IMPERIAL COLLEGE LONDON

CRYSTALLOGRAPHIC AND LIGAND BINDING STUDIES ON HUMAN SECRETORY RAB27A

BY

MOSTAFA JAMSHIDIHA

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

MAY, 2017

ABSTRACT

Rab27a is a member of Rab GTPase family of proteins, which plays a key role in cellular transport and docking of secretory vesicles at the plasma membrane. Rab27a is involved in transport/secretion of melanosomes and vesicles containing insulin, histamine, amylase, chemokines, metaloproteases and exosomes. Rab27a has been shown to promote the growth and invasion of multiple cancer types such as breast, lung, pancreatic and hepatocellular carcinoma. Evidence suggesting the significant role of Rab27a in multiple cancer types suggests Rab27a inhibition could potentially impact invasive primary tumors and metastases. The reliance of drug discovery on high-resolution structures prompted us to design suitable Rab27a constructs that crystallize but leave biologically relevant pockets of interest unoccupied. We used Rab27a-Slp2$_a$ complex and Rab27a homologue crystal structures to modify Rab27a into a suitable construct. In the first approach, I fused the Rab-binding domain (RBD) of an effector to Rab27a to promote crystallization. Crystals from this construct diffracted to a resolution better than 2 Å. The main advantage for this approach is leaving the SF4 pocket, which is an important interface for Rab27a-Slp2$_a$ interaction, empty. This pocket is not accessible in crystals obtained from Rab27a-Slp2$_a$ complex. In the second approach, I analyzed the crystal contact residues of Rab3a, a homologue of Rab27a. This information was used to make site-specific mutations based on residues important for crystallization of Rab3a, to assist the crystallization of Rab27a. This approach produced a construct that crystallizes consistently in short period of time, while leaving large surface of the interaction interface empty. Apart from crystallization, we carried out screening campaigns using differential scanning fluorimetry (DSF) and surface plasmon resonance (SPR) against Rab27a to identify ligands that bind to Rab27a-effector interaction interface. In the first approach, a fragment library using differential scanning fluorimetry (DSF) was screened against Rab27a. In the second method, a chemical library enriched with molecules complementary to the SF4 pocket was screened against Rab27a using surface plasmon resonance (SPR). At least one chemical series targeting the SF4 pocket was obtained from screening. Altogether, in this work, I have obtained crystals of Rab27a suitable for ligand interaction studies and ligands targeting Rab27a-effector interaction interface.

ACKNOWLEDGEMENTS

It has been a long journey since I left Iran in 2009 for a PhD in structural biology. There have been many great moments and some low moments. But it seems I've finally reached the end of an era. I would like to thank my supervisors Dr. Ernesto Cota and Dr. Edward Tate for giving me the opportunity to work in their labs. It has given me the chance to grow as a scientist. I would like to thank the magnificent Dr. Inma Perez for training me in the field of X-ray crystallography. I would like to thank Dr. Yingqi Xu for his assistance at the NMR facility. I would like to thank Dr. Marc Morgan for his assistance at the X-ray facility. I would like to thank Professor Steve Mathews for useful discussions on my project. I would like to give special thanks to Professor Randy Read and Dr. James murray for their help in a section of my X-ray data analysis. I would like to thank the groups of Dr. Nathan Brown and Dr. Rob Van Montfort groups for their assistance in our project. I would like to give special thanks to Dr. Martin Rowlands at ICR for his patient help in our project. I would like to thank my friends in the lab Jay, Sabine, Sarah, Masooma, Rhian, Xeni, Will, Max, Jimmy, John, Nessim for their help during my PhD. I also would like the opportunity and thank Professor Hans Vogel, Dr. Dowlatabadi, Dr. Hiroaki Ishida, Dr. Jessica Gifford, Dr. Leonard Nguyen and Dr. Subrata Paul for their help during my Masters at the University of Calgary. Being here at Imperial was not possible without their help in the initial stage of my science career. I would also like to thank Dr. Charlotte Sutherell for her help on the latter parts of my thesis. Living in London is expensive and would not be possible without the generous help of uncle Moez and his lovely wife Sirani, thank you guys. I want to sincerely thank my father and mother in laws, which have supported me in many different ways to help my studies here. I would like to thank my dad, sister and brother Mahdi for their support during studies. I would like to thank my friends Hamid, Esi and Sirvan for their support and companionship in the last 12 years since the undergraduate days. It has been a lot of fun being with you guys. I would like to thank my friend and life partner Fatemeh. She is an amazing person. She is always there for me in the hard times. I am so happy to be with her. It would definitely not be possible to reach anywhere close to where I am, without her unwavering support in the last 8 years. I would like to say a special hello to Ali, who joined us on 4th of January this year. We are so happy to have him besides us. Finally, I want to speak with my dearest mum. It is nearly 10 years since you left us. There isn't a day for me to not think about you. May you rest in peace.

To my dearest Fatemeh and my mother Zahra

Table of Contents

ABSTRACT 2

ACKNOWLEDGEMENTS 3

LIST OF FIGURES 8

LIST OF TABLES 11

LIST OF ABBREVIATIONS 12

DECLARATION OF ORIGINALITY 15

COPYWRITE DECLARATION 16

CHAPTER ONE: INTRODUCTION 17

1.1 Ras superfamily of proteins 17
1.2 Structural elements of Ras superfamily proteins 20
1.3 Conformational switch and dynamics of Ras superfamily proteins 22
1.4 Rab proteins and cellular trafficking 23
1.5 Rab proteins life cycle 26
1.6 Structural characteristics of Rab proteins 27
1.7 Rab proteins and the secretory pathway in cell physiology 30
1.8 Rab27a and the secretory pathway 30
1.9 Rab27a effectors and their role in the secretory pathway 33
1.10 Structural characteristics of Rab27a-effector interactions 37
1.11 Rab27a and its role in cancer 38
1.12 Targeting Ras proteins 40
1.13 Targeting Rab27a 43
1.14 Aims and objectives of project 44

CHAPTER TWO: MATERIALS AND METHODS 45

2.1 Materials 45
2.1.1 Chemicals 45
2.1.2 Media 45
2.1.3 Cell strains 45
2.1.4 Plasmids 46
2.2 Molecular biology 46
2.2.1 Cloning 46
2.2.2 Transformation 46
2.2.3 Agarose gel electrophoresis 46
2.3 Protein expression and purification 47
2.3.1 Rab27a expression and purification 47
2.3.2 Maltose-binding protein (MBP) fused Rab27a expression and purification 48
2.3.3 Rab27a-Slp1 complex expression and purification 48
2.3.4 Slp expression and purification 49
2.4 Selenomethionine labeling of Rab27a 49
2.5 TEV enzyme expression and purification 50
2.6 BirA enzyme expression and purification 50

2.7 Biotinylation of proteins **50**
2.7.1 Biotinylation of Rab27a and Slp 50
2.7.2 Validating biotinylation by western blot 51
2.8 Loading of nucleotides on Rab proteins **51**
2.8.1 Loading of GDP 51
2.8.2 Loading of Guanosine 5′ -[β , γ -imido]triphosphate] (GppNHp) 51
2.8.3 Validating nucleotide loading by NMR 52
2.8.4 Validating nucleotide loading by HPLC 52
2.9 Biophysical assays **52**
2.9.1 Differential scanning fluorimetry (DSF) 52
2.9.2 Surface plasmon resonance (SPR) 54
2.10 NMR spectroscopy **56**
2.10.1 General considerations 56
2.10.2 1D ^{1}H and 2D ^{1}H-^{15}N(or ^{13}C) NMR experiments 56
2.10.3 Ligand observed NMR experiments 56
2.10.4 Measuring K_D by NMR spectroscopy using 2D ^{1}H-^{15}N and ^{1}H-^{13}C NMR 59
2.10.5 Data processing and analysis 60
2.11 Ligand aggregation assay **60**
2.12 X-ray crystallography **61**
2.13 Computational screening of ICR vendor collection **62**

CHAPTER THREE: DESIGN AND CRYSTALLIZATION OF RAB27A CONSTRUCTS 64
3.1 Introduction **64**
3.2 Strategies for Rab27a crystallization **66**
3.3 Design of fusion constructs **67**
3.3.1 Slp2$_a$-Rab27a fusion crystallization 70
3.3.2 Designing new Slp-Rab27a fusion constructs 74
3.4 Designing crystallizable mutant Rab27a constructs **79**
3.4.1 MJ1Rab27a 84
3.4.2 MJ1-1Rab27a 87
3.4.3 MJ4-1Rab27a 89
3.5 Discussion **98**

CHAPTER FOUR: FRAGMENT-BASED LIGAND SCREENING OF RA27A USING DSF 101
4.1 Introduction **101**
4.2 DSF screening **103**
4.3 Analyzing binding using NMR spectroscopy **105**
4.4 Binding site identification by monitoring methionine residues using NMR **110**
4.5 Binding mode characterization using X-ray crystallography **113**
4.6 Discussion **117**

CHAPTER FIVE: DESIGN AND SCREENING OF A FOCUSED LIBRARY AGAINST RA27A USING SPR 121
5.1 Introduction **121**
5.2 Library design **123**

5.3 SPR screen **128**
5.3.1 Evaluating Rab27a activity on the chip 128
5.3.2 Assessing the quality of the SPR run 129
5.3.3 Clean-screen run to identify screen-interfering compounds 129
5.3.4 SPR results 129
5.4 Hit validation **131**
5.4.1 CCT326123 (and family of compounds) 133
5.4.2 CCT332162 136
5.4.3 CCT360813 138
5.4.4 CCT349808 139
5.4.5 CCT270661 140
5.5 Crystallography **141**
5.6 Discussion **141**

CHAPTER SIX: DISCUSSION 144

REFERENCES 154

APPENDIX 173

LIST OF FIGURES

Figure 1.1. Ras protein cycle.. 17
Figure 1.2. Ras downstream signaling pathways. 19
Figure 1.3. Key structural elements of Ras superfamily members. 21
Figure 1.4. Rab function and localization inside the cell. 25
Figure 1.5. Rab life cycle. 27
Figure 1.6. Sequence alignment of members of the Rab family. 28
Figure 1.7. Organization of the switch regions for Rab proteins.. 29
Figure 1.8. Role of Rab27a in melanosome transport. 32
Figure 1.9. Domain organization of Rab27a effectors. 35
Figure 1.10. Rab27a-Slp2$_a$ interaction interface.. 38
Figure 1.11. Direct inhibition of Ras proteins. 41
Figure 3.1. Crystallization of RadA mutant in different crystal forms. 66
Figure 3.2. Rab27a-Slp2a interaction.. 67
Figure 3.3. Interaction of the WF motif with SF4 pocket of Rab27a.. 68
Figure 3.4. Design of the Slp2$_a$-Rab27a fusion construct.. 69
Figure 3.5. Crystallization of Slp2$_s$-Rab27a fusion under different conditions.. 70
Figure 3.6. Slp2$_a$-Rab27a fusion structure analysis. 72
Figure 3.7. Y122 side chain orientation in Slp2$_a$-Rab27a fusion protein. 74
Figure 3.8. Sequence alignment of synaptotagmin-like proteins RBD$_{1\alpha}$. 75
Figure 3.9. Rab27a-Slp1 complex.. 77
Figure 3.10. Crystallization of Slp1-Rab27a fusion.. 78
Figure 3.11. Sequence alignment of multiple members of the Rab family in humans. 79
Figure 3.12. Substitution of the β_2-β_3 loop of Rab27a. 80
Figure 3.13. Designed mutation sites in Rab27a based on Rab3a crystal contacts. 81
Figure 3.14. MJ1Rab27a crystallization. 85
Figure 3.15. SF4 pocket occupancy In MJ1Rab27a. 86
Figure 3.16. MJ1-1Rab27a crystallization 88
Figure 3.17. Self-rotation function stereographic projection at different κ values calculated for dataset obtained from MJ4-1Rab27a crystals. 91
3.4.3.2 Anisotropy correction 92
Figure 3.18. Measuring anisotropy of MJ4-1Rab27a dataset. 92
Figure 3.19. Anisotropic diffraction of MJ4-1Rab27a crystals 93
Figure 3.20. Native Patterson map for MJ4-1Rab27a.. 94
Figure 3.21. Solving the structure of MJ4-1Rab27a using an ensemble structure. .. 95

Figure 3.22 96
Figure 3.23. NCS axes in MJ4-1Rab27a. 97
Figure 3.24. Evolution of mutant Rab27a constructs design. 100
Figure 4.1. Fragment-based ligand discovery of Rab27a. 102
Figure 4.2. DSF screening against Rab27a. 103
Figure 4.3. Hierarchical clustering of fragments hits. 104
Figure 4.4. Evaluating fragment binding to Rab27a using WaterLOGSY experiment.. 107
Figure 4.5. Evaluating fragment binding to Rab27a using STD and CPMG experiments. 108
Figure 4.6. ^{15}N, ^{1}H HSQC spectrum overlay of Rab27a in the presence and absence of fragments. 109
Figure 4.7. Assigning M93 and M185 methyl chemical shifts using mutagenesis .. 111
Figure 4.8. Monitoring fragment binding to Rab27a using ^{13}C, ^{1}H HSQC..... 112
Figure 4.9. Soaking $Slp2_a$-Rab27a fusion crystals with compounds 1, 3, 4, 12 and 15. 113
Figure 4.10. Fragment binding sites in $Slp2_a$-Rab27a crystals. 114
Figure 4.11. Characterization of fragments binding modes using $Slp2_a$-Rab27a fusion crystals.. 115
Figure 4.12. Comparison of ΔSFL-$Slp2_a$-Rab27a and $Slp2_a$-Rab27a crystal structures. 116
Figure 5.1. Screening Rab27a using a library enriched by shape similarity.. 123
Figure 5.2. Query molecule for shape-based screening. 124
Figure 5.3. Shape overlay of compounds to “WF” dipeptide using ROCS..... 126
Figure 5.4. Population distribution of chemical properties of ICR library before and after filtering by shape similarity screening.. 127
Figure 5.5. Demonstrating the binding of Rab27a-gppNHp to $Slp2_a$ RBD by use of SPR..... 128
Figure 5.6. Hits obtained through SPR screening..... 130
Figure 5.7. Compounds selected from SPR titrations for NMR follow up studies... 131
Figure 5.8. Evaluating aggregation of ligands using 1D ^{1}H NMR.. 132
Figure 5.9. Analyzing interaction of CCT326123 with Rab27a. 134
Figure 5.10. Analysis of CC326123 interaction with Rab27a at different concentrations..... 135
Figure 5.11. Naphthalen-1-ylmethyl sulfanyl derivatives obtained by SPR screening. 136
Figure 5.12. Analyzing interaction of CCT332162 with Rab27a. 137

Figure 5.13. Analyzing interaction of CCT360813 with Rab27a. 138
Figure 5.14. Analyzing interaction of CCT349808 with Rab27a. 139
Figure 5.15. Analyzing interaction of CCT270661 with Rab27a. 140
Figure 6.1. Structural overlay of HRas and multiple Rab members bound to GTP 145
Figure 6.2. Tagging lanthanide ions on Rab27a. 151

LIST OF TABLES

Table 1. 1. Rab3 family role in controlled secretion 31
Table 1. 2. Rab27a effector proteins function 34
Table 1. 3. Role of Rab proteins in tumor progression 39
Table 1. 4. Direct inhibition of Ras proteins 42

Table 2. 1. DMSO correction curve sample preparation table 54

Table 3. 1. Unit cell information from diffraction of $Slp2_a$-Rab27a in different conditions 71
Table 3. 2. Data Processing and Refinement Statistics for $Slp2_a$-Rab27a fusion 71
Table 3. 3. Data Processing and Refinement Statistics for Slp1-Rab27a complex 76
Table 3. 4. Unit cell information from diffraction of Slp1-Rab27a in different conditions 78
Table 3. 5. Designing different Rab27a mutant constructs for crystallization 82
Table 3. 6. Data Processing and Refinement Statistics for MJ1Rab27a 84
Table 3. 7. Data Processing Statistics for MJ1-1Rab27a 87
Table 3. 8. Data Processing Statistics for MJ4-1Rab27a 90
Table 3. 9. Estimation of number of molecules in ASU for MJ4-1 using Matthews coefficient 90

Table 5. 1. Filtering using physicochemical properties 123
Table 5. 2. Shape similarity metrics statistics 125

LIST OF ABBREVIATIONS

AP2: Adaptor Protein complex 2

Arf: ADP ribosylation factor

ASU: Asymmetric unit

Cog: Conserved Oligomeric Golgi

COP: Coat protein

CPMG: Carr-Purcell-Meiboom-Gil

DSF: Differential scanning fluorimetry

ECM: Extracellular matrix

EDC: 1-Ethyl-3-(3 dimethylaminopropyl) carbodiimide hydrochloride

EEA1: Early Endosome Antigen 1

EGFR: Epidermal growth factor receptor

ER: Endoplasmic Reticulum

ERK1/2: Extracellular-signal-regulated kinases 1/2

FBDD: Fragment-based drug discovery

FT: Farnesyltransferase

GAP: GTPase activating protein

GB1: protein G B1 domain

GDF: GDI Dissociation Factor

GDI: Guanine nucleotide dissociation inhibitors

GEF: Guanine-nucleotide exchange factor

GFP: Green fluorescent protein

GGT: Geranylgeranyl transferase

GppHp: Guanosine 5′-[β,γ-imido]triphosphate]

GS: Griscelli syndrome

HTS: High throughput screening

IP: Inositol Phosphate

IPTG: isopropyl β-D-1 thiogalactopyranoside

LB: Lysogenic broth

LLG: Log Likelihood Gain

MAP: Mitogen-Activated Protein

MBP : Maltose-binding protein

MMP: Matrix metaloprotease

MSA: Magnetic susceptibility anisotropy

NCS: Non-crystallographic symmetry

NEB: New England biolabs

NHS: N-hydroxysuccinimide

NK cells: Natural killer cells

NOE: Nuclear overhauser effect

PCR: Polymerase chains reaction

PCS: Pseudo-contact shifts

PI3K: Phosphoinositide 3 kinase

PLC : Phospholipase C

PPI: Protein-protein interaction

PRE: Paramagnetic relaxation enhancement

Rab: Ras-like proteins in brain

Ran: Ras like nuclear

Ras protein: Rat sarcoma protein

RBD: Rab-binding domain

REP: Rab escort protein

Rho: Ras homologous

ROCS: Rapid Overlay of Chemical Structures

RU: Response unit

SAR: Structure activity relationship

SER: Surface entropy reduction

SHD: Synaptotagmin homology domain

Slp: Synaptotagmin-like protein

Slac: Slp homologue lacking C2 domain

SPR: Surface plasmon resonance

STD: Saturation transfer difference

TBC domain: Tre 2/Bub2/Cdc16 domain

TEV: Tobacco Etch Virus

TGF: Transforming growth factor

TGN: Trans Golgi Network

TNCS: Translational non-crystallographic symmetry

VEGF: Vascular endothelial growth factor

WaterLOGSY: Water-ligand Observed via Gradient Spectroscopy

DECLARATION OF ORIGINALITY

The thesis presented here, contains data that is the result of my work during my PhD studies, and all else is appropriately referenced.

COPYWRITE DECLARATION

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the license terms of this work.

CHAPTER ONE: INTRODUCTION

1.1 Ras superfamily of proteins

Guanine nucleotide binding proteins are constituted of heterotrimeric G protein complexes and monomeric small GTPases that, together with their effectors, activators and inhibitors, are involved in almost all cellular functions. Most of the monomeric small GTPases are members of the Ras (rat sarcoma) superfamily of proteins [1]. Fundamental to the function of this superfamily of proteins is a GDP/GTP switch that controls the activity of these proteins. Primarily Ras proteins are bound to GDP and preserve an inactive conformation [2]. Binding to GTP induces a conformational change in the protein's structure leading to activation and effector interactions (Figure 1.1).

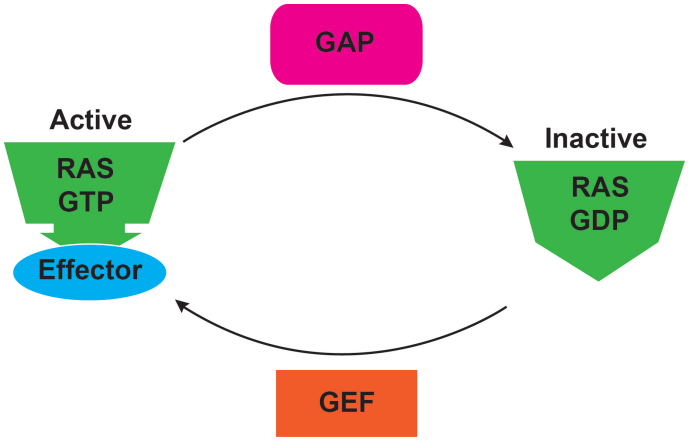


Figure 1.1. Ras protein cycle. All members of the Ras superfamily of proteins cycle between two active or inactive states through exchange of GDP and GTP. This conformational switch is facilitated by GTPase Activating Proteins (GAP) or Guanine nucleotide Exchange Factors (GEF).

Ras, Rho, Ran, Arf and Rab proteins are all members of the Ras superfamily of proteins, with each subfamily involved in different cellular functions. Ras subfamily of proteins known for their role in cancer, is constituted of H-Ras, K-Ras and N-Ras which are at the heart of a network of proteins involved in cell proliferation, growth and survival [3]. Initially, *RAS* genes (*K-RAS* and *H-RAS*) were thought as oncogenic viral genetic elements that can transform sarcoma cells [4, 5]. Separate studies demonstrated that genetic material responsible for transformation of NIH/3T3 mouse fibroblasts from chemically treated fibroblast cells is identical to the initially identified viral *H-RAS* and *K-RAS* [6]. Variants of *K-RAS* and *H-RAS* (and *N-RAS*) such as

G12V mutant were isolated from tumor cells and identified as responsible for disabling cellular control over Ras proteins in tumor cells [7]. DNA Sequencing confirmed the presence of several other *RAS* mutants across a whole range of cancer types. Intriguingly the high occurrence of *K-RAS* mutants in lung [8, 9], pancreatic [10] and colon carcinomas [11], make it the most frequently mutated oncogene in carcinogenic cells [12]. *K-RAS* also distinguishes itself from other forms of *RAS* as being an essential player of cell survival and functioning, as shown by knockout studies [13].

Ras proteins are activated by upstream signaling pathways and exert their function by activating crucial downstream kinase pathways that are not only involved in cell proliferation but also differentiation, organ development and neural plasticity. Activation of the epidermal growth factor receptor (EGFR) induces autophosphorylation of the intracellular domains triggering the sequential attraction of several proteins that lead to the exchange of GDP for GTP in Ras proteins and hence activate downstream pathways such as the MAP kinase, phosphoinositide 3-kinase (PI3K) and Phospholipase Cε (PLCε) pathways [14]. Raf kinase is a Ras effector that binds to GTP-bound Ras with high affinity. Once activated by RasGEF, Ras proteins activate the Raf kinase that in a series of cascades activate extracellular-signal-regulated kinases 1/2 (ERK1/2) which are a member of the Mitogen-Activated Protein (MAP) kinase family [3]. ERK will subsequently translocate to the nucleus where it phosphorylates certain transcription factors leading to cell proliferation or survival depending on cellular conditions. A schematic presentation of Ras upstream and downstream pathways is depicted in figure 1.2. Although being associated with several cancers, attempts to target the Ras pathway have not been very successful demonstrating the challenges in designing therapeutics against these proteins [12, 15].

Ras homologous (Rho) proteins are another member of the Ras superfamily that are involved in actin polymerization, cell shape and cell movement [16]. They facilitate cancer from cell survival to cell invasion and metastasis [17]. A key feature of the regulation of Rho proteins is guanine nucleotide dissociation inhibitors (GDIs) that extract inactive Rho from membranes and keep lipidated Rho solubilized in the cytosol [18, 19].

The Ras like nuclear (Ran), a member of the Ras superfamily are the most abundant small GTPase proteins inside the cell and are an indispensable component for all transport in and outside of the nucleus except for the export of

mRNA and ribosome subunits [20]. Ran-GTP binds to karyopherins bound to cargo inside the nucleus and induces the dissociation of the cargo. Subsequently it is transported with the karyopherin to the cytoplasm where RanGAP (GTPase activating protein) induces hydrolysis of GTP to GDP, leading to dissociation of Ran from karyopherins. Otherwise it can bind to karyopherins transporting cargos to the cytoplasm. Hydrolysis of GTP by RanGAP dissociates the ternary complex [21].

The Ras superfamily has two other major groups that are involved in the biogenesis and transport of vesicles between compartments. ADP ribosylation factor (Arf) proteins are the most divergent Ras superfamily members and regulate formation of coat protein (COP) I and II vesicles [22] while Ras-like proteins in brain (Rab) promote budding, transport and fusion of vesicles from donor to acceptor compartments (see below) [23].

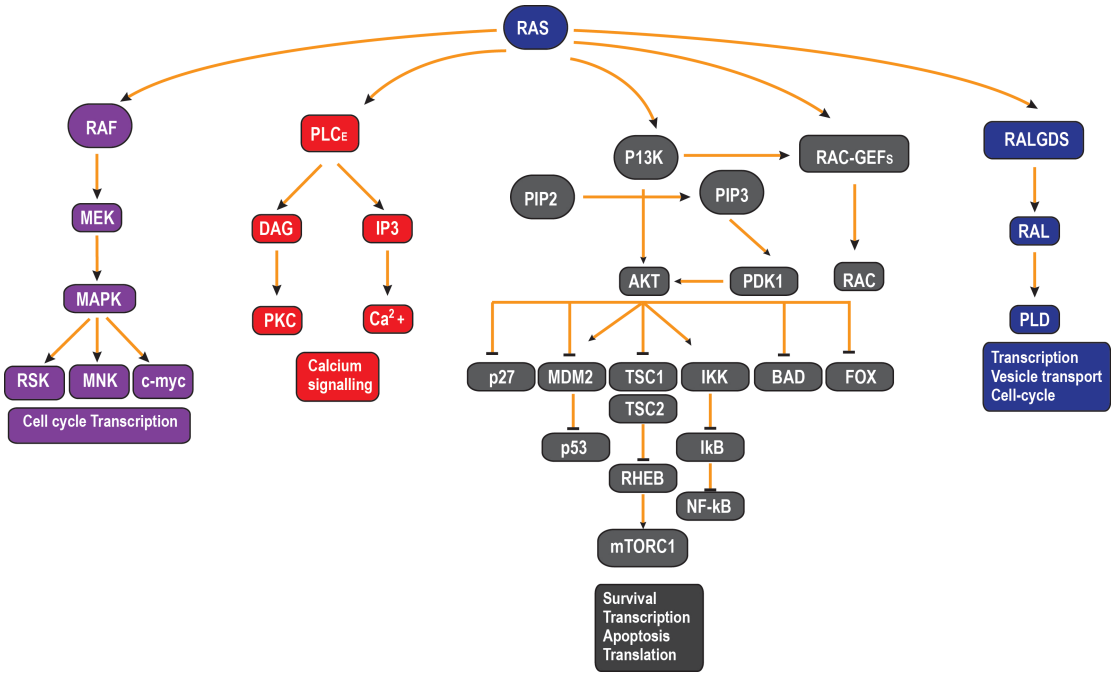


Figure 1.2. Ras downstream signaling pathways. MEK: Mitogen-activated kinase/ERK Kinase, ERK: Extracellular Regulated Kinase, MAPK: mitogen-activated protein kinase, MNK: MAPK-interacting kinases,RSK: Ribosomal S6 Kinase, PLCε: PhosphoLipase ε, DAG: diacylglycerol, IP3: Inositol trisphosphate, PKC: Protein kinase C, PI3K: PhosphoInositide 3-Kinase, AKT: v-Akt Murine Thymoma Viral Oncogene)/ PKB (Protein Kinase-B), PDK1: Phosphatidylinositol trisphosphate-Dependent Kinase 1, BAD: Bcl-2-Associated Death promoter protein, IKK: IκB kinase, TsC: Tuberous sclerosis proteins, MDM2: Mouse double minute 2, IκB: nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, RHEB: Ras homolog enriched in brain, mTORC1: mammalian target of rapamycin complex 1, RAC: Ras-related C3 botulinum toxin substrate 1, FOX: Forkhead box protein, RAL: Ras-like protien, RALGDS: RAL Guanine nucleotide Dissociation Stimulator ,PLD: PhosphoLipase D.

1.2 Structural elements of Ras superfamily proteins

A highly conserved fold of six β-strands surrounded by five α-helices and interconnected by ten loops is a common feature across the Ras superfamily (Figure 1.3). The nucleotide-binding pocket, which is formed by flexible loops, and a prenylation motif that is present in most Ras proteins (except Ran proteins), are key structural features that control the activity of Ras proteins. Five “G boxes”, are conserved structural elements that are involved in shaping the nucleotide binding and hence control the conformational switch (Figure 1.3) [1, 24]. The G1 box or the phosphate-binding loop is “aaaaGxxxxGKS/T”, were “a” is Leucine, Isoleucine or Valine. This sequence uses backbone nitrogen and lysine sidechains to directly interact and stabilize phosphate oxygen. The G2 box also named as the switch I region, contains a conserved Threonine that is crucial for stabilizing the GTP-bound conformation of Ras proteins through interacting with the γ-phosphate of GTP. The G3 box “DxxG” has a highly conserved aspartic acid residue that is involved in binding to a magnesium ion. This ion is essential for GTPase binding to any guanosine nucleotide. G4 and G5 are also involved in interactions with the nucleotide and stabilizing its binding to the protein [1, 25]. Apart from the nucleotide binding motifs, the majority of Ras proteins undergo lipid modification. Some members of the Ras superfamily such as the Ras and the Rho proteins bear a CaaX motif at their C-terminus and interact with Farnesyltransferase (FT) or GeranylGeranyl transferase I (GGT I) [3]. These enzymes covalently link an isoprenoid group to the cysteine in this motif, releasing the terminal three residues. On the other hand, Rab proteins have CC, CXC, CCX or CCXXX motifs and are sequestered to the target membrane by the transfer of GeranylGeranyl group by GGT II enzyme [23]. Finally, Arf proteins get myristoylated on their N-terminus and members of the Ran proteins do not get modified at all [1]. The nature of whether a protein is a substrate for FT or GGT I, is decided by residues downstream of cysteines in the prenylation motif (X). Although the rules are not exclusive, it is generally seen that the presence of a methionine, serine or a glutamine promotes farnesylation, while a leucine or isoleucine trigger geranylgeranylation [26]. Moreover, proteins such as KRas4B or NRas that are substrates for farnesylation, undergo geranylgeranylation in the absence of FT [27]. This promiscuity and complexity in prenylation lead to failed drug discovery attempts targeting FT in order to inhibit KRas pathway. This caused results obtained from clinical studies to be different in comparison to preclinical models [28, 29].

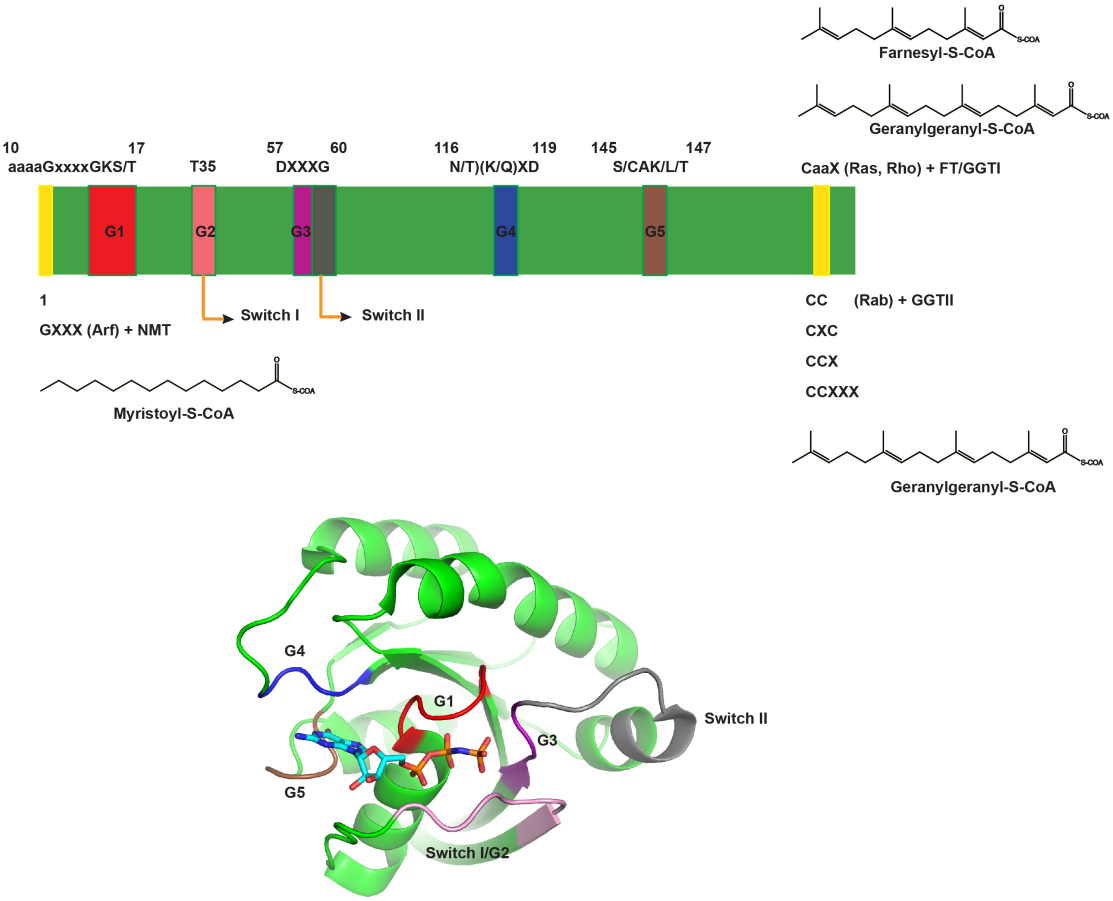


Figure 1.3. Key structural elements of Ras superfamily members. As illustrated, five regions labeled as G boxes, come together and form the nucleotide binding pocket. Switch I or G2 and Switch II that overlaps with G3, are highly dynamic loops that play an integral role in the conformational switch. NMT: N-myristoyltransferase, FT: Farnesyl transferase, GGT: Geranylgeranyl transferase. Protein structure was generated using structure of KRas-gppNHp (PDB code: 3GFT).

1.3 Conformational switch and dynamics of Ras superfamily proteins

The two loops that are part of the switch regions and form the nucleotide-binding pocket in Ras proteins (Figure 1.3), are associated with μs-ms motions. These loops fluctuate between different conformational states in the presence of bound GDP and GTP. This enables the protein to regulate its function by binding to different effector proteins [2, 30]. The original idea that these proteins exist in two rigid GDP(inactive)/GTP(active) states was questioned when evidence obtained by ^{31}P NMR on H-Ras demonstrated the presence of two major states for the GTP-bound protein [31]. High-resolution crystal structures of Ras mutants acting as locked conformational states have helped in the better understanding of their role [32, 33]. Crystallography data suggest state I adopts a nucleotide-free conformation resembling the conformation induced in Ras during nucleotide exchange, but state II echoes the structure of effector-bound Ras structures. This has been supported by the fact that state II has a much higher affinity for target effectors [31, 34]. The interplay between the two states and the exact physiological relevance of state I is still ambiguous. Moreover, recording the motions of the switch regions by NMR have revealed that, while loop II adopts a more rigid conformation in state I in comparison to the GDP state, loop I is extensively flexible, hence facilitating the binding to GTP [32, 34]. These data suggest, state I might guide the protein through a lower energy activation route to the active state II conformation [31, 34]. ^{31}P NMR has revealed that each of Ras isoforms equilibrates differently between state II and I [ratio of 1.3 for H-Ras and 0.072 for M-Ras] [35, 36]. Alanine scans on the nucleotide-binding loop have revealed that T35 in H-Ras (and the equivalent in other Ras proteins) located on the switch I loop interacts with γ-phosphate of GTP and Magnesium and stabilize state II conformation [36]. This has been confirmed through crystal structures obtained from wild-type and T35A mutant of H-Ras. Intriguingly, the crystal structures reveal the presence of two unique pockets in state I as a result of conformational rearrangements. Andrographolide derivatives have been shown to selectively stabilize state I through interaction with one of these pockets that is located in the proximity of the switch I region. This has brought up the idea that state I stabilizers could be used as potential new class of Ras inhibitors [34, 37]. Although the Ras superfamily has a highly conserved fold and this has been shown by several crystal structures from different members of this superfamily, dynamics data is restricted to K-Ras, H-Ras and N-Ras. Since several Rab proteins have been shown to be an attractive target for drug discovery programs, it is invaluable to study the dynamics of these proteins separately and understand their role in the function of

these proteins inside the cell.

1.4 Rab proteins and cellular trafficking

Trafficking of proteins inside the cell is accomplished through formation of vesicles from a donor compartment, transport by motor protein complexes, docking to acceptor compartments and finally fusion of the vesicle to the target membrane. Rab proteins have been shown to directly or indirectly be involved in all these cellular processes. The initial step in cargo trafficking is budding (Figure 1.4a). Arf proteins are known for their central role in recruiting coat protein complexes. These proteins induce the formation of vesicles [22] and are involved in cargo selection. Rabs are also involved in cargo selection, coat protein recruitment and the actual budding itself. Rab9, which is located on late endosome membranes, recruits tail-interacting protein of 47 kD (TIP47) and increases its affinity to bind to Mannose 6-phosphate receptors (MPRs). This process assists MPR selection for transport from late endosomes to the Trans Golgi Network (TGN) in the process of recycling MPRs [38]. Another Rab protein that is a key player in vesicle formation is Rab5 that is involved in endocytic vesicle formation, uncoating, movement and early fusion [39]. Rab5 is incorporated into the clathrin-coated pit during vesicle formation and is essential for the vesicle separation from the membrane. Rab1 has a similar function, enabling budding of COPI coated vesicles from the ER (Endoplasmic Reticulum) [40]. Upon release of the newly formed vesicle from the donor compartment, coat complexes have to be released from the vesicle as they can interfere with the fusion process (Figure 1.4a). There is evidence that Rab5 induces the dissociation of Adaptor Protein complex 2 (AP2) from endocytic vesicles by disrupting its interaction with Inositol Phosphate (IP) 4,5. This is accomplished by recruiting IP-3 kinase and IP phosphatase, which are all Rab5 effector proteins [39, 41]. Subsequently the vesicles are transported to the target membranes (Figure 1.4a). Rab4, 5, 6, 7, 8, 11 and 27 have been shown to be involved in the transport of vesicles using the actin and microtubule network through direct or indirect interaction with motor proteins [42]. While Rab6 and Rab14 use a direct binding to different kinesin isoforms for vesicular transport [43, 44], most Rab proteins use an adaptor protein to interact with the target motor protein. For example, Rab27a uses Melanophilin to interact with Myosin V_a and transport melanosomes on actin filaments around the plasma membrane periphery [45].

The final stages of vesicle transport are tethering and fusion [23]. Tethering is a vital step prior to fusion, ensuring the fidelity of the process (Figure 1.4a). All tethering

molecules fall into two categories: multi-protein complexes and coil-coil tethers [23]. Several protein complexes such as the exocyst (transport from Golgi to plasma membrane) and Conserved Oligomeric Golgi (COG) (retrograde transport within Golgi) have been identified that ensure the vesicles are docked to the right acceptor membranes [23].

Rab proteins have been shown to regulate the docking process by recruiting some of the tethering complexes like the Exocyst. Sec4, a Rab GTPase in yeast involved in exocytosis, has been shown to recruit the exocyst complex through interaction with Sec15 [46]. In mammalian cells Rab11 has been shown to interact with Sec15 [47]. Rab5 interacts with Early Endosome Antigen 1 (EEA1), a coiled-coil tethering molecule involved in early endosome fusion. In fact, Rab5 recruits IP phosphatases to generate IP-3P, a component needed to form the ternary complex with EEA1. This ternary complex forms a membrane domain ready to accept the incoming endosomal vesicle [48].

Fusion is last step in a cell trafficking event. SNARE and SM (Sec1/Munc18 like) proteins are the essential factors for the fusion process [49]. The SM-SNARE complex goes through an open and closed conformation. SM proteins assist fusion by acting as catalysts, by limiting diffusion and bringing the SNARE helices together in the right topological conformation. Four groups of SM proteins in humans control the fusion of vesicles with each targeted to a specific trafficking pathway [49]. Sly1 operates at the ER and Golgi, Munc18 at the plasma membrane, Vps45 with early endosomes while Vps33 is involved in late endosomes fusion [49]. Rab proteins recruit SM proteins to their appropriate membranes either through effector proteins or tethering complexes recruited by Rab proteins. For example, Rab5 recruits Vps45 using Rabenosyn5 and PI-3P, both Rab5 effectors [50]. Altogether, Rab proteins are key players of different aspects of cell trafficking. Moreover, Rabs play a major role in controlling the composition of membrane domains, through change of lipid composition and recruitment of different proteins, thus insuring the specificity of cargo flow inside the cell [51]. Figure 1.4b brings a summary of different Rab functions inside the cell.

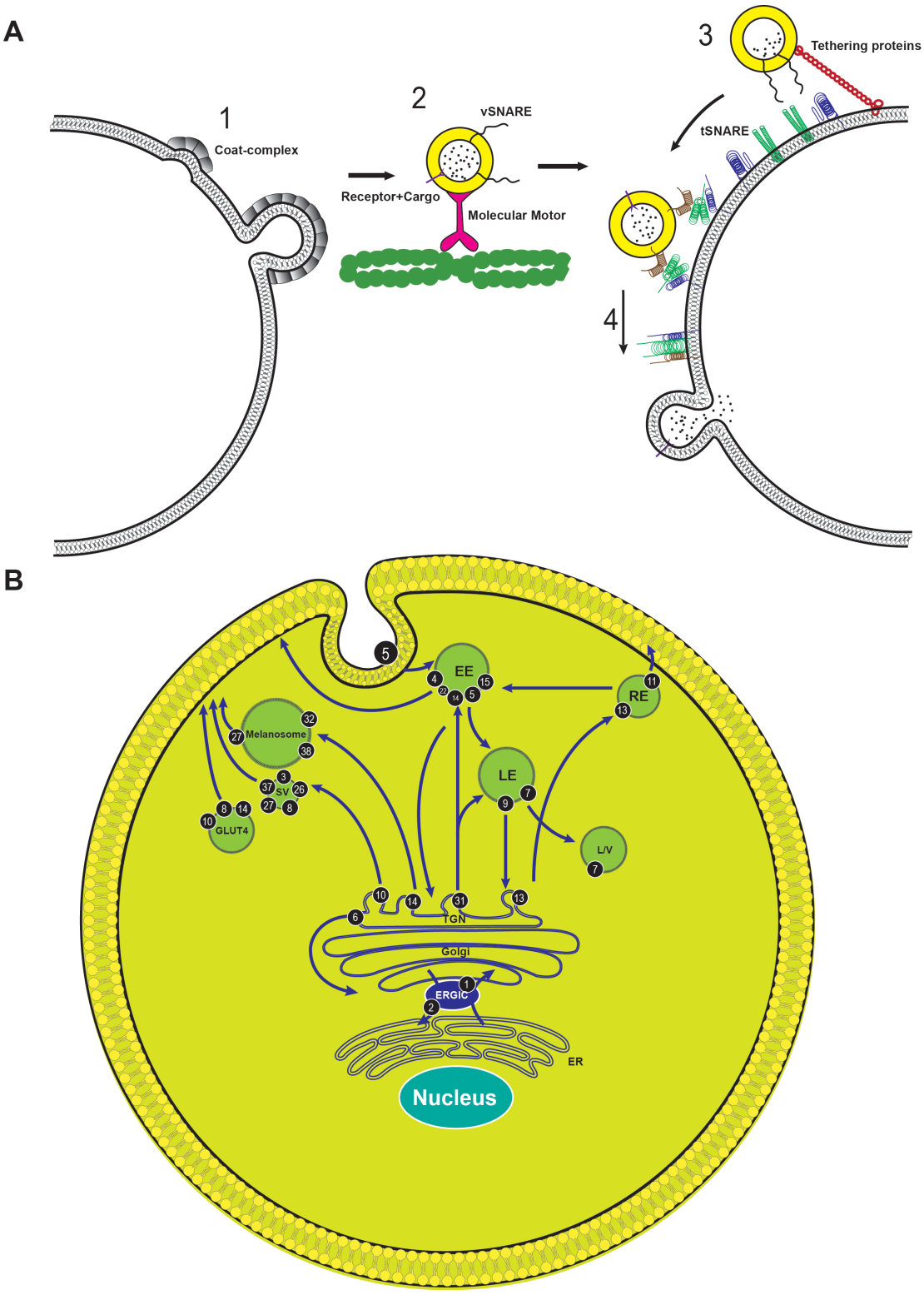


Figure 1.4. Rab function and localization inside the cell. A) Rab proteins act on different stages of vesicle transport including 1) budding 2) transport 3) docking and 4) fusion. B) Localization and cellular function of different Rabs is shown. Rabs on different membranes are responsible for transferring vesicles from donor to acceptor membranes. ERGIC: Endoplasmic Reticulum Golgi Intermediate Component, L/V: Lysosomes/Vacuoles, GLUT4: Glucose transporter type 4, EE: Early Endosomes RE: Recycled Endosome, LE: Late Endosome (LE), SV: Secretory Vesicles. Rab7 tethers LE to Interestingly Rab3, 8, 26, 27 and 37 are involved in the secretory pathway and are found on (SV) [23, 41].

1.5 Rab proteins life cycle

As explained earlier, similar to all Ras superfamily members, Rab proteins cycle between two active and inactive forms inside the cell (Figure 1.5). Upon expression, the protein binds to Rab escort protein (REP), which directs the protein to Rab GeranylGeranyl Transferase II (RabGGT). This enzyme adds a geranylgeranyl group to the C-terminus of Rab, after which it is transported to the membrane by REP [23]. Unlike other Ras superfamily members, Rabs do not have a specific motif for lipidation as it could be a CC, CXC, CCX or CCXXX motif [52] and it is the Rab-REP complex that brings the necessary binding selectivity towards GGTII [53]. REP interacts using C-terminus binding region (CBR) to a sequence of amino acids upstream of the prenylation motif of rabs named CBR interacting motif (CIM) [53]. The complex is presented to GGTII, which leads to a two-step prenylation of rab GTPases. Rab proteins similar to all Ras superfamily members have an equally high affinity for both GTP and GDP resulting in a very slow dissociation from the protein. Therefore, these proteins use Guanine-nucleotide exchange factors (GEF) and GTPase Activating Proteins (GAP) to cycle between different nucleotides. GEF catalyzes substitution of the GDP nucleotide bound to Rab with GTP. This induces a conformational change, which activates its interaction with downstream effector proteins that are involved in transport, tethering or fusing of the vesicle to target the membrane. After this process, GAP proteins accelerate the slow endogenous GTP hydrolysis by Rabs [23]. This promotes the dissociation of Rab proteins from its effectors. Subsequently the prenylated Rab-GDP binds tightly to the Guanine nucleotide Dissociation Inhibitor (GDI) that sequesters Rab from the membrane. Afterwards, GDI Dissociation Factors (GDF) such as Yip3 which is an integral membrane protein, helps to dissociate Rab from GDI and insert it into the membrane keeping Rab in the cycle (Figure 1.5) [23].

A key question in Rab life cycle is how are different Rabs allocated to their target membranes and what is the mechanism for their dissociation from high affinity complexes with REP or GDI. Extensive studies have revealed that GEFs recruit Rabs to their selective final membrane destinations [54]. The use of a mitochondrial-binding domain fused to GEFs associated with Rab1, 5 and 8 resulted in allocation of these proteins to mitochondrial outer membranes [55, 56]. This of course implies that GEFs themselves are transported to distinct membranes through certain mechanisms. Moreover, it has been suggested that the GEFs act as GDFs since they decrease the affinity of the Rab GTPase to REP or GDI, enabling their insertion

into the membrane [54]. This observation has brought up suggestions that unlike Ras subfamily proteins that need a GDF for separation from PDEδ [57], membranes dissociate Rab-REP or GDI complex. It is thought the dissociation is a result of trapping transient states that have the lipid group free of interaction with REP or GDI, by the membrane. Therefore, Rab proteins do not need a GDF [58].

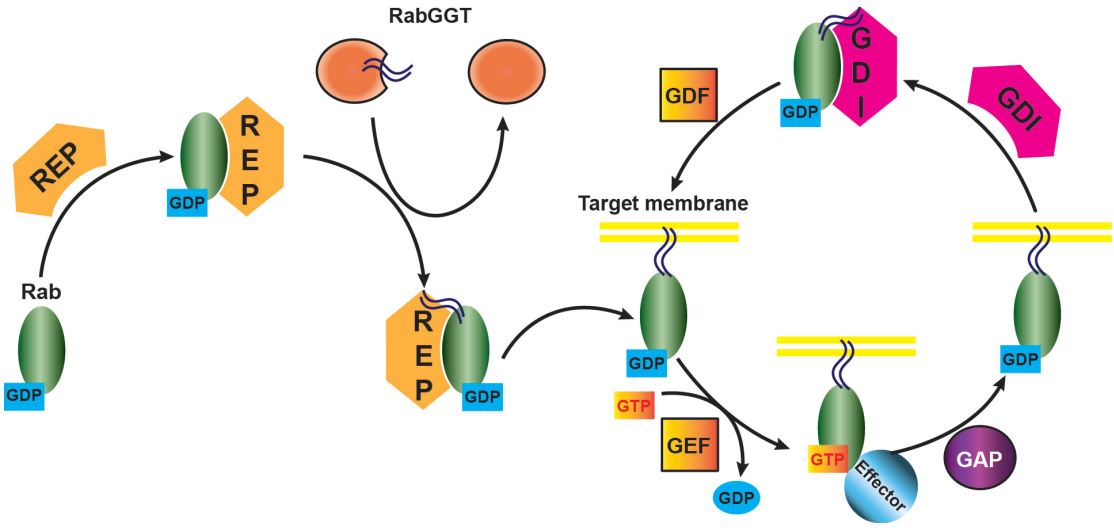


Figure 1.5. Rab life cycle. All Rab proteins follow a similar path in their life cycle inside the cell. REP carries Rab proteins through lipidation by GGTII. Subsequently, the Rabs are transferred to target membranes and GEFs assist the GTP loading process. This promotes effector binding. GAPs stimulate GTP hydrolysis. GDI extracts GDP-bound protein from the membrane delivering it once more to the donor membrane. REP: Rab escort protein, GGT: GeranylGeranyl Transferase, GEF: Guanine-nucleotide exchange factor, GAP: GTPase Activating Protein, GDI: Guanine nucleotide Dissociation Inhibitor, GDF: GDI Dissociation Factor

1.6 Structural characteristics of Rab proteins

All Rab proteins contain similar conserved structural motifs. Rabs are composed of five alpha helices and six beta sheets surrounding each other [59]. There are five motifs running through the switch one and two regions (Figure 1.6; highlighted in red) called RabF1-5. These motifs are conserved among Rab proteins, distinguishing them from other members of the Ras superfamily. Moreover, there are four regions (SF1-SF4) that differentiate Rab proteins from each other [59]. It has been shown that SF1, SF3 and SF4 create an interface used by Rab proteins to bind to effector proteins. SF3 is located in loop between α3 and β5 and is highly variable between different Rabs. SF3 and SF4 are involved in defining the specificity of Rab protein binding to target effectors [59].

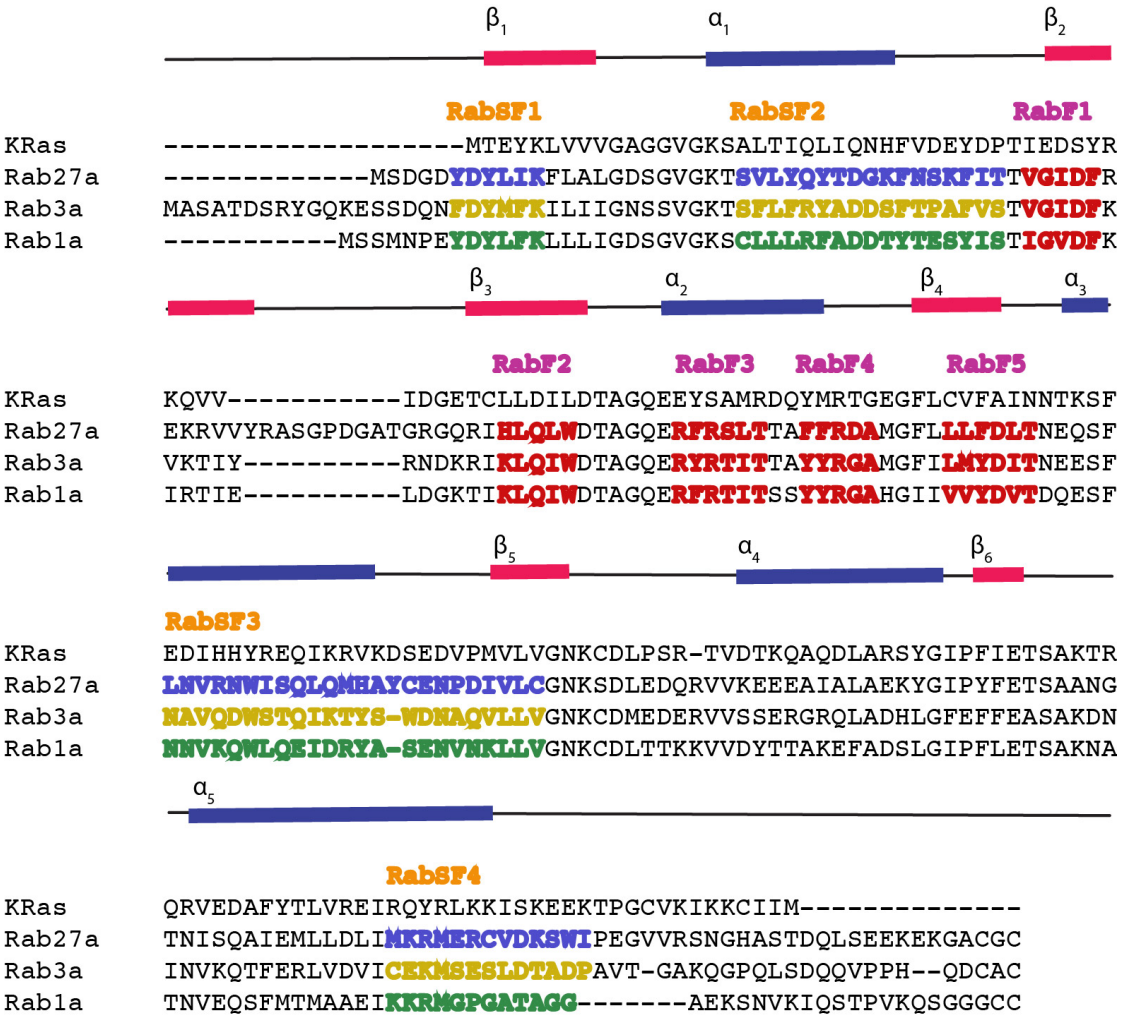


Figure 1.6. Sequence alignment of members of the Rab family. As illustrated by regions in red, RabF motifs are present with high similarity in all Rab proteins. The regions in blue, yellow and green termed SF1-4, are highly variable and are used to distinguish Rab proteins.

As explained before, Ras members have five G boxes that are involved in nucleotide binding. The G2 box also known as the switch I region spans from the loop into β2 region. This segment has a sequence of "T(I/V)G(I/V/A)(D/E)F" holds a highly conserved glycine among all Rab members not present in Ras [60]. The switch II region that overlaps with the G3 box starts from loop into α2 region (Figure 1.3 and 1.6). The conserved threonine in switch I and the glycine and aspartate in switch II provide interactions to stabilize the γ phosphate in GTP [60]. Moreover, the glycine in switch I, adopts unnatural Φ and Ψ angles pushing the sidechain of two conserved valine and isoleucine residues towards hydrophobic residues on the α2 helix of switch II. These interactions are tightened by a highly conserved tryptophan on β3 strand that stacks against them and two other conserved isoleucine and phenylalanine on β2 strand. All this together lead to a much more rigid switch region for Rab proteins compared to Ras (Figure 1.7). Formation of these interactions upon

binding of GTP to Rab proteins induce a change in the inter switch region and provide the interface to interact with effector molecules (Figure 1.7) [60].

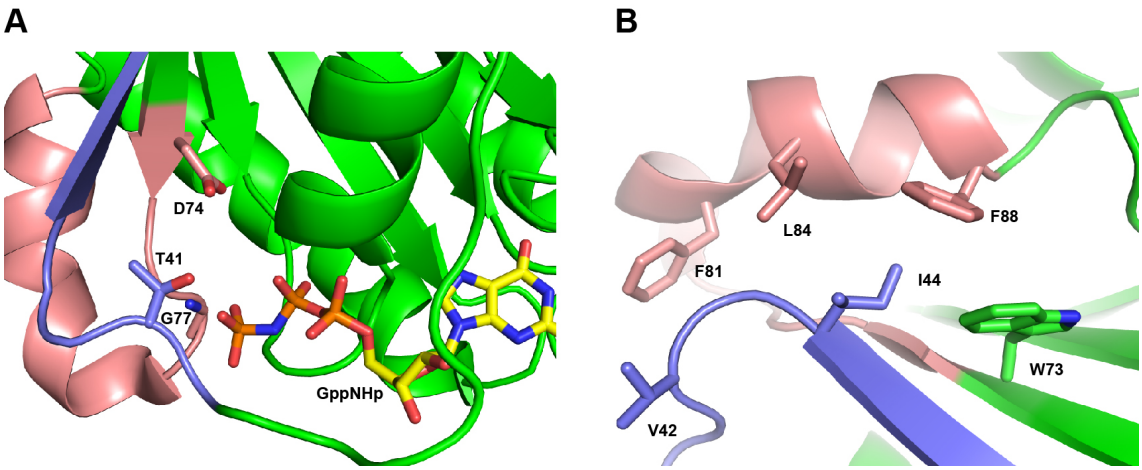


Figure 1.7. Organization of the switch regions for Rab proteins. A) The conserved threonine in switch I (highlighted in blue), aspartate and glycine in switch II (highlighted in pink) stabilize the phosphate in the nucleotide binding pocket. B) The phenylalanine and leucine in the switch II regions (pink) make extensive interactions with valine and isoleucine of switch I region (blue). Also a tryptophan from the β3 is stacked against the phenylalanine in the switch II and leucine in switch I region. These interactions are key in forming the interface required for effector molecule interactions. Images were generated using PDB entry 3BC1.

Since Rab effectors have to differentiate between the nucleotide binding states of the protein, it is not unexpected that all effectors have extensive interactions with the switch regions, overlapping with the binding site of REP, GDI, GEF and GAP [61, 62]. The hydrophobic nature of the switch region coupled to the conserved hydrophobic triad of F59, W76 and Y91 (numbering based on Rab3a), on the surface of Rab proteins has resulted in the hydrophobic nature of the effector proteins complementing their Rab counterparts [63]. The complex and different nature of Rab functions, require diverse range of Rab effectors. This brings up the key question as to how do Rabs selectively bind to their effectors in the pool of binders inside the cell in order to coordinate their function. Part of this selectivity comes from the fact that unlike GEF, GAP, REP and GDI, effector proteins extend their interaction into the variable SF3 and SF4 region as shown in several Rab complexes such as Rab3-rabphilin complex [64]. Moreover, it has been demonstrated that the conformation of the hydrophobic triad, affected by surrounding residues, is an important determinant in selective Rab-effector interaction [63].

1.7 Rab proteins and the secretory pathway in cell physiology

Apart from Rab1, Rab5, Rab6, Rab7 and Rab11 that are known as "housekeeping rabs" and are present in all cell types, all the other members of this family of proteins are expressed in specific cells or tissues [65]. The cell secretory system can be divided into two groups: 1- the constitutively active secretion system that involves membrane proteins or the extracellular matrix proteins in all cells 2- the regulated system [66]. Previously Rab3 isoforms had been extensively characterized for their role in regulated secretion [67, 68] but proteomics studies of organelle components involved in the regulated secretory pathway have identified multiple other Rabs not previously identified on the synaptic vesicles [69], natural killer (NK) cells secretory lysosomes [70], insulin secretory granules [71] and pancreatic acinar zymogen granules [72]. To complement proteomics studies and overcome the problem originated by trace vesicles from endosomes or Golgi transport, green fluorescent protein (GFP) tagged Rabs have been expressed in target cells to screen for Rab proteins that are specifically targeted to regulated secretory vesicles [73]. Among all 43 Rab proteins screened, Rab3(A/B/C/D), Rab26, Rab37, Rab27a and Rab27b have been shown to be exclusively involved as general regulators in trafficking of secretory vesicles [65]. Table 1.1 summarizes different Rab isoforms roles in regulated secretion.

1.8 Rab27a and the secretory pathway

Rab27a is unique in the Rab secretory family as its mutant forms have been correlated with a genetic disorder, Griscelli syndrome (GS); a recessive autosomal disorder known for hypopigmentation and lack of cytotoxic killing activity of T-cells in patients carrying the disease [74]. The molecular mechanisms of hypopigmentation that are related to defects in the transport of melanosomes in melanocytes have been extensively studied. Melanosomes are lysosome-related organelles that store melanin and protect our body against ultraviolet light. After production and maturation in melanocytes, the melanosomes are transported to the periphery of the cells using kinesin motor protein on the microtubule network. Prenylated Rab27a proteins that are sequestered to the melanosome membranes upon activation, bind to Rab-binding domains of melanophilin. Myosin Va motor protein and actin also bind to domains in the middle and c-terminal part of melanophilin respectively [75, 76].

Table 1. 1. Rab3 family role in controlled secretion

Rab isoform	Cellular function
Rab3A	Neurotransmitter release from synapses, hormone secretion from endocrine cells, sperm cell granule secretion [77, 78]
Rab3B	Hormone release from anterior pituitary cells, granule secretion from platelets, cell-surface transport of basolateral LDLR, basolateral transport of VSV-G to epithelial cells [79-81]
Rab3C	Neurotransmitter release from synapses [82]
Rab3D	Zymogen granule secretion from pancreatic acinar cells, Histamine release from mast cells, ACTH release, vWF secretion from endothelial cells [83-86]
Rab26	Amylase release from parotid acinar cells, α_2 receptor transport to plasma membrane [87, 88]
Rab27A	Melanosome transport, hormone secretion from pituitary cell; α and β pancreatic cells, histamine release from mast cells, granule secretion from platelets, neutrophils and CTLs, PSA secretion from prostate carcinoma cells, amylase secretion from cytotoxic T lymphocytes and parotid acinar cells, RANKL release from osteoblasts, exosome secretion, chemokine secretion for cell migration, Weibel-Palade body secretion by endothelial cells [89]
Rab27B	Pituitary hormone secretion, fusiform vesicles transport to the apical plasma membrane of umbrella cells, Zymogen granule secretion from pancreatic acinar cells, histamine release from mast cells, granule secretion from platelets, CTLs and endothelial cells, exosome secretion, RANKL release from osteoblasts, synaptic vesicle exocytosis, PLP transfer to oligodendrocyte membranes [89]
Rab37	Insulin release, histamine release form mast cells, secretion of TIMP1 as an inhibitor of metastasis [90, 91]

LDLR: low-density lipoprotein receptor, VSV-G: vesicular stomatitis virus glycoprotein ACTH: vWF: Von Willebrand factor, CTL: cytotoxic T lymphocytes, PSA: prostate specific antigen, RANKL: receptor activator of NF-κB ligand, PLP: proteolipid protein, TIMP1: inhibitor of metalloproteinase 1

This complex protein network is essential for the correct transfer of melanosomes from the microtubule network to the actin filaments (Figure 1.8) [92]. Any defects in Rab27a cause the loss of interaction with melanophilin and misplacement of melanosomes leading to silvery hair [93]. Apart from melanocytes, Rab27a is

involved in an array of endocrine and exocrine cell secretion system as summarized in table 1.1. Rab27a effectors are summarized in Table 1.2.

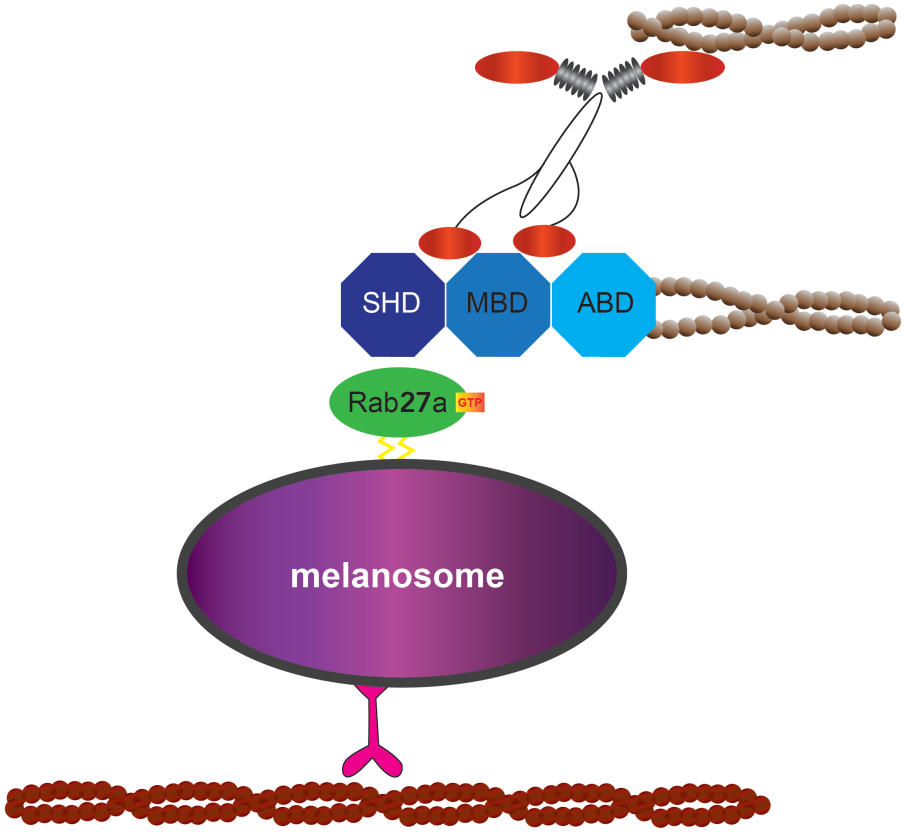


Figure 1.8. Role of Rab27a in melanosome transport. SHD: Slp homology domain, MBP: Myosin V_a binding domain, ABD: Actin binding domain. Rab27a binds to the SH domain of Melanophilin. Also Myosin V_a binds to this complex through the MBD domain. This ternary complex facilitates the transfer of melanosomes from the microtubules to the actin polymers in the periphery of the cells.

Moreover, Rab27a (and Rab27b) is also involved in the secretion of membrane vesicles from the cells [94]. Exosomes are circulating micro-vesicles with protein and RNA contents that have been shown to be involved in tumor aggression, angiogenesis, multidrug resistance and metastasis [95]. Recent studies have directly demonstrated that Rab27a overexpression leads to increase in tumor growth and invasion in breast cancer models. This effect has been correlated with its involvement in the docking of exosomes on to the plasma membrane through interaction with Slp4-a, an effector of Rab27a [94, 96, 97].

Rab27a has a very similar isoform, Rab27b. These two proteins usually share the same effectors and cellular localization, hence they are assumed to share redundancy in their function. However this is not the case in all secretory systems.

Differential phenotypes originating from knock out studies of both isoforms have suggested that the two proteins act at different stages of the secretory system. This has been shown for granule secretion form bone marrow-derived mast cells [98], neutrophils [99] and secretion of exosomes from HeLa cells [94]. In exosome secretion, Rab27a binds to Slp4-a, facilitating the docking process while Rab27b binds to slac2-b. Phenotype analysis of knock out mutants of Rab27b, indicate Rab27b functions upstream of Rab27a in exosome secretion [94].

1.9 Rab27a effectors and their role in the secretory pathway

Eleven effectors to this date have been identified to interact with Rab27a-GTP [89]. These effectors are multi-domain proteins that act by connecting the vesicles flagged by Rab27a (also Rab27b and sometimes Rab3) to motor proteins or facilitate the docking process [100]. These effectors can be divided into four groups based on their domains (Figure 1.9). The first group is the synaptotagmin-like proteins (Slp). Slps were the first Rab-binding proteins to be identified [101]. These proteins were recognized to contain a synaptotagmin homology domain (SHD) at the N-terminus and C2 domains (a type of calcium-binding motif) at the C-terminus [100]. The SHD domain constitutes two helical domains that are used by the Slp proteins to tightly bind to Rab27a [89]. Most of the Slp proteins contain two zinc-binding domains chelating the zinc ions through 4 separate cysteines (exceptions are Slp1 and Slp2-a that lack the zinc-holding motifs) [100].

Table 1. 2. Rab27a effector proteins function

Effector	Group	Function
Slp1/Exophilin7/JFC1	I	Amylase secretion by pancreatic acinar cells, dense granule secretion by platelets, granule exocytosis by cytotoxic T lymphocytes, insulin secretion from undocked granules by pancreatic β cells [102-105]
Slp2-a/Exophilin4	I	Peripheral melanosome distribution, mucus secretion by gastric cells, glucagon secretion by pancreatic cells [106-108]
Slp3-a/ Exophilin6	I	Exocytosis in PC12 cells [109]
Slp4-a/ Exophilin2/Granuphilin	I	Insulin secretion by pancreatic cells, amylase secretion by parotid acinar cells, RANKL release by osteoblasts, Exosome secretion [94, 110-112]
Slp5/ Exophilin9	I	RANKL release by osteoblasts, CFTR channel in epithelial cells, chemokine-induced cell migration [112-114]
Rabphilin-3a/ Exophilin1	I	Dense-core vesicle secretion by PC12 cells [115]
Slac2-a/ Exophilin3/Melanophilin	III	melanosome transport in melanocytes, histamine secretion by mast cell [45, 98]
Slac2-b/ Exophilin5	II	Exosome secretion [94]
Slac2-c/ Exophilin8/MyRIP	III	Melanosome transport in retina, Weibel-Palade body exocytosis by endothelial cells, dense-core vesicle exocytosis, amylase secretion by parotid acinar cells [116-119]
Noc2	II	Dense-core vesicle exocytosis, insulin secretion by pancreatic cells, amylase secretion by parotid acinar cells [120-122]
Munc13-4	IV	Tethering secretory lysosomes at the plasma membrane [123]
Coronin 1C	---	Regulating endocytosis of insulin secretory vesicles in pancreatic cells [124]
IQGAP-1	---	Regulating endocytosis of insulin secretory vesicles in pancreatic cells [125]

Slp4-a also known as granuphilin is a Rab27a effector that is specifically expressed in pancreatic beta cells and is associated with insulin secretion. It is suggested that granuphilin interacts with Syntaxin-1a or Munc18-1 using its C2 domains and helps to tether the secretory vesicles to the membrane prior to fusion [126]. The second group of effectors is Slp homologue lacking C2 domain2-b (Slac2-b) [127] and Noc proteins [128]. These proteins contain the SHD domain but lack the C2 domains. The third group consists of Slac2-a [45] and c [116]. These two proteins lack any C2 domains but contain an actin-binding domain plus a Myosin V_a binding domain used to interact with motor proteins and assist vesicle transport on actin filaments. The fourth group has only Munc13-4 as its only member. Munc13-4 doesn't have the SHD domains used by Slps to interact with Rab27a. Instead contains a Rab-binding domain constituted of beta strands C-terminal to a C2 domain [123, 129] (Figure 1.9).

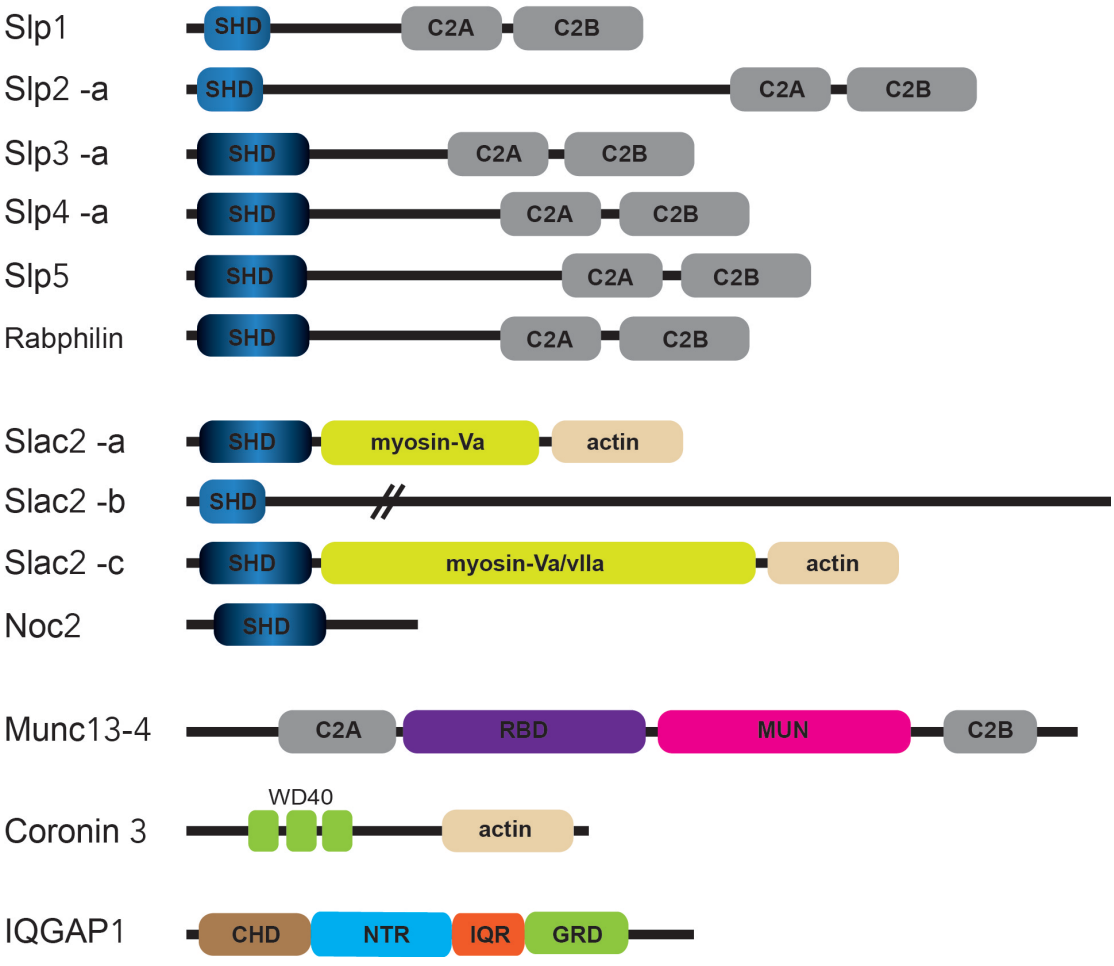


Figure 1.9. Domain organization of Rab27a effectors. SHD: Slp homology domain, RBD: Rab binding domain, CHD: Calponin homology domain, NTR: N-terminal repeats, IQR: IQ repeats, GRD: RasGAP related domain. All the SHD bearing effectors and Munc13-4 only recognize the GTP bound form of Rab27a while Coronin 3 and IQGAP1 only binding to GDP bound Rab27a. All the SHD domains except the ones in Slp1, Slp2-a and Slac-2b have zinc binding motifs.

Generally it is thought that Rab effectors target the GTP form, hence this conformation is widely considered as the “active” form. However two effectors have been identified that bind to Rab27a-GDP [89]. Coronin 3 (or 1C) and IQGAP1 have been shown to bind simultaneously to Rab27a-GDP in β-pancreatic cells and assist the internalization of membranes forming secretory vesicles [130]. Both proteins have different folds not resembling any of the Slp or Slac effectors. Coronin 1C has seven WD-domains forming a β-propeller structure [124]. IQGAP1 has a fold much similar to the Tre-2/Bub2/Cdc16 (TBC) domains present in almost all Rab GAPs but without the enzymatic activity of GAP proteins (Figure 1.9) [125]. Although it is thought Coronin 1C binds to Rab27a through its WD domains, the exact binding mode is not known. Table 1.2 summarizes the functions of these effectors inside the cell.

Rab27a controls the secretion of vesicles at two stages of transport and docking [89]. Based on the structural composition of the effectors it was originally thought that Rab27a uses C2 domain effectors to dock vesicles on plasma membranes as these domains bind phospholipids [89, 131]. Also it uses Myosin V_a containing effectors as an anchor to connect vesicles to motor proteins. A clear example of this notion is demonstrated in the transport of melanosomes, were Rab27a uses a two-step procedure using Slac2-a and Slp2-a [92, 132]. Myosin V_a binds to melanosomes through Slac2-a-Rab27a complex and transports the melanosomes to the periphery of the cells. Subsequently Rab27a-Slp2-a complex promotes docking of the vesicles using its C2 domain. It is clear now that this classification cannot be generalized. For example Slp1 and Slp3-a have been shown to interact with microtubule related kinesin-1 motor protein and facilitate the transport of vesicles in axons and immunological synapses [133, 134]. Moreover, in many cases Rab27a’s function inside the cell does not follow the sequential mode of activity. This has been seen in several cases such as β-pancreatic cells were the three effectors of Noc2, Slp4-a and Slac2-c, all interact with Rab27a modulating the secretion of granules. In this case, a cooperative rather than a sequential mode of action is identified for Rab27a-effector interactions [89].

1.10 Structural characteristics of Rab27a-effector interactions

Rab27a effectors have been separated into two groups based on their affinities. The high affinity group constitutes of Slp1, Slp2-a, Slp4-a, Slp5, Slac2-b, Noc2 and Rabphilin [135]. The low affinity group is composed of Slac2-a, Slac2-c and Slp3-a, suggesting a hierarchical structure for vesicle transport by Rab27a effectors. Structures of Rab27a-Slp2-a and Rab27b-Slac2-a, complemented with mutagenesis revealed in separate studies have lead to conclusions proposing similarities but also striking differences in the nature of the interactions [136, 137].

Rab27a makes use of a large accessible surface area to interact with effector protein; estimated around 1970Å^2 for Slp2-a and 2540 Å^2 for Slac2-a [137]. Moreover, the nature of the interaction suggests that the first long helical segment of Slp effectors packs against the hydrophobic surface of Rab27a, similar to the Rab3a-Rabphilin interaction. In Slp2-a interaction, the hydrophobic triad of Rab27a consisting of residues F46, W73 and F88 together with L9, V42, F81 and L84 make extensive contacts with I15, V18, L19 and L25 of RBD1 in Slp2-a (Figure 1.10) [137]. The importance of these residues is seen in GS patients where W73G mutations lead to the characteristic hypopigmentation as a result of Rab27a-Slac2-a interaction abrogation. Apart from the hydrophobic core, several hydrogen bonds and electrostatic interactions through charged residues provide additional contacts bringing selectivity to the effector interactions. Residue R35 in Slac2-a is highly conserved between all Rab3a and Rab27a effectors making a critical electrostatic contact with D7 (Figure 1.10). This interaction is also observed in Rab3-Rabphilin complex structure. In contrast, hydrogen bonds between Rab27a-D91 and R29 in Slac2-a (K in Slp2-a) plus Rab27a-Y6 with E32 (present in all Rab27a effectors) bring selectivity towards binding to Rab27a. In Rab3a, D91 is substituted by a glycine, while Y6 is F19, both changes failing to make the necessary hydrogen bonds. The absence of these interactions, prevent Rab3a binding to these effectors, making them selective towards Rab27a [137].

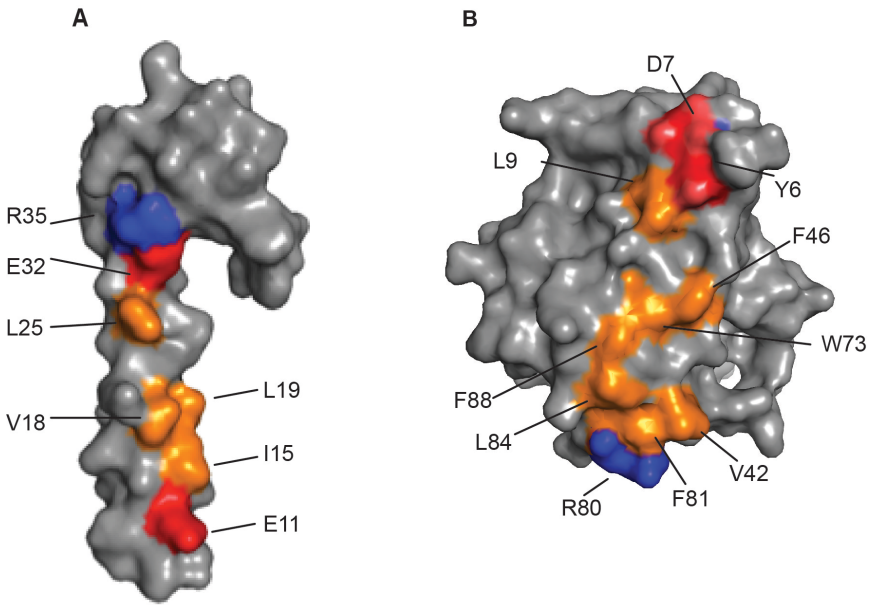


Figure 1.10. Rab27a-Slp2$_a$ interaction interface. Residues involved at the interaction interface of A) Slp2$_a$ and B) Rab27a are shown. As depicted a large hydrophobic patch contribute to the interaction between Rab27a and Slp2-a. Hydrogen bond/electrostatic interactions between R35, E32 of Slp2$_a$ and Y6 and D7 of Rab27a, provide selectivity towards Rab27a.

An intriguing nature of Rab27a-effector binding has been the role of α3/β5-SHD2 motif interaction. The α3/β5 loop as part of the SF3 region, is a highly variable region among Ras superfamily members. This loop together with the SF4 region form the SF4 pocket [59]. Furthermore, the SHD2 domain contains the conserved WF(Y) motif. Hence it is logical to assume, interaction between the pair is critical for binding effectors but also brings selectivity towards Rab27a. While this has proven to be the case for Slp2-a [137], Slp1 and Noc2 [138], it has been shown that this interaction plays no part towards the affinity or selectivity of Rab27a-Slac2-a complex formation. In fact deleting the “AYC” sequence of the α3/β5 loop in Rab27a does not disrupt its interaction with Slac2-a, nor does swapping it in Rab3a, give Rab3a Slac2-a binding capability [116]. All together, these evidence suggest multiple factors affecting the affinity and selectivity of Rab27a-effector interactions.

1.11 Rab27a and its role in cancer

Rab proteins are major components of cell trafficking; therefore it is no surprise that defects in their function has been seen in a range of disease such as genetic defects, infectious disease or cancer. Several neurological disorders such as Parkinson’s disease [139], Huntington [140] and the Carpenter syndrome [141] have been correlated with Rab malfunction. Also, it has been demonstrated that

intracellular pathogens such as *Salmonella*, *Chlamydia* and *Legionella* manipulate Rabs and associate with effectors involved in the endocytic pathways to avoid cellular lysosomal system and use the nutrients inside the cell for replication [142-144]. Moreover, overexpression of several Rab proteins has been shown in cancer cells highlighting the possibility that Rabs might have a key role in tumor progression [97]. This is summarized in table 1.3.

Table 1. 3. Role of Rab proteins in tumor progression

Role on cancer	Rab isoforms	Interacting/intermediate molecule
Integrin trafficking	Rab11/Rab25/Rab21	α6β4/α5β1/ α integrins [145-147]
Protease secretion	Rab8/Rab27a,b/Rab4	MT1-MMP/MMP9,2/procathepsin-L [96, 148-150]
Tumor–stromal cell communication	Rab27a	Exosomes [96]
Cell-cycle progression	Rab6c,Rab27b, Rab21	nd/ HSP90α/ α integrins [147, 151]
Resistance to anti-cancer drug resistance	Rab4/Rab8	P-gp/TMEM205 [152, 153]
Signaling antagonist	Rab23	Sonic Hedgehog pathway [154]

nd:not determined, MT1-MMP: membrane type 1-matrix metaloprotease, P-gp:P-glyco protein, TMEM205: transmembrane protein 205

Rab27a has been associated with two types of disease. One is the genetic disorder known as the Griscelli syndrome, a recessive autosomal disorder known for silvery hair (due to hypopigmentation) and immunodeficiency (due to lack of cytotoxic killing activity of T-cells) [155]. The second one is the promotion of growth and invasion of cancer cells. Rab27a is over expressed in multiple cancer such as breast [156], lung [157], pancreatic [158] and hepatocellular carcinoma [159] and its up regulation has been correlated with poor prognosis. Moreover, depleting Rab27a from cancer cells decreases the size of primary tumors and significantly reduces metastases [96, 160]. In breast cancer, Rab27a is overexpressed in tumor tissues compared to normal cells and also in metastatic lymph node obtained from breast cancer [161]. Moreover, depletion of Rab27a from human triple negative breast

cancer cell lines (MDA-MB-231 and BT549), diminishes degradation of the extracellular matrix, blocking a major step in metastasis (unpublished data from Chiara Recchi).

Rab27a promotes tumor growth and invasion on two levels. First, directly regulating the secretion of important soluble proteins from cancer cells such as chemokines and MMPs. These proteins are involved in the restructuring of the local extracellular matrix (ECM), which is vital for the growth and invasion of neighboring tissues [96]. The second level is triggering pro-tumoral cells such as neutrophils and myofibroblasts by modulating the secretion of exosomes [96]. Exosomes are small vesicles with diameters between 50-100 nm that are formed inside the cell from the endocytic vesicular system and are secreted from the cells by fusion with the plasma membranes. These micro-vesicles circulating in the blood stream, contain protein and RNA involved in tumor aggression, angiogenesis, multidrug resistance and metastasis [95]. Specifically, exosomal transforming growth factor-β1 (TGF-β1) complemented by the transmembrane proteoglycans are essential for differentiation of fibroblast cells to myofibroblasts [162]. This effect was not replaceable by soluble TGF-β1 and is also abolished by targeting Rab27a. Moreover, neutrophils, which play a pro-tumoral effect, are mobilized by exosomes and relocated to the cancerous tissue releasing granules (that has been shown Rab27a again is responsible for release) containing proangiogenic factors such as MMP9 and vascular endothelial growth factor (VEGF) [96]. Although this evidence suggest Rab27a is vital in tumor growth and metastasis its exact effect in remodeling the ECM and the immune system is not well understood and can vary from one cancer type to another.

1.12 Targeting Ras proteins

The dependence of Ras proteins on the recycling of their multi-stage regulatory cycle, means that blockade of any of the steps through direct inhibition of Ras or indirect inhibition by interacting with Ras partnering enzymes/effectors can lead to the arrest of their regulatory cycle, blocking their function (Figure 1.11) [163-165]. Direct inhibition methods have predominantly been focused on targeting the nucleotide-binding pocket, through stabilizing the GDP state or state I GTP conformation [37]. Ligands binding to pockets present in the GDP state often overlap with the GEF binding site, hindering the nucleotide exchange. Apart from this, direct blockage of GEF or effector interactions have been attempted through chemical

libraries or peptidomimetic approaches [166, 167]. These methods suffer from lack of affinity and selectivity. Indirect methods have mainly been towards impairing the spatio-temporal signaling of Ras proteins by inhibiting Farnesyl transferase (FT). Although, initial cellular data were very promising, several compounds that reached late stage clinical trials, showed to be ineffective. This was due to the compensatory prenylation of K-Ras by GGT in the absence of a functional FT [27, 28]. This demonstrates the importance of understanding the complexities associated with cell biology in target validation.

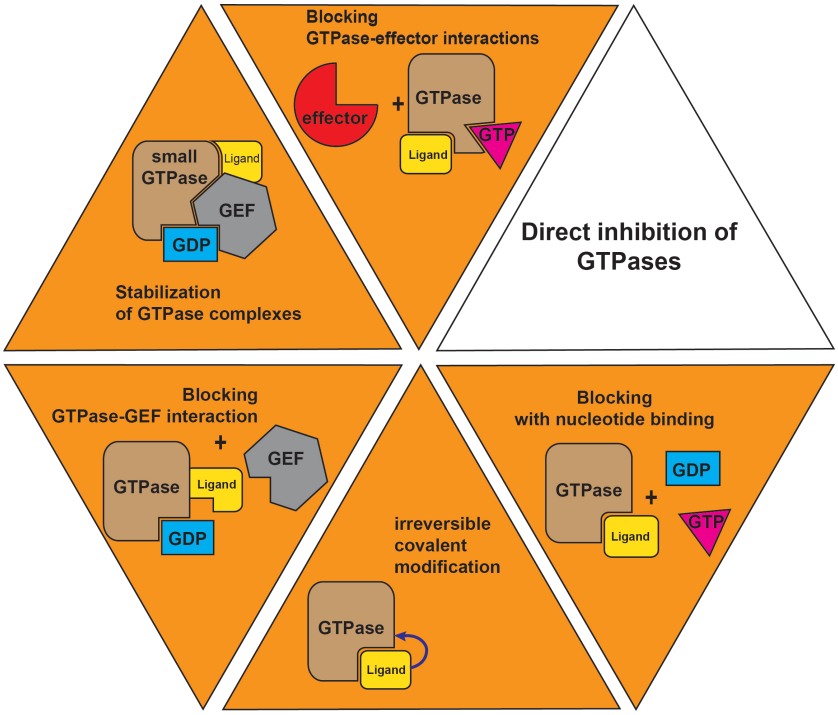


Figure 1.11. Direct inhibition of Ras proteins. Direct methods used to inhibit Ras members are shown. Ligands can interfere with nucleotide binding covalently or non-covalently through overlapping pockets or blocking direct binding. Also the cycle can be arrested by blocking GTPase-GEF binding, effector interaction or stabilizing complex with the GEF or GAP protein.

Several approaches have been taken towards direct Ras inhibition from a medicinal chemistry point of view. These methods can be classified into nucleotide analogues, in silico approaches, fragment-based ligand discovery, tethering approaches and peptidomimetics. Table 1.4 gives a summary on the approaches taken leading to potential direct binders of Ras proteins.

Table 1. 4. Direct inhibition of Ras proteins

Approach	Target	Comments	Ref.
Nucleotide analogue	Pan-inhibition	on-bead exchange of BODIPY-conjugated guanosine nucleotides	[168, 169]
In-silico methods	Rac1-GEF interaction	vHTS of >200000 compounds by docking with IC_{50} of ~50 μM	[170]
	M-Ras_{P40D}-GTP	Docking on a transient pocket present in state I GTP followed by similarity search	[37]
Fragment-based ligand discovery	K-Ras-GDP	11000 fragments screen by ^{15}N HSQC	[171]
Tethering	K-Ras_{G12C}-GDP	disulfide tethering screen	[172]
	K-Ras-GDP	Screen of cysteine mutants by a small thiol reactive groups attached to a ligand, appropriate match was taken for secondary screening	[173]
Peptidomimetics	K-Ras-GDP Rab8a-GppNHp	HBS method applied to a segment of SOS that binds K-Ras	[167]
		Stapled peptides that are segments of Rab8a effector using *i*+4, +7 alkene linkers	[166]

BODIPY: boron-dipyrromethene, vHTS: virtual high throughput screening, SOS: Son of Sevenless (Ras GEF)

Qi Sun et al. used a combination of fragment screening by protein-based NMR, SAR by catalogue and X-ray crystallography to identify and optimize fragments to ligands with approximate K_D of 200 μM that inhibit nucleotide exchange of K-Ras [171]. This work demonstrates that although the hits obtained through fragment screening for Ras proteins have affinities above the millimolar range, effective use of protein-based NMR can enable identifying these hits and their approximate binding

sites. Moreover, improvements in affinity and solubility by means of SAR by catalogue are important for obtaining crystal structures of the ligands. Ostrem et al. took advantage of a cysteine present close to the nucleotide binding site in K-Ras mutant G12C, to identify ligands that impair GDP to GTP exchange [172]. Analyzing the crystal structures revealed that the ligands bind to an allosteric site not visible in wild-type Ras. Although ligands obtained in this study are limited to only K-Ras G12C mutants, it illustrates the potential use of tethering to identify ligands targeting shallow binding sites in such proteins.

Although, significant research has been put towards Ras family ligand discovery, it suffers from several aspects. The majority of studies in this field fail to deliver high affinity ligands and most of the reported molecules have high micromolar affinities, which raise concerns whether the affinity recorded for such molecules were measured in soluble single molecule concentrations and not colloidal soluble aggregates. Moreover, as selectivity can be a major issue, especially in a large superfamily with high structural conservation, no systematic selectivity readings is done similar to the standards of research in kinase ligand discovery projects. Finally, studies lack comprehensive data to demonstrate the effect of such ligands on different parts of the protein's regulatory cycle. All together, steadily as our knowledge towards Ras proteins is improved due to high-resolution crystal structures obtained from different conformational states and dynamic data gathered by NMR spectroscopy, we walk away from the general idea that Ras GTPases are "undruggable" and targeting them gathers momentum.

1.13 Targeting Rab27a

Evidence suggesting the significant role of Rab27a in multiple cancer types suggests Rab27a inhibition could potentially impact invasive primary tumors and metastases. Moreover, the viable phenotype of Rab27a knockouts inferred from the redundant role of this protein in normal adults cells, implies inhibiting this protein would lead to low toxicity, hence making it an appropriate target for a potential cancer therapy. Since, most of drug discovery and target validation in the Ras superfamily has been towards Ras mutants, this information can be usefully implemented in similar projects targeting other members of the superfamily.

1.14 Aims and objectives of project

As most drug discovery projects rely heavily on structural information, it is imperative to have access to structural models, from crystallography or NMR, in order to be able to drive forward the medicinal chemistry side of a drug discovery project. Although a crystal structure of Rab27a is available in complex with Slp2-a RBD, Rab27a does not crystalize without an effector and attempts to crystalize it have failed. This project has two major targets 1) Develop models that can be used to study Rab27a-ligand interactions 2) Design and run screens feasible within the capabilities of our group to obtain potential hits and validate these ligands by means of NMR and/or crystallography.

CHAPTER TWO: MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals

All chemicals used throughout the experiments were purchased from Sigma unless otherwise indicated. Polyethylene glycols used in crystallography were purchased from Molecular Dimensions. All forms of guanosine analogues were purchased from Jena Biosciences. Primers for molecular biology were purchased from Invitrogen. Synthetic genes were purchased from GeneScript. All ligands were bought from Enamine.

2.1.2 Media

Lysogenic broth (LB): 10 g tryptone, 5 g yeast extract, 5 g NaCl in 1 litre of double-distilled H_2O (dd-H_2O). 2xTY media: 16 g tryptone, 10 g yeast extract, 5 g NaCl in 1 litre of dd-H_2O. Minimum media: 3.4 g Na_2HPO_4, 1.5 g KH_2PO_4, 0.25 g NaCl, 3 g D-glucose (^{13}C isotope labeled glucose if the intention was having ^{13}C labeled protein), 0.5 g NH_4Cl (^{15}N isotope labeled NH_4Cl if the intention was having ^{15}N labeled protein), 1 mL of 1 M $MgSO_4$, 0.5 mL of 10% thiamine, 300 µL of 1 M $CaCl_2$, 150 µL of 1 M $MnCl_2$, 90 µL of 1 M $FeSO_4$ in 1 litre of dd-H_2O. Seleno-methionine labeled media: M9 media without NH_4Cl plus 200 mg of each amino acid with the exception of cysteine and tyrosine, 120 mg/L of selenomethionine and 30 mg/L of methionine. All bacterial agar plates were prepared by addition of 1.5 g of agar per 100 mL of broth media.

2.1.3 Cell strains

All bacterial strains were purchased from New England Biolabs (NEB). Cloning procedures were done using *Escherichia coli* (E. coli) NEB5α strains. For all protein expression purposes commercially competent BL21(DE3) strains of E. coli were used.

2.1.4 Plasmids

pET15b was purchased from Invitrogen. pMAL-c2X and pRSF1 vectors were kindly provided by Dr. Steve Mathews laboratory. pGEX-6P1 vector was purchased from GE Healthcare. All Rab genes were kindly provided by Dr. Miguel Seabra. Jose Ramahlo at the University of Lisbon kindly provided the genes for all Rab27a effectors.

2.2 Molecular biology

2.2.1 Cloning

All polymerase chains reactions (PCR) were done using the Q5 PCR kit from NEB. All processes including linear DNA gel extraction and PCR purification, were accomplished using the appropriate Gene-Jet Thermo-scientific kits, according to the manufacturers instructions. Mutagenesis, additions, deletions or extensions in target genes were performed using the NEB mutagenesis kit. All subcloning of genes into appropriate vectors were performed restriction enzyme free using the Gibson assembly kit from NEB according to manufacturers instructions. Plasmid purification was performed using Gene-Jet Thermo-scientific kit and samples were sent for sequencing to SourceBioscience.

2.2.2 Transformation

Transformation of plasmids into bacterial strains was performed using a standard heat shock procedure: plasmids and cells were mixed and incubated on ice for 10 min, followed by 90 sec at 42°C, 3 min on ice, 400 µL LB added and incubated at 37°C for 1 hour and finally spread on LB agar plates with appropriate antibiotics.

2.2.3 Agarose gel electrophoresis

1% Agarose gels were prepared by heat assisted dissolving of agarose in Tris-Acetate-EDTA (TAE) buffer. Gels were run at 120 V and 2 Log DNA ladder purchased from Fermentas was used to estimate DNA band size. SYBR™ safe was purchased from Thermo-Fisher for DNA visualization.

2.3 Protein expression and purification

2.3.1 Rab27a expression and purification

All Rab27a constructs included human Rab27a accession number P51159, residues 1-192. Rab27a was expressed using a pET15b vector with N-terminal his-tag and Tobacco Etch Virus (TEV) enzyme cutting site separating the tag from the protein. Non-labeled proteins were prepared by growing cells in LB media unless otherwise mentioned. For labeling, minimum media with appropriate isotopes were used. For protein expression, plasmids were transformed to E. coli BL21 cells and spread on LB agar plates containing 100 mg/litre Ampicillin as resistance marker. Overnight culture was made from streaked plates into LB media containing 100 mg/litre Ampicillin and grown at 37°C. Main cultures were inoculated using the overnight cultures at 1% v/v. All cells were grown in appropriate media at 37°C and protein expression was induced [after bacterial turbidity measured by absorbance at 600 nm reached 0.7] using 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37°C for 3 hours. Subsequently the cells were centrifuged at 4k RPM for 10 min. The cells were re-suspended in a solution containing 500 mM NaCl, 10 mM imidazole, 5 mM $MgCl_2$ and 50 mM Tris at pH 8. Afterwards the cells were lysed using a cell disruptor at 25K psi and centrifuged at 15k RPM for 45 minutes. The supernatant was loaded on Ni^{2+}-NTA resin equilibrated with suspension buffer, washed extensively with the suspension buffer and eluted using a buffer containing 500 mM NaCl, 300 mM imidazole, 5 mM $MgCl_2$ and 50 mM Tris at pH 8. The purity of the protein was confirmed using SDS-PAGE gel electrophoresis. The protein was dialyzed for 6 hours against 100 mM NaCl, 5 mM $MgCl_2$ and 50 mM Tris, pH 8. Afterwards TEV protease was added to the protein solution at a molar ratio of 1/20 in the presence of 1 mM DTT overnight at 4°C to cleave off the histidine tag. The solution was dialyzed against 100 mM NaCl, 5 mM $MgCl_2$ and 50 mM Tris, pH 8 and re-run through Ni^{2+}-NTA resin to recapture TEV protease and any uncut protein. The flow through was collected, concentrated to 10 mg/mL and loaded on a superdex S-75 gel filtration column at a flow rate of 1 mL per min. The column was equilibrated with 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4 for NMR purposes or 150 mM NaCl, 5 mM $MgCl_2$ and 20 mM Tris at pH 8 for crystallography means. The peaks corresponding to Rab27a constructs were collected, concentrated to appropriate concentrations suitable for experiments and flash frozen using liquid nitrogen. Before gel filtration, there could be a step including biotinylation or loading of nucleotides depending on the experiments planned. All Rab27a constructs, including wildtype, mutants or fusions with Slp-2a

Rab27a-binding domain (RBD), were expressed and purified using methods described above. All Rab27a constructs that contained exposed cysteines at C123 and C188 were purified with in buffer containing an additional 0.1% betamercaptoethanol (βME) during Ni^{2+}-NTA steps and 1 mM DTT in all other purification steps (Appendix A Figure 1.A).

2.3.2 Maltose-binding protein (MBP) fused Rab27a expression and purification

MBP-fused Rab27a was used to study Rab27a-ligand interactions using NMR. Rab27a was cloned into a pMAL-c2X vector with MBP on the N-terminus of Rab27a with no enzyme-cutting site in between. The protein was expressed similar to pET15b Rab27a constructs. For harvesting and purification cells were re-suspended in a solution containing 500 mM NaCl, 5 mM $MgCl_2$ and 50 mM Tris at pH 8. Afterwards the cells were lysed using a cell disruptor at 25Kpsi and centrifuged at 15k RPM for 45 minutes. The supernatant was loaded on a 5 mL MBPTrap resin [from GE Healthcare] on an Akta Prime at 1mL per min flow rate. Next, the resin was washed using the same buffer until the absorbance fell below 30 mau. Finally the protein was eluted using 20 mM maltose, 500 mM NaCl, 5 mM $MgCl_2$ and 50 mM Tris at pH 8. The purity of the protein was confirmed using SDS-PAGE gel electrophoresis. The protein was concentrated to 10 mg/mL and loaded on a superdex S-200 gel filtration column equilibrated with 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4. The peaks corresponding to MBP-Rab27a were collected, concentrated to appropriate concentrations suitable for experiments and flash frozen using liquid nitrogen. (Appendix A Figure 2.A)

2.3.3 Rab27a-Slp1 complex expression and purification

Rab27a-Slp effectors complex were only formed using co-expression of both protein in E. coli cells. Slp1 RBD was cloned in a pET15b vector with a protein G B1 domain (GB1 domain) downstream of the TEV cutting site and N-terminal to the Slp1 protein. The Q78L mutant of Rab27a, which is locked in the GTP form, was cloned in a pRSF1 protein with a C-terminal his tag. In order to avoid the C-terminal his tag, a stop codon was placed at the end of the Rab27a translating genetic sequence. Subsequently BL21 cells were co-transformed with plasmids containing both genes. Overnight culture was made from streaked plates into LB media containing 100 mg/litre Ampicillin and 30 mg/litre Kanamycin and grown at 37°C.

Main cultures were inoculated using the overnight cultures at 1% v/v. All cells were grown in appropriate media at 37°C and protein expression was induced at 22°C overnight using 0.5 mM IPTG. Subsequent steps including harvesting and purification were similar to Rab27a purification as explained in sector 2.3.1 (Appendix A Figure 3.A).

2.3.4 Slp expression and purification

Slp proteins were prepared either by dissociating the complex of Rab27a-Slp complex or expressed separately with a GB1 tag.

2.3.4.1 Preparing Slp by dissociating the Rab27a-Slp complex

After preparing Rab27a-Slp complex according to 2.3.3, the complex was treated with 10 mM EDTA for 3 hours. Subsequently Slp was separated from Rab27a by gel filtrating the sample on a superdex S-75 column. The peaks corresponding to Slp were collected, concentrated to appropriate concentrations suitable for experiments and flash frozen using liquid nitrogen.

2.3.4.2 Expression and purification of GB1-Slp

GB1-Slp in a pET15b vector was expressed and purified at 18°C according to methods explained in section 2.3.1. Buffers did not contain $MgCl_2$ for Slp purification.

2.4 Selenomethionine labeling of Rab27a

Selenomethionine labeled Rab27a is produced by expressing protein in minimum media that instead of glucose has been enriched by 0.2 g of each amino acid (except tyrosine and cysteine) plus 125 mg of SeMet and 10 mg of Met per litre. The purification is similar to methods explained in section 2.3.1.

2.5 TEV enzyme expression and purification

TEV enzyme with an auto inactivation resistance mutation of S219V is cloned in a pRK793 vector with Ampicillin resistance with an auto-cleavable MBP tag at the N-terminus that is followed by a his-tag. The plasmid is co-transformed with a pRIL vector expressing the rare argU tRNA that contains a chloramphenicol resistance marker into BL21 cells. The protein is expressed at 30°C for 3 hours. Cells are harvested in a buffer containing 500 mM NaCl, 20 mM imidazole, 10% glycerol and 50 mM Tris at pH 8. Protein is purified using a Ni^{2+}-NTA resin and eluted off using a buffer containing 300 mM imidazole. The eluted protein was concentrated and desalted in a buffer containing 100 mM NaCl, 10% glycerol and 20 mM Tris at pH 8. Finally was protein was concentrated and split in 1 mg aliquots and frozen using liquid nitrogen and kept at -80°C (Appendix A Figure 4.A).

2.6 BirA enzyme expression and purification

BirA was expressed using a pET15b vector with an N-terminal his-tag. Protein was expressed at 18°C. Harvesting and purification was similar to methods described in section 2.3.1. Protein purity was confirmed using SDS-PAGE gel electrophoresis. After Ni^{2+}-NTA purification the protein was concentrated to 3 mL and buffer exchanged using a desalting column into with 30% glycerol, 150 mM NaCl and 20 mM Tris at pH 8. The protein was concentrated to 1 mg/mL, flash frozen using liquid nitrogen and stored at -80^{o}C.

2.7 Biotinylation of proteins

2.7.1 Biotinylation of Rab27a and Slp

Biotinylation of target proteins was enabled by the addition of an “avidity” tag with the 15-mer sequence of **“** GLNDIFEAQKIEWHE “ on the N-terminus of the protein of interest connected using a linker composed of “GSGS” sequence. For biotinylation purified protein of interest was concentrated to 50-100 µM and mixed with a 10X concentrated 0.5 M bicine buffer at pH 8.3 and 10X concentrated 100 mM ATP, 100 mM MgOAc, 500 µM D-Biotin. 2.5 ug of BirA enzyme is added for each 10 nmol of substrate protein. The reaction is incubated at RT for

4-6 hours. The sample is run through Ni^{2+}-NTA resin to capture BirA. The flow through was collected, concentrated and gel filtered using an appropriate column.

2.7.2 Validating biotinylation by western blot

Protein samples were run on a 12% pre-made gel from Generon using Tris-Glycine buffer (25 mM Tris, 192 mM Glycine, 0.1 % SDS at pH 8.4), transferred using Bio-Rad™ Trans-Blot® blotting system onto nitrocellulose transfer packs from Bio-Rad™. Blots were blocked using 3% w/v BSA dissolved in 100 mM NaCl, 0.05 % v/v TWEEN 20 and 50 mM phosphate buffer at pH 7.4. Next the blots were washed with 100 mM NaCl, 0.05 % v/v TWEEN 20 and 50 mM phosphate buffer at pH 7.4 for three time and incubated RT with anti-biotin horse radish peroxidase (HRP) conjugate antibody for 1 hr. Subsequently, chemiluminescent signal was detected using chemiluminescence western blotting kit from Roche (Appendix A Figure 5.A).

2.8 Loading of nucleotides on Rab proteins

2.8.1 Loading of GDP

A continuously inactive conformation of Rab27a is obtained by loading the protein with GDP. To achieve this, Rab27a was concentrated to 5.5 mg/mL in 150 mM NaCl, 10 mM EDTA, 20 mM Tris pH 8 buffer and incubated in the presence of 25 molar excess of GDP at RT for 3 hours. Subsequently the reaction is quenched by adding excess $MgCl_2$ and incubated on ice for 1 hour. Finally the sample is gel filtrated using an appropriate column equilibrated with the preferred buffer [174].

2.8.2 Loading of Guanosine 5′-[β,γ-imido]triphosphate] (GppNHp)

Continuously active conformation of Rab27a is obtained by loading the protein with non-hydrolysable GTP analogues such as GppNHp. To achieve this Rab27a was concentrated to 5.5 mg/mL [5 mg of Rab27a protein and not tag in case of presence] in 150 mM Nacl, 5 mM $MgCl_2$, 20 mM Tris pH 8 buffer. Next it was mixed with a 10X buffer containing 10 mM $ZnCl_2$ and 2 M $(NH_4)_2SO_4$ and 25 units of Antarctic phosphatase [purchased from NEB] plus 4 molar excess of GppNHp were added to the solution and incubated overnight at 4°C. Finally the sample was gel filtrated using an appropriate column [174].

2.8.3 Validating nucleotide loading by NMR

Complete nucleotide loading of Rab27a was evaluated by comparing the ^{1}H-^{15}N HSQC NMR spectrums of nucleotide analogue-loaded Rab27a with Rab27a. The complete shift in multiple peaks plus the disappearance of peaks corresponding to protons on labeled guanosine as a product of labeling media is a good indicator of nucleotide exchange (Appendix A Figure 6.A).

2.8.4 Validating nucleotide loading by HPLC

The sample of recombinant protein of known concentration was deproteinised using perchloric acid (PCA) precipitation. The neutralised supernatant was used directly for HPLC analysis (Ace C18 5 µm 150 x 4.6 mm column, 25 min isocratic run with 5% MeCN and 95% of 100 mM potassium phosphate, 10 mM tetrabutylammonium bromide at pH 6.5, column temp 40°C, flow rate 0.5 mL/min). Calibration curves were runs were performed for each nucleotide (GTP, GDP, GCP and GppNHp) and the linear correlation between nucleotide concentration and signal intensity was confirmed. (Appendix A Figure 7.A)[174].

2.9 Biophysical assays

2.9.1 Differential scanning fluorimetry (DSF)

DSF provides information on the stability of a protein by monitoring denaturation under temperature increments. Melting point (T_m) of a protein is calculated through measuring the change in the intensity of excitatory signal of a fluorescent dye that binds to exposed hydrophobic surfaces of a protein during temperature dependent denaturation. Ligands that bind to the protein would stabilize the folded state, therefore, increasing the Gibbs free energy of unfolding. This leads to an increase in T_m. The relationship between T_m, binding affinity and the free energy of unfolding and binding is complex and needs careful consideration before making comparisons between ligands [175, 176].

All samples for the DSF experiments contained 2 µM protein, 2 mM GppNHp, 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4 and 500X dilution of

SYPRO™ orange except for buffer screen experiments where the samples were under different buffer conditions in each well. For screening purposes a fragment library of 2500 compounds from the Ro3 Maybridge chemical library collection was used. Prior to screening master plates were spun at 1000 g for 1 min. Daughter plates were prepared by pipetting 0.4 µL of 100 mM fragment stocks into positions A/H2-11 of destination plates and diluted 10X with double-distilled water using a multi channel pipette. A/H1 and 12 were filled with 4 µL of 10% DMSO. Separately 1800 µL of a solution containing 2.5 µM of Rab27a-GppNHp and 1.25X of all other buffer components as described above were split between 8 tubes. 16 µL of this solution was pipetted into each well in the daughter plate then plates were sealed with a transparent cover. Final concentration in each well would have the protein at 2 µM and fragment at 2 mM. All DSF experiments were performed on a Mx3005P qPCR System. Experiments were recorded on a stepwise increment of temperature from 25-96°C in 71 cycles. SYPRO™ was used as the preferred fluorescent dye at an excitation and emission wavelength of 485 nm and 625 nm respectively [175]. Results obtained from these experiments were extracted into excel files using the MxPro software. Subsequently, data were fit to a sigmoidal Boltzmann equation using the Graphpad™ software [175]. The melting point of the protein, T_m, in each condition is defined as the 50% point of each curve and obtained by fitting the data using the sigmoidal equation as shown below were X is the temperature and a is the slope at the point of T_m:

$$I = I_{Min} + \frac{I_{Max} - I_{Min}}{1 + exp^{\left(\frac{T_M - X}{a}\right)}} \qquad (1)$$

For hit identification, initially control outliers were ruled out using median absolute deviation. The average and standard deviation between control samples was used to define the threshold for hit identification. Hits were identified using $T_m \geq$ AVG $T_{m\text{-control}}$ + 2×standarddeviation$_{Control}$+0.5 °C as cut off.

2.9.2 Surface plasmon resonance (SPR)

All SPR assays were performed on a T100 Biacore instrument with a T200 upgrade from GE Health Care and analyzed using the Biacore T200 evaluation software. All experiments were done using filtered and degassed buffer containing 10 mM HEPES, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM TCEP and 0.05% v/v Surfactant P20. For assays involving chemical molecules, fresh DMSO was used to adjust the organic solvent's content to 5% DMSO. All experiments were carried out by immobilizing biotinylated Rab27a on Neutravidin-coated CM5 chips. Prior to initialization the system was equilibrated with the new buffer using the prime command and the detectors were normalized towards the new chip using 70% glycerol. For Neutravidin immobilization, 0.1 M N-hydroxysuccinimide (NHS) and 0.4 M 1-ethyl-3-(3 dimethylaminopropyl) carbodiimide hydrochloride (EDC) were mixed and loaded on the chip for 7 min at 10 µL/min. Then 0.25 mg/mL Neutravidin, dissolved in acetate buffer with pH 4.5 was incubated on the chip for 7 min at 10 µL/min. Any remaining active groups were blocked by running 1 M ethanolamine (pH 8.5) for 7 min at 10 µL/min. Subsequently, biotinylated Rab27a were diluted to 80 µg/mL and 240 µg/mL respectively and used to load enough protein to increase the Response Unit (RU) by 10000 units on each cell. The remaining active Neutravidin molecules on the chip were inactivated using 0.29 mg/mL stock of PEG-Biotin (soluble form of biotin) by running through the chip for 120 s. Finally, the machine was left on manual mode to run overnight at 30 µL/min to wash unbound molecules off the chip.

Any run on the SPR machine was composed of 5 type of samples. 1) 50% DMSO for washing injection needle between injections 2) 10 startup buffer injections were preferred at the beginning of each experiment 3) solvent correction samples, which constitute a gradient of DMSO around the 5% which is used by the Biacore evaluation software to adjust for potential buffer mismatch as shown in Table 2.1.

Table 2. 1. DMSO correction curve sample preparation table

DMSO Conc. %	5.8	5.6	5.4	5.2	5.0	4.85	4.7	4.5
100% DMSO in µL	58	56	54	52	50	48.5	47	45
Buffer /no DMSO in µL	950	950	950	950	950	950	950	950

4) positive and negative control samples is used to demonstrate the integrity of the protein on the chip 5) analyte, which could be a protein, peptide or a ligand.

For sample preparation, protein stock solutions were diluted using a buffer with 5% v/v DMSO. The 5% v/v DMSO buffer was prepared by replacing 25 mL of a 500 mL buffer with DMSO. The 25 mL buffer without DMSO was used to dilute the ligands to their final concentration with 5% v/v DMSO content. For ligand screening, master plates were prepared with ligands at 10 mM stock concentration. 600 nl of each compound was dispensed using the Echo Acoustic Dispenser from Labcyte into 96 well plates. Afterwards ligands were diluted to 80 µL and a final concentration of 75 µM. Each plate was composed of 320 compounds, including 10 start up buffer samples, 9 controls (composed of 9 buffer controls and 9 positive controls) that were injected between every 40 cycles and 4 solvent corrections split through the entire run. N-terminal acetylated WF peptide with a free C-terminus was used as a control at 2 mM concentration. For ligand titrations, each ligand had a separate blank plus 1.25, 2.187, 5, 15, 35, 75 and 150 µM concentration test. Also, control samples were injected after every 3 analyte tests. All data analysis was carried out by Biacore™ evaluation software. Three steps of DMSO correction, molecular weight (MW) correction and a comparative adjustment using the positive and negative controls absolute responses were taken to prepare the raw data for comparison and analysis. Since DMSO has a significant effect on the refractive index, the effect of solvent variations between samples was adjusted using a DMSO correction curve prepared as explained in table 2.1. Subsequently, a MW correction was given to response for each sample using the MW for each compound and the MW of the “WF” di-peptide as the positive control. Finally, the response for each sample was normalized, by putting the response for the negative controls at zero and the positive controls at a 100.

2.10 NMR spectroscopy

2.10.1 General considerations

All NMR experiments were recorded at 298K unless otherwise indicated on a 600 MHz or 800 MHz Bruker AVANCE II spectrometer equipped with triple resonance inverse cryoprobes with a single axis z gradient. Generally all samples were made up in a buffer containing 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4 and 10% v/v D_2O otherwise indicated. For protein-ligand interaction studies d_6-DMSO up to 5% v/v was used depending on ligand solubility.

2.10.2 1D ^{1}H and 2D ^{1}H-^{15}N(or ^{13}C) NMR experiments

All 1D ^{1}H NMR experiments were collected using the zgespg pulse sequence. Water suppression was performed using the excitation sculpting method [177]. 1D NMR spectrums were collected with 64 scans, 8196 points and 16 ppm spectral width (SW) at 298K. 2D ^{1}H-^{15}N spectrums were collected at 298K using the hsqcetfpf3gpsi2 pulse sequence. All spectrums were collected using 8 scans, 2048 points and 16 ppm SW in the direct dimension and 140 points and 30 ppm SW for the ^{15}N channel with the center of the channel set at 119 ppm. 2D ^{1}H-^{13}C spectrums were collected at 298K either using the constant time HSQC pulse sequence of hsqcctetgpsp.2 [178] or the standard hsqcetgp pulse sequence. Both spectrums were collected using 8 scans, 2048 points and 16 ppm SW in the direct dimension and 312 points and 66 ppm SW for the ^{13}C channel with the center of the channel set at 40 ppm.

2.10.3 Ligand observed NMR experiments

Ligand observed NMR experiments provide certain advantages that make them indispensable tools for studying protein ligand interactions. These methods don't need labeled protein and, unlike protein observed NMR methods, there is no limit on the size of the proteins of interest to be studied [179]. The experiments are very short and the ability to screen cocktail of compounds makes them a good option for high throughput screening. Moreover, there is a wide affinity range of 100 nM to 10 mM that can be studies using these techniques. Ligand observed NMR experiments are classified into three categories [180]: direct ligand observation,

competitive assays and NOE-based experiments. Direct ligand observed NMR experiments are simple and fast 1D NMR experiments that are used to measure any observable that is changed upon addition of protein to a ligand solution. This observable can be the chemical shift, intensity, or the line width of ligand peaks [179, 180]. The change in each of these observables is correlated to the amount of free and bound form of the ligand that can be used to measure K_D. A major drawback of these experiments is the inability to monitor high affinity ligand-protein interactions due to the slow off rate of bound ligand. This results in the signal from the bound ligand decaying before the ligand is free in solution to be monitored by 1D NMR experiments. To overcome this problem it is suggested to run the experiments in a competitive format using a reporter (spy) molecule with lower affinity that occupies the same binding site as our suspected ligands of interest [181]. This competitive method can also be used to indirectly measure the K_D of high affinity interactions. Nuclear overhauser effect (NOE) based techniques use the fundamental concept of NOE in NMR to monitor protein ligand interactions. Saturation transfer difference (STD), waterLOGSY and Transfer-NOESY (Tr-NOESY) experiments fall into this category. STD and Tr-NOESY experiments can also provide structural information on the ligand binding-mode.

2.10.3.1 1D ^{1}H Carr-Purcell-Meiboom-Gil (CPMG) experiment

CPMG exploits the difference in T_2 relaxation of molecules with different molecular weight and tumbling time. The pulse sequence is comprised of a spin echo delay for 100 ms. Water saturation was achieved using presaturation during the relaxation delay for 3.5 ms. All samples for CPMG experiments were prepared in buffer containing 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4 and 10% v/v D_2O at 298 K otherwise indicated. The ligand/protein ratio was kept at 100/1, with 1 mM ligand concentration and 10 µM protein concentration.

2.10.3.2 Saturation transfer difference (STD)

2.10.3.2.1 General aspects of STD experiment

STD-NMR is based on saturating selected resonances of a protein that correspond to methyl proton chemical shifts that are usually not present in small molecules (0.7 ppm to -2 based on the spectrum of the ligands and protein) using

selective radiofrequency pulses. This leads to saturation of all the resonances in the protein through intramolecular ^{1}H-^{1}H cross relaxations, an effect known as spin-diffusion. If a ligand is bound on the surface (or pocket, cavity etc) of the protein, ligand protons will also saturate due to intermolecular ^{1}H-^{1}H cross relaxations [180, 182]. If the ligands off-rate kinetics is faster than the relaxation time, the saturation effect on the bound ligand will be observed in a decrease in signal proportional to the population of free ligands with saturated resonances. Another off-resonance spectrum that saturates a region far away from normal proton chemical shift distribution is acquired. The subtraction of the on-resonance from the off-resonance experiment yields a difference spectrum similar to a standard 1D ^{1}H proton spectrum. Only peaks that are in contact with the protein and are saturated will appear in the difference spectrum. STD quantification can be achieved by comparing the relative amount of signal decrease in the on-resonance experiment in comparison to the off-resonance experiment using the equation 2 below:

$$STD\ FACTOR = \frac{Io - I_{STD}}{Io} \qquad (2)$$

Where I_o is the intensity of the peak in the off-resonance and I_{STD} corresponds to the intensity of the peak in the on-resonance segment of the experiment [179].

2.10.3.2.2 Practical aspects of STD experiments

All samples for STD experiments were prepared in buffer containing 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4 in 100% D_2O at 298 K otherwise indicated. The ligand/protein ratio was kept at 100/1, with 1 mM ligand concentration and 10 µM protein concentration. stddiffesgp.3 was used for STD experiments [183]. This pulse is a pseudo 2D sequence for the saturation transfer difference using a Gaussian pulse train with a length of 49 ms and 50 dB of attenuation with an interpulse delay of 1 ms for saturation on f2 channel alternating between on and off resonance. The on-resonance frequency was set to 0.7 ppm and the off-resonance frequency was set at 40 ppm with a saturation time of 3 s and 64 scans. Water suppression was achieved using excitation sculpting with gradient and a spinlock was employed before acquisition to suppress protein signals

2.10.3.3 Water-ligand Observed via Gradient Spectroscopy (WaterLOGSY)

2.10.3.3.1 General aspects of WaterLOGSY experiment

The waterLOGSY experiment makes use of the different relaxation pathways that molecules with different tumbling times undergo, to discriminate bound from free ligands. The experiment is triggered by selectively saturating the water signal. This saturation leads to buildup of NOE signal in proximity ligands in solution, ligands bound to the protein and the protein itself. Water molecules on the surface of the protein have a much slower tumbling rate; hence ligands close to it tumble at a much slower rate compared to the free ligands. While the free ligands mainly relax through double-quantum pathway the bound molecules go through zero-quantum relaxation leading to opposite phased resonances in the NMR spectrum [179, 184].

2.10.3.3.2 Practical aspects of WaterLOGSY experiment

All samples for WaterLOGSY experiments were prepared in buffer containing 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4 and 10% v/v D_2O at 298K otherwise indicated. The ligand/protein ratio was kept at 100/1, with 1 mM ligand concentration and 10 µM protein concentration.The ePHOGSY pulse sequence with 128 scans was used for WaterLOGSY experiments. Selective water saturation was achieved using a 7.5 ms 180° selective Gaussian pulse. Solvent suppression was accomplished using two 2.5 ms 10% truncated Gaussian pulses. The relaxation and mixing times were 2.5 and 1.2 s, respectively.

2.10.4 Measuring K_D by NMR spectroscopy using 2D ^{1}H-^{15}N and ^{1}H-^{13}C NMR

The dissociation constant, K_D, is calculated through the concentration of the complex and free ligand and protein at equilibrium:

$$K_D = \frac{[P][L]}{[PL]} \qquad (3)$$

Since the total protein and ligand concentration are known, it is more practical to rearrange equation 3 using equation 4 and 5, yielding equation number 6 [185]:

$$P_T = [P] + [PL] \quad (4)$$

$$L_T = [L] + [PL] \quad (5)$$

$$[PL] = \frac{1}{2}\left[(K_D + L_T + P_T) - \sqrt{(-K_D - L_T - P_T)^2 - 4L_T P_T}\right] \quad (6)$$

In an interaction following fast exchange principles according to NMR time scale, K_{off}>>Δδ, the change in chemical shift (or any other NMR observable) is related to the fraction of bound protein and the maximum change which is obtained by subtracting the chemical shift of the fully saturated from the free state [185]. This data can be used to rearrange equation 6 into an equation that can be used to extract K_D using a non-linear regression fitting [185].

$$\Delta\delta_{\text{obs}} = \frac{\Delta\delta_{\max}}{2[P_T]}\left[(K_D + L_T + P_T) - \sqrt{(-K_D - L_T - P_T)^2 - 4L_T P_T}\right] \quad (7)$$

For K_D calculations using protein observed NMR, two samples were prepared with protein concentration set at 50 μM, one with no ligand and the other sample with the maximum soluble concentration of the ligand a buffer composed of 150 mM NaCl, 5 mM $MgCl_2$, 20 mM HEPES at pH 7.4, 10% v/v D_2O and final concentration of 5% v/v d_6-DMSO. Spectrums for the two samples were acquired according to section 2.10.2. Subsequently the sample with maximum soluble ligand was diluted using the other sample stepwise and spectrums were collected. Peaks with most chemical shift change in the ^1H dimension were chosen for analysis and K_D was calculated using a non-linear regression fitting to equation 7 using the Prism 7 software.

2.10.5 Data processing and analysis

All 1D experiments were processed and visualized using the Bruker TopSpin3.5 program. All 2D spectrums were processed using NMRPipe [186] and visualized by NMRView program [187].

2.11 Ligand aggregation assay

Ligand aggregation was evaluated using 1D NMR spectroscopy [188]. 100 mM d_6-DMSO stocks of organic molecules were prepared. 1.5 mL of 0.5 mM solution of each ligand in a buffer containing 150 mM NaCl, 5 mM $MgCl_2$, 20 mM

HEPES at pH 7.4, 10% v/v D_2O and final concentration of 5% v/v d_6-DMSO was prepared. Next each tube was centrifuged at 14000 RPM using a bench top centrifuge for 5 min. 490 µL was transferred into an NMR tube representing the highest soluble concentration. Another 490 µL was taken into another microfuge tube and mixed with 10 µL of tween 80 to yield a final concentration of 0.025% v/v TWEEN^R 80. This sample was transferred to another NMR tube representing the highest soluble concentration with detergent. Finally another 490 µL under went serial dilution four times with the same buffer containing 5% v/v d_6-DMSO. All samples were transferred to NMR tubes. 1D spectrums were collected with 64 scans at 298 K and processed using TopSpin 3.5 software.

2.12 X-ray crystallography

All proteins used for crystallization were prepared according to section 2.3.1 in a buffer containing 150 mM NaCl, 5 mM $MgCl_2$ and 20 mM Tris at pH 8. All Rab27a constructs fused to a segment of Slp2-a RBD were concentrated to 15 mg/mL for crystallization purposes. All MJRab27a constructs series or Rab27a constructs fused to other tags than Slp2-a were concentrated up to 25 mg/mL for potential crystallizations. Screens to find initial hits were carried out using the commercial screening plates available at Imperial College London X-ray suite. Details for these plates are available at http://www.imperial.ac.uk/x-ray-crystallography/crystallisation /crystallization-screens. Commercially available plates from Hampton research were utilized for additive and detergent screening. All screening and optimization of crystals were achieved using a sitting drop vapor diffusion method, adding 200 nl of mother liquor to 200 nl of protein in 96 well and 48 well MRC MAXI plates from Molecular Dimensions. All protein handling was carried out by Mosquito crystallization robot from ttplabtech. All Slp2-a fused Rab27a constructs were crystalized at 4°C, with crystals appearing after 12 hours and reaching maximum diameters in 4 days. MJRab27a series were crystalized at 20°C, with crystals appearing from 4 days to 3 months depending on the construct. Data collections were carried out at Diamond Light Source in Oxford using the i02, i04 or the i04-1 beamline at 100K, wavelength 0.9795 Å, using Pilatus detector. Data were collected at 0.2°-0.5° oscillations per image with a total oscillation up to 200° depending on the crystal space group. Data processing was carried out using Dials [189]. For MJ1Rab27a, MJ1-1Rab27a and MJ4-1Rab27a crystals, resolution bins with $CC_{1/2}$ above 0.5 and $\langle I \rangle / \langle \sigma(I) \rangle$ above 2 were used to obtain maximum

acceptable resolution. Initial phases were calculated using the molecular replacement program phaser. The coordinates of Rab27a from chain A of the Rab27a-Slp2$_a$ complex (PDB:3BC1) without the nucleotide and the magnesium ion were used as the search model [190]. Subsequently, the initial model generated by phaser was refined through an iterative cycle using COOT [191] and REFMAC5 [192]. Final model structures were validated using the Molprobity server [193] at http://molprobity.biochem.duke.edu. All structure images were prepared using Pymol [194]. Protein-ligand interaction diagram was generated using Maestro software from Schrodinger.

For Rab27a-ligand co-crystals, Slp2$_a$-Rab27a fusion crystals were soaked with these compounds at saturating concentrations in the presence of 5-10% DMSO for 24-48 hours. For model building, ligands were only modeled into the electron density after the protein had been refined to avoid any conformational bias, although the electron density for them would have been apparent from the primary model obtained by phaser. The ligands were built using COOT by providing the smiles string for each ligand. The CIF file containing the torsion and geometric restrains for each ligand was generated using Refmac5. The occupancy was set at 0.5 for atoms in ligands that were potentially present in two conformations.

2.13 Computational screening of ICR vendor collection

205185 compounds were retrieved from the ICR library collection. The compounds were filtered based on chemical properties including H-bond acceptors ≤10, H-bond donors ≤5, 20 ≤ non-H atom ≤40, polar surface area ≤140 Å^2, rotatable bonds≤10, 180≤MW≤500 and -5≤cLogP≤6 using Canvas from Schrodinger [195]. The remaining 177256 compounds were used as input for Omega2.5 from the OpenEye package for 3D conformer generation [196]. Torsional sampling was 5^o-60^o and a maximum of 200 conformers within 5 kcal/mol of the lowest energy conformer were generated with no limits on the number of stereocenters within the compound. Next, the shape of the conformers were compared to the query molecule which is the conformation of the “WF” peptide within the Slp2-a RBD in PDB code 3BC1 using Rapid Overlay of Chemical Structures (ROCS) software as part of the OpenEye package [197]. All conformers from each molecule were overlaid through the center of mass on the query molecule. The ROCS software overlays molecules to achieve maximum shape overlap for each conformer. After finding the best

orientation, two scores were given. One score was given for volume overlap or shape similarity and one separate score was given to the pharmacophore similarity named as the color force field. Therefore, when selecting the query molecule it is important to consider the pharmacophores important for the query molecule. The best conformer for each molecule was ranked using TanimotoCombo, which is the combination of Shape score and pharmacophoric features score using the Tanimoto metric after overlay of compounds solely based on their shape. 1117 Compounds with TanimotoCombo ≥ (2×Average+2SD) and ShapeTanimoto ≥ 0.75 were taken for screening by SPR. A score was also given using the Tversky metric for comparison. Equation 10 and 11 describes measuring similarity using the Tanimoto and Tversky metrics [198].

$$Similarity_{Tanimoto} = \frac{\sum_{i=1}^{N} x_{iA}\, x_{iB}}{\sum_{i=1}^{N}(x_{iA})^2 + \sum_{i=1}^{N}(x_{iB})^2 - \sum_{i=1}^{N} x_{iA} x_{iB}} \qquad (10)$$

$$Similarity_{Tversky} = \frac{c}{\alpha(a-c) + \beta(b-c) + c} \qquad (11)$$

CHAPTER THREE: DESIGN AND CRYSTALLIZATION OF RAB27A CONSTRUCTS

3.1 Introduction

The use of structural biology as a tool in drug discovery goes back to the 1980s, when several relevant drug targets such as HIV proteases were crystallized and interactions with substrate-based inhibitors were studied [199]. Over the years, the generation of crystal structures for clinically important proteins such as angiotensin-converting enzyme [200], carbonic anhydrase [201], neuraminidase [202] and thrombin [203] have facilitated the production of pharmaceutical drugs. Automation of several aspects of protein crystallography has had significant impact on its use in pharmaceutical companies. High-throughput liquid handling robots that can dispense volumes as small as 25 nl and premade crystallization screens in 96 well plates have enabled the search for new protein crystals to be accelerated [204]. Robust and precise robots such as ACTOR^{TM} have enabled researchers to mount and align crystals in the X-ray beam in an automated way, paving the path for screening hundreds of crystals [205]. Data processing and analysis has been facilitated through several software packages and automated pipelines at synchrotron facilities [206]. The overall automation of the whole process is especially vital when working with hard to crystallize proteins, such as membrane proteins, that might need to screen through hundreds of crystals to obtain a complete dataset [207]. Apart from studying protein-ligand interactions, structures have been used to obtain information on potential druggable binding sites. The structures are used in designing computational programs that exploit accessible pockets. The information obtained be used to guide virtual screening cascades that involve docking or molecular dynamics [208].

Apart from the obvious importance of structural biology and especially high resolution structures obtained by crystallography in drug discovery, disappointments from high-throughput screening (HTS) campaigns and the up rise of fragment-based drug discovery (FBDD) as a novel way of ligand discovery have brought structural biology more in the spotlight making it a first line screening tool for drug discovery [209]. Several companies such as Astex and Vernalis have used crystallography to screen against libraries of up to 2000 fragments enabling them to optimize weakly binding molecules into potent inhibitors [210]. This has the advantage of selecting hits based on binding modes rather than affinity. Also, the

detection of binders through the presence of electron density for potential hits accelerates the chemistry follow-up and validation through building a structure-activity relationship (SAR) around the hit. This is important for very weak binders used at millimolar concentrations that are maybe more prone to non-specific binding and false positives [204, 211].

Apart from crystallography, NMR has been another key tool in studying protein ligand interactions [212]. In fact, a Bcl_{xl} inhibitor was designed solely through an NMR strategy [213]. The broad range of progress made from labeling strategies, stronger magnets, the optimization and development of new pulse sequences have enabled researchers to use NMR as a tool for screening orthogonal hit validation. It has been used to generate structural models when a crystal structure was not attainable [214]. All together, structural biology has transformed drug discovery and become the heartbeat of the industry for the last two decades.

Since there are no published crystal structures of Rab27a on its own, one of the aims of this project has been the design of crystallisable Rab27a constructs that can be used to study the interaction of ligands obtained through ligand screening. A crystal suitable for drug discovery has certain specific criteria that might not be particularly important, if you are only interested in the structure of the protein from a biochemistry perspective. This chapter will focus on the different approaches that have been taken in order to obtain Rab27a crystals suitable for our purpose. Our target was to engineer a protein that can generate reproducible single crystals that do not suffer from any known pathologies which might occur in protein crystals, are resistant to a low percentage of organic solvents such as DMSO and other organic molecules, grow in a reasonable amount of time and diffract to a good resolution whilst leaving the pockets or surface of interest exposed for interaction studies. Un-obstructing the target pockets and surface can be a serious challenge when targeting protein interaction surfaces. Valkov et al. have neatly discussed this issue: they suggest that although due to the exposure of the surface one would think there is a higher chance for soaking the site with a ligand compared to enzyme active sites as it is more accessible, the large surface of contact up to 3000 Å means that there is a high chance a big portion of the surface will be involved in making crystal contacts, hence obstructing the surface of interest (Figure 3.1) [215]. Therefore, it is highly recommended to obtain multiple crystals in different crystal forms.

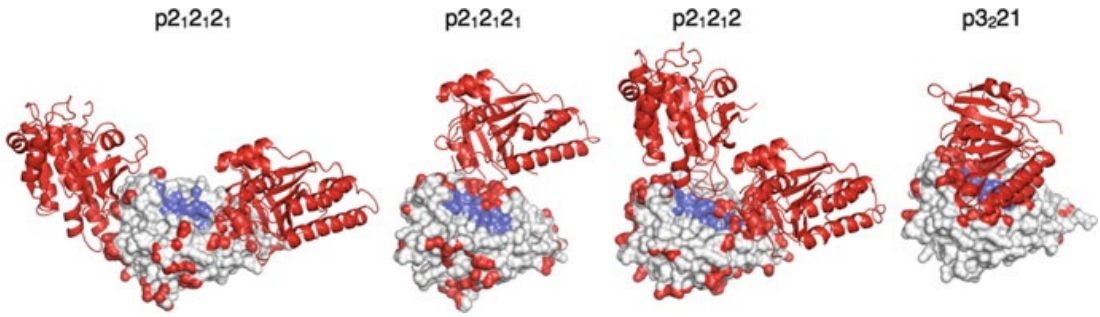


Figure 3.1. Crystallization of RadA mutant in different crystal forms. RadA interaction interface is colored in blue and residues involved in crystal contacts are colored in red. $P2_12_12_1$ space group has the least occluded binding site. Image was taken from [215] with permission. Permission is attached in appendix G.

3.2 Strategies for Rab27a crystallization

From the start of the project to March 2017, two structures have been deposited to the protein data bank (PDB), which correspond to Rab27a-$Slp2_a$ (3BC1) and Rab27b-$Slac2_a$ (2ZET). Chavas et al. had also previously reported their failed attempt to crystallize Rab27a-(gppNHp) without binding to its effector [216]. My own attempts to crystallize the protein at both 4 and 20^{o}C using the screening plates available at various concentrations ranging from 15-30 mg/mL did not yield any results. Moreover, tagging Rab27a with common proteins used to enhance the crystallization, such as maltose binding protein (MBP) (that have been optimized with surface entropy reduction (SER) mutations) did not help [217]. Therefore, I took two strategies:

1- To see whether there are possibilities to use the information from the reported Rab27a/b complex structures to generate fusion proteins between the protein and its effector, which still leave potential areas of the interaction interface unoccupied. These would potentially generate a construct that not only would crystallize but still be useful for studying the interaction of potential ligands from screening campaigns against Rab27a. This approach is described in detail in section 3.3.
2- To study the crystal contacts present in crystal structures of Rab27a homologues and evaluate the possibility of making certain mutations in Rab27a that would enable the protein to crystallize in the absence of its effector. These mutations would have to avoid regions that are important for

effector binding or at least the regions that are involved for ligand binding studies. Unlike the first approach, constructs from this strategy would potentially benefit from the absence of having a tag (which can be a segment of an effector or another protein domain), which can influence the binding mode of the ligands. This method is described in detail in section 3.4.

3.3 Design of fusion constructs

Because the crystals obtained through either fusion or mutagenesis approaches will be used for ligand interaction studies, it is important to leave hot spot regions of the interaction interface, that are important for effector binding and also attractable to ligands, unoccupied. Therefore, prior to our construct design, we initially evaluated Rab27a-Slp2$_a$ interaction interface. Slp2$_a$ Rab-binding domain (RDB) is made of α_1 (RBDα_1) and α_2 (RBDα_2) helical segments. RBDα_1 is a long helical segment covering the majority of interaction interface, which is flat, featureless, dominated by hydrophobic residues and unsuitable for specific ligand binding (Figure 3.2). In contrast, RBDα_2 has a gaff-like structure and uses a highly conserved tryptophan-phenylalanine (WF) motif to interact with the shallow SF4 pocket (Figure 3.2). Mutagenesis studies have revealed the importance of this motif for the Rab27a-Slp interactions [137].

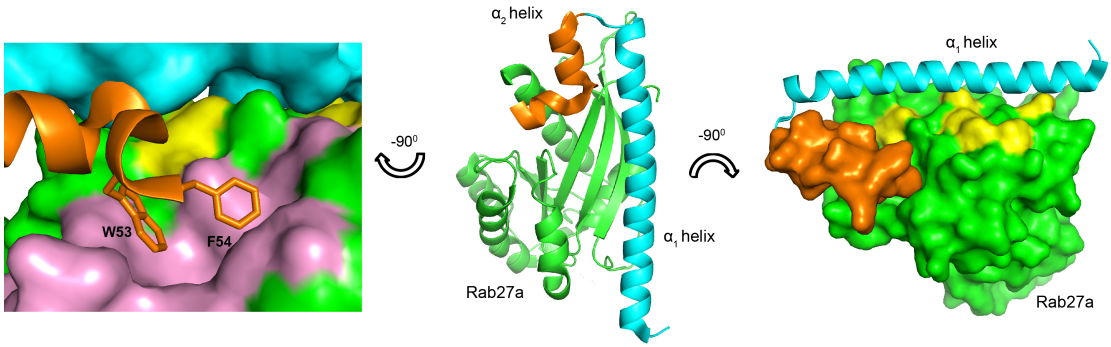


Figure 3.2. Rab27a-Slp2a interaction. Interaction between Rab27a (green) and Slp2$_a$ Rab binding domain (RBD) (cyan-brown) is shown. Slp2$_a$ RBD is composed of two helical segments RBDα_1 (cyan) and RBDα_2 (brown). RBDα_1 (cyan) interacts with a flat, featureless surface of Rab27a dominated by hydrophobic residues (yellow), while RBDα_2 (brown) interacts through a conserved motif of "WF" to the SF4 pocket (pink) in Rab27a.

Also, we assessed the ligandability of Rab27a using the FTMap web server [208]. This method determines the binding energy of molecules with different size, shape and polarity to identify binding pockets in a protein surface. Pockets that bind several molecules with a high rank are considered "consensus sites" or the "hotspots".

Analyzing the pockets in Rab27a structure in the absence of the effector suggest that after the nucleotide-binding pocket, the SF4 region is probably the most ligandable site on the Rab27-effector interface (Figure 3.3). Taking all these data into consideration we planned to design our constructs in a manner that leaves the SF4 pocket empty.

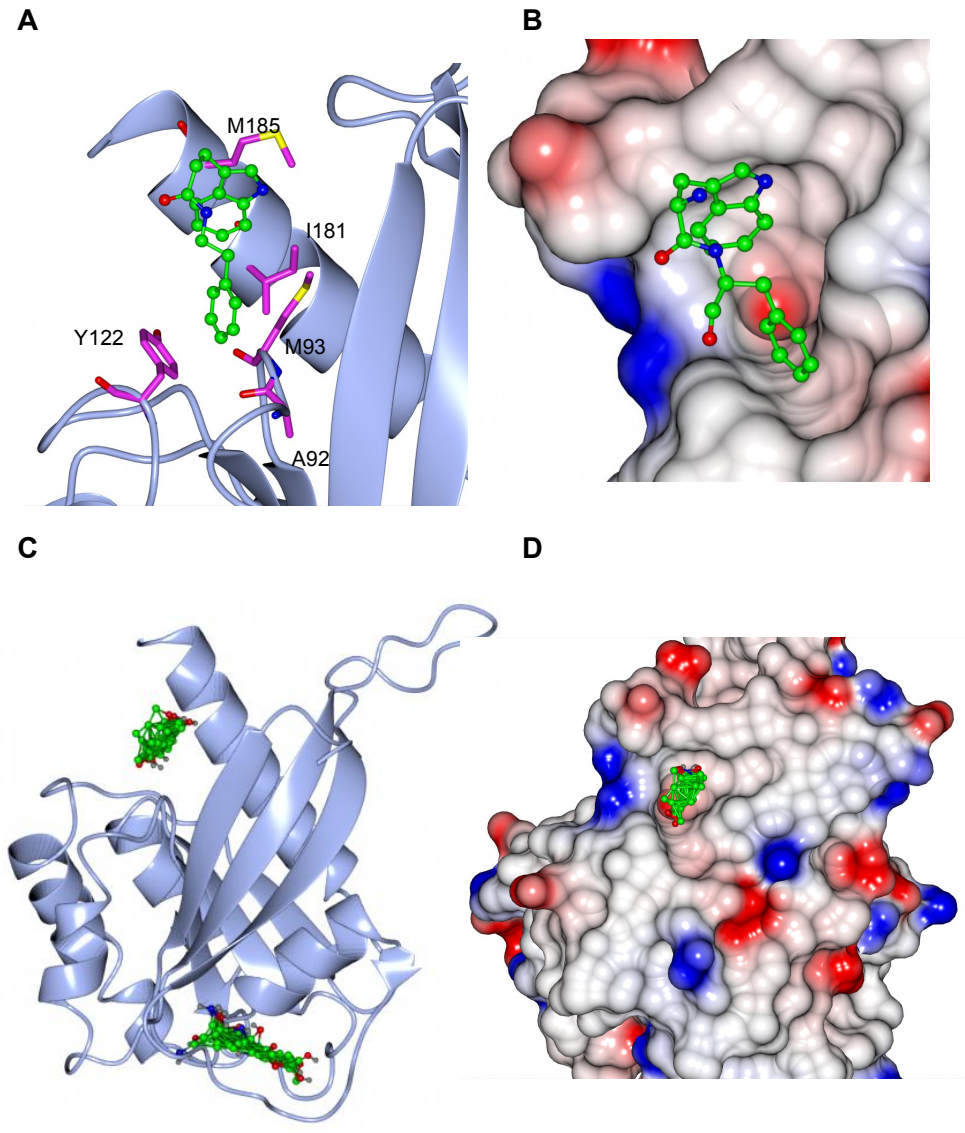


Figure 3.3. Interaction of the WF motif with SF4 pocket of Rab27a. A) green ball and stick represents WF motif while side chain residues A92, M93, Y122, I181 and M185 of Rab27a that interact with the WF motif are highlighted in purple B) Demonstrates the WF motif occupying the SF4 pocket. C) and D) depict results obtained from the FTMAP web server. Cluster of organic molecules occupying the nucleotide binding pocket and the SF4 are indicative of them being as hot spots.

Examining the Rab27a-Slp2$_a$ crystal structure, suggested that RBDα$_1$ of Slp2$_a$ is extensively involved in crystal contact formation and packing of the crystals, whether

it is between the two molecules in each asymmetric unit or the closely crystallographic symmetry related molecules. This suggested that any kind of fusion between RBDα$_1$ and Rab27a would facilitate crystallization (Figure 3.4).

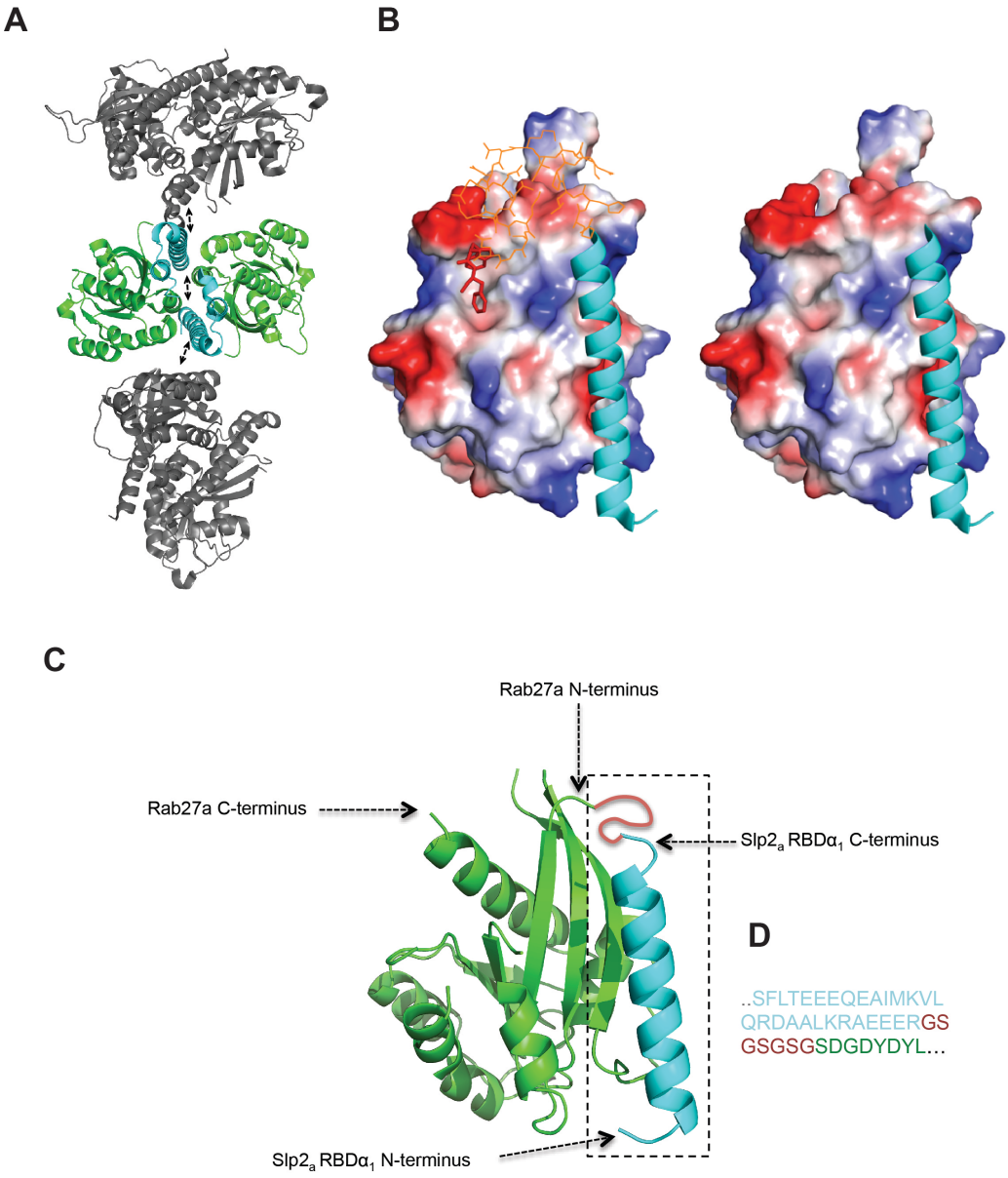


Figure 3.4. Design of the Slp2$_a$-Rab27a fusion construct. A) As demonstrated the two Slp RBDs (cyan) are involved in the formation of crystal contacts in each ASU and between symmetry related molecules (gray)(interaction shown using arrows) B) Surface representation of Rab27a bound to Slp2$_a$. RBDα$_1$ is presented in ribbon (cyan) while RBDα$_2$ is presented in lines (orange). The WF motif in RBDα$_2$ is highlighted in red. As demonstrated, a potential fusion between Rab27a and RBDα$_1$ would leave the SF4 unoccupied suitable for protein-ligand interaction studies. C) RBDα$_1$ of Slp2$_a$ (cyan) was connected to the N-terminus of Rab27a using a glycine-serine linker (red). D) The sequence of RBDα$_1$ used in the construct (cyan) followed by the linker (red) and the initial residues of Rab27a (green) are shown.

Moreover, due to the exclusion of RBDα$_2$, such a construct would also leave the important SF4 region accessible for occupation by ligands targeting this pocket. Designing such a construct was simplified by the fact that the N-terminus of Rab27a is in the vicinity of the C-terminal end of RBDα$_1$. This enabled us to fuse a truncated form of Slp2$_a$ containing RBDα$_1$ to the N-terminal part of Rab27a through a linker (Figure 3.4), as demonstrated by a 30a-aminoacid residue segment containing the RBDα$_1$, was fused to the N-terminal part of Rab27a using a glycine-serine linker (GSGSGSG). This protein was expressed, purified and used for crystallography purposes.

3.3.1 Slp2$_a$-Rab27a fusion crystallization

Crystallization screenings for the protein was set according to methods described in section 2.3.1 and 2.12. Unlike wild type Rab27a, the Slp2$_a$-Rab27a fusion crystallized across multiple plates in a variety of conditions at 4 ^{o}C. Three conditions with different habits labeled as A, B and C were taken forward for further optimization. The conditions for the crystallization of the crystals were optimized using a 0.3 pH unit gradient against 2% gradient of precipitant. This optimization led to fewer, larger and more reproducible crystals. The conditions obtained from the optimization and crystals resulting from these experiments are depicted in figure 3.5.

A	**B**	**C**
0.1 M Ammonium sulfate, 0.3 M Sodium formate, 3% w/v γ-PGA, 3% w/v PEG 20000 and pH 5 Cryo: Saturated formate Not DMSO resistant	0.1 M $LiCl_2$, 0.1 M HEPES, 20% w/v PEG400 and pH 7.5 Cryo: 25% glycerol Not DMSO resistant	20 mM $MgCl_2$, 50 mM bis-Tris, 15% 2-propanol and pH 6.8 Cryo: 25% glycerol DMSO resistant

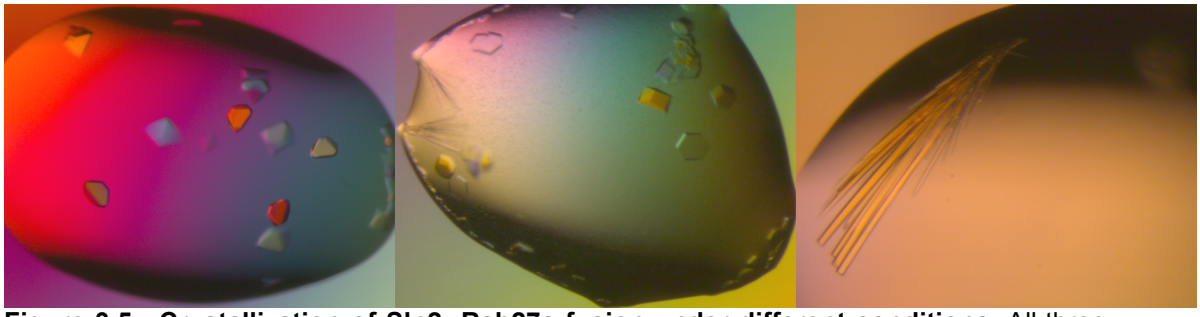


Figure 3.5. Crystallization of Slp2$_s$-Rab27a fusion under different conditions. All three conditions in A, B and C were optimized to find the correct buffer and precipitant conditions. All conditions were evaluated for resistance to organic solvent and only condition C was resistant to DMSO.

3.3.1.1 Slp2$_a$-Rab27a data collection and processing and refinement

The crystals for all three conditions were cryo-protected and shipped to the Diamond light source for data collection. Data were collected according to section 2.12. Crystals grown in conditions A and B only diffracted close to 3 Å while crystals in condition C, which was also resistant to up to 10% v/v DMSO, diffracted to 2.3 Å (Table 3.1). Therefore, the dataset obtained from condition C was taken forward for further analysis. All data analysis was carried out according to section 2.12. A single solution was obtained from phaser using the $P2_12_12_1$ space group. The R_{work} and R_{free} crystallographic factors for the final refined model were 0.19 and 0.22 respectively. All the statistics for the collected dataset are shown in table 3.2.

Table 3. 1. Unit cell information from diffraction of Slp2$_a$-Rab27a in different conditions

	Cell dimensions	Space group	ASU$_n$	Resolution
A	117.69 117.69 115.89 90 90 120	$P3_121$	2	3.07 Å
B	117.71 117.71 115.67 90 90 120	$P3_121$	2	3.11 Å
C	62.11 77.61 119.21 90 90 90	$P2_12_12_1$	2	2.32 Å

ASU$_n$: number of molecules in each asymmetric unit

Table 3. 2. Data Processing and Refinement Statistics for Slp2$_a$-Rab27a fusion

Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(Å) a = 62.11, b = 77.61, c = 119.21
Wavelength (A°)	0.97949
Resolution (A°)	58.85 - 2.32 (2.403 - 2.32)
Total reflections	437706
Unique reflections	40522
Multiplicity	10.8(8.9)
Completeness (%)	100.0(99.7)
<I>/<σ(I)>	19.5(3.1)
R-merge	0.092(0.67)
R-meas	0.102
Wilson B factor	36.39
$CC_{1/2}$	0.999
Refinement	
Reflections used in refinement	40483 (3976)
Reflections used for R_{free}	1945 (229)
R_{work} (%)	19.05 (24.20)
R_{free} (%)	22.42 (29.32)
Rmsd bond lengths (A°)	0.014
Rmsd bond angles (°)	1.41
Average B factors (A^{o2})/Number of atoms	
Macromolecules	38.3 /3280
Water molecules	36.28/238
Ligand non-H atoms gppnhp-Mg^{2+}	24.88/66
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	98.2
Ramachandran allowed region (%)	1.8
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.29

3.3.1.2 Slp2$_a$-Rab27a fusion structure analysis

The fusion protein orders in the crystal in a very similar manner to the complex structure (PDB:3BC1). Each asymmetric unit (ASU) is made of two fusion molecules related to each other by a non-crystallographic symmetry (NCS) two-fold axis (Figure 3.6). The electron densities are observable for the nucleotide, magnesium ion and the water molecules conserved in the binding pocket in the initial pre-refinement model. The backbone rmsd of Rab27a molecules between the two structures is at 0.21 Å. Each RBDα$_1$ in the fusion construct brings both Rab27a molecules together in each ASU (Figure3.6). Also, as expected this domain interacts with other Rab27a molecules in other ASUs, hence facilitating the formation of the crystal contact network.

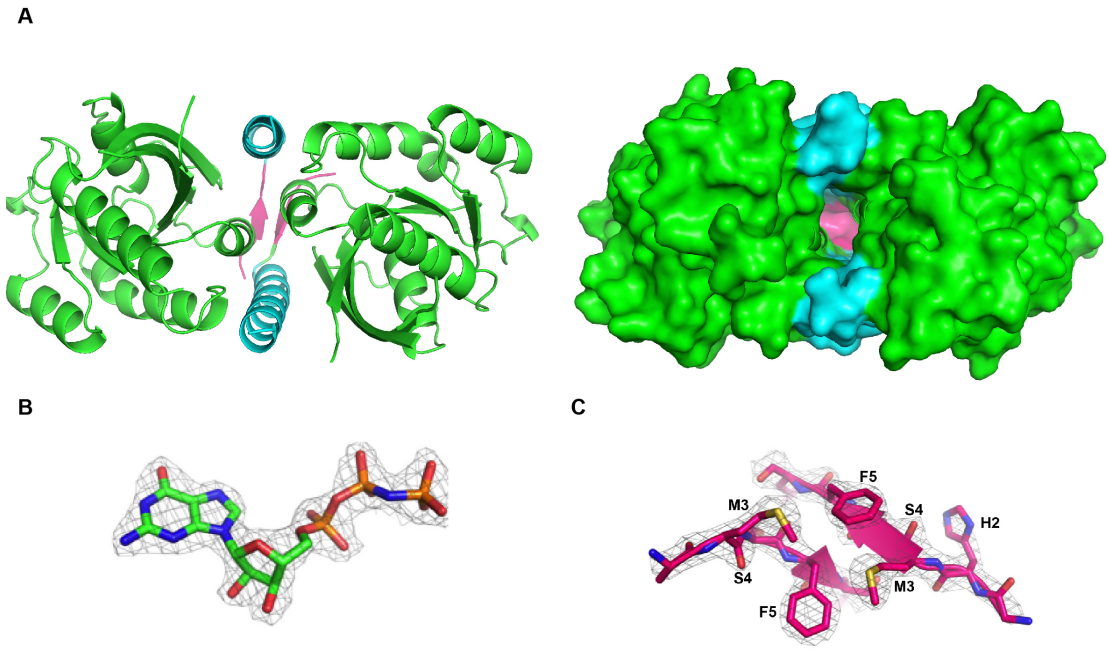


Figure 3.6. Slp2$_a$-Rab27a fusion structure analysis. A) Two Rab27a molecules (green) pack against each other and Slp2$_a$ RBDα$_1$ (cyan-pink) domains in the ASU. B) The electron density is completely present even for the nucleotide analogue bound to each Rab27a molecule in the 2F0-FC map. C) The five residues at the N-terminus of Slp2$_a$ RBDα$_1$ are highlighted in pink. These residues make several key interactions with their NCS related residues. As shown in A), the packing of two fusion molecules in each ASU generates an artificial pocket on the NCS lattice.

A key difference observed in this structure compared to the structure of Rab27a-Slp2$_a$ RBD complex (PDB:3BC1), is the formation of a beta sheet between the N-terminal segments of RBDα$_1$ (Figure 3.6) This difference is due to sequence

dissimilarity between our construct and the construct used for the complex structure. The construct used for the complex, had several residues originating from the thrombin enzyme-cleavage site. On the other hand, our construct has the five residues “HMSFL” at the N-terminus. “HM” are artifacts of the vector after TEV enzyme cutting site, but “SFL” are residues present in human Slp_{2a} sequence. These three residues form a beta strand under crystallographic conditions and are involved in making several contacts with their NCS related residues from the other RBDα_1 in the ASU (Figure 3.6). As a result of this interaction, a channel is formed between the two molecules in each ASU (Figure 3.6). This environment later proved to be problematic for ligand interaction studies (chapter 4).

Another difference between the fusion and complex structures is in the α_3-β_5 loop. This loop forms the SF4 pocket and is in direct contact with the “WF” motif of $Slp2_a$ RBDα_2 in the complex. The absence of RBDα_2 results in backbone and side chain variation in the α_3-β_5 loop between the two structures. The backbone rmsd of the loop in the two structures is 0.7. Additionally, Y122 side chain is completely flipped and is in contact with R90. This is in contrast to the Rab27a-$Slp2_a$ complex (PDB: 3BC1), were Y122 interacts with R184 (Figure 3.7). Moreover, detailed analysis of the fusion structure revealed that Y122 and R90 on chain A are in contact with D21 on RBDα_1 of chain B. This suggests that the RBDα_1 fused to Rab27a could be biasing the orientation of Y122 side chain on Rab27a under crystallographic conditions (Figure 3.7). This leads to a change in the shape of the SF4 pocket, which can affect the binding of ligands. Since this construct is used to study the binding mode of ligands to wild-type Rab27a, it is imperative to keep a similar model to the wild-type protein. Therefore, we mutated the D21 to A21 to see if we can remove this potential artifact from our construct. The new construct crystallized under similar conditions and produced crystals that would diffract to similar resolution (Appendix C) (Figure 3.7). Interestingly, we could not identify any electron density for both R90 and Y122 side chains. This implies that Y122 and R90 side chains are flexible and dynamic in the absence of an effector and the previous construct was biasing the orientation of Y122 side chain into a specific direction (Figure 3.7).

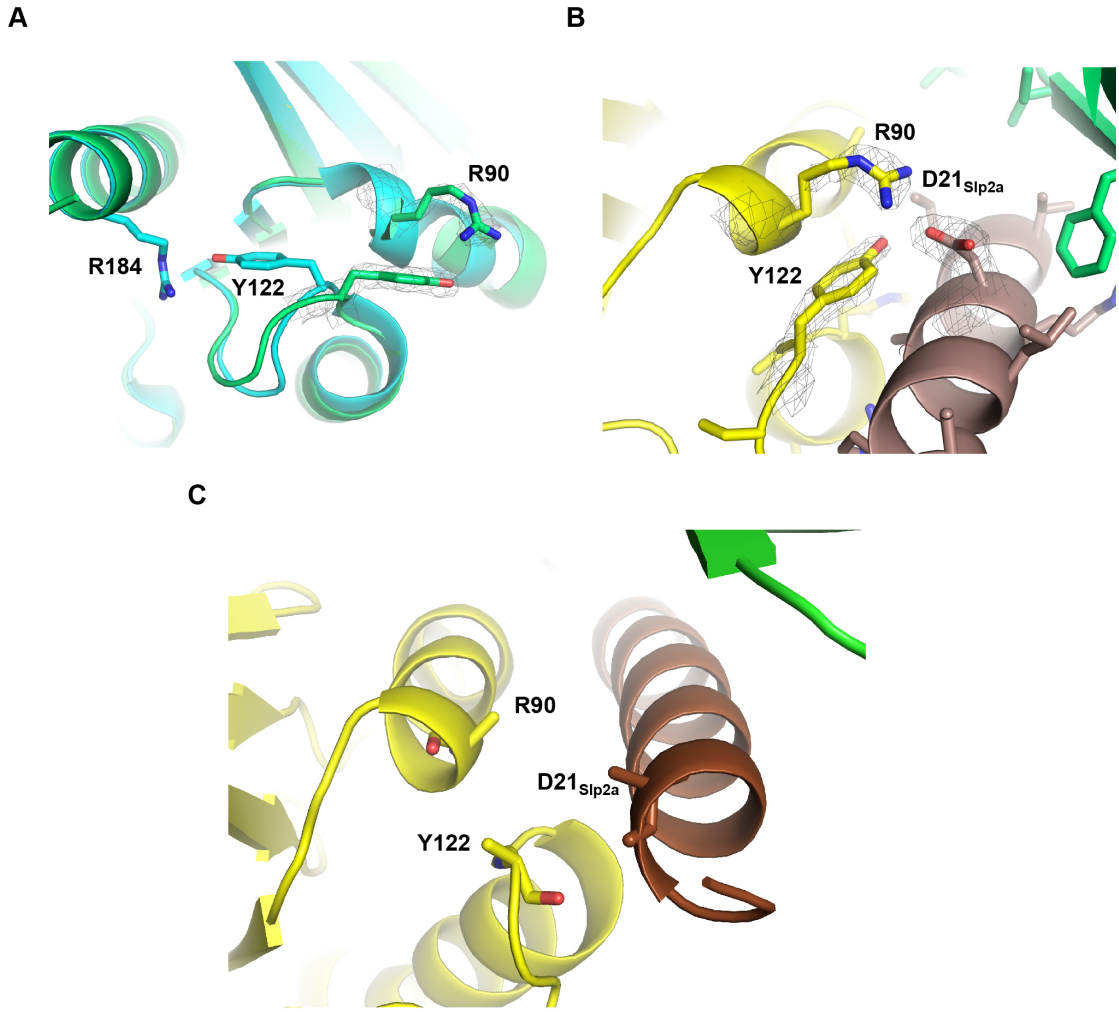


Figure 3.7. Y122 side chain orientation in $Slp2_a$-Rab27a fusion protein. A) Y122 orientation in the fusion construct (green) is 180° different from in wild-type Rab27a (cyan)-$Slp2_a$ complex. B) Y122 and R90 of $Slp2_a$-Rab27a fusion (yellow) in chain A interact with D21 of $Slp2_a$ RBDα_1 (brown) fused to Rab27a in chain B. This interaction is an artifact of crystallization but is having an effect on the characteristics of Y122, which is an important residue forming the SF4 pocket and interaction with the effector. C) No electron density is present for Y122 and R90 side chains in crystals of $Slp2_a$-Rab27a fusion constructs with a mutation on D21 of RBDα_1 of $Slp2_a$.

3.3.2 Designing new Slp-Rab27a fusion constructs

The $Slp2_a$-Rab27a fusion construct enabled us to crystallize Rab27a while leaving the SF4 pocket, which is an important interface for Rab27a-$Slp2_a$ interaction, unoccupied. On the other hand, a large artificial channel was formed between two fusion molecules in each ASU. This could interact with organic molecules irrespective of their true binding sites, therefore affecting the binding site and binding mode of chemical molecules. Moreover, the two-fold NCS symmetry between the two molecules in each ASU would hamper any attempts to study the binding mode

of ligands that have additional pharmacophores interacting with amino acids outside the SF4 pocket. Hence, we tried to design new fusion constructs that would crystallize in a different manner. Analyzing the published structures of Rab27-effectors [136, 137], it was clear that RBDα$_1$ for each effector is a major contributor in crystal packing. Therefore, we postulated that small variations in RBDα$_1$ sequence could affect the fusion crystallization. By changing the sequence of RBDα$_1$ fused to Rab27a, we tried to obtain crystals in different space groups with different number of molecules in each ASU. This would change the exposed or occluded surfaces of Rab27a for ligand interaction studies. Analyzing the sequence similarity of the RBD of Rab27a effectors revealed although RBDα$_1$ sequence is highly conserved between different effectors, there is variation between the residues that are not involved in the interaction between the effector and Rab27a (Figure 3.8). These residues are important in forming crystal contacts under crystallization conditions. In order to evaluate the effect of the α$_1$ sequence on the packing of the Rab27a constructs we initially crystallized the Rab27a-Slp1 complex.

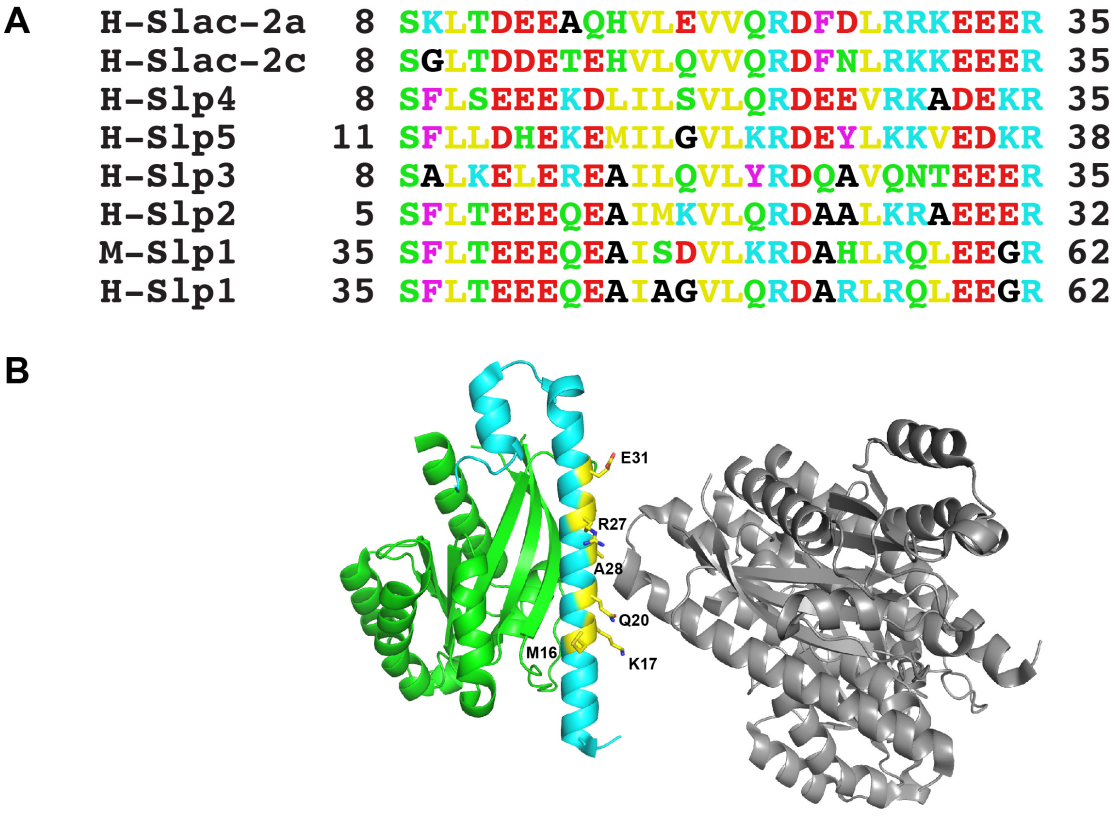


Figure 3.8. Sequence alignment of synaptotagmin-like proteins RBD$_{1α}$. H- is for human and M- is mouse species. As illustrated although RBD$_{1α}$ sequence is conserved, several residues not involved in the interaction with Rab27a make important interactions with symmetry related molecules under crystallization conditions. (amino acid color code: R,K:blue, D,E: red, F,W,Y: magenta, A,G,P: white, V,L,I,M,C: yellow, S,T,N,Q,H: green).

3.3.2.1 Rab27a-Slp1 complex crystal packing and structure

Rab27a-Slp1 complex was purified as explained in section 2.3.3. Crystallization screenings for the complex was set according to methods in section 2.3.1 and 2.12. Single untwinned crystals were formed 100 mM KCl, 20% PEG 4000 and 20 mM Bis-Tris pH 6.6 at 20^0C. All data collection processing, model building and refinement was done according to section 2.12. All the statistics for the collected dataset are in Table 3.3.

Table 3. 3. Data Processing and Refinement Statistics for Slp1-Rab27a complex

Data Collection	
Space group	$P2_12_12$
Unit cell parameters	(A°) a = 59.48, b = 112.64, c = 39.20
Wavelength (A°)	0.92
Resolution (A°)	56.32 - 2.00 (2.0 - 2.07)
Total reflections	131672
Unique reflections	18236
Multiplicity	7.2 (7.5)
Completeness (%)	98.7(98.7)
<I>/<σ(I)>	22.6(3.2)
R-merge	0.047 (0.58)
R-meas	0.055
Wilson B factor	32.49
$CC_{1/2}$	0.999
Refinement	
Reflections used in refinement	18231 (3976)
Reflections used for R_{free}	912 (89)
R_{work} (%)	18.15
R_{free} (%)	21.39
Rmsd bond lengths (A°)	0.014
Rmsd bond angles (°)	1.47
Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	34.09/3721
Water molecules	29.01/213
Ligand non-H atoms gppnhp-Mg^{2+}	24.18/66
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	97.9
Ramachandran allowed region (%)	2.1
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.41

Statistics for the highest-resolution shell are shown in parentheses.

Rab27a-Slp1 complex crystallizes in a $P2_12_12$ space group with only one molecule in each asymmetric unit. This is different from Rab27a-$Slp2_a$ complex and fusion structures that crystallize in a $P2_12_12_1$ space group with two molecules in each asymmetric unit. As a result, Rab27a-Slp1 heterodimer molecule did not have

another molecule packing against it through a two-fold symmetry axis. Therefore, a channel similar to the one seen in $Slp2_a$-Rab27a fusion was not formed. Therefore, we postulated a fusion between Slp1 and Rab27a would also crystallize in similar space groups and provide similar environment for ligand binding studies. Moreover, such a construct would be useful to study the binding mode of ligands that have additional pharmacophores interacting with amino acids outside the SF4 pocket (Figure 3.9).

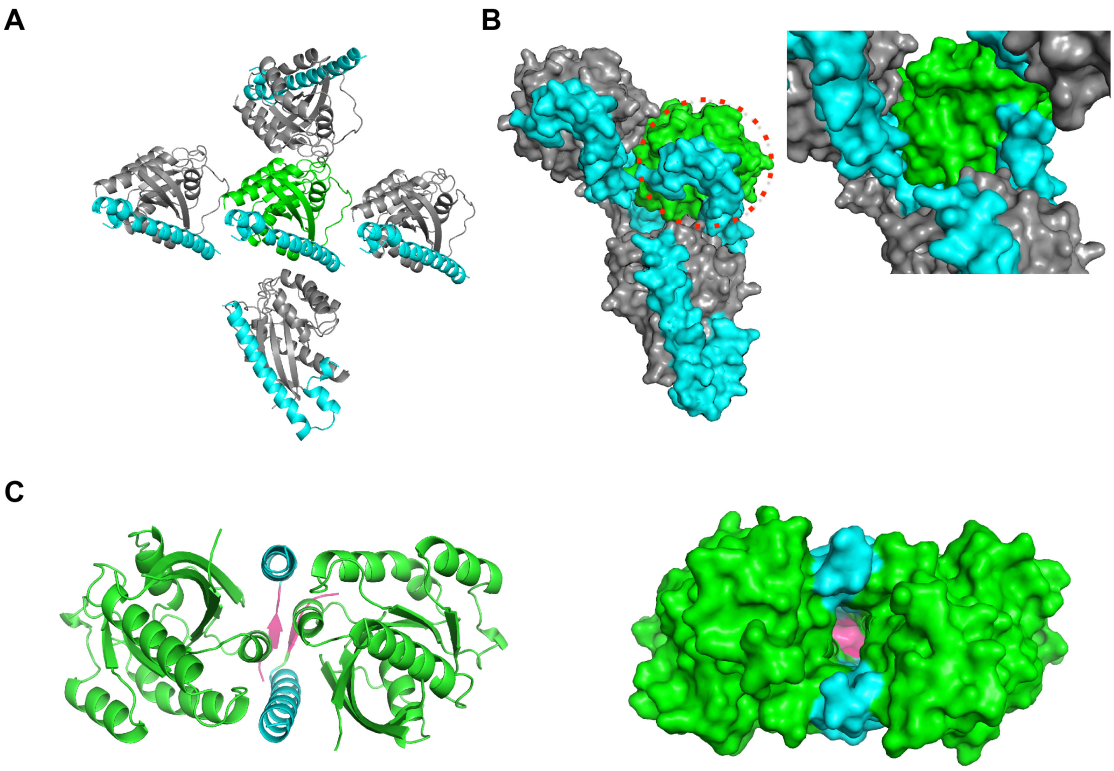


Figure 3.9. Rab27a-Slp1 complex. A) Rab27a (green)-slp1 (cyan) complex is shown. This heterodimer crystallizes different from Rab27a-$Slp2_a$ complex and fusion, were two molecules pack against each other on a two-fold symmetry axis (Rab27a (gray), Slp1(cyan)). B) Surface representation of Rab27a-Slp1 complex in contact with symmetry related molecules. If the Slp1-Rab27a fusion crystallizes in the same manner as the complex (shown by removing $RBD_{2\alpha}$ highlighted by red dotted circle), a more exposed surface area around the SF4 pocket would be accessible for ligand binding studies. C) Rab27a-$Slp2_a$ fusion is shown in ribbon and surface representations. As demonstrated, an artificial channel is formed between two fusion molecules packed against each other in each ASU.

3.3.2.2 Slp1-Rab27a fusion crystallization and structure

Slp1-Rab27a fusion was expressed and purified as explained in section 2.3.3. Crystallization screenings for the complex was set according to methods in section 2.3.1 and 2.12. Single crystals were obtained in different crystallization conditions as shown in figure 3.10. The construct crystallized in three different space groups of

$P3_121$, $P4_12_12$ and $P2_12_12_1$ (table 3.4). Although the Mathews coefficient predicted more than two molecules per ASU, phaser could only find two symmetry related molecules in each ASU in all three space groups. This is in contrast to the Rab27a-Slp1 complex structure, where only one molecule is present in each ASU and lacks a two-fold symmetry-related molecule (crystallographic or non-crystallographic). Slp1-Rab27a fusion crystallized in the same space group and the as the same unit cell as $Slp2_a$ fusion (Table 3.1 and Table 3.4). Therefore, Slp1-Rab27a fusion yielded an identical structural model to Slp2a-Rab27a fusion with similar packing arrangements. This could be due to the importance of $RBD_{\alpha 2}$ of Slp in providing crystal contacts. Therefore this strategy was deemed not useful until we investigate the effect of changes in the length of the protein fusion.

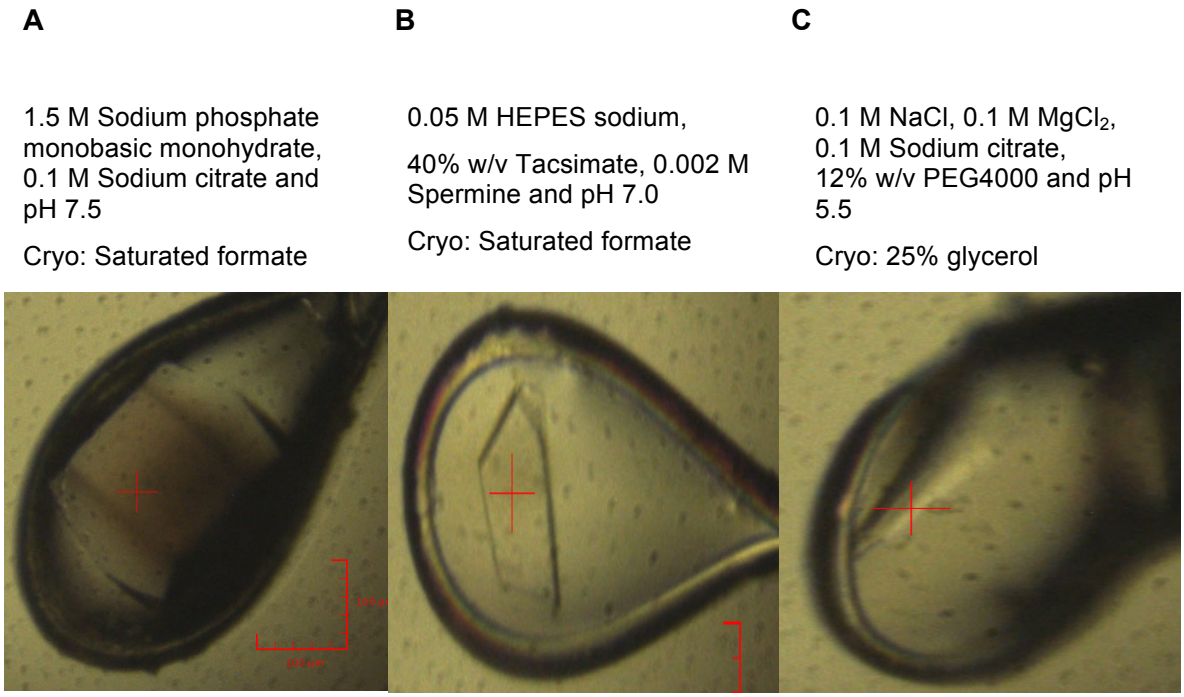


Figure 3.10. Crystallization of Slp1-Rab27a fusion. Fusion construct between Slp1 $RBD_{\alpha 1}$ and Rab27a crystallized in different conditions. Three conditions representing crystals packed in different space groups are shown. All three conditions in A, B and C were optimized to find the correct buffer and precipitant conditions.

Table 3. 4. Unit cell information from diffraction of Slp1-Rab27a in different conditions

	Cell dimensions	Space group	ASU_n	$Math.Coeff._n$	Resolution
A	120.12 120.12 109.83 90 90 120	$P3_121$	2	4	2.3 Å
B	117.71 117.71 115.67 90 90 120	$P4_12_12$	2	3	3.11 Å
C	62.11 77.57 120.51 90 90 90	$P2_12_12_1$	2	3	2.32 Å

ASU_n: number of molecules in each asymmetric unit obtained by phaser,
$Math.Coeff._n$: number of molecules in each asymmetric unit predicted by Mathews coefficient

3.4 Designing crystallizable mutant Rab27a constructs

Apart from the fusion strategy, a mutagenesis approach was taken towards designing crystallizable Rab27a constructs. Unlike the Rab27a-effector fusion, mutant constructs do not cover any part of the effector interaction interface. This is advantageous as no tags that could bias ligand binding are present in the construct. Of course, this implies that mutations have to be limited to residues that are not important for effector binding. All together constructs obtained through this approach can give us extra options to study ligand binding.

Three types of deletions and mutations were used to design the constructs. Initially, loops that are highly flexible and are not important for protein function were removed. Next, Rab3a-GppNHp crystal structure (PDB code: 3RAB) was analyzed and residues that were important for generation of electrostatic crystal contacts in Rab3a were introduced in Rab27a. The idea was that these mutated residues would recreate the contacts that can drive crystal formation in Rab3a. Finally, considering the surface entropy reduction (SER) theory, the sequence of Rab27a was analyzed using the SERp server at http://services.mbi.ucla.edu/SER/ to identify potential clusters of solvent-exposed residues such as glutamates and lysines and replaced by residues with side chains containing small conformational entropy [218]. This would potentially decrease the penalty for the decrease of entropy and provide a more thermodynamically suitable path during crystallization.

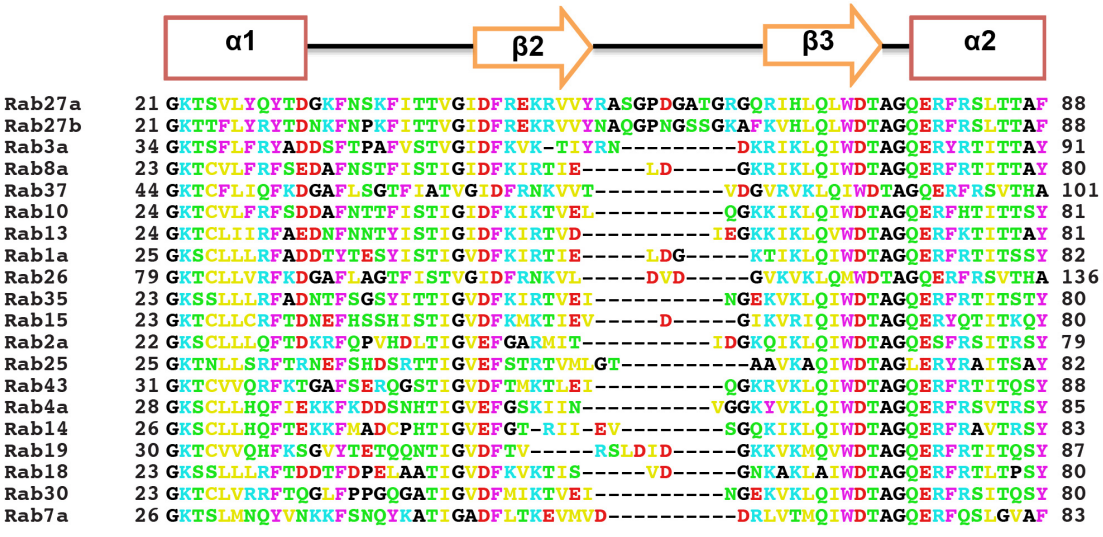


Figure 3.11. Sequence alignment of multiple members of the Rab family in humans. Rab27a,b have a loop between β_2 and β_3 strands that is much longer than other Rabs. This loop was substituted with a shorter sequence present in Rab3a. (amino acid color code: R,K:blue, D,E: red, F,W,Y: magenta, A,G,P: white, V,L,I,M,C: yellow, S,T,N,Q,H: green).

The β_2-β_3 loop in Rab27a (and b) is 12 amino acids long enriched with glycine, proline and serine residues. This loop is much longer than the short loop connecting the two beta strands in other Rab proteins (Figure 3.11 and Figure 3.12). No electron density was recognizable for this loop in neither the fusion or complex structures, indicating its flexible nature. As evident from Rab3a crystal structure (PDB code: 3RAB) the residues in this short loop (D68, K69) are involved in forming multiple interactions with symmetry related molecules (Figure 3.12). Therefore, as our first strategy we decided to replace the loop in Rab27a with the loop in Rab3a.

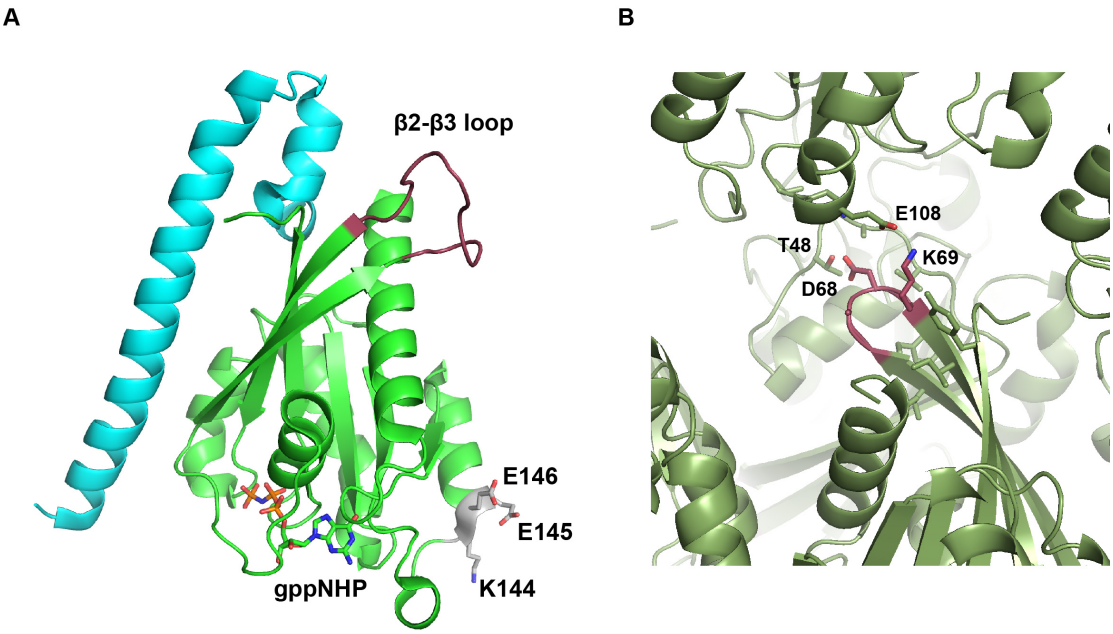


Figure 3.12. Substitution of the β_2-β_3 loop of Rab27a. A) Modeled β_2-β_3 loop (red) in a Rab27a (green)-Slp2$_a$ (cyan) complex structure. This loop was substituted with the shorter loop present in Rab3a (dark green) as shown in B. B) The β_2-β_3 loop in Rab3a (red) is involved in making important interactions with residues from symmetry related molecules in Rab3a crystals. Also as shown in A), K144, E145 and E146 (grey) were mutated to alanines to reduce penalty for conformation entropy reduction during crystallization.

Subsequently we analyzed the crystal structure of Rab3a (PDB code: 3RAB) and identified regions that are involved in making important electrostatic or hydrogen bond interactions with symmetry related molecules. Also, residues on the symmetry related molecules that were interacting with them were identified. Eventually, regions that were not on the effector interaction interface and their interaction partner was already present in wild-type Rab27a were taken for further studies. Instead of mutating a single residue, we tried to mimic regions involved in forming the network of crystal contacts in Rab3a, for Rab27a.

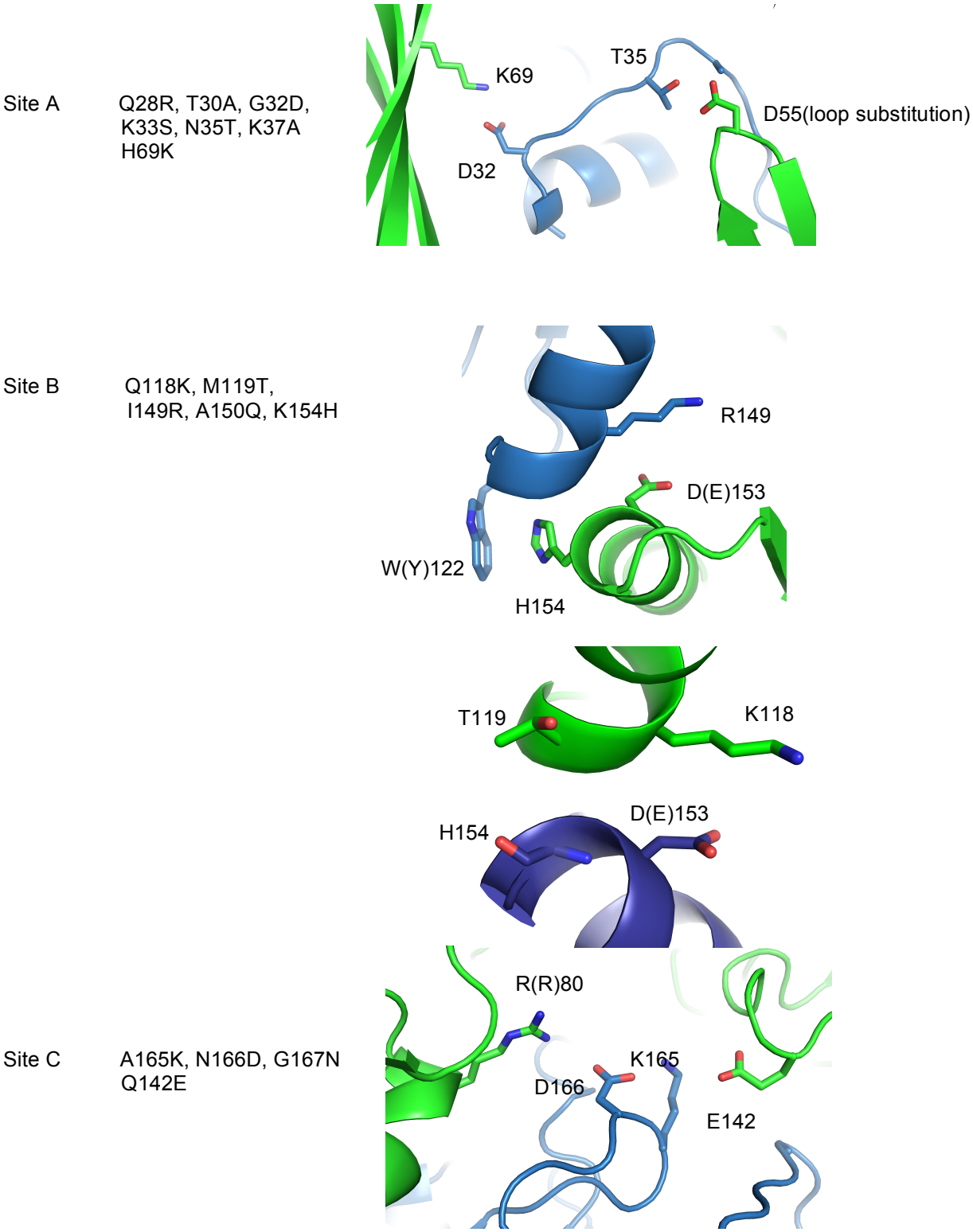


Figure 3.13. Designed mutation sites in Rab27a based on Rab3a crystal contacts. Three sites were identified for modification and labeled as A, B and C. Each site contains residues that are in contact with each other in Rab3a crystal structure (PDB code: 3RAB). Residues involved in contacts between symmetry related molecules in Rab3a for each site are depicted. Residues in parenthesis refer to Rab27a residues that did not need mutations due to functional similarity between the two residues. R(R)80 indicates R80 is identical in both Rab3a and Rab27a. Residue numbering is based on wild-type Rab27a sequence and not the mutants.

The reason behind this was the difficulty in rationally picking one mutation that would effectively orient Rab27a molecules to crystallize exactly like Rab3a. Also, if we had not obtained any crystals in our initial screens it would be difficult to go plan for more mutations, as we wouldn't know if this approach was effective at all. Three regions were targeted in total and are summarized in figure 3.13. All these mutations were not used in all designed constructs and not all of them gave positive results (this is discussed ahead and summarized in table 3.5).

As our final approach, using the SERp web server, we identified clusters of highly solvent exposed residues and mutated them to alanines. Interestingly, there was only one region on α_5-helix identified as a spot for modification. A cluster of K144, E145 and E146 was identified as a region with high conformational entropy suitable for modification (Figure 3.13A). The α_5 helix is not part of the switch or inter-switch, which are important for both nucleotide binding and effector interaction.

Table 3. 5. Designing different Rab27a mutant constructs for crystallization

	Loop sub.	Site A mod.	Site B mod.	Site C mod.	SER	Crystallization	Crystal info.
MJ0 Rab27a	+	−	−	−	−	−	−
MJ1 Rab27a	+	+	+	−	−	3-4 month	4 per ASU, $P3_121$, 3 Å
MJ2 Rab27a	+	+	+	+	−	3-4 days	Spherulite
MJ3 Rab27a	+	−	+	+	−	−	−
MJ4 Rab27a	+	−	+	−	−	−	−
MJ1-1 Rab27a	+	+	+	−	+	3-4 days	1 per ASU, $P4_232$, 3.2 Å
MJ2-1 Rab27a	+	+	+	+	+	3-4 days	Spherulite
MJ3-1 Rab27a	+	−	+	+	+	−	−
MJ4-1 Rab27a	+	−	+	−	+	3-4 days	16 per ASU , C2, 3 Å

mod.: modification, SER: surface entropy reduction [residues were mutated to alanines]

Table 3.5 summarizes the modifications made on each construct. The β_2-β_3 loop was removed on all constructs but was not enough solely, to induce crystallization in Rab27a. MJ1Rab27a was the first construct to crystallize, but only after 3-4 month. In order to accelerate the crystallization process, modifications were made at site C. Although this modification was made postulating more contacts will induce faster crystallization, MJ2Rab27a produced poorly diffracting particles such as spherulites (or even fiber like) in multiple conditions across several plates. None could be

optimized to diffractable crystals. Matrix Microseeding using seeds obtained from the spherulites or seeds from the MJ1Rab27a crystals did not yield significant results. Therefore, it was concluded site C modifications were not compatible with the other modifications. Triple alanine mutation on the “KEE” cluster for SER modifications was not enough to solely induce crystallization, improved the rate of crystallization in both MJ1-1Rab27a and MJ4-1Rab27a constructs. In total, three constructs, yielding diffractable crystals were obtained which are discussed in more detail in the following sections.

3.4.1 MJ1Rab27a

This construct was the first modified mutant that crystallized. Single crystals were formed in 0.12 Ethylene glycols (Molecular Dimensions), 20% PEG 500 MME, 10% PEG 20K, 100 mM Tris-Bicine buffer pH 8 after 3-4 month at 20 ^{0}C. All data collection processing, model building and refinement was done according to section 2.12. All the statistics for the collected dataset are in Table 3.6.

Table 3. 6. Data Processing and Refinement Statistics for MJ1Rab27a

Data Collection	
Space group	$P3_121$
Unit cell parameters	(A°) a = 120.99, b = 120.99, c = 111.90
Wavelength (A°)	0.92
Resolution (A°)	111.73 – 3.15 (2.99 - 3.04)
Total reflections	214807
Unique reflections	19535
Multiplicity	11.0(2.3)
Completeness (%)	100.0(100.0)
<I>/<σ(I)>	11.1(2.3)
R-merge	0.161 (1.093)
R-meas	0.169
Wilson B factor	54.42
$CC_{1/2}$	0.997
Refinement	
Reflections used in refinement	19532 (3976)
Reflections used for R_{free}	970 (106)
R_{work} (%)	26.76
R_{free} (%)	29.11
Rmsd bond lengths (A°)	0.017
Rmsd bond angles (°)	1.63
Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	55.73 /4957
Water molecules	49.95/45
Ligand non-H atoms gppnhp-Mg^{2+}	46.08/132
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	98.1
Ramachandran allowed region (%)	1.9
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	1.6

Statistics for the highest-resolution shell are shown in parentheses.

Systematic absence analysis suggested a high probability for both enantiomorphic space groups of $P3_121$ and $P3_221$. MR and refinement on data indexed in both space groups suggested that $P3_121$ is the correct space group. The final model has four molecules in each ASU. Interestingly, the new β_2-β_3 loop makes important contacts between molecules in the same asymmetric unit and crystallographic

symmetry related molecules (Figure 3.14). Also, mutated residues in site B regions interact with site B region of other molecules playing an important role in crystal formation (Figure 3.14).

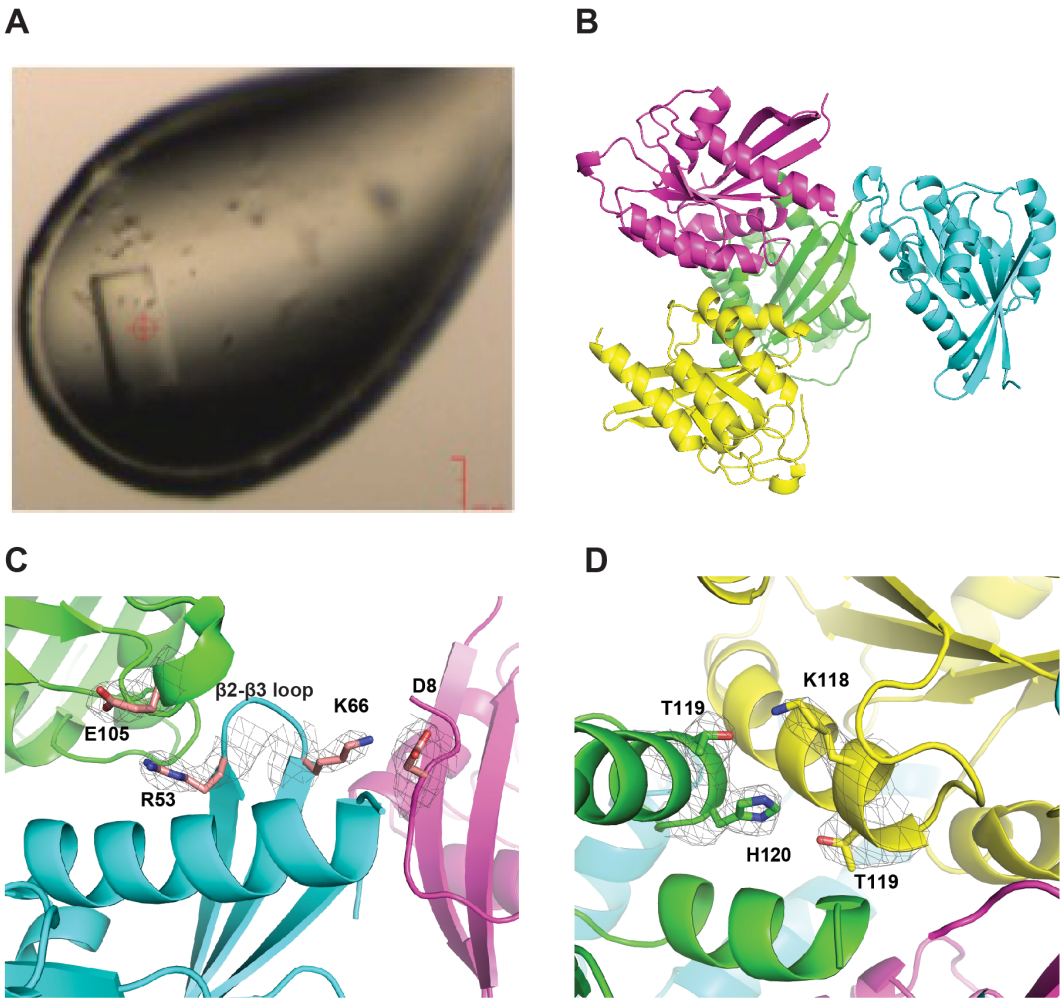


Figure 3.14. MJ1Rab27a crystallization. A) MJ1Rab27a crystals take 3-4 months to grow. B) The crystals have a $P3_121$ space group with four molecules in each ASU. C) The modified β_2-β_3 in MJ1Rab27a provides several key electrostatic interactions with crystallographic and NCS related molecules. D) K118 and T119 as a result of site B modifications, also interact with site B of other crystallographic symmetry related molecules providing extra contacts for crystal formation. Residue numbering is based on wild-type Rab27a sequence and not the mutants.

Analyzing the four molecules in each ASU reveals that the accessibility of the SF4 pocket is different for each molecule. SF4 pocket in chain A is completely obstructed by chain C. The pocket in chains C and D is accessible but is surrounded by other chains and symmetry related molecules. Interestingly, the SF4 pocket and the outer surface in chain B are both completely accessible. This makes chain B molecule completely suitable to study the binding mode of ligands with additional pharmacophores interacting with amino acids outside the SF4 pocket (Figure 3.15).

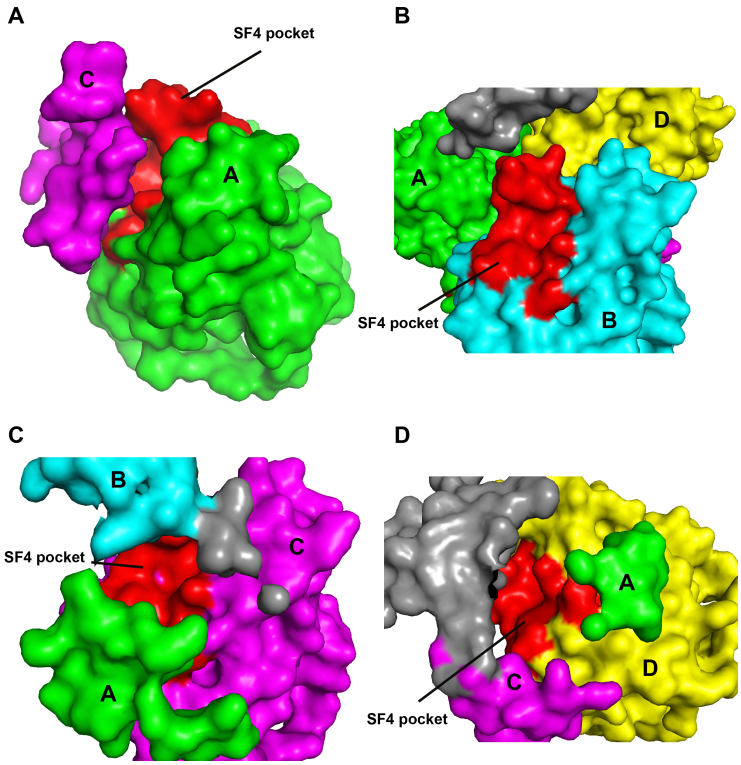


Figure 3.15. SF4 pocket occupancy In MJ1Rab27a. Surface representation of the SF4 pocket (red) for each chain in the ASU is shown in A), B), C) and D) respectively. Atoms in 12 Å proximity of each chain are depicted. Chain A (green), B (cyan), C (purple), D (yellow) and symmetry related molecules (gray). Symmetry related molecules are shown in gray. The SF4 pocket is completely occupied in chain A, while chain C and D are fairly accessible. Chains B is completely accessible, making it useful to study ligands that bind to the effector interaction interface.

MJ1Rab27a crystals provide an opportunity to analyze ligand binding to the effector interaction interface in different molecules of the ASU. The main problems associated with these crystals are the time needed for their growth and their diffraction, resolution limit. Therefore additive screens and micro seeding were attempted in order to obtain crystals with better diffraction that grow under a shorter amount of time. Neither of the two methods proved successful, hence other modifications were employed to obtain crystals with better characteristics.

3.4.2 MJ1-1Rab27a

Since it would take 3-4 month for MJ1Rab27a to crystallize, other modifications were made in order to accelerate the crystallization process. Therefore, site C modifications were made generating MJ2Rab27a and MJ3Rab27a constructs. These modifications did not lead to a positive output, only producing poorly diffracting spherulites (or fibers). Another approach taken was to make a triple alanine mutation based on the results obtained from the SERp server. K145, E146 and E147 were mutated to 145-AAA-147. The new construct, MJ1-1Rab27a, yielded single crystals in 20% v/v Ethylene glycol; 10 % w/v PEG 8000, 30 mM $MgCl_2$, 30 mM $CaCl_2$ and 100 mM HEPES pH 7.5 after 3-4 days at 20 ^{0}C. These crystals were taken for data collection. All data collection processing, model building and preliminary refinement was carried out according to section 2.12. All the statistics for the collected dataset are in table 3.7.

Table 3. 7. Data Processing Statistics for MJ1-1Rab27a

Data Collection	
Space group	$P4_232$
Unit cell parameters	(A°) a = 118.59, b = 118.59, c = 118.59
Resolution (A°)	46.34 – 3.20 (3.2-3.29)
Total reflections	804315
Unique reflections	18333
Multiplicity	43.87(7.5)
Completeness (%)	99.78(98.7)
<I>/<σ(I)>	11.2(1.7)
R-merge	0.454
R-meas	0.459
Wilson B factor	73.125
$CC_{1/2}$	0.998

The crystals diffracted to 3.2 Å. Analyzing the data revealed the presence of only one molecule in each ASU packed tightly against crystallographic symmetry related molecules (Figure 3.16). The SF4 pocket is exposed as shown in figure 3.16. Therefore, ligands targeting the SF4 pocket can bind without having their binding modes affected. Also, there is extra space in three directions to grow the fragments out of the SF4 pocket making this crystal a very suitable tool for structure-based ligand discovery (Figure 3.16). This characteristic, combined with the fact that the single crystals grow in under 4 days, devoid of crystal pathologies such as twinning in a cryo-protected condition make them more suitable for high-throughput ligand studies. The only drawback with the current crystals is their diffraction resolution limits, which is only up to 3.2 Å. Optimizing the crystal conditions to obtain bigger crystals, additive screens, micro seeding [219] and use of D_2O [220] are methods that could potentially improve the diffraction resolution of the crystals to levels below 3 Å.

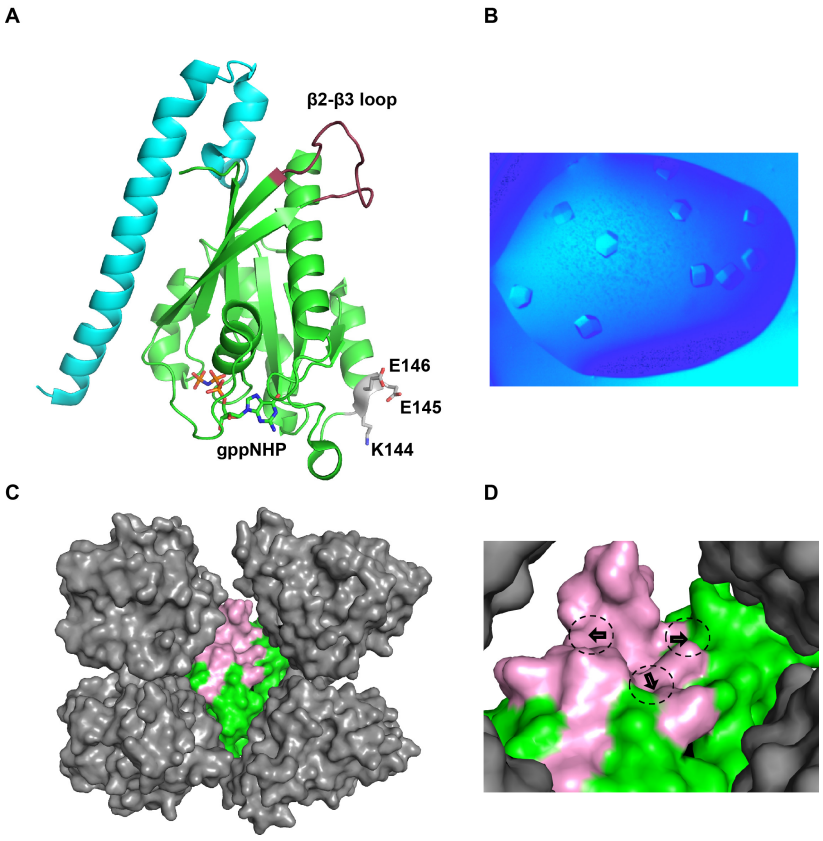


Figure 3.16. MJ1-1Rab27a crystallization. Rab27a (green), $Slp2_a$ (cyan), SF4 pocket (pink) A) 144-KEE-146 were mutated to alanines to reduce penalty for conformation entropy reduction during crystallization. B) Single crystals for MJ1-1Rab27a were obtained consistently in short time. C) Only one molecule is present in each ASU tightly packed against other symmetry related molecules (gray), though leaving the SF4 pocket (pink) exposed for ligand binding studies. D) The solvent accessibility of the SF4 pocket (pink) in MJ1-1Rab27a enables studying the binding mode of ligands that have additional pharmacophores interacting with amino acids outside the SF4 pocket.

3.4.3 MJ4-1Rab27a

Similar to MJ1-1Rab27a, MJ4-1Rab27a construct has K145, E146 and E147 mutated to 145-AAA-147. Single crystals were formed in 20% v/v Ethylene glycol; 10% w/v PEG 8000, 30 mM $MgCl_2$, 30 mM $CaCl_2$ and 100 mM HEPES pH 7.5 after 3-4 days at 20^0C. The statistics for data collection are brought in table 3.8. MR was done using phaser similar to other previous datasets, but phaser was initially unable to yield any model solutions. In order to solve the problem and use MR to obtain a model structure, we deciphered each of the problems separately and then used a manual mode of phaser to tackle the problem.

3.4.3.1 Cell content analysis

In order to complete a successful structure solving by MR, all components present in the ASU must be located in the correct orientation (rotation) and position (translation). Therefore, cell content analysis must be done prior to any calculations to reach an estimate of how many molecules are present and should be searched for in each ASU [221]. Since all the components contribute to the total scattering, large copy numbers in each ASU mean only a small fraction of diffraction comes from each component. Therefore, often there is only a weak or unsubstantial correlation between experimental and model amplitudes. Unlike traditional MR, maximum-likelihood based software phaser uses partial rotation and translation information obtained from the first component to obtain improved correlation (Log Likelihood Gain (LLG)) and signal to noise ratio (Z score) of both translation and rotation for the second component [221, 222]. This helps to discriminate between grid peaks and derive the best solution(s) with higher and more distinct signal to noise ratio. The Matthews coefficient forms the basis for cell content analysis in phaser. Matthews coefficient is defined as the crystal volume per unit protein molecular weight ($Å^3$/Dalton). Phaser uses an improved Matthews probability calculator which takes into account the probabilistic nature of the Matthews coefficient and resolution to predict the number of molecules in each ASU [223, 224].

Cell content analysis for MJ4-1Rab27a is shown in table 3.9. As demonstrated, the probability calculator does not have a single prediction with high probability but suggests that a large number of molecules from 16-24 could be present in each ASU.

Table 3. 8. Data Processing Statistics for MJ4-1Rab27a

Data Collection	
Space group	C2
Unit cell parameters	(A°) a = 130.33, b = 132.40, c = 230.38
Resolution (A°)	80.14 – 2.66 (2.66-2.88)
Total reflections	390926
Unique reflections	108964
Multiplicity	3.6(3.0)
Completeness (%)	99.9(99.6)
<I>/<σ(I)>	4.2(0.3)
R-merge	0.111
R-meas	0.116
Wilson B factor	60.05
$CC_{1/2}$	0.995(0.688)

Table 3. 9. Estimation of number of molecules in ASU for MJ4-1 using Matthews coefficient

Nmol/ASU	Matthews.Coeff	%solvent	P(tot)
11	4.34	71.71	0.00
12	3.98	69.13	0.01
13	3.68	66.56	0.01
14	3.41	63.99	0.02
15	3.19	61.42	0.03
16	2.99	58.84	0.05
17	2.81	56.27	0.07
18	2.65	53.70	0.10
19	2.52	51.13	0.12
20	2.39	48.56	0.13
21	2.28	45.98	0.13
22	2.17	43.41	0.12
23	2.08	40.84	0.09
24	1.99	38.27	0.06
25	1.91	35.69	0.03
26	1.84	33.12	0.01

Nmol: Number of molecules, P: probability

This is consistent with data obtained by the self-rotation function, suggesting at least 14-16 molecules are present in each ASU. The self-rotation function reveals the dihedral relationship between two or more identical molecules in each asymmetric unit. This function measures the overlap of the native Patterson function with an identical Patterson function rotated using a rotation matrix composed of ω, φ and κ spherical polar angles. The self-rotation function will have maxima at orientations were two Patterson functions are superimposed [225]. As illustrated in figure 3.17, there is a four-fold dihedral symmetry between the Patterson maps. The self-rotation function map suggests the presence of a four-fold NCS axis and 7-8 two-fold axes. The two-fold rotation axes can be identified from the peaks that are present at $\kappa=180^0$. The four-fold rotation axis is recognizable by a strong peak at both $\kappa=180^0$ and $\kappa=90^0$. All together, the data from both Matthews probability calculations and self-rotation functions, imply that a high number of molecules are present in each ASU. This could be a reason for the low Z score and LLG for search of the first component in phaser.

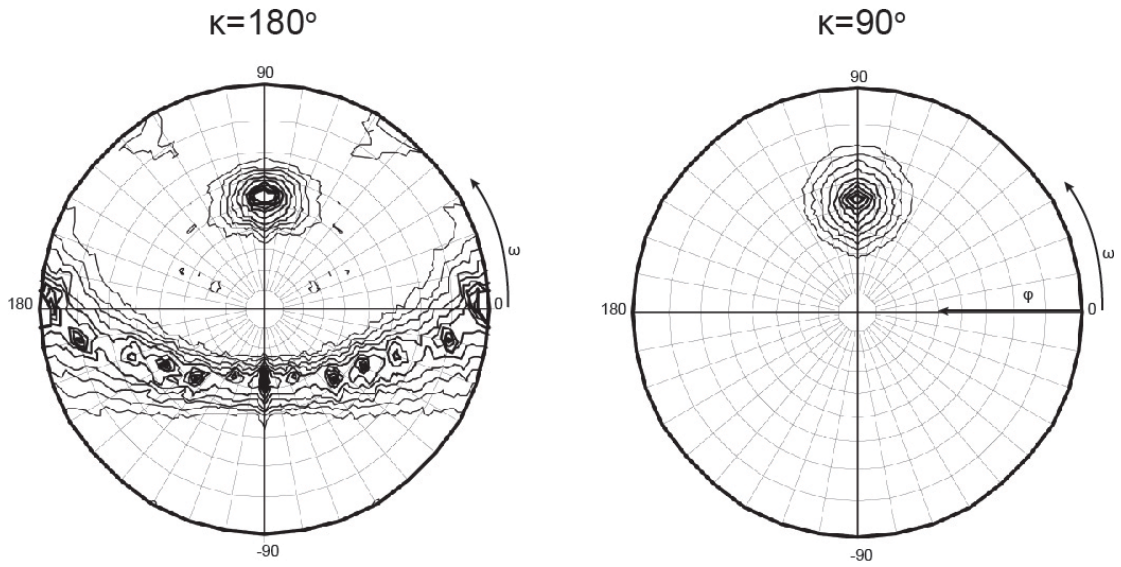


Figure 3.17. Self-rotation function stereographic projection at different κ values calculated for dataset obtained from MJ4-1Rab27a crystals. As depicted the maps at $\kappa=180^0$ and $\kappa=90^0$ predict the presence of 7-8 two-fold NCS and one four-fold NCS axis present in each ASU.

3.4.3.2 Anisotropy correction

When crystals diffract to different resolution limits in different dimensions the crystals are anisotropic. This is potentially due to the presence of different lattice contacts in different directions. Highly anisotropic data can affect maximum likelihood functions and therefore it is imperative to scale the intensities using anisotropic scaling factors. The anisotropy in each dataset is measured using anisotropic-delta B that is the difference between the two most different principal components of B factor. Each principal component is an exponential scaling factor used to normalize the data for each direction of the reciprocal space. Values above 50 $Å^2$ are considered as extremely anisotropic. The anisotropic Delta-B was measured for MJ4-1Rab27a dataset using the manual ANO mode of phaser (Figure 3.18).

A

```
phenix.phaser << eof > aniso.log
MODE ANO
ROOT aniso
HKLIN AUTOMATIC_DEFAULT_free.mtz
LABIN I=IMEAN SIGI=SIGIMEAN
END
eof
```

B

```
Principal components of anisotropic part of B affecting
observed intensities:
eigenB (Å²)     direction cosines (orthogonal coordinates)
131.590         0.7831   0.0000  -0.6219
 -61.889         0.0000   1.0000   0.0000
 -69.702         0.6219   0.0000   0.7831
```

Figure 3.18. Measuring anisotropy of MJ4-1Rab27a dataset. A) Script used to run phaser in anisotropy (ANO) mode. This mode calculates the anisotropic scaling factors in each direction and applies them to correct for the anisotropy in different dimensions of the reciprocal space. B) Converged principal components of B factor for each dimension after 8 cycles of iteration are shown. Anisotropic Delta-B factor, which is an indicator of how much the dataset is anisotropic, is 201. (131.590-(-69.702)).

As demonstrated this dataset is severely anisotropic with a Delta-B factor of 201 $Å^2$ (Figure 3.18), with the effective resolution 2.8 Å in one dimension and 5 Å in the other most divergent dimension. This is something rarely observed in other anisotropic datasets and could be the result of our protein engineering process (Figure 3.19).

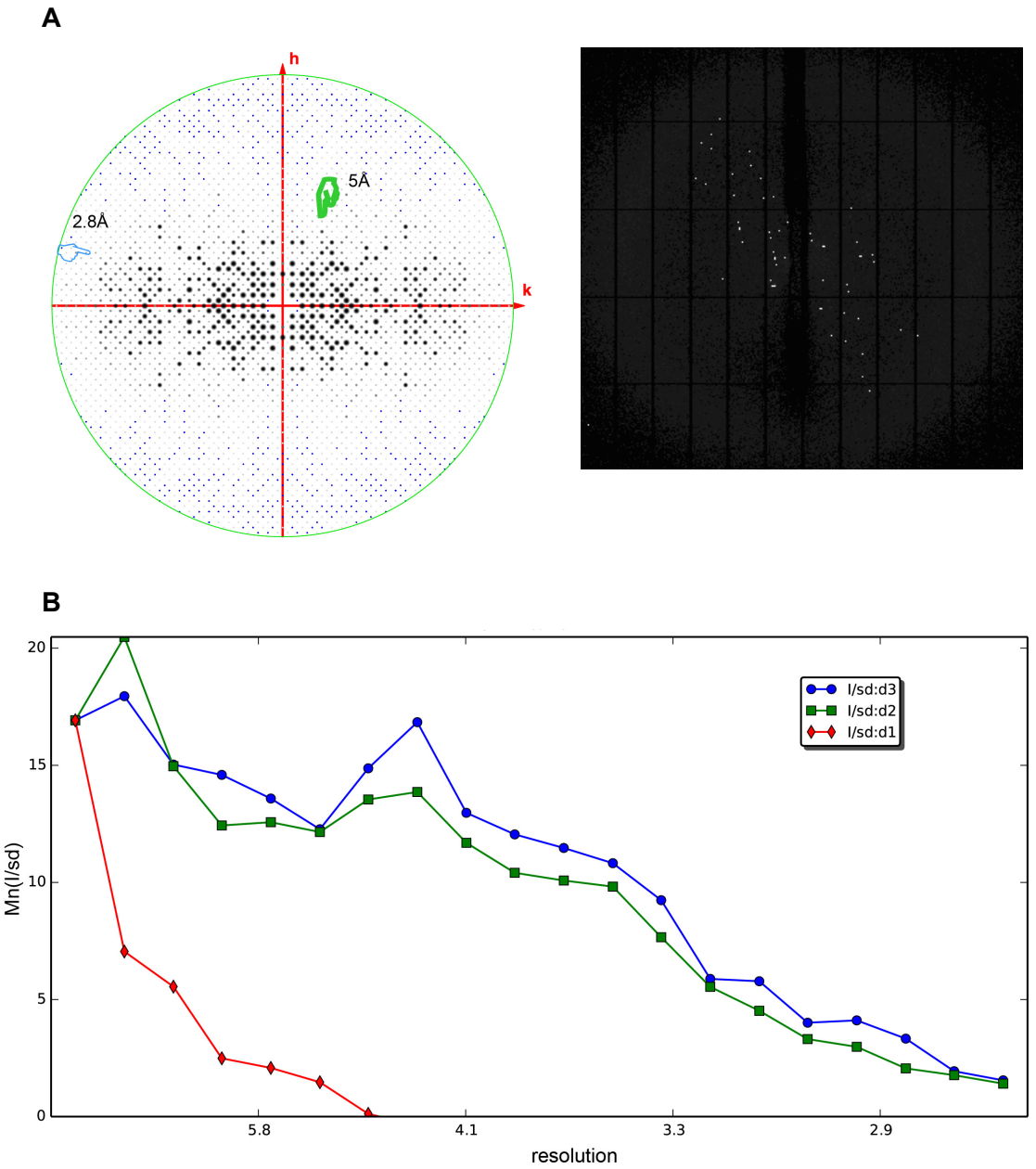


Figure 3.19. Anisotropic diffraction of MJ4-1Rab27a crystals. A) A diffraction spots (at l=0) and image demonstrating the severity of anisotropy on the dataset, with the crystal diffracting to 5 Å in one dimension and 2.8 Å in the other one. B) I/σ vs resolution for different directions shows the extremity of the anisotropy in diffraction between different crystal directions.

3.4.3.3 translational NCS (tNCS)

Translational NCS (tNCS) is the presence of two or more copies of the same molecule having the same orientation in the ASU. TNCS can lead to very weak or strong spot intensities therefore changing the Gaussian distribution of the structure factor intensities, an assumption made based on the central limit theorem by Wilson [226]. Modulation of spot intensities by tNCS is controlled by phase difference between tNCS related molecules (as a result of translation between the two or more molecules) and difference in the orientation of the copies. Native Patterson maps can reveal the presence of tNCS. Each map is composed of inter-atomic vectors connecting any two atoms in the ASU space to each other (every pair of atoms in space is represented by two vectors) with their tails translated to the origin of the map. In the presence of tNCS, all the vectors connecting all the equivalent atoms in both molecules will be identical (length and direction) resulting in a strong off origin peak (larger than 20% of the origin peak). Phaser versions 2.5.4 and after, use Maximum likelihood functions describing tNCS vectors as a result of pure translation and random subtle conformation and orientation changes (up to 10^{o}) to adjust expected intensities for each reflection by taking into account tNCS correction parameters. The automatic version of phaser can deal with simple cases of tNCS where a single tNCS vector correlates two groups. An understanding of the tNCS is needed for more complex situations, which includes a complete understanding of the cell content and identifying the tNCS vectors between the molecules [227].

Native Patterson map was calculated for MJ4-1Rab27a using the Patterson generate map software from CCP4. The dataset obtained for MJ4-1Rab27a shows a very strong off origin peak at (0.0, 0.023, 0.50) coordinates with a height 37% relative to the origin peak (Figure 3.20). These data indicate the presence of a strong tNCS in this dataset leading to a very wide range of intensity distribution.

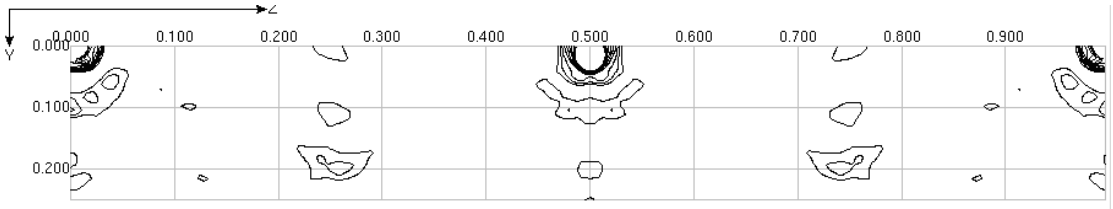


Figure 3.20. Native Patterson map for MJ4-1Rab27a. A slice of the Patterson map at X=0 is shown. As illustrated there is strong off origin peak at Y=0.023 and Z=0.5. This is a strong indicator for the presence of tNCS in the crystals.

3.4.3.4 Solving the MR problem

The approach for solving the MR problem was carried out with guidance and suggestions of Professor Randy Read from Cambridge University. In order to obtain a solution for the MR, an ensemble was generated using Rab27a(gppNHp), Rab27b(gdp), Rab27b(gppNHp) and Rab8a(gppNHp) structures from PDB deposits 3BC1, 2IEY, 2ZET and 4LHW respectively. The aligned structures were pruned to anything that does not superimpose, using Sculptor. This process mainly eliminates flexible loops leaving the conserved core fold intact. As shown in figure 3.21, the switch I, II and inter-switch regions, which are flexible and go under the most conformational change between GDP and GTP switch, are pruned out. MR was carried out using this model. Interestingly, although a reasonably clear solution was identified by phaser for a first pair of tNCS-related copies, but then Phaser insisted on placing further copies directly on top of the first two copies (Figure 3.21). This is potentially due to the high anisotropy of the dataset collected by diffracting the MJ4-1Rab27a crystals.

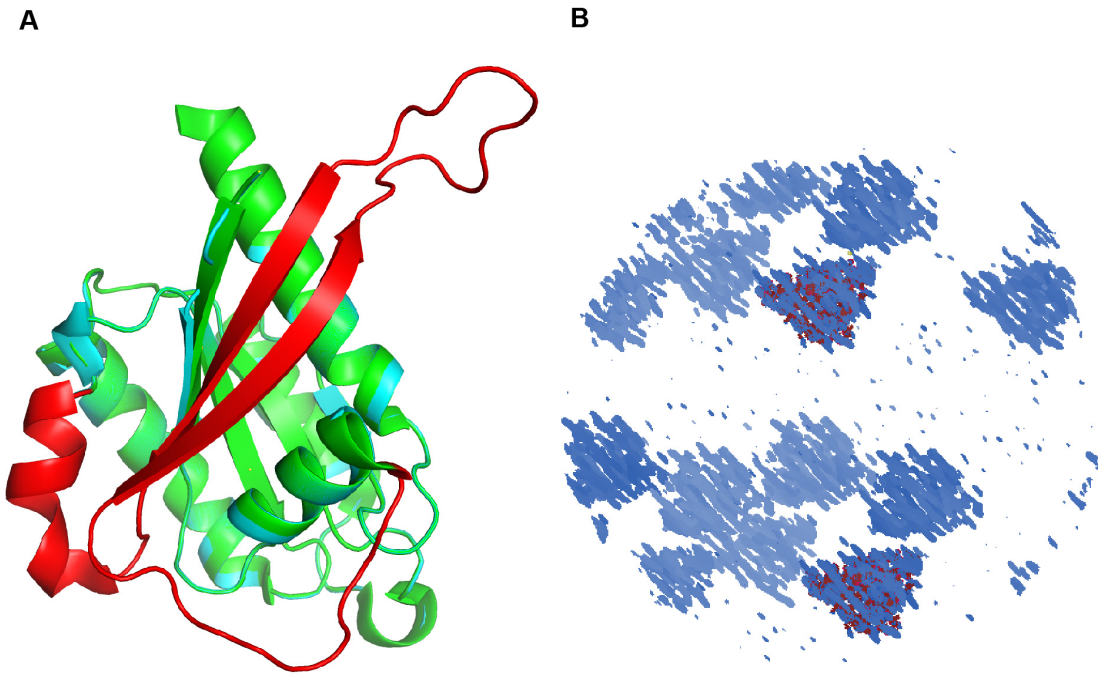


Figure 3.21. Solving the structure of MJ4-1Rab27a using an ensemble structure. A) Overlay of Rab27a (green) with an ensemble structure (cyan). The omitted region from Rab27a for MR is highlighted in red. This region contains the switch I, II and inter-switch region. B) As demonstrated MR only recognizes two tNCS related molecules (models highlighted in red) and tries to build molecules on top the two molecules only in the vertical dimension.

The combination of strong tNCS (Figure 3.20) and extremely high anisotropy with a Delta-B factor of 201 $Å^2$ (Figure 3.18) led to a very wide distribution of expected intensities, with many reflections expected to have extremely weak intensities based on the systematic effects of anisotropy and tNCS. In theory, the new intensity-based likelihood target in phaser should cope with these data. If the standard deviations of the systematically weak reflections are large compared to their expected intensities, they should essentially be ignored by the likelihood function. However, it is possible that there might be problems with the sigmas of these weak reflections or with the accuracy of the anisotropy and/or tNCS corrections applied to their expected intensities. In order to solve this problem, the anisotropic scaling factor and tNCS scaling factors were calculated using the ANO (anisotropy) mode and TNCS mode in phaser manual mode. Subsequently, using sftools from the CCP4 package, a new column was generated by multiplying the scaling factors of the ANO and tNCS modes. Afterwards, any reflection, which was scaled up with a multiplication factor of more than 10, was removed. This removed about 40% of the data. Experimenting different thresholds for the scaling factor suggested that a 50× cutoff still lead to a correct MR solution, but only with 20% of reflections lost. Using this trimmed data, phaser could successfully find 16 copies of the ensemble model (trying different possibilities from 12 to 20). This gave a model with sensible packing. Moreover, generation of an NCS-averaged map resulted in electron density corresponding to the nucleotide, suggesting that the model is correct (Figure 3.22).

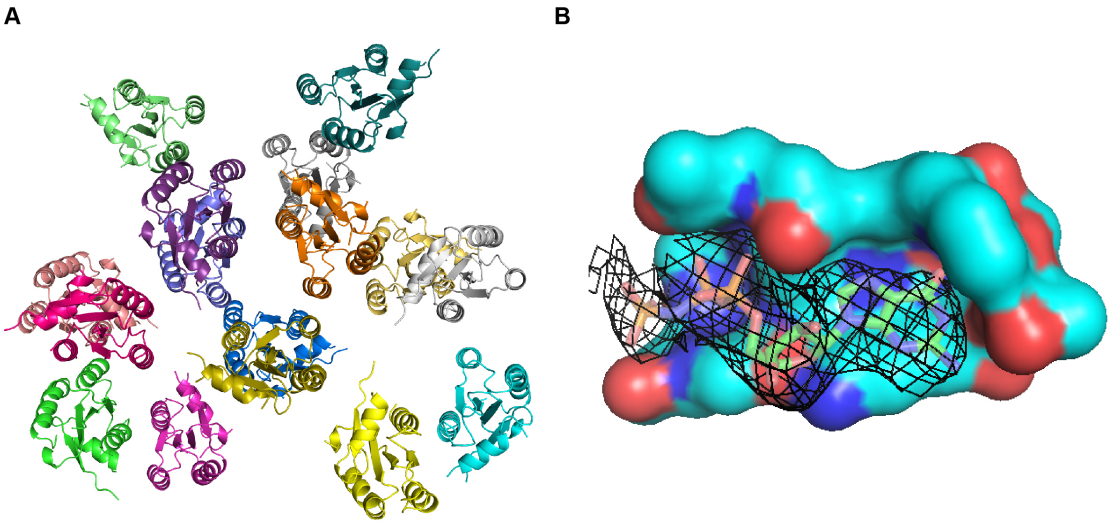


Figure 3.22. MJ4-1Rab27a model solution by MR. A) Modified version of MR found 16 molecules in each ASU, related to each other through NCS. B) electron density was recognizable for the nucleotide in the MR solution, suggesting that the model is correct. Rab27a nucleotide binding pocket in 5 Å proximity of the nucleotide is shown by a surface representation.

This model is consistent with data obtained by the self-rotation function (Figure 3.23). 8 pairs of molecules are related to each other by a two-fold rotation axis. Also, a four-fold rotation axis is evident from the structure. Finally, consistent with the native Patterson map, a translational symmetry between three groups of molecules is apparent in the model, (Figure 3.23).

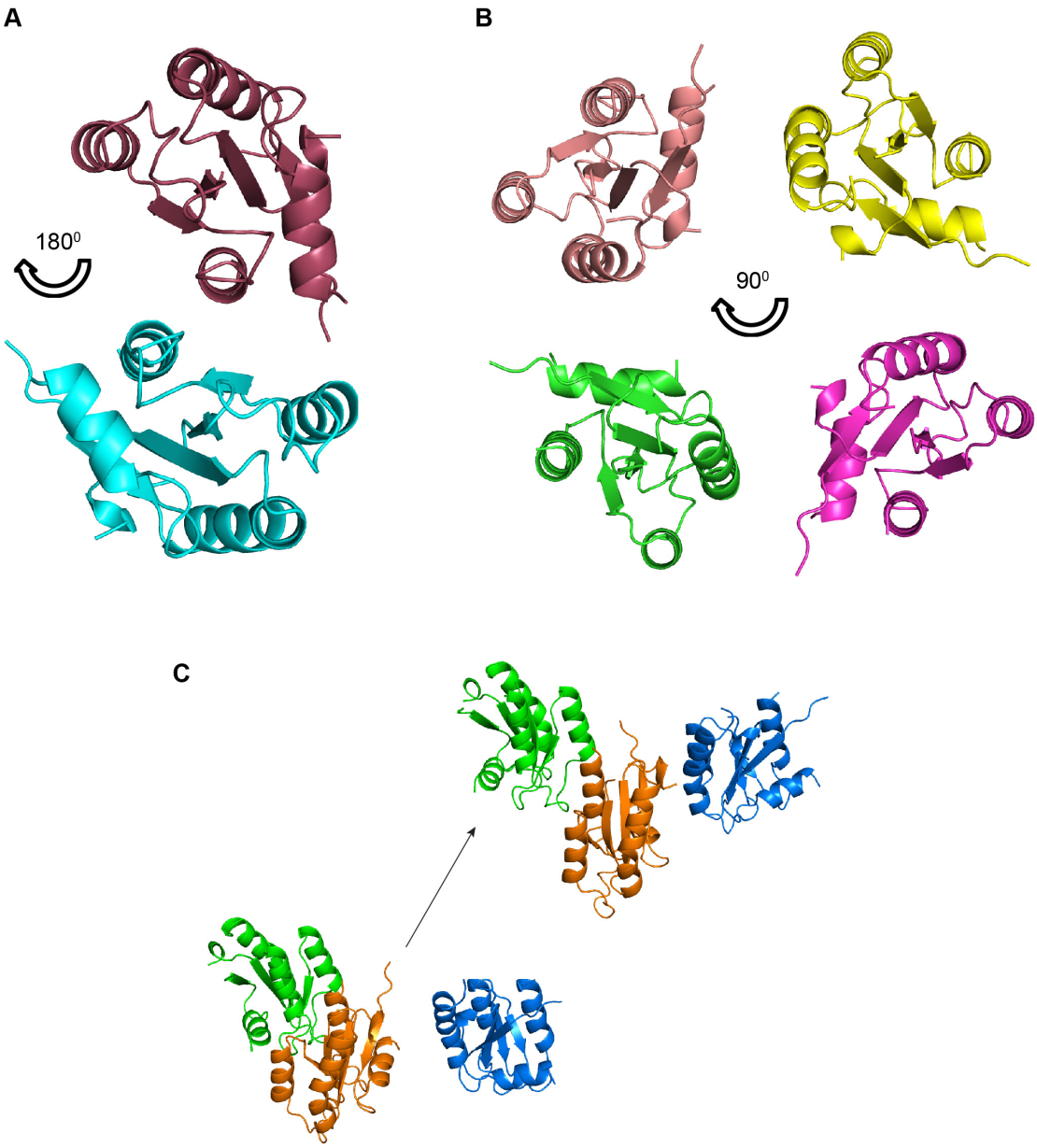


Figure 3.23. NCS axes in MJ4-1Rab27a. A) 8 two-fold NCS is present between 8 pair of molecules. One sample pair is shown. B) Similar to self-rotation function, a four-fold symmetry axis is present in the ASU. C) In corroboration with the native Patterson map, three group of molecules in each ASU are related to each other by translational vectors. The relationship for one group is shown.

Although we have a solution for this dataset from MR, in practice, the high copy number, the extreme anisotropy and strong tNCS make MJ4-1Rab27a an unsuitable construct for ligand binding studies. An interesting observation from the solved structure is the absence of decent number of crystal contacts in regions close to site A modification. This is potentially one of the main reasons for the extreme anisotropy in this crystal form. Therefore, site A modifications are potentially necessary as seen in MJ1-1Rab27a crystals (section 3.4.2)

3.5 Discussion

Availability of structural data can fire up any drug discovery project. Crystallography is the drive force of structure driven medicinal chemistry due to the fact that it gives you detailed information about ligand interaction with the protein side chain, enabling chemists to use that information to improve the potency of the binding molecules. As part of the bigger plan to obtain tight binding and selective ligands against Rab27a, we made attempts to obtain crystals of Rab27a. These crystals would be used to study the binding mode of ligands to Rab27a. Since in our hands, Rab27a did not crystallize on its own (in corroboration with previous reports [216]), we took two different approaches to obtain crystals of Rab27a.

First approach was inspired by the crystal structures available for Rab27a-Slp2$_a$, Rab27b-Slac2$_a$ and Rab3a-Rabphilin. Since the RBD of all the effectors for these proteins are made up of a long helix with a hook-like structure at the end, filling an important interaction site in the SF4 pocket [137], we made a fusion between the C-terminus of the effector's RBD$_{\alpha1}$ and N-terminus of Rab27a. This approach resulted in two benefits. Primarily, as evident from the complex crystal structure, RBD$_{\alpha1}$ provided extensive contacts with molecules within the ASU and crystallographic symmetry mates, therefore driving the crystallization of the construct. In fact, the Slp fusion constructs all crystallized in less than 12 hours in a variety of conditions, emphasizing the role of the RBDα1 in the fusion constructs for crystallization. Secondly, this construct would be ideal for studying ligands binding to the SF4 pocket. The crystals obtained from the Slp2$_a$-Rab27a fusion while diffract to a resolution suitable for studying ligand-binding modes in the SF4 pocket, have some drawbacks mainly due to the presence of RBD$_{\alpha1}$. Two Rab27a molecules crystallize against each other in each ASU generating an artificial pocket in the

crystal construct. This deep pocket can affect ligand interaction studies. Attempts were made to modify the construct by using the $RBD_{\alpha 1}$ from Slp1, in hope to change the arrangement of molecules in each ASU. Although the Rab27a-Slp1 complex inspired this approach, it did not bear the wanted results. The effect of this artificial pocket and the effect of direct removal of the residues involved in its formation are discussed in chapter 4.

The second approach, tried to reproduce interactions important for Rab3a crystal formation. Three sites away from the effector interaction interface were identified in Rab3a. These sites contained residues that were involved in generating crystal contacts through electrostatic and/or hydrogen bond interactions. Residues that were different in Rab27a (to Rab3a) were mutated to residues present in Rab3a. Several constructs were made with different mutation sites present in different constructs (Table 3.5). Of the three regions, site B modifications that were used to make MJ2Rab27a, MJ2-1Rab27a, MJ3Rab27a and MJ3-1Rab27a did not prove to be helpful in generating good quality crystals. MJ1Rab27a was the first construct to produce diffractable crystals. Four molecules were present in each ASU of crystals obtained for this construct. Chain B was completely accessible, making it ideal to study ligands binding to the effector interaction interface. The major problem with these crystals was that they suffered from slow growth. Therefore, attempts were made to improve MJ1Rab27a. In fact, constructs having site B modifications were made, in an attempt to increase the crystallization rate of MJ1Rab27a crystals. This approach did not give any positive results. Next, in order to improve MJ1Rab27a crystals, site A modifications were removed. This is because, some of the mutations in this site are located close to the nucleotide binding pocket. We hypothesized this might have affected the conformation homogeneity of the construct. Hence, MJ3Rab27a and MJ4Rab27a were made. This approach was not helpful either. Finally, based on the SER theory, a triple alanine mutation was made in MJ1,2,3 and 4Rab27a, anticipating that this would be helpful in generating crystallizable constructs. From the generated constructs, both MJ1-1Rab27a and MJ4-1Rab27a proteins crystallized. MJ4-1Rab27a crystallized in short amount of time but suffered from severe anisotropy and a pathologic condition known as translational NCS (tNCS). After solving the structure of MJ4-1Rab27a, it was evident that there are not enough crystal contacts in certain directions and site A modifications could be helpful. Indeed, MJ1-1Rab27a, which contains site A modifications, crystallized producing crystals that would have an isotropic diffraction pattern in the reciprocal space. The only drawback of MJ1-1Rab27a is that it diffracts to 3.15 Å. This can be improved,

potentially by changing cryoprotectant conditions, re-annealing, dehydration, microseeding or use of D_2O or any additive molecules in the future [219, 220]. This will constitute the next step in our optimization process.

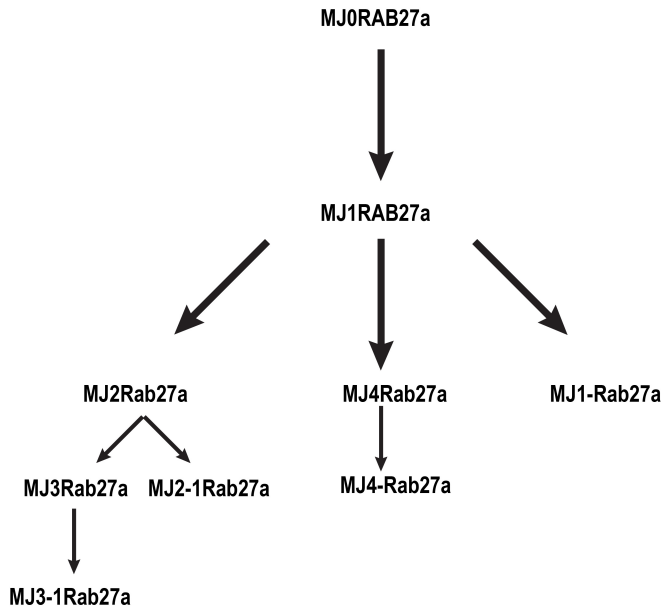


Figure 3.24. Evolution of mutant Rab27a constructs design. The evolution of mutant Rab27a constructs generated for crystallography is shown.

Apart from the approaches presented in this chapter, another method is taking advantage of another Rab protein currently known to crystallize on its own. In this method, the Rab protein would be used as a scaffold to generate crystals. This method can be very resource demanding, as multiple constructs with different variations need to be assessed. Moreover, the effects of minor conformational changes specific to a particular Rab on the shape of this pocket could cause complications in using such an approach. All together, we have used available structural information on Rab27 and homologues to produce crystallizable constructs of Rab27a that can be used for ligand interactions studies.

CHAPTER FOUR: FRAGMENT-BASED LIGAND SCREENING OF RA27A USING DSF

4.1 Introduction

Although protein-protein interactions (PPIs) have varied characteristics, some of the most interesting ones from a drug discovery perspective are among the most challenging targets due to the nature of the surfaces involved in these interactions [228]. Many PPIs, including Ras superfamily-effector interactions, cover large, flat, featureless areas with contrasting properties from small molecules present in chemical libraries [163]. The disparity is not facilitated by the fact that, most of the medicinal chemistry in the last two decades, has biased the commercially available chemical space towards the organic molecules complementary to well-defined pockets present in enzymes, receptors and ion-channels [229].

Apart from Peptidomimetic approaches that have been used [230], fragment-based drug discovery (FBDD) has been used as a successful approach to target PPIs [231]. In contrast to high throughput screening (HTS) 100,000s of compounds, fragment libraries are made of a few thousand small sized molecules often designed to have chemical groups that can engage in hydrogen bonds and Van der Waals interactions [209]. The “rule of 3“ (MW < 300, CLogP ≤ 3, hydrogen bond donor ≤ 3, hydrogen bond acceptor ≤ 3) is used as a general standard for designing fragment libraries. The distinct physicochemical properties of fragments, is key in developing fragments into high affinity binding molecules [209, 232]. FBDD enables synthetic chemists to control ligand characteristics by increasing affinity through adding functional groups that provide direction to ligand binding, as opposed to an increase in affinity without producing any specific interactions with the binding target [233]. Fragment linking approaches initially put FBDD on the spotlight by demonstrating the synergistic increase in affinity upon linking two fragments, binding to two binding sites close to each other [234]. Apart from linking, merging and growing are other approaches taken to develop fragments into high affinity lead compounds [235-237].

Various methods have been used to assess the weak binding of fragments to proteins. Structural biology approaches such as NMR [238, 239] and X-ray crystallography [240, 241] and biophysical methods including surface plasmon resonance (SPR) [242, 243], differential scanning fluorimetry (DSF) [244, 245], capillary electrophoresis [246], mass spectroscopy [247, 248],

microscale thermophoresis [249] and weak affinity chromatography [250] have all been used as screening tools for fragment-based drug discovery. Measuring direct binding of fragments to proteins at high concentrations leads to higher false positive rates in comparison to competitive screening methods. Hence, it is imperative to validate ligand binding using orthogonal methods or using a competitive binder with a known binding site [251]. X-ray crystallography and NMR, are cornerstone of validation, as they can provide information on the ligand binding site and binding mode [251]. Although, compounds can only be considered as true hits when a meaningful structure-activity relationship (SAR) can be built based on the proposed binding models [209].

In this work, a hierarchical approach was taken to screening Rab27a using a fragment library [244]. This procedure has been reported fruitful for several targets [252, 253]. Initially, screening was carried out using DSF, a fast, high throughput, and inexpensive procedure to obtain initial hits. Subsequently, selected hits were evaluated using the more time consuming but precise orthogonal methods of ligand-observed and protein NMR. Finally, the binding modes of potential hits were investigated using X-ray crystallography (Figure 4.1).

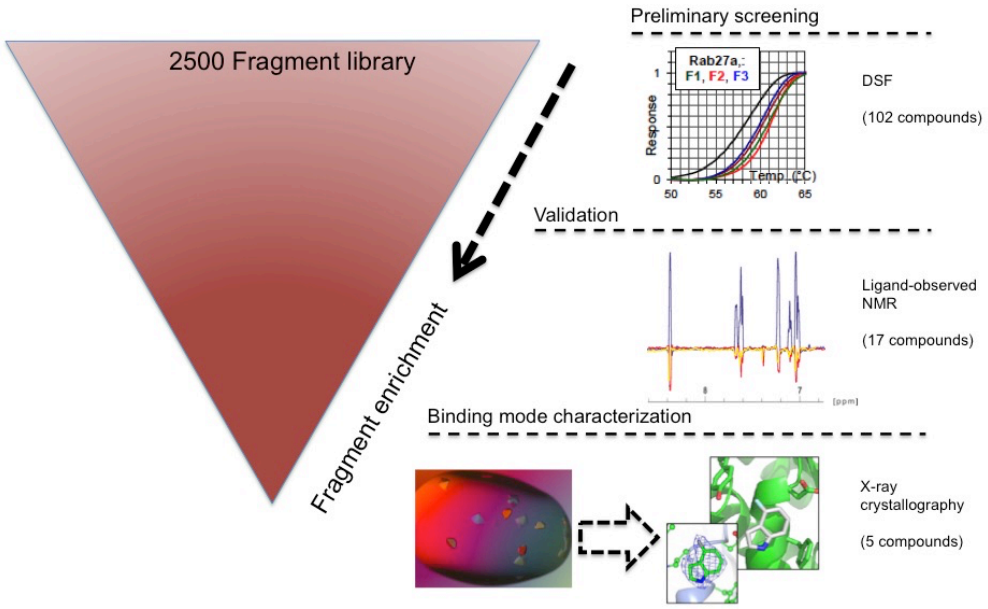


Figure 4.1. Fragment-based ligand discovery of Rab27a. Screening, validation and characterization of fragments was carried out using a hierarchical procedure. Differential scanning fluorimetry was used as a high throughput method to screen for initial hits leading to 102 compounds. The binding of 17 compounds was re-evaluated by NMR. Finally, five compounds were analyzed using X-ray crystallography.

4.2 DSF screening

Prior to running the screen, buffer conditions at different pH were evaluated to find conditions, which Rab27a is most stable at (Figure 4.2). As demonstrated, Rab27a was most stable in HEPES buffer pH 7.4. The screening was carried out and analyzed according to protocols in section 2.9.1. 102 compounds were identified as initial hits. Hits were identified using $T_m \geq$ AVG $T_{m\text{-control}}$ + 2×standarddeviation$_{Control}$+0.5°C as cut off. The majority of the compounds demonstrated no binding or a negative ΔT_m (Figure 4.2). Three compounds out of 102, had ΔT_m>10°C with poor quality denaturation curves. These compounds were eliminated from further analysis. A full table presenting the ΔT_m for each hit is in Appendix B.

A

Buffer/pH	T_m
PO_4^{3-}/4.4	34°C
PO_4^{3-}/5.4	35°C
PO_4^{3-}/6.4	36°C
PO_4^{3-}/7.4	43°C
Hepes/7.4	47.5°C
Tris/8.4	45.5°C
Tris/9.4	44.5°C

B

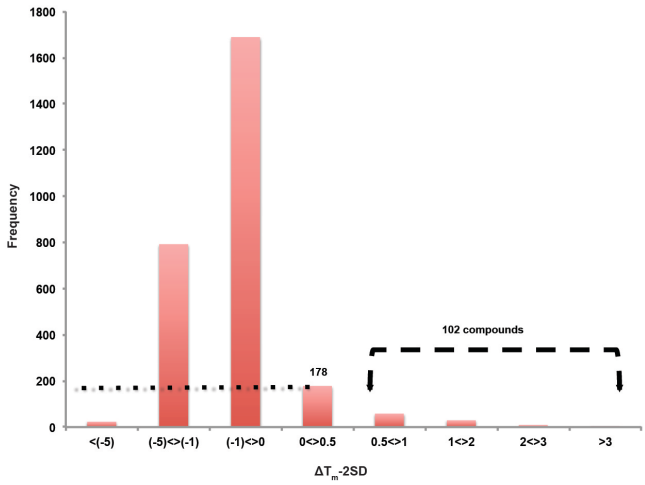


Figure 4.2. DSF screening against Rab27a. A) Stability of Rab27a was measured at different pH. Rab27a was most stable in HEPES buffer pH 7.4. All buffer conditions had 20 mM buffer, with 150 mM NaCl and 5 mM $MgCl_2$. B) As demonstrated, 102 compounds from the DSF induced shifts in T_m more than the AVG $T_{m\text{-control}}$ + 2×standarddeviation$_{Control}$+0.5 °C cut off. Ligands selected from these hits were taken for further validation.

Next, the hits were clustered using the software Scaffold Hunter 1.0 [254] into groups (Figure 4.3). Scaffold hunter clusters ligands by pruning sidechains of fragment structures to obtain the molecular framework also known as the murcko framework [255]. Subsequently, it removes the rings in an iterative process based on rules defined by Schuffenhauer at el. [256].

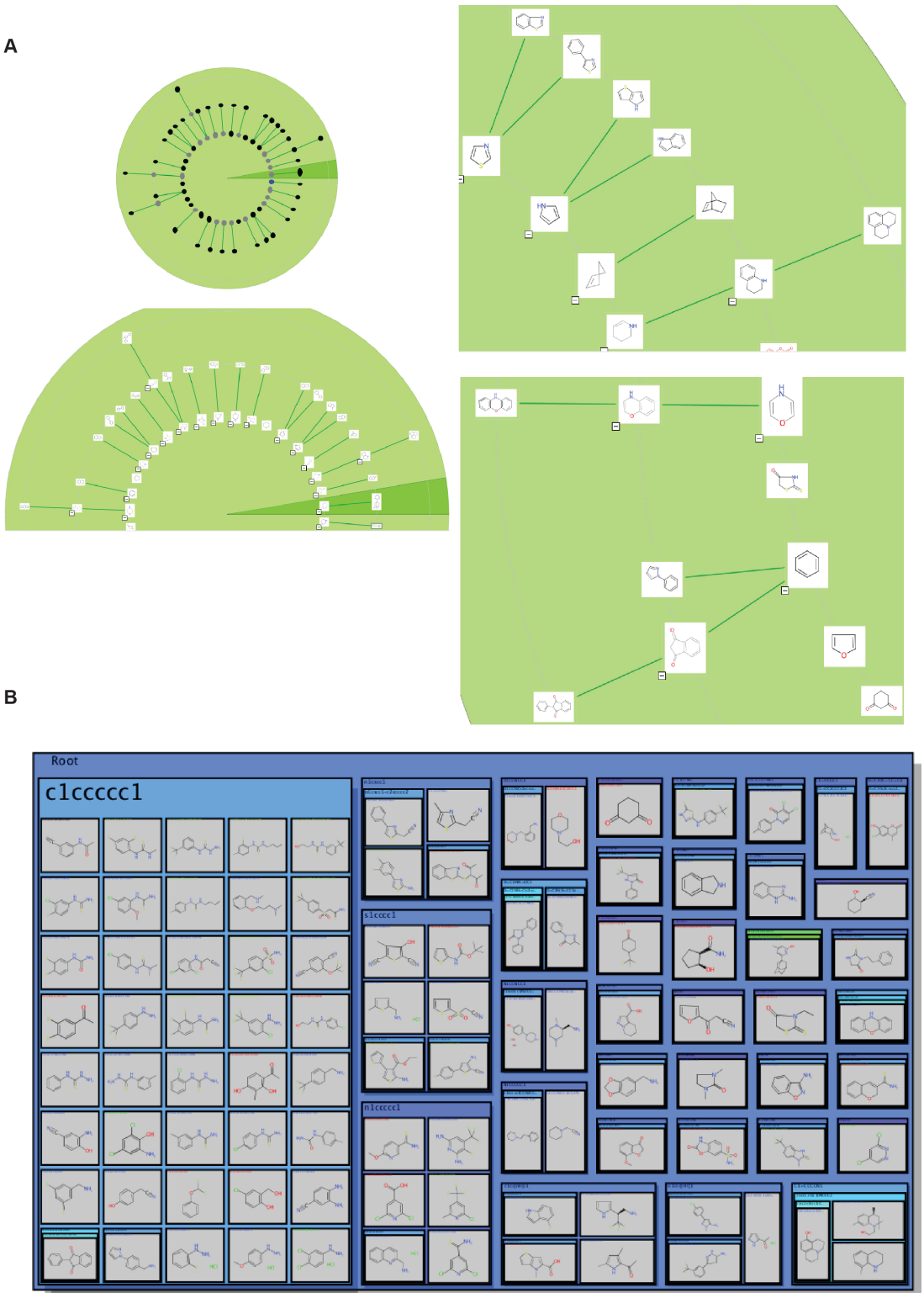


Figure 4.3. Hierarchical clustering of fragments hits. Rings were removed in an iterative manner using Scaffold Hunter 1.0 and clustered using the scaffold hierarchy. A) Demonstrates the scaffold tree. Each circle layer represents one layer. Each node represents one scaffold, which contains one or more compounds. Nodes connected to each other all are part of the same cluster. B) Compounds belonging to the same cluster are shown in separate clustered boxes.

Eventually, the fragments are classified using the scaffold hierarchy and presented using a tree map. Six clusters were identified with more than three members, with 25 hits remaining as singletons. 17 compounds from the major scaffolds and two of the singletons were selected based on availability in our lab and were taken for further analysis.

4.3 Analyzing binding using NMR spectroscopy

The high rate of false-positives from DSF demands validation of hits with orthogonal biophysical and structural methods. Ligand-observed and protein-based NMR have been shown to be used successfully to separate false positives from true binders (see section 2.10) [244]. For ligand observed NMR, WaterLOGSY, STD and CPMG experiments, which are three of the most common experiments, are used in a triple combination to evaluate ligand binding [244]. WaterLOGSY is the most sensitive of the three experiments but is also more prone to false-positives; therefore it is usually used in combination with the other two experiments especially in the absence of a competitive inhibitor. In WaterLOGSY, all the signals in a spectrum are phased relative to the water signal, which is considered as positive. In the absence of protein, the ligands should show negative signals (Figure 4.4, red curve). In the presence of protein, in case of interaction between the ligand and the protein, the signals corresponding to the ligand should be less negative or positive (Figure 5.5 e.g. A) [244]. CPMG, an experiment that takes advantage of the protein's tumbling, is carried out on two samples one with and one without the protein similar to the WaterLOGSY experiment. If a ligand does bind to the target protein, it should tumble at a slower rate leading to decrease in signal intensity in the spectrum [244]. The STD experiment, which is dependent on the spin diffusion effect, unlike the WaterLOGSY and CPMG experiments, is run on one sample in the presence of the protein. Each STD experiment has two scans, with one scan in the presence of proton saturation pulse and the other scan in the absence of any proton saturation. The difference between the two scans is presented as the difference spectrum. Any signal in such spectrum is indicative of protein-ligand interaction [257]. These methods are discussed in more detail in section 2.10.

All three experiments were carried out for the 17 selected ligands, according to protocols in section 2.10.3. Change in signals of ligand (in the presence and absence of protein) was only observable for the WaterLOGSY experiment and no change could be detected in the STD and CPMG experiments. As shown in figure

4.4, from the 17 compounds no signals were observable even in the 1D ^{1}H NMR spectrum of ligands **6** and **9.** This could be due to severe aggregation of the compounds [188]. This could also be as a result of compound degradation in the library master plates. Compounds **2** and **16** gave positive signals even in the absence of protein in the WaterLOGSY experiment, indicating ligand aggregation and therefore these ligands cannot be studied using ligand observed NMR. From the remaining 13 fragments, compounds **1**, **3**, **4**, **12** and **15** showed complete or partial reversal of signal phase in the presence of protein (Figure 4.4). These ligands were taken for further studies

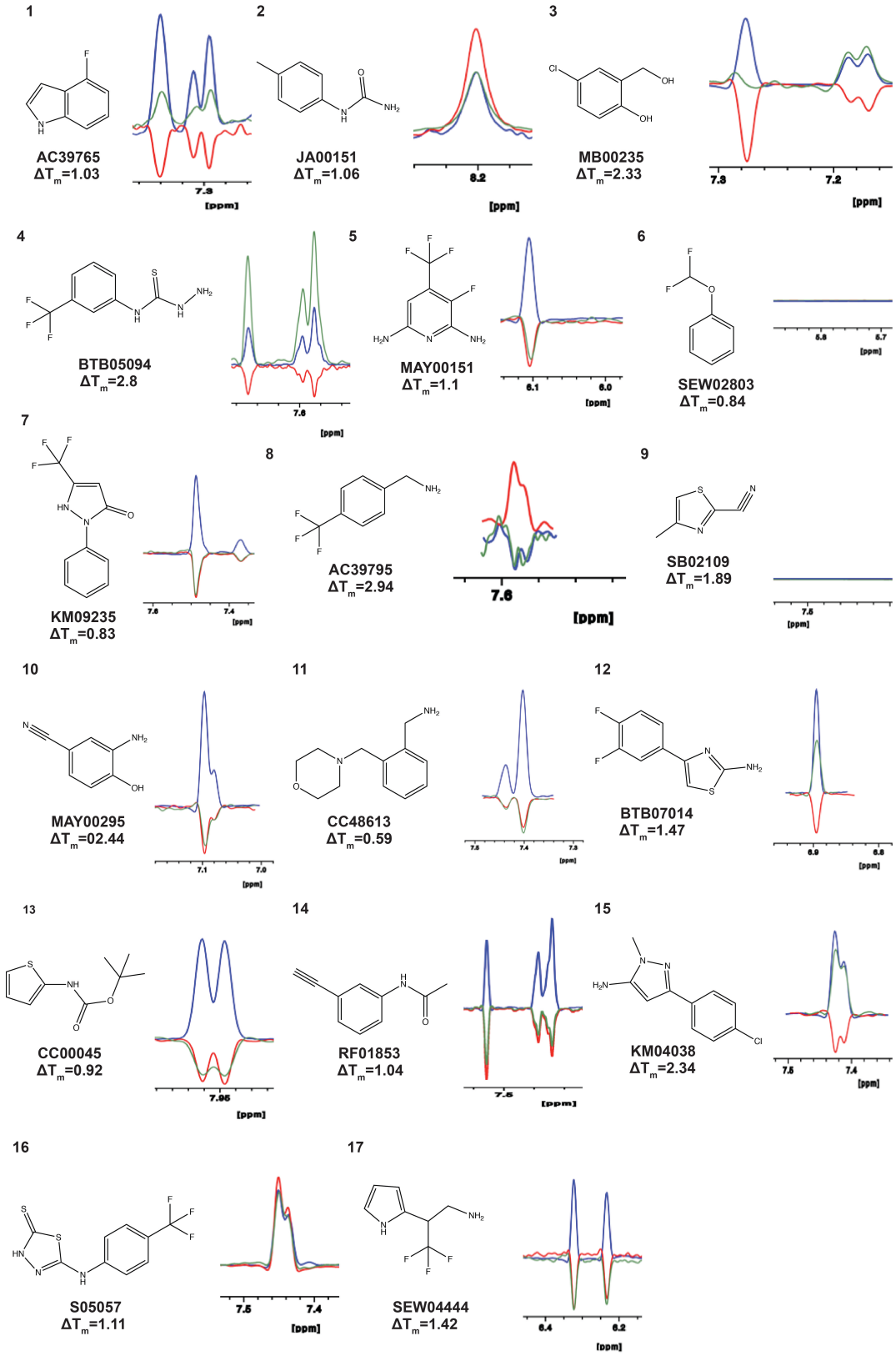


Figure 4.4. Evaluating fragment binding to Rab27a using WaterLOGSY experiment. Binding of cherry picked fragments for follow up studies was evaluated using the ligand-observed WaterLOGSY experiment. Blue curve corresponds to 1D ^{1}H of ligand in 20 mM HEPES, 5 mM $MgCl_2$ and 150 mM NaCl. Red curve corresponds to WaterLOGSY of 1 mM ligands in the absence of protein. Green curve corresponds to WaterLOGSY of ligands in the presence of 10 μM Rab27a. All spectrums were phased with respect to the water signal which was assigned a positive phase. In the absence of protein, the signal for any ligand should be negative. In the presence of protein, the negative signal should decrease or become positive for a binder. Only selected peaks from the aromatic protons for each ligand are shown.

In order to see if any change in signals could be observed in the STD and CPMG experiments for the 5 compounds showing binding with WaterLOGSY, modifications were made in the experimental set-up. The STD and CPMG were run at 5°C to decrease the tumbling rate and enhance the spin diffusion effect. The STD experiment was also carried out in 100% D_2O since it has been reported that residual water can suppress the spin diffusion effect in proteins and decrease experimental sensitivity by four-fold. From the five fragments tested, compounds **1**, **3** and **15** showed signals in the STD and the CPMG difference spectrums (Figure 4.5), while compounds **4** and **12** still showed no signals in the STD difference spectrum or any change in the CPMG acquired spectrums after the addition of protein.

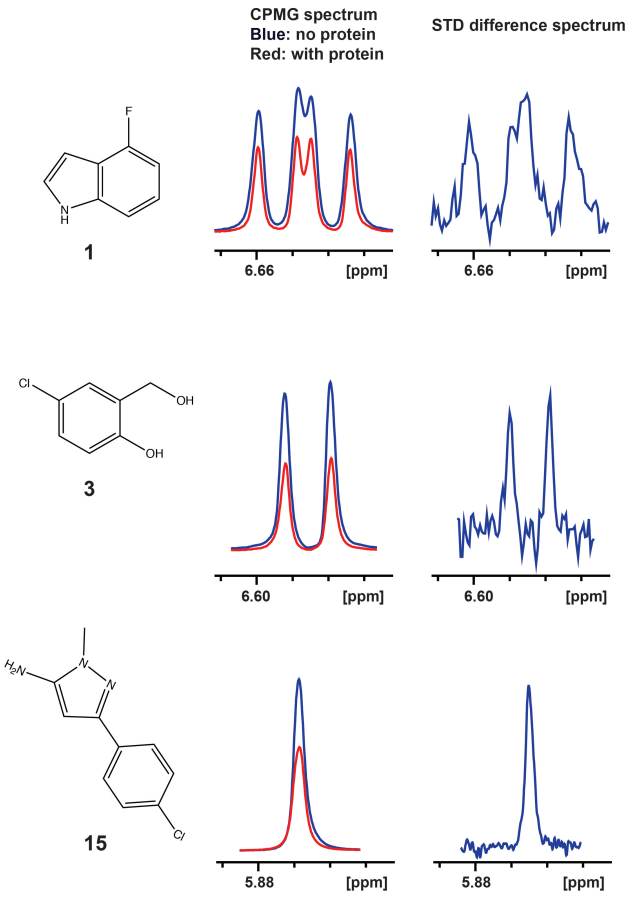


Figure 4.5. Evaluating fragment binding to Rab27a using STD and CPMG experiments.

Compounds **1**, **3**, **4**, **12** and **15** binding to Rab27a were evaluated using STD and CPMG experiments under modified conditions. Experiments were carried out at 5°C, with STD experiments run in 100% D_2O. As demonstrated, compounds **1**, **3** and **15** showed binding under new conditions but **4** and **12** did not show any binding.

To further confirm the data obtained from ligand-observed NMR, protein-observed NMR was used to evaluate binding of ligands to Rab27a. Figure 4.6 depicts ^{15}N HSQC spectrums of 100 µM Rab27a overlaid with Rab27a in the presence of 2 mM ligands. As illustrated, compound **4** did not induce any chemical shift change in the spectrum, thus corroborating the data form ligand-observed NMR experiments that this ligand does not bind to Rab27a. Compound **3**, **12** and **15** all induced slight shift in 5 peaks in the spectrum, with compound **12** having the least effect.

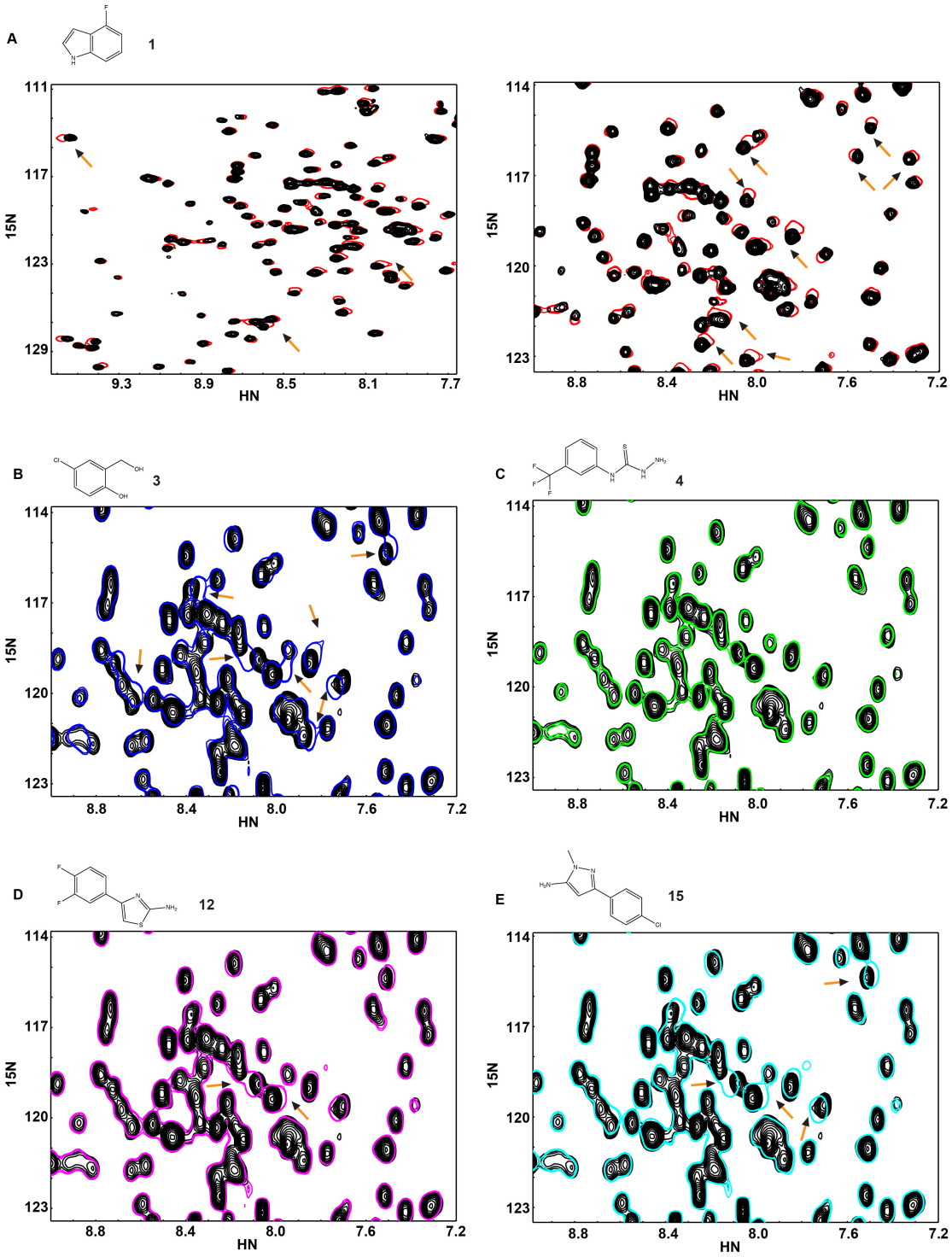


Figure 4.6. ^{15}N, ^{1}H HSQC spectrum overlay of Rab27a in the presence and absence of fragments. Compounds 1, 3, 4, 12 and 15 binding to Rab27a were evaluated using protein-observed NMR. As demonstrated, while compounds 1, 3 and 15 induced change in chemical shift of peaks in the spectrum, compounds 4 and compound 12 showed minute effects. Control spectrum containing 100 µM Rab27a+5% DMSO is shown in black. Overlapping spectrum refers to 100 µM Rab27a+ 2 mM ligand at 5% DMSO concentration.

In contrast to the other ligands, Compound **1**, which is an indole derivative, induced considerable change in multiple peaks. This could be that at high concentrations this ligand binds to multiple sites. An interesting observation from DSF screens in our lab, has been the frequency that compounds **4**, **12** and similar analogues showed up as binders to completely different targets. This is an indicator of the promiscuous binding nature of these ligands. Compound **4** and related analogues, all induced large shift in T_m, but did not demonstrate binding to any of the targets by NMR or other methods. Compound **12** is an amino-thiazole derivative. The amino-thiazoles have been shown to be promiscuous, non-specific binders to other target proteins [258].

4.4 Binding site identification by monitoring methionine residues using NMR

The lack of an assignment for the ^{15}N, ^{1}H NMR spectrum prevented us from using NMR for ligand binding site identification. Before taking the ligands for crystallography, we tried to obtain more information on the ligand binding sites through ^{13}C NMR.

As discussed in section 1.10 and 3.3, the SF4 pocket is an important region for Rab27a-Slp interactions. Moreover, analysis using FTMap web server proposed this site as the most ligandable site on the Rab27-effector interface. Similarities between the 4-fluoroindole (**1**) hit and the tryptophan side-chain of Slp, motivated us to see if any of compounds **1**, **3**, **12** or **15** bind to residues forming this pocket. In the absence of ^{15}N, ^{1}H HSQC assignment this is not possible. But analyzing the residues forming the SF4 pocket, indicated two methionine residues, M93 and M185 are present in the pocket (Figure 4.7). Therefore, we used peaks corresponding to M93 and M185 methyls in ^{13}C, ^{1}H HSQC spectrum to assess the binding of compounds to the SF4 pocket. For chemical shift assignment of M93 and M185, we took a two-step approach. In the first step, using a ^{13}C labeled sample, we collected a ^{13}C, ^{1}H spectrum using constant time HSQC pulse sequence of hsqcctetgpsp.2. Peaks corresponding to methionine methyls were identified, by having a 180^{o} phase difference from all the other ^{13}C-^{1}H peaks in the spectrum. Subsequently, using mutagenesis we assigned M93 and M185 methyl peaks (Figure 4.7).

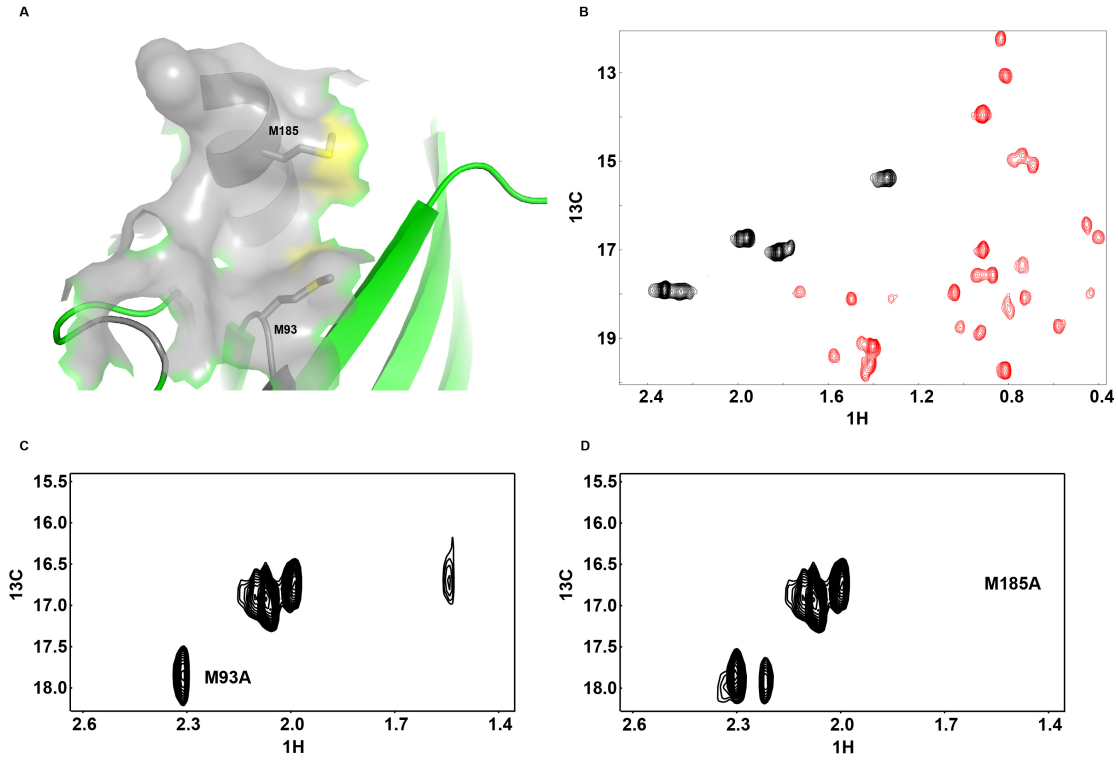


Figure 4.7. Assigning M93 and M185 methyl chemical shifts using mutagenesis. A) M93 and M185 sidechains are involved in forming the SF4 pocket. B) Peaks corresponding to Methionine methyls (shown in black) in 13C, 1H HSQC spectrums can be distinguished from other 13C-1H peaks (shown in red) using a constant-time HSQC. Methionine methyls will have 180^{o} phase difference with the other peaks in the spectrum. C, D) 13C, 1H HSQC Spectrums of M93A and M185A Rab27a respectively. In Each mutant the peak corresponding to the methionine residue disappears.

^{13}C, ^{1}H HSQC spectrums of 50 μM Rab27a were collected in the absence and presence of 2 mM concentrations of compounds **1**, **3**, **12** and **15**. From the fragments, only compound **1** and **15** induced chemical shift on M93 and slightly on M185. Compounds **3** and **12** did not induce any change (Figure 4.8). Subsequently, a titration using compounds **1** and **15** was carried out. While compound **1** demonstrated binding in a titratable manner with an estimated affinity of 2.1 mM, compound **15** displayed a complex binding behavior (Figure 4.8). This could be a result of non-specific binding, or binding to a near binding site and inducing a secondary effect on the methionine methyls.

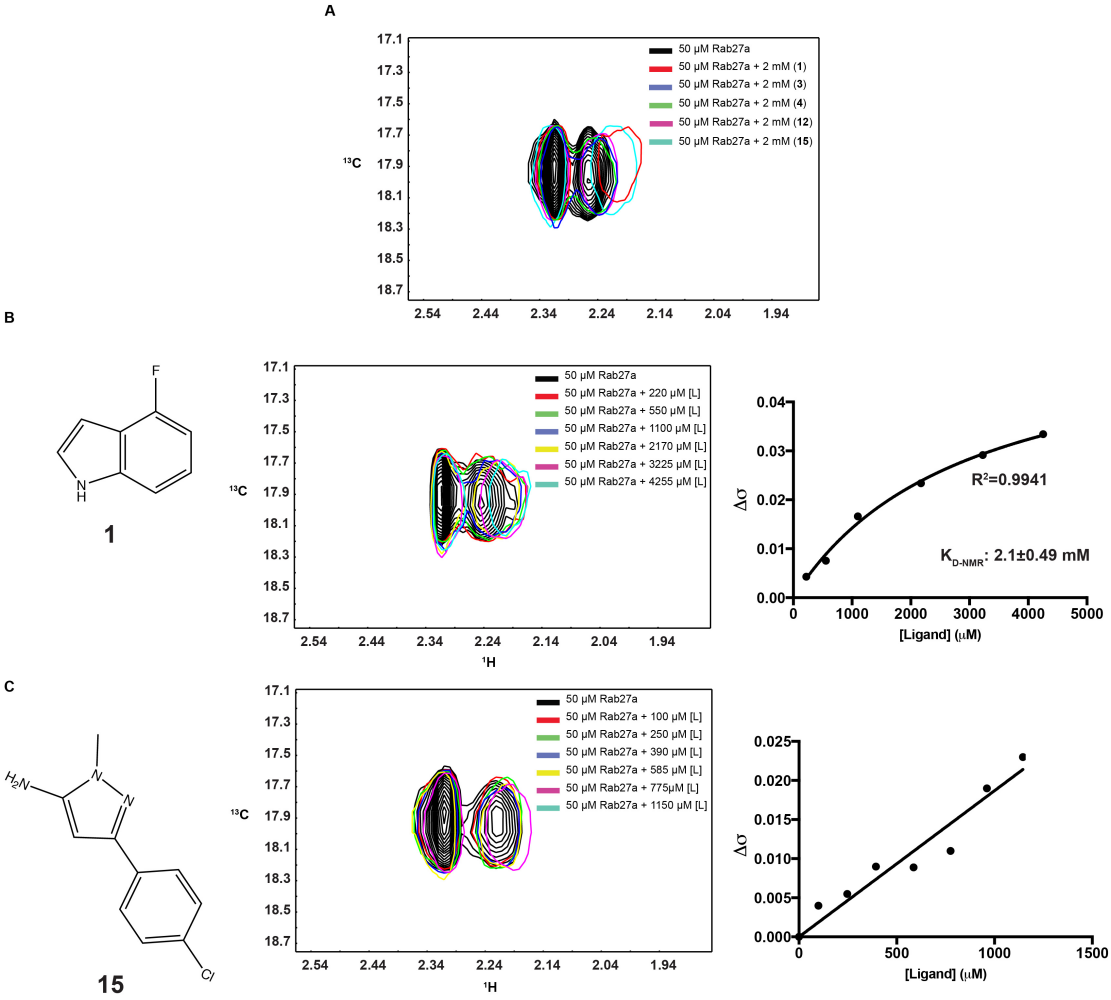


Figure 4.8. Monitoring fragment binding to Rab27a using ^{13}C, ^{1}H HSQC. Binding of fragments to Rab27a was monitored using M93 methyl chemical shift change, in ^{13}C, ^{1}H HSQC. A) compounds **1** (red), **3** (blue), **4** (green), **12** (magneta) and **15** (cyan), were added at 2 mM concentrations to 50 µM Rab27a. From the 5 fragments, only compounds **1** and **15** induced any change. There two compounds were titrated to measure K_D. B) Compound **1**, demonstrated a stoichiometric binding with a K_D=2.1±0.49 mM. C) Compound **15**, did not bind in a titratable manner, suggesting its effect at 2 mM is either non-specific or the effect on M93 is an indirect secondary effect.

4.5 Binding mode characterization using X-ray crystallography

The lack of a fully assigned ^{15}N, ^{1}H NMR spectrum prevented us from using NMR for ligand binding site identification. In order to identify the binding site and binding mode of compounds **1**, **3**, **12** and **15**, $Slp2_a$-Rab27a fusion crystals were soaked with these compounds at saturating concentrations in the presence of 5-10% v/v DMSO for 24-48 hours. All crystal preparation, data collection, processing, model building and refinement were carried out according to section 2.12 and 3.3.1. All the statistics for the collected datasets are available in Appendix C. Although, compounds **4** did not show any binding with NMR, it was also used for crystallization studies (Figure 4.9).

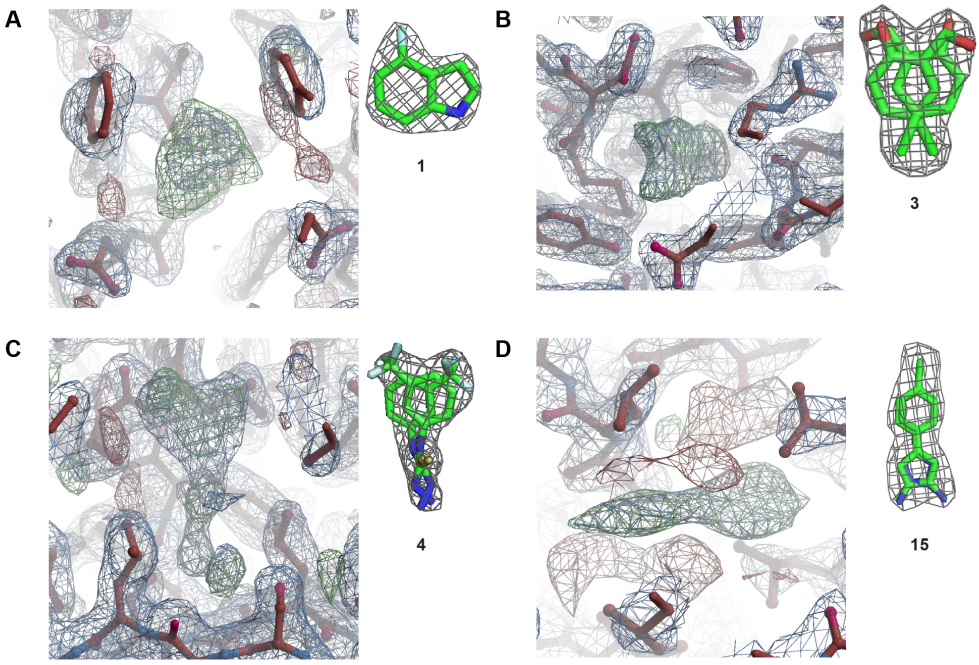


Figure 4.9. Soaking $Slp2_a$-Rab27a fusion crystals with compounds 1, 3, 4, 12 and 15. A, B, C and D demonstrate the presence of electron density (in green) from the F_o-F_C map for ligands in datasets obtained from diffracting crystals soaked with compounds **1**, **3**, **4** and **15**. Electron density was also observable for compound **12**, but data is not shown.

Interestingly, electron density was recognizable for all five compounds in the initial model obtained from phaser (Figure 4.9), even for compounds **4** that did not show any binding by protein-observed NMR. An initial observation on all the ligand binding sites is that, they all bind to residues in an artificial pocket. This pocket is located on the 2-fold NCS axis, between the α_2 helices of two Rab27a proteins and two $Slp2_a\,RBD_{1\alpha}$ domains in the fusion proteins (Figure 4.10). As explained in chapter 3, this pocket is an artifact of $Slp2_a$-Rab27a crystal form. Since the molecules all bind to residues located on a 2-fold symmetry axis, two molecules were built in each electron density representing conformations binding to both Rab27a proteins.

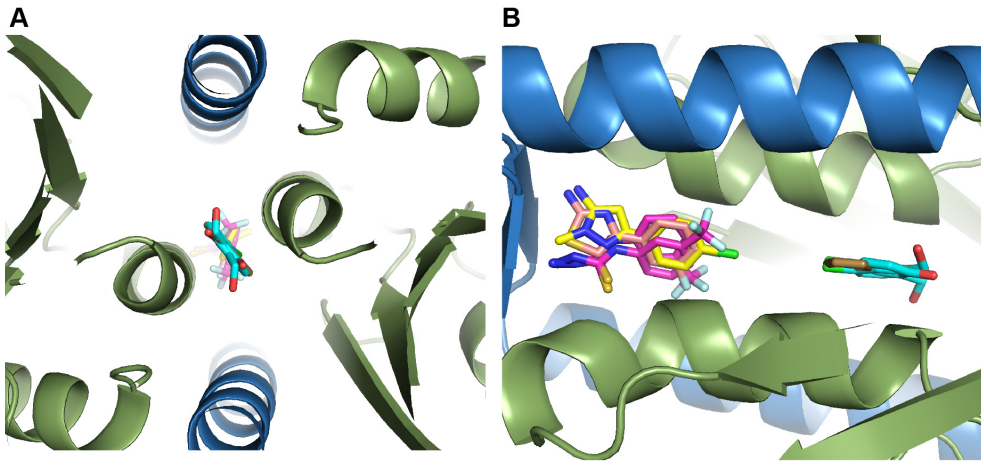


Figure 4.10. Fragment binding sites in $Slp2_a$-Rab27a crystals. Compounds from the DSF screen that were taken for crystallography, all occupied two sites in the $Slp2_a$-Rab27a fusion crystals. Both sites are located on the 2-fold NCS axis. Demonstration of the ligand binding site top view A) and side view B). The compounds bind to a pocket made of the α_2 helices of two Rab27a molecules (green) and the fused $Slp2_a\,RBD_{1\alpha}$ domains (blue). This artificial pocket is a consequence of construct design and crystal form, obscuring the real binding mode.

Ligand binding to protein surface on a symmetry axis is generally representative of a biased and unreliable model, not reflecting the true binding mode of the ligands [259]. Analyzing the binding sites, it is clear that the ligands cluster at two sites. Compounds **4**, **12** and **15** all interact with a hydrophobic environment created by L84, A87 and F88 of Rab27a in both chains and S2 (pdb numbering:-35) and L4 (pdb numbering:-36) of $Slp2_a\,RBD_{1\alpha}$ in the artificial pocket (Figure 4.10 and 4.11). This clearly demonstrates that the beta strand originating from the fused $Slp2_a\,RBD_{1\alpha}$, is influencing the binding mode of these compounds. In contrast, compounds **1** and **3**, occupy a very shallow site beneath the SF4 pocket. Compound **3**, interacts with D91 of one Rab27a molecules and Y122 of the other Rab27a molecule and has

its para-benzyl chlorine stacked against F88 of both Rab27a molecules (Figure 4.11). Compound **1** shows similar interactions with Y122 and F88. Both structures suggest such a binding mode that would not be possible or very unlikely in solution.

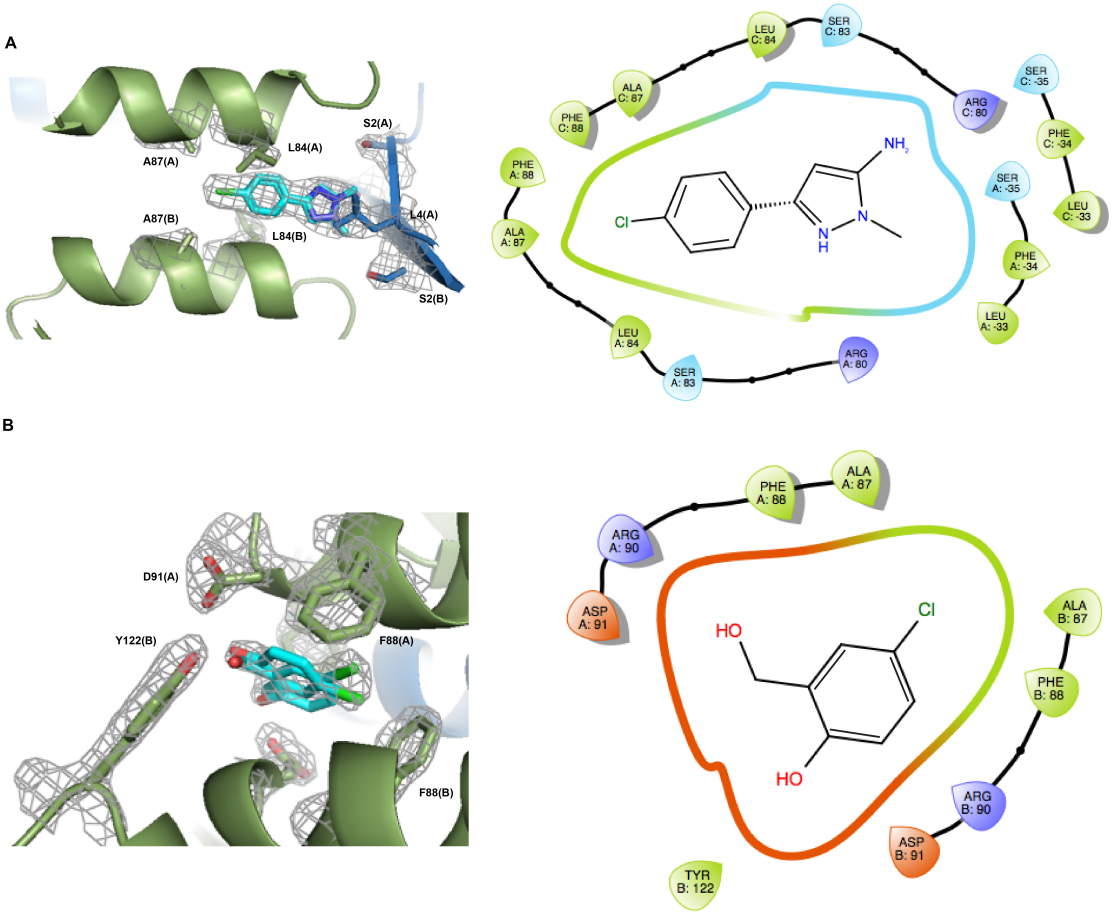


Figure 4.11. Characterization of fragments binding modes using Slp2$_a$-Rab27a fusion crystals. A) Demonstrates the Crystal structure of compound 15 with Slp2$_a$-Rab27a. The ligand binding-mode is artificial since it sits on a symmetry axis and also makes interactions with the Slp2$_a$ RBD$_{1\alpha}$ segment of the construct. Compounds 4 and 12 also adopted a similar artificial binding mode. B) Demonstrates the Crystal structure of compound 3 with Slp2$_a$-Rab27a. Similar to A, the compounds sit on the NCS axis interacting with two symmetrical proteins.

We concluded from the results, all the crystal structures showed artificial binding sites and binding modes. Therefore, several changes were made to the construct in an attempt to obtain the real binding sites for each ligand. As explained in chapter 3, section 3.3.1.2, Y122 and R90 on chain A are in contact with D21 on RBDα$_1$ of chain B in the fusion structure (Figure 3.7). This suggests that the RBDα$_1$ fused to Rab27a could be biasing the orientation of Y122 side chain on Rab27a under crystallographic conditions. Therefore, D21A mutation was made on the Slp2$_a$ RBD$_{1\alpha}$ to remove any imposed bias on Y122 (Figure 3.7). No electron density could be

identified for Y122 and R90 sidechains. We concluded, in the absence of an effector these residues are flexible. The second change was removing the "SFL" amino acids from $Slp2_a$ RBDα_1. These three residues form a beta strand under crystallographic conditions and are involved in making several contacts with their NCS related residues from the other RBDα_1 in the ASU (Figure 4.12). This was proposed to remove the artificially deep hydrophobic pocket formed between the two $Slp2_a$-Rab27a fusion molecules and thus eliminate artificial attraction of hydrophobic ligands into the pocket (Figure 4.12). The construct was designed, protein expressed according to methods in section 2.3.1 and crystalized under similar conditions to the original fusion construct. The new crystals diffracted to a resolution of 2.3 Å. All the statistics for the collected datasets are in Appendix C. The crystals were soaked with compounds **1**, **3**, **4**, **12** and **15** up to 7 days to maximum solubility. No recognizable electron density was identified for any of the ligands in datasets obtained from diffracting these crystals. This is potentially due to the weak binding of the fragments, with saturation levels not achieved within their solubility limits. These data re-confirm the idea that the previous obtained structures do not represent a true binding mode.

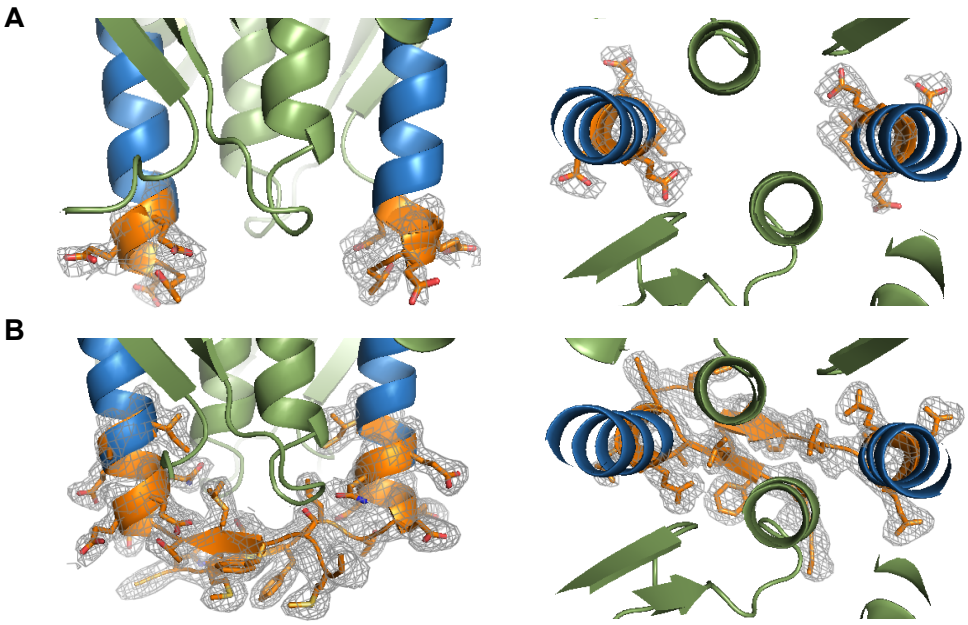


Figure 4.12. Comparison of ΔSFL-$Slp2_a$-Rab27a and $Slp2_a$-Rab27a crystal structures. Rab27a (green), $Slp2_a$ RBDα_1 (orange-blue). As illustrated in A) the N-terminal part of $Slp2_a$ RBDα_1 in ΔSFL-$Slp2_a$-Rab27a (orange) does not have the "SFL" amino acids [which form a short beta sheet between the NCS related molecules] in comparison to $Slp2_a$-Rab27a as shown in B). The absence of "SFL" in ΔSFL-$Slp2_a$-Rab27a, removes the artificially deep pocket formed between the two $Slp2_a$-Rab27a fusion molecules,. This eliminates artificial attraction of hydrophobic molecules into the pocket.

4.6 Discussion

Fragment-based ligand discovery has become a popular approach for drug discovery. Reasonable access to a small [sized] fragment library in academic institutes has enabled researchers to screen for binders using a range of biophysical and, if available, structural biology methods [237]. A drawback of this approach is the low affinity of hits, therefore heavily relying on structural data [NMR or crystallography] to understand binding modes and assist cycles of medicinal chemistry to obtain ligands with affinities in appropriate range to be studied for cellular effects [209].

Differential scanning fluorimetry (DSF) is a low cost, fast and high throughput method that monitors protein denaturation, using a fluorescent dye that has a change in fluorescence after binding to exposed hydrophobic patches of the proteins [244]. High false positive rates in DSF, signify the importance of validation using orthogonal methods [biophysical and structural] to insure filtering out the non-binders. Moreover, the inability to run the assay in a competition format using a reporter molecule or monitor the effect of molecules in a targeted manner towards the pocket of interest, reiterates the importance of using other experiments to identify ligand binding-sites [260].

Although Ras proteins have been targeted using fragment-based approaches using 1D/2D NMR and crystallography, screening Ras superfamily members using DSF has not been reported to date [163]. Since Ras proteins are highly dynamic proteins under going conformational change, NMR can detect transient pockets that form in solution whilst keeping the protein intact. In contrast, DSF is reliant on heating and denaturing of the protein. This raises questions over the protein nucleotide-binding state during the melting process and its effect on the conformation of the protein during denaturation and hence the presence of certain transient pockets. This can be a cause of false-positives or negatives in DSF experiments [261].

Nevertheless, since DSF was the only available method at our disposal for a fragment screening approach, we used a fragment library composed of 2500 compounds purchased from Maybridge chemicals to screen against Rab27a. 102 compounds were identified as hits. The compounds were clustered using an algorithm that applies an iterative removal of the rings based on rules defined by Schuffenhauer at el. and groups them using hierarchy of scaffolds [256].

The majority of the compounds were categorized in 6 groups. 17 compounds were selected, based on chemical diversity and taken for further validation by ligand-observed NMR. WaterLOGSY, STD and CPMG were used in combination to monitor ligand binding. Only 5 compounds from the selected compounds showed any response, only for the WaterLOGSY experiment. Both the STD and CPMG experiments did not record any effect. While WaterLOGSY is known to be a much more sensitive experiment, it is also known to be more prone to false-positives. If there was a positive control that could be used as a reporter molecule, the ligand binding could be evaluated more thoroughly and K_D could also calculated for each fragment through competition assays. Since the STD effect relies on spin diffusion, which can be modulated by temperature and solvent properties, the assay was repeated at 5^{o}C in 100% D_2O [WaterLOGSY experiment needs to be done in water since magnetization is transferred from water molecules to ligands]. This can improve spin diffusion by slowing the tumbling rate and also avoiding spin diffusion "leakage" due to dipole effects from H_2O protons [262]. Therefore, STD and CPMG were repeated under new experimental conditions for compounds **1**, **3**, **4**, **12** and **15**. These changes enabled us to detect signal in the difference spectrum of STD and CPMG for compounds **1**, **3** and **15**. The results were further confirmed by ^{15}N, ^{1}H HSQC NMR experiments. The results obtained from 2D NMR, also confirmed that compound **4** does not demonstrate any binding while compound **12** is a very weak binder. This is consistent with previous publications, reporting amino-thiazole and thio-semicarbazides as promiscuous non-specific binders [258]. This was also observed from experience within our research group, as several thio-semicarbazides appeared as initial hits binding to a range of proteins from fungal adhesins and bacterial chaperones to prenylation enzymes in the DSF assay. In fact the main group from our DSF screen cluster analysis is comprised of multiple benzyl thio-semicarbazide derivatives, which all induce a significant shift in the T_m in multiple proteins but do not show any binding activity in more precise methods such as protein-based NMR.

Similarities between the 4-fluoroindole and tryptophan side-chain prompted us to see if any of these ligands target the SF4 pocket. This was accomplished by monitoring change in methionine methyl chemical shift using ^{13}C, ^{1}H HSQC. M93 and M185 were assigned using mutagenesis. Only 4-fluoroindole (**1**) demonstrated a titratable binding to Rab27a. This result was confirmed by comparing the binding of 4-fluoroindole to Rab27a-$Slp2_a$ complex and to binding of Rab27a using WaterLOGSY experiment. The binding of an indole derivative to a

shallow pocket on a protein-protein interaction interface is not surprising as several proteins use tryptophan amino acids to interact with binding partners [263]. Scott *et al*. also demonstrated indole derivatives as the main hits from a DSF screen on Rad51. These fragments occupied a shallow pocket used to interact with the "FXXA" motif [231].

Studying the binding of the ligands to Rab27a by crystallography did not provide us any information on the true binding sites of these fragments. Ligand bound structures were originally obtained using $Slp2_a$-Rab27a fusion crystals. But a close investigation demonstrated that all the ligands were sitting on a NCS axis, casting doubt over the reality of the binding sites. Modifying the construct resulted in no electron density being identified for neither of the compounds. The low affinities of the ligands, in combination with solubility limitations potentially play an important factor, as complete occupancy could not be achieved.

Structural information is key in understanding the mode of fragments to proteins. Apart from crystallography, NMR is another route in acquiring information on ligand binding sites. A lot of effort in this project was put towards obtaining the assignment of the ^{15}N, ^{1}H HSQC spectrum of Rab27a. These were not successful (see chapter 6 for discussion on this matter). Nevertheless, in the absence of structural information other approaches can be taken to identify the binding sites of ligands. Mutagenesis can help in verifying potential binding sites of ligands. Similarity searching and SAR by catalogue is another popular approach in fragment based drug discovery that could assist in identifying key pharmacophores important for interaction and help understand the environment of a binding site. [256]. A similar approach could be taken for 4-fluoroindole (**1**). Scott et al. used a 42-fragment library of indole-like compounds for a SAR by catalogue approach to follow-up the hits obtained through DSF assay on Rad51.

The structural and dynamic characteristics of Rab27a suggest, finding hits for such a protein and modifying them into tight binding specific ligand is very challenging. Results obtained from DSF did not yield very promising results. This could be because DSF is not a suitable experiment to study ligand binding to a dynamic protein such as Rab27a. Therefore, NMR is probably a more suitable screening method for Rab27a. In collaboration with the Beatson institute we have recently, carried out a fragment screening by NMR. The results from this screen have produced diverse scaffolds that are currently under investigation through SAR by catalogue.

Although, the indole compound obtained from the DSF screen is not an attractive starting point for medicinal chemistry purposes, it could be an invaluable tool for future studies. As mentioned, rescreening the protein using a mini-library composed of indole derivatives might lead to more chemically attractive molecules. This compound could also be used as a reporter molecule in other screening campaigns such as ligand-based ^{1}H 1D NMR experiments. Moreover, the fluorine atom on the ligand provides another option, to exploit ^{19}F for ligand discovery purposes [264].

CHAPTER FIVE: DESIGN AND SCREENING OF A FOCUSED LIBRARY AGAINST RA27A USING SPR

5.1 Introduction

Rab27a interacts with Synaptotagmin-like proteins (Slps) using large hydrophobic patches and electrostatic interactions over 1970 Å surface [137]. Similar to other Ras proteins, Rab27a interaction interface is mostly flat and featureless, turning attempts to design ligands that can compete with effector binding into a daunting challenge. This is in contrast to some of the known protein-protein interaction interfaces, such as the Bcl-2/Bcl-xL, which has a well-defined groove [265]. Based on available structural data a unique characteristic of the Rab3 family of Rabs including Rab27a is the presence of a shallow pocket known as the SF4 pocket which accommodates interaction with two aromatic residues present in all relevant effectors [59, 137]. As shown in chapter 3 apart from the nucleotide-binding pocket, the SF4 pocket is one of the few available sites suggested by FTMap calculations as a potential ligandable pocket.

Computational methods have played a major role in drug discovery over the last twenty-five years, contributing to different aspects from ligand screening to binding mode characterization, library design, ligand optimization, pharmacokinetic and toxicological evaluation to name but a few [266]. Computer-aided ligand screening methodologies can be classified into structure-based and ligand-based methods [266, 267]. Structure-based methods calculate energetics of interactions between the protein and ligand using defined models and force fields [268, 269]. The structure of the protein could be provided from high-resolution crystal structures, NMR ensembles or homology models based on available similar structures. Virtual screening of ligands through docking hundreds of thousands of compounds on a defined site in a protein structure is a popular approach taken by several groups on various targets. Although this approach has been successful for several targets such as the GPCRs, it is partly due to the well-defined pockets available for these proteins [270]. Moreover, due to extended medicinal chemistry carried out on ligands attracting such targets, the chemical databases are enriched with organic molecules covering favorable chemical space for such targets. Apart from docking, ab-initio methods, molecular dynamic simulations and the recently suggested method dynamic undocking have been used to obtain initial hits for further structural and chemical characterization [266].

Ligand based methods are another approach that could be pursued in the absence or even presence of protein structural information [266]. Although protein-based methods have drawn a lot of attraction, a recent study has revealed that ligand-based methods in fact yield more potent compounds [271]. In the presence of activity information, machine-learning methods are used to predict the activity of newly synthesized compounds. In the absence of any activity information, search for new compounds could be based on pharmacophore mapping (which would be more efficient in the presence of several known binders) or similarity based search methods [198]. A similarity-based approach can be carried out even in the presence of only one known binder, with no pharmacophore information is required to be given as input data. Due to the use of a numerical scoring function, the degree of similarity can be varied based on experimental conditions. Any molecular descriptor such as 2D fingerprints, topological indices or physicochemical properties could be used in theory for similarity measurements. Maximum common subgraph and reduced graphs are other 2D approaches taken for similarity measurements [198].

Because molecular interactions are based on 3D recognition, ligand conformation can also be used for similarity searching [198]. Ligand shape is one of several 3D ligand-based approaches available in similarity methods. Shape-based similarity methods hinge on the idea that molecules with overlapping volumes have similar shape and molecules with similar shape could have similar activity. While the idea of comparing molecular shapes is trivial, calculating molecular shapes can be problematic. Describing shape with Gaussian mathematical models instead of hard-spheres has enabled a smooth shape description, paving the path for more accurate shape comparison [272].

Previous in-silico work done on Ras superfamily members has mainly focused on virtual high throughput screening (vHTS) through docking a large library of compounds on to a defined site using the crystal structure of interest [163]. Gao et al. used a combination of ligand fitting and flexible docking to find a ligand that inhibits Rac1-GEF interaction [273]. This ligand has a LogP of 5, suggesting a more thorough validation process should be taken before using it for cell-based studies. Rab27a interaction with melanophilin has been the subject of virtual screening by docking [274]. The results from this work are far from convincing, as no data proving the interaction of the hits with Rab27a or inhibition of interaction with effector is provided. Unlike other approaches taken for Ras superfamily members, we have taken a ligand-based method using shape-based similarity to find potential binders of Rab27a. Using shape similarity a small focused library was designed and

enriched with compounds that would exploit the chemical space complementary to the SF4 pocket present in Rab27a. The virtual screening was carried out using the software ROCS from the Openeye package. Ligands meeting defined criteria were taken for screening using SPR (chapter 2, section 2.13). Subsequently hits obtained from SPR were taken for further validation by means of NMR spectroscopy. A summary of the work is summarized in figure 5.1.

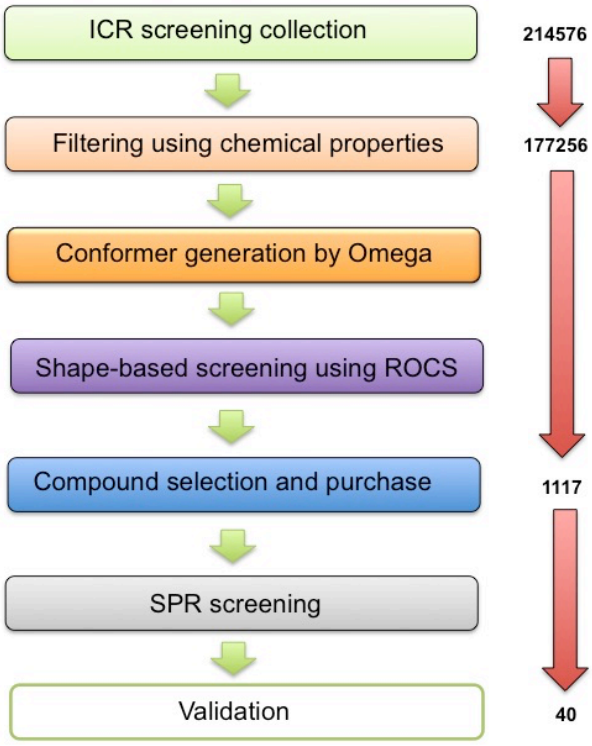


Figure 5.1. Screening Rab27a using a library enriched by shape similarity. As demonstrated, the original ICR library was filtered based on chemical properties similar to the Lipinski rule. Subsequently, conformers of each ligand generated by Omega were overlaid on the query molecule using ROCS. Compounds were selected based on shape similarity using the Tanimoto coefficient for SPR screening. 40 compounds were taken from the initial hits for further validation.

5.2 Library design

In order to prepare the molecules from the ICR vendor collection for screening, the library was initially subject of a filter based on the physicochemical properties of the molecules (Table 5.1). This step was mainly used to filter out thousands of drug molecule and peptides in the library that were not appropriate for hit to lead development. The range for the molecular weight and LogP were defined in a wide range to avoid losing any compound just based on MW or LogP.

Table 5. 1. Filtering using physicochemical properties

Molecular descriptor	Range	Molecular descriptor	Range
H-Acceptors	≤10	Rotatable bonds	≤ 10
H-Donors	≤5	Total molecular weight	180-500
Non-H atoms	20-40	cLogP	(-5)-(+5)
Polar surface area	≤ 140 $Å^2$		

Afterwards the remaining 177000 compounds were taken for screening. Since shape screening requires 3D conformation of molecules, Omega software from the Openeye package was used (chapter 2, section 2.13) to calculate 200 lowest energy conformations for every ligand in vacuum. Since no small molecule binders for Rab27a was available, we used the conformation adopted by "WF" in the crystal structure as the query conformation. Apart from shape, pharmacophores were defined for the query molecule. Analyzing the sequence similarity of $RBD_{2\alpha}$ in different Rab27a effectors it is evident that the presence of two aromatic residues is very conserved. Also, there is a water molecule that connects the nitrogen on the indole ring of tryptophan in the "WF" motf to L9 carbonyl in Rab27a [137]. Therefore, the aromatic rings and a hydrogen bond donor were defined as pharmacophore components of the query molecule (Figure 5.2).

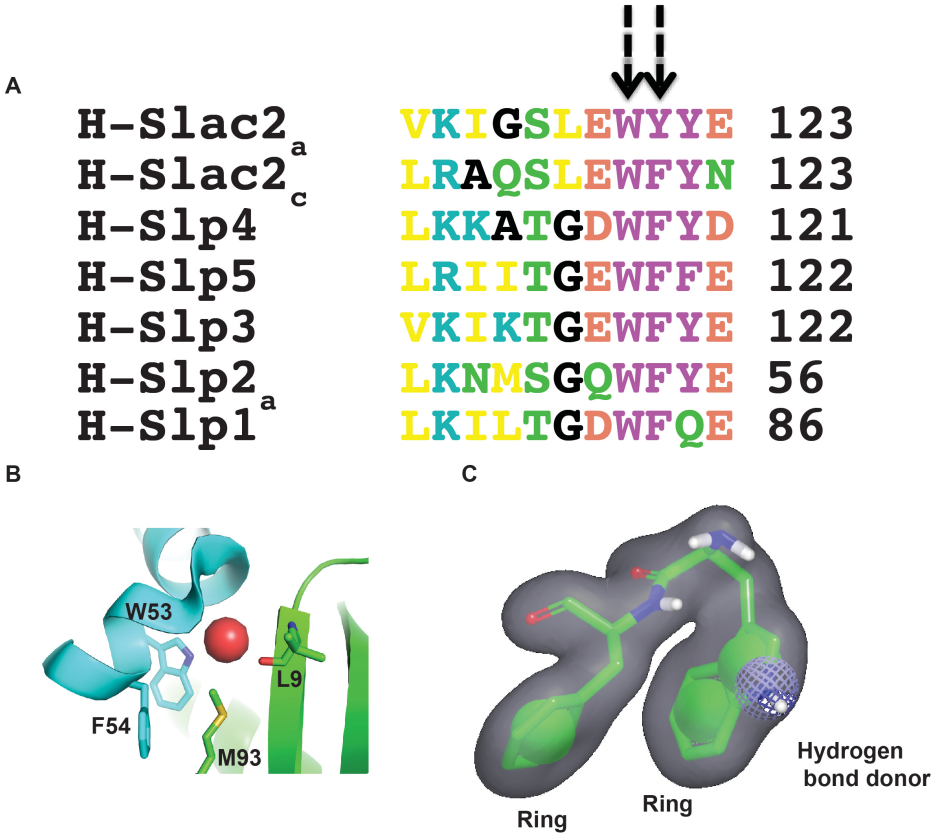


Figure 5.2. Query molecule for shape-based screening. A) Sequence analysis of $RBD_{2\alpha}$ of Rab27a effectors. The presence of two aromatic residues (W and F) that occupy the SF4 pocket is highly conserved. B) A water molecule connects the nitrogen on W53 of $Slp2_a$ RBD to L9 of Rab27a. Therefore, a hydrogen bond donor was defined as a characteristic of the query molecule. C) Demonstration of the query molecule. The shape in combination with two aromatic rings and a hydrogen bond donor, were defined as important aspects of the query molecule. (amino acid color code: R,K:blue, D,E: red, F,W,Y: magenta, A,G,P: white, V,L,I,M,C: yellow, S,T,N,Q,H: green).

Subsequently, shape-based screening was carried out on the library using ROCS software. Detail of the scoring method is given in section 2.13. In every cycle the conformers for each molecule were aligned on the query molecule through the center of mass for both molecules. Both the shape and color force field similarity were given scores using both the Tanimoto and Tversky metrics (chapter 2, section 2.13). A key difference between the two metrics is how they deal with dissimilarities between the query molecule and the ligand of study. In a Tanimoto metric, if the molecule of study has extra functional groups or pharmacophores, which are not defined as part of the query molecule, this will have a negative impact on the score but in a Tversky metric this has no effect. Table 5.2 summarizes the statistics of the virtual screening. As demonstrated, the combination of shape similarity score and color force field score using the Tanimoto metric can be used more efficiently at defining a cutoff to separate compounds for SPR screening. An additional restraint of minimum 0.75 for shape similarity using the Tanimoto metric was added. Values above this threshold represent shape similarity that is visible with the naked eye. At the end 1117 compounds were taken for SPR screening. Overlay of several compounds with the query molecule taken for the SPR screen are shown in figure 5.3.

Table 5. 2. Shape similarity metrics statistics

Title	S-Tanimoto	C-Tanimoto	Tanimoto Combo	S-Tversky	C-Tversky	Tversky Combo
Avg	0.206	0.574	0.789	0.399	0.688	1.087
SD	0.099	0.114	0.102	0.117	0.115	0.0014
Avg+2SD	0.405	0.803	0.994	0.633	0.91	1.089
Count for >	----	----	3509	----	----	42388
>+STanimoto>0.75			1117			-----

S: Shape, C: Color force field, Avg: Average, SD: Standard deviation, >: bigger than

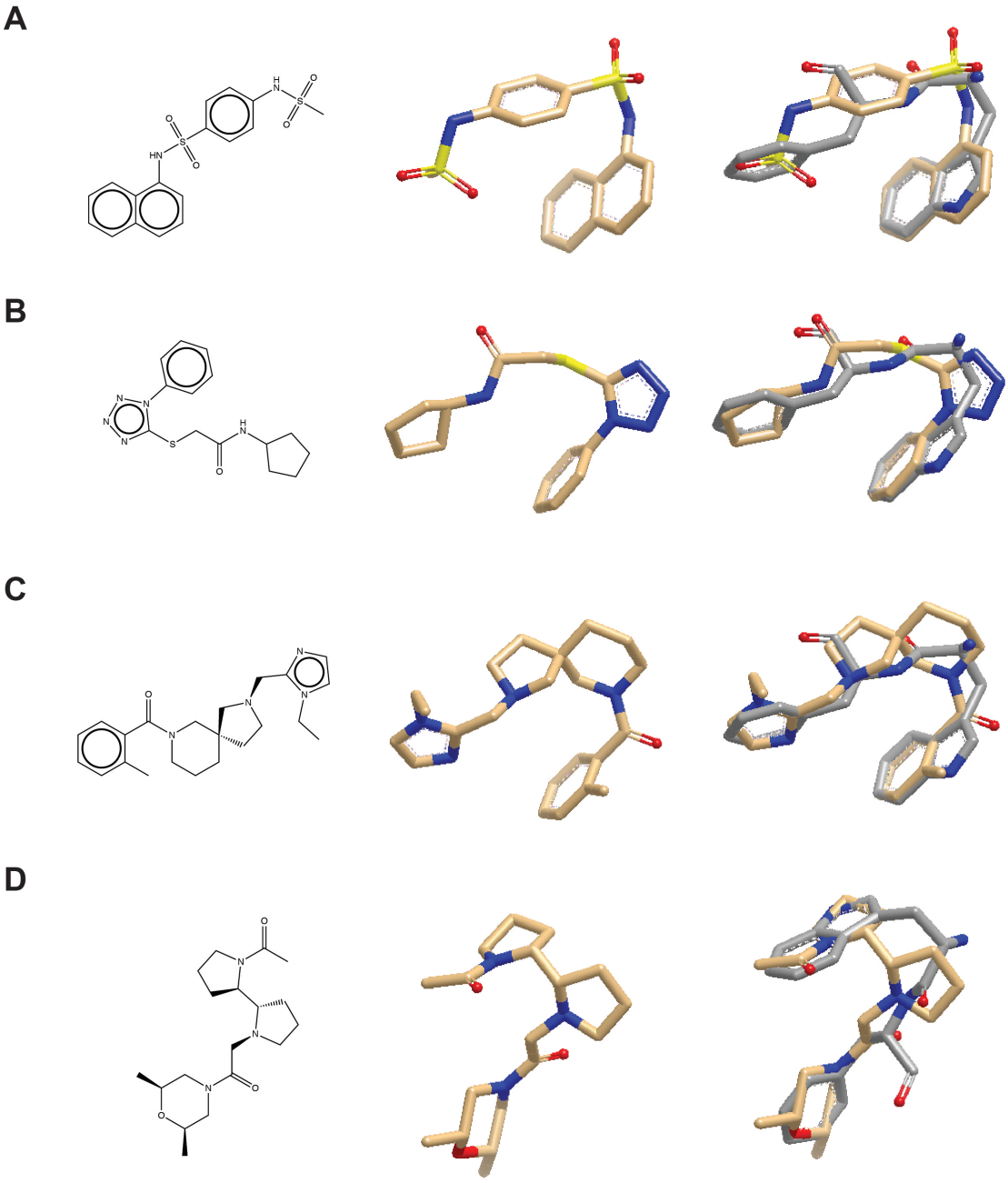


Figure 5.3. Shape overlay of compounds to "WF" conformation using ROCS. Overlay of four compounds with the "WF" conformation, as the query conformation is shown. Compounds were overlaid through the center of mass to obtain maximum shape complementarity.

Figure 5.4 shows the effect of library enrichment using shape similarity on different molecular descriptors by comparing the population distribution before and after screening. As expected, there is an increase in the percentage of molecules with MWof 300-350 as this is close to the MW of the query molecule. Also, although there is a slight increase in the percentage of hydrophobic molecules with 3<cLogP<4, the effect is negligible, indicating the shape screening enrichment is not biased

extensively towards molecules with high number of aromatic rings and greasy functional groups. Also there is a shift in the distribution for number of rotatable bonds, indicating molecules with too many rotatable bonds or too rigid which cannot accommodate the preferred conformation are filtered out.

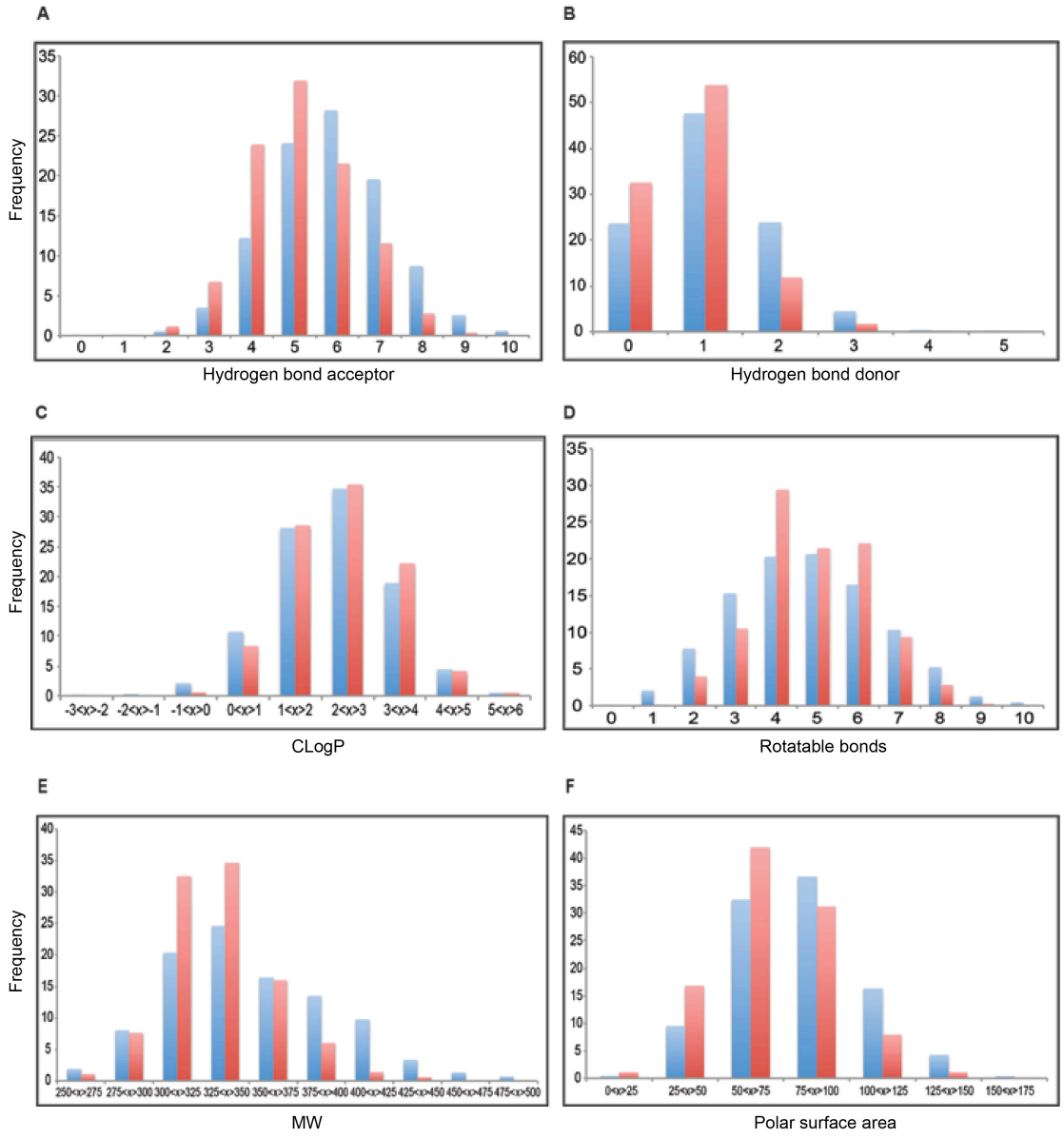


Figure 5.4. Population distribution of chemical properties of ICR library before and after filtering by shape similarity screening. Blue columns refer to the distribution of 117256 compounds from the ICR library prior to similarity screening. Red column show the distribution for the 1117 compounds selected after similarity screening from the initial 117256 compounds. A) Hydrogen bond acceptor B) Hydrogen bond donor C) cLogP D) Rotatable bonds E) Molecular weight F) Polar surface area.

5.3 SPR screen

"Running the screen" was done with split responsibility of Mostafa Jamshidiha and Dr. Charlotte Sutherell. Out of the 1117 compounds selected for SPR screen, only 1071 compounds were available. These compounds were screened in 4 plates according to procedures described in section 2.9.2.

5.3.1 Evaluating Rab27a activity on the chip

Biotinylated Rab27a was loaded on Neutravidin-loaded chips as explained in section 2.9.2. $Slp2_a$ RBD was used to show that Rab27a remains active after using the loading procedure mentioned in section 2.9.2. Figure 5.5, displays the interaction of $Slp2_a$ RBD with Rab27a and Slp2a-Rab27a fusion measured by SPR. As demonstrated, $Slp2_a$ RBD binds to Rab27a-gppNHp and also with a higher affinity in comparison to Slp2a-Rab27a fusion. This information verifies that Rab27a on the SPR chip surface is active. Also, $Slp2_a$ RBD does not show any binding to the reference cell.

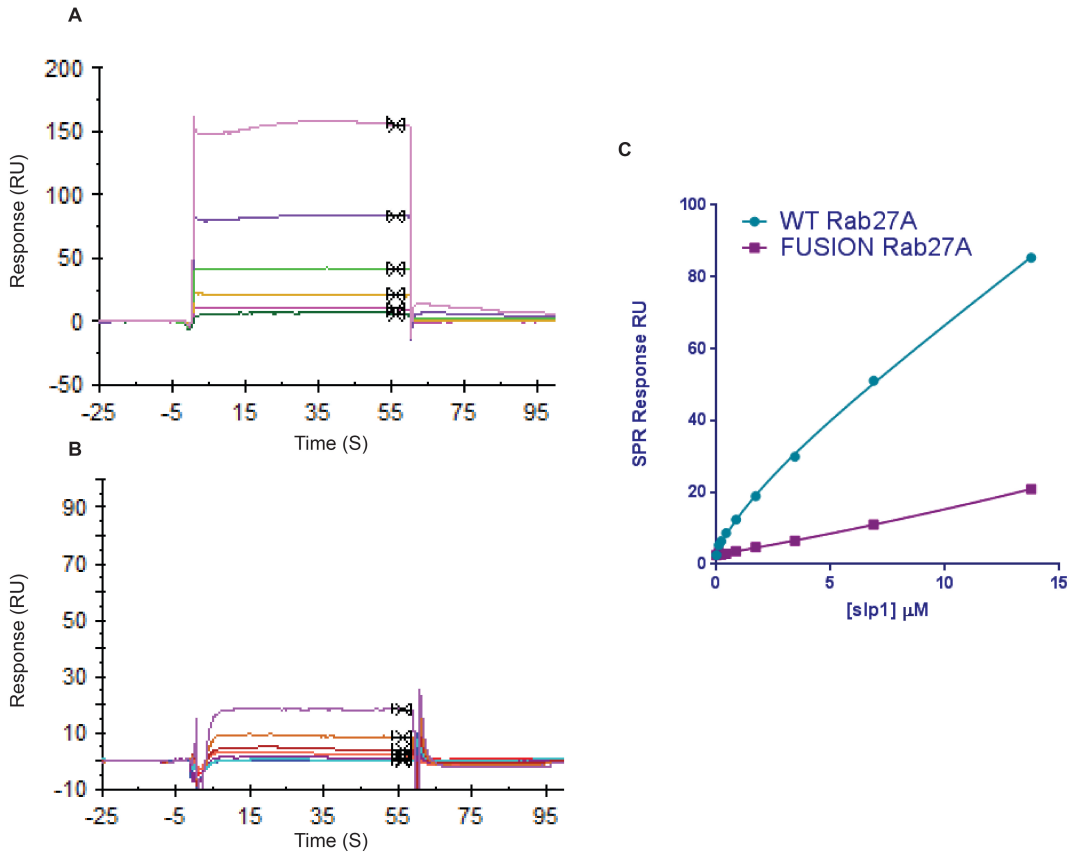


Figure 5.5. Demonstrating the binding of Rab27a-gppNHp to $Slp2_a$ RBD by use of SPR. A) shows the titration of $Slp2_a$ RBD on Rab27a-gppNHp loaded on a Neutravidin coated SPR chip while B) shows the titration of $Slp2_a$ RBD on Slp2a-Rab27a fusion. C) is the titration curve for each of A and B. As demonstrated the RBD binds to Rab27a with a much higher affinity tin comparison to the fusion construct.

5.3.2 Assessing the quality of the SPR run

After each SPR run, the stability of the protein on the chip was checked. This was evaluated through monitoring the decrease in baseline absolute response unit (RU) per cycle and consistent binding of the "WF" di-peptide to Rab27a. A decrease of less than 3 RU per cycle is considered as a stable baseline for ligand screening. The decrease in baseline absolute RU for a complete run of 430 cycles was 360. This translates into 1.19 RU per cycle, which is within the acceptable range during an SPR run. Also the "WF" di-peptide, was used during the screens to confirm the stability of the protein on the chip. The response from the injection every 40 cycles (9 in total for each plate) was consistent and associated with a stable sensorgram (Appendix D).

5.3.3 Clean-screen run to identify screen-interfering compounds

Prior to each screen, a clean-screen was carried out without the presence of DMSO correction. This was used to filter out compounds that bind to the chip irreversibly or have an off-rate much longer than the time between the cycles for two compounds, therefore affecting the results of the next samples. These compounds would produce a jump in the baseline absolute response curve against cycle number. None of the compounds used at 75 μM for screening had any negative impact on the protein or the chip during the SPR run (Appendix D).

5.3.4 SPR results

The screening was carried out in 4 plates. Due to difficulties with the SPR machine during the screening, plate numbers 2 and 3 had to be repeated. Because of the variation between the four plates, each plate was analyzed separately. For data analysis, the binding response, which is a report point at the end of each injection, was measured for each ligand. The cutoff for a hit was defined as average binding response+ 2×standard deviation and compounds with a binding response above the threshold were taken for further validation. In total 51 compounds from all the ligands run in 4 plates had a binding response above the defined cutoff. All the ligand structures and their sensorgrams from the screening are brought in Appendix E. A key point to consider when analyzing SPR data is the quality of the sensorgram. A poor sensorgram can lead to inaccurate values taken as the binding response at

the report point. This needs consideration, because after data extraction, every reading is considered as a numerical value and all are manipulated in a similar way. From the 51 compounds, 40 compounds were selected for further validation based on the shape of the sensorgrams. Figure 5.6 summarizes the structure and sensorgrams of a selected group from the 40 compounds.

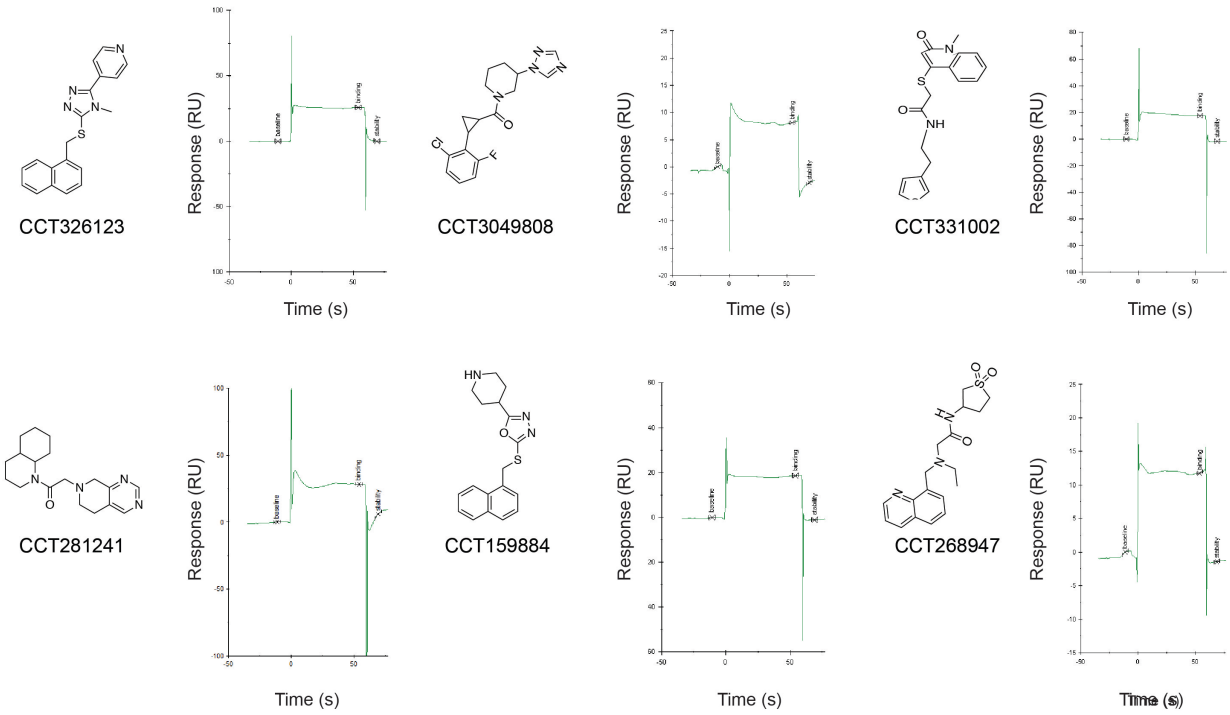


Figure 5.6. Hits obtained through SPR screening. Six compounds from the 40 compounds selected for further validation are shown here. These compounds show a stable line during the steady state of the sensorgram.

5.4 Hit validation

Validating the binding of hits obtained through screening to target protein is the most important step in a screening campaign, which requires use of multiple orthogonal methods. Here, we initially carried out a titration experiment using SPR. Compounds that demonstrated binding in a titratable manner with good quality sensorgrams were taken for analysis by NMR. Also, because of using high concentrations of ligands during SPR and NMR and limited solubility of the chemicals, we carried out aggregation assays using 1D ^{1}H NMR.

In the first step, 40 compounds were titrated against Rab27a-gppNHp by SPR. K_D and the ratio between theoretical R_{Max} and experimental R_{Max} was calculated for each ligand according to methods described in section 2.9.2. All the data from the titrations is brought in appendix F. Ligands with acceptable sensorgrams, K_D below 1 mM and $0.1 \leq R_{max}$ ratio ≤ 10 were taken for further studies. 13 ligands were selected in total, with 6 ligands belonging to the naphthalene-1-ylmethyl sulfanyl compounds and two ligands with the benzodiazol-2-one substructure. Figure 5.7 summarizes these molecules.

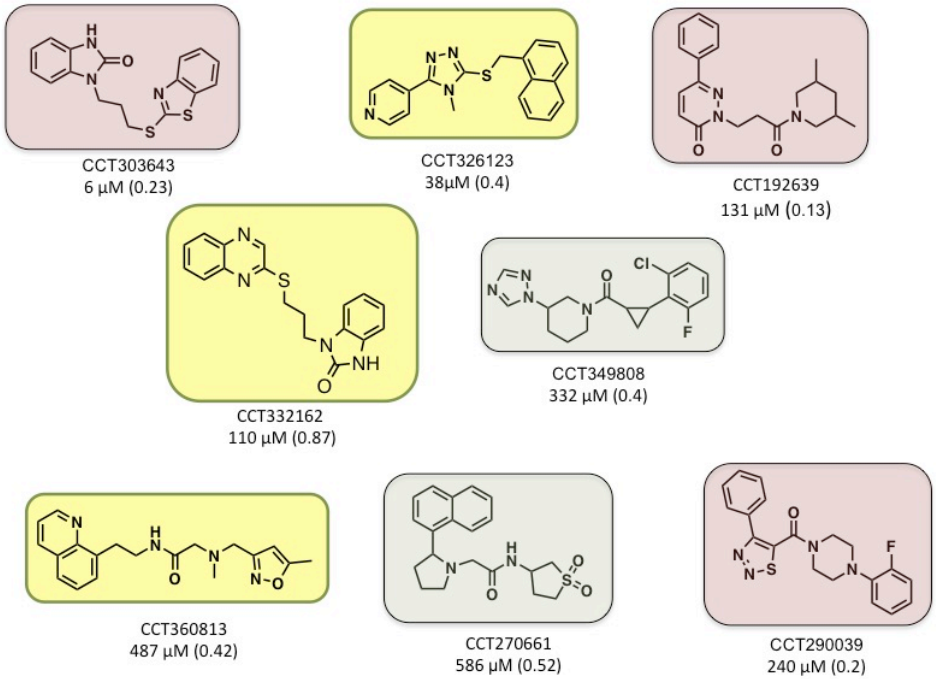


Figure 5.7. Compounds selected from SPR titrations for NMR follow up studies. 13 structures were obtained from the SPR titrations with acceptable sensorgrams, K_D below 1 mM and $0.1 \leq R_{max}$ ratio ≤ 10. 6 compounds are from the CCT326123 family. A box is used to highlight each compound. The color for each box demonstrates the level of aggregation for each compound under experimental conditions. This was measured by monitoring the severity of peak shift during a serial dilution experiment by 1D ^{1}H NMR. This assay was used as an indicator for formation of soluble aggregates of ligand in buffer conditions. No peak shift (gray), small peak shift (yellow), severe peak shift and broadening (red).

Whilst a K_D< 1 mM was used as a criteria for initial compound selection, the values were not used to compare the affinity between hits. This is because repeating of SPR assay and further evaluation by other methods was needed.

The hydrophobic nature of the SF4 pocket, enrichment of our library with compounds containing at least two aromatic rings and the poor quality of sensorgrams for several compounds raised concerns over SPR data (Appendix E). We were interested to see whether binding response in SPR or NMR experiment arises from single soluble molecules or soluble aggregates. Therefore, before taking any ligands for studying by protein-based NMR, we used ^{1}H NMR approach to evaluate formation of aggregates by these ligands at concentrations used for our experiments [188]. The experiments were done according to details in section 2.11. Figure 5.7 summarizes the outcome, in three different categories. Compounds CCT303643, CCT192639 and CCT290039 are in the first group displaying a low SPR R_{max} ratio and demonstrating peak broadening, poor signal quality, change in peak shape and change in chemical shift at different concentrations (Figure 5.7, 5.8).

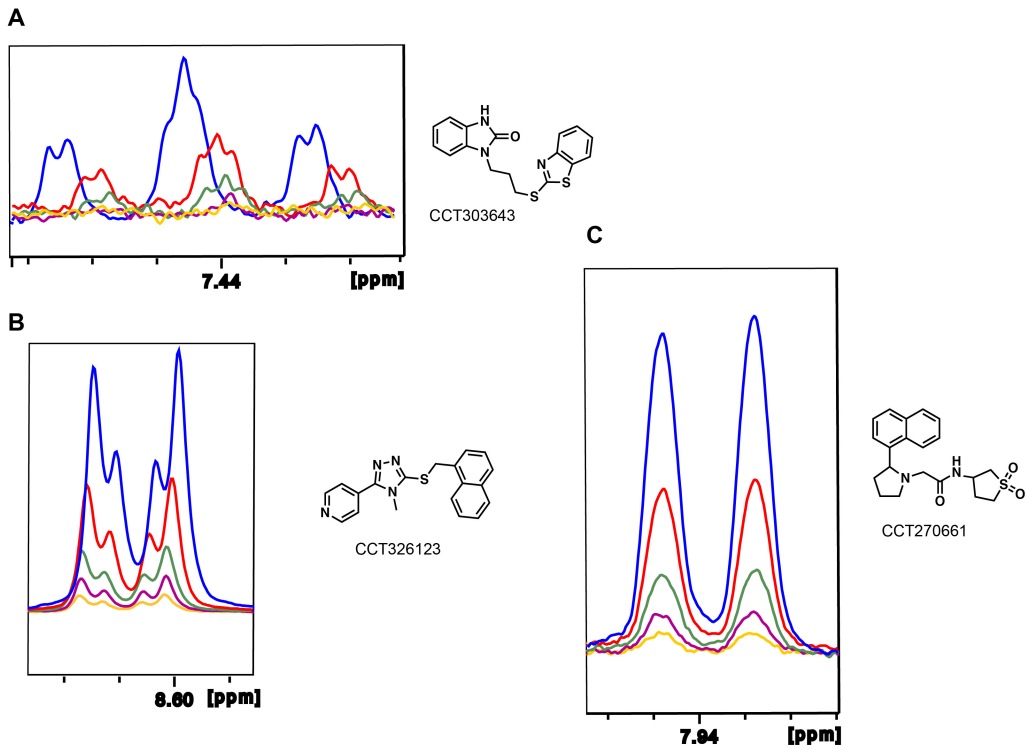


Figure 5.8. Evaluating aggregation of ligands using 1D ^{1}H NMR. Formation of soluble aggregates was evaluated for ligands obtained from SPR screening. Ligands from group 1, CCT303643, CCT192639 and CCT290039, as demonstrated in A) show severe peak change at different concentrations and have a poor NMR signal quality. Compounds CCT326123, CCT360813 in the second group demonstrate only a minor shift but show a good NMR spectrum. Compounds CCT349808 and CCT270661 are in third category. As displayed in C, these compounds show no change in peak shape, chemical shift or quality of the spectrum at different concentrations.

Compounds CCT326123, CCT360813 and CCT332162 are in the second category of compounds. These compounds demonstrate a good SPR R_{max} ratio and no change in the shape or number of the NMR peaks but do demonstrate slight change in chemical shift at different concentrations suggesting that they might form soluble aggregates that are not separable by centrifugation. Finally, CCT349808 and CCT270661 fall in the final category. These compounds have a K_D below 1 mM, have a good SPR R_{max} ratio and do not demonstrate any aggregation within their solubility limits under experimental conditions (Figure 5.7 and 5.8). Based on results obtained from the aggregation assay using 1D ^{1}H NMR, compounds from group 2 and 3 were taken for interaction studies using protein observed NMR.

5.4.1 CCT326123 (and family of compounds) (Group II)

CCT326123 and analogues were all synthesized by Dr. Charlotte Sutherell. Further evaluation of CCT326123 binding to Rab27a-gppNHp was carried out using NMR. Affinity was measured by measuring changes induced in M93 ^{13}C chemical shift in ^{13}C HSQC. Figure 5.9 summarizes information obtained for initial validation of CCT326123. Whilst, the ligand demonstrated binding through both SPR and NMR spectroscopy, there were major concerns whether the change in signal upon ligand titration in both assays stemed completely from " single soluble molecules" in a titratable and stoichiometric procedure. Figure 5.8 depicts some evidence towards formation of soluble aggregates. These aggregates were not separable from single soluble molecules by centrifugation (and in our case even use of detergents). Apart from this, we used the WaterLOGSY experiment to monitor change in the behavior of CCT326126 at different concentrations (Figure 5.10). We monitored the binding of CCT326123 to Rab27a at a ligand to protein ratio of 100:1 using 1 mM, 0.5 mM and 0.25 mM ligand concentrations. While at 1 mM ligand concentration, the addition of protein completely reverses the phase of ligand due to binding; such an effect was not seen at 0.5 mM and 0.25 mM concentrations. This is a very clear indication that at least " the ligand displays different behavior at different concentrations".

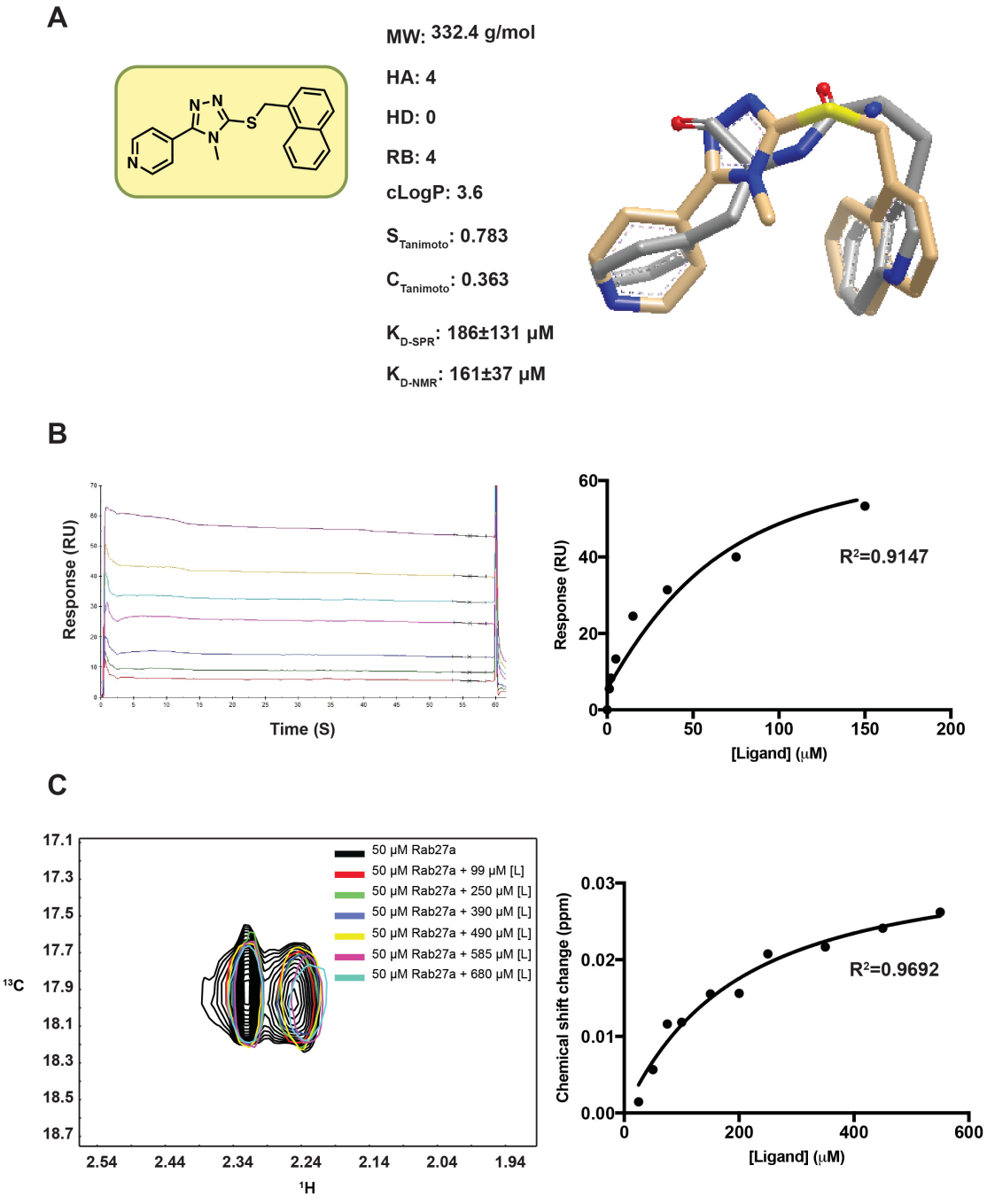


Figure 5.9. Analyzing interaction of CCT326123 with Rab27a. A summary of the data gathered for CCT326123 at different stages is given. A) Molecular descriptors, K_D determined by NMR and SPR and the shape overlay of CCT326123 (pink) with the query molecule (gray) is shown. B) Binding of CCT326123 to Rab27a was measured by SPR. Experiment was carried out three times for K_D measurements. C) Binding of CCT326123 to Rab27a was measured by monitoring changes in M93 ^{13}C chemical shift in ^{13}C HSQC.

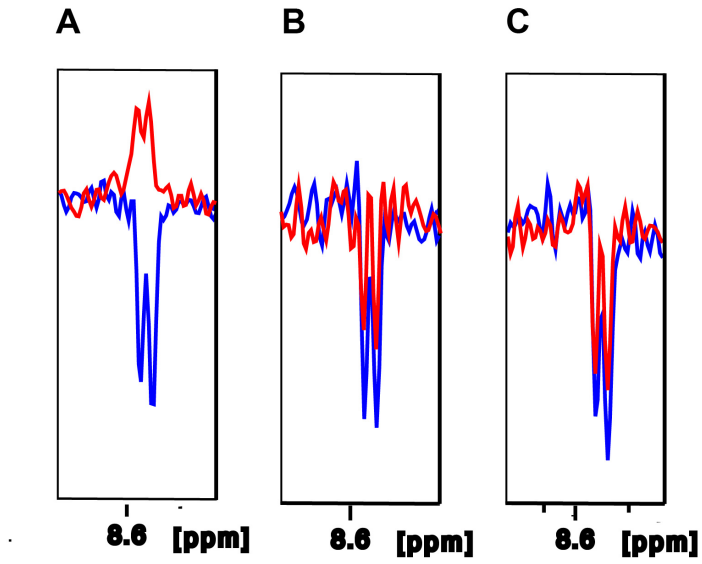


Figure 5.10. Analysis of CC326123 interaction with Rab27a at different concentrations. CCT326123 interaction with Rab27a was analyzed using WaterLOGSY experiment, in the absence (blue) and presence of 10 µM Rab27a. As demonstrated, while keeping the [ligand]/[protein] constant the ligand showed a more significant binding at 1 mM.

In addition to CCT326123, 5 other analogues of naphthalen-1-ylmethyl sulfanyl were in the initial selected hit pool (Figure 5.11 and appendix F). Preliminary data suggest the importance of the aromatic ring with a specially positioned hydrogen bond acceptor (pyridine ring) and a methyl on the triazole ring for the binding of naphthalen-1-ylmethyl sulfanyl derivatives. It is imperative to emphasize the fact that, it is very hard to make any conclusions based on this data, as it needs more replicates and evaluation by means of NMR and other methods. Moreover, since these ligands could potentially form aggregates, it could be very challenging to interpret SAR based on observed K_D from either NMR or SPR for this family of compounds.

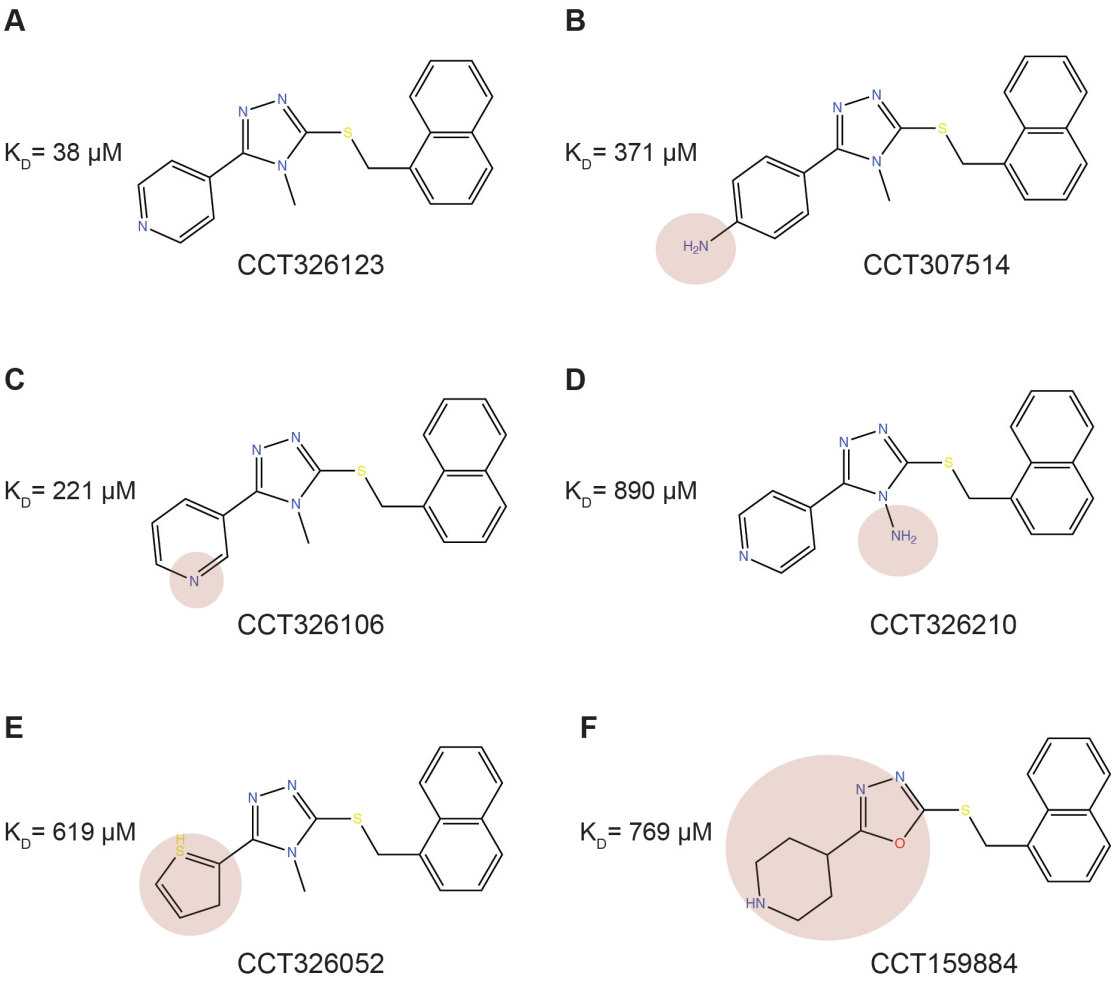


Figure 5.11. Naphthalen-1-ylmethyl sulfanyl derivatives obtained by SPR screening. Several compounds from naphthalen-1-ylmethyl sulfanyl derivatives demonstrated binding to Rab27a in the SPR screening. They are all analogues of CCT326123. None of these compounds displayed better binding than CCT326123. The data suggest that a methyl on the triazole ring and a spatially positioned hydrogen bond acceptor are very important for retaining activity.

5.4.2 CCT332162 (Group II)

Figure 5.12 demonstrates the SPR and NMR affinity measurements for CCT332162. While this compound demonstrated binding by SPR, no binding was observed by NMR. The compound showed low solubility under NMR experimental conditions. Precipitation was seen at concentrations as low as 50 μM for the compound. Therefore, it is possible binding seen in SPR is a result of precipitation rather binding of soluble single molecules to Rab27a. This could also be due to the effect of soluble aggregates.

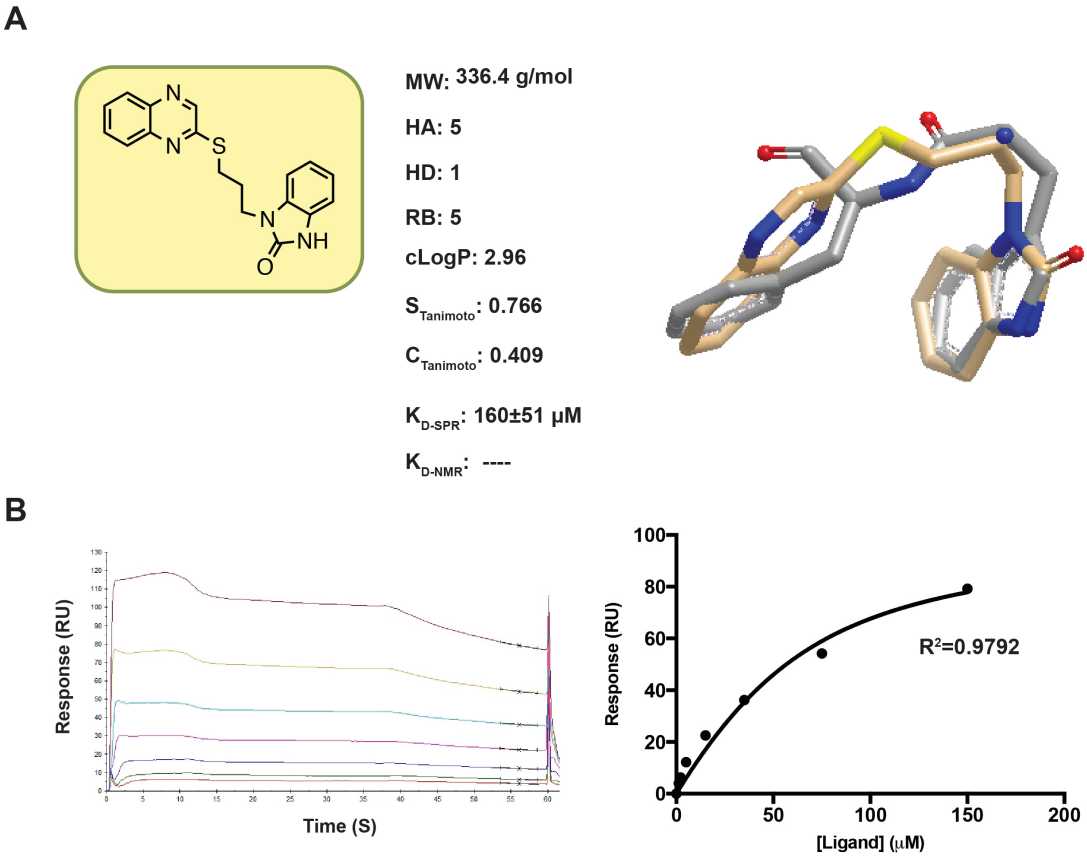


Figure 5.12. Analyzing interaction of CCT332162 with Rab27a. A summary of the data gathered for CCT332162 at different stages is given. A) Molecular descriptors, K_D determined by NMR and SPR and the shape overlay of CCT332162 (pink) with the query molecule (gray) is shown. B) Binding of CCT332162 to Rab27a was measured by SPR. Experiment was carried out three times for K_D measurements. C) Binding of CCT332162 to Rab27a was measured by monitoring changes in M93 ^{13}C chemical shift in ^{13}C HSQC. As demonstrated, no binding is observed even at 300 µM for CCT332162.

5.4.3 CCT360813 (Group II)

Figure 5.13 demonstrates the SPR and NMR affinity measurements for CCT360813. This compound demonstrated binding by SPR, but no binding was observed using NMR. While, 1D ^{1}H NMR demonstrated slight aggregation, the compound was soluble under NMR experimental conditions. Therefore, the signals recorded by SPR could be an artefact originated from aggregates. This is surprising as the 1D ^{1}H NMR spectrum gave a good signal to noise ratio. It could also infer that the ligand is a much weaker binder than what SPR is suggesting. Another explanation is that the ligand might be binding somewhere outside the SF4 pocket. Therefore, more analysis using ^{15}N HSQC is needed.

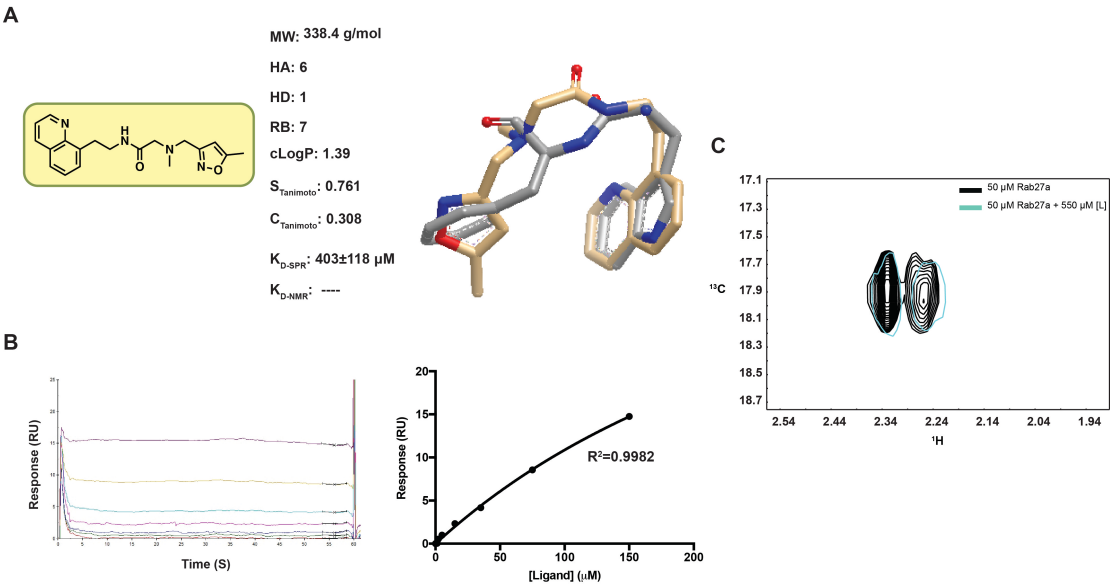


Figure 5.13. Analyzing interaction of CCT360813 with Rab27a. A summary of the data gathered for CCT360813 at different stages is given. A) Molecular descriptors, K_D determined by NMR and SPR and the shape overlay of CCT360813 (pink) with the query molecule (gray) is shown. B) Binding of CCT360813 to Rab27a was measured by SPR. Experiment was carried out three times for K_D measurements. C) Binding of CCT360813 to Rab27a was measured by monitoring changes in M93 ^{13}C chemical shift in ^{13}C HSQC. As demonstrated, no binding is observed even at 550 µM for CCT360813.

5.4.4 CCT349808 (Group III)

Figure 5.14 summarizes SPR and NMR affinity measurements for CCT360813. Similar to CCT332162 and CCT360813, This compound showed no binding by NMR. This comes as a surprise, because the compound did not show any aggregation or precipitation under experimental conditions. Since, this compound has three stereogenic centers, it is possible that only certain stereoisomers are active. Therefore, repeat of experiments with pure stereoisomers might be beneficial.

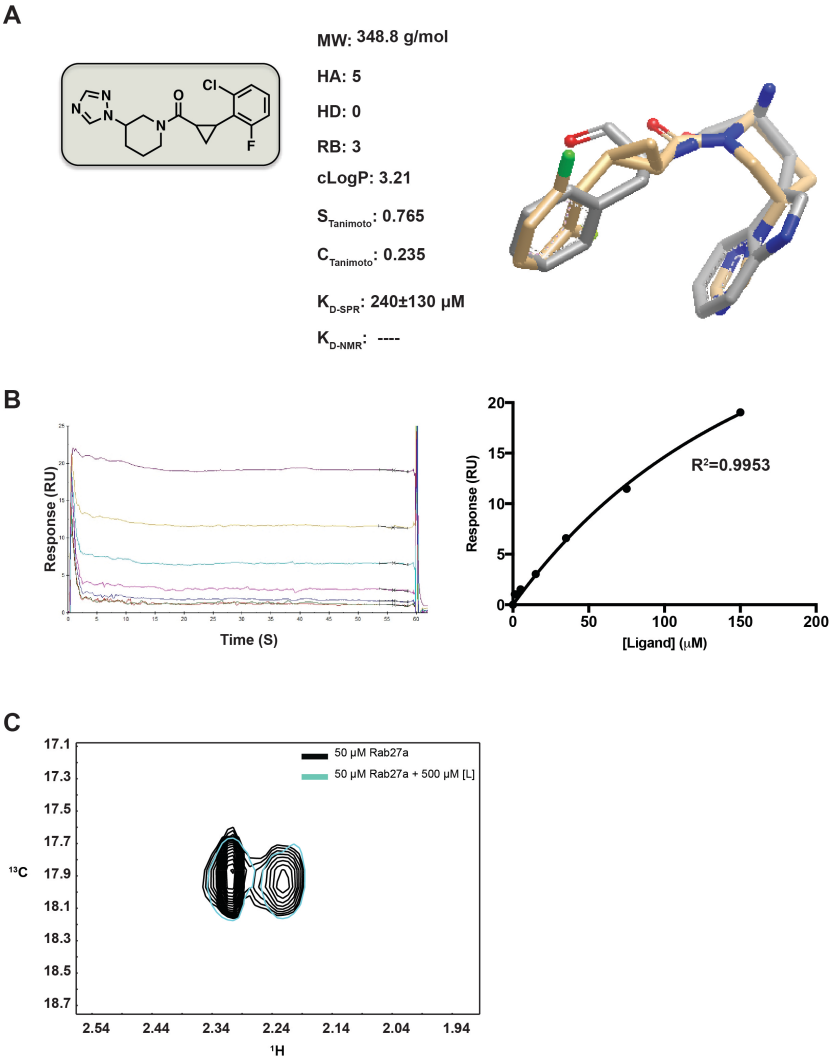


Figure 5.14. Analyzing interaction of CCT349808 with Rab27a. A summary of the data gathered for CCT349808 at different stages is given. A) Molecular descriptors, K_D determined by NMR and SPR and the shape overlay of CCT349808 (pink) with the query molecule (gray) is shown. B) Binding of CCT349808 to Rab27a was measured by SPR. Experiment was carried out three times for K_D measurements. C) Binding of CCT349808 to Rab27a was measured by monitoring changes in M93 ^{13}C chemical shift in ^{13}C HSQC. As demonstrated, no binding is observed even at 500 µM for CCT349808.

5.4.5 CCT270661 (Group III)

Figure 5.15 summarizes SPR and NMR affinity measurements for CCT270661. Interestingly, although this compound similar to CCT326123 is also a naphthalen-1-yl derivative, it does not show any aggregation (Figure 5.8). Also, this compound demonstrates binding with both SPR and NMR spectroscopy. Since, this compound also has two stereogenic centers, it is important to evaluate the potential difference in binding between different stereoisomers. Several analogues of this compound are to be purchased or synthesized for follow up studies.

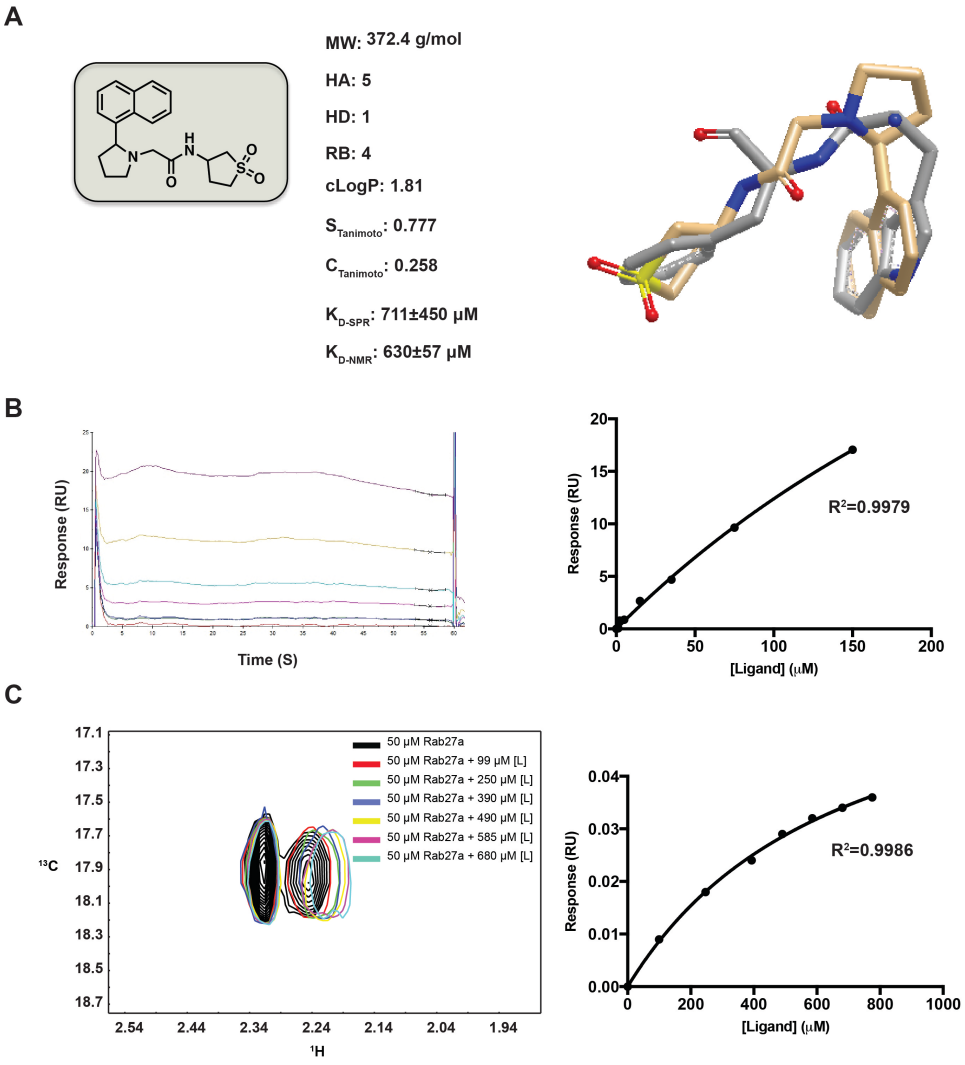


Figure 5.15. Analyzing interaction of CCT270661 with Rab27a. A summary of the data gathered for CCT270661 at different stages is given. A) Molecular descriptors, K_D determined by NMR and SPR and the shape overlay of CCT270661 (pink) with the query molecule (gray) is shown. B) Binding of CCT270661 to Rab27a was measured by SPR. Experiment was carried out three times for K_D measurements. C) Binding of CCT270661 to Rab27a was measured by monitoring changes in M93 ^{13}C chemical shift in ^{13}C HSQC.

5.5 Crystallography

Rab27a-ΔSlp2a fusion crystals were used to soak all five compounds analyzed by NMR. Unfortunately, no electron density was recognizable for any of the ligands. Since, CCT270661 is a chiral molecule, it might be important to separate chiral molecules from each other and re soak the crystals with each chiral molecule separately. Also the low solubility of these compounds could play a limiting factor in reaching saturation levels sufficient to observe electron density for any of the true binders.

5.6 Discussion

The featureless character of Rab27a-effector interaction interface has made targeting it very challenging. In this work, we used a ligand-based approach to design a library of compounds targeting the SF4 pocket. Using the conformation of the "WF" motif in $Slp2_a$ bound to Rab27a, a shape-based similarity method was used to enrich the library with compounds complementary to the SF4 pocket. The compounds in this focused library were screened against Rab27a using SPR. SPR provides several advantages as a screening tool. The need for minute amount of protein, high-throughput capability, high sensitivity with a broad range of affinity detection and the ability to use it for K_D measurements have made SPR a popular method in industry for screening small library of compounds. Although SPR has several benefits, similar to any method, it is not without complications. The biggest disadvantage of using SPR is the need to attach the protein to the surface. This procedure can lead to inactivation of the protein. Also, SPR is prone to artefacts. These artefacts could originate from variations in the machine e.g. temperature fluctuations, the protein loaded on the chip e.g. mass transfer effect or the analyte e.g. steric hindrance, non-specific binding or ligand aggregation. In our experience, non-specific binding of the analyte (ligands) to the protein due to aggregation/precipitation at the concentrations used for the assay was a major obstacle in analyzing data. This effect was probably exacerbated by the nature of compounds in our library, as the majority of compounds had aromatic rings and $2 < cLogP < 4$. In order to solve this problem two approach could be taken. One method is to design new libraries. Our library was not designed solely, based on the shape adopted by the "WF" dipeptide. We also put limitations on the library, by defining the presence of two aromatic rings in the query molecule.

One approach could be filtering out compounds that have more than one aromatic system. This way, the cLogP distribution of the library will shift towards more water-soluble molecules. The second option is to change the screening method. While most biophysical screening methods use high ligand concentrations, paramagnetic NMR and fluorine NMR provide the opportunity to screen against target protein at low protein and ligand concentrations. Jahnke et al. reported use of spin labeled proteins for screening weak binders to target proteins at ligand concentrations below 50 μM and protein concentrations as low as 1 μM [275]. This group later expanded this idea to use for second binding-site screening [276]. Running experiments at these concentrations would help avoid aggregation effects observed for lipophilic compounds at high concentrations. Apart from paramagnetic relaxation enhancement, fluorine NMR provides another method to screen or validate ligand binding [264]. Its selectivity and low background render the option of screening as many as 20 compounds in each NMR tube. Moreover, the large chemical shift range of fluorine, increase its sensitivity and resolution to small changes in the chemical environment. Therefore, ligand concentrations as low as 20 μM are sufficient for screening campaigns. Although, our library was not fluorinated, our collaborating chemists have synthesized several fluorinated analogues of the hits. If these analogues do show binding, we will take advantage of fluorine NMR to study ligand binding at low concentrations. This would help us overcome challenges of using compounds at high concentrations.

The initial results obtained from screening suggested we had multiple hits with K_D bellow 100 μM. More detailed analysis of data, suggested several hits were non-binders. For example, CCT303643 was initially identified as a binder through SPR with a K_D of approximately 6 μM. But solubility and 1D proton NMR analysis revealed this compound suffered from severe precipitation and aggregation at concentrations used for SPR. CCT360813 and CT332162 also, did not show any binding by NMR. From all the hits obtained through the initial screening by SPR, only CCT326123 and CCT270661 demonstrated binding by both SPR and NMR spectroscopy. CCT326123 demonstrated a complex behavior in interacting with Rab27a. WaterLOGSY and 1D ^{1}H NMR experiments both suggested the presence of aggregate particles in solution. These aggregates are a main cause of false positives in drug discovery. The importance of soluble aggregates has been reiterated in a recent letter by the editorial board of American chemical society (ACS) journals. The board has suggested all compounds reported for activity against a target for ACS journals should be evaluated for colloidal aggregations.

Similar to CCT326123, CCT270661 is also a naphthalen-1-yl derivative. Interestingly, formation of aggregates could not be detected for this compound by NMR spectroscopy. Also, CCT268947, which is an analogue of CCT270661 but with a more flexible linker, did not have a measurable K_D. While these initial data are promising, more analogues of CCT270661 need to be synthesized and tested. Moreover, since the main hypothesis for the shape-based approach has been the conformation of the bound ligand, it is important to exploit different methods to obtain information on the ligand binding mode. A combination of mutagenesis, NMR experiments such as saturation transfer difference, transfer-NOE and cross-linking by photo-activated head groups can be used to define ligand binding mode.

In conclusion, a shape-based similarity approach was used in search for ligands binding to the SF4 pocket. One hit (CCT270661) has been identified. This molecule demonstrates binding to Rab27a by both SPR and NMR spectroscopy. This compound could be used to study how chemical molecules bind to the SF4 pocket. This information can be used to design a new query molecule to redirect the chemical space of the library towards molecules complementary to the SF4 pocket.

CHAPTER SIX: DISCUSSION

With the emergence of Ras as one of the key regulators in several cancers [12], designing selective inhibitors of Ras with high affinity has been thought of as the “holy grail” of cancer [29]. While medicinal chemists have had much success with designing antagonists for ion-channels, G-protein coupled receptors (GPCRs), histamine receptors or kinases, designing Ras inhibitors has been very challenging. Apart from Ras, several other members of the Ras superfamily, especially members of the Rab family have been recognized as potential targets in various disease [97]. Experience gained from each protein can have an important implication on how we guide our research to study these proteins. Analyzing literature available for compounds designed as Ras and Rab inhibitors reveals some important notes suggesting certain facts need to be considered with regards to the nature of these proteins for experimental design. Understanding the causes of previous failed attempts can accommodate the planning and guiding of successful future projects.

Ras superfamily members have been shown to have picomolar affinity for GTP/GDP with a very slow off rate [277]. As a result, nucleotide exchange only happens in the presence of GEF/GAP or by other enzymatic procedures such as use of alkaline phosphatase or snake venom phosphodiesterase [174]. This is different to ATPases that have micromolar affinity for ADP, making design of competitive inhibitors for the ATP binding pocket more probable [278]. In a research conducted by Agola et al., binding of Boron-dipyrromethane (BODIPY)-GTP to nucleotide-free Rab7 loaded on GST was measured against an HTS library [279]. One hit was obtained through the HTS screen, which was a pan-interfering compound, binding to a range of Ras superfamily members. Although the researchers themselves demonstrate the inability of the suggested compound to remove the nucleotide from the binding pocket or increase its off rate, they suggest it acts as a direct cavity binder in the absence of any nucleotide. This assay design does not consider the nature of these proteins always cycling between GDP and GTP states, nor does it take into account the high concentration of GTP/GDP inside the cells. Moreover, apart from an indirect competition assay, no direct measurement is given for the binding of the ligand to the nucleotide-free protein. The researchers suggest aggregation of the protein at concentrations above 1 μM as a reason why no measurements can be made for binding in solution. This notion brings up the

question, whether loading the protein on the beads is sufficient to avoid aggregation in the absence of a nucleotide.

Another key characteristic of Ras superfamily members is the dynamic nature of the switch and inter-switch regions in these proteins. These movements control the function of the protein and directly effect our interpretation of data we collect through structural methods [34]. While high-resolution crystal structures can provide detailed information about side chain orientations, this information might not always mirror a complete image of a pocket or a certain residue in solution. For example, overlapping the structure of several Rab proteins and HRas in a GTP bound form available from the PDB it is evident that the loop comprising the switch I region, always crystallizes in a certain orientation (Figure 6.1). One must consider this as a crystal artifact, rather than a state truly representing the conformation of this loop in solution. This conformation could be present, only as a small potion of a large ensemble of conformations separated by low energy barriers. Therefore, in the author's opinion, one cannot make too many conclusions about the shape of pockets in the crystal structures, or use them for docking-based virtual HTS campaigns.

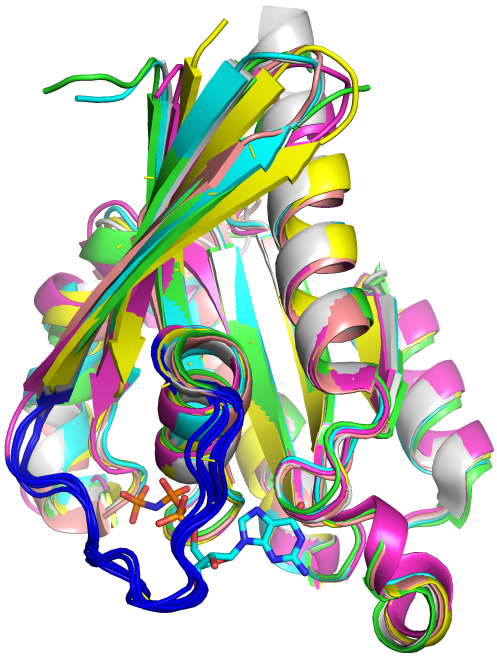


Figure 6.1. Structural overlay of HRas and multiple Rab members bound to GTP. An overlay of Hras (4RSG, cyan), Rab3a (3RAB, purple), Rab11(1YZK, pink), Rab5 (1Z07, yellow), Rab6 (1YZQ, gray), Rab27a (3BC1, green). As illustrated, the switch I loop highlighted in blue adopts a conserved rigid conformation. NMR studies on Ras proteins have shown this loop is flexible with an ensemble of conformations interconverting between each other.

In a research carried out by Hocker et al. and published in the peer reviewed journal of PNAS, several derivatives of andrographolides were docked computationally onto pockets that were solely observed in crystal structures of state I GTP-locked KRas [37]. The initial basis for this research was the observed inhibitory effect of these compounds on the MAPKinase pathway. Interestingly, they use the same cell-based assays as the only piece of experimental evidence to show that these ligands target such pockets and inhibit KRas. This work lacks the chain of experiments needed to correlate ligand binding to phenotypic assays. Moreover, the use of docking models (and suggested K_D based on docking energetics) for such claims is an over speculation that cannot be supported by the provided experimental evidence.

The third feature of Ras superfamily members important from a drug discovery perspective is the large surface area of interaction between them and their relevant effectors. For example, the area of interaction between Rab27a and $Slp2_a$ is 1970 Å with the residues important for the energetics of the interaction far apart from each other [137]. This has several implications. Primarily no chemical molecule can cover such large area; therefore designing a molecule that can effectively displace such as an effector with affinities within the low nanomolar range is a very challenging prospect. Second, designing α-helical peptides as inhibitors for such targets is not trivial. Methods such as hydrogen bond surrogates (HBS) or stapled peptides have been effectively used to design high affinity binders for proteins such as p53 [280, 281]. Unlike p53, the hotspots in Rab27a are distanced apart, meaning that peptidic inhibitors of 15-25 amino acids long, can only cover one hotspot region. This results in the peptides having micro molar affinities at best. For example, stapled peptides designed based on the sequence of Rab8 effectors has a K_D of 10 μM [282]. This is because Rab8 effector uses multiple motifs that come together in space to interact with Rab8. Therefore, stapled peptides designed based on the sequence of one of the effector motifs cannot break the micromolar affinity barrier.

Another element that needs consideration when working with Ras members is the nucleotide-binding state of your protein during the assay. While all published work in this research area mention if their protein is GDP bound or GTP bound, it is vital that the protein keeps a homogenous nucleotide state during assay run. This can be important especially when your protein is fixed to a solid state as in the case of SPR or beads used for flow cytometry. While methods such as NMR can be used to validate protein conformational homogeneity, other methods such as SPR or fluorescent bead-loaded proteins do not provide this information. This is a reason for

NMR being the method of choice for screening chemical libraries against these Rab/Ras proteins.

Since Ras superfamily members have few or any well-defined pockets (apart from the nucleotide-binding pocket which is occupied by GTP or GDP) HTS campaigns have not been successful. Fragment-based approaches have been key towards obtaining lead compounds. As a consequence of fragments low binding affinity, these ligands have to be used at concentrations close to millimolar in experimental conditions. This increases the chance of aggregation and obtaining false positives. Aggregation is aggravated by the large area of hydrophobic patches in proteins such as Rab27a. Analyzing Rab27a cycle, from expression that they bind to REP, to the point that they are transferred to their destination membrane, it is clear that it is always bound to another protein through large surface of hydrophobic patches [23]. This is one of the main reasons that studying these proteins on their own, is not trivial (in the absence of an effector or any other binding protein). This effect although not solely responsible, has been reflected in the high number of crystal structures and low number of NMR structures of Ras superfamily members deposited to the PDB. Conventional NMR experiments used for calculating 3D structures need sample concentrations of at least 300 μM or realistically 500 μM and beyond. These are minimum conditions in order to get reasonable signal in insensitive 3D experiments such as the HNCACB which is the least sensitive experiment among the four major pulse sequences of HNCACB, CBCACONH, HNCO and HNCACO used for peak assignment in ^{15}N HSQC spectrum.

At the beginning of this work, much of our knowledge on Rab27a structural characteristics was based on information gained from the Rab27a-Slp2$_a$ complex published structure (PDB: 3BC1) [137]. This structure provides information on one of the surfaces Rab27a uses to bind to target proteins. Apart from this, no other structural or dynamic data is available on Rab27a. For example, unlike KRas no experimental data similar to the ^{31}P spectrums collected for Ras is available to use for evaluating the population distribution of state I/II GTP in Rab27a (and other Rabs) [36]. Moreover, no crystal structure on the state I GTP or GDP bound form of Rab27a is available. But most importantly, no structural information such as the assignment for the 2D ^{15}N, ^{1}H NMR spectrum is available to study the dynamics of Rab27a. This piece of information could be used very effectively to study dynamics of Rab27a by monitoring backbone motions at different time-scales. Analyzing the change in the dynamic behavior of residues on the α_3-β_5 loop can help us better understand the shape and dynamics of the SF4 pocket. This information can be

used in designing chemical libraries that can exploit the chemical space in a better way. Moreover, it could be used to monitor ligand binding to the protein by obtaining information on ligand binding sites.

Noticeably both X-ray crystallography and NMR play an integral role in modifying initial weak fragments obtained through screening against Rab27a to potent lead compounds [283]. NMR especially is key since hits we obtained through NMR, DSF and SPR (or other means of fragment screening in the future) have K_D above the one millimolar. The low solubility and very weak affinity of these ligands limited the chances of obtaining ligand-bound crystal structures. It is usually assumed that the ligand concentration surrounding the crystals should be around 10× K_D in order to saturate the crystals with the ligands. This was not very feasible, as the solubility of the ligands obtained did not exceed 5 mM at best experimental conditions, while a range of 20-50 mM is needed for such weak binders. Therefore, NMR is usually the method of choice in the initial stages of hit to lead development. Information obtained through NMR can be used to increase the affinity of the ligands to the low micromole ranger, which are then more suitable for crystallographic studies. This approach has been a successfully used by other groups working with Ras proteins (Beatson institute-internal communications). We noticed from the beginning of the project that obtaining the amino acid specific assignment of the 2D ^{15}N, ^{1}H NMR spectrum for Rab27a is vital for the progress of the project. A lot of effort in the project was directed towards this aim but was not successful. This is partly due to two characteristics of Rab27a. Firstly, Rab27a is not stable at concentrations suitable for collecting 3D NMR experiments. A 500 µM sample of Rab27a in an NMR tube at 25^0C would precipitate after an hour leading to poor shimming and a sharp decrease in protein concentration. This led to a low signal to noise ratio and poor signal quality. This is potentially due to the large exposed hydrophobic surface of Rab27a in the absence of any binding partner. Moreover, although through extensive mutagenesis of several hydrophobic surface residues we were able to stabilize the protein to a point that the protein would be stable at experimental conditions for 12 hours, still the acquired spectrums would not yield good quality signals. This is due to the dynamic nature of the switch and inter-switch region of Rab27a (similar to other Ras superfamily members). The result of these motions is peak broadening and signal loss in the 3D spectrum. In fact, for our Rab27a construct with 190 residues only 130-140 peaks corresponding to backbone amide protons are present in the 2D ^{15}N, ^{1}H NMR spectrum, with no peaks present for 40-50 residues. This is a result of extensive peak broadening due to

intermediate exchange. This highlights the need for a different approach for the assignment of amide protons in the HSQC spectrum. The conventional method for assigning amide protons is reliant on 3D NMR experiments that transfer magnetization through bonds. In the case of a well-folded 10-15 kD protein, these experiments work fine and an assignment doesn't take more than a day or two, but for proteins with stability issues or dynamic behavior, it can be very challenging. Apart from methods reliant on magnetization transfer, methods reliant on selective labeling have been proposed for the assignment of the 2D ^{15}N, ^{1}H NMR spectrum. In work published by Parker et al. 4 samples with a variation on amino acid spin labels were used to assign the HSQC spectrum [284]. Since this method is only reliant on 2D spectrums, it can be very useful for unstable proteins such as Rab27a and other members of the Ras superfamily. The drawbacks of such a method are assignment degeneracy and expense. For Rab27a, this method should be able to assign 103 residues without ambiguities with other peaks assigned to a specific amino acid but without the exact residue number known. Some peaks will have two and some three or four probabilities. Also since this method is still reliant on 2D HSQC spectrums which is a magnetization transfer based experiment, still no information can be acquired for residues that have no representative peak in the 2D HSQC spectrum.

Paramagnetic NMR by use of paramagnetic ions such as lanthanides or cysteine reactive spin label molecules such as MTSL (*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methylmethanesulfonothioate), can provide another approach to gain distance and angular information on ligand binding to Rab27a. Also, it can be used to back calculate the assignment of the 2D ^{15}N, ^{1}H NMR spectrum for the available peaks in the spectrum [285]. Spin labels have been used to collect restraints for calculating membrane protein structures, intrinsically disordered proteins, protein-DNA complexes or even weak, transient domain-domain interactions. This is achieved by coupling spin labels with thiol-reactive groups to engineered cysteines on the surface of the protein. Wolfgang Jahnke et al. have discussed application of PRE in ligand discovery in detail [275, 286]. Due to the enhanced relaxation effects of spin labels, they can be used for primary/secondary screening or even K_D calculations. Since native Rab27a has two cysteines close to the SF4 pocket, this could prove to be a very effective method to screen for secondary binding sites close to the SF4 pocket. We already have carried out a screening using covalent inhibitors against Rab27a. Several hits have been identified and their structural characterization is on the way. Spin labels such as

MTSL have an isotropic magnetic susceptibility tensor and can only provide distance-dependent relaxation through paramagnetic relaxation enhancement (PRE). Therefore, multiple cysteine mutant constructs conjugated to a spin label with each mutant designed in a spatially separated position need to be used to provide a rough but acceptable estimate of ligand binding sites. This is an achievable target in this project, since the structure of Rab27a is known and can be used to design the position of the spin labels. The other way of using paramagnetic compounds is by using lanthanides that (except for Gadolinium) have a magnetic susceptibility anisotropy (MSA) [285]. Lanthanide MSA induces pseudo-contact shifts (PCS) in other nuclear spins. PCS is dependent on the distance and orientation between the metal ion and other nuclei. This information can be used as structural restraints for protein-ligand structure calculations [287]. Several approaches have been suggested for tagging lanthanides to proteins. One could conjugate lanthanide chelators to two closely related cysteines or use lanthanide binding peptides engineered in the protein sequence [285]. Figure 6.2 illustrates two possible approaches for tagging lanthanides to Rab27a. One strategy is to put a tag instead of the highly flexible loop between the β_2-β_3 strands as demonstrated for other proteins [288]. Another method could be to attach a highly rigid, high affinity chelator of lanthanides to two cysteines engineered on helix 3,4 or 5 [289]. An advantage of such methods is the possibility to make these modifications on the $Slp2_a$-Rab27a fusion construct that can crystallize. This way, the coordinates of the lanthanides can be measured very accurately, simplifying magnetic susceptibility tensor calculations.

X-ray crystallography was our main focus towards obtaining a structural model for studying Rab27a-ligand binding. Although a lot of effort was put towards obtaining suitable constructs, there is room for improvements in designing better models. The fusion constructs have limitations as mentioned before, but can provide a very good platform for studying the binding mode of ligands that target the SF4 pocket. Their use can be maximized if the ligands are covalently bound to nearby cysteine residues. This can help overcome limitations on ligands with low solubility and weak affinity during crystal soaking. Mutant Rab27a constructs were another approach that led to MJ1-1Rab27a. This construct crystallizes consistently in a short space of time and diffracts isotropically to 3.15 Å. This is a significant improvement on Rab27a, which would not crystallize on its own at all, however the resolution must be improved in order for it to be useful for ligand binding studies. This could potentially be improved by searching for additives, dehydration or micro-seeding of crystals. Another method that could be attempted is mutating the

residues of another Rab protein that crystallizes on its own to residues present in Rab27a in the SF4 pocket. This approach has been successful for other proteins such as Rad51 [290]. Finally, it is worth mentioning that the effort in our work has been towards the GTP-bound conformation of Rab27a while the GDP-bound conformation is another state that could be targeted for drug discovery.

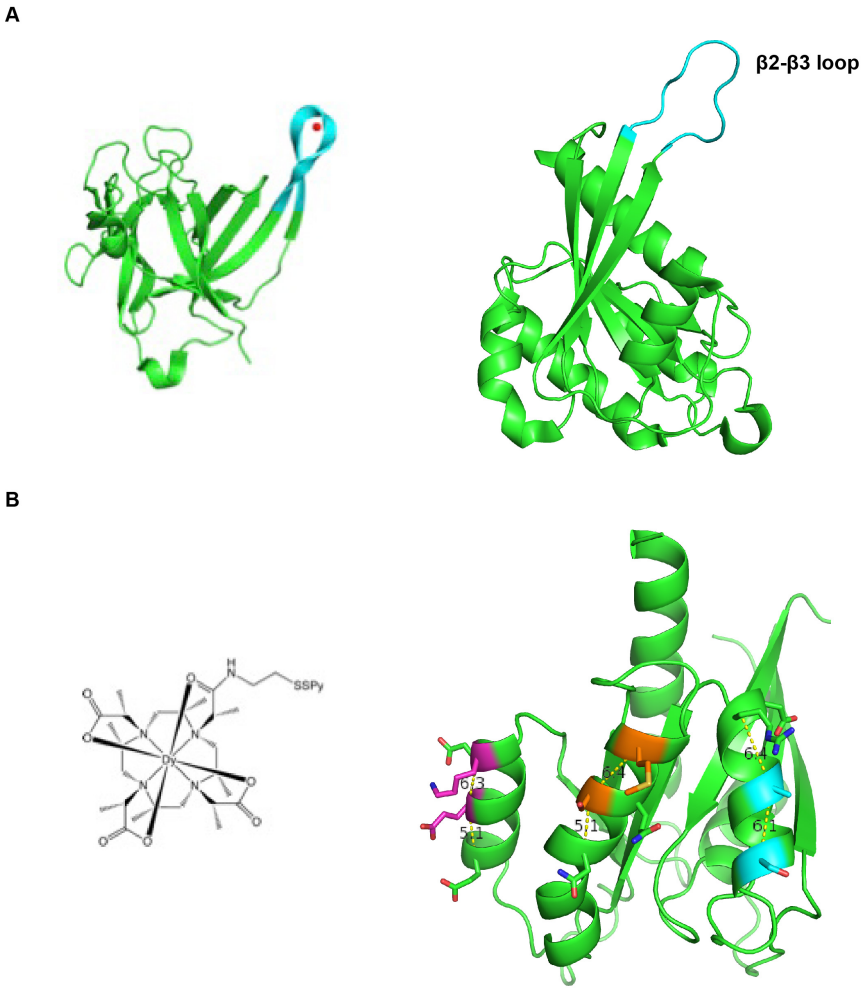


Figure 6.2. Tagging lanthanide ions on Rab27a. A) Lanthanide bindng tags in-between beta strands, similar the one used in Interlukin 1B (PDB:3LTQ, lanthanide ion in red) can be used to replace the loop between β_2-β_3 in Rab27a. Interlukin 1B is shown on the left and Rab27a on the right side of the figure. B) Highly rigid lanthanide chelator (on the left), can be tagged on two cysteines separated by three residues on the exposed side of a helix. Probable tagging sites are shown on the Rab27a structure on the right side of the figure.

Currently two papers published by Joung, J.Y. et al and Johnson, J.L. et al describe their efforts to target Rab27a-effector interactions [274, 291]. Joung et al. describe using a pharmacophore-based approach to target Rab27a/melanophilin interaction [274]. Ligands obtained from virtual screening were taken directly to phenotypic assay. While no experimental evidence was provided for Rab27a-ligand interactions, docking was used to prove these ligands block Rab27a-melanophilin interaction. The results from this work are very unconvincing, as they don't follow the chain of experiments that are necessary in order to take ligands from virtual screening to phenotypic assays. Ligands suggested in this paper were synthesized in our lab by Dr. Charlotte Sutherell and experimentally tested by means of NMR and SPR. The ligands did not show any binding in neither of the two assays. In the second work carried out by Johnson et al. several hits were obtained through time-resolved Förster resonance energy transfer (TR-FRET) experiment [291]. Hits obtained from the initial screening were taken for further cellular assays. No evidence was presented for the direct binding of the ligands to Rab27a. Moreover, putative ligands from this work were inherently suspicious of interfering with cellular assays used in this work. Ligands suggested in this paper were synthesized in our lab by Dr. Charlotte Sutherell and experimentally tested by means of NMR and SPR. The ligands did not show any binding in neither of the two assays.

In our attempt towards ligand screening we took two different approaches. One approach was using a focused chemical library targeting the SF4 pocket while the other approach was using a fragment library to screen against Rab27a for potential hotspots. The initial approach as explained in chapter 5 suffered from the absence of a chemical molecule as positive control. Therefore, we had to use the conformation adopted by the "WF" dipeptide of the Slp RBD in complex with Rab27a to design our library. While this was our starting point, our model and library design would improve a lot if we could have an iteration in library design and SPR screening to test different hypotheses with regards to ligand binding to the SF4 pocket. In this case, each time the shape similarity would be based on one of the major conformations adopted by the ligand through the simulations. Library design could benefit a lot from running MD simulations on CCT270661, a compound that was obtained from the SPR screen. This compound can provide invaluable information on the restraints that the SF4 pocket imposes on organic molecules. Apart from this, we could guide the library design towards compounds that mimic the desired shape but avoid having more than one aromatic ring in their structure. This could help to alleviate the compounds solubility and aggregation problems. Altogether, obtaining

structural information of the binding mode of CCT270661 to the SF4 pocket will be the next target in this project.

The second method we used to screen for binders was fragment screening by DSF. Apart from the shortcomings with DSF, the biggest challenge was the follow up of the validated hits. We discussed how NMR could play a major role in identifying the ligand binding sites. If NMR is not an option, SAR by catalogue could be a very effective way in exploring the chemical environment around ligand binding sites. This can help us build a model for each class of fragments and define pharmacophore features important for the binding of each ligand. Finally, since Rab27a does not have many well-defined pockets, it is worth considering screening complexes of Rab27a with its GEF or GAP protein, as more attractive pockets might form at the interface of interaction between Rab27a and its relevant GEF or GAP. EPI64 is Rab27a GAP, is approximately 45 kD and can be expressed in bacteria. Stabilizing Rab27a-EPI64 complex could be an interesting approach by decreasing the active form of Rab27a in the cell.

Since the start of the project, we have made major leaps towards obtaining crystals of Rab27a and ligands that bind to it. Recently, our group has finished screening a fragment library by NMR. Plans for a screening by crystallography at Diamond light source facility is also underway. Finally, a library of compounds with an acrylamide group, have been used to target pockets close to C123 and C188 in Rab27a, which are part of the SF4 pocket. We are currently working on acquiring crystal structures of ligands bound to Rab27a for hits obtained by these screening methods.

Altogether, designing inhibitors against Rab27a needs an approach different from conventional methods used for classic medicinal chemistry targets such as enzymes. The screening and structural methods both should meet the demands of the nature of the available pockets and the potential ligands obtained as initial hits. In this thesis I have tried to address some of these challenges and provide a platform for future studies.

REFERENCES

1. Wennerberg, K., K.L. Rossman, and C.J. Der, *The Ras superfamily at a glance.* J Cell Sci, 2005. **118**(Pt 5): p. 843-6.
2. Vetter, I.R. and A. Wittinghofer, *The guanine nucleotide-binding switch in three dimensions.* Science, 2001. **294**(5545): p. 1299-304.
3. Cox, A.D. and C.J. Der, *Ras family signaling: therapeutic targeting.* Cancer Biol Ther, 2002. **1**(6): p. 599-606.
4. Harvey, J.J., *An Unidentified Virus Which Causes the Rapid Production of Tumours in Mice.* Nature, 1964. **204**: p. 1104-5.
5. Kirsten, W.H. and L.A. Mayer, *Morphologic responses to a murine erythroblastosis virus.* J Natl Cancer Inst, 1967. **39**(2): p. 311-35.
6. Shih, C., et al., *Passage of phenotypes of chemically transformed cells via transfection of DNA and chromatin.* Proc Natl Acad Sci U S A, 1979. **76**(11): p. 5714-8.
7. Gibbs, J.B., et al., *Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules.* Proc Natl Acad Sci U S A, 1984. **81**(18): p. 5704-8.
8. Capon, D.J., et al., *Activation of Ki-ras2 gene in human colon and lung carcinomas by two different point mutations.* Nature, 1983. **304**(5926): p. 507-13.
9. Johnson, L., et al., *Somatic activation of the K-ras oncogene causes early onset lung cancer in mice.* Nature, 2001. **410**(6832): p. 1111-6.
10. Almoguera, C., et al., *Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes.* Cell, 1988. **53**(4): p. 549-54.
11. Forrester, K., et al., *Detection of high incidence of K-ras oncogenes during human colon tumorigenesis.* Nature, 1987. **327**(6120): p. 298-303.
12. Cox, A.D., et al., *Drugging the undruggable RAS: Mission possible?* Nat Rev Drug Discov, 2014. **13**(11): p. 828-51.
13. Johnson, L., et al., *K-ras is an essential gene in the mouse with partial functional overlap with N-ras.* Genes Dev, 1997. **11**(19): p. 2468-81.
14. Downward, J., *Targeting RAS signalling pathways in cancer therapy.* Nat Rev Cancer, 2003. **3**(1): p. 11-22.
15. Gysin, S., et al., *Therapeutic strategies for targeting ras proteins.* Genes Cancer, 2011. **2**(3): p. 359-72.
16. Etienne-Manneville, S. and A. Hall, *Rho GTPases in cell biology.* Nature, 2002. **420**(6916): p. 629-35.
17. Vega, F.M. and A.J. Ridley, *Rho GTPases in cancer cell biology.* FEBS Lett, 2008. **582**(14): p. 2093-101.
18. Scheffzek, K., et al., *Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form.* Nature, 1995. **374**(6520): p. 378-81.

19. Sasaki, T. and Y. Takai, *The Rho small G protein family-Rho GDI system as a temporal and spatial determinant for cytoskeletal control.* Biochem Biophys Res Commun, 1998. **245**(3): p. 641-5.

20. Weis, K., *Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle.* Cell, 2003. **112**(4): p. 441-51.

21. Kuersten, S., M. Ohno, and I.W. Mattaj, *Nucleocytoplasmic transport: Ran, beta and beyond.* Trends Cell Biol, 2001. **11**(12): p. 497-503.

22. Nie, Z., D.S. Hirsch, and P.A. Randazzo, *Arf and its many interactors.* Curr Opin Cell Biol, 2003. **15**(4): p. 396-404.

23. Hutagalung, A.H. and P.J. Novick, *Role of Rab GTPases in membrane traffic and cell physiology.* Physiol Rev, 2011. **91**(1): p. 119-49.

24. Bourne, H.R., D.A. Sanders, and F. McCormick, *The GTPase superfamily: conserved structure and molecular mechanism.* Nature, 1991. **349**(6305): p. 117-27.

25. Colicelli, J., *Human RAS superfamily proteins and related GTPases.* Sci STKE, 2004. **2004**(250): p. RE13.

26. Moores, S.L., et al., *Sequence dependence of protein isoprenylation.* J Biol Chem, 1991. **266**(22): p. 14603-10.

27. Whyte, D.B., et al., *K- and N-Ras are geranylgeranylated in cells treated with farnesyl protein transferase inhibitors.* J Biol Chem, 1997. **272**(22): p. 14459-64.

28. Cohen, S.J., et al., *Phase II and pharmacodynamic study of the farnesyltransferase inhibitor R115777 as initial therapy in patients with metastatic pancreatic adenocarcinoma.* J Clin Oncol, 2003. **21**(7): p. 1301-6.

29. Baines, A.T., D. Xu, and C.J. Der, *Inhibition of Ras for cancer treatment: the search continues.* Future Med Chem, 2011. **3**(14): p. 1787-808.

30. Ito, Y., et al., *Regional polysterism in the GTP-bound form of the human c-Ha-Ras protein.* Biochemistry, 1997. **36**(30): p. 9109-19.

31. Geyer, M., et al., *Conformational transitions in p21ras and in its complexes with the effector protein Raf-RBD and the GTPase activating protein GAP.* Biochemistry, 1996. **35**(32): p. 10308-20.

32. Araki, M., et al., *Solution structure of the state 1 conformer of GTP-bound H-Ras protein and distinct dynamic properties between the state 1 and state 2 conformers.* J Biol Chem, 2011. **286**(45): p. 39644-53.

33. Kobayashi, C. and S. Saito, *Relation between the conformational heterogeneity and reaction cycle of Ras: molecular simulation of Ras.* Biophys J, 2010. **99**(11): p. 3726-34.

34. Shima, F., et al., *Structural basis for conformational dynamics of GTP-bound Ras protein.* J Biol Chem, 2010. **285**(29): p. 22696-705.

35. Liao, J., et al., *Two conformational states of Ras GTPase exhibit differential GTP-binding kinetics.* Biochem Biophys Res Commun, 2008. **369**(2): p. 327-32.

36. Spoerner, M., A. Wittinghofer, and H.R. Kalbitzer, *Perturbation of the conformational equilibria in Ras by selective mutations as studied by 31P NMR spectroscopy.* FEBS Lett, 2004. **578**(3): p. 305-10.

37. Hocker, H.J., et al., *Andrographolide derivatives inhibit guanine nucleotide exchange and abrogate oncogenic Ras function.* Proc Natl Acad Sci U S A, 2013. **110**(25): p. 10201-6.

38. Carroll, K.S., et al., *Role of Rab9 GTPase in facilitating receptor recruitment by TIP47.* Science, 2001. **292**(5520): p. 1373-6.

39. Semerdjieva, S., et al., *Coordinated regulation of AP2 uncoating from clathrin-coated vesicles by rab5 and hRME-6.* J Cell Biol, 2008. **183**(3): p. 499-511.

40. Garcia, I.A., H.E. Martinez, and C. Alvarez, *Rab1b regulates COPI and COPII dynamics in mammalian cells.* Cell Logist, 2011. **1**(4): p. 159-163.

41. Stenmark, H., *Rab GTPases as coordinators of vesicle traffic.* Nat Rev Mol Cell Biol, 2009. **10**(8): p. 513-25.

42. Jordens, I., et al., *Rab proteins, connecting transport and vesicle fusion.* Traffic, 2005. **6**(12): p. 1070-7.

43. Echard, A., et al., *Interaction of a Golgi-associated kinesin-like protein with Rab6.* Science, 1998. **279**(5350): p. 580-5.

44. Ueno, H., et al., *KIF16B/Rab14 molecular motor complex is critical for early embryonic development by transporting FGF receptor.* Dev Cell, 2011. **20**(1): p. 60-71.

45. Strom, M., et al., *A family of Rab27-binding proteins. Melanophilin links Rab27a and myosin Va function in melanosome transport.* J Biol Chem, 2002. **277**(28): p. 25423-30.

46. Guo, W., et al., *The exocyst is an effector for Sec4p, targeting secretory vesicles to sites of exocytosis.* EMBO J, 1999. **18**(4): p. 1071-80.

47. Wu, S., et al., *Sec15 interacts with Rab11 via a novel domain and affects Rab11 localization in vivo.* Nat Struct Mol Biol, 2005. **12**(10): p. 879-85.

48. Simonsen, A., et al., *EEA1 links PI(3)K function to Rab5 regulation of endosome fusion.* Nature, 1998. **394**(6692): p. 494-8.

49. Chen, Y.A. and R.H. Scheller, *SNARE-mediated membrane fusion.* Nat Rev Mol Cell Biol, 2001. **2**(2): p. 98-106.

50. Nielsen, E., et al., *Rabenosyn-5, a novel Rab5 effector, is complexed with hVPS45 and recruited to endosomes through a FYVE finger domain.* J Cell Biol, 2000. **151**(3): p. 601-12.

51. Marnix Wieffer, M.P.M., and Marino Zerial, *Rab Proteins and the Organization of Organelle Membrane Domains*, in *Ras Superfamily Small G*

Proteins: Biology and Mechanisms 2, A. Wittinghofer, Editor. 2014, Springer. p. 17-38.

52. Seabra, M.C. and C. Wasmeier, *Controlling the location and activation of Rab GTPases.* Curr Opin Cell Biol, 2004. **16**(4): p. 451-7.

53. Guo, Z., et al., *Structures of RabGGTase-substrate/product complexes provide insights into the evolution of protein prenylation.* EMBO J, 2008. **27**(18): p. 2444-56.

54. Barr, F.A., *Review series: Rab GTPases and membrane identity: causal or inconsequential?* J Cell Biol, 2013. **202**(2): p. 191-9.

55. Gerondopoulos, A., et al., *BLOC-3 mutated in Hermansky-Pudlak syndrome is a Rab32/38 guanine nucleotide exchange factor.* Curr Biol, 2012. **22**(22): p. 2135-9.

56. Blumer, J., et al., *RabGEFs are a major determinant for specific Rab membrane targeting.* J Cell Biol, 2013. **200**(3): p. 287-300.

57. Ismail, S.A., et al., *Arl2-GTP and Arl3-GTP regulate a GDI-like transport system for farnesylated cargo.* Nat Chem Biol, 2011. **7**(12): p. 942-9.

58. Roger S. Goody, Y.W., and Aymelt Itzen, *Prenylation of RabGTPases, Their Delivery to Membranes, and Rab Recycling*, in *Ras Superfamily Small G Proteins: Biology and Mechanisms 2*, A. Wittinghofer, Editor. 2014, Springer. p. 3-16.

59. Pereira-Leal, J.B. and M.C. Seabra, *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): p. 1077-87.

60. Dumas, J.J., et al., *Structural basis of activation and GTP hydrolysis in Rab proteins.* Structure, 1999. **7**(4): p. 413-23.

61. Wu, Y.W., et al., *Membrane targeting mechanism of Rab GTPases elucidated by semisynthetic protein probes.* Nat Chem Biol, 2010. **6**(7): p. 534-40.

62. Localization, E.o.R.G.R.B.S.a.T.R.i.C.o.R.F.a., *Ras Superfamily Small G Proteins: Biology and Mechanisms 2*. 2014: Springer. 298.

63. Merithew, E., et al., *Structural plasticity of an invariant hydrophobic triad in the switch regions of Rab GTPases is a determinant of effector recognition.* J Biol Chem, 2001. **276**(17): p. 13982-8.

64. Ostermeier, C. and A.T. Brunger, *Structural basis of Rab effector specificity: crystal structure of the small G protein Rab3A complexed with the effector domain of rabphilin-3A.* Cell, 1999. **96**(3): p. 363-74.

65. Fukuda, M., *Regulation of secretory vesicle traffic by Rab small GTPases.* Cell Mol Life Sci, 2008. **65**(18): p. 2801-13.

66. Vazquez-Martinez, R. and M.M. Malagon, *Rab proteins and the secretory pathway: the case of rab18 in neuroendocrine cells.* Front Endocrinol (Lausanne), 2011. **2**: p. 1.

67. Takai, Y., et al., *Rab3A small GTP-binding protein in Ca(2+)-dependent exocytosis.* Genes Cells, 1996. **1**(7): p. 615-32.

68. Geppert, M. and T.C. Sudhof, *RAB3 and synaptotagmin: the yin and yang of synaptic membrane fusion.* Annu Rev Neurosci, 1998. **21**: p. 75-95.

69. Takamori, S., et al., *Molecular anatomy of a trafficking organelle.* Cell, 2006. **127**(4): p. 831-46.

70. Casey, T.M., J.L. Meade, and E.W. Hewitt, *Organelle proteomics: identification of the exocytic machinery associated with the natural killer cell secretory lysosome.* Mol Cell Proteomics, 2007. **6**(5): p. 767-80.

71. Brunner, Y., et al., *Proteomics analysis of insulin secretory granules.* Mol Cell Proteomics, 2007. **6**(6): p. 1007-17.

72. Rindler, M.J., et al., *Proteomic analysis of pancreatic zymogen granules: identification of new granule proteins.* J Proteome Res, 2007. **6**(8): p. 2978-92.

73. Tsuboi, T. and M. Fukuda, *Rab3A and Rab27A cooperatively regulate the docking step of dense-core vesicle exocytosis in PC12 cells.* J Cell Sci, 2006. **119**(Pt 11): p. 2196-203.

74. Menasche, G., et al., *Mutations in RAB27A cause Griscelli syndrome associated with haemophagocytic syndrome.* Nat Genet, 2000. **25**(2): p. 173-6.

75. Wu, X., et al., *Rab27a is an essential component of melanosome receptor for myosin Va.* Mol Biol Cell, 2002. **13**(5): p. 1735-49.

76. Wu, X.S., et al., *Identification of an organelle receptor for myosin-Va.* Nat Cell Biol, 2002. **4**(4): p. 271-8.

77. Henry, J.P., et al., *Role of Rab3a in neurotransmitter and hormone release: a discussion of recent data.* Biochem Soc Trans, 1996. **24**(3): p. 657-61.

78. Geppert, M., et al., *The role of Rab3A in neurotransmitter release.* Nature, 1994. **369**(6480): p. 493-7.

79. Tasaka, K., et al., *Rab3B is essential for GnRH-induced gonadotrophin release from anterior pituitary cells.* J Endocrinol, 1998. **157**(2): p. 267-74.

80. Karniguian, A., A. Zahraoui, and A. Tavitian, *Identification of small GTP-binding rab proteins in human platelets: thrombin-induced phosphorylation of rab3B, rab6, and rab8 proteins.* Proc Natl Acad Sci U S A, 1993. **90**(16): p. 7647-51.

81. Yamamoto, Y., et al., *Distinct roles of Rab3B and Rab13 in the polarized transport of apical, basolateral, and tight junctional membrane proteins to the plasma membrane.* Biochem Biophys Res Commun, 2003. **308**(2): p. 270-5.

82. Fischer von Mollard, G., et al., *Rab3C is a synaptic vesicle protein that dissociates from synaptic vesicles after stimulation of exocytosis.* J Biol Chem, 1994. **269**(15): p. 10971-4.

83. Ohnishi, H., et al., *Overexpression of Rab3D enhances regulated amylase secretion from pancreatic acini of transgenic mice.* J Clin Invest, 1997. **100**(12): p. 3044-52.

84. Tuvim, M.J., et al., *Rab3D, a small GTPase, is localized on mast cell secretory granules and translocates to the plasma membrane upon exocytosis.* Am J Respir Cell Mol Biol, 1999. **20**(1): p. 79-89.

85. Baldini, G., et al., *Expression of Rab3D N135I inhibits regulated secretion of ACTH in AtT-20 cells.* J Cell Biol, 1998. **140**(2): p. 305-13.

86. Knop, M., et al., *Rab3D and annexin A2 play a role in regulated secretion of vWF, but not tPA, from endothelial cells.* EMBO J, 2004. **23**(15): p. 2982-92.

87. Nashida, T., A. Imai, and H. Shimomura, *Relation of Rab26 to the amylase release from rat parotid acinar cells.* Arch Oral Biol, 2006. **51**(2): p. 89-95.

88. Li, C., et al., *Rab26 modulates the cell surface transport of alpha2-adrenergic receptors from the Golgi.* J Biol Chem, 2012. **287**(51): p. 42784-94.

89. Fukuda, M., *Rab27 effectors, pleiotropic regulators in secretory pathways.* Traffic, 2013. **14**(9): p. 949-63.

90. Ljubicic, S., et al., *The GTPase Rab37 Participates in the Control of Insulin Exocytosis.* PLoS One, 2013. **8**(6): p. e68255.

91. Tsai, C.H., et al., *Small GTPase Rab37 targets tissue inhibitor of metalloproteinase 1 for exocytosis and thus suppresses tumour metastasis.* Nat Commun, 2014. **5**: p. 4804.

92. Fukuda, M., T.S. Kuroda, and K. Mikoshiba, *Slac2-a/melanophilin, the missing link between Rab27 and myosin Va: implications of a tripartite protein complex for melanosome transport.* J Biol Chem, 2002. **277**(14): p. 12432-6.

93. Bahadoran, P., et al., *Characterization of the molecular defects in Rab27a, caused by RAB27A missense mutations found in patients with Griscelli syndrome.* J Biol Chem, 2003. **278**(13): p. 11386-92.

94. Ostrowski, M., et al., *Rab27a and Rab27b control different steps of the exosome secretion pathway.* Nat Cell Biol, 2010. **12**(1): p. 19-30; sup pp 1-13.

95. Kosaka, N., et al., *Dark side of the exosome: the role of the exosome in cancer metastasis and targeting the exosome as a strategy for cancer therapy.* Future Oncol, 2014. **10**(4): p. 671-81.

96. Bobrie, A., et al., *Rab27a supports exosome-dependent and -independent mechanisms that modify the tumor microenvironment and can promote tumor progression.* Cancer Res, 2012. **72**(19): p. 4920-30.

97. Recchi, C. and M.C. Seabra, *Novel functions for Rab GTPases in multiple aspects of tumour progression.* Biochem Soc Trans, 2012. **40**(6): p. 1398-403.

98. Singh, R.K., et al., *Distinct and opposing roles for Rab27a/Mlph/MyoVa and Rab27b/Munc13-4 in mast cell secretion.* FEBS J, 2013. **280**(3): p. 892-903.

99. Johnson, J.L., et al., *Rab27a and Rab27b regulate neutrophil azurophilic granule exocytosis and NADPH oxidase activity by independent mechanisms.* Traffic, 2010. **11**(4): p. 533-47.

100. Fukuda, M., *Versatile role of Rab27 in membrane trafficking: focus on the Rab27 effector families.* J Biochem, 2005. **137**(1): p. 9-16.

101. Fukuda, M., *Synaptotagmin-like protein (Slp) homology domain 1 of Slac2-a/melanophilin is a critical determinant of GTP-dependent specific binding to Rab27A.* J Biol Chem, 2002. **277**(42): p. 40118-24.

102. Saegusa, C., et al., *Expression of Rab27B-binding protein Slp1 in pancreatic acinar cells and its involvement in amylase secretion.* Arch Biochem Biophys, 2008. **475**(1): p. 87-92.

103. Neumuller, O., et al., *Synaptotagmin-like protein 1 interacts with the GTPase-activating protein Rap1GAP2 and regulates dense granule secretion in platelets.* Blood, 2009. **114**(7): p. 1396-404.

104. Holt, O., et al., *Slp1 and Slp2-a localize to the plasma membrane of CTL and contribute to secretion from the immunological synapse.* Traffic, 2008. **9**(4): p. 446-57.

105. Wang, H., et al., *The Rab27a effector exophilin7 promotes fusion of secretory granules that have not been docked to the plasma membrane.* Mol Biol Cell, 2013. **24**(3): p. 319-30.

106. Kuroda, T.S. and M. Fukuda, *Rab27A-binding protein Slp2-a is required for peripheral melanosome distribution and elongated cell shape in melanocytes.* Nat Cell Biol, 2004. **6**(12): p. 1195-203.

107. Saegusa, C., et al., *Decreased basal mucus secretion by Slp2-a-deficient gastric surface mucous cells.* Genes Cells, 2006. **11**(6): p. 623-31.

108. Yu, M., et al., *Exophilin4/Slp2-a targets glucagon granules to the plasma membrane through unique Ca2+-inhibitory phospholipid-binding activity of the C2A domain.* Mol Biol Cell, 2007. **18**(2): p. 688-96.

109. Tsuboi, T. and M. Fukuda, *The Slp4-a linker domain controls exocytosis through interaction with Munc18-1.syntaxin-1a complex.* Mol Biol Cell, 2006. **17**(5): p. 2101-12.

110. Gomi, H., et al., *Granuphilin molecularly docks insulin granules to the fusion machinery.* J Cell Biol, 2005. **171**(1): p. 99-109.

111. Fukuda, M., et al., *Slp4-a/granuphilin-a interacts with syntaxin-2/3 in a Munc18-2-dependent manner.* J Biol Chem, 2005. **280**(47): p. 39175-84.

112. Kariya, Y., et al., *Rab27a and Rab27b are involved in stimulation-dependent RANKL release from secretory lysosomes in osteoblastic cells.* J Bone Miner Res, 2011. **26**(4): p. 689-703.

113. Saxena, S.K., H. Horiuchi, and M. Fukuda, *Rab27a regulates epithelial sodium channel (ENaC) activity through synaptotagmin-like protein (SLP-*

5) and Munc13-4 effector mechanism. Biochem Biophys Res Commun, 2006. **344**(2): p. 651-7.

114. Colvin, R.A., et al., *Synaptotagmin-mediated vesicle fusion regulates cell migration.* Nat Immunol, 2010. **11**(6): p. 495-502.

115. Handley, M.T. and R.D. Burgoyne, *The Rab27 effector Rabphilin, unlike Granuphilin and Noc2, rapidly exchanges between secretory granules and cytosol in PC12 cells.* Biochem Biophys Res Commun, 2008. **373**(2): p. 275-81.

116. El-Amraoui, A., et al., *MyRIP, a novel Rab effector, enables myosin VIIa recruitment to retinal melanosomes.* EMBO Rep, 2002. **3**(5): p. 463-70.

117. Nightingale, T.D., et al., *Rab27a and MyRIP regulate the amount and multimeric state of VWF released from endothelial cells.* Blood, 2009. **113**(20): p. 5010-8.

118. Desnos, C., et al., *Rab27A and its effector MyRIP link secretory granules to F-actin and control their motion towards release sites.* J Cell Biol, 2003. **163**(3): p. 559-70.

119. Imai, A., et al., *The small GTPase Rab27B regulates amylase release from rat parotid acinar cells.* J Cell Sci, 2004. **117**(Pt 10): p. 1945-53.

120. Fukuda, M., E. Kanno, and A. Yamamoto, *Rabphilin and Noc2 are recruited to dense-core vesicles through specific interaction with Rab27A in PC12 cells.* J Biol Chem, 2004. **279**(13): p. 13065-75.

121. Cheviet, S., et al., *The Rab-binding protein Noc2 is associated with insulin-containing secretory granules and is essential for pancreatic beta-cell exocytosis.* Mol Endocrinol, 2004. **18**(1): p. 117-26.

122. Imai, A., et al., *Functional involvement of Noc2, a Rab27 effector, in rat parotid acinar cells.* Arch Biochem Biophys, 2006. **455**(2): p. 127-35.

123. Neeft, M., et al., *Munc13-4 is an effector of rab27a and controls secretion of lysosomes in hematopoietic cells.* Mol Biol Cell, 2005. **16**(2): p. 731-41.

124. Kimura, T., et al., *The GDP-dependent Rab27a effector coronin 3 controls endocytosis of secretory membrane in insulin-secreting cell lines.* J Cell Sci, 2008. **121**(Pt 18): p. 3092-8.

125. Kimura, T., et al., *Activated Cdc42-bound IQGAP1 determines the cellular endocytic site.* Mol Cell Biol, 2013. **33**(24): p. 4834-43.

126. Wang, J., et al., *Novel rabphilin-3-like protein associates with insulin-containing granules in pancreatic beta cells.* J Biol Chem, 1999. **274**(40): p. 28542-8.

127. Kuroda, T.S., et al., *The Slp homology domain of synaptotagmin-like proteins 1-4 and Slac2 functions as a novel Rab27A binding domain.* J Biol Chem, 2002. **277**(11): p. 9212-8.

128. Fukuda, M., *Distinct Rab binding specificity of Rim1, Rim2, rabphilin, and Noc2. Identification of a critical determinant of Rab3A/Rab27A recognition by Rim2.* J Biol Chem, 2003. **278**(17): p. 15373-80.

129. Shirakawa, R., et al., *Munc13-4 is a GTP-Rab27-binding protein regulating dense core granule secretion in platelets.* J Biol Chem, 2004. **279**(11): p. 10730-7.

130. Yamaoka, M., T. Ishizaki, and T. Kimura, *GTP- and GDP-Dependent Rab27a Effectors in Pancreatic Beta-Cells.* Biol Pharm Bull, 2015. **38**(5): p. 663-8.

131. Fukuda, M., *Rab27 and its effectors in secretory granule exocytosis: a novel docking machinery composed of a Rab27.effector complex.* Biochem Soc Trans, 2006. **34**(Pt 5): p. 691-5.

132. Fukuda, M. and T.S. Kuroda, *Slac2-c (synaptotagmin-like protein homologue lacking C2 domains-c), a novel linker protein that interacts with Rab27, myosin Va/VIIa, and actin.* J Biol Chem, 2002. **277**(45): p. 43096-103.

133. Arimura, N., et al., *Anterograde transport of TrkB in axons is mediated by direct interaction with Slp1 and Rab27.* Dev Cell, 2009. **16**(5): p. 675-86.

134. Kurowska, M., et al., *Terminal transport of lytic granules to the immune synapse is mediated by the kinesin-1/Slp3/Rab27a complex.* Blood, 2012. **119**(17): p. 3879-89.

135. Fukuda, M., *Distinct Rab27A binding affinities of Slp2-a and Slac2-a/melanophilin: Hierarchy of Rab27A effectors.* Biochem Biophys Res Commun, 2006. **343**(2): p. 666-74.

136. Kukimoto-Niino, M., et al., *Structural basis for the exclusive specificity of Slac2-a/melanophilin for the Rab27 GTPases.* Structure, 2008. **16**(10): p. 1478-90.

137. Chavas, L.M., et al., *Elucidation of Rab27 recruitment by its effectors: structure of Rab27a bound to Exophilin4/Slp2-a.* Structure, 2008. **16**(10): p. 1468-77.

138. Haynes, L.P., et al., *A direct inhibitory role for the Rab3-specific effector, Noc2, in Ca2+-regulated exocytosis in neuroendocrine cells.* J Biol Chem, 2001. **276**(13): p. 9726-32.

139. Gitler, A.D., et al., *The Parkinson's disease protein alpha-synuclein disrupts cellular Rab homeostasis.* Proc Natl Acad Sci U S A, 2008. **105**(1): p. 145-50.

140. Li, X., et al., *Mutant huntingtin impairs vesicle formation from recycling endosomes by interfering with Rab11 activity.* Mol Cell Biol, 2009. **29**(22): p. 6106-16.

141. Jenkins, D., et al., *RAB23 mutations in Carpenter syndrome imply an unexpected role for hedgehog signaling in cranial-suture development and obesity.* Am J Hum Genet, 2007. **80**(6): p. 1162-70.

142. Harrison, R.E., et al., *Salmonella impairs RILP recruitment to Rab7 during maturation of invasion vacuoles.* Mol Biol Cell, 2004. **15**(7): p. 3146-54.

143. Rzomp, K.A., et al., *Rab GTPases are recruited to chlamydial inclusions in both a species-dependent and species-independent manner.* Infect Immun, 2003. **71**(10): p. 5855-70.

144. Machner, M.P. and R.R. Isberg, *Targeting of host Rab GTPase function by the intravacuolar pathogen Legionella pneumophila.* Dev Cell, 2006. **11**(1): p. 47-56.

145. Yoon, S.O., S. Shin, and A.M. Mercurio, *Hypoxia stimulates carcinoma invasion by stabilizing microtubules and promoting the Rab11 trafficking of the alpha6beta4 integrin.* Cancer Res, 2005. **65**(7): p. 2761-9.

146. Caswell, P.T., et al., *Rab25 associates with alpha5beta1 integrin to promote invasive migration in 3D microenvironments.* Dev Cell, 2007. **13**(4): p. 496-510.

147. Pellinen, T., et al., *Small GTPase Rab21 regulates cell adhesion and controls endosomal traffic of beta1-integrins.* J Cell Biol, 2006. **173**(5): p. 767-80.

148. Bravo-Cordero, J.J., et al., *MT1-MMP proinvasive activity is regulated by a novel Rab8-dependent exocytic pathway.* EMBO J, 2007. **26**(6): p. 1499-510.

149. Hendrix, A., et al., *The secretory small GTPase Rab27B as a marker for breast cancer progression.* Oncotarget, 2010. **1**(4): p. 304-8.

150. Barbarin, A. and R. Frade, *Procathepsin L secretion, which triggers tumour progression, is regulated by Rab4a in human melanoma cells.* Biochem J, 2011. **437**(1): p. 97-107.

151. Young, J., J. Menetrey, and B. Goud, *RAB6C is a retrogene that encodes a centrosomal protein involved in cell cycle progression.* J Mol Biol, 2010. **397**(1): p. 69-88.

152. Ferrandiz-Huertas, C., A. Fernandez-Carvajal, and A. Ferrer-Montiel, *Rab4 interacts with the human P-glycoprotein and modulates its surface expression in multidrug resistant K562 cells.* Int J Cancer, 2011. **128**(1): p. 192-205.

153. Shen, D.W. and M.M. Gottesman, *RAB8 enhances TMEM205-mediated cisplatin resistance.* Pharm Res, 2012. **29**(3): p. 643-50.

154. Eggenschwiler, J.T., E. Espinoza, and K.V. Anderson, *Rab23 is an essential negative regulator of the mouse Sonic hedgehog signalling pathway.* Nature, 2001. **412**(6843): p. 194-8.

155. Wilson, S.M., et al., *A mutation in Rab27a causes the vesicle transport defects observed in ashen mice.* Proc Natl Acad Sci U S A, 2000. **97**(14): p. 7933-8.

156. Wang, J.S., et al., *Enhanced expression of Rab27A gene by breast cancer cells promoting invasiveness and the metastasis potential by secretion of insulin-like growth factor-II.* Mol Cancer Res, 2008. **6**(3): p. 372-82.

157. Li, W., et al., *Rab27A regulates exosome secretion from lung adenocarcinoma cells A549: involvement of EPI64.* APMIS, 2014. **122**(11): p. 1080-7.

158. Wang, Q., et al., *High expression of RAB27A and TP53 in pancreatic cancer predicts poor survival.* Med Oncol, 2015. **32**(1): p. 372.

159. Dong, W.W., et al., *Differential expression of Rab27A/B correlates with clinical outcome in hepatocellular carcinoma.* World J Gastroenterol, 2012. **18**(15): p. 1806-13.

160. Peinado, H., et al., *Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET.* Nat Med, 2012. **18**(6): p. 883-91.

161. Montel, V., et al., *Expression profiling of primary tumors and matched lymphatic and lung metastases in a xenogeneic breast cancer model.* Am J Pathol, 2005. **166**(5): p. 1565-79.

162. Webber, J., et al., *Cancer exosomes trigger fibroblast to myofibroblast differentiation.* Cancer Res, 2010. **70**(23): p. 9621-30.

163. Zhao, W., et al., *Direct Targeting of the Ras GTPase Superfamily through Structure-Based Design.* Curr Top Med Chem, 2016.

164. Spiegel, J., et al., *Small-molecule modulation of Ras signaling.* Nat Chem Biol, 2014. **10**(8): p. 613-22.

165. Cromm, P.M., et al., *Direct Modulation of Small GTPase Activity and Function.* Angew Chem Int Ed Engl, 2015. **54**(46): p. 13516-37.

166. Spiegel, J., et al., *Direct targeting of Rab-GTPase-effector interactions.* Angew Chem Int Ed Engl, 2014. **53**(9): p. 2498-503.

167. Patgiri, A., et al., *An orthosteric inhibitor of the Ras-Sos interaction.* Nat Chem Biol, 2011. **7**(9): p. 585-7.

168. Agola, J.O., et al., *A Competitive Nucleotide Binding Inhibitor: In Vitro Characterization of Rab7 GTPase Inhibition.* ACS Chem Biol, 2012. **7**(6): p. 1095-1108.

169. Hong, L., et al., *A Pan-GTPase Inhibitor as a Molecular Probe.* PLoS ONE, 2015. **10**(8): p. e0134317.

170. Cardama, G.A., et al., *Preclinical development of novel Rac1-GEF signaling inhibitors using a rational design approach in highly aggressive breast cancer cell lines.* Anticancer Agents Med Chem, 2014. **14**(6): p. 840-51.

171. Sun, Q., et al., *Discovery of Small Molecules that Bind to K-Ras and Inhibit Sos-Mediated Activation.* Angew Chem Int Edit, 2012. **51**(25): p. 6140-6143.

172. Ostrem, J.M., et al., *K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions.* Nature, 2013. **503**(7477): p. 548-551.

173. Sun, Q., et al., *A method for the second-site screening of K-Ras in the presence of a covalently attached first-site ligand.* Journal of Biomolecular NMR, 2014. **60**(1): p. 11-14.

174. Eberth, A. and M.R. Ahmadian, *In vitro GEF and GAP assays.* Curr Protoc Cell Biol, 2009. **Chapter 14**: p. Unit 14 9.

175. Niesen, F.H., H. Berglund, and M. Vedadi, *The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability.* Nat Protoc, 2007. **2**(9): p. 2212-21.

176. Holdgate, G.A. and W.H. Ward, *Measurements of binding thermodynamics in drug discovery.* Drug Discov Today, 2005. **10**(22): p. 1543-50.

177. Hwang, T.L. and A.J. Shaka, *Water Suppression That Works - Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients.* Journal of Magnetic Resonance Series A, 1995. **112**(2): p. 275-279.

178. Vuister, G.W. and A. Bax, *Resolution Enhancement and Spectral Editing of Uniformly C-13-Enriched Proteins by Homonuclear Broad-Band C-13 Decoupling.* Journal of Magnetic Resonance, 1992. **98**(2): p. 428-435.

179. Williamson, M.P., *Applications of the NOE in Molecular Biology.* Annual Reports on NMR Spectroscopy, 2009. **65**: p. 77-109.

180. Leung, I.K.H., *NMR spectroscopy – a simple yet powerful tool in chemical biology.* Chemistry in New Zealand, 2015: p. 42-50.

181. Wang, Y.S., D. Liu, and D.F. Wyss, *Competition STD NMR for the detection of high-affinity ligands and NMR-based screening.* Magn Reson Chem, 2004. **42**(6): p. 485-9.

182. Begley, D.W., et al., *Saturation transfer difference NMR for fragment screening.* Curr Protoc Chem Biol, 2013. **5**(4): p. 251-68.

183. Mayer, M. and B. Meyer, *Characterization of ligand binding by saturation transfer difference NMR spectroscopy.* Angewandte Chemie-International Edition, 1999. **38**(12): p. 1784-1788.

184. Dalvit, C., et al., *WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability.* J Biomol NMR, 2001. **21**(4): p. 349-59.

185. Markin, C.J. and L. Spyracopoulos, *Increased precision for analysis of protein-ligand dissociation constants determined from chemical shift titrations.* J Biomol NMR, 2012. **53**(2): p. 125-38.

186. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.

187. Johnson, B.A. and R.A. Blevins, *NMR View: A computer program for the visualization and analysis of NMR data.* J Biomol NMR, 1994. **4**(5): p. 603-14.

188. LaPlante, S.R., et al., *Compound aggregation in drug discovery: implementing a practical NMR assay for medicinal chemists.* J Med Chem, 2013. **56**(12): p. 5142-50.

189. Waterman, D.G., et al., *Diffraction-geometry refinement in the DIALS framework.* Acta Crystallogr D Struct Biol, 2016. **72**(Pt 4): p. 558-75.

190. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

191. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.

192. Winn, M.D., G.N. Murshudov, and M.Z. Papiz, *Macromolecular TLS refinement in REFMAC at moderate resolutions.* Methods Enzymol, 2003. **374**: p. 300-21.

193. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.

194. pymol.bibBibDesk334Schrodinger, LLCThe PyMOL Molecular Graphics System, Version 1.82015NovemberPyMOLPyMOL The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC.unpublished19/08/201022/12/2015.

195. Duan, J., et al., *Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods.* J Mol Graph Model, 2010. **29**(2): p. 157-70.

196. Hawkins, P.C. and A. Nicholls, *Conformer generation with OMEGA: learning from the data set and the analysis of failures.* J Chem Inf Model, 2012. **52**(11): p. 2919-36.

197. Hawkins, P.C., A.G. Skillman, and A. Nicholls, *Comparison of shape-matching and docking as virtual screening tools.* J Med Chem, 2007. **50**(1): p. 74-82.

198. ANDREW R. LEACH, V.J.G., *AN INTRODUCTION TO CHEMOINFORMATICS*. 2007: Springer. 255.

199. Appelt, K., *Crystal structures of HIV-1 protease-inhibitor complexes*

. Perspectives in Drug Discovery and Design, 1993. **1**(1): p. 23-48.

200. Natesh, R., et al., *Crystal structure of the human angiotensin-converting enzyme-lisinopril complex.* Nature, 2003. **421**(6922): p. 551-4.

201. Di Fiore, A., et al., *Crystal structure of human carbonic anhydrase XIII and its complex with the inhibitor acetazolamide.* Proteins, 2009. **74**(1): p. 164-75.

202. Wade, R.C., *'Flu' and structure-based drug design.* Structure, 1997. **5**(9): p. 1139-45.

203. Stubbs, M.T.B., W., *Crystal structures of thrombin and thrombin complexes as a framework for antithrombotic drug design.* Perspectives in Drug Discovery and Design, 1994. **1**(3): p. 431-452.

204. Blundell, T.L., H. Jhoti, and C. Abell, *High-throughput crystallography for lead discovery in drug design.* Nat Rev Drug Discov, 2002. **1**(1): p. 45-54.

205. Muchmore, S.W., et al., *Automated crystal mounting and data collection for protein crystallography.* Structure, 2000. **8**(12): p. R243-6.

206. Winn, M.D., et al., *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.

207. Moraes, I., et al., *Membrane protein structure determination - the next generation.* Biochim Biophys Acta, 2014. **1838**(1 Pt A): p. 78-87.

208. Halgren, T.A., *Identifying and characterizing binding sites and assessing druggability.* J Chem Inf Model, 2009. **49**(2): p. 377-89.

209. Erlanson, D.A., et al., *Twenty years on: the impact of fragments on drug discovery.* Nat Rev Drug Discov, 2016. **15**(9): p. 605-19.

210. Hartshorn, M.J., et al., *Fragment-based lead discovery using X-ray crystallography.* J Med Chem, 2005. **48**(2): p. 403-13.

211. Patel, D., J.D. Bauman, and E. Arnold, *Advantages of crystallographic fragment screening: functional and mechanistic insights from a powerful platform for efficient drug discovery.* Prog Biophys Mol Biol, 2014. **116**(2-3): p. 92-100.

212. Pellecchia, M., et al., *Perspectives on NMR in drug discovery: a technique comes of age.* Nat Rev Drug Discov, 2008. **7**(9): p. 738-45.

213. Petros, A.M., et al., *Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis.* J Med Chem, 2006. **49**(2): p. 656-63.

214. Emsley, J.W. and J. Feeney, *Forty years of Progress in Nuclear Magnetic Resonance Spectroscopy.* Progress in Nuclear Magnetic Resonance Spectroscopy, 2007. **50**(4): p. 179-198.

215. Valkov, E., et al., *Targeting protein-protein interactions and fragment-based drug discovery.* Top Curr Chem, 2012. **317**: p. 145-79.

216. Leonard M. G. CHAVAS, K.I., Masato KAWASAKI, Soichi WAKATSUKI, *Structural Insights into Rab27 Recruitment by its Effectors* Nihon Kessho Gakkaishi 2010. **51**(6): p. 320-326.

217. Moon, A.F., et al., *A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction.* Protein Sci, 2010. **19**(5): p. 901-13.

218. Goldschmidt, L., et al., *Toward rational protein crystallization: A Web server for the design of crystallizable protein variants.* Protein Sci, 2007. **16**(8): p. 1569-76.

219. D'Arcy, A., et al., *Microseed matrix screening for optimization in protein crystallization: what have we learned?* Acta Crystallogr F Struct Biol Commun, 2014. **70**(Pt 9): p. 1117-26.

220. Liu, B., et al., *Promoting crystallisation of the Salmonella enteritidis fimbriae 14 pilin SefD using deuterium oxide.* Biochem Biophys Res Commun, 2012. **421**(2): p. 208-13.

221. McCoy, A.J., *Solving structures of protein complexes by molecular replacement with Phaser.* Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 1): p. 32-41.

222. Storoni, L.C., A.J. McCoy, and R.J. Read, *Likelihood-enhanced fast rotation functions.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 3): p. 432-8.

223. Kantardjieff, K.A. and B. Rupp, *Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals.* Protein Sci, 2003. **12**(9): p. 1865-71.

224. Matthews, B.W., *Solvent content of protein crystals.* J Mol Biol, 1968. **33**(2): p. 491-7.

225. Drenth, J., *Principles of Protein X-Ray Crystallography*. Third edition ed. 2007: Springer.

226. Read, R.J., P.D. Adams, and A.J. McCoy, *Intensity statistics in the presence of translational noncrystallographic symmetry.* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 2): p. 176-83.

227. Read, R. *Molecular Replacement*. 2016; Available from: http://www.phaser.cimr.cam.ac.uk/index.php/Molecular_Replacement.

228. Scott, D.E., et al., *Small molecules, big targets: drug discovery faces the protein-protein interaction challenge.* Nat Rev Drug Discov, 2016. **15**(8): p. 533-50.

229. Villoutreix, B.O., et al., *A leap into the chemical space of protein-protein interaction inhibitors.* Curr Pharm Des, 2012. **18**(30): p. 4648-67.

230. Cunningham, A.D., N. Qvit, and D. Mochly-Rosen, *Peptides and peptidomimetics as regulators of protein-protein interactions.* Curr Opin Struct Biol, 2017. **44**: p. 59-66.

231. Scott, D.E., et al., *Using a fragment-based approach to target protein-protein interactions.* Chembiochem, 2013. **14**(3): p. 332-42.

232. Congreve, M., et al., *A 'rule of three' for fragment-based lead discovery?* Drug Discov Today, 2003. **8**(19): p. 876-7.

233. Hopkins, A.L., et al., *The role of ligand efficiency metrics in drug discovery.* Nat Rev Drug Discov, 2014. **13**(2): p. 105-21.

234. Shuker, S.B., et al., *Discovering high-affinity ligands for proteins: SAR by NMR.* Science, 1996. **274**(5292): p. 1531-4.

235. Keseru, G.M. and G.M. Makara, *Hit discovery and hit-to-lead approaches.* Drug Discov Today, 2006. **11**(15-16): p. 741-8.

236. *Fragments: past, present and future.* Drug Discov Today Technol, 2010. **7**(3): p. e147-202.

237. Hajduk, P.J. and J. Greer, *A decade of fragment-based drug design: strategic advances and lessons learned.* Nat Rev Drug Discov, 2007. **6**(3): p. 211-9.

238. Ma, R., et al., *Process of Fragment-Based Lead Discovery-A Perspective from NMR.* Molecules, 2016. **21**(7).

239. Harner, M.J., A.O. Frank, and S.W. Fesik, *Fragment-based drug discovery using NMR spectroscopy.* J Biomol NMR, 2013. **56**(2): p. 65-75.

240. Chilingaryan, Z., Z. Yin, and A.J. Oakley, *Fragment-based screening by protein crystallography: successes and pitfalls.* Int J Mol Sci, 2012. **13**(10): p. 12857-79.

241. Caliandro, R., et al., *Protein crystallography and fragment-based drug design.* Future Med Chem, 2013. **5**(10): p. 1121-40.

242. Navratilova, I. and A.L. Hopkins, *Fragment screening by surface plasmon resonance.* ACS Med Chem Lett, 2010. **1**(1): p. 44-8.

243. Perspicace, S., et al., *Fragment-based screening using surface plasmon resonance technology.* J Biomol Screen, 2009. **14**(4): p. 337-49.

244. Mashalidis, E.H., et al., *A three-stage biophysical screening cascade for fragment-based drug discovery.* Nat Protoc, 2013. **8**(11): p. 2309-24.

245. Ciulli, A., *Biophysical screening for the discovery of small-molecule ligands.* Methods Mol Biol, 2013. **1008**: p. 357-88.

246. Austin, C., et al., *Fragment screening using capillary electrophoresis (CEfrag) for hit identification of heat shock protein 90 ATPase inhibitors.* J Biomol Screen, 2012. **17**(7): p. 868-76.

247. Chen, X., et al., *A ligand-observed mass spectrometry approach integrated into the fragment based lead discovery pipeline.* Sci Rep, 2015. **5**: p. 8361.

248. Vivat Hannah, V., et al., *Native MS: an 'ESI' way to support structure- and fragment-based drug discovery.* Future Med Chem, 2010. **2**(1): p. 35-50.

249. Linke, P., et al., *An Automated Microscale Thermophoresis Screening Approach for Fragment-Based Lead Discovery.* J Biomol Screen, 2016. **21**(4): p. 414-21.

250. Duong-Thi, M.D., et al., *Weak affinity chromatography as a new approach for fragment screening in drug discovery.* Anal Biochem, 2011. **414**(1): p. 138-46.

251. Daniel A. Erlanson, W.J., *Fragment-based Approaches in Drug Discovery*. 2006: Wiley-VCH.

252. Sledz, P., et al., *High-throughput interrogation of ligand binding mode using a fluorescence-based assay.* Angew Chem Int Ed Engl, 2012. **51**(31): p. 7680-3.

253. Silvestre, H.L., et al., *Integrated biophysical approach to fragment screening and validation for fragment-based lead discovery.* Proc Natl Acad Sci U S A, 2013. **110**(32): p. 12984-9.

254. Wetzel, S., et al., *Interactive exploration of chemical space with Scaffold Hunter.* Nat Chem Biol, 2009. **5**(8): p. 581-3.

255. Bemis, G.W. and M.A. Murcko, *The properties of known drugs. 1. Molecular frameworks.* J Med Chem, 1996. **39**(15): p. 2887-93.

256. Schuffenhauer, A., et al., *The scaffold tree--visualization of the scaffold universe by hierarchical scaffold classification.* J Chem Inf Model, 2007. **47**(1): p. 47-58.

257. Krishnan, V.V., *Ligand screening by saturation-transfer difference (STD) NMR spectroscopy.* Current Analytical Chemistry, 2005. **1**(3): p. 307-320.

258. Devine, S.M., et al., *Promiscuous 2-aminothiazoles (PrATs): a frequent hitting scaffold.* J Med Chem, 2015. **58**(3): p. 1205-14.

259. Thomas G. Davies, M.H., *Fragment-Based Drug Discovery and X-Ray Crystallography.* 2013. **317**: p. 225.

260. Davis, B.J. and D.A. Erlanson, *Learning from our mistakes: the 'unknown knowns' in fragment screening.* Bioorg Med Chem Lett, 2013. **23**(10): p. 2844-52.

261. Douse, C.H., et al., *Targeting a dynamic protein-protein interaction: fragment screening against the malaria myosin A motor complex.* ChemMedChem, 2015. **10**(1): p. 134-43.

262. Mayer, M. and T.L. James, *Detecting ligand binding to a small RNA target via saturation transfer difference NMR experiments in D(2)O and H(2)O.* J Am Chem Soc, 2002. **124**(45): p. 13376-7.

263. Ma, B. and R. Nussinov, *Trp/Met/Phe hot spots in protein-protein interactions: potential targets in drug design.* Curr Top Med Chem, 2007. **7**(10): p. 999-1005.

264. Jordan, J.B., et al., *Fragment based drug discovery: practical implementation based on (1)(9)F NMR spectroscopy.* J Med Chem, 2012. **55**(2): p. 678-87.

265. Sattler, M., et al., *Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis.* Science, 1997. **275**(5302): p. 983-6.

266. Sliwoski, G., et al., *Computational methods in drug discovery.* Pharmacol Rev, 2014. **66**(1): p. 334-95.

267. Kapetanovic, I.M., *Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach.* Chem Biol Interact, 2008. **171**(2): p. 165-76.

268. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery: methods and applications.* Nat Rev Drug Discov, 2004. **3**(11): p. 935-49.

269. Meng, X.Y., et al., *Molecular docking: a powerful approach for structure-based drug discovery.* Curr Comput Aided Drug Des, 2011. **7**(2): p. 146-57.

270. Shoichet, B.K. and B.K. Kobilka, *Structure-based drug screening for G-protein-coupled receptors.* Trends Pharmacol Sci, 2012. **33**(5): p. 268-72.

271. Ripphausen, P., et al., *Quo vadis, virtual screening? A comprehensive survey of prospective applications.* J Med Chem, 2010. **53**(24): p. 8461-7.

272. Grant, J.A., M.A. Gallardo, and B.T. Pickup, *A fast method of molecular shape comparison: A simple application of a Gaussian description of molecular shape.* Journal of Computational Chemistry, 1996. **17**(14): p. 1653-1666.

273. Gao, Y., et al., *Rational design and characterization of a Rac GTPase-specific small molecule inhibitor.* Proc Natl Acad Sci U S A, 2004. **101**(20): p. 7618-23.

274. Joung, J.Y., et al., *Identification of novel rab27a/melanophilin blockers by pharmacophore-based virtual screening.* Appl Biochem Biotechnol, 2014. **172**(4): p. 1882-97.

275. Jahnke, W., S. Rudisser, and M. Zurini, *Spin label enhanced NMR screening.* J Am Chem Soc, 2001. **123**(13): p. 3149-50.

276. Jahnke, W., et al., *Second-site NMR screening and linker design.* Curr Top Med Chem, 2003. **3**(1): p. 69-80.

277. Schlichting, I., et al., *Biochemical and crystallographic characterization of a complex of c-Ha-ras p21 and caged GTP with flash photolysis.* Proc Natl Acad Sci U S A, 1989. **86**(20): p. 7687-90.

278. Mueller, M.P. and R.S. Goody, *Review: Ras GTPases and myosin: Qualitative conservation and quantitative diversification in signal and energy transduction.* Biopolymers, 2016. **105**(8): p. 422-30.

279. Agola, J.O., et al., *A competitive nucleotide binding inhibitor: in vitro characterization of Rab7 GTPase inhibition.* ACS Chem Biol, 2012. **7**(6): p. 1095-108.

280. Azzarito, V., et al., *Inhibition of alpha-helix-mediated protein-protein interactions using designed molecules.* Nat Chem, 2013. **5**(3): p. 161-73.

281. Bernal, F., et al., *A stapled p53 helix overcomes HDMX-mediated suppression of p53.* Cancer Cell, 2010. **18**(5): p. 411-22.

282. Cromm, P.M., et al., *Orthogonal ring-closing alkyne and olefin metathesis for the synthesis of small GTPase-targeting bicyclic peptides.* Nat Commun, 2016. **7**: p. 11300.

283. Lepre, C.A., J.M. Moore, and J.W. Peng, *Theory and applications of NMR-based screening in pharmaceutical research.* Chem Rev, 2004. **104**(8): p. 3641-76.

284. Parker, M.J., et al., *A combinatorial selective labeling method for the assignment of backbone amide NMR resonances.* J Am Chem Soc, 2004. **126**(16): p. 5020-1.

285. Koehler, J. and J. Meiler, *Expanding the utility of NMR restraints with paramagnetic compounds: background and practical aspects.* Prog Nucl Magn Reson Spectrosc, 2011. **59**(4): p. 360-89.

286. Jahnke, W., *Spin labels as a tool to identify and characterize protein-ligand interactions by NMR spectroscopy.* Chembiochem, 2002. **3**(2-3): p. 167-73.

287. Saio, T., et al., *An NMR strategy for fragment-based ligand screening utilizing a paramagnetic lanthanide probe.* J Biomol NMR, 2011. **51**(3): p. 395-408.

288. Barthelmes, K., et al., *Engineering encodable lanthanide-binding tags into loop regions of proteins.* J Am Chem Soc, 2011. **133**(4): p. 808-19.

289. Haussinger, D., J.R. Huang, and S. Grzesiek, *DOTA-M8: An extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy.* J Am Chem Soc, 2009. **131**(41): p. 14761-7.

290. Moschetti, T., et al., *Engineering Archeal Surrogate Systems for the Development of Protein-Protein Interaction Inhibitors against Human RAD51.* J Mol Biol, 2016. **428**(23): p. 4589-4607.

291. Johnson, J.L., et al., *Identification of Neutrophil Exocytosis Inhibitors (Nexinhibs), Small Molecule Inhibitors of Neutrophil Exocytosis and*

Inflammation: DRUGGABILITY OF THE SMALL GTPase Rab27a. J Biol Chem, 2016. **291**(50): p. 25965-25982.

APPENDIX

Appendix A

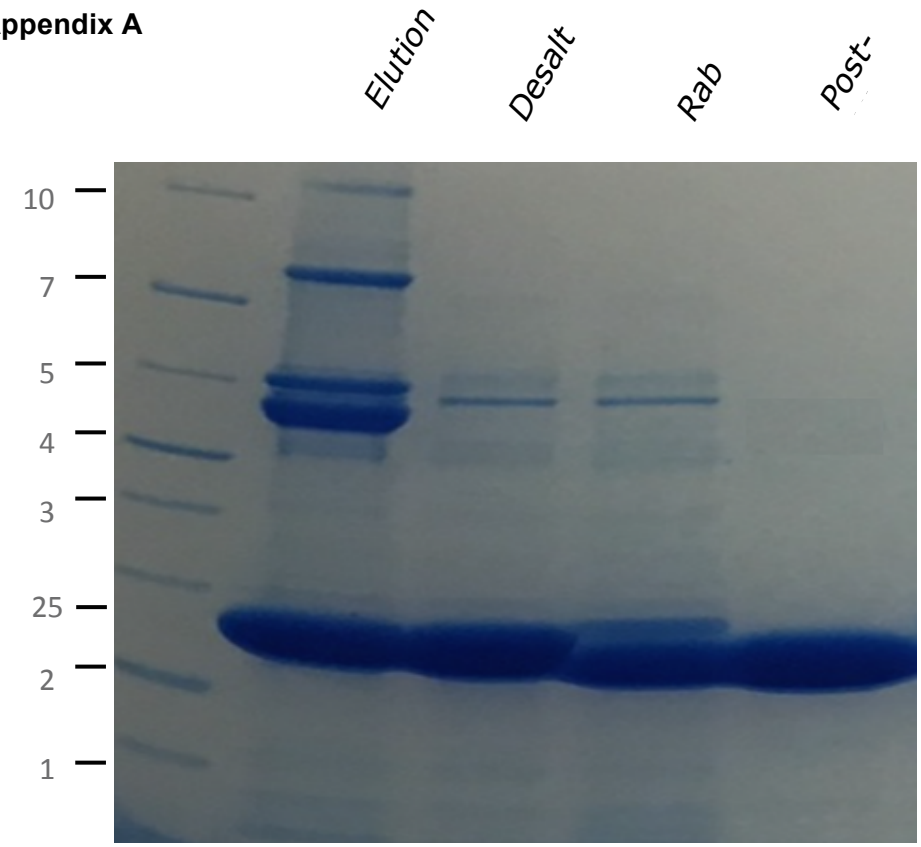


Figure 1.A. Purification of Rab27a using Ni^{2+}-NTA chromatography.

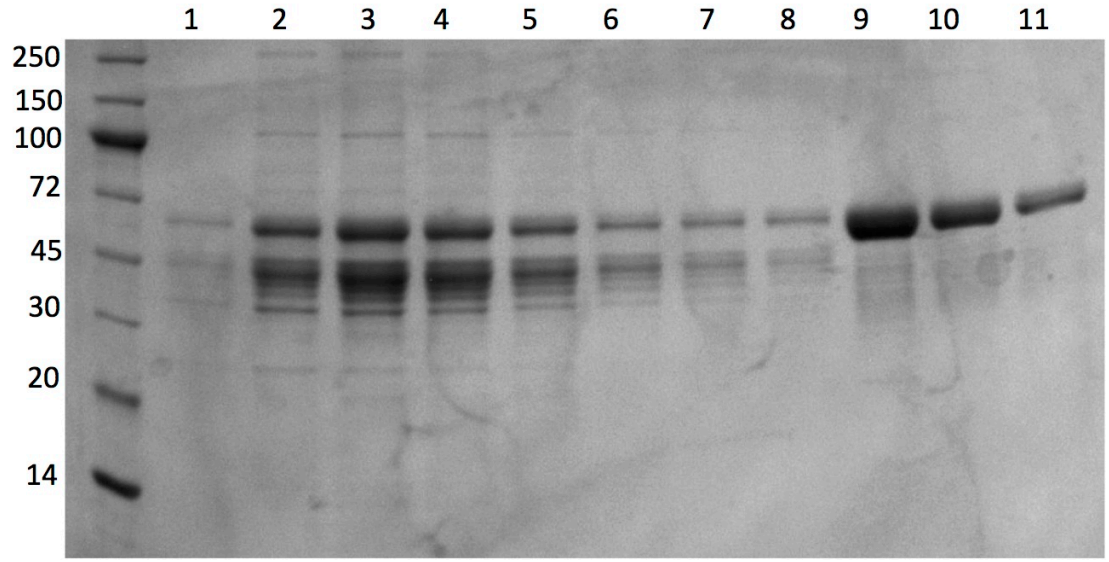


Figure 2.A. Purification of MBP-Rab27a. Protein was purified using amylose resin and gel filtered using a Superdex 200 column from GE. This blot shows the gel filtration result. Lane 9-11 were pooled and taken for crystallography.

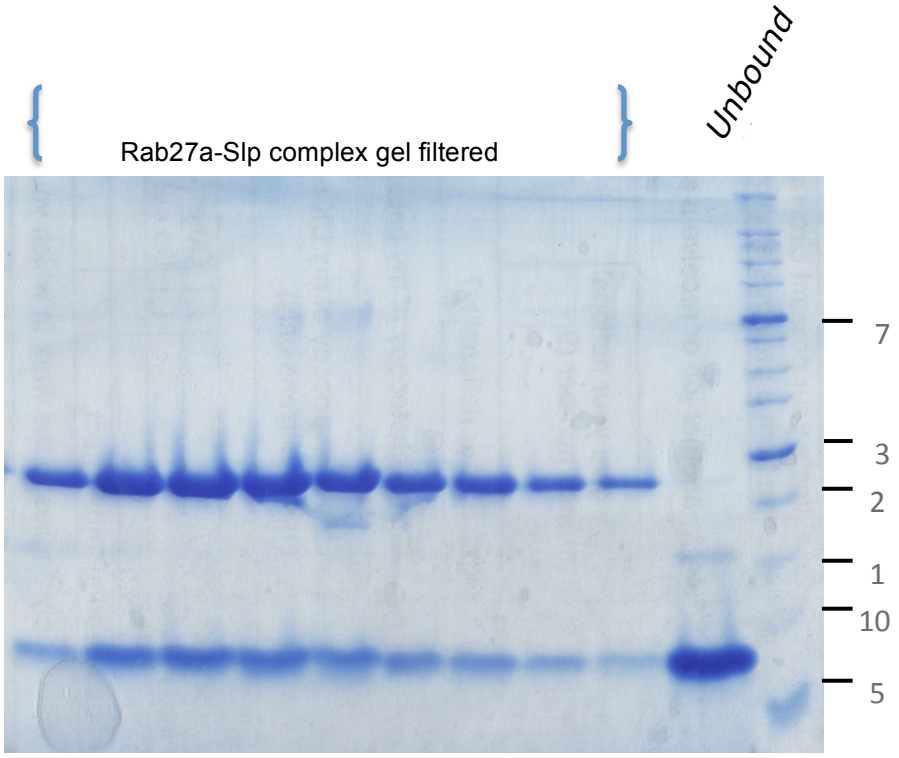


Figure 3.A. Purified Rab27a-Slp complex gel filteration. Rab27a-Slp complex purified by Ni^{2+}-NTA chromatography is gel filtered and separated from molecules not in the complex.

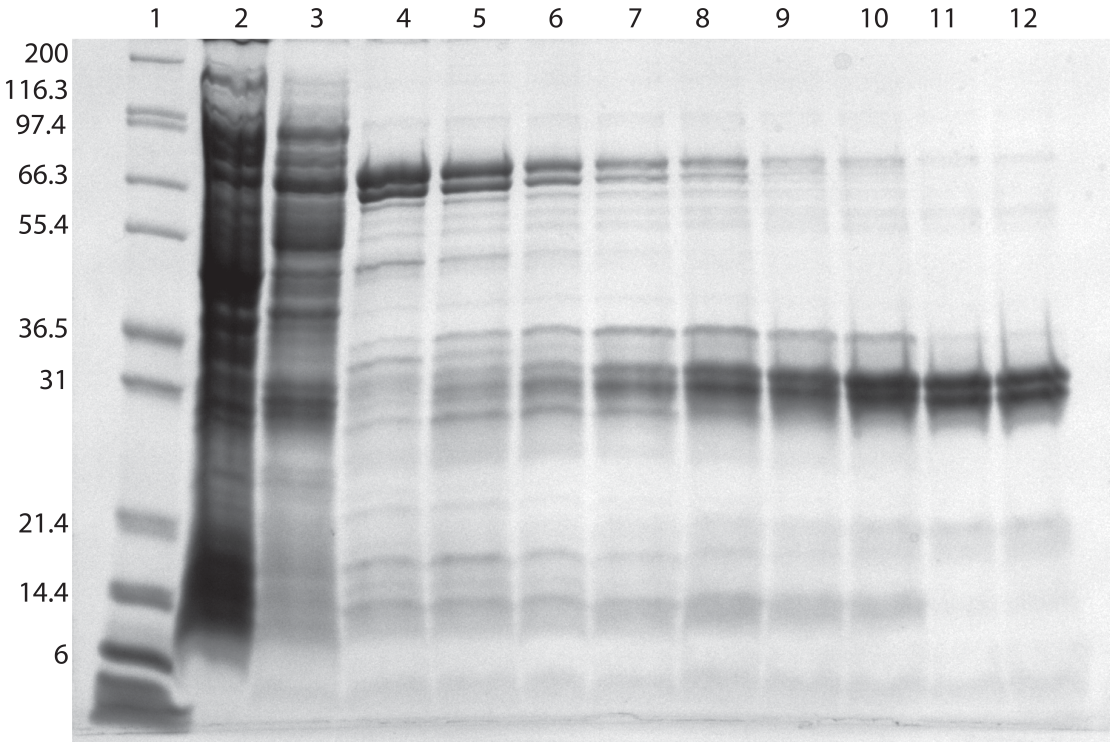


Figure 4.A. Purification of TEV enzyme. Lane 1: Ladder, Lane 2: Supernatant, Lane 3: Flow-through, Lane: 4-7: Wash, Lane: 8-12: Elution.

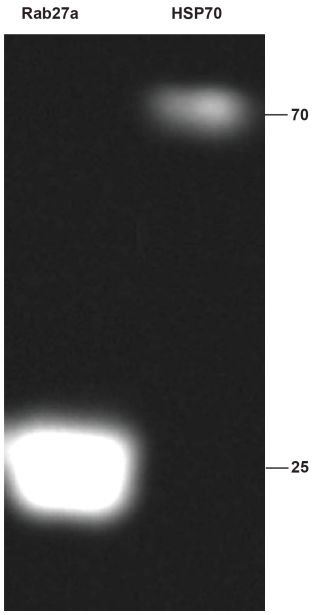


Figure 5.A. Validation of Rab27a biotinylation by western blot. Biotinylation of Rab27a proteins biotinylated according to methods described in section 2.7.1 is shown. A biotinylated HSP70 provided by Lab of Dr. Van Montfort at ICR was used as a positive control.

‘

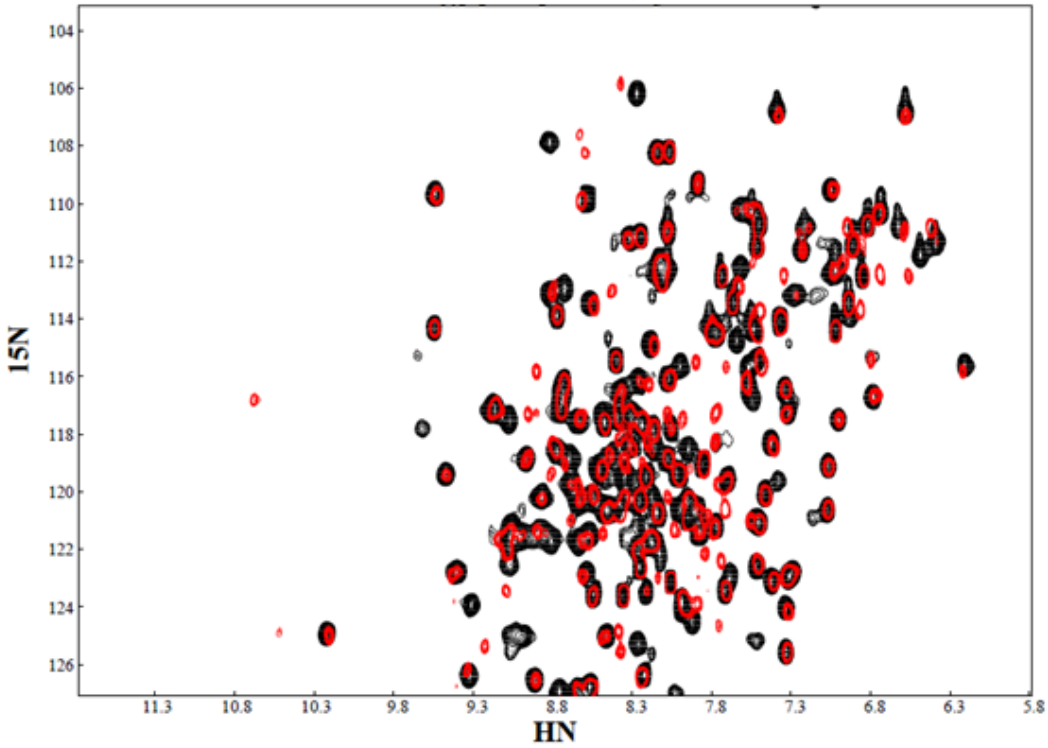
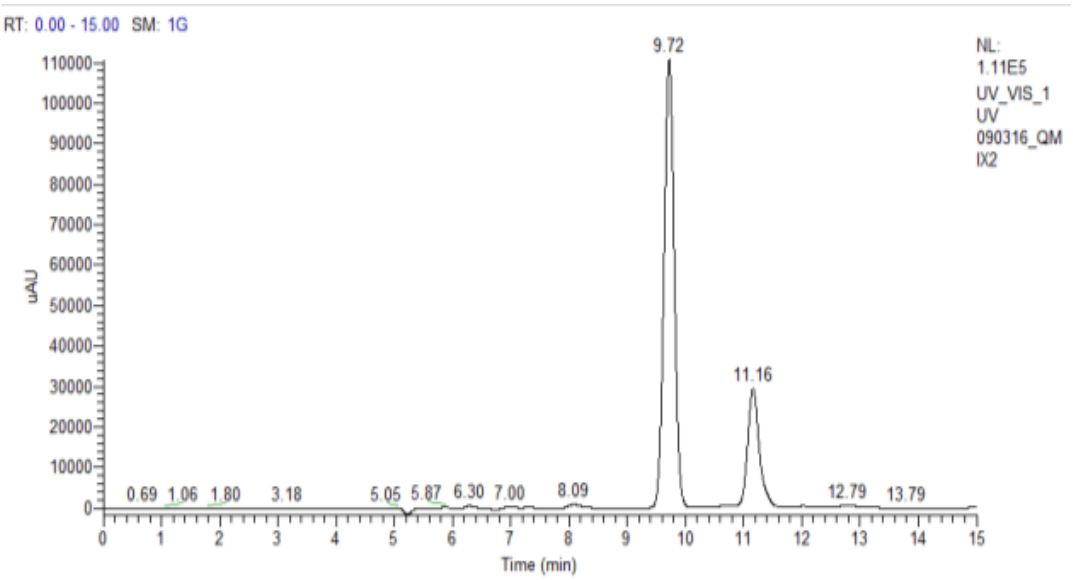


Figure 6.A. Overlay of ^{15}N- ^{1}H HSQC spectrums of purified Rab27a (in red) and purified Rab27a loaded with gpp(NH)p (in black).

A



B

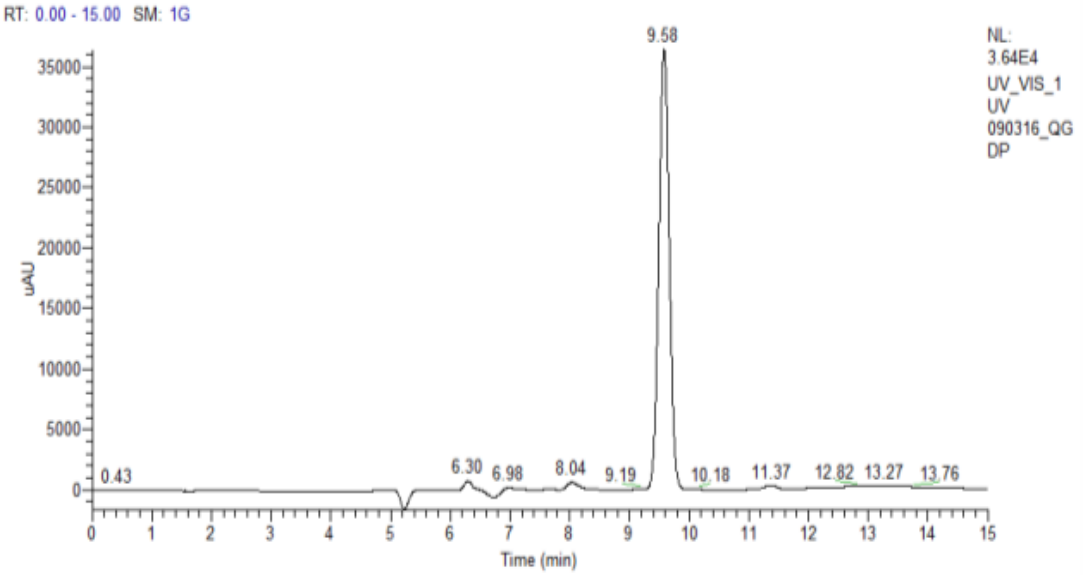
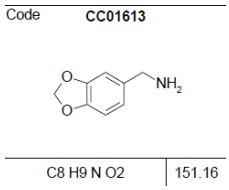
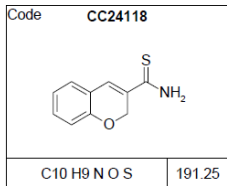
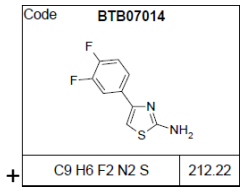
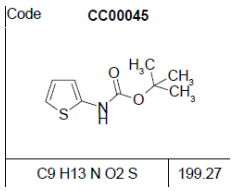
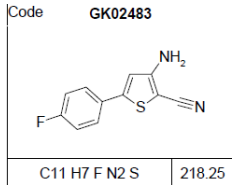
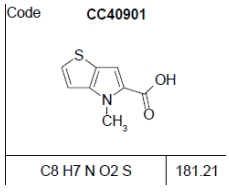
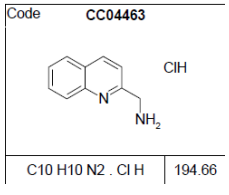
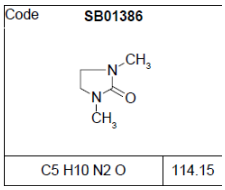
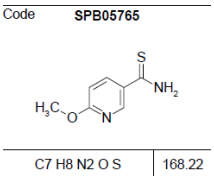
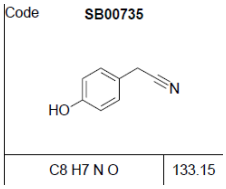
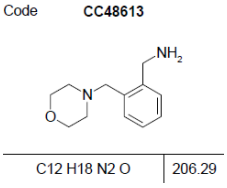
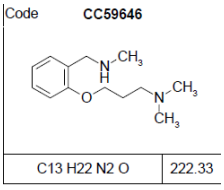
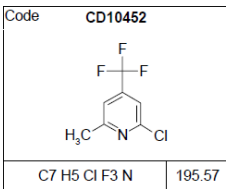
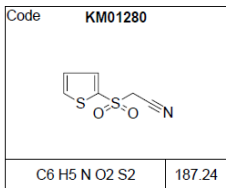
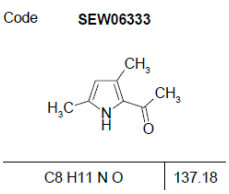
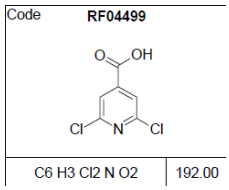
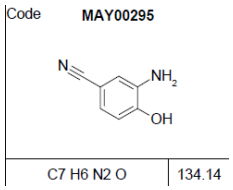
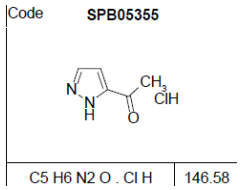
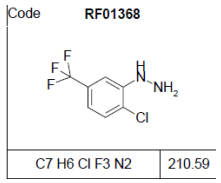
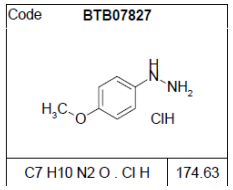
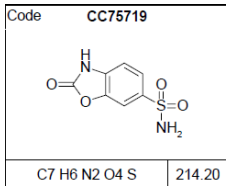
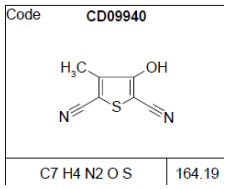
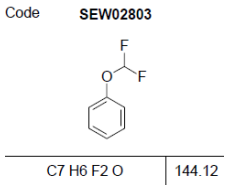


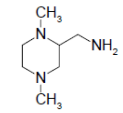
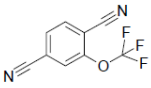
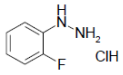
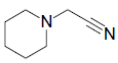
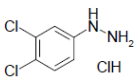
Figure 7.A. Validating nucleotide exchange by HPLC. A) Rab27a before nucleotide exchange. The majority of the Rab27a is bound to GDP. B) Rab27a after exchange with GDP. All the protein is bound to GDP.

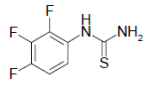
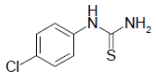
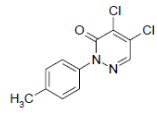
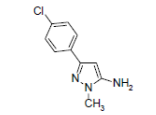
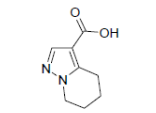
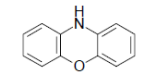
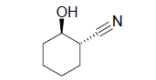
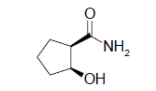
Appendix B. Hits obtained from fragment library screen against Rab27a using thermal-shift assay.

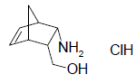
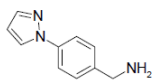
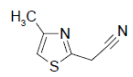
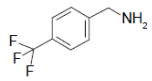
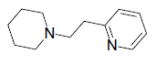
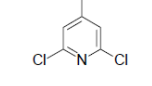
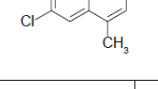
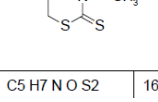
Plate Code	Tm	Δ>ΔTm+2SD+0.5	Maybridge code	Ligand structure
1A6	60.05	0.418165	CC01613	
1B9	60.08	0.448165	CC24118	
1F3	60.61	0.978165	BTB07014	
1F5	60.05	0.418165	CC00045	
2F7	60.26	0.214222	GK02483	
2G2	60.79	0.744222	CC40901	
3A11	59.52	0.364544	CC04463	

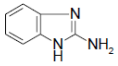
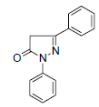
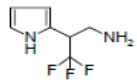
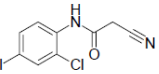
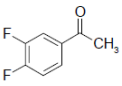
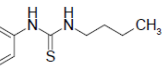
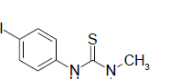
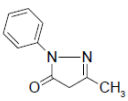
3C3	59.38	0.224544	SB01386	
3C6	59.17	0.014544	SPB05765	
3E2	59.59	0.434544	SB00735	
5A3	59.59	0.089914	CC48613	
5C11	59.64	0.139914	CC59646	
6F8	59.3	0.391703	CD10452	
7A11	58.83	0.139965	KM01280	
7B8	59.02	0.329965	SEW06333	

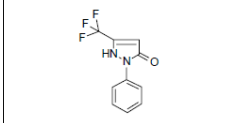
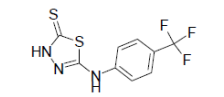
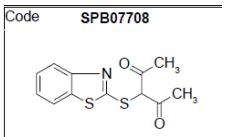
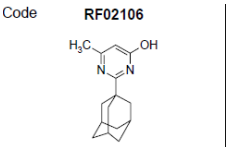
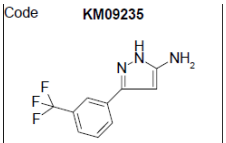
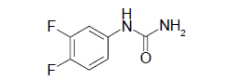
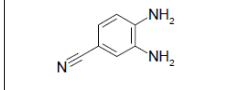
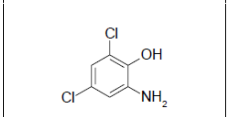
7C7	59.21	0.519965	RF04499	
7D2	60.64	1.949965	MAY00295	
7E8	58.77	0.079965	SPB05355	
7G6	60.01	1.319965	RF01368	
8D10	60.65	0.198057	BTB07827	
9A5	59.78	0.80567	CC75719	
9F10	60.0	1.02567	CD09940	
10A5	60.95	0.342313	SEW02803	

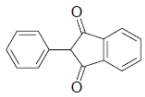
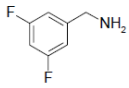
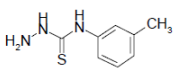
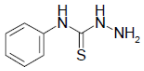
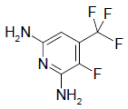
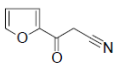
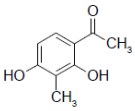
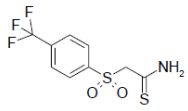
10A6	61.6	0.992313	CC13813	Code CC13813  C7 H17 N3 | 143.23
10A7	60.75	0.142313	BTB00733	Code BTB00733  C10 H13 N | 147.22
10A8	60.9	0.292313	TL00667	Code TL00667  C9 H3 F3 N2 O | 212.13
10C3	60.77	0.162313	AC12118	Code AC12118  C6 H13 N O2 | 131.17
10C4	60.74	0.132313	BTB07528	Code BTB07528  C6 H7 F N2 . Cl H | 162.59
11A9	59.82	0.642582	CD05168	Code CD05168  C7 H12 N2 | 124.19
11B7	60.9	1.722582	BTB08664	Code BTB08664  C6 H6 Cl2 N2 . Cl H | 213.49

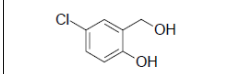
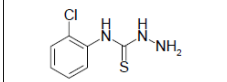
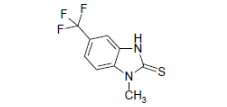
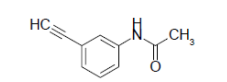
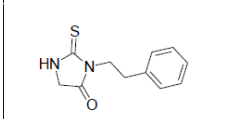
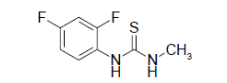
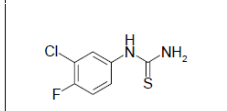
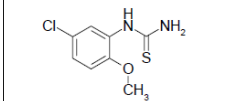
11B8	59.69	0.512582	S13513	Code **S13513**  C7 H5 F3 N2 S | 206.19
11D4	59.59	0.412582	SPB03349	Code **SPB03349**  C7 H7 Cl N2 S | 186.67
11D7	59.19	0.012582	BTB07813	Code **BTB07813**  C11 H8 Cl2 N2 O | 255.10
11G4	61.02	1.842582	KM04038	Code **KM04038**  C10 H10 Cl N3 | 207.66
12A9	59.42	0.19464	AC43267	Code **AC43267**  C8 H10 N2 O2 | 166.18
12C10	60.39	1.16464	AC13022	Code **AC13022**  C12 H9 N O | 183.21
13A10	61.5	0.15282	AC26582	Code **AC26582**  C7 H11 N O | 125.17
13A11	61.66	0.31282	AC27773	Code **AC27773**  C6 H11 N O2 | 129.16

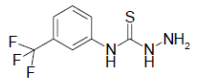
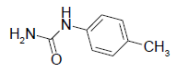
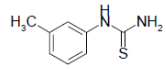
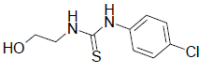
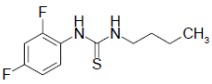
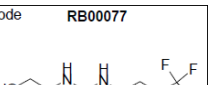
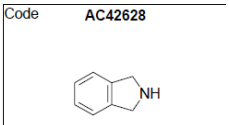
13C10	61.38	0.03282	AC26592	Code AC26592  C8 H13 N O . Cl H 175.66
15A6	60.19	1.100442	CC18413	Code CC18413  C10 H11 N3 173.22
15C5	60.48	1.390442	SB02109	Code SB02109  C6 H6 N2 S 138.19
17A3	59.54	1.036637	TL00150	Code TL00150  C8 H8 F3 N 175.15
18B6	58.6	0.115297	BTB01618	Code BTB01618  C12 H18 N2 190.29
18E4	58.51	0.025297	RF01360	Code RF01360  C6 H4 Cl2 N2 S 207.08
18G7	58.65	0.165297	RDR01464	Code RDR01464  C10 H7 Cl O3 210.61
18G9	58.56	0.075297	JFD03593	Code JFD03593  C5 H7 N O S2 161.25

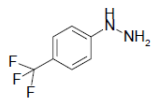
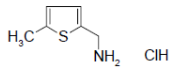
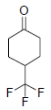
18H8	58.82	0.335297	BTB12099	Code BTB12099  C7 H7 N3 | 133.15
19C4	60.16	0.53783	BTB06198	Code BTB06198  C15 H12 N2 O | 236.27
20C8	60.34	0.923113	SEW04444	Code SEW04444  C7 H9 F3 N2 | 178.16
20G10	59.72	0.303113	BTB06601	Code BTB06601  C9 H6 Cl2 N2 O | 229.06
21A11	59	0.149718	BTB06035	Code BTB06035  C8 H6 F2 O | 156.13
21H7	59.23	0.379718	CD03731	Code CD03731  C11 H15 F N2 S | 226.32
23A11	59.27	0.4125	S01843	Code S01843  C9 H11 Cl N2 S | 214.72
23B4	59.28	0.4225	NRB05254	Code NRB05254  C10 H10 N2 O | 174.20

23C9	59.8	0.9425	KM00122	Code KM00122  C10 H7 F3 N2 O | 228.17
23D10	59.48	0.6225	S05057	Code S05057  C9 H6 F3 N3 S2 | 277.29
23E2	59.44	0.5825	SPB07708	Code SPB07708  C12 H11 N O2 S2 | 265.36
23F2	59.05	0.1925	RF02106	Code RF02106  C15 H20 N2 O | 244.34
23F6	59.19	0.3325	KM09235	Code KM09235  C10 H8 F3 N3 | 227.19
23G8	58.98	0.1225	CD02041	Code CD02041  C7 H6 F2 N2 O | 172.13
24B9	59.33	0.368886	SPB07644	Code SPB07644  C7 H7 N3 | 133.15
24B10	59.06	0.098886	BTB04021	Code BTB04021  C6 H5 Cl2 N O | 178.02

24F5	59.45	0.488886	SB00622	Code SB00622  C15 H10 O2 | 222.24
24H2	59.22	0.258886	AC39479	Code AC39479  C7 H7 F2 N | 143.14
25A4	60.42	0.915845	S00359	Code S00359  C8 H11 N3 S | 181.26
25A5	59.6	0.095845	SPB00239	Code SPB00239  C7 H9 N3 S | 167.24
25A6	60.11	0.605845	MAY00151	Code MAY00151  C6 H5 F4 N3 | 195.12
25E2	59.7	0.195845	CD05195	Code CD05195  C7 H5 N O2 | 135.12
25H4	59.6	0.095845	RH00145	Code RH00145  C9 H10 O3 | 166.18
26B2	58.83	0.037113	SPB02572	Code SPB02572  C9 H8 F3 N O2 S2 | 283.29

26F7	60.63	1.837113	MB00235	Code MB00235  C7 H7 Cl O2 | 158.58
26F11	60.49	1.697113	GK03258	Code GK03258  C7 H8 Cl N3 S | 201.68
26G6	59.28	0.487113	TG00013	Code TG00013  C9 H7 F3 N2 S | 232.23
27B4	59.52	0.548713	RF01853	Code RF01853  C10 H9 N O | 159.19
27C4	59.97	0.998713	RF03152	Code RF03152  C11 H12 N2 O S | 220.29
27D5	59.47	0.498713	RJF00130	Code RJF00130  C8 H8 F2 N2 S | 202.23
27E5	61.4	2.428713	RJF00180	Code RJF00180  C7 H6 Cl F N2 S | 204.65
27E6	60.13	1.158713	S12633	Code S12633  C8 H9 Cl N2 O S | 216.69

27E8	61.29	2.318713	BTB05094	Code BTB05094  C8 H8 F3 N3 S | 235.23
27E9	59.54	0.568713	JA00151	Code JA00151  C8 H10 N2 O | 150.18
27F4	59.1	0.128713	RH01166	Code RH01166  C8 H10 N2 S | 166.25
27G3	59.77	0.798713	RB00016	Code RB00016  C9 H11 Cl N2 O S | 230.72
27G11	59.04	0.068713	RJF01201	Code RJF01201  C11 H14 F2 N2 S | 244.31
27H3	61.05	2.078713	RB00077	Code RB00077  C10 H11 F3 N2 O S | 264.27
28B11	59.87	0.814338	AC42628	Code AC42628  C8 H9 N | 119.17

28C10	61.5	2.444338	AC39795	Code **AC39795**  C7 H7 F3 N2 | 176.14
28D11	59.59	0.534338	AC39765	Code **AC39765**  C8 H6 F N | 135.14
28G7	59.07	0.014338	CC65063	Code **CC65063**  C6 H9 N S . Cl H | 163.67
28G11	59.18	0.124338	BTB10356	Code **BTB10356**  C7 H9 F3 O | 166.14
28H7	58.42	0.094338	MO08501	Code **MO08501**  C4 H2 Cl2 N2 | 148.98
28H9	58.72	0.094338	AC11117	Code **AC11117**  C6 H8 O2 | 112.13

Appendix C. Processing and Refinement Statistics

Processing and Refinement Statistics for D21A-Slp2$_a$-Rab27a fusion

Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(A°) a = 61.71, b = 76.816, c = 117.818
Wavelength (A°)	0.97949
Resolution (A°)	37.26- 2.23 (2.31 - 2.23)
Total reflections	200712
Unique reflections	27129
Multiplicity	8.4(6.6)
Completeness (%)	99.8(99.2)
<I>/<σ(I)>	14.3(2.3)
R-merge	0.094 (0.902)
R-meas	0.109
Wilson B factor	40.13
$CC_{1/2}$	0.992
Refinement	
Reflections used in refinement	27112 (2003)
Reflections used for R_{free}	1341 (98)
R_{work} (%)	20.55
R_{free} (%)	23.37
Rmsd bond lengths (A°)	0.013
Rmsd bond angles (°)	1.50
Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	47.2 /3216
Water molecules	41.19/85
Ligand non-H atoms gppnhp-Mg^{2+}	29.15/66
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	97.2
Ramachandran allowed region (%)	2.8

Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.59

Statistics for the highest-resolution shell are shown in parentheses.

Processing and Refinement Statistics for $Slp2_a$-Rab27a fusion in complex with 4-fluoro indole

Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(A°) a = 61.95, b = 77.15, c = 119.52
Wavelength (A°)	0.97949
Resolution (A°)	47.24 - 2.09 (2.165 - 2.09)
Total reflections	250264
Unique reflections	34636
Multiplicity	9.3(7.5)
Completeness (%)	99.9(99.6)
<I>/<σ(I)>	14.5(3.0)
R-merge	0.078 (0.81)
R-meas	0.092
Wilson B factor	32.21
$CC_{1/2}$	0.996
Refinement	
Reflections used in refinement	34635 (3397)
Reflections used for R_{free}	1700 (183)
R_{work} (%)	20.06
R_{free} (%)	22.95
Rmsd bond lengths (A°)	0.014
Rmsd bond angles (°)	1.44
Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	34.55 /3280
Water molecules	33.56/241
Ligand non-H atoms gppnhp-Mg^{2+}	27.17/76
Φ/Ψ angles (%)	

Ramachandran Most favored region (%)	98.3
Ramachandran allowed region (%)	1.7
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.28

Statistics for the highest-resolution shell are shown in parentheses.

Processing and Refinement Statistics for $Slp2_a$-Rab27a fusion in complex with 5-chloro-2-hydroxybenzyl-alcohol

Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(A°) a = 61.48, b = 76.73, c = 119.31
Wavelength (A°)	0.97949
Resolution (A°)	58.85 - 2.32 (2.403 - 2.32)
Total reflections	250201
Unique reflections	34504 (3359)
Multiplicity	10.8(8.9)
Completeness (%)	99.7(99.3)
<I>/<σ(I)>	15.5(3.2)
R-merge	0.082 (0.77)
R-meas	0.096
Wilson B factor	36.39
$CC_{1/2}$	0.999
Refinement	
Reflections used in refinement	34504 (3359)
Reflections used for R_{free}	1698 (186)
R_{work} (%)	20.76
R_{free} (%)	23.1
Rmsd bond lengths (A°)	0.014
Rmsd bond angles (°)	1.44
Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	34.93 /3280
Water molecules	34.35/192

Ligand non-H atoms gppnhp-Mg^{2+}	26.62/86
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	98.5
Ramachandran allowed region (%)	1.5
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.0

Statistics for the highest-resolution shell are shown in parentheses.

Processing and Refinement Statistics for $Slp2_a$-Rab27a fusion in complex with 4-[3-trifluoromethyl-phenyl]-3-thiosemicarbazide

Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(A°) a = 62.81, b = 77.92, c = 119.87
Wavelength (A°)	0.97949
Resolution (A°)	48.9 - 2.02 (2.02 - 2.09)
Total reflections	277277
Unique reflections	39118
Multiplicity	7.1 (5.3)
Completeness (%)	100.0(99.7)
<I>/<σ(I)>	13.8(1.5)
R-merge	0.081 (0.86)
R-meas	0.094
Wilson B factor	36.47
$CC_{1/2}$	0.9995
Refinement	
Reflections used in refinement	39116 (3692)
Reflections used for R_{free}	1892 (202)
R_{work} (%)	21.11
R_{free} (%)	24.28
Rmsd bond lengths (A°)	0.013
Rmsd bond angles (°)	1.44
Average B factors ($A°^2$)/Number of atoms	
Protein non-H atoms	

Macromolecules	38.3 /3286
Water molecules	35.28/211
Ligand non-H atoms gppnhp-Mg^{2+}	33.81/96
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	99.0
Ramachandran allowed region (%)	1.0
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.58

Statistics for the highest-resolution shell are shown in parentheses.

Processing and Refinement Statistics for $Slp2_a$-Rab27a fusion in complex with 3-4-chlorophenyl-1-methyl-1H-pyrazol-5-amine

Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(A°) a = 61.3, b = 76.37, c = 119.96
Wavelength (A°)	0.97949
Resolution (A°)	28.44 - 2.52 (2.61 - 2.52)
Total reflections	231467
Unique reflections	19395
Multiplicity	7.2(6.9)
Completeness (%)	98.6(96.7)
<I>/<σ(I)>	9.4(2.0)
R-merge	0.112 (0.864)
R-meas	0.122
Wilson B factor	44.83
$CC_{1/2}$	0.999
Refinement	
Reflections used in refinement	19395 (1908)
Reflections used for R_{free}	970 (87)
R_{work} (%)	21.12
R_{free} (%)	24.14
Rmsd bond lengths (A°)	0.015
Rmsd bond angles (°)	1.48

Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	44.23 /3245
Water molecules	42.12/175
Ligand non-H atoms gppnhp-Mg^{2+}	30.45/80
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	98.2
Ramachandran allowed region (%)	1.8
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.26

Statistics for the highest-resolution shell are shown in parentheses.

Processing and Refinement Statistics for $Slp2_a$-Rab27a fusion in complex with 2-amino-4-3-4-dipluorophenyl-thiazole

Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(A°) a = 61.41, b = 76.95, c = 119.75
Wavelength (A°)	0.97949
Resolution (A°)	36.63 - 2.39 (2.475 - 2.39)
Total reflections	234511
Unique reflections	23050
Multiplicity	10.8(8.9)
Completeness (%)	100.0(99.7)
<I>/<σ(I)>	8.7(2.5)
R-merge	0.085 (0.81)
R-meas	0.087
Wilson B factor	40.32
$CC_{1/2}$	0.999
Refinement	
Reflections used in refinement	23049 (2265)
Reflections used for R_{free}	1139 (113)

R_{work} (%)	19.45
R_{free} (%)	22.32
Rmsd bond lengths (A°)	0.013
Rmsd bond angles (°)	1.46
Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	40.1 /3270
Water molecules	38.27/181
Ligand non-H atoms gppnhp-Mg^{2+}	29.26/98
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	98.2
Ramachandran allowed region (%)	1.8
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.87

Statistics for the highest-resolution shell are shown in parentheses.

Processing and Refinement Statistics for $\Delta Slp2_a$-Rab27a fusion

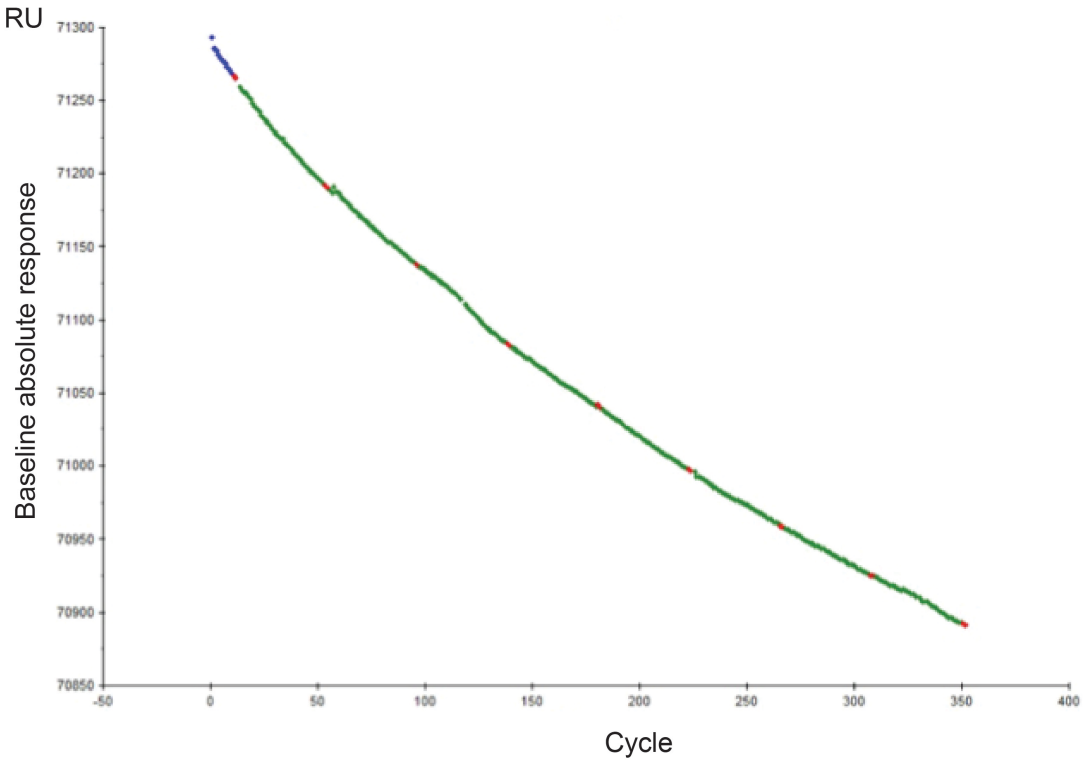
Data Collection	
Space group	$P2_12_12_1$
Unit cell parameters	(A°) a = 64.75, b = 67.89, c = 118.62
Wavelength (A°)	0.97949
Resolution (A°)	58.92 - 2.3 (2.382 - 2.3)
Total reflections	253324
Unique reflections	23637
Multiplicity	10.1(7.3)
Completeness (%)	99.0(98.1)
<I>/<σ(I)>	9.1(2.5)
R-merge	0.088 (0.90)
R-meas	0.101
Wilson B factor	41.85
$CC_{1/2}$	0.992
Refinement	

Reflections used in refinement	23635 (2120)
Reflections used for R_{free}	1124 (84)
R_{work} (%)	22.40
R_{free} (%)	25.52
Rmsd bond lengths (A°)	0.014
Rmsd bond angles (°)	1.51
Average B factors ($A^{\circ 2}$)/Number of atoms	
Protein non-H atoms	
Macromolecules	41.52 /3088
Water molecules	34.77/192
Ligand non-H atoms gppnhp-Mg^{2+}	29.59/66
Φ/Ψ angles (%)	
Ramachandran Most favored region (%)	97.0
Ramachandran allowed region (%)	3.0
Ramachandran outliers (%)	0.0
Rotamer outliers (%)	0.94

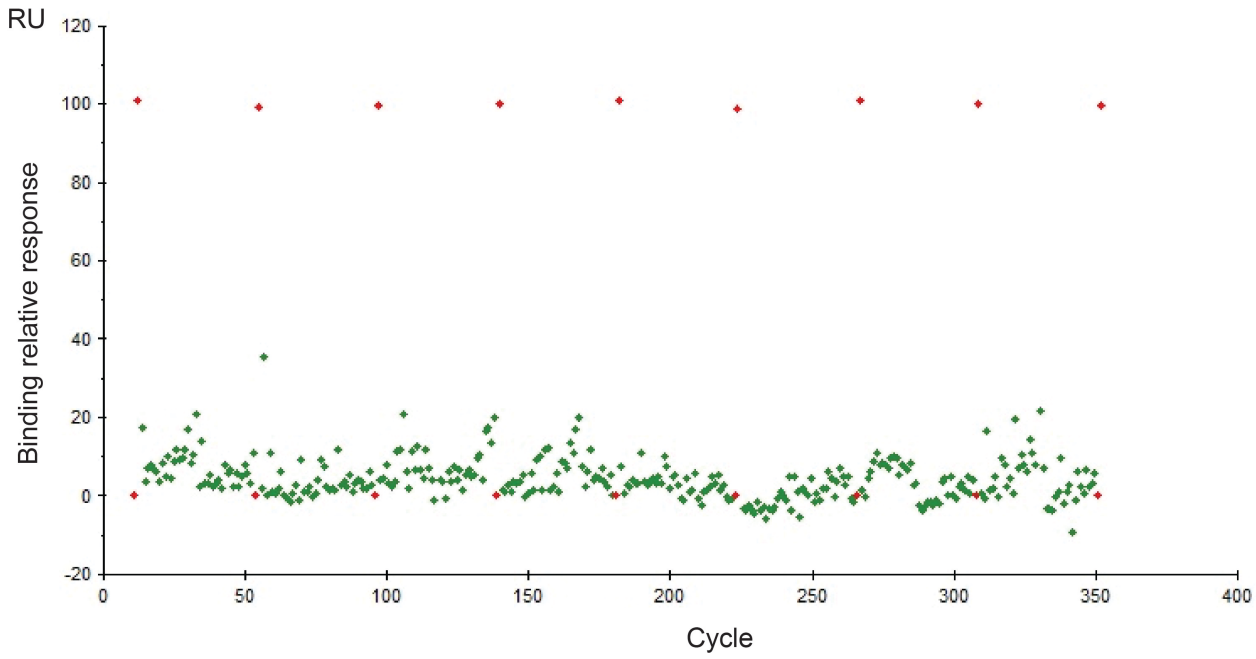
Statistics for the highest-resolution shell are shown in parentheses.

Appendix D. Assessing the quality of an SPR run. A) The baseline absolute response decreases with a RU of approximately 1 unit per cycle. Demonstrating that the protein does not wash of the chip and is stable over the complete run. B) Red dots indicate the response obtained from the WF dipeptide and buffer controls. As shown, the controls give consistent signal through out screen. C) The standard curve used for DMSO correction. As demonstrated, the curve encompasses a response range suitable for our experiments (the range between the two brown lines is our experimental responses). This curve is used to correct for the variation in the DMSO effect on refractive index.

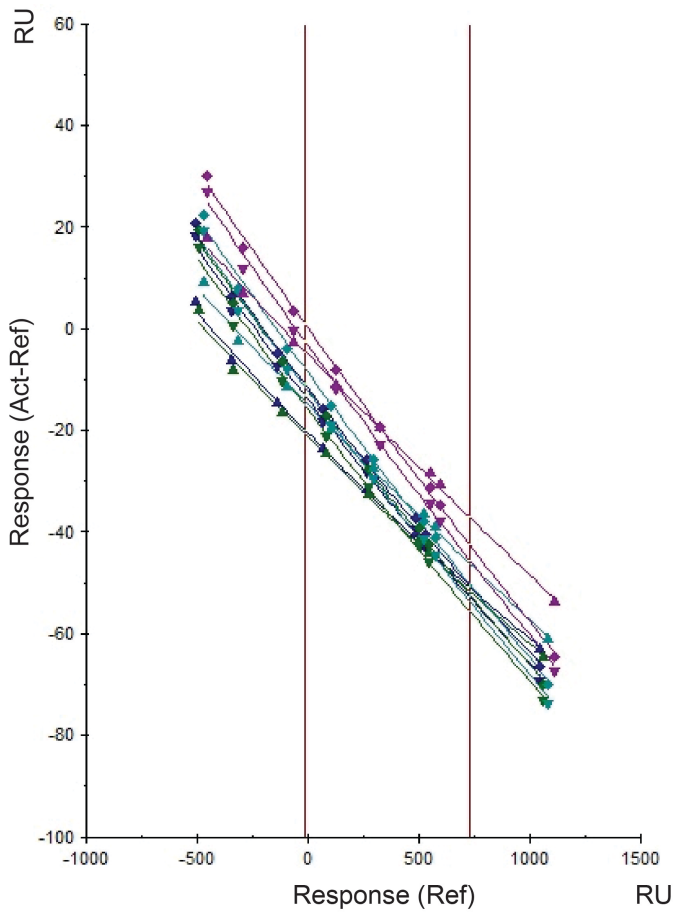
A



B


RU
120
100
80
60
40
20
0
-20
Binding relative response
0
50
100
150
200
250
300
350
400
Cycle

C


RU
60
40
20
0
-20
-40
-60
-80
-100
Response (Act-Ref)
-1000
-500
0
500
1000
1500
Response (Ref)
RU

Appendix E. Hits obtained from chemical library designed using shape-based similarity against Rab27a using SPR. These data display ligands with a response above the average + 2*standard deviation threshold. Yes and No in the first column refer to compounds taken for further validation.

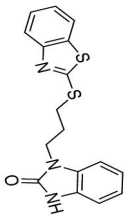
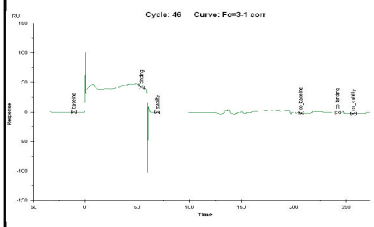
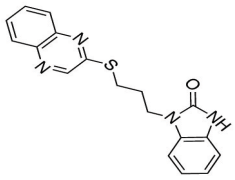
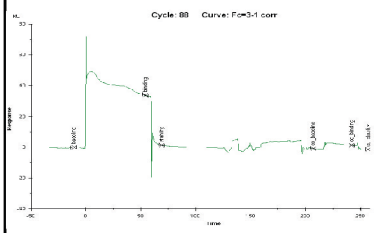
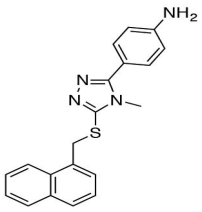
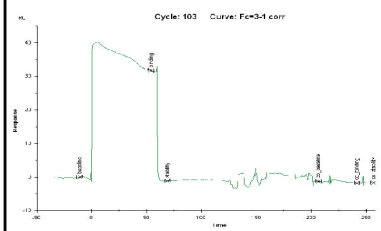
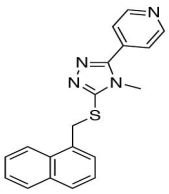
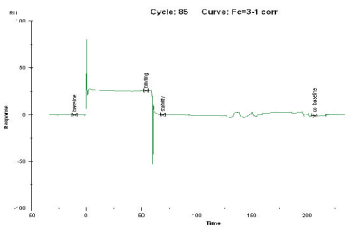
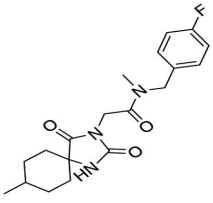
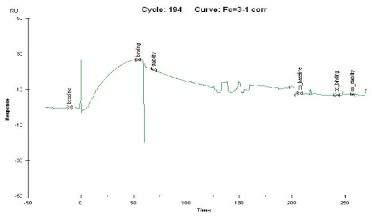
Plate	Code	Structure	RU	Sensorgram
1 **Yes**	CCT303643		57.5	
1 **Yes**	CCT332162		46.3	
1 **Yes**	CCT307514		42.0	
1 **Yes**	CCT326123		35.5	
1 **No**	CCT270787		33.3	

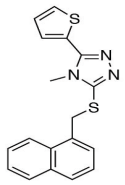
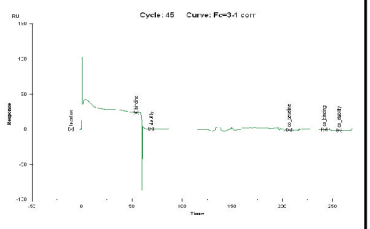
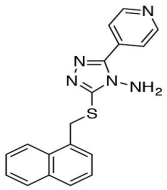
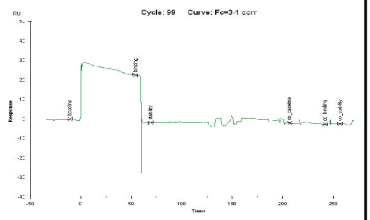
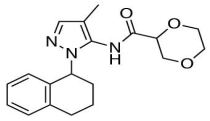
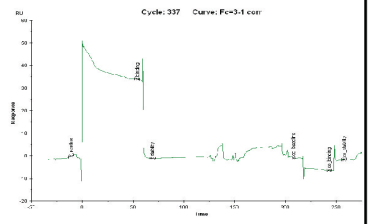
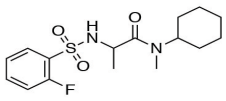
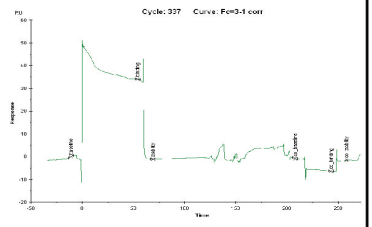
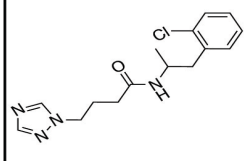
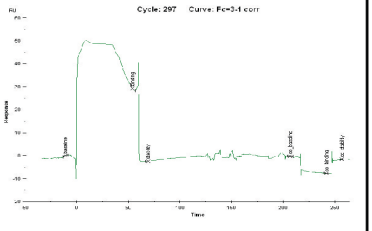
Plate	Code	Structure	RU	Sensorgram
1 **Yes**	CCT326052		33.2	
1 **Yes**	CCT326210		31.0	
2 **Yes**	CCT261567		42.9	
2 **Yes**	CCT268580		40.4	
2 **Yes**	CCT260742		38.2	

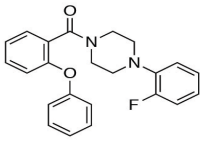
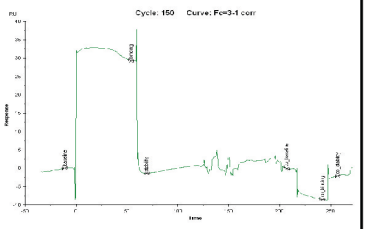
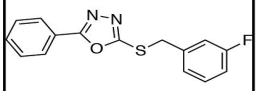
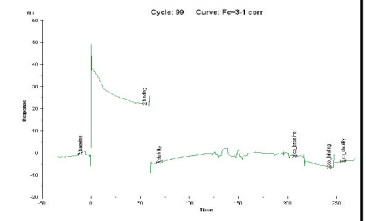
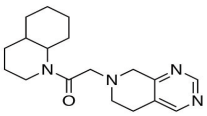
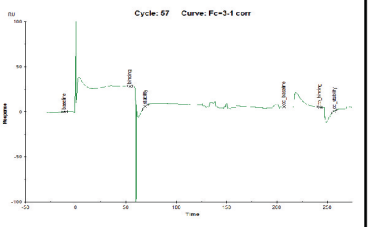
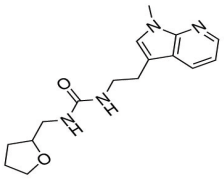
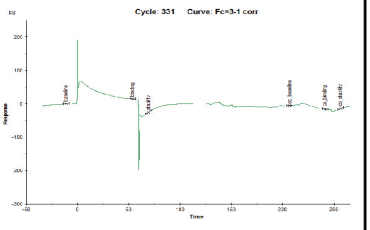
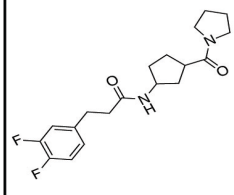
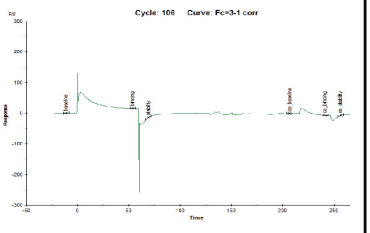
Plate	Code	Structure	RU	Sensorgram
2 **Yes**	CCT304167		35.9	
2 **Yes**	CCT289878		34.2	
3 **Yes**	CCT281241		33.6	
3 **Yes**	CCT352933		22.3	
3 **Yes**	CCT339292		21.3	

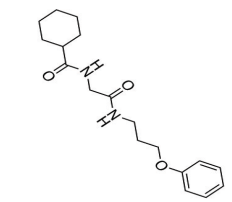
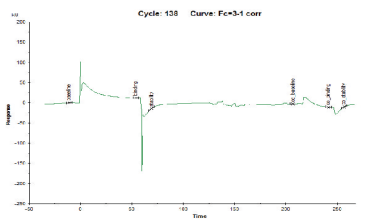
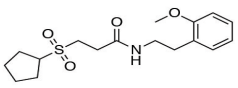
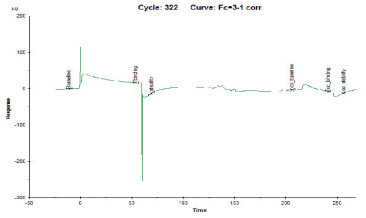
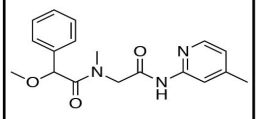
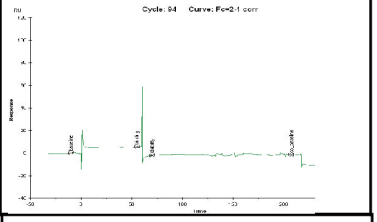
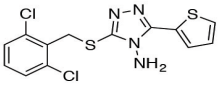
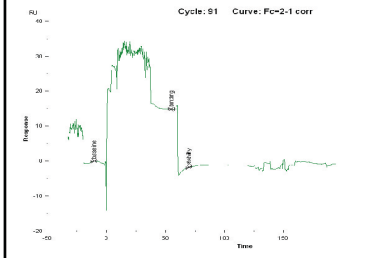
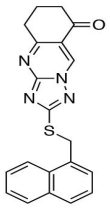
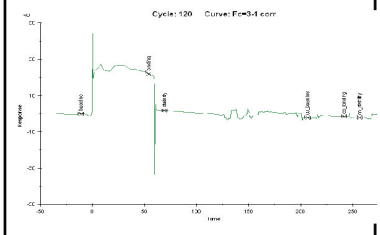
Plate	Code	Structure	RU	Sensorgram
3 **No**	CCT340691		20.0	
3 **Yes**	CCT347484		19.7	
4 **No**	CCT347823		61.0	
4 **Yes**	CCT290035		19.7	
1 **Yes**	CCT328985		27.8	

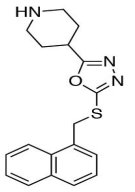
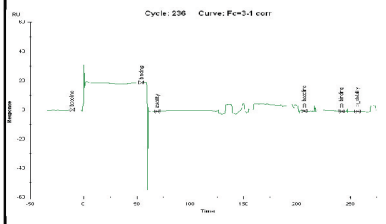
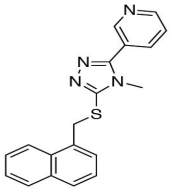
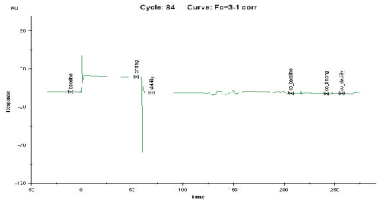
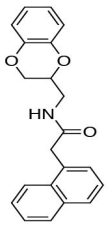
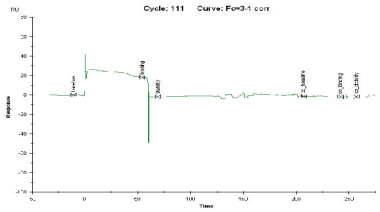
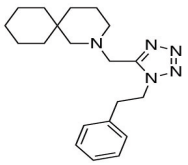
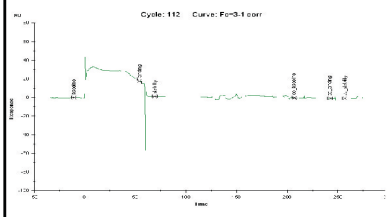
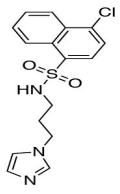
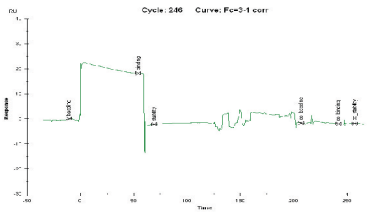
Plate	Code	Structure	RU	Sensorgram
1 **Yes**	CCT159884		26.2	
1 **Yes**	CCT326106		27.6	
1 **Yes**	CCT303093		24.9	
1 **Yes**	CCT226146		23.9	
1 **Yes**	CCT159894		23.0	

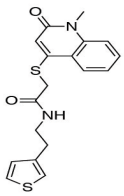
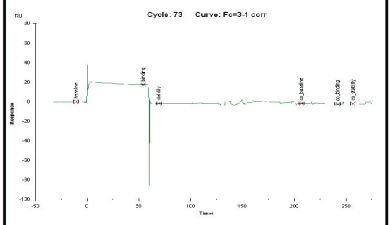
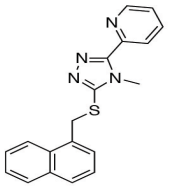
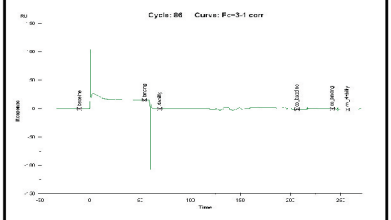
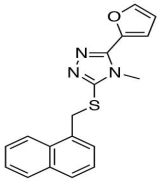
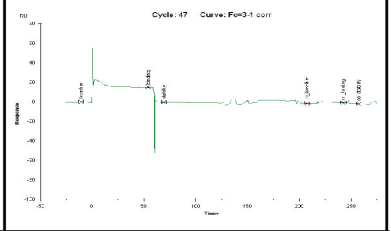
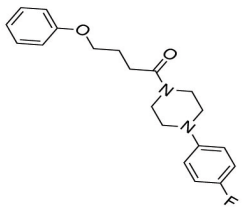
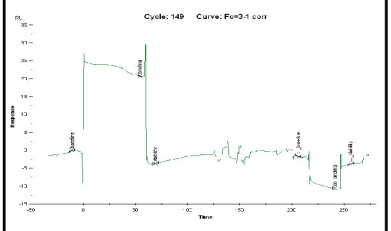
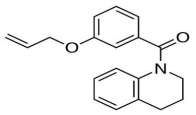
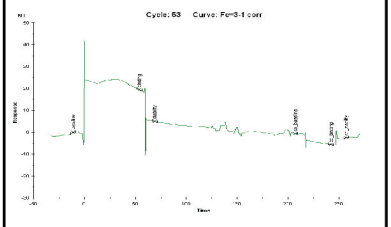
Plate	Code	Structure	RU	Sensorgram
1 **Yes**	CCT331002		23.0	
1 **No**	CCT326085		20.9	
1 **No**	CCT325785		21.8	
2 **Yes**	CCT304161		29.1	
2 **Yes**	CCT219413		29.0	

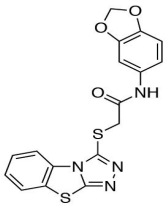
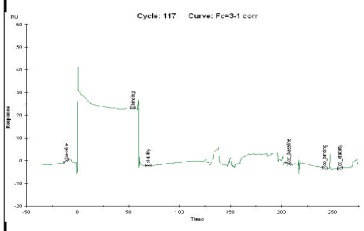
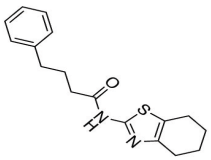
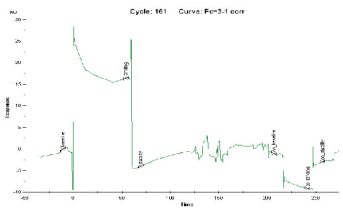
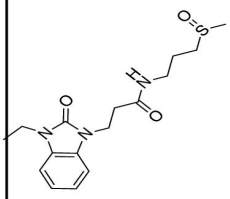
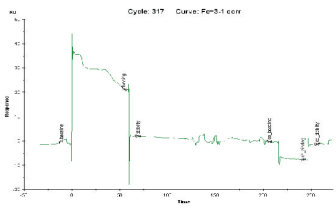
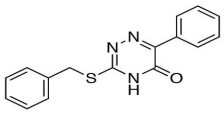
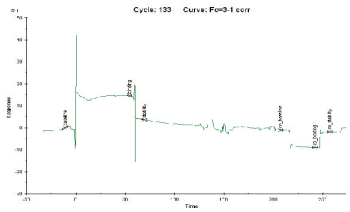
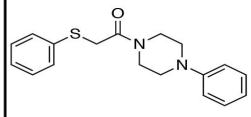
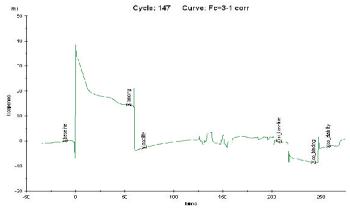
Plate	Code	Structure	RU	Sensorgram
2 **Yes**	CCT294941		26.4	
2 **Yes**	CCT306279		26.2	
2 **No**	CCT262910		25.2	
2 **Yes**	CCT300727		25.0	
2 **No**	CCT302652		23.5	

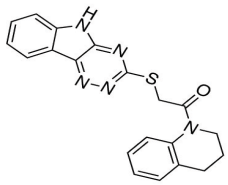
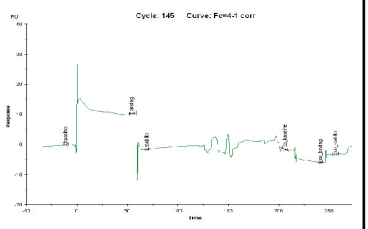
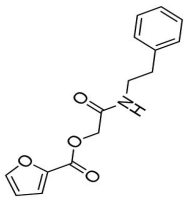
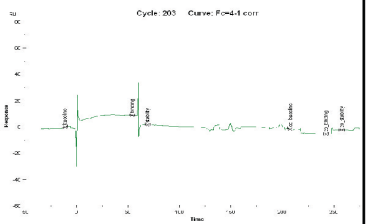
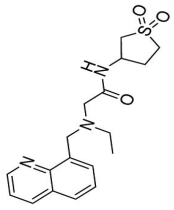
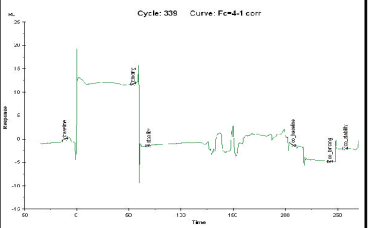
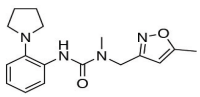
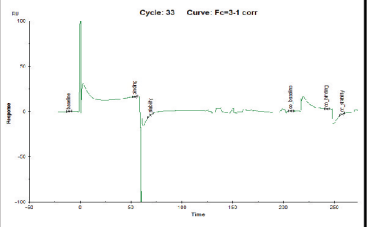
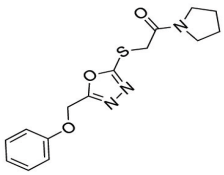
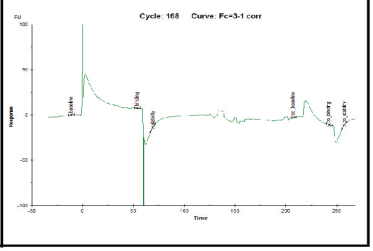
Plate	Code	Structure	RU	Sensorgram
2 **Yes**	CCT301984		18.6	
2 **Yes**	CCT316794		21.5	
2 **Yes**	CCT268947		17.6	
3 **Yes**	CCT279386		19.2	
3 **Yes**	CCT150291		18.4	

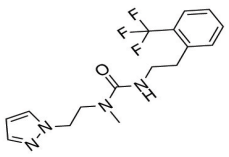
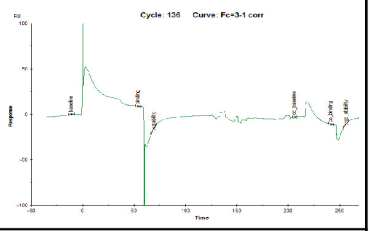
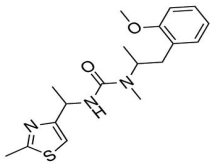
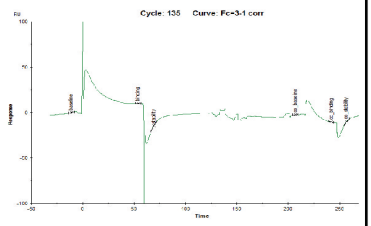
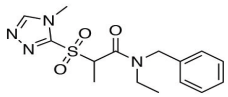
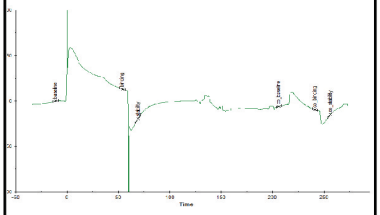
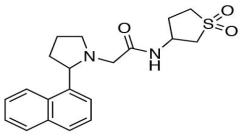
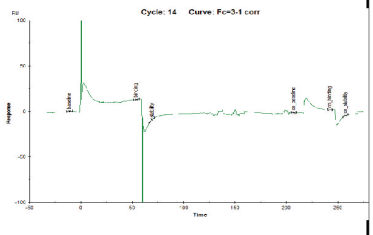
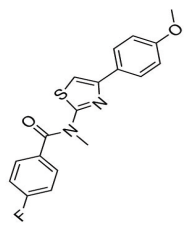
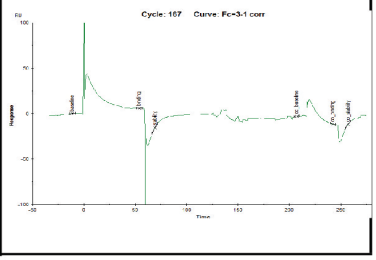
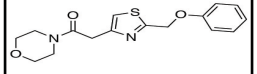
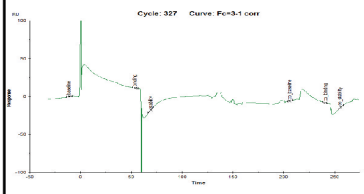
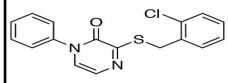
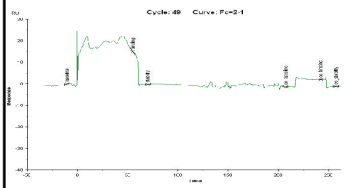
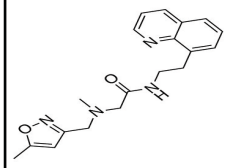
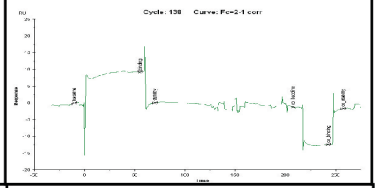
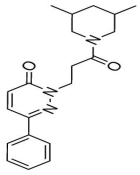
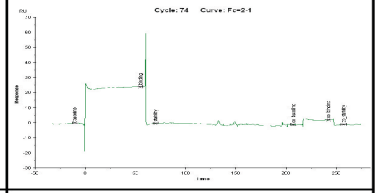
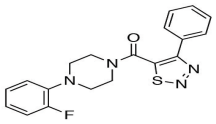
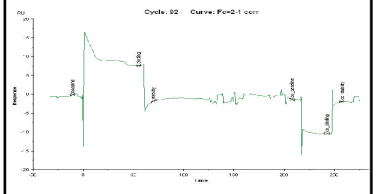
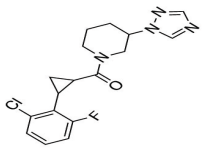
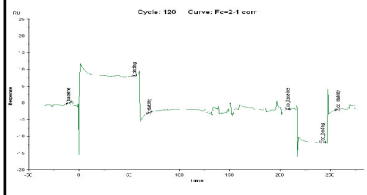
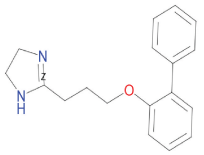
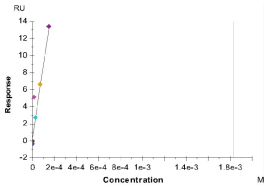
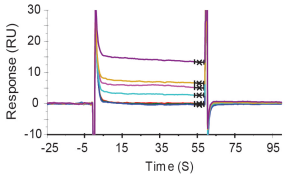
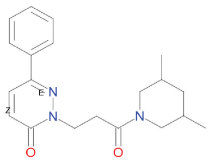
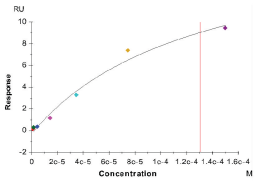
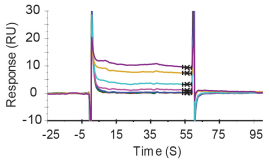
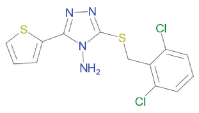
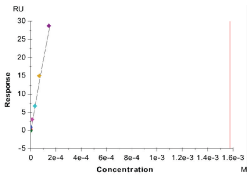
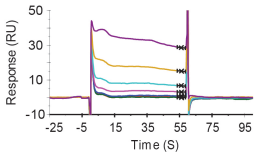
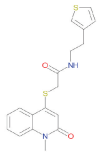
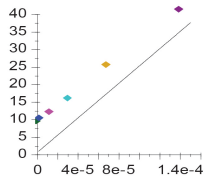
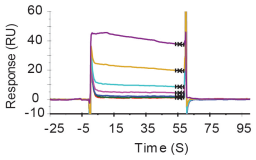
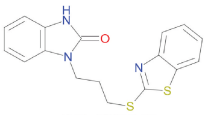
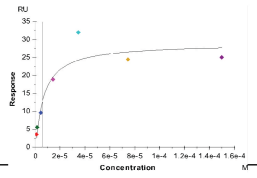
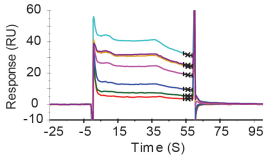
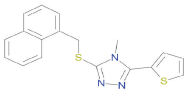
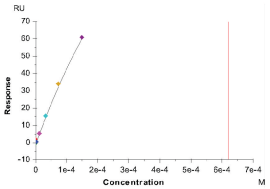
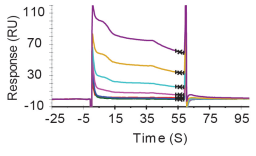
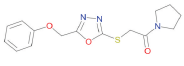
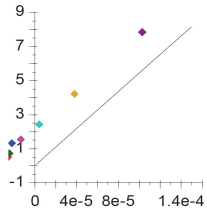
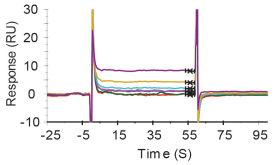
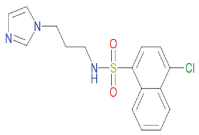
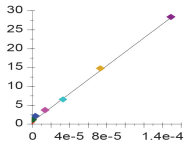
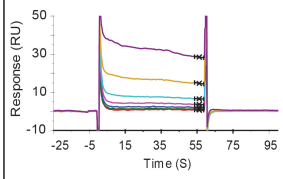
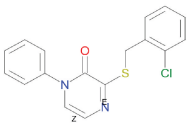
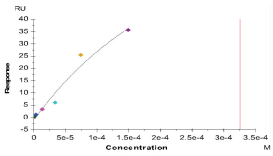
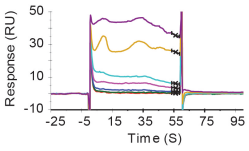
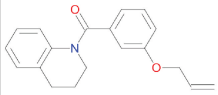
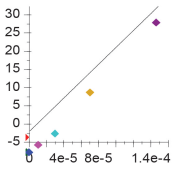
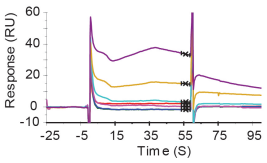
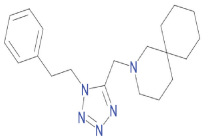
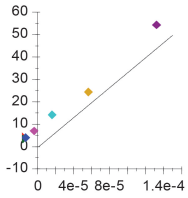
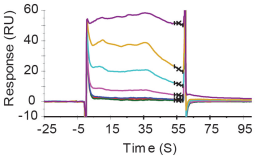
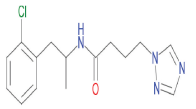
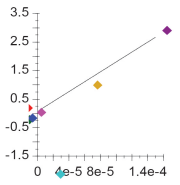
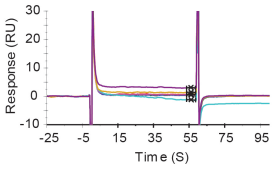
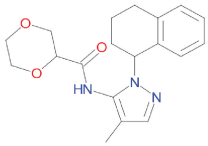
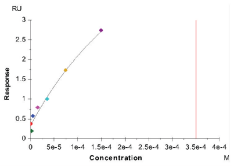
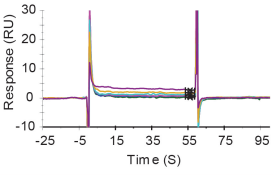
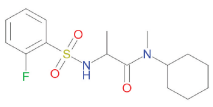
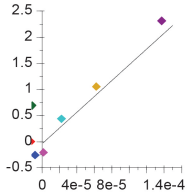
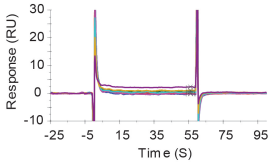
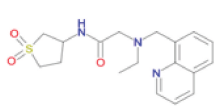
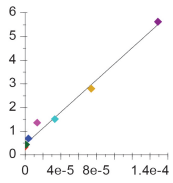
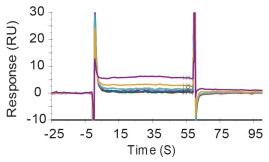
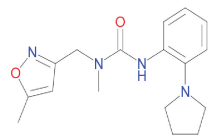
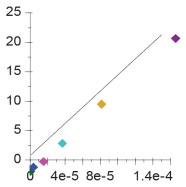
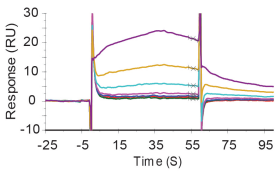
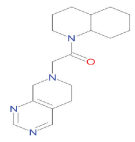
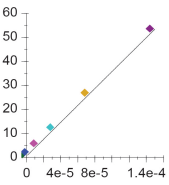
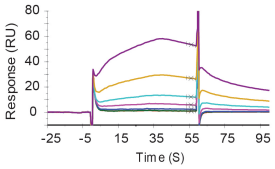
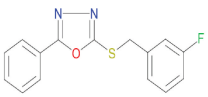
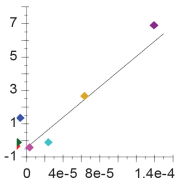
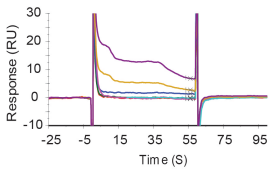
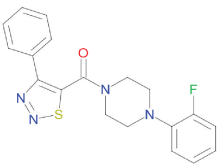
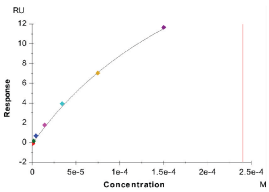
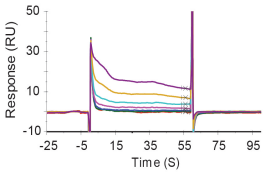
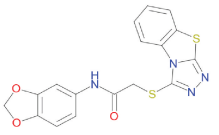
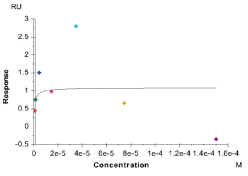
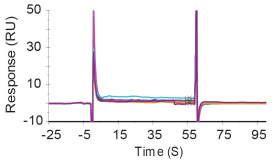
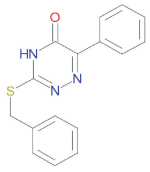
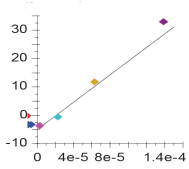
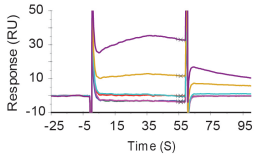
Plate	Code	Structure	RU	Sensorgram
3 **No**	CCT340500		17.6	
3 **No**	CCT340426		16.9	
3 **No**	CCT346341		16.6	
3 **Yes**	CCT270661		16.0	
3 **No**	CCT150153		15.4	

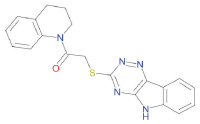
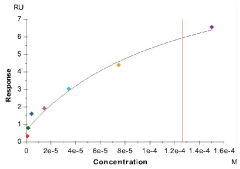
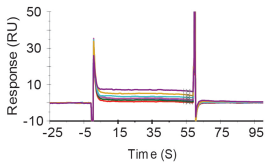
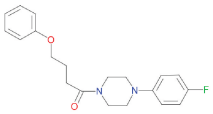
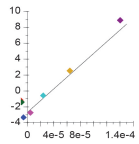
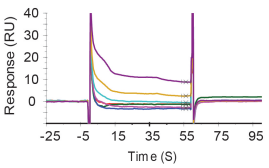
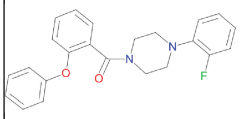
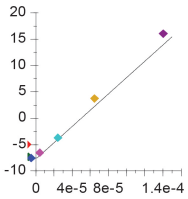
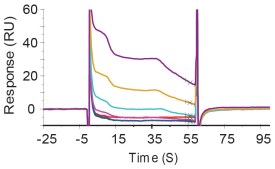
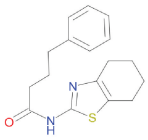
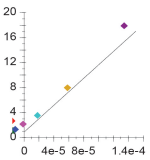
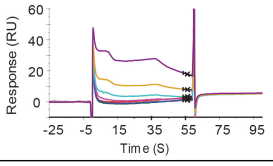
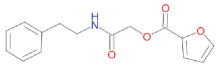
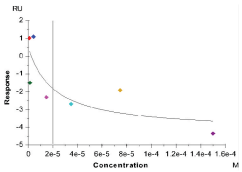
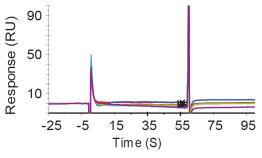
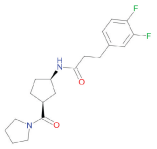
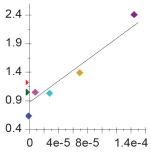
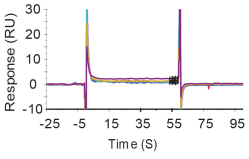
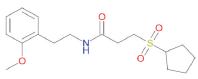
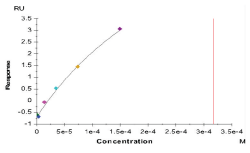
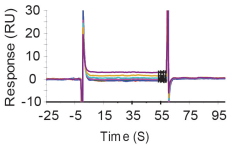
Plate	Code	Structure	RU	Sensorgram
3 **Yes**	CCT352530		14.8	
4 **Yes**	CCT188542		40.5	
4 **Yes**	CCT360813		34.4	
4 **Yes**	CCT192639		29.2	
4 **Yes**	CCT290039		9.3	
4 **Yes**	CCT349808		11.8	

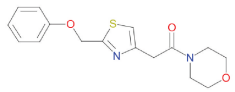
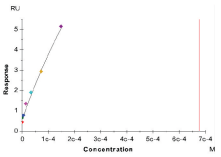
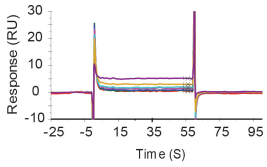
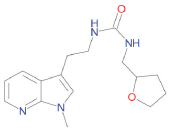
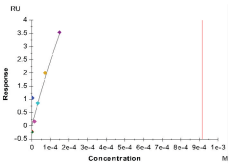
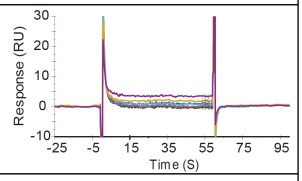
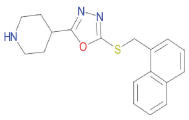
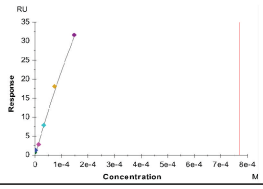
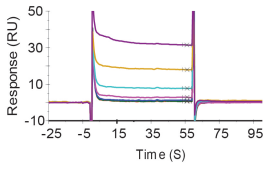
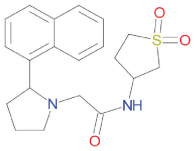
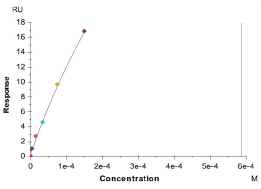
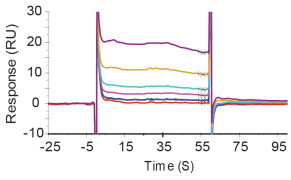
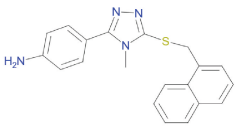
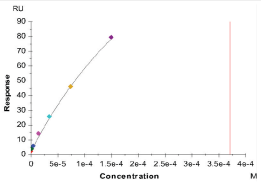
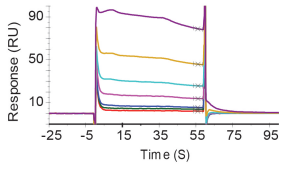
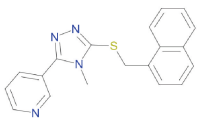
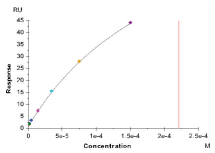
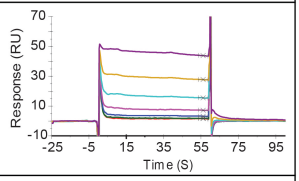
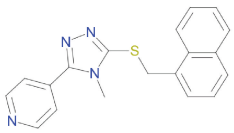
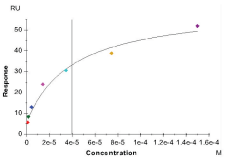
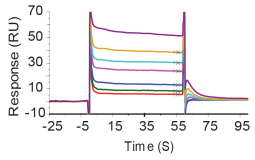
Appendix F. Validating the binding of ligands to Rab27a using SPR. As demonstrated, the binding of each hit to Rab27a is measured by titrating the ligands using method described in section 2.9.2. Compounds with good quality sensorgrams and measurable affinity were taken for further studies.

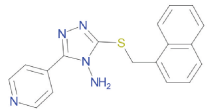
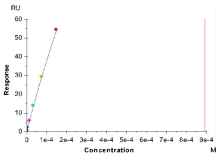
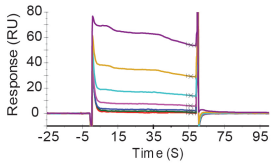
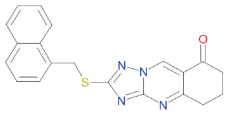
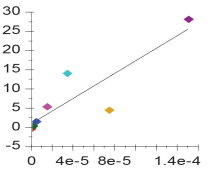
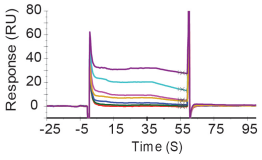
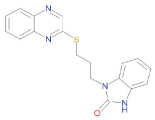
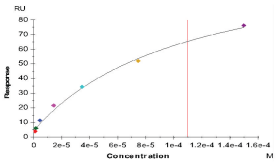
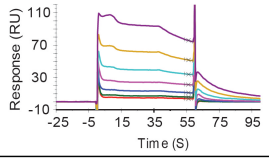
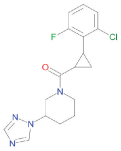
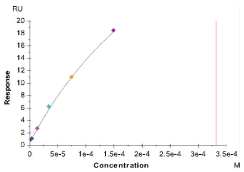
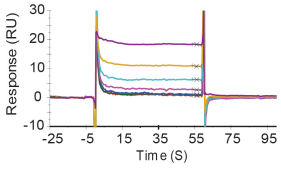
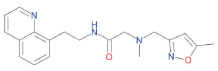
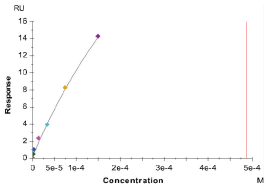
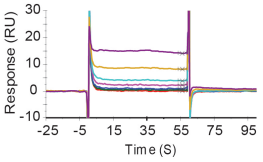
STRUCTURE	KD (μM)	ExpRmax/ ThrRmax	Langmuir Curve	Sensorgram
 CCT145783	1830	1.46		
 CCT192639	131	0.13		
 CCT290035	1580	2.23		
 CCT331002	>10000	1775		
 CCT303643	6	0.23		
 CCT326052	619	2.25		
 CCT150291	>10000	20.16		

 CCT159894	>10000	2643		
 CCT188542	326	0.87		
 CCT219413	>10000	80327		
 CCT226146	>10000	46030		
 CCT260742	>10000	24275		
 CCT261567	350	0.06		
 CCT268580	>10000	4324		

 CCT268947	>10000	3136		
 CCT279386	>10000	2256		
 CCT281241	>10000	5827		
 CCT289878	>10000	33844		
 CCT290039	240	0.2		
 CCT294941	1	0		
 CCT300727	>10000	94327		

 CCT301984	130	0.07		
 CCT304161	>10000	29169		
 CCT304167	>10000	31681		
 CCT306279	>10000	32132		
 CCT316794	20	-0.04		
 CCT339292	>10000	5678		
 CCT347484	317	0.08		

 CCT352530	675	0.19		
 CCT352933	914	0.2		
 CCT159884	769	1.44		
 CCT270661	586	0.52		
 CCT307514	371	1.83		
 CCT326106	221	0.77		
 CCT326123	39	0.4		

 CCT326210	890	2.71		
 CCT328985	>10000	74414		
 CCT332162	110	0.87		
 CCT349808	332	0.4		
 CCT360813	487	0.42		

Appendix G. License for figure 3.1 from Springer

SPRINGER LICENSE
TERMS AND CONDITIONS
Apr 10, 2017

This Agreement between mostafa jamshidiha ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

License Number	4052520037615
License date	
Licensed Content Publisher	Springer
Licensed Content Publication	Springer eBook
Licensed Content Title	Targeting Protein–Protein Interactions and Fragment-Based Drug Discovery
Licensed Content Author	Eugene Valkov
Licensed Content Date	Jan 1, 2011
Type of Use	Thesis/Dissertation
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	1
Author of this Springer article	No
Order reference number	
Original figure numbers	figure 6.5
Title of your thesis / dissertation	Designing structural biology tools for studying Rab27a
Expected completion date	May 2017
Estimated size(pages)	170
Requestor Location	mostafa jamshidiha imeprial college london london, SW7 2AZ United Kingdom Attn: mostafa jamshidiha
Billing Type	Invoice
Billing Address	mostafa jamshidiha imeprial college london london, United Kingdom SW7 2AZ Attn: mostafa jamshidiha
Total	0.00 USD

Terms and Conditions

Introduction
The publisher for this copyrighted material is Springer. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License
With reference to your request to reuse material on which Springer controls the copyright, permission is granted for the use indicated in your enquiry under the following conditions:

- Licenses are for one-time use only with a maximum distribution equal to the number stated in your request.
- Springer material represents original material which does not carry references to other sources. If the material in question appears with a credit to another source, this permission is not valid and authorization has to be obtained from the original copyright holder.
- This permission
• is non-exclusive
• is only valid if no personal rights, trademarks, or competitive products are infringed.
• explicitly excludes the right for derivatives.
- Springer does not supply original artwork or content.
- According to the format which you have selected, the following conditions apply accordingly:
• Print and Electronic: This License include use in electronic form provided it is password protected, on intranet, or CD-Rom/DVD or E-book/E-journal. It may not be republished in electronic open access.
• Print: This License excludes use in electronic form.
• Electronic: This License only pertains to use in electronic form provided it is password protected, on intranet, or CD-Rom/DVD or E-book/E-journal. It may not be republished in electronic open access.
For any electronic use not mentioned, please contact Springer at permissions.springer@spi-global.com.
- Although Springer controls the copyright to the material and is entitled to negotiate on rights, this license is only valid subject to courtesy information to the author (address is given in the article/chapter).
- If you are an STM Signatory or your work will be published by an STM Signatory and you are

requesting to reuse figures/tables/illustrations or single text extracts, permission is granted according to STM Permissions Guidelines: http://www.stm-assoc.org/permissions-guidelines/
For any electronic use not mentioned in the Guidelines, please contact Springer at permissions.springer@spi-global.com. If you request to reuse more content than stipulated in the STM Permissions Guidelines, you will be charged a permission fee for the excess content.
Permission is valid upon payment of the fee as indicated in the licensing process. If permission is granted free of charge on this occasion, that does not prejudice any rights we might have to charge for reproduction of our copyrighted material in the future.
-If your request is for reuse in a Thesis, permission is granted free of charge under the following conditions:
This license is valid for one-time use only for the purpose of defending your thesis and with a maximum of 100 extra copies in paper. If the thesis is going to be published, permission needs to be reobtained.
- includes use in an electronic form, provided it is an author-created version of the thesis on his/her own website and his/her university's repository, including UMI (according to the definition on the Sherpa website: http://www.sherpa.ac.uk/romeo/);
- is subject to courtesy information to the co-author or corresponding author.
Geographic Rights: Scope
Licenses may be exercised anywhere in the world.
Altering/Modifying Material: Not Permitted
Figures, tables, and illustrations may be altered minimally to serve your work. You may not alter or modify text in any manner. Abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of the author(s).
Reservation of Rights
Springer reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction and (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
License Contingent on Payment
While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Springer or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received by the date due, then any license preliminarily granted shall be deemed automatically revoked

and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Springer reserves the right to take any and all action to protect its copyright in the materials.

Copyright Notice: Disclaimer

You must include the following copyright and permission notice in connection with any reproduction of the licensed material: "Springer book/journal title, chapter/article title, volume, year of publication, page, name(s) of author(s), (original copyright notice as given in the publication in which the material was originally published) "With permission of Springer" In case of use of a graph or illustration, the caption of the graph or illustration must be included, as it is indicated in the original publication.

Warranties: None

Springer makes no representations or warranties with respect to the licensed material and adopts on its own behalf the limitations and disclaimers established by CCC on its behalf in its Billing and Payment terms and conditions for this licensing transaction.

Indemnity

You hereby indemnify and agree to hold harmless Springer and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

No Transfer of License

This license is personal to you and may not be sublicensed, assigned, or transferred by you without Springer's written permission.

No Amendment Except in Writing

This license may not be amended except in a writing signed by both parties (or, in the case of Springer, by CCC on Springer's behalf).

Objection to Contrary Terms

Springer hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and Springer (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing

and Payment terms and conditions, these terms and conditions shall control.
Jurisdiction
All disputes that may arise in connection with this present License, or the breach thereof, shall be settled exclusively by arbitration, to be held in the Federal Republic of Germany, in accordance with German law.
Other conditions:
V 12AUG2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.