use of NavMs F208L within co-crystallisation screens. Third, data collected from NavMs in the presence of CBD granted, for the first time within research, structural insight into the inhibitory effects of CBD, which have been observed via electrophysiological experiments (Fouda *et al*., 2020; Ghovanloo *et al*., 2021; Ghovanloo *et al*., 2018; Patel *et al*., 2016; Sait *et al*., 2020).

Please refer to methods section 2.7 for details on crystallisation trials, data collection and model building.

5.2 Results:

The results from this crystallography work contributed to the paper (Sait *et al.*, 2020).

5.2.1 Crystal screens with NavMs WT and NavMs F208L

Crystals for both NavMs WT and NavMs F208L, with and without CBD, formed across a range of conditions with NavMs at the previously optimised concentration (Sula *et al.* 2017) of 10mg/ml. Initial CBD concentrations of 2mM appeared to produce crystals thus no further optimisation of CBD concentration was carried out, since crystals with this concentration of CBD produced highly diffracting crystals which appeared to have electron density which could be ascribed to CBD (described in more detail in the following sections). Crystals ranged in size and shape (Fig 5.1), with NavMs WT having less recorded crystallisation events than NavMs F208L. The higher number of crystallisation events for NavMs F208L was expected since previous unpublished work in the Wallace lab had indicated that NavMs F208L crystallised more readily than NavMs WT.

NavMs WT and NavMs F208L both had more crystallisation events with CBD added versus without CBD added. This may indicate that CBD had a stabilising effect upon NavMs protein crystal structures.

Crystals for NavMs WT and NavMs F208L, with and without CBD, were fished, flash frozen in liquid nitrogen, and then sent off to synchrotrons for data collections.

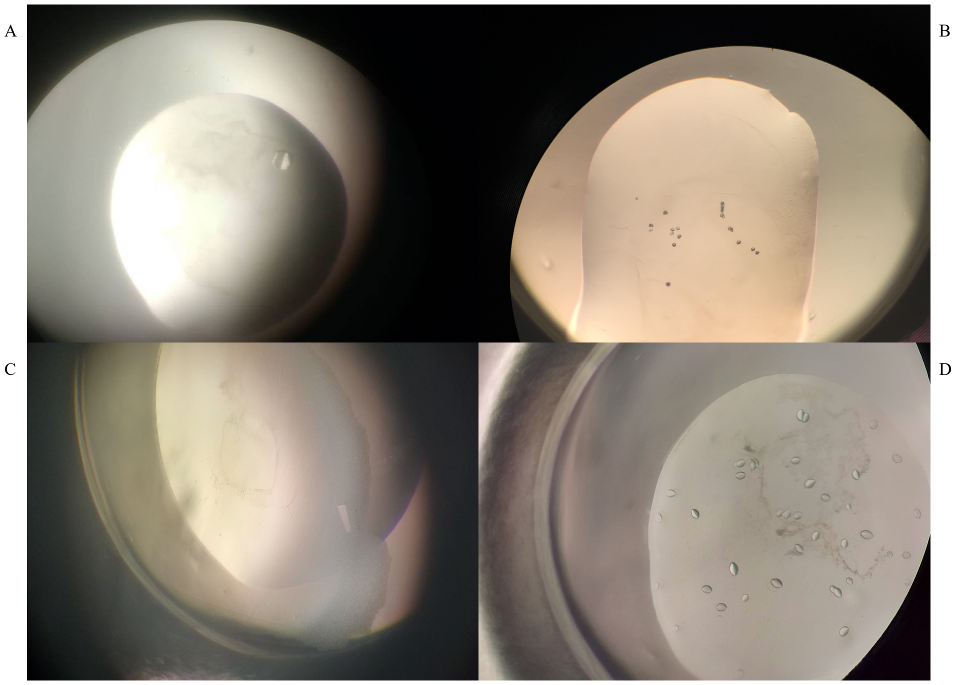


Figure 5.1: Example NavMs WT and NavMs F208L crystals with and without CBD added. *A) NavMs WT with CBD added (2 mM). B) NavMs WT without CBD added. C) NavMs F208L with CBD added (2 mM). D) NavMs F208L without CBD added.*

5.2.2 Data collections and initial data processing

Hundreds of crystals (NavMs WT and NavMs F208L) were screened at the synchrotron, and full data sets were collected from ~40 different crystals. Crystal data was collected on beamline P13 at the Electron Synchrotron (DESY, Germany); on beamline Proxima1 at the Soleil Synchrotron (France), and on beamlines IO3, IO4, and I24 at the Diamond Light Source (UK). For NavMs WT with and without CBD, none of the crystals diffracted to a high enough resolution (~3 Å) to warrant data collection. In addition, many of the NavMs WT and NavMs F208L crystals initially screened were not suitable for data collection due to ice damage, or were quickly revealing to be salt crystals rather than protein crystals.

After collecting data from NavMs F208L crystals with and without CBD, diffraction images for each separate data set went through integration and scaling via the use of XDS (Kabsch *et al*., 2010). Resolution cutoffs for datasets were adjusted in XDS via the analysis of the CC ½ and completeness values (see methods section 2.7.2 for details). Diffraction images for each dataset were then merged using Aimless (Evans and Murshudov, 2013). The deposited NavMs WT structure (PDB ID: 5HVX) was used as an initial phasing model for all diffraction datasets, using PHASER (McCoy *et al*., 2007), and initial rounds of refinement were carried out using REFMAC (Murshudov *et* *al*., 2011) and COOT (Winn *et al*., 2011), where COOT was used to analyse and improve the fit of the model to the calculated electron density, and REFMAC was used to feed in the Coot-adjusted model with the original diffraction data to create new phase estimates, and thus an adjusted and better-fitting model. All of these programmes are incorporated into the CCP4 suite of programmes (Winn *et al*., 2011). Because of the variety in both unit cell dimensions and resolution between crystals formed under the same conditions, as previously observed (Naylor *et al*., 2016; Sula *et al*., 2017), data from separate crystals were not merged.

Once initial processing was completed, COOT was used again but to investigate the models and electron density maps built for both NavMs F208L without CBD and NavMs F208L with CBD datasets. F_{obs}-F_{calc} electron density difference maps were employed to investigate whether datasets for NavMs F208L with CBD had any areas of unexplained electron density in comparison to NavMs F208L without CBD models that could possibly be explained via the presence of CBD (see theory appendix section 8.3.6 for details on what F_{obs}-F_{calc} electron density difference maps are). Within NavMs F208L with CBD datasets it was observed that there was a consistent area of electron density which was not explained by the model. CBD was fitted to this electron density using COOT, and this adapted model was used in further rounds of refinement to observe whether the inclusion of CBD improved the model, according to factors such as R_{free} and R_{work}. Continued refinement was monitored via the R_{free} and R_{work} values, and once the model could not be refined further other parameters were evaluated (see theory appendix section 8.3.7 for details of values studied during model assessment) to gage the quality of the models.

5.2.3 NavMs F208L without CBD and NavMs F208L with CBD structures

NavMs F208L without CBD and NavMs F208L with CBD deposited models: A final full-length refined structure of NavMs F208L without CBD was discerned to a resolution of 2.20 Å (Table 5.1) and was deposited into the PDB database (PDB ID: 6YZ2). A full-length refined structure of NavMs F208L with CBD was discerned to a resolution of 2.30 Å (Table 5.1) and was also deposited into the PDB database (PDB ID: 6YZ0). Both models had R factors 0.3 < (Table 5.1) indicating a good fit between the data and the model. In addition, RMS bond angles and bond lengths fell within acceptable values (less than 0.02 Å for bond lengths and less than 4° for bond angles), indicating that the models were biologically sensible as well as well-fitted.

NavMs F208L without CBD structure: Performing a structural alignment of the NavMs F208L and NavMs WT structure (PDB ID: 5HVX) (Fig 5.2) indicated that there are no discernible differences in structure between NavMs WT and NavMs F208L. This suggested that NavMs F208L was a viable alternative to NavMs WT for these crystallisation studies with CBD, since NavMs F208L crystallised at a higher rate and appeared to have no discernible difference in structure (root mean square deviation (RMSD)= 0.48 Å), thus observations drawn from NavMs F208L data were applicable to the NavMs WT structure as well.

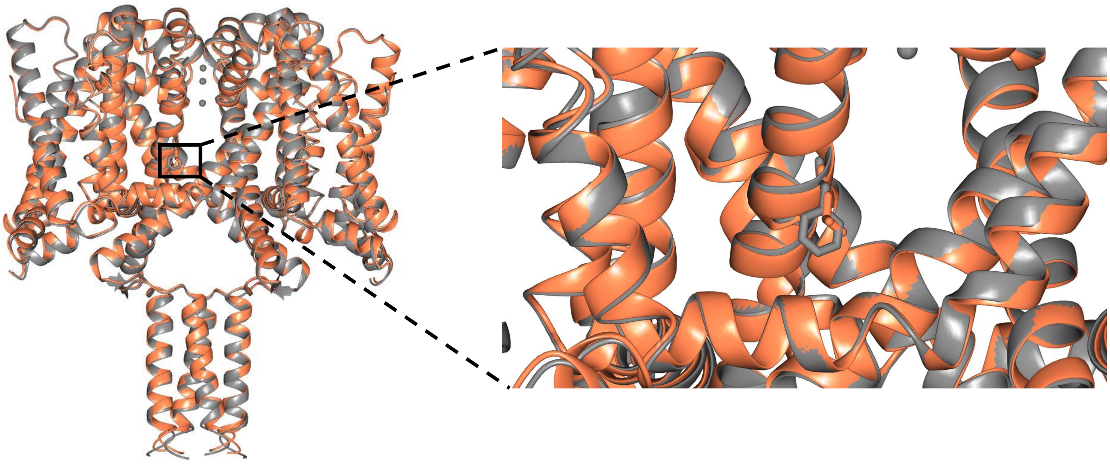


Figure 5.2: Structural alignment of NavMs WT and NavMs F208L (after (Sait et al., 2020)). *Depiction of the structural alignment of the refined structure for NavMs F208L (PDB ID: 6YZ2) (orange) and NavMs WT (PDB ID: 5HVX) (grey). From the structural alignment, there are no discernible differences in protein structure incurred due to the F208L point mutation. The L208 of NavMs F208L can be seen overlayed with the original F208 of NavMs WT in the zoomed in window. The similarity in structure indicates that NavMs F208L can be used within structural studies since it is structurally equivalent to NavMs WT. Figure produced using CCP4mg (McNicholas et al., 2011).*

Data collection	**NavMs F208L (PDB ID 6YZ2)**	**NavMs F208L-CBD (PDB ID 6YZ0)**
Wavelength (Å)	0.97624	0.97620
Space group	I422	I422
Unit-cell parameters		
a, b, c (Å)	108.94 ,108.94,209.37	110.4,110.4,207.8
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution range (Å)	104.68-2.20 (2.27-2.20)	103.94-2.30 (2.38-2.30)
Total number of observations	687781 (56721)	603236 (58839)
Total number unique	32388 (2771)	28708 (2770)
Completeness	100.0 (100.0)	99.3 (100.0)
Multiplicity	21.2 (20.5)	21.0 (21.2)
<I/σ(I)>	17.4 (2.2)	17.3 (1.0)
CC(1/2)	0.999 (0.950)	0.999 (0.877)
$R_{merge\ all}$	0.093 (1.20)	0.085 (3.803)
$R_{pim\ All}$	0.021 (0.278)	0.020 (0.877)
Solvent content (%)	72.80	73.64
Molecule per ASU	1	1
Wilson B factor (Å^2)	46.9	52.6
Refinement		
Resolution Range (Å)	28.8-2.2	48.1-2.3
R_{work}	0.24	0.23
R_{free}	0.26	0.26
Reflection, working	32388	27254
Reflection, free	1619	1376
Average B factor (all atoms)	75.0	95.3
RMS bond angle	0.82	0.90
RMS bond length (Å)	0.008	0.009
Ramachandran Analysis:		
Preferred region (%)	96.6	95.8
Allowed region (%)	3.0	4.2
Outliers (%)	0.4	0.0

Table 5.1: Crystal structure parameters for NavMs F208L without CBD (PDB ID: 6YZ02) and NavMs F208L with CBD (PDB ID: 6YZ0).

NavMs F208L CBD binding site: From comparisons of the NavMs F208L model electron density with and without CBD, the site of CBD binding could be clearly seen within the hydrophobic pocket structure due to a large area of unexplained electron density occurring within the NavMs F208L with CBD model (Fig 5.3), which CBD fitted (Fig 5.4-5.6). This hydrophobic pocket structure is known as the fenestration, which is created between the NavMs subunits and runs perpendicular to the channel (Montini *et al.*, 2018). This fenestration structure occurs four-fold within the NavMs structure due to NavMs being a tetramic structure. Fenestrations lie within the transmembrane region of the protein, and have been proposed to be the entry sites for hydrophobic drugs, which can gain entry to the interior of sodium ion channels (Hille, 1977). The occurrence of this large electron density attributable to CBD is consistent with and supports the proposal that large hydrophobic drugs gain entry to channels via the fenestration structure. This occurrence also supported one of the initial hypotheses of this research, which was that CBD would act to enter the channel via the fenestration structure since it too is a hydrophobic molecule like other hydrophobic drug counterparts.

The CBD molecule appears to be located at the end of the fenestration structure, and is consequently positioned very close to the channel central pore, indicating that it acts to block this structure just below the selectivity filter.

In the NavMs F208L structure without CBD there is also some electron density within the fenestration, but it is of a different size and shape (Fig 5.7), and has been attributed to detergent. Partial occupancy of the fenestration by detergent/lipid has been previously indicated via molecular dynamics simulations (Ulmschneider *et al*., 2013). In addition, trials to fit the electron density seen within the NavMs F208L structure with CBD using a detergent molecule instead clearly indicated that within the NavMs F208L structure where CBD is present, the electron density within the fenestration is occupied by CBD and not by a lipid since the electron density is only accounted for via the presence of the CBD molecule (Fig 5.7).

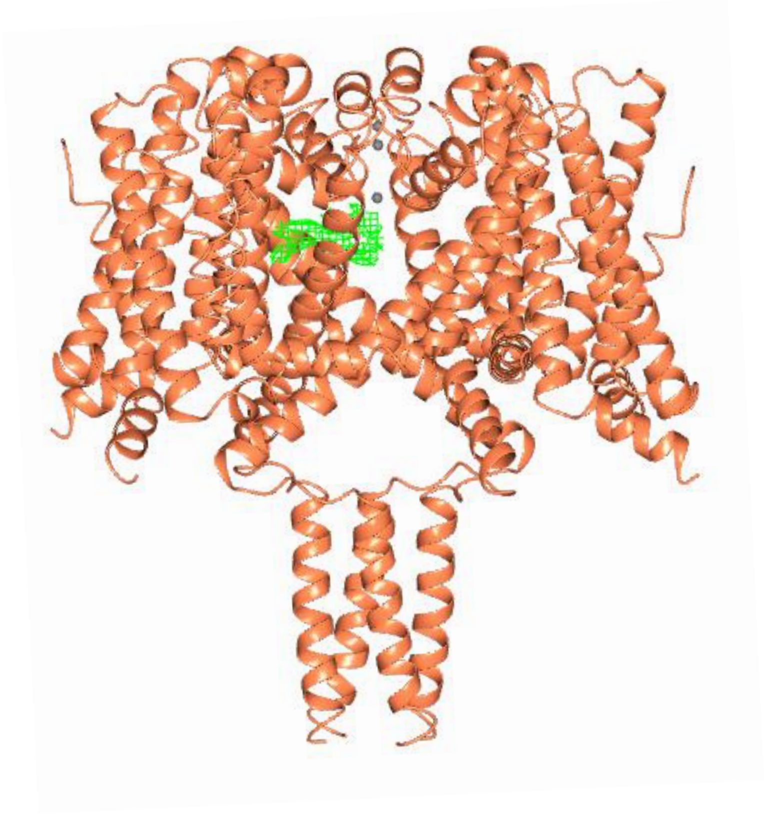


Figure 5.3: NavMs F208L structure with F_{obs}-F_{calc} electron density map indication of CBD within the fenestration. *The F_{obs}-F_{calc} electron density map (green mesh) was calculated at 3 sigma and indicated an area of electron density which the initial phasing model (PDB ID: 5HVX) left unaccounted for, and which was not observed in the NavMs F208L without CBD electron density model. Figure produced using CCP4mg (McNicholas et al., 2011).*

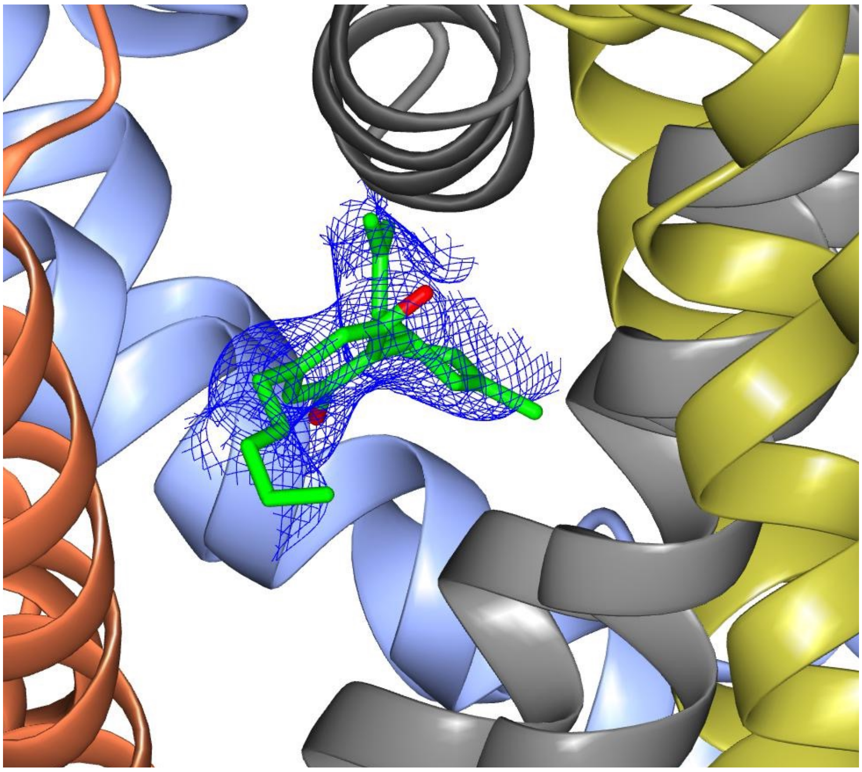


Figure 5.4: NavMs F208L structure with 2F_{obs}-F_{calc} electron density map indication of CBD within the fenestration. *Close up of the binding site for CBD once CBD was successfully fitted to the unaccounted-for electron density (see Fig 5.)within the model: the NavMs F208L structure is in the ribbon depiction and coloured by subunit chain (orange, grey, blue, and yellow). The 2F_{obs}-F_{calc} electron density map (blue mesh) for the CBD molecule (colouring as in A) stick depiction) was calculated at 0.7 sigma. Figure produced using CCP4mg (McNicholas et al., 2011). Cited from (Sait et al., 2020)*

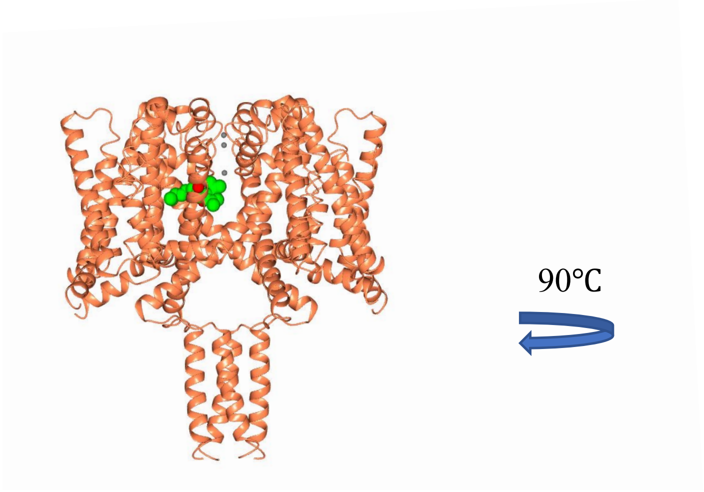


Figure 5.5: NavMs F208L structure with CBD (after (Sait et al., 2020)). *The refined crystal structure of NavMs F208L (orange ribbon) with a single CBD molecule (green space-filing depiction of carbons, with red representing oxygen atoms). The CBD molecule occurs deep within the fenestration, towards the hydrophobic central pore of the NavMs structure. Three sodium ions (grey spheres) are shown in the selectivity filter. Figure produced using CCP4mg (McNicholas et al., 2011).*

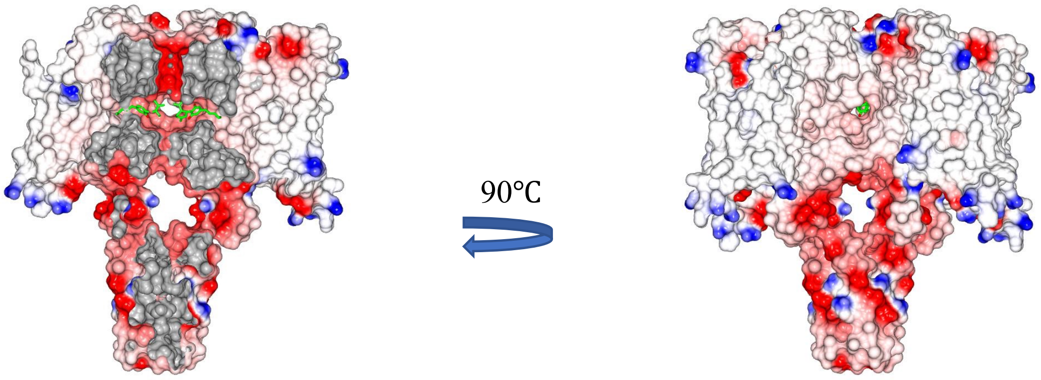


Figure 5.6: Space-filling model of NavMs F208L structure with CBD (after (Sait et al., 2020)). *Same orientations of the NavMs structure as in (Fig 5.), however, here the space filling model of NavMs is depicted, with colouring signifying electrostatics. CBD is depicted in green (like in A)) but is in stick form. On the left NavMs is sliced through to expose the CBD molecules within the fenestration structures. Two CBD molecules were used for clarity, and to indicate how CBD lies deep within the fenestration, blocking ion passage. On the right, there is a surface view of space filling structure of NavMs with a CBD molecule (coloured as in A)). The CBD molecule is only just about visible via the fenestration entry point. Figure produced using CCP4mg (McNicholas et al., 2011).*

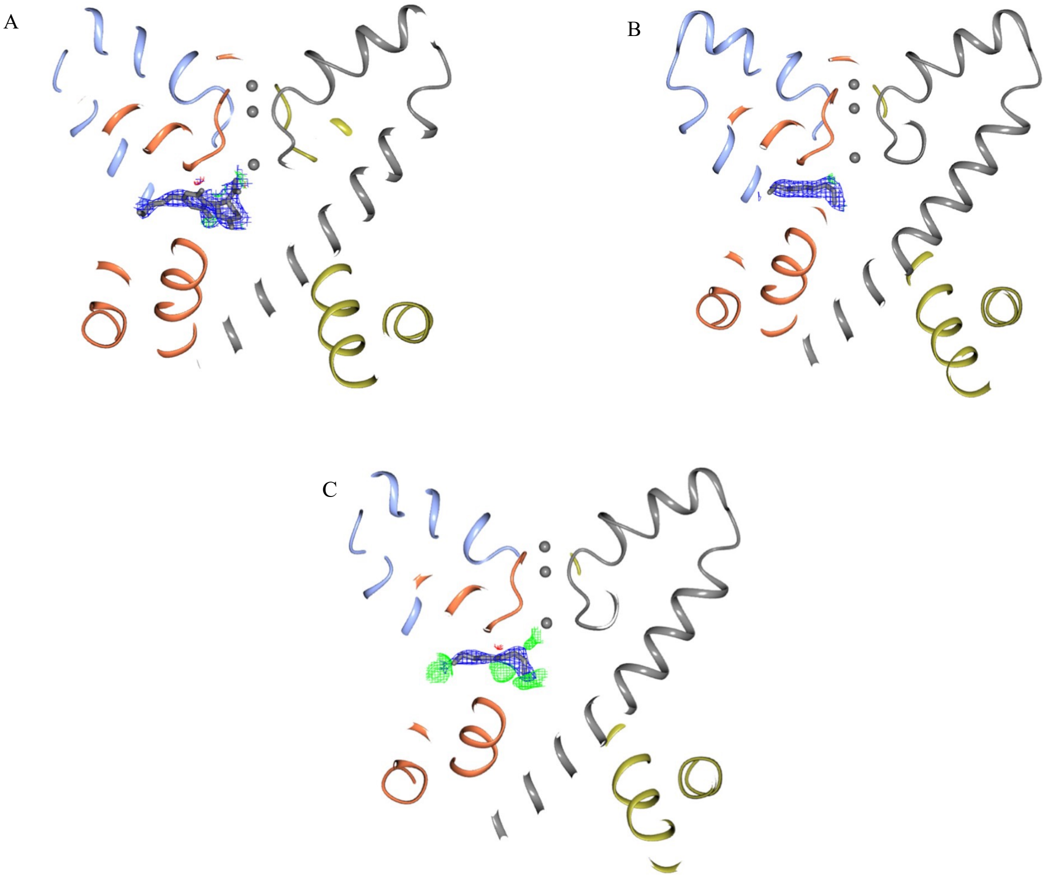


Figure 5.7: Electron density map comparisons for NavMs F208L with and without CBD. *A) Slice through ribbon depiction of NavMs F208L, coloured by chain (orange, grey, blue, and yellow). 2*F_{obs}*-*F_{calc} *map is contoured to 0.75 sigma (blue mesh) and indicates the presence of the CBD molecule. Three sodium ions are indicated by grey spheres. B) Slice through of NavMs F208L without CBD structure, with same depiction and colouring as A). 2*F_{obs}*-*F_{calc} *map contoured to 0.75 sigma in the fenestration and fitted with the aliphatic part of a lipid (*C_9H_{20}*) C). Slice through of the NavMs F208L model created from NavMs F208L with CBD data, using the same depiction and colouring as A) and B), but with the C9H20 structure fitted and refined into the structure instead of the CBD molecule. 2*F_{obs}*-*F_{calc} *map is contoured to 0.75 sigma (blue mesh) and the* F_{obs}*-*F_{calc} *map contoured to 3 sigma, with green mesh indicating unaccounted for positive electron density, and red indicating unaccounted for negative electron density. From this depiction, it becomes clear that the electron density seen within the fenestration of the NavMs F208L structure is due to CBD and not lipid, since fitting lipid leaves a lot of positive electron density unaccounted for, unlike in A). Cited from (Sait et al., 2020) and figure produced by Dr. Altin Sula using CCP4mg (McNicholas et al., 2011).*

Looking at the model and electron density constructed, it appears that it is feasible for there to be four CBD molecules (Fig 5.8), with each occupying a fenestration site. Using HOLE (Smart *et al*., 1996) simulations (Fig 5.9-5.10) indicated that although occupation by up to four CBD molecules is feasible, one is enough to block sodium ion passage. This provides a clear mechanism of channel inhibition by CBD; channel inhibition occurs via physical blocking of the ion passage.

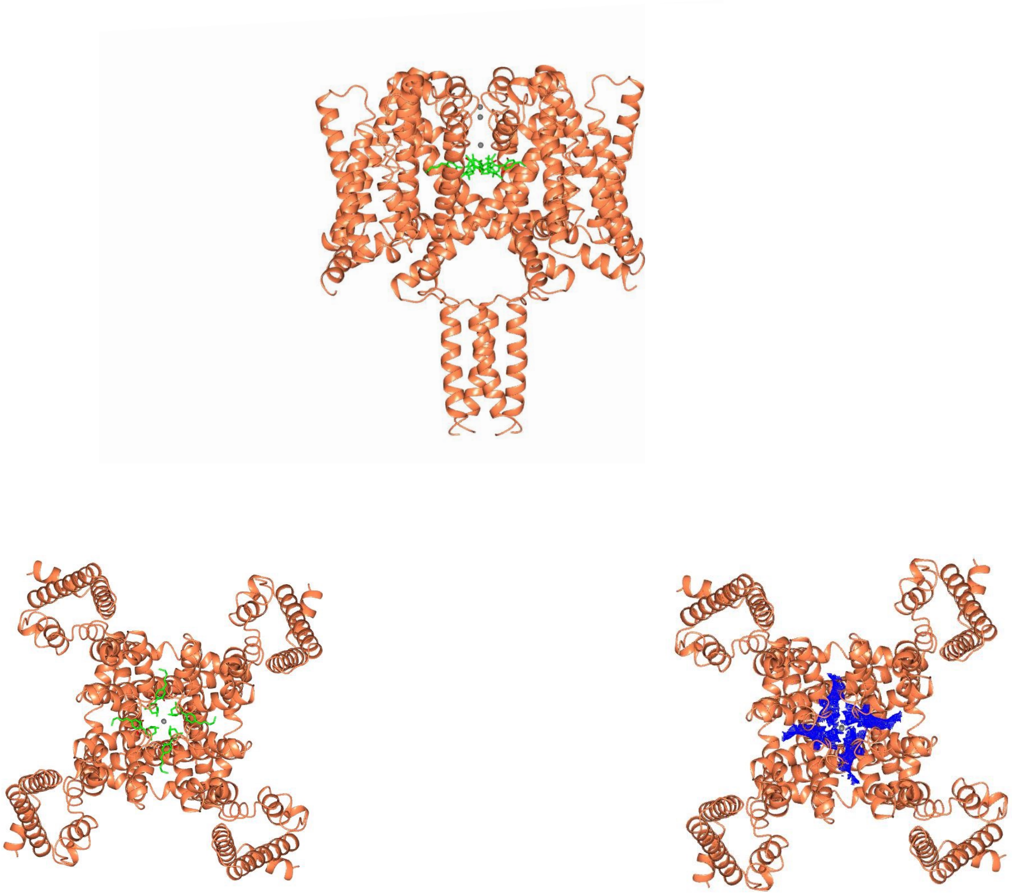


Figure 5.8: NavMs F208L with four CBD molecules fitted and pore diameters with and without CBD. *NavMs F208L depicted in the ribbon motif in the colour orange. NavMs F208L depicted with four CBD molecules (green) from the front (top), then NavMs F208L with four CBD molecules from the top (bottom left), and finally, NavMs F208L from the top view with the* $2F_{obs}$-F_{calc} *map (contoured to 0.75 sigma) for four CBD molecules displayed (bottom right). Cited from (Sait et al., 2020)*

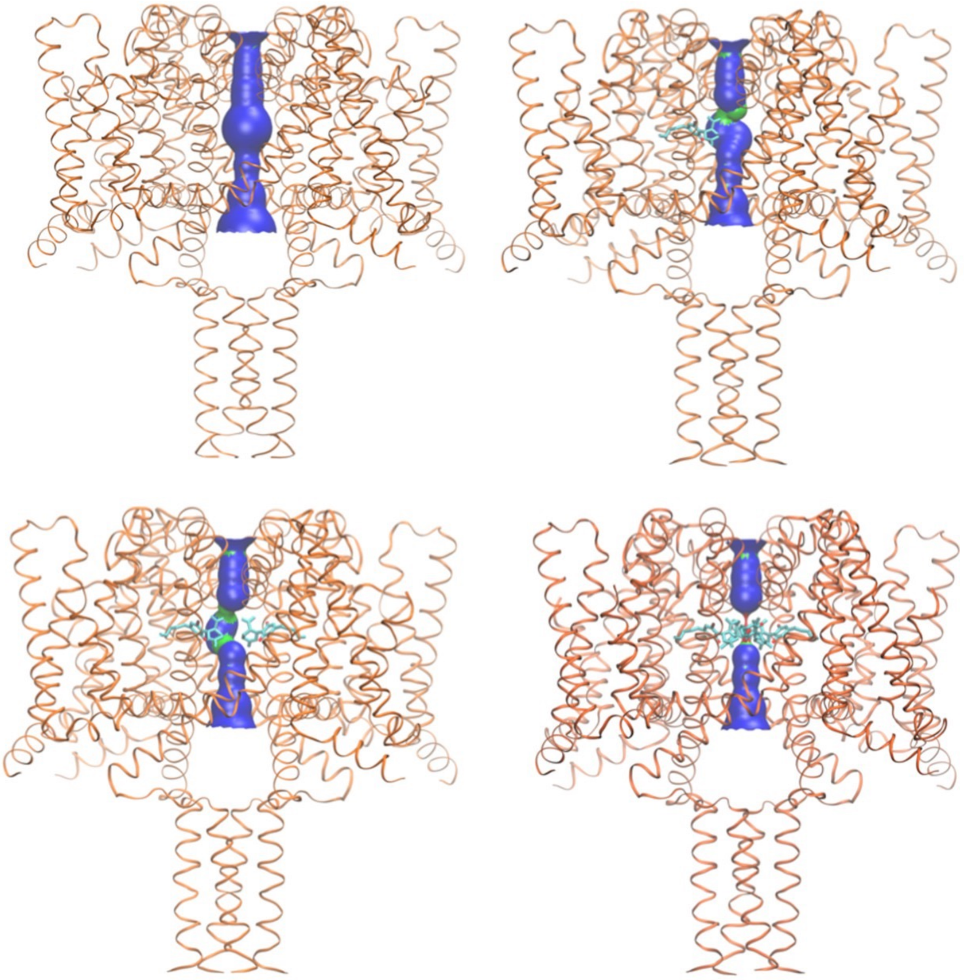


Figure 5.9: NavMs F208L pore diameters with and without CBD. *Depictions of pore dimensions for (from the left) NavMs F208L without CBD, with one CBD, with two CBD molecules and with four CBD molecules. Pore radii greater than 2.3 Å indicated as blue, and pore radii less than 2.3 Å indicated in green. The pore dimensions for NavMs F208L without any CBD molecules indicate that hydrated sodium ions could freely pass through the inner cavity, however, even with the presence of a single CBD molecule, occlusion of the pore occurs at the site where the fenestration meets the inner cavity. Four CBD molecules cause full occlusion of the hydrophobic cavity, consequently blocking sodium ion movement. Cited from (Sait et al., 2020) and images produced by Dr. Altin Sula using VMD software (Humphrey et al., 1996).*

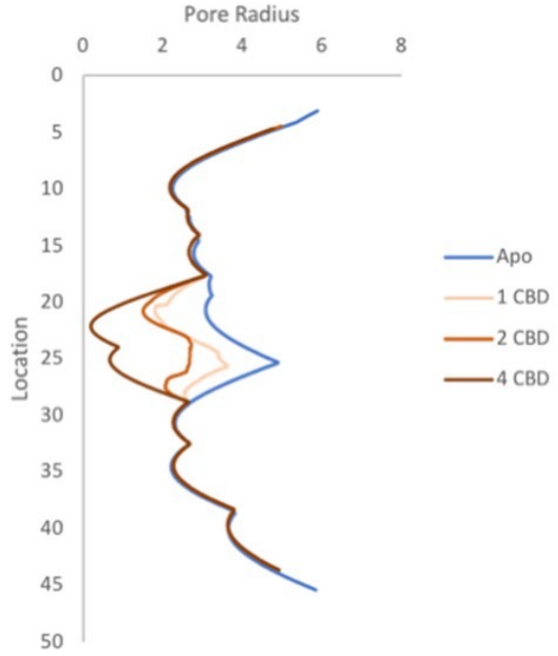


Figure 5.10: Sodium ion accessibility plots for NavMs F208L with zero to four CBD molecules fitted. *Accessibility plots for NavMs F208L without CBD (blue), with one CBD (coral), with two CBD molecules (red) and with four CBD molecules (black). Two CBD molecules and three CBD molecules present plots were present the plot for three CBD molecules is omitted. Where at least a portion of a pore is less than 2 Å, sodium ions will not be able to move across the channel (Naylor et al., 2016), thus these plots indicate that the presence of one CBD molecule is adequate to constrict a section of the pore to a radius of less than 2 Å, meaning that ion passage would not be able to occur. Cited from (Sait et al., 2020) and graph produced by Dr. Altin Sula using VMD software (Humphrey et al., 1996).*

Looking more closely at the binding site for CBD (Fig 5.11), the site is comprised of eleven residues that are contributed by three different subunits. When looking at the equivalent bonding site in human sodium ion channels, it is important to note that the equivalent residues that contribute to the binding site are contributed from three domains of the same polypeptide chain (Fig 5.12) since human sodium ion channels are pseudotetramers instead of true tetramers like their prokaryotic counterparts. The interactions between these residues and CBD are mostly hydrophobic in nature. However, the carbonyl of the main chain of M175 is involved in the formation of a hydrogen bond with the OH group present within the CBD molecule (Fig 5.11). This binding site occurs very close to binding site locations that have already been associated with known sodium channel blocking hydrophobic drugs. These drugs have been evidenced to inhibit sodium ion channel functioning in both the prokaryotic NavMs and prokaryotic sodium channel NavAb (Bagnéris *et al*., 2014; Gamal El-Din *et al*., 2018) (Fig 5.13), suggesting that this site is key to drug interactions in humans. In addition, both the residues T207 and F214 occur within the binding site, confirming our earlier hypothesis that went into forming NavMs mutants (see Chapter 3: Expression and Purification of NavMs Wild Type (WT), NavMs F208L and hypothesised binding site mutants), that CBD would bind in a similar location to other channel-blocking compounds. M204 is also observed within the binding site, which was interesting since alterations in this position could affect both CBD entry and CBD binding.

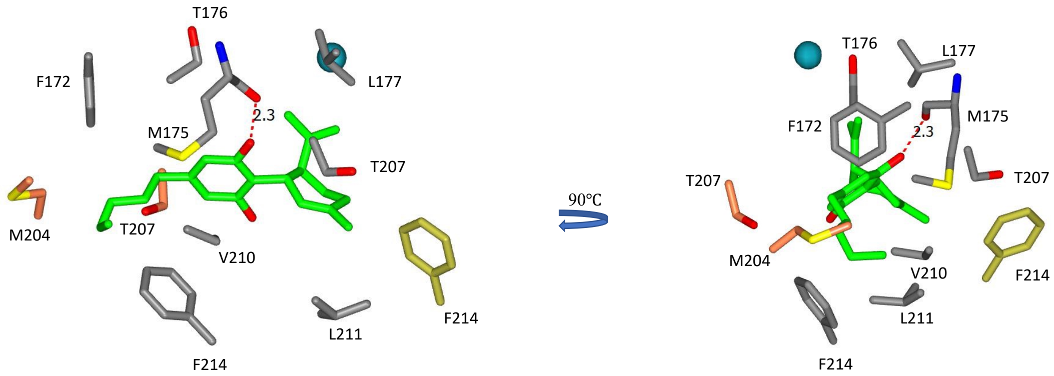


Figure 5.11: CBD binding site residues (after (Sait et al., 2020)). *A view of residues which lie within 4 Å of the CBD (green stick representation, with red representing oxygen atoms) location within the NavMs F208L structure. The binding site residues are shown in two orientations, with the second orientation representing a 90° rotation from the first orientation. Residues are coloured by domain (orange, grey, yellow for the contributing sub domains). The majority of interactions are hydrophobic in nature; however, as depicted via the red dashed line, the carbonyl group of the M175 backbone forms a hydrogen bond with the OH group of the CBD molecule. T207 and F214 both occur twice within the binding site, and have been previously been associated within drug binding (Bagnéris et al., 2014). T207 is later mutated in electrophysiology studies (see Chapter 6: Electrophysiology Experiments to Investigate the Effects of cannabidiol on NavMs Function). Residue atoms are coloured by atom type (carbon depicted as grey, orange, or yellow, depending on subdomain, oxygen as red, nitrogen as blue, and sulphur as yellow). A sodium ion is displayed as a turquoise sphere. Figure produced using CCP4mg (McNicholas et al., 2011).*

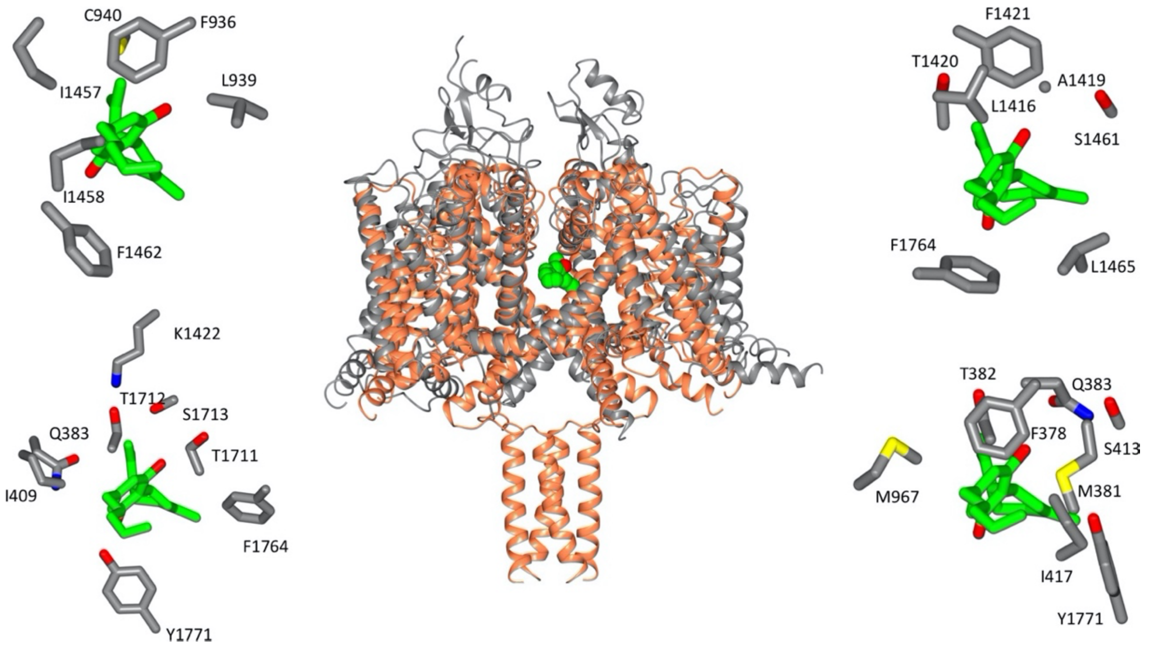


Figure 5.12: Structural alignment of NavMs F208L with CBD with human Nav 1.2 and equivalent binding sites within human Nav 1.2. *In the centre is a depiction of the structural alignment of the NavMs F208L structure with CBD (orange) and the structure of human Nav 1.2, which was discerned via cryo-EM (grey) (PDB ID: 6J8E). The RMSD for this structural alignment is 3.2Å. In the top left, are the equivalent binding residues which exist within 4 Å of the CBD molecule between human Nav 1.2 D1 and D2 are shown. In the top right, are the equivalent binding residues which exist within 4 Å of the CBD molecule between human Nav 1.2 D2 and D3 are shown. In the bottom left, are the equivalent binding residues which exist within 4 Å of the CBD molecule between human Nav 1.2 D3 and D4 are shown. In the bottom right, are the equivalent binding residues which exist within 4 Å of the CBD molecule between human Nav 1.2 D4 and D1 are shown. For the human Nav 1.2 residues depicted, atoms are coloured by atom type (carbon depicted as grey, oxygen as red, nitrogen as blue, and sulphur as yellow). Figure produced using CCP4mg (McNicholas et al., 2011). Cited from (Sait et al., 2020)*

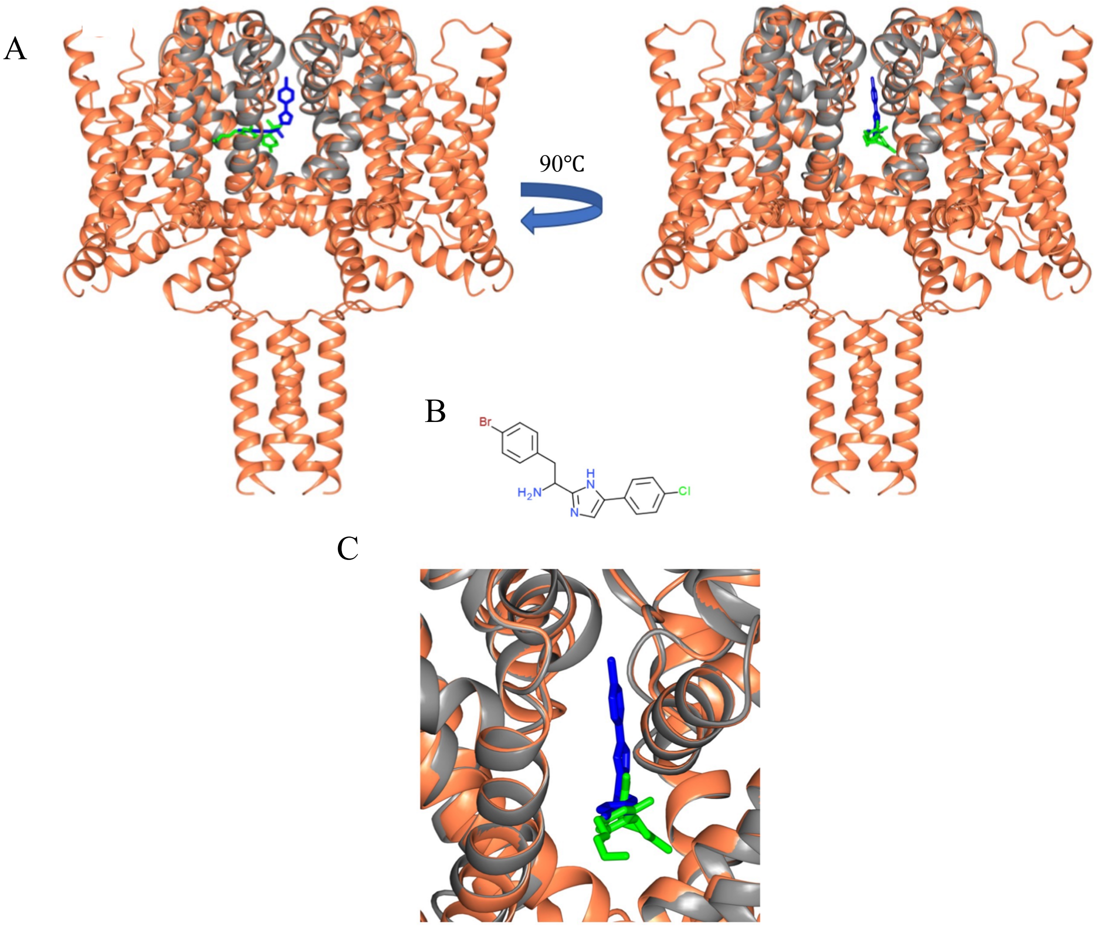


__Figure 5.13: Similarity between CBD binding site and analgesic compound binding site.__ A) Structural alignment of NavMs F208L (orange ribbon depiction) with CBD (green stick depiction) and the NavMs pore structure (grey ribbon depiction) with the drug PI1 (blue stick depiction) (Bagnéris et al., 2014). PI1 is an analgesic drug which inhibits ion movement through NavMs (Bagnéris et al., 2014). B) Close up of the locations for both CBD and PI1 within the structural alignment shown in A). This more clearly depicts the location of both compounds within the fenestration. C) Chemical structure of PI1. A) and B) produced using CCP4mg (McNicholas et al., 2011). C) produced using ChemDoodle (Todsen, 2014). Cited from (Sait et al., 2020)

Equivalent interactions with human Navs: One of the main focuses when investigating CBD and its functional effects was the human sodium ion channel Nav 1.1, due to its primary expression within the brain and association with epilepsy. Inserting CBD into the PDB structure for human Nav 1.1 (PDB ID: 7DTD) (Fig 5.14) and performing sequence alignments with NavMs (Fig 5.15-5.16) suggested that interactions between CBD and human Nav 1.1 would be very similar to human Nav 1.2 and NavMs, in that the CBD molecule would be interacting with a highly hydrophobic environment within the fenestration. As mentioned, in NavMs the T207 position has been associated with the binding of other hydrophobic drug compounds, such as analgesics. The equivalent residue in the human Nav 1.1 structure (F1774) (Fig 5.14) has been associated with the binding of sodium ion channel blocking compounds (Desaphy *et al*., 2010; Hanck *et al*., 2009; Liu *et al*., 2003), and was mutated in previous electrophysiology work, causing a drop in CBD binding affinity by a magnitude of ~2.5 (Ghovanloo *et al*., 2018).

The equivalent binding site residues, according to sequence alignments in human Nav 1.1-1.9 (coloured red in Fig 5.15-5.16), include both residues that are homologous in NavMs and human Navs, along with non-cognate residues. In most cases, the non-cognate residues are also variable across human Nav domains and human Navs (Fig 5.16). Looking at both sequence and structural alignments (Fig 5.12 and Fig 5.15-5.16), it is clear that despite there being non-cognate residues, the overall hydrophobic environment of the fenestration in which CBD would interact is maintained across the human Navs, suggesting that CBD could act to block multiple human Navs. This is supported by previous electrophysiology work, which found that CBD acts to inhibit multiple human Navs (Fouda *et al*., 2020; Ghovanloo *et al*., 2021; Ghovanloo *et al*., 2018).

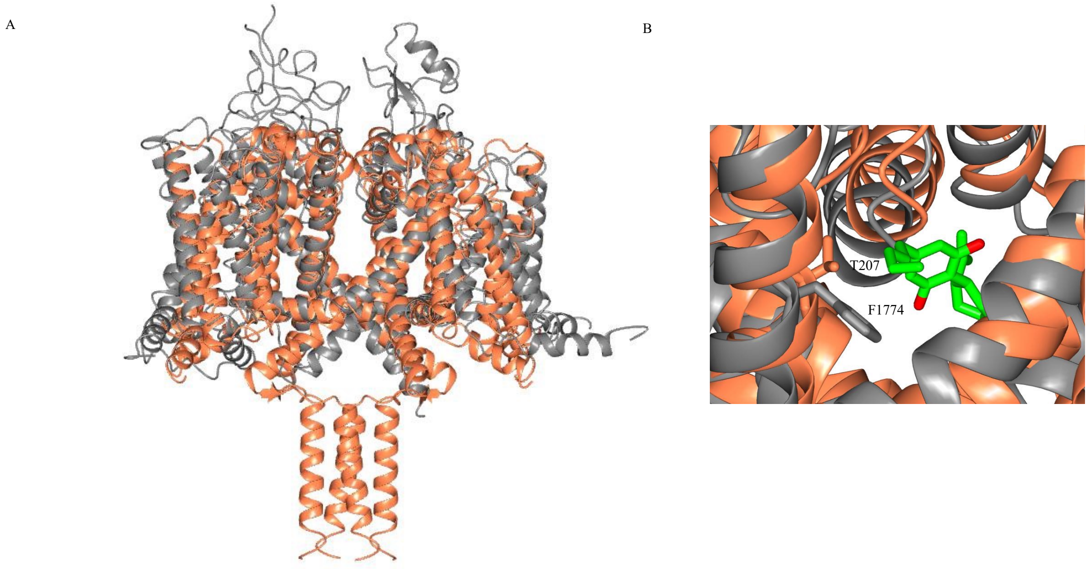
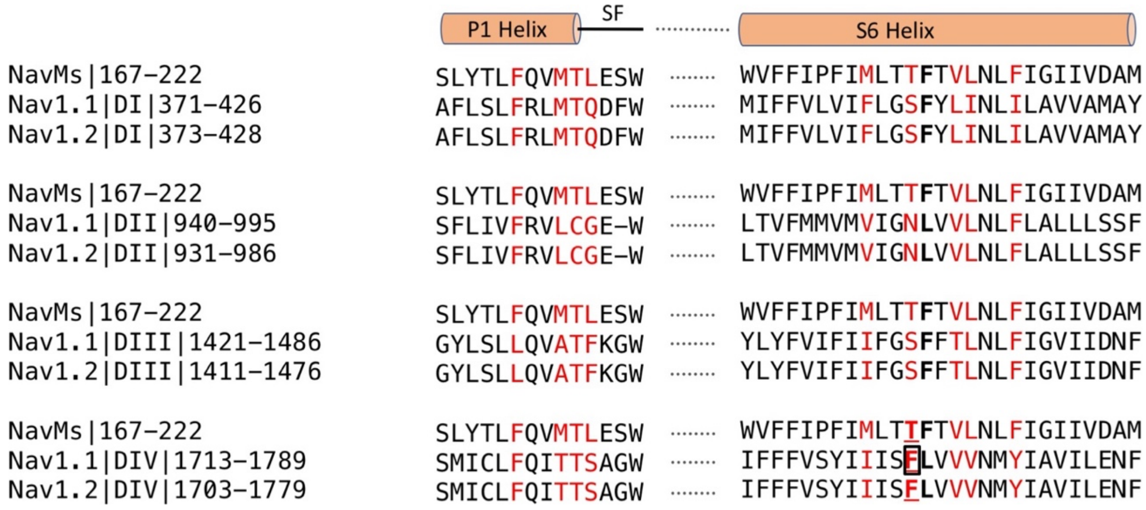


Figure 5.14: Structural alignment of NavMs F208L with CBD and human Nav 1.1. *A) Display of the overall structural alignment of the human Nav 1.1 structure (PDB ID: 7DTD) with the NavMs F208L with CBD structure (PDB ID: 6YZ0). The RMSD =3.30 Å for the alignment of the transmembrane regions. B) Close-up of overlay at the site of CBD binding (green representing carbon atoms and red representing oxygen atoms). Overall, human Nav 1.1, much like NavMs and human Nav 1.2, has a hydrophobic fenestration structure in which CBD could bind via hydrophobic interactions. In addition, the NavMs T207 residue corresponds to the human Nav 1.1 residue F1774, which has been previously identified as being involved in the binding of sodium ion channel blocking compounds (Desaphy et al., 2010; Hanck et al., 2009; Liu et al., 2003), and was also mutated within previous electrophysiology experiments exploring CBD binding and the inhibition of human sodium ion channels (Ghovanloo et al., 2018). Figure produced using CCP4mg (McNicholas et al., 2011).*



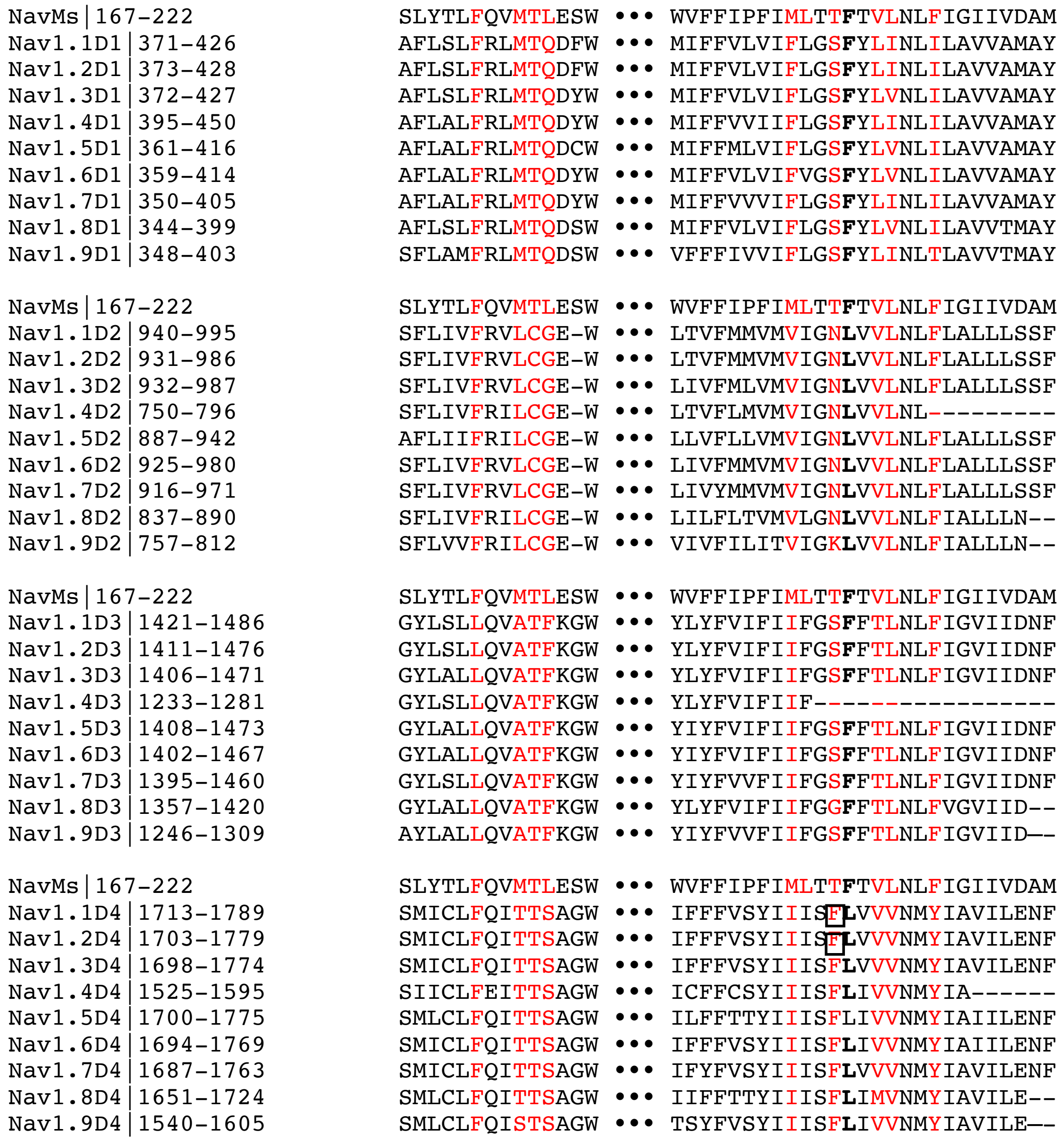
```
NavMs|167-222           SLYTLFQVMTLESW ······· WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1|DI|371-426       AFLSLFRLMTQDFW ······· MIFFVLVIFLGSFYLINLILAVVAMAY
Nav1.2|DI|373-428       AFLSLFRLMTQDFW ······· MIFFVLVIFLGSFYLINLILAVVAMAY

NavMs|167-222           SLYTLFQVMTLESW ······· WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1|DII|940-995      SFLIVFRVLCGE-W ······· LTVFMMVMVIGNLVVLNLFLALLLSSF
Nav1.2|DII|931-986      SFLIVFRVLCGE-W ······· LTVFMMVMVIGNLVVLNLFLALLLSSF

NavMs|167-222           SLYTLFQVMTLESW ······· WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1|DIII|1421-1486   GYLSLLQVATFKGW ······· YLYFVIFIIFGSFFTLNLFIGVIIDNF
Nav1.2|DIII|1411-1476   GYLSLLQVATFKGW ······· YLYFVIFIIFGSFFTLNLFIGVIIDNF

NavMs|167-222           SLYTLFQVMTLESW ······· WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1|DIV|1713-1789    SMICLFQITTSAGW ······· IFFFVSYIIISFLVVVNMYIAVILENF
Nav1.2|DIV|1703-1779    SMICLFQITTSAGW ······· IFFFVSYIIISFLVVVNMYIAVILENF
```

Figure 5.15: Sequence alignments of the CBD binding site residues of NavMs, with the same respective regions in human Nav1.1 and human Nav1.2. *Alignment carried out using Clustal Omega (Sievers and Higgins, 2018), and then manually edited. Red letters indicate the binding site residues within NavMs F208L mutant, and the equivalent residues in both human Nav 1.1 and human Nav 1.2. The 'F' in bold black lettering represents the location at which NavMs F208L is mutated, and also indicates the equivalent residues in human Nav 1.1 and human Nav 1.2, which are either F or L, depending on domain. The residues which have been identified in NavMs as binding residues, and their equivalents in human Nav 1.1 and human Nav 1.2, occur within the P1 pore helix, the selectivity filter loop, and the s6 helix. F1774 in human Nav 1.1 is boxed and is equivalent to T207 in NavMs. F1774 has been identified as being involved in the binding of anaesthetics, as well as appearing to reduce binding affinity for CBD when mutated, according to previous electrophysiology work (Ghovanloo et al., 2018). Cited from (Sait et al., 2020) and figure produced by Dr. Altin Sula using Microsoft PowerPoint (Microsoft Corporation, Albuquerque, New Mexico, USA).*



```
NavMs|167-222          SLYTLFQVMTLESW ••• WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1D1|371-426       AFLSLFRLMTQDFW ••• MIFFVLVIFLGSFYLINLILAVVAMAY
Nav1.2D1|373-428       AFLSLFRLMTQDFW ••• MIFFVLVIFLGSFYLINLILAVVAMAY
Nav1.3D1|372-427       AFLSLFRLMTQDYW ••• MIFFVLVIFLGSFYLVNLILAVVAMAY
Nav1.4D1|395-450       AFLALFRLMTQDYW ••• MIFFVVIIFLGSFYLINLILAVVAMAY
Nav1.5D1|361-416       AFLALFRLMTQDCW ••• MIFFMLVIFLGSFYLVNLILAVVAMAY
Nav1.6D1|359-414       AFLALFRLMTQDYW ••• MIFFVLVIFVGSFYLVNLILAVVAMAY
Nav1.7D1|350-405       AFLALFRLMTQDYW ••• MIFFVVVIFLGSFYLINLILAVVAMAY
Nav1.8D1|344-399       AFLSLFRLMTQDSW ••• MIFFVLVIFLGSFYLVNLILAVVTMAY
Nav1.9D1|348-403       SFLAMFRLMTQDSW ••• VFFFIVVIFLGSFYLINLTLAVVTMAY

NavMs|167-222          SLYTLFQVMTLESW ••• WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1D2|940-995       SFLIVFRVLCGE-W ••• LTVFMMVMVIGNLVVLNLFLALLLSSF
Nav1.2D2|931-986       SFLIVFRVLCGE-W ••• LTVFMMVMVIGNLVVLNLFLALLLSSF
Nav1.3D2|932-987       SFLIVFRVLCGE-W ••• LIVFMLVMVIGNLVVLNLFLALLLSSF
Nav1.4D2|750-796       SFLIVFRILCGE-W ••• LTVFLMVMVIGNLVVLNL---------
Nav1.5D2|887-942       AFLIIFRILCGE-W ••• LLVFLLVMVIGNLVVLNLFLALLLSSF
Nav1.6D2|925-980       SFLIVFRVLCGE-W ••• LIVFMMVMVIGNLVVLNLFLALLLSSF
Nav1.7D2|916-971       SFLIVFRVLCGE-W ••• LIVYMMVMVIGNLVVLNLFLALLLSSF
Nav1.8D2|837-890       SFLIVFRILCGE-W ••• LILFLTVMVLGNLVVLNLFIALLLN--
Nav1.9D2|757-812       SFLVVFRILCGE-W ••• VIVFILITVIGKLVVLNLFIALLLN--

NavMs|167-222          SLYTLFQVMTLESW ••• WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1D3|1421-1486     GYLSLLQVATFKGW ••• YLYFVIFIIFGSFFTLNLFIGVIIDNF
Nav1.2D3|1411-1476     GYLSLLQVATFKGW ••• YLYFVIFIIFGSFFTLNLFIGVIIDNF
Nav1.3D3|1406-1471     GYLALLQVATFKGW ••• YLYFVIFIIFGSFFTLNLFIGVIIDNF
Nav1.4D3|1233-1281     GYLSLLQVATFKGW ••• YLYFVIFIIF-----------------
Nav1.5D3|1408-1473     GYLALLQVATFKGW ••• YIYFVIFIIFGSFFTLNLFIGVIIDNF
Nav1.6D3|1402-1467     GYLALLQVATFKGW ••• YIYFVIFIIFGSFFTLNLFIGVIIDNF
Nav1.7D3|1395-1460     GYLSLLQVATFKGW ••• YIYFVVFIIFGSFFTLNLFIGVIIDNF
Nav1.8D3|1357-1420     GYLALLQVATFKGW ••• YLYFVIFIIFGGFFTLNLFVGVIID--
Nav1.9D3|1246-1309     AYLALLQVATFKGW ••• YIYFVVFIIFGSFFTLNLFIGVIID--

NavMs|167-222          SLYTLFQVMTLESW ••• WVFFIPFIMLTTFTVLNLFIGIIVDAM
Nav1.1D4|1713-1789     SMICLFQITTSAGW ••• IFFFVSYIIISFLVVVNMYIAVILENF
Nav1.2D4|1703-1779     SMICLFQITTSAGW ••• IFFFVSYIIISFLVVVNMYIAVILENF
Nav1.3D4|1698-1774     SMICLFQITTSAGW ••• IFFFVSYIIISFLVVVNMYIAVILENF
Nav1.4D4|1525-1595     SIICLFEITTSAGW ••• ICFFCSYIIISFLIVVNMYIA------
Nav1.5D4|1700-1775     SMLCLFQITTSAGW ••• ILFFTTYIIISFLIVVNMYIAIILENF
Nav1.6D4|1694-1769     SMICLFQITTSAGW ••• IFFFVSYIIISFLIVVNMYIAIILENF
Nav1.7D4|1687-1763     SMICLFQITTSAGW ••• IFYFVSYIIISFLVVVNMYIAVILENF
Nav1.8D4|1651-1724     SMLCLFQITTSAGW ••• IIFFTTYIIISFLIMVNMYIAVILE--
Nav1.9D4|1540-1605     SMLCLFQISTSAGW ••• TSYFVSYIIISFLIVVNMYIAVILE--
```

Figure 5.16: Sequence alignments of CBD-binding site regions of NavMs with corresponding regions in human Nav 1.1-1.9. *Alignments carried out as in (Fig 14), with Clustal Omega (Sievers and Higgins, 2018) and then manually annotated. Red lettering represents binding residues in NavMs, and equivalent residues in human Navs. The 'F' in bold black lettering represents the location at which NavMs F208L is mutated, and also indicates the equivalent residues in human Navs, which are either F or L, depending on domain. One can identify that, in comparison to NavMs binding residues, there are both cognate and non-cognate residues. Non-cognate residues appear to change across D1-4 as well as between human Navs. Taking this into account along with structural alignments (Fig 11 and 13) indicates that the overall hydrophobic fenestration environment is maintained across domains and across human Navs, suggesting that CBD would interact in a similar manner with other VGSCs, as the majority of CBD's interactions are hydrophobic in nature. F1774 in human Nav 1.1 is boxed, is equivalent to the T207 residue in NavMs, and appears to be homologous*

across all human Navs; the equivalent residue in human Nav 1.2 (F1764) (boxed) can be seen within Figure 5.12 as one of the equivalent binding residues within human Nav 1.2.

CBD binding to the TrpV2 channel in comparison to NavMs binding: The Transient receptor potential cation channel subfamily V member 2 (TRPV2) is associated with the non-specific movement of sodium and potassium cations across cell membranes, and in previous research, it has been suggested that CBD acts as an activator of this channel (Morelli *et al*., 2013; Qin *et al*., 2008). CBD appears to activate rat TRPV2 with an EC50 of 3.7 µM (Qin *et al*., 2008). Recent work, which sought to elucidate the structure of rat TRPV2 with CBD bound, resulted in a cryo-EM structure at a resolution of 3.2 Å (Pumroy *et al*., 2019). This structure indicates that CBD interacts with the hydrophobic cavity (fenestration) of the TRPV2 channel. Unfortunately, due to the lower resolution of the TRPV2 structure, it is not possible to perform a detailed analysis of what residues are involved in CBD binding. However, from the structure, it is possible to deduce that binding likely involves several hydrophobic side chains and requires partial refolding of the adjacent region of the protein (Pumroy *et al.*, 2019). Performing a structural alignment of the TRPV2 structure with CBD bound with the NavMs F208L structure with CBD bound (Fig 5.17) indicates that CBD interacts with TRPV2 within a similar region as in NavMs. However, the CBD molecule appears to insert itself deeper into the NavMs structure than the TRPV2 structure. The difference in CBD positionings could account for CBD appearing to be a blocker of sodium ion channels, but an activator of TRPV2.

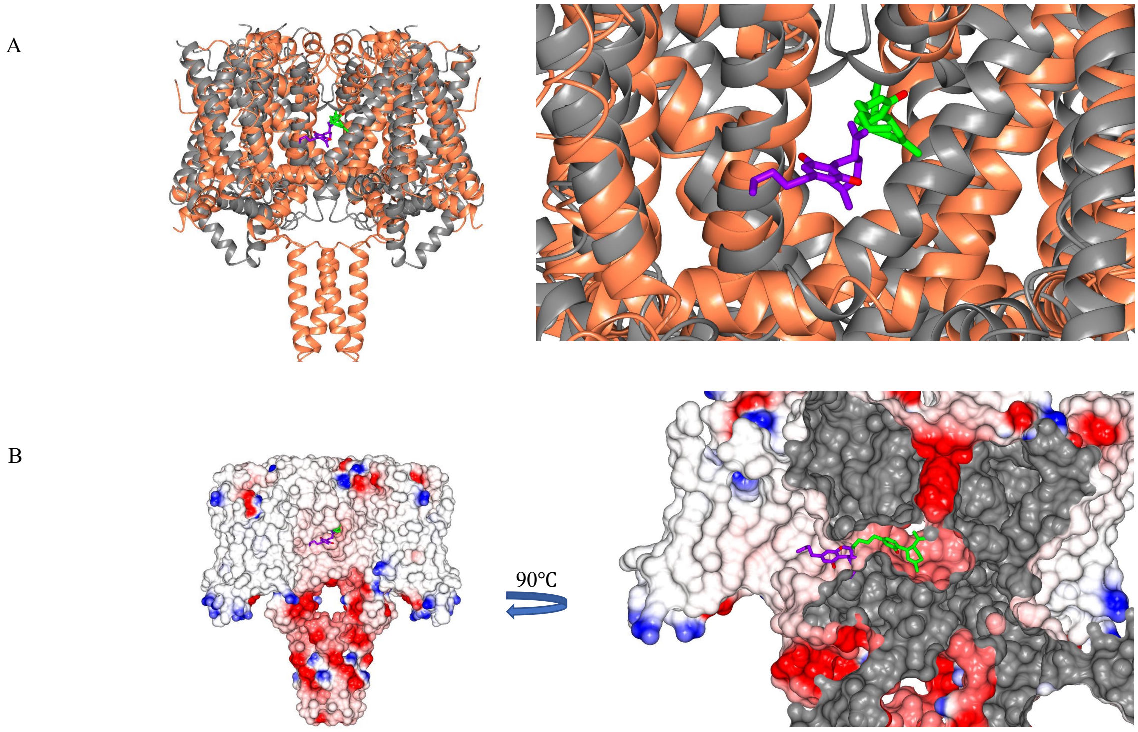


Figure 5.17: Structural alignment of NavMs F208L with rat TRPV2 in a CBD-bound state (after (Sait et al., 2020)). *Structural alignments were carried out in CCP4mg (McNicholas et al., 2011b) using NavMs F208L with CBD structure (PDB ID: 6YZ0) and the rat TRPV2 structure with CBD bound (PDB ID: 6U88), with the RMSD =4.2 Å. A) NavMs F208L in orange ribbon depiction, and TRPV2 in grey ribbon depiction. NavMs' CBD molecule (green for carbon atoms and red for oxygen atoms) appears deeper within the fenestration than for TRPV2. The right panel gives a closer, more detailed view of the CBD sites, indicating both similarities and differences in CBD location and orientation. B) Space filling depiction of NavMs F208L, with both the NavMs associated CBD molecule (same colouring as in A)), and the CBD molecule associated with the rat TRPV2 structure (same colouring as in A)). The right panel gives a close up and rotated look of the CBD molecules, highlighting how the CBD molecule in the NavMs F208L structure is deeper within the fenestration, with the CBD from the TRPV2 structure appearing more at the entrance of the fenestration, which could explain why CBD appears to act as an inhibitor of sodium ion channels but an activator of TRPV2 channels. Produced using CCP4mg (McNicholas et al., 2011).*

CBD versus THC binding: As already covered (see Chapter 1: Literature Review), the two main phytocannabinoids produced by the cannabis plant are THC and CBD, with THC being psychoactive, and CBD being non-psychoactive. Previous electrophysiology work has identified both CBD and THC as inhibitors of sodium ion channels, however they appear to have differences in calculated inhibition slopes (Hill slopes) (THC has a less steep Hill slope than that of CBD), suggesting functional effect differences (Ghovanloo *et al*., 2018). Both THC and CBD have very similar chemical structures, except that CBD has an extra free hydroxyl group on one of the rings, whilst in THC, the equivalent oxygen forms a closed pyran ring. Consequently, the NavMs F208L structure with CBD added was used to examine if there were any clues as to why THC and CBD appear to have these differences in inhibition slopes. As can be seen from (Fig 5.18), when placing the THC molecule in the same orientation and location within the CBD binding site of NavMs, THC can be physically accommodated, however, it appears to be missing the single hydrogen bond which CBD is able to form with the main chain carbonyl group of M175. This is due to the lack of that extra free hydroxyl group, as this hydroxyl group in CBD is what is responsible for forming the hydrogen bond. This lack of hydrogen bonding could explain the differences in inhibition slopes previously observed, since THC has less of an ability to form hydrogen bonds than its counterpart CBD due to this structural lack of an extra hydroxyl group.

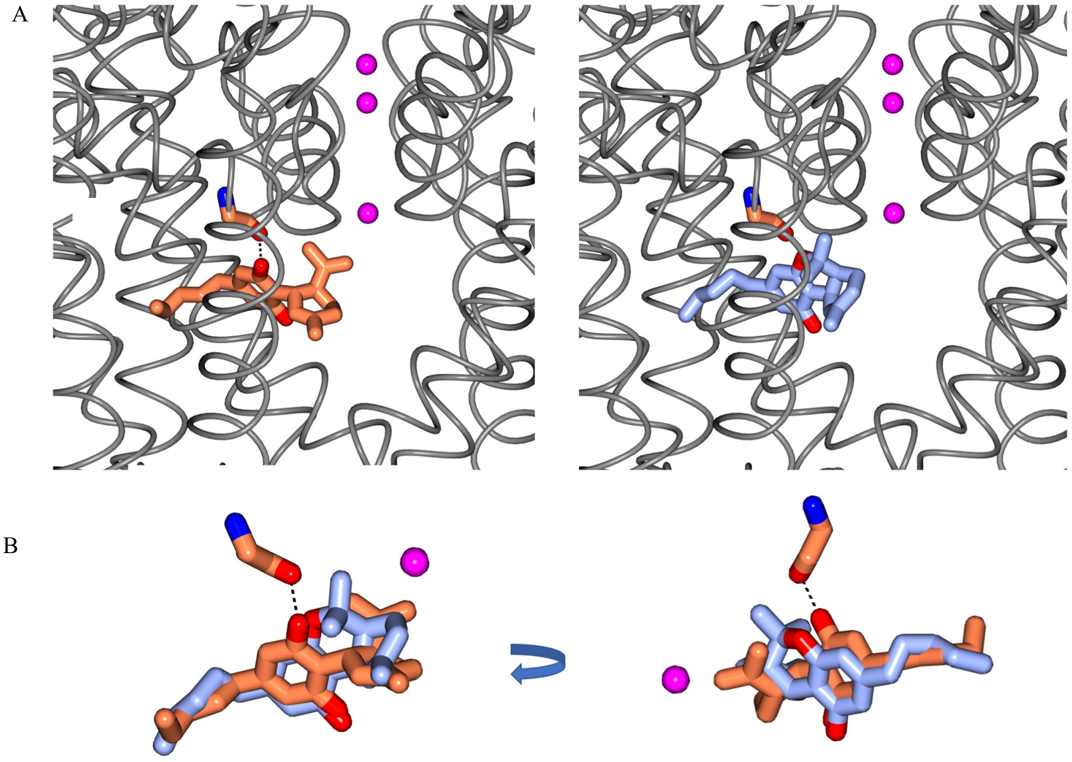


Figure 5.18: Structural comparisons of THC and CBD within NavMs F208L binding site. *A) The locations of CBD (left panel with orange stick representation for carbon atoms and red for oxygen atoms) and THC (right panel with light blue stick depiction for carbon atoms and red for oxygen atoms) within NavMs F208L (grey ribbon depiction), with sodium ions being represented by pink spheres. B) Depiction of the overlay of CBD and THC (colouring same as A)), indicating the similarities in orientation, but the lack of a hydrogen bond between M175 and THC due to the lack of the extra free hydroxyl group. Produced using CCP4mg (McNicholas et al., 2011). Cited from (Sait et al., 2020)*

Crystallisation trials with NavMs binding site mutants: After hypotheses were confirmed that CBD did indeed appear to bind within the fenestration, and involve the residues which had been previously selected for mutation, initial crystallisation trials were set up with these mutants (same protocols used as described in methods section 2.7.1), however, these trials either yielded no crystals or poorly diffracting crystals. Due to time and lab constraints thanks to the covid pandemic, it was not possible to perform further trials, so further work should involve trialling and optimising crystallisation screens for these mutants. Structural work with these mutants would reveal further insight into whether mutations at these positions affect CBD binding and/or entry into the NavMs channel.

5.3 Discussion:

This research provides for the first time an insight into the interaction between CBD and VGSCs.

From this research, several important findings were made. First, that crystallisation trials were possible with NavMs protein in the presence of CBD, that NavMs F208L yielded more crystals than NavMs WT, and those crystals tended to be to a higher resolution, indicating that NavMs F208L protein is a desirable candidate for use within drug binding investigations. In addition, analysis of NavMs F208L crystals with no drug present indicated no significant difference in structure when compared to NavMs WT, making it an excellent candidate model when wanting to investigate CBD binding, and indicating its possible future use within future drug-NavMs research using crystallography.

Second, data identifying CBD within the fenestration of the NavMs F208L protein was collected. This data is important because it provides the first structural evidence as to how CBD interacts with sodium ion channels, thus offering a mechanism by which it inhibits these channels. Binding appears to occur deep within the fenestration, which suggests CBD entry via the fenestration, leading to blockage of the channel pore, thus inhibiting sodium ion movement through the channel. This block could aid in explaining CBD’s apparent effectiveness as a treatment for epilepsies, since many epilepsies involve mutations within sodium ion channels. In addition, the general nature of CBD binding, along with previous research indicating that CBD acts to inhibit multiple human sodium ion channel isoforms, could help explain why CBD appears to have effects within other sodium ion channel related issues, such as pain and muscle spasticity. From this structural data, it was also possible to explore possibilities around why there are differences in inhibition slopes for THC and CBD, as well as to compare CBD binding in NavMs with CBD binding in TRPV2.

Chapter 6: Electrophysiology Experiments to Investigate the Effects of cannabidiol on NavMs Function

6.1 Introduction:

To explore whether CBD does functionally inhibit NavMs, as suggested by the crystal structure from Chapter 5, external collaborators (Dr. Mohammad-Reza Ghovanloo and Prof. Peter C Ruben from the Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Canada) were brought in to perform whole-cell voltage-clamp studies of cells transiently transfected with the NavMs gene.

Whole cell patch clamp is a form of electrophysiological experiment which involves recording currents through multiple channels simultaneously over a large span of cell membrane. This technique has already been successfully utilised to investigate the inhibitory effects of CBD on VGSCs (Fouda *et al.*, 2020; Ghovanloo *et al.*, 2018, 2021; Hill *et al.*, 2014; Mason and Cummins 2020; Patel *et al.*, 2016)

Inclusion of electrophysiological studies carried out by our collaborators allowed for functional analysis of CBD's effects on NavMs.

To study the binding site of CBD, studies were also carried out with the mutant T207A. The position T207 was selected because T207 appears within the proposed binding site for CBD and has been previously associated with drug binding (Bagnéris *et al.*, 2014).

Please refer to methods section 2.8 for details on electrophysiology experiment methodology.

6.2 Results:

Previous studies indicated little selectivity in CBD inhibition when it comes to human Navs and the prokaryotic sodium channel NaChBac (Ghovanloo *et al.*, 2018). In the work for this research, it was discerned that CBD does inhibit NavMs (Fig 6.1), but less potently (17.8 ± 0.5 μM) and with a slightly shallower Hill slope (1.5 ± 0.1) than human Nav channels that had been previously investigated. This moderate variation is consistent with previous observations of other sodium channels blockers (Bagnéris *et al.*, 2014). Overall, results indicated that CBD does inhibit NavMs in a similar fashion to human Navs and support the proposed binding of CBD within NavMs as a functionally relevant mechanism of action.

To investigate the proposed binding site within NavMs, CBD block was recorded for NavMs with the T207A mutation. T207 appears within the proposed binding site for CBD (see Fig 5.11) and has been previously associated with drug binding (Bagnéris *et al.*, 2014). The aim of mutating this threonine position to an alanine was to knock out this residue's ability to form hydrophobic interactions with the CBD molecule, since the alanine side chain is much smaller and would not reach as close to the CBD molecule. Measurements of the CBD block of NavMs T207A indicated that mutating T207 did result in a modest but significant reduction in CBD block (Fig 6.1), which correlates with the observations made of CBD binding location, CBD binding characteristics, and links the location of T207 to a functional effect.

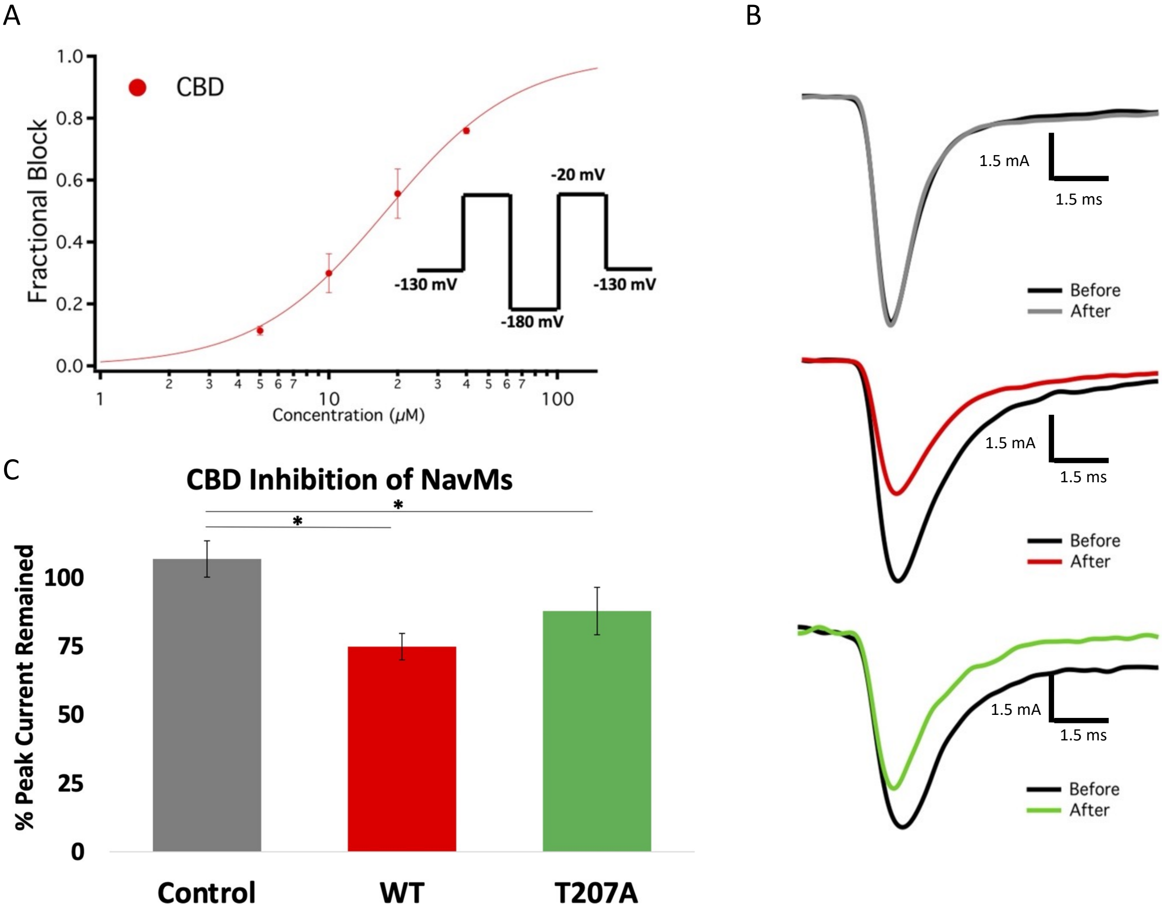


Figure 6.1: Electrophysiology experiments investigating CBD inhibition of NavMs. *A) Block by CBD was measured after ~six minutes wash and incubation in CBD. The IC50 for the inhibition of NavMs WT was calculated as 17.8 ± 0.5 µM, with an accompanying Hill slope of 1.5 ± 0.1. B) Sample traces for NavMs WT without CBD added (top), NavMs WT pre- and post- perfusion with 10 µM CBD (middle with pre-perfusion trace in black and post-perfusion trace in red) and NavMs T207A pre- and post- CBD perfusion with 10 µM CBD (bottom with pre-perfusion trace in black and post-perfusion trace in green). C) Bar graph depicting the percentage of peak sodium ion current remaining after perfusion with 10 µM CBD, with the control (NavMs WT without CBD) in grey, NavMs WT with CBD perfusion in red, and NavMs T207A with CBD perfusion in green. Number of repeats = 4-6. Statistical comparison against the control (grey): NavMs WT with CBD perfusion (red) = p<0.0027 and NavMs T207A with CBD perfusion (green) = p=0.0269. Cited from (Sait et al., 2020) and figures produced by Dr. Mohammad-Reza Ghovanloo and Prof. Paul Ruben). A) and B) were done using Igor Pro (Wavemetrics, Lake Oswego, Oregon, USA) and A)-C) were put together in Microsoft PowerPoint (Microsoft Corporation, Albuquerque, New Mexico, USA).*

6.3 Discussion:

Electrophysiology work was carried out to explore whether CBD inhibited NavMs in a way that was consistent with structural findings about the CBD binding site, and consequently its possible mode of action as an inhibitor.

Electrophysiology experiments highlighted that CBD does indeed act to block ion passage in NavMs, and that mutation of the T207 position had an effect on CBD's blockage of NavMs, thus linking T207 to a functional effect. Overall, these results compliment findings and proposed modes of action discerned during structural studies (see Chapter 5: NavMs WT and NavMs F208L Crystallisation Trials and X-ray Crystallography).

Chapter 7: Research Conclusions

Despite the wealth of electrophysiological and functional studies which evidence the interactions between VGSCs and CBD (Fouda *et al.*, 2020; Ghovanloo *et al.*, 2018, 2021; Le Marois *et al.*, 2020, Mason and Cummins, 2020; Patel *et al.*, 2016), interactions had not yet been studied or identified on a structural level. This lack of structural work previously meant that interactions between CBD and VGSCs were not understood on a structural level, causing a lack of knowledge of the inhibiting mechanism of CBD when interacting with sodium ion channels.

The aim of this doctoral research involved using X-ray crystallography, electrophysiology, and CD techniques to examine the interactions between CBD and VGSCs on a structural and functional level. NavMs was used as a model for human sodium channels, since it has been previously characterised, and has suitable binding and affinity kinetics for human sodium channel blockers, making it an ideal candidate model for studying CBD interactions, since previous functional research suggests CBD acts as a blocker of VGSCs.

7.1 Expression and purification of NavMs Wild Type (WT), NavMs F208L and hypothesised binding site mutants:

The prokaryotic channel NavMs was a great candidate model for studies into CBD and VGSC interactions, since the NavMs channel has evidenced sequence, structural and functional similarities to human VGSCs , as well as similar binding and affinities for sodium channel blockers, which CBD has been evidenced to be (Bagnéris *et al.*, 2013, 2014; Boiteux *et al.*, 2014; Lenaeus *et al.*, 2017; McCusker *et al.*, 2012; Nurani *et al.*, 2008; Oelstrom *et al.*, 2014; Payandeh *et al.*, 2011, 2012; Sula *et al.*, 2017, 2021; Zanatta *et al.*, 2019, Zhang *et al.*, 2012). In addition, NavMs is easier to produce than its human sodium ion channel counterparts due to the ability to use an *E. coli* expression system, cheap maintenance and production, and high-level protein production that creates suitable amounts of protein for structural studies.

Overall, NavMs WT and NavMs F208L protein were successfully produced and evidenced to be NavMs protein based upon matches between UV traces, SDS-PAGE gels, and immunoblotting with what has been previously seen and characterised for NavMs protein. Thanks to this work, research into the interactions between VGSC and

CBD could commence during this research, since abundant amounts of starting NavMs protein was prepared and ready for use within crystallisation screens. The fact that NavMs F208L was successfully produced, and produced in equal levels as NavMs WT, also gave insight into the NavMs F208L protein, in that it appears to maintain major folding patterns and an ability to form multimers, as well as being expressed equally as well as NavMs WT. Consequently, NavMs F208L can be produced in the same manner as NavMs WT, allowing for sufficient protein to be produced to carry out structural studies. These structural studies involved researching the effect of the F208L mutation on NavMs structure, as well as using NavMs F208L within crystallisation screens involving CBD, since previous work had indicated that NavMs F208L crystallises with a higher success rate than NavMs WT.

In addition, this research confirmed that hypothesised NavMs CBD binding site point mutants (NavMs M20WA/I/W, T207A, F214A) (selected on the hypothesis that CBD enters via the fenestration and would bind in a similar location to other previously investigated hydrophobic sodium channel blockers) can be successfully expressed and purified, allowing for the possibility of future structural experiments.

7.2 Circular dichroism spectroscopy of NavMs proteins:

CD can be a powerful tool when studying protein structure, function and/or interactions with ligands. Thermal melt CD is based upon the fact that secondary structural elements contribute to CD spectral signals, but when a protein is heated, it can begin to lose these secondary structural elements via unfolding, causing the CD spectrum to change. This change can be used to determine the effects of ligands on protein stability, thus giving an indicator of interactions between ligand and protein (Di Lella *et al.*, 2010; Fan *et al.*, 2007; Güler *et al.*, 2020; Huus *et al.*, 2006; Ireland *et al.*, 2018; Marada *et al.*, 2016; Metrick *et al.*, 2013; Szenczi *et al.*, 2006). In the case of this research, the aim of utilising CD, specifically thermal melt CD, was to analyse NavMs WT protein stability with and without CBD. These observations indicated if the NavMs WT protein was generally stable and folded with and without CBD present, thus signifying if NavMs was still a feasible model for crystallographic studies. These studies were also used to investigate whether the addition of CBD had any impact on the stability of the NavMs protein when compared to the protein without CBD, thus giving an indicator of binding interactions.

After collecting and processing data from NavMs WT samples with CBD added and comparing to data collected from NavMs WT without CBD added, CBD did not appear to have a significant effect on either NavMs WT protein secondary structure or on NavMs WT thermal stability. From these observations, it was hypothesised that the lack of effect on stability could due to the nature of CBD:VGSC interactions; thus crystallographic studies were needed to investigate CBD binding interactions, and consequently this hypothesis.

7.3 NavMs WT and NavMs F208L crystallisation trials and X-ray crystallography:

X-ray crystallography is a fundamental tool within research that investigates protein structures. In particular, X-ray crystallography allows investigations into protein-drug interactions (Aitipamula and Vangala, 2017), enabling insights into the mechanisms underpinning drug effects on a structural level. X-ray crystallography has already been successfully used to reveal structures of the full NavMs model protein with and without ligands , giving great insight and deepened understanding of sodium ion channel functioning, as well as the mechanisms behind sodium ion channel targeting ligands and drugs (Bagnéris *et al.*, 2013, 2014; Ke *et al.*, 2018; McCusker *et al*., 2012; Sula *et al*., 2017, 2021; Sula and Wallace, 2017; Ulmschneider *et al*., 2013).

The aim of including X-ray crystallography using the NavMs protein was to collect diffraction data of the NavMs protein with and without the presence of CBD so as to gain structural insight into the inhibitory effects of CBD. In addition, these experiments allowed for investigation of the hypothesis that CBD would act like other hydrophobic sodium channel blockers and enter via the fenestration and bind in a similar position within the channel.

The elucidation of the NavMs F208L-CBD structure showed a novel binding site deep within the NavMs fenestration, which is close to, but not coincident with, a known analgesic binding site. Identification of this binding site supported the initial hypothesis that CBD would enter the channel and act in a similar fashion to other known hydrophobic sodium channel inhibitors. Analysis of this binding site showed how CBD interacts with sodium ion channels, and provides a mechanistic explanation of CBD’s inhibition of sodium ion channels. Analysis of the binding site also reveals why CBD appears to be non-selective across human sodium ion channels, since binding seems to mostly rely on the existence of a hydrophobic environment in which the hydrophobic

CBD can readily enter the fenestration structure and form hydrophobic interactions with the respective residues. The evidenced structural underpinnings of CBD binding supported the hypothesis that the observed thermal melt CD results were due to the nature of VGSC:CBD binding interactions, since this type of binding (involving mostly hydrophobic interactions) would not alter NavMs stability.

In addition, this structure allowed for investigations into why THC and CBD appear to have differences in inhibition slopes when it comes to the inhibition of sodium ion channels. The structure also allowed for structural comparisons with the TRPV2-CBD structure, highlighting that CBD appears to bind around the fenestration site in TRPV2, albeit in a different area than in NavMs, which could explain why CBD acts as an activator of TRPV2 but an inhibitor of NavMs.

Additional structural investigations of NavMs F208L without CBD gave comparative insight into the structure of NavMs F208L versus NavMs WT, indicating that NavMs F208L is structurally unchanged from NavMs WT. This, in conjunction with the fact that NavMs F208L crystallises with a higher success rate, indicates that NavMs F208L can be used in other crystallisation and crystallography trials that wish to investigate other sodium ion channel targeting ligands and drugs.

It must be noted that initial crystallisation trials were set up with binding site mutants, however, these trials either yielded no crystals or poorly diffracting crystals. Due to time and lab constraints imposed by the Covid pandemic, it was not possible to perform further trials, so further future work would involve trialling and optimising crystallisation screens for these mutants.

7.4 Electrophysiology experiments to investigate the effects of cannabidiol on NavMs function:

To confirm that CBD does functionally inhibit NavMs, as suggested by the elucidated crystal structure, external collaborators (Dr. Mohammad-Reza Ghovanloo and Prof. Peter C Ruben from the Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Canada) performed whole-cell voltage-clamp studies of cells transiently transfected with the NavMs gene.

Electrophysiology experiments highlighted that CBD does indeed act to block ion passage in NavMs, and follow-up studies using NavMs T207A linked T207 to a functional effect, supporting its role within CBD binding.

Overall, these results complimented findings from structural studies.

7.5 Future research:

Future work that would aid further understanding of CBD binding could involve taking the NavMs binding mutants, optimising crystallisation, then performing crystallographic experiments. These experiments could involve further research into the interactions between NavMs and CBD, as well as general investigations into the effect of these mutations on NavMs structure. These mutants could also be used within other research areas that aim to investigate other ligands and their interactions with NavMs. The M204 mutants are interesting because they provide a range of fenestration sizes, and thus could be very useful for studying the effects of fenestration size on ligand entry and binding.

In addition, both NavMs and human Nav binding site mutants could be produced and used within further electrophysiology work to further our understanding of the binding site, and the relative importance of each binding site residue. In particular, it would be interesting if future work, both computational and lab-based, investigated human Nav fenestrations in more detail to identify residues that CBD could form hydrogen bonds with, and thus indicate the residues that may affect binding more drastically if mutated, consequently providing insight into what affects CBD binding most significantly within different human sodium ion channels.

7.6 Overall:

Looking over the body of work contained within this doctoral research project, it is clear that these investigations have achieved structural insight into how CBD interacts with VGSCs, which is vital to understanding how CBD interacts with VGSCs within the body and how CBD exerts its inhibitory effects upon these channels. In addition, this work highlighted and validated the use of NavMs F208L within future VGSC drug-binding research, since it crystallises with greater success than NavMs WT and appears to have little structural difference. In addition, this work has enabled the identification and production of binding site mutants that can be utilised within further structural work focusing on CBD binding, or even within other studies interested in these residues and their significance within fenestration dimensions, other ligand interactions, and/or general significance within the NavMs structure.

Overall, this work acts as a vital platform, and hopefully catalyst, for further continued research investigating CBD interactions; either further work to understand the interaction with VGSCs and these interactions' significance, or structural research investigating other possible CBD targets. This type of work will aid in understanding mechanistically how CBD is affecting targets, and thus aid in building a deeper knowledge of how CBD can be utilised within the medical field.

Chapter 8: Appendix of Methods Theory

8.1 Expression and Purification of NavMs Wild Type (WT) and NavMs F208L Methods Theory:

8.1.1 Bioinformatics

Overview: Bioinformatics involves utilising computational methods to aid in the understanding of large complex biological datasets. In structural biology, a major area of bioinformatics that aids in protein modelling are sequence alignments, specifically amino acid/protein sequence alignments. Alignment involves arranging sequences to identify similarities between each sequence involved. Alignment can aid in identifying regions of significance, be they evolutionary, structural, or functional (Gallery, 2005, p 692). There are two major types of sequence alignments, pairwise sequence alignments (PSA) and multiple sequence alignments (MSA). Pairwise alignments involve aligning two sequences and ascertaining the best matches between them, and thus regions of possible importance. Multiple sequence alignments involve the alignment of three or more sequences of a similar length. Multiple sequence alignments allow for homology detection across the sequences, enabling inferences about evolutionary relationships. Sequence alignments (both PSA and MSA) can fall under global alignment or local alignment methods. Global alignment involves calculating the similarity between sequences over their entire length. Local alignment involves identifying regions of similarity across sections of sequences. The issue with global alignment techniques, is that they run into issues when sequences are divergent across large spans of their sequences, only have local homology over sections within their sequences, and/or contain the presence of shuffled domains of similar sequence sections. Consequently, in these cases, local alignment is performed, allowing for local areas of similarity to be discerned (Chowdhury and Garai, 2017).

Computer algorithms designed to carry out alignments come in a variety of forms, including dynamic programming, heuristic programming and probabilistic models. Dynamic programming algorithms provide optimal alignment, however, once three or more sequences are included, the optimal alignment reached is often not biologically optimal. In addition, dynamic programming is complex, requires a lot computationally, and can suffer from high-dimensional and complexity issues. Consequently, MSA often utilise heuristic or probabilistic modelling algorithms, which are computationally more efficient than dynamic programming and provide feasible alignment within a short time frame, but do not offer a guarantee of finding the optimal matches (Chowdhury and Garai, 2017).

Databases and software: Uniprot is a comprehensive and open source of protein sequences and associated functional data (such as disease-related mutations) (Bateman, 2019). Consequently, sequences acquired from Uniprot can be used for alignments. EMBOSS NEEDLE (Madeira *et al.*, 2019) is an online tool for creating pairwise alignments, and employs the Needleman-Wunsch algorithm, which is a dynamic programming algorithm used to carry out global alignment between sequences (Needleman and Wunsch, 1970). Clustal Omega (Sievers *et al.*, 2011), a heuristic algorithm that is a multiple sequence alignment tool, uses seeded guide trees (a method of clustering which guides development of the order of sequence alignments (Sievers and Higgins, 2018)) and profile hidden Markov model techniques (a profiling tool which provides a consensus on the align moment, and includes set position-specific probabilities for inserts and deletions (Söding, 2005)). EXPASY translate allows for the translation of a DNA sequence into an amino acid sequence, making it useful when creating DNA primers and checking DNA sequences (Gasteiger *et al.*, 2003).

8.1.2 Constructs and cloning

Molecular cloning: Molecular cloning involves taking a target section of DNA sequence and inserting it into some form of genetic vector for propagation. The copies of the recombinant genetic vector can be utilised in the expression of the resulting protein encoded by the inserted DNA sequence, allowing for study of the protein's structure and function.

NavMs WT genetic cloning: The first instance of the NavMs gene being isolated from *Magnetococcus marinus* was in D.E. Clapham's lab, (McCusker *et al.*, 2012) and was donated to the Wallace lab. Dr. Claire Bagnéris is responsible for successfully cloning the NavMs gene into a Pet15b plasmid vector, along with a His-tag (McCusker *et al.*, 2012).

Site-directed ligase-independent mutagenesis cloning: To introduce site-specific point mutations into a gene construct, Site-directed Ligase-independent Mutagenesis (SLIM) can be used (Chiu *et al.*, 2004). SLIM works via ligase-independent cloning, which is a polymerase chain reaction (PCR) method involving primers which contain the point mutation of interest, and have complementary overhangs, allowing for ligation of the PCR-plasmid, and the consequent creation of plasmids with the desired mutation.

Gel electrophoresis: Gel electrophoresis involves loading DNA samples into wells inside one end of an agarose gel, which then has a current run across it by special apparatus. Because of DNA's inherent negative charge, it will migrate towards the cathode, and thus move through the gel. The speed of travel of a DNA fragment is usually inversely proportional to its length, thus allowing for DNA fragment identification via size. A DNA ladder, with DNA fragments of known lengths, is used alongside the samples, so that sample length can be determined (Carter and Shieh, 2015, pp 220-236). Consequently, gel electrophoresis can be used to validate successful cloning.

Transformation: As mentioned, bacterial transformation involves bacteria being able to take up foreign DNA from the environment. This phenomenon can be exploited to create bacteria containing plasmid constructs, allowing for up-scaled production. Transformation via heat shock is a simple technique used within labs to insert foreign plasmid material into bacteria. Heat shocking bacteria causes holes within the bacterial plasma membrane, allowing for foreign plasmid constructs to enter the bacterial cell (Froger and Hall, 2007).

Plasmid harvesting: To extract and purify plasmids, kits such as 'Wizard Plus SV Miniprep kit' are used. This kit works by lysing the cells via the action of sodium dodecyl sulfate (SDS) and alkaline buffers, and then loading the lysed product on to silica-based spin columns. These spin columns bind the DNA, which is washed and then eluted using nuclease-free water (Promega Corporation, 2009).

To determine final plasmid concentration a spectrophotometer is often used. A spectrophotometer can help determine average DNA concentrations by detecting how much ultraviolet (UV) light at 260 nm is absorbed by the DNA in a sample. The more light absorbed by the sample, the higher the DNA plasmid concentration.

DNA sequencing: DNA sequencing describes the process of attaining the order of nucleotides within a DNA sequence. Sequencing can be used to check plasmid construct sequences once they have been harvested and purified. Sanger sequencing is often used for sequencing DNA. Sanger sequencing is based upon the use of dideoxynucleotides, which block DNA polymerisation. These are fed in simultaneously with deoxynucleotides during DNA amplification and are usually labelled with a fluorophore. When a dideoxynucleotide is incorporated the DNA amplification stops. The DNA molecule lets off a signal which reports the last nucleotide incorporated (adenine, thymine, cytosine, or guanine). These DNA molecules of differing lengths are sorted via size, revealing the complete nucleotide sequence (Garrido-Cardenas *et al.*, 2017).

8.1.3 Protein expression and purification

Overview: Protein expression describes the synthesis of target proteins using expression systems. Expression systems can be mammalian, insect, bacterial, yeast, algal, or even cell free. Bacterial expression systems are often used because they are easy to maintain and genetically manipulate. In addition, bacterial expression systems are usually a cost-effective option.

Expression of a target protein within an expression system involves genetic transcription and translation of a gene, which has often been introduced into a host via a recombinant genetic plasmid. Transcription is the process of the target protein's DNA sequence being copied into a new molecule known as messenger ribonucleic acid (mRNA). This mRNA can then go on to the next step, termed translation, which involves synthesising the protein based upon decoding of the mRNA molecule by a ribosome (macromolecular machines found in cells which carry out protein synthesis).

Once protein has been successfully expressed, it must then be purified from the rest of the constituents. This step is known as protein purification and can involve several different steps depending upon the target protein in question. First, the protein must go through preliminary extraction from the expression system. The extraction phase can involve sonication, repeated freeze-thawing, homogenisation or permeabilisation of expression system cells via the use of detergents. Once a crude extraction has occurred, ultracentrifugation is often another step utilised in preliminary protein purification (especially in membrane protein purification), as it causes the formation of a pellet enriched with large dense particles (membrane in the case of membrane protein purification). This pellet can then be separated from the remaining supernatant and used in further purification steps. These further purification steps often involve one or more forms of chromatography, which shall be discussed in more detail in this section.

Expression Challenges: Expressing membrane proteins is a difficult process. Membrane proteins are generally toxic to cells used during over-expression and require either an amphipathic lipid or detergent to maintain folding and functionality. However, even if the tertiary structure of a membrane protein is preserved, maintenance of activity cannot be assured once the protein is out of the native membrane environment.

The ability to over-express and purify membrane proteins has progressed significantly (Schlegel *et al.*, 2010). The major difference in membrane protein expression techniques, compared to soluble protein protocols, is the isolation of membranes. Once cells have been lysed and centrifuged to remove cell debris, the soluble fraction is taken and spun in an ultracentrifuge (up to 40000 X gravitational force (g)) for two or more hours. This leads to the membrane fraction sedimenting out and forming a pellet, which is then collected and resolubilised in lipids or detergents, and then used within the next protein purification steps to attain the membrane protein.

Prokaryotic NavMs and protein expression: The advantage of expressing and purifying prokaryotic VGSCs instead of human VGSCs is that they can be expressed within *E. coli*. *E. coli* is the cheapest host available for protein expression. There are a range of commercial genetic vectors and a number of *E. coli* strains tailored for membrane protein over-expression (Gubellini *et al.*, 2011). Consequently, *E. coli* expression makes over-expressing several different membrane protein constructs significantly easier'.

Chromatography: Chromatography is the term used to describe the technique of separating a mixture which exists within a soluble mobile phase. The mobile phase containing the mixture is allowed to pass across a stationary phase, leading to the separation of the mixture constituents. Chromatography is used to purify target proteins. Affinity chromatography and size exclusion chromatography (SEC) are both different forms of chromatography. Affinity chromatography separates things out based upon a specific interaction between the molecules within the mobile phase and molecules within the stationary phase. SEC separates molecules within the mobile phase based on size.

Affinity chromatography can separate out membrane proteins based upon the presence of a His-tag on the target protein. When purifying a target membrane protein which has a His-tag, resolubilised membranes are passed through a HisTrap column (a nickel affinity stationary phase column), causing the His-tagged protein to be bound due to its high affinity for nickel. Non-specific binding is avoided by high concentrations of salt being present within the buffer. Bound protein is afterwards eluted with increasing amounts of imidazole, which works by competing with the His-tagged protein for interaction with Ni^{2+} binding sites.

SEC entails passing a mixture (such as a protein sample) through a column filled with a gel matrix consisting of porous sepharose beads. The smaller the molecule, the more space within the bead holes the molecule can occupy, causing elution time to be longer for smaller molecules. This allows for the separation of molecules by hydrodynamic radii.

Sodium dodecyl sulfate and polyacrylamide gel (SDS-PAGE): SDS-PAGE is a method that allows for separation and visualisation of protein based upon molecular weight. SDS-PAGE is used to validate protein samples after purification. SDS stands for sodium dodecyl sulfate. SDS is a surfactant which coats protein within a sample. This coating covers a protein's intrinsic charge, conferring similar mass:charge ratios to proteins in a sample.

SDS is a reducing agent and is found within sample buffer. Protein samples are mixed with this sample buffer. The reducing agent breaks di-sulfide bonds, disrupting secondary and tertiary protein structures, causing denaturing. Denatured proteins are evenly coated in negative charge by the SDS. The protein sample is then loaded on to a gel with a current running across it. The protein travels across the gel towards the cathode. The speed at which the protein travels through the gel is determined by its molecular weight, thus protein size can be determined using protein ladder markers.

Membrane proteins like NavMs often appear in a gel at a position that does not correspond with their actual molecular weight. This is explained by increased association with the SDS detergent molecules (Rath *et al.*, 2009). Banding and streaking are also often observed even within pure membrane protein samples like NavMs (Zhou *et al.*, 2000). The higher molecular weight bands, which indicate different multimeric states, suggest that these multimers are particularly stable.

Immunoblotting: Another method used for validating protein samples is immunoblotting. Immunoblotting is an analytical tool which is used to detect specific proteins via the use of antigens. It is a multi-step process starting with the separation of a protein sample by size via an SDS-PAGE. This is followed by transferring the proteins onto a membrane via blotting, and this membrane is then subjected to probing with antibodies. Immunoblotting can be a very sensitive tool, and is able to detect protein at very low concentrations. In addition, it can be very specific, thanks to the initial size separation due to the SDS-PAGE gel combined with the specificity of antigen binding.

8.2 Circular Dichroism Spectroscopy of NavMs Proteins Theory:

8.2.1 Overview

CD involves the differential absorbance of left- and right- polarised light by a chiral molecule. A plane-polarised light beam can be considered to consist of equal amounts of left- and right- circularly polarised light (Fig 8.1). The differential absorbance of left- and right- circularly polarised light by a sample of optically active molecules results in the plane-polarised light beam becoming elliptically polarised (Fig 8.2).

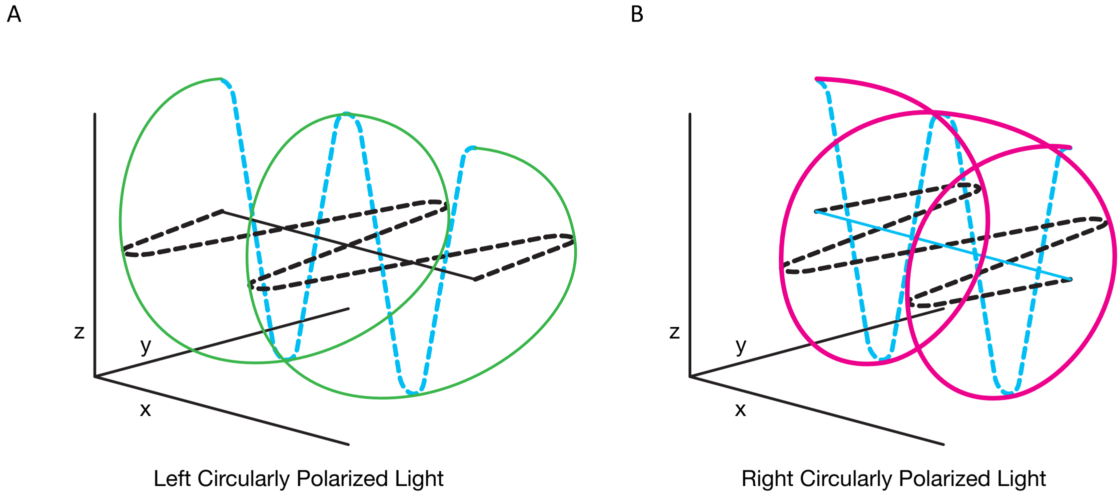


Figure 8.1: Circularly polarised light. *Light is an example of an electromagnetic wave, composed of perpendicular electric and magnetic fields. In both A) and B) the electric fields of two light waves are shown with ¼ wavelength difference between them using blue and black dotted lines. A) Visual representation of left-circularly polarised light (solid green line), which results from the constructive and destructive interference of the two light waves. B) Visual representation right-circularly polarised light (solid pink line), which results from the constructive and destructive interference of the two light waves. Produced using Affinity Designer (Serif Europe, Nottingham, UK).*

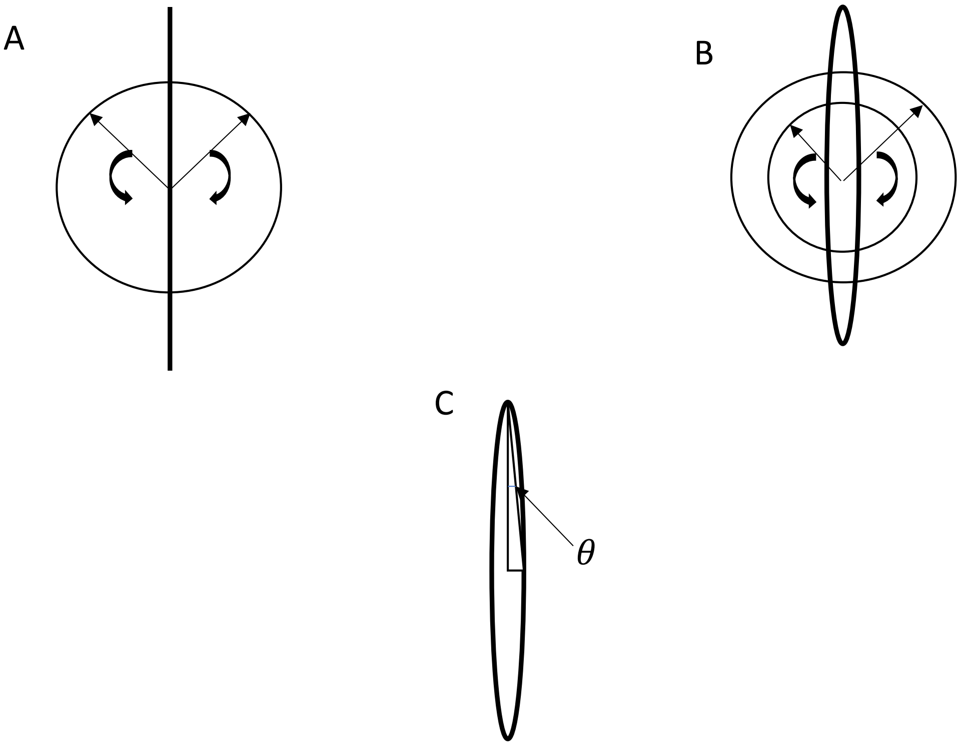


Figure 8.2: Electric vector components of plane-polarised light beam (after (Wallace and Janes, 2009)). *These views are of the electric vector components of a plane-polarised light beam consisting of left- and right- circularly polarised light along the direction of that beam's propagation. A) A visual representation of when the plane-polarised light beam passes through a sample that is achiral. Both left- and right-polarised light are equally absorbed, thus no CD signal is formed. B) Left- and right-circularly polarised light are unequally absorbed by a chiral sample, causing the resultant beam to have an elliptically polarised path. C) Visual representation of the ellipticity measure* θ*. Produced using Microsoft PowerPoint (Microsoft Corporation, Albuquerque, New Mexico, USA).*

Ellipticity is defined as the tangent of the ratio of the minor and major elliptical axes (Fig 8.2) and is measured in millidegrees. However, this measurement depends upon sample concentration, thus other measurements are usually used to report CD findings. A commonly used measure is Delta Epsilon, which is defined as:

$$\Delta\varepsilon = \varepsilon_L - \varepsilon_R = A_L - A_R/(dl)$$

Equation 8.1: Delta Epsilon definition: *Delta Epsilon is a commonly used measure within CD.* ε_L*and* ε_R *are left and right extinction coefficients for the left- and right-circularly polarised light and are usually calculated per residue in a macromolecule.* A_L *and* A_R *are the respective absorbances of left- and right- circularly polarised light by the sample. l is the pathlength (cm) of the sample that the plane-polarised light beam passes through. d is the molar concentration of the sample. Delta Epsilon values have units defined as* M^{-1} cm^{-1} *.*

Another measure used to report CD findings is mean residue ellipticity (MRE), which is defined as:

$$[\theta]_{MRE} = 3298\,\Delta\varepsilon$$

Equation 8.2: Definition of mean residue ellipticity: *Mean residue ellipticity has the units degrees* cm^2 $dmol^{-1}$ $residue^{-1}$ *and is directly linked to Delta Epsilon (*$\Delta\varepsilon$*).*

MRE can also be defined as:

$$[\theta]_{MRE} = (\theta \times 0.1 \times MRW)/cl$$

Equation 8.3: Second definition of mean residue ellipticity: *The direct measures of ellipticities (*θ*) can be converted into MRE via this equation. The mean residue weight of the sample is represented by MRW (measured in Daltons (Da)), c is the molar concentration of the sample, and l is the pathlength of the sample (measured in centimetres). MRW is calculated as the molecular weight of the sample divided by the residue number minus one.*

The ability to convert directly measured ellipticity values into $[\theta]_{MRE}$ or $\Delta\varepsilon$ is very useful because it allows for measures to be independent of pathlength and protein concentration, allowing for comparisons of measurements independent of their conditions, hence why these measures are used within the CD literature.

The chirality of a molecule often arises from asymmetry in the attachment of groups around a carbon atom. Amino acids are chiral in nature due to having four non-identical groups attached to a central ‘alpha’ carbon. Because amino acids are chiral, proteins made up of amino acids are also chiral in nature, meaning that CD signals are produced when plane-polarised light is passed through a sample.

The differential absorbance of left- and right- circularly polarised light by a chiral molecule can be detected if the chiral molecule contains groups which can absorb light, which are known as chromophores. A chromophore is defined as a self-contained region in a molecule where specific energy transitions can be attributed. For proteins, the primary chromophore of interest is the peptide bond. A CD signal is produced when an electron in a ground state is excited to another orbital (Wallace and Janes, 2009, pp 1-18). The key electron orbital transitions within a peptide bond are from n to π* and π to π*. Transitions from a n to a π* orbital cause a CD signal at ∽220 nm, whilst transitions from π to π* cause a signal at ∽ 190 nm. 190 nm and 220 nm fall within the far UV region of the light spectrum (~260 nm - ~190 nm). Most CD measurements involving biomacromolecules occur within the UV range of light (Ranjbar and Gill, 2009; Wallace and Janes, 2009, pp 1-18).

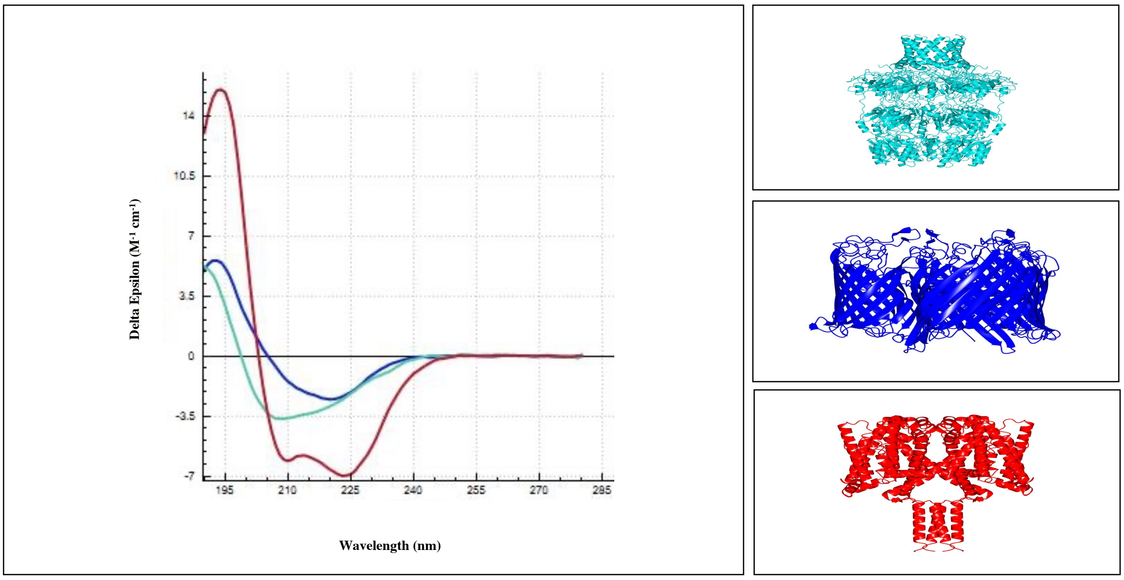


Figure 8.3: CD spectra for membrane proteins with different secondary structures. *The left panel of this figure displays the distinct spectral patterns for membrane proteins with different secondary structural types: mixed helical or beta sheet structures (turquoise: WZA translocon for capsular polysaccharides (Dong et al., 2006)), beta-barrel structures (blue: sucrose-specific porin (Forst et al., 1998)), and alpha-helical structures (red: NavMs prokaryotic full-length sodium channel (Ireland et al., 2018)). The CD spectra were obtained from spectral data deposited in the protein circular dichroism data bank (PCDDB) (Whitmore et al., 2017). PCDDB IDs: CD0000128000, CD0000127000, and CD0006141000 respectively. The side panels display the corresponding crystal structures of the spectra used within the main panel, with the structures displayed in a matching colour scheme. PDB IDs: 2J58, 1A0S, and 5HVX respectively. CDToolX (Miles and Wallace, 2018) was used to produce the main panel CD spectra figure, and CCP4mg (McNicholas et al., 2011) was used to produce the corresponding crystal structures.*

8.2.2 Thermal melt CD

Secondary structural elements, like α-folds and β-pleated sheets, produce characteristic CD spectral signals. However, when a protein is heated, it can begin to lose these secondary structural elements via unfolding, causing the CD spectrum to change. This spectrum change associated with changing temperature can be useful for determining the effects of mutations, ligands, and protein environments on protein stability. In addition, collecting CD data for a range of wavelengths increases the information that can be gathered about the unfolding of the protein in respect to temperature (Greenfield, 2007).

As mentioned, thermal melt studies have been employed to investigate the effects of mutations, ligands, and protein environment on the thermal stability of proteins. Understanding the effects of mutations on thermal stability can reveal what residues are important in the protein's reaction to heat stress, thus giving an indicator of how to target the engineering of proteins to improve their stability (Güler *et al.*, 2020). This knowledge can also give us an insight into how some proteins are able to act as thermal sensors (Marada *et al.*, 2016). Being able to observe an effect of a ligand on protein thermal stability provides evidence and insight into the interaction between the ligand and the protein sample, allowing for greater understanding of how the ligand elicits its effects upon the protein target (Di Lella *et al.*, 2010; Huus *et al.*, 2006). Understanding what environmental factors can optimise the thermal stability of a protein is important because it allows for optimisation of the storage, usage, and study of protein samples (Fan *et al.*, 2007; Metrick *et al.*, 2013; Szenczi *et al.*, 2006).

Thermal melt CD has been successfully implemented in studies involving the NavMs protein, which is pertinent to this current study. (Ireland *et al.*, 2018) utilised thermal melt CD to study the thermal stability of NavMs in different amphipathic environments. That study indicated that thermal CD is a useful tool when wanting to discern amphipathic environments for membrane proteins that maximise stability and aid in maintaining the protein's structure. Thermal melt CD was also used to study the binding of valproic acid (a commonly-used antiepileptic drug) to NavMs (Zanatta *et al.*, 2019). Thermal melt CD spectra were collected for NavMs with and without the presence of valproic acid. That study showed that valproic acid binds and destabilises NavMs. However, this destabilising effect was not discerned in the thermal CD spectra collected for the NavMs pore-only construct, indicating that valproic acid was likely to be binding and destabilising the NavMs protein in a region other than the pore, the classic binding site of many anticonvulsant drugs (Zanatta *et al.*, 2019).

8.2.3 Issues and solutions when using CD spectroscopy for studies of membrane proteins

CD has been utilised to study both soluble and membrane proteins. However, there are issues specific to membrane protein data collection that need to be considered. Primarily, membrane proteins do not form uniform isotropic solutions in aqueous environments, like soluble proteins do, but instead are contained within large anisotropic particles. The types of anisotropic particles used include liposomes (Jia *et al.*, 2020; Verchère *et al.*, 2017, pp 259-282), detergent micelles (Maliar *et al.*, 2020;

Yeliseev *et al*., 2020), unilamellar vesicles (Biner *et al*., 2016; Geertsma *et al*., 2008), nanodiscs (Dörr *et al*., 2014), amphipols (Kleinschmidt and Popot, 2015; Le Bon *et al*., 2018), and orientated lipid bilayers (Bürck *et al*., 2016). These anisotropic particles can cause various artefacts within collected CD data, which need to be accounted for when studying membrane proteins (Miles and Wallace, 2016).

The first issue to consider when using CD to study membrane proteins is called the ‘solvent shift effect’. This effect occurs due to membrane/detergent environments having different physical properties to those of aqueous solutions (membranes/detergents have lower dielectric constants than water). This difference in dielectric constants can influence the relative ground and excited states of the electronic transitions of peptide backbones in anisotropic environments when compared to their soluble protein counterparts in aqueous environments, which possess similar secondary structures (Abdul-Gader *et al*., 2011). To reduce this effect on the secondary structure analysis of membrane proteins, the SMP180 reference dataset was created for the study of membrane protein CD data (Abdul-Gader *et al*., 2011). It was previously observed that reference datasets based upon soluble proteins of known structure do not produce predictions of membrane protein secondary structures that are as reliable as for their soluble protein counterparts (Wallace, 2003). The SMP180 dataset includes both membrane and soluble protein synchrotron radiation circular dichroism (SRCD) spectra for proteins with known and deposited crystal structures, and has been evidenced to produce structural analyses of membrane proteins to a higher accuracy than its counterparts that only contain spectral data from soluble proteins (Abdul-Gader *et al*., 2011). SMP180 is available as an analysis tool in the DICHROWEB server (Whitmore and Wallace, 2004).

A second phenomenon associated with membrane protein CD data is the ‘differential light scattering effect’. This is caused when the suspended particles containing membrane protein have large dimensions compared to the wavelength of the light, which can cause a proportion of the beam to scatter away from the forward direction, reducing the amount of the beam reaching the detector, and creating the impression that more of the beam has been absorbed by the sample than there actually has. One method to try to mitigate the effects of this phenomenon is to reduce the size of the particles containing the membrane proteins, so they are smaller than the beam wavelength. The issue with this method is that it could affect the protein’s conformation due to the tight packing constraints imposed by the smaller particles (Seddon *et al*., 2004). A second option to combat scattering is to physically locate the sample as close to the beam detector as possible, allowing for much of the scattered light to be captured, and thus eliminating the scattering effect from sample collection (Wallace and Mao, 1984).

Another effect witnessed when collecting CD data for membrane proteins is the ‘differential absorption flattening effect’. In usual spectroscopic characterisation of solutions, there is the assumption of a uniform distribution of chromophores throughout the sample, thus the assumption is that, on average, the same number of proteins will be encountered along a beam’s journey through the sample. However, this assumption does not hold for membrane protein samples because membrane proteins are sequestered within particles, causing chromophores to be concentrated in discrete areas rather than uniformly dispersed. This causes a breakdown in the Beer-Lambert law (the concentration of a solution is directly proportional to the amount of light absorbed) (Wallace and Mao, 1984). In consequence, the CD signal for a membrane protein sample will be ‘flattened’ relative to the signal for a soluble protein with similar secondary characteristics that is uniformly dispersed (Wallace and Mao, 1984). The simplest solution to this issue is to try and increase the lipid-to-protein ratio (Mao and Wallace, 1984; Wallace and Mao, 1984), and thus cause an increased dispersion of the protein throughout the sample. However, lipids do contribute to the overall absorbance by the sample which causes less of the polarised light beam to reach through the sample to the detector. This leads to researchers having to find a compromise between the effects of detergent and the effects of particulate membrane protein solutions.

8.2.4 Instrumentation and secondary structure data analysis

When it comes to the instrumentation used for CD, most commercial lab CD instruments involve taking light produced by a suitable source, passing it through a monochromator, a linear polariser and then a polariser modulator, leading to the production of a beam which can rapidly shift between left- and right- circularly polarised light. This beam can then be passed through the sample, and absorbance recorded by a transmission detector (Wallace and Janes, 2009, pp 19-21).

DICHROWEB is one of several platforms that allows for the analysis of CD data via an online platform, allowing for CD spectral data to be studied for information about secondary structure elements, as well as being used to produce graphical analyses comparing calculated and experimental data (Whitmore and Wallace, 2004, 2008). DICHROWEB allows for the use of various deconvolution algorithms and reference datasets, which enable researchers to study the contribution of different secondary structure components to collected CD spectra for a given protein. These algorithms work on the assumption of linear independence and the additivity of different structural components contributing to the final complete spectra. DICHROWEB also provides deconvolution details, normalised root mean square deviation fit parameters, visual graphs of the experimental spectra data, calculated spectra data, and the difference between these experimental and calculated spectra (Whitmore and Wallace, 2008).

Another analysis tool is CDtoolX (Miles and Wallace, 2018). CDtoolX is a software that enables processing (averaging, baseline subtraction, zeroing, and unit conversion), display, comparisons, and singular value decomposition (SVD) analyses of CD spectra data. CDtoolX enables SVD analyses of structural fraction component contributions, which is useful when analysing thermal melt CD, because it allows for the analysis of changes in contributions of spectral species as a function of temperature changes (Miles and Wallace, 2018).

8.3 NavMs WT and NavMs F208L Crystallisation Trials and X-ray Crystallography Theory:

8.3.1 Overview

X-ray crystallography has been a fundamental technique within science for the determination and understanding of macromolecular structures, allowing for breakthroughs within various fields such as medicine, pharmaceuticals and materials science (Aitipamula and Vangala, 2017). X-ray crystallography is a technique which can discern three-dimensional molecular structures from crystals of a protein of interest. Crystals are exposed to an X-ray beam, and subsequent diffracted X-ray beam data is collected. Diffraction data is processed, revealing structural data about the protein forming the crystal.

8.3.2 Crystals and crystal formation

Production of protein crystals is a vital part of protein X-ray crystallography. A single protein molecule is a poor diffractor of X-rays. In a crystal the diffraction pattern of a protein becomes reinforced, allowing for the eventual determination of a three-dimensional structure.’

In X-ray crystallography, the key crystal structure concept is that crystals are built from small identical repeating units (defined as a unit cell) which stack within a regular three-dimensional lattice. The unit cell is formed of a smaller unit called the asymmetric unit. This asymmetric unit within protein X-ray crystallography is the protein molecule. The protein asymmetric unit forms the unit cell via symmetry operations.

To form protein crystals, protein must separate out of solution and assemble into a crystal lattice. This is achieved when protein is held within solution along with a precipitant. Once the solubility limit of the protein is exceeded within the solution, the solution becomes supersaturated and metastable. When both favourable conditions and nucleation sites are present, crystals have the potential to form (Fig 8.4).

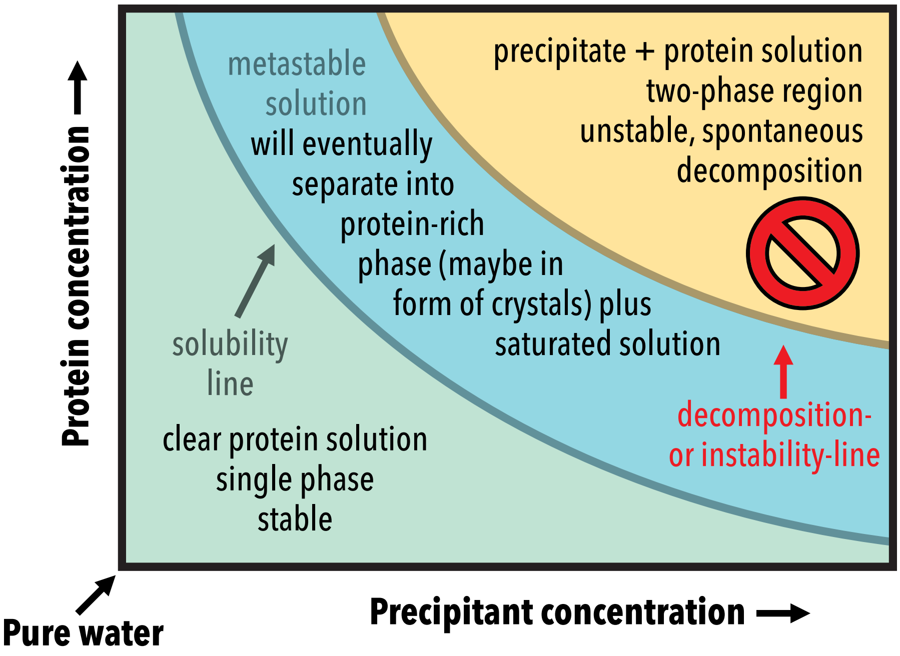


Figure 8.4: Solubility phase diagram (after (Rupp, 2009)). *This diagram gives a general visualisation of the relationship between precipitant and protein concentration within solution. Between the solubility and decomposition lines lies the metastable phase. The metastable phase represents the supersaturated protein solution. Produced using Affinity Designer (Serif Europe, Nottingham, UK).*

Most commonly, the vapour-diffusion method is used to encourage crystal formation. Protein is dissolved within an aqueous buffer containing some form of precipitant, such as polyethylene glycol (PEG), which is at a concentration just lower than that needed to precipitate the protein. Water from the buffer is carefully removed, leading the protein and precipitant concentration to rise, and eventually leading to protein precipitation.

Other methods for achieving crystallisation are batch crystallisation, dialysis, and free-interface diffusion. Batch crystallisation mixes highly concentrated solutions of protein and precipitant, immediately creating a supersaturated protein solution. Dialysis seals protein solution with a dialysis membrane, which allows the free movement of water and small molecules, but acts to block movement of protein and high molecular weight PEG. Free-interface diffusion brings protein and precipitant solutions into contact in a narrow vessel, letting them equilibrate against each other by diffusion.

Forming protein crystals is very sensitive, being affected by factors such as pH, temperature, and the relative concentrations of protein, salt, and precipitant. Membrane proteins have the added factor of also requiring the presence of a lipid or detergent for successful crystallisation. Optimal crystallisation conditions are difficult to predict, thus a multitude of conditions are often screened simultaneously. Commercial screens exist that allow for multiple conditions to be screened at once. Once successful crystallisation occurs, the specific condition leading to crystallisation can be recorded and further optimised.

8.3.3 X-ray diffraction

Bragg's law: When an X-ray beam hits a crystal, a fraction of the rays in the beam are diffracted. Most of these diffracted rays destructively interfere with each other, leading to these X-rays not being detected and recorded. A small percentage of the diffracted X-rays will constructively interfere, leading to the detection of those X-rays.

William Lawrence Bragg showed that X-ray diffraction and wave interference behaviour could be explained by treating diffraction by a crystal as if the X-rays were being reflected off sets of equivalent parallel planes. These planes can be drawn through the crystal lattice. Crystal lattice planes are determined by the faces of the unit cells, and unit cell faces are in turn determined by the unit cell dimensions, designated by three angles (α, β, and γ) and three edges (a, b, and c). Parallel planes within a crystal are designated by indices called Miller indices. With Miller indices, there are three indices (hkl). hkl together identify a particular set of parallel planes. h represents the number of planes per unit cell in the x direction. k represents the number of planes per unit cell in the y direction, and l represents the number of planes per unit cell in the z direction.

Bragg's law states that when waves hit parallel planes within a crystal lattice (with spacing d) at angle θ, and are then reflected by the same angle θ, constructive interference will occur when conditions are met (Equation 8.4). When constructive interference occurs, a detectable X-ray beam is produced, thus contributing to the crystal's X-ray diffraction pattern.

$$2dsin\theta = n\lambda$$

Equation 8.4: Mathematical equation for Bragg's law. *d represents the space between scatterers within the crystal, θ represents the incident angle of the X-ray beam, n represents an integer, and λ is the wavelength of the X-ray.*

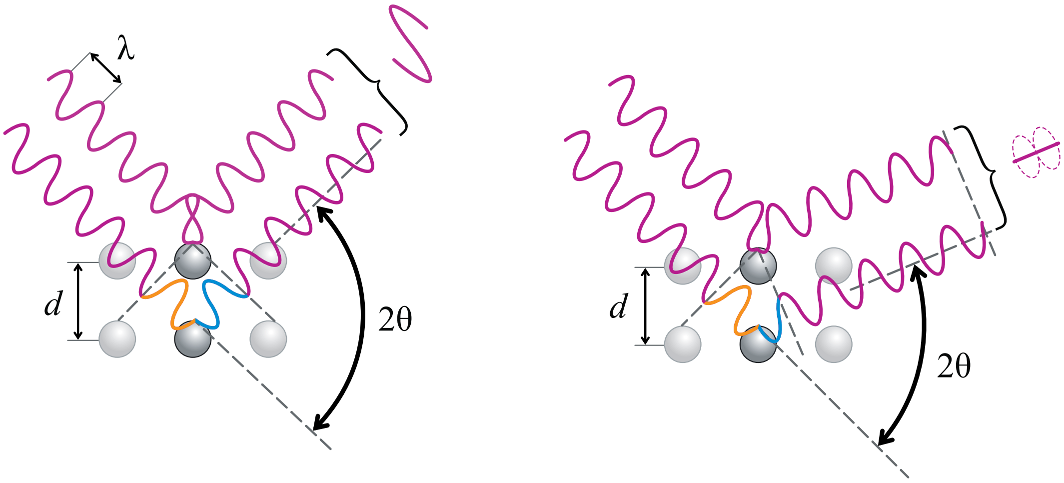


Figure 8.5: Constructive and destructive interference. *Constructive (left) and destructive (right) interference according to Bragg's law. Produced using Affinity Designer (Serif Europe, Nottingham, UK).*

Ewald sphere and reciprocal lattice: The concept of a reciprocal lattice is used to visualise the diffraction from a single crystal, as this crystal is re-orientated within an X-ray beam. The reciprocal lattice can be thought of as consisting of a set of points which are organised in a regular grid, with each point representing a possible diffraction. Points in the reciprocal lattice can be labelled with Miller indices (hkl). These reciprocal lattice indices (Miller indices) correspond to crystal planes from which diffraction could occur. The edges of the unit cells forming the reciprocal lattice are denoted by a*, b* and c*. These are reciprocals of the crystal unit cell edges a, b, and c (a*=1/a, b*=1/b and c*=1/c). The reciprocal lattice planes can be summarised by the vector equation: d*=ha* + kb* + lc*. d* is the reciprocal lattice spacing and is equivalent to the reciprocal of the planar spacing within the crystal (d) (d*=1/d). Consequently, unit cell dimensions of a real crystal can be calculated from the dimensions of the reciprocal lattice and associated diffractions.

The Ewald sphere aids in visualising the diffraction possibilities of a given crystal. The Ewald sphere is a sphere with a radius 1/ λ, where λ is the wavelength of the X-ray beam. The centre of this sphere is a point 1/ λ from a reciprocal lattice point which lies along the direction of the incident X-ray wave to a crystal. Any point in the reciprocal lattice that hits the surface of the sphere satisfies Bragg's law as a condition, indicating that diffraction of that X-ray wave can occur. Rotating the crystal will bring different reciprocal lattice points into contact with the Ewald sphere. The Ewald sphere is an important tool that describes what reflections can occur, due to meeting the conditions of Bragg's law, by using the relationship between the incident wave and the reciprocal lattice points.

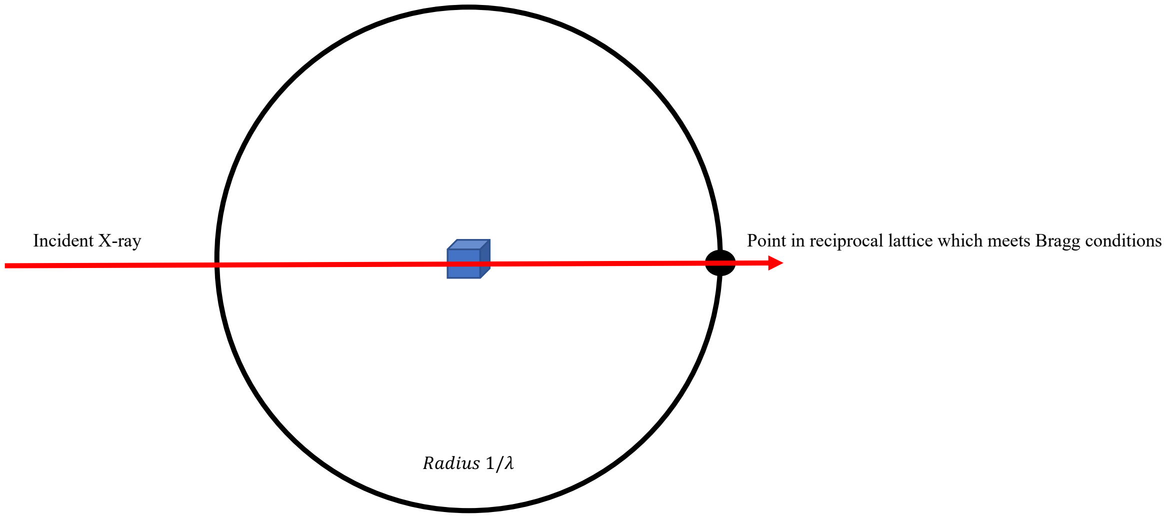


Figure 8.6: Diagram depiction of Ewald sphere. *The crystal is shown in the centre of the Ewald sphere which has a radius 1/λ, with λ representing the wavelength of the incident X-ray. A point in the reciprocal lattice which lies in the direction of the incident X-ray beam (red arrow) and satisfies Bragg's law is shown on the surface of the Ewald sphere as a black circle. Produced using Microsoft PowerPoint (Microsoft Corporation, Albuquerque, New Mexico, USA).*

Unit cell dimensions and symmetry: Since unit cells of a real crystal are symmetric, the reciprocal lattice is also symmetric. Consequently, certain sets of reflections will be equivalent. This is important because, in theory only, one member within a set of equivalent reflections needs to be measured. This knowledge helps reduce the magnitude of data collection required. On the other hand, including multiple members of an equivalent set can improve the accuracy of that reflection reading by taking an average of these equivalent reflection recordings.

When it comes to collecting reflection data, understanding the crystal's internal symmetry is highly important, since internal symmetry describes the symmetry of the contents within a unit cell. The internal symmetry of a unit cell is described by its space group, which is represented by a cryptic symbol. The cryptic symbol representing internal symmetry has a capital letter at the beginning, which represents the lattice type of the unit cell. There are thirteen different lattices which are termed Bravais lattices. There are three major lattice types: there is the primitive lattice (P), the internal lattice (I), and the face-centred lattice (F). The primitive lattice contains one lattice point at each corner of the cell. The internal lattice has an additional lattice point in the centre of the unit cell, and the face-centred lattice has an extra lattice point on all or some of the faces of the unit cell, but no internal lattice points. The capital letter at the front of the space group symbol is followed by symbols which represent the symmetry operations that can be carried out within the unit cell.

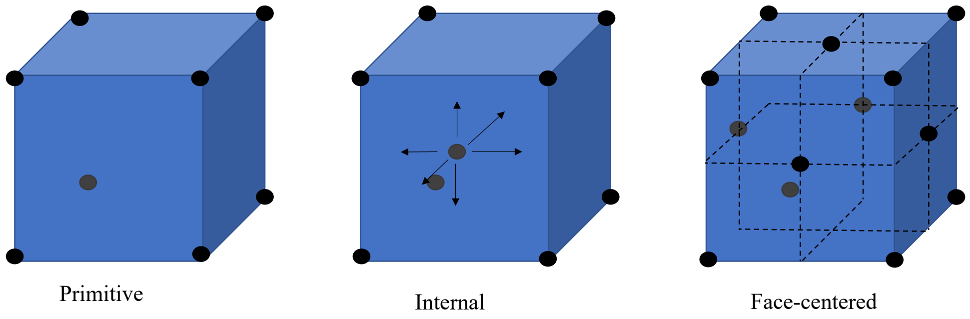


Figure 8.7: Depiction of primitive, internal, and face-centred lattices. *The primitive lattice (displayed on the left) contains one lattice point at each corner of the cell. The internal lattice (displayed in the middle) has a lattice point in the centre of the unit cell, and the face-centred lattice (displayed on the right) is shown with extra lattice points on all of the faces of the unit cell, but does not have any internal lattice points. Produced using Microsoft PowerPoint (Microsoft Corporation, Albuquerque, New Mexico, USA).*

Once unit cell dimensions, and the consequent lattice type, have been decided upon, the next step is to describe the internal symmetry of the unit cell. Within a unit cell, the smallest unit that possesses no symmetry operations is labelled the asymmetric unit and is often a single protein molecule. The arrangement of these proteins within a unit cell can be explained via symmetry operations. The symmetry operations for a protein can be broken down into translation, rotation, and screw axes, with screw axes being rotations and translations combined, meaning that the axes act simultaneously as the axis for rotation and translation of an asymmetric unit. The symbol nm represents an n-fold screw axis accompanied with a m/n translation. For example, a screw axis of 2_1 would see a 180° (360°/2) rotation of the protein molecule, in respect to the previous protein molecule, followed by a translation ½ of the axial length in the direction of the screw axis.

The symmetry pattern of a unit cell can cause distinct absences within the reflection data, which are useful because they aid in determining the space group of a unit cell.

Fourier transformations: Once the diffraction pattern data has been collected from a protein crystal, the next step is to convert this diffraction data into information about the real space inside of the unit cell. Each reflection in the diffraction data is the result of diffraction from all the molecules within the unit cell, causing a complex wave as a result. A simple wave can be represented by a periodic function (Equations 8.5-8.6).

$f(x) = F\cos 2\pi(hx + a)$ (8.5)

$f(x) = F\sin 2\pi(hx + a)$ (8.6)

Equations 8.5-8.6: Periodic functions representing a simple wave. *f(x) represents the height of the wave at any given point (x) along the wave's horizontal axis. x is a given point along the wave's horizontal axis, α represents the phase of the wave, h represents the wave's frequency, and F represents the amplitude of the wave.*

In (Equations 8.5-8.6), f(x) represents the height of the wave at any given point (x) along the wave's horizontal axis. Variable x and constant α are angles which are expressed as fractions of the wavelength;if x=1, then x is one wavelength away (2π radians or 360°) from the origin. α represents the phase of the wave, so if α =1/2 the wave has shifted ½ of a full wavelength (½ of 2π radians or 180°). F is a constant specifying the amplitude of the wave, and the constant h specifies the wavelength's frequency. As indicated by (Equations 8.5-8.6), a simple wave can be described by amplitude (F), frequency (h) and phase (α). Complex waves can be described via the summation of functions for simple waves, with the wavelengths for each simple wave being integral fractions of the complex wave. This summation to describe a complex wave is known as a Fourier series, and each simple wave function within this is called a Fourier term.

Each diffracted X-ray that is recorded is due to the contribution of all the molecules within a unit cell. The sum equation describing a diffracted X-ray is known as a structure factor equation, and the computed sum is known as the structure factor. One way of writing the structure factor sum is by having each term within the sum describing the diffraction by a single atom within the unit cell, leading to a term which contains the same number of terms as number of atoms within the unit cell (Equation 8.7).

$$Fhkl = fA + fB \ldots + fA' + fB' \ldots fF$$

Equation 8.7: Structure factor sum. *fA-fF represent the diffraction by single atoms within a unit cell. Fhkl is a summation of the diffraction contributions from every atom within the unit cell.*

The above function represents how each reflection recorded within the diffraction pattern consists of diffractive contributions from each atom within the unit cell of a protein crystal. Consequently, the structure factor (Fhkl) is a sum of multiple wave equations, and the structure factor equation is known as a Fourier sum. Another way of writing the structure factor equation is by dividing the electron density within a unit cell into small volume elements, termed $p(x, y, z)$. Within this re-imagined equation, each structure factor can be defined as the summation of diffractive contributions from each volume element within a unit cell.

The aim of studying diffraction data produced from X-ray crystallography is to discern the electron density within a unit cell, which can be achieved via the use of what is known as a Fourier transform. The Fourier transform describes the mathematical relationship between an object and its diffraction pattern. The transform allows conversion between the Fourier sum description of the reflections and the Fourier sum description of the electron density.

The Fourier transform demonstrates that for any given function f(x) there exists another function F(h) (Equation 8.8).

$$F(h) = \int_{-\infty}^{+\infty} f(x)e^{2\pi i(hx)}dx$$

Equation 8.8: Fourier transform. *F(h) and f(x) represent functions which are related together via the Fourier transform.*

F(h) is known as the Fourier transform of f(x) and has reciprocal units to f(x). (Equation 5) depicts that to get the Fourier transform of f(x) requires multiplying by $e^{2\pi i(hx)}$ and integrating the combined functions in respect to x. The Fourier transform is fully reversible; f(x) is the Fourier transform of F(h) and F(h) is the Fourier transform of f(x).

The Fourier transform can be applied to periodic functions which consist of any number of dimensions (Equations 8.9-8.10).

$$F(h,k,l) = \int_x \int_y \int_z f(x,y,z)e^{2\pi i(hx+ky+lz)}dx\,dy\,dz \quad (8.9)$$

$$f(x,y,z) = \int_h \int_k \int_l F(h,k,l)e^{-2\pi i(hx+ky+lz)}dh\,dk\,dl \quad (8.10)$$

Equations 8.9-8.10: Three-dimensional periodic functions which are the Fourier transforms of each other.

The function describing electron density is a periodic function consisting of complex three-dimensional wave equations. Three-dimensional waves have three frequencies along the x-, y-, and z- axes which are described by the three terms h, k, l. A general Fourier sum for a three-dimensional wave is given in the form below (Equation 8.11).

$$f(x,y,z) = \sum_h \sum_k \sum_l Fhkl\, e^{2\pi i(hx+ky+lz)}$$

Equation 8.11: General Fourier sum for a three-dimensional wave.

(Equation 8.11) describes how the complex three-dimensional wave (f (x, y, z)) is made up of terms for simple three-dimensional waves, whose frequency are h in the x direction, k in the y direction and l in the z direction. For each triplet set of frequencies, the associated three-dimensional wave has amplitude Fhkl and implicit phase a_{hkl}. The Fourier sum for the complex wave represented by (Equation 8.11) involves adding all of the terms with possible sets of h, k and l integer values.

The equations for both diffracted complex waves and electron density can be described by Fourier sums. These sums can be related via a Fourier transform, allowing for electron density to be calculated from the corresponding diffraction data.

$$Fhkl = \int_V p(x, y, z)e^{2\pi i(hx+ky+lz)}dV$$

Equation 8.12: Fourier sum for a structure factor Fhkl in respect to electron density within a unit cell.

(Equation 8.12) explains how a complex wave structure factor, Fhkl, can be written as the sum of the contributions from each small volume element (p (x, y, z)) within a unit cell. With the volume elements, the smaller these are, the more precise these averaged electron density volume elements become. These volume elements can, in effect, be made infinitesimally small, which means that the function for calculating the structure factor integrates all these electron density volume values. The end integral acts as a sum of all of the contributions of these electron density volume units, with the integral over V, the unit-cell volume, representing the integration over all the values for x, y, and z within the unit cell.

The Fourier transform operation is fully reversible; consequently, this means that the electron density can be discerned by doing a Fourier transformation of (Equation 9), leading to (Equation 8.13).

$$p(x, y, z) = \frac{1}{V}\sum_h \sum_k \sum_l Fhkl\ e^{-2\pi i(hx+ ky+ lz)}$$

Equation 8.13: Fourier sum for electron density in respect to Fhkl.

(Equation 8.13) is a sum rather than an integral like (Equation 8.12) because the structure factor Fhkl values are discrete reflection entities.

(Equation 8.12) suggests that we can calculate a full set of structure factors, and thus the diffraction pattern, from an electron density function. This back-transforming becomes useful within crystallography for a couple of reasons. Firstly, this computation can be used to obtain phases. As will be described later, the crystallographer begins with phase estimates, which are used to create structure factors, and consequently create a preliminary electron density model. From here, the crystallographer tries to improve the model, and from this new model, new structure factors can be estimated, leading to improved phase estimates. This process is iterated to gain improved estimates of the phases, and thus a better model of the electron density and associated molecule within the unit cell. Secondly, back-transforming allows for model assessment. By computing structure factor estimates via this back transformation process, the calculated structure factors will contain reflection intensity estimates (I_{calc}). Convergence of I_{calc} towards observed reflection intensity values (I_{obs}) suggests improvement within the electron density model.

8.3.4 Phase problem and molecular replacement

(Equation 8.13) describes how the electron density of a unit cell can be calculated via a Fourier sum incorporating structure factors. Since structure factors represent diffracted waves, they must specify frequency, amplitude, and phase. When calculating structure factors, amplitude is proportional to the square root of a diffracted X-ray's intensity. The frequency of a structure factor is equal to $1/d_{hkl}$, with d_{hkl} representing the plane spacing for a given set of parallel planes. The phase of a diffracted ray is not discernible from the diffraction pattern, and thus needs to be ascertained via other means, such as molecular replacement, in order to build a model of electron density.

Molecular replacement involves using the phases from the structure factors of a known protein structure, which is then used as a phasing model by placing the model into the unit cell of the target protein crystal structure and using this to calculate initial phase estimates. When the phasing model and target protein are isomorphous (i.e., the crystals belong to the same space group), the phases from the phasing model can be used directly with the diffraction pattern intensities of the target protein to calculate electron density for the target protein. Via back-transforming and following an iterative process, the phases should change from the original phasing model to the phases of the new target protein, leading to a model of the target protein. When the phasing model is not isomorphous with the target protein, the phasing model needs to be orientated and superimposed on to the target protein's unit cell, allowing for initial phase estimates which are most similar to those of the target protein. Once this is done, initial structure factors can then be calculated. From the target protein diffraction data, unit cell dimensions and symmetry can be calculated. Calculating unit cell dimensions and symmetry operations places an initial constraint on where the phasing model should be placed. Patterson maps are used as an efficient tool for calculating the best orientation of the phasing model within the target protein unit cell.

Patterson maps are interatomic vector maps created by squaring all structure factor amplitudes and setting all phases to zero. Patterson maps contain peaks for each atom relating to every other atom in a structure, with a large peak occurring where vectors relating atoms to themselves coalesce. It is possible to create a Patterson map for both the phasing model and the target protein since Patterson maps do not require phase information. If the target protein and phasing model are similar in structure, then their Patterson maps will also be similar. To calculate the optimal orientation of the phasing model, the origin of the Patterson map for the phasing model can be superimposed on to the origin of the target protein's Patterson map, followed by systematic rotation of the phasing model Patterson map.

Evaluation of the fit of the phasing model Patterson map on to the target protein Patterson map can be done via computation of a rotation function. The rotation function entails integrating the product of the target protein and the phasing model Patterson maps. When there are coincident peaks between two Patterson maps the product is large. Consequently, the more coincident peaks there are, the larger the integral product. The maximum integral value possible relates to the optimal rotation of the phasing model (Equation 8.14).

$$R(\phi,\varphi,\chi) = \int_{u,v,w} P^{target}(u,v,w)P^{model}\{(u,v,w)\times\ [\phi,\varphi,\chi]\}\,du\,dv\,dw$$

Equation 8.14: Rotation function for evaluating the orientation of the phasing model Patterson map on to the target protein Patterson map.

(Equation 8.14) indicates the value R at rotation angles (ϕ,φ,χ). (Equation 8.14) is an integral of the phasing model Patterson function $[P^{target}(u,v,w)]$ and the target protein Patterson function $[P^{model}(u,v,w)]$. (Equation 11) will reach a maximal value for R when the two Patterson maps have a lot of coincident peaks, indicating the optimal orientation of the phasing model within the target protein unit cell.

After optimal orientation has been reached, a translational search can be performed to optimise the location of the phasing model within the target protein unit cell allowing subsequent calculation of initial structure factors. One way to evaluate the location of the phasing model is via the correspondence between the observed structure factor amplitudes of the recorded data, and the calculated structure factor amplitudes from the model with the phasing model in a given location. This can be expressed as an R-factor, which is a measure of the overall agreement between these two sets of structure factor amplitudes.

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

Equation 8.15: R factor calculated as the summed difference between the observed structure factor amplitudes and calculated structure factor amplitudes divided by the summed observed structure factor amplitudes.

(Equation 8.15) is used to evaluate the location of the phasing model. It takes each reflection from the diffraction data and sums the difference in amplitude between the observed data set $|F_{obs}|$and the calculated amplitudes from the phasing model when it is placed in a particular location $|F_{calc}|$. This value difference is divided by the summed value of the observed structure factor amplitudes. For perfect agreement between $|F_{obs}|$ and $|F_{calc}|$, the summed differences would equal zero. Consequently, the smaller R is, the greater the agreement between the two datasets, and the better the location of the phasing model within the target protein's unit cell. R-values between 0.3-0.4 for protein crystals often give good estimates of phases for building an electron density model of the target protein.

Optimal orientation and location for a phasing model can be done using a program called PHASER (McCoy *et al.*, 2007). PHASER uses maximum likelihood methods to generate probability functions that help find the correct orientation and translation for the phasing model within the target unit cell.

8.3.5 Data collection

The goal of data collection is to determine the indices and intensities of as many X-ray diffraction reflections as possible. Collection must be as efficient as possible due to the damaging effect of X-rays upon contact with protein crystals.

Diffractometers: A diffractometer is the term used to describe the instrument used to collect X-ray data. This instrument measures the position and intensity of diffracted rays.

The major components which form a diffractometer are: (Rupp, 2009)

- X-ray source

- Optics to select monochromatic X-rays and focus them into a beam
- Goniostat- a device that orientates the crystal within the X-ray beam
- An X-ray detector (such as a charged-coupled device (CCD) detector)
- A cryocooler to keep crystals in a steam of nitrogen gas, helping reduce irradiation damage
- A microscope focused on diffractometer centre aiding alignment of crystal within the X-ray beam
- A sample-changing robot to allow rapid changing of crystals within the X-ray beam
- Data collection software

Synchrotrons as an X-ray source: Synchrotrons are a form of particle accelerator and are a powerful source of X-rays for crystallography. Synchrotrons consist of particle storage rings where electrons (or positrons) are circulated at velocities approaching the speed of light. The electrons are driven to these velocities by energy from radio-frequency transmitters. Velocities are maintained by strong magnets. When an electron is travelling within a synchrotron ring it emits energy in the form of X-rays. These X-rays are taken and used to form intense monochromatic X-ray beams that can be fired at crystals, leading to the production and collection of diffraction data.

Goniostats for rotating crystals: To diffract X-rays from all possible crystal planes, and collect a complete set of reflections, the crystal is rotated within the X-ray beam by a goniostat. The goniostat is allows for the application of the oscillation method, which involves rotating the crystal perpendicular to the X-ray axis.

Cryocrystallography: Cryocrystallography is crystallography done at low temperatures. Cryocrystallography offers better crystal diffraction (low temperatures increase crystal order), reduced radiation damage, and reduced water scattering. Additionally, cryocrystallography offers increased ease of transport, storage, and the re-use of crystals.

Cryoprotectants must be used to prevent ice crystals, which can damage protein crystals. Cryoprotectants include sugars like glucose, as well as the common precipitant polyethylene glycol (PEG).

Usually, cryocrystallography is carried out by picking up crystals in a circular loop, with the crystal remaining suspended within a thin film of solvent. The crystal is flash-frozen within liquid nitrogen, causing the crystal to become suspended within the now frozen solvent. The crystal is now ready for transportation, storage, and data collection.

Diffraction data processing: Processing of diffraction data images is performed by software, such as XDS (Kabsch *et al.*, 2010). Initial processing involves several main steps. The first step is indexing, which consists of determining the indices of each reflection per image frame, resulting in crystal geometry calculations. Integration follows, where the intensities of reflections per frame are recorded. Finally, reflections observed multiple times across frames are merged and scaled to give an average intensity value. Parameters, such as resolution, data completeness, anisotropy, and crystal mosaicity are calculated.

8.3.6 Model building and refinement

Overview: Building a preliminary model of the target protein involves using amplitudes from observed data, and phases computed from methods such as molecular replacement. This initial map may require improvement to allow for discernment of the target protein structure. Improvement of the map can be achieved via an iterative process termed bootstrapping. Bootstrapping involves taking clear features from the electron density map and adding these features to the phasing model. This new phasing map is used in the development of a subsequent new map. This process is repeated until the fit of the model with the calculated electron density can no longer be improved upon.

Refinement of the model is a vital aspect of model building. Refinement involves minimising the differences in structure amplitudes between the original observed data and the structure factor amplitudes predicted by the model. Refinement of protein models can be achieved via several approaches.

Initial map building and improvement: When calculating the initial electron density map for a target protein, the Fourier sum that describes this map is given by (Equation 8.16).

$$p(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l w_{hkl} \, |F_{obs}| \, e^{-2\pi i(hx+\,ky+lz-a'_{calc})}$$

Equation 8.16: Fourier sum for initial electron density map.

In (Equation 8.16) each structure factor hkl term has $|F_{obs}|$ amplitude which equates to the square root of the observed intensity ($(I_{hkl})^{1/2}$) from the collected data. The corresponding phase (a'_{calc}) comes from phasing methods such as molecular replacement. In addition, the term is weighted by a factor between zero and one, which is represented by w_{hkl}. This factor acts to weight phases in respect to the precision of said phases. The closer this weighting factor is to zero, the less reliable that phase measurement is, and multiplication by anything less than one reduces that given term's contribution to (Equation 8.16), reflecting the lack of precision in that term's phase.

Once the initial electron density model has been built, it can act as a model of the target protein. This electron density model function is modified to increase how realistic it is by including prior knowledge of the known properties of protein and water when they exist together in a crystal. The first step in modification often involves smoothing the map, which involves evaluating p (x, y, z) values along a grid of equally spaced points within the modelled unit cell. If p is negative, it is reassigned the value of 0. If p is positive, it is assigned the mean value of p for an area a given distance from the grid point in question. After this map smoothing, the amplitude of the map is increased so as to get the high density to low density ratio that agrees with the ratio of protein to solvent in a crystal, which is normally assumed to be 50:50. This adapted electron density model, and its associated function, can be used to compute new structure factors for the next round of map building. This map indicates what both the amplitudes and phases would be for the target structure if this model were correct. These newly calculated phases are collected and used with the original collected data intensities to carry out the next round of electron density map building.

Conversion to molecular model: After a given number of iterations of the map building process, there will be a point where a protein chain can be threaded through the electron density. Where the model is accurate and to a higher resolution, distinct individual solvent molecules may become visible within the electron density and can be added to the model. This process of adding a protein chain, and possibly other molecules, is termed map fitting, and involves the use of interactive computer graphics that display the protein model fitted to the model electron density. Map fitting allows for manipulation of the model to improve the model's fit to the electron density. As the protein model goes through manipulation, the computer program stores the protein chain atoms' positions as three-dimensional coordinates. This map of coordinates is used to share a protein structure with the scientific community. Programmes such as COOT (Emsley *et al.*, 2010) can be used to visualise calculated electron density maps and molecular models, allowing for exploration of the electron density and improvement of model fit.

Model bias and difference maps: The issue with switching over to a molecular model from an electron density model is the increased risk of introducing model biases into the electron density predictions for the target protein. To try and compensate for model biases and their influence on phase calculations, electron density map functions involving the subtraction of calculated structure factor amplitudes ($|F_{calc}|$) from observed data amplitudes from collected data ($|F_{obs}|$) are calculated. These maps show less influence from the model (Equation 8.17).

$$p(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l w_{hkl}\,(|F_{obs}| - |F_{calc}|)\,e^{-2\pi i(hx+ ky+lz-a'_{calc})}$$

Equation 8.17: F_{obs}-F_{calc} synthesis. *Electron density Fourier synthesis which aims to reduce the bias introduced by the model when computing the electron density map for the target protein.*

The electron density map computed from (Equation 8.17) is called a F_{obs}-F_{calc} map. This map will contain both positive and negative density. Regions of positive density indicate areas of electron density unaccounted for by the model, whilst negative density indicates areas of density suggested by the model that are not present within the unit cell. This map emphasises errors within the working model. This form of map allows for the detection of subtle errors in protein structure placement within the target protein model.

Another difference map, which allows for bias detection but still maintains the appearance of a full electron density contour map, is given by (Equation 8.18). This function forms the difference map known as the 2 F_{obs}-F_{calc} map.

$$p(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l w_{hkl}\,(2|F_{obs}| - |F_{calc}|)\,e^{-2\pi i(hx+ ky+lz-a'_{calc})}$$

Equation 8.18: 2 F_{obs}-F_{calc} synthesis. *Another electron density Fourier synthesis which* *reduces the bias introduced by the model when computing the electron density map for the target protein. This map reads more like a traditional electron density map, but still allows for the detection of bias introduced by the model when calculating electron density.*

Refinement overview: Refinement methods used during model production fall into two main categories: real space refinement and reciprocal space refinement. Real space refinement involves iteratively creating maps and building the resultant models. Reciprocal space refinement involves trying to improve the agreement of the model data with the original collected data. Reciprocal space refinement involves comparing calculated and original structure factor amplitudes ($|F_{calc}|$ versus $|F_{obs}|$). During model refinement, there is often a back and forth between these two forms of refinement.

Least squares: With least-squares refinement methods, the aim is to find a function (f(x)) that will fit one set of observations for one variable to another variable, in a way that minimises the summed squared differences between the observed values and calculated values (Equation 8.19).

$$D = \sum_i w_i[y^i - f(x_i)]^2$$

Equation 8.19: Summed squared differences using least squares. *Equation describes the calculation of the summed squared differences (D) between observed values (y^i)and calculated values $f(x_i)$. w_i acts as a weighting factor allowing for increased weighting for more reliable data.*

When using least squares in crystallography, the aim is to discern a set of atom positions within the molecular model which will minimise the squared differences between the calculated structure factor amplitudes ($|F_{calc}|$) and the originally observed structure factor amplitudes ($|F_{obs}|$) (Equation 8.20).

$$\phi = \sum_{hkl} w_{hkl}(|F_{obs}| - |F_{calc}|)^2_{hkl}$$

Equation 8.20: Least squares in crystallography. *This equation seeks to minimise the sum of squared differences between $|F_{obs}|$ and $|F_{calc}|$.*

Bayesian: A major issue with the least squares method of refinement is that it assumes that the models will improve via a linear path which will eventually end with a completely optimal model. However, there is uncertainty in these model choices, and these models can go on to bias future model building, causing the refinement to be sent down into a local minima rather than reaching the actual global minima representing the best fitted model. One way to avoid bias is to utilise Bayesian thinking. This involves expressing our understanding of current knowledge about a set of possible hypotheses (in crystallography this is a set of models for the target protein) via the use of prior probabilities. As data is collected, new knowledge is garnered, and, via an iterative process, those models with higher prior probability are pushed forward.

Molecular dynamics and energy: Prior knowledge of energy can be incorporated into the refinement process. This can involve applying least squares restraints on the overall model energy in an attempt to find the structure with the lowest energy within the current area of the model being manipulated.

Another method deployed in refinement involves incorporating molecular dynamics. The aim is to include knowledge of molecule movement based upon Newton's laws of motion within force fields, with the force fields representing covalent and non-covalent bonding. An energy term is fitted, and this term treats the model as decreasing in energy as the fit to the electron density is iteratively improved.

Constraints and restraints: As mentioned, refining a model can lead the model to land in local minima, meaning that the globally best fit model is at risk of never being reached. During the least squares refinement process, it is important to find a starting model that lands near the goal global minimum model to avoid the model landing in some other local minima. As a rough estimate, the greatest distance that a model can be from the global minimum model is represented by a value known as the radius of convergence. The radius of convergence is calculated as $d_{min}/4$, with d_{min} representing the lattice plane spacing that gave rise to the highest resolution reflection in the data set. Increasing the radius of convergence increases our chances of eventually finding the model which lands within the global minima, and thus represents the best fit between model and observed data. Constraints and restraints can help in increasing this radius of convergence. Constraints are fixed values for a given parameter within the molecular model, for example, constraining occupancies of atoms to one. Constraints consequently act to fix the relationships between parameters in a given model. Restraints are subsidiary conditions which are placed upon constraints within the model. These restraints can include things such as restraining bond lengths and angles within a specified range. Restraints aim to penalise any changes that make the model less realistic based upon our prior knowledge of protein folding.

During the refinement process, constraints and restraints are gradually lifted from the working model. Agreement between $|F_{obs}|$ and $|F_{calc}|$ is given priority in the refinement process, since the risk of falling into a local minima has been reduced.

Refinement parameters: As mentioned above, refinement of a model is monitored via the comparison of $|F_{obs}|$ and $|F_{calc}|$. As iterations of model building continue, there should be a convergence in values between $|F_{obs}|$ and $|F_{calc}|$ as the model moves towards being a better fit with the data. A common measure for following the convergence of $|F_{obs}|$ and $|F_{calc}|$ is the R-factor (Equation 8.15).

Another R measure used to capture model quality during the refinement process is known as the free R-factor (R_{free}). R_{free} is computed with a small randomly selected subset of intensities that are set aside before refinement. This measurement allows for the assessment of agreement, via cross-validation, between the calculated data and the observed data, with many crystallographers believing that R_{free} gives a less biased indication overall of how well the model fits the data. During the re-building and refinement process, both R_{free} and the other R-factor (usually termed R_{work}) should decrease, but if R_{work} becomes much lower than R_{free}, this is an indicator of a bad quality model, meaning the model does not fit the originally observed data. R-factors around the 0.15-0.3 mark are usually considered good, with R_{free} usually being slightly higher than R_{work}.

In addition to the R-factors described above, the convergence of the model with the actual structure of the target protein can be monitored via parameters which allow for the crystallographer to discern whether the model is chemically, stereochemically and conformationally realistic. However, it must be noted that in many refinement processes bond angles and lengths are often restrained during refinement rounds, meaning these parameters are less informative than their structural counterparts, which refine without the presence of restraints.

Refinement is carried out by programs like REFMAC (Murshudov *et al.*, 2011). These programs are statistical tools. REFMAC employs a maximum likelihood function to optimise the fit between the model and the data. REFMAC utilises factors such as bond distances, torsion angles, and temperature factors (B-factors) to assess the fit of the model to the observed data.

8.3.7 Assessing the model and model structure deposition

As mentioned, R-factors are used to track the refinement process, along with the allowance of constraints and restraints to be gradually lifted allowing the model to finish refinement freely. During the last few rounds, when constraints and restraints are lifted, values like the root mean square (rms) deviation in bond lengths, bond angles and conformational angles should settle at reasonable values. Usually, rms deviation values should not exceed 0.02 Å for bond lengths and should not exceed 4° for bond angles. In addition, peptide bonds are implied to be planar in nature, thanks to spectroscopic evidence, and knowing that peptide bonds should be planar allows for further scrutiny of the final model. ϕ is the conformational angle along the N-Cα bond of a peptide, and ψ is the conformational angle along the Cα-C bond of a peptide. ϕ and ψ are greatly restricted in their values due to steric repulsion, meaning they can only exist in certain allowed value pairs. ϕ and ψ values for each peptide in a model can be modelled using a Ramachandran diagram, allowing the crystallographer to see if any biologically unrealistic sections exist within the model in terms of conformational angles.

Another set of values which are valuable for assessing model fit are B-factors. B-factors calculated for each atom reflect both thermal motion and disorder. It is important to capture the movement of atoms within a crystal, e.g., by varying amino acid conformations and flexible protein regions, because these can play a part in lowering model resolution. A calculated B-factor acts as a measure of vibrational motion (Å^2). If B-factors are higher on the inside of a protein structure versus the outside, this suggests an issue with the model, since proteins tend to be more ordered and stable in their cores than their surfaces.

There are statistical values which can be helpful aids during model assessment. One value which is used to assess the significance of data used in a model is a value which describes the average signal-to-noise ratio of merged intensity values ($< \frac{I}{\sigma I} >$). I represents the merged intensity value, and σI represents the standard deviation in merged intensity data. $< \frac{I}{\sigma I} >$ is good to use when assessing resolution cut-off for a model, but it must be noted that this value can incur uncertainties due to uncertainties in the estimation of σI (Evans and Murshudov, 2013). Another value is the CC ½ value, which is the correlation coefficient between two random half data sets associated with the target protein. This value acts as an assessor of model resolution and has a range between one and zero. As values approach one, correlation approaches 100 %. At the lower relative resolution values for a given data set, correlation is generally close to one, but this falls sharply in value at higher resolutions for a data set, where intensities become weaker (Evans and Murshudov, 2013).

Once a model has been assessed and deemed ready for release to the scientific community, the model coordinates and associated metadata are deposited into the Protein Data Bank (PDB) (Berman *et al*., 2000), an open access repository for three-dimensional structural models of protein molecules.

Publication From Research

Sait LG*, Sula A*, Ghovanloo MR*, Hollingworth D, Ruben PC and Wallace BA (2020) **Cannabidiol interactions with voltage-gated sodium channels** *Elife* **9**: e58593.

Bibliography

Abdul-Gader A, Miles AJ and Wallace BA (2011) **A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy** *Bioinformatics* **27(12)**:1630–1636.

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

Ahrens J, Demir R, Leuwer M, De La Roche J, Krampfl K, Foadi N, Karst M and Haeseler G (2009) **The nonpsychotropic cannabinoid cannabidiol modulates and directly activates alpha-1 and alpha-1-beta glycine receptor function** *Pharmacology* **83**:217-222.

Aitipamula S and Vangala VR (2017) **X-Ray Crystallography and its Role in Understanding the Physicochemical Properties of Pharmaceutical Cocrystals** *Journal of the Indian Institute of Science* **97**:227–243.

Aizpurua-Olaizola O, Soydaner U, Öztürk E, Schibano D, Simsir Y, Navarro P, Etxebarria N and Usobiaga A (2016) **Evolution of the Cannabinoid and Terpene Content during the Growth of Cannabis sativa Plants from Different Chemotypes** *Journal of Natural Products* **79(2)**:324–331.

Alamry KA, Srivastava S, Shahbaaz M, Khan P, Gupta P, Syed SB, Azum N, Asiri AM, Islam A, Ahmad F and Hassan MI (2020) **Unravelling the unfolding pathway of human Fas-activated serine/threonine kinase induced by urea** *Journal of Biomolecular Structure and Dynamics* 1–10.

Antzelevitch C, Nesterenko V, Shryock JC, Rajamani S, Song Y and Belardinelli L (2014) **The role of late I Na in development of cardiac arrhythmias** *Handbook of experimental Pharmacology* **221**: 137–168.

Bagnéris C, Decaen PG, Hall BA, Naylor C E, Clapham DE, Kay CWM and Wallace BA (2013) **Role of the C-terminal domain in the structure and function of tetrameric sodium channels** *Nature Communications* **4**:2465.

Bagnéris C, DeCaen PG, Naylor CE, Pryde DC, Nobeli I, Clapham DE and Wallace BA (2014) **Prokaryotic NavMs channel as a structural and functional model for eukaryotic sodium channel antagonism** *Proceedings of the National Academy of Sciences of the United States of America* **111**:8428–8433.

Bagnéris C, Naylor CE, McCusker EC and Wallace BA (2015) **Structural model of the open–closed–inactivated cycle of prokaryotic voltage-gated sodium channels** *The Journal of General Physiology* **145 (1)**:5–16.

Baker MD, Chandra SY, Ding Y, Waxman SG and Wood JN (2003) **GTP-induced tetrodotoxin-resistant Na+ current regulates excitability in mouse and rat small diameter sensory neurones** *Journal of Physiology* **548:** 373-382.

Bateman A (2019) **UniProt: A worldwide hub of protein knowledge** *Nucleic Acids Research*, **47(D1)**:D506–D515.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN and Bourne PE (2000) **The Protein Data Bank** *Nucleic Acids Research* **28(1)**:235-242.

Biner O, Schick T, Müller Y and von Ballmoos C (2016) **Delivery of membrane proteins into small and giant unilamellar vesicles by charge-mediated fusion** *Federation of European Biochemical Societies Letters* **590(14)**:2051-2062.

Bisogno T, Hanuš L, De Petrocellis L, Tchilibon S, Ponde DE, Brandi I, Moriello AS, Davis JB, Mechoulam R and Di Marzo V (2001) **Molecular targets for cannabidiol and its synthetic analogues: Effect on vanilloid VR1 receptors and on the cellular uptake and enzymatic hydrolysis of anandamide** *British Journal of Pharmacology* **134(4)**:845-852.

Blake DR, Robson P, Ho M, Jubb RW and McCabe CS (2006) **Preliminary assessment of the efficacy, tolerability and safety of a cannabis-based medicine (Sativex) in the treatment of pain caused by rheumatoid arthritis** *Rheumatology* **45**:50-52.

Boiteux C, Vorobyov I, French RJ, French C, Yarov-YarovoyV and Allen TW (2014) **Local anesthetic and antiepileptic drug access and binding to a bacterial**

voltage-gated sodium channel *Proceedings of the National Academy of Sciences of the United States of America* **111(36)**: 13057-13062.

Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart O, Vonrhein C and Womack TO (2019) **BUSTER, version 2.10.3** Global Phasing Ltd, Cambridge, United Kingdom.

Bürck J, Wadhwani P, Fanghänel S and Ulrich AS (2016) **Oriented Circular Dichroism: A Method to Characterize Membrane-Active Peptides in Oriented Lipid Bilayers** *Accounts of Chemical Research* **49(2)**:184-192.

Campbell FA, Tramèr MR, Carroll D, Reynolds DJM, Moore RA and Mcquay HJ (2001) **Are cannabinoids an effective and safe treatment option in the management of pain? A qualitative systematic review** *British Medical Journal* **323(7303)**:13-16.

Carlini EA (2004) **The good and the bad effects of (-) trans-delta-9-tetrahydrocannabinol (Δ9-THC) on humans** *Toxicon* **44(4)**:461-467.

Carter M and Shieh J (2015) **Guide to Research Techniques in Neuroscience** *Academic Press.*

Caterina MJ, Rosen TA, Tominaga M, Brake AJ and Julius D (1999) **A capsaicin-receptor homologue with a high threshold for noxious heat** *Nature* **398(6726)**:436-441.

Chae PS, Rasmussen SGF, Rana RR, Gotfryd K, Kruse AC, Manglik A, Cho KH, Nurva S, Gether U, Guan L, Loland CJ, Byrne B, Kobilka BK and Gellman SH (2012) **A new class of amphiphiles bearing rigid hydrophobic groups for solubilization and stabilisation of membrane proteins** *Chemistry - A European Journal* **18(31)**:9485-9490.

Charalambous K, O'Reilly AO, Bullough PA and Wallace BA (2009) **Thermal and chemical unfolding and refolding of a eukaryotic sodium channel** *Biochimica et Biophysica Acta – Biomembranes* **1788(6)**:1279-1286.

Charalambous K and Wallace BA (2011) **NaChBac: The long lost sodium channel ancestor** *Biochemistry* **50(32)**:6742–6752.

Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS and Richardson DC (2010) **MolProbity: all-atom structure validation for macromolecular crystallography** *Acta Crystallographica* **D66**:12–21.

Chiu J, March PE, Lee R and Tillett D (2004) **Site-directed, Ligase-Independent Mutagenesis (SLIM): a single-tube methodology approaching 100 % efficiency in 4 h** *Nucleic Acids Research* **32**:e174.

Chowdhury B and Garai G (2017) **A review on multiple sequence alignment from the perspective of genetic algorithm** *Genomics* **109(5-6)**:419-431.

Costa B, TrovatoAE, Comelli F, Giagnoni G and Colleoni M (2007) **The non-psychoactive cannabis constituent cannabidiol is an orally effective therapeutic agent in rat chronic inflammatory and neuropathic pain** *European Journal of Pharmacology* **556(1-3)**:75-83.

Cox JJ, Reimann F, Nicholas AK, Thornton G, Roberts E, Springell K, Karbani G, Jafri H, Mannan J, Raashid Y, Al-Gazali L, Hamamy H, Valente EM, Gorman S, Williams R, McHale DP, Wood JN, Gribble FM and Woods CG (2006) **An SCN9A channelopathy causes congenital inability to experience pain** *Nature* **444**:894-898.

Cross JH, Devinsky O, Marsh E, Miller I, Nabbout R, Scheffer IE, Thiele EA, Laux L and Wright S (2017) **Cannabidiol (CBD) reduces convulsive seizure frequency in Dravet syndrome: results of a multi-centre, randomised, controlled trial** *Epilepsia* **58**:S12.

Cuñetti L, Manzo L, Peyraube R, Arnaiz J, Curi L and Orihuela S (2018) **Chronic Pain Treatment With Cannabidiol in Kidney Transplant Patients in Uruguay** *Transplantation Proceedings* **50(2)**:461-464.

de Kovel CGF, Meisler MH, Brilstra EH, van Berkestijn FMC, Slot R, van Lieshout S, Nijman IJ, O’Brien JE, Hammer MF, Estacion M, Waxman SG, Dib-Hajj SD and

Koeleman BPC (2014) **Characterization of a de novo SCN8A mutation in a patient with epileptic encephalopathy** *Epilepsy Research* **108(9)**:1511–1518.

de Lera Ruiz M and Kraus RL (2015) **Voltage-gated sodium channels: structure, function, pharmacology, and clinical indications** *Journal of Medicinal Chemistry* ***58*****(18)**:7093–7118.

Del Villar-Guerra R, Trent JO and Chaires JB (2018) **G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy** *Angewandte Chemie (International Edition in English)* **57(24)**:7171–7175.

Desaphy JF, Dipalma A, Costanza T, Bruno C, Lentini G, Franchini C, George A and Conte Camerino D (2010) **Molecular determinants of state-dependent block of voltage-gated sodium channels by pilsicainide** *British Journal of Pharmacology* **160(6)**:1521–1533.

Devane WA, Dysarz FA, Johnson MR, Melvin LS and Howlett AC (1988) **Determination and characterization of a cannabinoid receptor in rat brain** *Molecular Pharmacology* **34(5)**:605-613.

Devinsky O, Cross JH, Laux L, Marsh E, Miller I, Nabbout R, Scheffer IE, Thiele EA and Wright S (2017) **Trial of Cannabidiol for Drug-Resistant Seizures in the Dravet Syndrome** *New England Journal of Medicine* **376(21)**:2011–2020.

Devinsky O, Marsh E, Friedman D, Thiele E, Laux L, Sullivan J, Miller I, Flamini R, Wilfong A, Filloux F, Wong M, Tilton N, Bruno P, Bluvstein J, Hedlund J, Kamens R, Maclean J, Nangia S, Singhal NS, Wilson CA, Patel A and Cilio MR (2016) **Cannabidiol in patients with treatment-resistant epilepsy: an open-label interventional trial** *Lancet Neurology* **15(3)**:270-8.

Devinsky O, Nabbout R, Miller I, Laux L, Żołnowska M, Wright S and Roberts C (2018) **IP 110. Maintenance of Long-Term Safety and Efficacy of Cannabidiol Treatment in Dravet's Syndrome: Results of an Open-Label Extension Trial (GWPCARE5)** *Abstracts of the 44th Annual Meeting of the Society for Neuropediatrics* 49.

Devinsky O, Patel AD, Cross JH, Villanueva V, Wirrell EC, Privitera M, Greenwood SM, Roberts C, Checketts D, VanLandingham KE and Zuberi SM (2018) **Effect of**

cannabidiol on drop seizures in the lennox–gastaut syndrome *New England Journal of Medicine* **378(20)**:1888–1897.

Di Lella S, Martí MA, Croci DO, Guardia CMA, Díaz-Ricci JC, Rabinovich GA, Caramelo, JJ and Estrin DA (2010) **Linking the structure and thermal stability of β-galactoside-binding protein galectin-1 to ligand binding and dimerization equilibria** *Biochemistry* **49(35)**:7652-7658.

Dib-Hajj SD, Geha P and Waxman SG (2017) **Sodium channels in pain disorders: Pathophysiology and prospects for treatment** *Pain* **158(Supplement 1)**:S97-S107.

Dioun C (2017) **The Architecture of Illegal Markets: Towards an Economic Sociology of Illegality in the Economy** *Oxford University Press.*

Dong C, Beis K, Nesper J, Brunkan-LaMontagne AL, Clarke BR, Whitfield C and Naismith JH (2006) **Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein** *Nature* **444(7116)**:226-229.

Dörr JM, Koorengevel MC, Schäfer M, Prokofyev AV, Scheidelaar S, Van Der Cruijsen EA W, Dafforn TR, Baldus M and Killian JA (2014) **Detergent-free isolation, characterization, and functional reconstitution of a tetrameric K+ channel: The power of native nanodiscs** *Proceedings of the National Academy of Sciences of the United States of America* **111(52)**:18607-18612.

Eisenstein M (2015) **Showdown at the cannabis corral** *Nature* **525(7570)**:S15.

Elsohly MA, Gul W, Wanas AS and Radwan MM (2014) **Synthetic cannabinoids: analysis and metabolites** *Life Sciences* **97(1)**:78–90.

Emsley P, Lohkamp B, Scott WG and Cowtan K (2010) **Features and development of coot** *Acta Crystallographica* **D66**:486–501.

Escayg A, MacDonald BT, Meisler MH, Baulac S, Huberfeld G, An-Gourfinkel I, Brice A, LeGuern E, Moulard B, Chaigne D, Buresi C and Malafosse A (2000) **Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2** *Nature Genetics* **24(4)**:343-345.

Esposito G, Scuderi C, Savani C, Steardo L, De Filippis D, Cottone P, Iuvone T, Cuomo V and Steardo L (2007) **Cannabidiol in vivo blunts β-amyloid induced neuroinflammation by suppressing IL-1β and iNOS expression** *British Journal of Pharmacology* **151(8)**:1272-1279.

Estacion M, Gasser A, Dib-Hajj SD and Waxman SG (2010) **A sodium channel mutation linked to epilepsy increases ramp and persistent current of Nav1.3 and induces hyperexcitability in hippocampal neurons** *Experimental Neurology* **224(2)**:362-368.

Evans PR and Murshudov GN (2013) **How good are my data and what is the resolution?** *Acta Crystallographica* **D69**:1204–1214.

Faber CG, Lauria G, Merkies ISJ, Cheng X, Han C, Ahn HS, Persson AK, Hoeijmakers JGJ, Gerrits MM, Pierro T, Lombardi R, Kapetis D, Dib-Haj SD and Waxman SG (2012) **Gain-of-function Nav1.8 mutations in painful neuropathy** *Proceedings of the National Academy of Sciences of the United States of America* **109(47)**:19444-19449.

Fan H, Vitharana SN, Chen T, O'Keefe D and Middaugh CR (2007**) Effects of pH and polyanions on the thermal stability of fibroblast growth factor 20** *Molecular Pharmaceutics* **4(2)**:232-240.

Forst D, Welte W, Wacker T and Diederichs K (1998) **Structure of the sucrose-specific porin ScrY from Salmonella typhimurium and its complex with sucrose** *Nature Structural Biology* **5**:37-46.

Fouda MA, Ghovanloo MR and Ruben PC (2020) **Cannabidiol protects against high glucose-induced oxidative stress and cytotoxicity in cardiac voltage-gated sodium channels** *British Journal of Pharmacology* **177(13)**:1932-2946.

Frederick J, Buck ME, Matson DJ and Cortright DN (2007) **Increased TRPA1, TRPM8, and TRPV2 expression in dorsal root ganglia by nerve injury** *Biochemical and Biophysical Research Communications* **358(4)**:1059-1064.

Freeman AM, Petrilli K, Lees R, Hindocha C, Mokrysz C, Curran HV, Saunders R and Freeman TP (2019) **How does cannabidiol (CBD) influence the acute effects of**

delta-9-tetrahydrocannabinol (THC) in humans? A systematic review *Neuroscience and Biobehavioral Reviews* **107**:696-712.

Froger A and Hall JE (2007) **Transformation of Plasmid DNA into E. Coli using the heat shock method** *Journal of Visualised Experiments* **6**:253.

Fukuma G, Oguni H, Shirasaka Y, Watanabe K, Miyajima T, Yasumoto S, Ohfu M, Inoue T, Watanachai A, Kira R, Matsuo M, Muranaka H, Sofue F, Zhang B, Kaneko S, Mitsudome A and Hirose S (2004) **Mutations of Neuronal Voltage-gated Na+ Channel α1 Subunit Gene SCN1A in Core Severe Myoclonic Epilepsy in Infancy (SMEI) and in Borderline SMEI (SMEB)** *Epilepsia* **45(2)**:140-148.

Gamal El-Din TM, Lenaeus MJ, Zheng N and Catterall WA (2018) **Fenestrations control resting-state block of a voltage-gated sodium channel** *Proceedings of the National Academy of Sciences of the United States of America* **115**:13111–13116.

Gaoni Y and Mechoulam R (1964) **Isolation, Structure, and Partial Synthesis of an Active Constituent of Hashish** *Journal of the American Chemical Society* **86(8)**:1646-1647.

Garrido-Cardenas JA, Garcia-Maroto F, Alvarez-Bermejo JA and Manzano-Agugliaro F (2017) **DNA sequencing sensors: An overview** *Sensors (Switzerland)* **17(3)**:588.

Gaspar D, Lúcio M, Rocha S, Lima JLFC and Reis S (2011) **Changes in PLA 2 activity after interacting with anti-inflammatory drugs and model membranes: Evidence for the involvement of tryptophan residues** *Chemistry and Physics of Lipids* **164(4)**:292-299.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD and Bairoch A (2003) **ExPASy: The proteomics server for in-depth protein knowledge and analysis** *Nucleic Acids Research* **31(13)**: 3784–3788.

Geertsma ER, Nik Mahmood NAB, Schuurman-Wolters GK and Poolman B (2008) **Membrane reconstitution of ABC transporters and assays of translocator function** *Nature Protocols* **3(2)**:256-266.

Ghovanloo MR, Shuart NG, Mezeyova J, Dean RA, Ruben PC and Goodchild SJ (2018) **Inhibitory effects of cannabidiol on voltage-dependent sodium currents** *Journal of Biological Chemistry* **293**:16546–16558.

Ghovanloo MR, Choudhury K, Bandaru TS, Fouda MA, Rayani K, Rusinova R, PhaterpekarT, Nelkenbrecher K, Watkins AR, Poburko D, Thewalt J, Andersen OS, Delemotte L, Goodchild SJ and Ruben PC (2021) **Cannabidiol inhibits the skeletal muscle Nav1.4 by blocking its pore and by altering membrane elasticity** *Journal of General Physiology* **153(5)**: e202012701.

Gollery M (2005) **Bioinformatics: Sequence and Genome Analysis, 2nd Edition** 692 NY: *Cold Spring Harbor Laboratory Press*.

Gorrec F (2009) **The MORPHEUS protein crystallization screen** *Journal of Applied Crystallography* **42**:1035-1042.

Green B, Kavanagh D and Young R (2003) **Being stoned: A review of self-reported cannabis effects** *Drug and Alcohol Review* **22(4)**:453-460.

Greenfield NJ (2007) **Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions** *Nature Protocols* **1**:2527-2535.

Gubellini F, Verdon G, Karpowich NK, Luff JD, Boël G, Gauthier N, Handelman SK, Ades SE and Hunt JF (2011) **Physiological response to membrane protein overexpression in E. coli** *Molecular and Cellular Proteomics* **10(10)**: M111.007930.

Güler R, Svedmark SF, Abouzayed A, Orlova A and Löfblom J (2020) **Increasing thermal stability and improving biodistribution of VEGFR2-binding affibody molecules by a combination of in silico and directed evolution approaches** *Scientific Reports* **10**:18148.

Halford J, Marsh E, Mazurkiewicz-Beldzinska M, Gunning B, Checketts D, Roberts C and Thiele E (2018) **Long-term Safety and Efficacy of Cannabidiol (CBD) in Patients with Lennox-Gastaut Syndrome (LGS): Results from Open-label Extension Trial (GWPCARE5)(P1. 264)** *Neurology* **15**: 264.

Hanck DA, Nikitina E, McNulty MM, Fozzard HA, Lipkind GM and Sheets MF (2009) **Using lidocaine and benzocaine to link sodium channel molecular conformations to state-dependent antiarrhythmic drug affinity** *Circulation Research* **105(5)**:492-499.

Hayakawa K, Mishima K, Hazekawa M, Sano K, Irie K, Orito K, Egawa T, Kitamura Y, Uchida N, Nishimura R, Egashira N, Iwasaki K and Fujiwara M (2008) **Cannabidiol potentiates pharmacological effects of Δ9-tetrahydrocannabinol via CB1 receptor-dependent mechanism** *Brain Research* **1188**:157-164.

Heron SE, Crossland KM, Andermann E, Phillips HA, Hall AJ, Bleasel A, Shevell M, Mercho S, Seni MH, Guiot, MC, Mulley JC, Berkovic SF and Scheffer IE (2002) **Sodium-channel defects in benign familial neonatal-infantile seizures** *Lancet* **360(9336)**:851-852.

Hill AJ, Jones NA, Smith I, Hill CL, Williams CM, Stephens GJ and Whalley BJ (2014) **Voltage-gated sodium (NaV) channel blockade by plant cannabinoids does not confer anticonvulsant effects per se** *Neuroscience Letters* **566**:269–274.

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

Hosseinzadeh S, Eslami Moghadam M, Sheshmani S and Shahvelayati AS (2021) **New anticancer Pd and Pt complexes of tertamyl dithiocarbamate and DACH ligands against HT29 and Panc1 cell lines** *Journal of Biomolecular Structure and Dynamics* 1-11.

Huck V, Chen PC, Xu ER, Tischer A, Klemm U, Aponte-Santamaría C, Mess C, Obser T, Kutzki F, König G, Denis C, Gräter F, Wilmanns M, Auton M, Schneider SW, Schneppenheim R, Hennig J and Brehm MA (2020) **Gain-of-function variant p.Pro2555Arg of von Willebrand factor increases aggregate size through altering stem dynamics** *Thrombosis and Haemostasis* Online ahead of print.

Huizenga MN, Fureman BE, Soltesz I and Stella N (2018) **Proceedings of the Epilepsy Foundation’s 2017 Cannabinoids in Epilepsy Therapy Workshop** *Epilepsy and Behavior* **85**:237-242.

Hunter D, Oldfield G, Tich N, Messenheimer J and Sebree T (2018) **Synthetic transdermal cannabidiol for the treatment of knee pain due to osteoarthritis** *Osteoarthritis and Cartilage* **26 (Supplement 1)**:S26.

Huus K, Havelund S, Olsen HB, Van De Weert M and Frokjaer S (2006) **Chemical and thermal stability of insulin: Effects of zinc and ligand binding to the insulin zinc-hexamer** *Pharmaceutical Research* **23(11)**:2611-2620.

Ibeas Bih C, Chen T, Nunn AVW, Bazelot M, Dallas M and Whalley BJ (2015) **Molecular Targets of Cannabidiol in Neurological Disorders** *Neurotherapeutics* **12(4)**:699-730.

Ichikawa H and Sugimoto T (2000) **Vanilloid receptor 1-like receptor-immunoreactive primary sensory neurons in the rat trigeminal nervous system** *Neuroscience* **101(3)**:719-725.

Ireland SM, Sula A and Wallace BA (2018) **Thermal melt circular dichroism spectroscopic studies for identifying stabilising amphipathic molecules for the voltage-gated sodium channel NavMs** *Biopolymers* **109(8)**:e23067.

James JS (1998) **Medical marijuana: largest provider closed, some alternatives available** *AIDS Treatment News* **296**:1-7.

Jia Y, Zhao Y, Kusakizako T, Wang Y, Pan C, Zhang Y, Nureki O, Hattori M and Yan Z (2020) **TMC1 and TMC2 Proteins Are Pore-Forming Subunits of Mechanosensitive Ion Channels** *Neuron* **105(2)**:310-321.

Jiang R, Yamaori S, Okamoto Y, Yamamoto I and Watanabe K (2013) **Cannabidiol is a potent inhibitor of the catalytic activity of cytochrome P450 2C19** *Drug Metabolism and Pharmacokinetics* **28(4)**:332-338.

Jin H, Li P, Jin Y and Sheng L (2021) **Effect of sodium tripolyphosphate on the interaction and aggregation behavior of ovalbumin-lysozyme complex** *Food Chemistry* **352**:129457.

Johnson JR, Lossignol D, Burnell-Nugent M and Fallon MT (2013) **An open-label extension study to investigate the long-term safety and tolerability of THC/CBD oromucosal spray and oromucosal THC spray in patients with**

terminal cancer-related pain refractory to strong opioid analgesics *Journal of Pain and Symptom Management* **46(2)**:207-218.

Jurkat-Rott K, Holzherr B, Fauler M and Lehmann-Horn F (2010) **Sodium channelopathies of skeletal muscle result from gain or loss of function** *Pflugers Archiv European Journal of Physiology* **460(2)**:239-248.

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

Kabsch W (2010) **XDS** *Acta Crystallographica* **D66**:125-132.

Kaplan DI, Isom LL and Petrou S (2016) **Role of sodium channels in epilepsy** *Cold Spring Harbor Perspectives in Medicine* **6**:a022814.

Kato HE, Zhang F, Yizhar O, Ramakrishnan C, Nishizawa T, Hirata K, Ito J, Aita Y, Tsukazaki T, Hayashi S, Hegemann P, Maturana AD, Ishitani R, Deisseroth K and Nureki O (2012) **Crystal structure of the channelrhodopsin light-gated cation channel** *Nature* **482(7385)**:369-374.

Ke S, Ulmschneider MB, Wallace BA and Ulmschneider JP (2018) **Role of the interaction motif in maintaining the open gate of an open sodium channel** *Biophysical Journal* **115**:1920-1930.

Keegan N, Ridley H and Lakey JH (2010) **Discovery of biphasic thermal unfolding of OmpC with implications for surface loop stability** *Biochemistry* **49(45)**:9715-9721.

Khoury JM, de Castro Lourenço das Neves M, Roque MAV, de Brito Queiroz DA, de Freitas AAC, de Fátima Â, Moreira FA and Garcia FD (2019) **Is there a role for cannabidiol in psychiatry?** *The World Journal of Biological Psychiatry* **20(2)**:101-116.

Kibbe WA (2007) **OligoCalc: An online oligonucleotide properties calculator** *Nucleic Acids Research* **35**:W43-46.

Kim YP, Yeo KJ, Kim MH, Kim YC and Jeon YH (2010) **Structural characterization of the intra-membrane histidine kinase YbdK from Bacillus subtilis in DPC micelles** *Biochemical and Biophysical Research Communications* 391(3):1506-1511.

Klein BD, Jacobson CA, Metcalf CS, Smith MD, Wilcox KS, Hampson AJ and Kehne JH (2017) **Evaluation of Cannabidiol in Animal Seizure Models by the Epilepsy Therapy Screening Program (ETSP)** *Neurochemical Research* ***42*****(7)**:1939-1948.

Kleinschmidt JH and Popot JL (2015) **Folding and stability of integral membrane proteins in amphipols** *Archives of Biochemistry and Biophysics* **564**:327-343.

Kypr J, Kejnovská I, Renčiuk D and Vorlíčková M (2009) **Circular dichroism and conformational polymorphism of DNA** *Nucleic Acids Research* **37(6)**:1713-25.

Lai JK, Kubelka GS and Kubelka J (2018) **Effect of Mutations on the Global and Site-Specific Stability and Folding of an Elementary Protein Structural Motif** *The Journal of Physical Chemistry B* **122(49)**:11083-11094.

Lamba J, Paul S, Hasija V, Aggarwal R and Chaudhuri TK (2009) **Monitoring protein folding and unfolding pathways through surface hydrophobicity changes using fluorescence and circular dichroism spectroscopy** *Biochemistry (Moscow)* **74(4)**:393-398.

Laprairie RB, Bagher AM, Kelly MEM and Denovan-Wright EM (2015) **Cannabidiol is a negative allosteric modulator of the cannabinoid CB1 receptor** *British Journal of Pharmacology* **172(20)***:*4790-4805.

Laskowski RA, MacArthur MW, Moss DS and Thornton JM (1993) **PROCHECK: a program to check the stereochemical quality of protein structures** *Journal of Applied Crystallography* **26**:283-291.

Lastres-Becker I, Molina-Holgado F, Ramos JA, Mechoulam R and Fernández-Ruiz J (2005). **Cannabinoids provide neuroprotection against 6-hydroxydopamine toxicity in vivo and in vitro: Relevance to Parkinson's disease** *Neurobiology of Disease* **19(1-2)**:96-107.

Lauritsen KJ and Rosenberg H (2016) **Comparison of outcome expectancies for synthetic cannabinoids and botanical marijuana** *American Journal of Drug and Alcohol Abuse* **42(4)**:377-384.

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

Le Marois M, Ballet V, Sanson C, Maizières MA, Carriot T, Chantoiseau C, Partiseti M and Bohme GA (2020) **Cannabidiol inhibits multiple cardiac ion channels and shortens ventricular action potential duration in vitro** *European Journal of Pharmacology* **886**:173542.

Leehey M (2019) **A study of tolerability and efficacy of Cannabidiol on tremor in Parkinson's disease -Study results** *ClinicalTrials.gov* Retrieved June 10 2020 from https://clinicaltrials.gov/ct2/show/NCT02818777.

Leipold E, Liebmann L, Korenke GC, Heinrich T, Giesselmann S, Baets J, Ebbinghaus M, Goral RO, Stödberg T, Hennings JC, Bergmann M, Altmüller J, Thiele H, Wetzel A, Nürnberg P, Timmerman V, De Jonghe P, Blum R, Schaible HG, Weis J, Heinemann SH, Hübner CA and Kurth I (2013) **A de novo gain-of-function mutation in SCN11A causes loss of pain perception** *Nature Genetics* **45(11)**:1399-404.

Lenaeus MJ , Gamal El-Din TM, Ing C, Ramanadane K, Pomès R, Zheng N and Catterall WA (2017) **Closed and open states of a sodium channel** *Proceedings of the National Academy of Sciences* **114 (15)**:E3051-E3060.

Leweke FM, Piomelli D, Pahlisch F, Muhl D, Gerth CW, Hoyer C, Klosterkötter J, Hellmich M and Koethe D (2012) **Cannabidiol enhances anandamide signaling and alleviates psychotic symptoms of schizophrenia** *Translational Psychiatry* **2(3)**:e94.

Lewinter RD, Skinner K, Julius D and Basbaum AI (2004) **Immunoreactive TRPV-2 (VRL-1), A Capsaicin Receptor Homolog, in the Spinal Cord of the Rat** *Journal of Comparative Neurology* **470(4)**:400-408.

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 in English)* **60(20)**:11474-11480.

Linares M, Sun H, Biler M, Andréasson J and Norman P (2019) **Elucidating DNA binding of dithienylethenes from molecular dynamics and dichroism spectra** *Physical Chemistry Chemical Physics* **21**:3637-3643.

Liu G, Yarov-Yarovoy V, Nobbs M, Clare JJ, Scheuer T and Catterall WA (2003) **Differential interactions of lamotrigine and related drugs with transmembrane segment IVS6 of voltage-gated sodium channels** *Neuropharmacology* **44(3)**:413-422.

Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD and Lopez R (2019) **The EMBL-EBI search and sequence analysis tools APIs in 2019** *Nucleic Acids Research* **47(W1)**:W636-W641.

Maliar N, Okhrimenko IS, Petrovskaya LE, Alekseev AA, Kovalev KV, Soloviov DV, Popov PA, Rokitskaya TI, Antonenko YN, Zabelskii DV, Dolgikh DA, Kirpichnikov MP and Gordeliy VI (2020) **Novel pH-Sensitive Microbial Rhodopsin from Sphingomonas paucimobilis** *Doklady Biochemistry and Biophysics* **495**:342-346.

Manieri TM, Sensi SL, Squitti R and Cerchiaro G (2021) **Structural effects of stabilisation and complexation of a zinc-deficient superoxide dismutase** *Heliyon* **7(1)**:e06100.

Manzanares J, Julian M and Carrascosa A (2006) **Role of the Cannabinoid System in Pain Control and Therapeutic Implications for the Management of Acute and Chronic Pain Episodes** *Current Neuropharmacology* **4(3)**:239-257.

Mao D and Wallace BA (1984) **Differential Light Scattering and Absorption Flattening Optical Effects Are Minimal in the Circular Dichroism Spectra of Small Unilamellar Vesicles** *Biochemistry* **23(12)**:2667-2673.

Marada A, Karri S, Singh S, Allu PK, Boggula Y, Krishnamoorthy T, Guruprasad L, Babu V and Sepuri N (2016) **A Single Point Mutation in Mitochondrial Hsp70 Cochaperone Mge1 Gains Thermal Stability and Resistance** *Biochemistry* **55(51)**:7065-7072.

Mason ER and Cummins TR (2020) **Differential inhibition of human Nav1.2 resurgent and persistent sodium currents by cannabidiol and GS967** *International Journal of Molecular Sciences* **21(7)**:2454.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC and Read RJ (2007) **Phaser crystallographic software** *Journal of Applied Crystallography* **404**:658-674.

McCusker EC, Bagnéris C, Naylor CE, Cole AR, D'Avanzo N, Nichols CG and Wallace BA (2012) **Structure of a bacterial voltage-gated sodium channel pore reveals mechanisms of opening and closing** *Nature Communications* **3**:1102.

McNicholas S, Potterton E, Wilson KS and Noble MEM (2011) **Presenting your structures: the CCP4mg molecular-graphics software** *Acta Crystallographica* **D67**:384–394.

McPartland JM and Russo EB (2015) **Handbook of Cannabis** *Oxford University Press*.

Mecha M, Feliú A, Iñigo PM, Mestre L, Carrillo-Salinas FJ and Guaza C (2013) **Cannabidiol provides long-lasting protection against the deleterious effects of inflammation in a viral model of multiple sclerosis: A role for A2A receptors** *Neurobiology of Disease* **59**:141-150.

Mechoulam R and Shvo Y (1963) **Hashish-I. The structure of Cannabidiol** *Tetrahedron* **19(12)**:2073-2078.

Mechoulam R and Parker LA (2013) **The endocannabinoid system and the brain** *Annual Review of Psychology* **64**:21-47.

Metrick MA, Temple JE and Macdonald G (2013) **The effects of buffers and pH on the thermal stability, unfolding and substrate binding of RecA** *Biophysical Chemistry* **184**:29-36.

Mikuriya TH (1969) **Marijuana in medicine: past, present and future** *California medicine* **110**:34-40.

Miles AJ and Wallace BA (2016) **Circular dichroism spectroscopy of membrane proteins** *Chemical Society Reviews* **45**:4859-4872.

Miles AJ and Wallace BA (2018) **CDtoolX a downloadable software package for processing and analyses of circular dichroism spectroscopic data** *Protein Science* **27(9)**:1717-1722.

Mills B, Yepes A and Nugent K (2015) **Synthetic cannabinoids** *American Journal of the Medical Sciences* **350**:59-62.

Montini G, Booker J, Sula A and Wallace BA (2018) **Comparisons of voltage-gated sodium channel structures with open and closed gates and implications for state-dependent drug design** *Biochemical Society Transactions* **46**:1567-1575.

Moreau JJ (1972) **Du Hachisch et de l'Alienation Mentale: Etudes Psychologiques. English Edition** *Paris, New York: Librarie de Fortin Mason Raven Press*.

Morelli MB, Amantini C, Liberati S, Santoni M and Nabissi M (2013) **TRP channels: new potential therapeutic approaches in CNS neuropathies** *CNS and Neurological Disorders - Drug Targets* **12**:274-293.

Munro S, Thomas KL and Abu-Shaar M (1993) **Molecular characterization of a peripheral receptor for cannabinoids** *Nature* **365(6441)**:61-5.

Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA ,Winn MD Long F and Vagin AA (2011) **REFMAC5 for the refinement of macromolecular crystal structures** *Acta Crystallographica* **D67:**355–367.

Nabissi M, Morelli MB, Santoni M and Santoni G (2013) **Triggering of the TRPV2 channel by cannabidiol sensitizes glioblastoma cells to cytotoxic chemotherapeutic agents** *Carcinogenesis* **34**:48-57.

Nahar L, Onder A and Sarker SD (2019) **Cannabidiol (CBD) – An update** *Phytochemical Analysis* **31(2)**:135–46.

Najar MH, Chat OA, Bhat PA, Mir MA, Rather GM and Dar AA (2021) **Structural changes in trypsin induced by the bile salts: An effect of amphiphile hydrophobicity** *International Journal of Biological Macromolecules* **180**:121-128.

Naylor CE, Bagnéris C, DeCaen PG, Sula A, Scaglione A, Clapham DE and Wallace BA (2016) **Molecular basis of ion permeability in a voltage-gated sodium channel** *The European Molecular Biology Organisation Journal* **35**:820-830.

Needleman SB and Wunsch CD (1970) **A general method applicable to the search for similarities in the amino acid sequence of two proteins** *Journal of Molecular Biology* **48(3)**:443-453.

Newstead S, Ferrandon S and Iwata S (2008) **Rationalizing α-helical membrane protein crystallization** *Protein Science* **17(3)**:466-472.

Nurani G, Radford M, Charalambous K, O'Reilly AO, Cronin NB, Haque S and Wallace BA (2008) **Tetrameric bacterial sodium channels: Characterization of structure, stability, and drug binding** *Biochemistry* **47(31)**:8114-21.

Nurmikko TJ, Serpell MG, Hoggart B, Toomey PJ, Morlion BJ and Haines D (2007) **Sativex successfully treats neuropathic pain characterised by allodynia: A randomised, double-blind, placebo-controlled clinical trial** *Pain* **133(1-3)**:210-220.

O'Brien JE and Meisler MH (2013) **Sodium channel SCN8A (Nav1. 6): properties and de novo mutations in epileptic encephalopathy and intellectual disability** *Frontiers in Genetics* **4**:213.

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

O'Mahony AM, Cronin MF, McMahon A, Evans JC, Daly K, Darcy R and O'Driscoll CM (2014) **Biophysical and structural characterisation of nucleic acid complexes with modified cyclodextrins using circular dichroism** *Journal of Pharmaceutical Sciences* **103(5)**:1346-55.

O'Reilly AO, Charalambous K, Nurani G, Powl AM and Wallace BA (2008) **G219S mutagenesis as a means of stabilizing conformational flexibility in the bacterial sodium channel NaChBac** *Molecular Membrane Biology* **25(8)**:670-676.

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 Nav 1.2 by the m-conotoxin KIIIA** *Science* **363(6433)**: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 of the United States of America* **118(11)**:e2100066118.

Pan X, Li Z, Zhou Q, Shen H, Wu K, Huang X, Chen J, Zhang J, Zhu X, Lei J, Xiong W, Gong H, Xia B and Yan N (2018) **Structure of the human voltage-gated sodium channel Nav 1.4 in complex with β1** *Science* **362(6412)**:eaau2486.

Pandey R, Mousawy K, Nagarkatti M and Nagarkatti P (2009) **Endocannabinoids and immune regulation** *Pharmacological Research* **60(2)**:85-92.

Parker JL and Newstead S (2016) **Membrane protein crystallisation: Current trends and future perspectives** *Advances in Experimental Medicine and Biology* **922**: 61-72.

Parker JL and Newstead S (2012) **Current trends in α-helical membrane protein crystallization: An update** *Protein Science* **21(9)**:1358-1365.

Patching SG, Edara S, Ma P, Nakayama J, Hussain R, Siligardi G and Phillips-Jones MK (2012) **Interactions of the intact FsrC membrane histidine kinase with its pheromone ligand GBAP revealed through synchrotron radiation circular dichroism** *Biochimica et Biophysica Acta – Biomembranes* **1818(7)**:1595-1602.

Patel RR, Barbosa C, Brustovetsky T, Brustovetsky N and Cummins TR (2016) **Aberrant epilepsy-associated mutant Nav1.6 sodium channel activity can be targeted with cannabidiol** *Brain: A Journal of Neurology* **139**:2164-2181.

Patrick MH (2007) **Crystallography made crystal clear: A guide for users of macromolecular models, 3rd Edition** *Academic Press.*

Pauli CS, Conroy M, Vanden Heuvel BD and Park SH (2020) **Cannabidiol Drugs Clinical Trial Outcomes and Adverse Effects** *Frontiers in Pharmacology* **11**:63.

Payandeh J, Scheuer T, Zheng N and Catterall WA (2011) **The crystal structure of a voltage-gated sodium channel** *Nature* **475**:353-358.

Payandeh J, Gamal El-Din TM, Scheuer T, Zheng N and Catterall WA (2012) **Crystal structure of a voltage-gated sodium channel in two potentially inactivated states** *Nature* **486**:135–139.

Perucca P and Gilliam FG (2012) **Adverse effects of antiepileptic drugs** *The Lancet Neurology* **11(9)**:792-802.

Promega Corporation (2009) **Wizard Plus SV Minipreps DNA Purification System Quick Protocol**.

Psakis G, Saidijam M, Shibayama K, Polaczek J, Bettaney KE, Baldwin JM, Baldwin SA, Hope R, Essen LO, Essenberg RC and Henderson PJF (2009) **The sodium-dependent D-glucose transport protein of Helicobacter pylori** *Molecular Microbiology* **71(2)**:391-403.

Pumroy RA, Samanta A, Liu Y, Hughes TET, Zhao S, Yudin Y, Rohacs T, Han S and Moiseenkova-Bell VY (2019) **Molecular mechanism of TRPV2 channel modulation by cannabidiol** *eLife* **8**:48792.

Qin N, Neeper MP, Liu Y, Hutchinson TL, Lubin ML and Flores CM (2008) **TRPV2 is activated by cannabidiol and mediates CGRP release in cultured rat dorsal root ganglion neurons** *Journal of Neuroscience* **28**:6231-6238.

Raman IM and Bean BP (1997) **Resurgent sodium current and action potential formation in dissociated cerebellar Purkinje neurons** *Journal of Neuroscience* **17(12)**:4517-4526.

Raman IM, Sprunger LK, Meisler MH and Bean BP (1997) **Altered subthreshold sodium currents and disrupted firing patterns in Purkinje neurons of Scn8a mutant mice** *Neuron* **19(4)**:881-91.

Ranjbar B and Gill P (2009) **Circular dichroism techniques: Biomolecular and nanostructural analyses- A review** *Chemical Biology and Drug Design* **74(2)**:101-20.

Rath A, Glibowicka M, Nadeau VG, Chen G and Deber CM (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)**:1760-1765.

Rees E, Carrera N, Morgan J, Hambridge K, Escott-Price V, Pocklington AJ, Richards AL, Pardiñas AF; GROUP Investigators, McDonald C, Donohoe G, Morris DW, Kenny E, Kelleher E, Gill M, Corvin A, Kirov G, Walters JTR, Holmans P, Owen MJ and O'Donovan MC (2019) **Targeted Sequencing of 10,198 Samples Confirms Abnormalities in Neuronal Activity and Implicates Voltage-Gated Sodium Channels in Schizophrenia Pathogenesis** *Biological Psychiatry* **85(7)**:554-562.

Rosenberg EC, Patra PH and Whalley BJ (2017) **Therapeutic effects of cannabinoids in animal models of seizures, epilepsy, epileptogenesis, and epilepsy-related neuroprotection** *Epilepsy and Behavior* **70**:319-327.

Rupp B (2009) **Biomolecular crystallography: principles, practice, and application to structural biology** *Garland Science.*

Sait LG, Sula A, Ghovanloo MR, Hollingworth D, Ruben PC and Wallace BA (2020) **Cannabidiol interactions with voltage-gated sodium channels** *Elife* **9**:e58593.

Salamanca S and Chang JY (2005) **Unfolding and refolding pathways of a major kinetic trap in the oxidative folding of α-lactalbumin** *Biochemistry* **44(2)**:744-50.

Scalmani P, Rusconi R, Armatura E, Zara F, Avanzini G, Franceschetti S and Mantegazza, M (2006) **Effects in neocortical neurons of mutations of the**

Nav1.2 Na+ channel causing benign familial neonatal-infantile seizures *Journal of Neuroscience* **26(40)**:10100-10109.

Schlegel S, Klepsch M, Gialama D, Wickström D, Slotboom DJ and De Gier JW (2010) **Revolutionizing membrane protein overexpression in bacteria** *Microbial Biotechnology* **3(4)**:403-411.

Seddon AM, Curnow P and Booth PJ (2004) **Membrane proteins, lipids and detergents: Not just a soap opera** *Biochimica et Biophysica Acta – Biomembranes* **1666(1-2)**:105-17.

Sharma D, Singh A, Pathak M, Kaur L, Kumar V, Roy BG and Ojha H (2020) **DNA binding and antiradical potential of ethyl pyruvate: Key to the DNA radioprotection** *Chemico-Biological Interactions* **332**:109313.

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)**:1303-1308.

Shimosato G, Amaya F, Ueda M, Tanaka Y, Decosterd I and Tanaka M (2005) **Peripheral inflammation induces up-regulation of TRPV2 expression in rat DRG** *Pain* **119(1-3)**:225-232.

Sievers F and Higgins DG (2018) **Clustal Omega for making accurate alignments of many protein sequences** *Protein Science* **27(1)**:135-145.

Siligardi G, Hussain R, Patching SG and Phillips-Jones MK (2014) **Ligand- and drug-binding studies of membrane proteins revealed through circular dichroism spectroscopy** *Biochimica et Biophysica Acta – Biomembranes* **1838**:34-42.

Smart OS, Neduvelil JG, Wang X, Wallace BA and Sansom MSP (1996) **HOLE: A program for the analysis of the pore dimensions of ion channel structural models** *Journal of Molecular Graphics* **14(6)**:354-60.

Smart OS, Womack TO, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C and Bricogne G (2012) **Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER** *Acta Crystallographica* **D68***:*368-80.

Söding J (2005) **Protein homology detection by HMM-HMM comparison** *Bioinformatics* **21(7)**:951-60.

Stark T, Ruda-Kucerova J, Iannotti FA, D'Addario C, Di Marco R, Pekarik V, Drazanova E, Piscitelli F, Bari M, Babinska Z, Giurdanella G, Di Bartolomeo M, Salomone S, Sulcova A, Maccarrone M, Wotjak CT, Starcuk Z Jr, Drago F, Mechoulam R, Di Marzo V and Micale V (2019) **Peripubertal cannabidiol treatment rescues behavioral and neurochemical abnormalities in the MAM model of schizophrenia** *Neuropharmacology* **146**:212-221.

Stockings E, Zagic D, Campbell G, Weier M, Hall WD, Nielsen S, Herkes GK, Farrell M and Degenhardt L (2018) **Evidence for cannabis and cannabinoids for epilepsy: A systematic review of controlled and observational evidence** *Journal of Neurology, Neurosurgery and Psychiatry* ***89*****(7)**:741-753.

Striano P, Bordo L, Lispi ML, Specchio N, Minetti C, Vigevano F and Zara F (2006) **A novel SCN2A mutation in family with benign familial infantile seizures** *Epilepsia* **47**:218-220.

Sula A, Booker J, Ng LC, Naylor CE, DeCaen PG and Wallace BA (2017) **The complete structure of an activated open sodium channel** *Nature Communications* **8**:14205.

Sula A, DeCaen P, Naylor C, Zanatta G, Bagnéris C, Clapham DE, Pryde D and Wallace BA (2016) **Sodium Channel/Ligand Complex Structures as a Guide for Rational Drug Design** *Biophysical Journal* **110(3)**:33a.

Sula A, Hollingworth D, Ng LCT, Larmore M, DeCaen PG and Wallace BA (2021) **A tamoxifen receptor within a voltage-gated sodium channel** *Molecular Cell* **81(6)**:1160-1169.

Sula A and Wallace BA (2017) **Interpreting the functional role of a novel interaction motif in prokaryotic sodium channels** *Journal of General Physiology* **149**:613-622.

Szenczi Á, Kardos J, Medgyesi GA and Závodszky P (2006) **The effect of solvent environment on the conformation and stability of human polyclonal IgG in solution** *Biologicals* **34**:5-14.

Thiele EA, Marsh ED, French JA, Mazurkiewicz-Beldzinska M, Benbadis SR, Joshi C, Lyons PD, Taylor A, Roberts C and Sommerville K; GWPCARE4 Study Group (2018) **Cannabidiol in patients with seizures associated with Lennox-Gastaut syndrome (GWPCARE4): a randomised, double-blind, placebo-controlled phase 3 trial** *Lancet* **391(10125)**:1085-1096.

Todsen WL (2014) **ChemDoodle 6.0** *Journal of Chemical Information and Modelling* **54(8)**:2391-2393.

Touw M (1981) **The religious and medicinal uses of Cannabis in China, India and Tibet** *Journal of Psychoactive Drugs* **13(1)**:23-34.

Ulmschneider MB, Bagnéris C, McCusker EC, DeCaen PG, Delling M, Clapham DE Ulmschneider JP and Wallace BA (2013) **Molecular dynamics of ion transport through the open conformation of a bacterial voltage-gated sodium channel** *Proceedings of the National Academy of Sciences of the United States of America* **110**:6364-6369.

Urits I, Gress K, Charipova K, Habib K, Lee D, Lee C, Jung JW, Kassem H, Cornett E, Paladini A, Varrassi G, Kaye AD and Viswanath O (2020) **Use of cannabidiol (CBD) for the treatment of chronic pain** *Best Practice and Research Clinical Anaesthesiology* **34(3)**:463-477.

van Stokkum IHM, Spoelder HJW, Bloemendal M, van Grondelle R and Groen FCA (1990) **Estimation of protein secondary structure and error analysis from circular dichroism spectra** *Analytical Biochemistry* **191(1)**:110-8.

Verchère A, Broutin I and Picard M (2017) **Membrane Protein Structure and Function Characterization** *Humana Press*.

Volkers L, Kahlig KM, Verbeek NE, Das JHG, van Kempen MJA, Stroink H, Augustijn P, van Nieuwenhuizen O, Lindhout D, George AL, Koeleman BPC and Rook MB (2011) **Nav1.1 dysfunction in genetic epilepsy with febrile seizures-plus or Dravet syndrome** *European Journal of Neuroscience* **34(8)**:1268-75.

Wallace BA and Mao D (1984) **Circular dichroism analyses of membrane proteins: An examination of differential light scattering and absorption flattening**

effects in large membrane vesicles and membrane sheets *Analytical Biochemistry* **142(2)**:317-28.

Wallace BA (2003) **Analyses of circular dichroism spectra of membrane proteins** *Protein Science* **13**:100-112.

Wallace BA and Janes RW (2009) **Modern techniques for circular dichroism and synchrotron radiation circular dichroism spectroscopy** *IOS press*.

Wang Z, Lin Y, Liu W, Kuang P, Lao W, Ji Y and Zhu H (2019) **Voltage-Gated sodium Channels Are Involved in Cognitive Impairments in Parkinson's Disease-like Rats** *Neuroscience* **418**:231-243.

Ward SJ, McAllister SD, Kawamura R, Murase R, Neelakantan H and Walker EA (2014) **Cannabidiol inhibits paclitaxel-induced neuropathic pain through 5-HT 1A receptors without diminishing nervous system function or chemotherapy efficacy** *British Journal of Pharmacology* **171(3)**:636-645.

Whitmore L, Miles AJ, Mavridis L, Janes RW and Wallace BA (2017) **PCDDB: New developments at the protein circular dichroism data bank** *Nucleic Acids Research* **45(D1)**:D303-D307.

Whitmore L and Wallace BA (2008) **Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases** *Biopolymers* **89(5)**:392-400.

Whitmore L and Wallace BA (2004) **DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data** *Nucleic Acids Research* **32**:W668-73.

Whyte LS, Ryberg E, Sims NA, Ridge SA, Mackie K, Greasley PJ, Ross RA and Rogers MJ (2009) **The putative cannabinoid receptor GPR55 affects osteoclast function in vitro and bone mass in vivo** *Proceedings of the National Academy of Sciences of the United States of America* **106(38)**:16511-16516.

Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS and Richardson DC (2018) **MolProbity:**

More and better reference data for improved all-atom structure validation *Protein Science* **27**:293-315.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A and Wilson KS (2011) **Overview of the CCP4 suite and current developments** *Acta Crystallographica* **D67**:235-242.

Wong SS and Wilens TE (2017) **Medical cannabinoids in children and adolescents: A systematic review** *Pediatrics* **140(5)**:e20171818.

Yamaori S, Kushihara M, Yamamoto I and Watanabe K (2010) **Characterization of major phytocannabinoids, cannabidiol and cannabinol, as isoform-selective and potent** *Biochemical Pharmacology* **79(11)**:1691-8.

Yeliseev A, van den Berg A, Zoubak L, Hines K, Stepnowski S, Williston K, Yan W, Gawrisch K and Zmuda J (2020) **Thermostability of a recombinant G protein-coupled receptor expressed at high level in mammalian cell culture** *Scientific Reports* **10**:16805.

Zanatta G, Sula A, Miles AJ, Ng LCT, Torella R, Pryde DC, DeCaen PG and Wallace BA (2019) **Valproic acid interactions with the NavMs voltage-gated sodium channel** *Proceedings of the National Academy of Sciences of the United States of America* **116(52)**:26549-26554.

Zhang XY, Wen J, Yang W, Wang C, Gao L, Zheng LH, Wang T, Ran K, Li Y, Li X, Xu M, Luo J, Feng S, Ma X, Ma H, Chai Z, Zhou Z, Yao J, Zhang X and Liu JY (2013) **Gain-of-Function mutations in SCN11A cause familial episodic pain** *American Journal of Human Genetics* **93(5)**:957-966.

Zhang X, Ren W, Decaen P, Yan C, Tao X, Tang L, Wang J, Hasegawa K, Kumasaka T, He J, Wang J, Clapham DE and Yan N (2012) **Crystal structure of an orthologue of the NaChBac voltage-gated sodium channel** *Nature* **486(7401)**:130-4.

Zhou FX, Cocco MJ, Russ WP, Brunger AT and Engelman DM (2000) **Interhelical hydrogen bonding drives strong interactions in membrane proteins** *Nature Structural Biology* *7*(**2**):154-160.

Zuardi AW (2006) **History of cannabis as a medicine: A review** *Revista Brasileira de Psiquiatria* **28(2)**:153-7.