The dynamic structure of Rab46 – a novel endothelial protein

Sabina Daria Wiktor

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Medicine and Health

Leeds Institute of Cardiovascular and Metabolic Medicine

December 2023

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Sabina Daria Wiktor to be identified as Author of this work has been asserted by her in accordance with the Copyright, Designs and Patents Act 1988.

© 2023 The University of Leeds and Sabina Daria Wiktor

Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Alexander Breeze and Dr Lynn McKeown for all their help and guidance throughout the way.

I am grateful to Dr Brian Jackson, Dr Fatima Nadat-Skyrme and Laura Wilkinson Hewitt for showing me the ropes of protein expression and purification.

I would like to thank Dr Hugh Smith, Dr Arnout Kalverda, Dr Theodoros Karamanos and Dr Matthew Batchelor for their help with setting up NMR experiments and NMR data processing and analysis, as well as all the valuable discussions about this project. I am also grateful to Dr Denisa Hoxha for the helpful discussions about NMR data interpretation.

Further thanks to Pam Dossang for the help with organising the work placement at Charles River Laboratories. Here, I would also like to extend my gratitude to the whole CRL team for being so supportive over the last year.

Finally, I would like to thank my family and friends, and especially my partner, Kacper for all the support and encouragement throughout this PhD.

Abstract

Weibel-Palade bodies (WPBs) are endothelial-specific secretory organelles that serve a role of the vascular 'emergency kit'. WPBs store a cocktail of haemostatic, pro-inflammatory and pro-angiogenic mediators that can be released on demand to provide a rapid response to vascular injury. However, since untimely, inappropriate and excessive exocytosis of WPBs may contribute to the development and progression of atherosclerosis, secretion from endothelial cells has to be tightly coordinated.

Rab46 is a novel multi-domain member of Rab GTPase family of proteins that localizes to WPBs. It is a non-canonical Rab GTPase as, in addition to the conserved Rab domain, Rab46 contains a coiled-coil domain responsible for interactions with other proteins and an EF-hand domain which acts as a built-in Ca^{2+} sensor. Importantly, findings from McKeown laboratory demonstrated that Rab46 is a key regulator of context-dependent differential WPB trafficking. Since at the time when this study was commenced there had been no structural studies of Rab46, the aim of this project was to investigate the structure, dynamics, and cellular interaction partners of Rab46 using a combination of NMR spectroscopy complemented by a range of other biophysical methods.

As part of this study a system for overexpression of recombinant EF-hand, Coiled-coil and Rab domains of Rab46 was successfully established in *E. coli*, allowing for generation of high-quality protein samples for the downstream structural biology applications. The investigation of EF-hand domain properties using high-resolution solution NMR provided novel insights into the structural characteristics of the domain. Importantly, the triple resonance backbone assignment obtained for both apo and Ca^{2+}-saturated forms of EF-hand domain lays the foundation for further investigation of the Ca^{2+}-regulated functions of Rab46. Furthermore, investigation of the properties of recombinantly produced Rab domain provided preliminary evidence for an atypical GTPase profile of Rab46 characterised by high intrinsic nucleotide exchange rate of its Rab domain. Overall, the results presented in this thesis provide a solid starting point for future structural characterisation attempts of this atypical large Rab GTPase.

Table of contents

List of Figures

List of Tables

List of abbreviations

AGE	Advanced Glycation End Product
BEVS	Baculovirus Expression Vector System
BME	Basal Medium Eagle
BSA	Bovine Serum Albumin
cAMP	Cyclic Adenosine Monophosphate
CD	Circular Dichroism
CRP	C-reactive Protein
CSI	Chemical Shift Index
CSP	Chemical Shift Perturbation
Ct	Cycle Threshold
DTT	Dithiothreitol
EC	Endothelial Cell
EDTA	Ethylenediamine Tetraacetic Acid
FBS	Foetal Bovine Serum
FI	Fluorescence Intensity
FL	Full Length
FT	Flow Through
GAP	GTPase-Activating Protein
GDF	GDI-Displacement Factor
GDI	GDP Dissociation Inhibitor
GDP	Guanosine Diphosphate
GEF	Guanine Nucleotide Exchange Factor
GFP	Green Fluorescent Protein
GGTase	Geranylgeranyl transferase
GTP	Guanosine-5'-triphosphate
HRP	Horseradish Peroxidase
HSQC	Heteronuclear Single Quantum Coherence
IMAC	Immobilized Metal Ion Affinity Chromatography
IMS-MS	Ion-Mobility Spectrometry-Mass Spectrometry
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
LB	Luria-Bertani
MANT-GDP	N-Methylanthraniloyl Guanosine Diphosphate
MCS	Multiple Cloning Site

MOI	Multiplicity of Infection
MRW	Mean Residue Weight
MS	Mass Spectrometry
MTOC	Microtubule Organizing Centre
MW	Molecular Weight
MWCO	Molecular Weight Cutoff
Ni-NTA	Nickel-nitrilotriacetic Acid
NMR	Nuclear Magnetic Resonance
NO	Nitric Oxide
NOE	Nuclear Overhauser Effect
NUS	Non-uniform Sampling
PBS	Phosphate-buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene Glycol
pfu	Plaque-forming Unit
PM	Protein Marker
PMSF	Phenylmethylsulfonyl Fluoride
QAA	Quantitative Amino Acid analysis
REP	Rab Escort Protein
RMSD	Root Mean Square Deviation
ROS	Reactive Oxygen Species
RT	Room Temperature
RT-qPCR	Reverse Transcription-quantitative Polymerase Chain Reaction
SASA	Solvent Accessible Surface Area
SD	Standard Deviation
SDM	Site-directed Mutagenesis
SDS-PAGE	Sodium Dodecyl Sulphate–Polyacrylamide Gel Electrophoresis
SEC	Size-exclusion Chromatography
SMILE	Sparse Multidimensional Iterative Lineshape Enhanced
SNP	Single Nucleotide Polymorphisms
SOC	Super Optimal Medium with Catabolic Repressor
SOCE	Store-operated Ca^{2+} entry
SW	Spectral Width
TB	Terrific Broth

TBS-T	Tris-buffered Saline, 0.1% Tween ® 20
TCEP	Tris (2-carboxyethyl)phosphine
TCR	T-cell receptor
TEV	Tobacco Etch Virus
TNF-α	Tumour Necrosis Factor-α
TOI	Time of Infection
tRNA	Transfer RNA
TROSY	Transverse Relaxation Optimized Spectroscopy
v/v	volume/volume
vWF	von Willebrand Factor
WCL	Whole Cell Lysate
WPB	Weibel-Palade Body
wt	wild type
w/v	weight/volume

Chapter 1
Introduction

1.1. Background

1.1.1. The vascular endothelium

The endothelium is a layer of cells that lines the entire vascular system – from the largest arteries and veins to the capillaries and post-capillary venules (1,2). Since it was first identified over a century ago, it has been increasingly recognised that the vascular endothelium does not simply constitute an inert physical barrier between the blood and the vessel wall, as originally believed, but rather is a large, complex and dynamic 'organ', actively involved in the regulation of a variety of physiologically important processes such as control of the vascular tone, angiogenesis, blood coagulation, and leukocyte adhesion and extravasation, all of which play a crucial role in the wound healing process by minimizing blood loss, and preventing pathogens from spreading (1,3,4). At the same time, the ability of healthy vascular endothelium to maintain the balance between the endothelium-derived vasodilators and vasoconstrictors, growth promoting and inhibiting factors, and procoagulant and anticoagulant mediators, is key for sustaining the vasodilatory, anti-inflammatory and anti-thrombotic endothelial phenotype essential for the maintenance of normal vascular homeostasis.

1.1.2. Endothelial dysfunction

Given the multitude of crucial physiological processes facilitated by the vascular endothelium, it is not surprising that impaired endothelial function is implicated in the pathogenesis of multiple forms of cardiovascular disease (5). Endothelial dysfunction, defined as alterations in the functional phenotype of endothelial cells (ECs), manifests as impaired endothelium-dependent vasodilation, augmented endothelial permeability, and heightened pro-thrombotic and pro-inflammatory states, all of which promote the development and progression of atherosclerosis (5).

Endothelial injury and dysfunction can be induced by traditional cardiovascular risk factors such as hyperglycaemia, hypercholesterolaemia, hypertension, aging, and smoking, as well as non-traditional risk factors such as autoinflammatory and autoimmune diseases, and viral infections (6–8). Although adverse modulation of endothelial function can occur through multiple distinct mechanisms, a common denominator between many of these risk factors is the concomitant state of chronic inflammation and augmented generation of reactive oxygen species (ROS), and the resulting increase in oxidative stress which disturbs normal endothelial responses (6,9).

1.1.3. Cardiovascular disease risk factors in endothelial dysfunction

Reduced nitric oxide (NO) bioavailability is a hallmark of endothelial dysfunction (10). Besides its well-established role as a potent vasodilator, vascular NO has numerous anti-atherogenic and vasoprotective effects, which include regulation of vascular wall permeability, preventing platelet activation and aggregation, and inhibition of leukocyte recruitment and adhesion (11). In endothelial cells, NO is produced by endothelial nitric oxide synthase (eNOS) which utilises tetrahydrobiopterin (BH_4), NADPH and molecular oxygen to convert its substrate L-arginine to L-citrulline and NO (12). Pathological conditions associated with increased oxidative stress can lead to eNOS dysfunction and diminish NO bioavailability. For example, in diabetes and hypercholesterolaemia, the heightened oxidative stress attributed to increased blood sugar and oxidatively modified low-density lipoprotein (OxLDL) levels has been shown to result in oxidation of BH_4, an essential cofactor of eNOS (13,14), to catalytically incompetent dihydrobiopterin (BH_2), which antagonizes the effects of BH_4 by competitive binding to eNOS (15). As a consequence of the relative increase in the BH_2:BH_4 ratio, eNOS becomes uncoupled, causing a switch from NO to superoxide anion (O^{2-}) production, which not only results in limited NO synthesis, but also further exacerbates endothelial oxidative stress by increased production of ROS (16). Considering the many roles of NO in the vasculature mentioned above, the results of its compromised bioavailability are far beyond impaired endothelium-dependent vasodilation. For instance, given that NO has been reported to inhibit both Weibel-Palade bodies' exocytosis and von Willebrand factor-mediated platelet adhesion, defective NO synthesis may lead to increased plasma levels of von Willebrand Factor, and contribute to the pro-thrombotic state observed in atherosclerosis (17–19).

Inflammatory activation of endothelial cells is an important event underlying the initiation and progression of atherosclerosis (20). Chronic low-grade inflammatory state accompanying diabetes is associated with elevated levels of circulating inflammatory markers including interleukin-6 (IL-6), tumour necrosis factor-α (TNF-α) and C-reactive protein (CRP). Hyperglycaemia and the increased formation of advanced glycation end products (AGEs) under hyperglycaemic conditions have been both demonstrated to activate the ECs through sustained activation of nuclear transcription factor NF-κB (21,22), which results in enhanced expression of adhesion molecules including VCAM-1, ICAM-1, and E-selectin (23,24) contributing to increased endothelial permeability, and leukocyte recruitment. EC activation then leads to prothrombotic phenotype by upregulation of coagulation factors by inflammatory mediators including TNF-α and IL-6

(25,26). What is more, another important consequence of endothelial activation, mediated by to the effects of proinflammatory cytokines is decreased synthesis of endothelial NO (27). Furthermore, endothelial activation and consequent vascular injury trigger the release of pro-thrombotic and pro-inflammatory mediators such as von Willebrand factor and P-selectin from Weibel-Palade bodies (28). The resulting sustained proinflammatory and procoagulant state, characteristic of activated endothelium, aggravates vascular injury and endothelial dysfunction.

Since excessive, inadequate and prolonged endothelial cell activation contributes to endothelial dysfunction and, as such, to development of cardiovascular disease, ECs need to be equipped with appropriate molecular machinery to ensure that not only prompt but, crucially, adequate response to vascular injury is mounted.

1.1.4. Weibel-Palade bodies

Endothelial exocytosis is one of the first lines of defence against vascular trauma and storage of pre-made haemostatic and inflammatory mediators within secretory vesicles, known as Weibel-Palade bodies (WPBs), is one of the key strategies enabling ECs to quickly act upon vascular damage (2,29,30). WPBs are EC-specific organelles often referred to as ‘the vascular first aid kit’ (29). These large (1-3 μm) secretory granules contain a cocktail of haemostatic, pro-inflammatory and pro-angiogenic mediators that can be released on demand to rapidly respond to changes in the micro-environment resulting from vascular perturbation (31). Following stimulation, WPBs rapidly release the bioactive components into the circulation initiating processes necessary for vascular repair (31,32). The main cargo of WPBs is von Willebrand factor (vWF), a large multimeric glycoprotein crucial for WPB formation and platelet adhesion to the injured vessel wall – the process that initiates platelet plug formation to arrest bleeding (2,33). However, other WPB constituents such as P-selectin, angiopoietin-2, interleukin (IL)-8, eotaxin-3 and endothelin-1 (to name a few) also significantly contribute to the vascular repair process by triggering cellular responses orchestrating haemostasis and inflammation (34).

1.1.5. Differential exocytosis of WPBs

Of note, it has been demonstrated that endothelial cells accumulate distinct populations of WPBs which differ in the expression levels of the stored cargo (29). Markedly, the storage of P-selectin, a leukocyte receptor protein mediating early steps of leukocyte recruitment to the sites of inflammation, and angiopoetin-2, an angiogenic factor involved in regulation of cell migration required for wound closure, has been shown to be mutually

exclusive (35). Since angiopoietin-2 and P-selectin are sorted to distinct subpopulations of WPBs, context-dependent responses can be obtained by their selective release (34). Such differential exocytosis of WPBs represents a new layer of regulation enabling normal inflammatory response without promoting potentially dangerous thrombosis and vice versa (32). Recent studies showed how selective exclusion of subpopulations of WPBs from exocytosis can limit the release of physiologically irrelevant cargo and provided an invaluable insight into the molecular machinery involved in differential trafficking and exocytosis of WPBs, nevertheless, the exact mechanisms involved are yet to be elucidated (29,32,33).

Exocytosis of WPBs can be evoked in response to different physiological stimuli including physical injury like trauma or hypoxia and agonists such as epinephrine, thrombin and histamine (29,36–38). Depending on their mode of action, secretagogues triggering WPB exocytosis can be split into two distinct groups – agonists that act via activating cAMP signalling such as epinephrine and the ones that elicit their effects by increasing intracellular Ca^{2+} levels, like histamine and thrombin (34). The mechanisms by which ECs distinguish between seemingly homogeneous Ca^{2+} signals, like the ones induced by thrombin and histamine stimulation, to produce an agonist-appropriate cellular response are, however, unclear. The resulting agonist-induced WPB trafficking and exocytosis is a multistep process strictly controlled by a complex interplay between different molecular components of which members of Rab GTPase family of proteins are of particular importance (39).

1.1.6. Rab GTPases as master regulators of intracellular trafficking

Rab GTPases play a pivotal role in defining membrane identity and regulating intracellular trafficking events at all levels - from formation of transport vesicles to fusion with their target membranes (40). Rab GTPases are highly conserved G proteins constituting the largest family within the Ras superfamily of small GTPases. To date, 70 Rab proteins have been identified in humans, which demonstrates a plethora of membrane transport events in which they participate (41). Rab proteins are typically small, ranging from 20 to 30 kDa in size, however, examples of large, multi-domain Rabs have been described in the literature (42–44). The typical domain organisation and structural fold of Rab GTPases illustrated with an example of Rab3a GTPase is shown in Figure 1. Like other members of Ras superfamily, Rab proteins principally function as nucleotide-binding molecular switches, interconverting between the GDP-bound ‘off’ and GTP-bound ‘on’ states (45). Rab activation cycle and its regulation are illustrated in Figure 2.

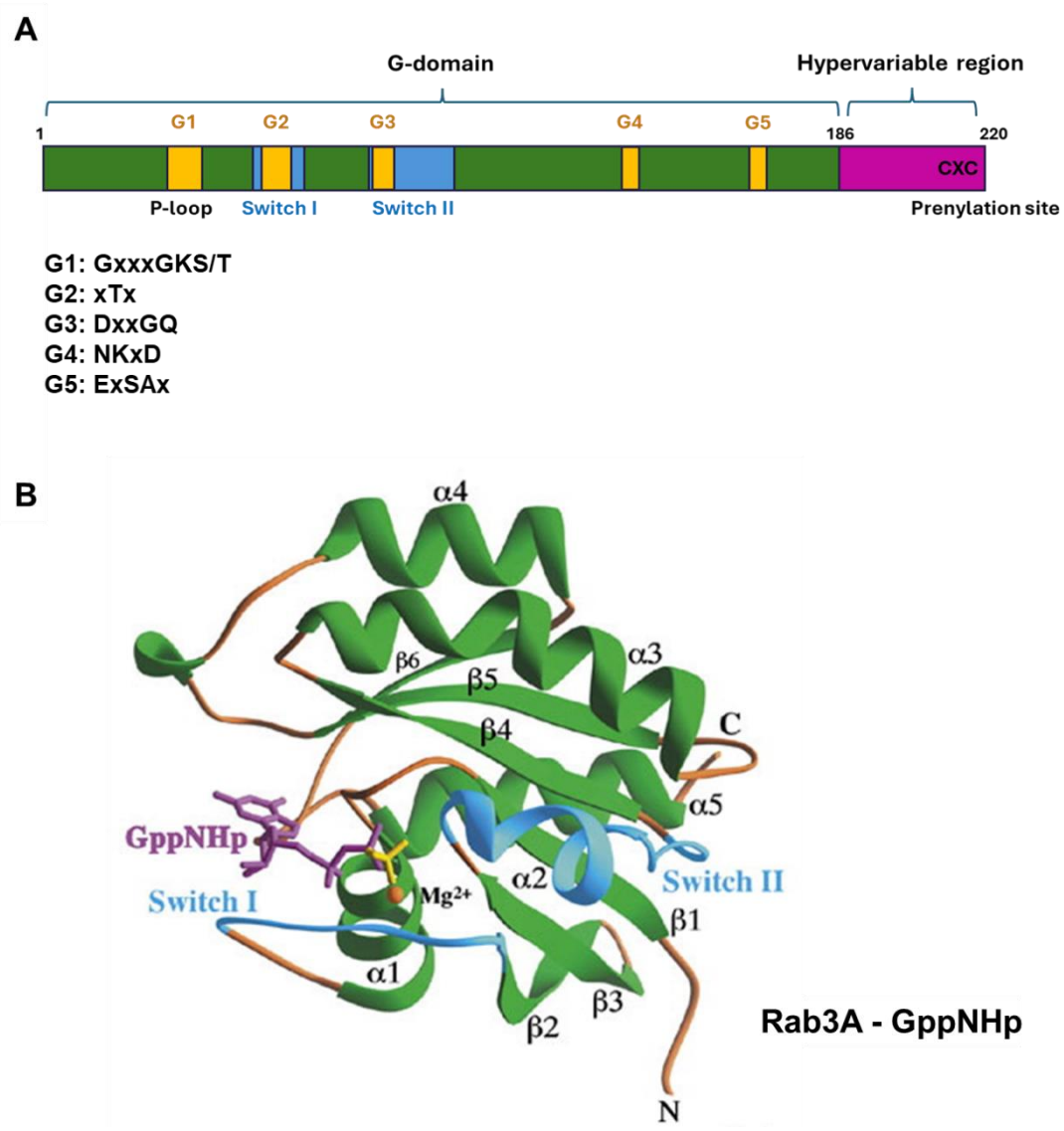

Figure 1 Domain organisation and structural fold of Rab GTPases. A) Schematic representation of the domain architecture of Rab3a GTPase (NP_002857.1). The conserved sequence motifs G1-G5 (G boxes) and Switch regions located within the G-domain of Rab3a (residues 1-186), as well as the C-terminal Hypervariable region (residues 186-220), and the prenylation site are highlighted. Consensus sequences of the five conserved G motifs containing residues involved in nucleotide binding are shown below the diagram. G1 box, also called P-loop/Walker A motif is involved in binding the phosphate moiety of nucleotides. The highly conserved K residue in the G1 box directly interacts with the γ-phosphate of GTP. The conserved S/T residue in the G1 box and T residue in G2 box/Switch I region, as well as the D residue in the G3 box/Switch II region do not make direct contacts with the nucleotide but coordinate Mg^{2+} which bridges the β- and γ-phosphate of GTP. The G3 box/Switch II region also contains the conserved catalytic Q residue, important for GTP hydrolysis. Binding of GTP γ-phosphate mediates

constraining interactions between the phosphate- and Mg^{2+}-binding motifs, which stabilises the Switch I and II regions in the ‘on’ conformation of Rab GTPase. GTP hydrolysis and release of Pi, in turn, destabilises the Switch regions, relaxing them into the ‘off’ conformation. Consequently, the pronounced structural differences between the ‘on’ and ‘off’ states allow for selective recognition of Rabs by their regulatory proteins and downstream effectors (41). The G4 and G5 motifs confer the specificity of the GTPase to guanine nucleotides – in the G4 motif the conserved N and D residues interact with the nitrogenous base of the nucleotide (with D residue forming a strong bifurcated hydrogen bond with the nitrogen atoms in the guanine base), while the conserved A residue of the G5 box interacts with O6 of guanine base, supporting specific nucleotide recognition (46). The location of the G-domain, G boxes, Switch regions, and Hypervariable region in Rab3a was identified by NCBI BLAST and NCBI Conserved Domains tools (47,48). B) Crystal structure of the GppNHp-bound G-domain of Rab3a (PDB: 3RAB, residues 15-186). The G domain of Rab proteins consists of a six-stranded mixed β-sheet (β1-β6), flanked by five α-helices (α1-α5). The segments corresponding to the Switch I and II regions are shown in blue. The Figure was adapted from Dumas *et al.* (49).

The degree of conformational change between active GTP-bound and inactive GDP-bound forms varies from one Rab protein to another (41). The nucleotide-dependent conformational rearrangements mainly involve Switch I and Switch II regions, nevertheless, more dramatic, global conformational changes have been reported in the literature for example for Rab28 (50). In the inactive, GDP-bound state, Switch regions tend to be disordered and protein activation by GTP binding causes them to undergo major changes to adopt well-ordered, structured conformation necessary for interactions with effector proteins (41).

Despite high level of conservation of the G domain and the overall structural similarities between the members of Rab family, Rabs are diverse and carry out distinct cellular functions by their specific interactions with partner and effector proteins, owing to individual structural differences (41). A study by Pereira-Leal and Seabra led to the identification of conserved Rab Family motifs (F1-F5), which distinguish Rabs from other members of the Ras superfamily of proteins, as well as Rab subfamily (SF1-SF4) motifs, predicted to define the sites of interactions with Rab effector proteins and enabling categorisation of Rabs into distinct subfamilies (51). Structural comparison of different

Rab proteins in their active state revealed the highest degree of heterogeneity within the Switch regions and the α3/β5 loop (spanning the SF3 region), with little change to the other parts of the protein, which provides insight into the structural determinants of specific effector protein recruitment by Rabs, enabling them to perform their respective functions (40).

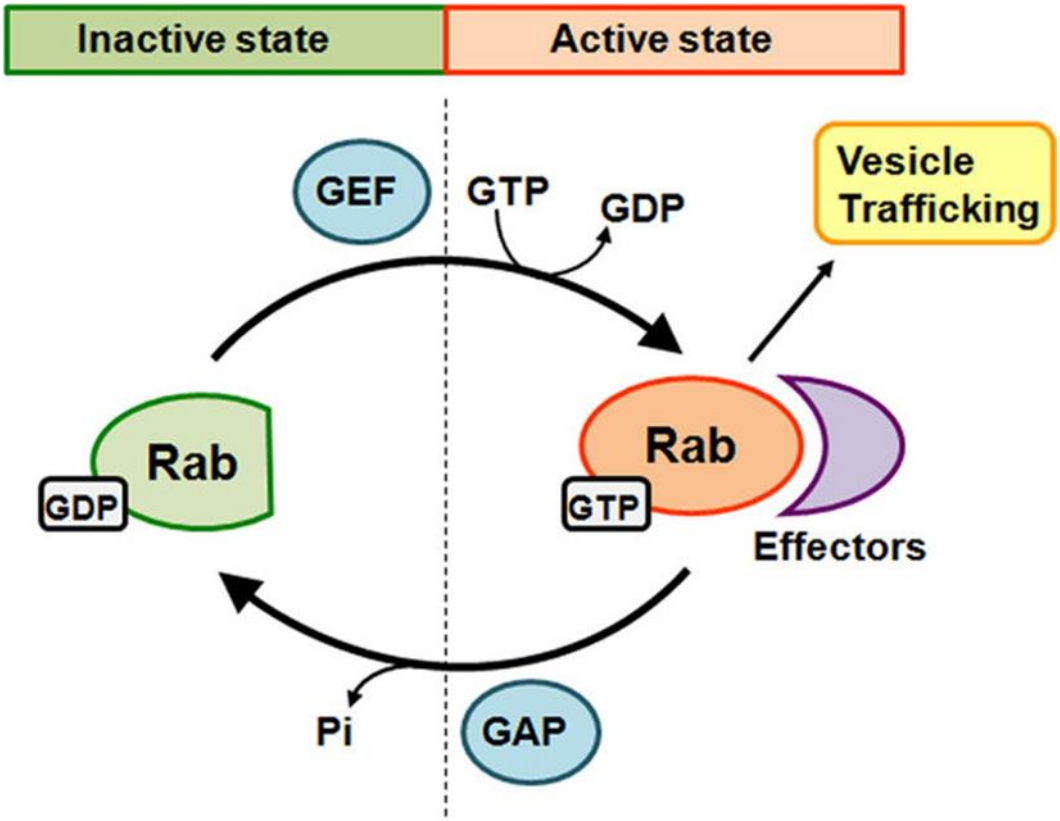

Figure 2 Rab GTPases activation cycle. Once attached to a membrane, Rab GTPases are activated by guanine nucleotide exchange factors (GEF) that catalyse the exchange of GDP to GTP. The GTP-bound, active form is then recognised by corresponding effectors (for example motor proteins) that carry out the diverse functions needed at each step of vesicular transport pathway (40). Rabs are then converted back to their inactive form by hydrolysis of GTP to GDP, the process facilitated by GTPase activating protein (GAP). As a result, Rab proteins can be recognised by GDP dissociation inhibitors (GDI) that remove them from the membrane and sequester them in the cytosol. GDI-displacement factor (GDF) then mediates directing Rab/GDI complex to specific membranes. Figure reprinted with permission from Wang *et al.* (52).

1.1.7. Control of Rab activity by regulatory proteins

Activity of Rabs in regulating membrane traffic is critically dependent on their ability to

reversibly associate with cellular membranes which is, in turn, enabled by prenylation, a lipid post-translational modification of the C-terminal cysteine residues (53). Rab prenylation motif typically contains two cysteine residues to which geranylgeranyl moieties are covalently attached by Rab geranylgeranyl transferase (RabGGT) (40).

Rabs are initially synthesised as soluble proteins in the cytosol. The newly translated Rab protein is recognised by Rab escort protein (REP) which presents it to RabGGT for the addition of one or, in most cases, two geranylgeranyl groups. Following prenylation and dissociation of RabGGT from the ternary Rab-REP-RabGGT complex, the prenylated, GDP-bound Rab is escorted by REP to its target membrane (54,55). Targeting of the Rab-REP complex to a specific membrane is mediated by the interaction with a membrane-localised guanine nucleotide exchange factor (GEF) which activates Rab by stimulation of the GDP-to-GTP exchange (56). Once Rab proteins are activated and stabilised on the membrane, their localisation may be further dictated by binding partners and effectors (57). Furthermore, the C-terminal hypervariable region of Rabs has also been demonstrated to contribute to membrane delivery specificity of certain Rab proteins (41,57).

1.1.8. Rab activation

Since the intracellular concentration of GTP is approximately ten times higher than of GDP, self-activation of GTPases can occur via spontaneous dissociation of GDP and subsequent intrinsic loading of GTP (58,59). Examples of atypical ‘fast cycling’ Rab GTPases characterised by weak intrinsic GTPase activity and high intrinsic nucleotide exchange activity, whose regulation appears to be independent from GEFs, have been described in the literature (59). For the majority of Rab GTPases, however, this process is typically slow and faster release of GDP is facilitated by GEFs which, by modifying the nucleotide-binding site of Rab, lower its nucleotide binding affinity and thus promote rapid loading of intracellular GTP, present in excess over GDP, and the consequent activation of the protein, enabling its interaction with effectors to execute specific functionalities (40)

1.1.9. Rab inactivation

Rab protein is then converted back to its inactive GDP-bound state, no longer capable of recruiting effectors, by GTPase activating protein (GAP). Finally, the inactive Rab is extracted from the membrane by GDP dissociation inhibitor (GDI) and stabilised and sequestered in the cytosol for another round of activation (40). On the new activation cycle, targeting of the Rab-GDI complex to specific membranes is mediated by GDI-

displacement factor (GDF). Self-inactivation of Rabs can occur through intrinsic GTPase activity by hydrolysing the bound GTP to GDP. Nevertheless, similarly to the GDP-to-GTP nucleotide exchange, intrinsic GTPase activity of small GTPases tends to be low (59), and therefore GTP hydrolysis has to be stimulated by GAPs to enable timely deactivation of Rabs, and switching off of the activities they coordinate.

Numerous Rab proteins, including Rab3, Rab15, Rab35 and Rab27, have been shown to be associated with WPBs and to regulate WPB exocytosis (39,60). The regulation of WPB exocytotic machinery and the role of selected Rab proteins in the process is depicted in Figure 3 (29). Markedly, Rab27 has a crucial role in mediating WPB anchorage to actin cytoskeleton to ensure complete maturation and multimerization of vWF prior to exocytosis (61). Other Rabs localising to WPBs include Rab33a and Rab37, whose functional importance at WPBs has not yet been elucidated (60), and a newly identified Rab46 that appears to play a pivotal role in regulation of context-dependent differential WPB exocytosis.

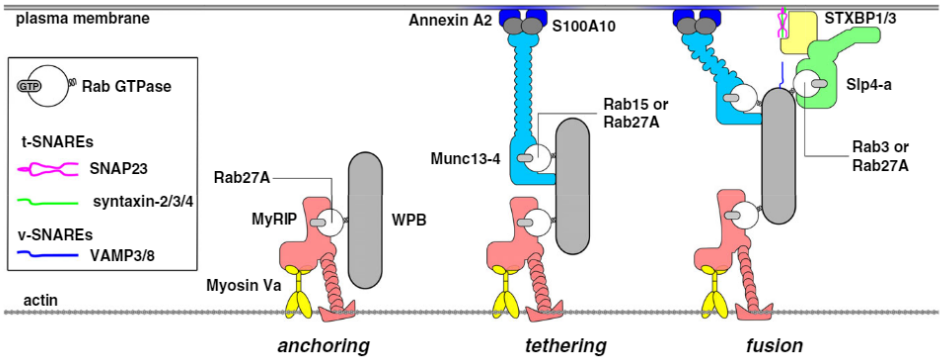

Figure 3 Exocytotic machinery of Weibel-Palade bodies. The different stages of regulated exocytosis of WPBs are mediated by Rab - effector complexes. A tripartite complex comprising Rab27a, MyRIP (Rab27a-specific effector) and Myosin Va mediates anchoring of WPBs to the cortical actin at the cell periphery to ensure complete maturation and multimerization of vWF prior to exocytosis. Secretagogue stimulation triggers an increased recruitment of the tethering factor Munc13-4 to WPBs and clustering of Munc13-4 at the sites of WPB-plasma membrane contact (62). Munc13-4, which can be recruited to WPBs in Rab-dependent (Rab27a/Rab15) or Rab-independent manner, supports acute WPB exocytosis by tethering WPBs to the fusion sites on the plasma membrane via the annexin A2-S100A10 complex. The Rab27a/Rab3-Slp4a complex docks WPBs and forms the link between the WPB and the plasma-membrane resident SNARE complex via members of the syntaxin-binding protein family (STXBP)

to mediate membrane fusion during WPB exocytosis. The figure was reprinted with permission from Schillemans *et al.* (29).

1.1.10. Rab46 GTPase

Rab46 (CRACR2A-a) is a novel member of Rab GTPase family recently described in endothelial cells and T-cells (43,63). Human CRACR2A (EFCAB4B) gene can be alternatively spliced to give two transcriptional variants (NM_032680.3 and NM_001144958.2). The short isoform of CRACR2A (CRACR2A-S, NP_116069.1) with a predicted molecular weight (MW) of 46 kDa, expression of which has not been demonstrated in endothelial cells (43), has been shown to regulate CRAC channel-mediated store-operated Ca^{2+} entry (SOCE) in T-cells, a process crucial for immune cells' proliferation and activation (64).

The long isoform of CRACR2A, Rab46, (NP_001138430.1) consists of 731 residues and is predicted to have a theoretical MW of 83 kDa. The endothelial variant of the protein (XP_006719084) contains 732 residues due to the presence of an extra serine residue (S425) between S424 and Q425. The short and long CRACR2A variants share 51% sequence identity, however, compared to the short isoform which contains two functional domains: EF-hand and coiled-coil, Rab46 is characterised by the presence of an additional Rab GTPase domain (Figure 4). Presence of both EF-hand and Rab domains suggests that Rab46 has a role in coupling changes in cytosolic Ca^{2+} concentration to cellular trafficking events in endothelial cells.

Structurally, Rab46 resembles two other non-canonical Rab GTPases, Rab44 and Rab45, that are both large-sized proteins (108 and 83 kDa respectively) characterised by presence of EF-hand, coiled-coil and Rab domains (42,44). Rab44 appears to have a role in osteoclast differentiation by modulating Ca^{2+} influx whereas Rab45 has been characterised as a self-associating GTPase and a potential tumour suppressor (42,44,65). Nevertheless, the number of studies of these non-canonical Rabs is limited and their exact modes of action and molecular mechanisms underlying their role in human diseases are unclear.

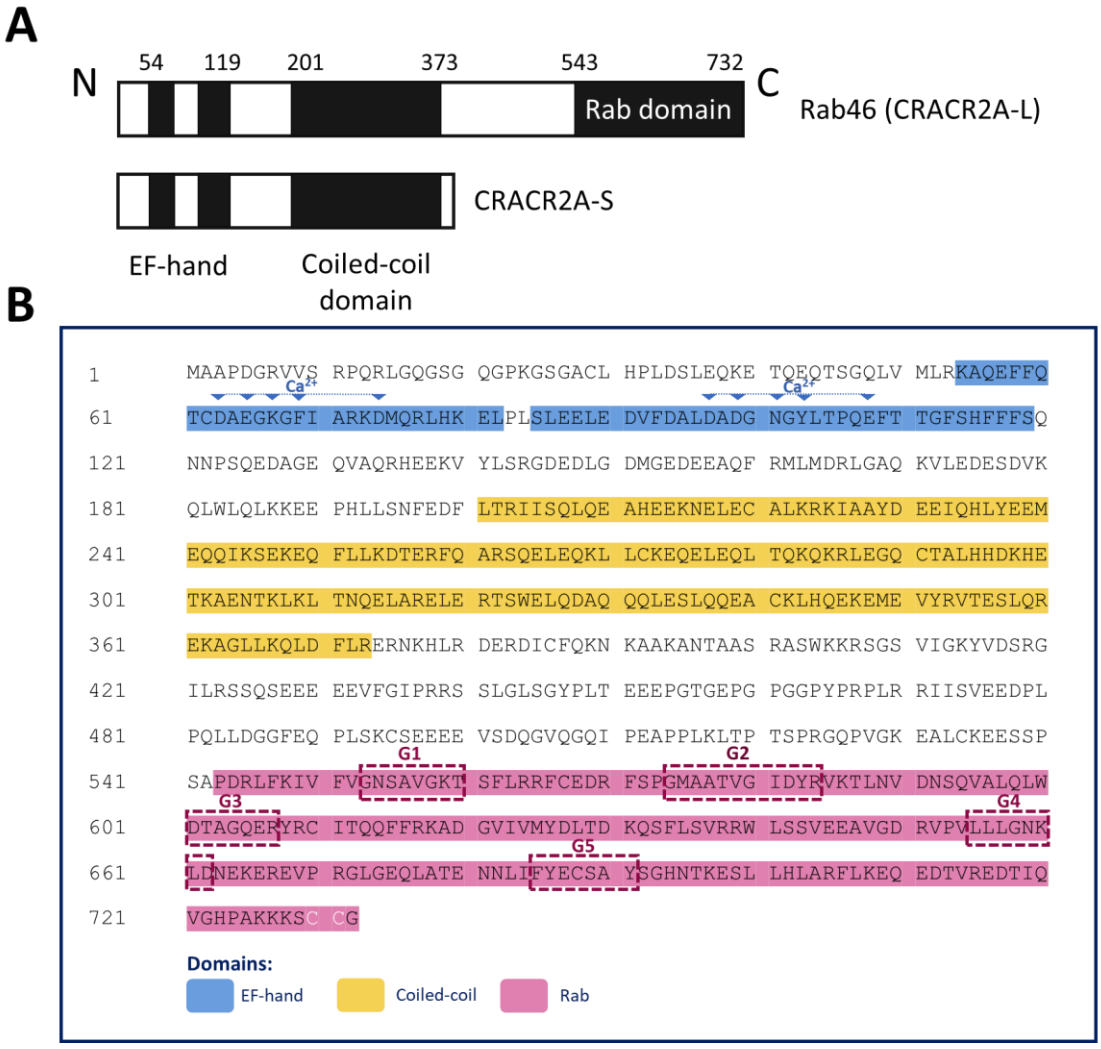

Figure 4 Sequence and functional domains of Rab46. (A) Comparison between Rab46 (CRACR2A-a) and short isoform of CRACR2A (CRACR2A-S). (B) Sequence of Rab46 and the locations of functional domains highlighting Ca^{2+} binding residues of the EF-hand domain and consensus sequences (G1-G5) of the Rab domain identified by NCBI BLAST and NCBI Conserved Domains tools (47,48). Two cysteine residues at the C terminus (in white) mark the prenylation site.

As presented in Figure 5, in addition to the three aforementioned functional domains, Rab46 is predicted to contain a long disordered region (A392-S538) linking the coiled-coil and Rab domains, spanning a large portion of proline-rich domain (PRD, M349 – S540), identified by Srikanth and colleagues (63). Disordered proline-rich motifs are common and often involved in protein-protein interactions (66). Indeed, in T-cells, interaction of Rab46 with GEF protein Vav1, crucial for its recruitment to the

immunological synapse following TCR stimulation, has been demonstrated to occur through the PRD of Rab46 (63).

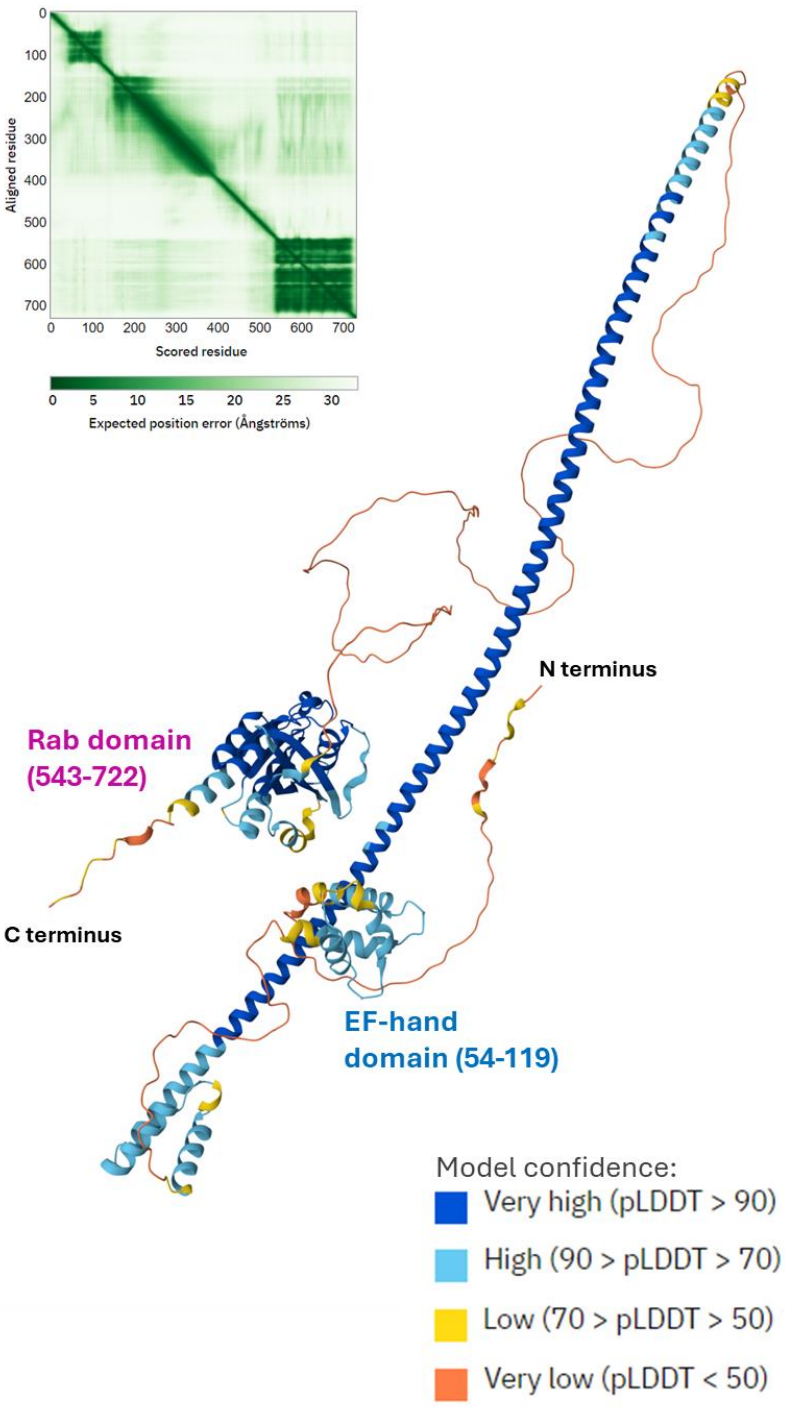

Figure 5 AlphaFold model of full length Rab46 structure. The AlphaFold-predicted structure of full length human Rab46 (AF-Q9BSW2-F1) is coloured according to pLDDT confidence values. The three clearly defined domains (EF-hand, Coiled-coil and Rab

domain) are separated by linkers with low-confidence scores. Predicted aligned error assessing the inter-domain accuracy of the model is shown in a box in the top left corner.

1.2. Functional domains of Rab46

1.2.1. EF-hand domain

EF-hand is the most common Ca^{2+}-binding motif. It is a helix-loop-helix motif, typically occurring as a pair of non-identical motifs to form a structurally stable EF-hand domain (67). Rab46 contains a single, N-terminal EF-hand domain and the locations of the predicted Ca^{2+}-binding residues are shown in Figure 4. As demonstrated by the recently reported crystal structure of Ca^{2+}-bound EF-hand domain of Rab46 (PDB: 6PSD), the protein possesses the expected helix-loop-helix fold (68). However, of the two EF-hand motifs of Rab46, only the second motif (EF2) has been shown to bind Ca^{2+} (68). In EF2 calcium binding loop of Rab46 (DADGNGYLTPQE), Ca^{2+} is coordinated by the side-chains of D97, D99, N101, E108, the main carbonyl oxygen of Y103, and a water molecule which is, in turn, coordinated by T105 (68). According to the traditional Ca^{2+}-binding loop nomenclature, these residues, involved in the pentagonal bipyramidal coordination of Ca^{2+} ion, correspond to positions X (D97), Y (D99), Z (N101), -Y (Y103), -X (T105) and -Z (E108) (68). Analysis of EF-hand motif sequences from different EF-hand containing proteins (Ca^{2+} binding and non Ca^{2+} binding) revealed that the amino acid residue crucial for Ca^{2+} binding at position -Z, coordinating the Ca^{2+} ion in a bidentate fashion, is substituted in the Ca^{2+}-binding loop (DAEGKGFIARKD) of the non-functional EF1 motif of Rab46 (Figure 6).

Since EF-hand proteins are involved in Ca^{2+} signalling, they participate in virtually all aspects of cell function (67). In most cases Ca^{2+}-binding to EF-hand proteins induces a major conformational change, although the extent of that change varies between EF-hand proteins and is intrinsic to their function, as conformational rearrangement is an important step in many Ca^{2+}-dependent processes (67). Based on their functions EF-hand proteins can be classified into two groups: Ca^{2+} sensors that transduce Ca^{2+} signals and Ca^{2+} signal modulators that play a role in regulation of duration of Ca^{2+} signals and the maintenance of Ca^{2+} homeostasis (69,70). One of the key differences between the proteins representing these two classes, aside from the distinct functions they perform, is the degree of conformational change they undergo upon Ca^{2+} binding. Numerous lines of evidence demonstrate that the degree of conformational rearrangement among EF-hand proteins can extend from very subtle and minor, as reported for Ca^{2+} buffering

protein calbindin, to substantial rearrangements involving complete reorientation of helices within EF-hand domain that occurs for example in calmodulin and other Ca^{2+} sensor proteins (71,72). Thus, compared to non-sensor EF-hand proteins, classical Ca^{2+} sensors typically show significantly more dramatic Ca^{2+}-induced changes. In calmodulin and related sensor proteins, apo forms typically occupy closed conformation characterised by antiparallel organisation of the helices of the EF-hand domain. Binding of Ca^{2+} induces rearrangements in the relative orientation of the helices to nearly orthogonal, which leads to transition to an open conformation and the exposure of large hydrophobic pocket that forms a site for target interaction.

The recent study by Lee and colleagues (68) revealed the role of EF-hand domain of Rab46 in modulating dynein motility, demonstrating its direct Ca^{2+}-dependent interaction with the conserved amphipathic helix of the dynein light intermediate chain-1 (LIC1, amino acid residues 433-458).

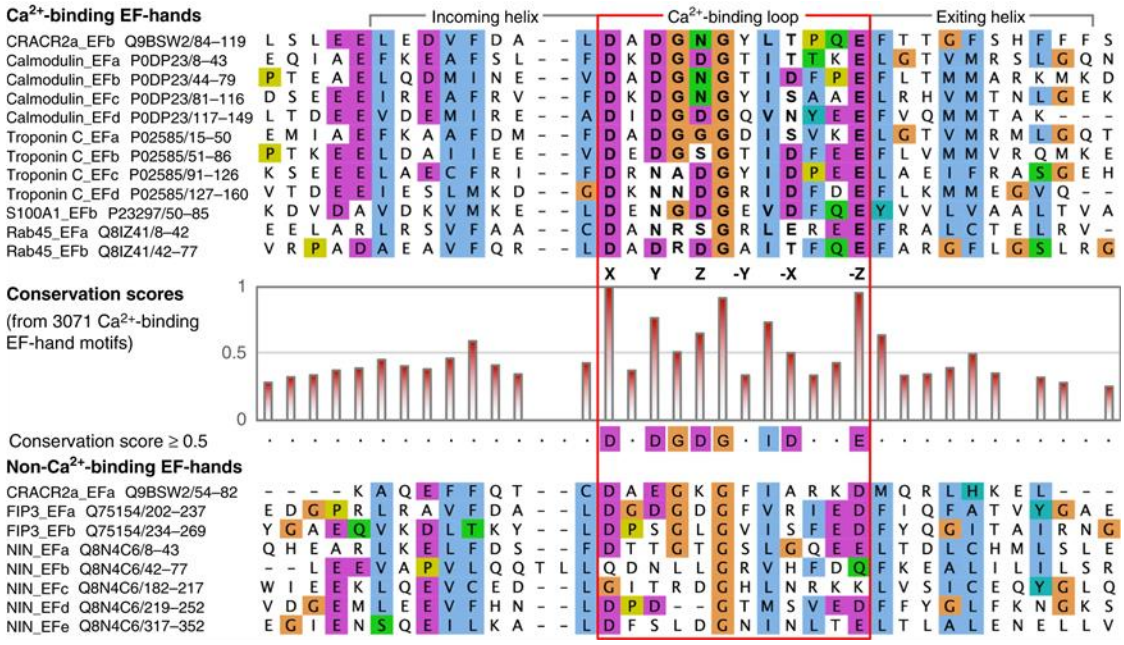

Figure 6 Conservation of Ca^{2+} ion-coordinating residues in Ca^{2+}-binding and non-Ca^{2+}-binding EF-hand motifs. The sequences of EF-hand motifs of Rab46 (CRACR2a) and other EF-hand containing dynein adaptors (FIP3, NIN and Rab45) were compared with the sequences of canonical EF-hands from calmodulin, skeletal muscle troponin C, and S100A1. The figure was reprinted with permission from Lee *et al.* (68).

1.2.2. Coiled-coil domain

Coiled-coils are highly abundant motifs often involved in mediating protein-protein

interactions (73). The presence of long Coiled-coil domains is a common feature of dynein adaptor proteins, and Coiled-coils of BICD2, BICDR1 and HOOK3 dynein adaptors have been demonstrated to be directly involved in recruiting dynein's tail to dynactin's filament to form a stable dynein-dynactin complex (74–76). Latest research suggests that Coiled-coil domain of Rab46 may be involved in binding of dynein in both endothelial cells and T-cells in, respectively, Ca^{2+}-independent and Ca^{2+}-dependent manner (33,74,77). Formation of a motile complex with the motor protein is, in turn, vital for initiation of vesicle trafficking along microtubules (33,74).

1.2.3. Rab domain

The analysis of amino acid sequence revealed that the C-terminal GTPase domain of Rab46 contains guanine nucleotide binding boxes (G1-G5) – i.e. consensus sequences conserved in Rab proteins (Figure 4), as well as the characteristic P-loop and Switch I and II regions (46,63). Additionally, as demonstrated by the superposition of the AlphaFold model of Rab domain of Rab46 with the crystal structure of Rab3a shown in Figure 7, the domain is predicted to have a typical GTPase fold, consisting of five α-helices and six β-strands interconnected by five loops responsible for guanine nucleotide and Mg^{2+} ion binding. The GTPase activity of wtRab46, assessed by Srikanth and colleagues by measuring its ability to hydrolyse GTP using a colorimetric malachite green reagent-based assay, confirmed that Rab domain of Rab46 was fully functional and that mutations of the catalytic residues (T559, Q604 and N658) result in impaired GTP binding and hydrolysis (63). The intrinsic GTP hydrolysis rate of the protein was, however, not determined.

1.3. Rab46 functions

In T-cells, Rab46 was shown to have a direct role in activation of the Ca^{2+} and JNK signalling following T-cell receptor (TCR) stimulation. Rab46 was found to be present in subsynaptic vesicles which, upon TCR stimulation, translocate from the Golgi into the immunological synapse to activate the mentioned signalling pathways which is, in turn, essential for the differentiation of T-helper cells (63). Moreover, recent research by Wang *et al.* further delineated the role of Rab46 in T-cells and provided evidence that Rab46 acts as a dynein adaptor that activates dynein-mediated transport in T-cells in a Ca^{2+}-dependent manner (74). The study by Lee and colleagues further supported Ca^{2+}-dependence of Rab46 interaction with dynein, demonstrating a direct interaction of the Ca^{2+}-bound EF-hand domain of Rab46 with a short helical fragment of dynein (68).

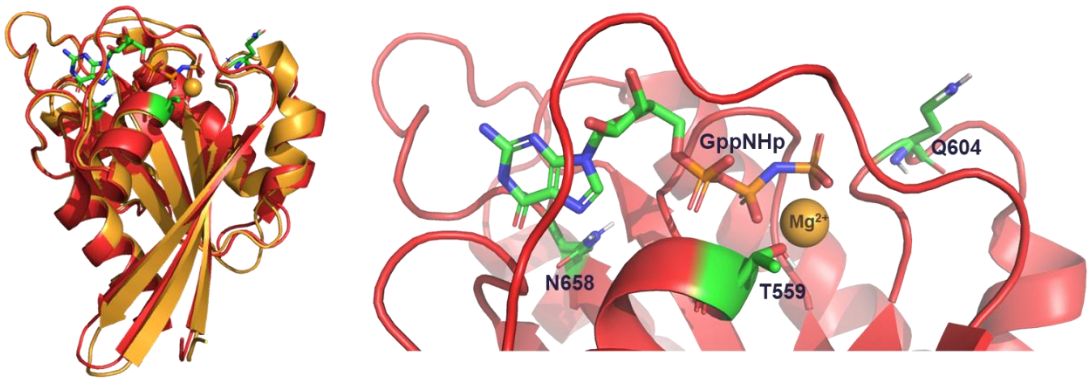

Figure 7 Superposition of AlphaFold model of the Rab domain of Rab46 and GppNHp-bound Rab3a crystal structure. AlphaFold model of the Rab domain of Rab46 (AF-Q9BSW2-F1, residues 542-709) and GppNHp-bound Rab3a crystal structure (PDB: 3RAB, residues 18-186) are shown in red and yellow, respectively. GppNHp and the residues involved in GTP binding and hydrolysis are shown in stick representation. Superposition of the two structures, and RMSD calculations were performed using PyMOL (78). RMSD = 0.733

Findings from McKeown laboratory demonstrated that Rab46 is a key regulator of differential WPB trafficking (33). In endothelial cells, Rab46 localizes to a subpopulation of WPBs devoid of P-selectin. The research has shown that, while stimulation with histamine causes mobilization of a subset of Rab46-negative WPBs carrying P-selectin to the cell surface, Rab46-positive WPBs storing angiopoietin-2 and other cargo irrelevant to inflammatory response are diverted to the microtubule organizing centre (MTOC), and this process is dependent on Rab46 (Figure 8, (33)). In this way, it is possible to obtain stimulus-appropriate inflammatory response whereas the release of prothrombotic factors can be minimised which, in turn, is important for maintaining healthy vasculature (33). In contrast to T-cells, in which Ca^{2+}-binding was demonstrated to be necessary for the interaction between Rab46 and dynein, in ECs Rab46 was found to interact with dynein in a Ca^{2+}-independent manner. Ca^{2+} binding to the EF-hand domain was, however, found to be required for the dispersal of Rab46-positive WPBs from the microtubule organising centre (33).

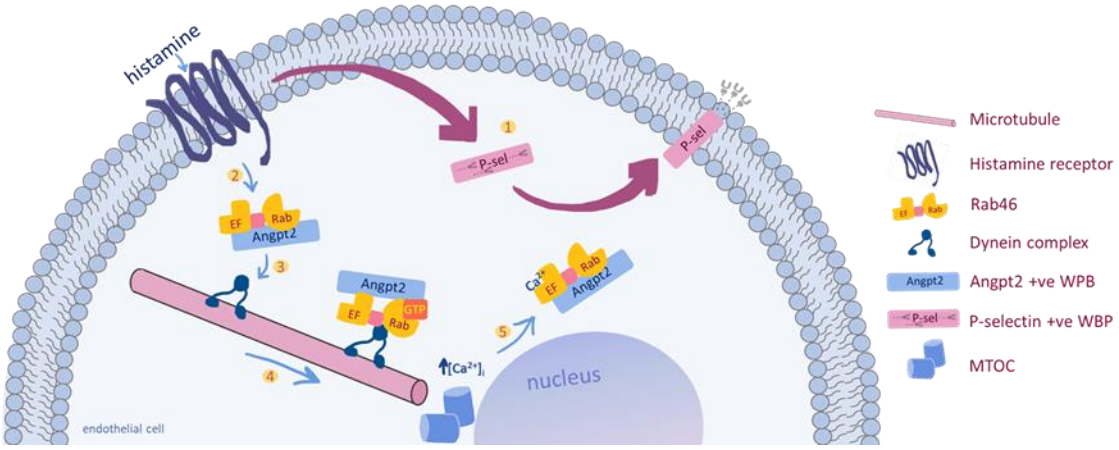

Figure 8 : Role of Rab46 in differential exocytosis of WPBs. (1) Following acute histamine stimulation, WPBs containing P-selectin are translocated to the cell surface in a Rab46-independent manner. (2) At the same time, stimulation with histamine triggers activation of Rab46 and (3) promotes interaction between GTP-bound Rab46 and dynein that is necessary to (4) evoke retrograde trafficking of a subset of WPBs containing angiopoietin-2 (Angpt2) and other non-inflammatory cargo, to the microtubule organising centre (MTOC). Subsequently, WPBs containing physiologically irrelevant cargo become anchored to the MTOC which minimizes the extensive response to vascular injury (5) Mobilization of intracellular Ca^{2+} induces Ca^{2+}-binding to the EF-hand domain of Rab46 and enables dissociation of the Rab46 positive WPBs from the perinuclear region. Figure adapted from Miteva et al. (33).

1.3.1. Rab46 in disease

The studies of the role of Rab46 in human disease are limited. Single nucleotide polymorphisms (SNPs) in CRACR2A gene (EFCAB4B) have been found to be associated with non-alcoholic fatty liver disease, heritable pulmonary arterial hypertension, as well as inflammatory conditions such as rheumatoid arthritis and periodontal disease (79–82). What is more, biallelic variants in CRACR2A (EFCAB4B) gene, resulting in one of the alleles encoding a point mutation in the coiled-coil domain of full length protein (E278D), and the other one encoding a truncated version of the protein, containing a point mutation in the EF-hand domain (R144G/E300*), have been linked to late onset combined immunodeficiency, characterised by loss of function in T-cells (83). The double mutation (R144G/E300*), resulting in the full length Rab46 no longer being expressed, has been demonstrated to lead to reduced cytokine production in T-cells as a consequence of profoundly decreased Ca^{2+} influx and impaired JNK

signalling (83,84).

Recently, three nonsynonymous SNPs in the EFCAB4B gene have been demonstrated to be associated with COVID-19 fatality (84). All three SNPs cause amino acid substitutions at positions showing high degree of conservation across species, which suggests an important role of the mutated residues in protein structure and function. Importantly, two of the identified SNPs, rs17836273 and rs36030417, result in missense mutations in EF-hand (A98T) and coiled-coil (H212Q) domains of Rab46, respectively. Since increased COVID-19 morbidity and mortality are linked to cardiovascular complications such as abnormal clotting, thrombosis and microvascular injury, inappropriate EC degranulation could be an important contributing factor determining the patient outcomes (84). Elevated levels of vWF, angiopoietin-2 and P-selectin reported in the sera of severely affected COVID-19 patients, have been found to be associated with in-hospital mortality (84–87). Exocytosis of these pro-thrombotic and pro-inflammatory factors from WPBs can be triggered by activated ECs as a result of endothelial damage associated with inflammation and hypoxia (86). Considering Rab46 function in both endothelial cell exocytosis and T-cell signalling, the missense variants identified in EFCAB4B may contribute to the pro-thrombotic and pro-inflammatory environment observed in severe COVID-19 cases through not yet elucidated molecular mechanisms.

1.4. Aim of the study

Storage of ready-made cargo primed for rapid release avoiding the need for transcription and translation is, indeed, a necessary strategy for vascular repair process. Nevertheless, it is also equally risky since, in addition to vWF, increased plasma levels of both P-selectin and angiopoietin-2 contribute to atherosclerotic plaque development and instability (88,89). Untimely, inappropriate and excessive endothelial activation and exocytosis may contribute to thrombotic disorders such as atherosclerosis which remains the major underlying cause of cardiovascular disease. Considering the role of Rab46 in regulation of differential WPB exocytosis, it is proposed that understanding of the 3D atomic structure and dynamics of Rab46 may contribute to the discovery of novel targets for therapeutic intervention in cardiovascular disease.

Given the anticipated flexible and dynamic nature of this large, multi-domain protein, Nuclear Magnetic Resonance (NMR) spectroscopy, which is particularly well-suited for characterisation of both structure and dynamics of proteins at atomic resolution (90), was chosen as the primary technique to explore the contribution of the agonist-induced local,

intradomain conformational changes, long-range interdomain interactions, and global structural rearrangements to the function of Rab46 GTPase.

Hypothesis: The dynamic regulation of Rab46 is dependent upon GTP hydrolysis and Ca^{2+} binding to distinct domains, and allosteric communication between these events affords potentially druggable intervention points.

Objectives:

1) Establish a recombinant *E. coli* expression system to clone, overexpress and purify distinct domains of Rab46.

2) Investigate structure, folding and cellular interaction partners of Rab46 using low-resolution structural and biophysical techniques.

3) Determine the domain structure, interactions and dynamics of Rab46 using high-resolution solution NMR and undertake studies to establish feasibility of NMR of Rab46 in endothelial cells or lysates.

Chapter 2

Materials and Methods

2.1. Molecular cloning

2.1.1. Polymerase Chain Reaction

A schematic representation of the cloning strategy is shown in Figure 9. pET32a-LIC and pET28a-LIC vectors used for bacterial expression of histidine-tagged recombinant proteins (Addgene plasmid #62310 and #26094 respectively) were a gift from Cheryl Arrowsmith. For the vector maps see Appendix. pET32a-LIC and pET28a-LIC plasmid DNA were digested with BseRI (NEB) to remove SacB gene stuffer sequence and create the following overhangs:

forward primer overhang (pET32a-LIC): 5' GTA TTT CCA GAG C --- 3'
forward primer overhang (pET28a-LIC): 5' GTT CCG CGT GGT AGT --- 3
reverse primer overhang (both vectors): 5' CAA GCT TCG TCA TCA --- 3'

Digest reactions containing 2 µL CutSmart® Buffer (NEB), 12 µL ddH2O, 5 µL plasmid DNA (0.5 µg) and 1 µL BseRI (NEB) were set up and incubated at 37ºC for 2 hr.

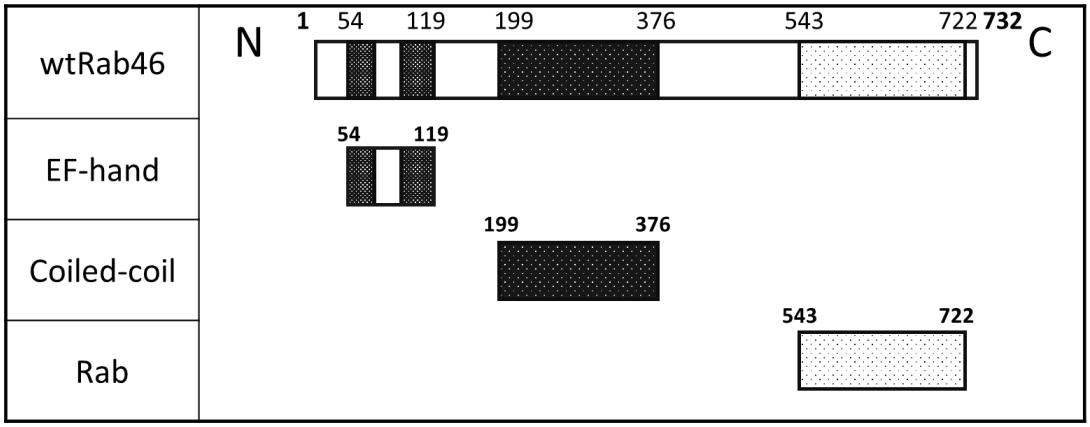

Figure 9 Schematic representation of cloning strategy. The diagram indicates fragments of protein coding sequence of Rab46 gene that were cloned into pET32a-LIC and pET28a-LIC expression vectors.

The primers used to amplify full length wtRab46 and its individual domains were designed so that ends homologous to overhangs left after BseRI digest were included in their sequence (see Table 1).

Table 1 Sequences of the PCR primers used for generation of constructs for protein expression.

Target	Sequence (5'-3')	Annealing temperature (ºC)	Product size (bp)
Full length Rab46 in pET32a-LIC	Fwd: GTATTTCCAGAGCGCTGCCCCTGACGG Rev: CAAGCTTCGTCATCAGCCACAGCAGGATTTCTTC	55	2196
EF-hand in pET32a-LIC	Fwd: GTATTTCCAGAGC GGCCAGCTAGTCAT Rev: CAAGCTTCGTCATCAGTTATTCTGGCTG	55	228
Coiled-coil in pET32a-LIC	Fwd: GTATTTCCAGAGCGACTTCCTGACCAGAA Rev: CAAGCTTCGTCATCAGTTCCTTTCCCTTAG	57	534
Rab in pET32a-LIC	Fwd: TTTCCAGAGCTCTGCCCCTGACCG Rev: CAAGCTTCGTCATCAGCCGACCTGAATGG	55	540
Full length Rab46 in pET28a-LIC	Fwd: GTTCCGCGTGGTAGTATGGCTGCCCCTGACG Rev: CAAGCTTCGTCATCAGCCACAGCAGGATTTCTTC	55	2196
EF-hand in pET28a-LIC	Fwd: GTTCCGCGTGGTAGTGGCCAGCTAGTCATG Rev: CAAGCTTCGTCATCAGTTATTCTGGCTG	55	228
Rab in pET28a-LIC	Fwd: GTTCCGCGTGGTAGTCCTGACCGGCTCTTCAAG Rev: CAAGCTTCGTCATCAGCCGACCTGAATGG	55	534
Full length Rab46 in pFastBac	Fwd: CTGTATTTTCAGGGCATGGCTGCCCCTGACGGG Rev: ACTTCTCGACAAGCTTCAGCCACAGCAGGATTTCTTC	55	2196

Each PCR reaction was performed in a final volume of 25 µL and contained: 12.5 µL CloneAmp HiFi PCR Premix (Takara Bio), 0.5 µL (350 ng) template DNA (wtRab46 in eGFP vector), 0.75 µL forward primer (0.3 µM), 0.75 µL reverse primer (0.3 µM) and 10.5 µL ddH_2O. The PCR temperature cycling conditions were as follows: initial denaturation at 98ºC for 1 min, 35 cycles of: denaturation at 98ºC for 10 s, annealing at 55ºC or 57ºC (as specified in Table 1) for 5 s, and elongation at 72ºC for 2 min, with the last cycle

followed by a 4-minute final extension step at 72ºC.

To remove any remainder methylated template DNA, PCR products were digested with 2 µL FastDigest DpnI enzyme (Thermo Scientific™) at 37ºC for 30 min.

2.1.2. Gel extraction

To ensure purity of the linearised vector and the individual inserts prior to In-Fusion reaction, DNA samples were purified using NucleoSpin® Gel extraction kit (Macherey-Nagel). The samples were run on a low-percentage agarose gel (0.6%) at 55 V for 1 hr to prevent melting of the gel which could lead to DNA denaturation. DNA bands were excised on a portable UV transilluminator, and purification was performed following the manufacturer's guidelines. Concentration of the purified DNA samples was estimated by measuring the absorbance at 260 nm on NanoDrop™ spectrophotometer.

2.1.3. In-Fusion reaction

In order to achieve optimal cloning efficiency, different insert to vector ratios were tested for each construct. The amount of insert DNA to be used in the In-Fusion reactions to obtain the desired insert to vector ratios was calculated using the following equation:

$$\frac{\text{amount of vector DNA in the reaction (ng)} \times \text{size of an insert (kb)}}{\text{size of the vector (kb)}} \times \frac{\text{insert}}{\text{vector}}$$

A 3:1 insert to vector ratio was used for full length wtRab46, Rab, and Coiled-coil domains, whereas a 5:1 ratio was used for the EF-hand domain due to its small size. Each In-Fusion reaction contained 2 µL of In-Fusion HD Enzyme Premix, 75 ng of purified plasmid DNA and an appropriate amount of insert DNA (calculated as described above). Reactions were adjusted to a final volume of 10 µL with ddH_2O and incubated at 50ºC for 15 min.

2.1.4. Transformation and isolation of plasmid DNA

The constructs obtained from In-Fusion reactions were transformed into Stellar competent cells (Takara Bio) following the manufacturer's protocol. 5 mL bacterial cultures in selective LB media containing 100 µg/mL ampicillin (pET32a-LIC) or 50 µg/mL kanamycin (pET28a-LIC), inoculated with a single fresh bacterial colony picked from the transformation plates for each construct, were grown overnight at 37ºC with shaking at 250 rpm. Overnight cultures were harvested by a 30-minute centrifugation at 4,000 × g and plasmid DNA was isolated using QIAprep Miniprep Kit (Qiagen®) as per manufacturer's instructions.

Successful cloning into pET28a-LIC and pET32a-LIC vectors was confirmed by a

diagnostic digest using XbaI restriction enzyme (Promega). Digest reactions containing 10 µL (0.5 – 1 µg) plasmid DNA, 0.5 µL (5 units) XbaI enzyme, 2 µL Buffer D (Promega) and 7.5 µL ddH2O were incubated at 37ºC for 2 hrs. Digested samples were analysed by agarose gel electrophoresis.

Additionally, the DNA sequences of recombinant plasmids were verified by Sanger sequencing with 'T7' forward and 'T7 Term' reverse universal primers (GENEWIZ). Due to its large size, full length wtRab46 construct in pET32a-LIC and pET28a-LIC vectors was sequenced with an additional forward sequencing primer complementary to the 5' end of Rab46 DNA sequence (see Appendix). Having confirmed that the DNA sequences of the expression constructs were correct, 250 mL bacterial cultures transformed with the constructs were grown overnight at 37ºC with shaking at 250 rpm, and plasmid DNA was extracted using EndoFree Plasmid Maxi kit (Qiagen®) following manufacturer's guidelines. The resulting amplified plasmid DNA was once more inspected by Sanger sequencing prior to protein expression attempts to ensure no unwanted mutations were introduced during the cloning process.

2.1.5. Site-directed mutagenesis

Site-directed mutagenesis (SDM) was primarily used to introduce amino acid substitutions in the EF-hand domain to generate constructs for expression of negative control protein samples for the calcium-binding studies. SDM was also used to modify the domain boundaries of Rab domain by deletion of a stretch of C-terminal residues by means of inverse PCR, in order to generate alternative expression constructs to address the issues with Rab domain solubility. Custom forward and reverse primers were designed to be complementary to the sites within the DNA sequence where the mutations were desired and included modifications to either substitute or delete parts of the resulting amino acid sequence. Sequences of mutagenic primers are shown in Table 2.

Table 2 Sequences of the mutagenic primers

Target	Sequence (5'-3')
$EF^{97DAD>AAA}$	Fwd: GTGTTTGATGCCCTGGCTGCTGCTGGCAATGGCTATCTG Rev: CAGATAGCCATTGCCAGCAGCAGCCAGGGCATCAAACAC
EF^{E108Q}	Fwd: TATCTGACCCCACAGCAGTTCACTACTGGATTT Rev: AAATCCAGTAGTGAACTGCTGTGGGGTCAGATA
Rab (aa 543-705)	Fwd: TGATAGTAGTGATGACGAAGCTTGCGGC Rev: CTACTATCACCTGGCCAGATGGAGCAG
Rab (aa 543-710)	Fwd: TGATAGTAGTGATGACGAAGCTTGCGGC Rev: CTACTATCATTGCTCCTTGAGGAACCTG

For generation of the majority of the constructs for mutant EF-hand domain expression, wtEF-hand domain in pET32a-LIC vector was used as a template DNA in PCR reactions. For the generation of the $EF^{97DAD>AAA,E108Q}$ mutant, EF^{E108Q} mutant construct in pET32a-LIC vector was used as a template in the PCR reaction with $EF^{97DAD>AAA}$ mutagenic primers. Site-directed mutagenesis of Rab domain was performed using Rab domain in pET32a-LIC plasmid as a template DNA in the reactions.

To minimise the risk of introducing unwanted secondary mutations, PCR mutagenesis was performed using high fidelity Phusion (Thermo Scientific®) DNA polymerase. PCR reactions containing 1 µL template DNA (150 ng), 0.5 µL forward primer (1 µM), 0.5 µL reverse primer (1 µM), 1 µL dNTPs (200 µM), 2 µL DMSO (4% (v/v)), 10 µL 5 × Phusion HF reaction buffer (Thermo Scientific®), 1 µL Phusion DNA polymerase and 34 µL ddH_2O were set up. Due to initially encountered issues with primer annealing, higher percentage of DMSO (8% (v/v)) was used for PCR reaction of the EF^{E108Q} mutant. The following PCR temperature cycling conditions were used for SDM: initial denaturation at 97ºC for 1 min, 18 cycles of denaturation at 97ºC for 30 s, annealing at 60ºC, and elongation at 72ºC for 3.5 min, with the last cycle followed by a 7-minute final extension step at 72ºC. Prior to transformation into competent *E. coli* cells, the PCR products were treated with Dpn1 to digest remainder template DNA as described for the generation of pET32a-LIC and

pET28a-LIC constructs above. Following PCR and Dpn1 digest, the mutated constructs were transformed into XL10-Gold® ultracompetent cells (Agilent) using a standard heat-shock protocol. Mutated plasmid DNA was amplified in 5 mL overnight *E. coli* cultures and isolated using QIAprep Miniprep Kit (Qiagen®) following manufacturer's instructions. Successful mutagenesis was confirmed by Sanger sequencing with 'T7' forward and 'T7 Term' reverse universal primers (GENEWIZ).

2.1.6. Full length Rab46 construct generation for insect cell expression

To generate the construct for insect cell expression of His_6-tagged full length Rab46, a gene encoding wtRab46 (aa 1-732) was inserted into pFastBacHT A vector (Invitrogen; for the vector map see Appendix). pFastBacHT A plasmid DNA was digested with SfoI and HindIII enzymes to remove the Multiple Cloning Site (MCS) region to be replaced with the DNA sequence encoding full length wtRab46, and create the following overhangs:

forward primer overhang after digest with SfoI 5' CTGTATTTTCAGGGC --- 3'

reverse primer overhang after digest with HindIII 5' AGCTTGTCGAGAAGT --- 3'

Digest reaction containing 5 µL CutSmart® Buffer (NEB), 33 µL ddH2O, 10 µL plasmid DNA (1 µg), 1 µL (10 units) SfoI (NEB) and 1 µL (20 units) HindIII (NEB), were incubated at 37ºC for 1 hr.

Similarly as in the case of other expression constructs generated using In-Fusion cloning method (Takara Bio), FL wtRab46 to be inserted in pFastBac HT A vector was amplified by PCR using primers designed so that ends homologous to overhangs left following digest with SfoI and HindIII restriction endonucleases were included in their sequence (Table 1). PCR reactions containing 12.5 µL CloneAmp polymerase, 0.5 µL (0.2 µM) forward primer, 0.5 µL (0.2 µM) reverse primer, 1 µL (250 ng) template DNA (wtRab46 in pET32a-LIC vector), 1 µL (4% (v/v)) DMSO and 9.5 µL ddH_2O were set up. Temperature cycling conditions were unchanged from those described in section 2.1.1 for the generation of constructs for bacterial expression using the CloneAmp HiFi PCR Premix (Takara Bio).

Purification of the linearised vector and insert DNA samples prior to In-Fusion reaction was performed using NucleoSpin® Gel extraction kit (Macherey-Nagel). In-Fusion reaction was performed as described above for the bacterial expression constructs, using a 3:1 insert to vector ratio. The reaction product was transformed into Stellar competent cells (Takara Bio) and plasmid DNA was isolated using QIAprep Miniprep Kit

(Qiagen®) following manufacturer's instructions. Successful cloning of the insert into pFastBacHT A vector was confirmed by Sanger sequencing using the universal pFastBacF and pFastBacR primers (GENEWIZ; for the primer sequences see Appendix), and the absence of any unwanted mutations in the protein coding sequence was confirmed by sequencing with an additional, Rab46-specific forward primer (see Appendix).

2.2. Generation of P1 and P2 recombinant baculovirus

2.2.1. Recombinant bacmid generation

Bacmid DNA was generated following the protocol described in the Bac-to-Bac® manual (Invitrogen, Life Technologies). Briefly, 100 µL of DH10Bac™ competent cells were transformed with 1 ng of Rab46 pFastBac plasmid DNA using a standard heat-shock protocol. Following addition of 900 µL of SOC medium, the cells were incubated for 4 hrs at 37ºC with shaking at 225 rpm and plated on selective LB agar plates containing 50 µg/mL kanamycin, 7 µg/mL gentamicin, 10 µg/mL tetracycline, 100 µg/mL Bluo-Gal and 40 µg/mL IPTG. The plates were incubated at 37ºC for 48 hrs, after which time white colonies (containing the recombinant bacmid) were restreaked on fresh, selective LB agar plates for blue-white screening (antibiotics, Bluo-Gal and IPTG concentrations indicated above), and incubated at 37ºC for 48 hrs to confirm the white phenotype of the picked colonies, and avoid picking the ones containing the unaltered, empty bacmid. True white colonies were then used to inoculate 2 mL LB cultures containing 50 µg/mL kanamycin, 7 µg/mL gentamicin, and 10 µg/mL tetracycline. Following the overnight incubation at 37ºC with shaking at 250 rpm, the cultures were harvested by centrifugation at 9,000 × g for 15 minutes and the bacmid DNA was isolated using the PureLink™ HiPure Plasmid Miniprep kit (Invitrogen) as per the instructions provided in the Bac-to-Bac® manual.

Successful bacmid transposition with the Rab46 pFastBac construct was confirmed by PCR using pUC/M13 universal primers (for the primer sequences see Appendix). PCR reactions containing 1 µL (100 ng) template DNA, 25 µL OneTaq® 2 × Master Mix, 1.25 µL (0.25 µM) forward pUC/M13 primer, 1.25 µL (0.25 µM) reverse pUC/M13 primer, and 21.5 µL ddH_2O were set up for each of the recombinant bacmid DNA samples. The following PCR temperature cycling conditions were used: initial denaturation at 94ºC for 30 s, 30 cycles of: denaturation at 94ºC for 30 s, annealing at 55ºC for 45 s, and elongation at 68ºC for 5 minutes, with the last cycle followed by a 5–minute final extension at 68ºC. PCR products were analysed by agarose gel electrophoresis to confirm their size matched the expected one of 4629 bp (corresponding to 2430 bp

bacmid, 2196 bp FL Rab46 insert and 3 bp stop codon).

2.2.2. *Sf9* cell culture

Spodoptera frugiperda (*Sf9*) cells were cultured in suspension in Gibco™ Sf-900™ II Serum Free Medium at 27ºC, with shaking at 120 rpm. Cell densities were maintained between 0.5×10^6 and 4.5×10^6 cells/mL, with cell count and viability (determined by trypan blue exclusion) regularly monitored on the Countess™ II FL automated cell counter (Life Technologies). *Sf9* cells were used until passage number 30.

2.2.3. P1 virus stock

To generate P1 FL Rab46 baculovirus, 5 mL of *Sf9* cells (at passage number 18) were seeded in a T25 flask at 1.0×10^6 cells/mL and incubated static at 27ºC for 30 mins to allow time for the cells to adhere to the bottom of the flask. For each transfection, 15 µL Bacmid DNA at 500 ng/µL were mixed with 300 µL of pre-warmed Sf-900 Serum Free media (Gibco™). In a separate sterile Eppendorf tube, 10 µL of Cellfectin™ II Reagent (Gibco™) were added to 100 µL pre-warmed Sf-900 Serum Free media (Gibco™). The bacmid DNA mix was added dropwise to the transfection reagent mix and incubated for 30 mins at room temperature. The transfection mix was then added dropwise to adherent *Sf9* cells. Cells were incubated at 27ºC for 24 hrs, after which time the media were replaced with 5 mL of fresh Sf-900 II Serum Free Medium (Gibco™). The transfected cells were incubated static at 27ºC for another 6 days, and the recombinant baculovirus released to the culture medium was harvested 7 days post-transfection by centrifugation of the culture medium at 500 × g for 5 mins. The cell pellet was discarded and the cleared supernatant containing the recombinant baculovirus (P1 stock) was collected and stored at 4ºC.

2.2.4. P2 virus stock

To amplify the obtained viral stock, 300 mL of *Sf9* cells at 1.5×10^6 cells/mL were infected with 1 mL of P1 Rab46 baculovirus. The infected cells were incubated at 27ºC with shaking at 120 rpm. Cells were monitored for signs of viral infection i.e. growth arrest and increased cell diameter on the Countess™ II FL automated cell counter. P2 virus was harvested 72 hours post-infection by centrifugation at 500 × g for 5 minutes, collecting the cleared supernatant. P2 baculovirus stock was supplemented with 2% heat-inactivated FBS for long-term storage, and stored protected from light at 4ºC.

2.2.5. Baculovirus titre estimation using droplet digital PCR

Recombinant baculoviral DNA was extracted from P2 Rab46 baculovirus stock sample for PCR analysis using Monarch genomic DNA purification kit, following the ‘Mammalian

Whole Blood (non-nucleated)' protocol (NEB). Viral DNA isolated from 100 µL of P2 baculovirus sample was serially diluted 10-fold from 10^{-1} to 10^{-7} prior to analysis by droplet digital PCR (ddPCR).

ddPCR was set up following ddPCR EvaGreen Supermix manual (Bio-Rad). Each reaction, performed in a final volume of 20 µL, contained 10 µL of 2 x QX200 ddPCR EvaGreen Supermix, 2 µL (100 nM) of forward GP64 primer, 2 µL (100 nM) of reverse GP64 primer, 2 µL of DNA template isolated in the previous step, and 4 µL RNase/DNase-free water. Reaction mixtures were prepared in a 96-well plate format for the automated droplet generation using QX200 Droplet Generator (Bio-Rad) as specified in the manufacturer's protocol, and the following cycling programme was performed: 95ºC for 5 mins, 40 cycles of: 95ºC for 30 s, and 60ºC for 1 min; followed by 4ºC for 5 min and 90ºC for 5 min.

Following thermal cycling, the droplets were quantified on the QX200 Droplet Reader (Bio-Rad). Data were analysed with QuantaSoft Software (Bio-Rad).

2.2.6. Titre estimation using RT qPCR

Viral titre was additionally quantified using baculoQUANT ALL-IN-ONE kit (Oxford Expression Technologies) following the manufacturer's protocol. DNA for the generation of a standard curve for the RT qPCR analysis was prepared by serial 10-fold dilution (virus ranging from 10^7 to 10^3 pfu/mL) of the internal viral standard (stock at 10^8 pfu/mL). Baculoviral DNA was then extracted from the pre-diluted viral standards, as well as from the undiluted and diluted (10-fold) recombinant budded P2 Rab46 baculovirus. 80 µL of each sample was spun down at 16,000 × g for 5 minutes to pellet the virus. Following careful removal of the supernatant, the viral pellet was resuspended in 20 µL of lysis buffer and mixed by vortexing. Viral DNA was then extracted using the following lysis programme on the thermal cycler: 65ºC for 15 min, 96ºC for 2 min, 65ºC for 4 min, 96ºC for 1 min, 65ºC for 1 min, and 96ºC for 30 sec.

PCR reactions containing 12.5 µL Low ROX Mix (Oxford Expression Technologies), 7.5 µL RNase free water, 3 µL probe/primers mix (Oxford Expression Technologies), and 2 µL of the isolated viral DNA were set up in a 96-well PCR plate. RNase free water and positive control viral DNA provided in the kit were used for setting up negative, and positive control reactions respectively. All samples were analysed in triplicate. RT qPCR was performed on the Rotor-Gene Q PCR cycler (Qiagen) using the following PCR cycle programme conditions: enzyme activation at 95ºC for 10 minutes and 40 cycles of denaturation at 95ºC for 10 s, and annealing/extension at 60ºC for 60 s. The equation

obtained from plotting the standard curve (Ct value against the virus quantity in pfu/mL) was used to calculate the titre of recombinant Rab46 baculovirus.

2.3. Protein Expression

2.3.1. Small-scale bacterial expression

The expression constructs in pET32a-LIC and pET28a-LIC vectors were transformed into multiple strains of competent bacteria (as detailed in Table 10) using standard heat-shock protocol, and plated on LB agar containing appropriate combination of antibiotics (see Table 3). 2 mL starter cultures were set up in selective LB media and grown at 37ºC with shaking at 220 rpm in a 96-well deep well plate covered with a gas permeable adhesive seal. Following the overnight incubation, 200 µL of the cultures were used to inoculate 1.8 mL cultures in selective LB media in a deep well plate. Cultures were grown at 37ºC with shaking at 220 rpm until OD_{600} reached ~0.6. Protein expression was then induced with 0.4 mM IPTG, and the cultures were incubated overnight at 25ºC with shaking at 180 rpm.

Bacterial cells were harvested by a 10-minute centrifugation at 4,000 × g. Clarified media were discarded, and protein purification method was performed on the Microlab® STAR Hamilton robot using MagneHis protocol (Promega). Following purification, protein samples were mixed with 2 × protein loading dye containing 10 mM DTT, boiled at 95ºC for 3 min, and analysed by SDS-PAGE. Polyacrylamide gels stained with InstantBlue™ (Expedeon) were imaged on the G:BOX gel documentation system (Syngene).

Table 3 *E. coli* strains tested for expression of recombinant Rab46 proteins

Bacterial strain	Antibiotics for expression of pET32a-LIC constructs	Antibiotics for expression of pET28a-LIC constructs
BL21 (DE3)	Ampicillin (100 µg/mL)*	Kanamycin (50 µg/mL)
Rosetta (DE3)	Ampicillin (100 µg/mL) Chloramphenicol (34 µg/mL)	Kanamycin (50 µg/mL) Chloramphenicol (34 µg/mL)
Rosetta 2 (DE3)	Ampicillin (100 µg/mL) Chloramphenicol (34 µg/mL)	Kanamycin (50 µg/mL) Chloramphenicol (34 µg/mL)
SHuffle T7 Express	Ampicillin (100 µg/mL) Spectinomycin (50 µg/mL)	Kanamycin (50 µg/mL) Spectinomycin (50 µg/mL)
Lemo21 (DE3)	Ampicillin (100 µg/mL) Chloramphenicol (34 µg/mL)	Kanamycin (50 µg/mL) Chloramphenicol (34 µg/mL)
ArcticExpress (DE3)	Ampicillin (100 µg/mL) Gentamicin (50 µg/mL)	Kanamycin (50 µg/mL) Gentamicin (50 µg/mL)
*Working concentration of each of the required antibiotics is given in the brackets.		

2.3.2. Batch purification of full length Rab46 expressed in bacteria

Due to no evidence of expression of full length wtRab46 at a 2 mL scale in the tested *E. coli* strains, the expression of the construct in pET32a-LIC vector was further tested at a 100 mL scale in Rosetta (DE3), Rosetta 2 (DE3), SHuffle T7 Express and ArcticExpress (DE3) cells. 10 mL overnight cultures were grown in LB media containing 100 µg/mL ampicillin and appropriate antibiotic for selection of each of the tested strains (see Table 3). Overnight starter cultures were inoculated into 100 mL LB media containing appropriate antibiotics. Protein expression in Rosetta, Rosetta 2 and SHuffle T7 Express strains was induced as previously described for the small-scale expression test with 0.4 mM IPTG, followed by incubation of the bacterial cultures at 18ºC overnight with shaking at 180 rpm. ArcticExpress cells inoculated with the overnight starter culture were grown for 3 hrs at 30ºC after which time, regardless of the obtained OD_{600}, expression was induced with 0.4 mM IPTG, and the cultures were incubated at 12ºC overnight with shaking at 180 rpm. The following morning cells were harvested by centrifugation at 4,000 × g for 20 mins. The resulting pellets were resuspended in 5 mL of lysis/wash buffer (50 mM Tris (pH 7.6), 300 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol and 0.075% (v/v) β-mercaptoethanol) and sonicated on ice for 1 min 40 s (10 cycles of 10 s on and 20 s off) at 60% amplitude using a probe sonicator (Fisher Scientific). The lysate of each sample was split into 1.5 mL Eppendorf tubes and clarified by centrifugation at top speed in a chilled benchtop centrifuge for 30 mins. The resulting supernatant was

incubated for 1.5 hrs with 100 µL (bed volume) Ni-NTA resin (Qiagen) pre-equilibrated in the lysis/wash buffer. Following, the incubation resin was collected by centrifugation at 500 × g for 1 min and washed twice with the wash buffer for 5 mins. Bound protein was eluted with a series of elution buffers (50 mM Tris (pH 7.6), 300 mM NaCl, 5% (v/v) glycerol, 0.075% (v/v) β-mercaptoethanol) containing increasing concentration of imidazole (80 mM, 160 mM, 240 mM, 320 mM, and 400 mM). Elution fractions were mixed with the 2 × protein loading dye, boiled at 95ºC for 3 mins, and analysed by SDS-PAGE.

2.3.3. Large-scale bacterial expression

Having found the optimal expression conditions for each of the target proteins, protein expression was performed large-scale. Proteins were typically expressed at a 2 L scale. 100 mL starter cultures in LB media containing the appropriate antibiotics (as detailed in Table 3) were grown at 37ºC with shaking at 220 rpm overnight. The following morning every 1 L of selective LB was inoculated with 20 mL of the overnight starter culture, and grown at 37ºC until OD_{600} reached ~0.6 when protein expression was induced with 0.4 mM IPTG. Cultures were incubated overnight at 18ºC or 25ºC (as indicated in Table 10) with shaking at 180 rpm and harvested by centrifugation (4,000 × g, 4ºC, 30 min) the following morning. Bacterial pellets were either processed immediately or stored at -20ºC.

2.3.4. Labelled expression of EF-hand proteins for NMR

Uniformly ^{15}N and ^{13}C-labelled wtEF-hand was expressed in BL21 (DE3) cells at a 2-3 L scale in M9 minimal media containing 42.3 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl and 18.7 mM NH_4Cl, supplemented with 1 mM $MgSO_4$, 0.1 mM $CaCl_2$ and 0.1% w/v glucose. 100 mL starter culture was set up in LB containing 100 µg/mL carbenicillin and, following an overnight incubation at 37ºC, used to inoculate 100 mL M9 starter culture. Large 500 mL M9 cultures, containing ^{15}N-NH_4Cl (Cambridge Isotope Laboratories) and, in case of production of double-labelled protein samples, U-^{13}C-glucose (Cambridge Isotope Laboratories) as the sole sources of nitrogen and carbon respectively, were inoculated with 5 mL of overnight M9 starters. Protein expression was induced at OD_{600} of 0.6 using 0.4 mM IPTG as described above for the production of unlabelled proteins. Cultures were incubated overnight at 25ºC with shaking at 180 rpm and harvested by centrifugation (4,000 × g, 4ºC, 30 min) the following morning.

2.3.5. Optimisation of ^{15}N-labelled Rab domain expression

Due to issues encountered with the expression of ^{15}N-labelled Rab domain in the standard M9 minimal media, the protocol for the labelled expression of the domain required optimisation. A small-scale expression test of labelled Rab domain was performed using Rosetta 2 cells (DE3) at a 10 mL scale in M9 medium containing 42.3 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl and 18.7 mM NH_4Cl, supplemented with 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.2% (w/v) glucose, 100 µg/mL carbenicillin, and 34 µg/mL chloramphenicol. Different additives to the growth medium including 5-10% LB broth, 0.5% (w/v) Celtone Base powder (Cambridge Isotope Laboratories), 1 × BME vitamins (Sigma-Aldrich), and 1 × trace elements (3.6 µM $FeSO_4$, 30.6 µM $ZnSO_4$, 1.6 µM $CuSO_4$, 0.5 µM $MnSO_4$, 40.5 nM $(NH_4)_6Mo_7O_{24}$ and 0.3 µM $Na_2B_4O_7$) were tested. Protein expression was induced with 0.4 mM IPTG at OD_{600} of 0.6, and the cultures were incubated overnight at 18ºC with shaking at 180 rpm. The following morning 1 mL sample was taken from each culture for analysis of protein expression. Cells were harvested by centrifugation (4,000 × g, 30 mins, 4ºC), and the resulting pellets were resuspended in 250 µL of lysis buffer (50 mM Tris (pH 7.6), 300 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol, 0.075% (v/v) β-mercaptoethanol). Cells were lysed by freeze-thawing (3 cycles of freezing at -80ºC and thawing at RT), and the lysate was clarified by centrifugation at maximum speed in a benchtop centrifuge at 4ºC. The resulting supernatant mixed with 2 × protein loading dye was boiled at 95ºC for 3 mins and analysed by SDS-PAGE to confirm the presence of the target protein in the soluble fraction.

2.3.6. Large-scale expression of ^{15}N-labelled Rab domain

Rosetta 2 cells (DE3) transformed with the Rab domain expression construct in pET32a-LIC vector were grown at a 2 L scale in the modified M9 medium (42.3 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 18.7 mM $^{15}NH_4Cl$,1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.2% (w/v) glucose, 0.5% (w/v) ^{15}N-Celtone Base powder (Cambridge Isotope Laboratories), 1 × BME vitamins (Sigma-Aldrich), and 1 × trace elements (3.6 µM $FeSO_4$, 30.6 µM $ZnSO_4$, 1.6 µM $CuSO_4$, 0.5 µM $MnSO_4$, 40.5 nM $(NH_4)_6Mo_7O_{24}$ and 0.3 µM $Na_2B_4O_7$). Protein expression induced with 0.4 mM IPTG at OD_{600} of 0.6 was performed overnight at 18ºC with shaking at 180 rpm.

2.3.7. Small scale expression test of FL Rab46 in Sf9 cells

Expression of full length wtRab46 in *Sf9* cells was tested at a 30 mL scale. Initially, *Sf9* cells at 1.5 × 10^6 cells/mL were infected with P2 Rab46 baculovirus at a single multiplicity of infection (MOI) of 7, and protein expression levels were investigated at different harvest timepoints post infection (24 hrs, 48 hrs, 72 hrs and 96 hrs). To further optimise

the expression conditions, different amounts of baculovirus (MOI of 3, 5, 7, and 10) were tested with a single harvest timepoint at 48 hrs post-infection. *Sf9* cells were harvested by 10-minute centrifugation at 1,000 × g. Successful target expression was confirmed by capture of the his-tagged protein using Ni-NTA beads (Qiagen). Briefly, cell pellet from small-scale expression test was resuspended in 1 mL of lysis buffer (50 mM Tris (pH 7.6), 500 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol and 0.075% (v/v) β-mercaptoethanol), and the sample was sonicated on ice for 1 min 40 s (20 cycles of 5 s on and 10 s off) at 60% amplitude using a probe sonicator. The lysate was clarified by a 30-minute centrifugation at top speed in a chilled benchtop centrifuge and the resulting supernatant was incubated for 1 hour with 100 µL (bed volume) Ni-NTA resin (Qiagen) pre-equilibrated in the lysis/wash buffer. The protein was purified in a batch mode as described above for FL Rab46 expressed at a 100 mL scale in *E. coli*, and the elution fractions were analysed by SDS-PAGE.

2.3.8. Large-scale insect cell protein expression

Based on the results from the small-scale expression screen, 3 L of *Sf9* cells at 1.5 × 10^6 cells/mL were infected with P2 Rab46 baculovirus at MOI of 5. Cell diameter and viability were measured on the automated cell counter (Life Technologies) before and post-infection with Rab46 P2 baculovirus. Cells were harvested 48 hrs post-infection by centrifugation at 1,000 × g for 10 mins at 4ºC (with medium acceleration and deceleration). Cell pellet was washed by gentle resuspension in ice-cold PBS and centrifuged at 1,000 × g for 10 mins at 4ºC. *Sf9* cell pellets were flash-frozen in liquid nitrogen and stored at -80ºC until purification.

2.4. Protein purification

2.4.1. Protein Isolation

All protein isolation and purification steps were performed either on ice or at 4ºC. Cell pellets were resuspended in pre-chilled Ni-NTA lysis buffer containing 50 mM Tris (pH 7.6), 300 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol and 0.075% (v/v) β-mercaptoethanol. Lysis buffer was added at 1 mL per 25 mL of original culture volume (i.e. 40 mL lysis buffer were used for each 1 L culture). To prevent proteolytic degradation of the proteins, Proteoloc™ Protease Inhibitor (Expedeon) was added at 400 µL per litre of original culture. For purification of FL Rab46 expressed in *Sf9* cells concentration of NaCl in the lysis buffer was increased to 500 mM to minimise non-specific binding of contaminating proteins to the nickel column (see Table 4). Additionally, 5 mM $CaCl_2$ and 5 mM $MgCl_2$ and additional protease inhibitors were included in the IMAC buffers for the purification of FL Rab46.

Table 4 Protein purification buffers

Buffer	Composition
IMAC lysis/wash	50 mM Tris (pH 7.6), 300 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol, 0.075% (v/v) β-mercaptoethanol
FL WT Rab46 IMAC lysis/wash	50 mM Hepes (pH 7.6), 500 mM NaCl, 15 mM imidazole, 5% (v/v) glycerol, 0.075% (v/v) β-mercaptoethanol, 2 mM PMSF, 2 mM benzamidine, 5 mM $MgCl_2$, 5 mM $CaCl_2$
FL WT Rab46 IMAC elution	50 mM Hepes (pH 7.6), 500 mM NaCl, 500 mM imidazole, 5% (v/v) glycerol, 0.075% (v/v) β-mercaptoethanol, 2 mM PMSF, 2 mM benzamidine, 5 mM $MgCl_2$, 5 mM $CaCl_2$
IMAC high-salt wash	50 mM Tris (pH 7.6), 500 mM NaCl, 20 mM imidazole, 5% (v/v) glycerol, 0.075% (v/v) β-mercaptoethanol
IMAC elution	50 mM Tris (pH 7.6), 300 mM NaCl, 500 mM imidazole, 5% (v/v) glycerol, 0.075% (v/v) β-mercaptoethanol
EF-hand SEC and storage	25 mM Hepes (pH 7.0), 150 mM NaCl, 1 mM TCEP, 0.1 % (w/v) NaN_3
Rab domain SEC and storage	25 mM Hepes (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$,5 mM TCEP, 0.1 % (w/v) NaN_3
FL WT Rab46 SEC and storage	25 mM Hepes (pH 7.6), 150 mM NaCl, 1 mM TCEP, 5 mM $CaCl_2$, 5 mM $MgCl_2$

Bacterial and insect cells were lysed by high pressure homogenisation (Avestin). Having equilibrated the system with the IMAC wash buffer, the sample was applied, and pressure was adjusted to either 22.5 kpsi or 7.5 kpsi for disruption of bacterial, and insect cells respectively. To reduce heat generation and subsequent protein denaturation or aggregation, the outlet of the homogenizer was connected to a cooling unit, and the sample was collected in a bottle kept on ice. To ensure complete lysis, cell suspension was passed through the system twice until homogenous lysate was obtained. Cell debris was removed by centrifugation (35,000 × g, 4ºC, 45 mins), and the clarified supernatant was collected.

2.4.2. Immobilised Metal Ion Affinity Chromatography (IMAC)

Histidine-tagged proteins were purified by IMAC using an ÄKTA Pure chromatography system (Cytiva). The supernatant from the previous step was applied on either 1 mL (FL Rab46 expressed in *Sf9* cells) or 5 mL (all other proteins expressed in *E. coli*) HisTrap™ FF nickel column (Cytiva), pre-equilibrated with 5 column volumes (CV) of IMAC wash buffer (see Table 4). Following sample application, the column was washed in IMAC wash, and high-salt buffers (5 CV each), and the target protein was eluted over 20 CVs with 0%-100% gradient of elution buffer containing 500 mM imidazole. Fractions were analysed by SDS-PAGE and the ones containing the protein of interest were pooled and concentrated using 20 mL Pierce™ Protein Concentrators (Thermo Scientific™) with either 3 K MWCO (wtEF hand, EF-hand mutant proteins, and Rab domain) or 30 K MWCO (full length Rab46). To cleave the affinity tag, his-tagged TEV protease was added to the sample at a 1:50 mass ratio, and the protein was dialysed overnight in a dialysis cassette (Pierce) with an appropriate MWCO (3 K for EF-hand and Rab domains, and 30 K for full length Rab46) at 4ºC against IMAC wash buffer prepared without imidazole. The cleaved tag and TEV protease were removed in a reverse IMAC step on the 5 mL HisTrap column pre-equilibrated in IMAC wash buffer, collecting the flow through containing the untagged protein. To reduce the sample volume prior to SEC and achieve high resolution separation, the untagged protein was concentrated using 3 K or 30 K MWCO 20 mL Pierce™ Protein Concentrator (Thermo Scientific™).

2.4.3. Size Exclusion Chromatography (SEC)

Concentrated protein was further purified on either HiLoad Superdex 75 pg (wt EF-hand domain, EF-hand mutant proteins, and Rab domain) or HiLoad Superdex 200 pg (full length Rab46) 16/600 or 26/600 column (Cytiva) pre-equilibrated with 1.5 CV of appropriate SEC buffer (see Table 4). Following SDS-PAGE analysis, fractions containing pure protein were pooled and concentrated using 20 mL Pierce™ Protein Concentrator (Thermo Scientific™). Concentration of Rab domain and FL Rab46 were determined by measuring the absorbance of the protein samples at 280 nm in triplicate on NanoDrop spectrophotometer. Molar extinction coefficient values of Rab domain and FL Rab46 used for calculation of protein concentration were 18450 M^{-1} cm^{-1} and 46870 M^{-1} cm^{-1} respectively. Due to absence of tryptophan residues, the concentration of EF-hand proteins could not be measured using conventional spectrophotometric methods (A_{280nm}) and therefore was routinely determined using DC protein assay (Bio-Rad) with standard curves generated using BSA. Concentrated protein samples were flash frozen in liquid nitrogen and stored at -80ºC.

2.4.4. Ca^{2+}-free protein preparation

Ca^{2+}-free EF-hand domain was prepared by dialysis of the concentrated protein pool from the tag removal step against SEC buffer containing 25 mM EDTA. Ca^{2+}-free sample was purified by gel filtration on the HiLoad Superdex 75 pg column as described above. Following SEC, to ensure complete removal of EDTA for downstream applications, the concentrated sample was further buffer-exchanged into EDTA-free SEC buffer using 7K MWCO Zeba™ Spin column (Thermo Scientific™). Additionally, prior to calcium-binding experiments, the protein was exhaustively dialysed against EDTA-free SEC buffer, with the dialysis buffer replaced at least 5 times.

2.4.5. Western blotting

Recombinant FL Rab46 protein samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane (Millipore) using a semi-dry Pierce Power Blotter transfer system (Thermo Scientific). The membrane was incubated in 5% (w/v) non-fat milk prepared in TBS-T (20 mM Tris (pH 7.5), 145 mM NaCl, 0.5% Tween-20) for 1 hr at RT to prevent non-specific antibody binding and reduce the background. Following the blocking step, the membrane was incubated overnight at 4ºC with primary rabbit EFCAB4B antibody (Proteintech) diluted at 1:1000 in 5% (w/v) non-fat milk prepared in TBS-T. The following morning, it was washed 4 times for 10 mins in TBS-T, and subsequently incubated for 1 hr at RT with HRP-conjugated donkey anti-rabbit secondary antibody (Jackson ImmunoResearch) diluted at 1:10,000 in 5% (w/v) non-fat milk in TBS-T. The membrane was then washed 4 times in TBS-T (each wash was performed for 10 mins) prior to visualisation with Pierce ECL Western Blotting detection reagent (Thermo Scientific) on the G:Box gel documentation system (Syngene).

2.5. Mass Spectrometry

Intact mass, peptide mapping and Ion-Mobility Spectrometry-Mass Spectrometry (IMS-MS) experiments, and subsequent analysis were carried out by Dr Rachel George from Leeds Mass Spectrometry facility. Proteins in the MS experiments were used at 10 µM. For the studies of the effect of calcium binding samples were buffer exchanged into 100 mM ammonium acetate using 75 uL Zeba Spin Desalting Columns, 7K MWCO (Thermo Scientific) according to manufacturer instructions. Ammonium acetate was supplemented with 20 µM calcium acetate (Sigma) where appropriate.

2.6. Nuclear Magnetic Resonance (NMR)

NMR experiments were performed using 250-500 µM ^{15}N- and $^{15}N^{13}C$-labelled protein samples prepared in SEC buffer (as indicated in Table 4) containing 5% (v/v) D_2O,

loaded into 5 mm or 3 mm Wilmad tubes. Unless stated otherwise spectra were acquired at 25ºC on 750 MHz Bruker spectrometer equipped with a TCI cryoprobe.

2.6.1. Triple resonance backbone assignment

Experiments for backbone assignment of EF-hand domain were performed with 500 µM $^{15}N^{13}C$-labelled EF-hand protein either dialysed against EDTA, as described above for preparation of Ca^{2+}-free protein samples or saturated with 10 mM $CaCl_2$. Backbone assignments of apo and Ca^{2+}-bound EF-hand domain were obtained from analysis of ^{1}H-^{15}N HSQC, HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCA(CO)CB, and HNCACB spectra. All 3D spectra were acquired employing a non-uniform sampling (NUS) scheme (25%) in the indirect ^{13}C and ^{15}N dimensions and were reconstructed by SMILE (Sparse Multidimensional Iterative Lineshape Enhanced) NUS processing algorithm interfaced with NMRPipe software (91). Acquisition parameters for each of the experiments for triple resonance backbone assignment are listed in Table 5. Spectra were processed using NMRpipe/NMRDraw (92), and resonance assignment was performed manually using the CcpNMR analysis software version 2.5 (93). Chemical shift assignment quality and referencing were validated using PANAV software (94). Secondary chemical shifts and chemical shift index (CSI) (95) used to determine secondary structure elements from the $^{13}C\alpha$ and $^{13}C\beta$ shifts were calculated using CcpNMR analysis v2.5 (93).

2.6.2. CS-Rosetta structure prediction

Backbone assignment chemical shift data exported in NMRPipe (TALOS) format were used as an input into Chemical Shift-Rosetta web server (CS-Rosetta version 3.3, available at: https://csrosetta.bmrb.io (96)) for prediction of the apo and Ca^{2+}-bound EF-hand structures. Models were generated using the default parameters. The lowest-energy CS-Rosetta model of apo EF-hand and crystal structure of Ca^{2+}-bound EF-hand domain (PDB: 6PSD) were pre-processed and energy-minimised using Protein Preparation Wizard (Maestro software, Glide, Schrödinger). Superposition of the apo and Ca^{2+}-bound protein structures and RMSD calculations were performed using PyMOL (78). Solvent accessible surface area (SASA) of the lowest energy CS-Rosetta structures of apo and Ca^{2+}-bound EF-hand, and the published crystal structure of Ca^{2+}-bound EF-hand (PDB: 6PSD) was calculated using UCSF Chimera software version 1.16 (97). Hydrophobic residues on the surface of CS-Rosetta apo and crystal Ca^{2+}-bound structures were coloured red according to the Eisenberg hydrophobicity scale using ‘Color_h’ script in PyMOL (78,98).

Table 5 Acquisition parameters for triple resonance backbone assignment NMR experiments.

HSQC	Pulse sequence: hsqcetfpf3gpsi		
	^{1}H	SW (Hz)	11961.722
		Points	2048
	^{15}N	SW (Hz)	1818.843
		Points	256
	Ns		4
HNCA	Pulse sequence: B_HNCA		
	^{1}H	SW (Hz)	10482.180
		Points	1280
	^{15}N	SW (Hz)	1818.843
		Points	76
	^{13}C	SW (Hz)	4887.586
		Points	96
	Ns		16
HN(CO)CA	Pulse sequence: B_HNCOCA		
	^{1}H	SW (Hz)	10482.180
		Points	1280
	^{15}N	SW (Hz)	1818.843
		Points	76
	^{13}C	SW (Hz)	4887.586
		Points	96
	Ns		16
HNCO	Pulse sequence: B_HNCO		
	^{1}H	SW (Hz)	10482.180
		Points	1280
	^{15}N	SW (Hz)	1818.843
		Points	96
	^{13}C	SW (Hz)	1880.406
		Points	116
	Ns		8
HN(CA)CO	Pulse sequence: IBS_Best_HNCACO		
	^{1}H	SW (Hz)	10504.202
		Points	1280
	^{15}N	SW (Hz)	1818.843

		Points	76
	^{13}C	SW (Hz)	1880.406
		Points	116
	Ns		32
HNCACB	Pulse sequence: B_HNCACB		
	^{1}H	SW (Hz)	10482.180
		Points	1280
	^{15}N	SW (Hz)	1818.843
		Points	76
	^{13}C	SW (Hz)	11286.682
		Points	160
	Ns		48
HN(CO)CACB	Pulse sequence: cbcaconhgp3d		
	^{1}H	SW (Hz)	10504.202
		Points	2048
	^{15}N	SW (Hz)	1818.843
		Points	76
	^{13}C	SW (Hz)	11286.682
		Points	148
	Ns		16

2.6.3. 1D NMR

Prior to the calcium binding NMR experiments, 1D ^{1}H NMR spectra were acquired with water suppression by excitation sculpting using zgesgp pulse sequence (Topspin 3.2, Bruker) with 4 scans to assess whether residual EDTA was successfully removed from the EF-hand domain protein samples. 1D Spectra were processed using NMRpipe/NMRDraw (92) and overlaid for comparison using CcpNMR analysis software v2.5 (93).

2.6.4. NMR Ca^{2+} titration experiment

Ca^{2+} binding to EF-hand domain was evaluated by serial addition of $CaCl_2$ to the 250 µM ^{15}N-labelled protein sample prepared in SEC buffer (Table 4) containing 5% (v/v) D_2O. To minimize sample dilution during Ca^{2+} titration experiment, increasing concentrations of $CaCl_2$ stock solutions (2 mM, 4 mM, 8 mM, 25 mM, 50 mM, 100 mM, and 1000 mM) prepared in SEC buffer were used for obtaining the samples for consecutive ^{1}H-^{15}N HSQC experiments. Spectra were acquired using hsqcetfpf3gpsi pulse sequence (Topspin 3.2, Bruker) employing 2048 points in the direct ^{1}H dimension, and 256 points

in the indirect ^{15}N dimension. Each HSQC experiment was recorded with 8 scans and spectral widths of 11961.722 Hz and 1818.843 Hz for ^{1}H and ^{15}N dimensions, respectively. Spectra were processed using NMRpipe/NMRDraw (92) and analysed with CcpNMR analysis software v2.5 (93), copying the assignments for apo and Ca^{2+}-bound EF-hand domain obtained from the triple resonance backbone assignment experiments described above.

Chemical shift perturbations for the backbone NH peaks ($\Delta\delta_{NH}$) were calculated using the equation below. A scaling factor of 0.14 was used for the ^{15}N shifts (99).

$$\Delta\delta_{\mathrm{NH}} = \sqrt{\frac{1}{2}[\Delta\delta_{\mathrm{H}}^2 + (0.14 \times \Delta\delta_{\mathrm{N}}^2)]}$$

CSP values above mean and mean + SD were mapped on the published crystal structure of Ca^{2+}-bound EF-hand domain (PDB: 6PSD) in yellow, and orange respectively using PyMOL software (78).

2.6.5. HSQC experiments at lower pH

The effect of lowering the pH on the Ca^{2+}-saturated EF-hand domain was monitored using 250 µM ^{15}N-labelled EF-hand protein sample saturated with 5 mM $CaCl_2$, prepared in SEC buffer containing 5% (v/v) D_2O. To prevent an increase in ionic strength during titration, protein samples at lower pH were prepared by buffer-exchange into SEC buffer (25 mM Hepes, 150 mM NaCl, 1 mM TCEP, 0.1 % (w/v) NaN_3) prepared at pH of 6.0 and 6.5, using Zeba Spin Desalting Columns, 7K MWCO (Thermo Scientific). ^{1}H-^{15}N HSQC spectra were acquired using hsqcetfpf3gpsi pulse sequence (Topspin 3.2, Bruker) with 8 scans, 2048 × 256 data points in the direct and indirect dimensions respectively, and spectral widths of 11961.722 Hz for ^{1}H, and 1818.843 Hz for ^{15}N dimensions. Spectra were processed and analysed using NMRPipe/NMRDraw and CcpNMR analysis v2.5 software (92,93). CSPs were calculated using the equation described in section 2.6.4. CSP values above mean and mean + SD were mapped on the published crystal structure of Ca^{2+}-bound EF-hand domain (PDB: 6PSD) in salmon and red, respectively, using PyMOL software (78).

2.6.6. ZZ exchange experiments

^{1}H-^{15}N ZZ exchange spectra were acquired using 500 µM ^{15}N-labelled EF-hand domain sample in SEC buffer saturated with 10 mM $CaCl_2$, containing 5% (v/v) D_2O. Three separate experiments with different mixing times (50, 250 and 600 ms) were recorded using hsqcexf3gpwgph pulse sequence (Topspin 3.2, Bruker) with 32 scans, 2048 × 256

complex points and spectral widths of 11961.72 Hz, and 1818.182 Hz for the ^{1}H and ^{15}N dimensions, respectively. Spectra were processed and analysed using NMRPipe/NMRDraw and CcpNMR analysis software (92,93).

2.6.7. Relaxation experiments

^{15}N relaxation measurements were performed using 500 µM ^{15}N-labelled apo and Ca^{2+}-saturated (10 mM $CaCl_2$) EF-hand domain samples prepared in SEC buffer containing 5% (v/v) D_2O. ^{15}N T1 and T2 experiments were acquired with 8 scans, 1792 × 320 points, and spectral widths of 11961.722 Hz and 1818.182 Hz in the direct and indirect dimensions, respectively.

^{15}N T1 and T2 relaxation time constants were derived from ^{15}N T1 and T2 spectra recorded with different relaxation delay values. For T1 measurements ten spectra were acquired with relaxation delays of 0, 20, 120, 60, 160, 90, 40, 10, 20, 90 ms. For the T2 measurements, eleven spectra were recorded with relaxation delays of 10, 4, 6, 8, 14, 20, 2, 12, 16, 6, and 12 ms. Spectra were processed using NMRPipe and analysed using CcpNMR analysis v2.5 software (93), and custom scripts. R1 and R2 rate constants were extracted using in-house written scripts including a Monte-Carlo error estimation with 1000 steps.

^{1}H-^{15}N heteronuclear NOE experiments were recorded in an interleaved fashion (reference and saturated spectra) using hsqcnoef3gpsi pulse sequence (Topspin 3.2, Bruker) with 16 scans, 1792 × 640 points, and spectral widths of 11961.722 Hz and 1818.182 Hz for ^{1}H and ^{15}N dimensions respectively. Spectra were processed using a custom processing script and NMRPipe software (92). ^{1}H-^{15}N HetNOE values were calculated in CcpNMR analysis v2.5 as the ratio of peak intensities in the saturated to the reference spectra (93).

2.6.8. Amide RDC measurements

Aligned ^{15}N-labelled apo EF-hand sample at 500 µM was prepared in SEC buffer containing 12.5 mg/mL *Pf1* phage (ASLA Biotech) and 5% (v/v) D_2O. An equivalent protein sample without *Pf1* phage was prepared for the experiment under isotropic conditions. $^{1}J_{NH}$ couplings were measured from the spectra acquired using the ARTSY pulse sequence (Artsyf3gpphrc_prot, i.e. the experiment version for protonated samples) with 96 scans, 2048 × 640 points and spectral widths of 11961.722 Hz and 1818.843 Hz in the ^{1}H and ^{15}N dimensions respectively. Reference and attenuated spectra recorded in an interleaved manner were split in Topspin 3.2 (Bruker).

Reference and attenuated spectra for the samples under isotropic and aligned conditions

were processed in NMRPipe (92). The noise level in the spectra was estimated using NMRDraw (92). $^{1}J_{NH}$ and associated error values were extracted from isotropic and aligned experiments using an analysis script from Fitzkee and Bax (100). Backbone amide $^{1}D_{NH}$ RDCs in apo EF-hand domain were calculated by taking the difference in $^{1}J_{NH}$ couplings in aligned and isotropic media. The error on the experimental RDCs was calculated from the signal-to-noise ratio of the ARTSY spectra using the following equation:

$$\text{RDC error} = \sqrt[2]{\text{error\_aligned}^2 + \text{error\_isotropic}^2}$$

The obtained RDC restraints table and the lowest-energy CS-Rosetta structure of apo EF-hand with added protons (prepared using the online tool available at: https://spin.niddk.nih.gov/bax/) were then used to calculate the alignment tensor and produce a plot of calculated versus experimental RDCs using Xplor-NIH (101).

2.6.9. NMR experiments with ^{15}N-labelled Rab domain

A ^{1}H-^{15}N HSQC experiment with Rab domain was acquired at 25ºC using 250 µM ^{15}N-labelled sample prepared in SEC buffer containing 5% (v/v) D_2O. The protein used in the experiment was not loaded with any nucleotides or GTP/GDP nucleotide analogues. The ^{1}H-^{15}N HSQC spectrum was acquired using hsqcetfpf3gpsi pulse sequence (Topspin 3.2, Bruker) with 32 scans, 2048 × 256 data points in direct and indirect dimensions, and spectral widths of 11961.722 Hz for ^{1}H, and 1818.843 Hz for ^{15}N dimensions.

A ^{1}H-^{15}N TROSY experiment with Rab domain was performed at 15ºC using 60 µM ^{15}N-labelled protein sample loaded with 20-fold molar excess of GppNHp (Sigma-Aldrich) prepared in SEC buffer containing 5% D2O. ^{15}N-labelled Rab domain loading with GppNHp (Sigma-Aldrich) was performed as described in sections 2.9.1 and 2.9.2 for the preparation of MANT-GDP loaded Rab domain samples. Following loading with GppNHp, and prior to the NMR experiment, the ^{15}N-labelled protein was buffer-exchanged back into SEC buffer on a NAP-5 column (Cytiva) following the procedure described in section 2.9.3.

The ^{1}H-^{15}N TROSY spectrum was acquired using IBS_BTROSY pulse sequence (Topspin 3.2, Bruker) with 512 scans, 1280 × 192 points in direct and indirect dimensions respectively, and spectral widths of 10504.202 Hz for ^{1}H and 1818.182 Hz for ^{15}N dimensions.

^{1}H-^{15}N HSQC and TROSY spectra were processed and analysed using NMRPipe/NMRDraw, and CcpNMR analysis v2.5 software (92,93).

2.7. Isothermal Titration Calorimetry (ITC)

Ca^{2+} titration was performed on MicroCal iTC200 system (Malvern Panalytical) by sequential injections every 120 seconds of 2 µL of 1 mM $CaCl_2$ solution into the sample cell containing 80-100 µM protein prepared in SEC buffer at 25ºC. Data fitting and calculation of binding parameters was performed using MicroCal ITC-ORIGIN Analysis software.

2.8. Circular Dichroism (CD)

The secondary structure content of EF-hand and Rab domain proteins was assessed by Far-UV CD spectroscopy. Protein samples for CD were prepared by exhaustive dialysis against phosphate buffer compatible with CD measurements. EF-hand (wt and mutant) and Rab domain protein samples were prepared in 50 mM sodium phosphate buffer (pH 7.2) containing 50 mM NaCl, and 50 mM sodium phosphate buffer (pH 8.0), respectively.

Far-UV CD spectra (180-260 nm) were acquired on a Chirascan Plus Spectrometer (Applied Photophysics) at 20ºC using 200 µL protein samples at 0.2 mg/mL, loaded into 1 mm quartz cuvettes. Two consecutive spectra were recorded for each sample between 180-260 nm with 1 nm increments, speed of 1 nm/s and 2 nm bandwidth. Average taken of the two consecutive readings was normalised by subtraction of the background signal obtained from acquiring the spectrum of the sample buffer alone.

The concentration of each sample estimated using DC protein assay was corrected by calculating the concentration from Beer-Lambert law ($A= \varepsilon \times c \times l$) using pathlength (l) of 0.1 cm, absorbance (A) value recorded at 208 nm, and extinction coefficient (ε) at 205 nm (peptide bond absorption) estimated using the 'Protein Parameter Calculator' tool (102).

Mean residue weight (MRW) for each protein was calculated as the molecular weight (in Da) divided by the number of backbone amides (i.e. number of amino acids – 1). To correct for sample concentration effect, CD data in millidegrees (mdeg) were converted to mean residue ellipticity (Δε) following the method described in Greenfield (2007) (103). First, ellipticity [θ] was calculated by multiplying the measured CD spectrum by mean residue weight and dividing the product by the optical pathlength and protein concentration, as illustrated by the equation below:

$$[\theta]\ (deg \times cm^2/dmol) = \frac{\text{CD}\ (mdeg) \times \text{MRW}\ (\text{Da})}{\text{pathlength}\ (mm)\ \times\ \text{concentration}\ (mg/mL)}$$

Mean residue ellipticity (Δε) was then calculated by dividing the obtained ellipticity values by a factor of 3,298.

$$\Delta\varepsilon = \frac{[\theta]}{3{,}298}$$

Secondary structure content was calculated using DichroWeb online tool (104) using the CDSSTR method and the reference set SP175, optimised for the wavelength range from 175 to 260 nm (104,105).

2.8.1. Thermal Denaturation

A stepped temperature ramping CD experiment was performed on a Chirascan Plus Spectrometer (Applied Photophysics) using 200 µL apo and Ca^{2+}-bound (saturated with 5 mM $CaCl_2$) wt EF-hand domain samples prepared at 0.2 mg/mL in 50 mM sodium phosphate buffer (pH 7.2), loaded into 1 mm quartz cuvettes. Measurements for the apo and Ca^{2+}-bound EF-hand domain samples were performed simultaneously utilising a rotating 6-cell carousel included in the spectrometer setup. Temperature was increased by 1ºC in 2-minute intervals and the readings were taken following stabilisation of the sample temperature. Two consecutive spectra were recorded for each sample between 180-260 nm with 1 nm increments, speed of 1 nm/s and 2 nm bandwidth. Analysis of the spectra and calculation of the thermal transition midpoints (T_m) values from the CD profiles obtained during temperature ramping were performed using Global 3 software (Applied Photophysics).

2.9. Nucleotide Exchange Experiment

2.9.1. Rab domain sample preparation

The MANT-GDP nucleotide exchange assay was performed following the published method (106). Prior to nucleotide loading the purified Rab domain was buffer-exchanged into low-Mg^{2+} buffer (Table 6) on a NAP-5 column (Cytiva). Briefly, following the column equilibration in 3 CVs of low-Mg^{2+} buffer, 200 µL of Rab domain sample at 120-200 µM (depending on the batch) was applied to the top of the column and, once the sample entered the column bed completely, 500 µL of low-Mg^{2+} buffer (stacker volume) was added to the column, and the flow through (not containing any protein) was collected. The buffer-exchanged protein sample was then eluted with 1 mL of low-Mg^{2+} buffer and collected as 100 µL fractions. Concentration of each of the fractions was measured using Bradford reagent (Sigma-Aldrich), with the standard curve generated using BSA. The elution fraction found to contain the highest concentration of the buffer-exchanged protein was used for the nucleotide loading step.

Table 6 Composition of buffers for the nucleotide exchange assay

Buffer	Composition	Volume needed for a single experiment
Low-Mg^{2+}	20 mM Hepes (pH 7.5), 50 mM NaCl, 0.5 mM $MgCl_2$	100 mL
2 × MANT-GDP loading	20 mM Hepes (pH 7.5), 50 mM NaCl, 0.5 mM $MgCl_2$, 10 mM EDTA, 2 mM DTT, 20-fold molar excess MANT-GDP (depending on the protein concentration following exchange into low-Mg^{2+} buffer)	100 µL
Nucleotide exchange	40 mM Hepes (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 2 mM DTT	100 mL

2.9.2. MANT-GDP loading

100 µL of Rab domain buffer-exchanged into low-Mg^{2+} buffer was mixed with 100 uL of 2 × MANT-GDP loading buffer (Table 6) containing 20-fold molar excess of MANT-GDP triethylammonium salt solution (Sigma-Aldrich). Reaction was set up in a light protective black 0.5 mL Eppendorf tube and was incubated static in a benchtop thermomixer (Starlab) at 20ºC for 90 mins, after which time it was stopped by addition of 2 µL of 1 M $MgCl_2$ (to obtain a final $MgCl_2$ concentration of 10 mM). The terminated reaction was then incubated static at 20ºC for another 30 mins.

2.9.3. Removal of unbound MANT-GDP

To remove the unbound MANT-GDP the 200 µL sample was buffer-exchanged on a NAP-5 column (Cytiva) pre-equilibrated in 3 CVs of nucleotide exchange buffer (Table 6), as per the instructions described above for sample preparation for the nucleotide loading. 100 µL elution fractions containing the buffer-exchanged Rab domain loaded with MANT-GDP were kept on ice protected from light, in black 0.5 mL Eppendorf tubes. Concentration of each elution fraction was measured using Bradford reagent (Sigma) with the standard curve generated using BSA.

2.9.4. Determination of loading efficiency

MANT-GDP standards were prepared by two-fold serial dilution (40 µM to 0.156 µM) of free MANT-GDP (Sigma-Aldrich) in nucleotide exchange buffer. The fluorescence intensity (FI) of MANT-GDP standards, MANT-GDP loaded Rab domain samples, and

nucleotide exchange buffer blank was measured in a final volume of 15 µL in a low-volume, flat-bottom, black 384-well plate (Greiner), at 25ºC on the POLARstar OPTIMA microplate reader (BMG LABTECH), by reading the signal every 45 sec, 10 times in total, with excitation and emission wavelengths of 360 nm and 460 nm, respectively.

FI values from all ten timepoints for each sample were averaged and normalised by subtraction of the background fluorescence value obtained from measurement of the buffer-only control. MANT-GDP standard curve was obtained by plotting the measured FI of free MANT-GDP samples against their concentration. Since the signal of MANT-GDP typically increases approximately 2-fold upon binding to a protein (106), the fluorescent signal of the Rab domain samples loaded with MANT-GDP was first divided by two, and then the standard curve equation was used to estimate the concentration of bound MANT-GDP. Loading efficiency was then calculated using the formula below:

$$\text{Loading efficiency} = \frac{\text{Molar concentration of bound MANT} - \text{GDP}}{\text{Molar concentration of Rab domain sample}} \times 100\%$$

2.9.5. Nucleotide exchange assay

Having estimated the loading efficiency, 14 µL MANT-GDP loaded Rab domain or nucleotide exchange reaction buffer samples were transferred to a low-volume, flat-bottom, black 384-well plate (Greiner). For each sample, four technical replicates were prepared. GppNHp (Sigma-Aldrich) was reconstituted in 50 mM Hepes buffer (pH 7.5) to obtain 100 mM stock concentration. A stock solution of GppNHp at a molar concentration 1,500-fold higher than the one of Rab domain to be used in an assay (i.e. 37.5 mM GppNHp for 25 µM MANT-GDP loaded Rab domain) was prepared in nucleotide exchange buffer. The fluorescence signal stability was confirmed by measuring the FI of the prepared samples for 150 sec at 25ºC prior to initiation of the nucleotide exchange experiment.

The experiment was started by addition of 1 µL GppNHp working solution to the assay samples (to obtain 100-fold final molar excess over Rab domain). 1 µL of nucleotide exchange buffer was added in place of GppNHp to the control samples. FI signal was measured at 25ºC on the POLARstar OPTIMA microplate reader (BMG LABTECH), reading the signal every 50 sec for 1 hr with excitation and emission wavelengths of 360 nm and 460 nm, respectively. The assay was repeated three times, with Rab domain protein samples from two different batches tested in each experiment.

2.9.6. MANT-GDP assay data analysis

Data were normalised by subtracting the background intensity of the nucleotide

exchange buffer from the FI of each assay sample. Relative fluorescence was calculated as a ratio of each FI value to the FI at the first time point after addition of either GppNHp or reaction buffer. The observed rate constant value (K_{obs}) was calculated in GraphPad Prism (107) by fitting the data with nonlinear regression, using the ‘one phase exponential decay’ equation.

Chapter 3
Protein Production

3.1. System for overexpression of recombinant full length Rab46 and its individual domains

3.1.1. Background

Since there had been no published method for the production of recombinant Rab46 protein or its individual functional domains available when this study was commenced, protocols had to be developed and optimised from scratch.

Protein constructs were initially designed for bacterial overexpression to ensure that large quantities of target proteins are obtained to allow their biophysical and structural characterisation. Being a low cost, quick and easy to scale up culture method, bacterial expression has clear advantages as a system for production of recombinant proteins and hence *E. coli* is typically the expression host of first choice. There is a wide variety of engineered *E. coli* strains available to facilitate production of difficult protein targets offering solutions to challenges associated with target toxicity, poor solubility or misfolding as a result of incorrect disulphide bond formation. Moreover, bacterial expression is well suited for production of uniformly and non-uniformly isotopically labelled proteins for NMR spectroscopy studies, which was an important consideration for this project.

The different protein constructs used in this study are summarised in Table 7. Constructs for expression of individual domains were designed not only as a backup strategy in case of the full length Rab46 proving difficult to express or purify, but also to enable study of the protein at the domain level as there had been no studies characterising functional domains of Rab46 GTPase when the project started. The domain boundaries used for construct design were identified using the NCBI CD-Search tool (48) (accession no: XP_006719084.1). Since cutting a domain ‘too short’ as a result of the difficulties associated with establishing the precise domain boundaries may lead to loss of its stability (108), the constructs for expression of individual functional domains of Rab46 were designed to include amino acid residues beyond the predicted domain boundaries to maximise the chances of obtaining stable proteins.

Table 7 Summary of constructs for bacterial expression of full length Rab46 and its individual functional domains

Construct name	Residues	MW tagged protein	MW untagged protein
FL Rab46 in pET32a-LIC	Met1 - Gly732	100.63 kDa	83.37 kDa
EF-hand in pET32a-LIC	Gly47 - Asn122	26.07 kDa	8.81 kDa
Coiled-coil in pET32a-LIC	Asp199 - Asn376	38.79 kDa	21.52 kDa
Rab in pET32a-LIC	Pro543 - Gly722	37.89 kDa	20.63 kDa
FL Rab46 in pET28a-LIC	Met1 - Gly732	85.31 kDa	83.42 kDa
EF-hand in pET28a-LIC	Gly47 - Asn122	10.75 kDa	8.86 kDa
Rab in pET28a-LIC	Pro543 - Gly722	22.57 kDa	20.69 kDa

3.1.2. Molecular cloning

As shown in Table 7, Rab46 constructs were generated using two expression vectors. Both pET28a-LIC and pET32a-LIC are low copy number plasmids, which can be a helpful feature in reducing leaky expression prior to protein expression induction with IPTG which, in case of target protein being toxic to the host cell, can lead to limited cell growth and subsequently, to decreased protein yields. While pET28a-LIC vector contains only a small cleavable N-terminal his-tag, pET32a-LIC vector additionally encodes cleavable N-terminal thioredoxin- and S-tags. These additional tags offer a strategy for overcoming issues with protein solubility as well as an alternative means for protein purification and detection of the expressed target by immunoblotting (antibodies against S-tag). Thus, aiming to find expression conditions yielding highest amounts of soluble protein for each target, the two vectors were tested simultaneously. Constructs were designed so that tags were restricted to the N-terminus of recombinant proteins to allow their straightforward removal for structural biology applications by proteolytic cleavage.

The expression constructs with N-terminal tags were produced as described in section 2.1. As demonstrated in Figure 10 (Panel A), pET32a-LIC plasmid digested with BseRI endonuclease gave the expected band pattern confirming complete linearisation of the vector.

DNA inserts encoding FL Rab46, EF-hand and Rab domains were successfully amplified by PCR (Figure 10, Panel B) and, following digestion of template DNA with DpnI and its subsequent gel purification, DNA of sufficiently high purity was obtained for insertion into the expression vector by In-Fusion method.

Successful insert incorporation was verified by diagnostic digest with XbaI restriction endonuclease (Figure 11) and by DNA sequencing. The same workflow was followed in case of constructs in pET28a-LIC vector (gel data not shown).

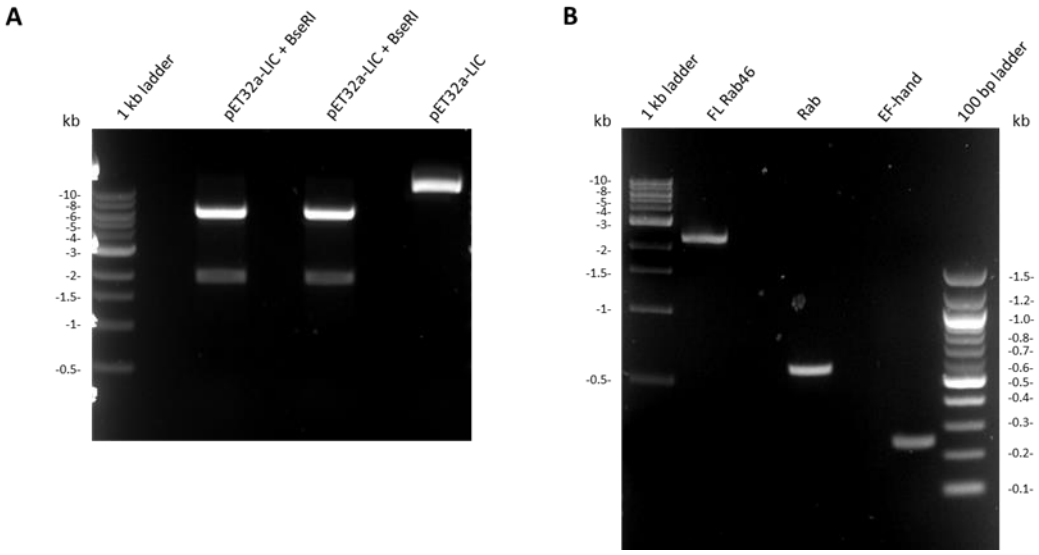

Figure 10 Generation of constructs for expression of FL Rab46, EF-hand and Rab domains. (A) Agarose gel of pET32a-LIC vector digested with BseRI. Complete digestion of pET32a-LIC was achieved as demonstrated by the presence of two bands in ‘pET32a-LIC + BseRI’ lanes. The sizes of obtained DNA fragments match the ones expected for the linearised pET32a-LIC vector (6.5 kb) and the excised SacB stuffer sequence (1.4 kb). ‘pET32a-LIC’ lane contains non-digested plasmid as a negative control. (B) Agarose gel of DpnI-digested and purified PCR products. Sizes of the obtained products match the expected sizes of 2.2 kb for FL Rab46, 0.5 kb for Rab domain and 0.2 kb for EF-hand domain.

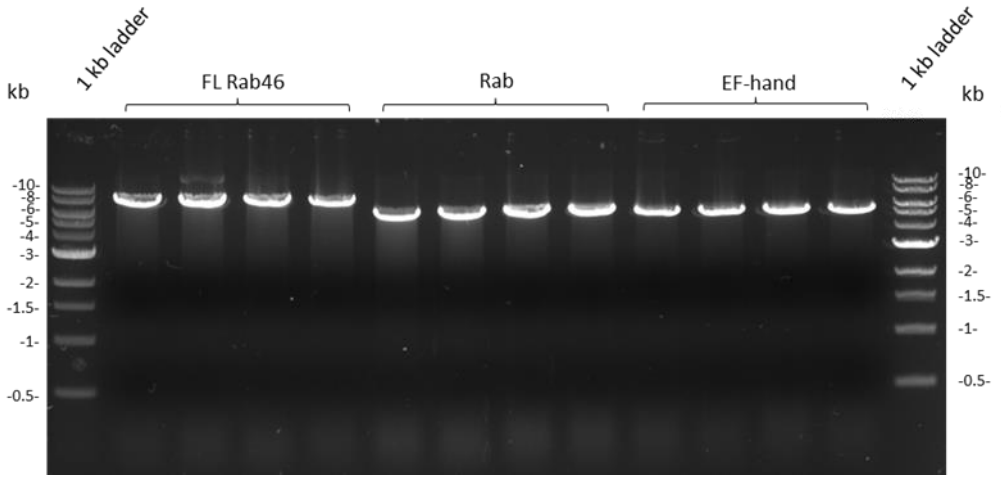

Figure 11 Diagnostic digest of Rab46 expression constructs with XbaI. Plasmid DNA of Rab46 constructs in pET32a-LIC vector was digested with XbaI enzyme at 37ºC for 2 hrs. For each of the three constructs, DNA isolated from 4 individual colonies was tested. The observed bands appear to have correct sizes, roughly matching the expected ones of 8.7 kb, 7 kb, and 6.7 kb for FL Rab46, Rab domain and EF-hand constructs respectively.

3.1.3. Coiled-coil domain

Unfortunately, generation of the construct for expression of the Coiled-coil domain proved to be much less straightforward than of the other Rab46 domains. Despite extensive troubleshooting efforts to optimise the PCR of Coiled-coil by modifying the reaction components and thermal cycling conditions, the product could not be successfully amplified. Summary of the tested PCR conditions is shown in Table 8.

Table 8 Conditions tested for optimisation of Coiled-coil domain PCR

Gradient PCR using different annealing temperatures	No product amplification
Addition of 2 - 5% DMSO	No product amplification
CloneAmp vs GoTaq vs Phusion polymerase	No product amplification
3-step vs 2-step PCR	No product amplification
Touchdown PCR	No product amplification
Primer redesign	Successful product amplification

Eventually, a decision was made to attempt redesign of the Coiled-coil domain PCR primers aiming to resolve issues with primer annealing to the template DNA. Given the results of small-scale expression screen described later in this thesis (section 3.2.1), the new primers were designed to obtain expression construct in pET32a-LIC vector only. In order to ensure better primer characteristics, i.e. lower the primer melting temperature (Tm) and propensity to form secondary structures during PCR, the domain boundaries were extended from Leu201-Arg373 (original primers) to Asp199 - Asn376. As shown in Table 9, the new pair of primers exhibited somewhat better characteristics, as evaluated by OligoEvaluator tool (Sigma Aldrich).

Table 9 Primers for Coiled-coil domain cloning

Name	**Primer Sequence**	**Tm**	**Secondary structure**
Original Fwd (pET28a-LIC)	GTTCCGCGTGGTAGTCCTGACCGGCTCTTCAAG	82.4ºC	moderate
Original Rev (pET28a-LIC)	CAAGCTTCGTCATCACCTTAGGAAATCC	71.4ºC	very weak
Re-designed Fwd (pET32a-LIC)	GTATTTCCAGAGCGACTTCCTGACCAGAA	72.8ºC	Weak
Re-designed Rev (pET32a-LIC)	CAAGCTTCGTCATCAGTTCCTTTCCCTTAG	72.7ºC	very weak

As shown in Figure 12, following primer redesign Coiled-coil was successfully amplified and inserted into pET32a-LIC using the same PCR and In-Fusion reaction conditions as for the other Rab46 domains (described in section 2.1).

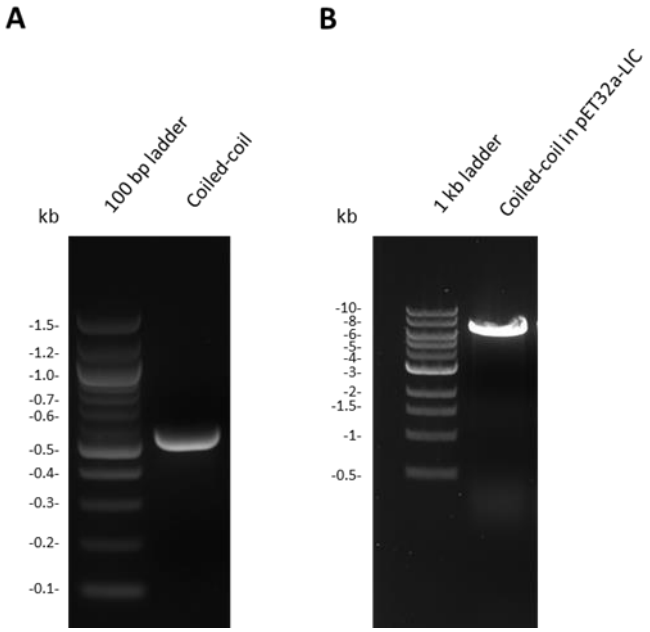

Figure 12 Generation of Coiled-coil domain expression construct. (A) Agarose gel of DpnI-digested and purified Coiled-coil PCR product. The size of the obtained band matches the expected one of 0.5 kb. (B) Diagnostic digest of Coiled-coil expression construct with XbaI. Coiled-coil in pET32a-LIC vector was digested with XbaI enzyme at 37ºC for 2 hrs. The observed band appears to have the correct size of 7 kb.

3.2. Protein expression

3.2.1. Small Scale Expression Screen

To find conditions yielding largest amounts of soluble protein, expression of the pET28a-LIC and pET32a-LIC constructs was tested in multiple *E. coli* strains at a 2 mL scale. Since the constructs for expression of human Rab46 proteins were not codon-optimised, Rosetta(DE3) and Rosetta2(DE3) strains containing additional tRNAs for the codons rarely used in *E. coli* were included in the screen, in addition to the standard BL21(DE3) strain routinely used for protein production. Expression of the Rab46 constructs was also tested in SHuffle T7 Express strain engineered to enhance correct folding of protein targets containing multiple disulphide bonds, and Lemo21 (DE3) strain which allows for the protein expression levels to be tuned by addition of L-rhamnose, to alleviate potential issues with target toxicity. Although Rab46 is an intracellular protein and hence no disulphide bond formation is expected in Rab46 constructs, disulphide bond isomerase (DsbC) constitutively expressed by the SHuffle T7 Express strain, has been demonstrated to act as a chaperone that can assist in the folding of proteins that do not

require disulphide bonds (109). The 2 mL cultures were grown in LB media supplemented with appropriate antibiotics in a 96-well plate format. Following induction with 0.4 mM IPTG, protein expression was performed overnight at 25ºC. His-tagged proteins captured using magnetic Ni-NTA beads were analysed by SDS-PAGE (Figure 13).

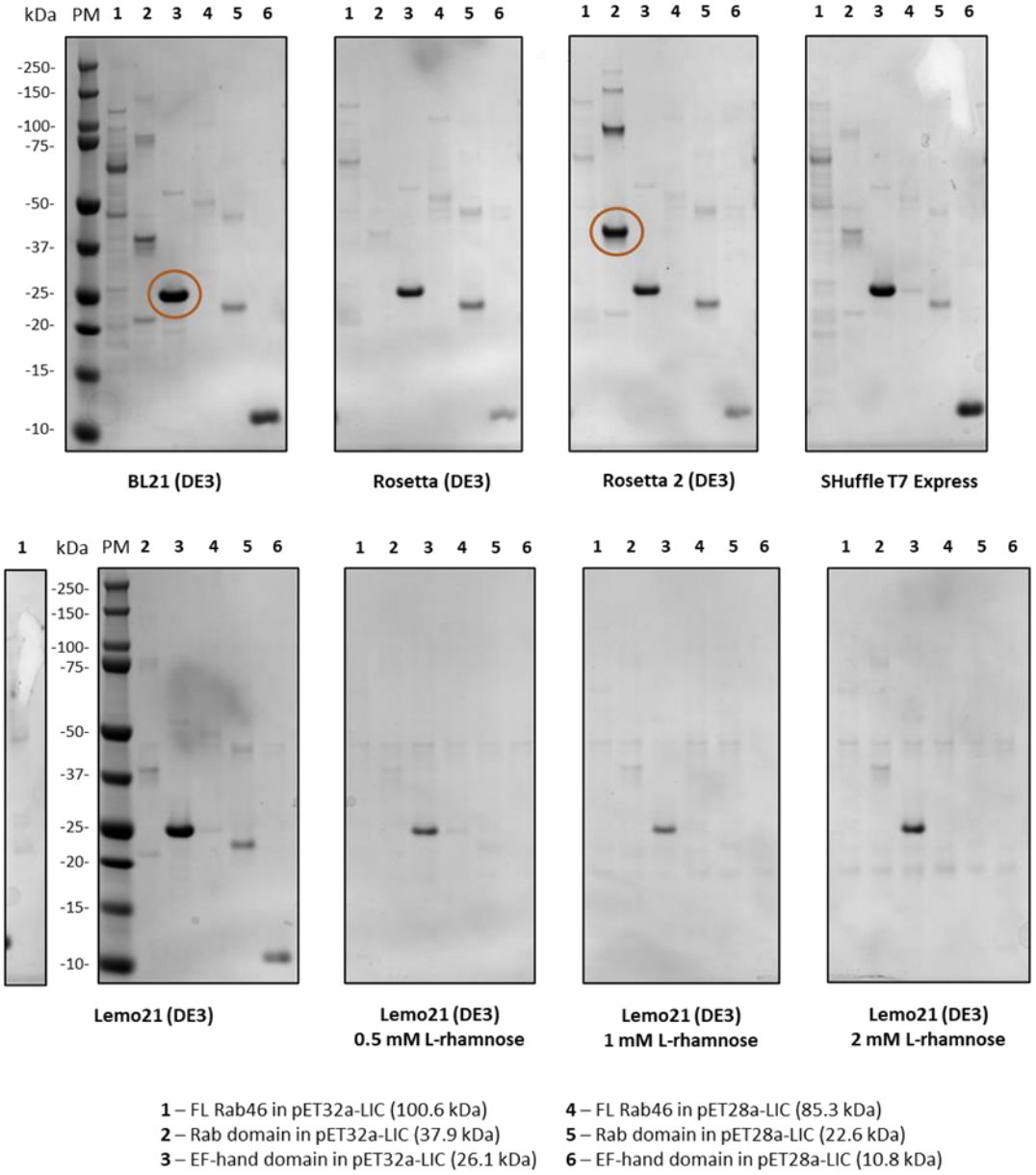

Figure 13 Small-scale expression of Rab46 in different *E. coli* strains. Lane numbering and the expected molecular weights of each of the constructs are indicated in the figure legend below the protein gels. The conditions identified as optimal for expression of EF-hand and Rab domains are marked with red circles. PM – protein marker.

Unfortunately, there was no clear evidence for successful overexpression of the full length Rab46 protein under any condition tested in the small-scale expression screen (Figure 13). Although there are multiple bands present in the lanes corresponding to full length protein expressed using either pET32a-LIC or pET28a-LIC based constructs (conditions 1 and 4), their molecular weights do not match the expected protein sizes of 100.6 kDa and 85.3 kDa respectively. Nevertheless, the band pattern observed for conditions 1 and 4 appears to be consistent across the different strains tested, with the most prominent band seen at approximately 65 kDa for condition 1 and 50 kDa for condition 4, potentially suggesting either proteolytic degradation of the full length protein or early termination of protein translation.

The expression of the two functional domains of Rab46, however, proved to be much more successful. As shown in Figure 13, EF-hand in pET32a-LIC vector (sample 3) was expressed under every condition tested. The highest expression of the target was observed in BL21(DE3), SHuffle T7 Express and Lemo21(DE3) strain in the absence of L-rhamnose. EF-hand in pET28a-LIC vector (sample 6) also demonstrated a good level of expression in BL21 and SHuffle T7 Express cells. However, the bands obtained for pET32a-LIC construct appeared markedly stronger, and hence a decision was made to use it for large-scale expression of EF-hand domain protein in BL21 cells.

Compared to EF-hand, Rab domain constructs showed moderately lower expression levels across all tested *E. coli* strains. Nevertheless, while the expression levels of Rab domain in pET32a-LIC (sample 2) varied considerably between the strains, a band at approximately 38 kDa, corresponding to the tagged protein, could be seen for all conditions tested. The protein was also successfully expressed using Rab in pET28a-LIC construct (sample 5) under majority of conditions. Of note, the expression levels from pET28a-LIC construct appeared less variable across the strains, in contrast with pET32a-LIC construct. Of all conditions tested, unquestionably, the best expression level was obtained using pET32a-LIC construct in Rosetta 2 (DE3) cells, and this was the condition taken forward for the large-scale production of Rab domain.

Therefore, based on the obtained results it may be concluded that the experiment not only enabled identification of the optimal expression conditions of EF-hand and Rab domains but also demonstrated a clear advantage of the pET32a-LIC over pET28a-LIC vector-based constructs in terms of enabling production of higher yields of soluble recombinant proteins. While the decision for selection of the optimal expression condition for each of the targets was based solely on this criterion, an important caveat of pET32a-

LIC expression, which was not considered at the time, is the considerable size of the combined affinity and solubility tag (17.4 kDa) on the pET32a-LIC-based constructs. The tag is therefore nearly as large as Rab (20.5 kDa) and twice as large as EF-hand (8.8 kDa) domain, which, when cleaved post affinity capture, corresponds to respectively 45% and 66% of the generated protein being effectively thrown away. Particularly in the case of EF-hand domain, high expression of which was still demonstrated using the construct in pET28a-LIC vector in BL21 and SHuffle T7 Express strains, expression using pET32a-LIC construct may be worth reconsidering. Since EF-hand domain produced from pET28a-LIC construct contains a significantly smaller affinity tag (2 kDa), even with slightly lower expression levels achieved per litre of bacterial culture, more protein may be recovered following the purification procedure. Hence, going forward, and especially with regard to expensive isotopic labelling for NMR, using pET28a-LIC constructs for the expression of EF-hand protein may prove more optimal.

3.2.2. Small-scale expression of Coiled-coil domain

Following optimisation of the Coiled-coil domain PCR conditions and successful generation of the expression construct in pET32a-LIC vector, described in section 3.1.3, protein expression was tested small-scale at 25ºC in the same set of bacterial strains as the one used for the other Rab46 domains. As demonstrated in Figure 14, Coiled-coil protein was successfully expressed in every strain tested, although addition of 0.5 mM L-rhamnose appeared to significantly lower its expression levels in Lemo21 (DE3) strain. Despite high background observed in lane 4, SHuffle T7 Express was identified as optimal strain for large-scale production of Coiled-coil demonstrating levels of target protein expression appreciably higher than any other tested strain.

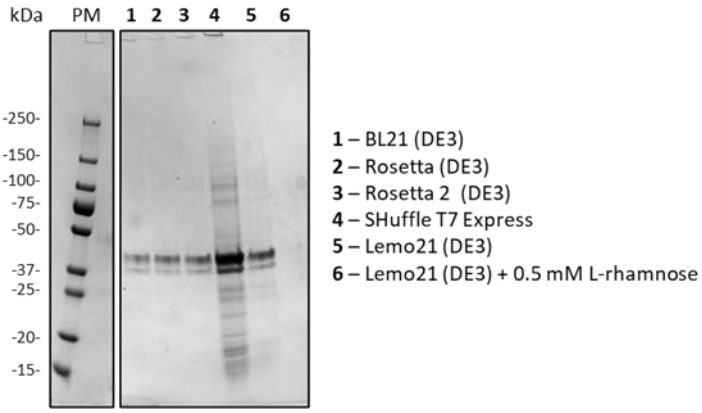

Figure 14 Small-scale expression of Coiled-coil domain. Lane numbering corresponding to different bacterial strains used in the screen is indicated in the figure legend on the right-hand side of the protein gel. PM – protein marker.

3.2.3. Batch expression of FL Rab46

Since expression of FL Rab46 could not be confirmed in the bacterial strains tested in the small-scale expression screen, it was anticipated that the level of expression from 2 mL cultures might have been too low for the protein to be detected by Coomassie dye staining. Furthermore, it was speculated that presence of the bands at the lower than expected molecular weight seen for FL Rab46 constructs in small-scale expression test, could be due to protein degradation. Therefore, expression of pET32a-LIC FL Rab46 construct was examined at larger, 100 mL scale and at a lower temperature (18ºC and 12ºC) using a modified set of bacterial expression host cells. In addition to Rosetta (DE3), Rosetta 2 (DE3) and SHuffle T7 Express cells already tested at 2 mL scale, ArcticExpress (DE3) strain, designed to facilitate eukaryotic protein expression at low temperature (12ºC) was included in the experiment.

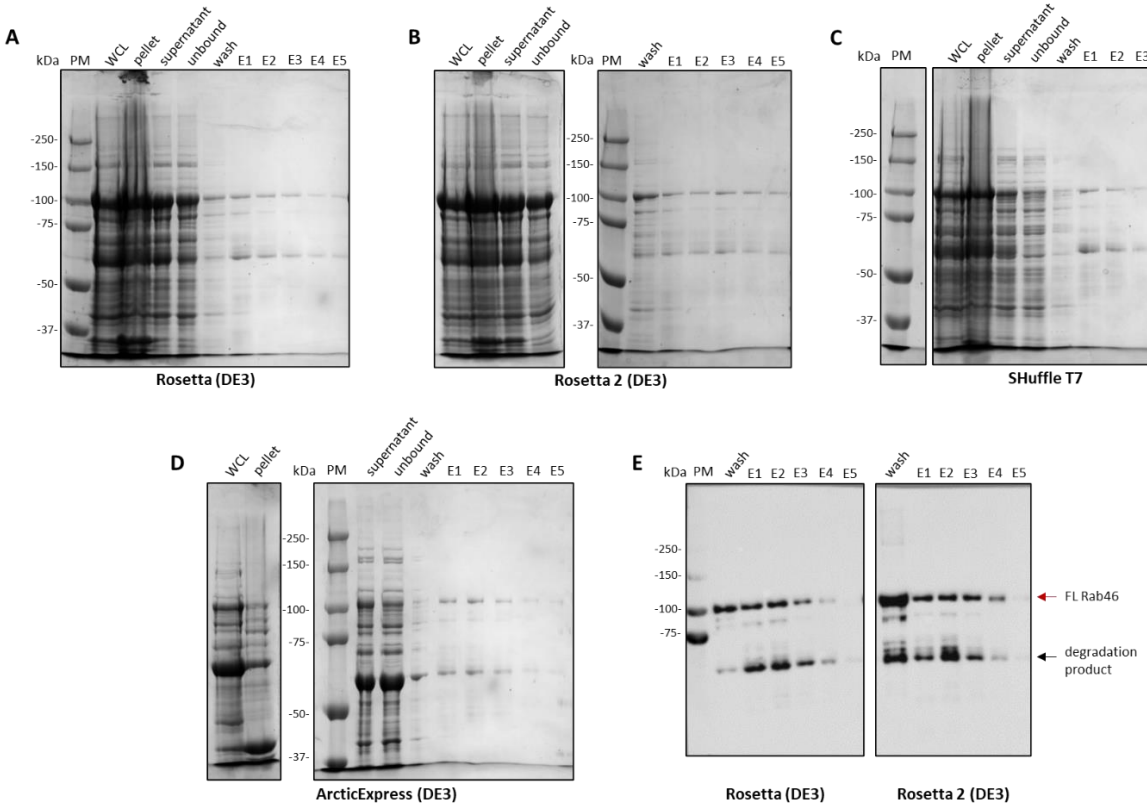

Figure 15 SDS-PAGE analysis of pET32a-LIC FL Rab46 construct expressed at 100 mL scale in different *E. coli* strains. (A-D) Coomassie-stained SDS-PA gels for expression test of FL Rab46 in different *E. coli* strains as indicated below the gels. (E) Western blot analysis of FL Rab46 expressed in Rosetta and Rosetta 2 strains using Rab46-specific antibody. PM – protein marker; WCL – whole cell lysate; E1, E2, E3, E4, and E5 – samples eluted with respectively 80 mM, 160 mM, 240 mM, 320 mM, and 400 mM imidazole present in the elution buffer.

His-tagged FL Rab46 protein was purified in batch mode using Ni-NTA resin. The samples collected at consecutive steps of affinity purification were analysed by SDS-PAGE (Figure 15). A band at approximately 100 kDa was observed for every tested strain, indicating successful expression of FL Rab46. However, a lower molecular weight band at approximately 65 kDa was also present under all conditions, consistent with the result from the previous small-scale expression screen. Furthermore, a considerable proportion of the full length Rab46 was present in the unbound and wash fractions across different strains tested, suggesting that the expressed protein did not bind the nickel resin very tightly. While expression level was comparatively low across all the strains, slightly stronger protein bands were seen in wash and elution lanes for Rosetta and Rosetta 2 cells. The samples from these two most promising conditions were further analysed by western blotting using Rab46-specific antibody, successfully confirming the identity of the 100 kDa band as full length Rab46. Importantly, as the lower molecular weight band was also recognised by the Rab46-specific antibody, western blot analysis provided evidence that the 65 kDa protein was not an unrelated contaminant but a product of Rab46 degradation. This finding, in turn, supported the claim that the most prominent band observed in small-scale screen most likely corresponded to the degraded Rab46 protein. Hence, the potential explanation as to why no band of the correct size was observed for the FL Rab46 at small scale, could be that, since the expression test was performed at a higher temperature, it might have led to a complete proteolytic degradation of the target protein.

Nonetheless, based on the results of the screen, expression in Rosetta 2 cells at 18ºC was identified as suitable for large-scale production of full length Rab46. Considering the fact that FL Rab46 expressed in *E. coli* appeared to be prone to degradation, it was decided that additional protease inhibitors would be included in the lysis buffer to avoid proteolysis of the target protein during large-scale purification.

The summary of conditions determined as optimal for bacterial expression of the Rab46 proteins is shown in Table 10.

Table 10 Optimal *E. coli* strains and expression conditions for production of Rab46 proteins.

Protein	Expression vector	Optimal *E. coli* strain	Antibiotics	Expression temperature
FL Rab46	pET32a-LIC	Rosetta™ 2 (DE3)	Ampicillin (100 µg/mL) Chloramphenicol (34 µg/mL)	18ºC
EF-hand	pET32a-LIC	BL21 (DE3)	Ampicillin (100 µg/mL)	25ºC
Coiled-coil	pET32a-LIC	SHuffle T7 Express	Ampicillin (100 µg/mL) Spectinomycin (50 µg/mL)	25ºC
Rab	pET32a-LIC	Rosetta™ 2 (DE3)	Ampicillin (100 µg/mL) Chloramphenicol (34 µg/mL)	25ºC

3.3. Large-scale expression and purification of Rab46 proteins

3.3.1. EF-hand domain

Unlabelled and isotope-labelled EF-hand domain was expressed at a 2 L scale using conditions identified in section 3.2.1, summarised in Table 10. No additives were required in M9 minimal media for growth of the bacteria transformed with EF-hand expression construct. Protein samples at purity greater than 95%, suitable for structural biology applications, were obtained in a relatively straightforward manner, in a three-step purification – an initial affinity capture step, followed by reverse affinity chromatography to remove the fusion tag cleaved with TEV protease, and a final Size Exclusion Chromatography (SEC) step to ensure high purity and homogeneity of the protein samples. As demonstrated by SEC elution profile of EF-hand domain (Figure 16, panel C) the retention volume of peak 2 at approximately 87 mL on a 16/600 S75 column, is consistent with the EF-hand protein being a monomer (8.8 kDa). Since peak 1 at approximately 78 mL, most likely corresponding to the EF-hand dimer, is not completely resolved from peak 2, care was taken not to pool the dimer and monomer fractions together.

Typically, good yields of approximately 10 mg and 4 mg of the final purified protein per litre of bacterial culture were obtained for unlabelled and isotopically labelled EF-hand domain respectively. Protein identity as EF-hand domain was unambiguously confirmed by mass spectrometry peptide mapping analysis with 100% sequence coverage (Figure 17). An extra serine residue was present at the N terminus following removal of the affinity tag with TEV protease.

Since EF-hand domain does not contain any tryptophan residues, accurate determination of the protein concentration was not possible using standard spectrophotometric methods, relying on measuring the sample absorbance at 280 nm. Therefore, protein concentration of EF-hand samples was routinely estimated using colorimetric DC protein assay (BioRad) using BSA as a protein concentration standard. However, as protein concentration needs to be precisely determined to ensure that reliable results can be obtained from biophysical and structural characterisation efforts, suitability of DC protein assay and BSA standard for accurate determination of EF-hand concentration was assessed by comparing the DC protein assay values with the result obtained from quantitative amino acid analysis (QAA, AltaBioscience). The concentration of the provided EF-hand sample, estimated at 5.2 mg/mL using DC protein assay, was determined as 5.27 mg/mL using QAA hence confirming satisfactory accuracy of the DC protein assay.

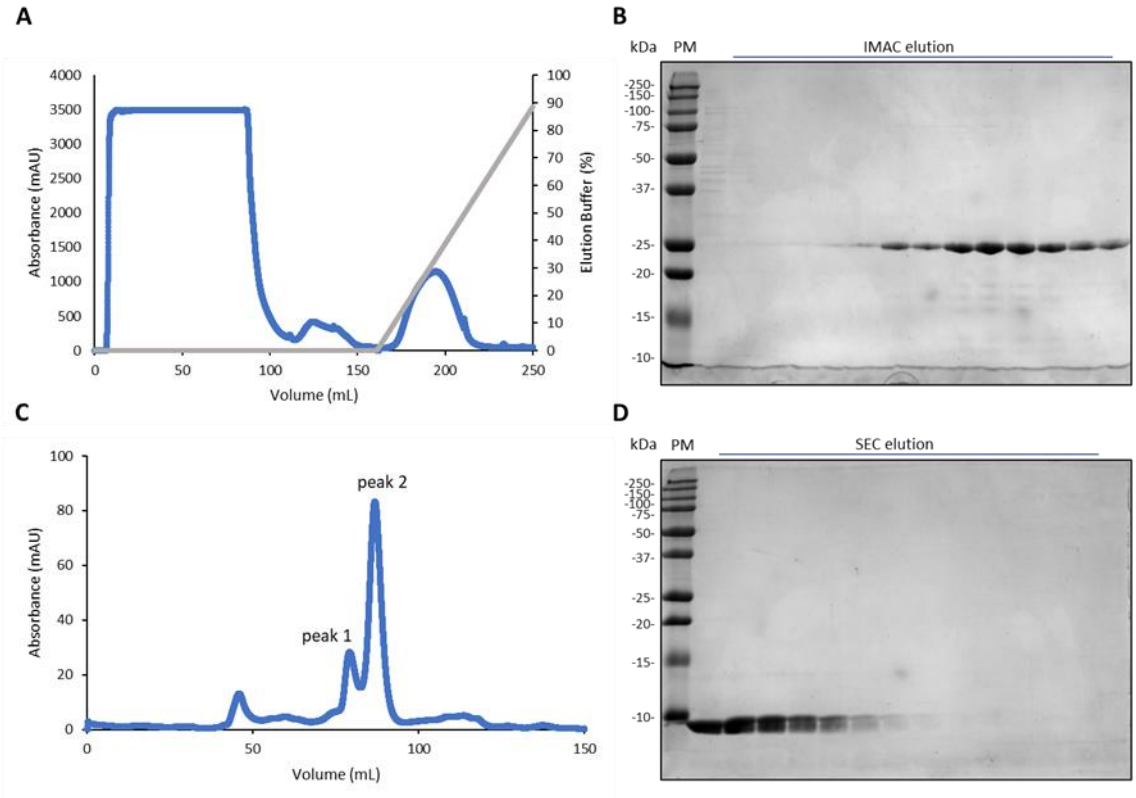

Figure 16 Large-scale expression and purification of EF-hand domain. (A) Chromatogram for affinity capture of unlabelled EF-hand domain. Elution buffer contained 400 mM imidazole. (B) SDS-PAGE analysis of his-tagged, unlabelled EF-hand domain after the capture step. (C) Chromatogram for SEC of unlabelled EF-hand domain. (D) SDS-PAGE analysis of unlabelled EF-hand domain after SEC. PM – protein marker.

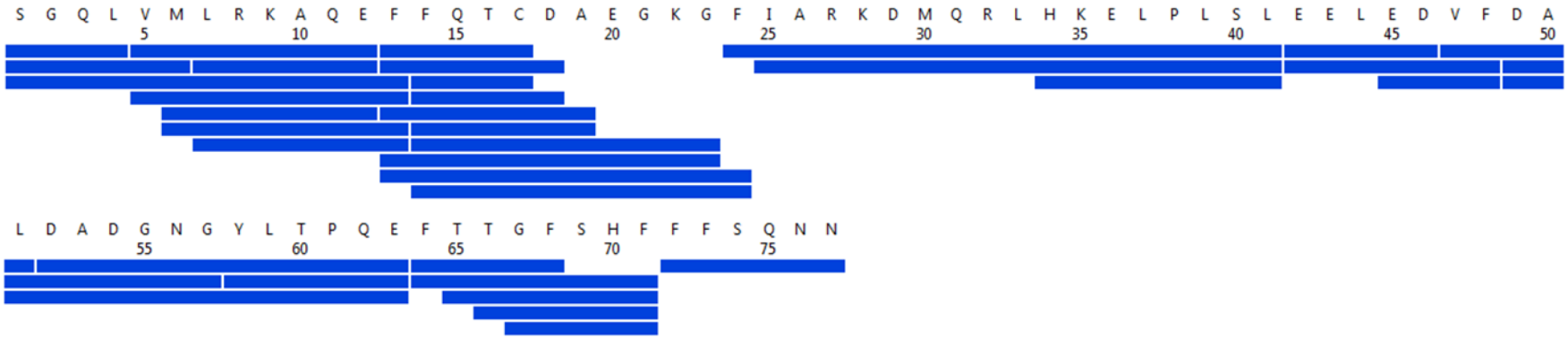

Figure 17 Peptide mapping of EF-hand domain. Schematic diagram illustrating sequence coverage map for peptide ID analysis of EF-hand domain. 100% sequence coverage was achieved, successfully confirming the protein identity.

3.3.2. Large-scale expression and purification of Rab domain

Similarly to EF-hand domain, unlabelled Rab domain was expressed and purified without issues at a 2 L scale. As demonstrated by SEC elution profile in Figure 18 (panel C) the retention volume of peak 3 at approximately 72 mL on a 16/600 S75 column, corresponds to monomeric Rab domain (20.5 kDa). Care was taken to only pool the monomer fractions to avoid mixing them with potential other oligomerisation states of Rab (peaks 1 and 2). Following the same three-step purification strategy as for EF-hand, yields of 20 mg of the final, untagged, 95% pure protein per litre of bacterial culture were typically obtained. Protein identity was successfully confirmed by mass spectrometry peptide mapping analysis, showing 97% sequence coverage (Figure 19).

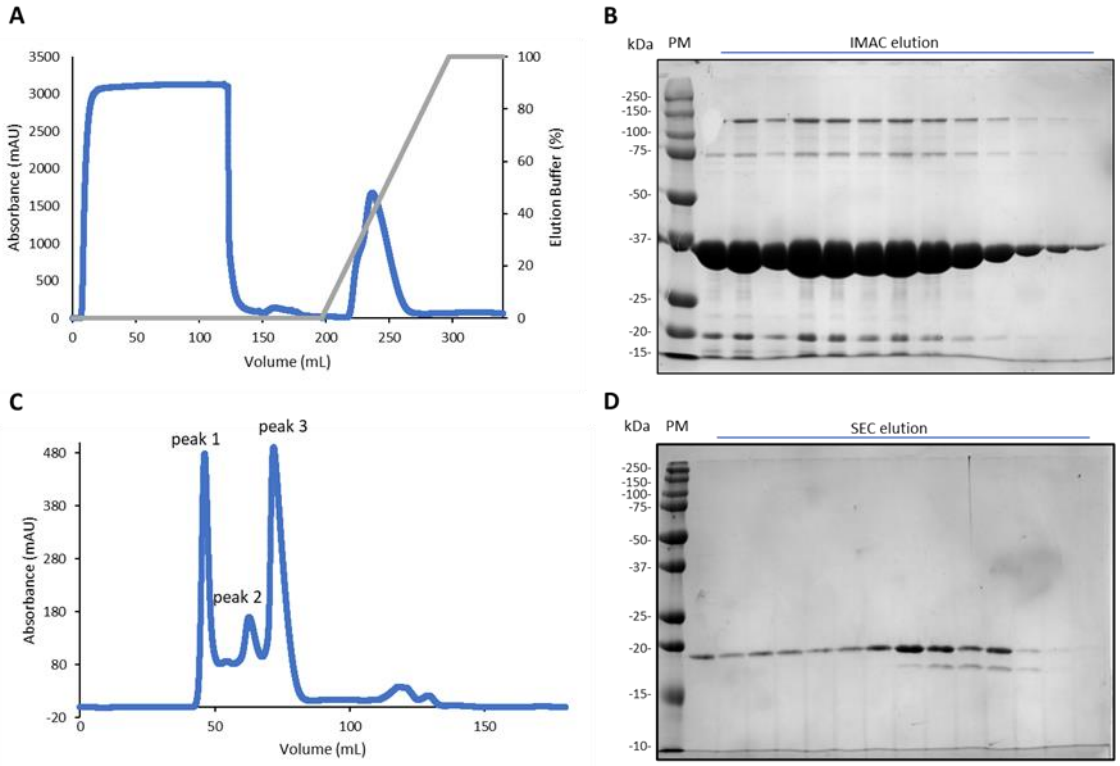

Figure 18 Large-scale expression and purification of Rab domain. (A) Chromatogram for affinity capture of unlabelled Rab domain. Elution buffer contained 400 mM imidazole. (B) SDS-PAGE analysis of his-tagged, unlabelled Rab domain after the capture step. (C) Chromatogram for SEC of unlabelled Rab domain. (D) SDS-PAGE analysis of unlabelled Rab domain after SEC. PM – protein marker.

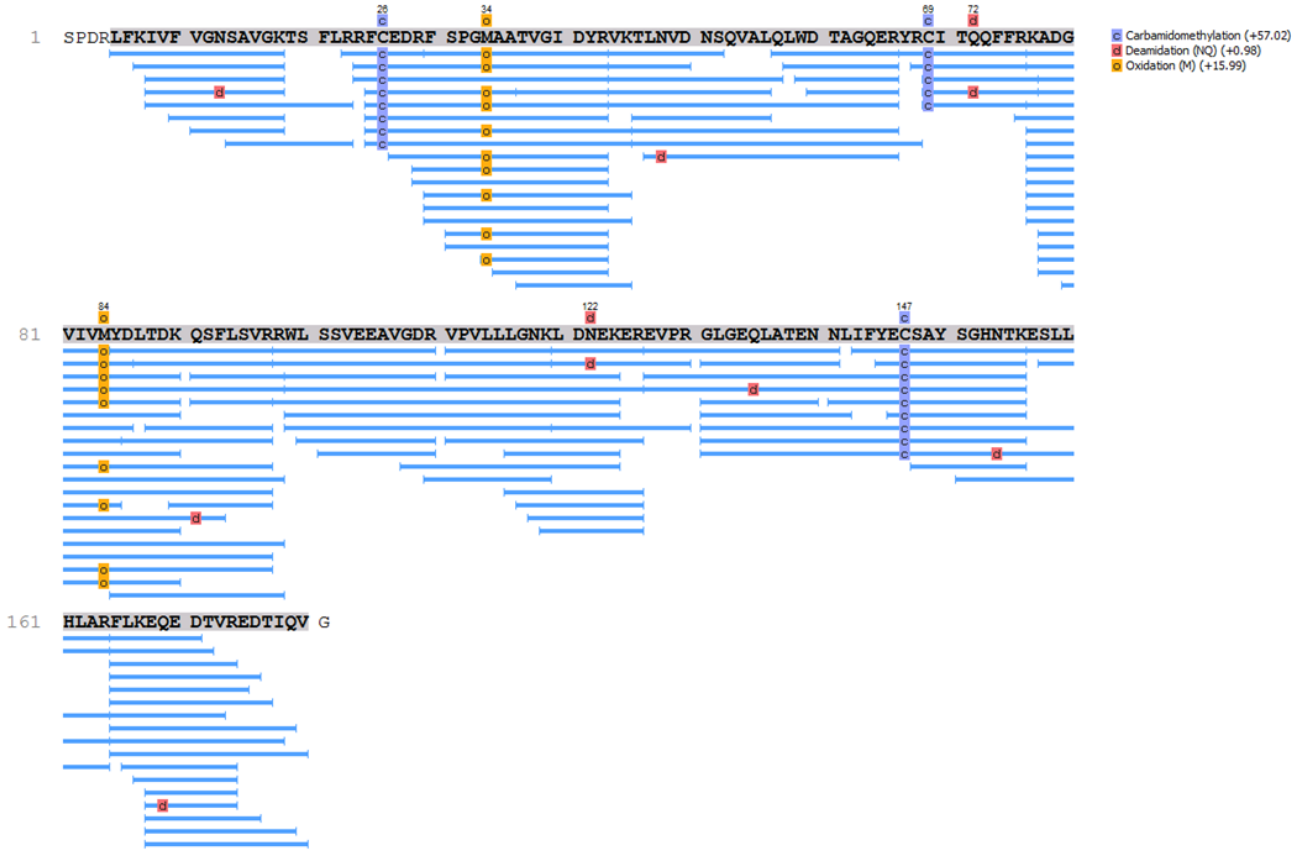

Figure 19 Peptide mapping of Rab domain. Schematic diagram illustrating sequence coverage map for peptide ID analysis of Rab domain. 97% sequence coverage was achieved, successfully confirming the protein identity. The identified post-translational modification sites are highlighted as per legend on the right-hand side of the diagram.

3.3.3. Labelled expression of Rab domain

Despite the high yields of unlabelled Rab domain, expression of ^{15}N-labelled Rab domain was not successful (Figure 20, panel A). Rosetta 2 cultures transformed with Rab domain expression construct exhibited poor growth in M9 minimal medium, which appeared to be the main reason for the dramatically low expression of the target protein. Therefore, various modifications to the expression protocol, described in section 2.3.5, were tested, including growing the overnight starter cultures in a nutritionally richer TB medium and addition of 5-10% LB broth to M9 medium. While the former did not appear to have any effect, the latter helped to increase the expression levels of the target protein. Nevertheless, yields comparable to unlabelled Rab domain were only achieved when the composition of M9 medium was adjusted by addition of Celtone Complete Medium, BME vitamins and trace elements, and increasing the glucose concentration. Following optimisation of the M9 medium composition, ~18 mg of 95% pure, ^{15}N-labelled Rab domain were obtained from 1 L of bacterial culture.

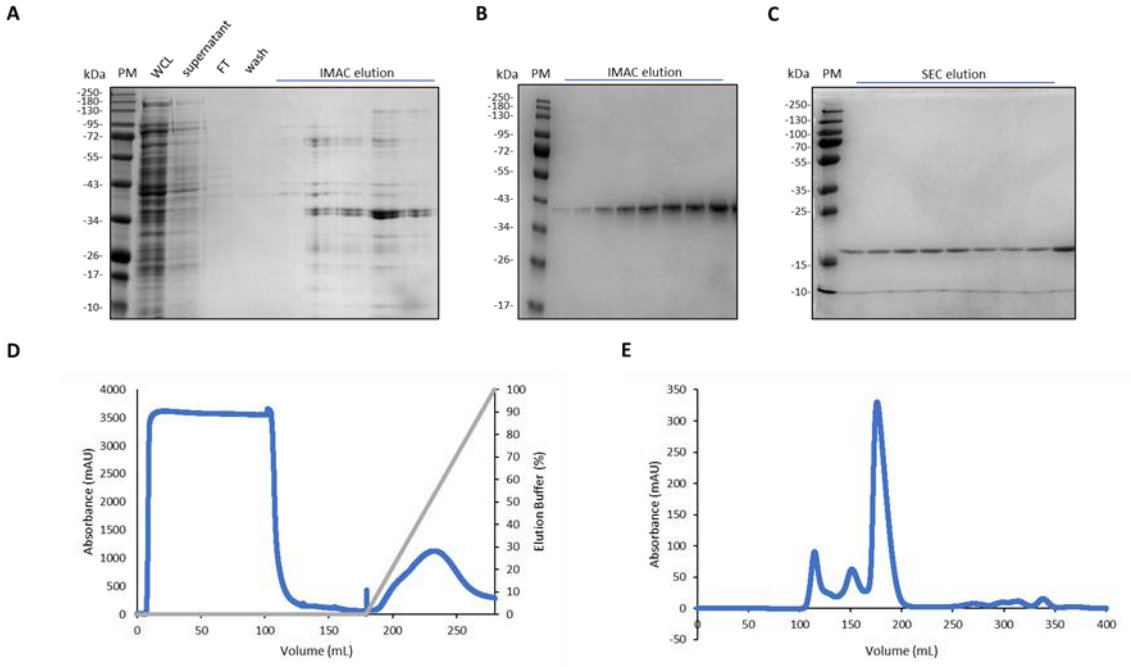

Figure 20 Large-scale expression and purification of ^{15}N-labelled Rab domain. (A) Failed expression attempt of ^{15}N-labelled Rab domain in M9 minimal medium. (B) SDS-PAGE analysis of his-tagged, ^{15}N-labelled Rab domain after the capture step, following the optimisation of M9 medium composition. (C) SDS-PAGE analysis of ^{15}N-labelled Rab domain after SEC. (D) Chromatogram for affinity capture of ^{15}N-labelled Rab domain. (E) Chromatogram for SEC of ^{15}N-labelled Rab domain. PM – protein marker**.**

3.3.4. Large-scale expression of full length Rab46

Based on the result of the 100 mL expression screen, an attempt was made to express and purify the full length Rab46 protein from 8 L Rosetta 2 cells. In addition to the standard broad-spectrum protease inhibitors, PMSF and benzamidine were included in the lysis buffer and great care was taken to keep the sample cold throughout the purification procedure to minimise proteolytic degradation of recombinant Rab46. Unfortunately, despite these extra measures, the protein still suffered from significant proteolysis. As shown in Figure 21, right after affinity capture step a significant amount of lower molecular weight product was present in the protein sample. Despite some degree of separation of the low molecular weight contaminants from the FL protein on the S200 column, the protein was not stable and precipitated during the concentration step. A sample of poor quality and purity was obtained following the large-scale purification.

Optimisation of the expression and purification of FL Rab46 from bacteria was attempted with no success. Different buffer composition and incorporation of an extra ion exchange step in the purification method were tested but failed to improve the final sample quality. Thus, further attempts to produce full length Rab46 protein in *E. coli* were aborted and a decision was made to switch to a different expression system, aiming to produce a more stable, well-behaved protein for biophysical and structural characterisation.

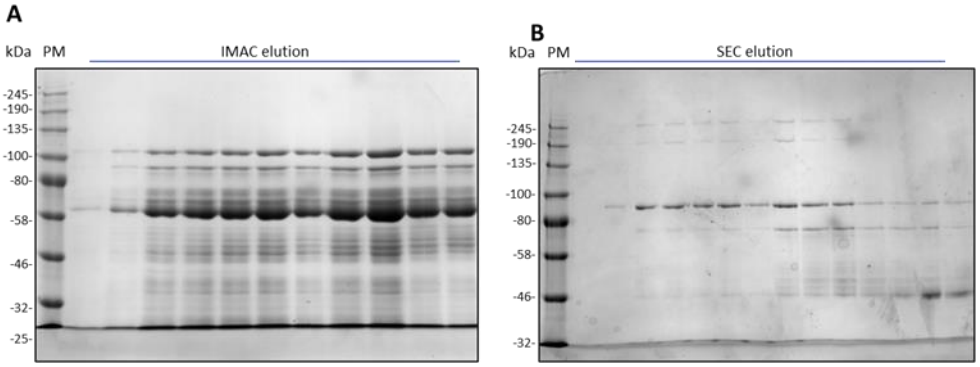

Figure 21 Large-scale bacterial expression of FL Rab46. SDS-PAGE analysis of FL Rab46 purified from 8 L Rosetta 2 cells after (A) affinity capture and (B) SEC.

3.3.5. Insect cell expression of full length Rab46

Since production of the full length Rab46 proved unsuccessful in *E. coli,* a different strategy to obtain the protein for structural characterisation had to be developed. One of the important limitations of *E. coli* expression system is its inability to provide complex

post-translational modifications, which are often crucial for stability and functionality of many protein targets. Therefore, baculovirus expressing vector system (BEVS) was chosen to address the challenges with full length Rab46 production. Being based on insect cell lines, BEVS not only facilitates obtaining majority of post-translational modifications found in mammalian cells, but also offers a possibility to be adapted for production of labelled proteins for NMR studies of Rab46 interdomain interactions.

3.3.6. Molecular cloning

A construct for production of his-tagged FL wt Rab46 in insect cells was generated using the same strategy as for the pET32a-LIC and pET28a-LIC constructs described in section 3.1.2. Bacmid DNA for generation of recombinant baculovirus was produced following Bac-to-Bac manufacturer's protocol.

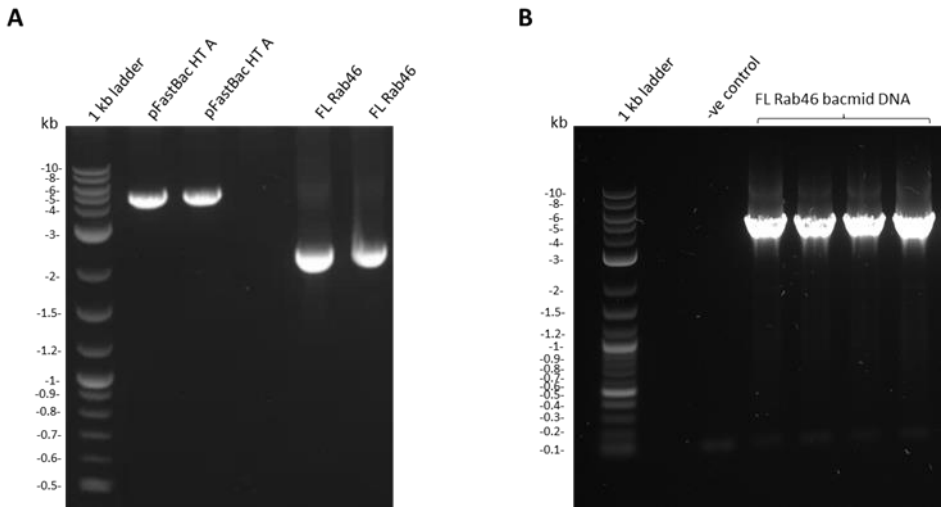

Figure 22 Generation of construct for insect cell expression of FL Rab46. Agarose gel of pFastBacHT A vector linearised with SfoI and HindIII, and DpnI-digested, gel purified FL Rab46 PCR product. Sizes of the obtained products match the expected sizes of 4.7 kb for linearised pFastBacHT A plasmid and 2.2 kb for PCR-amplified FL Rab46. (B) Diagnostic PCR confirming successful bacmid transposition with FL Rab46 construct using pUC/M13 universal primers. The observed band appears to have the correct size of 4.6 kb.

As demonstrated in Figure 22, pFastBacHT A plasmid was successfully digested with SfoI and HindIII restriction endonucleases, and FL Rab46 was amplified by PCR to produce a construct for generation of recombinant bacmid DNA. Successful bacmid

transposition with FL Rab46 pFastBac construct was verified by diagnostic PCR using pUC/M13 universal primers (Figure 22, panel B).

3.3.7. Quantification of FL Rab46 baculovirus titre

FL Rab46 P1 and P2 recombinant baculovirus were generated in *Sf9* cells. For better accuracy, prior to expression attempts the viral titre was estimated using two different quantification methods – droplet digital PCR (ddPCR) and RT qPCR using GP64 primers (Figure 23).

The analysis demonstrated high titre of the recombinant baculovirus of 1.95×10^9 pfu/mL and 2.98×10^9 pfu/mL as determined by RT qPCR and ddPCR respectively. Since the obtained results are within the same order of magnitude, there is a good level of agreement between the viral titre estimation by the two methods.

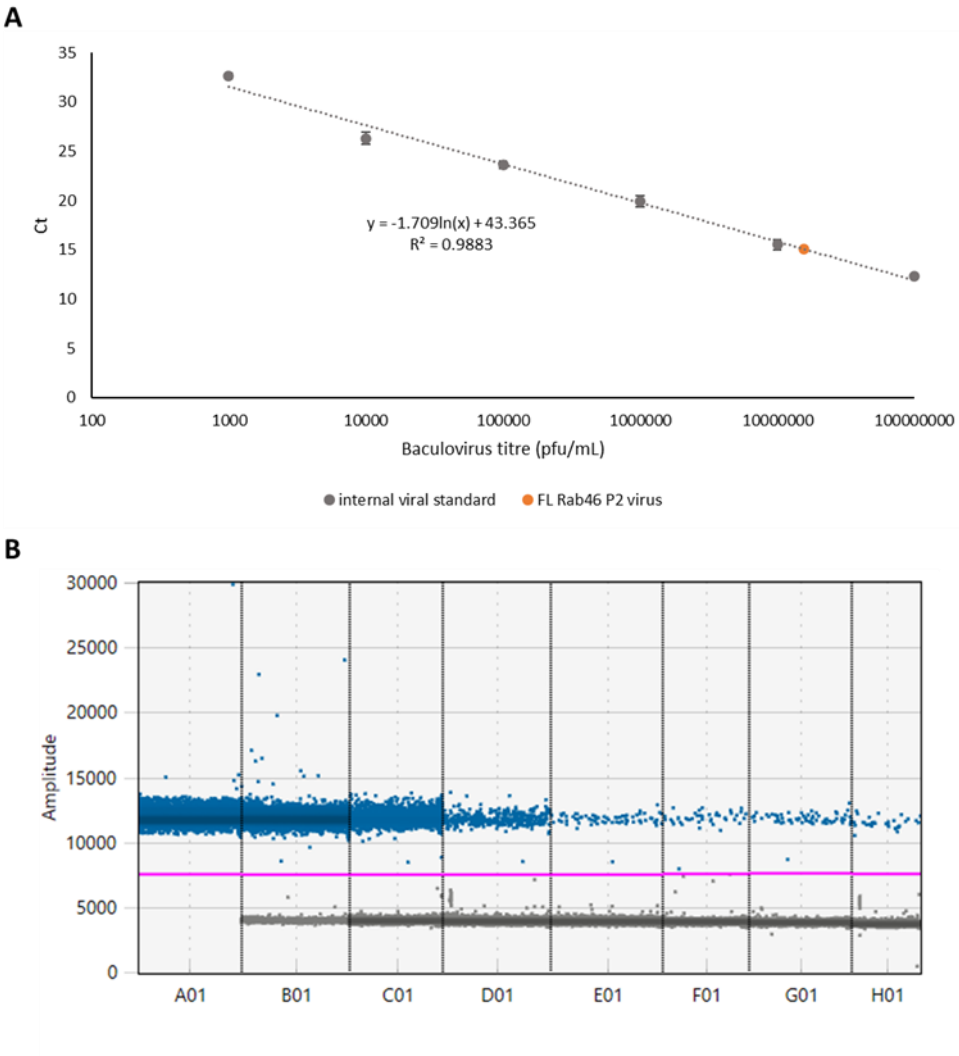

Figure 23 Quantification of FL Rab46 baculovirus titre using RT qPCR and droplet digital PCR. (A) RT qPCR plot for quantification of FL Rab46 baculovirus titre using GP64 primers. Standard curve was generated using the internal viral standard included in the baculoQUANT kit (B) Diagram for droplet digital PCR quantification of P2 FL Rab46 baculovirus titre using GP64 primers. Labels A-G on the y-axis refer to serially diluted isolated viral DNA (10^{-1}-10^{-7}). Sample H is the negative control (no viral DNA).

3.3.8. Small-scale insect cell expression test

To find optimal conditions for expression of FL Rab46, expression was tested in 30 mL *Sf9* cultures. Different harvest points post infection (time of infection – TOI) were first tested using a single multiplicity of infection (MOI) of 7. As demonstrated in Figure 24, panel A, there was no evidence for protein expression at 24 hrs, and the expression only became apparent at 48 hrs post infection. At 72 hrs the expression was substantially decreased, and completely lost at 96 hrs. Having found the optimal harvest point, different amounts of virus (MOI) of 3, 5, 7 and 10 were tested to see if the expression levels of the protein can be enhanced. While time of harvest was of significance, no noticeable difference was observed between different MOIs (Figure 24, panel B)

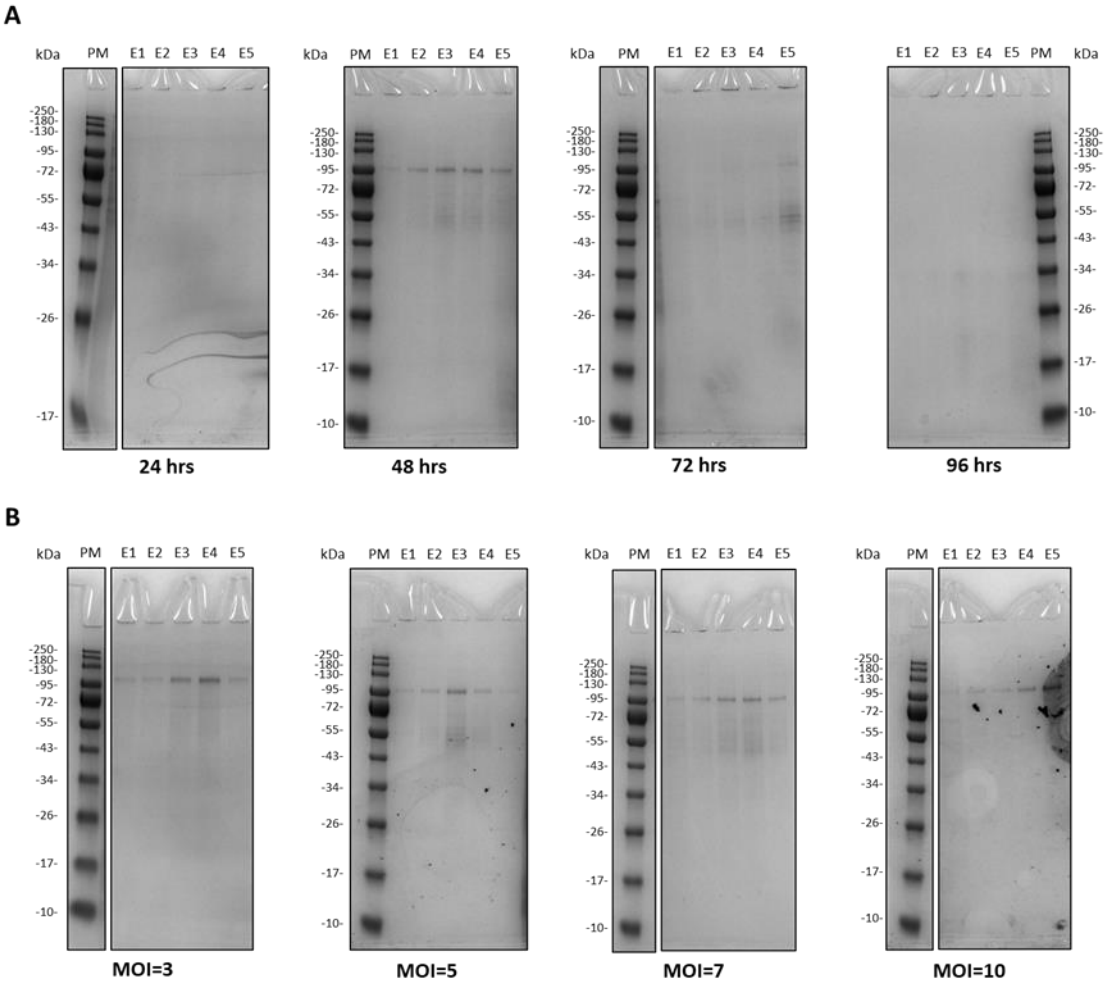

Figure 24 Small-scale expression test of FL Rab46 in *Sf9* cells. (A) Expression levels of FL Rab46 at different harvest points. (B) Expression levels of FL Rab46 at different MOI. PM – protein marker; E1, E2, E3, E4, and E5 – samples eluted with respectively 80 mM, 160 mM, 240 mM, 320 mM, and 400 mM imidazole present in the elution buffer.

3.4.9. Large-scale insect cell expression of FL Rab46

FL Rab46 was expressed in *Sf9* cells using MOI of 5 at a 3 L scale. The histidine-tagged protein was first captured using 1 mL HisTrap column (Figure 25, panels A and B). Since there was a lot of contaminants present in the protein sample after the affinity capture step, it was further purified by SEC. Although SDS-PAGE analysis revealed that the more distinct peak 1 in the SEC elution profile (Figure 25, panel C) contained the protein of interest, retention volume of 40 mL corresponds to the void volume of the S200 16/600 gel filtration column. Elution in the void volume is explained by the molecular weight of the protein being larger than the cutoff size for the column, which in case of Superdex 200 column corresponds to a MW of 600 kDa, indicating that the portion of Rab46 protein eluting in peak 1 was likely aggregated. The rest of non-resolved peaks in SEC, eluting within the separation range of the column, labelled as ‘peak 2’ fractions, contained FL Rab46 not successfully separated from the contaminating proteins.

Western blotting analysis successfully confirmed the protein identity as full length Rab46. However, attempts to confirm mass by Mass Spectrometry failed, likely due to the issues with protein aggregation.

Compared to the full length protein expressed in bacteria, the purity of the protein produced in insect cells was significantly improved. Nevertheless, protein stability was poor, and the sample was prone to aggregation and precipitation. Various buffer additives and modifications of the purification method were tested aiming to obtain a more stable protein sample. Unfortunately, the attempts proved unsuccessful with the protein consistently eluting in the void volume of the S200 gel filtration column. No suitable sample of full length Rab46 protein was therefore obtained for structural characterisation.

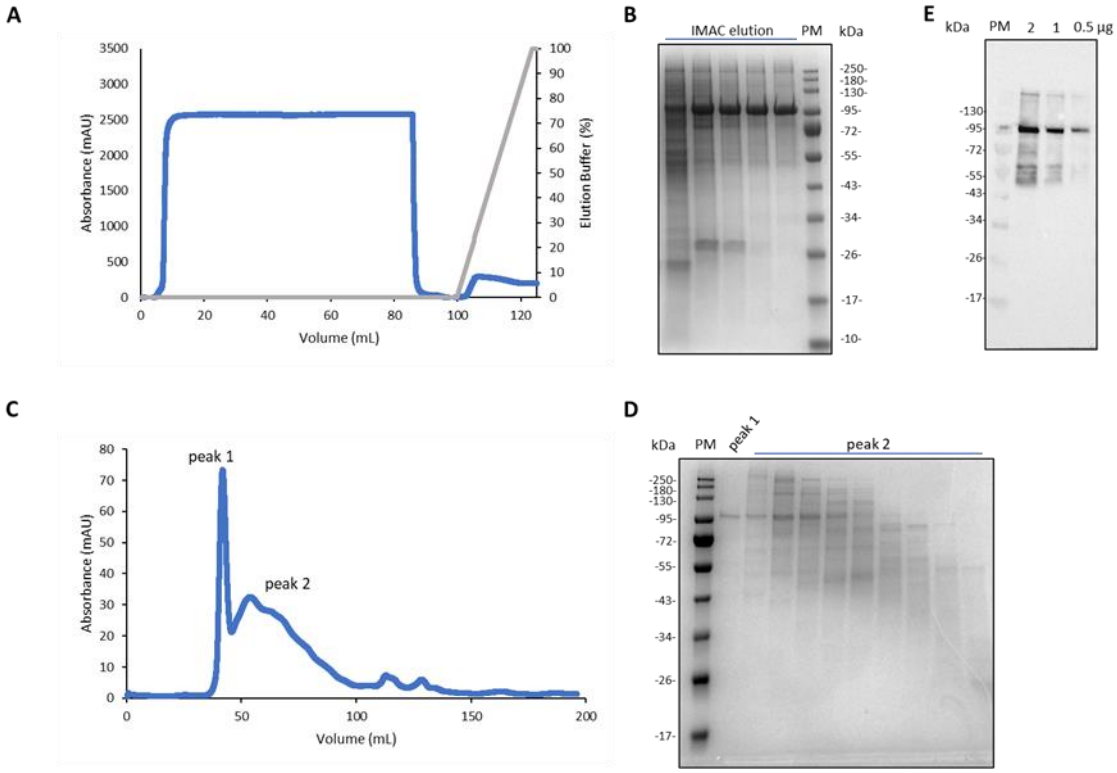

Figure 25 Large-scale insect expression and purification of FL Rab46. (A) Chromatogram for affinity capture of his-tagged FL Rab46. (B) SDS-PAGE analysis of his-tagged Rab46 after the capture step. (C) Chromatogram for SEC of his-tagged FL Rab46. (D) SDS-PAGE analysis of his-tagged FL Rab46 after SEC. (E) Western blot analysis of FL Rab46 expressed in *Sf9* cells using Rab46-specific antibody. PM – protein marker.

Chapter 4

Characterisation of EF-hand domain of Rab46

4.1. Introduction

Since at the time this study was commenced there had been no structural studies of Rab46 GTPase, the aim of this project was to investigate its structure, interactions and dynamics. However, due to the obstacles encountered with the production of the full length protein described in Chapter 3, the project eventually focused on the structural and biophysical characterisation of the isolated functional domains of Rab46. Although the EF-hand motif is a common, extensively studied structural motif, it has been demonstrated to exhibit a great diversity of structural responses to Ca^{2+}-binding, domain organisation and target interaction properties in different EF-hand containing proteins. Considering the role of Ca^{2+} binding in the Rab46-dependent WPB dispersal from the MTOC, understanding of the structural and Ca^{2+}-binding properties of EF-hand domain of Rab46 is of interest, to gain more insight into the Ca^{2+}-dependent regulation of Rab46 activity.

4.2. Protein fold and backbone assignment of EF-hand domain

4.2.1. Circular Dichroism

The recombinant wt EF-hand domain purified from *E. coli* (expression of which was described in Chapter 3) was confirmed to be folded by Far-UV Circular Dichroism (CD) spectroscopy. CD spectroscopy is invaluable for obtaining initial structural information for a protein of interest, enabling quick examination of its secondary or tertiary structure, as well as conformational changes and stability in solution. Far-UV protein CD spectra measured between 180 – 240 nm, arising predominantly from differential absorption of left- and right-circularly polarised light by peptide bonds, provide information about the types of secondary structure elements present in a protein (110). Different secondary structure elements exhibit characteristic absorption values, which, in turn, gives rise to distinctive CD spectra. For instance, spectra of primarily α-helical proteins are characterised by two minima at 208 nm and 222 nm, and a maximum at 193 nm. Thus, the acquired Far-UV CD spectrum of apo wt EF hand (Figure 26), plotted as molar circular dichroism (Δε) against wavelength, clearly indicates high α-helical content, as expected for an EF-hand domain consisting of two helix-loop-helix EF-hand motifs. α-helical content calculated from the CD spectrum using DichroWeb tool (Table 11, (104)) was 63% (48% corresponding to regular and 15% to distorted helix i.e. referring to the

two residues at either ends of α-helix), which is in good agreement with the secondary structure prediction from PsiPred tool (Figure 26, (111)).

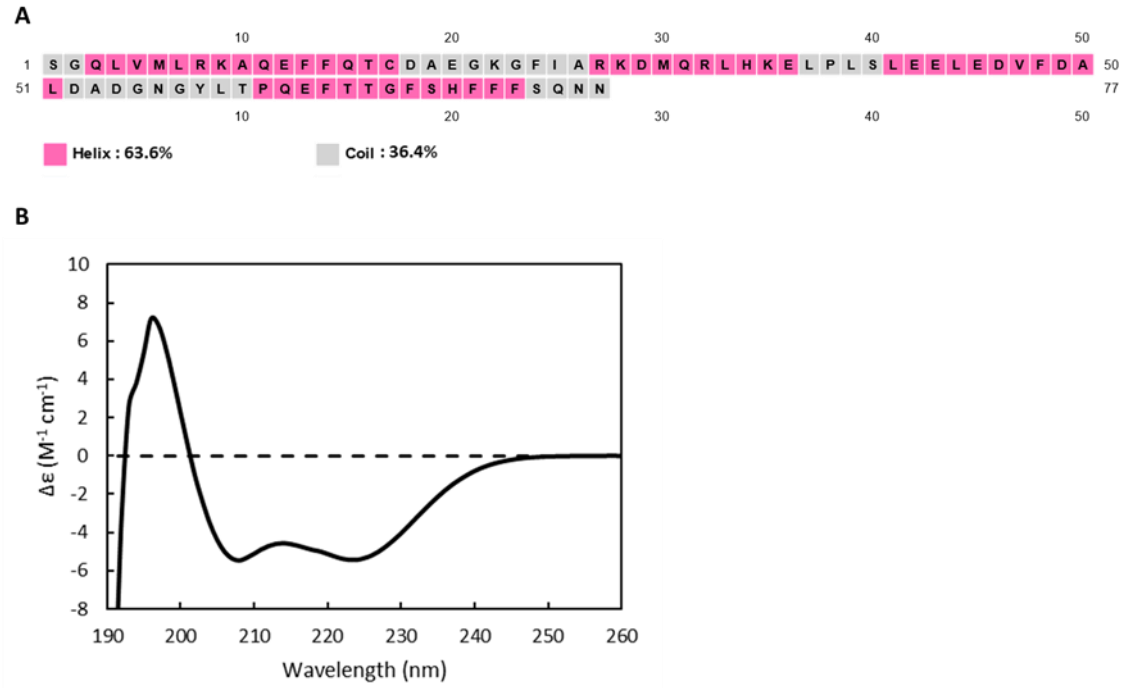

Figure 26 Far-UV CD of wt EF-hand domain. A) Secondary structure of wt EF-hand domain predicted using PsiPred tool (111). Calculated percentage of each of the secondary structure elements is shown below the sequence plot. B) Far-UV CD spectrum of apo wt EF-hand domain in 50 mM sodium phosphate buffer (pH 7.2) containing 50 mM NaCl. The spectrum was acquired over the wavelength range of 260 – 180 nm. Data are plotted as molar circular dichroism (Δε) versus wavelength.

Table 11 Summary of the secondary structure content of wt EF-hand domain calculated using DichroWeb CDSSTR method (104).

Helix1	Helix2	Strand1	Strand2	Turns	Unordered
48%	15%	5%	5%	9%	18%

4.2.2. ^{1}H-^{15}N HSQC of wt EF-hand domain

Having confirmed the recombinantly produced EF-hand domain was folded, isotopically labelled samples were produced for further characterisation of the protein by NMR spectroscopy. NMR spectroscopy is a powerful biophysical technique for studying protein structure and dynamics (112). One of the indisputable strengths of NMR is that it enables monitoring of the protein dynamics, its structural changes and interactions at an atomic level. A simple 2D Heteronuclear Single-Quantum Coherence (HSQC) experiment not only provides information on whether the protein of interest is folded but also whether it is stable (and hence suitable for further experiments e.g. for the structure determination) and subject to conformational rearrangements. Since ^{1}H-^{15}N HSQC experiment provides the correlations between the chemical shift of nitrogen and the attached amide proton, the cross peaks observed in an HSQC spectrum, to a large extent, correspond to backbone amide groups in the protein (and thus one signal is expected for each amino acid residue, with the exception of the N-terminal and proline residues which do not appear in an HSQC due to exchange with the solvent and absence of an amide proton respectively). What is more, additional peaks are observed in the HSQC for side-chain groups of residues containing nitrogen-bound protons such as tryptophan, asparagine or glutamine. While arginine and lysine side-chain groups also contain nitrogen-bound protons, their side-chain NH often exchange too fast or are extremely broadened and outside the ^{15}N spectral width, and thus are not always observed. Likewise, histidine indole NH peaks do not appear in HSQC spectrum as, in protonated histidine residues, these are beyond the excitation bandwidth when the carrier is optimised for the backbone NH groups, and thus are typically not excited.

^{1}H-^{15}N HSQC spectra of apo and Ca^{2+}-saturated EF-hand were acquired to investigate the effect of calcium-binding on the wtEF-hand domain of Rab46. While spectra of both apo and Ca^{2+}-saturated EF-hand show good chemical shift dispersion consistent with well-folded proteins (Figure 27), the appearance of the two spectra is appreciably different. The results of the experiment, therefore, provided initial evidence that EF-hand is not only functional and binds calcium, but also undergoes a conformational change upon calcium binding, as illustrated by widespread chemical shift perturbation.

As EF-hand domain of Rab46 is a 77-residue protein containing two prolines, and no tryptophan residues (which could be confused with downfield backbone NH signals), approximately 74 cross peaks are expected in its HSQC spectrum. Compared to the apo form, there are far more signals observed in the spectrum of the Ca^{2+}-saturated EF-hand (72 and 106 peaks respectively, excluding the obvious side-chain peaks in the top right

corner of the spectrum), which may be indicative of Ca^{2+}-bound form exhibiting slow conformational exchange between different states. The possibility of the sample not being fully saturated was examined. Addition of $CaCl_2$ up to 40 mM final concentration (80-fold molar excess), however, did not have any effect on the number of peaks in the spectrum (data not shown). While other possible explanations for higher than expected number of peaks include presence of impurities in the sample or protein degradation, the extra peaks are only apparent following addition of $CaCl_2$ to the sample, with the number of signals observed for apo form being in good agreement with that expected for the domain. Furthermore, protein degradation does not seem to be the cause for the greater number of peaks appearing in the spectrum as, typically, it leads to generation of additional peaks in the downfield ^{15}N region of the HSQC, which is not the case for the spectrum of Ca^{2+}-bound EF-hand. Lastly, SDS-PAGE analysis of the sample did not demonstrate presence of any contaminating proteins, and hence it is likely that the additional peaks are indeed due to the Ca^{2+}-bound EF-hand domain existing in multiple conformations. Moreover, as Ca^{2+}-saturated form is characterised by generally broader signals than the apo form, it might further support the hypothesis that the protein is sampling a different conformation or a transiently oligomeric state.

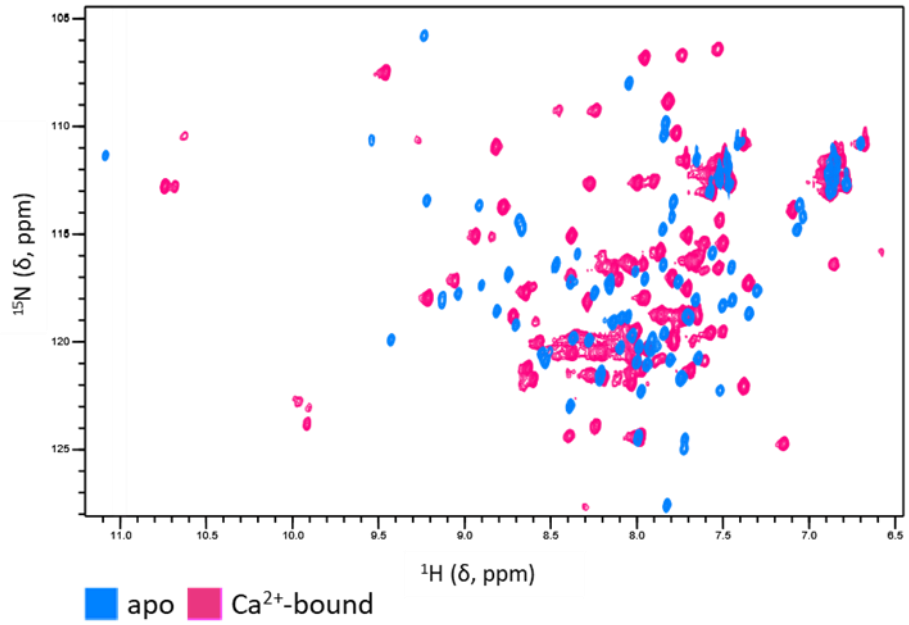

Figure 27 EF-hand domain conformational changes upon calcium binding. Overlay of two unassigned ^{1}H–^{15}N HSQC spectra of apo and Ca^{2+}-saturated forms of EF-hand domain of Rab46. The spectra were recorded using 0.5 mM EF-hand domain uniformly labelled with ^{15}N in buffer containing 25 mM Hepes (pH 7.0), 150 mM NaCl, 1 mM TCEP and 0.1% NaN_3. The data were acquired on a 750 MHz Bruker spectrometer at 25ºC. The apo form of EF-hand domain was prepared by dialysis against 25 mM EDTA, whereas the Ca^{2+}-bound form was saturated with 5 mM $CaCl_2$. The colour code of the two spectra is indicated in the bottom left corner.

4.2.3. Triple resonance backbone assignment of apo and Ca^{2+}-bound EF-hand

Resonance assignment is a prerequisite for the NMR studies of protein structure, dynamics and protein-ligand interactions. Although a good deal of information about the overall state of the EF-hand protein, its calcium binding properties and the resulting conformational change could be obtained from an unassigned HSQC spectrum, completion of resonance assignment was essential to enable a more comprehensive analysis and interpretation of the observed results. Considering that the spectra of apo and Ca^{2+}-bound EF-hand differed quite significantly, the resonances were assigned for both forms independently. While it is often possible to trace assignment from one form to the other if the binding is in fast exchange (by means of performing a ligand titration into the protein), the NMR calcium titration experiment (described in section 4.4 below) demonstrated a mixed exchange regime with a combination of resonances in fast,

intermediate and slow exchange, rendering this approach not possible. A standard set of triple resonance experiments for backbone chemical shift assignment (^{1}HN, ^{15}N, ^{13}Cα, ^{13}Cβ, ^{13}CO) including HNCO, HN(CA)CO, HNCA, HN(CO)CA, HN(CO)CACB, and HNCACB was therefore collected using ^{13}C,^{15}N-labelled apo and Ca^{2+}-saturated EF-hand samples.

In principle, triple resonance backbone assignment of a protein of interest is typically based on the HNCACB and HN(CO)CACB spectra, with HNCACB experiment providing strong correlations for each NH group to the $Cα_i$ and $Cβ_i$ chemical shifts (in antiphase to one another) of the amino acid residue the given NH group belongs to (i), and weaker correlations of the NH group with the $Cα_{i-1}$ and $Cβ_{i-1}$ of the preceding (i-1) residue. HN(CO)CACB experiment which correlates the NH group with the $Cα_{i-1}$ and $Cβ_{i-1}$ carbons only is then used in conjunction with HNCACB to enable identification of the 'i' and 'i-1' peaks. Since the length of the delay τCC (tauCC) used in the HNCACB experiment was adjusted to enhance sensitivity/select for Cβ correlations, HNCA and HN(CO)CA experiments were recorded to compensate for loss of sensitivity for the α-carbons in the HNCACB spectrum. In addition, HNCO and HN(CA)CO spectra were acquired to obtain assignment of CO chemical shifts and to further help establish sequential connectivity between neighbouring amino acid residues. The characteristic values of Cα and Cβ chemical shifts were then used to identify the amino acid types and to guide the sequence-specific assignment. Since, to date, there have been no NMR studies of Rab46 or any of its functional domains, this is the first report of backbone resonance assignment of apo and Ca^{2+}-saturated EF-hand domain of Rab46.

4.2.4. Apo EF-hand domain

Triple resonance experiments for backbone assignment of apo EF-hand domain were acquired on a 750 MHz Bruker spectrometer at 25ºC, using 0.5 mM sample dialysed against NMR buffer (25 mM Hepes [pH 7.0], 150 mM NaCl, 1 mM TCEP, 0.1% NaN_3) containing 25 mM EDTA. Figure 28 shows representative sequential assignment walk used to obtain the backbone chemical shift assignments of apo EF-hand. The numbering of the residues in the strip plots and assigned HSQC spectra corresponds to an isolated EF-hand domain i.e. G2 residue represents G47 in the full length Rab46 protein. The assigned ^{1}H-^{15}N HSQC is presented in Figure 29. 94.6% of the amino acid sequence could be unambiguously assigned for the apo EF-hand domain (70 out of 74 expected residues). The missing assignments were for residues G2, K22, G23, and H70. In summary, 70 ^{15}N, 75 ^{13}Cα, 69 ^{13}Cβ, and 75 ^{13}CO chemical shifts were assigned for apo EF-hand. Chemical shift referencing and resonance assignment were assessed using

PANAV software (94). No mis-assigned resonances were identified neither for apo nor Ca^{2+}-bound EF-hand domain (described in section 4.2.8).

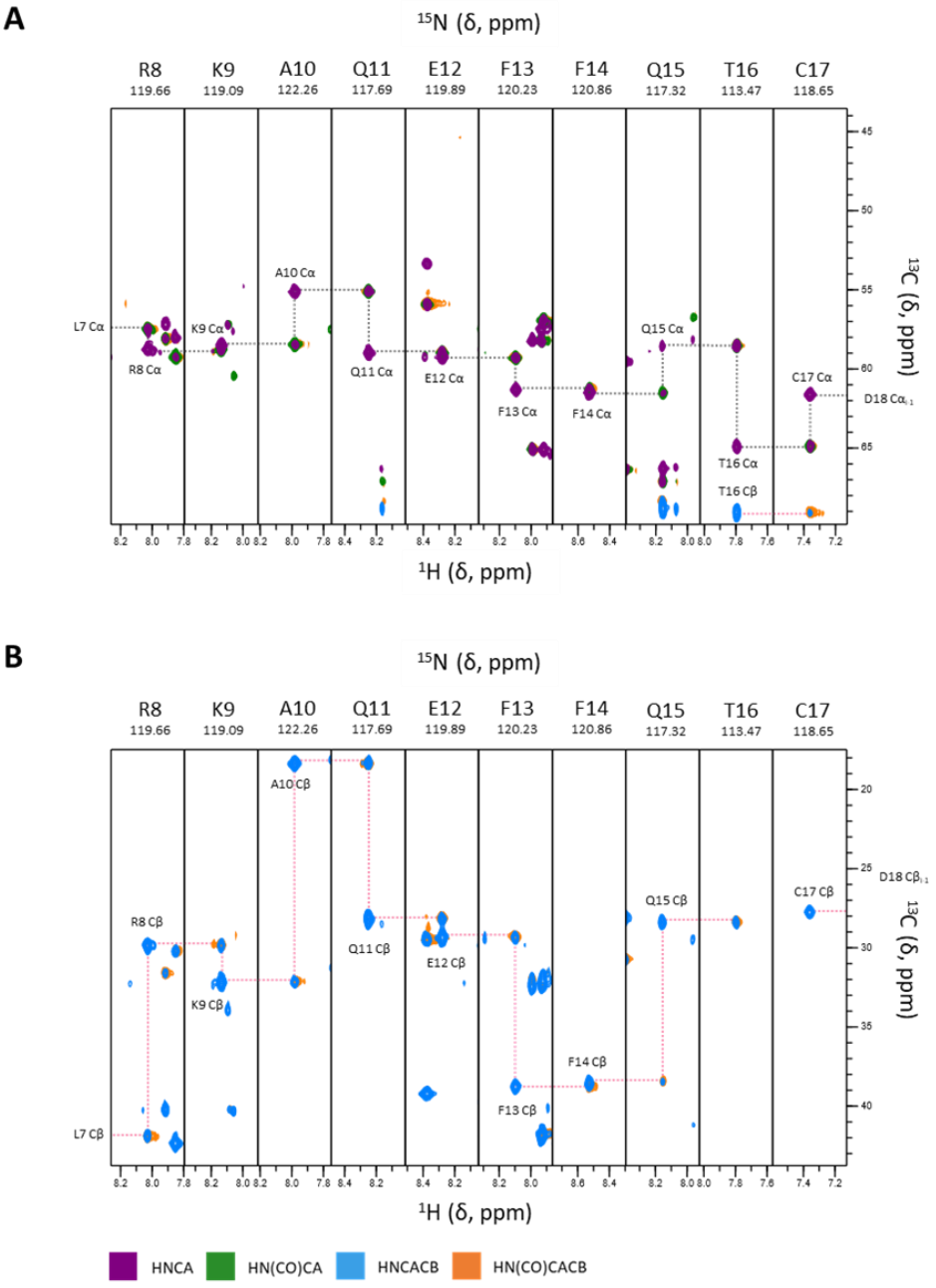

Figure 28 Representative sequential backbone walk for residues R8-C17 of the isolated apo EF-hand domain. (A) Overlay of strip plots from the HNCA and HN(CO)CA spectra showing inter- and intra-residue Cα correlations. (B) Overlay of strip plots from the HNCACB and HN(CO)CACB spectra showing inter- and intra-residue Cβ correlations. The colour code of the four 3D NMR experiments shown in the strips is indicated in the bottom left corner. Grey and pink dashed lines indicate sequential walk for α- and β-carbons respectively.

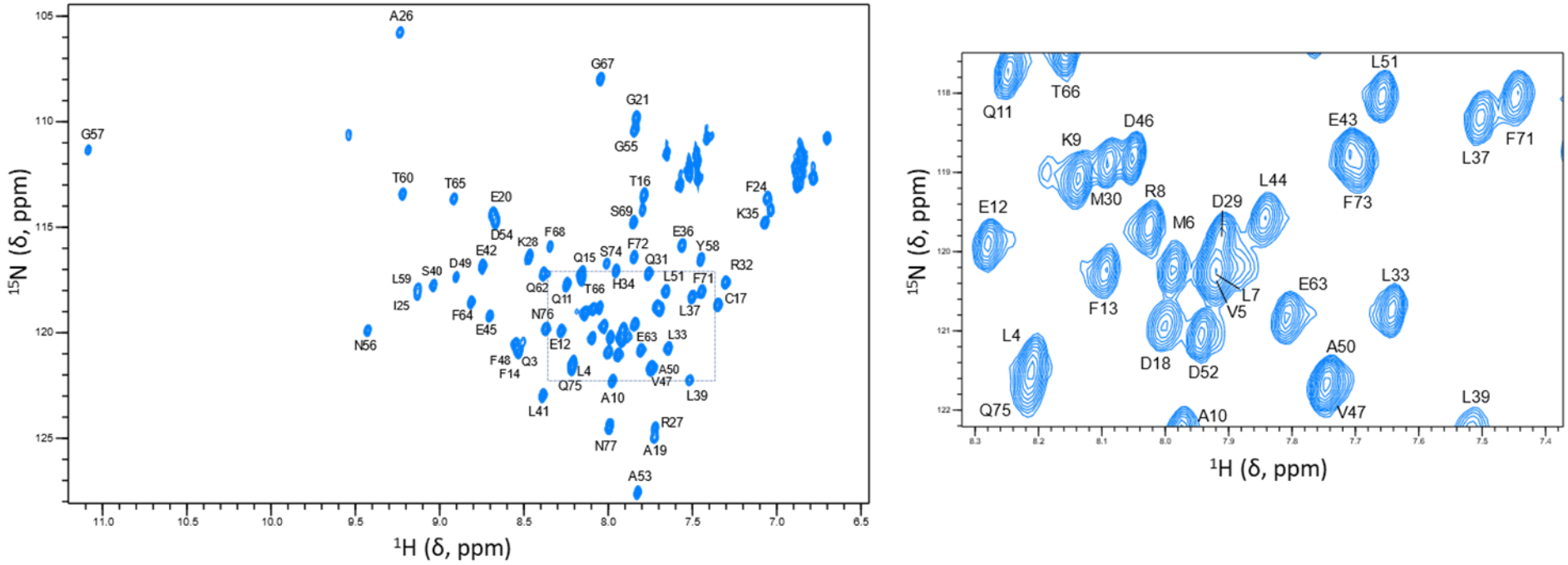

Figure 29 Assigned ^{15}N-^{1}H HSQC spectrum of apo EF-hand domain. The assigned backbone amide cross-peaks were labelled with single amino acid letter code and the residue position within the isolated EF-hand domain. Close-up view of the assignment of the most crowded region of the spectrum is shown in a box on the right-hand side of the HSQC spectrum.

4.2.5. Chemical Shift Index for apo EF-hand

The residue-specific ^{13}Cα and ^{13}Cβ chemical shifts can vary significantly depending on the local chemical environment each amino acid residue is found in. The direction of chemical shift deviation from the expected random coil value can provide indication about whether a stretch of residues occurs within α-helix or β-strand. Comparison of the obtained chemical shifts with the random coil values of the same amino acid type from the literature can therefore inform about the secondary structure. Hence, having obtained the triple resonance backbone assignment of apo EF-hand, ^{13}Cα and ^{13}Cβ chemical shifts were used to calculate Chemical Shift Index (CSI) and determine the sequence-specific secondary structure of apo EF-hand in solution. Secondary chemical shifts (Δδ), defined as the difference between the observed chemical shift and the one expected for a random coil, were calculated for each of the assigned residues with CcpNMR analysis software (Figure 30, (93)). Since Cα atoms of residues within α-helices experience a downfield shift with regard to Cα atoms located in random coil regions, Cα carbons in helices tend to have positive secondary chemical shifts. For Cβ atoms, in turn, the opposite is true, and hence for the residues within α-helices, negative secondary chemical shifts are expected. The obtained secondary chemical shifts were combined to determine the CSI values using CcpNMR analysis (93) (calculated according to the method described in (95)). Using this method built into the CcpNMR analysis software (93), residues within α-helices and β-strands are characterised by CSI values of -1 and 1, respectively. Hence, overall as shown in Figure 30 , the ^{13}Cα and ^{13}Cβ secondary chemical shifts and the CSI values obtained for apo EF-hand in solution show little divergence from the location of the secondary structure elements predicted using PsiPred (111).

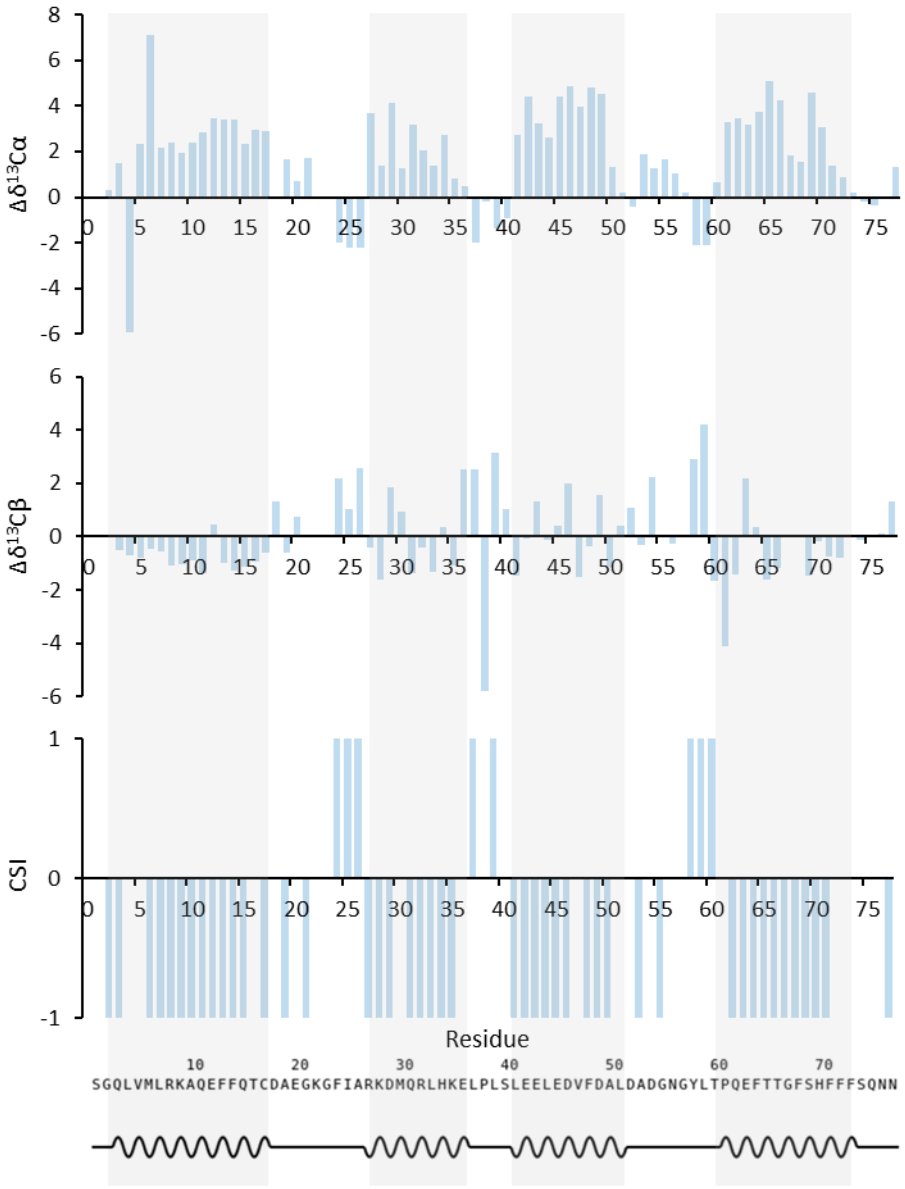

Figure 30 Sequence-specific analysis of the secondary structure of apo EF-hand in solution. The first two graphs demonstrate the secondary chemical shifts (Δδ) of ^{13}Cα and ^{13}Cβ atoms. The third graph shows the chemical shift index (CSI) calculated from ^{13}Cα and ^{13}Cβ secondary chemical shifts. Schematic of the predicted secondary structure elements in EF-hand domain is shown at the bottom of the figure.

4.2.6. CS-Rosetta structure model of apo EF-hand

The crystal structure of EF-hand domain of Rab46 (PDB: 6PSD) was published during the course of this research project (68). Nevertheless, the study by Lee and colleagues (68) reported the structure of Ca^{2+}-saturated EF-hand only and provided no insight into Ca^{2+}-free state of the domain. Hence, encouraged by the evidence from the aforementioned NMR experiment indicative of the apo EF-hand adopting a different conformation to the Ca^{2+}-bound form, an attempt was made to try to obtain the structure of apo to allow direct comparison between the two forms. Since the conventional approach to structure determination using NMR relies on labour-intensive complete side-chain resonance assignment (113) and obtaining a large structural restraint dataset from different NMR experiments (including pair distance, dihedral angle and orientation restraints), in the interest of time (with the time constraints imposed by the global pandemic of coronavirus disease 2019 for example) backbone chemical shift data were used as an input to CS-Rosetta server (version 3.3) for prediction of the apo EF-hand structure. As the models generated by CS-Rosetta appeared to converge, with the ten lowest-energy models clustering within 2 Å from the lowest-energy structure (the averaged Cα rmsd of the ten lowest-energy models was 1.21 ± 0.21 Å), the results were deemed to be a meaningful representation of the apo EF-hand structure (see 'Convergence Results for CS-Rosetta models' in Appendix). The superposition of the lowest-energy CS-Rosetta structure (residues 2-72) and the published crystal structure (PDB: 6PSD; residues 2-74) of EF-hand domain is presented in Figure 31. Although the structural alignment of the two forms did not show a dramatic difference in protein conformation upon calcium binding, changes to the relative orientations of the four helices and the calcium binding loops are apparent (reported RMSD = 2.66 Å).

In order to validate the apo EF-hand model, agreement of the lowest-energy CS-Rosetta structure with experimental backbone amide $^{1}D_{NH}$ residual dipolar couplings (RDCs) was assessed. $^{1}D_{NH}$ RDCs, measured using protein samples in weakly aligning media, provide orientational information about backbone amide bond vectors with respect to the external magnetic field and are therefore commonly used for structure refinement and validation purposes (114). Alignment was first attempted using PEG($C_{12}E_5$)/Hexanol mixture, however, this approach proved unsuccessful. Since both EF-hand and *Pf1* phage are negatively charged at the pH of the experiment, *Pf1* was considered to be an appropriate choice, mitigating risk of the protein interacting with charged alignment medium. Apo EF-hand successfully aligned in a liquid crystalline suspension of 12.5 mg/mL filamentous bacteriophage *Pf1*. ^{2}H quadrupolar splitting of 12.1 Hz was obtained

indicating quite a strong degree of alignment, nonetheless, the observed slight resonances' broadening was deemed acceptable, and that condition was taken forward. An ARTSY experiment (100) was acquired from 0.5 mM ^{15}N-labelled sample at 25ºC on a Bruker 750 MHz spectrometer. Backbone amide $^{1}D_{NH}$ RDCs in apo EF-hand domain were then calculated using spectral ARTSY data by taking the difference in $^{1}J_{NH}$ couplings between the attenuated and reference spectra. The plot showing agreement between the set of 53 experimental RDCs and those back-calculated from the CS-Rosetta structure of apo EF-hand is shown in Figure 31 (panel C). RDCs for residues either overlapped in the spectra or showing big error values (Q3, V5, A19, L33, E36, L37, L39, E43, V47, A50, Y58, F71) were excluded from the analysis. Furthermore, RDCs for residues F73, S74, Q75, N76 and N77 were also not taken into account as this C-terminal part of the protein was not modelled in CS-Rosetta structure. The RDC R-factor of 34.4% (Figure 31, panel C) indicates fair agreement between the (relatively few) remaining experimental RDCs and the CS-Rosetta structure of apo EF-hand (115).

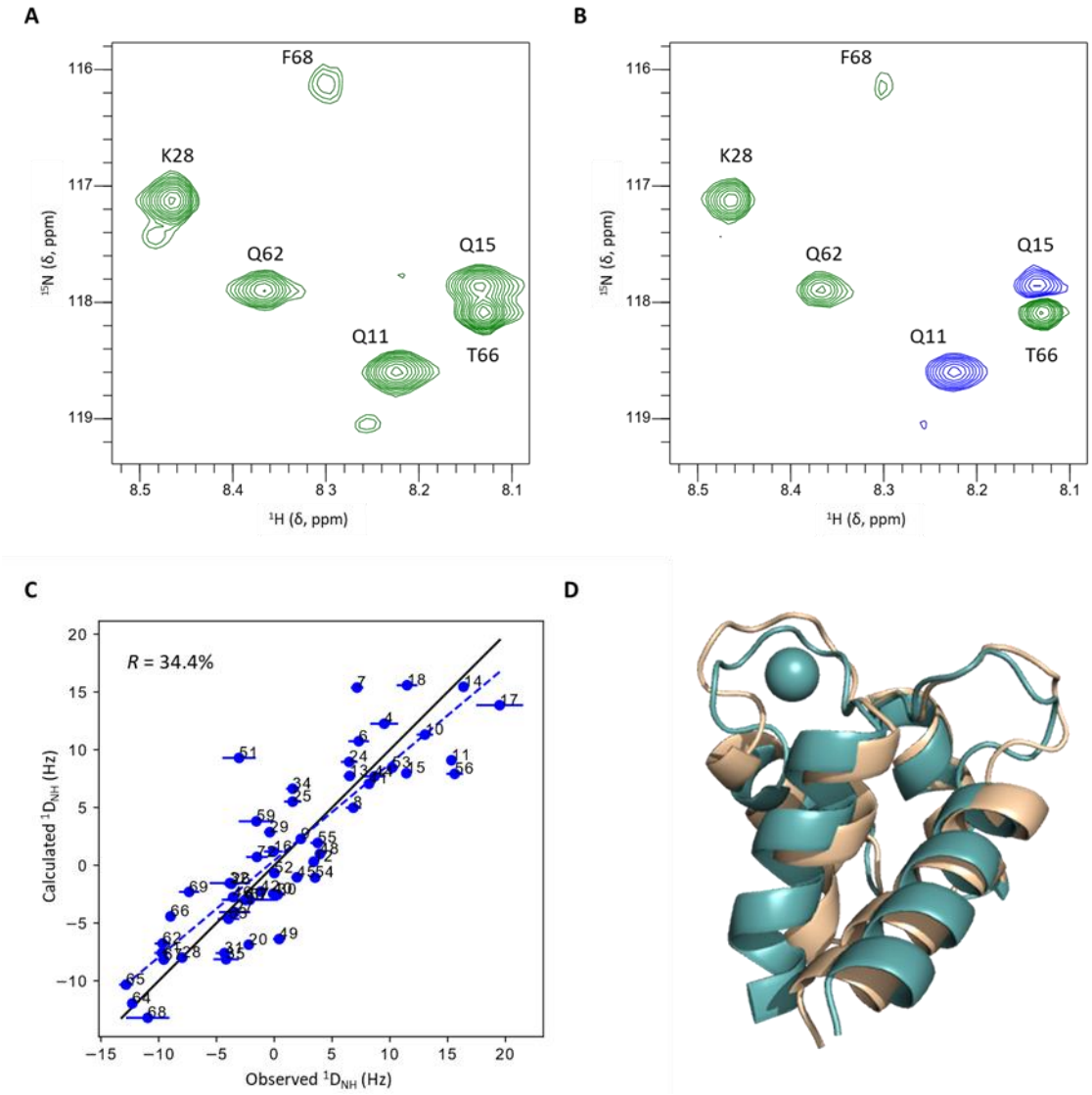

Figure 31 CS-Rosetta structure of apo EF-hand domain. Small regions of reference (A) and attenuated (B) ARTSY spectra recorded at 750 MHz using 0.5 mM sample aligned in 12.5 mg/mL Pf1 phage. Peak positions determined using the reference spectrum were copied directly onto the attenuated spectrum. Positive and negative cross peaks are coloured green and blue respectively. (C) Agreement between the experimental ARTSY-derived $^{1}D_{NH}$ RDCs for apo EF-hand and the RDCs back-calculated from the CS-Rosetta structural model. The error on the experimental RDCs was computed from the signal-to-noise ratio of the ARTSY spectra. Overlapped residues and the ones showing largest error values were excluded from the analysis. R-factor is indicated in the top left corner of the plot. (D) Superposition of CS-Rosetta apo EF-hand structure (beige) and the published Ca^{2+}-bound crystal structure (PDB: 6PSD; teal). RMSD = 2.66.

4.2.7. Downfield shifted glycine residues in apo and Ca^{2+}-saturated EF-hand domain

Of note, backbone assignment revealed that, compared to other glycine residues in apo EF-hand, the amide proton of G57 is distinctly shifted downfield (11.1 ppm). Although the cross peak of G57 α-carbon is weak and buried in noise in HNCA spectrum (Figure 32), and thus on its own may not appear to be particularly convincing, its chemical shift is consistent with ^{13}C α-carbon of a glycine residue, which typically is in the range of 43 to 47 ppm, more upfield than ^{13}C α-carbon chemical shifts of other amino acid types. Furthermore, the α- and β-carbon peaks of the preceding residue (N56) appear strong in HN(CO)CACB spectrum (Figure 32), speaking in favour of the peak belonging to G57. Since the α-carbon of the preceding N56 also produces a weak peak in HN(CO)CA spectrum (despite having a strong peak in HNCA, Figure 32), it is likely a result of conformational change broadening. Nevertheless, to examine the possibility of the suspected G57 cross peak being a folded side-chain peak, HSQC experiment was repeated with a wider spectral window (data not shown). As the peak did not change its position in the spectrum, the experiment provided supporting evidence for the downfield-shifted peak belonging to G57. Interestingly G57 amide proton was also found to be downfield shifted in the Ca^{2+}-bound form, albeit not as strongly as in apo (Table 12). These observations raised a question of what kind of chemical environment G57 must be found in both apo and Ca^{2+}-saturated forms of EF-hand domain.

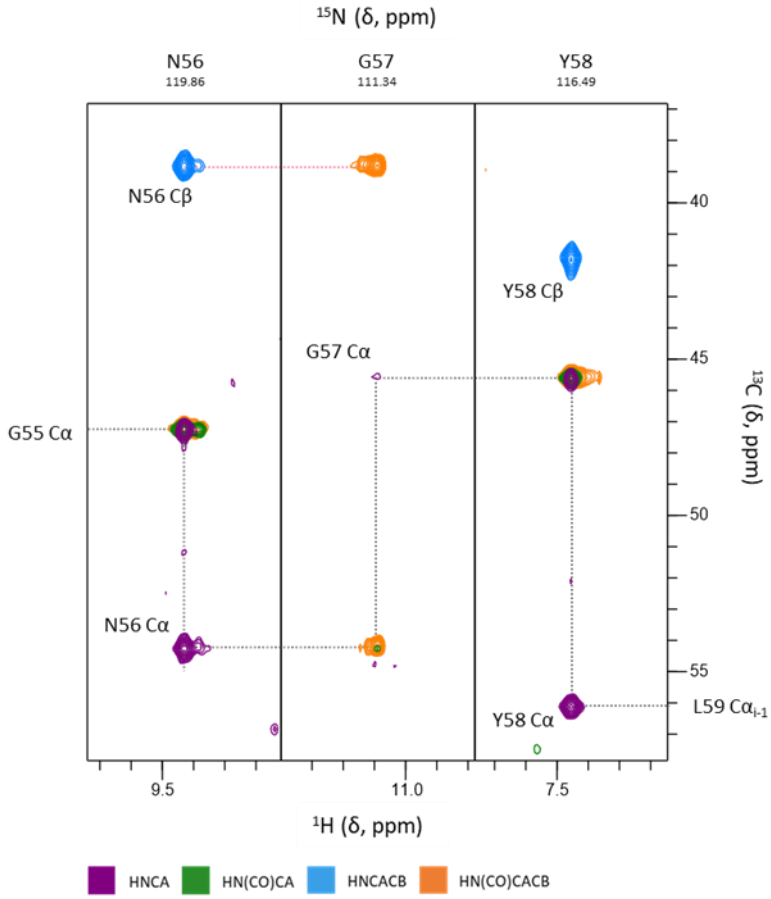

Figure 32 Sequential backbone walk for residues N56-Y58 of the isolated apo EF-hand domain. Overlay of strip plots from the HNCA and HN(CO)CA spectra showing inter- and intra-residue Cα correlations, and HNCACB and HN(CO)CACB spectra showing inter- and intra-residue Cβ correlations. The colour code of the four 3D NMR experiments shown in the strips is indicated in the bottom left corner. Grey and pink dashed lines indicate sequential walk for α- and β-carbons respectively.

Closer examination of the assigned spectra of apo (Figure 29) and Ca^{2+}-saturated EF-hand (Figure 35), revealed that G57, although the most downfield-shifted, was not the only glycine residue showing a dramatic downfield ^{1}HN chemical shift change. As shown in Table 12, G23 residue, not present in the apo EF-hand spectrum, also experiences downfield shift in Ca^{2+}-saturated spectrum. Perhaps, had the G23 cross peak not been broadened beyond detection, it would also be expected to exhibit downfield shift in the spectrum of apo EF-hand.

Table 12 Glycine residues in apo and Ca^{2+}-bound EF-hand domain

Glycine residue	^{1}HN chemical shift in apo (ppm)	^{1}HN chemical shift in Ca^{2+}-bound (ppm)
G2	not present	not present
G21	7.8	7.8
G23	not present	10.6
G55	7.9	7.8
G57	11.1	10.7
G67	8.1	7.5

The location of each of the six glycine residues within the structure of EF-hand domain was therefore examined. The N-terminal G2 is not present in the HSQC spectra of both forms of EF-hand likely due to exchange broadening. G67, located within an α-helix, was found not to be shifted downfield neither in apo nor in Ca^{2+}-bound EF-hand spectra (Table 12). The remaining four glycine residues (G21, G23, G55 and G57) in EF-hand domain of Rab46 are located within the two calcium binding loops. Notably, G23 and G57, i.e. the two glycine residues exhibiting downfield shift, are the second glycine residues within the two calcium-binding motifs. Conversely, the cross peaks of the first glycine residues in the EF-hand motifs (G21 and G55) are not shifted downfield (Figure 29 and Figure 35). A more careful structural analysis of the calcium-binding loops using the crystal structure of EF-hand domain (PDB: 6PSD) showed that the amide protons of all four glycine residues within the loops are hydrogen-bonded to either D18 (G21 and G23) or D52 (G55 and G57) residues. However, based on the distance between the atom pairs, hydrogen bonds formed between G23 and G57 amide protons and OD2 of their respective aspartate residue appear to be stronger (1.9 and 2.1 Å) than the H-bonds formed between G21 and G55 and OD1 of the corresponding aspartate (2.4 and 2.7 Å), which could, to some extent, explain the downfield chemical shift of the former two glycine residues (Figure 33, (116)). What is more, the ring current from aromatic residues can also lead to substantial shifts of the NMR resonances (117). An immediate observation is that both downfield-shifted glycine residues are directly adjacent to amino acids containing aromatic R-groups, F24 and Y58 (Figure 33, Panel C). The distances measured from the centre of F24 and Y58 aromatic rings to amide protons of G23 and G57 are 5.9 Å and 6.2 Å respectively. However, G23 and G57 are also in close spatial proximity to F14 and F48 (5.9 Å and 5.7 Å respectively), which, in turn, are H-bonded to D18 and D52 – the aforementioned aspartate residues forming H-bonds with the glycine

residues in the calcium-binding loops. Of note, the amide protons of G23 and G57 are approximately in the plane of the F14 and F48 aromatic rings, which is not the case for the aromatic rings of F24 and Y58 (Figure 33). Thus, considering the downfield direction of the G23 and G57 amide chemical shifts, it is plausible that the observed extreme shifts could be an effect of hydrogen bonding of the glycine residues to D18 and D52, additionally influenced by the ring current from the neighbouring F14 and F48 residues (117). Nevertheless, the distances between the rings of the aromatic residues and the glycine amide protons do not imply large effects.

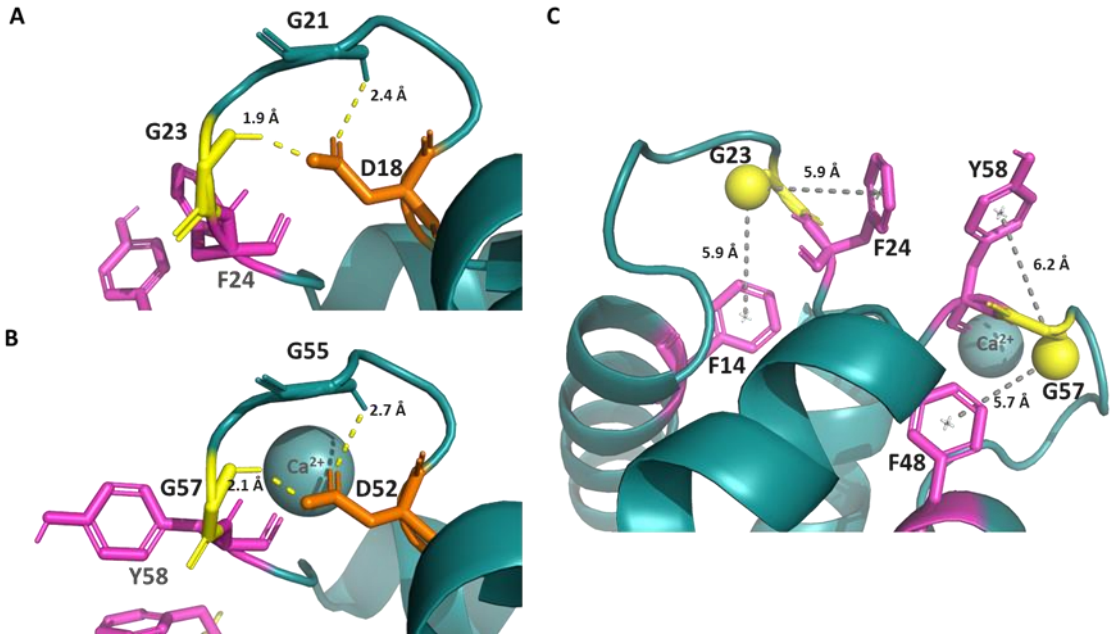

Figure 33 Structural analysis of downfield-shifted glycine residues in EF-hand domain. Closeups on the first (A) and second (B) calcium binding loop of EF-hand. (C) Aromatic residues in the vicinity of downfield-shifted glycine residues. Amide protons of downfield-shifted glycine residues shown as spheres. Distances are shown as grey dashed lines. Hydrogen bonds are represented as yellow dashed lines.

4.2.8. Ca^{2+}-saturated EF-hand

Similarly to apo EF-hand domain, triple resonance experiments for backbone assignment of Ca^{2+}-bound EF-hand domain were acquired on a 750 MHz Bruker spectrometer at 25ºC using 0.5 mM sample saturated with 5 mM $CaCl_2$. Figure 34 shows representative sequential assignment walk for residues T16-I25 used to obtain the backbone chemical shift assignments of Ca^{2+}-saturated EF-hand domain. The assigned ^{1}H-^{15}N HSQC is presented in Figure 35. As opposed to apo form (Figure 29), which

showed little signal overlap, HSQC spectrum of Ca^{2+}-saturated EF-hand contains several areas of overlapped resonances, which proved to be slightly more challenging to assign. Nevertheless, 95.9% of the sequence could be unambiguously assigned for the Ca^{2+}-bound EF-hand domain (71 out of 74 expected residues). The missing assignments in the HSQC spectrum were for residues G2, Q3, and H70. In summary, 71 ^{15}N, 75 Cα, 70 Cβ, and 74 CO chemical shifts were assigned for Ca^{2+}-bound EF-hand domain.

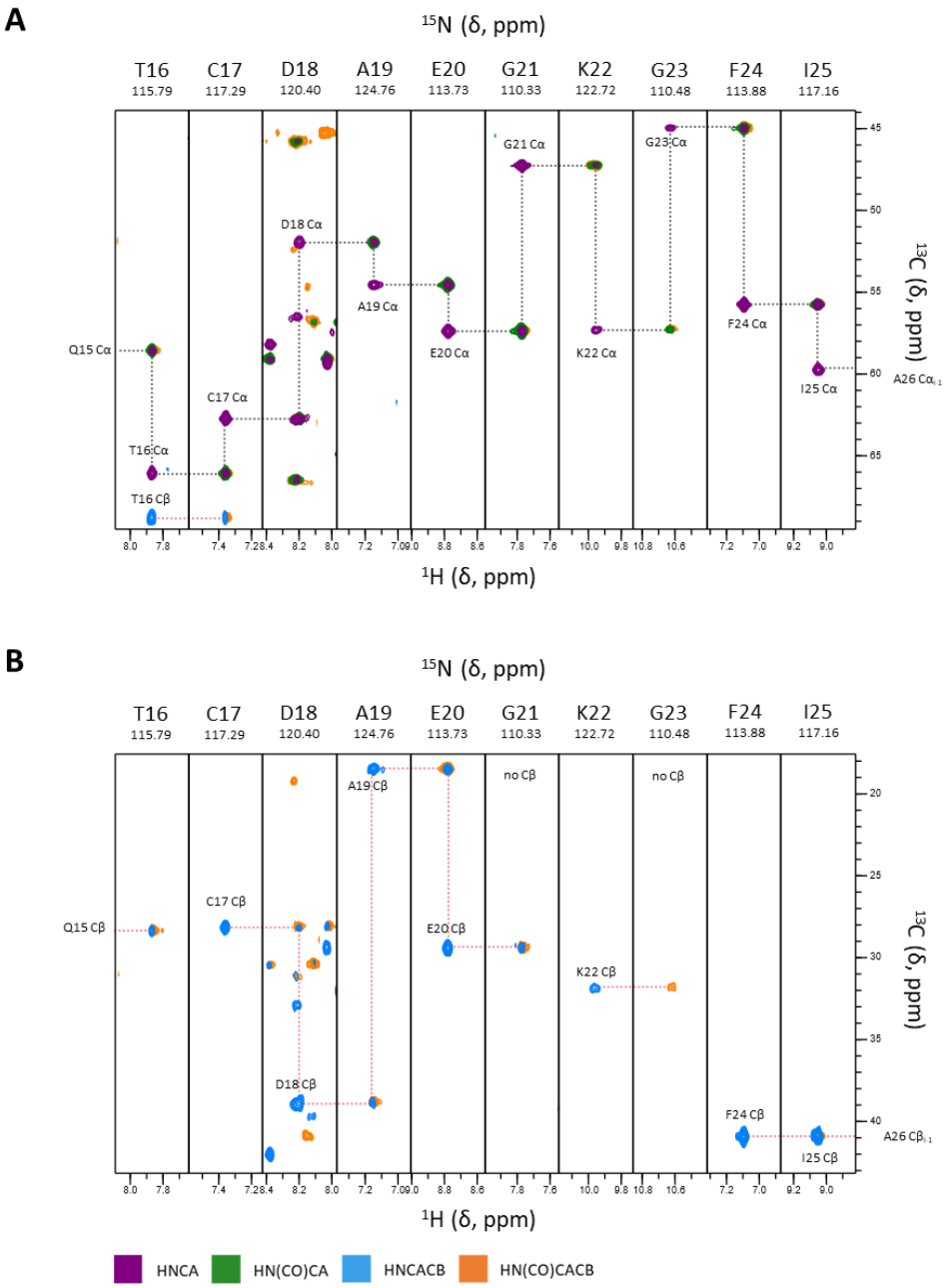

Figure 34 Representative sequential backbone walk for residues T16-I25 of the isolated Ca^{2+}-saturated EF-hand domain. (A) Overlay of strip plots from the HNCA and HN(CO)CA spectra showing inter- and intra-residue Cα correlations. (B) Overlay of strip plots from the HNCACB and HN(CO)CACB spectra showing inter- and intra-residue Cβ correlations. The colour code of the four 3D NMR experiments shown in the strips is indicated in the bottom left corner. Grey and pink dashed lines indicate sequential walk for α- and β-carbons respectively.

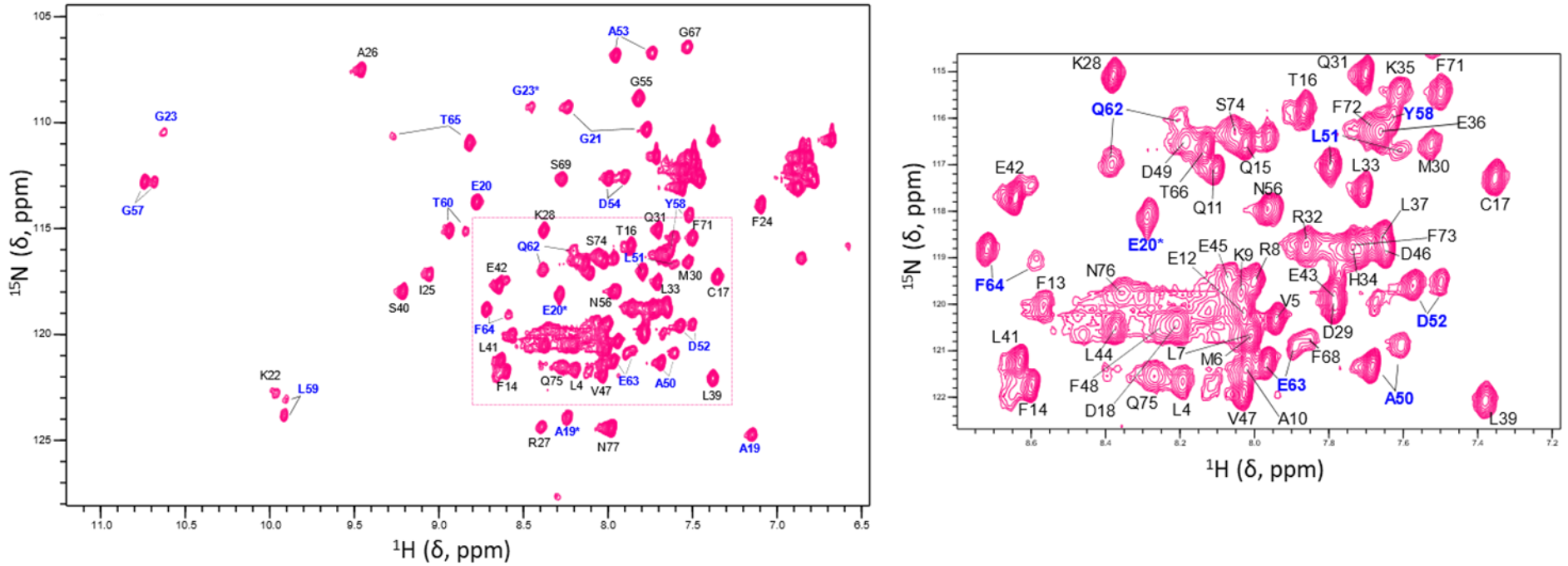

Figure 35 Assigned ^{15}N-^{1}H HSQC spectrum of Ca^{2+}-saturated EF-hand domain. The assigned backbone amide cross-peaks were labelled with single amino acid letter code and the residue position within the isolated EF-hand domain. Close-up view of the assignment of the most crowded, overlapped region of the spectrum is shown in a box on the right-hand side of the HSQC spectrum. Peak doublets are labelled in dark blue and the satellite peaks in distant peak pairs are marked with asterisks.

4.2.9. Satellite peaks in Ca^{2+}-bound EF-hand spectra

Importantly, backbone assignment revealed that the aforementioned (section 2.6.1) extra peaks seen in the HSQC spectrum of Ca^{2+}-bound EF-hand are due to peak doubling of certain amino acid residues (Figure 35, peak doublets labelled in dark blue). The additional peaks were found to be typically lower in intensity (with the exception of G23), however, the ratio of intensity of main to satellite peak was found to be inconsistent between the residues for which doubling was observed (see Appendix). Furthermore, in the sequential walk, satellite peaks could only be followed for short stretches of amino acid sequence as the connectivity was lost at the end of those stretches. Moreover, some of the 'main' peaks (G55, N56 and E63) appeared to be shared with the 'satellite' regions in the sequential walk. An example of sequential walk for the additional peaks is demonstrated in Figure 36. Interestingly, analysis of the location of these satellite peaks within EF-hand domain demonstrated they were primarily localised near the two calcium-binding loops. This could imply e.g. different ways of coordinating the calcium ion, however, one of the loops has been shown not to be functional. Nevertheless, it is possible that doubling of other residues (located further away from the calcium-binding loops) also occurred and these resonances happened to be buried within the heavily overlapped, crowded region of the spectrum.

4.2.10. ZZ exchange spectroscopy

The additional peaks in the ^{1}H-^{15}N HSQC spectrum, as well as two distinct tracts of connectivity in the 3D spectra for backbone resonance assignment, indicate on the Ca^{2+}-bound form of EF-hand being present in solution in two different conformations. To address the question of whether the observed peak doubling is due to an actual conformational heterogeneity or a slow interconversion between two states e.g. exchange between for example ligand-free and ligand-bound or open and closed form of EF-hand, ZZ-exchange NMR experiment was performed (Figure 37). ZZ-exchange spectroscopy provides insight into slow exchange processes between different molecular states, occurring at a ms-s timescale (118). Typically, a range of mixing times expected to overlap with the anticipated timescale of the interconversion process is used in the experiment. Mixing times of 0.05, 0.25 and 0.6 s were therefore tested in the experiment with Ca^{2+}-saturated EF-hand domain. The acquired ZZ exchange spectra overlaid perfectly with the reference HSQC spectrum of Ca^{2+}-bound EF-hand (Figure 37). No exchange (cross) peaks were observed in ZZ-exchange spectra in the intersections of the diagonal (auto) peak pairs of the minor and major state, which could speak in favour of sample heterogeneity, and not interconversion between two

conformations. Nevertheless, it is possible for the exchange to be too slow to be picked up by the ZZ-exchange experiment as could be the case e.g. for cis-trans proline isomerisation which would also manifest as peak multiplication in the NMR spectra (119,120). Due to the differences in chemical shift environment of the trans and cis proline isomers, as well as the slow exchange between the two forms, proline isomerisation results in doubling of both the given proline and its neighbouring residues' cross peaks (121). There are two proline residues in EF-hand domain of Rab46 – P38 and P61, with P61 indeed being located in the vicinity of doubled peaks (Figure, Panel B).

Proline isomerisation, however, does not seem to explain doubling of the peaks for the residues within the first Ca^{2+}-binding loop (A19, E20, G21 and G23), for which, with the exception of G21, the largest distances between the peaks of minor and major form are observed (see Appendix). Such big differences (1.56 ppm for G23 and 1.2 ppm for A19) likely reflect a significant change in the chemical environment these residues are found in, potentially associated with the loop adopting a different conformation (67). Considering the major change in the position of G23 peak, from the downfield region of the spectrum (10.6 ppm in ^{1}H dimension) in the major state to the central region (8.45 ppm) in the minor state, the conformational rearrangement of the loop could potentially lead to D18 moving away from G23 or to a change in the relative orientation between G23 and the aromatic ring of F14, minimising the downfield shift potentially arising from the influence of strong hydrogen bonding and aromatic ring current.

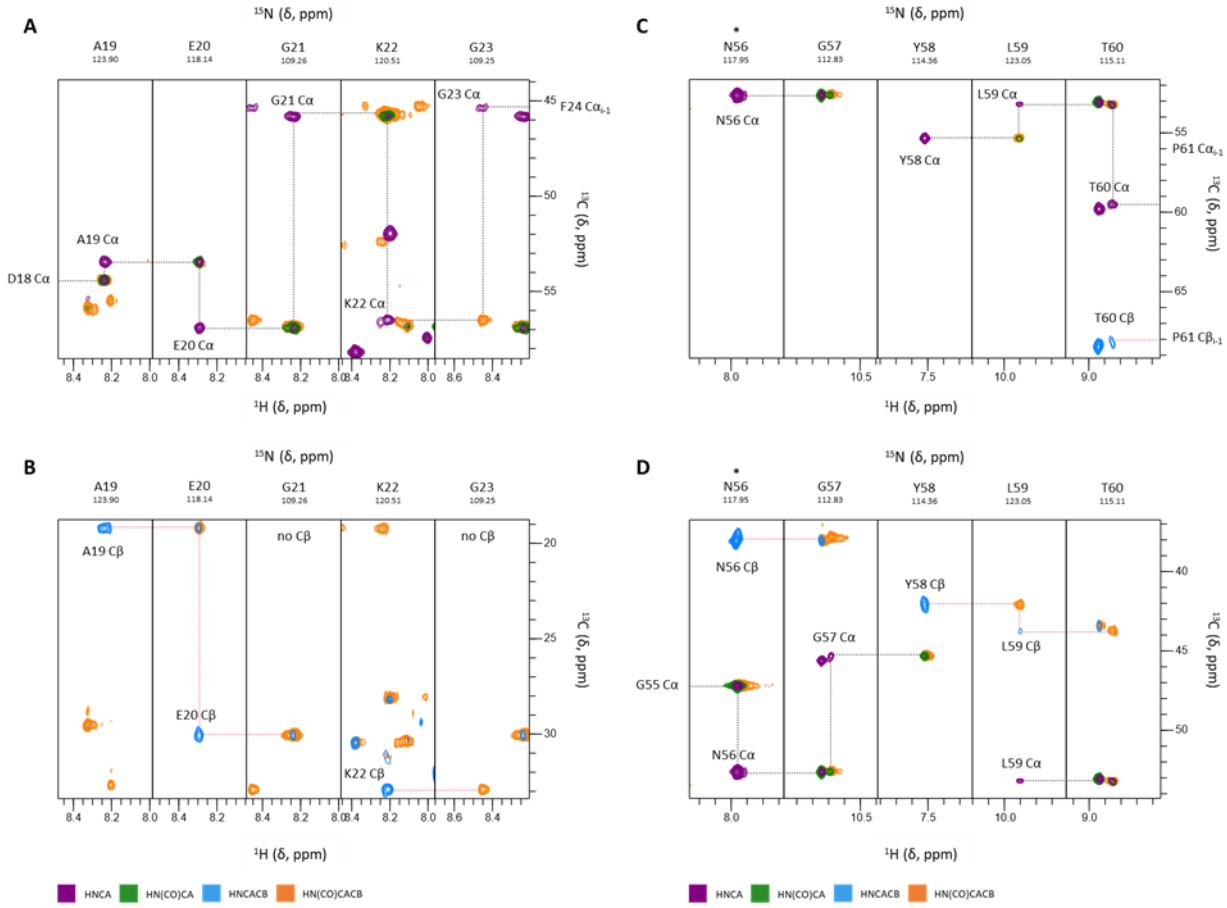

Figure 36 Representative sequential backbone walk for peak doublets observed near first and second calcium-binding loop. (A, C) Overlay of strip plots from the HNCA and HN(CO)CA spectra showing inter- and intra-residue Cα correlations. (B, D) Overlay of strip plots from the HNCACB and HN(CO)CACB spectra showing inter- and intra-residue Cβ correlations. The colour code of the four 3D NMR experiments shown in the strips is indicated in the bottom left corner. Grey and pink dashed lines indicate sequential walk for α- and β-carbons respectively.

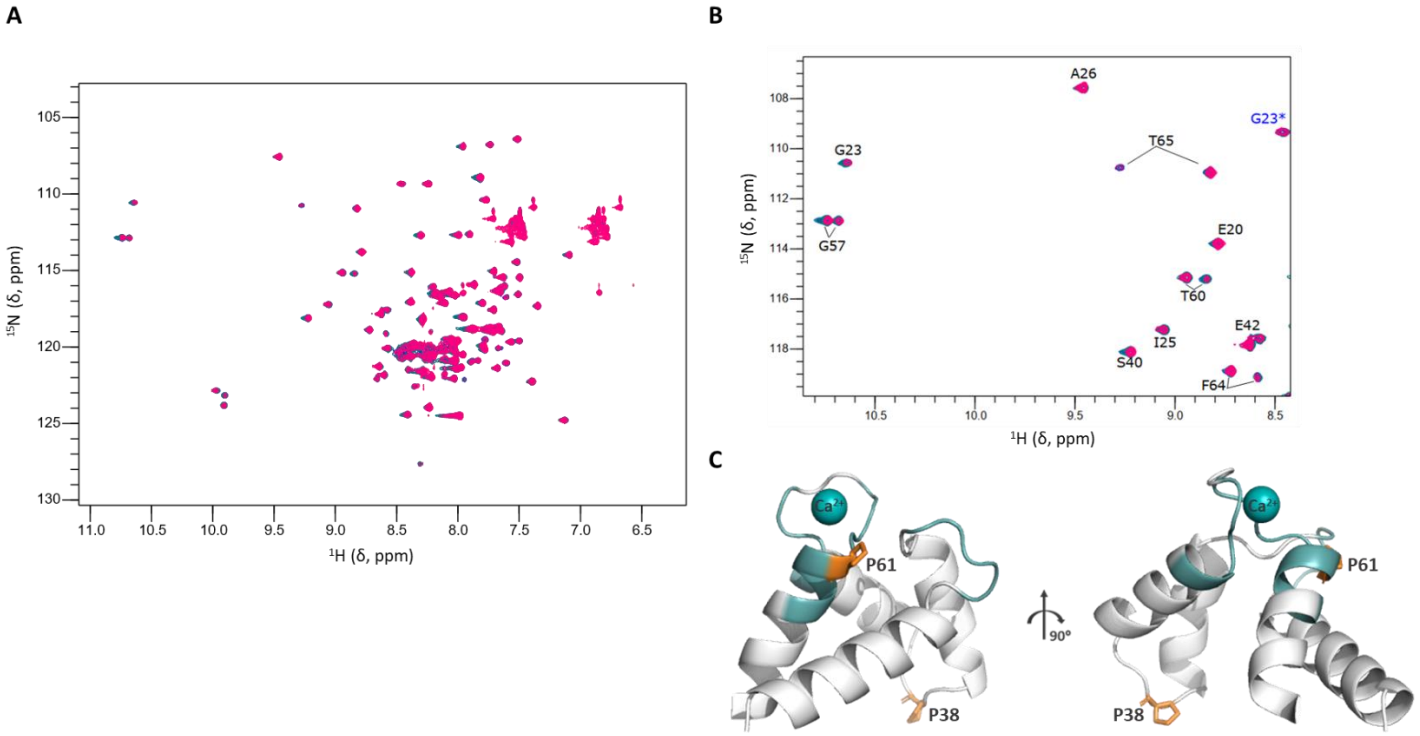

Figure 37 Overlay of ZZ-exchange spectra. (A) Overlay of ZZ-exchange spectra and (B) close-up view of the region containing multiple doubled peaks. Reference ^{1}H–^{15}N HSQC spectrum is shown in magenta. Spectra acquired using mixing times of 0.05, 0.25 and 0.6 s are shown in teal, dark blue and purple respectively. (C) Double peaks' locations (coloured teal) mapped on the crystal structure of EF-hand domain. Proline residues are shown in stick representation and coloured orange.

4.2.11. ^{1}H-^{15}N HSQC at lower pH

Since the pH-induced changes in the free energy of the protein can promote the population of alternative protein conformations (122), additional HSQC experiments were performed with the Ca^{2+}-bound EF-hand protein at a lower pH to examine whether a pH-dependent equilibrium exists between the minor and major conformational states. ^{1}H–^{15}N HSQC spectra were acquired for samples exchanged into NMR buffer at pH 6.5 and 6.0, as well as for the reference sample at pH 7.0. Due to the sample availability limitations, the experiments were acquired at a lower sample concentration (250 μM), resulting in a slight decrease in cross peak intensities (compared to the previously described experiments). Nevertheless, peak doublets from the two conformational states of Ca^{2+}-bound EF-hand domain, although weaker, are still visible in ^{1}H-^{15}N HSQC spectrum enabling analysis of the effect of solution pH on the two forms of Ca^{2+}-bound EF-hand. The pH-induced chemical shift perturbations (CSPs) in the HSQC spectrum of Ca^{2+}-saturated EF-hand domain are shown in Figure 38. The largest shifts occur for the N-terminal residues and the ones in the vicinity of H34 and H70 (Figure 38, Panels D and E). The chemical shift changes observed for these residues are in fast exchange (Figure 38, Panel B), consistent with protonation/deprotonation of the ionizable groups of histidine residues. Of note, although pH changes generally do not seem to lead to significant shifts of the peak doublets, the said peak pairs experience pH-induced changes in intensity (Figure 38, Panel C). Interestingly, at pH 6.0 vast majority of the peaks of the minor conformation of Ca^{2+}-bound EF-hand appears to increase in intensity (on average 2.2 times compared to their intensity at pH 7.0; see Appendix), while the intensity of the major conformation peaks is decreased (on average 1.3 times compared to intensity at pH 7.0) in majority of cases, with the exception of Y58 and Q62, for which an increase in intensity of the major form peaks at a lower pH is observed (Figure 38, Panel C). This finding may suggest that the population of the two conformational states of Ca^{2+}-bound EF-hand is pH-dependent, with low pH seemingly promoting the population of the minor conformation.

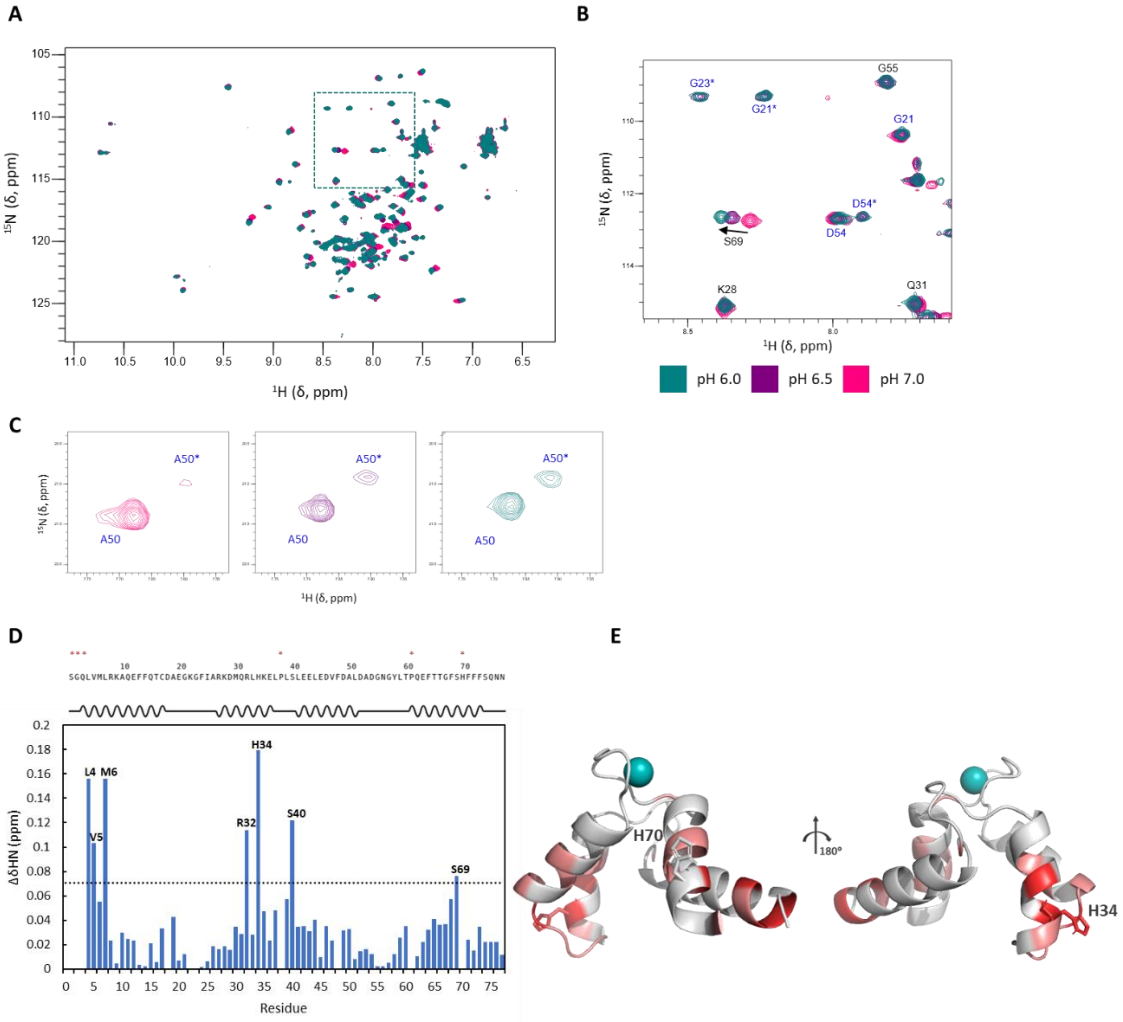

Figure 38 Effect of lowering the pH on the Ca^{2+}-bound EF-hand domain. (A) Overlay of the ^{1}H–^{15}N HSQC spectra of Ca^{2+}-saturated EF-hand domain at different pH of the NMR buffer. The colour code of the spectra is indicated in the bottom right corner of panel B. (B) Close-up view illustrating the example pH-induced chemical shift perturbations. (C) pH-induced intensity changes of A50 peak doublet. Peak of the minor conformation is marked with an asterisk (D) pH-induced CSPs plotted against residue number of EF-hand domain. Cut off = mean + SD. Missing residues are marked with red asterisks above the secondary sequence plot. (E) Residues experiencing largest shifts mapped on the crystal structure of EF-hand domain. H34 and H70 shown in stick representation. Residues with CSPs above mean are coloured salmon while residues with CSPs above mean+SD are coloured red.

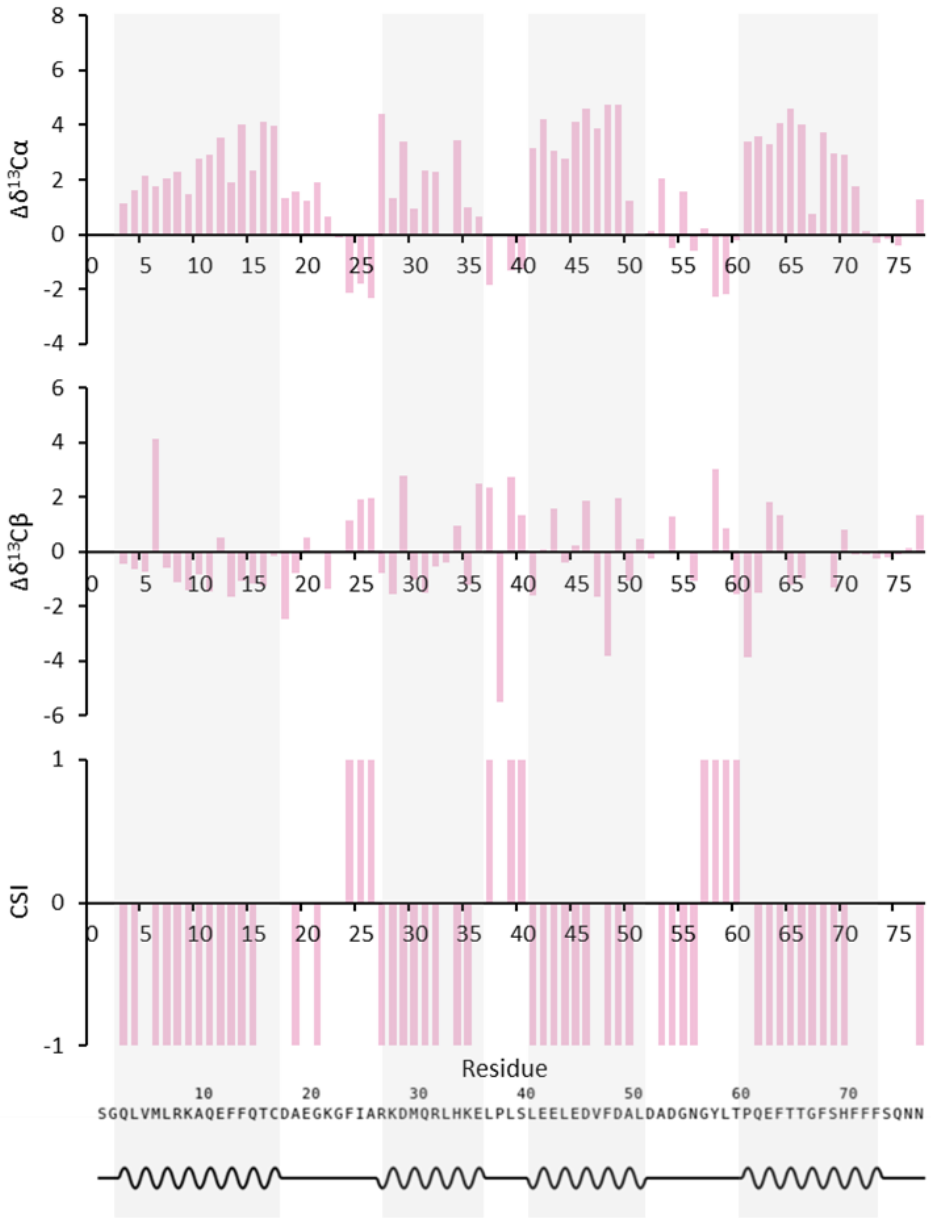

Figure 39 Sequence-specific analysis of the secondary structure of Ca^{2+}-bound EF-hand in solution. The first two graphs demonstrate the secondary chemical shifts (Δδ) of 13Cα and 13Cβ atoms. The third graph shows the chemical shift index (CSI) calculated from 13Cα and 13Cβ secondary chemical shifts. Schematic of the predicted secondary structure elements in EF-hand domain is shown at the bottom of the figure.

4.2.12. Chemical Shift Index for Ca^{2+}-saturated EF-hand

CSI for the Ca^{2+}-bound EF-hand was calculated using Cα and Cβ chemical shifts obtained from the backbone assignment as described for the apo form in section 4.2.5. In the case of peak doublets, the chemical shifts of the major form peaks were used. As demonstrated in Figure 39, overall the CSI is in good agreement with the predicted location of secondary structure elements. Interestingly CSI for both apo (Figure 30) and Ca^{2+}-saturated (Figure 39) forms shows a consistent pattern of differences from the expected secondary structure in the loop regions, e.g. for residues 24-26 or 58-60 with CSI indicating on the residues being located within β-strand structure.

CSI for second set of peaks of the minor Ca^{2+}-bound conformation was found to be generally the same with the exception of residues T60 and F64 for which the CSI value changed respectively from 1 and -1 to 0, indicating loss of ordered structure (see Appendix).

4.3. Effect of calcium binding on EF-hand domain

4.3.1. Calcium and EF-hand domain stability

To assess the effect of calcium binding on the stability of the recombinant EF-hand domain, thermal denaturation of apo and Ca^{2+}-bound protein was monitored by Far-UV CD spectroscopy. No difference was observed in the appearance of the apo and Ca^{2+}-saturated CD spectra at 20ºC (Figure 40). Flattening of the characteristic negative peaks at 208 nm and 222 nm in the CD spectrum denotes the gradual loss of α-helical structure and indicates the temperature-induced unfolding of the EF-hand protein. Melting temperatures (T_m) of apo and Ca^{2+}-bound EF-hand, were estimated from the thermal denaturation data assuming single transition (i.e. from folded to unfolded state) using Global 3 software (Figure 40). T_m of apo and Ca^{2+}-bound EF-hand were calculated to be 51.2ºC and 56.7ºC respectively. The addition of calcium, therefore, seems to have a stabilising effect on the protein, leading to apparent increase in the protein's melting temperature of 5.5ºC. Nevertheless, the determined T_m values should be treated with caution as these were derived from a single CD temperature ramp experiment.

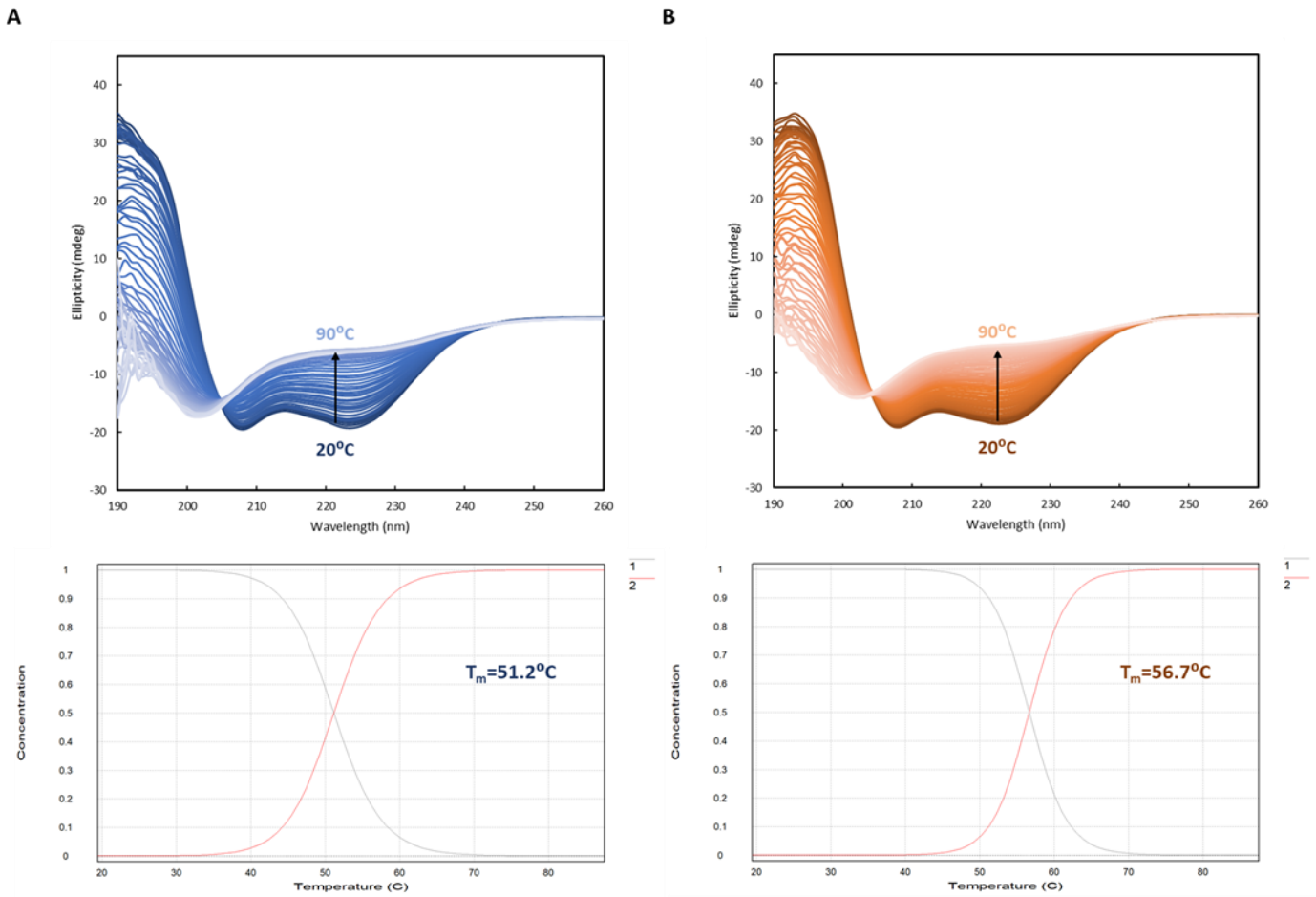

Figure 40 Thermal melt plots of apo (A) and Ca^{2+}-bound (B) EF-hand domain. Overlaid CD spectra measured from 260 nm to 180 nm over a temperature range of 20ºC-90ºC. Proportion of folded (1) to unfolded (2) state plotted against temperature and the derived T_m value for each form is shown at the bottom of each thermal denaturation plot.

4.3.2. EF-hand domain of Rab46 adopts an open conformation upon Ca^{2+} binding

Although the structural alignment of the Ca^{2+}-saturated and apo EF-hand structures (Section 4.2.6, Figure 31) demonstrated differences to the relative orientations of the α-helices pairs of the two EF-hand motifs, that observation on its own does not inform about the effect of such conformational change on the protein function. To further investigate the implications of the conformational rearrangement occurring upon calcium binding, solvent accessible surface area (SASA) was calculated for the lowest energy CS-Rosetta apo structure and the published crystal structure of Ca^{2+}-bound EF-hand domain using Chimera 1.16 software (97) (Table 13). Compared to apo, Ca^{2+}-bound EF-hand was found to have a larger SASA (4550.88 Å and 5079.67 Å respectively), suggesting that the calcium-bound form adopts a more open conformation.

Table 13 Solvent accessible surface area (SASA) comparison of apo and calcium-bound EF-hand domain.

EF-hand structure	Ca^{2+}-bound (X-ray, PDB: 6PSD)	Apo (CS Rosetta)
SASA	5079.67 Å	4550.88 Å

Furthermore, analysis of the surfaces of the two forms revealed that calcium binding leads to exposure of a greater number of hydrophobic residues on the protein surface (Figure 41), which results in the formation of a hydrophobic pocket, permitting interactions with binding partners.

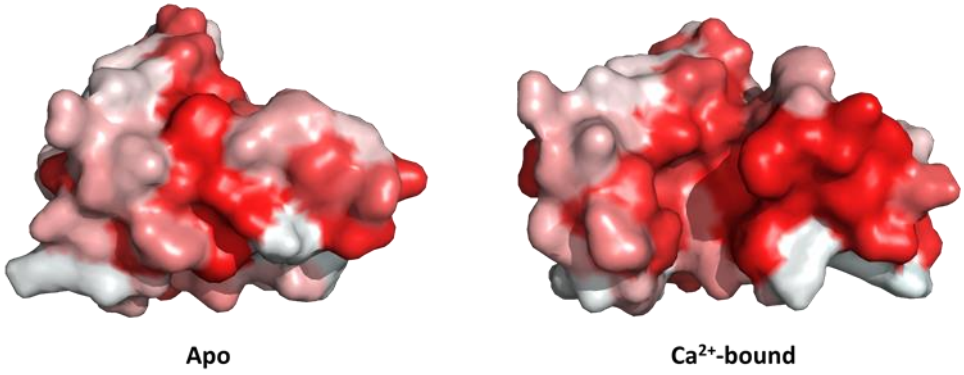

Figure 41 Comparison of the surfaces of apo and Ca^{2+}-bound EF-hand structures. The two structures were aligned and hence are shown in the same orientation. Hydrophobic residues on the surface of CS-Rosetta apo and crystal Ca^{2+}-bound structures were coloured red in PyMOL (78).

Additionally, backbone chemical shifts measured for Ca^{2+}-saturated EF-hand (major form) were used to obtain a model of the Ca^{2+}-bound protein using CS-Rosetta server to allow comparison between the two predicted 'solution structures'. The SASA of the lowest-energy CS-Rosetta model of Ca^{2+}-bound EF-hand was calculated to be 4750.96 Å, greater than the one of apo, consistent with the domain opening, yet smaller than the one of Ca^{2+}-bound crystal structure. Nevertheless, unlike apo, the CS-Rosetta structure of Ca^{2+}-bound EF-hand has not been validated by any additional experiments. Furthermore, as non-protein components (e.g. ligands or cofactors) are not taken into account in the CS-Rosetta model-building protocol, the calcium ion is missing from the structure, which can additionally contribute to the differences between the crystal and CS-Rosetta structures.

4.3.3. IMS-MS

Ion Mobility Spectrometry-Mass Spectrometry (IMS-MS) was used to experimentally validate whether the structural changes occurring upon calcium binding indeed lead to opening of the EF-hand domain of Rab46. Since the quaternary protein structure is preserved in the gas phase, IMS-MS is a particularly useful approach for obtaining information on intact proteins, their composition and conformational dynamics (123). The MS and IMS-MS spectra of apo and Ca^{2+}-bound EF-hand are shown in Figure 42. The shift to higher charge states observed for the Ca^{2+}-bound form (Panels A and C) likely results from exposure of a greater number of protonation sites and is suggestive of either partial unfolding or opening of the Ca^{2+}-bound EF-hand domain. Therefore, the obtained IMS-MS result appears to speak in favour of the apo and Ca^{2+}-saturated EF-hand of Rab46 existing in compact and open conformations respectively, in agreement with the SASA analysis (124).

A concerning finding, however, is that despite the MS experiment being performed with a sample dialysed against EDTA, Ca^{2+} (addition of ~38 Da in the spectrum) appears to be present in the apo sample (Panels B and D). The presence of contaminating Ca^{2+} can be introduced by buffers, salts or purified water and is a common problem in Ca^{2+} binding studies (125). While the low amounts of contaminating Ca^{2+} likely have a negligible effect when high concentrations of the proteins are used (e.g. for NMR or ITC experiments), the effect of background Ca^{2+} will be more pronounced when lower protein concentration is used, as is in the case of MS analysis, for which EF-hand domain was used at low micromolar (10 μM) concentration. Therefore, it is likely that the peaks with a mass corresponding to Ca^{2+}-bound EF-hand observed in the apo MS spectrum arise from reagent or instrument contamination with trace amounts of calcium.

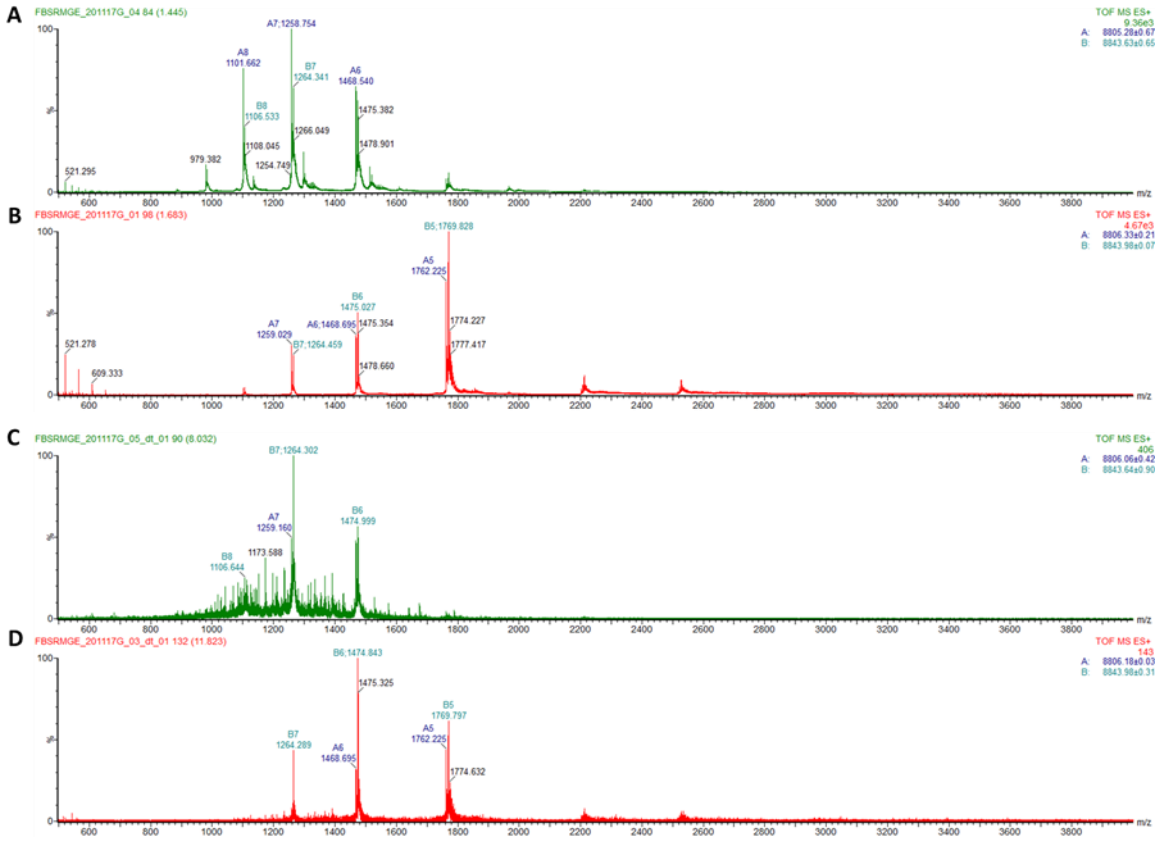

Figure 42 Native MS analysis of EF-hand domain. (A, C) MS and IMS-MS profiles of Ca^{2+}-bound EF-hand in 100 mM ammonium acetate supplemented with 20 µM calcium acetate (2-fold excess). (B, D) MS and IMS-MS profiles of apo EF-hand in 100 mM ammonium acetate.

4.4. Ca^{2+} titration experiment

4.4.1. Intermediate state

An NMR calcium titration experiment was performed to observe the gradual changes to the protein conformation upon Ca^{2+} binding and derive binding parameters by measuring chemical shift perturbations (CSPs) in NMR spectra acquired at increasing concentrations of Ca^{2+}. As illustrated by Figure 44, the experiment revealed presence of an intermediate Ca^{2+}-binding step at a 1:5 ligand to protein ratio (dark blue spectrum), characterised by an incomplete set of NMR signals and severe peak broadening, indicative of chemical exchange on an intermediate timescale relative to the NMR chemical shift timescale (Figure 43). As proteins are highly dynamic molecules, they do not simply adopt one structure upon ligand binding. Instead, they tend to continuously

sample an equilibrating set of conformations and the dynamics of these interconversions can be slow, intermediate, or fast on the NMR timescale (112).

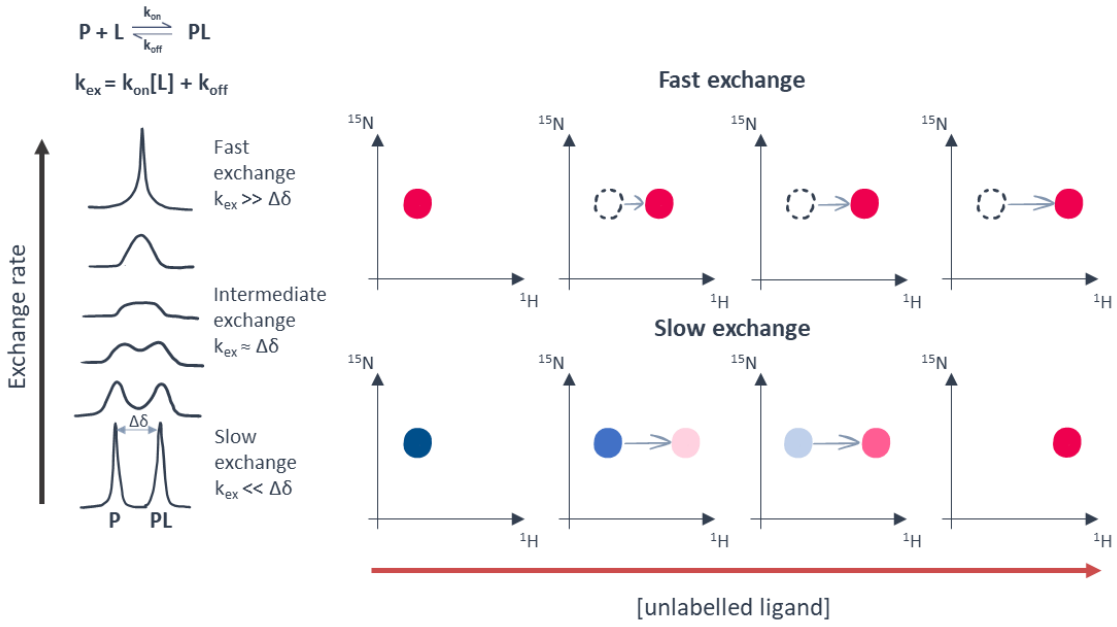

Figure 43 Using CSPs to characterise ligand binding. In the presence of two interconverting species (e.g. ligand-free and ligand-bound protein), the NMR signal varies as a function of the exchange rate (k_{ex}). When the exchange is fast, i.e. the difference in chemical shifts between the species in exchange (Δδ) is smaller than the exchange rate (k_{ex}), only an averaged chemical shift is observed**.** When the exchange is slow, i.e. Δδ is larger than k_{ex}, chemical shifts corresponding to both species are observed. When the exchange is intermediate, i.e. Δδ is equal to k_{ex}, chemical shifts of the exchanging species become broadened. The figure was adapted from Bourgeois *et al.* (126). P – protein, L – ligand, k_{on} -rate constant of association, k_{off} -rate constant of dissociation.

Such broadening of signals beyond detection, as observed for EF-hand at 1:5 ligand to protein ratio (Figure 44), often suggests interconversion between two or more states on a µs-ms timescale, which could correspond to, for example, switching between two different conformations or between Ca^{2+}-bound and Ca^{2+}-free states. As the calcium concentration increases, new peaks appear in different chemical environments with respect to the apo form, which eventually leads to reorganisation of the spectrum (Figure 44 and Figure 45), corresponding to the conformational change described in the previous sections. As demonstrated in Figure 46 (Panel B), the biggest shifts occur for the

residues within the calcium binding loop and E, and F helices of EF2, as well as F helix of EF1 motifs, likely in agreement with these getting re-arranged upon the domain adopting a more open conformation. Furthermore, the observation of smaller CSPs for the residues within the EF1 Ca^{2+} binding site appears to be in agreement with this site not being functional and with EF-hand domain containing a single Ca^{2+} ion binding site in the EF-loop of EF2 motif.

4.4.2. Chemical shift perturbations at increasing calcium concentration

The initial NMR titration experiments demonstrated virtually no CSPs at lower calcium concentrations, followed by a sudden large shift as the concentration increased, giving rise to sigmoidal appearance of the titration curves plotted for the residues for which chemical shift changes occurred in fast exchange (data not shown). Though sigmoidal appearance of the binding curves can be suggestive of cooperative binding, considering that one of the calcium-binding sites of EF-hand of Rab46 was demonstrated not to be functional (68), this observation raised a question whether it was a real effect or a result of some sort of systematic error at the lower concentration points. It was quickly realised that the likely explanation for the observed pattern was the presence of residual EDTA in the protein samples following dialysis to remove Ca^{2+} bound to the EF-hand domain following the purification. Indeed, EDTA presence was also confirmed by 1D NMR (data not shown). Realisation of the issue with residual EDTA in the Ca^{2+}-free samples led to modification of the purification protocol to perform the Ca^{2+} removal step prior to SEC. Furthermore, exhaustive dialysis and multiple additional buffer exchange steps of the EF-hand samples using PD10 or Zeba spin columns were introduced in the sample preparation process prior to calcium binding studies, due to EDTA being notoriously difficult to remove (127). Calcium removal step, however, could not be avoided as the spectra obtained from the first NMR experiments (data not shown), using the purified protein from which calcium was not actively removed, demonstrated large degree of heterogeneity, indicating on some of the protein existing in apo and some in Ca^{2+}-bound forms following the expression and purification. Therefore, going forward, considering the difficulties in ensuring effective EDTA removal (127) it may be worthwhile to consider using Chelex 100 resin as an alternative means for Ca^{2+} removal from the protein preparations prior to Ca^{2+} binding studies (128).

Though the amounts of residual EDTA were successfully lowered by the optimisation of the sample preparation protocol, the binding curves obtained from following CSPs in the repeated experiment were still somewhat sigmoidal (Figure 45), likely resulting from incomplete removal of EDTA. This, in turn, explains the peaks not changing their

positions significantly at low calcium concentrations (perhaps up to the point EDTA present in the sample is completely saturated) up to 50 µM when a major change occurs, causing a considerable number of signals (Figure 44, Panel A) to disappear and the remaining population of peaks to shift. The missing peaks start to reappear at a 1:1 ratio of ligand to protein (250 µM $CaCl_2$,Figure 44). At Ca^{2+} concentrations above 50 µM, a proportion of cross peaks shifts substantially in different directions as demonstrated in Figure 44 and Figure 45, with the shift direction indicated with the black arrows. The observed gradual movement towards the fully ligand-saturated protein peaks indicates that the residues are involved in fast exchange at the NMR chemical shift timescale. However, simultaneously with this movement, new peaks gradually increasing in intensity appear in the spectrum (Figure 44), which is characteristic of intermediate or slow exchange processes. Interestingly, the peaks for which slow exchange is observed, are predominantly the 'satellite peaks' belonging to the aforementioned minor Ca^{2+}-bound state of EF-hand, while the peaks of the major form experience fast exchange.

Therefore, these data indicate that the exchange regime differs for different resonances, depending on local dynamics and on the chemical shift differences of those resonances between the states that are exchanging relative to the exchange rates. A mixture of fast, slow and intermediate exchange regimes can be characteristic of two-step binding reactions. Nevertheless, knowing there cannot be any cooperativity between the two calcium binding sites, this is unlikely to be an explanation for the observed pattern.

An important implication of such 'mixed exchange' regime is that it significantly complicates the analysis, as the CSPs are intractable for some residues due to line broadening and signal disappearance, which was the reason for having to independently assign both the apo and Ca^{2+}-bound forms. Furthermore, quantification of the dissociation constant using CSPs is far less straightforward outside the fast exchange regime, as low to intermediate exchange rates result in line broadening. Since there is no good way to distinguish CSPs due to direct binding of the ligand from CSPs due to allosteric change, using CSPs to obtain the K_d value for Ca^{2+} binding to EF-hand domain could lead to inaccuracies (129). Hence, a decision was made to derive Ca^{2+}-binding parameters using Isothermal Titration Calorimetry (ITC).

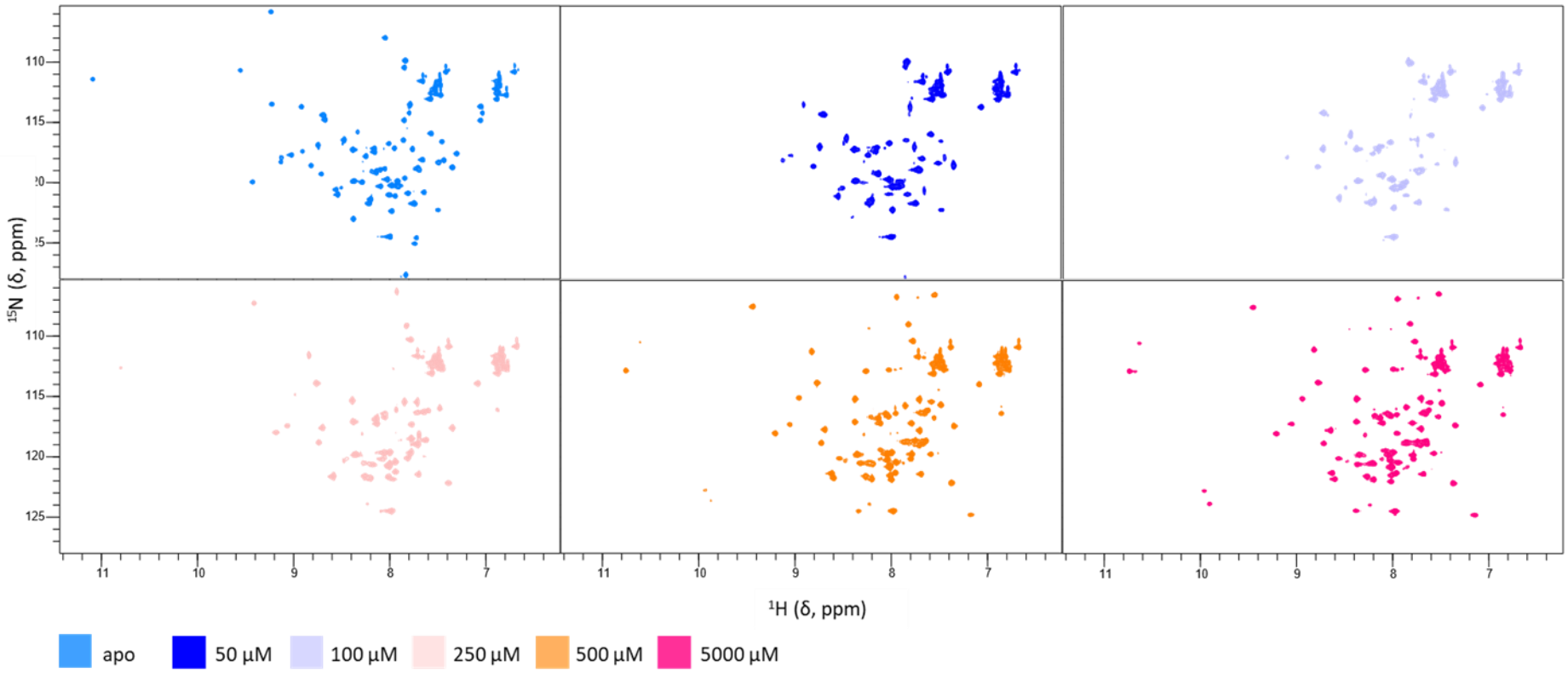

Figure 44 ^{1}H–^{15}N HSQC spectra of EF-hand domain of Rab46 at increasing Ca^{2+} concentration. The colour code of the six spectra is indicated in the bottom left corner. The spectra were recorded from a 0.25 mM solution of recombinant EF-hand uniformly labelled with 250 µM ^{15}N-labelled EF-hand in buffer containing 25 mM Hepes (pH 7.0), 150 mM NaCl, 1 mM TCEP and 0.1% NaN_3. The data were acquired on a 750 MHz Bruker spectrometer at 25ºC.

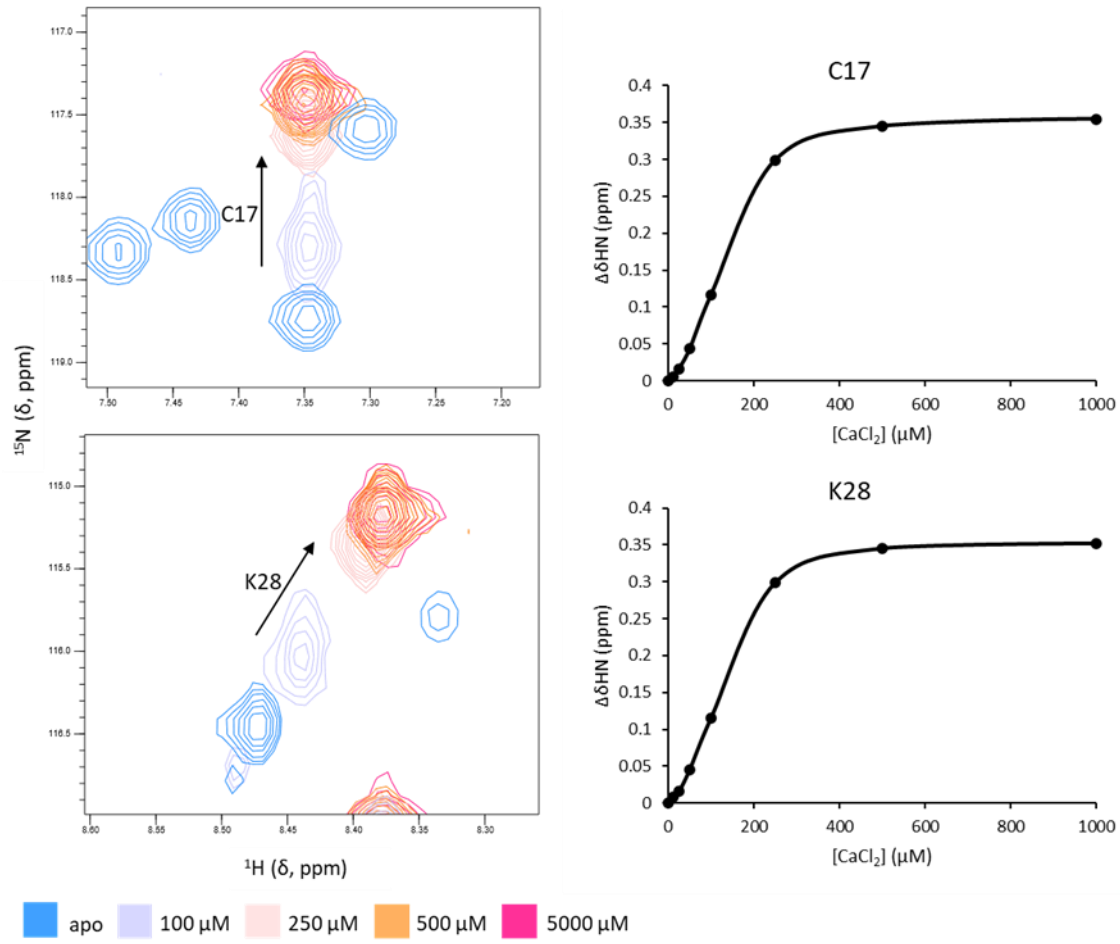

Figure 45 CSPs for example residues in fast exchange regime. The colour code of the six spectra (some of the titration points have been omitted for a clearer picture) is indicated in the bottom left corner. Plots of weighted CSPs at increasing concentration of calcium (12.5, 25, 50, 100, 250, 500 and 1000 μM) for the two residues are shown on the right-hand side of the figure.

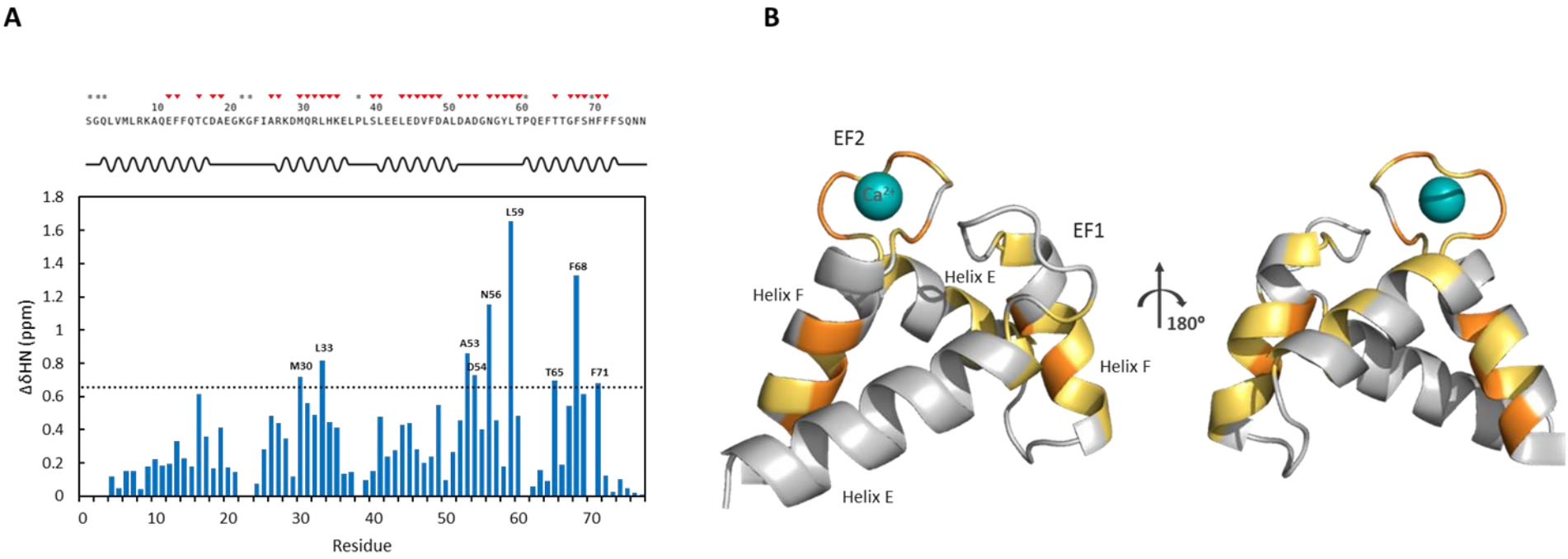

Figure 46 Analysis of CSPs occurring upon calcium binding to the EF-hand domain. (A) NMR CSP plots for calcium binding to EF-hand. Weighted CSPs between the apo and fully saturated EF-hand are plotted against the amino acid residue. Cut-off of 0.66 ppm (mean + SD) is shown as the dotted line with the residues for which CSPs above the cut-off value were observed individually labelled on the plot. Residues which disappear during the titration are marked with red arrows above the secondary structure plot. Residues which are missing in either apo or Ca^{2+}-saturated spectra are marked with grey asterisks. (B) Residues experiencing largest shifts mapped on the crystal structure of EF-hand domain. CSPs above mean and mean + SD are coloured in yellow and orange respectively.

4.4.3. Isothermal Titration Calorimetry

ITC was performed to derive the thermodynamic parameters of Ca^{2+} binding to EF-hand domain. As demonstrated by Figure 47, binding of Ca^{2+} appears to be a single endothermic event at a molar ratio of n = 1.14 ± 0.01 which corresponds to a relatively low-affinity binding of a single Ca^{2+} ion with a K_d of 10.12 µM. The average K_d value obtained from 6 independent ITC experiments was 11.98 ± 0.90 µM. The result obtained from ITC is, therefore, in agreement with only one Ca^{2+} binding loop of EF-hand domain being functional.

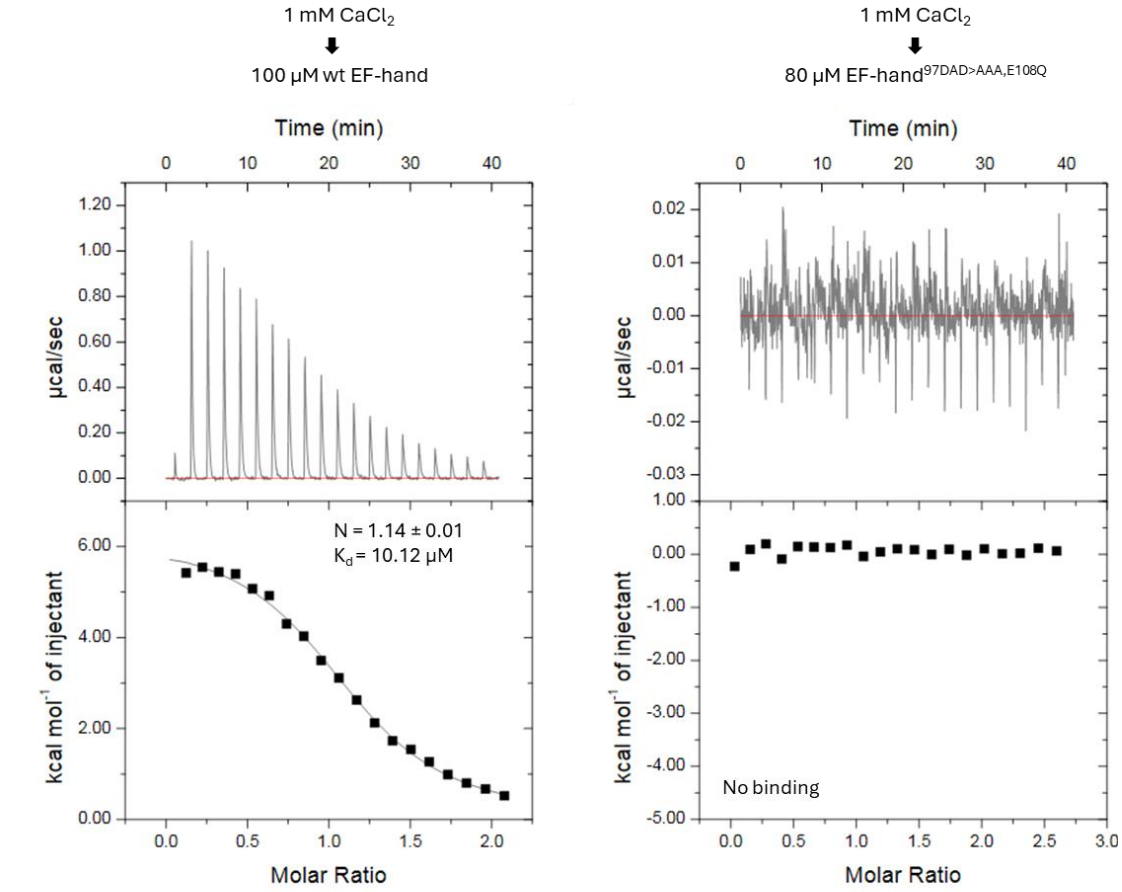

Figure 47 Binding isotherm of EF-hand domain of Rab46. Binding parameters are shown in the box in the bottom panel. The data were fitted with 'one set of sites' model and the binding parameters were calculated using MicroCal ITC-ORIGIN Analysis Software. Dissociation constant (K_d), calculated as a reciprocal of the association constant K calculated from the fit, was found to be 10.12 µM. EF-hand$^{97DAD>AAA,E108Q}$ mutant was used as a negative control for calcium binding studies by ITC. Data shown for the wt EF-hand domain are representative of n=6 independent ITC experiments repeated with similar results.

4.4.4. Dynamics of apo and Ca^{2+}-bound EF-hand domain

Proteins are inherently dynamic molecules, undergoing a variety of structural transitions which, depending on the complexity of the motion, occur on different timescales, ranging from atomic vibrations on the fs-ps timescale and fast loop motions (ps-ns), to more global motions such as the ones associated with conformational changes or protein folding/unfolding, occurring on the µs-s timescale (Figure 48). For that reason, protein dynamics is an important factor defining the protein functional repertoire (130). The major advantage offered by NMR spectroscopy over other types of structural and biophysical methods is that it enables studies of dynamic processes at atomic level over a wide range of timescales.

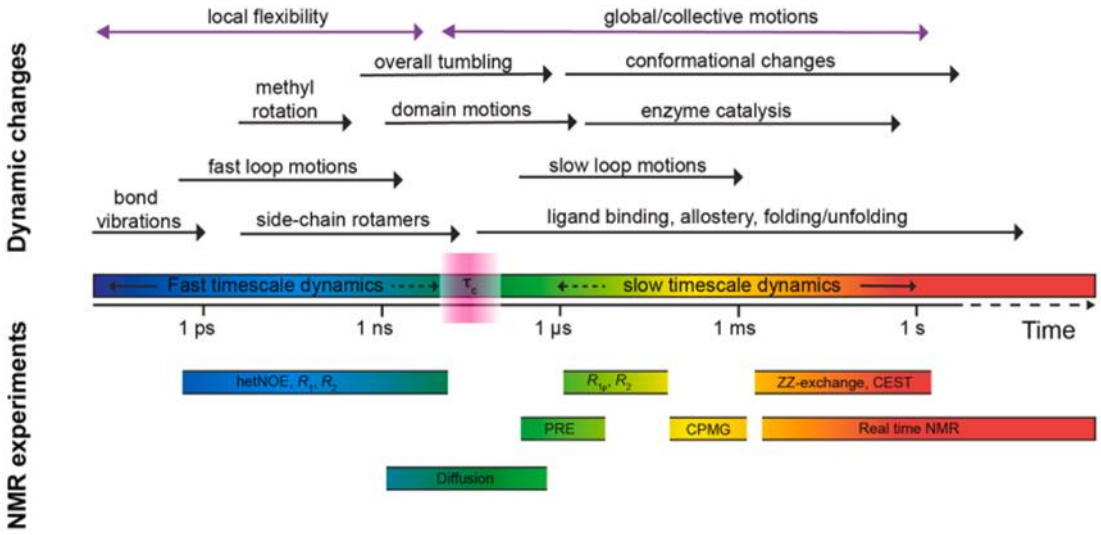

Figure 48 Timescales of dynamic processes accessible to NMR experiments. NMR methods which can be used to discern protein dynamics at different timescales are shown. The figure was reprinted with permission from Kawale and Burmann (130).

As backbone internal motions are traditionally examined by ^{15}N NMR relaxation experiments, T1, T2 and hetNOE experiments, informing about the amplitude of backbone motion on sub-nanosecond timescale, were recorded for apo and Ca^{2+}-bound EF-hand to investigate the effect of Ca^{2+} binding on the dynamic properties of the protein. The obtained plots of the ^{15}N spin-lattice or longitudinal (R_1) relaxation rate, ^{15}N spin-spin or transverse (R_2) relaxation rate, and heteronuclear NOE (hetNOE) are shown in Figure 49. Since hetNOE values in the range of 0.6–0.8 are expected in folded, rigid proteins, the mean hetNOE values of 0.74 ± 0.14 and 0.77 ± 0.13 for apo and Ca^{2+}-bound EF-hand respectively provide another confirmation of the protein being folded. Lower

hetNOE and R_1 values observed for the N- and C-terminal residues indicate fast (ps) motions, consistent with the N- and C-termini being flexible. The R_1 plot shows a comparable trend of values for both apo and Ca^{2+}-bound forms, with somewhat reduced values of R_1 for the residues in the flexible loop connecting the EF1 and EF2 motifs and the ones in the EF2 calcium binding loop, likely arising from fast movements of the two loop regions. Nevertheless, the R_1 and hetNOE plots imply very little change in fast timescale dynamics between apo and Ca^{2+}-bound EF-hand domain. The slightly increased R_2 values in the C-terminal F helix of EF2 are potentially suggestive of this region showing dynamics on a μs-ms timescale when Ca^{2+} is bound, however that finding ought to be confirmed by additional experiments.

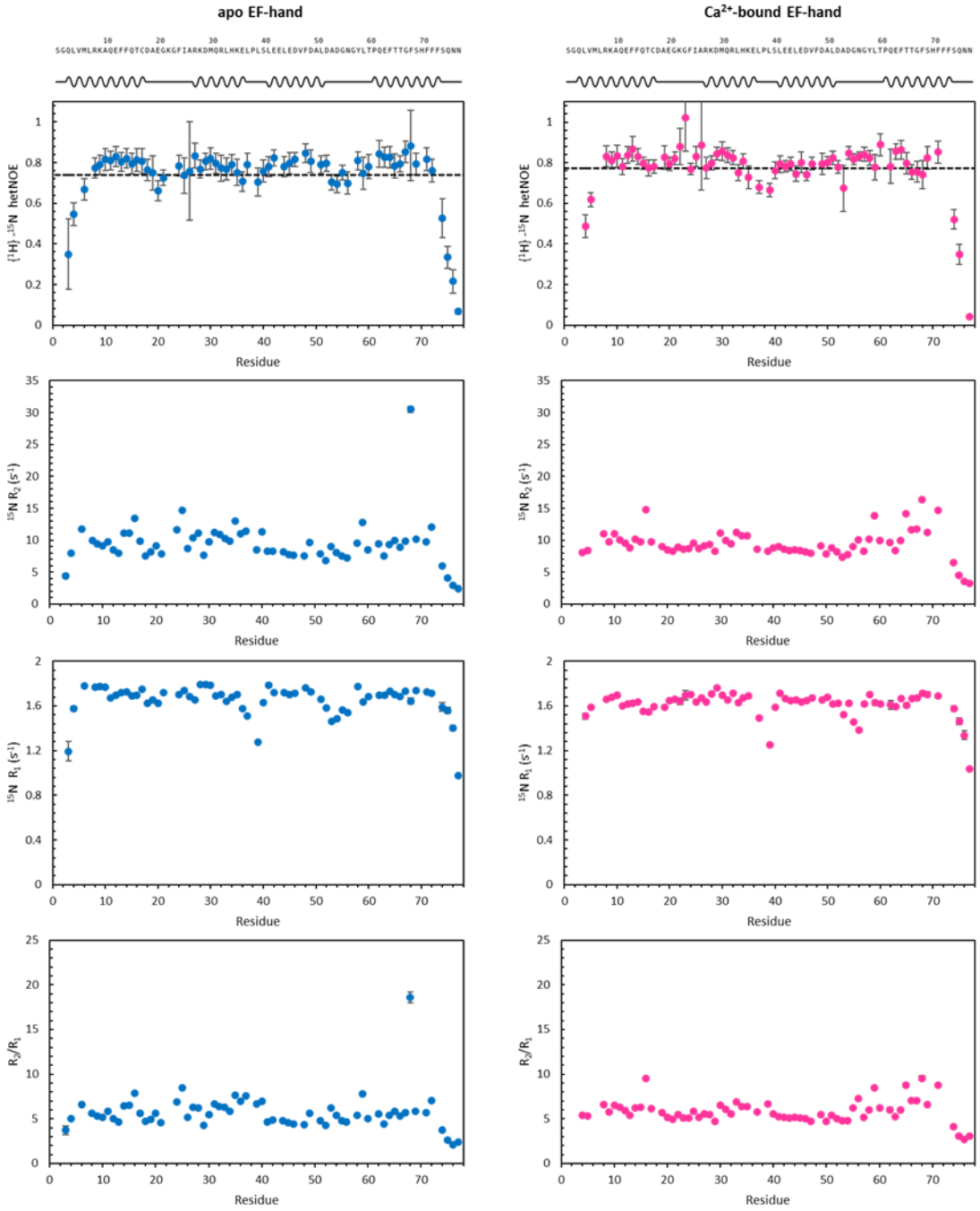

Figure 49 Relaxation properties of apo and Ca^{2+}-bound EF-hand domain. ^{15}N relaxation parameters of EF-hand domain measured at 750 MHz and 25°C. Schematic of the predicted secondary structure elements in EF-hand domain is shown at the top of the relaxation plots. HetNOE, R_2, R_1 and R_2/R_1 plots for apo and Ca^{2+}-bound EF-hand are shown in blue and pink respectively.

Chapter 5

Characterisation of Rab domain of Rab46

5.1. Introduction

Following optimisation of its expression in isotopically labelled media for NMR (described in Chapter 3), Rab domain was the second functional domain of Rab46 for which biophysical and structural characterisation was attempted. Rab domain is the GTPase domain responsible for binding and hydrolysis of GTP, which is central to the function of Rab46 as a molecular switch, alternating between an inactive GDP-bound and an active GTP-bound states. Considering the role of Rab46 GTPase in the context-dependent differential trafficking of WPBs, an understanding of the structural properties of its GTPase domain, as well as the intrinsic and extrinsic mechanisms controlling the levels of inactive and active states of the protein, is of interest to provide insight into its mode of action, and the timing of cellular processes regulated by Rab46.

5.2. Protein fold of Rab domain

5.2.1. Circular Dichroism

Structurally, the GTPase domain of Rab proteins is characterised by a conserved fold consisting of six-stranded mixed β-sheet, flanked by five α-helices, typical of Ras superfamily of proteins (131). Secondary structure content of the recombinant Rab domain of Rab46 purified from *E. coli* (expression of which was described in Chapter 3) was examined by Far-UV CD spectroscopy. Far-UV CD spectrum of Rab domain was measured between 180 nm and 260 nm, and the spectral data were plotted as molar circular dichroism (Δε) against wavelength (Figure 50), in the same manner as the CD spectrum of EF-hand domain (section 4.2.1). As described in section 4.2.1 and illustrated by the example of EF-hand domain, CD spectra of α-helical proteins are characterised by two negative bands, at 208 nm and 222 nm, and a single positive band at 193 nm. Hence, the contribution from the α-helices is immediately apparent upon visual inspection of the CD spectrum of Rab domain (Figure 50, Panel B). Discerning the contribution from the β-sheet structures, as well as an accurate estimation of the β-sheet content from a CD spectrum is, however, a more challenging task as, in contrast to α-helical proteins, proteins containing β-sheets demonstrate significantly larger spectral diversity, depending on the twist and the parallel-antiparallel orientation of the strands in a β-sheet (132). The percentage content of α-helices and β-strands calculated from the acquired CD spectrum of Rab domain using DichroWeb tool (104) was 34% and 18% respectively (Table 14), lower than predicted using PsiPred tool (111) for both secondary

structure elements (37% and 24.9% for α-helix and β-strand respectively, Figure 50, Panel A). The observed discrepancy between the calculated and predicted secondary structure elements' content is, to some extent, likely to be explained by inaccuracies in β-sheet content estimation associated with the aforementioned higher degree of diversity of β-sheet structures' spectral profiles. However, another possible explanation is that the observed difference can be attributed to the nucleotide-bound state of Rab domain. Since recombinantly expressed Rab GTPases are typically purified in a GDP-bound form, and the Rab domain was not loaded with any nucleotide prior to the CD experiment, the protein used in the experiment would be expected to be predominantly in a GDP-bound state. Although the degree of the nucleotide-dependent conformational change varies among Rab proteins, the Switch I and II regions of the GTPase domain generally tend to be disordered in the 'off' GDP-bound conformation (41). This, however, would have not been taken into account in the PsiPred (111) secondary structure prediction, as indeed illustrated by the diagram in Figure 50, where Switch I and II regions (the predicted location of which was identified using NCBI blast) are predicted to be partially structured – with five residues within Switch I (G580-R584 – residue numbering corresponds to the full length protein) predicted in a β-strand, and eleven residues in Switch II (E606-F616) predicted in an α-helix. Given that these would most likely not be expected to be structured in the GDP-bound form of Rab46, the adjusted predicted content of α-helix and β-strand was calculated, based on the PsiPred (111) prediction, as 30.9% and 22.1%, respectively. Although these values still do not perfectly match the experimentally determined secondary structure content, Far-UV CD experiment with Rab domain provides supporting evidence for the protein possessing the expected α/β Rossman fold characteristic of the small GTPase family.

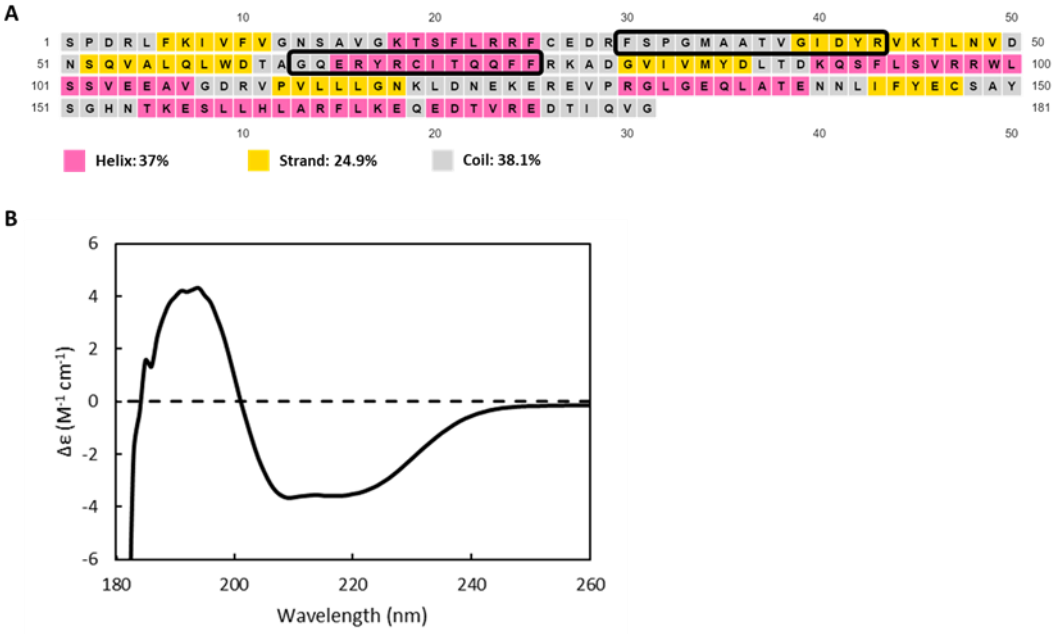

Figure 50 Far-UV CD of Rab domain. A) Secondary structure of Rab domain predicted using PsiPred tool (111). Calculated percentage of each of the secondary structure elements is shown below the sequence plot. Residues within Switch I and II regions are indicated in the secondary structure diagram with black boxes. B) Far-UV CD spectrum of Rab domain in 50 mM sodium phosphate buffer (pH 8.0). The spectrum was acquired over the wavelength range of 260 – 180 nm. Data are plotted as molar circular dichroism (Δε) versus wavelength.

Table 14 Summary of the secondary structure content of Rab domain calculated using DichroWeb CDSSTR method (104).

Helix1	Helix2	Strand1	Strand2	Turns	Unordered
19%	15%	11%	7%	13%	35%

5.2.2. ^{1}H-^{15}N HSQC of Rab domain

Having optimised the expression of isotopically labelled Rab domain, described in section 2.3.5, ^{1}H-^{15}N HSQC experiment was acquired to assess the protein's suitability for triple resonance backbone assignment, aiming to gain further insight into its structural and biophysical properties. The protein used in the HSQC experiment was not pre-

loaded with any non-hydrolysable nucleotide analogue and hence the spectrum demonstrated in Panel A of Figure 51 would be expected to represent the GDP-bound state of Rab domain. As expected for a well-folded protein, the ^{1}H-^{15}N HSQC spectrum of the (presumed) GDP-bound Rab domain shows a good chemical shift dispersion. Since the Rab domain of Rab46 comprises 181 amino acids, four of which are proline residues, approximately 176 signals are expected in its HSQC spectrum. Using CcpNMR analysis software (93), the number of cross peaks in the HSQC spectrum of Rab domain was roughly estimated at 190 peaks, excluding the obvious side-chain peaks in the top right corner of the spectrum. Of note, peak intensities vary significantly across the HSQC spectrum of the (presumed) GDP-bound Rab domain. The observed higher than expected number of signals, together with a combination of strong and weak peak intensities, may be suggestive of slow conformational exchange between different states (e.g. monomer and dimer), or conformational heterogeneity of GDP-bound Rab domain. Moreover, the presence of a number of overlapped, broader resonances in the central region of the spectrum could correspond to the protein sampling a different conformation. The fact that the switch regions are expected to be unordered and mobile in the GDP-bound form of Rab proteins could therefore offer one potential explanation for the observed signal intensity pattern. It is also plausible that the produced Rab domain sample is in fact heterogeneous and the observed difference in signal intensities arises from the presence of two or more differently populated states – e.g. a more populated GDP-bound and a less populated GTP-bound Rab. These hypotheses would, however, have to be tested by means of performing additional NMR experiments (e.g. ZZ-exchange should we be able to assign the protein) and Mass Spectrometry experiments to examine the nucleotide-bound state of the purified recombinant Rab domain.

An important limitation of the recombinantly produced Rab domain protein, revealed by the NMR experiment was that, under the experimental conditions, and likely due to the high protein concentration used in the experiment (250 µM i.e. 5.1 mg/mL), the sample experiences heavy precipitation, despite relatively short duration of the experiment. To address the sample solubility issues, ^{1}H-^{15}N HSQC experiments were also acquired with Rab domain at lower temperatures (10ºC and 15ºC; data not shown), nevertheless, it did not seem to alleviate precipitation of the Rab domain sample. Thus, as it stands, the protein proved not to be suitable for recording experiments for triple-resonance backbone assignment.

5.2.3. ^{1}H-^{15}N TROSY of GppNHp-loaded Rab domain

To examine whether the recombinant Rab domain protein can be successfully stabilised by locking it in an active, GTP-bound conformation, the protein was loaded with 20-fold molar excess of GppNHp – a non-hydrolysable analogue of GTP. Aiming to improve the spectral resolution and reduce the signal crowding in the middle region of the spectrum of the GppNHp-loaded Rab domain sample, a ^{1}H-^{15}N TROSY experiment was acquired at 15ºC instead of a ^{1}H-^{15}N HSQC. Due to the sample availability limitations and sample losses associated with buffer-exchange prior, and post loading with GppNHp, the spectrum was acquired from a 60 µM sample (1.2 mg/mL) with an increased number of scans. Therefore, though some similarities between the spectra of GDP- and GppNHP-loaded Rab domain are apparent (Figure 51), the two spectra are not directly comparable due to the differences in the experimental setup, including both sample conditions, and acquisition parameters, resulting in cross peaks in the two spectra being shifted relative to each other in both the F2 and F1 domains. Therefore, the spectra were not overlaid, and are presented next to one another in Figure 51 instead. As demonstrated in Panel B of Figure 51, the spectrum of GppNHp-loaded Rab is characterised by little signal overlap and somewhat better signal dispersion than the ^{1}H-^{15}N HSQC spectrum of the (presumed) GDP-bound Rab domain. Furthermore, compared to the ^{1}H-^{15}N HSQC of GDP-Rab, the signal intensity appears more uniform across the ^{1}H-^{15}N TROSY spectrum of GppNHp-Rab, though some weaker cross peaks are also present. As speculated in case of the GDP-bound Rab, the presence of weaker cross peaks could be suggestive of either slow exchange between different conformational states, or conformational heterogeneity of the protein sample. Moreover, the number of peaks in the TROSY spectrum of GppNHp-loaded Rab calculated in CcpNMR software (93) was found to be lower than expected (159 peaks), which could potentially arise from conformational exchange on an intermediate timescale, resulting in some of the resonances becoming broadened beyond detection. Indeed, residues located within Switch regions were reported to undergo conformational exchange on the µs-ms timescale, resulting in broadening of the resonances beyond detection (59). Nevertheless, having an unassigned spectrum of the protein, it is impossible to draw conclusions on where the missing cross peaks are located within the Rab domain. On the other hand, the observed variability in peak intensities, taken together with the reduced number of peaks in the spectrum of GppNHp-loaded Rab, raises a question of how efficient the nucleotide loading was. Hence, as in the case of the GDP-bound sample, it would be extremely useful to obtain Mass Spectrometry confirmation of the nucleotide-bound state, to determine the loading efficiency. Considering the evidence presented for MANT-GDP-

loaded Rab domain described below in section 5.2.6, it is likely that the loading was not particularly efficient and the method would require further optimisation, perhaps by introducing an additional purification step, to allow obtaining a homogenous GppNHp-bound protein sample.

Similarly as for the presumed GDP-bound sample, a considerable amount of precipitation was observed in the NMR tube for the GppNHp-loaded Rab domain following the experiment, likely due to the longer acquisition time required to compensate for the used lower concentration of the protein. Therefore, the described NMR experiments demonstrated that, under current conditions, obtaining the experiments for triple resonance backbone assignment of Rab domain would simply not be feasible due to the encountered issues arising from limited sample stability, high aggregation propensity, together with the potential conformational heterogeneity of the produced protein samples.

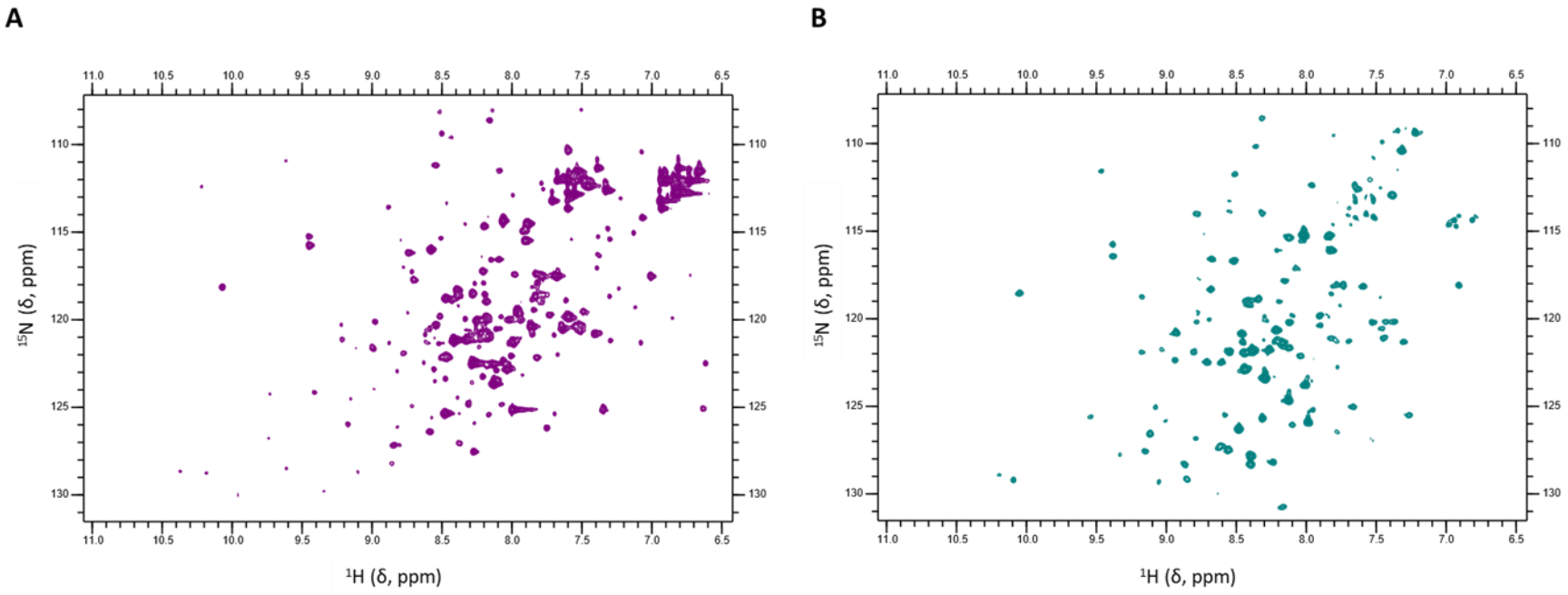

Figure 51 ^{1}H-^{15}N HSQC and TROSY spectra of Rab domain. (A) ^{1}H-^{15}N HSQC spectrum of GDP-Rab domain. The spectrum was recorded using 250 µM Rab domain uniformly labelled with ^{15}N, prepared in a buffer containing 25 mM Hepes (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 1 mM TCEP and 0.1% NaN_3. (B) ^{1}H-^{15}N TROSY spectrum of GppNHp-Rab domain. The spectrum was recorded using 60 µM Rab domain uniformly labelled with ^{15}N, loaded with 20-fold molar excess of GppNHp and prepared in a buffer containing 25 mM Hepes (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 1 mM TCEP and 0.1% NaN_3. The data were acquired on a 750 MHz Bruker spectrometer at 25ºC (^{1}H-^{15}N HSQC) and 15ºC (^{1}H-^{15}N TROSY).

5.2.4. Optimisation of Rab domain construct

Since the modification of buffer composition by varying ionic strength, pH and including additives such as arginine and glutamic acid did not appear to have any effect on the poor solubility of Rab domain in a hanging drop buffer screen (data not shown), an attempt was made to try and optimise the Rab domain expression construct to address the issue with limited protein solubility. The intrinsically disordered hypervariable C-terminal region of Rab proteins, which is often removed for structure determination purposes (133), was already not present in the original Rab domain construct spanning residues from P543 to G722. Two different truncation mutants (P543-705R and P543-710Q) of the original Rab domain construct in pET32a-LIC vector were therefore designed based on the alternative NCBI Blast (47) predictions of the domain boundaries. These Rab domain mutants therefore contain a shorter C-terminal region, part of which is predicted to be disordered, however, a small part of the C-terminal helix is also removed in both constructs. The expression of the two constructs was first tested in Rosetta (DE3) and Rosetta 2 (DE3) strains and, having confirmed higher level of soluble protein expression in Rosetta 2 (DE3) cells (data not shown), the proteins were expressed at a 2 L scale. Unfortunately, the two constructs did not purify well large-scale, with both P543-R705 and P543-Q710 appearing to suffer from even worse solubility issues than the original Rab domain construct. For the P543-R705 construct, not only was the expression level lower than that of the original P543-G722 construct, but also a significant amount of the protein was found to be present in the insoluble fraction (Figure 52). The amount and purity of P543-R705 protein after IMAC was, therefore, deemed unsuitable to proceed with further purification steps. In case of P543-Q710 construct, the expression level was higher and the yield after the capture step, though not as good, was somewhat closer to the original Rab domain construct. Nevertheless, the protein heavily precipitated during the tag cleavage step performed by dialysis with his-tagged TEV protease and, due to the associated sample losses, the purification of P543-Q710 was also aborted at this stage. However, it is important to mention that the large-scale purification of the two alternative Rab domain constructs was only attempted once, and it would potentially be beneficial to try to repeat the expression, and purification of P543-Q710, potentially testing modified buffer conditions. Nevertheless, from the SDS-PAGE analysis it is apparent that both constructs are considerably dirtier than the original Rab domain construct (described in 3.3.2) following the capture step (Figure 52), and therefore optimisation of these may not be worth pursuing any further.

Based on the obtained results it appears that the C-terminal helix, which was partly truncated in the new Rab domain constructs, may be of importance for the protein stability. Alternative constructs should be explored and surface mutagenesis (e.g. introducing mutations to the solvent-accessible cysteine residues as, for example, C610 residue is surface exposed in the AlphaFold model of Rab46) may be beneficial to reduce nonspecific protein aggregation and precipitation and increase the chances of obtaining a well-behaved protein for triple resonance assignment, and the subsequent structural characterisation. Lastly, ensuring the sample homogeneity by means of optimising the nucleotide loading protocol to successfully lock the Rab protein in either its inactive or active conformation may also translate to its improved stability and suitability for structural studies.

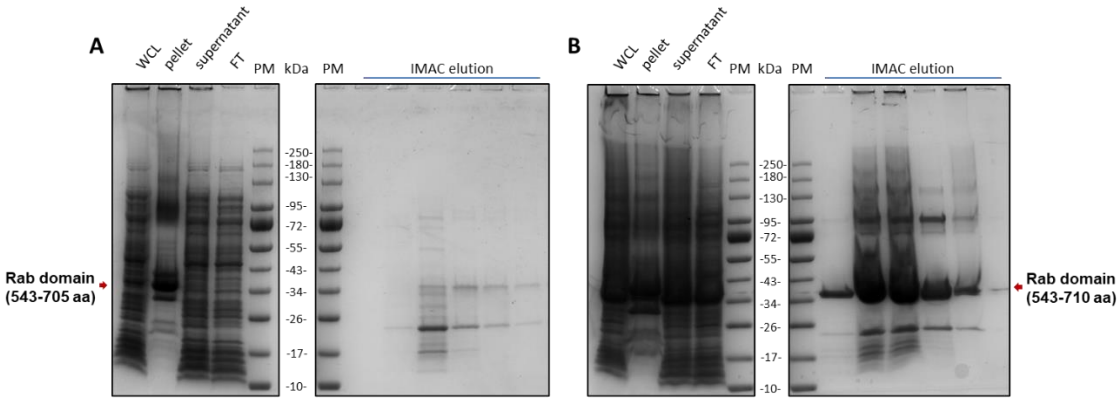

Figure 52 SDS-PAGE analysis of large-scale affinity capture of the alternative Rab domain expression constructs. A) Affinity capture of his-tagged P543-R705 Rab domain construct. B) Affinity capture of his-tagged P543-Q710 Rab domain construct. WCL – Whole-cell lysate, FT – Flow through, PM – Protein marker.

5.2.5. Rab domain GTPase activity

To assess whether the recombinantly produced Rab domain protein is functional, the colorimetric malachite green reagent-based assay (Sigma-Aldrich) was used to examine its intrinsic GTPase activity. Unfortunately, since malachite green assay relies on the measurement of the release of inorganic phosphate following GTP hydrolysis, the method is highly sensitive to impurities, and the experiments with Rab domain failed due to issues with free phosphate present both in the sample buffer, and the purified protein preparation, resulting in elevated background signal across the whole assay plate. Although the hydrolysis rates of Rab GTPases can vary significantly among different

members of the Rab family of proteins, they generally tend to be very slow, unless assisted by Rab-specific GTPase-activating proteins (Rab GAPs) (41,133). Hence, the signal from the inorganic phosphate liberated as a result of GTP hydrolysis during a relatively short time course of the experiment would be obscured by high background levels unless the issue with the reagents' contamination with free phosphate is addressed.

5.2.6. Guanine nucleotide exchange

To monitor the nucleotide exchange of GDP to GTP by recombinantly produced Rab domain of Rab46, a fluorescent MANT-GDP-based nucleotide exchange experiment was performed, following the method described by Kanie and Jackson (2018) (106). Rab domain was first exchanged into a buffer containing low concentration of Mg^{2+} (0.5 mM), and subsequently loaded with 20-fold molar excess of MANT-GDP in the presence of 10 mM EDTA. The unbound MANT-GDP was then removed from the protein sample by buffer-exchange, and the loading efficiency was calculated with the standard curves generated using 0.156 μM – 40 μM MANT-GDP, as a ratio of the molar concentration of Rab-bound MANT-GDP to the molar concentration of Rab domain protein sample. The average loading efficiency from three performed nucleotide exchange experiments was 35.12 ± 2.74%.

The nucleotide exchange reaction was initiated by addition of GppNHp (or buffer for the control samples) to obtain a 100-fold molar excess of the GTP analogue over Rab domain protein, and the fluorescence signal was measured every 50 s for 65 minutes. The plot in Figure 53 shows the representative relative fluorescence curves for GppNHp and control Rab domain samples, as well as the curve normalised by taking the ratio of GppNHp sample fluorescence to the buffer sample fluorescence, to account for the observed gradual decrease in MANT-GDP fluorescence intensity over the time course of the experiment.

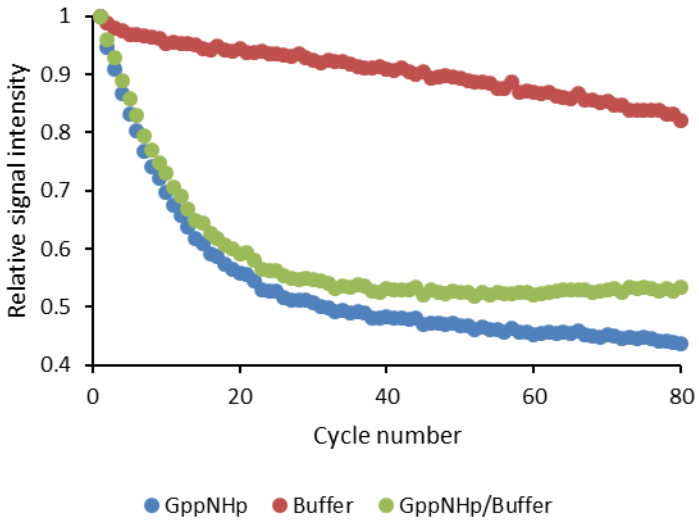

Figure 53 Relative fluorescence of MANT-GDP-loaded Rab domain samples during the nucleotide exchange experiment. Data shown in the plot was averaged from four technical replicates and is representative of a single experiment (n=1).

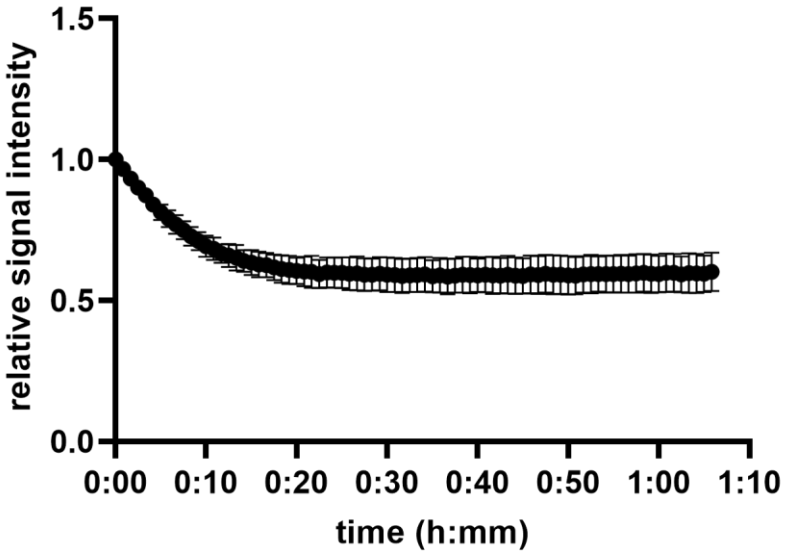

Figure 54 Plot of MANT-GDP fluorescence decay during the time course of nucleotide exchange with GppNHp. Nucleotide exchange was monitored by measuring the fluorescent signal during the release of MANT-GDP from Rab domain induced by addition of 100-fold molar excess of GppNHp. Data shown in the plot are averaged from three independent nucleotide exchange experiments. Two batches of Rab domain protein were tested in each experiment with four technical replicates performed for each sample. Error bars represent SD (n=3).

The plot for the nucleotide exchange experiment showing the averaged data obtained from three independent nucleotide exchange experiments is demonstrated in Figure 54. The observed nucleotide exchange rate constant value (K_{obs}) was calculated as $2.35 \times 10^{-3}\ s^{-1}$ by fitting the data to single exponential function in GraphPad Prism software (107). Of note, the calculated K_{obs} value would indicate relatively high intrinsic GDP/GTP exchange activity of Rab domain of Rab46, higher than that of traditional GEF-dependent GTPases (134). However, since MANT-nucleotides have been reported to alter rates of both hydrolysis and nucleotide exchange, these findings would have to be validated by additional experiments. Nevertheless, the nucleotide exchange experiment not only demonstrated that the recombinantly produced Rab domain is functional, but also provided initial evidence of its (potentially) unique GTPase profile (should the obtained result be confirmed by additional experiments), which, given the fact that the concentration of GTP in the cell is 5 to 10 times higher than GDP, could imply that, in the absence of regulatory mechanisms, Rab46 may have the capacity to self-activate and persist in an active state in cells.

Chapter 6

Discussion

Rab46 is an important regulator of T-cell signalling and differential WPB trafficking. The exact molecular mechanisms underlying its function, largely dictated by its structural and dynamic properties as well as interactions with its binding partners, however, remain poorly understood. At the time this project was commenced there had been no structural studies of this large, multidomain Rab GTPase. What is more, the number of studies of the other non-canonical, large Rab GTPases is also limited and thus very little is known about the molecular mechanisms underlying their modes of action. Understanding of the structural properties of Rab46 GTPase, as well as the intrinsic and extrinsic mechanisms controlling the levels of inactive and active states of the protein, is therefore of interest to provide insight into its mode of action, and the timing of cellular processes regulated by Rab46. Furthermore, while the roles and mechanisms of this novel protein in human disease are still being unravelled, considering the important function of Rab46 in regulation of differential WPB exocytosis, understanding of the 3D atomic structure and dynamic regulation of Rab46 may contribute to the discovery of novel targets for therapeutic intervention in cardiovascular disease. In this this study, we present novel insights into the structural characteristics of the two functional domains of Rab46, EF-hand and Rab domain. Crucially, we demonstrate, for the first time, evidence supporting conformational change of the EF-hand domain of Rab46 upon Ca^{2+} binding, leading to the Ca^{2+}-bound domain adopting an open conformation. Furthermore, our findings reveal existence of more than one conformational state of the Ca^{2+}-bound EF-hand domain in solution. Finally, we describe initial evidence pointing to an atypical GTPase profile of Rab46 characterised by high intrinsic nucleotide exchange rate of its GTPase domain. Altogether, the results presented in this thesis shed light on some of the distinctive properties of Rab46 and provide a solid starting point for future structural characterisation of this atypical large Rab GTPase.

6.1. Recombinant protein production for structural characterisation

As described in Chapter 3 of this thesis, protocols for overexpression of EF-hand, Coiled-coil and Rab domains of Rab46 in *E. coli* were successfully developed from scratch. Besides the conditions deemed optimal for the large-scale expression of individual Rab46 domains, additional conditions offering alternative expression routes, should these be required in the future studies, have been identified in the small-scale expression screen. The subsequent assessment of the recombinantly produced domains of Rab46

by Circular Dichroism, Mass Spectrometry and NMR Spectroscopy demonstrated that the established expression and purification system was suitable for generation of high-quality protein samples for structural and biophysical investigations.

Unfortunately, despite exhaustive attempts at producing full length Rab46 protein, no suitable sample could be obtained for structural characterisation. In line with the findings presented in this thesis, the observation that FL Rab46 protein expressed in *E.coli* was prone to degradation was also reported by Wang *et al.* (74). In their study, published during the course of this project, Wang and colleagues used a pET28 vector-based construct to express a His-Strep-GFP-tagged Rab46 in BL21-CodonPlus (DE3)-RIPL strain (74), broadly equivalent to Rosetta or Rosetta 2 strains utilised for expression of FL Rab46 described in this thesis. Presence of both his- and Strep-tag on the construct used by Wang and colleagues allowed efficient purification of the protein by initial capture on a nickel column, followed by another affinity capture step on a StrepTrap column, and final polishing by SEC. Compared to the method described in this thesis, this tandem purification strategy allowed obtaining a significantly purer protein sample (74), nevertheless the protein was still reported to suffer from degradation (74). Going forward it would be of interest to further investigate the identity of the truncated 65 kDa product of proteolytic degradation observed during expression of the full length Rab46 in *E. coli* (described in Chapter 3 of this thesis), for example by means of peptide mapping of the protein band excised from a polyacrylamide gel. Since extended, unstructured protein sequences are not well protected from proteolysis, it is plausible that the 65 kDa protein corresponds to a stable fragment of Rab46, resulting from proteolytic cleavage of the exposed, unstructured regions of the protein, which, if confirmed, could potentially prove to be very useful for designing an expression construct for probing the inter-domain interactions within Rab46 GTPase.

While problems with the full length Rab46 expression in *E. coli*, a system which is typically better suited for production of short or single domain proteins (135), were to some extent expected, expression in a eukaryotic system was anticipated to help overcome the issues around poor protein stability. Nevertheless, suboptimal construct design was likely the predominant factor contributing to the limited success of the FL protein production in *Sf9* cells. Both the *E. coli* and *Sf9* FL Rab46 expression constructs used in this study were designed to encode hexahistidine-tagged full length wild type protein. It would have been perhaps a more practical strategy to simultaneously explore multiple construct variants generated e.g. by truncating the predicted unstructured regions of the protein and/or introducing different fusion tags, enabling to pursue

alternative purification strategies, to increase the success rate of expression and purification of the target protein. Furthermore, co-expression of the full length Rab46 in the presence of a binding partner to assemble in a multiprotein complex could offer a solution to overcome the issues with poor target stability. Multidomain proteins participating in large macromolecular complexes, with inter- and/or intradomain flexibility, exhibit a strong tendency for aggregation (136). Therefore, purification of such proteins, containing disordered regions involved in transient complex formation, is often challenging due to their heterogeneity, poor solubility and instability (136). Interactions of the protein with a binding partner may induce partial folding of the disordered regions. Thus, since Rab46 is predicted to contain a large, disordered region (PRD domain) demonstrated to be involved in interaction with a Rho/Rac guanine nucleotide exchange factor Vav1 in T-cells (63), co-expression with a binding partner such as Vav1, dynein heavy chain (DHC), or light intermediate chain-1 (LIC-1) could be a potential strategy for stabilisation of recombinant Rab46 (68,77). Expression of such complex could be performed in *Sf9* cells utilising pFastBac Dual vector which allows simultaneous expression of two protein targets using Bac-to-Bac baculovirus system. Although significantly more expensive than labelled protein expression in *E. coli*, isotope labelling is possible to be achieved in insect cells (137). Therefore, provided that sufficient amounts of stable protein could be obtained from insect cell expression following its optimisation, going forward NMR complemented by lower-resolution structural techniques such as small-angle X-ray and neutron scattering (SAXS/SANS) could potentially still be utilised to investigate the structural and dynamic properties of full length Rab46 GTPase.

6.2. Characterisation of EF-hand domain

Binding of calcium to the EF-hand domain of Rab46 has been demonstrated to be necessary for dispersion of WPBs anchored to MTOC in response to histamine stimulation (33). The results demonstrated in Chapter 4 of this thesis represent an initial attempt to characterise the EF-hand domain of Rab46 GTPase in solution and provide evidence that the domain is fully functional and subject to a significant conformational rearrangement upon Ca^{2+} binding. The near-complete triple resonance backbone assignment of both apo and Ca^{2+}-bound forms of EF-hand domain, presented for the first time in this study, provides a solid foundation for future characterisation of the solution structure and dynamics of the domain. Furthermore, the obtained backbone assignment will enable residue-specific tracking of interactions, for example allowing to validate and further investigate the interaction between EF-hand domain and LIC1 in solution (68).

Importantly, the findings presented for the EF-hand domain are of relevance to function of both CRACR2A-S and Rab46 as the sequence of EF-hand domain is shared between the two isoforms of CRACR2A.

6.2.1. Ca^{2+}-induced opening of EF-hand domain of Rab46

Classical Ca^{2+} sensors show profound Ca^{2+}-induced conformational rearrangements, typically resulting in a transition from a closed to an open domain conformation. Previous studies characterised the short isoform of Rab46 as a cytosolic Ca^{2+} sensor (64) and therefore, the major conformational rearrangement of its EF-hand domain is expected, as it represents a crucial step, necessary for target recognition and coupling of transient elevation of intracellular Ca^{2+} levels with appropriate regulatory processes. CD thermal denaturation experiment with EF-hand domain revealed a stabilising effect of the Ca^{2+}-induced conformational rearrangement, however it would be beneficial to repeat the experiment to gain more confidence in the derived T_m values. CS-Rosetta model of apo EF-hand domain obtained using the backbone chemical shifts offered an initial insight into the extent of the conformational change induced by Ca^{2+} binding. Additionally, the NMR Ca^{2+} titration experiment and the subsequent CSPs analysis allowed to further pinpoint the location of residues involved in Ca^{2+} induced conformational change to the calcium binding loop and E, and F helices of EF2, as well as F helix of EF1 motifs.

Comparison of solvent accessible surface area (SASA) of apo and Ca^{2+}-bound CS-Rosetta generated structures, and a crystal structure of Ca^{2+}-bound EF-hand provided evidence for the Ca^{2+}-induced domain opening. This finding was additionally supported by IMS-MS analysis, demonstrating a shift to higher charge states in the spectra of Ca^{2+}-bound form likely as a result of exposure of a greater number of protonation sites, consistent with the domain adopting a more open conformation. Nevertheless, the results of SASA analysis should be treated with caution as the CS-Rosetta structure of apo EF-hand used in the SASA comparison has not been extensively validated (and Ca^{2+}-bound CS-Rosetta model has not been validated at all). While fair agreement between experimentally derived RDCs and RDCs back-calculated from the apo EF-hand model was observed, it is important to note that a large number of RDCs was excluded for the residues whose resonances were overlapped in the spectra. Measuring additional sets of RDCs using different alignment media would therefore be beneficial to reduce ambiguity and to further validate and/or refine the obtained model of apo EF-hand (138). Furthermore, an important caveat in making comparisons between the crystal and CS-Rosetta structures is that, in the crystal structure, Ca^{2+}-bound EF-hand domain is complexed with LIC1 helix (68). While domain opening in the presence of a target

sequence without the requirement for Ca^{2+} pre-activation has been reported for some EF-hand proteins, the hydrophobic cleft enabling interaction with LIC1 helix appears to form only upon calcium binding (68), as apo EF-hand has been shown not to bind LIC1. Ideally, solution structures of Ca^{2+}-free and bound EF-hand forms, not complexed with any binding partners would be compared to characterise the conformational change occurring purely as an effect of Ca^{2+} binding to the domain. Nevertheless, while the presented results would require further investigation, they provide initial evidence for the domain adopting a closed conformation in apo form and an open conformation upon Ca^{2+} binding, as would be expected for a Ca^{2+} sensor protein (139).

6.2.2. Downfield-shifted glycine residues

The unexpectedly large downfield chemical shifts of G23 and G57 residues in EF-hand domain spectra were speculated to be an effect of hydrogen bonding to D18 and D52 residues, potentially amplified by a ring current from the neighbouring aromatic residues. It appears, however, that the markedly large low-field chemical shift of glycine residues at this position in the Ca^{2+} binding loop (i.e. connecting Z and -Y positions) is not so uncommon feature of the EF-hand-type Ca^{2+} binding site. Examples of similarly downfield-shifted glycine residues hydrogen-bonded to aspartate residues have been reported in the literature for Ca^{2+}-bound forms of EF-hand domains of calmodulin (140–142). In the case of EF-hand domain of Rab46, however, downfield-shifted glycine residues are present in the spectra of both apo and Ca^{2+}-bound states, potentially suggesting a difference in the conformation of the Ca^{2+}-binding loop between apo forms of EF-hand domains of Rab46 and calmodulin.

6.2.3. Two conformational states of Ca^{2+}-bound EF-hand domain in solution

Remarkably, EF-hand domain of Rab46 exhibited two structural forms in the Ca^{2+}-bound state, as indicated by peak doubling in the ^{1}H-^{15}N HSQC spectrum and two distinct tracts of connectivity in the 3D spectra acquired for the triple resonance backbone assignment of the protein. The doubled peaks were found to be predominantly localised around the Ca^{2+}-binding loops, though it is plausible that the peaks of residues located elsewhere in the protein could be broadened beyond detection or buried in the overlapped region of the spectrum. Additional experiments were performed to explore the nature of the observed peak doubling and whether it was due to conformational heterogeneity or slow interconversion between two or more states. Although the result of ZZ-exchange experiment appeared to support the notion of conformational heterogeneity, it was noted that the interconversion between different states could occur too slowly to be picked up by the ZZ-exchange experiment. If the possibility for the observed pattern being due to

slow proline isomerisation was to be examined, peptidylprolyl isomerase A could potentially be used to catalyse cis-trans proline isomerisation to probe whether exchange peaks appear in the intersections of auto peak pairs.

It was further speculated that the big differences in the weighted distances between some of the peaks of the major and minor conformational states (particularly G23 and A19), likely reflecting a significant change in the chemical environment these residues are in, may be associated with the Ca^{2+}-binding loop adopting a different conformation, resulting in the hydrogen bond between G23 and D18 being weakened or destroyed in the minor form manifesting by an upfield shift of the G23 peak. Importantly, investigation of the effect of pH on the Ca^{2+}-bound EF-hand domain, monitored by ^{1}H-^{15}N HSQC experiment, provided evidence for the minor conformational state of EF-hand becoming more populated at low pH. Examples of similar peak doubling patterns in ^{1}H-^{15}N HSQC spectra of Ca^{2+}-binding proteins have been reported in the literature (143,144). For instance, presence of structured and unstructured forms of STIM luminal domain at low Ca^{2+} concentration was manifested by peak doubling in the ^{1}H-^{15}N HSQC spectrum and two tracts of connectivity in the 3D spectra (143) in the same way as observed for Ca^{2+}-bound EF-hand domain of Rab46. Interestingly, in an NMR study of mouse Type II cadherin-8 (8EC1), a non-EF-hand calcium binding protein, low pH and high calcium concentration have been demonstrated to favour dimer formation, resulting in peak doubling in the ^{1}H-^{15}N HSQC spectrum of 8EC1 (144). Thus, in light of the observation of the low pH promoting the population of the minor form of EF-hand of Rab46, it is plausible that the observed peak doubling corresponds to monomer- and dimer-state resonances in slow exchange on the chemical shift timescale. While this could be a potential explanation for the peak pattern observed in the Ca^{2+}-bound EF-hand spectra, it is somewhat speculative and would require further investigation. It would therefore be interesting to further examine the pH dependency of the minor conformational form of Ca^{2+}-bound EF-hand domain by performing additional NMR experiments and employing other techniques such as analytical SEC to probe the oligomeric state of EF-hand domain at different pH values.

6.2.4. Ca^{2+}-binding affinity of EF-hand domain

EF-hand proteins display a wide range of Ca^{2+} binding affinities (145,146). ITC analysis of Ca^{2+} binding to EF-hand domain of Rab46 revealed a single, relatively low-affinity Ca^{2+}-binding site (K_d =11.98 ± 0.90 μM), consistent with only one EF-hand motif (EF2) being involved in Ca^{2+} binding. However, it is important to note that the affinity can be significantly different for full length Rab46, as interdomain contacts are likely to influence

Ca^{2+} binding (146). Of note, ITC analysis suggested an endothermic nature of Ca^{2+} binding, likely associated with the Ca^{2+}-induced conformational rearrangement of the EF-hand domain. Importantly, the obtained binding isotherm, as well as the derived binding parameters, are in agreement with the ones reported by Lee and colleagues for the MBP-tagged EF-hand domain of Rab46 (68).

6.3. Characterisation of Rab domain

Further to characterisation of EF-hand domain described in Chapter 4, Chapter 5 focused on the Rab domain of Rab46. Importantly, the results presented in this Chapter provide preliminary evidence that Rab domain of Rab46 is functional and has an intrinsic GDP/GTP exchange activity, higher than that of traditional GEF-dependent GTPase. This finding may additionally imply that the nucleotide binding to Rab46 is weaker than the pM-level K_d values of other small GTPases such as Ras (147). Binding of nucleotides with micromolar affinities and ability of a small GTPase to spontaneously exchange GDP to GTP may imply absence of a traditional GEF and suggest involvement of other forms of GTPase regulatory factors (106).

The intrinsic hydrolysis rates of Rab GTPases are typically slow. Therefore, while the intrinsic GTPase activity of Rab46 would have to be first experimentally determined to allow drawing strong conclusions, the observed high intrinsic nucleotide exchange rate points to a unique GTPase profile of Rab domain, which could imply that, in the absence of regulatory mechanisms, Rab46 GTPase may have the capacity to self-activate and persist in an active state in cells. Examples of such fast cycling Rab GTPases characterised by low GTPase and high nucleotide-exchange activities include Rab4a and RABL2B (59,134). Nevertheless, as mentioned above, determination of the rate of hydrolysis of Rab46 and validation of the obtained nucleotide exchange rate would be required to classify Rab46 as a fast cycling GTPase with confidence. On that note, as MANT-nucleotides have been demonstrated to alter the nucleotide exchange rates (148) it would be advisable to explore other experimental routes. HSQC experiment-based NMR methods have been used with success to determine GTP hydrolysis and nucleotide exchange rates of small GTPases (59,149). Nevertheless, in order to pursue such approaches suitable protein samples, stable enough to withstand the long NMR experiments at high protein concentration, would have to be first obtained to enable triple-resonance backbone assignment of the GDP-bound and GTP-bound Rab domain. Since the Rab domain sample used in this study proved unsuitable for recording experiments for triple-resonance backbone assignment, alternative constructs or methods for locking and stabilising the GTPase in either its active or inactive state would

have to be explored to increase the chances of obtaining a well-behaved protein sample for biophysical and structural characterisation.

Lastly, an important consideration throughout the investigation of Rab domain properties, applicable to interpretation of both CD and NMR results, is its unconfirmed nucleotide-bound state. While recombinantly expressed Rab GTPases are typically purified in a GDP-bound form (50), Rab6A is an example of Rab protein isolated as a mixture of GDP- and GTP-bound forms owing to its extremely low intrinsic GTPase activity (133). Hence, a possibility that a heterogeneous sample of Rab domain of Rab46 was obtained post purification cannot be excluded without additional experiments. It would therefore be extremely useful to obtain native mass spectra to investigate the nucleotide-bound state of the purified protein as well as the loading efficiency of the GppNHp-loaded Rab samples (150).

6.4. Concluding remarks

Given the important role of Rab46 GTPase in regulation of T-cell signalling and differential WPB exocytosis, this thesis aimed to gain better understanding of its structure and dynamic regulation. Overall, the results presented in this thesis provide initial insight into some of the structural properties of Rab46 and offer a solid starting point for its future structural characterisation efforts.

An *E. coli*-based system for overexpression of recombinant EF-hand, Coiled-coil and Rab domains of Rab46 was successfully established. The investigation of EF-hand domain properties using high-resolution solution NMR complemented by additional biophysical techniques including CD, MS and ITC, provided novel insights into its structural characteristics. Importantly, the triple resonance backbone assignment obtained for both apo and Ca^{2+}-saturated forms of EF-hand domain lays the foundation for further investigation of the Ca^{2+}-regulated functions of Rab46.

While the ability to characterise Rab domain using NMR-based approaches was hindered by issues around protein solubility, the protein was demonstrated to possess its expected biological activity by MANT-GDP-based nucleotide exchange experiment. In that regard, this study provides preliminary evidence for an atypical GTPase profile of Rab46 as implied by high intrinsic nucleotide exchange rate of its Rab domain, however this finding would have to be validated by additional experiments.

Unfortunately, lack of success with production of the full length protein hindered attempts at characterisation of the agonist-induced intradomain conformational changes, long-range interdomain interactions, and global structural rearrangements in Rab46 GTPase.

Furthermore, due to the time constraints created by the outbreak of COVID-19 pandemic as well as -80ºC freezer fault in August 2021 leading to losses in protein samples, studies to establish feasibility of NMR of Rab46 in endothelial cells were not undertaken.

Future work could therefore focus in the first place on optimisation of the constructs for expression of full length Rab46 and Rab domain to find conditions yielding protein samples suitable for downstream structural biology applications. Methods for introducing isotopically-labelled Rab46 domains into ECs would then have to be developed with the ultimate aim of employing in-cell NMR approaches to observe the dynamic interactions of Rab46 in a physiological environment.

Appendix

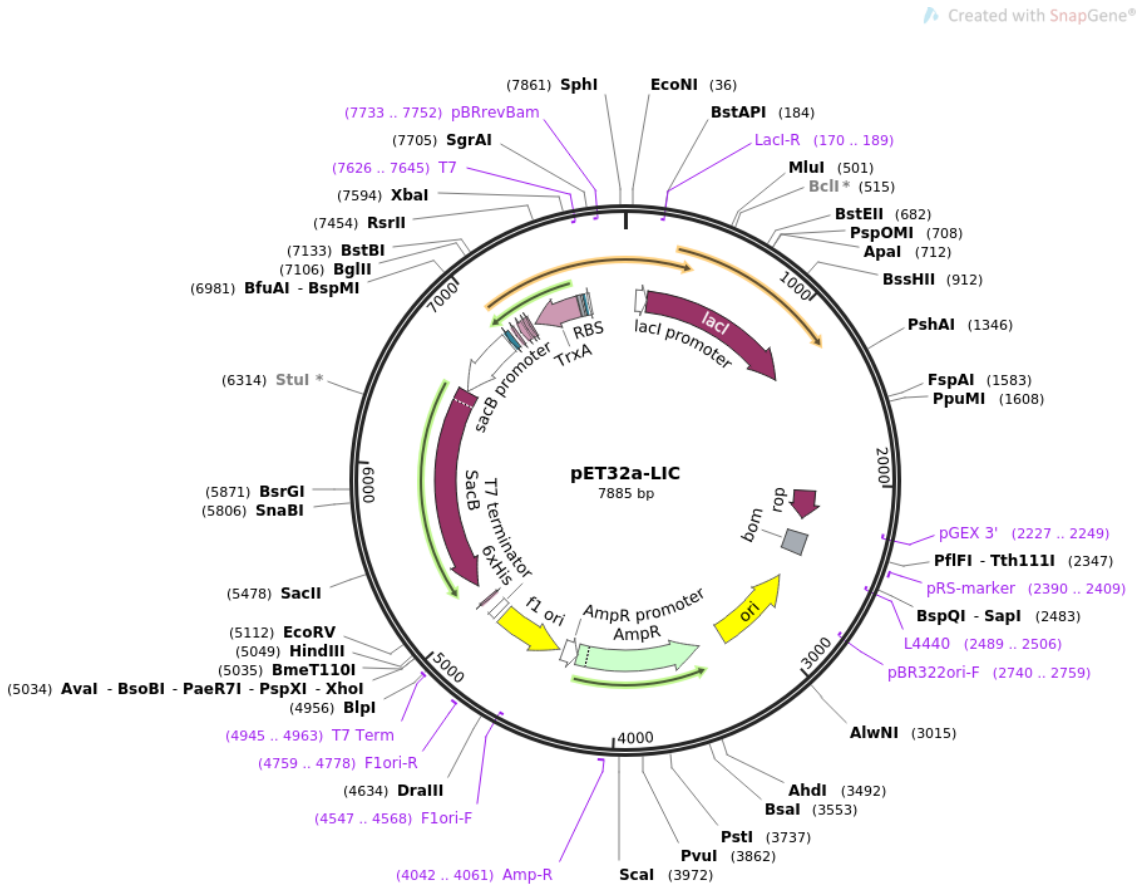

Figure 55 Vector map of pET32a-LIC

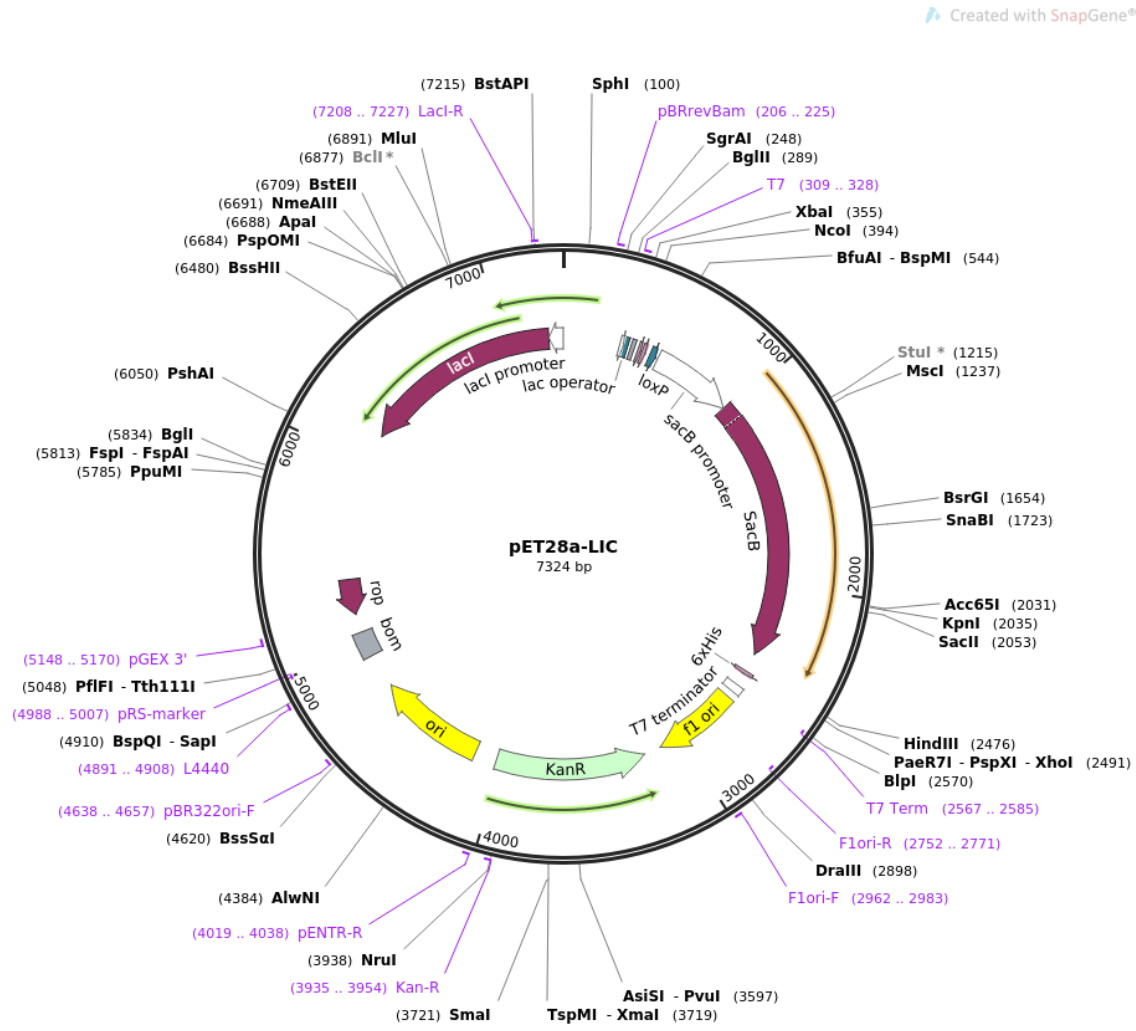

Figure 56 Vector map of pET28a-LIC

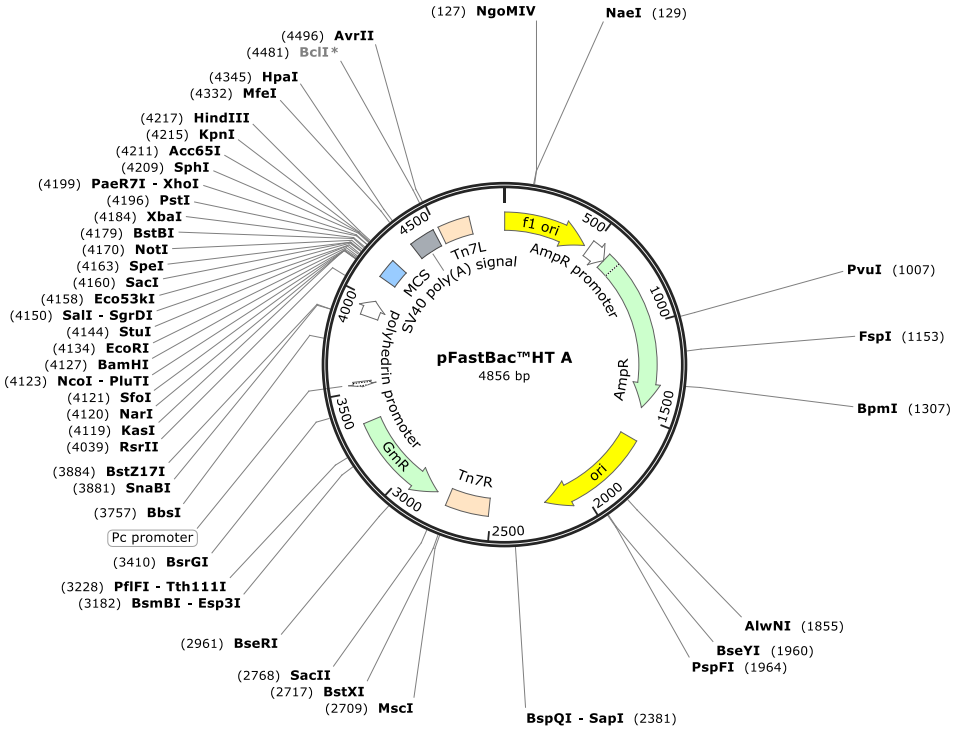

Figure 57 Vector map of pFastBac HT A

Table 15 PCR and DNA sequencing primers

Primer name	Sequence (5’-3’)
T7	TAATACGACTCACTATAGGG
T7 Term	GCTAGTTATTGCTCAGCGG
pFastBacF	GGATTATTCATACCGTCCCA-
pFastBacR	CAAATGTGGTATGGCTGATT
Rab46 Forward	ATGGCTGCCCCTGACG
GP64 Forward	CGGCGTGAGTATGATTCTCAAA
GP64 Reverse	ATGAGCAGACACGCAGCTTTT
pUC/M13 Forward	CCCAGTCACGACGTTGTAAAACG
pUC/M13 Reverse	AGCGGATAACAATTTCACACAGG

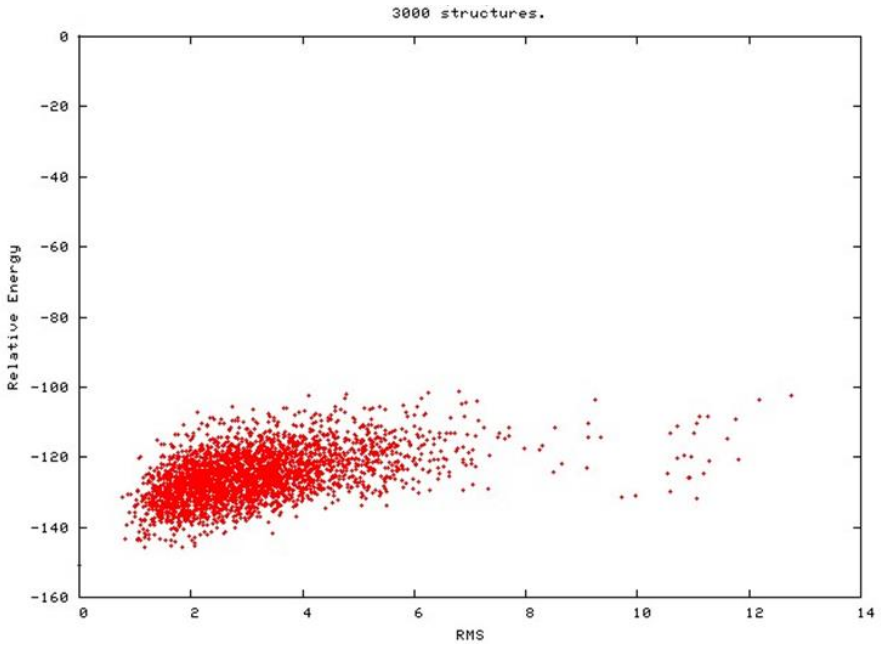

Figure 58 Convergence results for CS-Rosetta models of apo EF-hand. The models generated using CS-Rosetta converged, with ten lowest-energy models clustering within 2.0 Å of the lowest-energy structure.

Table 16 Relative distance and peak intensity between the peaks doubled in the ^{1}H-^{15}N HSQC spectrum of Ca^{2+}-bound EF-hand domain

Residue	Intensity ratio (major: minor)	Weighted distance (ppm)
19Ala	1.13	0.801428561
20Glu	1.77	1.201518874
21Gly	2.92	0.436982993
23Gly	0.79	1.568825961
49Asp	1.94	0.981835803
50Ala	4.07	0.127453052
51Leu	5.12	0.158136932
52Asp	2.01	0.053007174
53Ala	2.42	0.154452524
54Asp	2.03	0.066211625
57Gly	2.39	0.043098158
58Tyr	1.42	0.433948387
59Leu	2.16	0.187485055
60Thr	3.03	0.066927793
62Gln	1.73	0.276540765
63Glu	3.74	0.097159083
64Phe	6.17	0.100555902
65Thr	4.68	0.328689854

Table 17 Intensities of peaks doubled in the ^{1}H-^{15}N HSQC spectrum of Ca^{2+}-bound EF-hand domain recorded at different pH values

Residue	Intensity at pH 6.0		Intensity at pH 7.0		Ratio (pH 6.0 : pH 7.0)	
	Major	Minor	Major	Minor	Major	Minor
19Ala	14892400	23919500	23426400	8341200	0.63571	2.867633
20Glu	34980500	25647600	55318400	11770400	0.632348	2.178991
21Gly	30099100	22833800	51885500	6265820	0.580106	3.644184
23Gly	4230670	13772300	7914310	4096760	0.53456	3.361754
49Asp	13259800	7783170	19299000	2332050	0.687072	3.33748
50Ala	47104400	11397200	66361200	4288180	0.709818	2.657818
51Leu	35991800	5341900	51765700	2205890	0.695283	2.421653
52Asp	29551700	12174000	34924700	4433050	0.846155	2.746191
53Ala	33849200	12451700	42818800	6934480	0.790522	1.795621
54Asp	33083200	13977000	42791300	6084200	0.773129	2.297262
57Gly	21207100	7922440	33609700	6583850	0.630982	1.203314
58Tyr	52240300	13908000	26771400	5516180	1.951347	2.52131
59Leu	10413300	4624230	12881800	3098940	0.808373	1.492197
60Thr	23614700	6806520	26979600	2958060	0.87528	2.301008
62Gln	49313800	11944000	27100900	5305000	1.819637	2.251461
63Glu	54043400	9097380	53807000	5163980	1.004393	1.761699
64Phe	31413800	483432	42045100	2451130	0.747145	0.197228
65Thr	22078900	1156160	29978400	912605	0.736494	1.266879

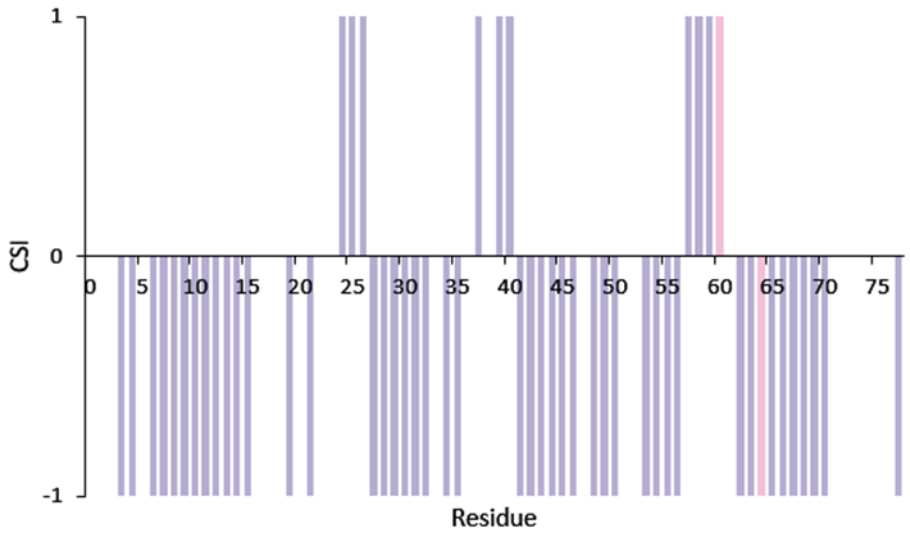

Figure 59 Comparison of CSI between the major (pink) and minor (blue) forms of Ca^{2+}-saturated EF-hand domain

References

1. Bouïs D, Hospers GAP, Meijer C, Molema G, Mulder NH. Endothelium in vitro: A review of human vascular endothelial cell lines for blood vessel-related research. Angiogenesis. 2001;4:91–102.

2. Torisu T, Torisu K, Lee IH, Liu J, Malide D, Combs CA, et al. Autophagy regulates endothelial cell processing, maturation and secretion of von Willebrand factor. Nat Med. 2013;19(10):1281–7.

3. Deanfield JE, Halcox JP, Rabelink TJ. Endothelial function and dysfunction: Testing and clinical relevance. Vol. 115, Circulation. 2007. p. 1285–95.

4. Xu S, Ilyas I, Little PJ, Li H, Kamato D, Zheng X, et al. Endothelial Dysfunction in Atherosclerotic Cardiovascular Diseases and Beyond: From Mechanism to Pharmacotherapies. Pharmacol Rev. 2021;73(3):924–67.

5. Sun HJ, Wu ZY, Nie XW, Bian JS. Role of Endothelial Dysfunction in Cardiovascular Diseases: The Link Between Inflammation and Hydrogen Sulfide. Front Pharmacol. 2020;10.

6. Hadi HAR, Carr CS, Al Suwaidi J. Endothelial Dysfunction: Cardiovascular Risk Factors, Therapy, and Outcome. Vasc Health Risk Manag. 2005;1(3):183.

7. Castellon X, Bogdanova V. Chronic Inflammatory Diseases and Endothelial Dysfunction. Aging Dis. 2016;7(1):81.

8. Prasad M, Leon M, Lerman LO, Lerman A. Viral Endothelial Dysfunction: A Unifying Mechanism for COVID-19. Vol. 96, Mayo Clinic Proceedings. Elsevier Ltd; 2021. p. 3099–108.

9. Amiya E. Interaction of hyperlipidemia and reactive oxygen species: Insights from the lipid-raft platform. World J Cardiol. 2016;8(12):689.

10. Janaszak-Jasiecka A, Płoska A, Wierońska JM, Dobrucki LW, Kalinowski L. Endothelial dysfunction due to eNOS uncoupling: molecular mechanisms as potential therapeutic targets. Cell Mol Biol Lett. 2023;28(1).

11. Drożdż D, Drożdż M, Wójcik M. Endothelial dysfunction as a factor leading to arterial hypertension. Pediatric Nephrology. 2023;38(9):2973–85.

12. Förstermann U, Sessa WC. Nitric oxide synthases: Regulation and function. Vol. 33, European Heart Journal. 2012.

13. Crabtree MJ, Smith CL, Lam G, Goligorsky MS, Gross SS. Ratio of 5,6,7,8-tetrahydrobiopterin to 7,8-dihydrobiopterin in endothelial cells determines glucose-elicited changes in NO vs. superoxide production by eNOS. American Journal of Physiology-Heart and Circulatory Physiology. 2008;294(4):H1530–40.

14. Bowers MC, Hargrove LA, Kelly KA, Wu G, Meininger CJ. Tetrahydrobiopterin attenuates superoxide-induced reduction in nitric oxide. Vol. 3, Frontiers in Bioscience. 2011.

15. Sugiyama T, Levy BD, Michel T. Tetrahydrobiopterin Recycling, a Key Determinant of Endothelial Nitric-oxide Synthase-dependent Signaling Pathways in Cultured Vascular Endothelial Cells. Journal of Biological Chemistry. 2009;284(19):12691–700.

16. Förstermann U, Xia N, Li H. Roles of vascular oxidative stress and nitric oxide in the pathogenesis of atherosclerosis. Vol. 120, Circulation Research. Lippincott Williams and Wilkins; 2017. p. 713–35.

17. Matsushita K, Morrell CN, Cambien B, Yang SX, Yamakuchi M, Bao C, et al. Nitric oxide regulates exocytosis by S-nitrosylation of N-ethylmaleimide-sensitive factor. Cell. 2003;115(2):139–50.

18. Roberts W, Michno A, Aburima A, Naseem KM. Nitric oxide inhibits von Willebrand factor-mediated platelet adhesion and spreading through regulation of integrin alpha(IIb)beta(3) and myosin light chain. J Thromb Haemost. 2009;7(12):2106–15.

19. Smith NL, Chen MH, Dehghan A, Strachan DP, Basu S, Soranzo N, et al. Novel associations of multiple genetic loci with plasma levels of factor VII, factor VIII, and von willebrand factor: The charge (cohorts for heart and aging research in genome epidemiology) consortium. Circulation. 2010;121(12):1382–92.

20. Mussbacher M, Schossleitner K, Kral-Pointner JB, Salzmann M, Schrammel A, Schmid JA. More than Just a Monolayer: the Multifaceted Role of Endothelial Cells in the Pathophysiology of Atherosclerosis. Vol. 24, Current Atherosclerosis Reports. Springer; 2022. p. 483–92.

21. Piga R, Naito Y, Kokura S, Handa O, Yoshikawa T. Short-term high glucose exposure induces monocyte-endothelial cells adhesion and transmigration by

increasing VCAM-1 and MCP-1 expression in human aortic endothelial cells. Atherosclerosis. 2007;193(2):328–34.

22. Bierhaus A, Schiekofer S, Schwaninger M, Andrassy M, Humpert PM, Chen J, et al. Diabetes-Associated Sustained Activation of the Transcription Factor Nuclear Factor-κB. Diabetes. 2001;50(12):2792–808.

23. Mussbacher M, Salzmann M, Brostjan C, Hoesel B, Schoergenhofer C, Datler H, et al. Cell Type-Specific Roles of NF-κB Linking Inflammation and Thrombosis. Front Immunol. 2019;10.

24. Tabit CE, Chung WB, Hamburg NM, Vita JA. Endothelial dysfunction in diabetes mellitus: Molecular mechanisms and clinical implications. Vol. 11, Reviews in Endocrine and Metabolic Disorders. 2010. p. 61–74.

25. Del Turco S, Basta G, Lazzerini G, Chancharme L, Lerond L, De Caterina R. Involvement of the TP receptor in TNF-α-induced endothelial tissue factor expression. Vascul Pharmacol. 2014;62(2):49–56.

26. Szotowski B, Antoniak S, Poller W, Schultheiss HP, Rauch U. Procoagulant soluble tissue factor is released from endothelial cells in response to inflammatory cytokines. Circ Res. 2005;96(12):1233–9.

27. Yan G, You B, Chen SP, Liao JK, Sun J. Tumor necrosis factor-α downregulates endothelial nitric oxide synthase mRNA stability via translation elongation factor 1-α 1. Circ Res. 2008;103(6):591–7.

28. Theofilis P, Sagris M, Oikonomou E, Antonopoulos AS, Siasos G, Tsioufis C, et al. Inflammatory mechanisms contributing to endothelial dysfunction. Biomedicines. 2021;9(7).

29. Schillemans M, Karampini E, Kat M, Bierings R. Exocytosis of Weibel–Palade bodies: how to unpack a vascular emergency kit. Vol. 17, Journal of Thrombosis and Haemostasis. Blackwell Publishing Ltd; 2019. p. 6–18.

30. Metcalf DJ, Nightingale TD, Zenner HL, Lui-Roberts WW, Cutler DF. Formation and function of Weibel-Palade bodies. J Cell Sci. 2008;121(Pt 1):19–27.

31. van Mourik JA, de Wit TR, Voorberg J. Biogenesis and exocytosis of Weibel-Palade bodies. Histochem Cell Biol. 2002;117(2):113–22.

32. Nightingale TD, McCormack JJ, Grimes W, Robinson C, Lopes da Silva M, White IJ, et al. Tuning the endothelial response: differential release of exocytic cargos from Weibel-Palade bodies. Journal of Thrombosis and Haemostasis. 2018;16(9):1873–86.

33. Miteva KT, Pedicini L, Wilson LA, Jayasinghe I, Slip RG, Marszalek K, et al. Rab46 integrates Ca2+ and histamine signaling to regulate selective cargo release from Weibel-Palade bodies. Journal of Cell Biology. 2019;218(7):2232–46.

34. Rondaij MG, Bierings R, Kragt A, Van Mourik JA, Voorberg J. Dynamics and plasticity of Weibel-Palade bodies in endothelial cells. Arterioscler Thromb Vasc Biol. 2006;26(5):1002–7.

35. Fiedler U, Scharpfenecker M, Koidl S, Hegen A, Grunow V, Schmidt JM, et al. The Tie-2 ligand Angiopoietin-2 is stored in and rapidly released upon stimulation from endothelial cell Weibel-Palade bodies. Blood. 2004;103(11):4150–6.

36. Galbusera M, Zoja C, Donadelli R, Paris S, Morigi M, Benigni A, et al. Fluid Shear Stress Modulates von Willebrand Factor Release From Human Vascular Endothelium. Blood. 1997;90(4):1558–64.

37. Lowenstein CJ, Morrell CN, Yamakuchi M. Regulation of Weibel-Palade body exocytosis. Trends Cardiovasc Med. 2005;15(8):302–8.

38. Pinsky DJ, Naka Y, Liao H, Oz MC, Wagner DD, Mayadas TN, et al. Hypoxia-induced exocytosis of endothelial cell Weibel-Palade bodies. A mechanism for rapid neutrophil recruitment after cardiac preservation. Journal of Clinical Investigation. 1996;97(2):493.

39. Zografou S, Basagiannis D, Papafotika A, Shirakawa R, Horiuchi H, Auerbach D, et al. A complete rab screening reveals novel insights in Weibel-Palade body exocytosis. J Cell Sci. 2012;125(20):4780–90.

40. Hutagalung AH, Novick PJ. Role of Rab GTPases in Membrane Traffic and Cell Physiology. Physiol Rev. 2011;91(1):119–49.

41. Pylypenko O, Hammich H, Yu IM, Houdusse A. Rab GTPases and their interacting protein partners: Structural insights into Rab functional diversity. Small GTPases. 2018;9(1–2):22–48.

42. Shintani M, Tada M, Kobayashi T, Kajiho H, Kontani K, Katada T. Characterization of Rab45/RASEF containing EF-hand domain and a coiled-coil motif as a self-associating GTPase. Biochem Biophys Res Commun. 2007;357(3):661–7.

43. Wilson LA, McKeown L, Tumova S, Li J, Beech DJ. Expression of a long variant of CRACR2A that belongs to the Rab GTPase protein family in endothelial cells. Biochem Biophys Res Commun. 2015;456(1):398–402.

44. Yamaguchi Y, Sakai E, Okamoto K, Kajiya H, Okabe K, Naito M, et al. Rab44, a novel large Rab GTPase, negatively regulates osteoclast differentiation by modulating intracellular calcium levels followed by NFATc1 activation. Cellular and Molecular Life Sciences. 2018;75(1):33–48.

45. Zhen Y, Stenmark H. Cellular functions of Rab GTPases at a glance. J Cell Sci. 2015;128(17):3171–6.

46. Sanghavi HM, Rashmi R, Dasgupta A, Majumdar S. G-domain prediction across the diversity of G protein families. bioRxiv. 2019;

47. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

48. Yang M, Derbyshire MK, Yamashita RA, Marchler-Bauer A. NCBI's Conserved Domain Database and Tools for Protein Domain Analysis. Curr Protoc Bioinformatics. 2020;69(1):e90.

49. Dumas JJ, Zhu Z, Connolly JL, Lambright DG. Structural basis of activation and GTP hydrolysis in Rab proteins. Structure. 1999 Apr;7(4):413-s2.

50. Lee SH, Baek K, Dominguez R. Large nucleotide-dependent conformational change in Rab28. FEBS Lett. 2008;582(29):4107–11.

51. Pereira-Leal JB, Seabra MC. The mammalian Rab family of small GTPases: definition of family and subfamily sequence motifs suggests a mechanism for functional specificity in the Ras superfamily. J Mol Biol. 2000;301(4):1077–87.

52. Wang S, Hu C, Wu F, He S. Rab25 GTPase: Functional roles in cancer. Vol. 8, Oncotarget. 2017.

53. Pereira-Leal JB, Hume AN, Seabra MC. Prenylation of Rab GTPases: Molecular mechanisms and involvement in genetic disease. Vol. 498, FEBS Letters. 2001. p. 197–200.

54. Vázquez-Martínez R, xMalagón R. Rab proteins and the secretory pathway: The case of Rab18 in neuroendocrine cells. Front Endocrinol (Lausanne). 2010;2(JAN):1–9.

55. Corbeel L, Freson K. Rab proteins and Rab-associated proteins: Major actors in the mechanism of protein-trafficking disorders. Vol. 167, European Journal of Pediatrics. 2008. p. 723–9.

56. Blümer J, Rey J, Dehmelt L, Maze T, Wu YW, Bastiaens P, et al. RabGEFs are a major determinant for specific Rab membrane targeting. Journal of Cell Biology. 2013;200(3):287–300.

57. Li F, Yi L, Zhao L, Itzen A, Goody RS, Wu YW. The role of the hypervariable C-terminal domain in Rab GTPases membrane targeting. Proc Natl Acad Sci U S A. 2014;111(7):2572–7.

58. Bos JL, Rehmann H, Wittinghofer A. GEFs and GAPs: Critical Elements in the Control of Small G Proteins. Vol. 129, Cell. Elsevier B.V.; 2007. p. 865–77.

59. Génier S, Létourneau D, Gauthier E, Picard S, Boisvert M, Parent JL, et al. In-depth NMR characterization of Rab4a structure, nucleotide exchange and hydrolysis kinetics reveals an atypical GTPase profile. J Struct Biol. 2020;212(1).

60. Biesemann A, Gorontzi A, Barr F, Gerke V. Rab35 protein regulates evoked exocytosis of endothelial Weibel–Palade bodies. Journal of Biological Chemistry. 2017;292(28):11631–40.

61. Nightingale TD, Pattni K, Hume AN, Seabra MC, Cutler DF. Rab27a and MyRIP regulate the amount and multimeric state of VWF released from endothelial cells. Blood. 2009;113(20):5010–8.

62. Chehab T, Santos NC, Holthenrich A, Koerdt SN, Disse J, Schuberth C, et al. A novel Munc13-4/S100A10/annexin A2 complex promotes Weibel–Palade body exocytosis in endothelial cells. Mol Biol Cell. 2017 Jun 15;28(12):1688–700.

63. Srikanth S, Kim K Do, Gao Y, Woo JS, Ghosh S, Calmettes G, et al. A large Rab GTPase encoded by CRACR2A is a component of subsynaptic vesicles that transmit T cell activation signals. Sci Signal. 2016;9(420).

64. Srikanth S, Jung HJ, Kim K Do, Souda P, Whitelegge J, Gwack Y. A novel EF-hand protein, CRACR2A, is a cytosolic Ca2+ sensor that stabilizes CRAC channels in T cells. Nat Cell Biol. 2010;12(5):436–46.

65. Nakamura S, Takemura T, Tan L, Nagata Y, Yokota D, Hirano I, et al. Small GTPase RAB45-mediated p38 activation in apoptosis of chronic myeloid leukemia progenitor cells. Carcinogenesis. 2011;32(12):1758–72.

66. Alcantara J, Stix R, Huang K, Connor A, East R, Jaramillo-Martinez V, et al. An Unbound Proline-Rich Signaling Peptide Frequently Samples Cis Conformations in Gaussian Accelerated Molecular Dynamics Simulations. Front Mol Biosci. 2021;8.

67. Grabarek Z. Structural basis for diversity of the EF-hand calcium-binding proteins. J Mol Biol. 2006;359(3):509–25.

68. Lee IG, Cason SE, Alqassim SS, Holzbaur ELF, Dominguez R. A tunable LIC1-adaptor interaction modulates dynein activity in a cargo-specific manner. Nature Communications 2020 11:1. 2020;11(1):1–13.

69. Aitio H, Laakso T, Pihlajamaa T, Torkkeli M, Kilpeläinen I, Drakenberg T, et al. Characterization of apo and partially saturated states of calerythrin, an EF-hand protein from S. erythraea : A molten globule when deprived of Ca 2+ . Protein Science. 2001;10(1):74–82.

70. Nelson MR, Thulin E, Fagan PA, Forsén S, Chazin WJ. The EF-hand domain: A globally cooperative structural unit. Protein Science. 2002;11(2):198–205.

71. Chazin WJ. Relating form and function of EF-hand calcium binding proteins. Acc Chem Res. 2011;44(3):171–9.

72. Skelton NJ, Kördel J, Akke M, Forsén S, Chazin WJ. Signal transduction versus buffering activity in Ca(2+)-binding proteins. Nat Struct Biol. 1994;1(4):239–45.

73. Liu J, Rost B. Comparing function and structure between entire proteomes. Protein Science. 2001;10(10):1970–9.

74. Wang Y, Huynh W, Skokan TD, Lu W, Weiss A, Vale RD. CRACR2a is a calcium-activated dynein adaptor protein that regulates endocytic traffic. Journal of Cell Biology. 2019;218(5):1619–33.

75. Urnavicius L, Lau CK, Elshenawy MM, Morales-Rios E, Motz C, Yildiz A, et al. Cryo-EM shows how dynactin recruits two dyneins for faster movement. Nature. 2018;554(7691):202–6.

76. Urnavicius L, Zhang K, Diamant AG, Motz C, Schlager MA, Yu M, et al. The structure of the dynactin complex and its interaction with dynein. Science. 2015;347(6229):1441–6.

77. Pedicini L, Wiktor SD, Simmons KJ, Money A, McKeown L. Affinity-based proteomics reveals novel binding partners for Rab46 in endothelial cells. Scientific Reports 2021 11:1. 2021;11(1):1–16.

78. The PyMOL Molecular Graphics System, Version 2.3.2. Schrödinger LLC;

79. Chalasani N, Guo X, Loomba R, Goodarzi MO, Haritunians T, Kwon S, et al. Genome-wide association study identifies variants associated with histologic features of nonalcoholic fatty liver disease. Gastroenterology. 2010;139(5).

80. Barozzi C, Galletti M, Tomasi L, De Fanti S, Palazzini M, Manes A, et al. A Combined Targeted and Whole Exome Sequencing Approach Identified Novel Candidate Genes Involved in Heritable Pulmonary Arterial Hypertension. Sci Rep. 2019;9(1).

81. Mitsunaga S, Hosomichi K, Okudaira Y, Nakaoka H, Kunii N, Suzuki Y, et al. Exome sequencing identifies novel rheumatoid arthritis-susceptible variants in the BTNL2. J Hum Genet. 2013;58(4):210–5.

82. Bevilacqua L, Navarra CO, Pirastu N, Lenarda R Di, Gasparini P, Robino A. A genome-wide association study identifies an association between variants in EFCAB4B gene and periodontal disease in an Italian isolated population. J Periodontal Res. 2018;53(6):992–8.

83. Wu B, Rice L, Shrimpton J, Lawless D, Walker K, Carter C, et al. Biallelic mutations in calcium release activated channel regulator 2A (CRACR2A) cause a primary immunodeficiency disorder. Elife. 2021;10:72559.

84. Wang D, Wiktor SD, Cheng CW, Simmons KJ, Money A, Pedicini L, et al. EFCAB4B (CRACR2A) genetic variants associated with COVID-19 fatality. medRxiv. 2022;

85. Philippe A, Chocron R, Gendron N, Bory O, Beauvais A, Peron N, et al. Circulating Von Willebrand factor and high molecular weight multimers as markers of endothelial injury predict COVID-19 in-hospital mortality. Angiogenesis. 2021;24(3):505–17.

86. Villa E, Critelli R, Lasagni S, Melegari A, Curatolo A, Celsa C, et al. Dynamic angiopoietin-2 assessment predicts survival and chronic course in hospitalized patients with COVID-19. Blood Adv. 2021;5(3):662–73.

87. Fenyves BG, Mehta A, Kays KR, Beakes C, Margolin J, Goldberg MB, et al. Plasma P-selectin is an early marker of thromboembolism in COVID-19. Am J Hematol. 2021;96(12):10.

88. Chan JMS, Monaco C, Wylezinska-Arridge M, Tremoleda JL, Cole JE, Goddard M, et al. Imaging vulnerable plaques by targeting inflammation in atherosclerosis using fluorescent-labeled dual-ligand microparticles of iron oxide and magnetic resonance imaging. J Vasc Surg. 2018;67(5):1571-1583.e3.

89. Theelen TL, Lappalainen JP, Sluimer JC, Gurzeler E, Cleutjens JP, Gijbels MJ, et al. Angiopoietin-2 blocking antibodies reduce early atherosclerotic plaque development in mice. Atherosclerosis. 2015;241(2):297–304.

90. Alderson TR, Kay LE. NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. Vol. 184, Cell. Elsevier B.V.; 2021. p. 577–95.

91. Ying J, Delaglio F, Torchia DA, Bax A. Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. J Biomol NMR. 2017;68(2):101–18.

92. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995;6(3):277–93.

93. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins. 2005;59(4):687–96.

94. Wang B, Wang Y, Wishart DS. A probabilistic approach for validating protein NMR chemical shift assignments. J Biomol NMR. 2010;47(2):85–99.

95. Wishart DS, Sykes BD. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J Biomol NMR. 1994;4(2):171–80.

96. Lange OF, Rossi P, Sgourakis NG, Song Y, Lee HW, Aramini JM, et al. Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. Proc Natl Acad Sci U S A. 2012;109(27):10873–8.

97. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—A visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–12.

98. PyMOLWiki. Color h. [Internet]. [cited 2023 Dec 20]. Available from: https://pymolwiki.org/index.php/Color_h

99. Williamson MP. Using chemical shift perturbation to characterise ligand binding. Vol. 73, Progress in Nuclear Magnetic Resonance Spectroscopy. Elsevier B.V.; 2013. p. 1–16.

100. Fitzkee NC, Bax A. Facile measurement of 1H-15N residual dipolar couplings in larger perdeuterated proteins. J Biomol NMR [Internet]. 2010;48(2):65. Available from: http://spin.niddk.nih.gov/bax/pp

101. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. Journal of Magnetic Resonance. 2003;160(1):65–73.

102. Anthis NJ, Clore GM. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. Protein Sci. 2013;22(6):851–8.

103. Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. Nature Protocols 2007 1:6. 2007;1(6):2876–90.

104. Miles AJ, Ramalli SG, Wallace BA. DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. Protein Sci. 2022;31(1):37–46.

105. Compton LA, Johnson WC. Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. Anal Biochem. 1986;155(1):155–67.

106. Kanie T, Jackson P. Guanine Nucleotide Exchange Assay Using Fluorescent MANT-GDP. Bio Protoc. 2018;8(7).

107. GraphPad Prism version 9. GraphPad Software, Boston, Massachusetts USA;

108. Batey S, Nickson AA, Clarke J. Studying the folding of multidomain proteins. Vol. 2, HFSP Journal. 2008. p. 365–77.

109. Chen J, Song JL, Zhang S, Wang Y, Cui DF, Wangt CC. Chaperone Activity of DsbC. Journal of Biological Chemistry. 1999;274(28):19601–5.

110. Kelly S, Price N. The use of circular dichroism in the investigation of protein structure and function. Curr Protein Pept Sci. 2000;1(4):349–84.

111. Buchan DWA, Jones DT. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Res. 2019;47(W1):W402–7.

112. Kleckner IR, Foster MP. An introduction to NMR-based approaches for measuring protein dynamics. Biochim Biophys Acta. 2011;1814(8):942–68.

113. Raman S, Lange OF, Rossi P, Tyka M, Wang X, Aramini J, et al. NMR structure determination for larger proteins using backbone-only data. Science. 2010;327(5968):1014–8.

114. Chiliveri SC, Robertson AJ, Shen Y, Torchia DA, Bax A. Advances in NMR Spectroscopy of Weakly Aligned Biomolecular Systems. Chem Rev. 2022;122(10):9307–30.

115. Bax A. Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Science. 2003;12(1):1–16.

116. Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature. 2015;519(7541):106–9.

117. Baskaran K, Wilburn CW, Wedell JR, Koharudin LMI, Ulrich EL, Schuyler AD, et al. Anomalous amide proton chemical shifts as signatures of hydrogen bonding to aromatic sidechains. Magnetic Resonance. 2021;2(2):765–75.

118. Kawale AA, Burmann BM. Chapter 5 - Advanced NMR spectroscopy methods to study protein structure and dynamics. In: Saudagar P, Tripathi T, editors. Advanced Spectroscopic Methods to Study Biomolecular Structure and Dynamics. Academic Press; 2023. p. 125–52.

119. Sarkar P, Reichman C, Saleh T, Birge RB, Kalodimos CG. Proline cis-trans isomerization controls autoinhibition of a signaling protein. Mol Cell. 2007;25(3):413–26.

120. Alderson TR, Benesch JLP, Baldwin AJ. Proline isomerization in the C-terminal region of HSP27. Cell Stress Chaperones. 2017;22(4):639.

121. Sebák F, Szolomájer J, Papp N, Tóth GK, Bodor A. Proline cis/trans Isomerization in Intrinsically Disordered Proteins and Peptides. Front Biosci (Landmark Ed). 2023;28(6).

122. Kukić, Farrell D, Søndergaard CR, Bjarnadottir U, Bradley J, Pollastri G, et al. Improving the analysis of NMR spectra tracking pH-induced conformational changes: Removing artefacts of the electric field on the NMR chemical shift Predrag. Proteins: Structure, Function and Bioinformatics. 2010;78(4):971–84.

123. Ben-Nissan G, Sharon M. The application of ion-mobility mass spectrometry for structure/function investigation of protein complexes. Curr Opin Chem Biol. 2018;42:25–33.

124. Lanucara F, Holman SW, Gray CJ, Eyers CE. The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. Nature Chemistry 2014 6:4. 2014;6(4):281–94.

125. Reece KL, Moss RL. Removal of contaminating calcium from buffer solutions used in calcium binding assays. Anal Biochem. 2007;365(2):274–6.

126. Bourgeois B, Worman HJ, Zinn-Justin S, Bourgeois B, Worman HJ, Zinn-Justin S. NMR Spectroscopy as a Tool to Provide Mechanistic Clues About Protein Function and Disease Pathogenesis. Magnetic Resonance Spectroscopy. 2012;

127. Mónico A, Martínez-Senra E, Cañada FJ, Zorrilla S, Pérez-Sala D. Drawbacks of Dialysis Procedures for Removal of EDTA. PLoS One. 2017;12(1).

128. Vorherr T, James P, Krebs J, Enyedi A, McCormick DJ, Penniston JT, et al. Interaction of calmodulin with the calmodulin binding domain of the plasma membrane Ca2+ pump. Biochemistry. 1990;29(2):355–65.

129. Williamson MP. Using chemical shift perturbation to characterise ligand binding. Prog Nucl Magn Reson Spectrosc. 2013;73:1–16.

130. Kawale AA, Burmann BM. Characterization of backbone dynamics using solution NMR spectroscopy to discern the functional plasticity of structurally analogous proteins. STAR Protoc. 2021;2(4).

131. Wittinghofer A, Vetter IR. Structure-Function Relationships of the G Domain, a Canonical Switch Motif. Annu Rev Biochem. 2011;80:943–71.

132. Micsonai A, Bulyáki É, Kardos J. BeStSel: From Secondary Structure Analysis to Protein Fold Prediction by Circular Dichroism Spectroscopy. Methods Mol Biol. 2021;2199:175–89.

133. Bergbrede T, Pylypenko O, Rak A, Alexandrov K. Structure of the extremely slow GTPase Rab6A in the GTP bound form at 1.8A resolution. J Struct Biol. 2005;152(3):235–8.

134. Kanie T, Abbott KL, Mooney NA, Plowey ED, Demeter J, Jackson PK. The CEP19-RABL2 GTPase Complex Binds IFT-B to Initiate Intraflagellar Transport at the Ciliary Base. Dev Cell. 2017;42(1):22-36.e12.

135. Saul J, Petritis B, Sau S, Rauf F, Gaskin M, Ober-Reynolds B, et al. Development of a full-length human protein production pipeline. Protein Science. 2014;23(8):1123–35.

136. Levy N, Eiler S, Pradeau-Aubreton K, Maillot B, Stricher F, Ruff M. Production of unstable proteins through the formation of stable core complexes. Nat Commun. 2016;7.

137. Franke B, Opitz C, Isogai S, Grahl A, Delgado L, Gossert AD, et al. Production of isotope-labeled proteins in insect cells for NMR. J Biomol NMR. 2018;71(3):173–84.

138. Fattori J, Rodrigues FHS, Pontes JGM, Paula Espíndola A, Tasic L. Monitoring Intermolecular and Intramolecular Interactions by NMR Spectroscopy. Appl NMR Spectrosc. 2015;3:180–266.

139. Grabarek Z. Insights into modulation of calcium signaling by magnesium in calmodulin, troponin C and related EF-hand proteins. Biochim Biophys Acta. 2011;1813(5):913–21.

140. Hikichi K, Ikura M, Minowa O. Hydrogen Bonding in the Carboxyl-Terminal Half-Fragment 78-148 of Calmodulin As Studied by Two-Dimensional Nuclear Magnetic Resonance. Biochemistry. 1985;24(16):4264–9.

141. Ikura M, Minowa O, Yazawa M, Yagi K, Hikichi K. Sequence-specific assignments of downfield-shifted amide proton resonances of calmodulin Use of two-dimensional NMR analysis of its tryptic fragments. FEBS Lett. 1987;219(1):17–21.

142. Ogura K, Okamura H, Katahira M, Katoh E, Inagaki F. Conformational dynamics of yeast calmodulin in the Ca2+-bound state probed using NMR relaxation dispersion. FEBS Lett. 2012;586(16):2548–54.

143. Enomoto M, Nishikawa T, Back SI, Ishiyama N, Zheng L, Stathopulos PB, et al. Coordination of a Single Calcium Ion in the EF-hand Maintains the Off State of the Stromal Interaction Molecule Luminal Domain. J Mol Biol. 2020;432(2):367–83.

144. Miloushev VZ, Bahna F, Ciatto C, Ahlsen G, Honig B, Shapiro L, et al. Dynamic properties of a type II cadherin adhesive domain: implications for the mechanism of strand-swapping of classical cadherins. Structure. 2008;16(8):1195–205.

145. Babini E, Bertini I, Capozzi F, Chirivino E, Luchinat C. A Structural and Dynamic Characterization of the EF-Hand Protein CLSP. Structure. 2006;14(6):1029–38.

146. Drmota Prebil S, Slapšak U, Pavšic M, Ilc G, Puž V, De Almeida Ribeiro E, et al. Structure and calcium-binding studies of calmodulin-like domain of human non-muscle α-Actinin-1. Sci Rep. 2016;6.

147. Zhang Z, Gao R, Hu Q, Peacock H, Matthew Peacock D, Dai S, et al. GTP-state-selective cyclic peptide ligands of K-ras(G12D) block its interaction with Raf. ACS Cent Sci. 2020;6(10):1753–61.

148. Mazhab-Jafari MT, Marshall CB, Smith M, Gasmi-Seabrook GMC, Stambolic V, Rottapel R, et al. Real-time NMR Study of Three Small GTPases Reveals That Fluorescent 2′(3′)-O-(N-Methylanthraniloyl)-tagged Nucleotides Alter Hydrolysis and Exchange Kinetics. Journal of Biological Chemistry. 2010;285(8):5132–6.

149. Zhao Q, Fujimiya R, Kubo S, Marshall CB, Ikura M, Shimada I, et al. Real-Time In-Cell NMR Reveals the Intracellular Modulation of GTP-Bound Levels of RAS. Cell Rep. 2020;32(8).

150. Moghadamchargari Z, Huddleston J, Shirzadeh M, Zheng X, Clemmer DE, M. Raushel F, et al. Intrinsic GTPase Activity of K-RAS Monitored by Native Mass Spectrometry. Biochemistry. 2019;58(31):3396–405.