Understanding the mechanism of regulation of a small Ras-like GTPase, MglA, driving *Myxococcus xanthus* motility

A thesis

Submitted in partial fulfilment of the requirements

Of the degree of

Doctor of Philosophy

By

Sukanya Chakraborty

20162012



INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Dedicated to Thammu, Ma and Kakai

Certificate

Certified that the work incorporated in the thesis entitled " **Understanding the mechanism of regulation of a small Ras-like GTPase, MglA, driving *Myxococcus xanthus* motility** " submitted by **Sukanya Chakraborty** was carried out by the candidate, under my supervision.

The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Gayathri

Dr. Gayathri Pananghat

Date: 06.03.2023

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas have been included; I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.



Sukanya Chakraborty

20162012

Date: 06.03.2023

Acknowledgements

I would like to express my heartfelt gratitude to my supervisor, Dr. Gayathri Pananghat, for all the time, guidance and support she provided me throughout my PhD journey. She has been a constant source of inspiration through her persistence, vast experience and in-depth knowledge of the subject. I would like to thank her for teaching me how to critique scientific results and experimentation. I am also deeply indebted to her for helping me persevere through my personal life issues.

I would like to sincerely thank my advisory committee members, Prof. Jayant Udgaonkar, Prof. Ruchi Anand and Dr. Nishad Matange, for their valuable time and inputs through the yearly meetings. I would also like to thank Prof. Saikrishnan Kayarat, Dr. Radha Chauhan and Dr. Janesh Kumar for the constructive discussions in the group meetings. I would like to thank Prof. Lotte Sogaard Andersen from MPI, Marburg, for providing me with the RomY construct and finding time for several informative discussions. Special thanks to Dr. Krishan Gopal from IMTech Chandigarh for initiating the collaborative work on MglC. Thanks also go to Dr. Tâm Mignot for providing valuable insights on this work.

None of this work would have been possible without the constant help from my lab mates and seniors. I would like to especially thank Dr. Jyoti Baranwal and Manil Kanade for introducing me to this project and the associated techniques. I would like to thank Jazleena, Puja, Ranjana, Akshata, Kushan, Tushar, Pallavi and Dilsha for their assistance at various stages of my work. In addition, I am grateful to all the past and present lab mates, especially Shrikant, Shekhar, Jaya, Mrinmayee, Joyeeta, Vani, Apurba, Soumyajit, Puja, Rinku, Saurabh and Vaishnavi for making working in the lab a wonderful experience. I would like to thank all the past and present members of Prof. Saikrishnan's lab, especially Neha, Ishtiyaq, Mahesh, Vishal, Sujata, Pratima, Om, Vinayak, Sanket, Rushik, Akhilesh, Akash, Ufaq, Nikki, Ashwin and Geetanjali for all their help and support. I would like to thank Suman, Preeti and Rupam from Prof. Jayant's lab for their help with mass spectrometry and CD spectroscopy. I would like to thank Shikha and Susovan from Dr. Sudha Rajamani's lab for their help with HPLC experiments. I would also like to thank Yashwant, Rupali and Yamini from Dr. Amrita Hazra's lab for prompt help with reagents.

I would like to thank Sangeeta and Jyotsna from Dr. Radha Chauhan's lab in NCCS for their help with SEC-MALS experiments.

I would like to thank the Synchrotron facilities at Diamond Light Source, UK; ESRF, Grenoble; Elettra Sincrotrone Trieste, Italy and, macromolecular crystallography facility at IISER Pune, NCCS and NCL for crystal diffraction and data collection.

I would like to thank my juniors, Puja, Ranjana and Rinku for proofreading my thesis.

I would like to thank all the Bio office personnel in IISER Pune, especially Sandeep, Kalpesh, Piyush, Mrinalini, Sachin, Hariprasad, Rupali, Sandeep and Mahesh sir for their assistance with reagents, instruments and the relevant official procedures. I am grateful to the entire department of Biology and all the past and present chairs for successfully maintaining an open lab structure, conducive to fruitful research. I would also like to thank the housekeeping personnel for their daily help with washing and autoclaving.

I would like to thank the administrative staff, especially Swapnil, Alok, Sayalee, Ranjan, Tushar, Sneha and Nayana for providing me with all the necessary formal help for the completion of this degree and obtaining fellowships.

I am extremely grateful to IISER Pune, Infosys Foundation, Department of Biotechnology, Government of India and Biochemical Society, UK, for providing financial assistance during my course and for attending conferences.

My stay at IISER has been a rewarding experience with respect to nurturing my love for music and drama. I would like to thank all the past and present members of Aroha club, especially Harsha, Shridhar, Anuvind, Pratyush, Avisikta, Rutwik, Neelay, Farhan, Lakshmi, Rohit, Nagananda, Kshitij, Arya, Kwanit, Vineet and Akash for the wonderful jamming sessions and performances we staged over the last few years. I am grateful to all the past and present members of the Drama club, especially Amogh, Shweta, Chrisil, Sourajit, Avadhoot, Sameer, Dhriti and Anwesh for the fun-filled rehearsals and shows we performed all over Pune.

I received incessant support from my seniors and friends which helped me retain my sanity through the tough phases of this journey. I would like to thank all my batchmates, especially Krishnendu, Soumya, Kaveri, Aparna, Himani, Gauri, Joy, Indra, Saikat and Pulak for their constant support. I would like to specially mention Vani, a friend who has inspired me and

helped me through thick and thin. I would like to extend my regards to Arunabha, Saptashwa, Aslam, Basudev, Anweshi, Arijeet, Suddhasatwa, Kushan, Mrinmayee, Amogh, Kartik and Shrihar for a memorable stay at IISER Pune.

I would be remiss in not mentioning my school and college friends, Aparupa, Sreejib, Dishari, Pratim, Arnab, Swastik and Parijat for being there by my side over all these years.

Lastly, I would like to thank Projjwal Kanjilal for being a pillar of strength, through his constant motivation and care. I would like to thank my family, especially Ma and Kakai for all their love and blessings and for pushing me to dream bigger.

Sukanya

Table of Contents

List of Figures **2**

List of Tables **6**

Abbreviations **8**

Abstract **10**

Synopsis **12**

Chapter 1: Introduction **18**

1.1 Background 20

1.2 Biological oscillators: Their organization and importance 20

1.3 The cell polarity oscillator in *Myxococcus xanthus* 22

1.4 Common mechanisms of small Ras-like GTPase regulation 25

1.5 Types of motilities in *Myxococcus xanthus* 28

1.6 Spatial regulation of motility 31

1.7 Current model of polarity reversal 38

1.8 Motivation behind the study 41

Chapter 2: Materials and methods **48**

2.1 Introduction 50

2.2 Molecular Cloning 50

2.3 Protein Expression 58

2.4 Protein Purification 61

2.5 Protein Quality Check 64

2.6 Activity Assays 67

2.7 Binding assays 70

2.8 Protein Crystallography 73

Chapter 3: Characterization of a non-canonical Walker B motif in MglA 82

3.1 Introduction 84

3.2 Results 85

3.2.1 Identification of the probable Walker B residues in MglA *85*

3.2.2 Cloning and purification of MglA Walker B mutants *87*

3.2.3 Validation of the key Walker B residue *88*

3.2.4 Comparison with Walker B in other P-loop NTPases *90*

3.2.5 Regulation of the conformation of Walker B residues *92*

3.3 Conclusion 95

Chapter 4: Mechanism of MglA activation by an asymmetrically interacting MglB dimer 98

4.1 Introduction: 100

4.2 Results: 102

4.2.1 Design of the linked constructs for a functional asymmetric MglA MglB complex *102*

4.2.2 Cloning and purification of the asymmetric MglA MglB complexes *103*

4.2.3 Differential stability of the linked complexes *104*

4.2.4 Identifying the nucleotide bound to the linked complexes *109*

4.2.5 Effect of MglB Ct-helices in GTP hydrolysis rates of MglA *109*

4.2.6 Effect of MglB Ct-helices on nucleotide exchange by MglA *111*

4.2.7 Role of N-terminal β-strand of $MglB_1$ *113*

4.2.8 Role of Asn-37 and Arg-98 for MglB GAP activity *116*

4.3 Conclusion 117

Chapter 5: Characterization of RomX, a novel protein in *Myxococcus xanthus* polarity module 122

5.1 Introduction 124

5.2 Results 125

5.2.1 Cloning and purification of RomX constructs *125*

5.2.2 RomX is monomeric in solution *126*

5.2.3 RomX is composed of alpha helices *127*

5.2.3 Structure determination of RomX *128*

5.2.4. Features of RomX structure *130*

5.2.5 Interaction of RomX with MglA *133*

5.2.6 RomX acts as MglA GEF *134*

5.3 Conclusion 135

Chapter 6: Functional dissection of multiple domains of the response regulator, RomR **138**

6.1 Introduction 140

6.2 Results 142

6.2.1 Cloning and purification of RomR domains *142*

6.2.2 Interaction of RomR domains with RomX and MglA *144*

6.2.3 Effect of RomR domains on MglA activity *150*

6.2.4 Interaction between RomR-REC and RomX *152*

6.2.5 Characterizing the C-terminal output domain *155*

6.3 Conclusion 159

Chapter 7: Preliminary investigation of MglB interactors and their role in *Myxococcus xanthus* polarity reversal **162**

7.1 Introduction 164

7.2 Results 165

7.2.1 Fold Identification of RomY *165*

7.2.2 Expression and Purification of RomY *168*

7.2.3 Structure determination of RomY^{AC} *168*

7.2.4 MglB Ct helix is critical for stimulating the co-GAP activity of RomY *171*

7.2.5 Expression and Purification of MglC *173*

7.2.6 MglC interacts with RomR via the C-terminal region *173*

7.3 Conclusion 174

Chapter 8: Conclusion and future directions **176**

8.1 Major Findings 178

8.2 Lacunae in the previous models of polarity reversal 180

8.3 A revised model of the polarity module 180

8.4 Future Directions 184

List of Figures

Figure 1.1. Oscillations in biological systems 21
Figure 1.2. *Myxococcus xanthus*: Developmental stages and motility 23
Figure 1.3. Polarity reversals in *Myxococcus xanthus* 24
Figure 1.4. A model for the coupling between the Frizilator and the MglA/B oscillator 25
Figure 1.5. MglA, a member of the small Ras-like GTPase family 27
Figure 1.6. S and A-motility apparatus in *M. xanthus* 29
Figure 1.7. Chemosensory apparatus in *M. xanthus* 32
Figure 1.8. Cascade of proteins driving polarity reversal 33
Figure 1.9. Interface between MglA and MglB 35
Figure 1.10. Interaction of MglA with downstream effectors 38
Figure 1.11. Gated relaxation oscillator model 39
Figure 1.12. Current model of *M. xanthus* polarity reversal 40

Figure 2.1. RF cloning strategy 51
Figure 2.2. Strategy for cloning linked MglAB constructs 53
Figure 2.3. Schematic of Ni-NTA binding to hexahistidine tag 61
Figure 2.4. CD spectroscopy 65
Figure 2.5. NADH coupled GTPase assay 68
Figure 2.6. Nucleotide exchange assay 69
Figure 2.7. Isothermal Titration calorimetry 71
Figure 2.8. Oswald-Miers phase diagram 74
Figure 2.9. Protein structure determination 78

Figure 3.1. $G3^{Alt}$ motif as the alternate Walker B motif in the MglA family 85
Figure 3.2. Characterization of the probable Walker B mutants 87
Figure 3.3. Identification of Asp-58 of MglA as the key Walker B motif 89
Figure 3.4. Walker B conformations in P-loop NTPases 91
Figure 3.5. Conformational changes in Walker B across GTP to GDP-bound states 93

Figure 4.1. Asymmetric features of MglA and MglB complex 101
Figure 4.2. Design of the linker between MglA and $MglB_1$ 102
Figure 4.3. Co-expression strategy for purification of asymmetric MglA MglB complexes 104

Figure 4.4. Purification of asymmetric complexes 106
Figure 4.5. C-terminal residues of MglB contribute to differential stability of the linked complexes. 108
Figure 4.6. HPLC profiles of the linked complexes 109
Figure 4.7. Constructs with $MglB_1$ helix stimulate higher GTP hydrolysis rates of MglA. . 110
Figure 4.8. $MglB_1$ Ct helix stimulates nucleotide exchange on MglA 112
Figure 4.9. $MglB_1$ N terminal β-strand and $MglB_2$ Asn-37/Arg-98 are essential for MglB GEF and GAP activity, respectively 115
Figure 4.10. Mechanism of dual GAP and GEF activity of MglB 118

Figure 5.1. RomX is the newly identified GEF for MglA in the polarity module 125
Figure 5.2. Expression and purification of RomX constructs 126
Figure 5.3. RomX is composed of alpha helices 127
Figure 5.4. Structure determination of RomX using anomalous diffraction 128
Figure 5.5. Structural superposition of RomX and Atlastin 133
Figure 5.6. Binding studies of RomX and MglA 134
Figure 5.7. RomX acts as a MglA GEF 135
Figure 5.8. AlphaFold model of MglA with RomX 136

Figure 6.1. *In-vivo* studies with RomR domains 140
Figure 6.2. Domain-wise dissection of RomR using different constructs 143
Figure 6.3. Interaction of RomR-REC^{125} with RomX and MglA 145
Figure 6.4. Interaction of RomR-REC^{141} with RomX and MglA 146
Figure 6.5. Interaction of RomR-Ct with RomX 147
Figure 6.6. Interaction of RomR-REC^{141}Ct with RomX and MglA 148
Figure 6.7. Interaction of RomR-Ct with MglA 149
Figure 6.8. Activity assays of RomR constructs with RomX and MglA 151
Figure 6.9. Interaction between RomR-REC and RomX 153
Figure 6.10. Interface between RomR-REC and RomX 155
Figure 6.11. C-terminal tail of RomR 156
Figure 6.12. Analyzing the disordered linker of RomR 159
Figure 6.13. AlphaFold model of MglA, RomX and RomR-REC 160

Figure 7.1. Fold prediction of RomY 166

Figure 7.2. Comparison of RomY fold with Roadblock and longin domains 167
Figure 7.3. Crystallization of RomY$^{\Delta C}$ 169
Figure 7.4. Crystal structure of RomY$^{\Delta C}$ 170
Figure 7.5. Interaction of RomY with MglA MglB complex 172
Figure 7.6. Interaction of MglC with RomR 173

Figure 8. Current model of polarity reversals in *M. xanthus* 182

List of Tables

Table 2.1: List of constructs and their respective primers used for cloning..... 54
Table 2.2: Expression conditions for constructs used in the study..... 58
Table 2.3: Recipe for minimal media for expressing selenomethionine labelled proteins..... 60
Table 2.4: Summary of crystallization trials 75

Table 4.1: Summary of observed k_{cat} for linked and unlinked complexes 111
Table 4.2: Summary of rates of nucleotide association and dissociation for the linked and unlinked MglA MglB complexes 113
Table 4.3: Summary of observed k_{cat} for MglB mutants with MglA 114
Table 4.4: Summary of rates of nucleotide association and dissociation for the unlinked MglA MglB complexes with the MglB mutants..... 116

Table 5.1: Data Collection and Refinement Statistics of RomX..... 129
Table 5.2: List of homologous structures of RomX identified by DALI 131

Table 8: Summary of polarity module proteins and their functions..... 183

Abbreviations

A-motility	Adventurous gliding motility
Amp	Ampicillin
ASCE	Additional strand conserved glutamate
ATPase	ATP hydrolase
CD	Circular dichroism
CHis	C-terminal hexahistidine tag
Ct	C terminal
DALI	Distance matrix alignment
E. coli	*Escherichia coli*
Frz	Frizzy proteins
GAP	GTPase activating protein
GDP	Guanosine diphosphate
GEF	Gunanine nucleotide exchange factor
GMPPCP	Guanosine-5'-[(β,γ)-methyleno]triphosphate
GMPPNP	Guanosine 5′-[(β,γ)-imido]triphosphate
GTP	Guanosine triphosphate
GTPase	GTP hydrolase
HPLC	High Pressure Liquid Chromatography
IDR	Intrinsically disordered region
ITC	Isothermal titration calorimetry
Kan	Kanamycin
K_D	Dissociation constant
k_{obs}	Observed turnover number
k_{off}	Rate of dissociation
k_{on}	Rate of association
LS	Light scattering
λ_{max}	Emission maximum
M. xanthus	*Myxococcus xanthus*
MAD	Multiple wavelength anomalous diffraction
MALS	Multi angle light scattering
mant	(N-methylanthraniloyl)

MglA	Mutual gliding protein A
MglB	Mutual gliding protein B
$MglB^{\Delta Ct}$	MglB with C-terminal truncation
MglC	Mutual gliding protein C
NADH	Nicotinamide adenine dinucleotide
NHis	N-terminal hexahistidine tag
NTPase	Nucleotide triphosphate hydrolase
OD_{600}	Optical density at 600 nm
PCR	Polymerase chain reaction
pLDDT	Predicted local distance difference test
REC	Receiver domain
RF	Restriction free
RI	Refractive index
RMSD	Root Mean Square Deviation
RomR	Regulator of motility response regulator
RomX	Regulator of motility X
RomY	Regulator of motility Y
$RomY^{\Delta C}$	RomY with C-terminal truncation
S-motility	Social gliding motility
SAD	Single wavelength anomalous diffraction
SDS	Sodium dodecyl sulphate
SEC	Size exclusion chromatography
SIMIBI	Signal recognition particle, MinD, and BioD
SRP	Signal recognition particle
T4P	Type IV pili
TRAFAC	Translation factor

Abstract

Polarity reversal in the soil bacterium *Myxococcus xanthus* is regulated by a cascade of proteins depending on extracellular signals. MglA, a prokaryotic small Ras-like GTPase, is the master regulator of this motility. MglA GTP hydrolysis is accelerated by its cognate GTPase activating protein (GAP), MglB which also acts as the Guanine nucleotide exchange factor (GEF). My work aims to dissect this dual GAP and GEF activity of MglB in light of the asymmetric interaction of MglB dimer with MglA. We observe that the C-terminal helix of the $MglB_1$ protomer interacts with MglA opposite the nucleotide-binding pocket, allosterically stimulating nucleotide exchange. The GAP activity is driven by the MglB dimeric interface, which orients the MglA catalytic residues. Interaction of MglB orients a non-canonical Walker B motif in MglA important for nucleotide binding. We also characterize RomY, a newly identified co-GAP of MglB.

RomR-RomX, a complex of two proteins, has also been shown to act as GEF for MglA, making it interesting to understand the regulation of these two different GEF modules. Domain-wise characterisation of RomR shows the N-terminal receiver domain interacting with RomX and consequently, we identify the interface. The crystal structure of RomX reveals it to be an alpha-helical protein, strikingly similar to the stalk domain of the GTPase atlastin. We observe the GEF activity of RomX without any contribution from the RomR receiver domain. We compare this with the MglB GEF activity, which aids in distinguishing the mechanisms of MglA activation by MglB and RomR-RomX. RomR has also been shown to interact with MglC, which also interacts with MglB. We observe that the C-terminal domain of RomR interacts with MglC, which plays a role in recruiting MglB. Fitting all these pieces into the puzzle, we finally propose an updated model of the mechanism of cell polarity maintenance and reversal, characterizing *Myxococcus xanthus* motility.

Synopsis

Understanding the mechanism of regulation of a small Ras-like GTPase, MglA, driving *Myxococcus xanthus* motility

Name: Sukanya Chakraborty
Registration number: 20162012
Supervisor: Dr. Gayathri Pananghat
Department: Biology, IISER Pune, India
Program: Integrated PhD
Date of registration: August 2016

Motility in the soil bacterium *Myxococcus xanthus* is characterized by frequent reversal of cell polarity depending on extracellular signals[1]. MglA, a prokaryotic small Ras-like GTPase, is the master regulator of this motility pattern[2]. MglA-GTP localises to the leading pole of the cell and relocates as the cell reverses polarity, whereas MglA GDP is diffused in the cytoplasm[3]. The localization and function of MglA are regulated by multiple interacting proteins, which together constitute the cell polarity oscillator in this bacterium.

MglA GTP hydrolysis is accelerated by its cognate GTPase activating protein (GAP), MglB, an Rbl/LC7 protein, which also acts as the Guanine nucleotide exchange factor (GEF)[4–7]. MglB is present at both cell poles, with a higher concentration at the lagging pole. The GAP activity of MglB is further stimulated by RomY at the lagging pole, which together helps exclude MglA from this pole[8]. The structure of the MglA-MglB complex shows that a dimer of MglB interacts asymmetrically with one molecule of MglA. The Rbl/LC7 domain of MglB contributes to the GAP activity, whereas the C-terminal helix contributes to the GEF activity[7]. However, the exact mechanism and relevance of the MglB GEF activity during polarity reversals are unclear.

In addition to MglB, another complex of two proteins, RomR-RomX, has also been shown to act as GEF for MglA[9]. RomR-RomX localizes similarly to MglB in the cell. RomR is a multidomain protein with an N-terminal response receiver domain followed by a long-disordered linker and a C-terminal domain[10]. The dynamics of RomR help in regulating the frequency of cell polarity oscillations[11,12]. RomR has been shown to stimulate the GEF activity of RomX, which is a protein of novel fold. RomR-RomX together has been proposed to recruit MglA GTP at the leading pole. There are several unresolved questions pertaining to the

localization and GEF activity of the RomR-RomX complex. Further, it is interesting to understand the relevance of the two GEF modules, MglB and RomR-RomX, regulating MglA localization during polarity reversals.

MglB has also been shown to interact with another Rbl/LC7 protein, MglC, which also interacts with RomR[13]. MglB and MglC have been shown to form a 2:4 complex[14]. It localizes similarly to MglB with a major cluster at the lagging cell pole. But the exact function of MglC in the polarity reversal module remains to be deciphered.

In this thesis, we have attempted to dissect the mechanism of regulation of MglA by its interacting partners and their relevance in the polarity reversal module.

Chapter wise summary

Chapter 1: Introduction

In this chapter, we have performed a comprehensive review of the literature concerning the cell polarity oscillator in *Myxococcus xanthus*, the proteins orchestrating cell polarity reversals and their role in driving cell motility. We have detailed the current opinions regarding the sequence of events driving reversals and the unanswered questions in the field, some of which we have tried to address in this thesis.

Chapter 2: Materials and Methods

This chapter enlists the techniques for biochemical and structural studies used in this work.

Chapter 3: Characterization of a non-canonical Walker B motif in MglA

In this chapter, we experimentally validated that Asp-58 of MglA, present in an alternate G3 ($G3^{Alt}$) motif, acts as a non-canonical Walker B. It is important for maintaining the Mg^{2+} coordination at the nucleotide-binding pocket. We performed an exhaustive study of other P-loop NTPases to compare the position of Walker B and the associated conformational changes regulated across the GDP and GTP-bound states. The $G3^{Alt}$ motif is a part of the MglB interface, which allosterically regulates the positioning of the Walker B. The presence of a Thr-78 at the canonical Walker B position in MglA, enabled us to identify an equivalent conserved threonine in ASCE ATPases. Further, we proposed that across evolution, Thr-78 of MglA has

been replaced with aspartate in the eukaryotic successors, to achieve the function of nucleotide binding.

Chapter 4: Mechanism of MglA activation by an asymmetrically interacting MglB dimer

We attempted to dissect the dual GAP and GEF activity of MglB, regulating the MglA nucleotide-bound state. We designed a unique strategy to generate MglA complexed to asymmetric MglB dimers with C-terminal helix truncations. We confirmed that the C-terminal helix of MglB, opposite to the nucleotide-binding pocket of MglA, allosterically drives GDP to GTP exchange on MglA. We designed mutants of MglB which confirmed that the GAP activity of MglB is driven by the dimeric interface of the MglB Roadblock domain, with a major contribution from the protomer near the Switch residues of MglA. Hence, we confirmed that the interaction of the $MglB_2$ protomer with MglA helps in orienting the catalytic residues of the GTPase in a conformation to drive GTP hydrolysis by MglA. Identification of distinct GAP and GEF domains of MglB enabled us to propose a model of how these activities can be spatially regulated across the cell poles.

Chapter 5: Characterization of RomX, a novel protein in *Myxococcus xanthus* polarity module

We determined the structure of RomX, a newly identified protein with no characterized homologs. It was found to be structurally similar to the stalk domain of atlastin, a member of the dynamin family GTPases. It was observed that the interaction between MglA and RomX was weak and transient. We further observed the nucleotide exchange activity on MglA by RomX and compared it with that observed with MglB.

Chapter 6: Functional dissection of multiple domains of the response regulator, RomR

Further, we perform a domain-wise dissection of RomR and its function. We characterized the RomR/RomX complex, mediated by the RomR N-terminal REC domain and the residues critical for forming the interface. However, as reported earlier, we did not observe the additional stimulation of RomX GEF activity in the presence of RomR. Using the information on the RomR-RomX complex, we generated models with MglA and found that MglA was captured in the GDP state, as expected for a GEF. Finally, we inspected the long-disordered linker of RomR and compared it with a dataset of RomR-like proteins, and observed that these regions mostly form disordered globules and can expand based on additional interactions.

Chapter 7: Preliminary investigation of MglB interactors and their role in *Myxococcus xanthus* polarity reversal

To fill the gaps in understanding the polarity module, we initiated the characterization of MglB interactors, which might play a role in regulating its GAP and GEF activity. We determined the structure of RomY, the MglB GAP activator. We also observed that the C-terminal helix of MglB is critical for RomY binding and activity. Further, we observed that MglC interaction with RomR is mediated through the C-terminal domain of RomR, with no contribution from the receiver domain. This shows how RomR, MglC and MglB might form a heteromeric complex to recruit these proteins to the lagging pole.

Chapter 8: Conclusion and future directions

In this chapter, we presented an updated picture of the polarity reversal module, converging the existing literature with the major findings from the thesis. We also highlighted the missing links in the model, which require further experimental validation in the future.

List of publications:

1. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to polarity oscillations in bacterial motility

Jyoti Baranwal, Sebastian L'Hospice, Manil Kanade, **Sukanya Chakraborty**, Priyanka Gade, Shrikant Harne, Tâm Mignot, Pananghat Gayathri; ***PLOS Biology, 17(9), e3000459 (2019)***

2. A Distinct Motif in a Prokaryotic Small Ras-Like GTPase Highlights Unifying Features of Walker B Motifs in P-Loop NTPases

Manil Kanade, **Sukanya Chakraborty**, Sanket Shelke, Pananghat Gayathri; ***Journal of Molecular Biology (Cover article: Volume 432, Issue 20, Pages 5544-5564) (2020)***

3. Mechanism of GTPase activation of a small Ras-like GTPase MglA by an asymmetric dimer of MglB

Sukanya Chakraborty, Manil Kanade, Pananghat Gayathri; ***bioRxiv 2023.03.05.531159; doi: https://doi.org/10.1101/2023.03.05.531159***

4. Dissecting the mechanism of nucleotide exchange by a multidomain response regulator RomR and a novel prokaryotic GEF RomX

Sukanya Chakraborty, Akshata Kotwal, Pananghat Gayathri ***(Manuscript under preparation)***

5. Fitting pieces into the puzzle: A picture of cell polarity maintenance and reversals in *Myxococcus xanthus*

Sukanya Chakraborty, Pananghat Gayathri ***(Review under preparation)***

References

1. Leonardy, S., Bulyha, I. & Søgaard-Andersen, L. Reversing cells and oscillating motility proteins. *Mol Biosyst* **4**, 1009 (2008).
2. Hartzellt, P. & Kaiser, D. Function of MglA, a 22-Kilodalton Protein Essential for Gliding in *Myxococcus xanthus*. *J. Bacteriol.* vol. 173 (1991).
3. Nan B, Bandaria JN, Guo KY, Fan X, Moghtaderi A, Yildiz A, Zusman DR.The polarity of myxobacterial gliding is regulated by direct interactions between the gliding motors and the Ras homolog MglA. *Proc Natl Acad Sci U S A* **112**, E186-93 (2015).
4. Miertzschke M, Koerner C, Vetter IR, Keilberg D, Hot E, Leonardy S, Søgaard-Andersen L, Wittinghofer A. Structural analysis of the Ras-like G protein MglA and its cognate GAP MglB and implications for bacterial polarity. *EMBO J*. 2011 Aug 16;30(20):4185-97. doi: 10.1038/emboj.2011.291.
5. Zhang, Y., Franco, M., Ducret, A. & Mignot, T. A bacterial Ras-like small GTP-binding protein and its cognate GAP establish a dynamic spatial polarity axis to control directed motility. *PLoS Biol* **8**, (2010).
6. Galicia C, Lhospice S, Varela PF, Trapani S, Zhang W, Navaza J, Herrou J, Mignot T, Cherfils J. MglA functions as a three-state GTPase to control movement reversals of *Myxococcus xanthus. Nat Commun*. 2019 Nov 22;10(1):5300. doi: 10.1038/s41467-019-13274-3.
7. Baranwal J, Lhospice S, Kanade M, Chakraborty S, Gade PR, Harne S, Herrou J, Mignot T, Gayathri P. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. *PLoS Biol.* 2019 Sep 27;17(9):e3000459. doi: 10.1371/journal.pbio.3000459.
8. Szadkowski, D., Carreira, L. A. M. & Søgaard-Andersen, L. A bipartite, low-affinity roadblock domain-containing GAP complex regulates bacterial front-rear polarity. *PLoS Genet* **18**, (2022).
9. Szadkowski D, Harms A, Carreira LAM, Wigbers M, Potapova A, Wuichet K, Keilberg D, Gerland U, Søgaard-Andersen L. Spatial control of the GTPase MglA by localized RomR-RomX GEF and MglB GAP activities enables *Myxococcus xanthus* motility. *Nat Microbiol*. 2019 Aug;4(8):1344-1355. doi: 10.1038/s41564-019-0451-4.
10. Leonardy, S., Freymark, G., Hebener, S., Ellehauge, E. & Søgaard-Andersen, L. Coupling of protein localization and cell movements by a dynamically localized response regulator in *Myxococcus xanthus*. *EMBO J.* **26**, 4433–4444 (2007).
11. Zhang, Y., Guzzo, M., Ducret, A., Li, Y.-Z. & Mignot, T. A dynamic response regulator protein modulates G-protein–dependent polarity in the bacterium *Myxococcus xanthus*. *PLoS Genet* **8**, e1002872 (2012).
12. Guzzo M, Murray SM, Martineau E, Lhospice S, Baronian G, My L, Zhang Y, Espinosa L, Vincentelli R, Bratton BP, Shaevitz JW, Molle V, Howard M, Mignot T. A gated relaxation

oscillator mediated by FrzX controls morphogenetic movements in *Myxococcus xanthus*. *Nat Microbiol*. 2018 Aug;3(8):948-959. doi: 10.1038/s41564-018-0203-x.

13. McLoon AL, Wuichet K, Häsler M, Keilberg D, Szadkowski D, Søgaard-Andersen L. MglC, a Paralog of *Myxococcus xanthus* GTPase-Activating Protein MglB, Plays a Divergent Role in Motility Regulation. *J Bacteriol*. 2015 Nov 16;198(3):510-20. doi: 10.1128/JB.00548-15.
14. Kapoor, S., Kodesia, A., Kalidas, N., Ashish & Thakur, K. G. Structural characterization of *Myxococcus xanthus* MglC, a component of the polarity control system, and its interactions with its paralog MglB. *JBC* **296**, (2021).

Chapter 1: **Introduction**

1.1 Background

All living systems are bestowed with regulatory circuits that largely influence their most essential survival processes[1]. Mostly, such circuits function by fine-tuning their gene expression based on specific cues and orchestrating events like cell division, chromosome segregation, cell cycle progression and cell polarity specification[2–6]. All of these require the integration of external signals to the cellular interior to execute responses. This enables the cell to modulate these circuits with respect to the changing environment and ultimately facilitate survival (Fig 1.1 A). These circuits operate over time and space to form intracellular patterns [7]. Often, these show periodicity in properties giving rise to temporal and spatial oscillatory patterns. Unlike temporal oscillators like circadian rhythms, the spatial oscillatory mechanisms have been relatively less explored, especially in the context of prokaryotic systems. One of the notable examples of such a phenomenon is the event of polarity reversal in the bacterium *Myxococcus xanthus*, which plays a key role in driving its motility, critical to maintaining its life cycle[6,8]. The core of this oscillatory system consists of proteins which exhibit unique subcellular localisation patterns determining the front-rear cell polarity, which in turn, is tightly regulated by external chemosensory cues[9,10]. In terms of the localisation and identity of proteins and also their regulatory mechanisms, this event resembles the phenomenon of cell migration observed in immune cells or metastatic cancer cells in eukaryotes.

1.2 Biological oscillators: Their organization and importance

One of the features defining a living system is the capability to respond to changing environments and consequently adapt. One of the mechanisms to elicit such a response is by periodically modulating gene expression or localization over spatial and temporal scales or, more appropriately, by generating oscillations[7,11]. Oscillations not only provide a way to entrain the system with the periodic variations in the environment but also can transmit a greater amount of information in terms of frequency and amplitude variations rather than a steady response[12] (Fig 1.1 B). From a mechanistic point of view, oscillations are periodic variations about a central tendency which switches between two distinctly defined states. From a biological perspective, a central tendency would be something like the mid-cell positional coordinates or the optimum activity of a protein. Such a response can be executed to give rise to a spectrum of different states, each of which is relevant in terms of the optimal functioning of an organism in a fluctuating environment.

In the case of an ideal oscillator, it behaves harmonically in space and time. However, in real situations, such a system is usually dampened by the external resisting forces, which consequently decrease the amplitude of oscillations, a phenomenon more commonly known as dampening[13,14]. However, if there is a source of energy that is supplied at a frequency equivalent to the oscillation frequency, it can be transformed into a sustained oscillation with the same initial amplitude till infinite time points despite the presence of external dampening forces. Such a phenomenon occurs in biological systems where energy obtained from bond breakage by cleaving macromolecules is used to drive such sustained oscillatory circuits over time[15].

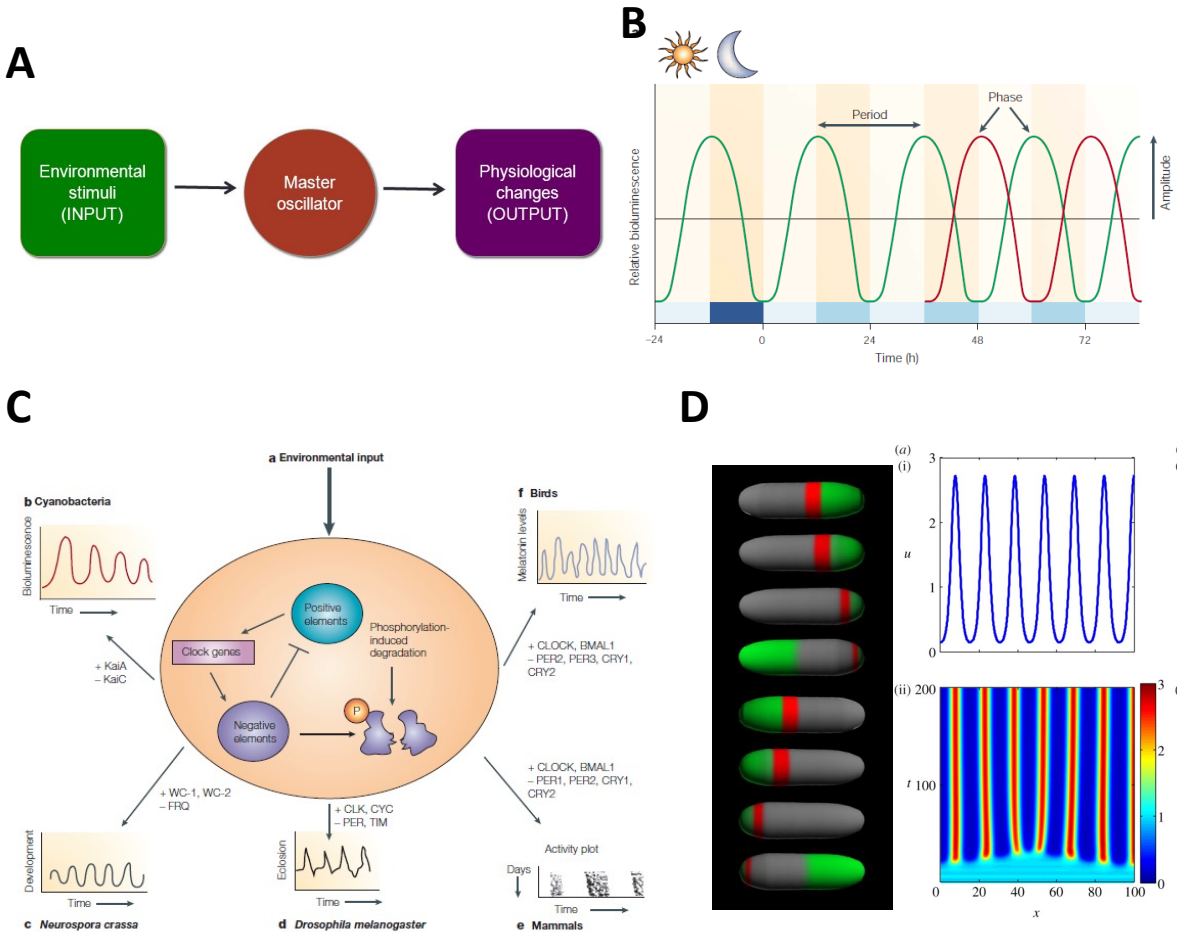


Figure 1.1. Oscillations in biological systems

A. Components constituting biological oscillators

B. Pattern of circadian rhythm maintained by external stimuli. Sustained oscillations are maintained under constant conditions. Measured features: period (the duration between peaks or troughs), the amplitude (the level of expression measured from the midline to either the peak or trough), and the phase (the relative positioning of the curve about a specific time point) (Reproduced from Golden & Canales, 2003[12]).

C. Common mechanisms driving circadian oscillators. The clock genes drive these oscillators under the influence of positive and negative feedback from regulatory elements (Reproduced from Bell-Pedersen *et al.*, 2005[16]).
D. Spatial oscillatory patterns formed by MinD/E (green) and MinC (red) across cell poles, which function to position the cell division machinery at the mid-cell. On the right are simulated Turing patterns characterizing diffusion-driven morphogen gradients observed during embryonic development (Reproduced from Meinhardt & de Boer, 2001 and Maini *et al*., 2012[13,17]).

Biological oscillators can be broadly categorized into temporal and spatial oscillators[7]. Temporal oscillators include events such as the accumulation or activity of proteins or other metabolites with respect to temporal cycles such as the circadian rhythm[12,16,18,19] (Fig 1.1 C). On the other hand, spatial oscillators involve variation in the specific localization of a protein or a macromolecule in space to determine different subcellular positions to form distinct patterns or gradients, which spatially regulate an event[1,5,20–28] (Fig 1.1 D). All of these oscillatory mechanisms can be broadly stripped down to a basic structure. It consists of an input sensory signal, which regulates the oscillator with respect to external cues. This feeds into the core oscillator consisting of the biochemical machinery, which generates the oscillations depending on the input. This oscillator communicates with downstream signaling pathways, manifesting the functional output.

1.3 The cell polarity oscillator in *Myxococcus xanthus*

Myxoccoccus xanthus is a rod-shaped, Gram-negative, soil-inhabitant bacterium with unique motility patterns[29–32]. It utilizes both Type IV pili-dependent and independent mechanisms to drive its movement. It goes through a complex life cycle regulated by the availability of nutrients[33–35] (Fig 1.2). Under nutrient-abundant conditions, the bacteria collectively swarm to obtain nutrients through predation or acting as saprotrophs (Fig 1.2 Stage I). Under conditions of intermediate nutrient availability, they display a rippling behavior characterized by swarms with alternate regions of high and low cell densities (Fig 1.2 Stage II). Under nutrient-limiting conditions, the cells collectively sporulate and aggregate to form multicellular fruiting bodies (Fig 1.2 Stage III, IV). All these three stages are characterized by differently regulated motility patterns[31].

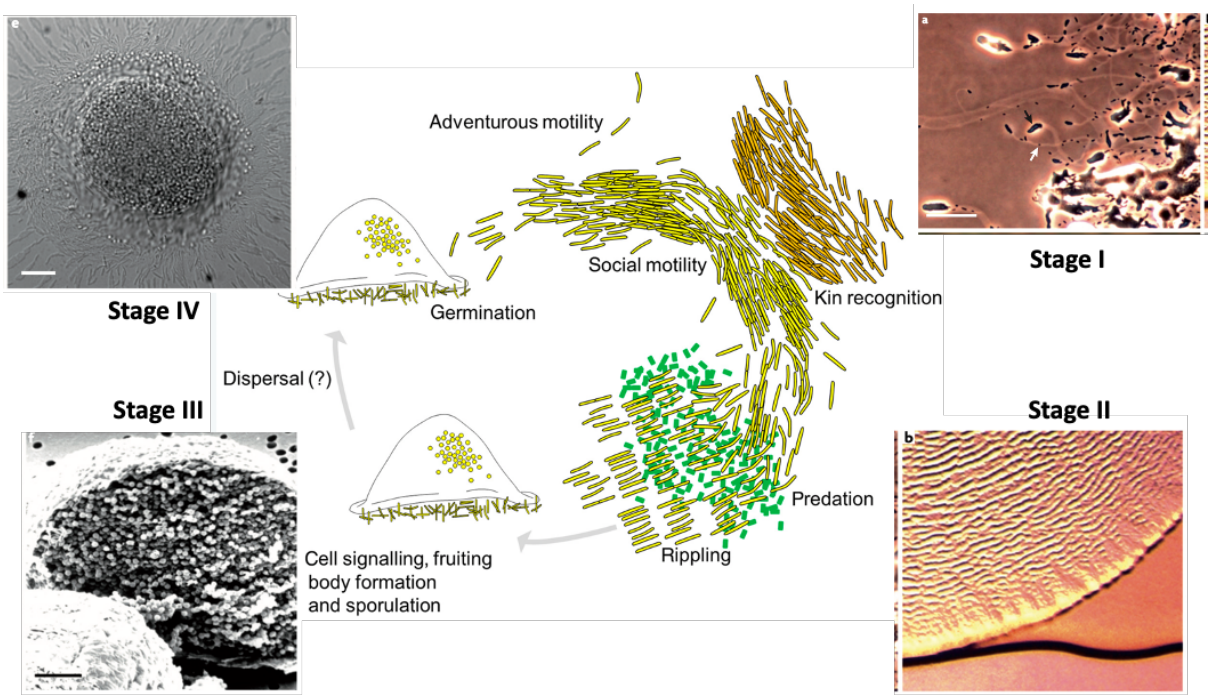


Figure 1.2. *Myxococcus xanthus*: Developmental stages and motility
The central panel represents the complex life cycle of *Myxococcus xanthus* through different phases of nutrient availability. (Reproduced from Vos, 2021[29]). All scale bars in individual panels are 15 µm (Reproduced from Zusman *et al.*, 2007[31]).
Stage I. Swarms of cells migrate as groups exhibiting S or social motility (black arrow), with individual cells exploring the environment using A or adventurous motility (white arrow).
Stage II. Cells move in ripples exhibiting predatory behaviour. Reversals are often stimulated at this stage to regulate the cell density of ripples, modulated by extracellular signals. Ripples have a wavelength of ~100 µm.
Stage III. Cross-section of a fruiting body showing the enclosed spores.
Stage IV. Optical slice of a fruiting body showing central core containing spores and the periphery has to germinate individual rod-shaped cells.

Rod-shaped *Myxococcus* cells have a specified front-rear polarity that defines their direction of movement[34,36,37] (Fig 1.3 A). The basis of the polarity is due to the specific localization of polar-situated proteins. Motility patterns, as mentioned before, are characterized by frequent reversal of cell polarity[9,38]. This causes movement of the cell in the opposite direction associated with the re-localization of these polarly localized proteins (dynamic proteins) (Fig 1.3 B). These proteins that oscillate pole to pole form the polarity module essential for driving cell motility.

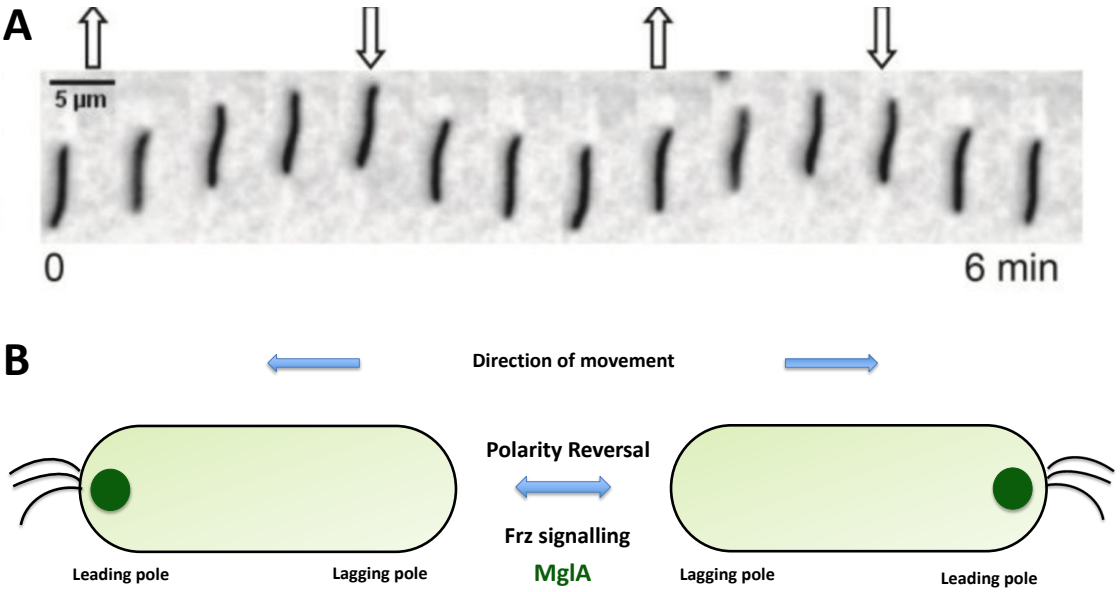


Figure 1.3. Polarity reversals in *Myxococcus xanthus*
A. Time-lapse microscopy at 40 s intervals of a single *M. xanthus* cell. Each reversal event is marked by a change in the direction of movement (indicated by arrows). This is a hyper-reversing mutant, with reversals occurring every 2 mins (Reproduced from Kaimer *et al.*, 2014[39]).
B. Polarity reversal is characterized by swapping of the leading and lagging poles of the cell associated with the assembly of T4P at the leading pole. The GTPase MglA also switches across respective lagging poles. The frequency of reversals is regulated by Frz chemosensory module (adapted from Baranwal *et al.*, 2019[40]).

The cell polarity oscillator in *Myxococcus* consists of two inter-connected oscillatory circuits (Fig 1.4) that together fine-tune the frequency of reversal events[41]. This is regulated by the extracellular signals, which act as a cue for nutrient availability, in turn facilitating the appropriate motility behavior for the particular stage of the life cycle. MglA, a bacterial GTPase, lies at the centre of regulation of the polarity reversals and cell motility[42]. Apart from the MglA regulated central polarity module consisting of several dynamic polar localized proteins, there is a chemosensory-like apparatus, more commonly known as a 'frizilator', which is also a separate oscillator that relays messages from the extracellular environment. It senses external diffusible signals or contact-dependent signals from the extent of cellular crowding[31,39,43]. The central polarity oscillator was recently found to operate without any external input at its intrinsic frequency due to the probable feedback from the downstream proteins. All three modules driving the *Myxococcus* motility will be subsequently detailed. Both oscillators should be capable of persisting independently, but the relevant physiological output is manifested only when they are coupled. Functionally such a coupled oscillatory

system enables the bacterium to switch from low reversal frequencies, which facilitate a wandering behavior, to a high reversal frequency associated with homing or staying close to a nutrient source.

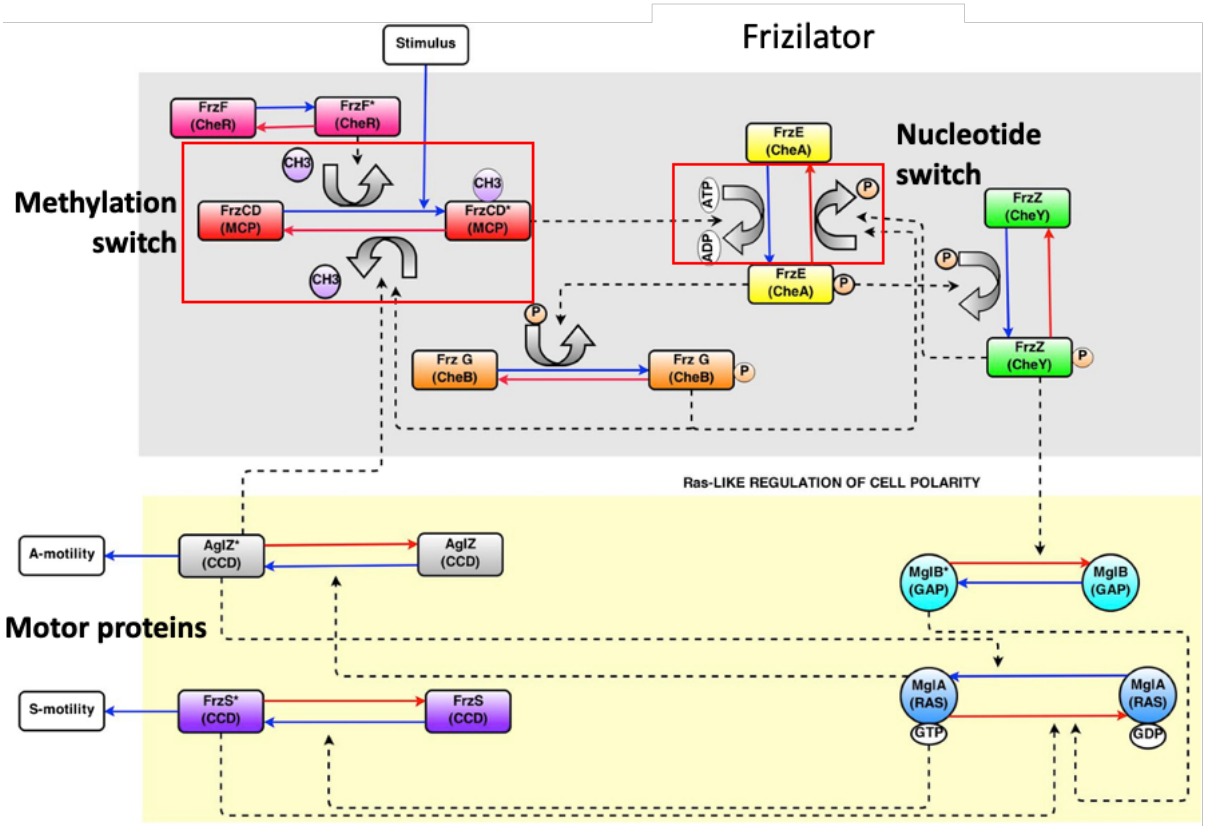


Figure 1.4. A model for the coupling between the Frizilator and the MglA/B oscillator
The Frz proteins sense the external stimulus via a methylation-dependent switch which triggers a cascade of downstream phosphorylation events which feed onto the MglA/B oscillator. The downstream motors also provide feedback (Reproduced from Eckhert *et al.*, 2014[41]).

1.4 Common mechanisms of small Ras-like GTPase regulation

MglA is a prokaryotic GTPase which acts as the master regulator of polarity reversals and cell motility[44,45]. MglA belongs to the family of small Ras-like GTPases, representing a class of hydrolase enzymes that hydrolyze GTP to GDP. These GTPases go through a GTPase cycle switching from their active GTP state to the inactive GDP state. The switching of Ras GTPases has been implicated in several phenomena like cell division, migration, growth, etc. These GTPases are recognized by the presence of the conserved G1-G5 motifs[46–49] (Fig 1.5 A). The G1 motif consists of a phosphate-binding P-loop with its "GXXXXGKS/T" motif[50,51]. G2 consists of a conserved threonine which is critical for nucleotide binding and hydrolysis. G3

motif with its characteristic "DXXG" motif is known to interact with the ɣ-phosphate oxygen, with the aspartate interacting with a water molecule, which maintains the Mg^{2+} coordination at the active site[52]. The G2 and G3 motifs constitute the Switch I and Switch II loops, which undergo conformational change as the GTPase switches from its GTP to GDP-bound states (Fig 1.5 B). G4 domain provides the guanine base specificity through its "N/TKXD" motif. G5 domain orients the G4 loop and also assists base recognition. MglA is the first member of the prokaryotic small Ras-like GTPase to be characterized[42]. It is an ancestor to the Arf-Sar-Gα lineage of small Ras-like GTPases, which has been extensively studied in eukaryotes. It is a 22-kDa protein with a core G-domain consisting of six β-strands and five α-helices[53] (Fig 1.5 E). It contains the aforementioned conserved motifs that drive its function of nucleotide binding and hydrolysis (Fig 1.5 D). To understand the mechanism of MglA regulation, it is important to focus on the common mechanism of regulation of small Ras-like GTPases[46–48,54].

GTPase activating proteins (GAPs) typically bind to the GTPase in the GTP-bound state (Fig 1.5 C). It stimulates the GTP hydrolysis rates, thus resulting in a fast flipping of the GTPase switch from its "ON" GTP state to its "OFF" GDP state. The conserved catalytic residues include a glutamine that orients the catalytic water molecule for nucleophilic attack and an arginine that stabilizes the negative charge on the phosphate (Fig 1.5 B). These residues must be optimally positioned in space to trigger hydrolysis. Usually, in small Ras-like GTPases, one of these residues is provided by the GAP, and the other is present in the GTPase at the active site[55,56].

Small Ras-like GTPases have a higher affinity to the product GDP than the substrate GTP (Fig 1.5 C). Hence it requires guanine nucleotide exchange factors (GEFs) which bind to the GTPase in the GDP-bound state and consequently weaken the affinity for GDP. This could be achieved by inducing a conformational change in the Switch loops causing the GTPase to release GDP[57]. A GEF could also displace the bound GDP by sterically occluding the nucleotide at the binding pocket. GEFs typically stabilize the nucleotide-free pocket of the GTPase till GTP occupies it. GTP binding causes the GEFs to lose their affinity for the GTPase.

Additionally, GEF activity is regulated by guanine nucleotide exchange inhibitors or GDIs, which bind to the GTPase in the GDP state, competitively inhibiting the binding of GEFs (Fig 1.5 C)[58]. This retains the GTPase in the "OFF" state. Further, when GDFs, or guanine nucleotide dissociation factor, binds to the GDI, they lose their affinity for the GTPase[58]. This releases the GTPase resulting in the consequent binding of GEFs, switching the GTPase back

to its GTP-bound "ON" state. In the active state, GTPases interact with effectors activating downstream signalling pathways regulating vital cellular processes. MglA, a *Myxococcus xanthus* small Ras-like GTPase, is also regulated by various such regulators, together driving motility of this bacterium.

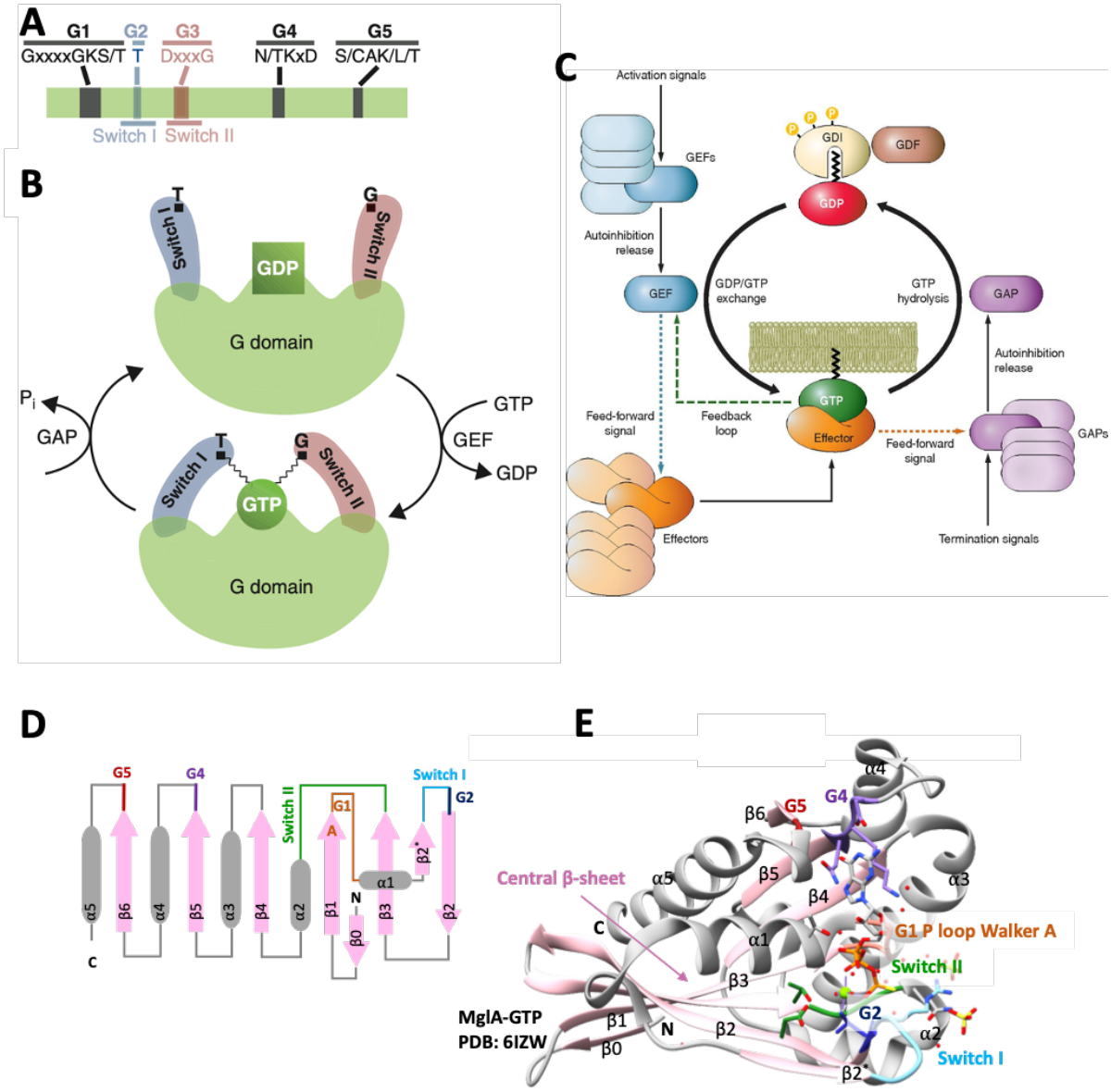


Figure 1.5. MglA, a member of the small Ras-like GTPase family

A. Conserved motifs, G1-G5 and the switch loops of a prototypic G-domain. The conserved residues belonging to each domain are labelled (Reproduced from Bulyha *et al.*, 2011[57]).

B. The GTPase cycle of a canonical small Ras-like GTPase with the activity of GAPs and GEFs is shown. The conformational changes in the Switch loops across the GDP and GTP states and the

respective contacts between the ɣ-phosphate and the conserved Thr and Gly residues are highlighted (Reproduced from Bulyha *et al.*, 2011[57]).
C. GTPase cycle of small Ras-like GTPases and the regulation by GAPs, GEFs, GDIs and GDFs. The activity of GEFs and GAPs are under feedback loops (by the broken line), where the GTPase regulates the GEF, and feed-forward signalling (dotted lines), where GEFs select their respective effectors (shown as a stack, in light orange), which binds to the active GTPase. Effectors can activate specific GAPs through feed-forward signalling (violet) (Reproduced from Cherfils and Zeghouf, 2013[58]).
D. Topology diagram of MglA, a small Ras-like GTPase. The β-strands and α helices are shown in pink and grey, respectively. G1 (orange), G2 (navy blue), G4 (purple), G5 (maroon), Switch I (cyan), and Switch II (green) are labelled.
E. Structure of MglA GTP (PDB ID 6IZW) with the secondary structures labelled in black. The conserved motifs in the G-domain are highlighted as per panel D.

1.5 Types of motilities in *Myxococcus xanthus*

To facilitate the different motility patterns characterizing the *Myxococcus* social behaviour across different stages of its complex life cycle, it exhibits two different motility patterns. Unlike flagellated bacteria with movement speeds at the rates of μm s^{-1}, the overall speed of movement for these cells is low at rates of μm min^{-1}. The bacteria can move in groups demonstrating social or S-motility or twitching motility powered by the Type-IV pili (T4P) at the leading cell pole[30,32,35,39,59–61]. The bacteria also exhibit adventurous motility or A-motility or gliding motility of individual cells away from the coordinated swarms (Fig 1.2 A). This is driven by the presence of motility complexes that attach the cell to the substratum and drive movement, much like the eukaryotic focal adhesion complexes. It is important to note that these two types of motilities are encoded by separate sets of genes, encoding separate motility machinery with distinct organizations and mechanisms.

1.5.1 T4P dependent S-motility

The T4P is assembled at the leading cell pole towards the direction of cell movement. It extends and attaches to the substratum or the surfaces of adjacent cells in the collective aggregates[62]. Consequently, it retracts, thereby pulling the cell forward, generating a twitching movement towards the site of attachment. The force generated by a single pilus can be up to 150 pN and hence can be concluded as a robust motor driving movement. Mutagenesis studies revealed a cluster of genes which encode the constituents of the T4P and power the extension and retraction[63–65]. The cluster encoding the T4P genes is shown in Fig 1.6 A with their corresponding functions. Some of these remain stationary, whereas others are dynamic and

relocate across the poles during cellular reversals. PilB and PilT are ATPases which exhibit dynamic polar localization across leading and lagging poles, respectively[36,66] (Fig 1.6 C). Their coordinated action results in the disassembly of the pili at the old leading pole and reassembly at the opposite pole.

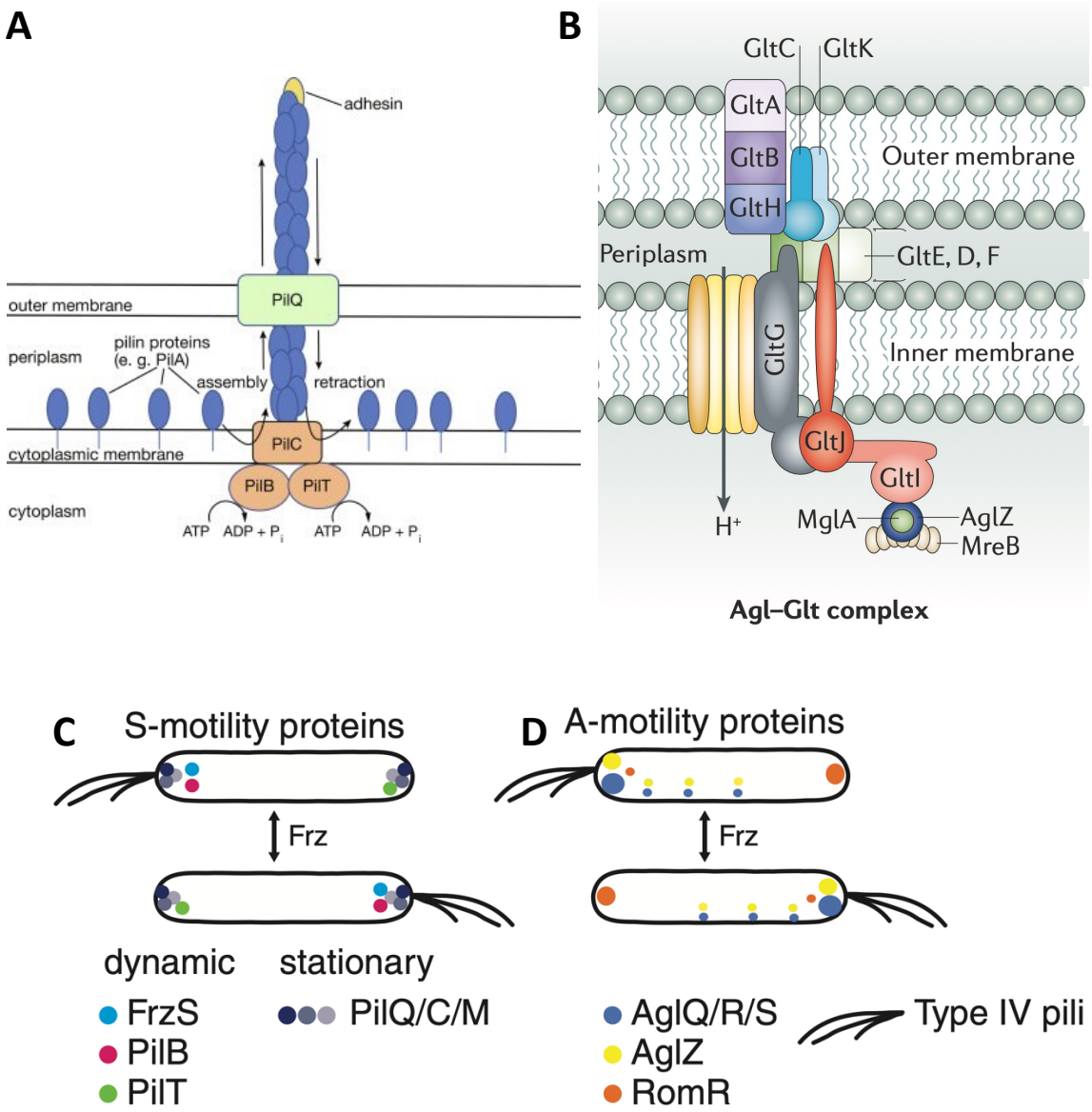


Figure 1.6. S and A-motility apparatus in *M. xanthus*
A. Simplified model of the type IV pilus machinery spanning the bacterial cell envelope. The pilus is constituted of PilA (blue) and extends through an outer membrane channel comprised of PilQ proteins. Adhesins (yellow) are present at the distal tip. An ATPase at the base pulls PilA monomers from the inner membrane and inserts them into the base of the pilus. ATP hydrolysis causes disassembly of the pilus, which causes retraction of the pilus and recycles the PilA proteins to the cytoplasmic membrane. Assembly and disassembly are driven by two ATPases (PilB and PilT, respectively) that function together with PilC (Reproduced from McBride, 2019[59]).

B. The architecture of the Agl/Glt complex driving A-motility. The localization of AglQ, AglR, AglZ, GltD, GltF, and MglA has been determined by fluorescence microscopy. The localization of GltD, GltE, GltG, and GltH has been determined by fractionation experiments. Direct protein interactions have only been shown to occur between AglR and GltG and AglZ and MreB. The Agl motor initiates a conformational change of the GltG (causing a power stroke), which would be further transduced to the cell surface by cognate Glt proteins. MglA is sandwiched between the GltI and AglZ, communicating further with MreB to drive cell motility. (Reproduced from Agrebi *et al.*, 2015[67]).
C, D. The dynamically localized S and A-motility proteins, respectively, during cellular reversals induced by the Frz chemosensory system. The stationary proteins do not relocate during reversals. The respective colour representing the localized proteins are labelled below (Reproduced from Bulyha *et al.*, 2011[57]).

1.5.2 Agl-Glt complex driven A-motility

AglZ was the first identified gliding motility protein in *M. xanthus*. It was observed that gliding motility is driven by periodically positioned clusters of molecular motors which attached the cell to the substratum. They were similar in function to the eukaryotic focal adhesion complexes[68–71]. These complexes were static with respect to the substratum and were assembled towards the leading cell pole and disassembled as they approached the lagging pole of the gliding cell (Fig 1.6 D). Further, the AglRQS complex was identified to constitute the gliding motor, powering the propulsion of the cell through proton motive force. Further, a host of other proteins were identified, which spanned across the cytoplasm to the outer cell membrane to generate forces to propel the cell forward when coupled to the AglRQS motors[72] (Fig 1.6 B). These were named the Gliding transducers or Glt proteins (as shown in Fig 1.6 B), and they interact with AglR.

The periodic arrangement of gliding motors correlates with their interaction with MreB, an actin-like filament-forming cytoskeletal protein[71,73,74]. The strong interaction of the focal adhesion complexes with the substratum generates force to propel the bacteria. The adhesins on the outer-membrane region of these complexes attach strongly to the substrates. Further, these complexes trace the path of MreB in a counter-clockwise fashion, thereby propelling the cell forward in a clockwise helical trail[75]. Such a phenomenon generates drag forces transversely towards the lagging cell pole. This causes a resultant push on the substrate behind and a forward thrust to the cell.

1.6 Spatial regulation of motility

M. xanthus needs to strictly define its direction of motility under contrasting scenarios concerning aggregation to form fruiting bodies or foraging for nutrients or predation. This is regulated by modulating the frequency of the cell polarity oscillations or polarity reversal events. The cell polarity oscillation is powered by a GTPase, MglA, which localizes dynamically to the leading cell pole and plays a central role in specifying the front-rear cell polarity and its regulation. The MglA oscillator is coupled to the Frizilator, making the system sensitive to extra-cellular cues. The output of this coupled oscillator, via MglA, further feeds onto downstream effectors resulting in cell motility.

1.6.1 Frizilator, the chemosensory module

The Frizilator comprising the Frz proteins is the chemosensory module showing homology to the chemosensory module consisting of the Che proteins in other bacteria[31,39]. The function of chemoreception is performed by FrzCD, a methyl-accepting chemosensory protein, with assistance from FrzE, a histidine kinase[61]. Upon activation from FrzCD, FrzE transfers the phosphoryl switch across its N-terminal kinase and C-terminal response regulator domains downstream to FrzZ and FrzX (Fig 1.7). The autophosphorylation of FrzE inhibits its kinase domain, thereby preventing the constitutive phosphorylation of downstream proteins. FrzZ consists of two response regulator domains over one in FrzX, which communicates to the central polarity module proteins[76]. FrzZ mirrors MglA localization at the leading cell pole, and FrzX is present at the lagging pole. A concerted action of these regulators transmits the signal to the polar localized proteins initiating the reversal switch (Fig 1.8A). However, the exact mechanism of how these proteins trigger reversal remains to be investigated.

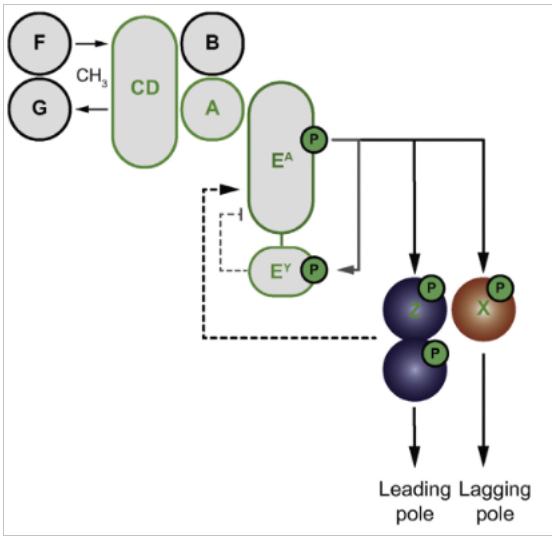


Figure 1.7. Chemosensory apparatus in *M. xanthus*

Schematic representation of the polarity inversion module comprising methyl-accepting chemotaxis protein (MCP) FrzCD, the CheA-like histidine kinase FrzE with an N-terminal kinase domain (FrzEA) and a C-terminal receiver domain FrzEY, the CheW-like coupling proteins FrzA and FrzB, FrzF and FrzG with the response regulators FrzZ and FrzX. The core Frz proteins are indicated in green. The phosphotransfer from $FrzE^{A}$ under low signalling is shown in grey, and the phosphotransfer at higher signalling levels is in black lines. The feedback on $FrzE^{A}$ by $FrzE^{Y}$~P is in dashed grey and by FrzZ~P in dashed black (Reproduced from Carreira *et al*., 2022[77])

1.6.2 The Mgl/Rom polarity reversal module

Mutagenesis screens revealed the identity of two proteins critical to both the S and A motility of *M. xanthus*. These were named mutual gliding or Mgl proteins. Deleting either MglA or MglB affected the motility of the cells, showing that these genes were critical for driving motility [38,42]. Further, there was a regulator of motility response regulator protein, namely RomR, that was identified to be an essential gene to generate polarity reversals[20,78]. Consequently, new components of the polarity module, namely MglC, RomX and RomY, were identified. All these proteins show unique patterns of polar localization, which are dynamic and relocate as the cell reverses polarity (Fig 1.8 B).

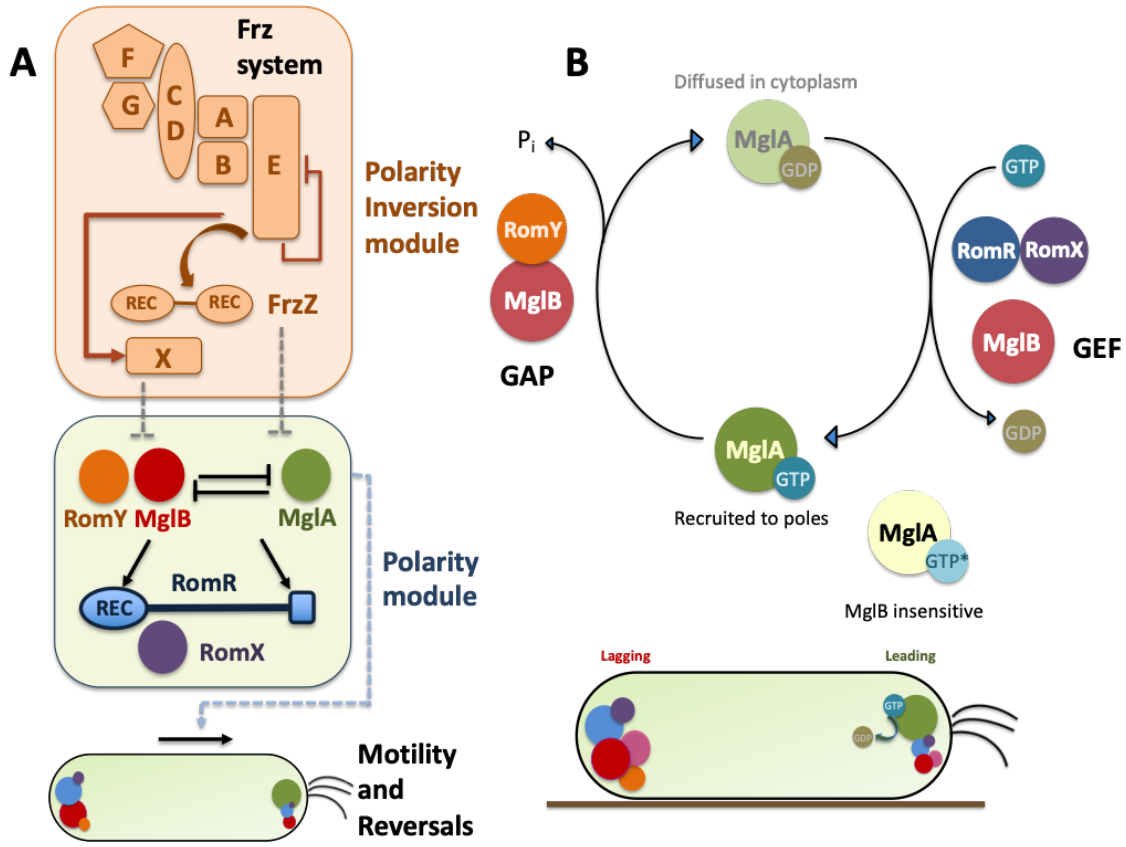


Figure 1.8. Cascade of proteins driving polarity reversal
A. Schematic showing the polarity inversion or the Frz chemosensory module showing the respective Frz proteins. It communicates with the central polarity module consisting of proteins defining cell polarity and driving cell polarity oscillations. MglA, the master regulator, activates downstream events driving motility and reversals (Adapted from Keilberg *et al*., 2013[32]).
B. GTPase cycle of MglA and its regulators from the polarity module. The lower panel shows the localization of these proteins at the poles of the polarized cell.

1.6.2 A. MglA, a prokaryotic small Ras-like GTPase

MglA lies at the centre of the polarity reversal module, acting as the master regulator driving polarity reversals (Fig 1.8 A). Similar to other small Ras-like GTPases, it switches between the active GTP-bound and inactive GDP-bound states. This is associated with conformational change at regions near the nucleotide-binding pocket (also referred to as the Switch regions), which affects its localization and interaction with other proteins driving motility and reversals[40,79,80]. The active GTP bound state localizes to the leading cell pole and relocates to the other pole when the reversal switch is initiated. When the bound GTP is hydrolysed, the protein falls off from the pole in the GDP-bound state and diffuses in the cytoplasm. MglA was also crystallized in a mixed state where Switch I was in the inactive state and Switch II in the

active state[80]. Such a conformation was termed as MglA-GTP* state, and this species of MglA was proposed to dominate at the leading pole. In the active state, it interacts with SgmX, an effector and also plays a role in the correct polar localization of PilB and PilT ATPases, all of which are critical to S-motility[81]. It also interacts with AglZ in the GTP state, localizing to the gliding motors, manifesting the cell movement[6,72,79]. Like other small-Ras like GTPases, MglA has a very slow intrinsic GTP hydrolysis rate. This results in the requirement of a GTPase activating protein (GAP) which can assist GTP hydrolysis, thereby causing the GTPase to Switch from its "ON" GTP state to the "OFF" GDP state (Fig 1.8 B). Further, MglA has a high affinity to GDP over GTP. This necessitates the requirement of a GEF that assists MglA in losing its affinity towards GDP, which consequently helps GTP binding. Hence GEFs are essential to reactivate MglA from the "OFF" GDP state back to its "ON" GTP state. The prospective regulators of MglA will be discussed further in the subsequent sections.

1.6.2 B. MglB, a dual GAP and GEF for MglA

Along with MglA, another protein, MglB, was identified to be critical for regulating cell polarity reversals in *M. xanthus*[8,79] (Fig 1.8 A). It is present in the same operon upstream of MglA in the *M. xanthus* genome. MglB is localized in a bipolar asymmetric fashion in polarized cells, with a bigger cluster at the lagging cell pole (Fig 1.8 B). Deletion of MglB retains cell motility but causes cells to reverse with a higher frequency causing a loss of directionality[9,35,45,82–85]. It was observed to act as the cognate GAP (GTPase activating protein) for MglA and accelerates the intrinsic slow GTP hydrolysis activity of MglA. MglB consists of a Roadblock domain, a domain commonly associated with regulators of eukaryotic small Ras-like GTPases[40,53,80]. In addition, MglB consists of a C-terminal extension adorned with acidic residues. Biochemical assays showed that the presence of the C-terminal region further accelerated the overall GTP hydrolysis rates of MglA[40].

Structural studies of MglA complexed with MglB revealed a complex where one molecule of MglA interacted with two molecules of MglB (Fig 1.9 A)[40,53,80]. Initial structural studies of MglA/MglB from *Thermus thermophilus* revealed that the Roadblock domain of MglB interacted with MglA in the GTP-bound active form[53]. When MglA switched from its GDP-bound inactive state to the GTP-bound active state, there was an associated conformational change, similar to other small Ras-like GTPases. In addition, there was a flipping of the β2 strand of MglA along with a registry shift of 2 amino acids. This caused a screw-like movement (rotation + translation by 2 residues) of the β2 strand, which reoriented the constitutive

residues. This resulted in the formation of the MglB binding interface (Fig 1.9 B). This shift in the β-strand, also known as the 'β-screw movement' resulted in the transition of the MglA nucleotide binding pocket from a closed to an open conformation along with the repositioning of catalytic residues driving GTP hydrolysis[53]. Further, it was observed that the C-terminal tail of MglB assisted this conformational change causing the strongly bound GDP to dissociate and consequently favour GTP binding[40]. Hence it was concluded that the C-terminal tail of MglB assisted nucleotide exchange, acting as a GEF for MglA. The mechanism of this dual activity needs to be further elucidated.

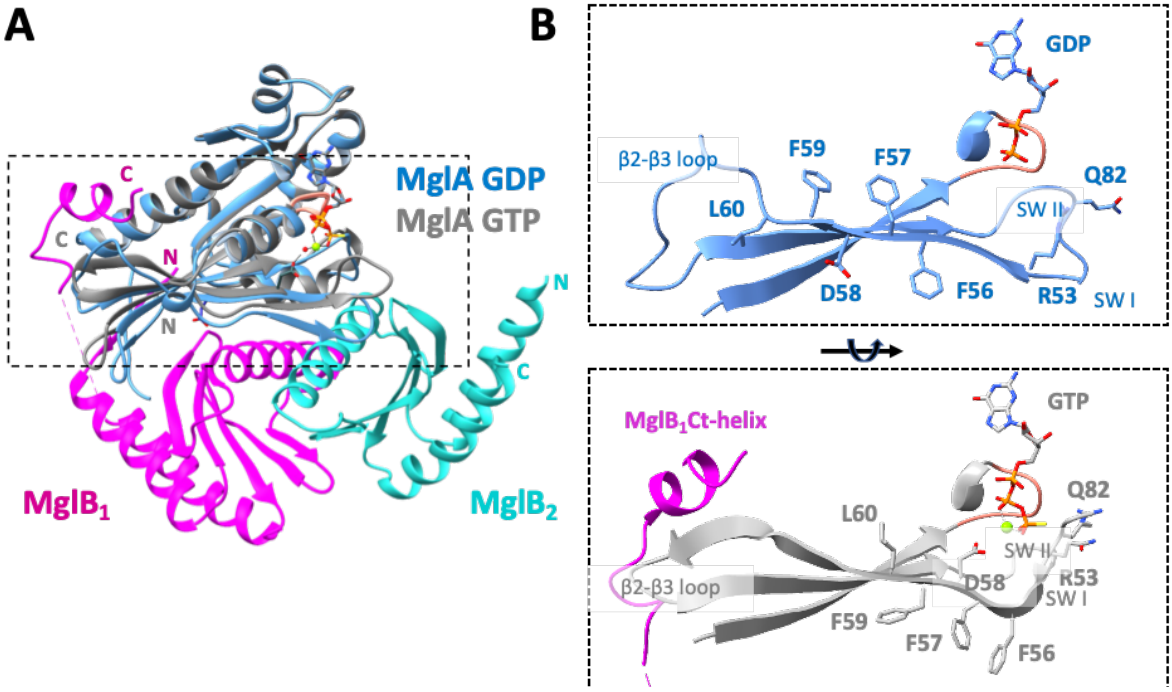


Figure 1.9. Interface between MglA and MglB
A. Crystal structure of MglA in complex with MglB dimer (PDB ID 6IZW[40]) captured bound to a GTP analog. MglA-GTP state (grey) is superposed with the GDP state (PDB ID 5YMX) (blue). N and C termini of each chain are labelled.
B. MglA/MglB interface, as highlighted by the dashed lines in A. The β-screw movement is characterized by the flipping of the β2 strand and the reorientation of residues at the active site and the interface. These residues are highlighted in the MglA GDP structure in the top panel and the MglA-MglB complex bound to GTP in the lower panel.

1.6.2 C. RomR/RomX, the GEF module

RomR was identified as a regulator of motility response regulator protein essential for driving A-motility in *M. xanthus*[20,78,86]. Like MglB, it localizes in a bipolar asymmetric fashion with a bigger cluster at the lagging cell pole (Fig 1.8 A, B). The pole-to-pole oscillation of RomR was

dependent on MglA. It acts downstream of the Frz pathway regulating A-motility, probably by regulating the other candidates of the polarity module. RomR is a multi-domain protein with an N-terminal response receiver domain (1-116 residues) and a C-terminal output domain (117-420). The output domain consists of a long-disordered region followed by a conserved C-terminal tail. Upon deletion of the response receiver domain, the cellular reversals were impaired and showed that it is essential that RomR relocalisation during reversals. The C-terminal output domain was necessary for the correct polar localization of RomR[86].

Recently, another novel protein, RomX, was identified, mirroring the genomic distribution of RomR[77,87]. It also localized majorly at the lagging pole in a bipolar asymmetric distribution like RomR (Fig 1.8 A, B). Deletion of RomX affected A-motility similar to RomR deletion. RomX, together with RomR, was shown to interact with MglA in the GTP-bound state. However, RomX was shown to exert a GEF activity on MglA, which was further accelerated in the presence of RomR. Together with RomR, RomX was shown to act as an active recruiter of MglA to the leading cell pole. It also co-localizes with MglA-GTP at the gliding motility complexes driving A-motility.

1.6.2 C. Other interactors of MglB

Recently other partners of MglB were identified as a part of the polarity module. A homolog of MglB, named MglC, was identified, which localizes dynamically to the lagging cell pole[88–90](Fig 1.8 B). MglC deletion retained motility but caused the cells to be hypo-reverse compared to the wild-type cells. MglC, like MglB, was a constitutive dimer mediated through the extension of their Roadblock domains. It was observed that MglC interacted with MglB and RomR in the polarity module. MglC was also found to be important for the polar localization of MglA and MglB. However, the mechanism of regulation of polarity reversal frequency by MglC is unclear.

Recently, RomY was identified as an uncharacterized protein in the *Myxococcus* genome which generally co-occurs with the Rom proteins. It was seen to localize to the lagging pole in an MglB-dependent fashion[91]. Deletion of RomY had a similar effect as MglB deletion on motility, causing the cells to hyper-reverse, causing no net directional movement. Biochemically, it was observed to stimulate MglB GAP activity, further affecting the localization of MglA-GTP (Fig 1.8 B).

1.6.3 Downstream effector proteins

As mentioned earlier, MglA in the GTP-bound active state interacts with both S and A-motility machinery to orchestrate the movement of the cell in the direction of the leading pole of the cell. The T4P assembly at the leading pole is stimulated by the interaction of MglA with SgmX (Social Gliding Mutant X)[81]. MglA-GTP interacts with the C-terminus TPR domain (12-14 tetratricopeptide repeats) of SgmX (Fig 1.10 B). This results in a conformational change in SgmX, which is essential for its localization at the leading pole, thereby activating the T4P machinery.

The presence of MglA-GTP at the leading pole stimulates the correct localization of PilB and PilT ATPases, powering the T4P machinery. These ATPases are recruited to the poles by a MglA paralog, SofG (Social motility function GTPase)[57,92,93]. It contains N and C-terminal extensions over the Ras domain. SofG directly interacts with bactofilin BacP, a polymeric bacterial cytoskeletal protein forming large clusters at subpolar regions (Fig 1.10 A). Interestingly, it was observed that MglB also acted as a GAP for SofG[93]. It is reported that SofG shuttling over BacP recruits PilB and PilT towards the lagging cell pole and further interaction with MglA-GTP sorts PilB and PilT to lagging and leading poles, respectively. But the mechanism and the exact order of events to this phenomenon still need to be clarified.

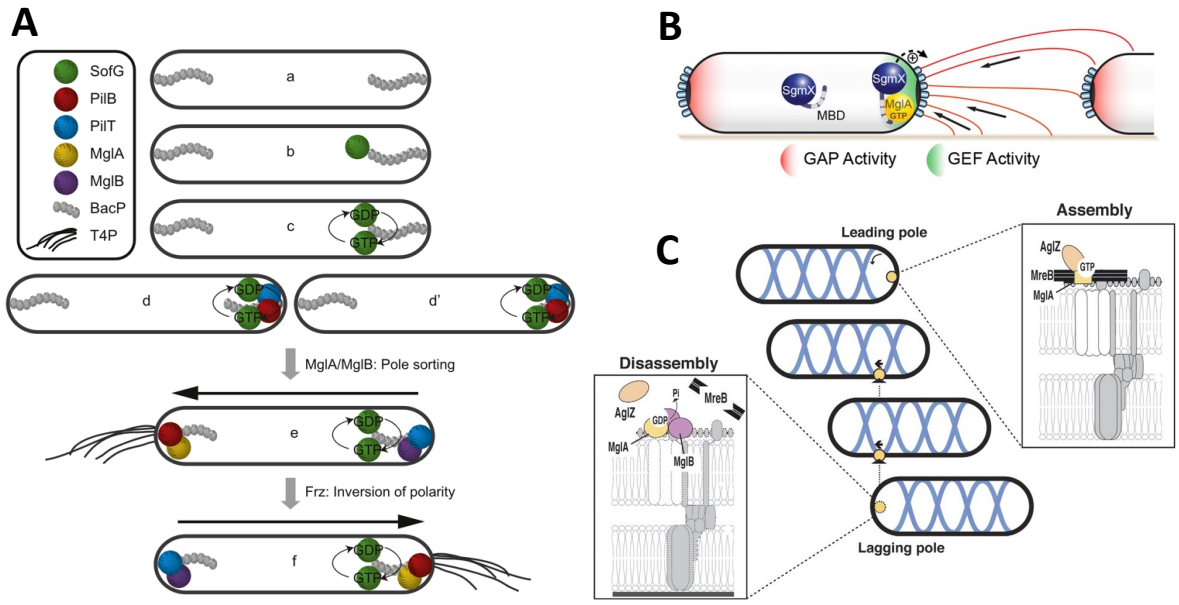


Figure 1.10. Interaction of MglA with downstream effectors
A. Schematic showing the model for regulation of dynamic polarity of S-motility proteins, PilB and PilT by SofG and MglA. The respective proteins are shown with the colour key on the left. Subpolarly localized SofG recruits PilB and PilT to the pole and MglA consequently sorts them to the leading and lagging poles, respectively (Reproduced from Bulyha *et al*., 2013[92]).
B. Model for T4P activation by MglA interaction with SgmX. MglA GTP is localized at the leading pole (yellow) (due to the GEF activities of MglB and RomR/X (GEF Activity, Green; GAP Activity, Red). The SgmX C-terminal TPR Domain (MglA Binding Domain) inhibits SgmX polar localization, which, when interacts with MglA-GTP, reveals a polar binding site that recruits SgmX to the pole activating T4P assembly (Reproduced from Mercier *et al.,* 2020[81]).
C. The interaction between MglA-GTP with AglZ and MreB drives gliding motility. The active complex translocates to the lagging pole where MglB dissociates MglA from the MreB cytoskeleton due to GAP activity, disassembling the motility complex. The MreB tracks are shown in blue. (Reproduced from Treuner-Lange *et al.*, 2015[74]).

To drive A-motility, MglA-GTP interacts with the Agl/Glt motility complexes, specifically interacting with AglZ at the cytoplasmic side of the complex. MglA-GTP further contacts the cytoskeletal protein, MreB, thereby transmitting the forces from the focal adhesion complexes to drive the cell movement (Fig 1.10 C). Due to the abundance of MglA-GTP at the leading pole, the Agl/Glt complexes are recruited at this pole. As the complexes approach the lagging pole, the GAP activity of MglB flips the MglA switch to the GDP state, causing disassembly of the complexes[74]. Hence, the concerted activity of MglA and MglB is critical for establishing front-rear polarity and regulating the frequency of reversals.

1.7 Current model of polarity reversal

The extracellular signal acting as a cue for nutrient availability causes methylation of the MCP, FrzCD, which in turn phosphorylates response regulator FrzA and activates histidine kinase FrzE. Upon activation, FrzE autophosphorylates and consecutively activates the two response regulators, FrzZ and FrzX, which interface with the polarity module. FrzE autophosphorylation also causes its auto-inhibition, thereby acting as a phosphate sink preventing constitutive activation of the Frz signalling. Phosphorylated FrzZ ($FrzZ^P$) further activates upstream FrzE kinase, amplifying the FrzE activity.[31,39]

$FrzZ^P$ localizes at the leading cell pole and phosphorylated FrzX ($FrzX^P$) is present at the lagging pole[76] (Fig 1.11 A). Wild-type cells reverse every 10-15 minutes, with each reversal event lasting 30-60 seconds until the new polarity is established. It was observed that MglA was the first protein to relocalise during reversal and is inverted within 45 seconds [74,88]. This causes an intermediate co-localization of MglA and MglB at the old lagging pole till MglB

completes its re-localization by 90 seconds. RomR takes the longest time of around 160 seconds to reverse (Fig 1.11 B).

Based on these timescales, a model was proposed for the sequence of events regulating a reversal (Fig 1.11 A)[76]. According to this model, $FrzX^P$ localizes at the lagging pole in an MglB-dependent manner, which subsequently displaces MglB from this pole. $FrzZ^P$ localizes to the leading pole and causes unbinding of MglA from this pole by some unknown mechanism. At the lagging pole, when MglB is displaced, the RomR-RomX activity dominates to recruit MglA-GTP to this pole. This is facilitated by the slow dynamics of RomR, which further helps to introduce a refractory period when the module is insensitive to further Frz signalling[6,76,82].

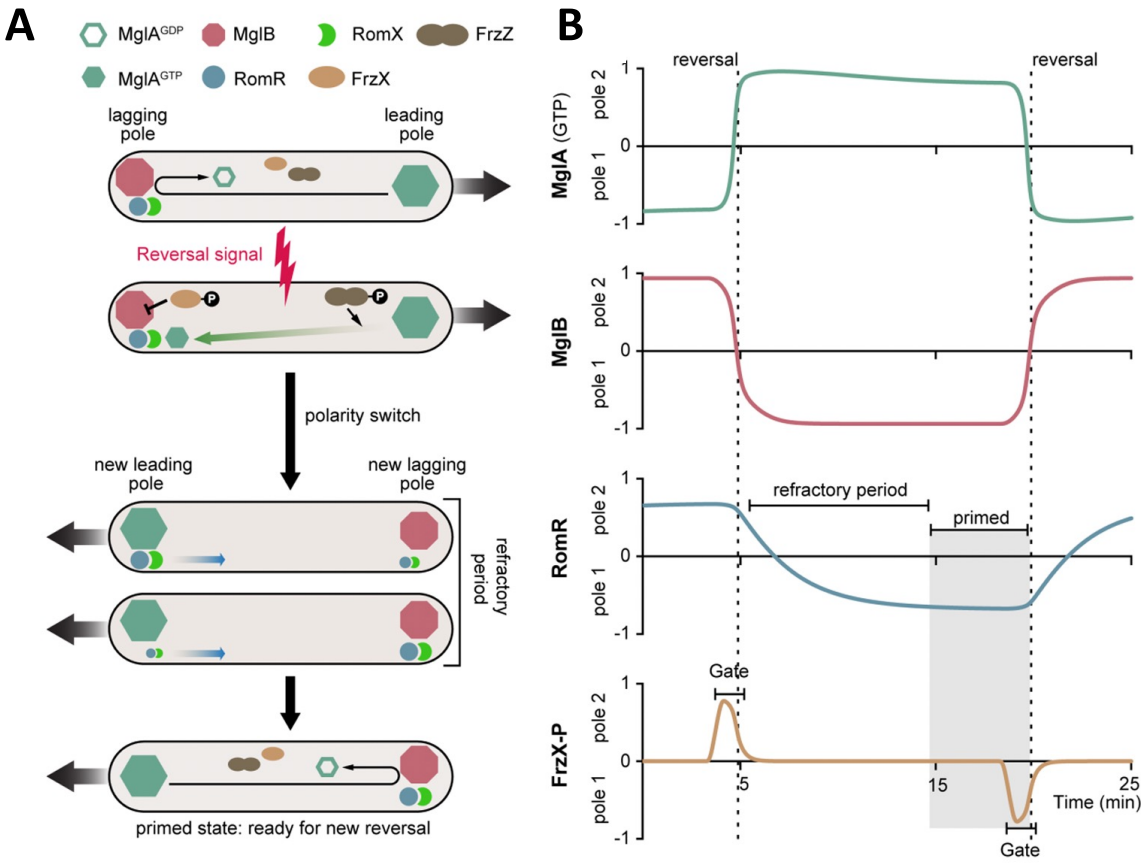
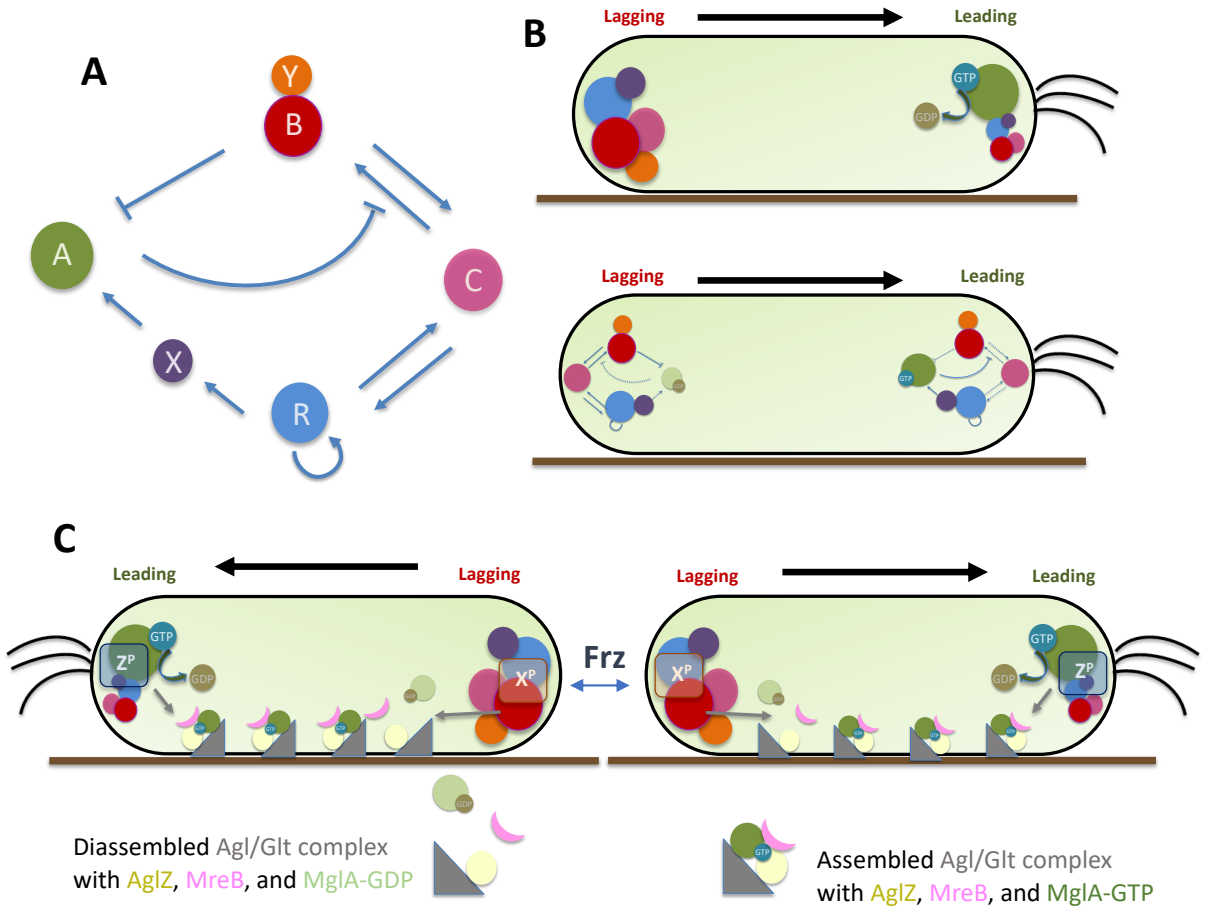


Figure 1.11. Gated relaxation oscillator model
A. Order of events during reversal as per initial model. MglA-GTP is present at the leading pole and MglB, RomRX is present at the lagging pole. Under signalling from FrzE, FrzX~P accumulates at the lagging pole. RomRX recruits MglA to the new leading pole and consequently localizes at the lagging pole. RomR and FrzX~P regulate the refractory period before the next reversal is initiated (Reproduced from Herrou and Mignot, 2020[6]).

B. Simulated profiles of MglA, MglB, RomR and FrzX~P dynamics during the reversal cycle (Reproduced from Guzzo *et al.*, 2020[76]).

The more recent model takes into account the GAP and GEF activities of MglB/RomY and RomR/RomX, respectively, to elucidate the sequence of events[77,87,90,91] (Fig 1.12 A, B, C). At the lagging cell pole, MglB, RomY, RomR, and RomX, are all present abundantly. However, the MglB/RomY GAP activity dominates over the RomRX GEF activity (Fig 1.12 B). This displaces MglA-GTP from the lagging cell pole and is recycled to the cytoplasmic MglA-GDP pool. Simultaneously, MglA-GTP is recruited at the leading pole owing to the GEF activity of the RomR/RomX complex. Due to the absence of RomY at this pole, MglB GAP activity is overshadowed. MglA-GTP* population is also proposed to accumulate at this pole which was shown to be insensitive to the GAP stimulation of MglB. During reversals, FrzX^{P} localization at the lagging pole probably inhibits the MglB GAP activity, causing the RomR/RomX GEF to dominate (Fig 1.12 C). This results in the loss of MglB and recruitment of MglA-GTP at this pole. Consequently, at the other pole, there is recruitment and activation of MglB/RomY due to the presence of FrzZ^{P}. The GAP activation displaces the MglA-GTP from this new lagging pole. This activates the MglB/RomR positive feedback loop at this pole with further contribution from MglC (Fig 1.12 A). This leads to the establishment of a reversed front-rear polarity axis and marks the completion of the reversal event.

Figure 1.12. Current model of *M. xanthus* polarity reversal
A. Schematic of the regulatory circuit among the polarity module proteins (A, B, C denote MglA, MglB, MglC, R, X, Y denote RomR, RomX, RomY) (Adapted from Carreira *et al.*, 2022).
B. Localization and regulatory circuits (top and bottom, respectively), operating at the leading and lagging poles of the reversing cell (Adapted from Carreira *et al.*, 2022[77]).
C. Localization of the FrzZ and FrzX from the chemosensory module and effectors driving polarity reversals. The Agl/Glt complexes are shown at the base of the cell (Adapted from Carreira *et al.*, 2022[77]).



1.8 Motivation behind the study

Certain aspects of the current model of polarity reversal of *M. xanthus* require further validation. Several of these proteins, such as RomR, RomX, MglC, RomY and MglB, co-localize in the cell, especially at the lagging cell pole. However, the exact mechanisms driving their localization or regulatory functions are unknown. Despite reports of the GEF activity of MglB[40], it is not accounted for in any of the aforementioned models. Hence, the claim for RomR-RomX as the sole active GEF module, is debated. It is not entirely clear why RomR-RomX GEF activity is solely relevant at the leading pole, where it is present in lesser concentrations as compared to the lagging pole. Similarly, for the GAP activity leading to MglA-GTP exclusion from the lagging pole, there is a requirement for a simultaneous GEF activity. Moreover, RomR-RomX has been shown to form a complex with MglA-GTP and not with MglA-GDP, quite unlike the properties of a GEF, which has a higher affinity for the GTPase in the GDP-bound state. All these questions pose the need for biochemical and structural studies to decipher the nature of these polar localized proteins and how these fit into

the polarity reversal modules. Such a systematic enquiry will validate the order of events driving the cell polarity oscillations and their role in the motility of *M. xanthus*. Moreover, it could reflect unique regulatory mechanisms prevalent in the ancestral prokaryotic GTPase MglA and how they evolved to function in their eukaryotic counterparts.

This thesis majorly focuses on the central module generating cell polarity oscillations in *M. xanthus*. Starting from MglA, the master regulator, we have investigated the conformational changes in the active site as it switches across GDP and GTP states and how it could be critical for its regulatory functions and also its prospective binding with other interactors (GEFs, GAPs and effectors). Further, MglB which acts as a dual GAP and GEF interacts with MglA and regulates its localization and dynamics. Hence, we have attempted to dissect the GAP and GEF activity of MglB and proposed a model of how it could play a role in determining the nucleotide-bound state of MglA. We have initiated the investigation on newly identified MglB interactors, RomY and MglC, paving the way for future validation on how these proteins functionally fit into the polarity module.

The other module comprising RomR and its partner RomX occupies a central position in the polarity module, apparently competing with the MglB-driven GAP module to regulate MglA localization. However, there needs to be knowledge of how RomRX interacts and how their GEF function is regulated. We have attempted to structurally and biochemically characterize RomX, a protein of novel fold. Further, we have performed a domain-by-domain dissection of RomR to determine each of their function. Consequently, we have attempted to capture the interaction of RomR with RomX which throws light on their probable mechanism of MglA regulation. With all these results, we have proposed an updated model of how the polarity module proteins orchestrate the cell polarity oscillations.

References

1. Cheong, R. & Levchenko, A. Oscillatory signaling processes: The how, the why and the where. *Curr Opin Gen Dev.* vol. 20 665–669 Preprint at https://doi.org/10.1016/j.gde.2010.08.007 (2010).
2. Nguyen, L. K. & Kholodenko, B. N. Feedback regulation in cell signalling: Lessons for cancer therapeutics. *Sem Cell Dev* vol. 50 85–94 https://doi.org/10.1016/j.semcdb.2015.09.024 (2016).
3. Raskin, D. M. & de Boer, P. A. J. Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli MATERIALS AND METHODS. *Cell Biol.* vol. 96. (1999).

4. Meinhardt, H. & de Boer, P. A. J. Pattern formation in Escherichia coli: A model for the pole-to-pole oscillations of Min proteins and the localization of the division site. *PNAS December* vol. 4 doi10.1073 pnas.251216598 (2001).
5. Ebersbach G, Ringgaard S, Møller-Jensen J, Wang Q, Sherratt DJ, Gerdes K. Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. *Mol Microbiol*. 2006 Sep;61(6):1428-42. doi: 10.1111/j.1365-2958.2006.05322.x.
6. Herrou, J. & Mignot, T. Dynamic polarity control by a tunable protein oscillator in bacteria. *Curr Opin Cell Biol* **62**,.
7. Lenz, P. & Søgaard-Andersen, L. Temporal and spatial oscillations in bacteria. *Nat Rev Microbiol* vol. 9 565–577 Preprint at https://doi.org/10.1038/nrmicro2612 (2011).
8. Mignot, T., Merlie, J. P. & Zusman, D. R. Regulated pole-to-pole oscillations of a bacterial gliding motility protein. *Science (1979)* **310**, 855–857 (2005).
9. Keilberg, D. & Søgaard-Andersen, L. Regulation of bacterial cell polarity by small GTPases. *Biochemistry* **53**, 1899–1907 (2014).
10. Bulyha, I., Hot, E., Huntley, S. & Søgaard-Andersen, L. GTPases in bacterial cell polarity and signalling. *COMICR* vol. 14 726–733 https://doi.org/10.1016/j.mib.2011.09.001 (2011).
11. Goldbeter, A. Oscillations and waves of cyclic AMP in *Dictyostelium*: A prototype for spatiotemporal organization and pulsatile intercellular communication. *Bull Math Biol* **68**, 1095–1109 (2006).
12. Golden, S. S. & Canales, S. R. Cyanobacterial circadian clocks — timing is everything. *Nat Rev Microbiol* **1**, 191–199 (2003).
13. Maini, P. K., Woolley, T. E., Baker, R. E., Gaffney, E. A. & Lee, S. S. Turing's model for biological pattern formation and the robustness problem. *Interface Focus* **2**, 487–496 (2012).
14. Turing, A. M. The chemical basis of morphogenesis. *Bull Math Biol* **52**, 153–197 (1990).
15. Novák, B. & Tyson, J. J. Design principles of biochemical oscillators. *Nat Rev Mol Cell Biol* vol. 9 981–991 Preprint at https://doi.org/10.1038/nrm2530 (2008).
16. Bell-Pedersen D, Cassone VM, Earnest DJ, Golden SS, Hardin PE, Thomas TL, Zoran MJ. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nat Rev Genet.* 2005 Jul;6(7):544-56. doi: 10.1038/nrg1633.
17. Meinhardt, H. & de Boer, P. A. Pattern formation in Escherichia coli: a model for the pole-to-pole oscillations of Min proteins and the localization of the division site. *PNAS* **98**, 14202–7 (2001).
18. Cohen, S. E. & Golden, S. S. Circadian Rhythms in Cyanobacteria. *MMBR* **79**, 373–385 (2015).
19. Egli, M. Intricate protein-protein interactions in the cyanobacterial circadian clock. *JBC* **289**, 21267–21275 (2014).
20. Leonardy, S., Bulyha, I. & Søgaard-Andersen, L. Reversing cells and oscillating motility proteins. *Mol Biosyst* **4**, 1009 (2008).
21. Wang, X., Montero Llopis, P. & Rudner, D. Z. Bacillus subtilis chromosome organization oscillates between two distinct patterns. *PNAS* **111**, 12877–12882 (2014).
22. Sgro AE, Schwab DJ, Noorbakhsh J, Mestler T, Mehta P, Gregor T. From intracellular signaling to population oscillations: bridging size- and time-scales in collective behavior. *Mol Syst Biol*. 2015 Jan 23;11(1):779. doi: 10.15252/msb.20145352.
23. Schäfer E, Tarantola M, Polo E, Westendorf C, Oikawa N, Bodenschatz E, Geil B, Janshoff A. Chemotaxis of *Dictyostelium discoideum*: collective oscillation of cellular contacts. *PLoS One*. 2013;8(1):e54172. doi: 10.1371/journal.pone.0054172.
24. Ebersbach, G. & Gerdes, K. Bacterial mitosis: Partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Mol Microbiol* **52**, 385–398 (2004).

25. Park KT, Wu W, Battaile KP, Lovell S, Holyoak T, Lutkenhaus J. The Min oscillator uses MinD-dependent conformational changes in MinE to spatially regulate cytokinesis. *Cell*. 2011 Aug 5;146(3):396-407. doi: 10.1016/j.cell.2011.06.042.
26. Cao, Y., Lopatkin, A. & You, L. Elements of biological oscillations in time and space. *Nat Struct Mol Biol* vol. 23 1030–1034 Preprint at https://doi.org/10.1038/nsmb.3320 (2016).
27. Lutkenhaus, J. & Sundaramoorthy, M. MinD and role of the deviant Walker A motif, dimerization and membrane binding in oscillation. *Mol Microbiol* **48**, 295–303 (2003).
28. Tsai TY, Choi YS, Ma W, Pomerening JR, Tang C, Ferrell JE Jr. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science*. 2008 Jul 4;321(5885):126-9. doi: 10.1126/science.1156951.
29. Vos, M. *Myxococcus xanthus*. *Trends Microbiol* vol. 29 562–563 https://doi.org/10.1016/j.tim.2021.03.006 (2021).
30. Mauriello EM, Zusman DR. Polarity of motility systems in *Myxococcus xanthus*. *Curr Opin Microbiol*. 2007 Dec;10(6):624-9. doi: 10.1016/j.mib.2007.09.012.
31. Zusman, D. R., Scott, A. E., Yang, Z. & Kirby, J. R. Chemosensory pathways, motility and development in *Myxococcus xanthus*. *Nat Rev Microbiol* vol. 5 862–872 Preprint at https://doi.org/10.1038/nrmicro1770 (2007).
32. Keilberg, D. Regulation of motility and polarity in *Myxococcus xanthus*. *Annu. Rev. Microbiol* 145 (2013).
33. Zhang, Y., Ducret, A., Shaevitz, J. & Mignot, T. From individual cell motility to collective behaviors: Insights from a prokaryote, *Myxococcus xanthus*. *FEMS Microbiol Rev* vol. 36 149–164 Preprint at https://doi.org/10.1111/j.1574-6976.2011.00307.x (2012).
34. Kaiser, D. & Warrick, H. Transmission of a signal that synchronizes cell movements in swarms of *Myxococcus xanthus*. *PNAS* **111**, 13105–13110 (2014).
35. Islam, S. T. & Mignot, T. The mysterious nature of bacterial surface (gliding) motility: A focal adhesion-based mechanism in *Myxococcus xanthus*. *Semin Cell Dev Biol* **46**, 143–154 (2015).
36. Spormann, A. M. & Kaiser, D. Gliding mutants of *Myxococcus xanthus* with high reversal frequencies and small displacements. *J Bacteriol* **181**, 2593–2601 (1999).
37. Kaiser, D., Robinson, M. & Kroos, L. Myxobacteria, polarity, and multicellular morphogenesis. *Cold Spring Harb Perspect Biol* **2**, 1–26 (2010).
38. Herrou, J. & Mignot, T. Dynamic polarity control by a tunable protein oscillator in bacteria. *Curr Opin Cell Biol* **62**,.
39. Kaimer, C., Berleman, J. E. & Zusman, D. R. Chemosensory signaling controls motility and subcellular polarity in *Myxococcus xanthus*. *COMICR* vol. 15 751–757 Preprint at https://doi.org/10.1016/j.mib.2012.10.005 (2012).
40. Baranwal J, Lhospice S, Kanade M, Chakraborty S, Gade PR, Harne S, Herrou J, Mignot T, Gayathri P. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. *PLoS Biol.* 2019 Sep 27;17(9):e3000459. doi: 10.1371/journal.pbio.3000459.
41. Eckhert, E., Rangamani, P., Davis, A. E., Oster, G. & Berleman, J. E. Dual biochemical oscillators may control cellular reversals in *Myxococcus xanthus*. *Biophys J* **107**, 2700–2711 (2015).
42. Hartzellt, P. & Kaiser, D. Function of MglA, a 22-Kilodalton Protein Essential for Gliding in *Myxococcus xanthus*. *J Bacteriol* vol. 173 (1991).
43. Igoshin, O. a, Goldbeter, A., Kaiser, D. & Oster, G. A biochemical oscillator explains several aspects of *Myxococcus xanthus* behavior during development. *PNAS* **101**, 15760–15765 (2004).

44. Nan B, Bandaria JN, Guo KY, Fan X, Moghtaderi A, Yildiz A, Zusman DR. The polarity of myxobacterial gliding is regulated by direct interactions between the gliding motors and the Ras homolog MglA. *Proc Natl Acad Sci U S A* **112**, E186-93 (2015).
45. Leonardy S, Miertzschke M, Bulyha I, Sperling E, Wittinghofer A, Søgaard-Andersen L. Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. *EMBO J* **9**, 2276–2289 (2010).
46. Simanshu, D. K., Nissley, D. v. & McCormick, F. RAS Proteins and Their Regulators in Human Disease. *Cell* vol. 170 17–33 Preprint at https://doi.org/10.1016/j.cell.2017.06.009 (2017).
47. Vetter, I. R. Ras Superfamily Small G Proteins: Biology and Mechanisms 1. 25–51 (2014) doi:10.1007/978-3-7091-1806-1.
48. Wuichet, K. & Søgaard-Andersen, L. Evolution and diversity of the ras superfamily of small GTPases in Prokaryotes. *Genome Biol Evol* **7**, 57–70 (2014).
49. Kinoshita, K., Sadanami, K., Kidera, A. & Go, N. Structural motif of phosphate-binding site common to various protein superfamilies: all-against-all structural comparison of protein-mononucleotide complexes. *Protein Eng.* vol. 12 https://academic.oup.com/peds/article/12/1/11/1513093 (1999).
50. Saraste, M., Sibbald, P. R. & Wittinghofer, A. The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**, 430–434 (1990).
51. Dever, T. E., Glynias, M. J. & Merrick, W. C. GTP-binding domain: three consensus sequence elements with distinct spacing. *PNAS* **84**, 1814–1818 (1987).
52. John J, Rensland H, Schlichting I, Vetter I, Borasio GD, Goody RS, Wittinghofer A. Kinetic and structural analysis of the Mg(2+)-binding site of the guanine nucleotide-binding protein p21H-ras. *J Biol Chem*. 1993 Jan 15;268(2):923-9.
53. Miertzschke M, Koerner C, Vetter IR, Keilberg D, Hot E, Leonardy S, Søgaard-Andersen L, Wittinghofer A. Structural analysis of the Ras-like G protein MglA and its cognate GAP MglB and implications for bacterial polarity. *EMBO J*. 2011 Aug 16;30(20):4185-97. doi: 10.1038/emboj.2011.291.
54. Dong, J. H., Wen, J. F. & Tian, H. F. Homologs of eukaryotic Ras superfamily proteins in prokaryotes and their novel phylogenetic correlation with their eukaryotic analogs. *Gene* **396**, 116–124 (2007).
55. Mishra, A. K. & Lambright, D. G. Invited review: Small GTPases and their GAPs. *Biopolymers* vol. 105 431–448 Preprint at https://doi.org/10.1002/bip.22833 (2016).
56. Nirwan N, Singh P, Mishra GG, Johnson CM, Szczelkun MD, Inoue K, Vinothkumar KR, Saikrishnan K. Hexameric assembly of the AAA+ protein McrB is necessary for GTPase activity. *Nucleic Acids Res*. 2019 Jan 25;47(2):868-882. doi: 10.1093/nar/gky1170
57. Bulyha, I., Hot, E., Huntley, S. & Søgaard-Andersen, L. GTPases in bacterial cell polarity and signalling. *COMICR* vol. 14 726–733 Preprint at https://doi.org/10.1016/j.mib.2011.09.001 (2011).
58. Cherfils, J. & Zeghouf, M. Regulation of Small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* **93**, 269–309 (2013).
59. McBride, M. J. Nonflagellar Bacterial Motility. *Encyclopedia of Microbiology* 319–328 (2019) doi:10.1016/B978-0-12-801238-3.66188-3.
60. Sun, M., Wartel, M., Cascales, E., Shaevitz, J. W. & Mignot, T. Motor-driven intracellular transport powers bacterial gliding motility. *PNAS* **108**, 7559–7564 (2011).
61. Mauriello, E. M. F., Mignot, T., Yang, Z. & Zusman, D. R. Gliding Motility Revisited: How Do the Myxobacteria Move without Flagella? *MMBR* **74**, 229–249 (2010).
62. Wall, D. & Kaiser, D. Type IV pili and cell motility. *Mol Microbiol* vol. 32 01–10 Preprint at https://doi.org/10.1046/j.1365-2958.1999.01339.x (1999).

63. Nudleman, E., Wall, D. & Kaiser, D. Polar assembly of the type IV pilus secretin in *Myxococcus xanthus*. *Mol Microbiol* **60**, 16–29 (2006).
64. Bulyha I, Schmidt C, Lenz P, Jakovljevic V, Höne A, Maier B, Hoppert M, Søgaard-Andersen L. Regulation of the type IV pili molecular machine by dynamic localization of two motor proteins. *Mol Microbiol*. 2009 Nov;74(3):691-706. doi: 10.1111/j.1365-2958.2009.06891.x.
65. Bischof, L. F., Friedrich, C., Harms, A., Søgaard-Andersen, L. & van der Does, C. The Type IV pilus assembly ATPase PilB of *Myxococcus xanthus* interacts with the inner membrane platform protein pilc and the nucleotide-binding protein PilM. *JBC* **291**, 6946–6957 (2016).
66. Jakovljevic, V., Leonardy, S., Hoppert, M. & Søgaard-Andersen, L. PilB and PilT are ATPases acting antagonistically in type IV pilus function in *Myxococcus xanthus*. *J Bacteriol* **190**, 2411–2421 (2008).
67. Agrebi, R., Wartel, M., Brochier-Armanet, C. & Mignot, T. An evolutionary link between capsular biogenesis and surface motility in bacteria. *Nat Rev Microbiol* vol. 13 318–326 Preprint at https://doi.org/10.1038/nrmicro3431 (2015).
68. Nan, B. & Zusman, D. R. Uncovering the mystery of gliding motility in the myxobacteria. *Annu Rev Genet* **45**, 21–39 (2011).
69. Nan, B., Mauriello, E. M. F., Sun, I. H., Wong, A. & Zusman, D. R. A multi-protein complex from *Myxococcus xanthus* required for bacterial gliding motility. *Mol Microbiol* **76**, 1539–1554 (2010).
70. Nan B, Bandaria JN, Moghtaderi A, Sun IH, Yildiz A, Zusman DR. Flagella stator homologs function as motors for myxobacterial gliding motility by moving in helical trajectories. *Proc Natl Acad Sci U S A.* 2013 Apr 16;110(16):E1508-13. doi: 10.1073/pnas.1219982110.
71. Mauriello EM, Mouhamar F, Nan B, Ducret A, Dai D, Zusman DR, Mignot T. Bacterial motility complexes require the actin-like protein, MreB and the Ras homologue, MglA. *EMBO J*. 2010 Jan 20;29(2):315-26. doi: 10.1038/emboj.2009.356.
72. Sliusarenko, O., Zusman, D. R. & Oster, G. The motors powering A-motility in *Myxococcus xanthus* are distributed along the cell body. *J Bacteriol* **189**, 7920–7921 (2007).
73. van Teeffelen, S., S. Wang, L. Furchtgott, K.C. Huang, N.S. Wingreen, J.W. Shaevitz, and Z. Gitai. 2011. The bacterial actin MreB rotates, and rotation depends on cell-wall assembly. *PNAS USA*. 108:15822–15827. https://doi.org/10.1073/pnas.1108999108
74. Treuner-Lange A, Macia E, Guzzo M, Hot E, Faure LM, Jakobczak B, Espinosa L, Alcor D, Ducret A, Keilberg D, Castaing JP, Lacas Gervais S, Franco M, Søgaard-Andersen L, Mignot T. The small G-protein MglA connects to the MreB actin cytoskeleton at bacterial focal adhesions. *J Cell Biol.* 2015 Jul 20;210(2):243-56. doi: 10.1083/jcb.201412047.
75. Faure LM, Fiche JB, Espinosa L, Ducret A, Anantharaman V, Luciano J, Lhospice S, Islam ST, Tréguier J, Sotes M, Kuru E, Van Nieuwenhze MS, Brun YV, Théodoly O, Aravind L, Nollmann M, Mignot T. The mechanism of force transmission at bacterial focal adhesion complexes. *Nature*. 2016 Nov 24;539(7630):530-535. doi: 10.1038/nature20121.
76. Guzzo M, Murray SM, Martineau E, Lhospice S, Baronian G, My L, Zhang Y, Espinosa L, Vincentelli R, Bratton BP, Shaevitz JW, Molle V, Howard M, Mignot T. A gated relaxation oscillator mediated by FrzX controls morphogenetic movements in *Myxococcus xanthus*. *Nat Microbiol*. 2018 Aug;3(8):948-959. doi: 10.1038/s41564-018-0203-x.
77. Carreira, L. A. M., Szadkowski, D., Müller, F. & Søgaard-Andersen, L. Spatiotemporal regulation of switching front–rear cell polarity. *COCEBI* vol. 76 Preprint at https://doi.org/10.1016/j.ceb.2022.102076 (2022).
78. Zhang, Y., Guzzo, M., Ducret, A., Li, Y.-Z. & Mignot, T. A Dynamic Response Regulator Protein Modulates G-Protein–Dependent Polarity in the Bacterium *Myxococcus xanthus*. *PLoS Genet* **8**, e1002872 (2012).

79. Zhang, Y., Franco, M., Ducret, A. & Mignot, T. A bacterial ras-like small GTP-binding protein and its cognate GAP establish a dynamic spatial polarity axis to control directed motility. *PLoS Biol* **8**, (2010).
80. Galicia C, Lhospice S, Varela PF, Trapani S, Zhang W, Navaza J, Herrou J, Mignot T, Cherfils J. MglA functions as a three-state GTPase to control movement reversals of *Myxococcus xanthus. Nat Commun*. 2019 Nov 22;10(1):5300. doi: 10.1038/s41467-019-13274-3.
81. Mercier R, Bautista S, Delannoy M, Gibert M, Guiseppi A, Herrou J, Mauriello EMF, Mignot T. The polar Ras-like GTPase MglA activates type IV pilus via SgmX to enable twitching motility in *Myxococcus xanthus*. *Proc Natl Acad Sci U S A*. 2020 Nov 10;117(45):28366-28373. doi: 10.1073/pnas.2002783117.
82. Dinet, C. & Mignot, T. Unorthodox regulation of the MglA Ras-like GTPase controlling polarity in *Myxococcus xanthus*. *FEBS Lett* (2022) doi:10.1002/1873-3468.14565.
83. Mignot, T. & Noellmann, M. New insights into the function of a versatile class of membrane molecular motors from studies of *Myxococcus xanthus* surface (gliding) motility. *Microbial Cell* **4**, 98–100 (2017).
84. Luciano J, Agrebi R, Le Gall AV, Wartel M, Fiegna F, Ducret A, Brochier-Armanet C, Mignot T. Emergence and modular evolution of a novel motility machinery in bacteria. *PLoS Genet*. 2011 Sep;7(9):e1002268. doi: 10.1371/journal.pgen.1002268.
85. Wartel M, Ducret A, Thutupalli S, Czerwinski F, Le Gall AV, Mauriello EM, Bergam P, Brun YV, Shaevitz J, Mignot T. A Versatile Class of Cell Surface Directional Motors Gives Rise to Gliding Motility and Sporulation in *Myxococcus xanthus*. *PLoS Biol* **11**, (2013).
86. Leonardy, S., Freymark, G., Hebener, S., Ellehauge, E. & Søgaard-Andersen, L. Coupling of protein localization and cell movements by a dynamically localized response regulator in *Myxococcus xanthus*. *EMBO J* **26**, 4433–4444 (2007).
87. Szadkowski D, Harms A, Carreira LAM, Wigbers M, Potapova A, Wuichet K, Keilberg D, Gerland U, Søgaard-Andersen L. Spatial control of the GTPase MglA by localized RomR-RomX GEF and MglB GAP activities enables *Myxococcus xanthus* motility. *Nat Microbiol*. 2019 Aug;4(8):1344-1355. doi: 10.1038/s41564-019-0451-4.
88. McLoon AL, Wuichet K, Häsler M, Keilberg D, Szadkowski D, Søgaard-Andersen L. MglC, a Paralog of *Myxococcus xanthus* GTPase-Activating Protein MglB, Plays a Divergent Role in Motility Regulation. *J Bacteriol*. 2015 Nov 16;198(3):510-20. doi: 10.1128/JB.00548-15.
89. Kapoor, S., Kodesia, A., Kalidas, N., Ashish & Thakur, K. G. Structural characterization of *Myxococcus xanthus* MglC, a component of the polarity control system, and its interactions with its paralog MglB. *JBC* **296**, (2021).
90. Carreira, L., Szadkowski, D., Lometto, S., Hochberg, G. K. A. & Søgaard-Andersen, L. Molecular basis and design principles of a system for switchable 3 front-rear polarity and directional migration 4 5 doi:10.1101/2022.12.09.519731 (2022).
91. Szadkowski, D., Carreira, L. A. M. & Søgaard-Andersen, L. A bipartite, low-affinity roadblock domain-containing GAP complex regulates bacterial front-rear polarity. *PLoS Genet* **18**, (2022).
92. Bulyha I, Lindow S, Lin L, Bolte K, Wuichet K, Kahnt J, van der Does C, Thanbichler M, Søgaard-Andersen L. Two small GTPases act in concert with the bactofilin cytoskeleton to regulate dynamic bacterial cell polarity. *Dev Cell*. 2013 Apr 29;25(2):119-31. doi: 10.1016/j.devcel.2013.02.017.
93. Kanade, M., Singh, N. B., Lagad, S., Baranwal, J. & Gayathri, P. Dual specificity of a prokaryotic GTPase-activating protein (GAP) to two small Ras-like GTPases in *Myxococcus xanthus*. *FEBS J.* **288**, 1565–1585 (2021).

Chapter 2: **Materials and methods**

2.1 Introduction

The thesis focuses on the biochemical and structural characterization of the polarity module proteins in *Myxococcus xanthus*. To achieve the objectives we followed the basic workflow consisting of the following steps:

1. Cloning the respective genes in a suitable vector
2. Standardization of protein expression and purification
3. Biochemical assays
    1. Activity assays
    2. Interaction studies
4. Structural studies
    1. Protein crystallization
    2. Structure determination
    3. Structure analysis

In this chapter, I will be discussing these general methodologies in terms of principles and detailed protocols. Specific details pertaining to individual experiments and proteins will be detailed in the corresponding chapters.

2.2 Molecular Cloning

There are several techniques used to insert the gene of interest under a suitable promoter in a plasmid vector which is consequently used to express proteins in suitable bacterial strains. Most methods employ the use of restriction enzymes to generate complementary cuts between the vector and the insert which is consequently ligated. In this study, in addition to the commonly used restriction-ligation-based strategies, I have majorly employed PCR-based restriction enzyme-free strategies[1].

Restriction free cloning (RF cloning)

Principle

Here the gene to be cloned is PCR amplified flanked by priming sites complementary to vector sequences (Fig 2.1). Further, this product is used as a megaprimer in a linear amplification reaction using a suitable vector template. This generates a nicked, circular DNA molecule which is preferentially retained in a *DpnI* digestion reaction, eliminating the methylated parent template. This is further transformed in *E. coli* competent cells and plasmids isolated from the

colony are screened for a positive clone. This is achieved by colony PCR based strategies with primers distinguishing positive or negative clones or by observing correct-sized insert release using suitable restriction enzymes.

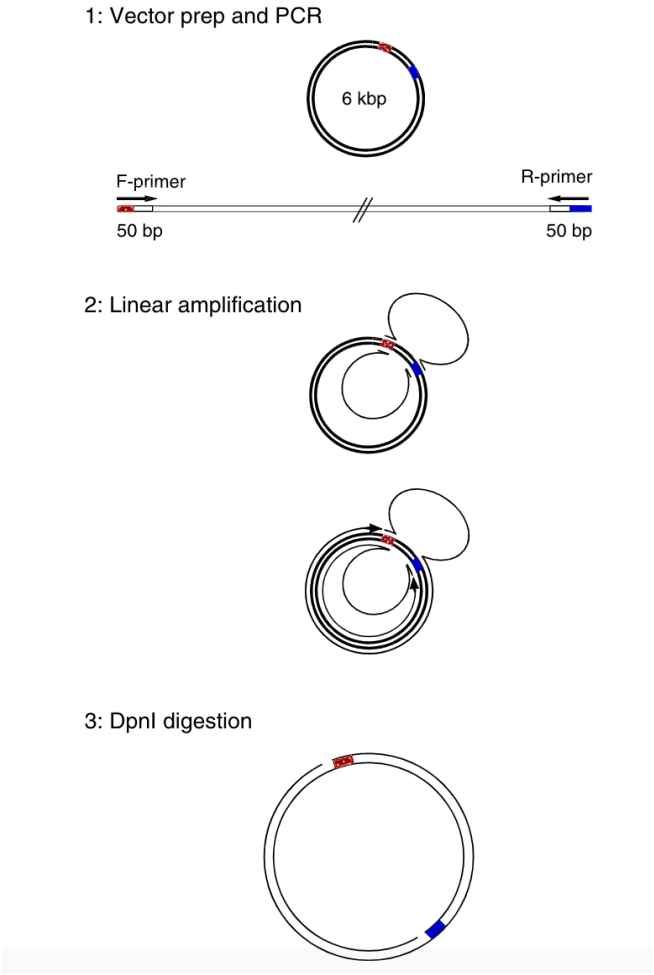


Figure 2.1. RF cloning strategy
(Reproduced from van den Ent and Lowe 2005[1]).

Methodology

2.2.1 MglA Walker B mutants and MglB GAP/GEF mutants (performed with Manil Kanade[2]) MglA and MglB constructs were used as reported by Baranwal *et al.,* 2019[3] using suitable primers and cloned into the pHis17-Amp^{R} (Addgene plasmid #78201). For generating the MglA^{D58A} and MglA^{T78A}, mutations were incorporated in MglA using primers listed in Table 2.1. MglB$^{\Delta N}$ and MglB^{N37R98A} mutations were incorporated in MglB using primers listed in Table 2.1. The clones were subjected to *DpnI* (New England Biolabs Inc.) digestion followed by transformation. Positive clones were selected on ampicillin-containing plates and checked

by the consequent release of correct-sized fragments following double digestion with *NdeI* and *BamHI* (New England Biolabs Inc.). All clones were confirmed by sequencing.

2.2.2 Linked constructs of MglA/MglB

Restriction-free cloning [30] was used to generate the constructs of MglA-link-MglB with C-terminal hexahistidine tag in *pHis17-Kan^R* (*mglA-link-mglB-H$_6$,* refer Addgene plasmid #78201 for vector backbone, *amp^R* cassette replaced with kanamycin resistance using specific primers: pHis17 *Kan^r*). Further MglA-link-MglB$^{\Delta Ct}$ with C-terminal hexahistidine tag in *pHis17-Kan^R* was cloned (*mglA-link-mglB$^{\Delta Ct}$-H$_6$*). The primers used are enlisted in Table 3. Similarly, MglB and MglB$^{\Delta Ct}$ were cloned with a C-terminal Strep-tag in pHis17-*Amp^R*. MglA and MglB constructs used were as reported in Baranwal, et al., 2019 [3]. The clones were subjected to *DpnI* (New England Biolabs Inc.) digestion followed by transformation. Positive clones were selected on plates containing suitable antibiotic (kanamycin and ampicillin respectively) and checked by the consequent release of correct-sized fragments following double digestion with *NdeI* and *BamHI* (New England Biolabs Inc.). All the clones were confirmed by sequencing.

<u>Cloning strategy of the co-expression construct</u>

Firstly, MglB-Strep and MglA with the intervening operon sequence was cloned (*mglB-Strep-Operon-mglA)*. The operon sequence was synthesized using forward and reverse primers (Table 3) using overlap extension. This product was used as a megaprimer to be inserted in *pHis17 mglA-link-mglB-H$_6$* construct as the template at the 5' end of the *mglA* sequence. MglA has an internal *XhoI* restriction site. We performed double digestion using the internal *XhoI* and *HindIII* (on the vector backbone at the 3' end of the gene) restriction sites on *mglB-Strep-Operon-mglA* and *mglA-link-mglB-H$_6$* constructs to generate the vector and insert, respectively, for the ligation reaction (enzymes from New England Biolabs Inc.). *mglB-Strep-Operon-mglA* vector treated with TSAP (Promega) and the digested *mglA-link-mglB-H$_6$* insert was ligated at the MglA *XhoI* restriction site using T4 DNA Ligase, (New England Biolabs Inc). The clones were transformed and screened similarly using double digestion with *NdeI* and *HindIII*. This generated the {AB^{WT}} construct.

For {AB$_1$$^{\Delta Ct}B_2$}, *mglB$^{\Delta Ct}$-Strep-Operon-mglA* was PCR amplified using the same overlap extension primers as before, using *mglB$^{\Delta Ct}$-Strep* construct. This product was digested using

N-terminal *NdeI* and internal *XhoI* and inserted into a similarly digested *mglA-link-mglB-H*$_6$ vector. {$AB_1B_2^{\Delta Ct}$}, and {$AB^{\Delta Ct}$} were generated using combinations of *NdeI* and *XhoI* digested inserts of *mglB-Strep-Operon-mglA* and *mglB*$^{\Delta Ct}$*-Strep-Operon-mglA*, respectively, and ligated to a digested *NdeI* and *XhoI mglA-link-mglB*$^{\Delta Ct}$*-H*$_6$ template.

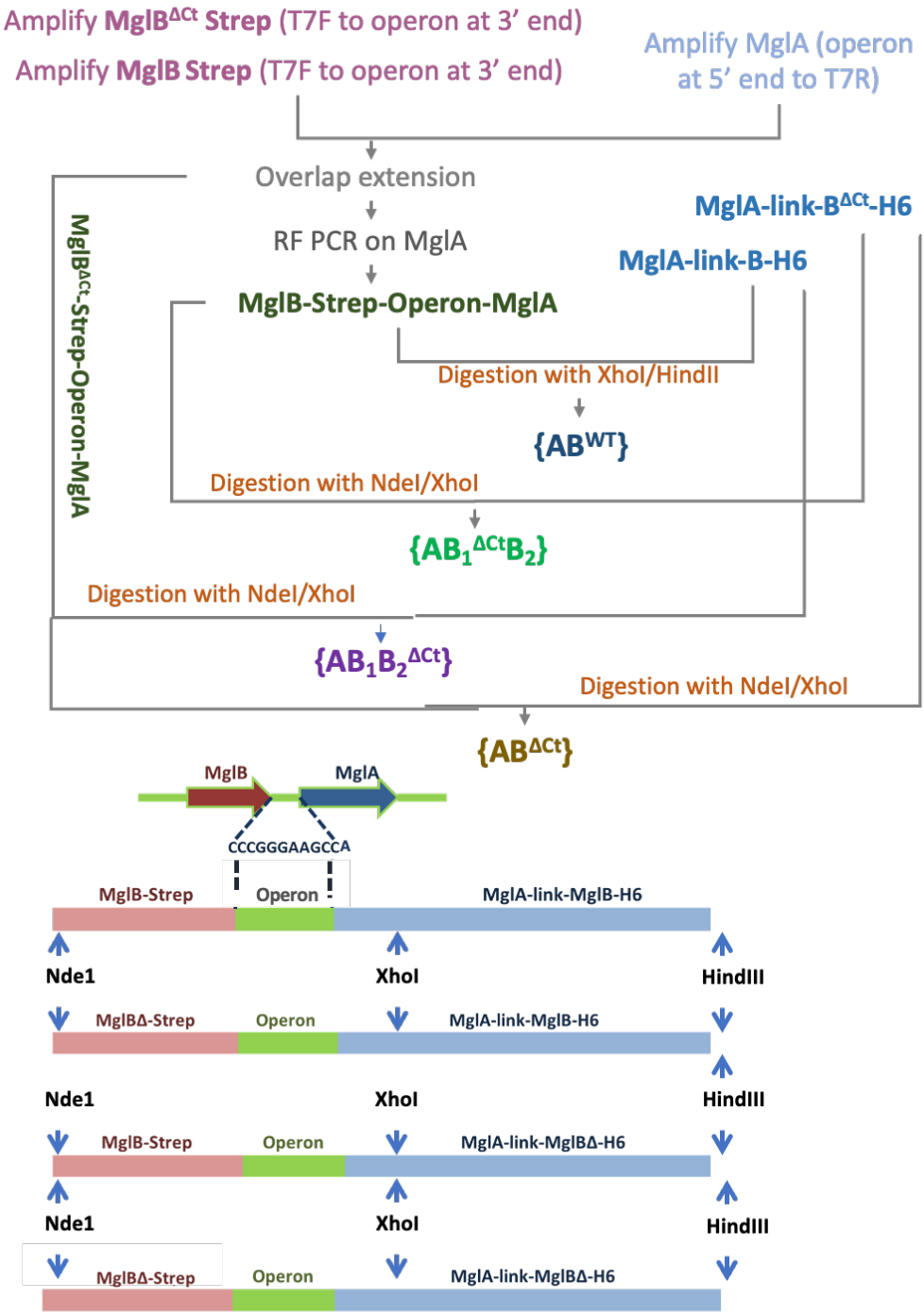


Figure 2.2. Strategy for cloning linked MglAB constructs

2.2.3 RomR and RomX constructs

For the REC domain of RomR, residues 1-116 of the RomR gene was amplified from *romR-pETPhos* (obtained from Dr. Tam Mignot's lab) construct. It was cloned in *pHis17-Kan*^{R} vector with a C-terminal hexahistidine tag using respective primers (Table 2.1). Further, a construct with N-terminal hexahistidine tag was cloned. Since the protein was impure, further 1-125 and 1-141 residues of the *romR* gene were cloned in the *pHis17-Kan*^{R} vector with a C-terminal hexahistidine tag. In an attempt for constructing a co-expression system with RomX, RomR REC (1-125) was cloned with a C-terminal Strep-tag as well.

Further, for the C-terminal, residues 331-420 *romR* gene were cloned in the *pHis17-Kan*^{R} vector with a C-terminal hexahistidine tag (insert amplified from *romR* gene synthesized from BioTechDesk). Residues 398-420 were cloned with an N-terminal GST tag in the *pHis17-Amp*^{R} vector (GST was amplified from the *pGEX-4T-1* vector and inserted in *pHis17-Amp*^{R}. Further a shorter construct of RomR was constructed with the REC domain (1-141) followed by C-terminal ordered region (331-420) in the *pHis17-Kan*^{R} vector with a C-terminal hexahistidine tag

RomX gene was amplified from *M. xanthus genomic* DNA (obtained from Dr. Tam Mignot's lab) and cloned into *pHis17-Amp*^{R} vector with both N and C-terminal hexahistidine tags. Further, it was also cloned with N and C-terminal Strep tags. A similar strategy as mentioned above was also used to generate interface mutants of RomR and RomX[4].

Table 2.1: List of constructs and their respective primers used for cloning

Plasmid Constructs	**Primer Sequence**	**Template**	**Resistance**	**Purpose**
MglA and MglB constructs				
pHis17 mglB$^{\Delta N}$	MglB ΔN-F 5'GTTTAACTTTAAGAAGGAGATATACATATG TACGAAGAGGAGTTCACC3' MglB H6-R 5'GCTTTTAATGATGATGATGATGATGGG3'	*pHis17 mglB*	Amp^R	MglB$^{\Delta N}$ mutant with C-terminal hexahistidine tag
pHis17 mglB^{N37R98A}	MglB N37A-F 5'CCTCGTCGACAAGGCCGGGCAGCTCATCTC C3' MglB R98A-R 5'CGACCAGCACGACCGCGCTGCCGACGATGG 3'	*pHis17 mglB*	Amp^R	MglB^{N37R98A} mutant with C-terminal hexahistidine tag

Plasmid Constructs	**Primer Sequence**	**Template**	**Resistance**	**Purpose**
pHis17 mglA. mglB, mglB$^{\Delta Ct}$	Refer [3]		Amp^{R}	MglA and MglB, MglB$^{\Delta Ct}$ with C-terminal hexahistidine tag
pHis17 mglA ^{D58A}	MglA D58A-F 5′-CCGCACGCTCTTCTTCGCCTTCCTGCCGCTGTCGC-3’ MglA H6-R 5’GCTTTTAATGATGATGATGATGATGGG3’	*pHis17 mglA*	Amp^{R}	MglA^{D58A} mutant with C-terminal hexahistidine tag
pHis17 mglA ^{T58A}	MglA D58A-F 5′-CGCTTCCACCTGTACGCGGTGCCCGGTCAGGTC-3′ MglA H6-R 5’GCTTTTAATGATGATGATGATGATGGG3’	*pHis17 mglA*	Amp^{R}	MglA^{T78A} mutant with C-terminal hexahistidine tag
pHis17 mglA-link-mglB H6	MglA-link-B-F 5'CCTCACCGAGCTCAAGAAGGGTGGTGGTTCCGGCACGCAACTGGTGATGTACG3' MglB H6-R 5’GCTTTTAATGATGATGATGATGATGGG3’	*pHis17 mglA*	Kan^{R}	MglA linked to MglB with a short GS linker, with C-terminal hexahistidine tag
pHis17 mglA-link-mglB$^{\Delta Ct}$ H6	MglA-link-B-F 5'CCTCACCGAGCTCAAGAAGGGTGGTGGTTCCGGCACGCAACTGGTGATGTACG3' MglB$^{\Delta Ct}$ H6-R 5’GCTTTTAATGATGATGATGATGATGGGATCCAGGACTGTCAGTCTTCT3’	*pHis17 mglA*	Kan^{R}	MglA linked to MglB$^{\Delta Ct}$ with a short GS linker, with C-terminal hexahistidine tag
pHis17 mglB Strep	MglB-F 5’GTTTAACTTTAAGAAGGAGATATACAT3’ MglB-Strep-R 5'TTATTTTTCGAACTGAGGATGAGACCAGGATCCCTCGCTGAAGAGGTTGTCG3'	*pHis17 mglB*	Amp^{R}	MglB with C-terminal Strep tag
pHis17 mglB$^{\Delta Ct}$ Strep	MglB-F 5’GTTTAACTTTAAGAAGGAGATATACAT3’ MglB-Strep-R 5'TTATTTTTCGAACTGAGGATGAGACCAGGATCCAGGACTGTCAGTCTTCTTC3'	*pHis17 mglB*	Amp^{R}	MglB$^{\Delta Ct}$ with C-terminal Strep tag
pHis17 mglB-Strep-Operon-mglA	T7-F 5’TAATACGACTCACTATAGGG3’ Strep-Operon-MglA-R 5’GGATGAGTAATTGATGAAGGACATTGGCTTCCCGGGTTATTTTTCGAACTGAGGATGAG3’ Strep Operon MglA-F 5’FCTCATCCTCAGTTCGAAAAATAACCCGGGAAGCCAATGTCCTTCATCAATTACTCATCC3’ T7-R 5’GCTAGTTATTGCTCAGCGG3’	*pHis17 mglA*	Kan^{R}	To clone linked constructs of MglA-B

Plasmid Constructs	**Primer Sequence**	**Template**	**Resistance**	**Purpose**
pHis17 {AB^{WT}}	Strep Operon MglA- 5'FCTCATCCTCAGTTCGAAAAATAACCCGGGAAGCCAATGTCCTTCATCAATTACTCATCC3' Strep Operon MglA-R 5'GGATGAGTAATTGATGAAGGACATTGGCTTCCCGGGTTATTTTTCGAACTGAGGATGAG3'	*pHis17 mglA*	Kan^{R}	Co-expression vector expressing MglA-link-MglB-H6 and MglB-Strep, for generating asymmetric MglA MglB complexes
pHis17 Kan^{R}	pHisK-F 5'GCTAAAGTTGTTAAAGACTTTAAATTGCCGCGCGGCAGCCACATG3' pHisK-R 5'GTATATATGAGTAAACTTGGTCTGACAGTTAGAAAAACTCATCGAGCATC3'	*pHis17*	Kan^{R}	Convert pHis17 amp resistance vector into kan resistance vector
RomR and RomX constructs				
pHis17 romR REC116 H6	RomR-F 5'GTTTAACTTTAAGAAGGAGATATACATATGCCCAAGAATCTGCTG3' RomR-116-H6-R 5'GCTTTTAATGATGATGATGATGATGGGATCCCACCTTGTCGAGCAGCACCTG3'	*pHis17 mglA-link-mglB H6*	Kan^{R}	REC domain of RomR (residues 1-116) with C-terminal hexahistidine tag
pHis17 romR REC125 H6	RomR-F 5'GTTTAACTTTAAGAAGGAGATATACATATGCCCAAGAATCTGCTG3' RomR-125-H6-R 5'GCTTTTAATGATGATGATGATGATGGGATCCGTTGGACTTCTGGCCGACCAGCGCCTT3'	*pHis17 romR REC116 H6*	Kan^{R}	REC domain of RomR (residues 1-125) with C-terminal hexahistidine tag
pHis17 romR REC125 Strep	RomR-F 5'GTTTAACTTTAAGAAGGAGATATACATATGCCCAAGAATCTGCTG3' RomR-125-Strep-R 5'GCTTTTAATGATGATGATGATGATGGGATCCGTTGGACTTCTGGCCGACCAGCGCCTT3'	*pHis17 mglB Strep*	Amp^{R}	REC domain of RomR (residues 1-125) with C-terminal Strep-tag
pHis17 romR REC141 H6	RomR-F 5'GTTTAACTTTAAGAAGGAGATATACATATGCCCAAGAATCTGCTG3' RomR-125-H6-R 5'GCTTTTAATGATGATGATGATGATGGGATCCCTGAGGAGCCGCGTGGCGCACCTGCGT3'	*pHis17 romR REC116 H6*	Kan^{R}	REC domain of RomR (residues 1-141) with C-terminal hexahistidine tag
pHis17 romR REC141Ct H6	RomR-REC141Ct-F 5'ACGCAGGTGCGCCACGCGGCTCCTCAGCCGTCCATCAGCATCGAGGACTCGCTGCCC3' RomR-H6-R 5'GCTTTTAATGATGATGATGATGATGGGATCCGTGCTGGGTCTCTC3'	*pHis17 romR REC141 H6*	Kan^{R}	REC domain of RomR (residues 1-116) linked to C-terminal (residues 331-420) with C-terminal hexahistidine tag

Plasmid Constructs	**Primer Sequence**	**Template**	**Resistance**	**Purpose**
pHis17 romR 331Ct H6	RomR-331Ct-F 5'GTTTAACTTTAAGAAGGAGATATACATATGCCGTCCATCAGCATCGAGGACTCGCTG3' RomR-H6-R 5'GCTTTTAATGATGATGATGATGATGGGATCCGTGCTGGGTCTCTC3'	*pHis17 mglA-link-mglB H6*	Kan^R	C-terminal of RomR (residues 331-420) with C-terminal hexahistidine tag
pHis17 romR REC E108A H6	RomR-F 5'GTTTAACTTTAAGAAGGAGATATACATATGCCCAAGAATCTGCTG3' RomR-125-H6-R 5'CTTCACCTTGTCGAGCAGCACCTGGCTCGCGAAGGGCTTGGTGAC3'	*pHis17 romR REC125 H6*	Kan^R	REC domain mutant, $RomR^{E108A}$ (residues 1-125) with C-terminal hexahistidine tag
pHis17 romX H6	RomX-F 5'GTTTAACTTTAAGAAGGAGATATACATATGACGGACGAGGAAAAGGTC3' RomX-H6-R 5'GCTTTTAATGATGATGATGATGATGGGATCCCCAGATCTTCGACTTCAC3'	*pHis17 mreB H6*	Amp^R	RomX with C-terminal hexahistidine tag
pHis17 H6 romX	RomX-H6-F 5'ATATACATATGCGTGGCCATCATCATCATCATCATGGAGGAACGGACGAGGAAAAGGTC3' RomX-R 5'TTTAATGATGATGATGATGATGGGATCCTTACCAGATCTTCGACTTCAC3'	*pHis17 H6 mglB*	Amp^R	RomX with N-terminal hexahistidine tag
pHis17 romX Strep	RomX-F 5'GTTTAACTTTAAGAAGGAGATATACATATGACGGACGAGGAAAAGGTC3' RomX-Strep-R 5'TTATTTTTCGAACTGAGGATGAGACCAGGATCCCCAGATCTTCGACTTCACATG3'	*pHis17 mglB Strep*	Amp^R	RomX with C-terminal Strep tag
pHis17 Strep romX	RomX-Strep-F 5'CTTTAAGAAGGAGATATACATATGCGTGGCTGGTCTCATCCTCAGTTCGAAAAAGGAGGAACGGACGAGGAAAAGGTC3' RomX-R 5'TTTAATGATGATGATGATGATGGGATCCTTACCAGATCTTCGACTTCAC3'	*pHis17 mreB H6*	Amp^R	RomX with N-terminal Strep tag
pHis17 H6 romX L75A,R76A	RomX-F 5'GTTTAACTTTAAGAAGGAGATATACATATGACGGACGAGGAAAAGGTC3' RomX-L75A R76A-H6-R 5'CGACTTCACATGCGCCTTCGACGCCGCGACGATGTCGTTGATCGC3'	*pHis17 H6 romX*	Amp^R	$RomX^{L75A,R76A}$ with N-terminal hexahistidine tag

2.3 Protein Expression

The cloned plasmids were transformed in BL21 strains of *E. coli* to standardise the expression of the respective proteins. These strains are specifically designed for expressing recombinant proteins at sufficient levels. This is achieved by the presence of genetic markers in the strains which prevent protein degradation. The inducible protein expression enables the successful expression of toxic proteins with minimal effect on cells. The two strains used for this study are:

i. Bl21DE3: T7 promoter-based expression systems with lambda DE3 lysogen.
ii. BL21AI: Chromosomal insertion at T7 RNA polymerase into the ***ara*B** locus of the ***ara*BAD** operon, regulating T7 RNAP via the arabinose-inducible *ara*BAD promoter, deficient in proteases like *lon* and *ompT*.

Heat shock was given to the cells with plasmids at 42ºC for 90 seconds. A patch of colonies was inoculated in the respective antibiotic-containing media (1X LB with 0.1mg/ml ampicillin or 0.05 mg/ml kanamycin) and was kept in shaking conditions at 37ºC. 1% v/v of the primary culture was used to inoculate the secondary culture. The cultures were induced with 0.2% L-arabinose once it reached the exponential phase of growth (between 0.6-0.8 OD_{600}). The cultures were incubated at respective post-induction temperatures for the respective duration. The samples were subjected to 15% SDS-PAGE to observe the overexpressed band of protein of interest. The constructs and their suitable expression conditions are detailed in Table 2.2.

Table 2.2: Expression conditions for constructs used in the study

Construct	**Strain Transformed**	**OD_{600} for induction**	**Post induction condition**	**Purification Strategy**
MglA/MglA mutants	Bl21DE3	0.5-0.6	30°C, 5 hours	HisTrap (2.4.1) SEC (2.4.3)
MglB/MglB mutants	Bl21AI	0.6-0.8	18°C, 10-12 hours	HisTrap (2.4.1) Anion Exchange (2.4.4)

$\{AB^{WT}\}$, $\{AB_1B_2^{\Delta Ct}\}$, $\{AB_1^{\Delta Ct}B_2\}$	Bl21DE3	0.6-0.8	18°C, 12-14 hours	HisTrap (2.4.1) StrepTrap (2.4.2) SEC (2.4.3)
$\{AB^{\Delta Ct}\}$	Bl21AI	0.6-0.7	18°C, 10-12 hours	HisTrap (2.4.1) StrepTrap (2.4.2)
RomR constructs	Bl21AI	0.6-0.8	18°C, 12-14 hours	HisTrap (2.4.1) Anion Exchange (2.4.4)
RomX constructs	Bl21AI	0.6-0.8	18°C, 12-14 hours	HisTrap (2.4.1) SEC (2.4.3)

Numbers in brackets for the last column refers to the section describing the detailed methodology for these steps.
All proteins were stored in the final buffer A50 (50 mM NaCl, 50 mM Tris pH 8.0).

Expression of selenomethionine labelled proteins:

The minimal media for growing selenomethionine labelled RomX cultures was prepared using the following recipe[26]:

Table 2.3: Recipe for minimal media for expressing selenomethionine labelled proteins

	Components to be autoclaved (mixed in 900 ml MQ)	**Amounts in grams for 1 L culture**	**Components to be filter sterilized (mixed in 100 ml MQ)**	**Amounts in grams for 1 L culture**
1.	Alanine	0.500	D-Glucose	10.000
2.	Arginine HCl	0.580	$MgSO_4.7H_2O$	0.250
3.	Aspartic acid	0.400	$FeSO_4$	4.2 mg
4.	L- Cystine	0.033	H_2SO_4	8 µl
5.	Glutamic acid (Na)	0.750	Thiamine	0.005
6.	Glutamine	0.333	Selenomethionine	50 mg
7.	Glycine	0.540		
8.	Histidine	0.060		
9.	Isoleucine	0.230		
10.	Leucine	0.230		
11.	Lysine HCl	0.420		
12.	Phenylalanine	0.133		
13.	Proline	0.100		
14.	L- Serine	2.083		
15.	Threonine	0.230		
16.	Tyrosine	0.170		
17.	Valine	0.230		
18.	Adenine	0.500		
19.	Guanosine	0.670		
20.	Thymine	0.170		
21.	Uracil	0.500		
22.	Sodium Acetate	1.500		
23.	Succinic Acid	1.500		
24.	Ammonium Cl	0.750		
25.	Sodium Hydroxide	0.850		
26.	Dibasic K_2HPO_4	10.500		

The autoclaved (900 ml) and sterilized (100ml) media were mixed and supplemented with 0.1mg/ml ampicillin. The primary culture was grown using same protocol as normal cultures. When the culture reached exponential phase, per 1 L culture, 1 ml primary culture was aliquoted in a sterile 1.5 ml tube and centrifuged at 3300g at 4°C for 10 minutes. The supernatant with LB media was discarded and the pellet was resuspended pellet with 1 ml Le Master's medium which was used to inoculate the 1 L secondary culture. Cultures were

incubated at 37°C until OD reached around 0.5. 100 mg each of lysine HCl, phenylalanine and threonine, 50 mg each of leucine, isoleucine and valine and 50 mg L-selenomethionine was weighed and added to the culture. Culture was induced with 0.2% L-arabinose and incubated at 18°C for 14 hours.

2.4 Protein Purification

2.4.1 Hexahistidine -tag-based affinity purification

Principle

In this method, the interaction between divalent metal ions from the matrix, in this case, Ni^{2+}, and the imidazole ring of the histidine tag of protein is utilized to purify proteins (Fig 2.3). Commercially available matrices of nickel-nitrilotriacetic acid (Ni-NTA) are coupled to resins[5]. These matrices coordinate metal ions via four of its coordination sites whereas two remaining sites are exposed to interact with histidine residues in the affinity tag. Binding occurs at neutral to slightly basic pH (7.5 – 8). Elution of proteins is achieved by competing for the binding of histidine with higher concentrations of imidazole (100-500 mM).

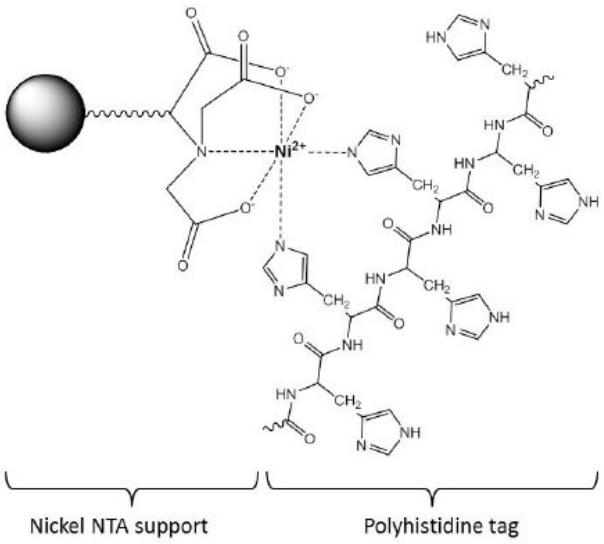


Figure 2.3. Schematic of Ni-NTA binding to hexahistidine tag

(Reproduced from Moser and Magdelin, 2012[6]).

Methodology

Harvested cells were resuspended in the lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl and 10% glycerol) at 4°C. Consequently, the samples were centrifuged at 39,000 g for 45 minutes. The lysate was loaded on an equilibrated 5 ml HisTrap HP (GE Healthcare) column. The loading buffer used (Buffer A) was 200 mM NaCl, 50 mM Tris pH 8.0 and the elution buffer (Buffer B) used was 200 mM NaCl, 50 mM Tris pH 8.0, and 500 mM Imidazole. Protein was eluted using a stepwise gradient of elution buffer (2%, 5%, 10%, 20%, 50%, 100% of Buffer B). Consequently, the fractions were subjected to dialysis to the final storage buffer (50 mM NaCl, 50 mM Tris pH 8.0) and concentrated further using centrifugal concentrators to be used for further steps.

2.4.2. Strep-tag-based affinity purification

Principle

This method utilizes the phenomenon of molecular mimicry, where an engineered peptide (WSHPQFEK) or Strep-tag present in fusion with proteins, binds to streptavidin which is conjugated with the resin under mild buffer conditions[7]. Streptavidin naturally binds biotin with high affinity at the same site where the Strep-tag binds, thereby eluting the bound proteins when subjected to low concentrations of biotin.

Methodology

Lysate/pooled fractions from the previous step were centrifuged at 39,000 g for 15 minutes and the supernatant was loaded on a 5 ml StrepTrap HP column (GE Healthcare). The binding buffer used was 150 mM NaCl, 50 mM Tris pH 8.0 and the elution buffer used was 150 mM NaCl, 50 mM Tris pH 8.0, 2.5 mM desthiobiotin. Elution buffer was injected into the column in 2 rounds of 5 ml each (Fraction I and II, respectively). Consequently, the fractions were subjected to dialysis to the final storage buffer (50 mM NaCl, 50 mM Tris pH 8.0) and concentrated further using centrifugal concentrators to be used for further steps.

2.4.3 Size exclusion chromatography

Principle

It is a method for separating molecules based on size. The smaller molecules are trapped within the pores of the matrix hence the bigger-sized molecules elute first. The elution volume (Ve) linearly declines with the log of the molecular hydrodynamic volume[8]. Standards of known molecular weight are used for calibrating the columns and the standard curve is used to determine the size of the protein. Usually, blue dextran (2000 kDa) is used to determine the void volume (Vo). (Ve/Vo) is plotted against the log of the MW with the standards of known molecular weight. This standard curve is used to determine the molecular weight of globular proteins using the elution volume.

Methodology

Superdex-75 column (GE Healthcare) was equilibrated with 50 mM NaCl, 50 mM Tris pH 8 (A50) buffer. Approximately 2-3 mg of protein (volume less than 500 µL) was injected into the column (GE Healthcare). UV absorbance at 280 nm was observed to monitor the elution of the protein. The respective fractions were pooled and concentrated and stored in final buffer A50 (50 mM NaCl, 50 mM Tris pH 8.0).

2.4.4 Anion Exchange Chromatography:

Principle

Depending on the pH of the buffer, the surface residues of amino acids ionize, imparting a net positive or negative charge. At low pH, as amino groups are protonated, protein molecules carry a net positive charge unlike at a higher pH, carboxylic acid groups are deprotonated and the surface of the protein molecules is negatively charged. At the isoelectric point (pI), which is a specific pH value, the protein is electrically neutral. Based on the pH of the buffer, the positive or negatively charged protein molecules bind to the oppositely charged resin. Elution is performed by changing the ionic strength or pH of the elution buffer.

Methodology

Protein purified from earlier steps was injected in MonoQ 4.6/100 PE (GE Healthcare). Buffers used for binding and elution were Buffer A (50 mM Tris pH 8.0, 50 mM NaCl) and Buffer B (50 mM Tris pH 8.0, 1 M NaCl), respectively. A linear gradient of Buffer A ranging from 0% to 50% Buffer B over 20 column volumes was injected and the fractions containing the protein were pooled and diluted to final storage buffer conditions (50 mM NaCl, 50 mM Tris pH 8.0). Protein was concentrated and stored.

2.5 Protein Quality Check

2.5.1 Bradford Assay: Measuring protein concentration

Principle

The principle of this assay is that the binding of protein molecules to Coomassie brilliant blue G-250 dye under acidic conditions results in a colour change from brown to blue[9]. The dye forms a strong, noncovalent complex with non-polar pockets of the denatured protein and consequently, with the amino groups through electrostatic interactions. When the dye binds to the protein, it causes a shift in the absorbance from 465 nm to 595 nm, causing a colour change from brown to blue. The intensity is proportional to the concentration of the protein-dye complex.

Methodology

Bradford reagent (Biorad) was added to BSA standard concentrations ranging from 0.1 – 1 mg/ml and was used to obtain the standard curve (5 µl BSA+250 µl reagent) in a 96-well plate. Protein dilutions were also added and mixed in the same ratio. The absorbance measurements were taken at a wavelength of 595 nm (Bradford, 1976). The standard curve was used to estimate the concentration of protein.

2.5.2 CD spectroscopy: Checking protein secondary structure content

Principle

Circular dichroism (CD) spectroscopy is a type of light absorption spectroscopy where a difference in absorbance of right- and left-circularly polarized light by a protein sample is measured[10] (Fig 2.4 A) It is sensitive to the relative abundance of secondary structures in proteins, e.g., α-helices, β-strands, and the resultant CD spectrum is a linear combination of each secondary structure type (Fig 2.4 B). A complex CD spectrum between 260 and approximately 180 nm can be deconvoluted and compared with spectra from pure secondary structural elements to determine the percentage of these constituent elements in proteins and peptides.

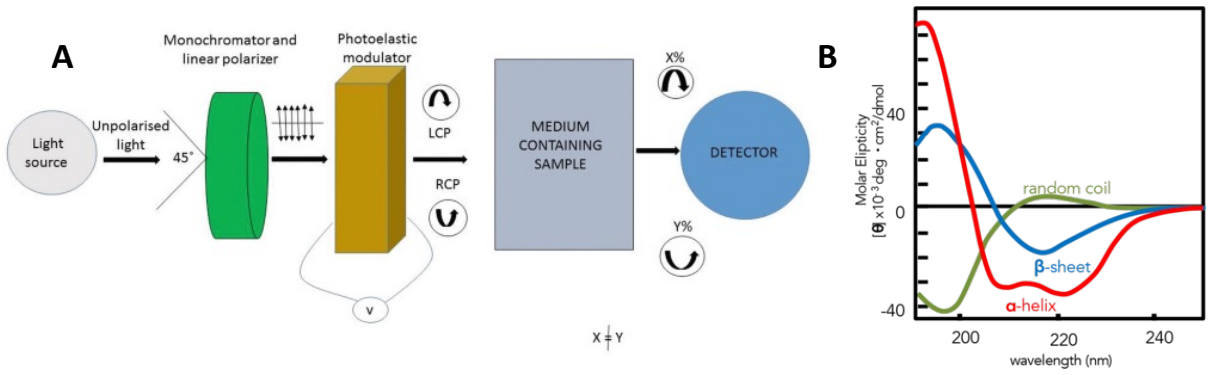


Figure 2.4. CD spectroscopy

A. Design of a CD spectrophotometer (Reproduced from Subadini *et al*., 2022[10]).
B. Typical CD spectra of pure secondary structure elements (Reproduced from CMI, adapted from N. Greenfield 1969).

Methodology

Far-UV CD spectra were collected using a Jasco J-815 spectropolarimeter with a protein concentration of 10 µM in a 1 mm cuvette, using a scan speed of 50 nm/min, a digital integration time of 2 s, and a bandwidth of 1 nm. Spectra could be collected between 200 to 250 nm under permissive high-tension voltage. Each spectrum was averaged for over 30 scans. Spectra were deconvoluted and analyzed using BeStSel[11].

2.5.3 Size exclusion chromatography coupled with Multi-angle light scattering (SEC-MALS): For estimating molar masses

Principle

Molar mass interpretation from SEC is subject to errors due to the erratic movement of non-globular proteins. In SEC-MALS the differential refractive index detector measures the concentration of the sample, whereas the light scattering detector measures simultaneously the light scattered at several angles. The light scattering (LS) and refractive index (RI) measurements can be used to measure the molecular weight (MW) of the eluted fraction using the equation,

MW= K'(LS)/(RI), where K' K_{RI}/[K_{LS}(dn/dc)]

Here, K_{RI} and K_{LS} calibration constants and dn/dc is the increment of refractive index with concentration[12].

Methodology

Superdex 200/75 Increase 10/300 GL column was used for SEC which was connected to an Agilent HPLC unit having an 18-angle light scattering detector (Wyatt Dawn HELIOS II) and a refractive index detector (Wyatt Optilab T-rEX). The experiments were performed at room temperature. The column was equilibrated with A50 (50 mM NaCl, 50 mM Tris pH 8.0) buffer at 0.4 ml/min. BSA at 2 mg/ml was used for the calibration of the system. The purified proteins (approximately 100 μM, 110 μL) were consequently loaded to estimate the molecular weight of the eluted peaks. The Zimm model implemented in ASTRA software was used for the curve fitting and estimation of molecular weights. GraphPad Prism was used to average molar mass from fitted plots.

2.5.4 High-pressure liquid chromatography to estimate nucleotide bound to protein

Principle

Here, proteins are denatured and filtered to isolate the nucleotide bound to the protein. The nucleotides are run on an ion exchange column to distinguish them based on retention times owing to the difference in the negative charge of the phosphate group. The area under the curve

is a measure of the concentration of nucleotide bound and this could be compared with standards with known concentrations of pure nucleotides for further quantification.

Methodology

The protein samples are diluted with buffer A50 and then heat denatured at 75 °C. The heated protein sample was spun at 21,000g for 10 min. Then the supernatant was filtered with a 0.22-µm cellulose acetate filter (Corning), and the sample was loaded on the DNAPac PA200 ion exchange column (Thermo Fisher). Buffer A (2 mM Tris pH 8.0) was used as a binding buffer. The runs were performed at a flow rate of 0.4 ml/min, with linear gradients of 0% to 20% Buffer B (2 mM Tris (pH 8.0), 1 M NaCl) for 3 CV (column volumes), 20%–50% Buffer B for 3 CV and 100% B for 2 CV. The sample injection volume was 100 µl. The absorbance at 260 nm, indicative of the presence of GDP, was plotted against conductivity. The GDP standard was obtained with 100 µl of a 500 µM filtered GDP solution.

2.6 Activity Assays

2.6.1 NADH coupled GTPase assay

Principle

It is a coupled enzymatic assay where the conversion of NADH to NAD^+ is coupled to the amount of GDP released[13]. As GTP to GDP conversion due to hydrolysis progresses, pyruvate kinase (PK) utilizes GDP/ADP and phosphoenol pyruvate to pyruvate, regenerating GTP[14]. Further, lactate dehydrogenase (LDH) utilizes NADH and pyruvate to produce NAD^+ and lactate (Fig 2.5). This results in a decrease in NADH absorbance at 340 nm which is equivalent to the amount of GDP released.

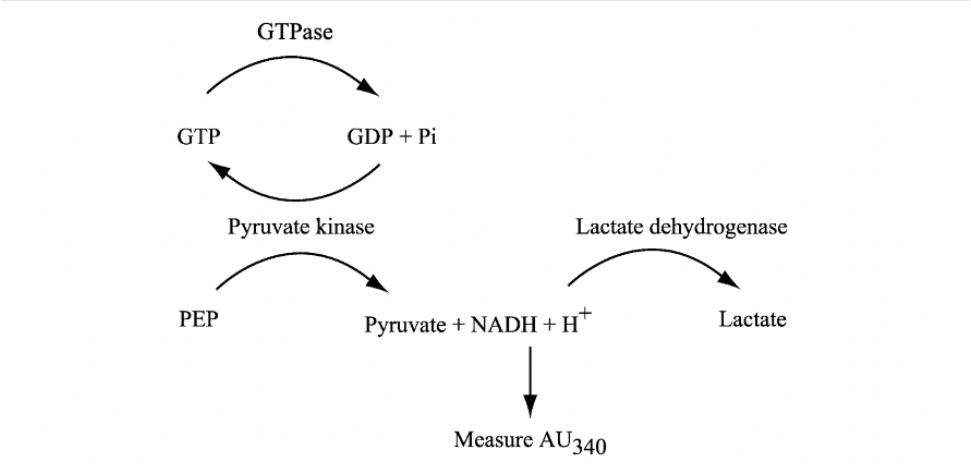


Figure 2.5. NADH coupled GTPase assay
(Reproduced from Ingerman and Nunnari, 2005[13])

Methodology

A master mix was prepared in A50 buffer (50 mM NaCl, Tris pH 8.0) containing 600 µM NADH, 1 mM phosphoenol pyruvate, 5 mM $MgCl_2$, 1 mM GTP and pyruvate kinase/lactate dehydrogenase enzyme mix (~25 U/ml). All components were mixed to a 200 µl reaction volume and added to Corning 96 well flat bottom plate. The reactions were initiated by adding purified complexes or a mix of proteins in respective ratios to a final concentration of 10 µM MglA. A decrease in NADH absorbance was measured at 340 nm using a Varioskan Flash (4.00.53) multimode plate reader. The absorbance was measured every 20 s for 7200 s. The initial time point and absorbance of buffer components were subtracted from all readings. NADH absorbance was converted to GDP produced using a slope obtained from a standard curve containing known concentrations of NADH. GraphPad Prism was used for data analysis and plotting the k_{obs} values.

2.6.2 Nucleotide exchange assay

Principle

Mant (N-methylanthraniloyl)-GDP and *mant*-GTP are fluorescent GDP/GTP analogues[15,16]. They are sensitive to the environment which is modified by binding to the protein. The chemical modification of the ribose moiety results in an increase in the fluorescent emission upon binding to proteins. However, it does not affect binding affinities as compared to

unlabelled nucleotides. Hence, an increase in the fluorescence intensity can be a cue for nucleotide binding to proteins (Fig 2.6). The bound nucleotide can also be competed out with unlabelled nucleotides and the dissociation kinetics can be measured by the rate of decrease in fluorescence intensity.

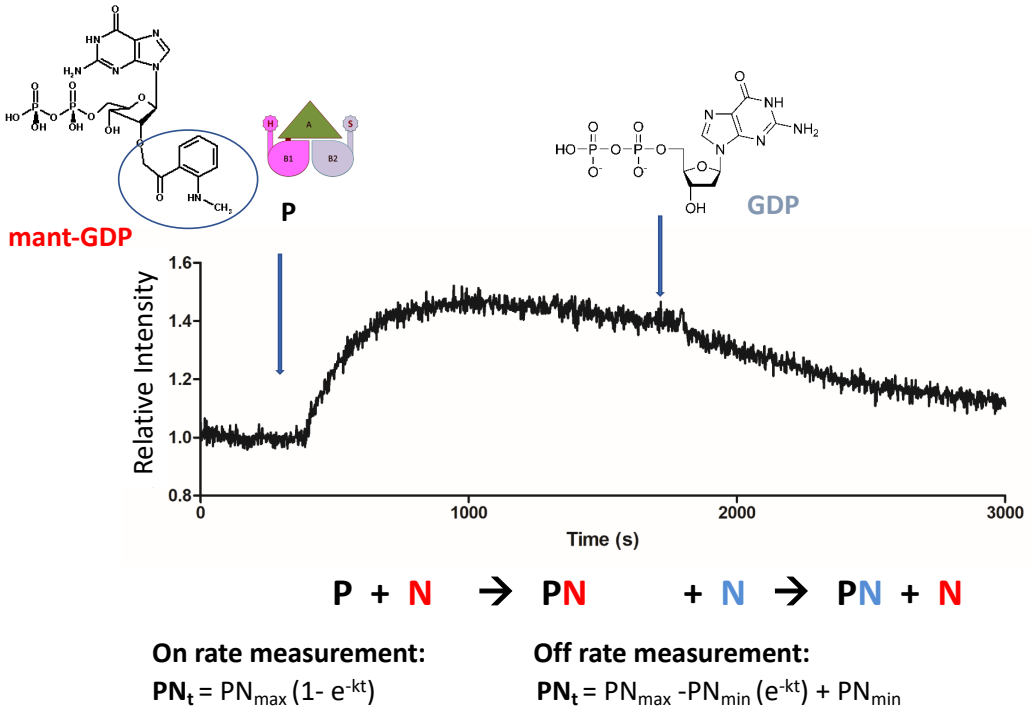


Figure 2.6. Nucleotide exchange assay
On the top is the experimental design for monitoring nucleotide association and dissociation kinetics. The chemical structure of fluorescent GDP analog is shown.
Below, is the description of equations used for obtaining rates of association and dissociation

Methodology

The kinetic measurements for the linked complexes were performed on Fluoromax-4 (Horiba) where the intensity of fluorescence emission by *mant*-labelled GDP/GMPPNP (Jena Bioscience, Germany) at 440 nm was monitored after the excitation at 360 nm (protocol similar to the one used in11). The sample volume was 200 µl in a quartz cuvette (10 × 2 mm path length) and excitation and emission slit widths of 2 nm. *mant*-GDP/GMPPNP (final concentration 800 nM) was present in buffer A50 (50 mM Tris pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$). The protein mix, i.e., a final concentration of 3 µM of the MglA with respective proteins was added in the cuvette at 400 seconds after stabilization of the signal from only *mant*-nucleotide. Consequently, the fluorescence was recorded for 1,400 seconds where the

increase in fluorescence intensity reflected the nucleotide-binding kinetics. At 1,800 seconds, the *mant*-labelled nucleotide was completed out with an excess of unlabelled GDP/GTP (final concentration 500 μM), resulting in the release of *mant*-labelled nucleotide from the protein which manifested as a decrease of fluorescence intensity. For plotting the relative intensities from the measurements, each value was divided by the average of the first 200 readings (400 seconds). These accumulations and decay reactions were fitted to exponential binding equations as given below in GraphPad Prism to estimate the k_{on} and k_{off} values.

$P + N_1 \rightarrow PN_1 + N_2 \rightarrow PN_2 + N_1$, where P represents a protein, N1 is the labelled nucleotide, N2 is the unlabelled nucleotide, and PN denotes the protein-nucleotide complex.

For estimation of k_{on}, $PN_t = PN_{max}\ (1 - e^{-kt})$.

For estimation of k_{off}, $PN_t = PN_{max} - PN_{min}\ (e^{-kt}) + PN_{min}$

Here, PN_t represents the amount of the complex at time t, PN_{max} is the maximum amount of the complex upon the association, and PN_{min} is the minimum amount of the complex after dissociation.

2.7 Binding assays

2.7.1 Analytical size exclusion chromatography

The principle and methodology are the same as size exclusion chromatography. When studying protein interaction in presence of nucleotides, the column is equilibrated with A50 supplemented with 0.1 mM GDP/GTP (Sigma Aldrich) and 2 mM $MgCl_2$. The proteins to be injected are mixed into a 200 μl volume with 1 mM GDP/GMPPNP and 2 mM $MgCl_2$ and injected. The formation of protein complexes results in an increase in mass causing elution volumes to shift. This is further confirmed by running the eluted fractions on SDS PAGE. Since most of the proteins used in this study run aberrantly on SEC, no interpretation of molar masses was made using a standard curve for elution volumes. The molecular weights of complexes were confirmed using SEC-MALS.

2.7.2 Isothermal Titration Calorimetry

Principle

As proteins interact, energy is absorbed or released which is measured by an isothermal calorimeter (ITC). The instrument consists of a reference cell connected to a sample cell (Fig 2.7). The heat change associated with the interacting molecules results in a change in the temperature of the sample cell, which is restored as the reference cell temperature by the instrument. The energy added or removed from the sample cell is measured and consequently, enthalpy (ΔH), entropy (ΔS) and stoichiometry (n) of the molecules are calculated.

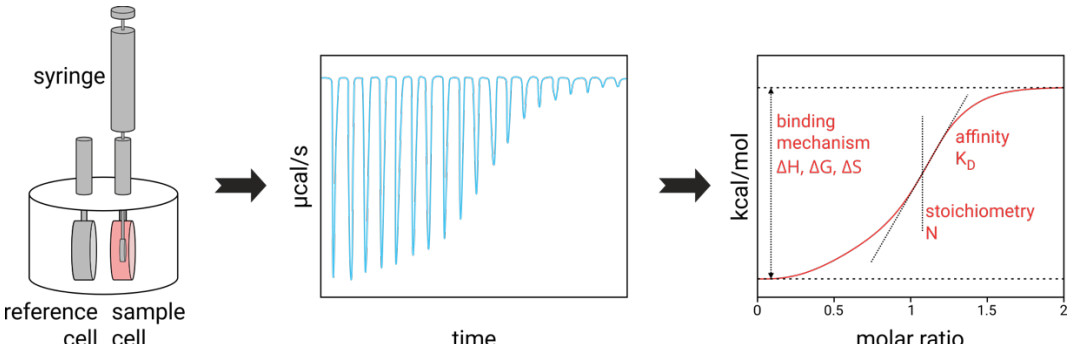


Figure 2.7. Isothermal Titration calorimetry
From left to right, simplified architecture of an ITC instrument, a typical ITC thermogram for an exothermic reaction and calculation of thermodynamic parameters associated with protein-protein interactions using ITC (Reproduced from 2bind.com).

Methodology

The proteins (RomR-REC and RomX) used for interaction studies were dialysed into 20 mM phosphate buffer (K_2HPO_4+KH_2PO_4), pH 8, overnight and the exact concentrations were estimated using Bradford assay. ITC measurements were carried out by titrating 300 µM of RomX (N-terminal hexahistidine tagged construct) into 20 µM of RomR-REC in the sample cell using the MicroCal PEAQ-ITC, using a stirring speed of 750 rpm. 19 injections, 2 seconds each were made. Data for each experiment was modelled using a one-site binding model provided in the MicroCal PEAQ- ITC analysis software with fitted offset control. Error in K_D and N was calculated as the standard error of the mean using three independent experiments.

2.7.3 Fluorescence Anisotropy measurements

Principle

When plane polarized light is incident onto free fluorescent molecules which are tumbling randomly in solution, the emitted light is depolarized. The binding of a ligand/protein to the fluorescent molecule results in a reduction of the random tumbling motion, causing the emitted light to remain polarized. Fluorescence anisotropy is a measure of this difference in the extent of depolarization. Higher anisotropy can be used as a proxy to monitor binding events and can be consequently used to measure binding affinities.

Methodology

To estimate the binding of MglB to nucleotide-bound MglA mutants, fluorescently-labelled nucleotide *mant*-GDP (m-GDP) or *mant*-GMPPCP (m-GCP) (Jena Bioscience) was used (protocol similar to the one used by Baranwal *et al*.[3]). The same methodology was used to monitor the binding between MglA and RomX. The change in anisotropy was measured in Fluoromax-4 (Horiba), in a cuvette with 200 µl sample volume at 25 °C, and excitation and emission slit widths of the 5-nm bandpass. A mix of 400 nM *mant*-labelled nucleotide in buffer (50 mM Tris (pH 8.0), 50 mM NaCl and 2 mM $MgCl_2$) with 2 µM of MglA (all final concentrations) was titrated with increasing concentrations of MglB/RomX. Each reading corresponds to 10 averaged single-point anisotropy values. The initial value of anisotropy of the labelled nucleotide with MglA was subtracted from all the readings. GraphPad Prism was used to plot the values against the concentration of the MglB dimer(since MglA binds to one dimer of MglB) or RomX and to fit the data to the binding equation for the estimation of K_D. The equation used to fit the data points for one site-specific binding is,

$$y=B_{max}\times x\div(K_D+x),$$

where y represents the bound fraction (relative fluorescence anisotropy-reading), B_{max} is the anisotropy value at saturation (obtained by fitting the data to the equation), x is the concentration of MglB dimer and K_D is the binding constant. The relative fluorescence anisotropy values were calculated by dividing all the readings by the max value and subsequently plotted against MglB dimer/RomX concentration. Each data point is an average

of at least three independent measurements, and the standard error is shown in which each data represents three repeats. Error bars represent standard error.

2.7.4 Tryptophan Fluorescence Assay

Principle

Tryptophan is one of the rarely occurring amino acids in proteins. It is known to possess intrinsic fluorescence[17]. Changes in the tryptophan microenvironment cause a change in the Trp fluorescence spectrum and result in a shift of the maximum fluorescence peak as well as a variation of the fluorescence intensity. An exposed tryptophan in solution has an emission maximum (λ_{max}) of around 360 nm whereas when it is buried, it shifts to lower wavelengths (blue shift) along with a change in fluorescence intensity.

Methodology

The experiment was performed in 20 mM phosphate buffer (K_2HPO_4+KH_2PO_4), pH=8. The sample volume was 200 µl in a quartz cuvette (10 × 2 mm path length) and excitation and emission slit widths of 2 nm. The excitation spectra were recorded between 200-300 nm. The emission spectra were recorded at the excitation wavelength of 287 nm and the emission spectra were recorded between 300-450 nm. 10 µM RomX was added to the buffer and the spectra were recorded. Further, it was titrated with 1-20 µM RomR in steps of 1 µM each. The emission maxima (λ_{max}) and the corresponding fluorescence intensity (in CPS, counts per second) were recorded and plotted against the concentration of RomR to estimate a binding curve.

2.8 Protein Crystallography

2.8.1 Protein Crystallization

Principle

Protein crystals consist of periodic arrangements of molecules in space stabilized by intermolecular contacts, giving rise to a crystal lattice[18,19]. The constituent unit cells can be used to generate the crystal using symmetry operations. An asymmetric unit is the smallest component of a unit cell. The lattice type and its corresponding symmetry are defined by its space group. Protein solutions under suitable pH, temperature, salt and precipitant

concentrations, separate from the solution and form stable periodic arrays, yielding protein crystals. A critical step to initiate this process is nucleation. For this, the protein has to gradually transition from undersaturation to supersaturation, with a specific concentration aiding nucleation (Fig 2.8). This is aided by gradual changes in protein and precipitant relative concentrations. This is achieved by several techniques, one of them being vapour diffusion. In this arrangement, protein mixed with the crystallization condition is present in a droplet sealed in a closed chamber containing a reservoir of the same condition. Vapour diffusion between the droplet and chamber results in a concentration of protein in the droplet, causing it to transition to the nucleation zone, forming crystals. Sitting drop and hanging drop are two such techniques of protein crystallization through vapour diffusion.

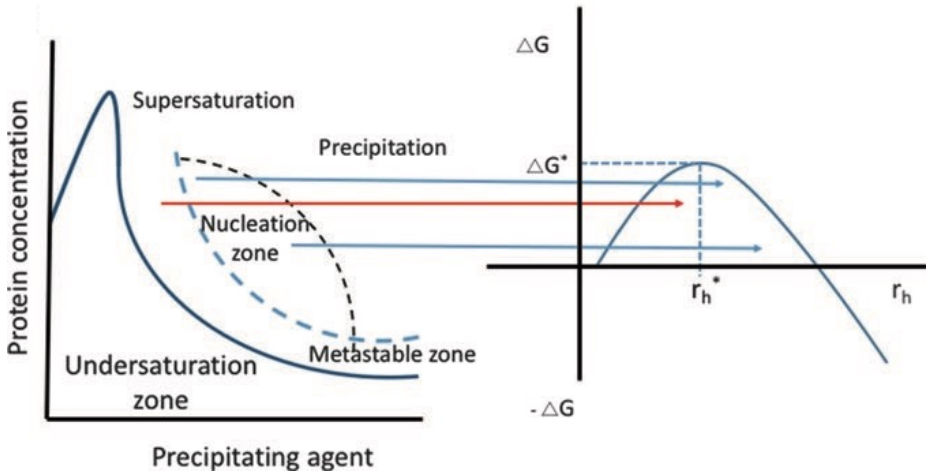


Figure 2.8. Oswald-Miers phase diagram
Nucleation occurring at the supersaturated zone (Reproduced from Moreno, 2017[19]) with the associated energetics

Methodology

Several conditions (about 1400) from commercially available screens (Molecular dimensions, Hampton Research, Jena Biosciences, Table 2.4) were used to prepare protein crystallization sitting drop plates (96 well MRC plate, SWISS-SCI) using Mosquito © robotic dispenser (TTP Labtech) using 100 nl protein solution with 100 nl conditions. Several drop ratios (protein: condition) and protein concentrations were checked. In some cases, additive conditions were added with parent conditions (5 µl and 45 µl, respectively) where hits were obtained. The reproducible conditions were used to set up scaled-up drops (drop size 2-5 µl), in hanging and

sitting drop plates. Finally obtained crystals were flash frozen in liquid nitrogen in the parent condition supplemented with cryoprotectants (20% glycerol, ethylene glycol or PEG E 400).

Table 2.4: Summary of crystallization trials

Construct	**Purpose**	**Crystallization trials**	**Final optimized condition**	**Result**
$\{AB^{WT}\}$ + GTP-ɣS	To ensure formation of functional MglA MglB complex	Main screens 10-15	200 mM Ammonium sulphate, 8% PEG4K	Needle shaped crystals, optimized and data collected
$\{AB_1{}^{\Delta Ct}B_2\}$ +GDP	To determine structure of MglA MglB complex bound to GDP	Main screens 1-15	No hits	-
$\{AB_1B_2{}^{\Delta Ct}\}$ +GDP	To determine structure of MglA MglB complex bound to GDP	Main screens 1-15 Additive Screen	40 mM potassium phosphate, 16% PEG 8K, 20% glycerol	Hexagonal crystals, did not diffract
RomR-REC^{125}	To determine structure of RomR-REC domain	Main screens 1-15 Additive Screen	No hits	-
RomR-REC^{141}	To determine structure of RomR-REC domain	Main screens 10-15	No hits	-
RomR-REC^{125} + RomX (N term His)	To determine structure of RomR-REC domain with RomX	Main screens 1-15 Additive Screen	No hits	-
RomR-REC^{141} + RomX (N term His)	To determine structure of RomR-REC domain with RomX	Main screens 1-15	No hits	-
RomR-REC^{141}Ct + RomX (N term His)	To determine structure of RomR-REC domain with RomX	Main screens 10-15	No hits	-
RomX (C term His)	To determine structure of RomX	Main screens 10-15	5 mM $MgCl_2$, 50 mM HEPES Na pH 7, 25% PEG MME 550	Needle shaped crystals, optimized and data collected
RomY	To determine structure of RomY	Main screens 1-15	8% PEG 3000, 0.1 M KH_2PO_4 +Na_2HPO_4 pH 7	Needle shaped crystals, optimized and data collected

Main Screen	*Catalogue Number*
1	MD1-30
2	MD 1-31
3	MD 1-32
4	MD 1-60
5	JBScreen HTS-II
6	JBScreen HTS-I
7	MD 1-47
8	MD1-40
9	HR2-130
10	HR2-134
11	HR2-248
12	HR2-139
13	HR2-131
14	Rigaku Wizard 1008647
15	Rigaku Wizard 1008648

2.8.2 Data collection and structure refinement

Principle

X-rays are used to produce the diffraction pattern as they hit the electrons in the crystals because their wavelength λ is typically the same order of magnitude (1–100 Å) as the spacing d between planes in the crystal. The diffraction data is recorded as a collection of images. The intensities of the diffraction spots are estimated which provides the structure factor amplitudes ($|F_{hkl}|$) of each reflection (corresponding to unique h, k and l values) in the reciprocal space. Further phase information for the structure factors is obtained through various methods such as molecular replacement, isomorphous replacement using heavy atoms, and anomalous scattering based techniques. The structure factor amplitudes and phase estimates are used to calculate the electron density map using the inverse Fourier transform function and to finally determine the crystal coordinates by the interpretation of the electron density map (Fig 2.9).

Phase information for the structure factors were obtained from the following two methods for the structures determined in this thesis.

A. **Molecular replacement:** Phase information of a homolog of the protein under study, with at least 20-25% sequence identity is used to build a model which is refined to determine the experimental structure.
B. **Single wavelength anomalous diffraction/Multi-wavelength anomalous diffraction (SAD/MAD):** Proteins with no structural homologs are expressed with heavy atom modifications to specific amino acids, e.g., selenomethionine which is then used for phasing[18,19]. An absorption scan of the protein helps select multiple wavelengths, to diffract the crystal. The largest changes in the anomalous scattering factors are at the absorption edges, hence the data is collected at the peak (λ_1) for single anomalous dispersion (SAD) and additionally at inflexion (λ_2) and remote ($\lambda_{3,\ 4}$) for multiple anomalous dispersion (MAD). This data is used to determine the phase of the heavy atom substructure, which can be used to calculate the phase information for the protein structure.

Using these methods the phase information is extracted. This, together with the amplitude information of the diffraction data is used to generate an electron density map. It can be represented as a simple map of continuous density $(2F_o-F_c)$ and the difference maps (F_o-F_c). The backbone atoms and the amino acid residue side chains are built into the electron density map, iteratively, to reduce the values of R_{work} and R_{free} indicating an overall improvement in the quality of the electron density maps. Finally, the model is validated by checking the primary sequence of the protein chain, by checking the real space fit of the side chains into the electron density map and inspecting the prevalent rotamers of the side chains[20]. The Ramachandran plot is used to validate the model by ensuring that there are no residues in the the disallowed conformations. The structure is complete when all the refinement statistics such as R_{work}, R_{free}, Ramachandran outliers and RMSD of all bond lengths and angles are within the permissible range.

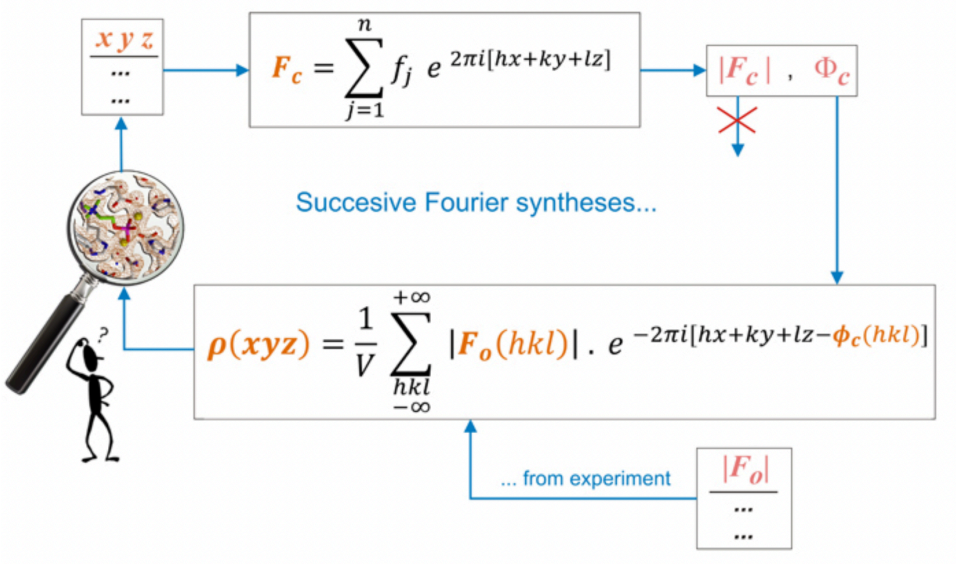


Figure 2.9. Protein structure determination
Schematic showing successive steps of phasing and structure refinement (Reproduced from Cristalografía y Biología Estructural, CSIC)

Methodology

Diffraction data were collected from European Synchrotron Radiation Facility (ESRF), Grenoble, Elettra Sincrotrone Trieste and Diamond Light Source, UK. Data was reduced and scaled using iMOSFLM[20] and AIMLESS[21], respectively, using the CCP4i2[22] package. In the same package, PHASER[23] was used to perform a molecular replacement. For the RomY structure, the AlphaFold model was used for molecular replacement. PHENIX[24] was used for refinements and the model was built in iterative cycles using Coot[25]. RomX anomalous data were collected at a wavelength of 0.9793 with *f'*=8.03 and *f''*=3.85. Autosol and Autobuild wizards of PHENIX[26] were used to build the model which was finally refined to satisfactory R_{work} and R_{free} values. The refinement statistics for each structure were tabulated to validate the quality of the refined data.

2.9 Structure prediction using AlphaFold

Principle

AlphaFold is a web-based tool recently developed to predict 3D structures of proteins based solely on the amino acid sequence[27]. This method utilizes algorithms to predict proteins structures based on the evolutionary, physical and geometric constraints of protein folding. It

applies machine learning where all the physical and biological properties of protein structures are integrated with information from structure databases and multi-sequence alignments to predict near-accurate protein structures.

Methodology

Using the ColabFold pipeline, AlphaFold structure prediction was performed. Sequence of the desired protein was used as the input. For generating multimeric complexes, the multiple chains of identical proteins were fed, based on the stoichiometry. ColabFold was executed with default settings. Out of the five predicted models, Predicted Local Distance Difference Test (pLDDT) scores were used to evaluate the reliability of different parts of the model ((>90, high accuracy; 70-90, good accuracy; 50-70, low accuracy; <50, not interpretable)[27].

References

1. van den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *J Biochem Biophys Methods* **67**, 67–74 (2006).
2. Manil Kanade. Biochemical characterization of two prokaryotic small Ras-like GTPases and their common effector. (IISER, 2019).
3. Baranwal J, Lhospice S, Kanade M, Chakraborty S, Gade PR, Harne S, Herrou J, Mignot T, Gayathri P. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. *PLoS Biol.* 2019 Sep 27;17(9):e3000459. doi: 10.1371/journal.pbio.3000459.
4. Costa, G. L., Bauer, J. C., McGowan, B., Angert, M. & Weiner, M. P. Site-directed mutagenesis using a rapid PCR-based method. *Methods Mol Biol* **57**, 239–248 (1996).
5. Bornhorst JA, Falke JJ. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol*. 2000;326:245-54. doi: 10.1016/s0076-6879(00)26058-8.
6. Magdeldin, S. & Moser, A. Affinity Chromatography: Principles and Applications. www.intechopen.com (200AD).
7. Schmidt, T. G. M., Rgen Koepke, J., Frank, R. & Skerra, A. Molecular Interaction Between the Strep-tag Affinity Peptide and its Cognate Target, Streptavidin. *J. Mol. Biol* vol. 255 (1996).
8. Hong, P., Koza, S. & Bouvier, E. S. P. A review size-exclusion chromatography for the analysis of protein biotherapeutics and their aggregates. *J. Liq. Chromatogr. Relat.* vol. 35 2923–2950 Preprint at https://doi.org/10.1080/10826076.2012.743724 (2012).
9. Kielkopf, C. L., Bauer, W. & Urbatsch, I. L. Bradford assay for determining protein concentration. *Cold Spring Harb Protoc* **2020**, 136–138 (2020).
10. Subadini, S., Hota, P. R., Behera, D. P. & Sahoo, H. Circular Dichroism Spectroscopy: Principle and Application. *Optical Spectroscopic and Microscopic Techniques* 19–33 (2022) doi:10.1007/978-981-16-4550-1_2.
11. Micsonai A, Wien F, Bulyáki É, Kun J, Moussong É, Lee YH, Goto Y, Réfrégiers M, Kardos J. BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315–W322 (2018).

12. Folta-Stogniew, E. & Williams, K. R. Determination of Molecular Masses of Proteins in Solution: Implementation of an HPLC Size Exclusion Chromatography and Laser Light Scattering Service in a Core Laboratory & METHODS REVIEWS. *JBT* vol. 10 (1999).
13. Ingerman, E. & Nunnari, J. A Continuous, Regenerative Coupled GTPase Assay for Dynamin-Related Proteins. *Methods Enzymol* **404**, 611–619 (2005).
14. Kiianitsa, K., Solinger, J. A. & Heyer, W. D. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Anal Biochem* **321**, 266–271 (2003).
15. Remmers, A. E., Posner, R. & Neubig, R. R. Fluorescent guanine nucleotide analogs and G protein activation. *JBC* **269**, 13771–13778 (1994).
16. Tan, Y. & Sun, Q. A mant-GDP dissociation assay to compare the guanine nucleotide binding preference of small GTPases. *Bio Protoc* **11**, (2021).
17. Sindrewicz, P. *et al.* Intrinsic tryptophan fluorescence spectroscopy reliably determines galectin-ligand interactions. *Sci Rep* **9**, (2019).
18. McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallographica Section F:Structural Biology Communications* vol. 70 2–20 Preprint at https://doi.org/10.1107/S2053230X13033141 (2014).
19. Moreno, A. Advanced methods of protein crystallization. in *Methods Mol Biol* vol. 1607 51–76 (Humana Press Inc., 2017).
20. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281 (2011).
21. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *urn:issn:0907-4449* **69**, 1204–1214 (2013).
22. Potterton L, Agirre J, Ballard C, Cowtan K, Dodson E, Evans PR, Jenkins HT, Keegan R, Krissinel E, Stevenson K, Lebedev A, McNicholas SJ, Nicholls RA, Noble M, Pannu NS, Roth C, Sheldrick G, Skubak P, Turkenburg J, Uski V, von Delft F, Waterman D, Wilson K, Winn M, Wojdyr M. CCP 4 i 2: The new graphical user interface to the CCP 4 program suite. *Acta Crystallogr D Struct Biol* **74**, 68–84 (2018).
23. Bunkóczi G, Echols N, McCoy AJ, Oeffner RD, Adams PD, Read RJ. Phaser.MRage: Automated molecular replacement. *Acta Crystallographica Section D: Biological Crystallography* vol. 69 2276–2286 (2013).
24. Terwilliger, T. *et al.* PHENIX Wizards and Tools The PHENIX project *Computational Crystallography Initiative (LBNL) Los Alamos National Lab (LANL)*.
25. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
26. Guerrero SA, Hecht HJ, Hofmann B, Biebl H, Singh M. Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. *Appl Microbiol Biotechnol.* 2001 Sep;56(5-6):718-23. doi: 10.1007/s002530100690.
27. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D.Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2

Chapter 3: **Characterization of a non-canonical Walker B motif in MglA**

3.1 Introduction

MglA, like other small Ras-like GTPases, belongs to the family of P-loop NTPases. NTPases are enzymes which hydrolyse nucleotide triphosphates (e.g., ATP or GTP) to diphosphates (ADP or GDP). They are broadly subdivided into Kinase-GTPase and ASCE (additional strand catalytic glutamate) ATPase divisions [1,2]. Further, the GTPase division is subdivided into two classes, namely, *TRAFAC* (*Tra*nslation *Fac*tors) and *SIMIBI* (*Si*gnal recognition particle, *Mi*nD and *Bi*oD).

All these aforementioned P-loop NTPases have certain conserved motifs which enable their function of binding and hydrolysing NTPs (Fig 3.1 A). One of them is the Walker A motif which is essential for binding and positioning the phosphates of NTP and hence known as the P-loop or phosphate loop [3,4]. The T/S of the consensus GXXXXGK[T/S] motif is involved in direct Mg^{2+} coordination[2,5]. The other conserved motif is the Walker B motif (consensus sequence "hhhh[**D/E**]") which has a conserved aspartate or glutamate which contributes to maintaining Mg^{2+} co-ordination [2,5]. This interaction may be direct or indirect via a water molecule. The "hhhh" portion of the motif consists of hydrophobic residues which form β-strand with the D/E present at the C-terminal end of the motif. The Walker B acidic residue is critical for nucleotide hydrolysis and even sometimes essential for nucleotide binding [6–8].

MglA does not possess a conserved acidic residue in the equivalent of the Walker B motif in small Ras-like GTPases (G3 motif or a conserved "DxxG" motif) [9,10] (Fig 3.1 B). Instead, it has a threonine or a glycine in the G3 motif (residue 78) [11,12]. However, the structural superposition of MglA in GTP bound state with eukaryotic Ras revealed that the Asp-58 residue of MglA which lies in the adjacent strand to the conserved Walker B strand was capable of indirectly maintaining the Mg^{2+} co-ordination via a water molecule (Fig 3.1 D,E). Since this was a non-canonical Walker B equivalent candidate, the role of this residue was validated experimentally by carrying out biochemical assays in the presence of MglB, its cognate GAP and atypical GEF[12–16]. Further, we investigated the Walker B residues and their conformations in representatives from different classes of P-loop NTPases. We identified conserved features pertaining to the role of Walker B aspartates and glutamates in maintaining Mg^{2+} coordination and their positions in the structures of different GTPases and ATPase classes. The conformation of Asp-58 in the newly identified Walker B motif of MglA was seen to switch across the GDP and GTP bound states, which could probably be acting as an allosteric regulator for nucleotide binding. We identified similar examples from the P-loop GTPases where the

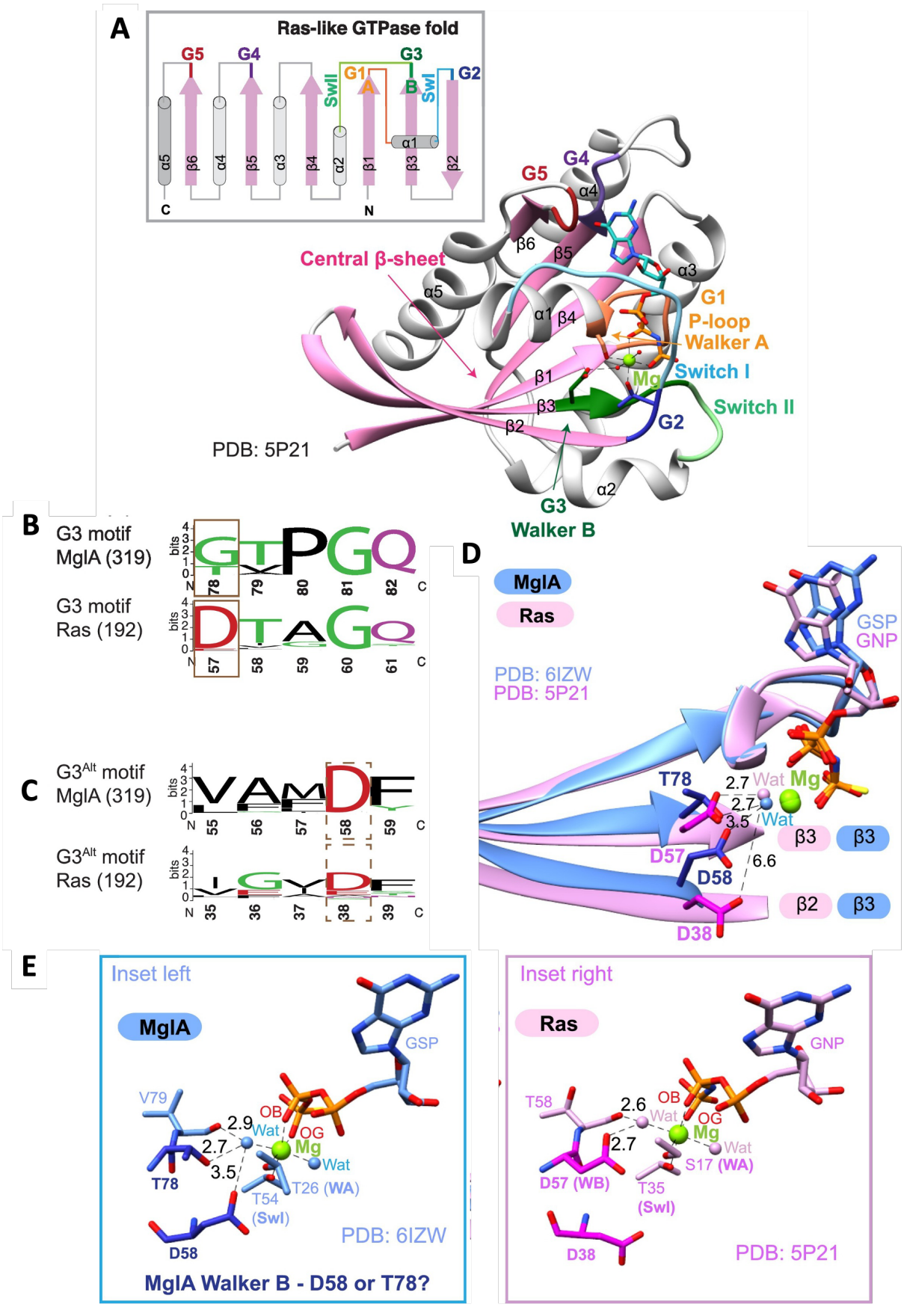
switch from GDP to GTP states was associated with a conformational change in Walker B residues, which was critical for the GTPase function.

3.2 Results

3.2.1 Identification of the probable Walker B residues in MglA

MglA, a family of prokaryotic small Ras-like GTPases, has been divided into 5 classes phylogenetically [11]. The Asp-58 is a residue in a conserved motif "LFFDF" as present in *Myxococcus* MglA, MglA and other group 1 members, while in other groups, it is part of a conserved "VAMDF" motif (Fig 3.1 C)[17,18]. This motif is a part of the β-strand at the binding interface of MglB [12,14,15]. We identified this motif as an alternative to the conserved G3 of Walker B motif present in small Ras-like GTPases and referred to this as the $G3^{Alt}$ motif, of which the Asp-58 was a part of (Fig 3.1 D). Hence, we compared the role of the conserved Thr-78 at the G3 motif and the non-canonical Asp-58 at the $G3^{Alt}$ motif in driving GTP hydrolysis and nucleotide binding in MglA to identify the functional Walker B motif in the MglA family (Fig 3.1 D,E).

Figure 3.1. $G3^{Alt}$ motif as the alternate Walker B motif in the MglA family
A. Topology of small Ras-like GTPases (top) with β-strands (pink) and α-helices (grey), G1 (orange), G2 (dark blue), G3 (dark green), G4 (purple), G5 (dark red), Switch I (Sw I; light blue) and Switch II (Sw II; light green) motifs highlighted and in the H-Ras structure (below, PDB ID: 5P21) with the central β-sheet shown in pink.
B. WebLogos showing the alignment of G3 motifs in prokaryotic MglA family and eukaryotic small Ras-like GTPases. The number of sequences in the alignment is shown in brackets. The Asp-57 in the G3 motif of Ras and the corresponding residue in MglA is highlighted in brown boxes. The residue numbering corresponds to MglA (top) and H-Ras (bottom).
C. WebLogos showing the alignment of $G3^{Alt}$ motifs in prokaryotic MglA and eukaryotic small Ras-like GTPases. The number of sequences in the alignment is shown in brackets. Asp-58 in the $G3^{Alt}$ motif of MglA and the corresponding residue in Ras are highlighted in brown dashed boxes. The residue numbering corresponds to MglA (top) and H-Ras (bottom).
D. Structural superposition of MglA bound to GTP-γS (blue; PDB ID: 6IZW) and H-Ras bound to GMPPNP (pink; PDB ID: 5P21). Thr-78 and Asp-58 of MglA and Asp-57 and Asp-38 of H-Ras are shown in sticks (dark blue and magenta, respectively). Green spheres denote Mg^{2+}. Distances, in Å, from the residues to the coordinating water molecules and secondary structure elements are labelled.
E. Insets show MglA and H-Ras structures (left and right, respectively) where the residues involved in Mg^{2+} coordination are shown in sticks. Interacting waters (Wat), Switch I (SwI), Walker A (WA), Walker B (WB) and O-β and O-γ of the nucleotides are highlighted.


A
Ras-like GTPase fold
G5
G4
G3
G1
G2
SwII
SwI
α5
β6
α4
β5
α3
β4
α2
β1
α1
β3
β2
C
N
Central β-sheet
PDB: 5P21
P-loop
Walker A
Switch I
Switch II
Walker B
B
G3 motif
MglA (319)
G3 motif
Ras (192)
C
G3Alt motif
MglA (319)
G3Alt motif
Ras (192)
D
MglA
Ras
PDB: 6IZW
PDB: 5P21
GSP
GNP
T78
D57
D58
D38
Wat
Mg
β3
β2
E
Inset left
MglA
V79
T78
D58
T54
(SwI)
T26 (WA)
GSP
PDB: 6IZW
MglA Walker B - D58 or T78?
Inset right
Ras
T58
D57 (WB)
T35
(SwI)
S17 (WA)
GNP
D38
PDB: 5P21

3.2.2 Cloning and purification of MglA Walker B mutants

To validate the role of Asp-58 and Thr-78 as the probable Walker B acidic residues, we mutated these respective residues to alanines (Fig 3.1 C). The resultant constructs were named $MglA^{D58A}$ and $MglA^{T78A}$, respectively (Methods: 2.2.1). The proteins were purified using the C-terminal hexahistidine tag (Ni-NTA affinity chromatography), followed by size exclusion chromatography (with Manil Kanade) (Methods: 2.3). The proteins were observed to be well-folded and monomeric from the size exclusion chromatography profiles (Fig 3.2 A).

To investigate the nucleotide bound to these mutants, we performed high-pressure liquid chromatography using the DNAPac PA200 ion exchange column (Fig 3.2 B, Methods 2.5.4). It was observed that both the mutants, as well as wild type MglA, came bound to GDP, without GDP being added at any stage of purification of these proteins. Qualitatively, it was observed that $MglA^{D58A}$ contained significantly less amount of bound GDP as compared to $MglA^{T78A}$ and wild-type MglA, although the same concentrations and volumes of protein were used to prepare the extracts for these runs (Fig 3.2 B).

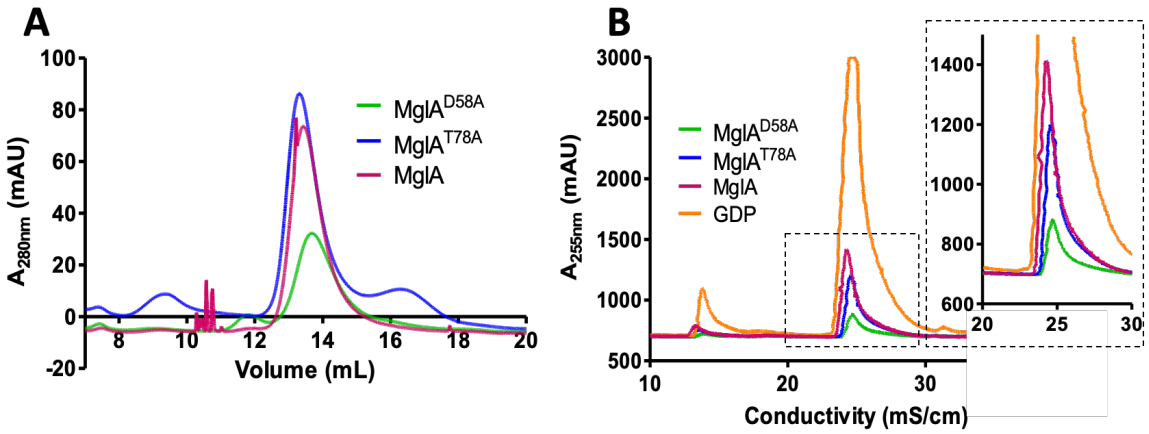


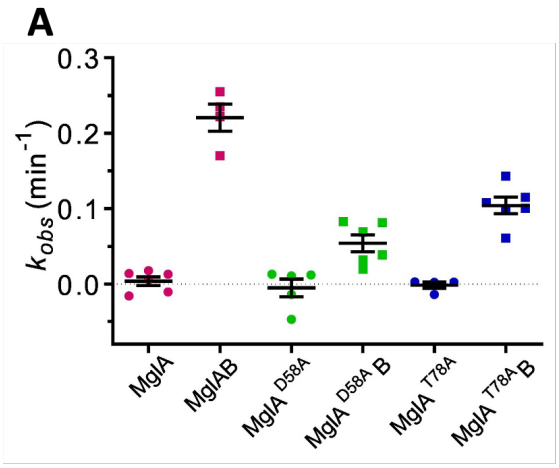
Figure 3.2. Characterization of the probable Walker B mutants

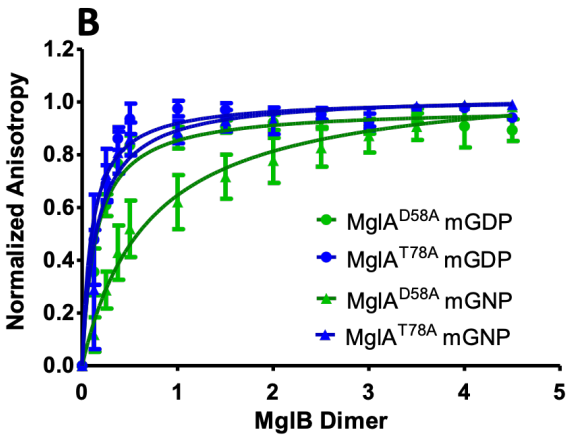
A. Analytical Superdex 75 profiles of MglA (magenta), $MglA^{D58A}$ (green) and $MglA^{T78A}$ (blue). B. HPLC profile of extracts from MglA (magenta), $MglA^{D58A}$ (green) and $MglA^{T78A}$ (blue). GDP used as a standard is shown in orange.

3.2.3 Validation of the key Walker B residue

Using the NADH-coupled GTPase assay, we quantified the rate of GDP released for the MglA mutants (with Manil Kanade, Methods: 2.6.1). Since the intrinsic hydrolysis activity of MglA is very low, we used the MglA GAP, MglB, purified and mixed with the MglA constructs to determine the effect of these mutations on GTP hydrolysis rates of MglA [14]. We observed that in presence of MglB, both $MglA^{D58A}$ and $MglA^{T78A}$ mutants showed two to three-fold lower GTPase activity compared to that of wild-type MglA (comparing the average k_{obs} from multiple repeats) (Fig 3.3 A). Since the catalytic activity of MglA mutants was not completely abolished, we checked the effect of a typical Walker B mutation on other small Ras-like GTPases. We observed that in HRas, like many other GTPases, the Walker B mutation reduces the GTP hydrolysis rates but does not completely abolish it [8]. Our observations were consistent with this result but due to the same effect in both the mutants, we could not identify the Walker B residue from this assay.

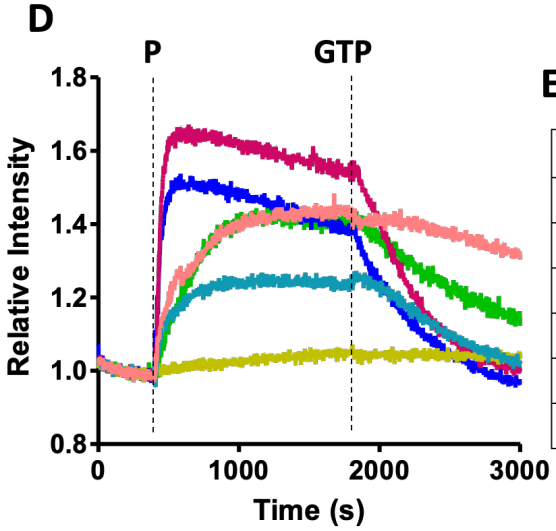
Since we performed the GTPase assay in presence of MglB, the observed decrease in GTP hydrolysis rates could also be possible due to the effect of the mutations on the affinity of MglB to the MglA mutants. Consistent with the observations in wild-type MglA and MglB, MglB binds MglA in both the GDP and GTP bound states, indicating the occurrence of the β-screw movement which results in the formation of the MglB binding interface [14]. We consequently went ahead to check the effect of these mutations on the MglB binding affinity in the presence of GDP and GMPPCP, a non-hydrolysable analogue of GTP (Fig 3.3 B). We performed fluorescence anisotropy measurements of the mutants with *mant*- labelled nucleotides upon titration with increasing amounts of MglB (Methods: 2.7.3). The affinities of MglB for the *mant*-GDP bound state of $MglA^{D58A}$ and $MglA^{T78A}$ (0.15 ± 0.02 μM and 0.10 ± 0.02 μM, respectively) were comparable to the value for MglA (0.13 ± 0.06 μM) (Fig 3.3 C). However, the affinities of MglB in the *mant*-GMPPCP bound states were lesser for $MglA^{D58A}$ (0.70 ± 0.13 μM) than $MglA^{T78A}$ (0.15 ± 0.03 μM) which was similar to wild-type MglA (0.10 ± 0.01 μM), implying that the interaction between MglB and MglA driven by GTP binding was reduced in the $MglA^{D58A}$ mutant.





C

	K_d (mant-GDP) (μM)	K_d (mant-GMPPCP)(μM)
MglA^{D58A}	0.15±0.02	0.7±0.1
MglA^{T78A}	0.10±0.02	0.15±0.03
MglA	0.13±0.06	0.10±0.01



E

	k_{on} (μM^{-1}s^{-1})	k_{off} (s^{-1})
MglA^{D58A} +MglB (Mg^{2+})	0.005 ± 0.002	0.001 ± 0.0002
MglA^{T78A} +MglB (Mg^{2+})	0.02 ± 0.007	0.003 ± 0.002
MglA + MglB (Mg^{2+})	0.02 ± 0.003	0.002 ±0.0001
MglA^{D58A} +MglB	Ambiguous	Ambiguous
MglA^{T78A} +MglB	0.01 ± 0.001	0.0009 ± 0.0003
MglA + MglB	0.006 ± 0.0007	Ambiguous

Figure 3.3. Identification of Asp-58 of MglA as the key Walker B motif

A. Comparison of k_{obs} values for MglA (dark pink), MglA^{D58A} (green) and MglA^{T78A} (blue). Circles and squares represent the activities in the absence and presence of MglB, respectively. The mean and standard error is shown in black lines (long and short horizontal lines, respectively). B. Fluorescence anisotropy measurements for MglB titrated against MglA^{D58A} bound to *mant*-GDP (green triangles) and *mant*- GMPPNP (green circles), and MglA^{T78A} bound to *mant*-GDP (blue triangles) and *mant*-GMPPNP (blue circles). K_D values and the respective standard errors are tabulated in panel C.

D. Kinetics of nucleotide-binding to MglA and mutants based on the increase in fluorescence of *mant*-GDP upon binding to protein after 400 seconds (point of addition labelled as P with a dotted line), followed by addition excess of unlabelled GTP at 1800 s (point of addition labelled as GTP with a dotted line). The k_{on} and k_{off} values and the respective standard errors are tabulated in panel E.

We further monitored fluorescence changes upon binding of the MglA mutants to *mant*-GDP, followed by the decrease in fluorescence upon the addition of an excess of GTP (Fig 3.3 D, Methods: 2.6.2). We compared the association and dissociation rates (k_{on} and k_{off} rates, respectively) for the mutants in the presence of MglB with that of MglA with and without Mg^{2+}(Fig 3.3 E). The hypothesis was that both Asp-58 and Thr-78 play an important role in Mg^{2+} coordination when bound to MglB (Fig 3.3 E). To validate the role of Mg^{2+}, we performed the same set of assays in the presence and absence of Mg^{2+}.

In presence of an excess of Mg^{2+} (5 mM), both k_{on} for *mant*-GDP (800 nM) binding and k_{off} for *mant*-GDP release upon addition of excess GTP was comparable for $MglA^{T78A}$ to that of the wildtype MglA (Fig 3.3 E). MglA showed a markedly slower rate of association of *mant*-GDP in absence of Mg^{2+} and very slow exchange with GTP. This observation was similar for the $MglA^{T78A}$ mutant. The binding kinetics of $MglA^{D58A}$ with *mant*-GDP in the presence of Mg^{2+} matched with that of MglA in the absence of Mg^{2+}, which shows that the Mg^{2+} coordination was disrupted in the mutant. Hence, the $MglA^{D58A}$ mutant appeared to be defective in nucleotide exchange compared to that of $MglA^{T78A}$. This validates the role of Asp-58 of the $G3^{Alt}$ motif playing the role of a functional Walker B residue as compared to Thr-78 of the G3 motif.

3.2.4 Comparison with Walker B in other P-loop NTPases

We performed a comprehensive structural analysis to compare the position of Walker B residue in other families of P-loop GTPases and ATPases (with Manil Kanade). We chose representative candidates from three divisions of P-loop NTPases, namely TRAFAC, SIMIBI (GTPases) and ASCE (ATPases). All these structures had a conserved core consisting of a central β-sheet surrounded by α-helices. We performed a superposition of these structures onto the structure of HRas p21 (PDB: 5P21, bound to GMPPNP) and numbered the strands in the central β-sheet (Fig 3.4 A). The strand preceding the P-loop was numbered 1 and the consequent strands were numbered according to their order of occurrence. “P” or “A” refers to the orientation (parallel or antiparallel) with respect to the previous strand.

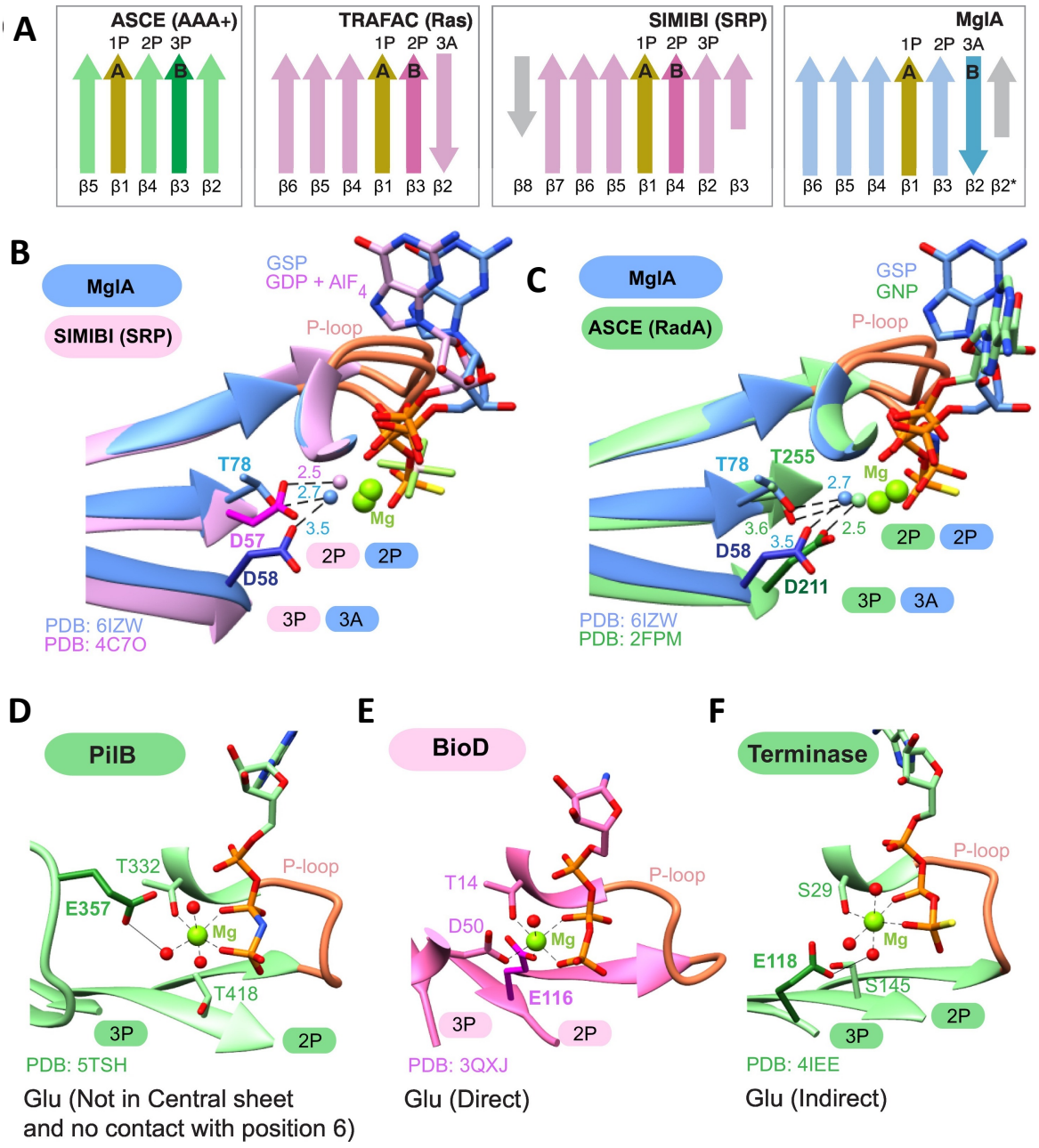


Figure 3.4. Walker B conformations in P-loop NTPases

A. Central β sheet topology diagram of major P-loop NTPase families. ASCE (green; representative: AAA+), TRAFAC (pink; representative H-Ras), SIMIBI (pink; representative: SRP) and MglA (blue), with Walker A and Walker B related strands highlighted as A and B, respectively. Strands are labelled according to the nomenclature of the fold (below) and our nomenclature (above).

B. Superposition of MglA (blue; PDB: 6IZW) and SIMIBI (representative SRP; pink, PDB: 4C7O). Asp-58 (sticks; dark blue) of MglA is present in strand 3A, while Asp-190 (sticks; purple) of SRP is present in strand 2P. Distances between residues and position 6 water are labelled in black dotted lines.

C. Superposition of MglA (blue; PDB: 6IZW) and ASCE (representative RadA; green; PDB: 2FPM). Asp-211 of RadA (sticks; dark green) is present in strand 3P. (d–f) Variation in residues for magnesium coordination in P-loop NTPases. Residues involved in Mg^{2+} coordination are shown in stick representation. Walker B residue is highlighted in a darker shade (dark green or purple). Examples of

coordination by glutamate outside the central β sheet and no contact with position 6 water (PilB; PDB ID: 5TSH; (D)), direct coordination by glutamate (BioD; PDB ID: 3QXJ; (E)) and indirect coordination by glutamate (terminate; PDB ID: 4IEE; (F)) are shown.

The members of TRAFAC or SIMIBI classes had the Walker B residue at the tip of strand 2, the canonical Walker B position. This strand 2 was parallel to the rest of the strands of the β-sheet (Fig 3.4 B). However, in the ASCE family, Walker B was present in strand 3, similar to the position in MglA (Fig 3.4 C). The only difference between the two was that this strand was parallel in the ASCE members whereas, in MglA, this strand was antiparallel with respect to the central β-sheet.

Interestingly, an exception was observed in the PilB/PilT subfamily of ATPases where instead of being a part of the central β-sheet, the Walker B glutamate was present on a loop interacting with water molecules which in turn coordinates the Mg^{2+} ion at the ATP binding pocket [19]. It was observed that Walker B aspartates generally interacted with a water molecule thereby indirectly satisfying the Mg^{2+} coordination. However, in the case of glutamates, the residue directly interacted with Mg^{2+} thereby satisfying the coordination. There were only two exceptions identified in the GP2 terminus of bacterial virus SF2 (PDB ID: 4IEE) and SV40 helicase (PDB ID: 1SVM), where the Walker B glutamate indirectly contacted the Mg^{2+} via a water molecule (Fig 3.4 E, F). However, the presence of indirect contact via water with Mg^{2+} in the case of Walker B aspartates was ubiquitous with no exceptions.

3.2.5 Regulation of the conformation of Walker B residues

As mentioned earlier, the $G3^{Alt}$ motif of MglA, of which Walker B Asp-58 forms a part of the interface between MglA and MglB. This strand flips and undergoes a β-screw movement upon GTP binding and this conformation is stabilised by interaction from MglB. This conformational change positions the Asp-58 in a position favourable for coordinating the Mg^{2+} residue at the active site, thereby acting as an allosteric regulator for GTP binding (Fig 3.5 A). From our previous analysis of the structural superposition of representative P-loop NTPases, we looked for cases where a conformational change in the Walker B residue existed between the diphosphate (GDP/ADP) and triphosphate (ADP/ATP) bound states.

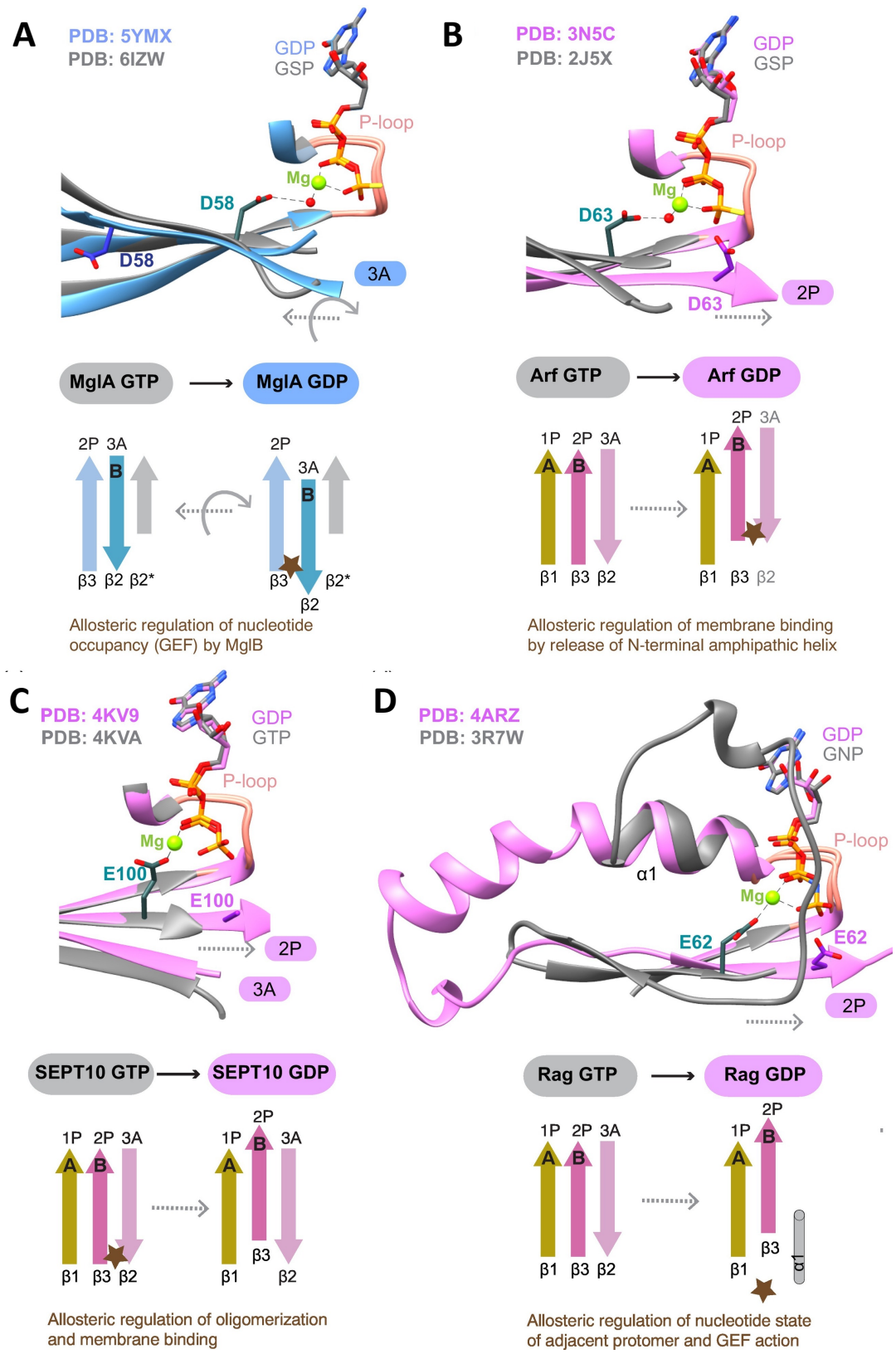
In the case of Arf GTPases, a member of the small Ras-like GTPases (TRAFAC class), we observed a conformational change across GDP and GTP bound states, where in the GDP state, a relative movement of the 2P and 3A strands causes the Walker B aspartate to face away from

the Mg^{2+} co-ordinating water molecule, to which it interacts in the GTP state (Fig 3.5 B). This relative movement of strands 2 and 3 across the GTP and GDP states causes a global conformational change in the molecule causing the release of an N-terminal amphipathic helix which is involved in recruiting it to membranes [20–22]. Hence this conformational change contributes to the allosteric regulation of membrane binding of Arf GTPase, which is relevant for its function.

Interestingly, for the Septin family, another member of the TRAFAC class, strand 2 undergoes movement between the GDP and GTP states, which causes the Walker B glutamate to face the Mg^{2+} at the catalytic pocket (Fig 3.5 C). This conformational change acts as the allosteric regulator of oligomerization and membrane binding [23].

Further, in the case of heterodimeric Rag GTPases (RagA/RagC and Gtr1p/Gtr2p), the RagA or RagB forms a heterodimer with RagC or RagD [24,25]. At any point in time, either RagA/RagB or RagC/RagD binds GTP while the other subunit remains bound to GDP. This alternate nucleotide conformation across the heterodimer is regulated by conformational changes in the α_1 helix of one of the protomers, due to steric clashes from the GTP-bound state of the other protomer. This movement of the α_1 helix causes movement of strand 2, of which the Walker B glutamate is a part, which also undergoes a conformational change across GDP and GTP states (Fig 3.5 D).

Figure 3.5. Conformational changes in Walker B across GTP to GDP-bound states
Structural superposition of the GTP-equivalent (active) and GDP bound (inactive) states (A–D), highlighting the movement of Walker B between the two states, and the schematic representation of the movement.
A. MglA-GSP (gray; PDB ID: 6IZW) and MglA-GDP (blue; PDB ID: 5YMX)
B. Arf6-GTP (gray; PDB ID: 2J5X) and Arf6-GDP (pink; PDB ID: 3N5C)
C. SEPT10-GTP (grey; PDB ID: 4KV9) and SEPT10-GDP (pink; PDB ID: 4KVA). E100 labelled in the GDP conformation is truncated till the CB atom in the crystal structure.
D. RagC-GTP (grey; PDB ID: 3R7W) and RagC-GDP (pink; PDB ID: 4ARZ).


A
PDB: 5YMX
PDB: 6IZW
GDP
GSP
P-loop
Mg
D58
D58
3A
MglA GTP
MglA GDP
2P 3A
B
β3 β2 β2*
2P
3A
B
β3
β2*
β2
Allosteric regulation of nucleotide occupancy (GEF) by MglB
B
PDB: 3N5C
PDB: 2J5X
GDP
GSP
P-loop
Mg
D63
D63
2P
Arf GTP
Arf GDP
1P 2P 3A
A
B
β1 β3 β2
2P 3A
1P
B
A
β1 β3 β2
Allosteric regulation of membrane binding by release of N-terminal amphipathic helix
C
PDB: 4KV9
PDB: 4KVA
GDP
GTP
P-loop
Mg
E100
E100
2P
3A
SEPT10 GTP
SEPT10 GDP
1P 2P 3A
A
B
β1 β3 β2
2P
1P
B
3A
A
β3
β1
β2
Allosteric regulation of oligomerization and membrane binding
D
PDB: 4ARZ
PDB: 3R7W
GDP
GNP
P-loop
α1
Mg
E62
E62
2P
Rag GTP
Rag GDP
1P 2P 3A
A
B
β1 β3 β2
2P
1P
B
A
β3
β1
α1
Allosteric regulation of nucleotide state of adjacent protomer and GEF action

3.3 Conclusion

This study identifies a new non-canonical Walker B motif in the MglA family of prokaryotic small Ras-like GTPases, which is a part of the $G3^{Alt}$ motif ([L/V][F/A][F/M]D). The Walker B aspartate is positioned by MglB binding which in turn acts as an allosteric regulator of nucleotide binding. Further, the Thr-78 present in the canonical G3 motif also plays a role in the catalytic activity of MglA. This relative position of threonine in G3 and Asp in the $G3^{Alt}$ motifs, respectively, is similar to what is observed in the ASCE family of ATPases. This is a feature of convergent evolution where MglA, despite belonging to the TRAFAC class, consists of similar residues as the ASCE ATPases at the nucleotide-binding pocket assisting nucleotide binding and hydrolysis.

A recent analysis of the catalytic mechanisms of nucleotide hydrolysis highlights the importance of the H-bond network between the Ser/Thr of Walker A and the Asp/Glu of Walker B[26,27]. The active site residues (Walker A and Walker B) hold the NTP at the catalytic pocket with an eclipsed conformation between β and ɣ-phosphates. The binding of the NTPase to an activator (GAP), which inserts a catalytic Arg/Lys in the active site, induces a twist in the ɣ-phosphate, thereby weakening the bond between β and ɣ-phosphate. This causes the ɣ-phosphate to transition into a more planar conformation. Further, the ɣ-phosphate oxygen, which coordinates the Mg^{2+}, is displaced. This causes a net increase in the local positive charge. The GAP insertion also constricts the catalytic pocket, causing a swift relocation of a proton from Ser/Thr of Walker A to Asp of Walker B, which are now connected by a shorter and lower barrier H-bond. The resultant Walker A alkoxide abstracts a proton from the catalytic water leaving the OH^- to attack the ɣ-phosphate of the NTP, triggering hydrolysis. Consequently, the negative charge on the β-phosphate abstracts a proton from Ser/Thr of Walker A which is further neutralized when the proton from Walker B is restored. Such a mechanism requires proximity and a low-barrier H-bond network between Walker A and Walker B. Similarly, the non-canonical Asp-58 and the T-78 in the canonical Walker B position in MglA, or other ASCE ATPases, could also contribute to such a H-bond network to drive catalysis. This could explain why Thr-78 was critical for MglA catalytic activity in presence of its activator, MglB.

It has been understood that the Arf/Sar/Gα lineage of GTPases has evolved from the MglA family [11,28,29] and during the course of evolution has evolved the G3 and $G3^{Alt}$ motifs to facilitate nucleotide binding. It is probable that eventually, in the prokaryotic GTPases, the Thr in the G3 motif was replaced with Asp to satisfy the Mg^{2+} coordination facilitating nucleotide

binding. Interestingly, the feature of conformational changes in Walker B as a regulator of nucleotide binding is also observed in Rag and Arf families, both of which have been believed to have evolved from their ancestor, MglA. In this context, it further becomes interesting to dissect the mechanism of MglB binding to MglA and its consequent activation.

References

1. Leipe, D. D., Koonin, E. v. & Aravind, L. Evolution and classification of P-loop kinases and related proteins. *J Mol Biol* **333**, 781–815 (2003).
2. Leipe, D. D., Wolf, Y. I., Koonin, E. v. & Aravind, L. Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol* **317**, 41–72 (2002).
3. Kinoshita, K., Sadanami, K., Kidera, A. & Go, N. Structural motif of phosphate-binding site common to various protein superfamilies: all-against-all structural comparison of protein-mononucleotide complexes. *Protein Eng* vol. 12 https://academic.oup.com/peds/article/12/1/11/1513093 (1999).
4. Saraste, M., Sibbald, P. R. & Wittinghofer, A. The P-loop — a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**, 430–434 (1990).
5. Wendler, P., Ciniawsky, S., Kock, M. & Kube, S. Structure and function of the AAA+ nucleotide-binding pocket. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1823**, 2–14 (2012).
6. Nirwan N, Singh P, Mishra GG, Johnson CM, Szczelkun MD, Inoue K, Vinothkumar KR, Saikrishnan K. Hexameric assembly of the AAA+ protein McrB is necessary for GTPase activity. *Nucleic Acids Res*. 2019 Jan 25;47(2):868-882. doi: 10.1093/nar/gky1170
7. Chiang P, Sampaleanu LM, Ayers M, Pahuta M, Howell PL, Burrows LL. Functional role of conserved residues in the characteristics secretion NTPase motifs of the *Pseudomonas aeruginosa* type IV pilus motor proteins PilB, PilT and PilU. *Microbiology (N Y)* **154**, 114–126 (2008).
8. John J, Rensland H, Schlichting I, Vetter I, Borasio GD, Goody RS, Wittinghofer A. Kinetic and structural analysis of the Mg(2+)-binding site of the guanine nucleotide-binding protein p21H-ras. *J Biol Chem*. 1993 Jan 15;268(2):923-9.
9. Bourne HR, Sanders DA, McCormick F. The GTPase superfamily: conserved structure and molecular mechanism. *Nature*. 1991 Jan 10;349(6305):117-27. doi: 10.1038/349117a0.
10. Dever, T. E., Glynias, M. J. & Merrick, W. C. GTP-binding domain: three consensus sequence elements with distinct spacing. *Proceedings of the National Academy of Sciences* **84**, 1814–1818 (1987).
11. Wuichet, K. & Søgaard-Andersen, L. Evolution and diversity of the ras superfamily of small GTPases in Prokaryotes. *Genome Biol Evol* **7**, 57–70 (2014).
12. Miertzschke M, Koerner C, Vetter IR, Keilberg D, Hot E, Leonardy S, Søgaard-Andersen L, Wittinghofer A. Structural analysis of the Ras-like G protein MglA and its cognate GAP MglB and implications for bacterial polarity. *EMBO J*. 2011 Aug 16;30(20):4185-97. doi: 10.1038/emboj.2011.291.
13. Zhang, Y., Franco, M., Ducret, A. & Mignot, T. A bacterial ras-like small GTP-binding protein and its cognate GAP establish a dynamic spatial polarity axis to control directed motility. *PLoS Biol* **8**, (2010).
14. Baranwal J, Lhospice S, Kanade M, Chakraborty S, Gade PR, Harne S, Herrou J, Mignot T, Gayathri P. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell

polarity oscillations in bacterial motility. *PLoS Biol.* 2019 Sep 27;17(9):e3000459. doi: 10.1371/journal.pbio.3000459.

15. Galicia C, Lhospice S, Varela PF, Trapani S, Zhang W, Navaza J, Herrou J, Mignot T, Cherfils J. MglA functions as a three-state GTPase to control movement reversals of *Myxococcus xanthus. Nat Commun*. 2019 Nov 22;10(1):5300. doi: 10.1038/s41467-019-13274-3.
16. Leonardy S, Miertzschke M, Bulyha I, Sperling E, Wittinghofer A, Søgaard-Andersen L. Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. *EMBO J* **9**, 2276–2289 (2010).
17. Manil Kanade. Biochemical characterization of two prokaryotic small Ras-like GTPases and their common effector. *MS Thesis* (IISER, 2019).
18. Kanade, M., Chakraborty, S., Shelke, S. S. & Gayathri, P. A Distinct Motif in a Prokaryotic Small Ras-Like GTPase Highlights Unifying Features of Walker B Motifs in P-Loop NTPases. *J Mol Biol* **432**, 5544–5564 (2020).
19. McCallum, M., Tammam, S., Khan, A., Burrows, L. L. & Lynne Howell, P. The molecular mechanism of the type IVa pilus motors. *Nat Commun* **8**, (2017).
20. Pasqualato, S., Renault, L. & Cherfils, J. Arf, Arl, Arp and Sar proteins: A family of GTP-binding proteins with a structural device for 'front-back' communication. *EMBO Rep* **3**, 1035–1041 (2002).
21. Pasqualato, S., Ménétrey, J., Franco, M. & Cherfils, J. The structural GDP/GTP cycle of human Arf6. *EMBO Rep* **2**, 234–238 (2001).
22. Hillig RC, Hanzal-Bayer M, Linari M, Becker J, Wittinghofer A, Renault L. Structural and Biochemical Properties Show ARL3-GDP as a Distinct GTP Binding Protein. *Structure* **8**, 1239–1245 (2000).
23. Zeraik AE, Pereira HM, Santos YV, Brandão-Neto J, Spoerner M, Santos MS, Colnago LA, Garratt RC, Araújo AP, DeMarco R. Crystal structure of a *Schistosoma mansoni* septin reveals the phenomenon of strand slippage in septins dependent on the nature of the bound nucleotide. *JBC* **289**, 7799–7811 (2014).
24. Anandapadamanaban M, Masson GR, Perisic O, Berndt A, Kaufman J, Johnson CM, Santhanam B, Rogala KB, Sabatini DM, Williams RL. Architecture of human Rag GTPase heterodimers and their complex with mTORC1. *Science*. 2019 Oct 11;366(6462):203-210. doi: 10.1126/science.aax3939.
25. Shen, K., Choe, A. & Sabatini, D. M. Intersubunit Crosstalk in the Rag GTPase Heterodimer Enables mTORC1 to Respond Rapidly to Amino Acid Availability. *Mol Cell* **68**, 552-565.e8 (2017).
26. Kozlova, M. I., Shalaeva, D. N., Dibrova, D. v. & Mulkidjanian, A. Y. Common Mechanism of Activated Catalysis in P-loop Fold Nucleoside Triphosphatases—United in Diversity. *Biomolecules* **12**, (2022).
27. Kozlova, M. I., Shalaeva, D. N., Dibrova, D. v. & Mulkidjanian, A. Y. Common Patterns of Hydrolysis Initiation in P-loop Fold Nucleoside Triphosphatases. *Biomolecules* **12**, (2022).
28. Dong, J. H., Wen, J. F. & Tian, H. F. Homologs of eukaryotic Ras superfamily proteins in prokaryotes and their novel phylogenetic correlation with their eukaryotic analogs. *Gene* **396**, 116–124 (2007).
29. Klinger, C. M., Spang, A., Dacks, J. B. & Ettema, T. J. G. Tracing the Archaeal Origins of Eukaryotic Membrane-Trafficking System Building Blocks. *Mol Biol Evol* **33**, 1528–1541 (2016).

Chapter 4: **Mechanism of MglA activation by an asymmetrically interacting MglB dimer**

4.1 Introduction:

MglA, the prokaryotic small Ras-like GTPase forms a complex with MglB, identified to be its cognate GAP[1,2]. The stoichiometry of interaction is 1:2. Since MglA is an asymmetric molecule, the interaction with MglB dimer is also asymmetric (PDB ID: 6IZW) (Fig 4.1A) [3,4]. The two MglB protomers (referred to as $MglB_1$ and $MglB_2$), interact with different MglA residues.

An interesting aspect of this asymmetric interaction between MglA and MglB (PDB ID: 6IZW) is that the C-terminal α-helix (Ct-helix) of $MglB_1$ interacts with the α_5 helix and the extended β_2-β_3 loop of MglA (Fig 4.1 A, B)[4]. We observed that the Ct-helix of MglB drives the GEF activity of MglB. A C-terminal truncated construct of MglB (last 20 amino acids of MglB truncated, termed as $MglB^{\Delta Ct}$) interacts with MglA only in the GTP bound state, unlike the full-length MglB which binds MglA both in the GTP and GDP bound states[4]. The C-terminal region of MglB is intrinsically flexible, as observed in the $MglB_2$ protomer, which is disordered in the structure of the MglA MglB complex. However, in the $MglB_1$ protomer, the Ct-helix was stabilised upon interaction with MglA.

The N-terminal residues of the two MglB protomers adopt different secondary structures in the MglA-MglB asymmetric complex. It is observed that in $MglB_1$, residues 2-7 constitute a β-strand that forms hydrogen bonds with the β_0 strand of MglA extending the central β-sheet in the MglA core[4] (Fig 4.1 C). However, residues 1-7 are disordered in the $MglB_2$ protomer. In $MglB_1$, the first helix of the Rbl/LC7 fold extends from Glu-11 whereas, in $MglB_2$, the helix initiates from Tyr-8 (Fig 4.1 A). It is clear that the N-terminal β-strand of $MglB_1$ is a part of the binding interface of MglB to MglA and hence, this strand might play a role in orienting MglB to MglA.

The conformations of Switch I and Switch II loops of MglA are reoriented upon interaction with MglB. Consequently, Thr-54, which completes the Mg^{2+} coordination in MglA, and Arg-53, which is the catalytic residue positioned near the nucleotide-binding pocket, are optimally oriented to stimulate hydrolysis by MglA (Fig 4.1 D). This mechanism of indirect GAP activity where the active site residues are not provided by the GAP but instead, a conformational change of the active site residues drives hydrolysis, is unique for the GAP MglB[1,3,4]. If the MglA-GTP structure with and without MglB (PDB ID: 6IZW, 6H17) are compared, it is observed that Switch-I takes a different conformation in the absence of MglB[3]. Switch 1 of MglA comes near

Asn-37 and Arg-98 of $MglB_2$, and hence these interactions might have a plausible role in the regulation of MglB GAP activity (Fig 4.1 D).

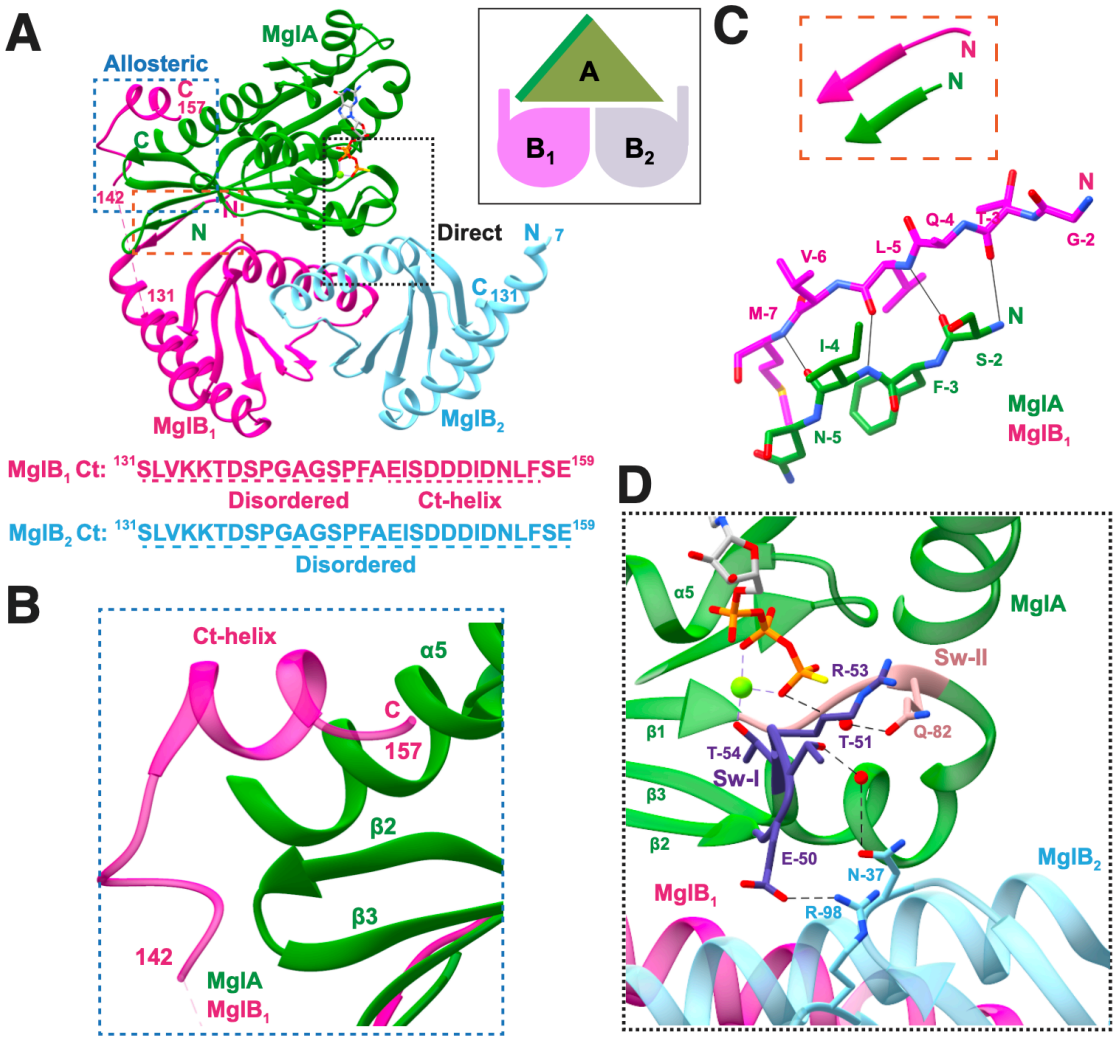


Figure 4.1. Asymmetric features of MglA and MglB complex

A. Structure of the *M. xanthus* MglA-MglB complex (PDB ID: 6IZW) (MglA in green, $MglB_1$ in magenta and $MglB_2$ in light blue) highlighting the $MglB_1$ Ct-helix interaction with MglA (blue dotted box) labelled as allosteric interaction and $MglB_2$ Ct-helix interaction with MglA nucleotide-binding pocket (black dotted box) labelled as direct interaction. The residues at the C and N-termini ('C' and 'N' respectively) and the breaks of protein chains (residue numbers) are labelled for all the chains. The amino acid sequence of the Ct-region of the two MglB protomers is shown below with the disordered residues in the crystal structure and the Ct-helix sequence labelled. In the inset, a schematic representation, as depicted in later figures, of the MglA MglB complex is shown.

B. Interaction of $MglB_1$ Ct- helix (magenta) (start and end residue numbers labelled) with MglA (green) helix binding pocket. MglA α_5 helix and the β_2-β_3 loop have been labelled.

C. Interaction between $MglB_1$ N-terminal β-strand and β_0-strand of MglA which extends the MglA β-sheet constituting the central Ras fold.

D. Interaction between Asn-37 and Arg-98 residues of the $MglB_2$ protomer with Switch-I region (dark purple) of MglA. Switch-I and Switch-II residues are labelled in dark purple and peach respectively.

The dual GAP and GEF activities of MglB need to be regulated to control the relative pool of MglA GTP and GDP-bound states during polarity reversals within the cell. In this chapter, we dissect the roles of $MglB_1$ and $MglB_2$ to regulate the dual GAP and GEF activity of MglB, in light of its asymmetric interaction with MglA. We propose a model of the GEF and GAP active regions of MglA-MglB interacting interfaces and regulating these interactions in space and time plays a role in the relative localisation of MglA-GTP in determining polarity axis in *Myxococcus xanthus*.

4.2 Results:

4.2.1 Design of the linked constructs for a functional asymmetric MglA MglB complex

To dissect the asymmetric interaction between the two MglB protomers with respect to the interaction of the C-terminal residues, we attempted to generate an asymmetric dimer of MglB, in which only one of the MglB protomers (either $MglB_1$ or $MglB_2$) possessed the C-terminal residues. Since the MglB dimer asymmetry is a consequence of its interaction with MglA, we designed a construct linking the C-terminus of MglA with the N-terminus of MglB (Fig 4.2). The C-terminus of MglA was near the N-terminus of the $MglB_1$ protomer in the crystal structure (PDB ID 6IZW). So, we designed a short flexible linker to allow the formation of a functional complex. Hence, a Gly-Ser linker (GS) between MglA and MglB was chosen to link MglA to $MglB_1$. The length was rigid and was designed so as to prevent excessive free rotation between MglA and $MglB_1$. This was also consistent with the crystal packing in the MglA-GDP structure, where the C-terminal hexahistidine tag formed a β-strand which continued the central β-sheet of MglA (PDB ID: 5YMX; Fig 4.2).

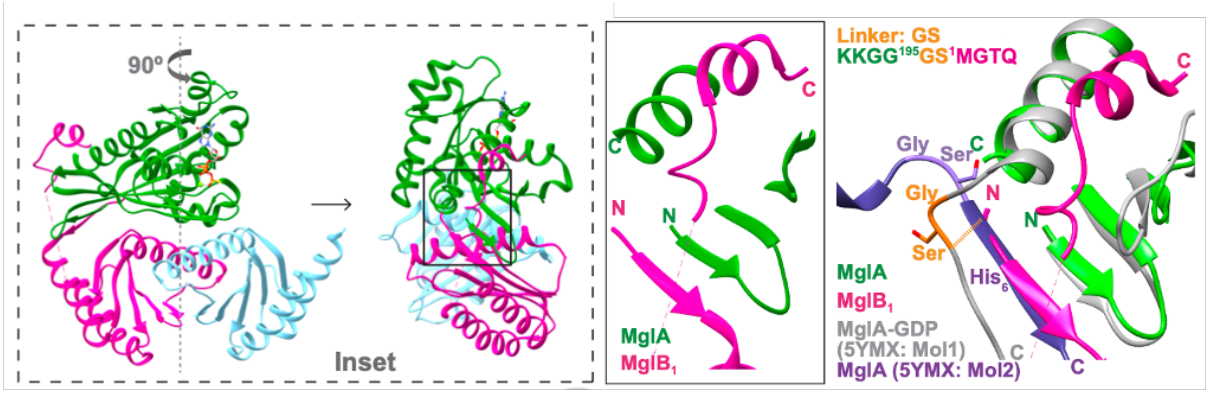


Figure 4.2. Design of the linker between MglA and $MglB_1$
Design of a linked construct based on the orientation of N- and C-termini of MglA and MglB. (Insets). Structure of the *M. xanthus* MglA MglB complex (PDB ID: 6IZW; left; box in dashed outline) rotated

sideways and zoomed in (right; box in solid outline) to highlight the respective termini positions. A short Gly-Ser linker (light orange) between the C-terminus of MglA and the N-terminus of $MglB_1$ will serve to connect MglA and MglB. The position of a similar Gly-Ser linker from the C-terminal end of MglA to its C-terminal hexahistidine tag (purple; as observed in the crystal packing of MglA-GDP structure PDB ID: 5YMX) superposed with the N-terminal β-strand of $MglB_1$ is shown to demonstrate the feasibility of interaction.

We attempted purification of the MglA-link-$MglB_1$ construct, without co-expression with MglB, but the protein was insoluble. This is probably because a dimeric MglA-link-$MglB_1$ might be hindered due to a steric clash between the linked MglAs. This will result in an exposed MglB dimer interface which is unstable, probably leading to an insoluble protein. To resolve this, we adopted a co-expression strategy from a single plasmid where $MglB_2$ and MglA-link-$MglB_1$ were present together.

4.2.2 Cloning and purification of the asymmetric MglA MglB complexes

We adopted a strategy where co-expression from a single plasmid was carried out. $MglB_2$ and MglA-link-$MglB_1$ respectively were expressed as an operon, designed based on the *mglBA* operon in the *Myxococcus xanthus* genome (Fig 4.3 A, Methods: 2.2.2). In such a co-expression system, an MglB protomer could interact with the exposed MglB dimeric interface of the MglA-link-$MglB_1$. A complex containing a linked MglA-link-$MglB_1$ construct is shown by an enclosure within '{' & '}' (for example, $\{AB^{WT}\}$; (Fig 4.3 A), while the absence of '{' and '}' represents a complex formed by mixing different purified proteins (eg. AB^{WT}).

MglA-link-$MglB_1$ was cloned with a C-terminal hexahistidine tag (GSHHHHHH) and $MglB_2$ was with a C-terminal Strep-tag (WSHPQFEK) [5]. This aided in the purification of the intact linked complexes using sequential affinity chromatography (Ni-NTA and StrepTrap respectively) (Fig 4.3 B, Methods: 2.3). This strategy was further used for the purification of the C-terminal truncated constructs of $MglB_1$, and $MglB_2$, respectively, to generate the asymmetric complexes comprising MglB with and without the C-terminal residues ($\{AB_1^{\Delta Ct}B_2\}$, $\{AB_1B_2^{\Delta Ct}\}$; Fig 4.3 A). As controls, $\{AB^{WT}\}$, a linked construct of wild type MglB with MglA, and $\{AB^{\Delta Ct}\}$, a linked construct of C-terminal truncated MglB with MglA, were also purified. Any intermediate event of dissociation would elute the MglB dimer as a separate fraction from the Strep-Trap column due to the higher affinity conferred by the presence of two Strep tags over the linked complex with one Strep-tag.

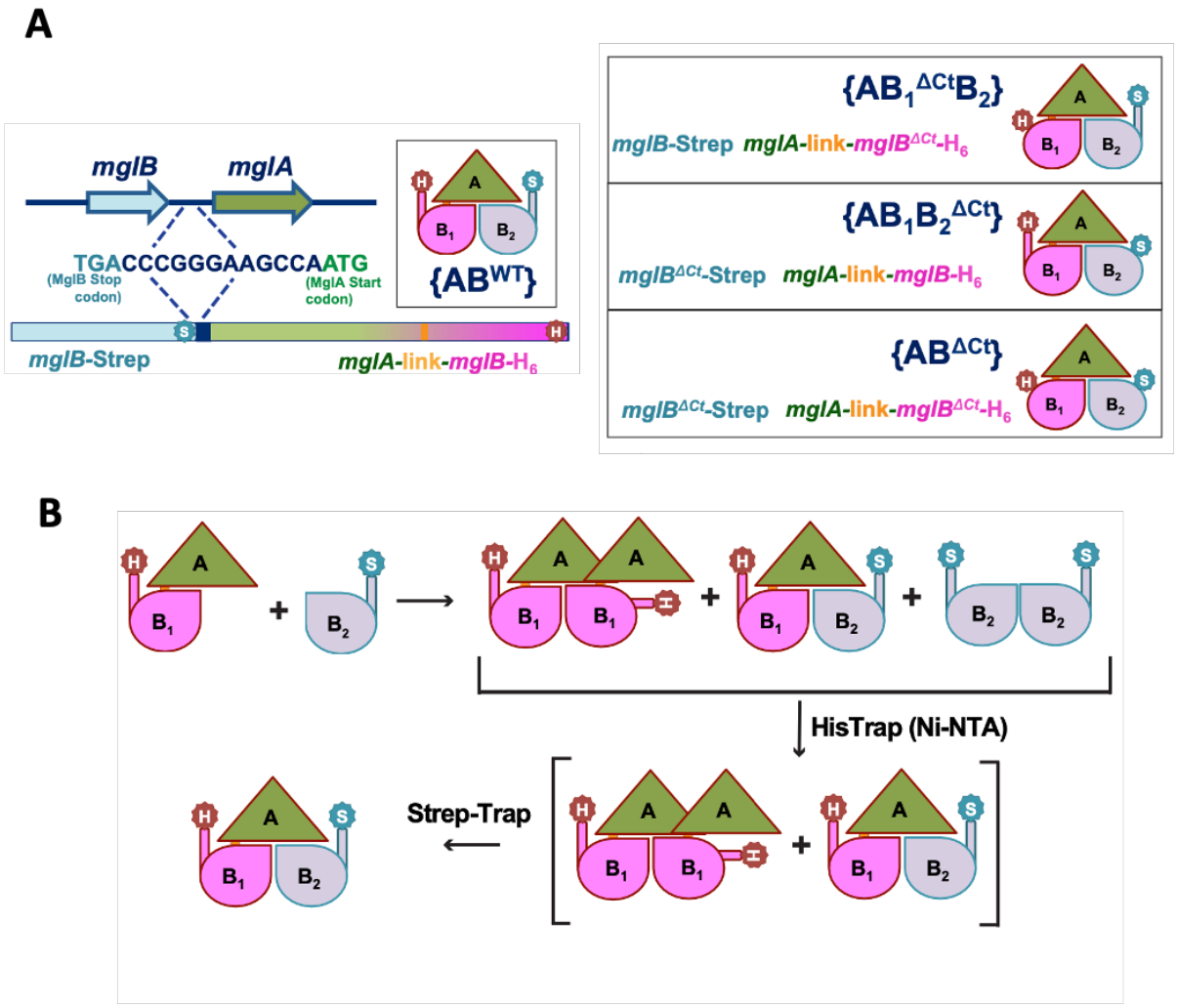


Figure 4.3. Co-expression strategy for purification of asymmetric MglA MglB complexes
A. Strategy for co-expressing constructs of $MglB_2$ (grey) and MglA-link-$MglB_1$ (green-pink) (MglA: green; Linker: yellow; $MglB_1$: pink) with an intervening DNA sequence from the operon (dark blue). The possible combinations of complexes thus formed are shown. The linked constructs include $\{AB^{WT}\}$ and those with respective truncations of $MglB_1$ and $MglB_2$ C-terminal residues ($\{AB_1{}^{\Delta Ct}B_2\}$, $\{AB_1B_2{}^{\Delta Ct}\}$ and $\{AB^{\Delta Ct}\}$ respectively). $MglB_2$ (outlined in blue) and MglA-link-$MglB_1$ (outlined in dark red) are tagged with Strep (S: WSHPQFEK) and hexahistidine tag (H) respectively.
B. Strategy of purification of the asymmetric complexes through sequential affinity chromatography steps (Ni-NTA and StrepTrap).

4.2.3 Differential stability of the linked complexes

Sequential affinity-based chromatography (Ni-NTA followed by StrepTrap) was exploited to purify $\{AB^{WT}\}$, wherein in the second StrepTrap elution step, the majority of the linked complexes were isolated with the first injection of D-desthiobiotin (fraction I) with a negligible amount of MglB dimers (Fig 4.4 A) at later fractions (fraction II; MglB dimers with higher

affinity to the column due to the presence of two Strep-tags). We further proceeded in purifying the rest of the linked constructs with C-terminal deletions. $\{AB_1B_2^{\Delta Ct}\}$ profiles were similar to $\{AB^{WT}\}$ through the purification steps (Fig 4.4 C). Interestingly, StrepTrap fraction II of elution for $\{AB_1^{\Delta Ct}B_2\}$ and $\{AB^{\Delta Ct}\}$ purifications (Fig 4.4 B, D) contained higher proportions of MglB as compared to fraction I where MglB co-eluted with MglA-link-$MglB_1$. This suggests a possible shuffling of MglB protomers in the complexes after Ni-NTA elution.

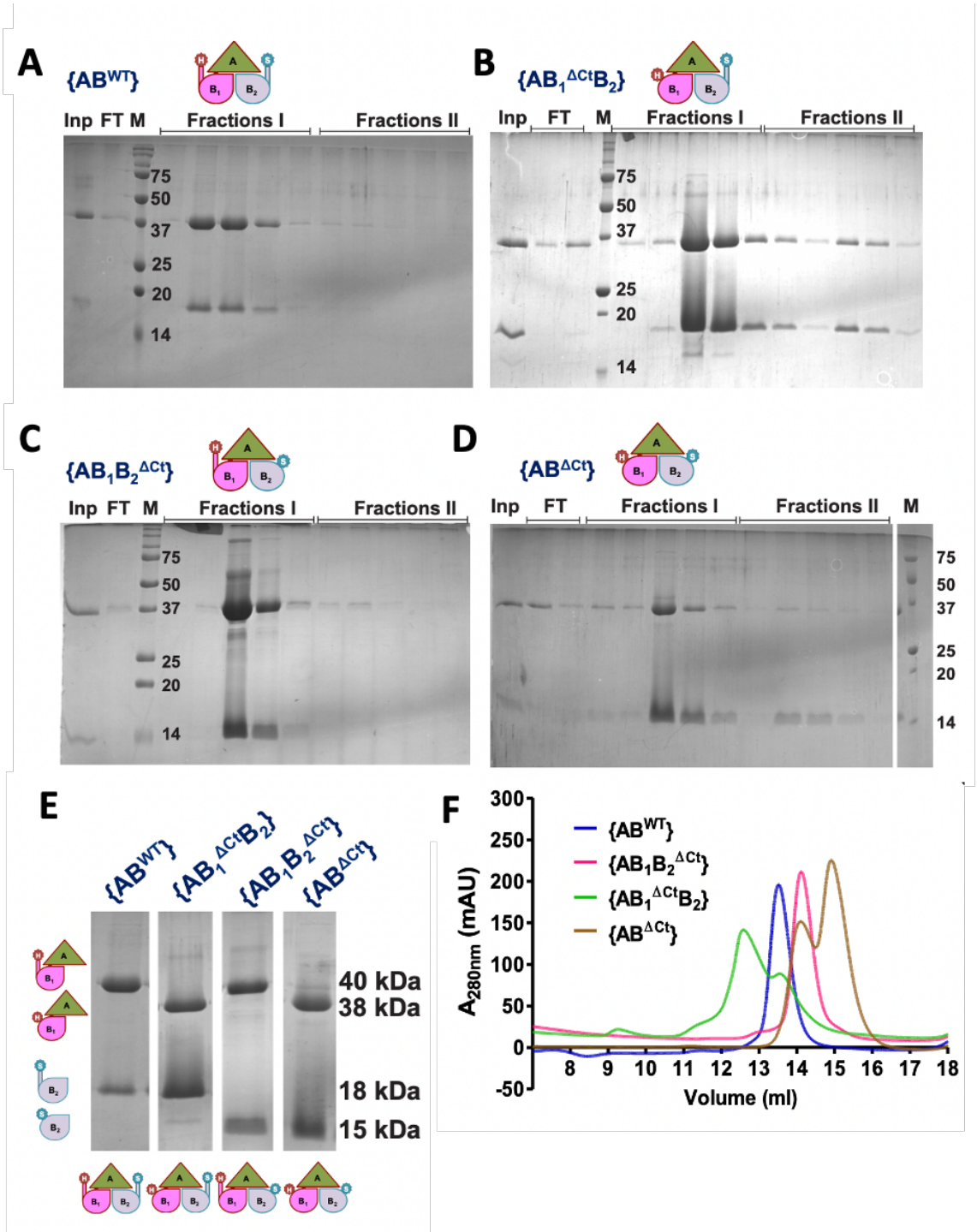


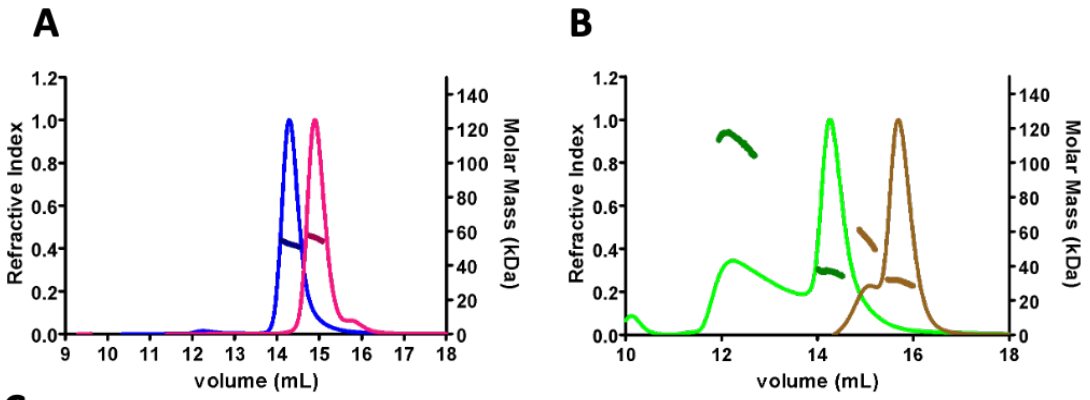
Figure 4.4. Purification of asymmetric complexes
A, B, C, D. SDS-PAGE showing the StrepTrap elution profiles of the linked complexes of $\{AB^{WT}\}$, $\{AB_1{}^{\Delta Ct}B_2\}$, $\{AB_1B_2{}^{\Delta Ct}\}$ and $\{AB^{\Delta Ct}\}$ respectively. The first lane labelled as Inp is the load (obtained after Ni-NTA purification), and the lane labelled FT represents flow through (unbound protein). M is the protein ladder with the molecular weights (in kDa) labelled. Fractions I and II represent the 5 eluted fractions of 1 ml each from each injection of 5 ml of 2.5 mM desthiobiotin-containing elution buffer. The major constructs eluted as 'fractions I' are represented via the schematic.
E. SDS PAGE showing the purified complexes from the linked constructs (cropped images of single lanes from figure panels A – D compiled for ease of comparison). The upper band corresponds to MglA-link-$MglB_1$ and the lower band is to $MglB_2$. Corresponding schematics of each construct are shown on the left and top of the lanes, and the sizes are labelled on the right.
F. Analytical size exclusion chromatography (Superdex 200) profiles for the linked complexes: $\{AB^{WT}\}$ (blue), $\{AB_1B_2{}^{\Delta Ct}\}$ (magenta), $\{AB_1{}^{\Delta Ct}B_2\}$ (green) and $\{AB^{\Delta Ct}\}$ (orange).

MglB moves aberrantly in SEC due to the presence of the C-terminal residues (Fig 4.4 E). To estimate the molecular weights of the elution peaks conclusively, size exclusion chromatography coupled with multi-angle light scattering studies (SEC-MALS; Fig 4.5 A, B) was employed (Methods: 2.5.3). In the case of $\{AB^{WT}\}$, a single peak eluted at 14.2 ml (Fig 4.5 A) corresponding to a molar mass of around 53 kDa, corresponding to the expected mass of the complex (tabulated in Fig 4.5 C). Similarly, in the case of $\{AB_1B_2{}^{\Delta Ct}\}$, a single-homogenous peak eluted at 14.9 ml, corresponding to an average molar mass of 56 kDa (Fig 4.5 A). The differences in elution volumes of $\{AB^{WT}\}$ and $\{AB_1B_2{}^{\Delta Ct}\}$ could be due to the absence of aberrant movement without the $MglB_2$ Ct-helix in $\{AB_1B_2{}^{\Delta Ct}\}$, despite a negligible difference in molecular weights of the two complexes. A similar difference in elution volumes has been observed between AB^{WT} vs $AB^{\Delta Ct}$ complexes as well, despite the molecular weights being very close[4].

However, the $\{AB_1{}^{\Delta Ct}B_2\}$ complex separated into two peaks in SEC (Fig 4.4 D). The first peak (peak 1; eluted at 12.2 ml; Fig 4.5 B) corresponded to a molar mass of 112 kDa correlating to the mass of a dimeric $\{AB_1{}^{\Delta Ct}B_2\}$ complex. However, the peak was not symmetric and showed a long tail till the second peak (peak 2; eluted at 14.2 ml; Fig 4.5 B) correlating with a full-length MglB dimer (36 kDa) (Fig 4.4D). This tailing nature of the peak corresponded to a decreasing trend in the molar mass distribution, which could not be reliably estimated beyond 12.6 ml. This probably indicated the presence of a heterogeneous mix of monomeric and dimeric $\{AB_1{}^{\Delta Ct}B_2\}$ complexes. This dissociation of protomers of the complex was further accelerated in $\{AB^{\Delta Ct}\}$ which had a major $MglB^{\Delta Ct}$ dimer peak (labelled as peak 2; eluted at 15.6 ml) and a smaller peak corresponding to $\{AB^{\Delta Ct}\}$ (peak 1; eluted at 15.1 ml with a molar

mass of approximately 55 kDa) (Fig 4.5 B). These results observed during purification and the SEC-MALS studies indicate that the $MglB_1$ Ct-helix is critical for stabilising the complex consisting of MglA-link-$MglB_1$ and $MglB_2$. This is in line with our earlier observations of the interaction of MglB with MglA in the GDP-bound state only in the presence of the MglB C-terminal residues. From the observation of a stable complex only when $MglB_1$ C-terminal helix is present, this stabilization effect is probably a contribution from the $MglB_1$ C-terminal helix while the absence of it destabilises the complex.

The presence of elution peaks corresponding to respective $MglB_2$ dimers (either MglB or $MglB^{\Delta Ct}$ in $\{AB_1{}^{\Delta Ct}B_2\}$ and $\{AB^{\Delta Ct}\}$ purifications, respectively) shows that in the absence of the $MglB_1$ helix, the linked MglA MglB complex dissociates and releases free $MglB_2$ protomers. Moreover, when the C-terminal helix of $MglB_1$ is absent, MglA might not be restricted to an orientation such that it occupies the interaction interface of the MglB dimer, but is free to move around the linker. The steric hindrance of MglB dimer due to an incompatible conformation, especially in the $MglB_2$ protomer, with the MglA-GDP conformation, might result in rotation of MglA along the linker and away from the MglB interface, destabilization of the MglB dimer interface and its consequent dissociation. The dissociated MglB protomers might form dimers or *trans* oligomers (as observed in $\{AB_1{}^{\Delta Ct}B_2\}$ complex) with MglA-link-$MglB_1$ in solution. Fig 4.5 D lists different *trans* complex species that could exist in solution, out of which option 'iii' depicts a plausible confirmation following a rotation of MglA away from the MglB interface, and probably unique to $\{AB_1{}^{\Delta Ct}B_2\}$. Interestingly, these might represent the dimeric $\{AB_1{}^{\Delta Ct}B_2\}$ complexes consistent with the observations in SEC-MALS. Since the dimeric complexes are observed only in $\{AB_1{}^{\Delta Ct}B_2\}$ and not in $\{AB^{WT}\}$ and $\{AB_1B_2{}^{\Delta Ct}\}$, we conclude that the $MglB_1$ Ct-helix interaction with MglA is essential for the formation of stable and functional linked MglA MglB complexes.



C

Construct	Peak 1 (AB complex)			Peak 2 (B_2 dimer)		
	Elution Volume (ml)	Expected molar mass (kDa)	Observed molar mass (kDa)	Elution Volume (ml)	Expected molar mass (kDa)	Observed molar mass (kDa)
$\{AB^{WT}\}$	14.2	58.5	53	NA	-	-
$\{AB_1B_2^{\Delta Ct}\}$	14.9	56.1	56	NA	-	-
$\{AB_1^{\Delta Ct}B_2\}$	12.2	56.1	112.9	14.2	36.7	36
$\{AB^{\Delta Ct}\}$	15.1	53.7	55.6	15.6	31.9	31

D

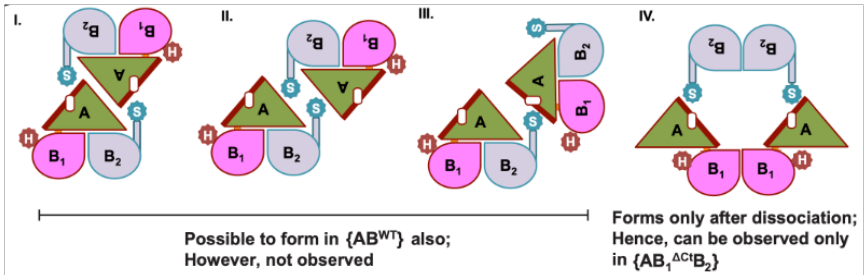


Figure 4.5. C-terminal residues of MglB contribute to differential stability of the linked complexes.

A, B. SEC-MALS (Superdex-200 profiles for SEC carried out at room temperature ~ 25 °C) for the linked constructs: $\{AB^{WT}\}$ blue and $\{AB_1B_2^{\Delta Ct}\}$ magenta in panel A and $\{AB_1^{\Delta Ct}B_2\}$ green and $\{AB^{\Delta Ct}\}$ brown in panel B. The left and the right y axes represent the refractive index and molar mass (kDa) respectively for both plots. Molar masses are represented by thicker darker lines as compared to refractive indices.

C. Table enlisting the elution volumes and corresponding average molar masses (expected and observed; Peaks 1 and 2 correspond to AB_1B_2 complex and B_2 dimer, respectively as labelled in panels A and B) for the linked complexes as concluded from SEC-MALS.

D. Possible *trans* complexes in solution

4.2.4 Identifying the nucleotide bound to the linked complexes

The stability of the linked complex may be related to the presence of bound nucleotides. To analyse the relative amounts of nucleotide bound to the linked complexes, we performed HPLC analysis of the extracts of the respective purified protein samples (Methods: 2.5.4). We observed that all the linked complexes came intrinsically bound to GDP, however, the amount of bound GDP was different despite being injected in the same molar ratios of protein samples processed (100 µM, 100 µl) (Fig 4.6). The constructs with the Ct-helix deleted in $MglB_1$, $\{AB_1^{\Delta Ct}B_2\}$ and $\{AB^{\Delta Ct}\}$ contained lesser amounts of bound GDP compared to those with Ct-helix on $MglB_1$ ($\{AB^{WT}\}$ and $\{AB_1B_2^{\Delta Ct}\}$). This confirmed that the absence of the $MglB_1$ Ct-helix reduced the concentration of nucleotide bound linked complex in solution. Interestingly, this also correlates with the stability of the complex observed, as described in the previous section. Hence $MglB_1$ Ct-helix was important for stabilising the GDP-bound MglA-MglB complexes.

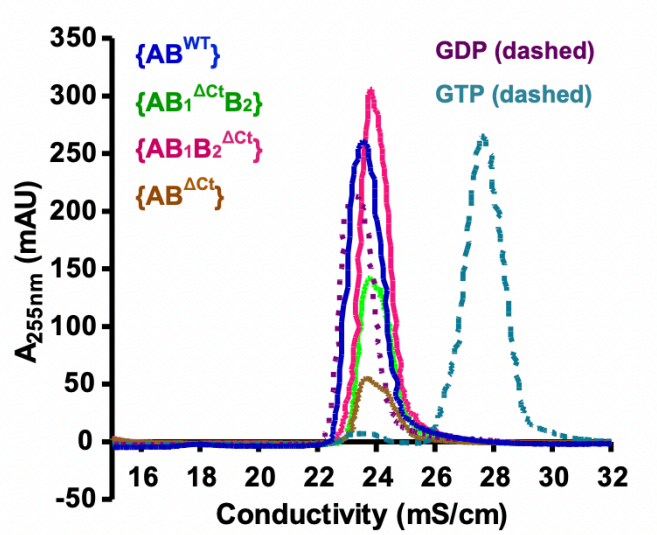


Figure 4.6. HPLC profiles of the linked complexes
$\{AB^{WT}\}$ (blue), $\{AB_1B_2^{\Delta Ct}\}$ (magenta), $\{AB_1^{\Delta Ct}B_2\}$ (green) and $\{AB^{\Delta Ct}\}$ (orange). GDP and GTP used as standards are shown in dotted lines in purple and cyan respectively.

4.2.5 Effect of MglB Ct-helices in GTP hydrolysis rates of MglA

The purified asymmetric complexes of MglA and MglB were used to quantify the GTP hydrolysis rates of MglA and the relative contribution from the C-terminal helices of MglB (Methods: 2.6.1). $\{AB^{WT}\}$ showed similar k_{obs} values compared to the unlinked AB^{WT} complex

which confirms the formation of a functional linked MglA MglB complex using our strategy (Fig 5A, Table 4.1). We observed that $\{AB_1B_2^{\Delta Ct}\}$ exhibited GTP hydrolysis rates comparable to that of AB^{WT} (linked and unlinked). This confirms that the C-terminal helix of the $MglB_2$ protomer does not contribute to accelerating the overall GTPase activity of MglA (Fig 4.7, Table 4.1).

Both $\{AB_1^{\Delta Ct}B_2\}$ and $\{AB^{\Delta Ct}\}$ complexes showed reduced GTP hydrolysis rates, both of which had the $MglB_1$ Ct-helix truncated. Interestingly, this activity was equivalent to the activity of the unlinked complex of MglA and $MglB^{\Delta Ct}$ ($AB^{\Delta Ct}$) where the Ct-helix was absent in the MglB dimer (Fig 4.7, Table 4.1). Hence, the asymmetric linked complexes probably mimicked the corresponding unlinked MglA MglB complexes. Therefore, this comparable activity of unlinked MglA and $MglB^{\Delta Ct}$ complex and linked $\{AB_1^{\Delta Ct}B_2\}$ and $\{AB^{\Delta Ct}\}$, suggests that the concentrations of the MglA were similar in all these cases. Hence the reduced activity of $\{AB_1^{\Delta Ct}B_2\}$ and $\{AB^{\Delta Ct}\}$ suggests that the decrease is due to the absence of $MglB_1$ Ct-helix. Thus, the presence of the $MglB_1$ helix was preferentially critical for accelerating the MglA GTP hydrolysis rate, and not the $MglB_2$ Ct-helix.

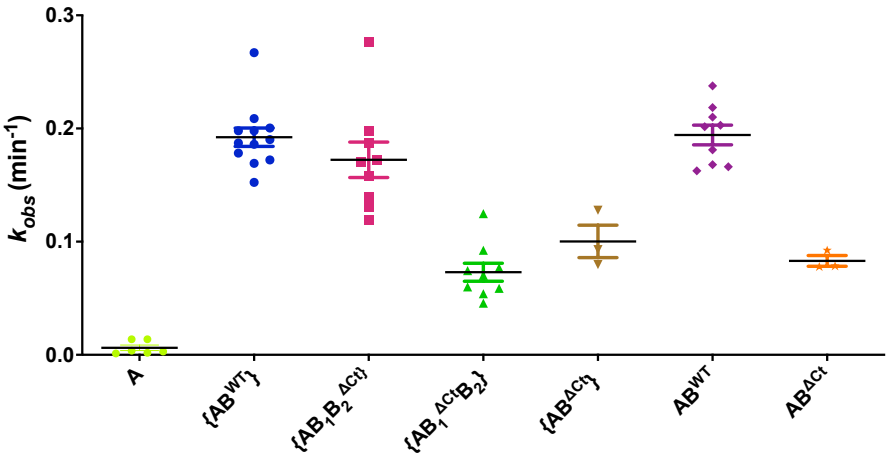


Figure 4.7. Constructs with $MglB_1$ helix stimulate higher GTP hydrolysis rates of MglA.
Comparison of k_{obs} values of $\{AB^{WT}\}$ (blue), $\{AB_1^{\Delta Ct}B_2\}$ (green), $\{AB_1B_2^{\Delta Ct}\}$ (magenta), $\{AB^{\Delta Ct}\}$ (brown), unlinked AB^{WT} (purple), $AB^{\Delta Ct}$ (coral). The mean and 95% confidence intervals are shown by long and short horizontal lines, respectively for each sample. {} denotes linked complexes.

Table 4.1: Summary of observed k_{cat} for linked and unlinked complexes

(k_{obs} in min^{-1}, k_{obs} here is the amount of GDP released per minute per µM enzyme, for details, see Materials & Methods), 'n' denotes the number of repeats.

Construct	k_{obs} (min^{-1})
MglA	0.009 ± 0.003 (n=3)
MglB	0.02 ± 0.002 (n=3)
$\{AB^{WT}\}$	0.19 ± 0.008 (n=13)
$\{AB_1^{\Delta Ct}B_2\}$	0.07 ± 0.008 (n=9)
$\{AB_1B_2^{\Delta Ct}\}$	0.16 ± 0.016 (n=10)
$\{AB^{\Delta Ct}\}$	0.10 ± 0.014 (n=3)
AB^{WT}	0.19 ± 0.02 (n=7)
$AB^{\Delta Ct}$	0.08 ± 0.005 (n=3)

4.2.6 Effect of MglB Ct-helices on nucleotide exchange by MglA

We performed nucleotide exchange assays with fluorescently labelled *mant*-GDP/GMPPNP, using the purified functional linked complexes (Fig 4.8, Methods: 2.6.2). We started with *mant*-GDP exchange assays where we monitored *mant*-GDP binding to the linked complexes (labelled as Stage I) followed by its dissociation (labelled as Stage II) with excess unlabelled GDP (Fig 4.8 A). $\{AB^{WT}\}$ and $\{AB_1B_2^{\Delta Ct}\}$ showed immediate association with *mant*-GDP (k_{on} = 2.7 $\mu M^{-1}min^{-1}$ and 3.2 $\mu M^{-1}min^{-1}$, respectively). The rates of association for $\{AB_1^{\Delta Ct}B_2\}$ were slower. The rate was similar to the unlinked AB^{WT} complex (k_{on} = 0.9 $\mu M^{-1}min^{-1}$). $\{AB^{WT}\}$ and $\{AB_1B_2^{\Delta Ct}\}$ showed higher rates of exchange than unlinked AB^{WT} (Table 4.2). This is due to the association of MglB to MglA owing to the presence of the linker which accelerates the chances of MglB Ct-helix association with MglA. $\{AB^{\Delta Ct}\}$ exhibited the slowest k_{on} of 0.45 $\mu M^{-1}min^{-1}$ owing to the absence of the C-terminal residues in both the linked $MglB_1$ and $MglB_2$ protomer (or in the dissociated $MglB^{\Delta Ct}$ dimer). This was similar to that observed in unlinked $AB^{\Delta Ct}$ (0.6 $\mu M^{-1}min^{-1}$). $\{AB_1^{\Delta Ct}B_2\}$ shows similar k_{on} as AB^{WT} (0.92 ± 0.01 $\mu M^{-1}min^{-1}$ and 1.08 $\mu M^{-1}min^{-1}$, respectively). This is either due to the presence of the $MglB_2$ Ct-helix or due to a probable contribution from the *trans* activity by the dissociated full-length $MglB_2$ dimer

(Fig 4.8 A). The rates of exchange for dissociation of *mant*-GDP by adding unlabelled GDP (k_{off}) were again highest for {AB^{WT}} and {AB$_1$B$_2^{\Delta Ct}$} (0.26 min^{-1} and 0.33 min^{-1}, respectively) followed by {AB$_1^{\Delta Ct}$B$_2$} (0.18 min^{-1}). {AB$^{\Delta Ct}$} (k_{off} = 0.09 min^{-1}) was slower as the exchange of GDP was not facilitated in the absence of the MglB Ct-helix, similar to unlinked AB$^{\Delta Ct}$.

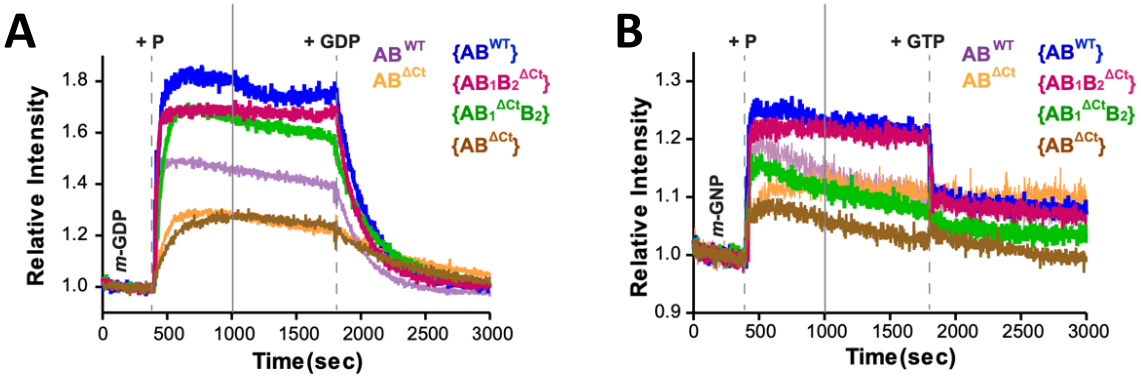


Figure 4.8. MglB$_1$ Ct helix stimulates nucleotide exchange on MglA
A. Kinetics of increase in *mant*-GDP fluorescence (region labelled as *m*-GDP) upon adding {AB^{WT}} (blue), {AB$_1^{\Delta Ct}$B$_2$} (green), {AB$_1$B$_2^{\Delta Ct}$} (magenta), {AB$^{\Delta Ct}$} (brown), AB^{WT} (purple), AB$^{\Delta Ct}$ (coral), at 400 seconds (marked by the dashed line labelled '+P'; Stage I), followed by a competition of *mant*-GDP by adding excess unlabelled GDP at 1800 seconds (marked by the dashed line labelled '+GDP'; Stage II).
B. Kinetics of increase in *m*-GMPPNP fluorescence (region labelled as *m*-GNP) upon adding {AB^{WT}} (blue), {AB$_1^{\Delta Ct}$B$_2$} (green), {AB$_1$B$_2^{\Delta Ct}$} (magenta), {AB$^{\Delta Ct}$} (brown), AB^{WT} (purple), AB$^{\Delta Ct}$ (coral), at 400 seconds (marked by the dashed line labelled '+P'; Stage I), followed by a competition of mant-GMPPNP by adding excess unlabelled GTP at 1800 seconds (marked by the dashed line labelled '+GTP'; Stage II).
The phases I and II, which represent association and dissociation for k_{on} and k_{off} estimation respectively (tabulated in Table 4.2A), are demarcated by the solid line.

Further, we investigated the *mant*-GMPPNP exchange rates of the linked complexes followed by dissociation with excess unlabelled GTP (Fig 4.8 B). {AB^{WT}} and {AB$_1$B$_2^{\Delta Ct}$} showed instant association and dissociation rates. Hence k_{on} and k_{off} could not be estimated. This is probably due to the presence of the linked MglB$_1$ Ct-helix in both these complexes which stabilized the complexes and drove efficient nucleotide exchange. If the MglB$_2$ helix had any probable contribution to the exchange activity, the {AB$_1$B$_2^{\Delta Ct}$} would not show such active nucleotide exchange as {AB^{WT}}. {AB$_1^{\Delta Ct}$B$_2$}, without a Ct-helix in the MglB$_1$ protomer, exhibited measurable exchange kinetics with k_{on} and k_{off} values of 2.2± 0.5 μM^{-1}min^{-1} and 0.18 ± 0.1 min^{-1}, respectively. This was again similar to the unlinked AB^{WT} (Table 4.2). The interaction of unlinked MglB with MglA and also the *trans* activity from the dissociated MglB

dimers with the linked MglA, are both diffusion limited, resulting in similar exchange rates. This is unlike the interaction from the linked $MglB_1$ Ct-helix which showed fast exchange kinetics ($\{AB^{WT}\}$ and $\{AB_1B_2^{\Delta Ct}\}$). $\{AB^{\Delta Ct}\}$ exhibited exchange rates similar to unlinked $AB^{\Delta Ct}$. However, we observed an inconsistent rise in fluorescent intensity signal upon protein addition across the replicates. This is probably due to the lack of stability of the complex during the time scale of this experiment due to the absence of any stabilization from the MglB-Ct helix (no *trans* stabilization like in $\{AB_1^{\Delta Ct}B_2\}$). Unlike in GTPase assays (where a GTP concentration of 1 mM was used), a low concentration of *mant*-GMPPNP (800 nM during association) failed to stimulate GTP binding and stabilization for this complex. Hence, we failed to obtain estimates of k_{on}. k_{off} was negligibly low (0.07 min^{-1}). All these results indicate that in $\{AB^{WT}\}$ and $\{AB_1B_2^{\Delta Ct}\}$, the exchange of GDP with GTP is driven by the $MglB_1$ Ct-helix, which interacts allosterically away from the nucleotide-binding pocket of MglA. In $\{AB_1^{\Delta Ct}B_2\}$, this effect is probably exerted in *trans* from the dissociated MglB dimer.

Table 4.2: Summary of rates of nucleotide association and dissociation for the linked and unlinked MglA MglB complexes
(k_{on}($\mu M^{-1}min^{-1}$) and k_{off}(min^{-1}) respectively).

Construct	*mant* GDP-GDP exchange		*mant* GMPPNP-GTP exchange	
	k_{on} ($\mu M^{-1}min^{-1}$)	k_{off} (min^{-1})	k_{on} ($\mu M^{-1}min^{-1}$)	k_{off} (min^{-1})
$\{AB^{WT}\}$	2.7 ± 0.1 (n=3)	0.26 ± 0.003 (n=3)	ND (n=3)	ND (n=3)
$\{AB_1^{\Delta Ct}B_2\}$	0.92 ± 0.01 (n=3)	0.18 ± 0.004 (n=3)	2.2 ± 0.5 (n=3)	0.18 ± 0.1 (n=3)
$\{AB_1B_2^{\Delta Ct}\}$	3.2 ± 0.11 (n=3)	0.33 ± 0.004 (n=3)	ND (n=3)	ND (n=3)
$\{AB^{\Delta Ct}\}$	0.48 ± 0.02 (n=3)	0.09 ± 0.002 (n=3)	ND (n=3)	0.066 ± 0.03 (n=3)
AB^{WT}	1.08 (n=1)	0.18 (n=1)	2.3 (n=1)	0.28 (n=1)
$AB^{\Delta Ct}$	0.70 (n=1)	0.04 (n=1)	1.5 (n=1)	ND (n=1)

4.2.7 Role of N-terminal β-strand of $MglB_1$

As mentioned earlier, we observe an asymmetric interaction between the N-terminal of $MglB_1$ and $MglB_2$ protomers. In $MglB_1$, the 2-7 residues constitute a β-strand that interacts with the

β_0-strand of MglA. There might be a possible role of this strand in orienting the Ct-helix of $MglB_1$ to MglA. To investigate this, we deleted these corresponding residues in MglB and the construct hence designed was named $MglB^{\Delta N}$ (Fig 4.9 A). These experiments were performed using the unlinked MglB constructs, since the linker strategy was not compatible with the N-terminal residue deletion. We expressed and purified $MglB^{\Delta N}$ mutant and confirmed that this protein was well folded using size exclusion chromatography (Fig 4.9 B, C, D). We further looked at the GTP hydrolysis rates of MglA in presence of $MglB^{\Delta N}$ (Methods: 2.6.1). In presence of this mutant MglB, MglA showed reduced overall GTPase activity as compared with $MglB^{WT}$ addition. The k_{obs} were similar to that of $MglB^{\Delta Ct}$ (Table 4.3, Fig 4.9 E) which had only GAP and no GEF activity due to the absence of the Ct helix.

Table 4.3: Summary of observed k_{cat} for MglB mutants with MglA

(k_{obs} in min^{-1}, k_{obs} here is the amount of GDP released per minute per μM enzyme, for details, see Materials & Methods), 'n' denotes the number of repeats.

Construct	**k_{obs} (min^{-1})**
MglA	0.009 ± 0.003 (n=3)
MglB	0.02 ± 0.002 (n=3)
$MglAB^{\Delta N}$	0.08 ± 0.006 (n=4)
$MglAB^{N37R98A}$	0.11± 0.004 (n=4)

In fluorescently labelled *mant* GDP exchange assays (Methods: 2.6.2), we observed that k_{on} and k_{off} rates of MglA in presence of $MglB^{\Delta N}$ mutant (0.8 $\mu M^{-1}min^{-1}$ and 0.09 min^{-1}, respectively) are similar to MglA alone which proved that the GEF activity of $MglB^{\Delta N}$ was affected as compared to $MglB^{WT}$ (Fig 4.9 F). To confirm that this abrogated GEF activity of the mutant was not due to reduced affinity to MglA, we performed the same assay with an excess of MglB and its mutant (MglA: MglB in 1:4 ratio) and obtained a similar trend (Fig 4.9 G). We went ahead to perform the same assays with fluorescently labelled *mant*-GMPPNP where MglA in presence of $MglB^{\Delta N}$ indeed showed a slower k_{on} rate similar to $MglB^{\Delta Ct}$ (0.8 $\mu M^{-1}min^{-1}$) (Fig 4.9 H, Table 4.4). Hence, we conclude that the N-terminal β-strand in MglB facilitates the optimal interaction of the $MglB_1$ Ct helix with MglA thereby promoting the allosteric GEF activity on MglA.

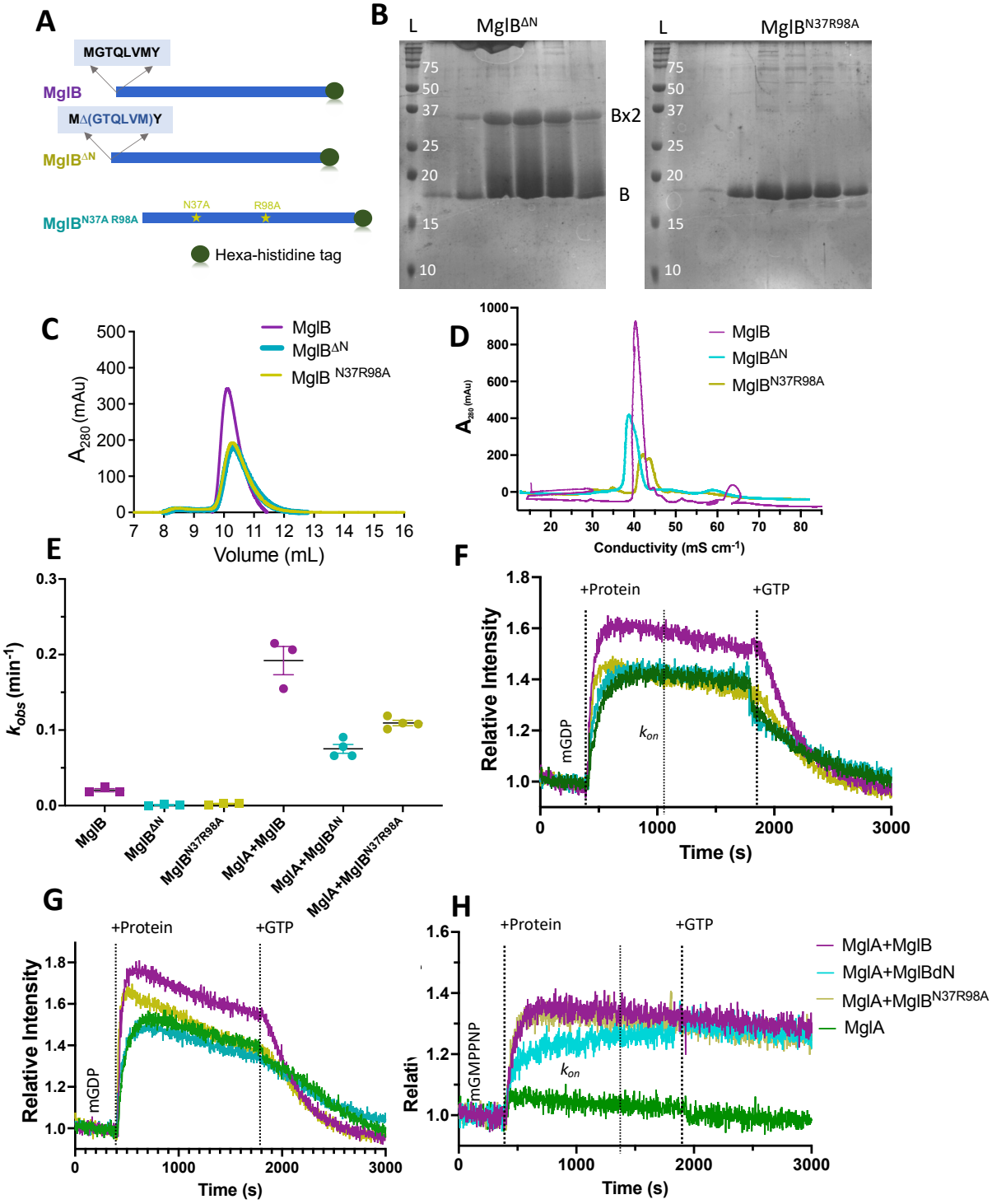


Figure 4.9. $MglB_1$ N terminal β-strand and $MglB_2$ Asn-37/Arg-98 are essential for MglB GEF and GAP activity, respectively

A. Schematic describing the mutant constructs of MglB designed for this study

B. SDS PAGE showing the purified proteins of the MglB mutants after MonoQ purification. L describes the ladder. B and 2xB correspond to bands of monomeric and dimeric MglB, respectively.

C. Size exclusion chromatography (Superdex-75) profiles of mutants of MglB (purple), $MglB^{\Delta N}$ (cyan) and $MglB^{N37\ R98A}$ (ochre).

D. Anion exclusion chromatography (MonoQ) profiles of mutants of MglB (purple), $MglB^{\Delta N}$ (cyan) and $MglB^{N37\ R98A}$ (ochre).
E. Comparison of k_{obs} values of MglB control (square) with $MglB^{WT}$ (purple) with MglB asymmetry mutants $MglB^{\Delta N}$ (cyan) and $MglB^{N37R98A}$ (ochre) and respective constructs with MglA (circles).
The mean and 95% confidence intervals are shown by long and short horizontal lines, respectively for each sample.
F, G. Kinetics of increase in *mant*-GDP fluorescence (region labelled as *m*-GDP) upon adding MglA (dark green), unlinked AB^{WT} (purple), MglA with MglB mutants, $MglB^{\Delta N}$ (cyan) and $MglB^{N37R98A}$ (ochre) at 400 seconds (marked by the dashed line labelled '+P'; Stage I), followed by a competition of *mant*-GDP by adding excess unlabelled GDP at 1800 seconds (marked by the dashed line labelled '+GDP'; Stage II). The plot in G is the same assay with a 1:4 ratio of MglA: MglB.
H. Kinetics of increase in *mant*-GMPPNP fluorescence (region labelled as *m*-GMPPNP) upon adding MglA (dark green), unlinked AB^{WT} (purple), MglA with MglB mutants, $MglB^{\Delta N}$ (cyan) and $MglB^{N37R98A}$ (ochre) at 400 seconds (marked by the dashed line labelled '+P'; Stage I), followed by a competition of *mant*-GMPPNP by adding excess unlabelled GTP at 1800 seconds (marked by the dashed line labelled '+GTP'; Stage II).
The phases I and II, which represent association and dissociation for k_{on} and k_{off} estimation respectively (tabulated in Table 4.2B and 4.2C), are demarcated by the solid line.

Table 4.4: Summary of rates of nucleotide association and dissociation for the unlinked MglA MglB complexes with the MglB mutants
(k_{on}($\mu M^{-1}min^{-1}$) and k_{off}(min^{-1}) respectively).

Construct	*mant* GDP-GDP exchange				*mant* GMPPNP-GTP exchange
	k_{on} ($\mu M^{-1}min^{-1}$) 1:2	**k_{off} (min^{-1}) 1:2**	**k_{on} ($\mu M^{-1}min^{-1}$) 1:4**	**k_{off} (min^{-1}) 1:4**	**k_{on} ($\mu M^{-1}min^{-1}$) 1:2**
MglA	0.7 (n=1)	0.1 (n=1)	0.8 (n=1)	0.06 (n=1)	ND
MglAB	1.3 (n=1)	0.2 (n=1)	2.2±0.2 (n=3)	0.1±0.03 (n=3)	1.8±0.5 (n=3)
$MglAB^{\Delta N}$	0.8±0.1 (n=3)	0.09±0.01 (n=3)	1.0±0.1 (n=3)	0.08±0.01 (n=3)	0.8±0.08 (n=3)
$MglAB^{N37R98A}$	2.7±0.4 (n=3)	0.2±0.03 (n=3)	2.7±0.5 (n=3)	0.2±0.04 (n=3)	1.5±0.08 (n=3)

4.2.8 Role of Asn-37 and Arg-98 for MglB GAP activity

As evident from the structure of MglA MglB complex in the GTP bound state (PDB ID 6IZW), MglB interaction with MglA reorients the Switch-I region of MglA in a conformation driving active hydrolysis thereby functioning as a GAP. We observed that Asn-37 and Arg-98 were

interacting with the Switch-I residues in this conformation which might be important in conferring the GAP activity to MglB. To confirm this hypothesis, we designed a double mutant of MglB, where both Asn-37 and Arg-98 were mutated to alanines ($MglB^{N37R98A}$) (Fig 4.6A). The mutant protein was purified identically to the wild-type MglB and we observe no effect of the mutations on the folding or oligomeric status of the protein as evident from size exclusion chromatography and MonoQ elution profile (Fig 4.6 B, C, D). We observed a reduction in the GTP hydrolysis rates of MglA in presence of $MglB^{N37R98A}$ mutant (Fig 4.6 E, Methods: 2.6.1). To rule out the possibility of this reduced overall GTPase rates due to an impaired GEF activity, we performed a nucleotide exchange assay with *mant* GDP and competed it out with unlabelled GDP (Fig 4.6 F). We observed that the k_{on} rate of MglA with $MglB^{N37R98A}$ was higher than with MglB (2.7 $\mu M^{-1}min^{-1}$ compared to 1.3 $\mu M^{-1}min^{-1}$) (Table 4.2 B). However, the k_{off} rate was similar to $MglAB^{WT}$ (0.2 min^{-1}). We performed the same assay with *mant* GMPPNP and competition with unlabelled GTP, we observed that the k_{on} rate of MglA with $MglB^{N37R98A}$ (1.5 $\mu M^{-1}min^{-1}$) was comparable to that observed for $MglAB^{WT}$ (1.8 $\mu M^{-1}min^{-1}$) (Fig 4.6 H). Hence we confirmed that the $MglB^{N37R98A}$ had an intact GEF activity and the reduced GTPase rates of MglA were indeed due to an impaired GAP activity only. All these observations validate that the dimeric interface of MglB predominantly consisting of the $MglB_2$ protomer is the GAP active region of MglB whereas the Ct-helix of the $MglB_1$ protomer is the GEF active region.

4.3 Conclusion

This study adopts a unique approach to dissect the functional relevance of the asymmetric interaction between MglA and MglB. We observed that the asymmetric linked complexes of MglA MglB with Ct-helix truncations, showed reduced stability when the $MglB_1$ Ct-helix was deleted. In these unstable complexes, we also observed the dissociated $MglB_2$ protomers from the complex to associate with each other to form dimers which were consequently capable of interacting with MglA in *trans*. We compared the activity of the linked complexes with that of a mix of MglA and MglB/$MglB^{\Delta Ct}$. While the activity of MglA in presence of $MglB_2$-Ct truncation was similar to that in presence of MglB, $MglB_1$-Ct truncation resulted in activity similar to $MglB^{\Delta Ct}$. Hence, we show that the $MglB_1$ Ct-helix acts via the helix binding pocket on MglA which lies opposite to the nucleotide-binding pocket, whereas the C-terminal residues of $MglB_2$ protomer do not play a role in the GEF activity (Fig 4.10).

Our conclusions were further validated when we observed the role of the N-terminal β-strand of MglB in orienting the $MglB_1$ Ct-helix of MglB onto MglA. This β-strand of the $MglB_1$ is an

extension of the Rbl/LC7 fold of MglB. This strand in the $MglB_1$ protomer interacts with the β_0 strand of the central β-sheet of MglA, unique to the MglA family. In $MglB_2$, the first α-helix of Rb/LC7 is longer compared to $MglB_1$. If $MglB_1$ also had a continuous longer α-helix, it would hinder the interaction of the Ct-helix of $MglB_1$ onto MglA. Instead, it forms a β-strand continuous with the central β-sheet of the MglA Ras fold. This interaction helps to avoid clashing of the N-terminal region of MglB and Ct-helix of MglB and further aids in the optimal interaction of the $MglB_1$ Ct helix with MglA, stimulating the allosteric GEF activity.

We observed the role of Asn-37 and Arg-98 residues of MglB in orienting the Switch-I residues of MglA thereby contributing partially to the MglB GAP activity. Additionally, we attribute the MglB dimeric interface constituting majorly of $MglB_2$, as the GAP active region of MglB. MglB contains a roadblock domain (RD), an ancestor of the longin domain, a domain common in the GEFs of the small Ras-like GTPase family [6]. MglB RD forms a symmetric dimer to form an extended β-meander platform, which interacts asymmetrically with MglA. We observe that the RD platform contributes to the GAP activity whereas the extension to this domain, functions through an independent allosteric mechanism involving the C-terminal helix (Fig 4.10).

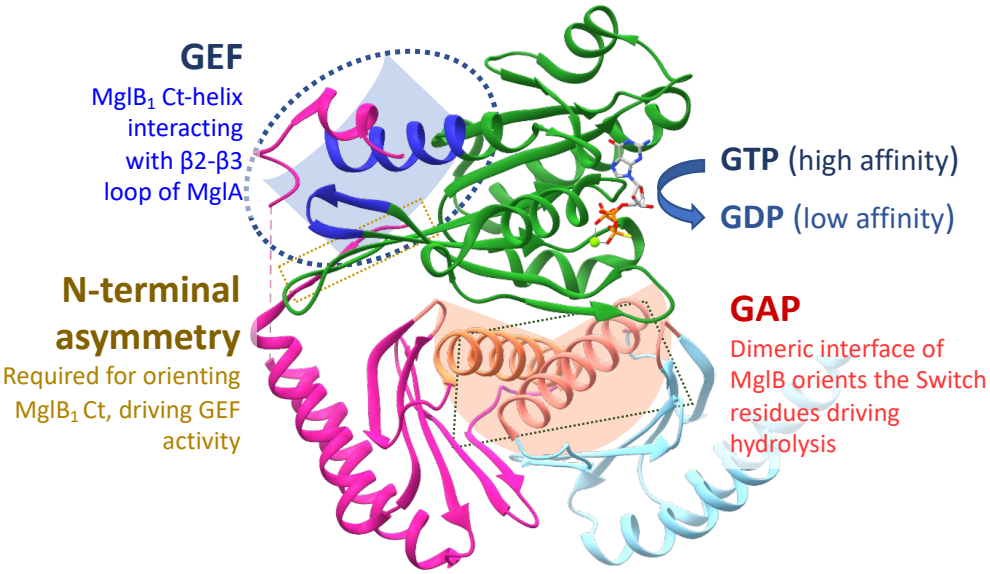


Figure 4.10. Mechanism of dual GAP and GEF activity of MglB
The asymmetric complex of MglA (green) with $MglB_1$ (magenta) and $MglB_2$ (light blue) with interfaces annotated with the GAP and GEF functions of MglB. The C-terminal helix of $MglB_1$ interacts with the MglA helix binding pocket (blue) imparting the GEF activity. The N-terminal region of $MglB_1$ (yellow)

interacts with the central β-sheet of MglA and helps orient the $MglB_1$ Ct helix into the GEF pocket. The dimeric interface of MglB (orange) stabilizes MglA and promotes the GAP function.

Hence we report a unique mechanism of regulation where the Ct-helix of $MglB_1$ interacts with the helix binding pocket consisting of the β_2-β_3 loop of MglA which is the allosteric regulator site, stabilizing the flipped conformation of MglA following the β-screw movement. This conformation favours the nucleotide-binding pocket to remain in an 'open' conformation ready to bind GTP and causes the GDP to dissociate. Consequently, when GTP binds, it binds tightly to the catalytic pocket which is further hydrolysed by the GAP activity of MglB. It is further interesting to decipher, how this dual GAP and GEF activities of MglB are resolved in the cell. Disengaging the Ct-helix of MglB is essential to separate the GAP and GEF activities of MglB, thereby retaining MglB as a GAP only. In such a disengaged state, which could be driven by post-translational modifications or other interaction partners within the cell, MglB loses its GEF function and also detaches from MglA once GTP hydrolysis takes place. There could be a possible contribution from other participants of this module, for example, competition from the RomR-RomX proteins and stimulation from RomY or interaction with MglC which might help resolve this mechanism of how the GAP and GEF activities are compartmentalised at the two poles in the cell.

References:

1. Miertzschke M, Koerner C, Vetter IR, Keilberg D, Hot E, Leonardy S, Søgaard-Andersen L, Wittinghofer A. Structural analysis of the Ras-like G protein MglA and its cognate GAP MglB and implications for bacterial polarity. *EMBO J.* 2011 Aug 16;30(20):4185-97. doi: 10.1038/emboj.2011.291.
2. Zhang, Y., Franco, M., Ducret, A. & Mignot, T. A bacterial Ras-like small GTP-binding protein and its cognate GAP establish a dynamic spatial polarity axis to control directed motility. *PLoS Biol* **8**, (2010).
3. Galicia C, Lhospice S, Varela PF, Trapani S, Zhang W, Navaza J, Herrou J, Mignot T, Cherfils J. MglA functions as a three-state GTPase to control movement reversals of *Myxococcus xanthus. Nat Commun*. 2019 Nov 22;10(1):5300. doi: 10.1038/s41467-019-13274-3.
4. Baranwal J, Lhospice S, Kanade M, Chakraborty S, Gade PR, Harne S, Herrou J, Mignot T, Gayathri P. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. *PLoS Biol.* 2019 Sep 27;17(9):e3000459. doi: 10.1371/journal.pbio.3000459.

5. Aule, T. G., Smith, M. A., Pearson, A. R., Knowles, P. F. & Mcpherson, M. J. Probing the Molecular Mechanisms in Copper Amine Oxidases by Generating Heterodimers. *ChemBioChem* **16**, 4, 23 (2015)
6. Levine TP, Daniels RD, Wong LH, Gatta AT, Gerondopoulos A, Barr FA. Discovery of new longin and Roadblock domains that form platforms for small GTPases in Ragulator and TRAPP-II. *Small GTPases* vol. 4 62–69 Preprint at https://doi.org/10.4161/sgtp.24262 (2013).

Chapter 5: **Characterization of RomX, a novel protein in *Myxococcus xanthus* polarity module**

5.1 Introduction

The prokaryotic small Ras-like GTPase was identified and characterized with MglB, its cognate GAP[1–3]. However, for over two decades, a GEF for MglA remained to be identified, a necessary step towards obtaining a complete picture of the polarity reversal oscillator in *M. xanthus*. Simultaneous to our report of the GEF activity of MglB[4], a new regulator in the polarity reversal module, namely RomX was identified to act as a GEF for MglA along with RomR[5]. Having characterized the dual GAP and GEF activities of MglB in Chapter 4, we ventured to characterize this newly identified GEF module for MglA and investigate its mechanism of MglA regulation.

RomX co-occurs with its partner, RomR across genomes. In *M. xanthus* it was reported that the RomX polar localization is regulated by RomR[5]. RomR interacts with RomX which further interacts with MglA GTP to form a heteromeric complex and helps recruit MglA GTP to the leading pole[5]. RomX has been shown to act as the active guanine nucleotide exchange factor (GEF) for MglA (Fig 5.1 A). Together with RomR, it was observed to accelerate the GDP to GTP exchange on MglA at a faster rate than RomX alone (Fig 5.1 B). Unlike MglB, there was no GAP stimulation of MglA by RomX or RomR. However, unlike typical GEFs, the RomR-RomX complex does not interact with MglA in the GDP state. In addition to its function at the leading pole, RomR-RomX also co-localises with MglA GTP in the Agl/Glt focal adhesion complexes driving gliding motility. The GEF activity of RomR/RomX is assumed to be exerted to recruit MglA-GTP at or near the leading pole of the cell. However, it occurs in a bipolar asymmetric cluster, with a bigger cluster at the lagging pole in a polarized cell.

All these observations raise certain questions, which need to be addressed. Firstly, why does an annotated GEF not interact with the GTPase in the GDP state and interact only with the GTP state? To answer this, we have attempted to biochemically and structurally characterize RomX and its interaction with MglA. Further, why does a bigger cluster of RomR-RomX exist at the lagging cell pole when the activity is predominantly necessary at the leading cell pole for recruiting more of MglA-GTP? Moreover, why does the system have two GEFs, MglB and RomRX? To answer this, we have attempted to assay the effect of RomX on GTPase and nucleotide exchange kinetics of MglA and compare it with that of MglB. This, along with the existing information on the regulation of polarity reversal, as described in Chapter 1, has helped us to understand the consequence of their simultaneous presence in MglA regulation.

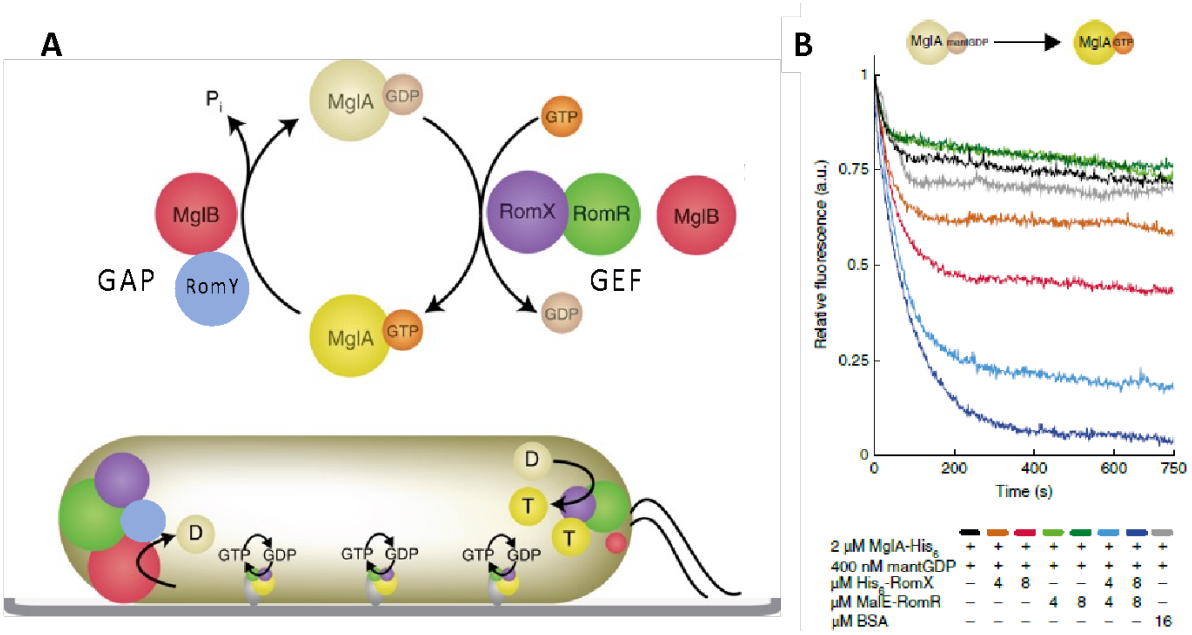


Figure 5.1. RomX is the newly identified GEF for MglA in the polarity module
A: RomX shown in purple was proven to act as a GEF for MglA along with RomR, its partner. The lower panel shows the bipolar asymmetric localization of RomX with a larger cluster at the lagging pole (Adapted from Szadkowski et. al, 2019 [1]).

B: RomX stimulates the release of mant-GDP with cold GTP on MglA in a concentration dependant fashion which is further stimulated in presence of RomR (Data from Szadkowski et. al, 2019[1]).

5.2 Results

5.2.1 Cloning and purification of RomX constructs

The *romX* gene (MXAN_3350) was amplified from *Myxoccoccus xanthus* DK1622 genomic DNA and cloned in pHis17 (Methods: 2.2.3). Constructs were designed with a C-terminal hexahistidine (GSHHHHHH) tag and Strep-tag (GSWSHPQFEK) and an N-terminal hexahistidine tag (MRHHHHHHM) and Strep-tag (MRGWSHPQFEKGG). To standardize expression, the aforementioned constructs were transformed in BL21AI cells and induced at OD_{600} 0.8 with 0.2% L-Arabinose and grown overnight at 18°C (Fig 5.2 A). The C-terminal Strep tagged construct could not be expressed despite multiple trials at different post-induction temperatures (37°C, 30°C, 18°C) and different expression strains (Bl21DE3, Rosetta DE3, C43). This was probably because the Trp from Strep-tag sterically clashed with the N-terminal Trp of RomX.

The hexahistidine tagged constructs were purified using Ni-NTA chromatography followed by size exclusion chromatography (Superdex 75) (Fig 5.2 A, Methods: 2.4). The trace impurities

post Ni-NTA or StrepTrap purification eluted at the Superdex 75 void (8 ml) and a pure protein peak was obtained at 13.3 ml (Fig 5.2 A, B).

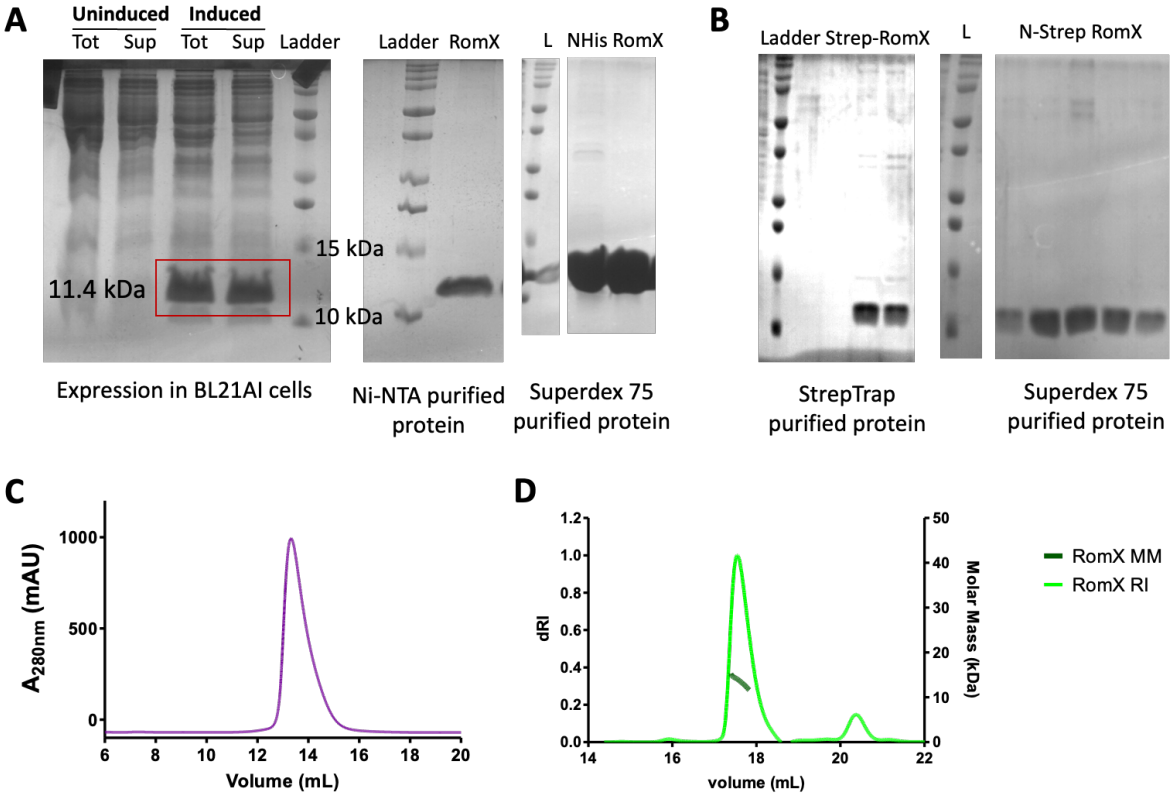


Figure 5.2. Expression and purification of RomX constructs

A: SDS PAGE for expression of N-His RomX. The expressed protein band at 11.4 kDa in induced cultures is highlighted in red. The right panels are SDS PAGE for Ni-NTA purified and Superdex 75 purified protein.

B: SDS PAGE for purification of N-Strep RomX. The left and right panels are for StrepTrap and Superdex 75 purified proteins, respectively.

C: Superdex 75 analytical profile for hexahistidine tagged RomX. The protein elutes at 13.8 ml.

D: Superdex 200 SEC MALS run for RomX. The protein elutes at 17.7 ml and the observed molar mass is 13.6 kDa (±1.964%).

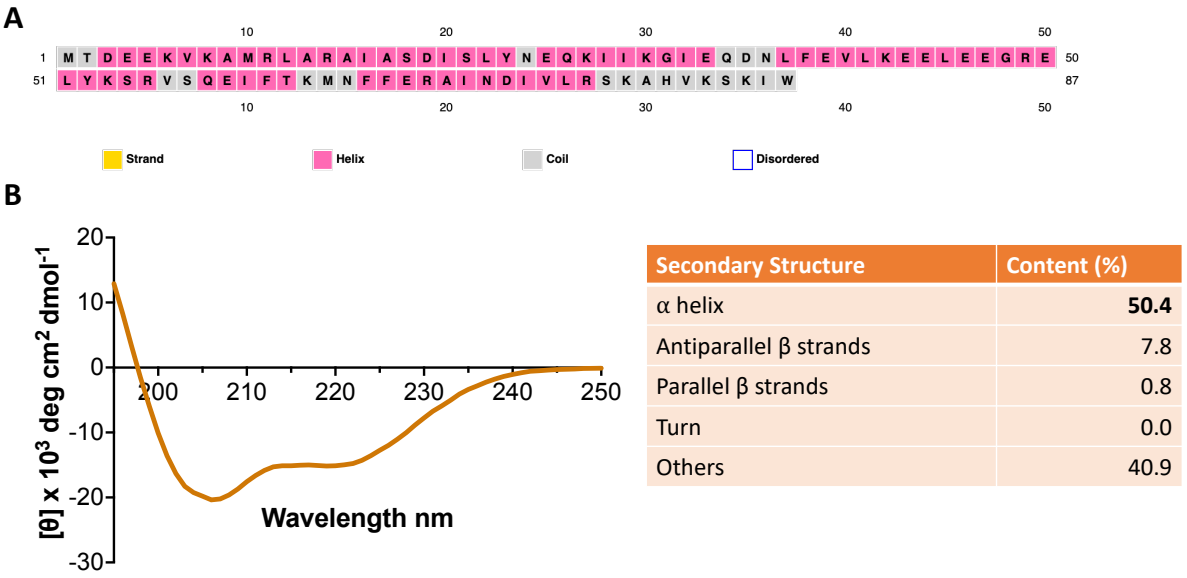
5.2.2 RomX is monomeric in solution

RomX despite being an 11.4 kDa protein, eluted at 13.3 ml, a volume higher than the expected monomeric elution volume (for Superdex 75 it is expected to elute around 15 ml) (Fig 5.2 C). Secondary structure prediction from the RomX sequence shows it is an alpha-helical protein probably capable of forming elongated coils. The elongated shape probably causes it to elute erroneously in SEC as per molar mass (Fig 5.3 A). To resolve this, it was subjected to SEC-MALS to calculate the exact molar mass (Methods: 2.5.3). Here, Superdex 200 SEC column

was used. The observed molar mass was around 13.6 kDa whereas the expected molar mass was around 11.4 kDa (Fig 5.2 D). Hence, it was concluded that RomX was monomeric in solution.

5.2.3 RomX is composed of alpha helices

Elution volumes from SEC and SEC-MALS experiments validated that the RomX structure probably consisted of alpha helices. To confirm this CD spectroscopy was performed using 10 µM RomX in 20 mM phosphate buffer, pH 8 in a 1-mm path length cuvette (Methods: 2.5.2). Buffer containing Tris pH=8 showed large high-tension voltage leading to noisy CD signal. The spectrum obtained had peaks at around 221 nm and 206 nm. When analyzed using BeStSel[6], it showed a high propensity of alpha helices which confirmed that RomX was alpha-helical and monomeric in solution (Fig 5.3 B).



Secondary Structure	Content (%)
α helix	**50.4**
Antiparallel β strands	7.8
Parallel β strands	0.8
Turn	0.0
Others	40.9

Figure 5.3. RomX is composed of alpha helices
A: Secondary structure prediction of RomX using PSIPRED shows it to be composed primarily of alpha helices.

B: The CD spectrum of the RomX protein on the left. Data was fed on to BeStSel server to obtain the percentage of estimated secondary structure content. The fitting shows a major percentage of alpha helices (50.4%), tabulated in right.

5.2.3 Structure determination of RomX

The constructs of RomX with C-terminal and N-terminal hexahistidine tags, respectively, were purified and subjected to crystallization screens (Methods: 2.8). There were multiple hits obtained for C-His-RomX but none for N-His-RomX. This shows that the position of the hexahistidine tag might affect crystal packing and hence, consequently affect the possibility of better crystallization of the protein.

After multiple rounds of optimization, the condition which gave the best reproducible crystals was with 25% PEG MME 550, 5 mM $MgCl_2$ and 50 mM HEPES Na pH 7. The crystals were long needle-shaped and the best crystals were obtained in 24 well hanging drop plates in a 1:0.5 protein: condition ratio. A high concentration of C-His-RomX of around 0.8-1 mM was used for crystallization. Several different cryoprotectants were used out of which 20% glycerol gave the best diffraction spots.

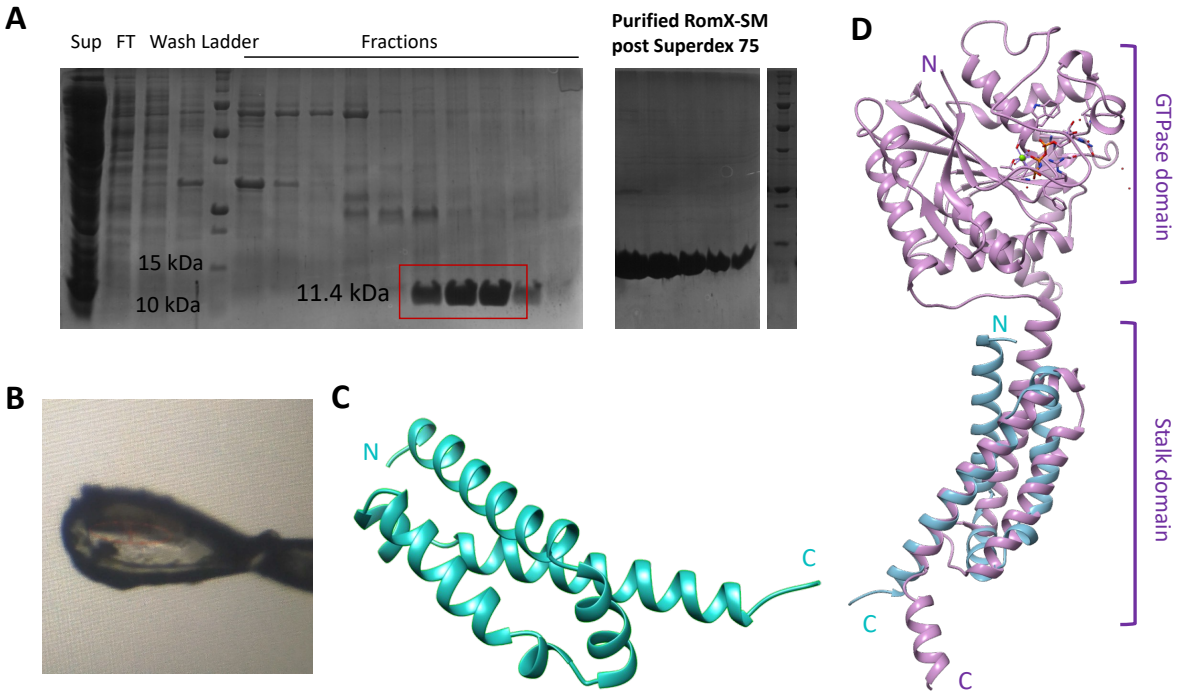


Figure 5.4. Structure determination of RomX using anomalous diffraction

A: SDS PAGE for purification of C-His RomX grown in minimal media supplemented with Selenomethionine. The protein was purified using Ni-NTA followed by Superdex 75. The profiles were the same as earlier purifications.

B. Needle-shaped crystal of RomX mounted on a 0.2 mm cryoloop, during data collection.

C: Crystal structure of C-His RomX solved with SAD (for details refer to Chapter 2). There was one molecule in the asymmetric unit.

D: Superposition of RomX structure on the stalk domain of Atlastin (4IDN), a member of the dynamin family GTPase.

Table 5.1: Data Collection and Refinement Statistics of RomX

Collected at	Diamond Light Source
Wavelength	0.9793
Data Collection Statistics	
Space group	$P2_12_12_1$
a,b,c (Å)	33.3, 56.1, 68.6
α, β, γ (°)	90, 90, 90
Resolution (Å)	1.9
No. of unique reflections	266856 (8682)
R_{merge} (%)	11.5 (117.3)
R_{pim} (%)	3.2 (41.1)
CC_{half}	0.999 (0.699)
Mean I/sigmaI	19.1(2.0)
Completeness (per cent)	99.1(86.6)
Refinement Statistics	
Resolution (angstrom)	29.955-1.88
No. of reflections	19678
R_{work}/R_{free} (%)	18.91/22.16
Average B factor ($Å^2$)	33.71
Wilson B factor ($Å^2$)	27.67
RMS deviation	
Bond lengths (Å)	0.003
Bond angles (°)	0.480
Ramachandran Map	
Favoured (%)	97.8
Allowed (%)	2.2
Outliers (%)	0

*Values in parentheses denote the last resolution shell

Since RomX is a novel protein with no known structural homologs, we had to employ anomalous diffraction techniques using selenomethionine-labelled protein to determine the phase of the diffraction data to solve the crystal structure. Further, the protein was expressed

in minimal media supplemented with selenomethionine and was consequently purified (Fig 5.4 A). The protein was crystallized in the same condition (Fig 5.4 B). The anomalous scattering data was obtained in Diamond Light Source, UK and we obtained an overall resolution of 1.9 Å for the best crystal (Fig 5.2 C). The phases were estimated using the anomalous signal, and the structure was successfully determined (Fig 5.2 C, Table 5.1).

5.2.4. Features of RomX structure

RomX was expectedly found to be an alpha-helical protein. It primarily comprised a three-helix bundle (Fig 5.4 C). There was one single molecule in the asymmetric unit and crystal-packing analysis further confirmed that it was monomeric in solution. PDBePISA analysis showed no probable interface for higher-order assemblies. The single tryptophan at the C-terminal aided in rendering 280 nm absorbance peaks for protein detection unlike in RomR.

The structure of RomX was given as an input to the DALI server and it returned several hundred hits for proteins with one or more chains in the same fold as RomX, from within the Protein Data Bank. The top 50 structures in the DALI output were superimposed onto the RomX structure using the software UCSF Chimera. The RMSD values of all the superimposed structures were recorded. Based on the RMSD values obtained a total of 19 structures were shortlisted for further analysis (Table 5.2). Many of the shortlisted structures were GTPases belonging to the atlastin family and, consequently, had a GTPase domain, and a stalk domain. The stalk domain was the part that superimposed onto RomX owing to the same fold (Fig 5.4 D). Interestingly, the stalk domain of atlastin in humans and Sey1p in yeasts are the proteins most structurally similar to RomX, whereas their associated domains are functionally similar to MglA (GTP hydrolysis), the protein with which the RomR-RomX complex interacts.

Atlastin showed 3 conformations owing to the binding of 3 different ligands to them - 1) GDP-bound state, 2) GDP-AlF4-bound state (transition state), and 3) GNP-bound state[7,8]. So, the 3 structures from each of these classes were shortlisted based on the highest resolution and lowest RMSD values (upon superimposition with RomX), to study the change in conformation caused by the movement of domains and helices (with Mr. Kushan Lahiri). Central axes were drawn through the stalk domain and the GTPase domain to find the angles between them (H1-H2, H2-H3, H1-H3, in Table 5.2) and look for movement between the two domains (GTPase stalk domain angle, Fig 5.5 A, Table 5.2). On superimposing the three phases of the atlastin molecules, it was found that the GDP-bound state (6B9D) was conformationally different from

the GNP-bound (4IDN)[4], and GDP-AlF4-bound[3] (6B9F) states, mainly in terms of the angle between the central axes that were drawn through the GTPase and the stalk domains (Fig 5.5 B). The conformational change between the GDP-bound and GTP-bound states of atlastin occurs in the orientation of the GTPase domain with respect to the stalk domain (observed as a movement in the GTPase domain when superposed with the stalk domains). This conformational change does not exist between GTP- bound and GDP-AlF_4-bound (transition) states.

Table 5.2: List of homologous structures of RomX identified by DALI

Chain	Name	Organism	Ligands bound	Z score	RMSD	Nres	%id	Resolution	RMSD while superposition	GTPase Stalk Domain Angle	H1-H2	H2-H3	H1-H3
6b9f-A	ATLASTIN-1;	Homo sapiens	GDP, ALF, GOL, MG	7.6	3.1	411	13	1.90 Å	0.816Å	38.2°	23.5°	21.5°	20.0°
6b9d-A	ATLASTIN-1;	Homo sapiens	GDP, GOL, MG	6.9	2.7	379	14	1.95 Å	0.707Å	63.1°	28.1°	23.3°	25.9°
4idn-A	ATLASTIN-1;	Homo sapiens	GNP, MG	6.9	2.6	417	14	2.25 Å	0.755Å	35.3°	23.6°	22.6°	17.1°
5cb2-A	PROTEIN SEY1;	Candida albicans SC5314	GNP. MG	7.9	2.7	656	10	2.9 Å	1.140Å	Very long stalk domain, no interaction with GTPase domain	27.0°	24.2°	20.0°
5ca9-A	PROTEIN SEY1;	Candida albicans SC5314	GDP, ALF, MG	7.7	2.8	656	14	2.8 Å	0.778Å	Same as above	23.7°	24.0°	18.1°
5ca8-A	PROTEIN SEY1;	Candida albicans SC5314	GDP, MG	6.8	2.9	636	15	2.30 Å	0.809Å	Same as above	22.7°	21.3°	15.9°
5v7p-A	PROTEIN-S-ISOPRENYLCYSTEINE O-METHYLTRANSFERASE;	Tribolium castaneum	SAH, MPD, UND, D10	7.7	3.7	281	4	2.3 Å	1.188 Å	No GTPase domain	23.6°	22.6°	18.5°
3swh-A	PROTEIN UNC-13 HOMOLOG A;	Rattus norvegicus	-	7.2	3.8	295	10	2.65 Å	1.166Å	No GTPase domain	15.1°	4.3°	16.9°
1ldj-A	CULLIN HOMOLOG 1;	Homo sapiens	ZN	7.1	2.9	725	9	3.0 Å	0.334Å	No GTPase domain	16.3°	20.7°	34.2°
1z8u-C	ALPHA-HEMOGLOBIN STABILIZING PROTEIN;	Homo sapiens	HEM	6.9	3	91	10	2.40 Å	1.265 Å	No GTPase domain	29.0°	25.2°	15.6°
6k1z-A	GUANYLATE-BINDING PROTEIN 1;	Homo sapiens	FAR	6.8	2.5	544	11	2.31 Å	0.793Å	No GTPase domain	29.8°	26.4°	15.1°
6wb9-3	ENDOPLASMIC RETICULUM MEMBRANE PROTEIN COMPLEX SU	Saccharomyces cerevisiae W303	POV, NAG	6.7	2.9	197	4	3.00Å	1.474Å	No GTPase domain	17.0°	28.0°	19.5°
6rvy-A	NEUTRAL AMINO ACID TRANSPORTER B(0);	Homo sapiens		6.7	3.5	430	5	4.13Å	0.712Å	No GTPase domain	29.2°	44.3°	39.3°
6me0-C	T.EL4H RNA;	Thermosynechococcus elongatus	MG, NA	6.7	3.7	414	9	3.60Å	0.718Å	No GTPase domain	18.7°	22.1°	22.6°
2odm-B	UPF0358 PROTEIN MW0995;	Staphylococcus aureus subsp. aureus MW2	MSE	6.6	4.9	83	10	2.24Å	0.552Å	No GTPase domain	13.5°	7.9°	5.9°
5l90-A	CYTOCHROME P450;	Bacillus megaterium DSM 319	HEM	6.6	2.8	385	6	2.55Å	0.844Å	No GTPase domain	60.7°	42.6°	27.6°
4l8i-A	RSV EPITOPE SCAFFOLD FFL_005;	Synthetic construct		6.6	4.2	114	11	2.00Å	0.58Å	No GTPase domain	7.1°	8.8°	4.6°
6tyi-E	BIOPOLYMER TRANSPORT PROTEIN EXBB;	Escherichia coli K-12	PGT, PEV	6.5	3	224	5	3.30Å	1.342Å	No GTPase domain	29.2°	30.2°	32.6°

There might be a possibility for RomX to interact with the MglA at a location similar to the linkage between the GTPase and the stalk domains of Atlastin. Since a GEF typically interacts with the GTPase in the GDP-bound state, Atlastin-GDP (6B9D) was chosen as the structure of reference while superimposing. There is a high propensity for Atlastin to dimerize in the GDP-AlF_4-bound (transition) state, through both the GTPase and stalk domains. However, in the GDP bound state, the stalk domains are separated in space and it dimerizes only through its

GTPase domain (Fig 5.5 C). Hence, when two RomX chains are superposed on the Atlastin-GDP dimer, clearly the monomers are far apart with no scope for dimerization. This indicates, that it is probably a monomer of RomX which interacts with MglA in the GDP-bound state to act as a GEF. The figure also highlights that the connection between the stalk domain and the GTPase in Atlastin is away from the nucleotide-binding pocket, unlike interactions observed in typical GEFs, which sit near the nucleotide-binding pocket competing out the bound GDP (Fig 5.5 D). This offers a possibility for RomX to act near the MglA allosteric pocket, the same site where the MglB-Ct GEF binds (shown in Chapter 4), indicating some possible mechanisms of competition between the two GEF modules.

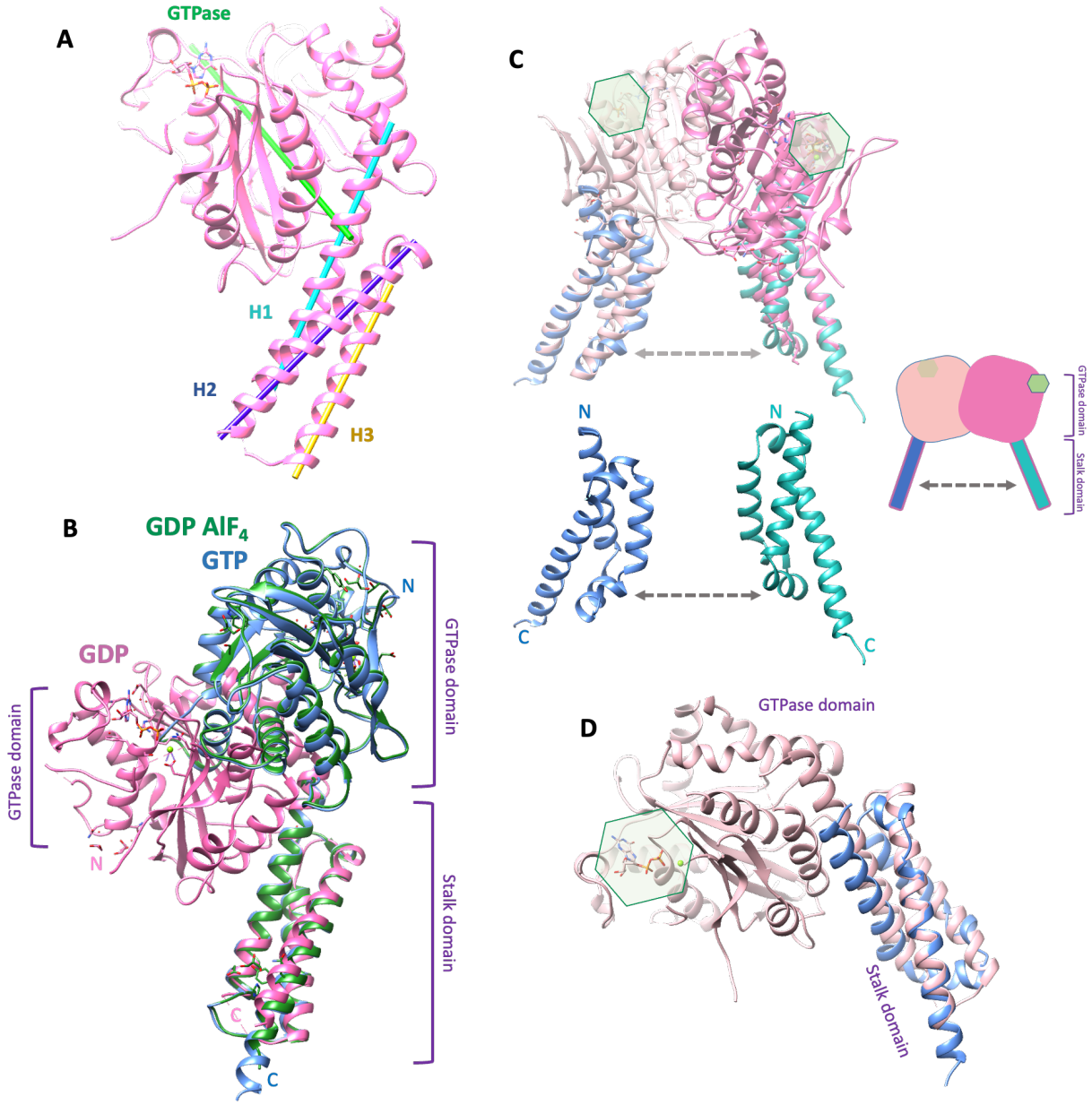


Figure 5.5. Structural superposition of RomX and Atlastin
A: Structure of atlastin in GDP (6B9D) boud state in pink. The axes were drawn through the GTpase domain (green) and the 3 helices in the stalk domain (H1: cyan, H2: blue, H3: ochre). The angles between these axes were calculated and tabulated in Table 5.2.

B: Superimpositions of atlastin molecules. The stalk domain is used to superpose the 3 structures of Atlastin corresponding to 3 nucleotide states, GTP state (4IDN) in blue, GDP AlF_4 transition state (6B9F) in green and GDP (6B9D) in pink. The N and C termini are labelled in respective colours.

C: Superposition of RomX structures on the stalk domains of Atlastin (6B9D) dimer in GDP state (chain A and B in hot and light pink, respectively). The green hexagon denotes the bound GDP. The separation in space between the 2 RomX molecules in blue is highlighted in the lower panel where the Atlastin dimer is hidden.

D: Superposition of RomX on the stalk domain of Atlastin (6B9D) in GDP state, showing that the GDP binding pocket (green hexagon) is away from the connection between the stalk domain and GTPase domain

5.2.5 Interaction of RomX with MglA

To confirm the previous hypotheses, we tried to capture the interaction of MglA with RomX in presence of the relevant nucleotide. Analytical SEC was carried out using MglA and RomX in presence of GTP and GDP respectively (Figure 5.6 A, B, Methods: 2.7.1)). MglA and RomX eluted at 13.6 ml and 14 ml, respectively. These volumes were very close to conclusively separate the peaks. Since there was no shift of the resultant peak of a mix of MglA and RomX towards a lower elution volume, we could not prove any interaction between the two proteins in the presence of either nucleotide (Fig 5.6 A, B). This was probably because the interaction was too weak to be captured by SEC. We further performed a more sensitive, fluorescence anisotropy measurement with *mant*-labelled GTP and GDP where we titrated increasing concentrations (1-12 μM) of RomX to 2 μM MglA. Preliminary results did not show any increase in anisotropy of MglA upon RomX addition. Hence, we could not conclusively estimate if RomX binds to MglA either in presence of GTP or GDP (Figure 5.6 C, D).

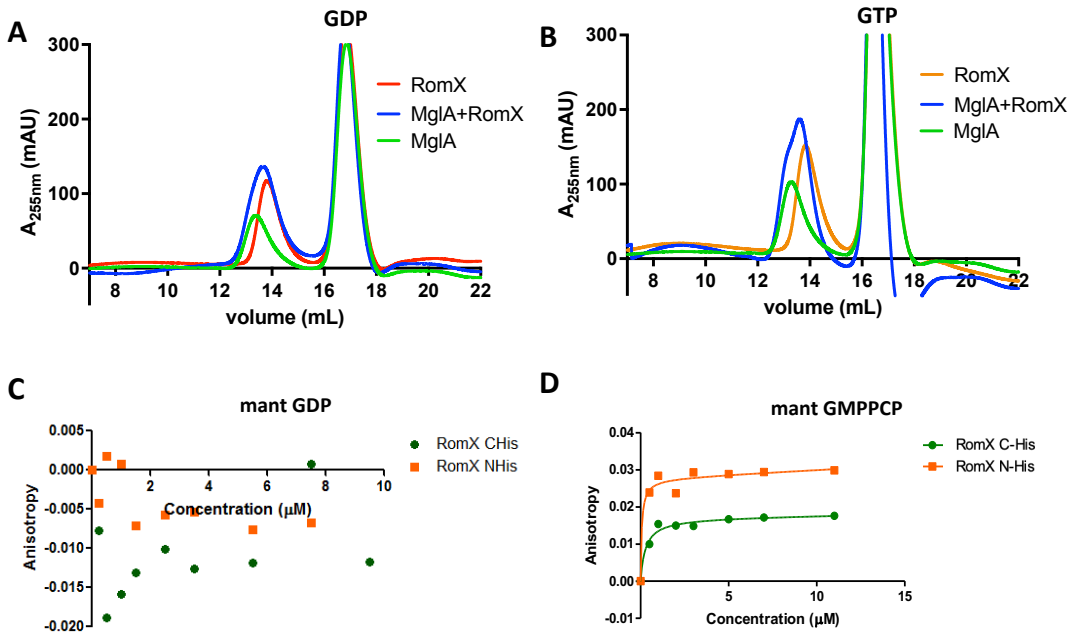


Figure 5.6. Binding studies of RomX and MglA

A, B: Analytical Superdex 75 to detect binding of RomX to MglA in presence of GDP and GTP, respectively. No significant binding was observed as MglA (magenta) and RomX (blue) eluted at very close volumes between 13-14 ml and there was no obvious shift in the elution volume at the peak of the MglA+RomX (green).

C, D: Representative plots for fluorescence anisotropy studies to quantify binding of RomX to MglA in presence of *mant*-GDP and *mant*-GMPPCP respectively. There was no binding observed with *mant*-GDP, whereas, in presence of *mant*-GMPPCP, MglA precipitated out in the cuvette, providing no conclusive and consistent results.

5.2.6 RomX acts as MglA GEF

Since we could not capture a direct interaction of RomX with MglA, we tried to estimate the effect of RomX on MglA activity, as an indirect readout for the probable weak transient interaction. There was no stimulation in the MglA GTP hydrolysis rates upon the addition of RomX (Fig 5.7 A, Methods: 2.6.1). The results were consistent even when higher concentrations of RomX (1:4 MglA: RomX) were used. MglA+MglB (1:2) was used as a positive control which showed the GAP stimulation of MglA.

Nucleotide exchange assays were performed in which *mant*-GMPPNP association was monitored (Methods: 2.6.2). It was observed that in the presence of RomX, MglA-GDP consistently showed active nucleotide exchange and showed binding with mant-GMPPNP (Fig 5.7 A). This demonstrated that RomX exhibited a nucleotide exchange activity on MglA,

similar to earlier reports. However, the binding is not stable, as observed with MglB. This probably results in an intrinsic decrease in fluorescence intensity shortly after the association. This also confirms that RomX indeed interacts with MglA but the interaction is probably transient and low-affinity.

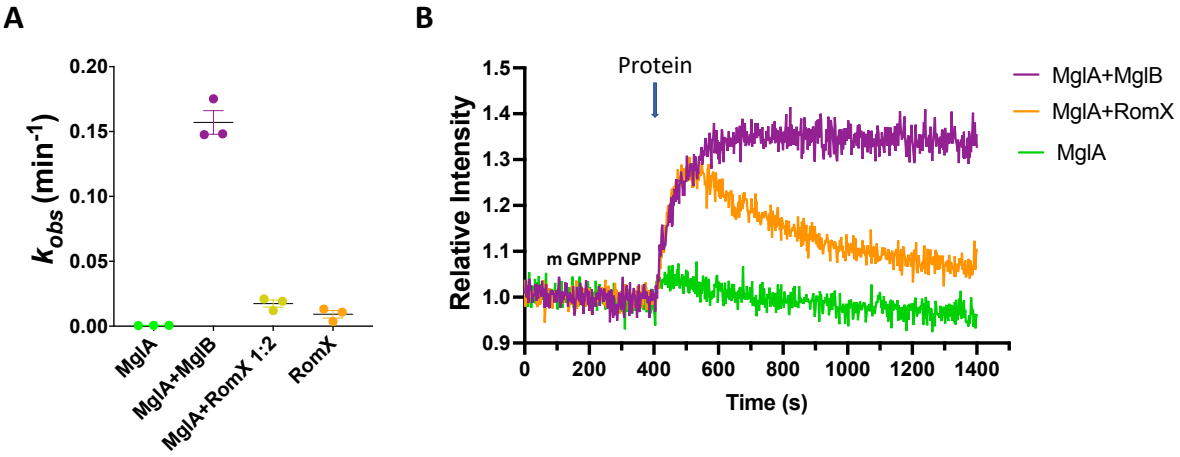


Figure 5.7. RomX acts as a MglA GEF

A: NADH coupled GTPase assay data showing that MglB is an active GAP but not RomX. The black line represents the mean and the coloured bars represent the standard error.

B: Representative plots (from n=5 repeats) showing *mant*-GMPPNP association assay of MglA. At 400 s, the protein was added, shown by the arrow. RomX stimulates the release of GDP on MglA and uptake of *mant*-GMPPNP in a 1:2 ratio (MglA 3 µM: RomX 6 µM), whereas MglA alone does not show exchange. In presence of MglB, there is a steady signal for binding, unlike in presence of RomX, there is a sharp decline.

5.3 Conclusion

The newly identified member of the polarity module, RomX, was characterized biochemically and structurally. It was shown to be an alpha-helical protein which is monomeric in solution. It is interesting to note that a fold similar to RomX is associated with a GTPase domain in human Atlastin, a member of the dynamin family of GTPases. In Atlastin, the GTPase domain and stalk domain is linked by amino acid residues, thereby forming one continuous chain. That is not the case in *Myxococcus xanthus*, where the GTPase equivalent MglA and the stalk equivalent RomX are not linked together but probably only interact with each other in a 1:1 ratio.

Based on the interaction studies between MglA and RomX, it seems likely they interact weakly. To predict the interface between MglA and RomX, we used AlphaFold to generate a complex of MglA with RomX. Interestingly, AlphaFold prediction shows RomX to interact with MglA in the GDP state, typical for a GEF (Fig 5.8 A,B). We indeed observe the GEF activity of RomX on MglA, but the GTP association is not stabilized, as observed with MglB. We closely inspected the interface residues between RomX and MglA (Fig 5.8 C). There are arginines (Arg-11,14) of RomX and glutamates (Glu-108,111,115) of MglA which come in close proximity and could be holding the complex together. Moreover, Asp-19 of RomX inserts into the MglA GDP binding pocket, near the Switch II, which could be possibly the residue driving the GEF activity. This is similar to the mechanism of eukaryotic GEFs which inserts an acidic residue to destabilize GDP binding[9]. The RomX-MglA interface and the consequent mechanism of RomX GEF activity could be validated in future by mutating these aforementioned residues.

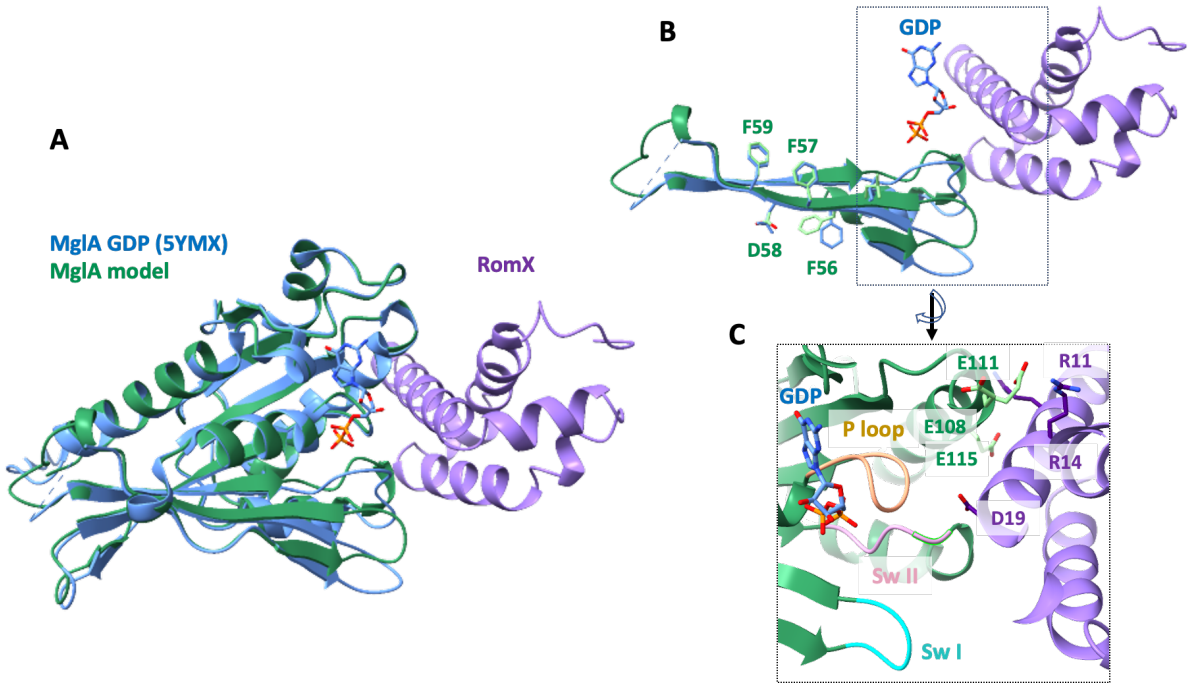


Figure 5.8. AlphaFold model of MglA with RomX

A: RomX was seen to interact with MglA in the GDP state, shown by the superposition of MglA GDP (PDB ID: 5YMX) in blue, with the model in green. RomY is shown in purple.

B: The GDP state of MglA is highlighted by the conformation of the β2 strand of MglA with the "LFFDF" motif in the buried conformation (Refer to Chapter 3).

C: Interaction interface between RomX and MglA. The interface residues are highlighted (MglA residues in green and RomX residues in purple). MglA P-loop, Switch I, Switch II are highlighted in orange, pink and cyan, respectively.

The interface between MglA and RomX was not predicted to be as extensive as that between MglA and MglB. Hence the interaction of RomX to MglA could be further stabilized in the presence of other proteins like RomR, with which RomX is known to form a complex. Hence, in the next chapter, we investigate the partner of RomX, RomR and attempt to further characterize the effect of the RomR-RomX complex in regulating MglA activity.

References

1. Leonardy S, Miertzschke M, Bulyha I, Sperling E, Wittinghofer A, Søgaard-Andersen L. Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. *EMBO J* **9**, 2276–2289 (2010).
2. Miertzschke M, Koerner C, Vetter IR, Keilberg D, Hot E, Leonardy S, Søgaard-Andersen L, Wittinghofer A. Structural analysis of the Ras-like G protein MglA and its cognate GAP MglB and implications for bacterial polarity. *EMBO J*. 2011 Aug 16;30(20):4185-97. doi: 10.1038/emboj.2011.291.
3. Zhang, Y., Franco, M., Ducret, A. & Mignot, T. A bacterial Ras-like small GTP-binding protein and its cognate GAP establish a dynamic spatial polarity axis to control directed motility. *PLoS Biol* **8**, (2010).
4. Baranwal J, Lhospice S, Kanade M, Chakraborty S, Gade PR, Harne S, Herrou J, Mignot T, Gayathri P. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. *PLoS Biol.* 2019 Sep 27;17(9):e3000459. doi: 10.1371/journal.pbio.3000459.
5. Szadkowski D, Harms A, Carreira LAM, Wigbers M, Potapova A, Wuichet K, Keilberg D, Gerland U, Søgaard-Andersen L. Spatial control of the GTPase MglA by localized RomR-RomX GEF and MglB GAP activities enables *Myxococcus xanthus* motility. *Nat Microbiol*. 2019 Aug;4(8):1344-1355. doi: 10.1038/s41564-019-0451-4.
6. Micsonai A, Wien F, Bulyáki É, Kun J, Moussong É, Lee YH, Goto Y, Réfrégiers M, Kardos J. BeStSel: A web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315–W322 (2018).
7. O'Donnell, J. P., Byrnes, L. J., Cooley, R. B. & Sondermann, H. A hereditary spastic paraplegia-associated atlastin variant exhibits defective allosteric coupling in the catalytic core. *JBC* **293**, 687–700 (2018).
8. Hu, X., Wu, F., Sun, S., Yu, W. & Hu, J. Human atlastin GTPases mediate differentiated fusion of endoplasmic reticulum membranes. *Protein Cell* **6**, 307–311 (2015).
9. Toma-fukai, S. & Shimizu, T. Structural Insights into the Regulation Mechanism of Small GTPases by GEFs. *Molecules* 1–31 (2019).

Chapter 6: **Functional dissection of multiple domains of the response regulator, RomR**

6.1 Introduction

In the previous subsection, we have shown that RomX acts as a GEF for MglA, but the interaction between them is weak and transient. RomX has been shown to interact with its partner RomR, which together have been shown to act as two of the key regulators of *M. xanthus* polarity reversals[1]. RomR might play a role in stabilizing the interaction of RomX on MglA, thereby accelerating its GEF activity, as reported earlier[1]. Having deciphered the mechanism of GEF stimulation of MglA by MglB described in Chapter 4, it is important to understand the mechanism of RomX-RomR mediated GEF activity. Hence, we have tried to dissect the interaction between RomX and RomR and further attempted to understand how this protein complex interacts with MglA.

RomR has an N-terminal response receiver domain (RomR-REC; Fig 6.1 A [2]) and a C-terminal output domain [2]. The output domain consists of a long unstructured linker region followed by a C-terminal helical tail (hereby referred to as RomR-Ct). RomR localizes similarly to MglB and RomX, in a bipolar asymmetric fashion with a larger cluster at the lagging pole (Fig 6.1 B). *In vivo* studies of motility behaviour and localization studies show that in the absence of the RomR-REC, the output domain (residues 117-420) was unable to switch across poles and the cells were defective in gliding motility[2]. Further, in the absence of the output domain, the RomR-REC (residues 1-140), did not localize to the cell poles and remained diffused, causing defects in A-motility[2]. This showed that the receiver domain helps recruit RomR to the poles and the output domain is important for RomR dynamics during reversals[2].

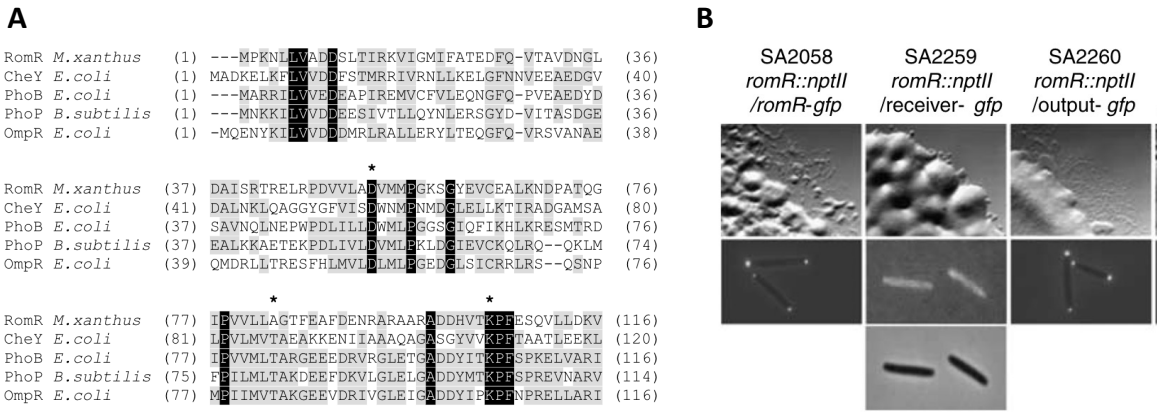


Figure 6.1. *In-vivo* studies with RomR domains
A: Sequence alignment of the RomR N-terminal receiver domain with recognized receiver domains. Asterisks indicate conserved residues, signature of receiver domains. Black and grey residues show 100% and 60-70% identity, respectively (Reproduced from Leonardy *et al*., 2007[2]).

B. Motility phenotypes of a *romR* deletion mutant when supplemented with different RomR domains tagged to GFP. The top panels show the motility of the cells in 1.5% CTT agar media and the lower panel shows fluorescence microscopy images of the cells from the respective cultures. Cells with receiver domain has diffused RomR localization. With only the output domain, RomR was static at the poles with no reversals (Reproduced from Leonardy *et al*., 2007[2]).

RomR bridges with the chemosensory Frz module and consequently acts in concert with MglA and MglB to orchestrate polarity reversals [2–4]. It has been debated if RomR goes through a phosphorylation-mediated switch to regulate reversals like other response regulators[3,5,6]. The output of the Frz module is delivered by a kinase which phosphorylates one of its downstream regulators, FrzX, which acts at the lagging cell pole[3]. Phosphorylation of Asp-53 of RomR-REC was shown to be important for RomR dynamics[7]. However, it was also shown that the Frz signals could be directly communicated to MglA raising questions about the importance RomR phosphorylation in driving polarity reversals[3]. Upon receiving the trigger from phosphorylated FrzX, RomR has been shown to initiate the polarity reversal but shows slower dynamics than MglA/MglB, thereby forming a 'gated relaxation oscillator'[8]. Further, it promotes correct polar localization of MglB at the lagging pole via MglC, another homodimeric Roadblock domain protein like MglB, which plays a yet unknown role in the polarity module[9,10]. Apart from its role as a polar recruitment factor, RomR has been shown to amplify the GEF activity or RomX on MglA[1]. Consequently, RomR-RomX together helps recruit MglA at the lagging pole, resulting in the conversion of the lagging pole to the leading pole[11]. However, molecular details of neither RomR/RomX nor MglA/RomR/RomX interaction have been characterized. RomR, being a multidomain protein, is interesting to decipher how it interacts with the polarity module proteins and regulates the activity of MglA driving motility. Hence in this chapter, we attempt to determine the exact functional domain boundaries of RomR and how these individual domains interact with RomX and MglA to drive polarity reversals.

6.2 Results

6.2.1 Cloning and purification of RomR domains

RomR Receiver (REC) domain

Sequence alignment with other response receivers identified residues 1-116 as the conserved N-terminal receiver domain (Fig 6.1 A). Previous *in-vivo* studies were performed using a construct with 1-140 residues, which was longer than the predicted REC domain (Fig 6.1 B). To determine the exact boundary of the REC domain, we cloned a construct with residues 1-116 or RomR (RomR-REC^{116}) (Methods: 2.2.3) (Fig 6.2 A, B, C). RomR-REC^{116} with C-terminal hexahistidine tag was insoluble, however, when expressed with N-terminal hexahistidine tag the protein was retained in the soluble fraction. Further, this protein was attempted to be purified using multiple trials of hexahistidine tag based affinity chromatography, ammonium sulphate precipitation, anion exchange chromatography and size exclusion chromatography. However, protein remained impure.

We further investigated the secondary structure prediction using PSIPRED and observed that the 116th residue lay in the middle of a helix in the REC domain which might cause the protein to partially unfold and consequently aggregate. To resolve this, we further cloned the REC domain with residues 1-125 and 1-141 (Fig 6.2 C). These constructs were well soluble and were purified using hexahistidine tag-based affinity chromatography followed by size exclusion chromatography (Fig 6.2 D, Methods: 2.4). Though the protein appeared to be reasonably pure in SDS-PAGE gel, background activity was observed in the GTPase assays indicating the presence of contaminating enzymes. Hence the protein was subjected to ion exchange chromatography using MonoQ after which it was obtained with high purity. It eluted between 37-42 mS cm^{-1} (Fig 6.2 F).

C-terminal (RomR-Ct) domain

C terminal region of RomR (RomR-Ct, residues 331-420) was cloned expressed in the same conditions as RomR-REC constructs (Methods: 2.2.3) (Fig 6.2 C). It was purified as a 12 kDa protein using Ni-NTA, SEC (Superdex 75) and MonoQ. A shorter construct of RomR, RomR-REC^{141}Ct linking the REC domain (1-141 residues) and the C terminal (331-420) was also designed and expressed and purified using the same aforementioned steps (Fig 6.2 C, E).

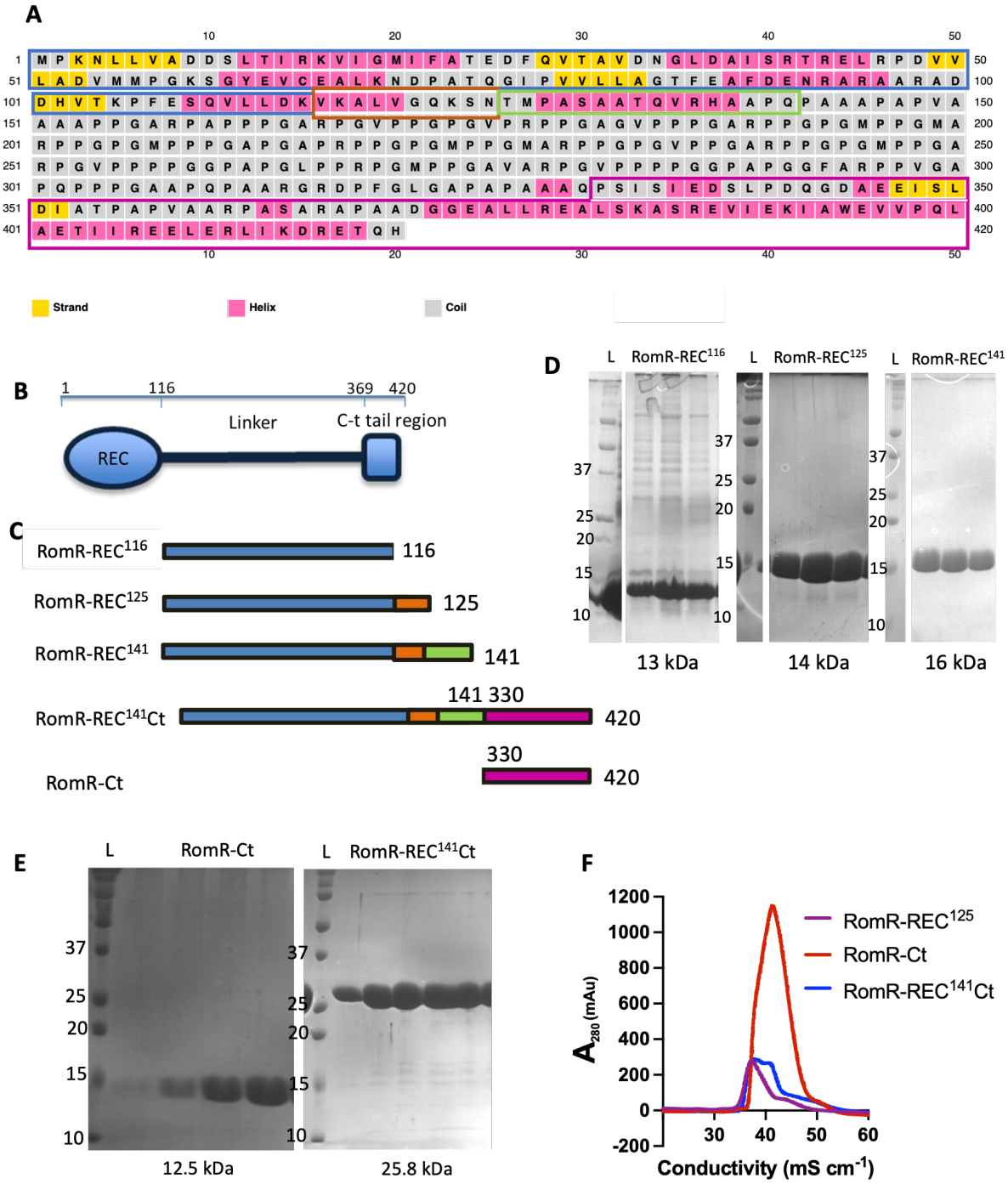


Figure 6.2. Domain-wise dissection of RomR using different constructs

A: Secondary structure prediction of RomR using PSIPRED. The three different boundaries of the REC domain are highlighted in blue (116), orange (125) and green (141), respectively. The C-terminal domain is highlighted in magenta.

B. Schematic showing the domains of RomR, N-terminal REC domain, a long-disordered linker followed by a C-terminal helical domain.

C: The different domain-wise dissected constructs of RomR that were designed for this study. The colours of the REC and the C-terminal domains are as per A. The residue numbers of the boundaries are labelled.
D. SDS PAGE showing purified proteins of the three REC domain constructs of RomR after Superdex 75 (SEC) step.
E. SDS PAGE showing purified proteins of the two C-terminal domain constructs of RomR after MonoQ (anion exchange) step.
F. MonoQ profile for RomR constructs. RomR-REC^{125} , RomR-REC^{141}Ct, RomR-Ct in purple, blue and red, respectively.

6.2.2 Interaction of RomR domains with RomX and MglA

For the interaction studies, analytical gel filtration chromatography was performed with RomR-REC^{125} and RomX in presence of MglA (Methods: 2.7.1). No coelution was observed as there were 2 distinct peaks around 12.1 ml corresponding to the RomR-REC^{125}+RomX complex and at 13.1 ml to MglA, as verified by SDS PAGE (Fig 6.2 A). This observation was consistent even when the same experiment was performed in the presence of GTP and GDP respectively (Fig 6.3 A, B, C). A similar trend was observed for RomR-REC^{141} (Fig 6.4 A, B, C).

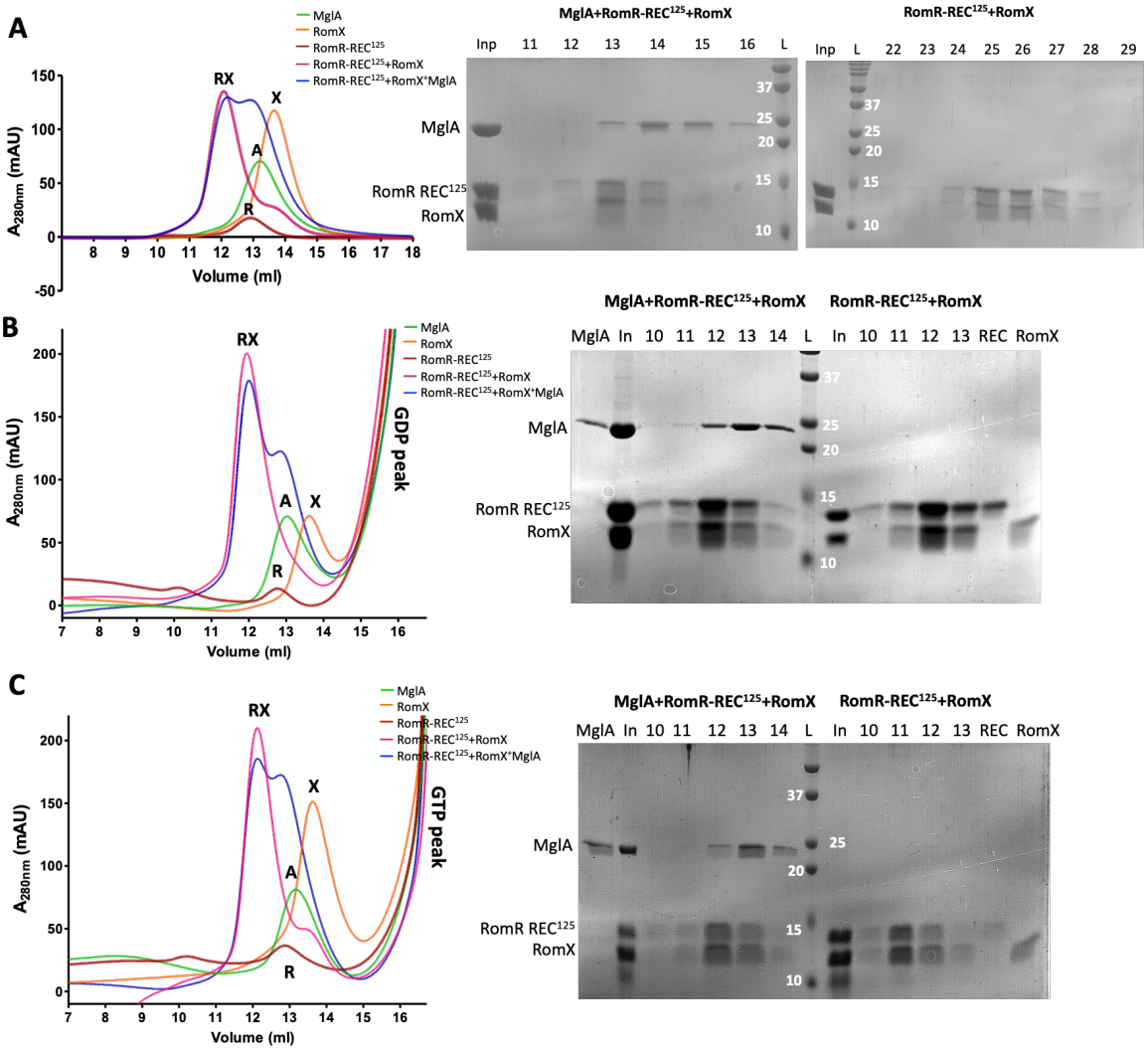


Figure 6.3. Interaction of RomR-REC^{125} with RomX and MglA

A, B, C: Analytical Superdex 75 runs showing the interaction of RomR-REC^{125} with RomX and MglA without nucleotide (A), with GDP (B) and with GTP (C). The axes are cut to mask the high peaks of GDP and GTP elution at 18 ml, as shown in the figure. The right panels show the corresponding SDS PAGE images of the respective runs. The lane numbers indicate the volumes of elution as per the chromatograms (In: Input, L: Ladder, molecular weights are labelled) and the proteins corresponding to the bands are labelled. The corresponding peaks in the chromatograms are also labelled.

MglA (green), RomX (orange), RomR-REC^{125} (brick red), RomR-REC^{125} + RomX (magenta) and MglA+ RomR-REC^{125} + RomX (blue)

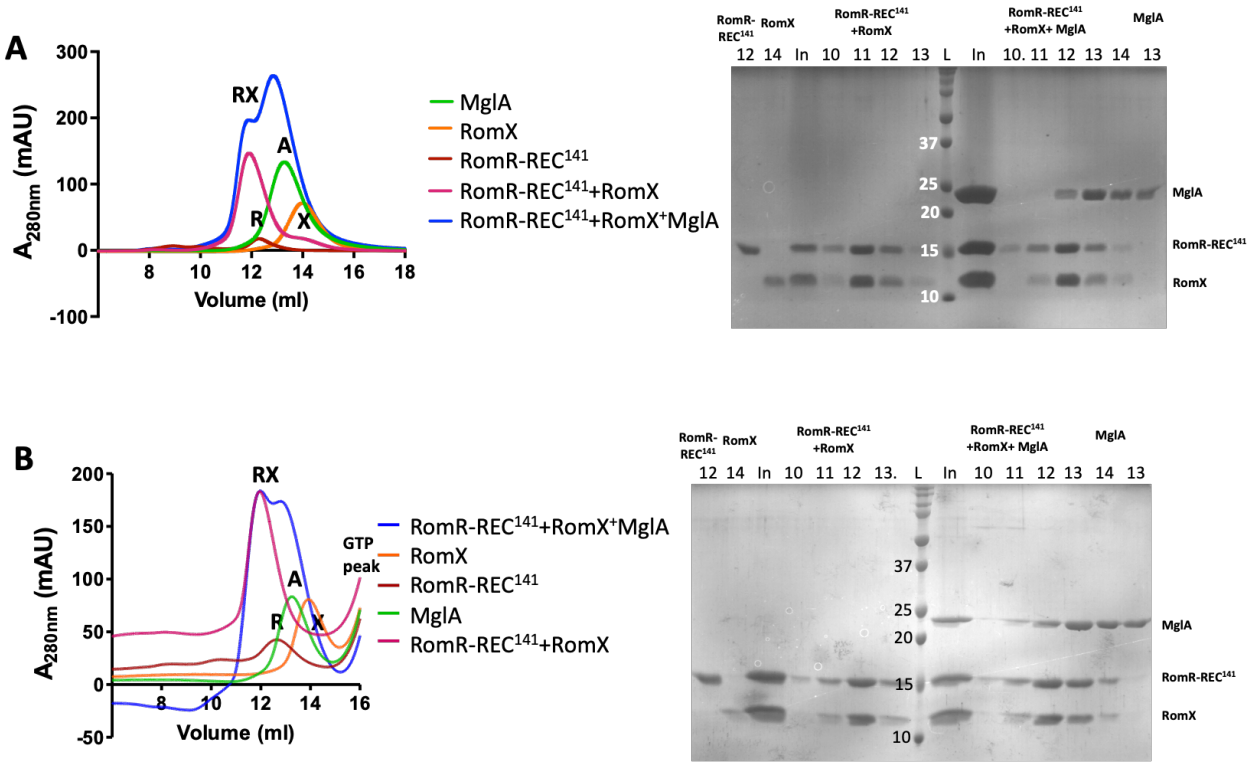


Figure 6.4. Interaction of RomR-REC^{141} with RomX and MglA

A, B, C: Analytical Superdex 75 runs showing the interaction of RomR-REC^{141} with RomX and MglA without nucleotide (A), with GTP (B). The axes are cut to mask the high peak of GTP elution at 18 ml, as shown in the figure. The right panels show the corresponding SDS PAGE images of the respective runs. The lane numbers indicate the volumes of elution as per the chromatograms (In: Input, L: Ladder, molecular weights are labelled) and the proteins corresponding to the bands are labelled. The corresponding peaks in the chromatograms are also labelled.

MglA (green), RomX (orange), RomR-REC^{141} (brick red), RomR-REC^{141} + RomX (magenta) and MglA+ RomR-REC^{141} + RomX (blue)

It was observed that the RomR-Ct does not interact with RomX based on SEC (Superdex 200 column). There was a clear separation of peaks observed at 12.3 ml for RomR-Ct and 17.6 ml for RomX (Fig 6.5). This confirms that the REC domain of RomR is involved in interaction with RomX and not the C-terminal helical domain. The shorter linked RomR-REC^{141}Ct interacted with RomX and eluted as a single peak at 10.5 ml in Superdex 200, possibly solely due to the presence of the REC domain. This was further confirmed in the SDS PAGE where we observe co-elution of RomR-REC^{141}Ct with RomX (Fig 6.6 A).

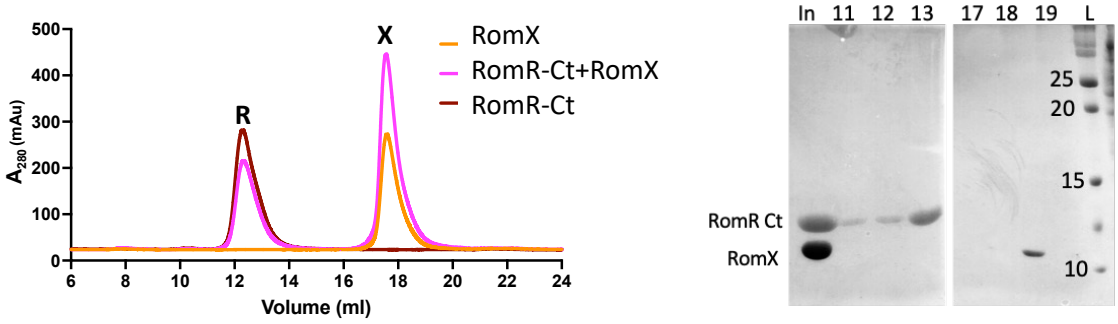


Figure 6.5. Interaction of RomR-Ct with RomX

Analytical Superdex 200 run showing no interaction of RomR-Ct with RomX. The lower panel shows the corresponding SDS PAGE image. The lane numbers indicate the volumes of elution as per the chromatograms (In: Input, L: Ladder) and the proteins corresponding to the bands are labelled. The corresponding peaks in the chromatograms are also labelled.

RomX (orange), RomR-Ct (brick red), $RomR^{REC-Ct}$ + RomX (magenta)

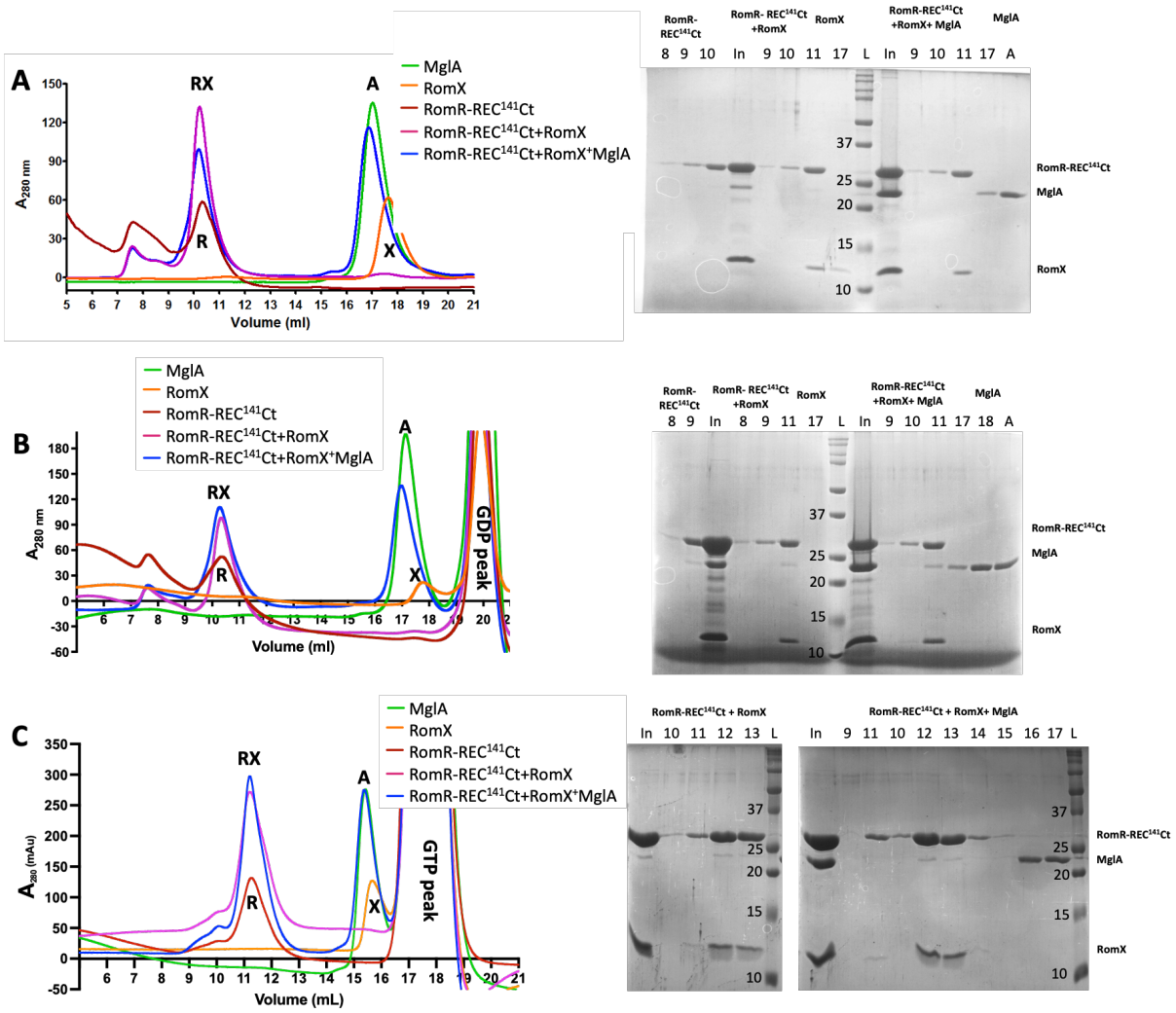


Figure 6.6. Interaction of RomR-REC^{141}Ct with RomX and MglA

A, B, C: Analytical Superdex 200 runs showing the interaction of RomR-REC^{141}Ct with RomX and MglA without nucleotide (A), with GDP (B) and with GTP (C). The axes are cut to mask the high peaks of GDP and GTP elution, as shown in the figure. The right panels show the corresponding SDS PAGE images of the respective runs. The lane numbers indicate the volumes of elution as per the chromatograms (In: Input, L: Ladder) and the proteins corresponding to the bands are labelled. The corresponding peaks in the chromatograms are also labelled.
MglA (green), RomX (orange), RomR- REC^{141}Ct (brick red), RomR- REC^{141}Ct + RomX (magenta) and MglA+ RomR- REC^{141}Ct + RomX (blue)

Much like in the case of the RomR-REC domain constructs, the RomR-REC^{141}Ct -RomX complex did not interact with MglA in the SEC column to form a heteromeric complex in the presence or absence of nucleotides, namely GDP and GTP (Fig 6.6 A, B, C). There was a clear separation of peaks for MglA at 17 ml and RomR-REC^{141}Ct-RomX at 10 ml and this was confirmed with SDS PAGE. Hence, we cannot observe the formation of a MglA and RomR-

RomX complex in SEC. This is probably because the interaction is mediated solely by RomX with no contribution from RomR constructs under study, and is hence low-affinity and transient. Expectedly the observation was similar for RomR-Ct, which neither interacted with RomX, nor with MglA in the presence or absence of nucleotides (Fig 6.7 A, B, C).

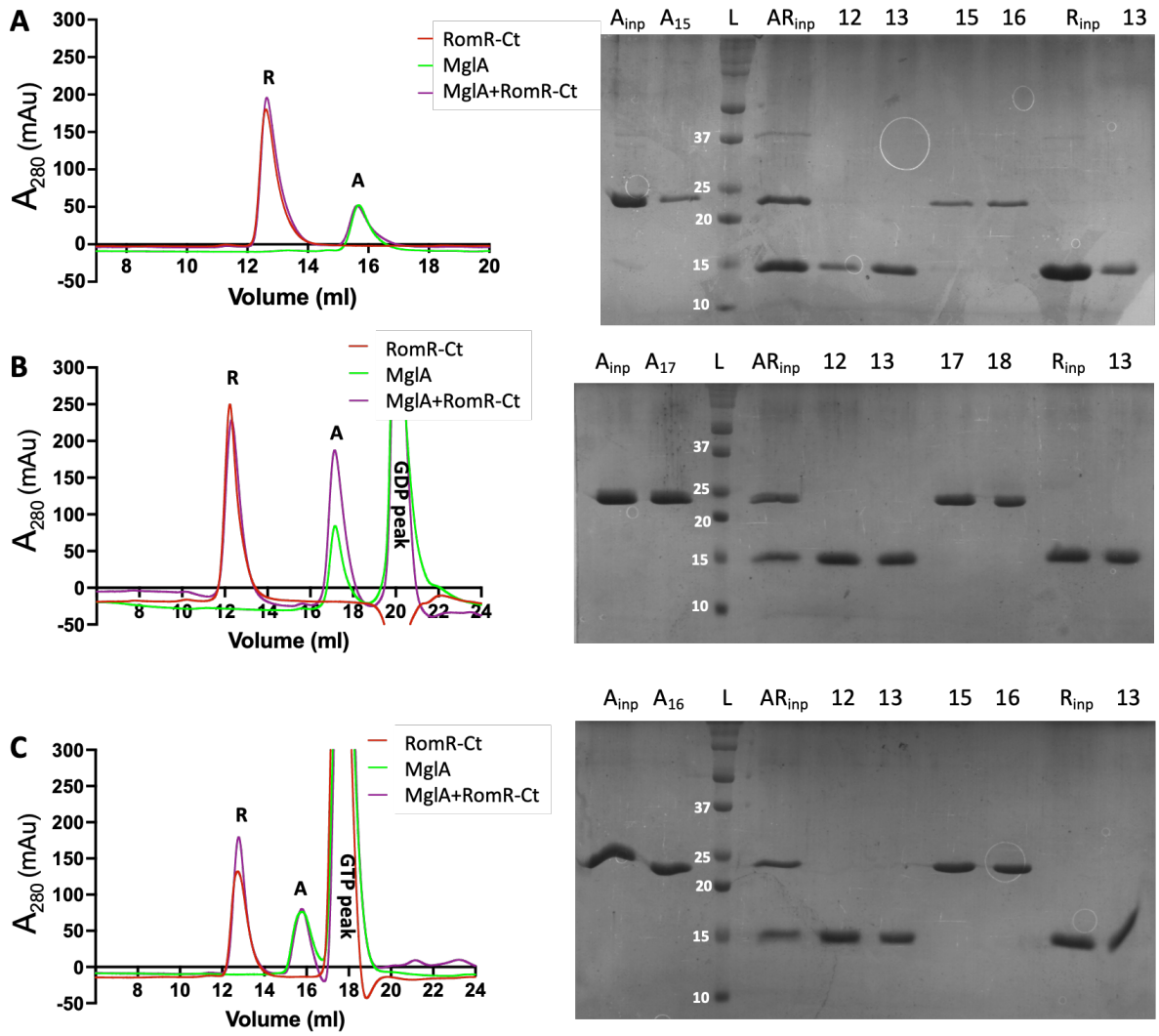


Figure 6.7. Interaction of RomR-Ct with MglA

A, B, C: Analytical Superdex 200 runs showing the interaction of RomR-Ct with MglA without nucleotide (A), with GDP (B) and with GTP (C). The axes are cut to mask the high peaks of GDP and GTP elution at 18 ml, as shown in the figure. The right panels show the corresponding SDS PAGE images of the respective runs. The lane numbers indicate the volumes of elution as per the chromatograms (In: Input, L: Ladder) and the proteins corresponding to the bands are labelled. The corresponding peaks in the chromatograms are also labelled.
MglA (green), RomR-Ct (brick red), and MglA+ RomR-Ct(purple)

6.2.3 Effect of RomR domains on MglA activity

Similar to the results described in Chapter 5, we could not capture the interaction of the RomR constructs with MglA in presence of RomX in analytical SEC studies. However, previous reports show the interaction of RomR and RomX with MglA in the GTP-bound state, quite unlike that of a GEF. We could not reproduce this observation and hence proceeded to check the role of the aforementioned RomR constructs in MglA activation. Firstly, analysed the effect of all these constructs on the GTPase activity of MglA (Methods: 2.6.1). It was observed that RomR-REC^{125} and RomX alone or as a complex did not accelerate the GTP hydrolysis rates by MglA. The result was identical with RomR-REC^{141} with and without RomX (Fig 6.8 A). Further, RomR-Ct did not stimulate the GTPase activity of MglA as observed in the NADH-coupled GTPase assay. The observation was similar for RomR-REC^{141}Ct with or without RomX (Fig 6.8 A). This indeed confirms that there is no binding and activation of MglA in the GTP bound state by RomR-RomX complex.

In Chapter 5, we observed the GEF activity of RomX on MglA. It has been shown that this activity is further stimulated in the presence of RomR. Since we observe the interaction of RomR-REC with RomX, we wanted to see if the presence of RomR-REC constructs stimulates the GEF activity of RomX. We studied the association kinetics, monitoring the exchange of GDP bound to MglA using *mant*-GMPPNP (Methods: 2.6.2). RomX induced a fast nucleotide exchange on MglA, similarly with and without RomR-REC^{125}. RomR-REC^{125} alone did not affect MglA activity. Similarly, neither RomR-Ct nor RomR-REC^{141}Ct drove *mant*-GMPPNP exchange with MglA-GDP (Fig 6.8 C, D). When RomR-Ct and RomR-REC^{141}Ct along with RomX were added to MglA, the rates were similar to RomX alone. This showed that it was RomX alone which drives nucleotide exchange on MglA without any contribution from the RomR REC domain (Fig 6.8 B). RomX forms a complex with RomR-REC and MglA implying that RomX is possibly sandwiched between its two interactors, with no direct interaction between MglA and RomR-REC.

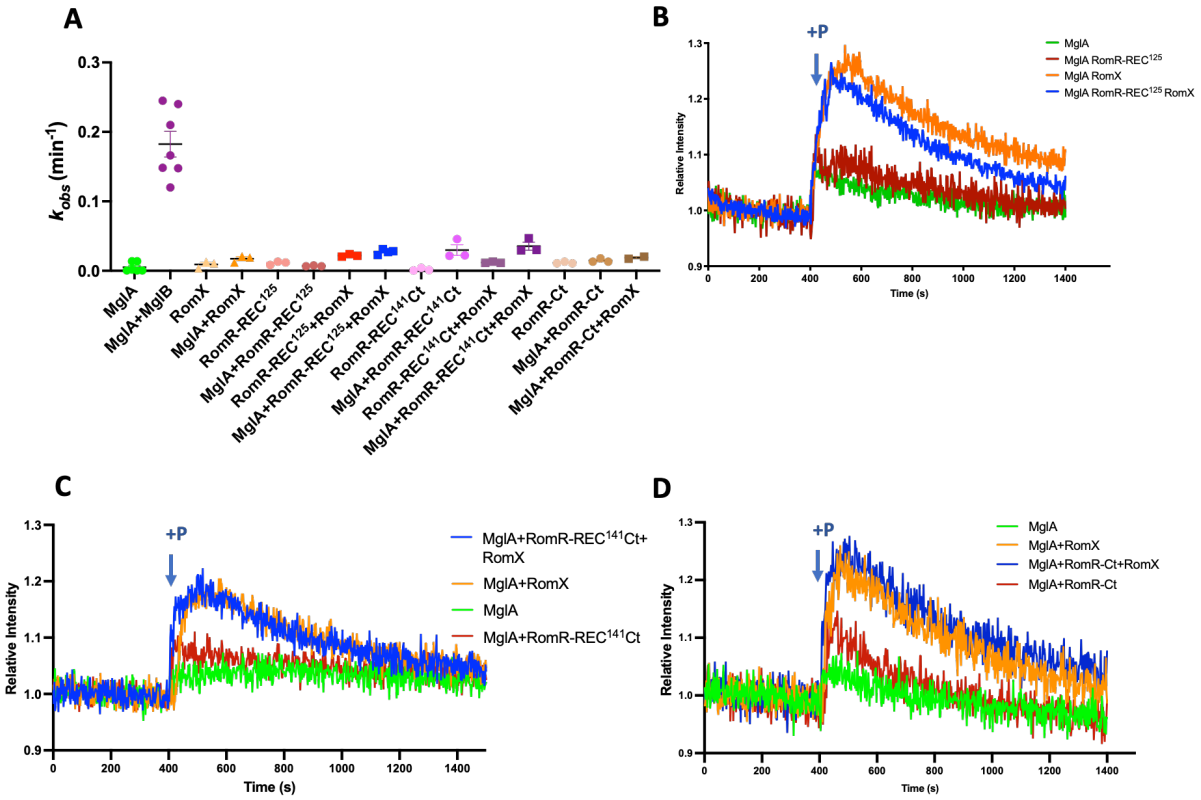


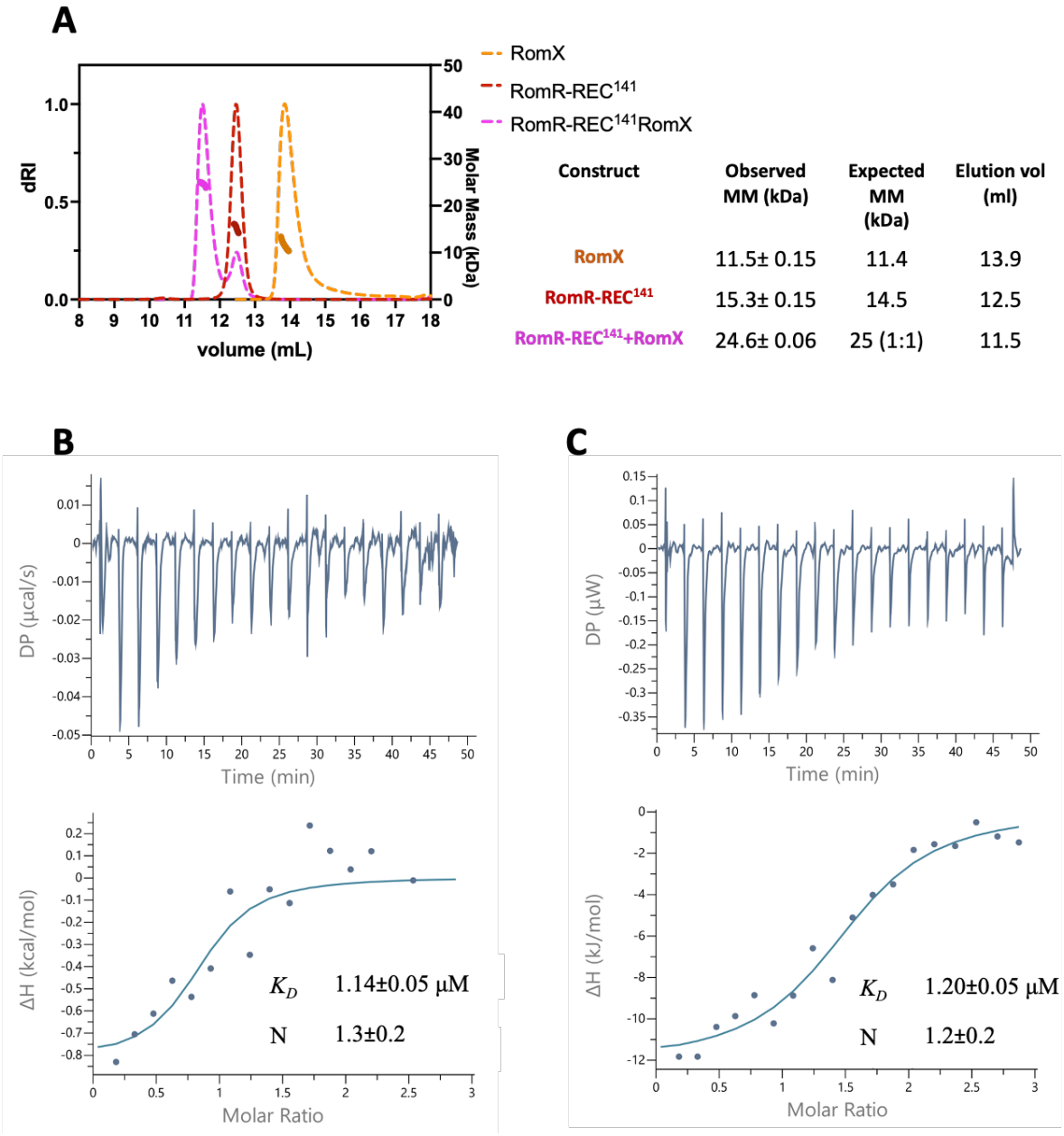
Figure 6.8. Activity assays of RomR constructs with RomX and MglA

A: NADH coupled GTPase assay data showing that none of the RomR constructs with or without RomX shows GAP activity as observed with MglB. The black line represents the mean and the coloured bars represent the standard error.
B, C and D: Representative plots showing *mant*-GMPPNP association assays of MglA with RomR REC constructs, RomR-REC^{141}Ct and RomR-Ct domains, respectively. The trend was similar for all 3 cases. RomX (orange) stimulates release of GDP on MglA and uptake of *mant*-GMPPNP in a 1:2 ratio and the rates are similar with RomX+ RomR (blue) in the same 1:2 ratio (MglA 3 µM : RomRX 6 µM), whereas RomR alone (brick red) does not stimulate exchange. n=3 repeats performed for each MglA, MglA+RomX, MglA+RomR-REC^{125}, MglA+RomR-REC^{125}+RomX, MglA+RomR-REC^{141}Ct, MglA+RomR-REC^{141}Ct+RomX, MglA+RomR-Ct, MglA+RomR-Ct+RomX, runs.

Similarly in the nucleotide exchange assays, neither RomR^{Ct} nor RomR$^{REC\text{-}Ct}$ drove *mant*-GMPPNP with GDP exchange on MglA Fig 6.8 C, D). When the same assays were performed in the additional presence of RomX, the exchange rates were similar to that observed with RomX alone. This showed that RomX alone drove nucleotide exchange on MglA. There were no contributions from RomR REC or C-terminal domains on nucleotide exchange on MglA.

6.2.4 Interaction between RomR-REC and RomX

To quantify the interaction between RomR REC^{141} and RomX and to determine the stoichiometry of the complex, we performed size exclusion chromatography coupled with multi-angle light scattering to determine the molar mass of the complex (Fig 6.9 A). In Superdex 75, there was a shift in the peak of the complex from 12.5 ml for RomR REC and 14 ml for RomX to 11.5 ml. The observed molar mass was 24.6 kDa which corresponded to a 1:1 complex of RomR REC (14.5 kDa) and RomX (11.4 kDa).



Construct	Observed MM (kDa)	Expected MM (kDa)	Elution vol (ml)
RomX	11.5± 0.15	11.4	13.9
RomR-REC^{141}	15.3± 0.15	14.5	12.5
RomR-REC^{141}+RomX	24.6± 0.06	25 (1:1)	11.5

Figure 6.9. Interaction between RomR-REC and RomX

A: Superdex 75 coupled with multi-angle light scattering to show the interaction of RomR-REC and RomX. The molar mass of the corresponding peaks is plotted on the right y-axis. RomX is in orange, RomR-REC^{141} is in red and complex is in magenta. The observed and expected molar mass and the respective elution volumes are tabulated in right.
C, D. ITC thermograms showing the titration of RomX with RomR-REC^{125} and RomR-REC^{141}, respectively. The top panels show the raw data of exothermic heat pulses with time and the lower panels show the corresponding differential binding curve fitted to a single site binding model.

Further to quantify the affinity between RomR-REC and RomX, isothermal titration calorimetry was performed (Fig 6.9 B, Methods: 2.7.2). The data was plotted to a one-site binding model using the MicroCal PEAQ ITC Analysis software with a fitted offset control. The same titration experiment was performed with RomR REC^{141} and the results were similar (Fig 6.9 C). We obtained an average K_D of 1.14 ± 0.05 μM for RomR REC^{125} and RomX and 1.2 ± 0.5 μM for RomR REC^{141}. The stoichiometry of interaction N was similar for both RomR REC^{125} and RomR REC^{141}, 1.3 ± 0.2 and 1.2 ± 0.2, respectively. This confirms that RomR REC interacts with RomX in a 1:1 ratio.

To characterize the interface between the RomR-REC domain and RomX, an AlphaFold model of the complex was generated using one chain each of RomX and RomR (1-125 residues) (Fig 6.10 A). The prediction of the RomR-REC domain was expected to be reliable as the REC domain is prevalent and conserved. RomX model superposed well with the experimentally obtained structure with an RMSD of 0.8 Å. Looking closer into the interface, it was observed that Glu-108 of RomR formed a salt bridge interaction with Arg-76 of RomX. Further, the Val-17 or RomR and Leu-75 of RomX formed a hydrophobic core into which Trp-87 of RomX also tucks in when in complex with RomR-REC. This tryptophan is otherwise exposed in solution in the monomeric structure of RomX.

Since Trp-87 is the sole tryptophan residue in the RomRX complex which shows a conformational transition across the monomeric RomX and RomRX complex, it could be used as a tool to monitor RomRX complex formation. Since this was a hypothesis based on the AlphaFold prediction of the structure of the complex, it was validated by monitoring the fluorescence of the single Trp-87 in RomX. If the Trp-87 is exposed in monomeric RomX and is buried in the RomRX complex interface, it would show a shift in the wavelength and intensity of the Trp emission spectrum maxima. To investigate this, 10 μM of RomX is titrated with increasing concentration of RomR-REC (Fig 6.5 E, Methods: 2.7.4). It was observed that

there was indeed a decrease in the Trp fluorescence intensity and a gradual blue shift in the emission spectra maxima from 360 nm to around 330 nm (Fig 6.10 G, H). This indicated a change in the environment of the Trp as the RomRX complex formed. This validates that Trp 87 is one of the residues which forms the RomRX complex interface.

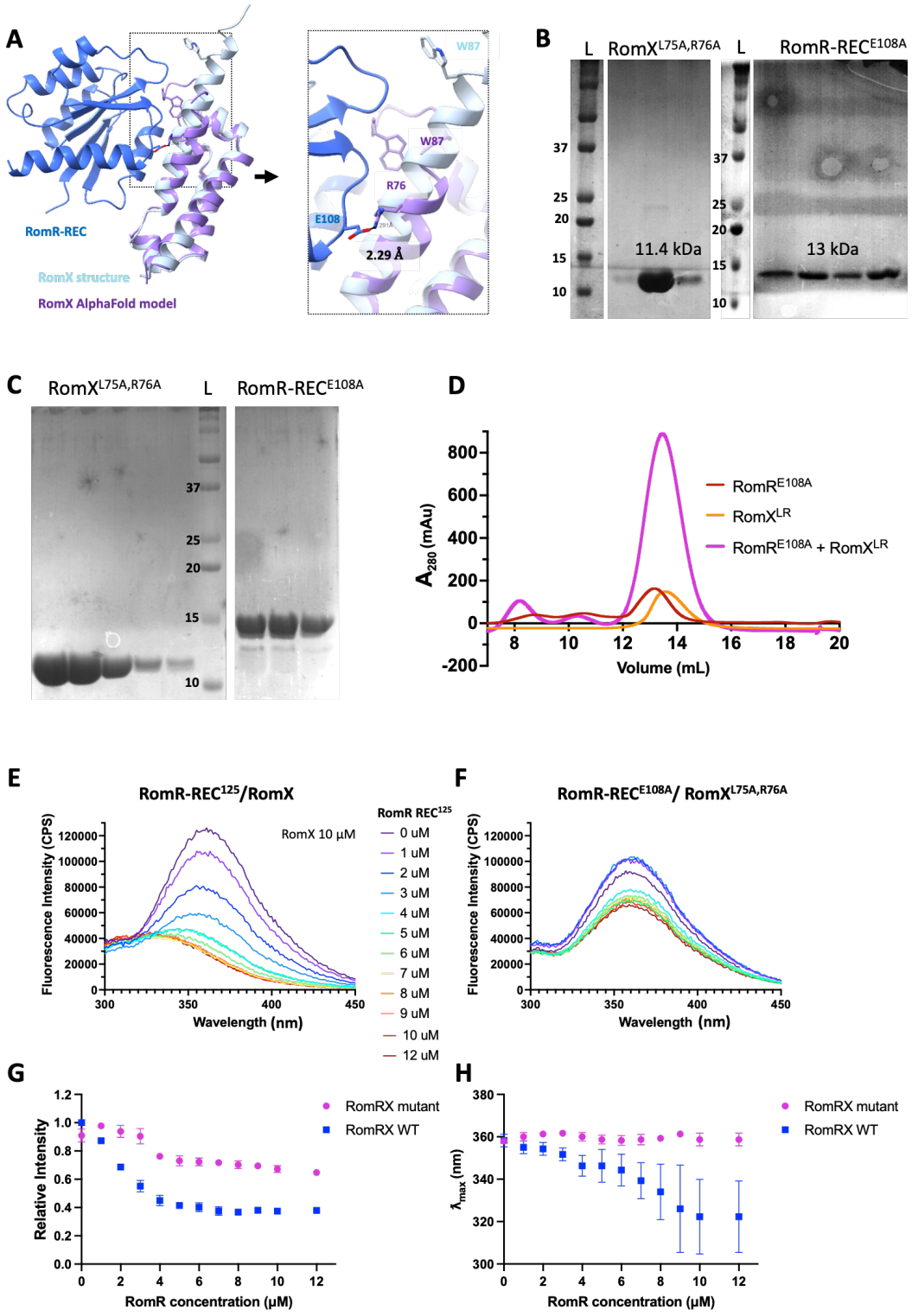


Figure 6.10. Interface between RomR-REC and RomX

A: AlphaFold model of RomR-REC (blue) and RomX complex. Experimentally obtained RomX structure in purple is superposed on the model in light blue. In right, the interface of the complex is expanded showing the probable interface residues, W87 and R76 in RomX and E108 in RomR. The interacting distance between RomR E108 and RomX R76 is highlighted. The conformational change of RomX W87 across the monomeric (purple) and complex states (light blue) is also shown
B, C. SDS PAGE showing the purified proteins of $RomX^{LR}$ and RomR-REC^{E108A} mutants using Ni-NTA and Superdex 75, respectively.
D. Analytical Superdex 75 of $RomX^{LR}$(orange) and RomR-REC^{E108A} (brick red) mutants. The co-injection of the mutants (magenta), shows no shift to a lower elution volume, as observed in WT proteins (Fig 2 A).
E, F. Tryptophan fluorescence spectra to determine the complex formation RomR-REC and RomX in WT and mutants (left and right, respectively). RomX is titrated with increasing concentrations of RomR-REC as shown in the legend.
G, H. Relative fluorescence intensity at emission maxima λ_{max} and the corresponding λ_{max} values are plotted against concentration of RomR-REC to determine the shift in fluorescence spectra as well as quenching due to complex formation, due to change in the environment of the Trp87 at the RomRX interface. The mean of 3 replicates is plotted with the error bars indicating standard errors.

Based on our information regarding the interface, we designed interface mutants of RomR and RomX, namely RomR-REC^{E108A} and $RomX^{L75A,R76A}$ (L75, R76 both mutated to A) (Fig 6.10 B, C). These were purified using Ni-NTA followed by Superdex 75. There was no shift in the peak of the resultant complex of RomRX as observed in the case of RomR-REC^{125}+RomX (Fig 6.10 D, Fig 6.3 A). However, since the elution volumes were very close, the individual peaks could not be resolved.

Further, Trp fluorescence measurements were performed using the interface mutants of RomR-REC and RomX (Fig 6.10 B, C). $RomX^{L75A,R76A}$ mutant was titrated with increasing concentrations of RomR-REC^{E108A}. It was observed that there was no blue shift observed as in the case of wild-type constructs of RomR-REC^{125} and RomX (Fig 6.10 F). There was a slight decrease in the intensity of the emission maxima at higher concentrations of RomR-REC (Fig 6.10 G, H). This was probably due to the crowding of RomR-REC molecules which minorly masked the Trp-87 fluorescence signal. This confirms that RomR-REC Glu-108, RomX Arg-76 and Trp-87 are the key interface residues holding the RomRX complex.

6.2.5 Characterizing the C-terminal output domain

From the previous section, we determine that the N-terminal residues 1-125 of RomR form an independently folded REC domain which interacts with RomX. The C-terminal output domain is characterized by a long disordered region consisting of repeat sequences enriched in proline

and glycine residues. Residues 368-420 were predicted to be alpha-helical. As mentioned in the previous section, RomR C-terminal domain (RomR-Ct), was purified and studied. It was also subjected to CD spectroscopy in 20 mM phosphate buffer, similar conditions as mentioned for RomX (Methods: 2.5.2). The spectrum obtained had peaks at around 221 nm and 206 nm. When analyzed using BeStSel, it showed a 100% propensity for alpha helices which confirmed that RomR^{Ct} was alpha-helical (Fig 6.11 A). According to AlphaFold prediction, it is composed of a long alpha helix, which was bent in the middle at Val-398, Val-399 positions (Fig 6.11 B).

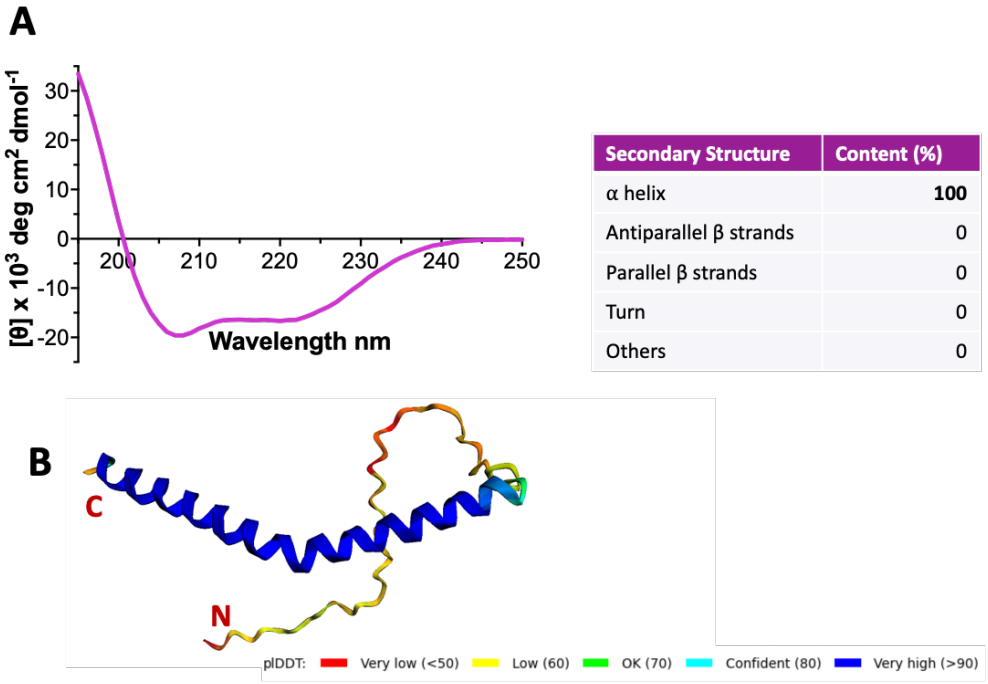


Figure 6.11. C-terminal tail of RomR
A: CD spectrum of the RomR-331Ct protein was fed on to BeStSel server to obtain the percentage of estimated secondary structure content. The fitting shows a 100% alpha-helical content.
B. AlphaFold prediction of the RomR-Ct colored according to pLDDT scores with N and C termini labeled.

The proline-rich linker region of RomR (117-368) was analysed for the 10 other RomR sequences presented by Leonardy and Søgaard Andersen, 2007[2]. We observed that there is little to no sequence conservation in the low complexity region across different RomR homologs. However, there could exist a conservation of properties. Hydrophobicity and molecular size plots did not show much conservation of pattern across the sequences. We further went ahead and exploited CIDER[12], a web server for analysing intrinsically disordered regions (IDRs) in proteins (with Ms. Akshata Kotwal). When fed with the IDR sequences, the server provides a set of parameters used to predict the kinds of conformations the IDRs might

adopt in solution. The parameters include $\varkappa$ (extent of charged amino acid mixing), Ω (pattern of charged + proline residues), hydropathy, FCR (fraction charged residues) etc. All this information is used to generate a diagram of states where FCR and IDR conformations are correlated. The following regions of the plot indicate the different conformations the IDRs might adopt[13].

1. **Weak polyampholytes & polyelectrolytes**: These sequences form globules.
2. **Janus sequences**: Globules or coils depending on other factors like (salt concentration, ligand binding, cis-interactions etc.)
3. **Strong polyampholytes**: Coils, hairpins and chimaeras (depends on κ)
4. **Negatively charged strong polyelectrolytes**: Polyanionic repulsion-driven rods or coils
5. **Positively charged strong polyelectrolytes**: Polycationic repulsion-driven rods or coils

Hence, the presence of clusters in the diagram of states indicates some conserved properties rendered by the presence of these IDRs in RomR-like proteins. We identified more RomR-like proteins from the sequence databases (Fig 6.12 A). The C-terminal region of RomR (Q1D3Z1, 369-420) was used for a BLAST search. Amongst the three domains of RomR, expectedly, the N-terminal (1-116) and C-terminal (369-420) were found to be conserved. Given the ubiquity of the response receiver domain, we chose the C-terminal domain for our search. We also exploited the overlapping genomic distribution of RomR and RomX and filtered the RomR sequences which co-occurred with RomX. As a response regulatory domain is typically more than 100 amino acids long, and proteins smaller than 100 amino acids were discarded. The 136 proteins obtained from above were used for a domain search using SMART [14]. 123 proteins had a receiver domain and a low complexity region. We compared the conservation of the C-terminal domain of proteins in this dataset with the 31 proteins obtained from Keilberg D and Søgaard-Andersen L (2014) [4] and observed a conserved WEVV motif at the C-terminal (Fig 6.12 B). Finally, we obtained a new data set of 123 RomRs with two characteristic features, a. Response regulatory domain of nearly 120 amino acids followed by a disordered linker and b. a C-terminal domain with a WEVV motif.

When the diagram of states was plotted for this dataset, all the proteins were found clustered in either of the two particular regions in the plot (Fig 6.12 C). Most of the IDRs including the one in RomR formed globules (FCR≤0.25). A few of them extended to the regions predicted

for coils (0.25≤FCR≤0.35). The clustering of these proteins in the plot suggests a common behaviour of IDRs of RomR-like proteins in solution. IDRs with similar physical properties are present in the N-terminal domain of α-synuclein (1-100 residues), which is intrinsically disordered but forms amphipathic helices on membrane binding[15] This indicates that the IDR of RomR might play a role in stabilising the interaction of RomR, specifically to the poles recruiting it with either the polar membrane or other interacting proteins at the cell poles. This is in line with the *in-vivo* observations, where we see the role of the output domain with the IDR for RomR polar recruitment. RomR is known to form heteromeric complexes like RomR-RomX-MglA or RomR-MglC-MglB which could be stabilised by additional interactions from the RomR-IDR, which gets structured upon complex formation. This feature of RomR IDRs could be validated experimentally in the future by comparing RomR-REC^{Ct} construct with wild-type RomR.

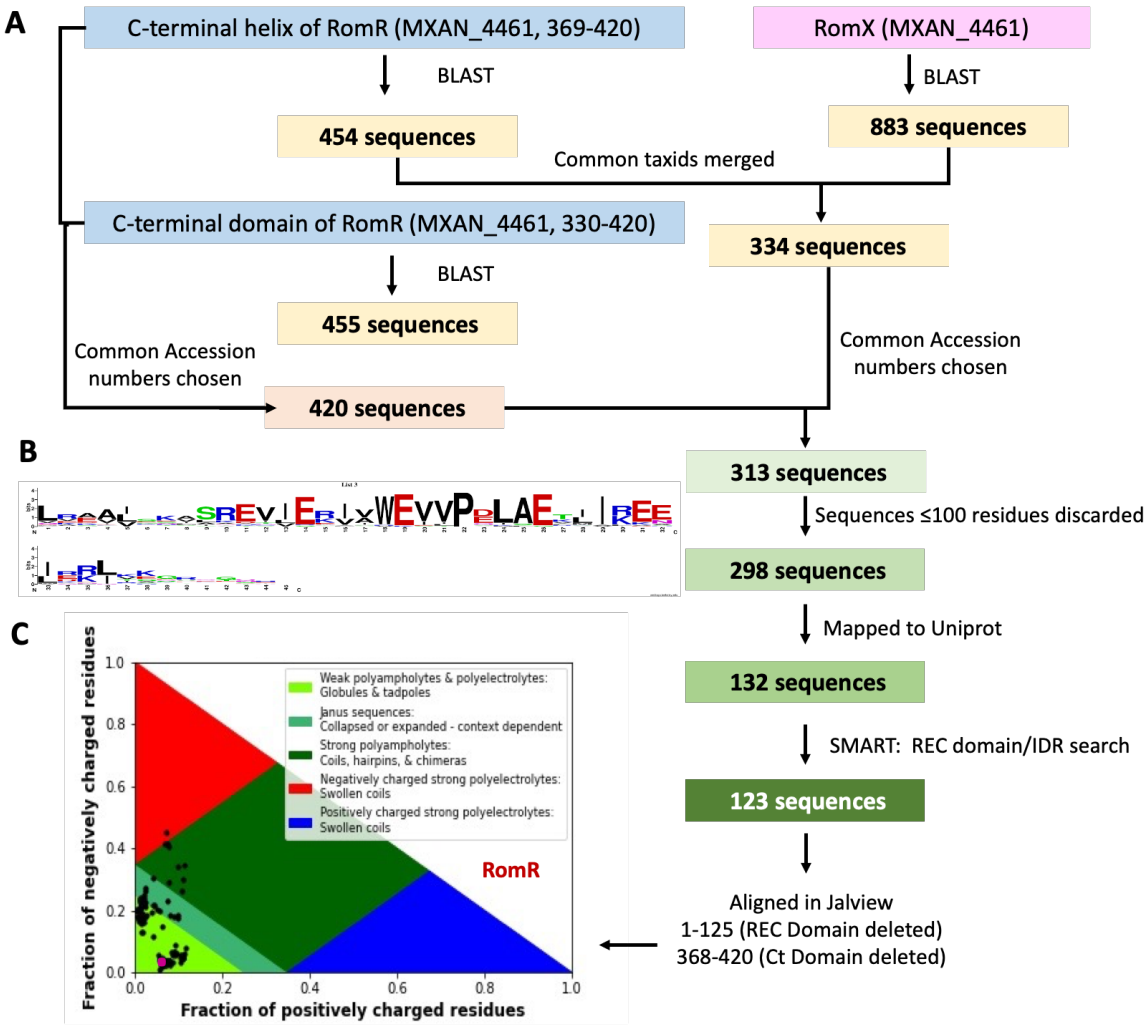


Figure 6.12. Analyzing the disordered linker of RomR

A: Steps used to generate the dataset of 123 RomR-like proteins. The final output is used to plot C.
B. Web logos for the conserved C-terminal domain of the RomR-like proteins. The WEVVP was chosen as the characteristic feature for identifying a RomR-like protein.
C. Diagram of states for the IDRs of 123 RomR-like proteins identified showing the fraction of positively and negatively charged residues in the disordered linker indicating the formation of globules and coils depending on the environment. RomR is highlighted in red.

6.3 Conclusion

RomR was shown to be one of the key players of the polarity module by receiving signals from the chemosensory module and initiating the reversal switch. It was proven to show slower dynamics than MglA and MglB, thereby regulating the time and frequency of polarity reversals[8]. We dissected the protein domains of RomR and observed that the REC domain of RomR interacts with RomX strongly in a 1:1 ratio. Salt bridge interaction between E108 of RomR and R76 of RomX was predicted based on the AlphaFold structure and experimentally validated for being critical to hold the complex together. The C-terminal Trp of RomX tucks into the interface when it binds to RomR. However, no direct interaction was observed between the RomR-RomX complex and MglA in analytical SEC. This is probably because the interaction is transient and has a low affinity. Due to the dilution effect of the SEC column, the interaction was probably not captured. We further generated an AlphaFold model of the RomRX complex with MglA (Fig 6.13). In line with our proposal, RomX is indeed sandwiched between RomR and MglA. MglA has been captured in its GDP-bound conformation, as predicted with RomX (Fig 5.8). The interface of RomX and MglA is not extensive, but probably nucleotide dependant as it sits very close to the nucleotide-binding pocket of MglA. The narrow interface explains the weak and transient interaction of RomR-RomX with MglA which we failed to capture through our experiments. According to this model, we do not capture any direct interaction between RomR and MglA, explaining why we do not see any effect of RomR domains in accelerating the GEF activity of RomX. It will be interesting to further investigate the residues driving the direct GEF activity of RomX at the MglA nucleotide-binding pocket and compare it with the allosteric GEF activity observed with MglB. Further, additional interactions from the IDR of RomR and its role in stabilization need to be studied in detail.

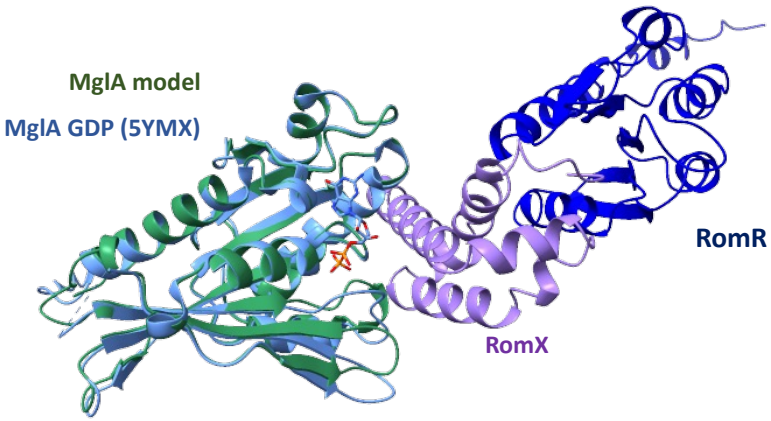


Figure 6.13. AlphaFold model of MglA, RomX and RomR-REC
RomX is seen to interact with MglA in the GDP state. It acts near the nucleotide-binding pocket of MglA. The RomX MglA interface is not extensive, unlike the RomX RomR-REC interface. It is sandwiched between MglA and RomR.

The C terminal domain of RomR does not interact with RomX. There was no direct interaction with MglA too that was observed. This extends the probability for an interaction of the C-terminal domain of RomR with other regulators like MglB or MglC in addition to the interaction of the REC domain with RomX. The long intrinsically disordered region might act as a bridge between these two domains. Hence RomR seems to be a key player in the polarity module where the multiple domains of the protein interface with different regulatory partners which in turn regulates MglA driving polarity reversals and motility. In the next chapter, we have investigated the interaction of MglB and RomR with additional interacting proteins to generate a complete picture of the polarity reversal event.

References

1. Szadkowski D, Harms A, Carreira LAM, Wigbers M, Potapova A, Wuichet K, Keilberg D, Gerland U, Søgaard-Andersen L. Spatial control of the GTPase MglA by localized RomR-RomX GEF and MglB GAP activities enables *Myxococcus xanthus* motility. *Nat Microbiol*. 2019 Aug;4(8):1344-1355. doi: 10.1038/s41564-019-0451-4.
2. Leonardy, S., Freymark, G., Hebener, S., Ellehauge, E. & Søgaard-Andersen, L. Coupling of protein localization and cell movements by a dynamically localized response regulator in *Myxococcus xanthus*. *EMBO J* **26**, 4433–4444 (2007).
3. Zhang, Y., Guzzo, M., Ducret, A., Li, Y.-Z. & Mignot, T. A Dynamic Response Regulator Protein Modulates G-Protein–Dependent Polarity in the Bacterium *Myxococcus xanthus*. *PLoS Genet* **8**, e1002872 (2012).
4. Leonardy, S., Bulyha, I. & Søgaard-Andersen, L. Reversing cells and oscillating motility proteins. *Mol Biosyst* **4**, 1009 (2008).

5. Mignot, T., Merlie, J. P. & Zusman, D. R. Regulated pole-to-pole oscillations of a bacterial gliding motility protein. *Science (1979)* **310**, 855–857 (2005).
6. Stock AM, Robinson VL, Goudreau PN. Two-component signal transduction. *Annu Rev Biochem.* 2000;69:183-215. doi: 10.1146/annurev.biochem.69.1.183.
7. Leonardy S, Miertzschke M, Bulyha I, Sperling E, Wittinghofer A, Søgaard-Andersen L. Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. *EMBO J* **9**, 2276–2289 (2010).
8. Guzzo M, Murray SM, Martineau E, Lhospice S, Baronian G, My L, Zhang Y, Espinosa L, Vincentelli R, Bratton BP, Shaevitz JW, Molle V, Howard M, Mignot T. A gated relaxation oscillator mediated by FrzX controls morphogenetic movements in *Myxococcus xanthus*. *Nat Microbiol*. 2018 Aug;3(8):948-959. doi: 10.1038/s41564-018-0203-x.
9. McLoon AL, Wuichet K, Häsler M, Keilberg D, Szadkowski D, Søgaard-Andersen L. MglC, a Paralog of *Myxococcus xanthus* GTPase-Activating Protein MglB, Plays a Divergent Role in Motility Regulation. *J Bacteriol*. 2015 Nov 16;198(3):510-20. doi: 10.1128/JB.00548-15.
10. Kapoor, S., Kodesia, A., Kalidas, N., Ashish & Thakur, K. G. Structural characterization of *Myxococcus xanthus* MglC, a component of the polarity control system, and its interactions with its paralog MglB. *JBC* **296**, (2021).
11. Carreira, L. A. M., Szadkowski, D., Müller, F. & Søgaard-Andersen, L. Spatiotemporal regulation of switching front–rear cell polarity. *COCEBI* vol. 76 Preprint at https://doi.org/10.1016/j.ceb.2022.102076 (2022).
12. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G. & Pappu, R. v. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys J* **112**, 16–21 (2017).
13. Bianchi, G., Longhi, S., Grandori, R. & Brocca, S. Relevance of electrostatic charges in compactness, aggregation, and phase separation of intrinsically disordered proteins. *Int. J. Mol. Sci.* vol. 21 1–30 Preprint at https://doi.org/10.3390/ijms21176208 (2020).
14. Schultz, J., Copley, R. R., Doerks, T., Ponting, C. P. & Bork, P. *SMART: a web-based tool for the study of genetically mobile domains*. *Nucleic Acids Res* vol. 28 http://SMART.embl-heidelberg.de (2000).
15. Das, T. & Eliezer, D. Membrane interactions of intrinsically disordered proteins: The example of alpha-synuclein. *Biochim Biophys Acta Proteins Proteom* **1867**, 879 (2019).

Chapter 7: **Preliminary investigation of MglB interactors and their role in *Myxococcus xanthus* polarity reversal**

7.1 Introduction

MglB, one of the key regulators of MglA, localizes in a bipolar asymmetric distribution with the major cluster at the lagging pole. It was observed that there is a dual GAP and GEF activity of MglA contributed by the Roadblock domain and the C-terminal helix respectively [1]. Recently, a couple of proteins, RomY and MglC, have been identified in the polarity module, which interacts with MglB and localize in an MglB-dependent fashion [2,3]. They are prospective regulators of MglB which in turn affect the localization of MglA-GTP at the leading pole, thereby establishing a front-rear polarity axis.

RomY was recently shown to interact with MglB and localizes at the lagging pole and switch during polarity reversals [2]. It was proposed to form a low-affinity complex with MglB. This was concluded based on the localization of RomY at the lagging pole where the MglB concentration is higher[2,4]. RomY was observed to interact with MglB in presence of MglA-GTP to form a heteromeric complex. RomY was shown to stimulate the GAP activity of MglB, thereby excluding MglA-GTP from the lagging cell pole. At this pole, RomY probably causes the MglB GAP activity to dominate over the RomR-RomX GEF activity, which is necessary for the correct localization of MglA. However, there is no structural information on RomY and how it forms a complex with MglB along with MglA to activate the GAP activity. Knowledge of the stoichiometry of the MglA/MglB/RomY complex, binding affinities and the binding interfaces would help understand the mechanism of a dominant GAP at the lagging pole which is critical for the maintenance of front-rear polarity axis of *M. xanthus*.

Further, another component of the polarity module, namely, MglC has also been identified as an MglB interactor [5,6]. MglC was shown to be important for the polar localization of MglA and MglB. MglC is another protein with the Roadblock domain fold and interacts with MglB to form an extended Roadblock domain platform, similar to several eukaryotic small Ras-like GTPase regulators[7]. SAXS studies show a 2:4 complex between MglC and MglB. RomR has been shown to drive MglC localization in the cell in a similar bipolar asymmetric fashion, with a higher concentration at the lagging pole. But there is no information on the exact mechanism of MglC recruitment to the poles. Moreover, as of yet, there is no information on the function of MglC in the polarity module. Hence it is important to investigate the mechanism of polar recruitment of MglC. The interaction of MglC with MglB and RomR needs to be further dissected to understand the exact order of events leading to the similar localization of all these proteins in a polarized cell. MglB interacts with MglC via its Rbl/LC7 domain, which is also

the MglB GAP active domain. Therefore, MglC could also probably play a role in regulating MglB GAP activity at the leading and lagging poles in turn regulating MglA localization and function.

7.2 Results

7.2.1 Fold Identification of RomY

Initially, the annotated sequence for RomY was MXAN_5749 as shown in Fig 7.1 A. Later it was reannotated to MXAN_RS27865 (residues 71-258 as per sequence in Fig 7.1 A). We performed a secondary structure prediction using PsiPred of the whole frame as per MXAN_5749 and also fed the sequence to SwissModel. We obtained a model similar to dihydro neopterin aldolase (PDB ID: 3V9O), for residues ranging from 73-163 of RomY. The sequence alignment between the aldolase and RomY model is shown in Fig 7.1 B.

Further, we compared the topology of aldolase with longin (ββαβββαα) and Roadblock (αββαβββ) folds. Aldolase has a ββααββ topology, which is a circularly permuted incomplete longin domain fold (Fig 7.2 A, B). Consequently, we did a structure-based sequence alignment using PROMALS3D where we observed that the MXAN_5749 gene was predicted to have a longin domain fold whereas the reannotated MXAN_RS27865 had an incomplete longin domain fold (Fig 7.2 C). This was interesting as multiple longin/Roadblock domains can interact to form multimers and such a scaffold is known to be prevalent in the regulation of small GTPases[7].

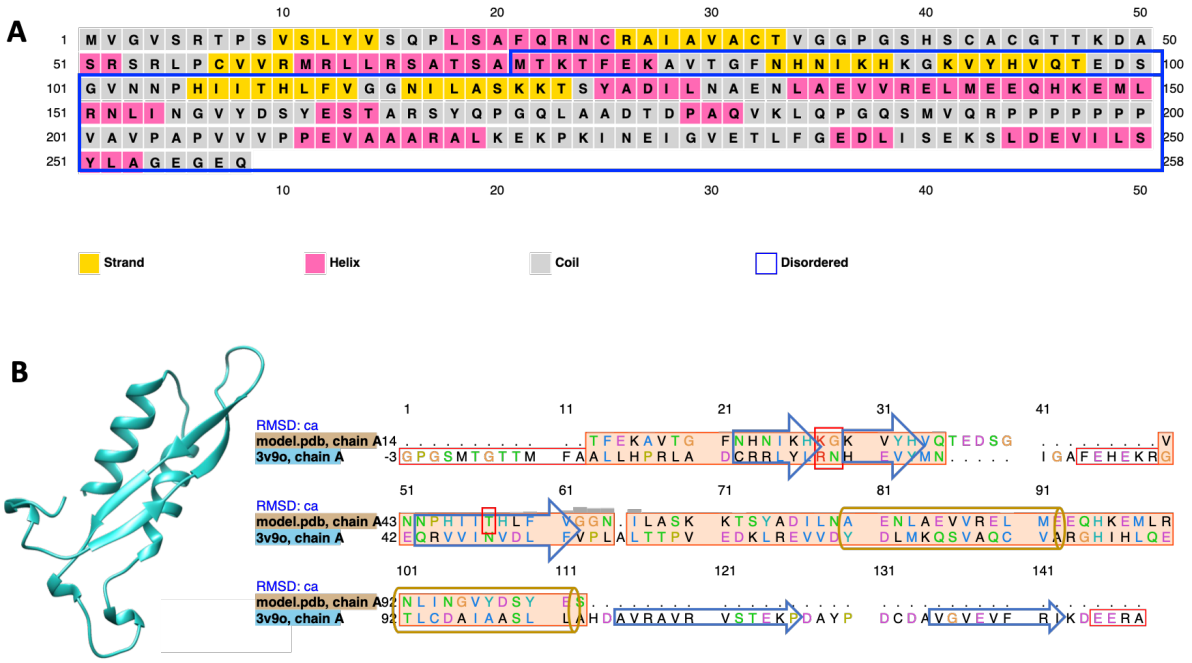


Figure 7.1. Fold prediction of RomY
A. PSIPRED result showing the secondary structure prediction of RomY as annotated in MXAN_5749. The pink highlighted residues form alpha helices, yellow form β strands and the grey residues are loops connecting the secondary structures. The sequence corresponds to MXAN_5749 and the blue highlighted sequence is of reannotated MXAN_RS27865.
B. SWISS-MODEL prediction using the RomY sequence as in A. It shows a hit with dihydro neopterin aldolase (PDB ID 3V9O) and the structure of the model is shown on left. In right, is the structure-based sequence alignment of the RomY model with aldolase (using match-align in UCSF Chimera) with secondary structures highlighted. The blue arrows indicate β strands and the yellow cylinders denote alpha helices.

We performed a BLAST search against PDB with the RomY sequence and did not obtain any significant hits. Hence it was proposed to be a novel protein with no similarity with experimentally solved structures. Hence, we attempted to generate an AlphaFold model of the reannotated MXAN_RS27865. Earlier reports show that the N-terminal region of RomY (1-89 residues) forms a folded domain and the C-terminal region is disordered[2]. This N-terminal domain ($RomY^{\Delta C}$) was predicted to be the functional domain of RomY, interacting with MglA and MglB. Hence, we modelled RomY from residues 1-89 (sequence from MXAN_RS27865) and inspected the arrangement of secondary structures. The model was composed of 4 consecutive β-strands followed by an α-helix. This was a different fold (ββββα) as compared to the aldolase fold (ββααββ) (Fig 7.2 C).

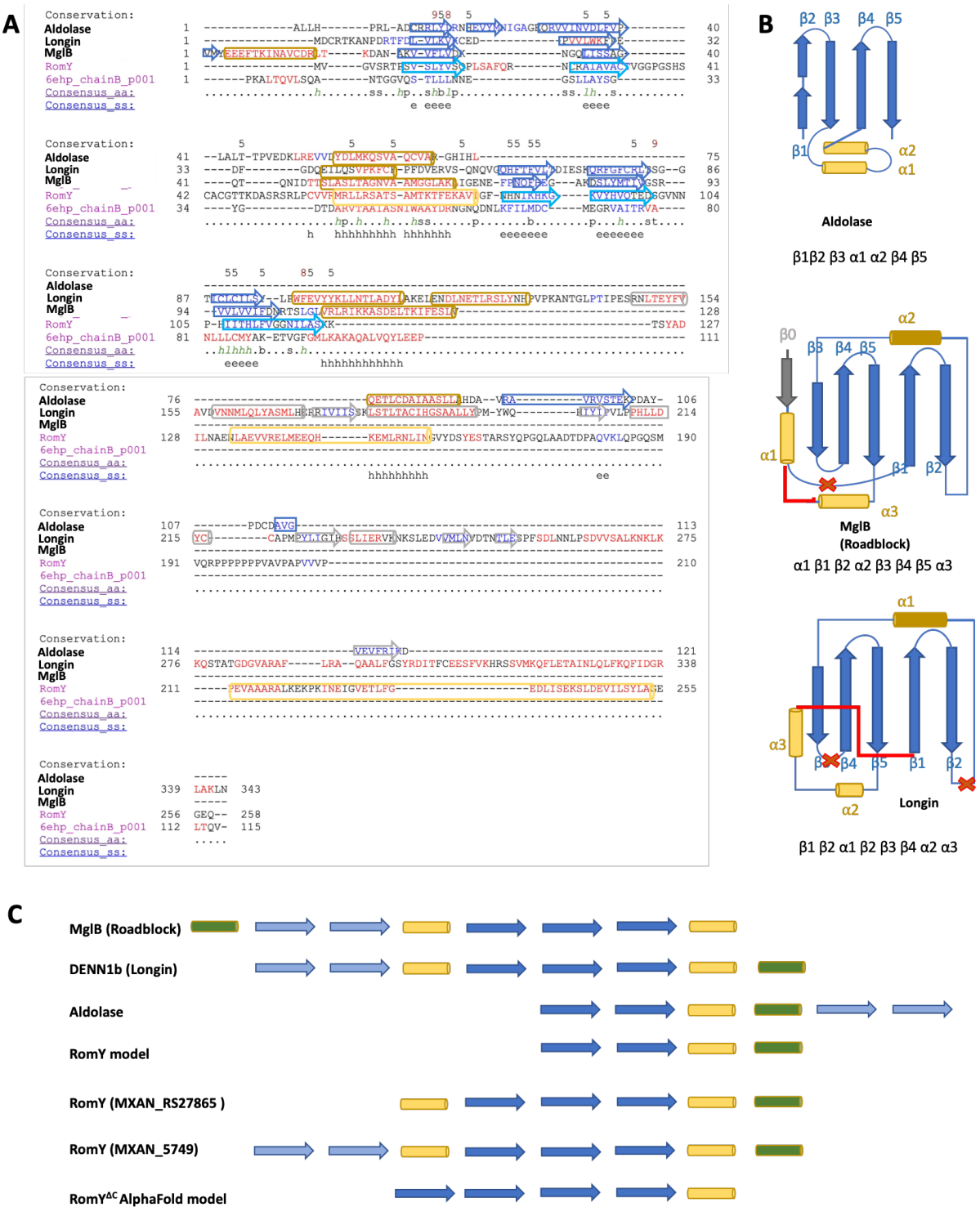


Figure 7.2. Comparison of RomY fold with Roadblock and longin domains

A. Structure-based sequence alignment of RomY using PROMALS3D analysis, where the sequence of RomY is aligned with DENN1b (a longin domain protein) and MglB (a Roadblock domain protein) and aldolase. The structures highlighted where the blue arrows indicate β strands and the yellow cylinders denote alpha helices.

B. Topology diagrams for aldolase, Roadblock and longin domains, respectively, are shown. The secondary structures are labelled and numbered in the order of their occurrence in the protein sequence (N-terminal to C-terminal). The arrangement of secondary structure elements for each fold is described below the topology diagram.
C. The arrangement of secondary structures in DENN1b (a longin domain protein) and MglB (a Roadblock domain protein) and aldolase. This is further aligned with the SWISS-MODEL predicted RomY model and the arrangement of the secondary structures in the 2 respective annotated sequences of RomY (MXAN_5749 and MXAN_RS27865) followed by the AlphaFold model of $RomY^{\Delta C}$.

7.2.2 Expression and Purification of RomY

We obtained the plasmid containing 1-89 residues of RomY (MXAN_RS27865) with an N-terminal hexahistidine tag, very generously shared by Dr Dobromir Szadkowski and Dr Lotte Sogaard Andersen (for details, refer to $RomY^{N}$ construct in Szadkowski *et al.*, 2022[2]). The protein, named $RomY^{\Delta C}$ henceforth, was expressed and purified using the N-terminal hexahistidine tag where it was eluted at an imidazole concentration of 250 mM (Fig 7.3 A) (performed along with Ms Puja Kumari, Methods: 2.4). Further, this protein was subjected to Superdex 75 where it eluted at 11.7 ml (Fig 7.3 A, B). To check the oligomeric status of RomY in solution, we performed MALS to determine the molar mass of the eluted protein (Fig 7.3 C). The peak fraction corresponded to a molar mass of 12.5 ± 0.02 kDa which corresponds to the monomeric mass of 12.6 kDa. This shows that $RomY^{\Delta C}$ is monomeric in the solution.

7.2.3 Structure determination of $RomY^{\Delta C}$

As mentioned earlier, RomY is a novel protein with no sequence similarity with any structurally characterized proteins. Hence, we proceeded to crystallize the protein to solve the structure of the purified $RomY^{\Delta C}$. We subjected the protein to crystallization trials by screening multiple conditions containing varying salt, buffer and precipitants (Methods: 2.8). We used concentrated protein above 10 mg/ml concentration to further optimize the hits. We obtained the final optimized needle-shaped crystals with 15 mg/ml of purified protein in the crystallization buffer, which consisted of 8% PEG 3000, 0.1 M KH_2PO_4+Na_2HPO_4 pH 7 (Fig 7.3 D), at a 1:1 protein:condition ratio. We used different cryoprotectants out of which 20% PEG E 400 gave the best results.

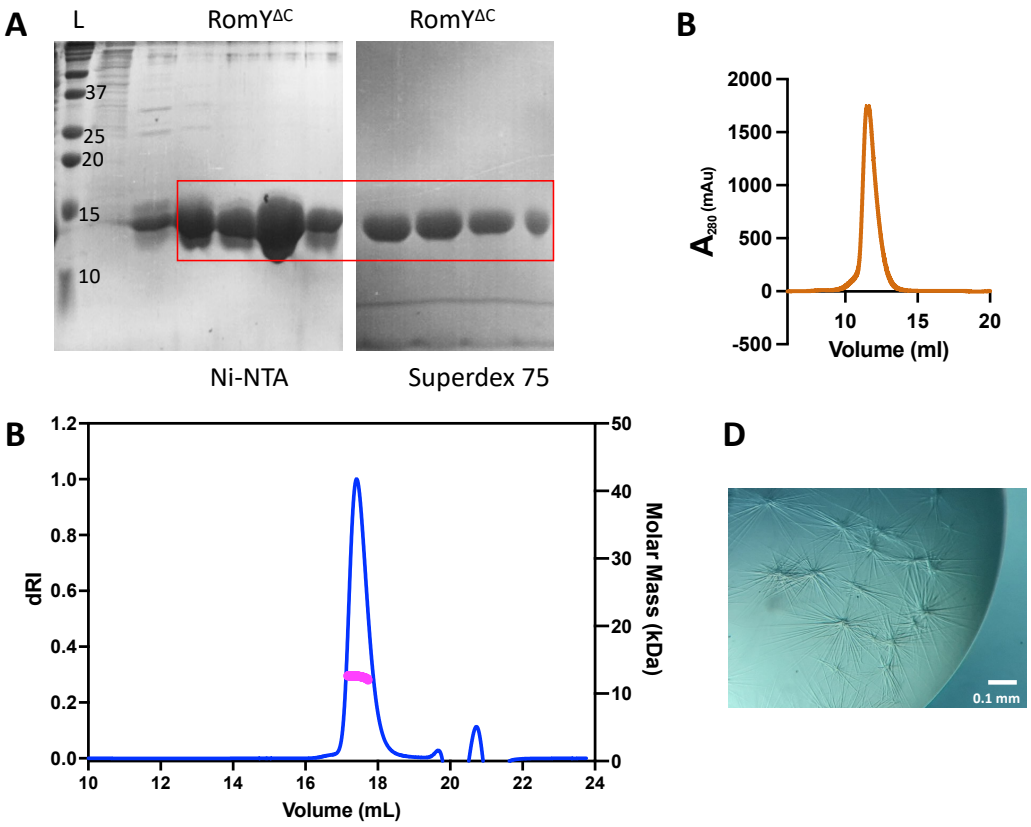


Figure 7.3. Crystallization of RomY$^{\Delta C}$
A. SDS PAGE showing the purified RomY$^{\Delta C}$ protein using His-tag-based affinity purification (Ni-NTA) followed by SEC (Superdex 75). L denotes the ladder with bands labelled in kDa and the purified RomY$^{\Delta C}$ band is highlighted in red.
B. Superdex 75 profile of RomY$^{\Delta C}$.
C. SEC-MALS profile showing the chromatogram (dRI in the left y-axis) for RomY$^{\Delta C}$ in blue and the molar mass of the eluted peak in pink (right y-axis).
D. Needle-shaped crystals of RomY$^{\Delta C}$. The scale bar is shown in the bottom right.

We diffracted the crystals in the aforementioned conditions and obtained diffraction data from ESRF under the automated remote data collection strategy (ID30A-1 / MASSIF-1 beamline). The quality of the collected data is tabulated in Table 7.1. Residues 6-89 of RomY$^{\Delta C}$ was predicted with high confidence by AlphaFold, which was also well correlated with the PSIPRED prediction (Fig 7.4 A). Hence, we used the AlphaFold model generated using the sequence of the purified protein to perform molecular replacement to solve the structure of RomY$^{\Delta C}$ (Fig 7.4 B). We finally refined the structure to a resolution of 1.5 Å to a R_{work}/R_{free} of 0.2080/0.2226. There was a single molecule in the asymmetric unit, with the crystal packing

confirming its monomeric nature. PDBePISA analysis showed there was no propensity of this domain for further oligomerization.

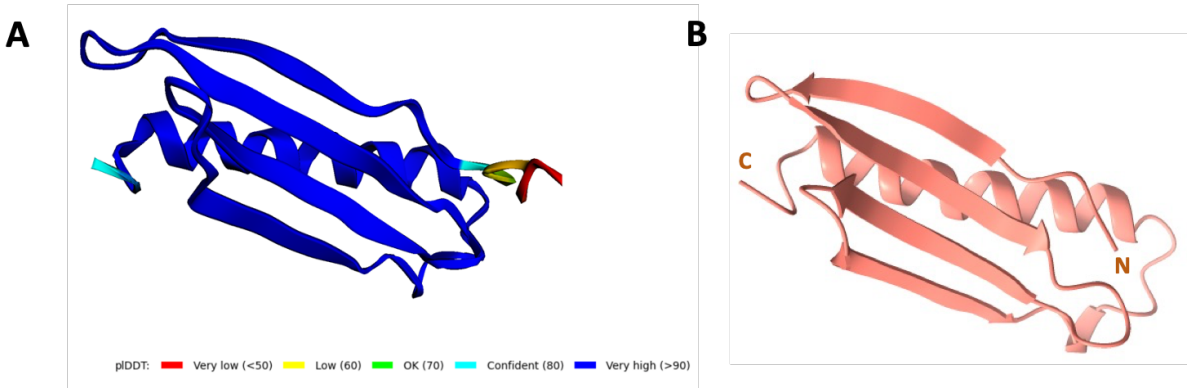


Figure 7.4. Crystal structure of RomY$^{\Delta C}$
A. AlphaFold model of RomY$^{\Delta C}$ coloured according to the predicted lDDT scores.
B. The structure of RomY$^{\Delta C}$ as obtained with the N and C termini labelled. N-terminal forms a β-sheet with 4 strands followed by a α helix.

The structure consisted of 4 consecutive β-strands followed by an α-helix, confirming this to be a fold different from the longin domain (Fig 7.4 B). Gly-11 to Ser-64 formed a continuous stretch of four consecutive antiparallel β-strands, which formed a β-sheet. This is followed by an α-helix which further continues to form a long-disordered region. The C-terminal disordered region was not essential for the localization or GAP activity of RomY[2]. Hence, the RomY$^{\Delta C}$ structure could be used to model a functionally relevant MglA/MglB/RomY complex to determine the mechanism of GAP stimulation by RomY.

Table 7.1. Data Collection and refinement statistics of RomY$^{\Delta C}$

(last resolution shell values are enclosed in parentheses)

Collected at	European Synchrotron Radiation Facility (ESRF)
Wavelength	0.965
Data Collection Statistics	
Space group	$P2_12_12_1$
a,b,c (Å)	32.2, 42.8, 115.5
α, β, γ (°)	90, 90, 90
Resolution (Å)	1.3 (1.3)
No. of unique reflections	140224 (38622)
R_{merge} (%)	6.3 (105.6)
R_{pim} (%)	3.6 (63.7)
CC_{half}	0.998 (0.473)
Mean I/sigmaI	8.4 (1.0)
Completeness (percent)	96.3 (98.9)
Refinement Statistics	
Resolution (angstrom)	57.7-1.46
No. of reflections	26821
R_{work}/R_{free} (%)	0.2080/0.2226
Average B factor (Å^2)	20.7
Wilson B factor (Å^2)	16.77
RMS deviation	
Bond lengths (Å)	0.006
Bond angles (°)	0.856
Ramachandran Map	
Favoured (%)	97.06
Allowed (%)	2.94
Outliers (%)	0

7.2.4 MglB Ct helix is critical for stimulating the co-GAP activity of RomY

Further, we went ahead to check the GTPase stimulation of MglB using the purified RomY$^{\Delta C}$ (Fig 7.5 A, Methods: 2.6.1)). According to Szadkowski *et al*., 2022, RomY$^{\Delta C}$ was as active as

the full-length protein[2]. We indeed observed that RomY$^{\Delta C}$ stimulated the GTPase activity of MglA in presence of MglB to 1.5 times that of MglB alone. To clarify the stoichiometry of interaction of the MglA:MglB:RomY$^{\Delta C}$ complex, we performed the GTPase assay with 2 different ratios of MglA:MglB:RomY$^{\Delta C}$, i.e., 1:2:1 and 1:2:2. We observed similar levels of GAP activity with both ratios, demonstrating that 1:2:1 is probably the functional stoichiometry and is sufficient for stimulating the GAP activity of MglB on MglA (Fig 7.5 B).

Further, we performed the same assay using a 1:2:2 ratio of MglA:MglB$^{\Delta C}$: RomY$^{\Delta C}$, using MglB with a C-terminal truncation, which is a GAP only. We observe that RomY$^{\Delta C}$ fails to stimulate the GAP activity of MglB$^{\Delta C}$. This shows that the C-terminal helix of MglB is important for either RomY interaction with the MglA:MglB complex or the consequent GAP stimulation of MglB on MglA. AlphaFold prediction of the complex also suggests the possibility of the C-terminal helix of $MglB_2$ interacting with RomY.

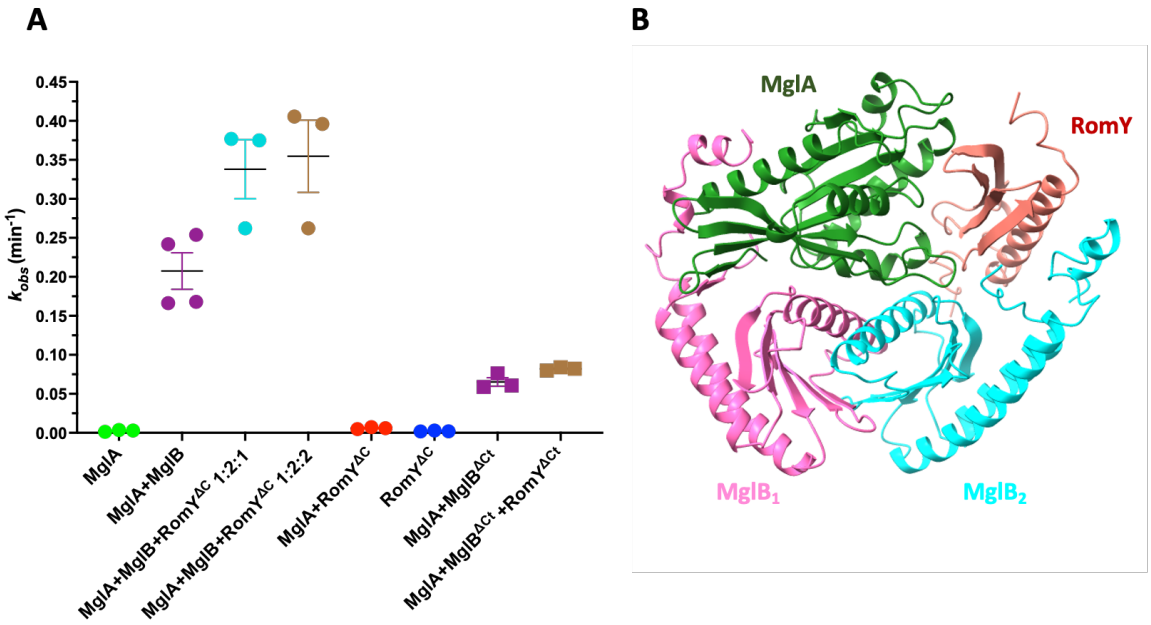


Figure 7.5. Interaction of RomY with MglA MglB complex
A. Comparison of k_{obs} values for MglA (green circle), MglA+MglB 1:2 (purple circle), MglA+MglB+ RomY$^{\Delta C}$ 1:2:1 (cyan circle), 1:2:2 (brown circle), MglA+RomY$^{\Delta C}$ (red circle), RomY$^{\Delta C}$ (blue circle), MglA+ MglB$^{\Delta Ct}$ (purple square), MglA+ MglB$^{\Delta Ct}$ + RomY$^{\Delta C}$ 1:2:2 (brown square). The mean is shown in black lines and the standard error in coloured long and short horizontal lines.
B. AlphaFold model of MglA (green), $MglB_1$ (pink) $MglB_2$ (cyan) and RomY (salmon).

7.2.5 Expression and Purification of MglC

A construct containing *pET-Duet-A-TEV-mglC*, as reported in Kapoor *et al*., 2021, was shared generously by Dr Krishan Gopal from IMTech Chandigarh, to perform interaction studies with MglA, MglB and RomR. We expressed the protein using the same protocol as in Kapoor *et al.*, 2021[5] and purified the protein using HisTrap following the size exclusion chromatography using Superdex 75 (Fig 7.6 A, Methods: 2.4). The protein was purified and eluted at 11.5 ml.

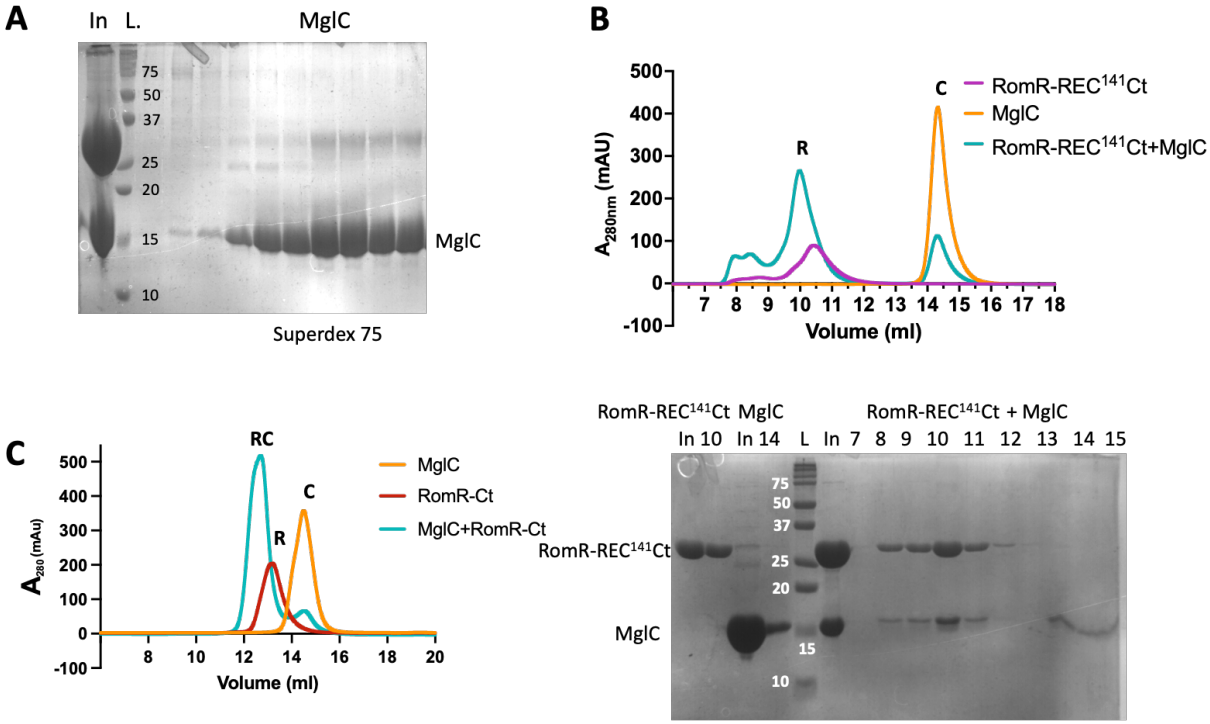


Figure 7.6. Interaction of MglC with RomR

A. SDS PAGE showing the purified MglC protein using SEC (Superdex 75). In denotes input, L denotes ladder with bands labelled in kDa and the purified MglC band is labelled.

B. Analytical Superdex 200 runs showing the interaction of $RomR^{REC\text{-}Ct}$ with MglC. The lower panels show the corresponding SDS PAGE of the respective runs. The lane numbers indicate the volumes of elution as per the chromatograms (In Input, L: Ladder) and the proteins corresponding to the bands are labelled. The corresponding peaks in the chromatograms are also labelled.

$RomR^{REC\text{-}Ct}$ (purple), $RomR^{REC\text{-}Ct}$ + MglC (cyan) and MglC (orange)

C. Analytical Superdex 200 runs showing the interaction of $RomR^{REC\text{-}Ct}$ with MglC. The peaks in the chromatograms are labelled as $RomR^{Ct}$ (R; red), $RomR^{Ct}$ + MglC (RC; cyan) and MglC (C; orange).

7.2.6 MglC interacts with RomR via the C-terminal region

It has been already reported by Kapoor *et al.*, 2021, using SAXS studies that MglB and MglC form a strong complex in a 4:2 ratio [5,6]. We further checked for MglC interaction using our short-linked construct of RomR ($RomR\text{-}REC^{141}Ct$, as studied in Chapter 6) using analytical

SEC on Superdex 200 column (Methods: 2.7.1). We observed that there was indeed an interaction between RomR-REC^{141}Ct and MglC (Fig 7.6 B). We performed the same experiment using the RomR^{Ct} construct containing the C-terminal domain of RomR. We observed a shift in the peak of the complex to 12.7 ml, whereas RomR^{Ct} and MglC, individually, eluted at 13.3 ml and 14.6 ml, respectively (Fig 7.6 C). We could not resolve the two proteins in SDS PAGE as they both run at around 13 kDa. Further structural and biochemical studies can throw light on the mechanism and implication of the interaction between RomR and MglC.

7.3 Conclusion

We have performed preliminary structural and biochemical characterization of RomY. In cells, the activity of RomY is only exerted at the lagging cell pole where MglB concentrations are high. Here, it binds MglA-GTP in the presence of MglB stimulating fast GTP hydrolysis and excluding MglA-GDP to the cytoplasm. The interaction of RomY could play a role in modifying the affinity of MglB to MglA in the GTP and GDP bound states, which otherwise is comparable. These functions could help regulate the dominant GAP activity of MglB at the lagging pole. The absence of RomY and MglB interaction at the leading pole could reduce the GAP activity, and the concerted GEF activity of RomX and MglB could help generate and retain abundant pools of MglA-GTP at the leading pole.

Based on our results and the AlphaFold model of MglA complexed with RomY and MglB, it is possible that the C-terminal extension of MglB might play a role in stabilizing the RomY interaction with MglA in the GTP bound state, stimulating hydrolysis. We provide experimental validation of this model by demonstrating that the C-terminal helix of MglB is necessary for the stimulation of GAP activity by RomY. Further studies with the C-terminal asymmetric constructs as described in Chapter 4, could provide an excellent system to validate the exact mechanism of RomY binding and activation. Consequently, mutational studies between RomY and MglA/B interface remain to be explored.

We observe that in addition to the interaction of MglB with MglC, the C-terminal tail of RomR also binds to MglC. RomR at the REC domain binds to RomX with the long-disordered linker acting as a link between the C-terminal tail which interacts with MglC. MglC did not have a known function in the polarity module. Based on our preliminary observation, it seems that RomR via its C-terminal domain interacts with MglC, which further recruits MglB at the

lagging cell pole. Since MglC and MglB interact in a 2:4 ratio, it generates positive feedback causing active recruitment of MglB in high concentrations during polarity reversals. MglB has comparable affinities for MglC[5] and MglA (in the range of 0.2-0.4 μM). Based on docking studies of MglB dimer with MglC and the structure of MglB dimer with MglA, it is predicted that there exists a conformational difference in the MglB dimer in its MglC and MglA bound states[6], which needs further validation. It is interesting to explore the possible role of this conformational difference of MglB in regulating its interaction with MglA, and its consequent GAP activation, which could be regulated by MglC. Further, it will be interesting to compare the relative affinities of MglB and RomR to MglC and how these 3 proteins form a polar scaffold, regulating the time scales of polarity reversals.

References

1. Baranwal J, Lhospice S, Kanade M, Chakraborty S, Gade PR, Harne S, Herrou J, Mignot T, Gayathri P. Allosteric regulation of a prokaryotic small Ras-like GTPase contributes to cell polarity oscillations in bacterial motility. *PLoS Biol.* 2019 Sep 27;17(9):e3000459. doi: 10.1371/journal.pbio.3000459.
2. Szadkowski, D., Carreira, L. A. M. & Søgaard-Andersen, L. A bipartite, low-affinity roadblock domain-containing GAP complex regulates bacterial front-rear polarity. *PLoS Genet* **18**, (2022).
3. McLoon AL, Wuichet K, Häsler M, Keilberg D, Szadkowski D, Søgaard-Andersen L. MglC, a Paralog of *Myxococcus xanthus* GTPase-Activating Protein MglB, Plays a Divergent Role in Motility Regulation. *J Bacteriol*. 2015 Nov 16;198(3):510-20. doi: 10.1128/JB.00548-15.
4. Leonardy S, Miertzschke M, Bulyha I, Sperling E, Wittinghofer A, Søgaard-Andersen L. Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. *EMBO J* **9**, 2276–2289 (2010).
5. Kapoor, S., Kodesia, A., Kalidas, N., Ashish & Thakur, K. G. Structural characterization of Myxococcus xanthus MglC, a component of the polarity control system, and its interactions with its paralog MglB. *JBC* **296**, (2021).
6. Carreira, L., Szadkowski, D., Lometto, S., Hochberg, G. K. A. & Søgaard-Andersen, L. Molecular basis and design principles of a system for switchable 3 front-rear polarity and directional migration 4 5 doi:10.1101/2022.12.09.519731 (2022).
7. Levine TP, Daniels RD, Wong LH, Gatta AT, Gerondopoulos A, Barr FA. Discovery of new longin and Roadblock domains that form platforms for small GTPases in Ragulator and TRAPP-II. *Small GTPases* vol. 4 62–69 Preprint at https://doi.org/10.4161/sgtp.24262 (2013).

Chapter 8: **Conclusion and future directions**

8.1 Major Findings

The study attempts to investigate the crosstalk among multiple players of the polarity module and their mechanism of activation of MglA, the master regulator driving polarity reversals and motility. In Chapter 1, we presented a model of polarity reversal in *M. xanthus* based on existing literature. We obtained the following results from our work, based on which we attempt to update the model and highlight the open ends to be investigated in future studies.

- **MglA Asp-58 is the non-canonical Walker B motif**

We validated the newly identified Walker B motif in MglA important for maintaining the Mg^{2+} coordination at the nucleotide-binding pocket[1]. Asp-58 present in the $G3^{Alt}$ motif acts as a non-canonical Walker B. We performed an exhaustive study of other P-loop NTPases to compare the position of Walker B and the associated conformational changes regulated across the GDP and GTP-bound states. The $G3^{Alt}$ motif is a part of the MglB interface, which allosterically regulates the positioning of the Walker B. The presence of a Thr-78 at the canonical Walker B position which plays a MglA activity enabled us to identify an equivalent conserved threonine in ASCE ATPases, which could play a role in driving nucleotide hydrolysis. Further, we propose that across evolution, the ancestral Thr-78 of MglA has probably been replaced with an Asp in the eukaryotic successors, Arf/Sar/Gα, to achieve the function of Mg^{2+} coordination essential for nucleotide binding. The allosteric regulation of Walker B, and thereby hydrolysis, probably plays a major role in switching MglA through various stages of active and inactive conformations during the polarity reversals.

- **$MglB_1$ Ct-helix allosterically drives nucleotide exchange on MglA**

Having explored the role of MglB in the regulation of MglA nucleotide binding and hydrolysis, we attempted to dissect the dual GAP and GEF activity of MglB, regulating the MglA nucleotide bound state, which is relevant to its subcellular localization. We designed a unique strategy to generate MglA complexed to asymmetric MglB dimers with C-terminal helix truncations. We confirmed that the C-terminal helix of MglB, opposite to the nucleotide-binding pocket of MglA, allosterically drives GDP to GTP exchange on MglA. We designed mutants of MglB which confirmed that the GAP activity of MglB is driven by the dimeric interface of the MglB Roadblock domain, with a major contribution from the protomer near the Switch residues of MglA. Hence, we confirm that interaction with specific residues of the $MglB_2$ protomer with MglA assists in orienting the catalytic residues of the GTPase in a

conformation to drive GTP hydrolysis by MglA. This reinforces the mechanism of indirect GAP activity, as proposed for MglB unlike other eukaryotic GAPs, which directly contribute catalytic residues to the GTPase. We also decipher a unique mechanism of allosteric GEF activity, unlike canonical GEFs which interact at the nucleotide-binding pocket, competitively driving out GDP. Identification of distinct GAP and GEF domains of MglB enables us to propose a model of how these activities can be spatially regulated across the functionally relevant dominant GAP activity at the lagging pole and GEF activity at the leading pole.

- **RomX alone acts as MglA GEF, with no contribution from RomR-REC**

We structurally and biochemically characterize RomX, a newly identified player of the polarity module. It was found to be structurally similar to the stalk domain of atlastin, a member of the dynamin family GTPases. We characterized the nucleotide exchange activity of RomX. However, the interaction was found to be weak and transient. Further, we perform a domain-wise dissection of RomR and its function. We characterized the RomR/RomX complex, mediated by the RomR N-terminal REC domain and the residues critical for forming the interface. However, we did not observe the additional stimulation of RomX GEF activity in presence of RomR, as reported earlier[2]. Using the information on the RomR-RomX complex, we generated models with MglA and found it was captured in the GDP state, as expected for a GEF. Finally, we inspected the long-disordered linker of RomR and compared it with a dataset of RomR-like proteins, and observed that these regions mostly form disordered globules and can expand based on additional interactions. Hence it can act as a bridge between two independently interacting N and C-termini of RomR.

- **MglB C-terminal region is important for RomY-mediated GAP stimulation**

To fill the gaps in understanding the polarity module, we initiated the characterization of MglB interactors, which might play a role in regulating its GAP and GEF activity. We determined the structure of RomY, the MglB GAP activator. We also observed that the C-terminal helices of the MglB dimer are important for RomY-associated GAP stimulation. Based on this information, we proposed a model in line with Szadkowski *et al*., 2022, where additionally the C-terminal helix of the MglB protomer near the MglA catalytic pocket could be preferentially critical for RomY binding and activity.

- **MglC interacts with RomR C-terminal and MglB**

We observed that MglC, another MglB-like Roadblock domain protein also interacts with RomR. The interaction is mediated through the C-terminal domain of RomR, with no contribution from the receiver domain. Further insights on heteromeric complexes of MglB/MglC/RomR, MglA/MglB/RomY and close inspection of conformations and binding interfaces could give us more information on relevant complexes at the cell poles.

8.2 Lacunae in the previous models of polarity reversal

Earlier models do not incorporate a precise role of RomR in regulating MglA and MglB localization. Having dissected the GAP and GEF active domains of MglB, we propose the $MglB_1$ Ct helix as a regulatory handle, spatially resolving the GAP activity at the lagging pole and the GEF activity at the leading pole. Attaching to the polar scaffolds via MglB Ct-helix at the lagging pole could further help recruit MglB. Moreover, the interaction of MglB with MglC via the Rbl/LC7 domain might help regulate the GAP activity at different stages of reversals. This clarifies the function of MglC in the polarity module.

Earlier models present a fuzzy picture of RomR and its role in regulating MglA and MglB localization. Having dissected the multiple domains of RomR, we clarify how it recruits the other proteins at the lagging pole. It recruits RomX through the N-terminal receiver domain. Via the C-terminal domain, it recruits MglC, which further recruits MglB. Hence, the concentration of RomR determines the net concentration of all these interacting proteins at the leading and lagging poles. The weak interacting interface between RomX and MglA could be the possible reason behind the dominant MglB/RomY GAP activity at the lagging cell pole. However, a low-grade GEF activity of RomX together with the active GAP activity of MglB/RomY help exclude MglA from the lagging pole. At the leading pole, a concerted GEF activity of RomX and MglB-Ct plays a role to recruit MglA.

8.3 A revised model of the polarity module

Summarizing our results and interpretations, we propose an updated picture of the polarity module and the probable sequence of events during a polarity reversal. Under normal conditions, the Frz signalling is low, and hence only low concentrations of $FrzX^P$ and $FrzZ^P$ exist within the cell. This results in a basal reversal frequency. A polarized cell has a sufficient concentration of RomR at the lagging cell pole. It acts as a scaffold to recruit MglC via its C-

terminus and RomX via its REC domain (Fig 8 A). MglC further recruits MglB (Fig 8 B, step 1). When MglB is present in high concentration, it is capable of interacting with MglA-GTP and RomY, forming a heteromeric complex. This causes a conformational transition of MglB from MglC binding to a MglA-RomY binding conformation (Fig 8 B, step 2). Upon detachment from MglC, there is a need for an additional recruitment factor for MglB, which keeps it localized to the lagging pole. Hence, in this condition, MglB, RomY, RomX, and MglC, in addition to RomR, all are accumulated at the lagging pole.

Lagging pole maintenance

Hence, at the lagging pole, the GAP activity of MglB is stimulated by RomY which together functions to exclude MglA-GTP from the lagging pole by converting it to MglA GDP, which has to diffuse in the cytoplasm. Here the GAP domain of MglB is free to bind MglA-GTP with further stabilization from RomY (Fig 8 B, step 2). The GEF active allosteric helix of MglB ($MglB_1$-Ct) could be sequestered by additional unknown interactions, thereby inhibiting MglB binding to MglA-GDP, releasing MglA. This interaction via the C-terminal helix could also keep MglB recruited at the pole in absence of its interaction with MglC. RomX localization to the lagging pole could be an effect of its interaction with RomR. But due to the transient, low-affinity interaction of RomX with MglA GDP, the GEF activity provides input of MglA-GTP to the active MglB RomY GAP for GTP hydrolysis. A basal level of RomX GEF activity is important for generating local pools of MglA-GTP, which is immediately converted to MglA-GDP and recycled to the cytoplasm by the MglB/RomY GAP activity. $MglB_2$-Ct helix might play a role in assisting the GAP stimulation of RomY.

Leading pole maintenance

At the leading cell pole, all the proteins in the lagging pole, except RomY are present, but in lesser concentrations. The absence of RomY is probably due to lesser concentration of MglB, insufficient for its recruitment, causing a net reduction in the MglB GAP activity. MglA-GTP is recruited at the pole which requires an active GEF activity (Fig 8 B, step 2). This could be contributed by the concerted activity of RomX and MglB-Ct, which dominates over the MglB GAP activity. A concerted activity of RomX at the MglA active site and MglB-Ct at the allosteric site could together exert a dominant GEF activity. Additional interaction from MglC and negligible RomY interactions could contribute to an occlusion of the GAP domain of MglB to interact with MglA, but the MglB C-terminal remains free to exert its GEF activity. Further, the presence of effectors like SgmX at the leading pole could help sequester more amount of

MglA-GTP at this pole (Fig 8 C). A GTPase inactive state (MglA-GTP* [7]) conformation could also maintain MglA in the GTP state at the leading pole.

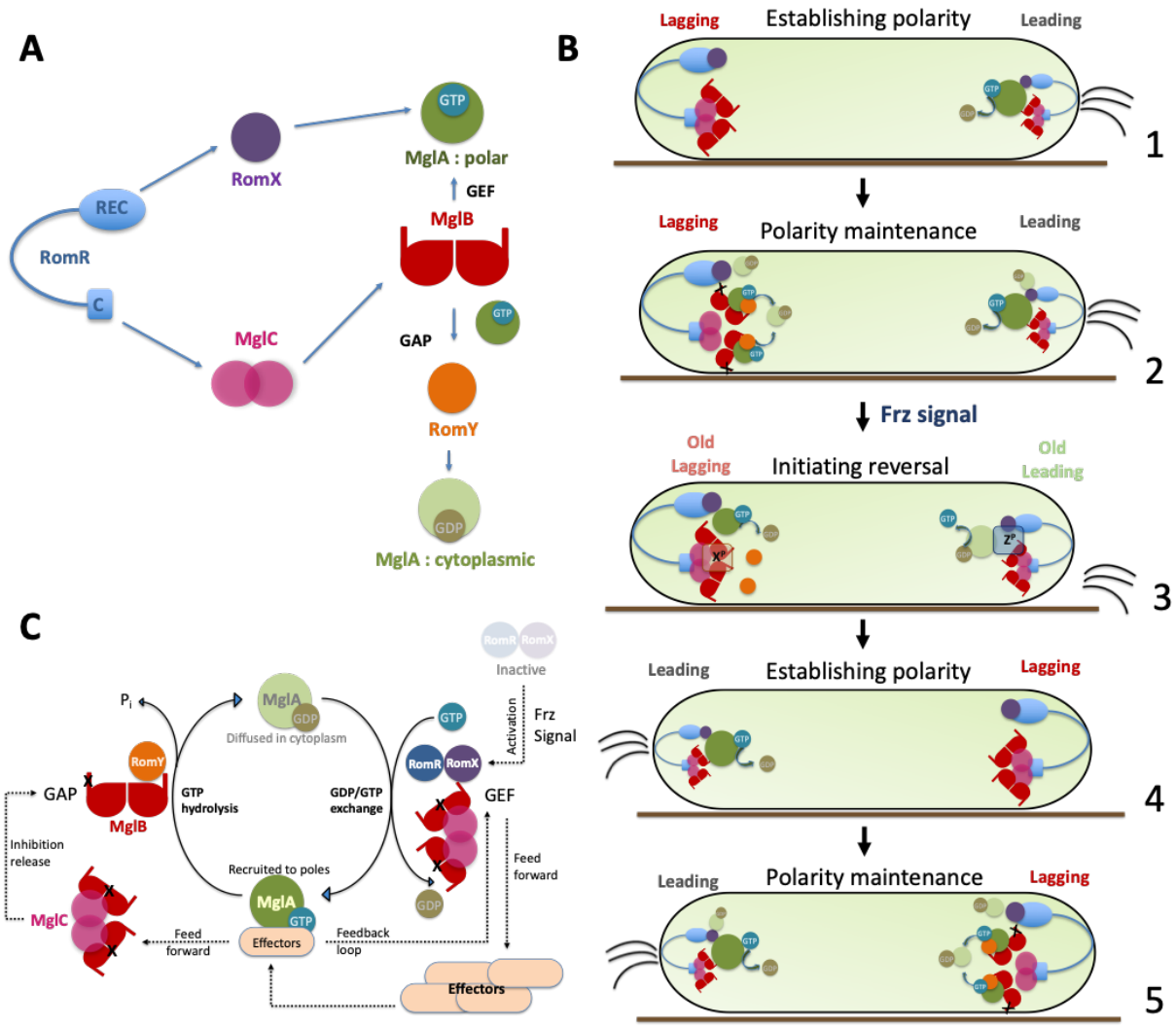


Figure 8. Current model of polarity reversals in *M. xanthus*

A. Regulatory circuit of polarity module proteins

B. Order of events during polarity reversals (steps 1-5)

C. Regulation of MglA GTPase cycle by GAPs and GEFs with speculated positive and inhibitory feedback mechanisms shown in dotted lines.

Polarity reversal event

Upon signalling from the Frz module, $FrzX^P$ accumulates at the lagging pole. It is unknown how $FrzX^P$ triggers the reversal. One of the first events in a reversal is the dissociation of MglB and its localization towards the opposite pole. This could be achieved by $FrzX^P$ by breaking

the MglB polar recruitment mediated via its C-terminal helix. Consequently, both MglB and RomY fall off from this pole (Fig 8 B, step 3). The remaining amount of MglB binds to MglC, in a GAP inactive conformation. But due to the slow dynamics of RomR, RomX is retained. This stimulates the same order of events as the old leading pole, where the GEF activity of RomX and MglB-Ct dominates and eventually causes net recruitment of MglA-GTP at this pole. Simultaneously, $FrzZ^{P}$ localizes to the old leading pole and causes MglA to fall off, thereby reducing the length of the refractory period produced by RomR. This could be due to the basal GAP activity of MglB alone which has slowly accumulated at this pole. Consequently, as RomR slowly accumulates at the old leading pole, it progressively recruits more MglC and MglB (Fig 8 B, step 4). Once RomR and MglB have accumulated sufficiently, RomY is recruited at the old lagging pole and stimulates the MglB GAP activity further (Fig 8 B, step 5). This marks the end of the reversal event where the cell is polarized and ready to receive the next stimulus for reversal.

Table 8: Summary of polarity module proteins and their functions

Protein	Function	Evidence	Protein	Function	Evidence
Polarity Maintenance					
Lagging Pole			*Leading Pole*		
RomR	Recruits MglC via Ct	This study (Ch 7); Carreira *et al.*[2]	**RomR**	Recruits MglC via Ct	This study (Ch 7); Carreira *et al.*[2]
	Recruits RomX via REC	This study (Ch 6)		Recruits RomX via REC	This study (Ch 6)
MglC	Recruits MglB	Mcloon *et al.*[3] Kapoor *et al.*[4]	**MglC**	Inactivates MglB GAP	Needs to be validated
MglB	MglA GAP activity at high concentration, with RomY	Szadkowski *et al.*[5]	**MglB**	MglB GEF	This study (Ch 4)
	GEF helix inactivated	Needs to be validated			
RomY	MglB GAP stimulation	Szadkowski *et al.*[5]	**RomX**	MglA GEF	Szadkowski *et al.*[6] This study (Ch 5)
	Role of MglB-Ct in GAP stimulation	This study (Ch 7)			
RomX	MglA GEF	Szadkowski *et al.*[6] This study (Ch 5)			

Protein	Function	Evidence	Protein	Function	Evidence
Polarity reversal					
Lagging pole to leading pole			*Leading pole to lagging pole*		
FrzX^P	Displaces MglB	Needs to be validated	**FrzZ^P**	Displaces MglA	Needs to be validated
RomR	Recruits MglA GTP via RomX GEF	Szadkowski *et al.*[6] This study (Ch 6)	**RomR**	Recruits MglC via Ct	This study (Ch 7) Carreira *et al.*[2]
				Recruits RomX via REC	This study (Ch 7)
MglC	Inactivates MglB GAP	Needs to be validated	**MglC**	Recruits MglB	Mcloon *et al.*[3] Kapoor *et al.*[4]
MglB	MglB GEF	This study (Ch 4)	**MglB**	MglA GAP activity at high concentration, with RomY	Szadkowski *et al.*[5]
RomX	MglA GEF	Szadkowski *et al.*[6] This study (Ch 5)	**RomX**	MglA GEF	Szadkowski *et al.*[6] This study (Ch 5)

8.4 Future Directions

A revised model for polarity reversals was proposed taking into account all the observations from this study and the existing work in the field. However, there is a requirement for additional validation to complete the whole regulatory circuit and explain the order of events. There are questions such as how exactly the FrzX^P and FrzZ^P initiate the reversal at lagging and leading poles, respectively. This requires detailed structural and biochemical characterization of these Frz output proteins and their interaction with relevant players, such as MglB, MglA and RomR and other interacting proteins, many of which are yet to be identified.

We have proposed a model where dominant GAP and GEF activities at the lagging and leading poles, respectively, explain the localization pattern of MglA. This is due to the spatially resolved GAP and GEF activities of MglA, differentially regulated at the two poles. MglB shows similar switching timescales as MglA (45-60 seconds), as compared to other proteins

like RomR. This probably allows an active regulation of MglA by the spatially resolved GAP and GEF activities within similar timescales, by MglB. This could be the relevance of having this dual activity coupled to the same protein. However, we still require experimental validation to explain how this dual activity is regulated by the concerted activity of all the polarity module proteins. Quantification of MglA GTPase and nucleotide exchange activities (using the techniques employed in this study) in the simultaneous presence of MglB and RomX or MglB and RomR or MglB with RomR/RomX or RomY, respectively, could give us a clear picture of which activity dominates when all these proteins are present together at the cell poles. Quantification of binding affinities between MglA and its interactors, like RomX or RomY, could also throw light on which of the protein competitively excludes the other to bind and exert its regulation on MglA. The structures of the newly identified proteins, RomX and RomY could be used to identify interacting interfaces of the relevant protein complexes through further structural and mutational studies. Our study has also provided certain tools such as the asymmetric MglA/MglB complexes, or the domain-wise dissected constructs of RomR, which could be used to inspect such multi-faceted interaction of this complex regulatory network.

References

1. Kanade, M., Chakraborty, S., Shelke, S. S. & Gayathri, P. A Distinct Motif in a Prokaryotic Small Ras-like GTPase Highlights Unifying Features of Walker B Motifs in P-Loop NTPases. *J Mol Biol* **432**, 5544–5564 (2020).
2. Carreira, L., Szadkowski, D., Lometto, S., Hochberg, G. K. A. & Søgaard-Andersen, L. Molecular basis and design principles of a system for switchable 3 front-rear polarity and directional migration 4 5 doi:10.1101/2022.12.09.519731 (2022).
3. McLoon AL, Wuichet K, Häsler M, Keilberg D, Szadkowski D, Søgaard-Andersen L. MglC, a Paralog of *Myxococcus xanthus* GTPase-Activating Protein MglB, Plays a Divergent Role in Motility Regulation. *J Bacteriol*. 2015 Nov 16;198(3):510-20. doi: 10.1128/JB.00548-15.
4. Kapoor, S., Kodesia, A., Kalidas, N., Ashish & Thakur, K. G. Structural characterization of Myxococcus xanthus MglC, a component of the polarity control system, and its interactions with its paralog MglB. *JBC* **296**, (2021).
5. Szadkowski, D., Carreira, L. A. M. & Søgaard-Andersen, L. A bipartite, low-affinity roadblock domain-containing GAP complex regulates bacterial front-rear polarity. *PLoS Genet* **18**, (2022).
6. Szadkowski D, Harms A, Carreira LAM, Wigbers M, Potapova A, Wuichet K, Keilberg D, Gerland U, Søgaard-Andersen L. Spatial control of the GTPase MglA by localized RomR-RomX GEF and MglB GAP activities enables *Myxococcus xanthus* motility. *Nat Microbiol*. 2019 Aug;4(8):1344-1355. doi: 10.1038/s41564-019-0451-4.
7. Galicia C, Lhospice S, Varela PF, Trapani S, Zhang W, Navaza J, Herrou J, Mignot T, Cherfils J. MglA functions as a three-state GTPase to control movement reversals of *Myxococcus xanthus. Nat Commun*. 2019 Nov 22;10(1):5300. doi: 10.1038/s41467-019-13274-3.