Biochemical and Structural studies on SofG, a GTPase involved in motility of *Myxococcus xanthus*

A thesis submitted towards partial fulfillment of the requirements of BS-MS Dual Degree Program

By

Ningthoujam Birjeet Singh

20111073

Under the guidance of

Dr. Gayathri Pananghat

Department of Biology

Indian Institute of Science Education and Research (IISER) Pune

Thesis Advisor: **Dr. Sunando Datta,** IISER Bhopal

Certificate

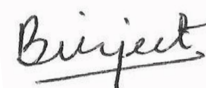
This is to certify that this dissertation entitled "Biochemical and Structural studies on SofG, a GTPase involved in motility of *Myxococcus xanthus*" towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research (IISER), Pune represents the research carried out by Ningthoujam Birjeet Singh at Indian Institute of Science Education and Research (IISER), Pune under the supervision of Dr. Gayathri Pananghat, INSPIRE Faculty, Department of Biology, IISER Pune during the academic year 2015 - 2016.

Date: 28th March 2016

Dr. Gayathri Pananghat
INSPIRE Faculty
IISER Pune

Declaration

I hereby declare that the matter embodied in the report entitled "Biochemical and Structural studies on SofG, a GTPase involved in motility of *Myxococcus xanthus*" are the results of the investigations carried out by me at the Department of Biology, Indian Institute of Science Education and Research (IISER), Pune, under the supervision of Dr. Gayathri Pananghat and the same has not been submitted elsewhere for any other degree.



Date: 28th March 2016

Ningthoujam Birjeet Singh

20111073

BS-MS student

IISER Pune

Abstract

Motility in bacteria is a fundamental process. Molecular mechanism behind novel modes of bacterial motility such as social motility and adventurous gliding motility in the social bacterium, *Myxococcus xanthus*, has not been fully elucidated. Recent studies in motility of *Myxococcus xanthus* showed the use of GTPases in deciding the polarity of the bacteria which leads to change in direction during movement. SofG is expected to be a GTPase that helps in polar localisation of PilB and PilT and thus help in pili formation. Towards understanding the mechanism of action of SofG, *in vitro* studies on *Myxococcus xanthus* SofG were carried out by cloning, optimising the protein purification protocol and thus purifying the protein from heterologous expression in *E. coli*. The oligomeric status of the purified SofG is a homogenous monomer, as observed from size exclusion chromatography. Circular dichroism spectroscopy showed that the purified protein exhibits secondary structure and also folded well. GTPase activity assays demonstrated that SofG is indeed a GTPase as it hydrolysed GTP. Structural studies have been initiated with setting up of crystallisation trials of the homogeneous protein sample. This could lead to crystallisation of the protein and determination of its crystal structure. These biochemical, biophysical and structural characterisation of SofG will contribute towards understanding the molecular mechanism of *Myxococcus* motility.

Table of Contents

Page No.

Chapter 1: Introduction 6

1.1 *Myxococcus xanthus* as a model system and its motility 6

1.2 Polarity and cell reversal in *M. xanthus* 7

1.3 SofG and its expected role in motility 9

1.4 Small Ras-like GTPases 11

1.5 Similarity between SofG and MglA 13

1.6 Objectives 13

Chapter 2: Materials and Methods 15

2.1 Design of SofG constructs 15

2.2 Cloning 16

2.3 Expression and solubility check of SofG in *E. coli* 18

2.4 Protein purification protocols 20

2.5 Malachite green assay 23

2.6 Crystallisation trials of SG-60 24

2.7 Binding assay using *mant*-labeled nucleotides 25

2.8 Circular dichroism spectroscopy 26

2.9 High Performance Liquid Chromatography (HPLC) 26

Chapter 3: Results and Discussion 27

3.1 Cloning 27

3.2 Optimisation of expression and solubility of constructs 28

3.3 Standardization of protein purification 31

3.4 Biochemical assays: Estimation of GTP hydrolysis 38

3.5 Size Exclusion Chromatography: Monomeric state of SG-60 39

3.6 Crystallisation trials 40

3.7 Binding assay using *mant*-labeled nucleotides 40

3.8 Circular Dichroism (CD) spectroscopy of SG-60 41

3.9 High Performance Liquid Chromatography (HPLC) to detect bound nucleotide 42

Chapter 4: Conclusion and Future Prospects 44

References 46

Abbreviations

CHAPS - 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate

DTT - Dithiothreitol

EDTA - Ethylene Diamine Tetra Acetate

GAP - GTPase Activating Protein

GDP - Guanosine Di Phosphate

GTP - Guanosine Tri Phosphate

HEPES - 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IPTG - Isopropyl β- D- thio Galactoside

LB - Luria-Bertani

Mant-GDP - (2'-(or-3')-O-(N-Methylanthraniloyl) Guanosine 5'-Diphosphate

Mant-GTP - (2'-(or-3')-O-(N-Methylanthraniloyl) Guanosine 5'-Triphosphate

MglA - Mutual gliding motility protein A

MglB - Mutual gliding motility protein B

Ni-NTA - Nickel- Nitrilotriacetic acid

OD_{600} - Optical density at 600 nm

PAGE - Poly Acrylamide Gel Electrophoresis

PCR - Polymerase Chain Reaction

PEG - polyethylene glycol

RPM – rotations per minute

SDS - Sodium Dodecyl Sulphate

T4P – Type IV pili

List of Figures

Figure No.	Figure Title	Page No.
1.1	Life cycle of *Myxococcus xanthus*	7
1.2	Polarity in *M. xanthu*s	8
1.3	Reversal of *M. xanthus* and localisation of proteins involved in it	9
1.4	Model proposing the function of SofG and BacP	10
1.5	Small GTPase regulatory cycle	11
1.6	Structural features of small Ras-like GTPase	12
1.7	Sequence alignment of SofG with MglA showing conserved motifs	13
2.1	Sequence alignment between *Myxococcus* MglA and SofG	15
2.2	Secondary structure prediction of SofG using PsiPred	16
2.3	Schematic representation of restriction-free (RF) cloning	19
2.4	Gel filtration run profile of markers	22
2.5	Standard graph obtained from different markers	23
2.6	Standard curve for phosphate estimation	24
3.1	Amplification of SofG by PCR and clone verification of SG-H6	27
3.2	Clone verification of SG-18 and SG-60	27
3.3	Clone verification of SG-277 and SG-60-277	28
3.4	Expression check of SG-60	29
3.5	Expression check of SG-60-277	29
3.6	Expression check of SG-277	30
3.7	Expression check of SG-60 at 15°C	30
3.8	Purification profile for Ni-NTA affinity purification of SG-60	32
3.9	Run profile of ion exchange chromatography using MonoS	34
3.10	Run profile of size exclusion chromatography using Superdex 75	35
3.11	Superdex 75 run profile of SG-60 after addition of GDP	37
3.12	Comparison of iP released by Malachite green assay of SG-60	38

3.13	Superdex 75 run profile of SG-60 with plot of markers	39
3.14	Needles and small crystals from screening	40
3.15	Mant-GDP binding assay of SG-60	41
3.16	CD spectrum of SG-60	42
3.17	HPLC run of nucleotides	42
3.18	HPLC run of SG-60	43

List of Tables

Table No.	Table title	Page No.
2.1	Constructs designed for cloning	17
2.2	Reaction mixture used for PCR amplification	17
2.3	Reaction conditions for PCR reaction	17
2.4	Primers used for cloning different constructs	18
2.5	Molecular weight markers for analytical gel filtration and their parameters	22
3.1	Expression and solubility status of various constructs	31

Acknowledgements

I would like to express my deepest gratitude to my mentor Dr. Gayathri Pananghat for giving me the opportunity to work in her group. It is a great pleasure to work in her group and her constant guidance, support and encouragement without which this work would not have been possible. I have learnt a lot from her and also I am grateful to her for the scientific freedom and guidance I was provided throughout the project.

I would like to thank Dr. Saikrishnan Kayarat for his valuable inputs, comments and helpful suggestions on this project. I really appreciate those lab meetings for clearing doubts and concepts.

Special thanks to Dr. Sunando Datta, IISER Bhopal for being a member of my thesis advisory committee and providing me with inputs.

I would like to thank Jyoti Baranwal, Priyanka Gade, Mahesh Chand, Neha Nirwan and Saket Bagde for assisting me in various parts of the project, experiments, assays and also encouraging me throughout the project. I really enjoyed your friendship and random funny talks. I also thank Shrikant Harne, Manasi Kulkarni, Ishtiyaq Khan, Rajnandani Kashyap, Vishal Adhav, Sujata Sharma, Manil Kanade, Sharvari Tendulkar, Gyan Mishra, Dheerendra Singh, Yaikhomba Mutum, Sonal Lagad and other lab members for various discussions, supporting me and maintaining a fun, lively atmosphere in the lab.

I thank all my friends and family for showing faith and providing confidence during my entire IISER life. I am also very grateful to INSPIRE for providing me stipend throughout my entire duration at IISER, Pune.

1. Introduction

Locomotion or motility is one of the main characteristics of bacteria. Motility affects or influences a number of functions in bacteria like growth and density of population (Lauffenburger et al., 1982), predation (Berleman and Kirby, 2009), virulence (Josenhans and Suerbaum, 2002), etc. Motility of bacteria are most commonly mediated by appendages such as flagella, pili etc. (Jarrell and McBride, 2008). A prokaryotic system with minimal number of components might be a tractable model system to understand motility, without the redundancy present in eukaryotic systems.

1.1. *Myxococcus xanthus* as a model system and its motility

Myxococcus xanthus is a gram negative rod-shaped delta proteobacteria. It is commonly found in the top soil and has a genome size of 9.45 Mb (Chen et al, 1990). *M. xanthus* exhibits complex life cycle and so undergoes complex developmental processes (Figure 1.1). Depending on the nutritional status, there are two different cellular patterns. In nutrient rich condition, they move together as swarms, while in nutrient poor conditions, they form fruiting bodies called as myxospore (Konovalova et al., 2010). These myxospores are heat-resistant and can re-germinate during suitable environmental conditions. Thus, they exhibit two types of motility systems – (i) Social motility (S-motility) or twitching motility: It is a type IV pili dependent motility and is responsible for movement of cells in group and (ii) Adventurous gliding motility (A-motility): It is facilitated by complex of proteins analogous to focal adhesion complex and is responsible for movement of cells when isolated from the group (Mauriello and Zusman, 2007). Due to its complex life cycle characteristics, *Myxococcus* is considered a model organism to study differentiation, multicellularity, and novel modes of bacterial motility, etc.

In S-motility, the motility is achieved by extension and retraction of pili which brings the cell to move towards the direction of pili (Li et al., 2003). These extensions and retractions of pili are energized by PilB and PilT which are cytoplasmic ATPases (Jakovljevic et al., 2008). In A-motility, there are no external appendages involved and the cell moves by gliding on the surface (Nan and Zusman, 2011). This motility mechanism remains elusive as there are not much changes associated with the

morphology of the cell. The machinery for adventurous gliding motility have been observed by localization of fluorescently labelled components within the cell.

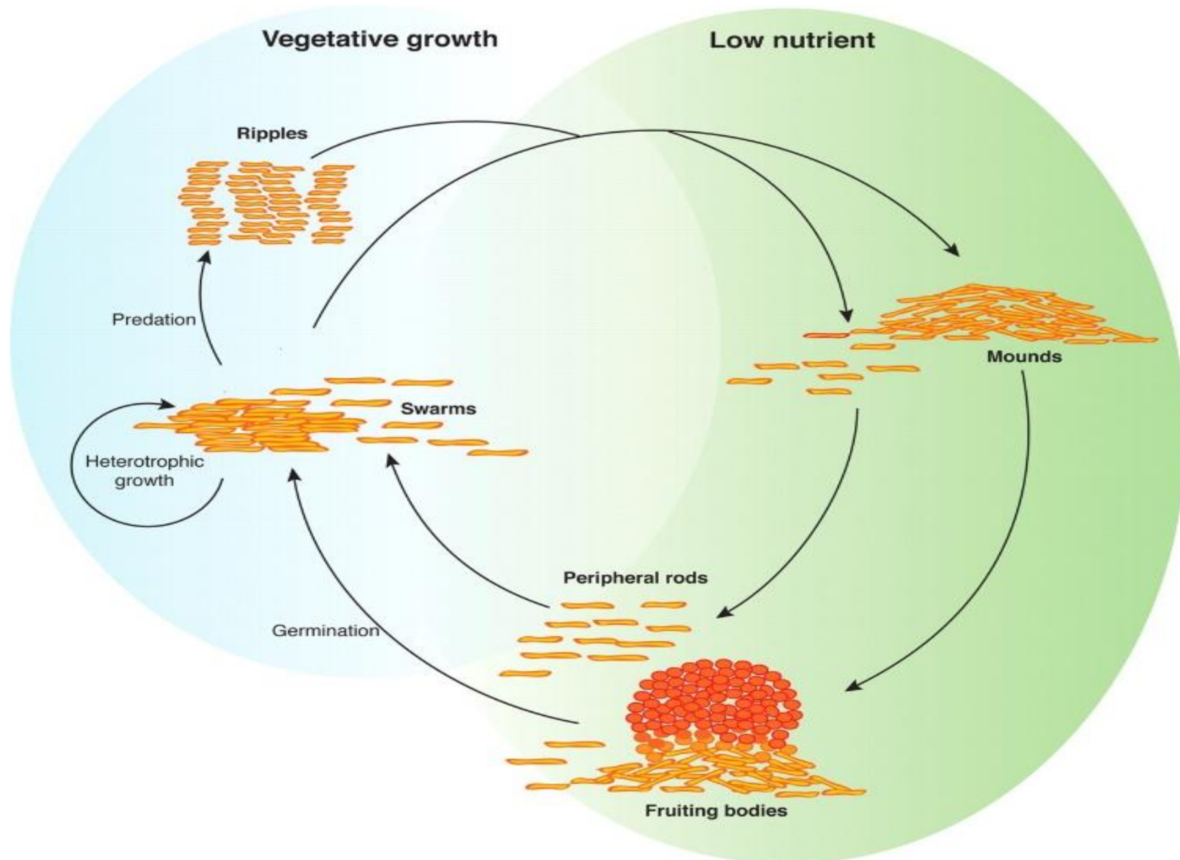


Figure 1.1: Life cycle of *Myxococcus xanthus* (Adapted from Mauriello et al., 2010)

Gliding and twitching motility has periodic reversals and these reversals help them to redirect their movement in response to stimulus. There is also distinct polarity in the cell and which changes frequently during periodic reversals (Mauriello and Zusman, 2007).

1.2. Polarity and cell reversal in *M. xanthus*.

Polarity arises because of arrangement of different components of the cell in particular locations, especially the poles. Bacteria have a very simple system as compared to eukaryotes. However, they regulate many processes by localizing proteins at different parts of the cell at different points of time. This leads to the creation of polarity to the cell and this polarity changes from time to time depending on the external environmental condition. This polarity creates distinct leading and

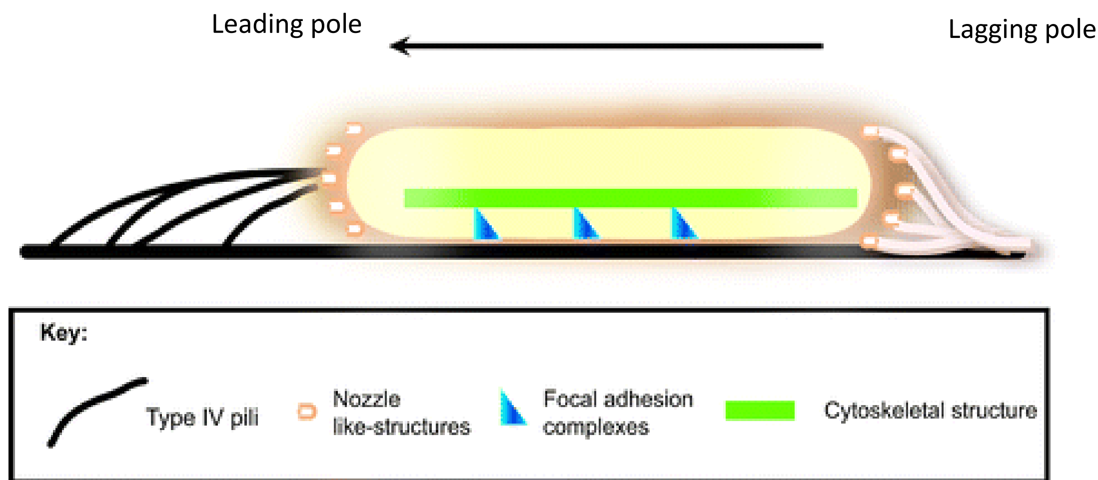


Figure 1.2: Polarity in *M. xanthu*s. (Adapted from Leonardy et al., 2008)

lagging poles in the bacteria. In *Myxococcus xanthus*, polarity is based on the position of pili and the direction of movement of the cell (Figure 1.2). The pole which has pili is known as the leading pole as it points to the direction of motion while the other opposite pole which does not have pili is known as lagging pole. Different types of proteins are found in the two poles. And this polarity is maintained by many proteins and among that small Ras-like GTPases of *Myxococcus xanthus*, MglA and SofG play a major role (Leonardy et al., 2010; Bulyha et al., 2013).

One of the unique features of *M. xanthus* is the frequent reversal of the poles which allow the cell to redirect in respond to attractant or repellent (Shi et al., 1993). This reversal happens at an interval of approximately every 7–8 minutes. During this reversal, the old leading cell pole becomes the new lagging pole and so this leads to the inversion of the cell (Blackhart and Zusman, 1985). This reversal is controlled by the Frz chemosensory system (Blackhart and Zusman, 1985). During reversal, Type IV pili (T4P) disassembles at the leading pole and the previous lagging pole becomes the new leading pole while the T4P assembles at this new leading pole (Mignot et al. 2005). So between reversals, the position of PilB, an ATPase which catalyses the extension of T4P and FrzS protein changes from previous leading pole to new leading pole and PilT, which energise the retraction of T4P changes from previous lagging pole to new lagging pole (Figure 1.3; Bulyha et al., 2009). This results in a change in the direction of movement of the bacteria (Figure 1.3).

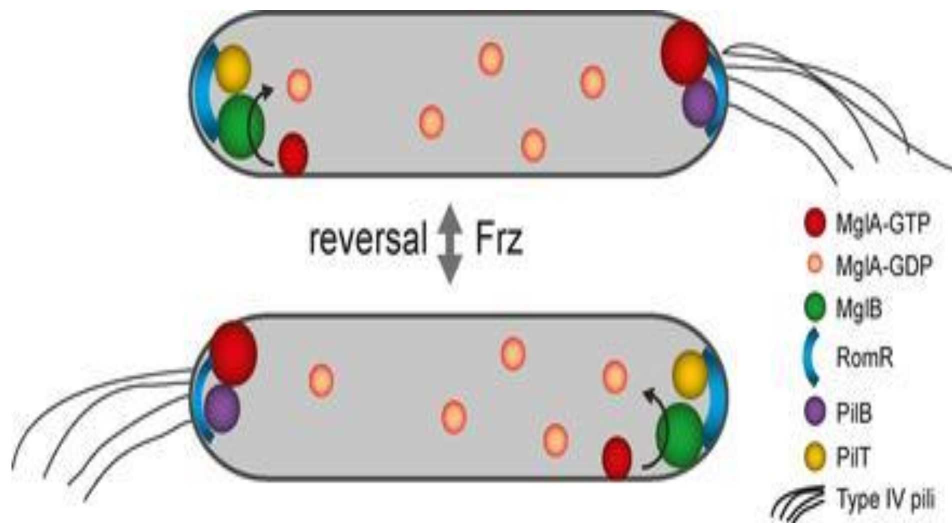


Figure 1.3: Reversal of *M. xanthus* and localisation of proteins involved in it. (Adapted from Treuner-Lange and Søgaard-Andersen, 2014)

In maintaining the polarity before and after reversal of the cell, there are many proteins involved and among them MglA, a small Ras-like GTPase and MglB, its cognate GAP, are important ones which help in defining the leading and lagging polarity axis. MglA in GTP-bound form is restricted to the leading pole only and MglB to the lagging pole only. This helps in defining the definite leading/lagging polarity axis (Leonardy et al., 2010). For reversals, the Frz system has to induce the relocation of MglA-GTP to the lagging pole and so this results in an inversion of the leading/lagging polarity axis (Leonardy et al., 2010). So GTPases are important in maintaining polarity and during reversals.

1.3. SofG and its expected role in motility

In a recent study by Bulyha et al. (Bulyha et al., 2013), they came across a new protein (MXAN_6703) which is expected to have a major role in the motility of *Myxococcus xanthus*. It is also a small Ras-like GTPase which was named as SofG (S-motility function GTPase). It is expected to interact with BacP (bactofilin, a novel cytoskeletal protein present in bacteria) which is required for the polar localization of PilB and PilT (Bulyha et al., 2013). It has all the conserved motifs of a GTPase ie. P-

loop, switch I, switch II, G4, etc. (see next section on GTPase motifs). Preliminary studies showed that deletion mutation of SofG (Δ*sofG*) exhibited defect in S-motility. It also suggested about the role of another protein BacP, which is a bactofilin cytoskeletal protein, in social motility. So, a model was suggested which includes the function of SofG and BacP (Figure 1.4).

The model suggested that SofG gets recruited on BacP to form a subpolar cluster (Figure 1.4). Then SofG shuttled over BacP in a GTPase-dependent manner to nearest pole and back and this results in the polar localisation of PilB and PilT. The shuttling of SofG might occur by depolymerisation and re-polymerisation of BacP.

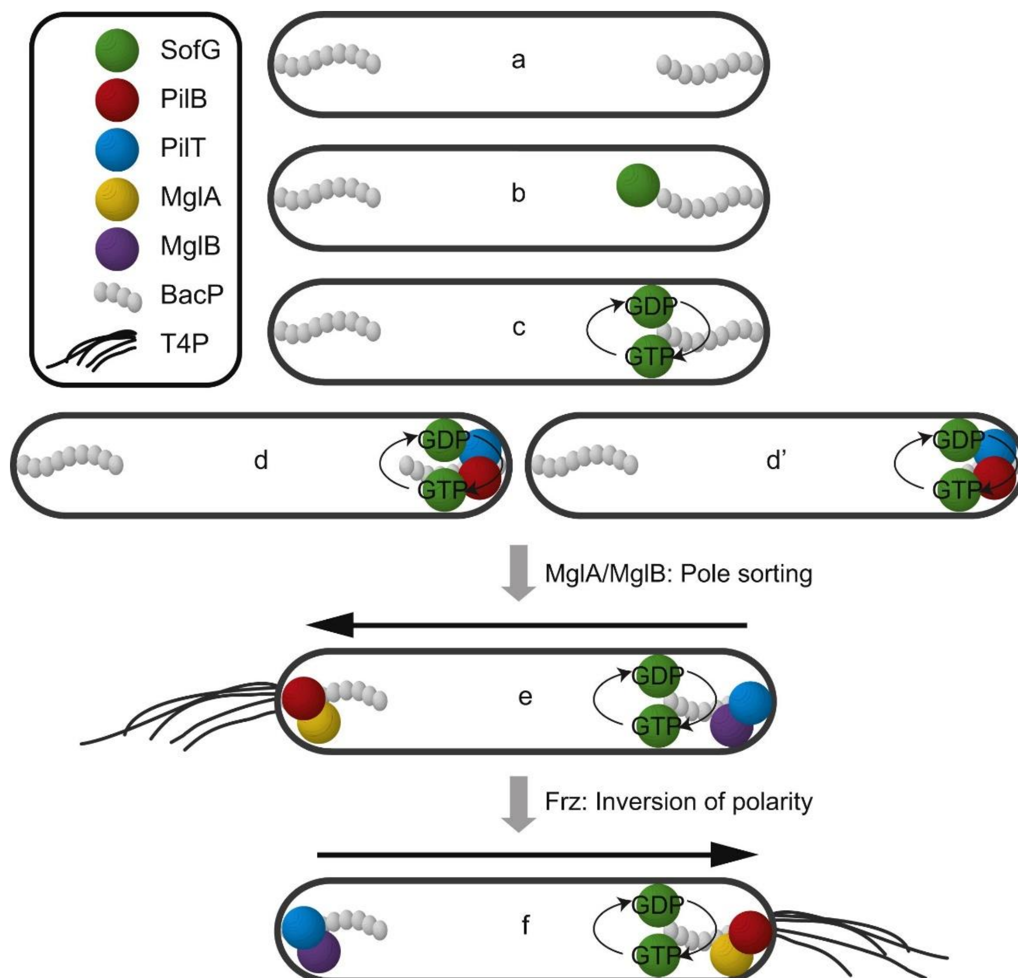


Figure 1.4: Model proposing the function of SofG and BacP in promoting dynamic polarity of PilB and PilT localization in a step-wise manner. (Adapted from Bulyha et al., 2013).

After bringing PilB and PilT to the same pole, MglA and its cognate GAP, MglB might sort them to different correct poles (Zhang et al., 2010). So MglA function with respect to PilB and PilT polarity might depend on SofG.

1.4. Small Ras-like GTPases

Two major players in the motility and polarity reversals of *Myxococcus*, MglA and SofG, belong to the family of small Ras-like GTPases. This section gives an introduction to the general features of small Ras-like GTPases.

GTPases comprise a large family of hydrolase enzymes which can bind to GTP and hydrolyse it. GTPase acts in a cyclic type of action in which they have an active form and an inactive form (Figure 1.5). They are also known as molecular switches as they can change their interactions depending on the GTP or GDP-bound conformations, and hence exist in the 'on' of 'off' state according to the requirement of the cell (Vetter and Wittinghofer, 2001). They are important in a wide range of cellular processes which includes movement and signalling events, intracellular transport, etc. (Heider et al., 2010).

Small GTPases are also known as Ras superfamily of GTPases and it is divided into five families: Rab, Rho, Ran, Arf (Sar), and Ras proteins (Wennerberg et al., 2005). Members of these families of proteins share a common core structure which is called as the G-domain that consists of five alpha helices and six beta sheets (Figure 1.6A). An important feature within these G-domains is the conserved structural feature which are known as the switch I and switch II regions, in which the major conformational changes take place upon binding of GTP and during hydrolysis (Heider et al., 2010).

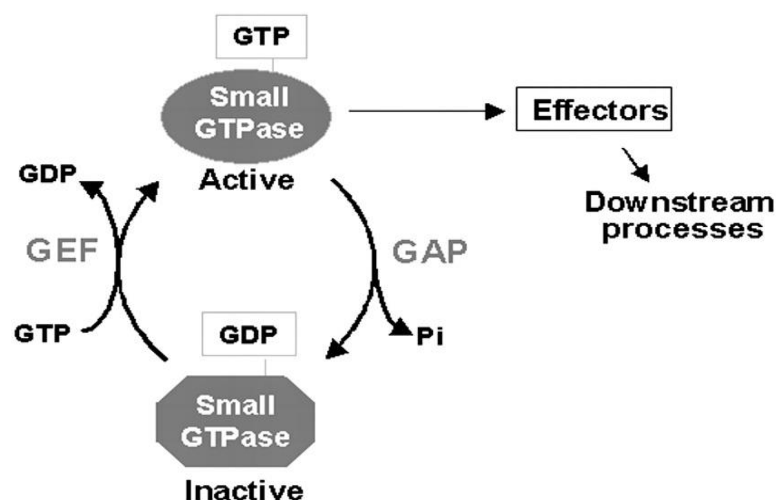


Figure 1.5: Small GTPase regulatory cycle (Adapted from Nielsen et al, 2008)

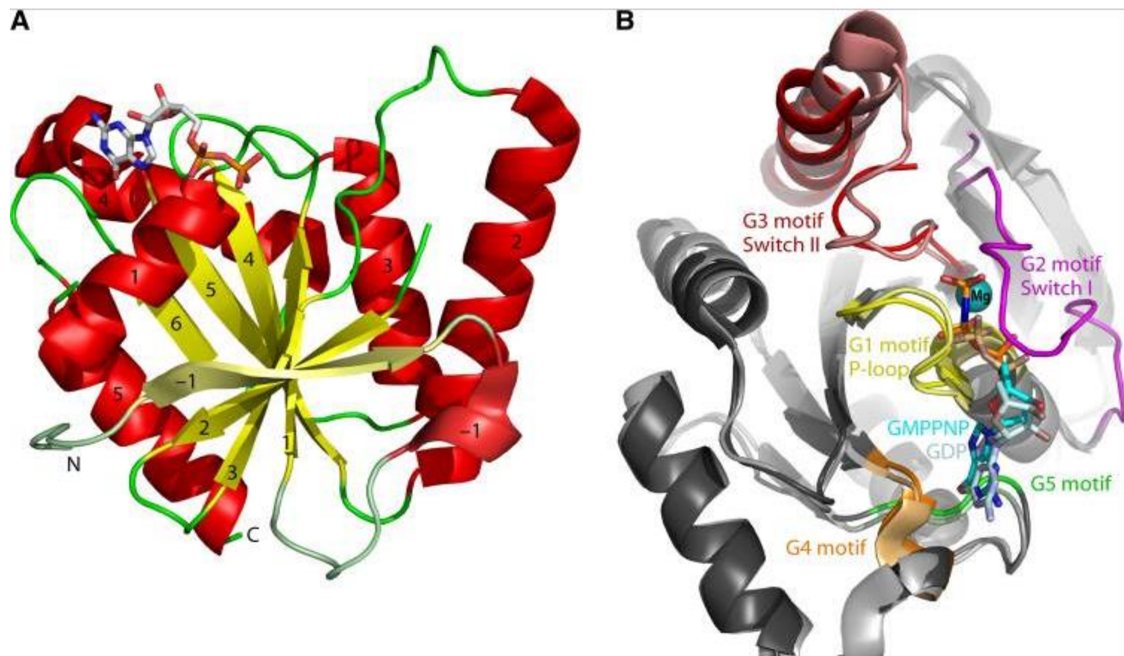


Figure 1.6: Structural features of small Ras-like GTPases. **(A)** Tertiary structure of small Ras-like GTPase highlighting the secondary structure elements (PDB ID: 1SVI). **(B)** Conserved GTPase motifs. (Adapted from Verstraeten et al. 2011).

The G domain is also further characterized by the presence of G motifs (G1 through G5) (Figure 1.6B), which are the conserved amino acid sequences. G1 [GX_4GK(S/T)] is also known as the P-loop or Walker A motif. It is involved in the binding of the phosphates of GTP and GDP (Paduch et al., 2001). G2 region is a highly conserved motif and contains a conserved threonine residue which is responsible for Mg^{2+} binding (Paduch et al., 2001; Hwang and Inouye, 2001). G3 motif comprises of DX_2G motif which is also involved in coordination to Mg^{2+} and binding to the γ-phosphate (Verstraeten et al., 2011). It is also known as Walker B motif. The G2 and G3 loops has the maximum change during binding and exchange of nucleotide and so they are known as Switch I and Switch II respectively. (Verstraeten et al., 2011). The G4 loop comprises the NKXD motif and has Lys and Asp residues which directly interacts with the nucleobase (Hwang and Inouye, 2001). G5 motif is involved in interactions with the guanine moiety of the substrate and is not strictly conserved across GTPases (Ruzheinikov et al., 2004).

1.5. Similiarity between SofG and MglA

SofG shares about 45% sequence identity with *Myxococcus xanthus* MglA (Figure 1.7). So, it is highly expected that it might have similar fold to that of MglA. Hence, it was suggested that SofG is a paralog of MglA (Bulyha et al., 2013).

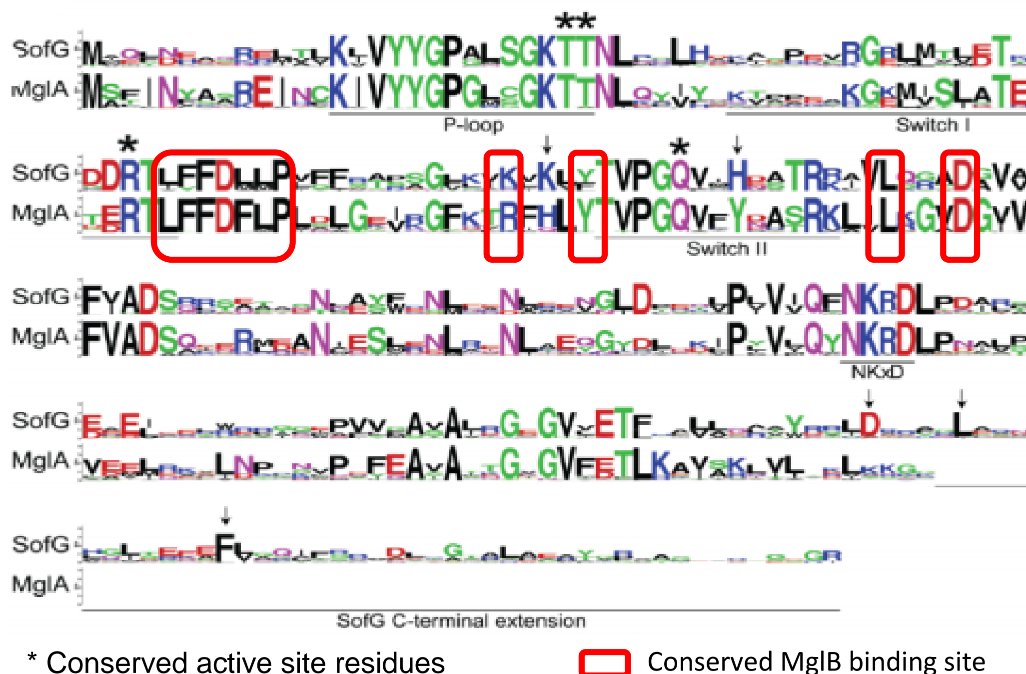


Figure 1.7: Sequence alignment of SofG with MglA showing conserved motifs of GTPases (Adapted from Bulyha et al., 2013).

1.6. Objectives

The molecular basis of how the two GTPases MglA and SofG act in concert to position the respective motility machinery and regulate cell polarity is currently not known. So, a structural and biochemical approach towards studying of SofG might help in understanding how these proteins function and how they interact with other proteins towards maintaining the polarity and assist in motility of the bacteria.

The specific objectives of the project are:

i) Cloning and overexpression of SofG in *E. coli*. Relevant constructs suitable to facilitate overexpression and protein purification will be designed.

ii) Standardization of protein purification of SofG constructs.

iii) Biochemical characterization of SofG through GTPase activity measurements.

iv) Structure determination of SofG using X-ray crystallography.

The following chapters of the thesis provide a description of the methodology adopted and detailed experimental procedures (Chapter 2), results obtained and their interpretation (Chapter 3) and conclusions and future directions of the work (Chapter 4).

2. Materials and Methods

The following chapter describes the various experimental methods followed during the work reported in the thesis.

2.1. Design of SofG constructs

Before starting cloning, constructs were designed based on information from literature, sequence databases, sequence alignment and secondary structure prediction. All the sequence alignments and secondary structure prediction were carried out using avaialble web servers such as ClustalO (Chenna et al., 2003) and PSIPRED (Jones DT., 1999).

SofG (UNIPROT number: MXAN_6703) shares 45% sequence identity with MglA (a small Ras like GTPase) in *Myxococcus*. Based on the sequence alignment between SofG and MglA (Figure 2.1), it was observed that SofG consists of extra 60 amino acids at the N-terminus, and ~50 amino acids at the C-terminus, compared to MglA. Constructs of SofG with and without the extra amino acids were designed by aligning with the MglA sequence. Based on sequence alignment with MglA (Figure 2.1), as the corresponding sequence lies in the mid portion, constructs shown in Table 2.1 were designed. Further constructs were designed based on secondary structure prediction (Figure 2.2).

```
CLUSTAL O(1.2.1) multiple sequence alignment


SofG            MRSRISSNRSCVQAIPRMSIASTPGTLPAFRVGPLPTIKPVVGREGVKLLRSTVSQSPTL
MglA            ----------------------------------------------------------MS


SofG            QLNHAQRELTLKIVYYGPGLSGKTTNLRHLHAKASPEVRGRLLTVETHDDRTLFFDLLPV
MglA            FINYSSREINCKIVYYGPGLCGKTTNLQYIYNKTAAETKGKLISLSTETDRTLFFDFLPL
                 :*::.**:. *********.******:::: *:: *.:*:*:::.*. *******:**:

SofG            FFSTSSGFKVKVKLFTVPGQVIHNATRRIVLQGADAVVFIADSRRSATADNNAYWRNLQE
MglA            SLGEIRGFKTRFHLYTVPGQVFYDASRKLILKGVDGVVFVADSQIERMEANMESLENLRI
                 :.   ***.:.:*:******:::*:*:::*:*.*.***:***: .    *    .**:

SofG            NMKENNLDPSQVPVVIQFNKKDLPDARTDAEIEESRRRGGEAVVGAVALRGEGVLETFHA
MglA            NLAEQGYDLNKIPYVIQYNKRDLPNAVTVEEMRKALNHRNIPEYQAVAPTGVGVFDTLKA
                *: *:  *  .::* ***:**:***:* *  *:.:: .:       ***  * **::*::*

SofG            VAQAAYRRLDMRAHLARNLGLTEAEFLGQIFRRMDLTGTALASMYGRAAGEARSGEGR
MglA            VAKLVLTELKKGG---------------------------------------------
                **: .  .*.  .
```

Figure 2.1: Sequence alignment between *Myxococcus* MglA and SofG obtained using Clustal Omega

Secondary Structure Map

Feature predictions are colour coded onto the sequence according to the sequence feature key shown below.

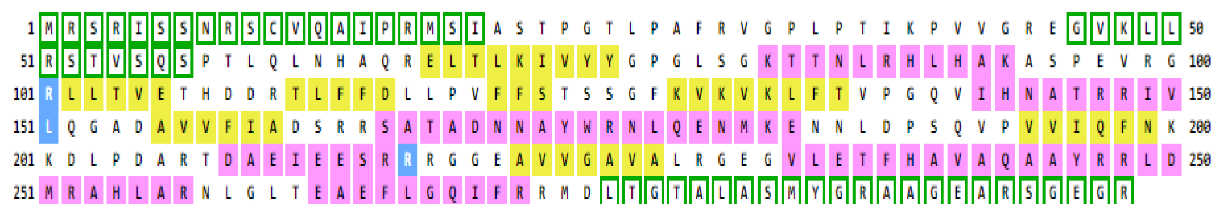


Figure 2.2: Secondary structure prediction of SofG using PSIPRED.

Table 2.1: Constructs designed

Construct description	Construct name
Full length untagged	SG-U
C-terminal $(His)_6$ full length	SG-H6
18 to 298 $(His)_6$	SG-18
60 to 298 $(His)_6$	SG-60
1 to 258 $(His)_6$	SG-258
60 to 258 $(His)_6$	SG-60-258
1 to 277 $(His)_6$	SG-277
60 to 277 $(His)_6$	SG-60-277

2.2. Cloning

Cloning is the technique of introducing our gene of interest into a vector. Once the gene is inserted into the vector, then this vector can be introduced into host organism and can make the host organism produce our protein of interest. To enable large-scale expression and purification of our protein of interest, essential for biochemical and structural studies, the host organism used in this study is *Escherichia coli*. This section primarily deals with strategies and protocols used for cloning of different SofG constructs into pHis17 vector.

2.2.1. Polymerase Chain Reaction (PCR) amplification of *sofG*

To start with cloning, first of all we need to have huge amount of our gene of interest (insert). So, our gene of interest was amplified from genomic DNA of *Myxococcus xanthus* (DSMZ; catalogue number 21863). This amplification was done by using a technique called polymerase chain reaction (PCR). The reaction mixture and the conditions used for PCR amplification are given in Tables 2.2 and 2.3.

Table 2.2: Reaction mixture used for PCR amplification

Components	Volume (µl)	Final Concentration
Genomic DNA	0.5	100 ng/ 50 µl
Forward primer	1	400 nM
Reverse primer	1	400 nM
Reaction Buffer	5	10x
AccuPrime DNA polymerase	0.5	
MilliQ water	42	
Total	50	

Table 2.3: Reaction cycle for PCR

Temperature (°C)	Time
97	5 minutes
97	30 seconds
58	30 seconds
68	55 seconds
68	2 minutes
4	∞

After amplifying the gene, two different methods can be employed to insert the gene to vector i.e. Restriction free method and Restriction digestion-ligation method. The primers for PCR were designed such that cloning by both methods were possible. However, all the clones were obtained by restriction free cloning method only and is described in the next section. The forward and reverse primers employed are given in Table 2.4.

Table 2.4: Primers used for cloning different constructs

Primer	Primer sequence (5'-3')	Length
MXSG-F	ATGAGGAGCCGGATTTCGTCGAACC	25
MXSG-R	TCGCCCTTCTCCGCTGCGCGCTTCAC	26
SG-F	GTTTAACTTTAAGAAGGAGATATACATATGAGGAGCC GGATTTCG	45
SG-H6R	GCTTTTAATGATGATGATGATGATGGGATCCTCGCCC TTCTCCGCTG	47
SG-18F	GTTTAACTTTAAGAAGGAGATATACATATGAGCATCGC CTCCACGCCTG	49
SG-60F	GTTTAACTTTAAGAAGGAGATATACATATGTTGCAACT CAACCATGCCC	49
SG-258 H6R	GCTTTTAATGATGATGATGATGATGGGATCCGTTGCG CGCCAGGTGGGCGCGC	53
SG-277 H6R	GCTTTTAATGATGATGATGATGATGGGATCCCGTGAG GTCCATGCGCC	48

2.2.2. Restriction free method of cloning (RF cloning)

This is a method of cloning which does not require any restriction enzyme and is restriction site independent (van den Ent and Löwe, 2006). It is a PCR-based method of incorporating the gene of interest into the vector. For this, primers were designed in such a way that both 3' and 5' ends should have vector complementary sequence. A schematic diagram of how restriction free cloning works and the steps involved are summarised in Figure 2.3. For verification of clones and for carrying out restriction-based cloning if required, genes were inserted between NdeI and BamHI sites. The positive clones were confirmed using NdeI and BamHI digestion and subsequently by sequencing.

2.3. Expression and solubility check of SofG in *E. coli*

In order to find out the optimal condition for the expression of protein, different strains of *E. coli*, different temperatures and different stages of induction were checked. The strains of *E. coli* used for expression check were: BL21AI, BL21DE3, Rosetta, C41 and C43. Induction of culture was checked at 0.4, 0.6 and 0.8 $O.D_{600}$. The growing cultures were incubated at different temperatures of 30°C, 25°C and 18°C.

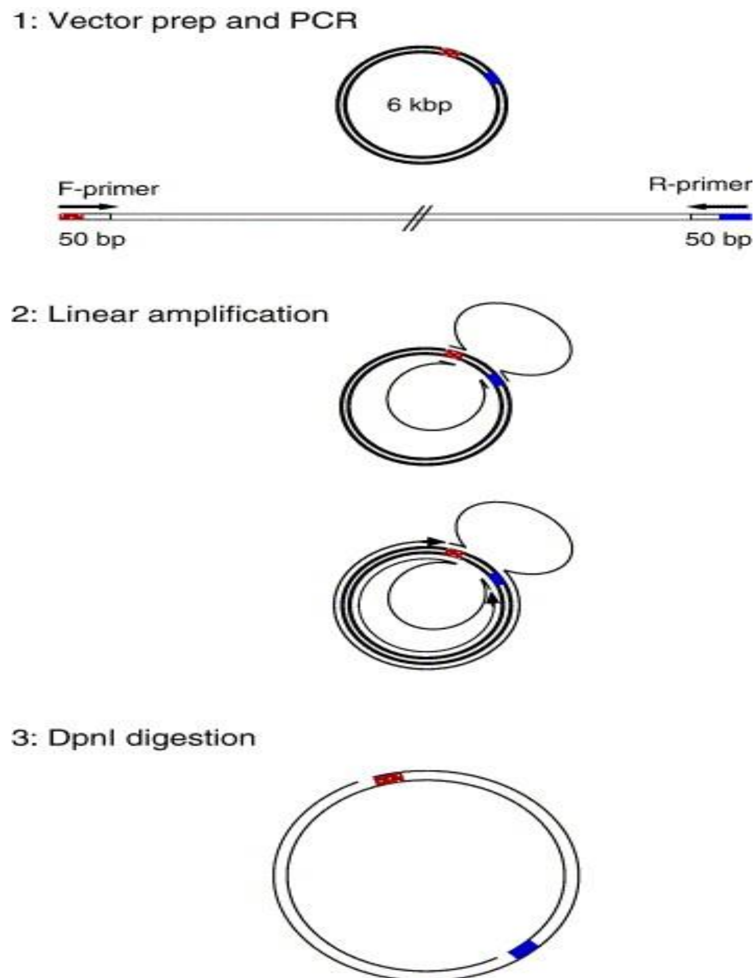


Figure 2.3: Schematic representation of restriction-free (RF) cloning. (Adapted from van den Ent and Löwe, 2006)

SofG constructs were transformed into each strain of *E.coli* cells by chemical transformation and plated on LB-Agar plates containing Ampicillin (100 µg/ml). A patch of colonies were inoculated in a 5 ml of sterilised LB broth containing ampicillin. The inoculated culture were kept at 37 °C until the desired $O.D_{600}$ reached. The culture was then induced by adding inducer like L-arabinose (SRL chemicals) keeping the final concentration of 0.2% for BL21AI and L-arabinose plus IPTG (Sigma; final concentration of 0.5 mM) for BL21DE3. After induction, the cultures were incubated at 30 °C for about 6 hours, about 8 hours for 25 °C or 12-14 hours at 18 °C by shaking at 200 RPM.

The cells were harvested by centrifugation and then supernatant was discarded. Then pelleted cells were resuspended in lysis buffer depending on the size of pellet (around 500 µl for 5 ml culture). The cells were sonicated by using a sonicator (VibraCell) at an amplitude of 50% for a pulse cycle of one minute (1 seconds on, 3

seconds off). After taking out an aliquot of 10 µl for total cell lysate, the cells were centrifuged at high speed and supernatant was collected. The amount of protein in the total cell lysate provides information about the overexpression while the protein in the supernatant fraction gives an estimate of the soluble fraction.

All samples were heated at 100°C with 2x SDS loading dye for 10 minutes and then run on a 12% resolving SDS PAGE.

2.4. Protein purification protocols

Following standardisation of the expression and solubility, larger volume cultures were grown, and the protein was obtained in the soluble lysate. Since SofG has been cloned with a hexa-histidine tag, the first step of purification was by Ni-NTA Affinity chromatography. All the chromatography experiments were carried out on AKTA Prime chromatography system (GE LifeSciences).

2.4.1. Affinity Chromatography: Ni-NTA

Since the protein was C-terminal hexa-histidine tagged, the first round of purification is an affinity chromatography technique i.e. Ni-NTA (Nickel Nitrilotriacetic acid). Nickel ions have capacity to bind to a stretch of histidine residues. So, the histidine-tagged protein can bind to the Ni-NTA column. But, those proteins which do not have an exposed histidine stretch will not bind to the column with high affinity. In order to elute the protein from the column, imidazole was used in the buffer. Imidazole has a similar structure to that of histidine and it can compete with histidine for binding with Ni^{2+} in the column. So at high concentrations of imidazole, the protein elutes out of the column.

A brief description of the standardised protocol is given below. The cell pellet was resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl, 10% Glycerol) in around 180-200 ml for 6 litre culture. Then it was sonicated at amplitude of 60% for a cycle of 6 minutes (1 second on, 3 second off). Then the lysate was centrifuged at High Speed centrifuge (Beckman Coulter) using JA25.50 rotor at 39191 xg for 45 minutes at 4°C. In the meantime, the 5ml Ni-NTA column (HisTrap HP, GE Lifesciences) was equilibrated with Buffer A (50 mM Tris pH 8.0, 200 mM NaCl). After the centrifugation got over, the cell lysate was loaded onto the Ni-NTA column. After loading, the column was washed with at least 6 column volumes (C.V.) each of

2%, 5% and 10% of Buffer B (Buffer A containing 500 mM imidazole) to remove proteins that bind non-specifically. The bound protein was eluted by washing the column with 6 C.V. each of 20% B, 50% B and 100% B, and the eluted sample was collected as 5 ml fractions. Alternate fractions were run on 12% resolving SDS-PAGE to check the purity of the protein.

2.4.2. Ion exchange chromatography

Both MonoS and MonoQ runs were carried out using AKTA Purifier system of GE Healthcare Lifesciences.

After Ni-NTA, fractions containing our protein of interest were pooled together and transfered to a dialysis bag (Thermo Scientific Snakeskin™ Dialysis tubing, 10K MWCO). Then the protein was dialysed for 2 hours in Buffer A25 (50 mM Tris pH 8.0, 25 mM NaCl) in order to remove excess salt and imidazole. Centrifugation was carried out in a High Speed centrifuge (Beckman Coulter) using JA25.50 rotor at 39191 xg for 30 minutes at 4°C. The protein was then filtered using 0.22 µm filter. Then protein was loaded to 1.7 ml MonoS or MonoQ column (GE Lifesciences) which was already equilibrated with Buffer A25, using a 50-ml SuperLoop (GE Lifesciences). After loading, elution was done using a linear gradient of Buffer A25 (50 mM Tris pH 8.0, 25 mM NaCl)/H25 (50 mM HEPES pH 6.5, 25 mM NaCl) and A1000 (50 mM Tris pH 8.0, 1 M NaCl)/H1000 (50 mM HEPES pH 6.5, 1 M NaCl) till 100% of buffer A1000/H1000, over 20 column volumes. If protein was bound and eluted at a particular percentage of elution buffer (particular ionic strength depending on the affinity), then it could be observed from the UV peak, and further the fractions were checked on SDS-PAGE.

2.4.3. Gel filtration Chromatography

Gel filtration chromatography was done in both AKTA Prime and AKTA Purifier. Gel filtration chromatography was done as a part of the protein purification as the last step to remove imidazole, and also as an analytical tool for estimating the oligomeric status of SofG. The protocol is as follows. The 24-ml Superdex 75 (GE Healthcare Lifesciences) was equilibrated with Buffer A50 (50 mM Tris pH 8.0, 50 mM NaCl). Then the concentrated protein was loaded using a 500 µl loop. After loading,

fractions were collected and when protein of our interest gets eluted, we could see the peak of UV.

Table 2.5: Molecular markers in analytical gel filtration and their parameters

Marker	Molecular Weight (kDa)	Volume of elution (V_e)	V_e/V_o	log molecular weight
Blue dextran	2000	8.5	1	-
Carbonic anhydrase	29	12.04	1.416471	1.462398
Albumin	66	9.83	1.156471	1.819544
Cytochrome C	12.4	15.8	1.858824	1.093422

For determining oligomeric state of SG-60, different standard molecular markers (Table 2.5) were loaded to Superdex 75 column (Figure 2.4). Then the standard graph was plotted for determining the size of SG-60 (Figure 2.5).

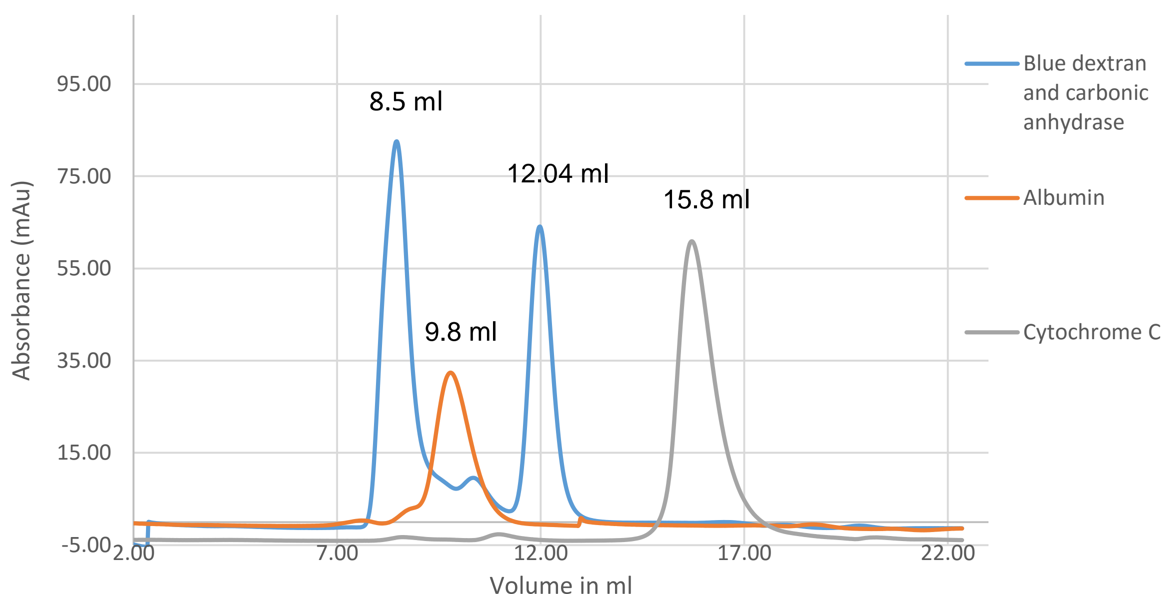


Figure 2.4: Gel filtration run profile of markers. The elution volumes are 8.5 ml (blue dextran), 9.8ml (albumin) and 12.04 ml carbonic anhydrase 15.8 ml (cytochrome C).

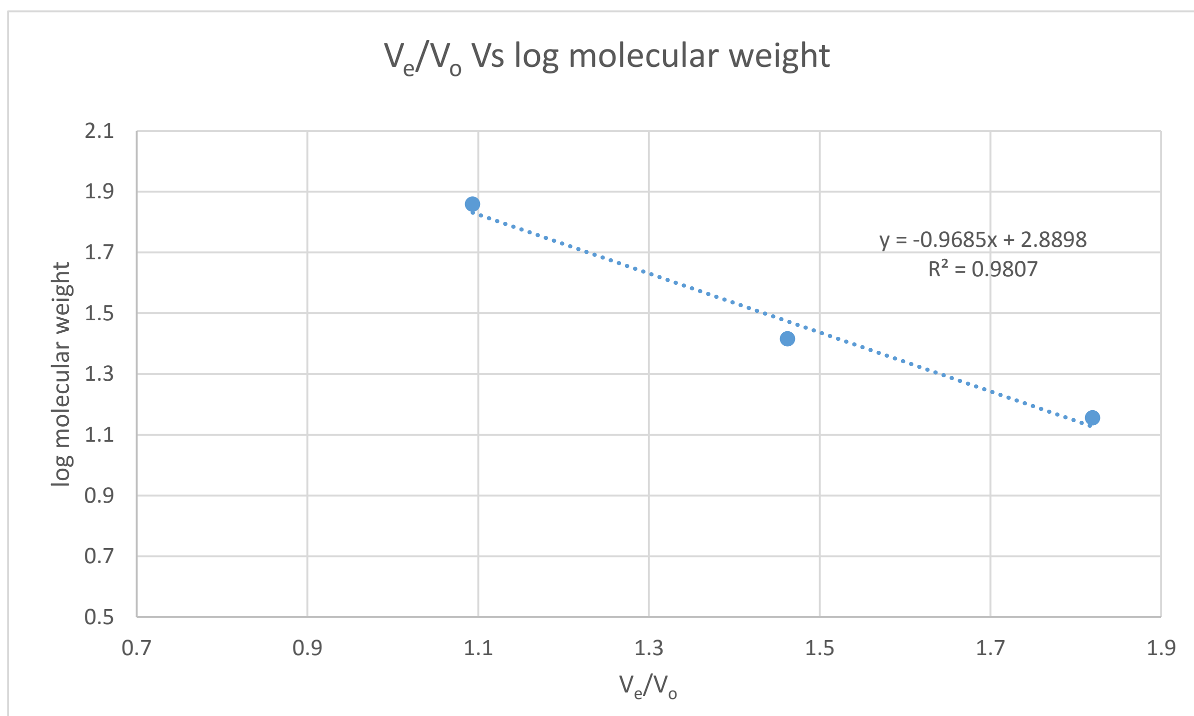


Figure 2.5: Standard graph obtained from different markers

2.5. Malachite green assay

Malachite green assay was used for measuring free inorganic phosphate (iP) in a reaction, since GTPase reaction involves conversion of GTP to GDP and phosphate. The principle behind the assay is that inorganic phosphate released in the reaction reacts with ammonium molybdate and forms a yellow-coloured complex. This complex reacts with malachite green to form a green-coloured complex. This can be measured at 630 nm and the amount of iP can be quantified (Geladopoulos et al., 1991). Reaction mix was prepared with buffer (50 mM Tris pH 8.0, 50 mM NaCl and 5 mM $MgCl_2$), 0.5 mM GTP, MilliQ water and protein, and incubated at 30°C for different time points and stopped by heating at 65°C for 10 minutes. Buffer only, buffer containing 0.5 mM GTP and buffer with protein were used as controls to account for the inorganic phosphate, if any, which was contained in the reaction mixtures.

Volume of reaction for each time point was 30µl. All the reactions were done in triplicates. 25 µl of each reaction was taken in a 96 well-plate, and 50 µl of Malachite green solution (containing working concentration of 2.68 mM Malachite green, 1.5%

ammonium molybdate and 0.176% tween 20) was added to each time point of reaction and control. After addition of malachite green solution, it was incubated for 10 minutes and absorbance was measured at 630 nm using plate reader (Varioskan™). Different concentrations (ranging from 20 μM to 200 μM of sodium phosphate solution) was used for plotting standard phosphate control plot (Figure 2.6).

Calculations

Blank was subtracted from each of the standards and samples. The standard curve was plotted by phosphate in nano moles vs absorbance at 630 nm. Using the value of slope, amount of iP in the samples were calculated.

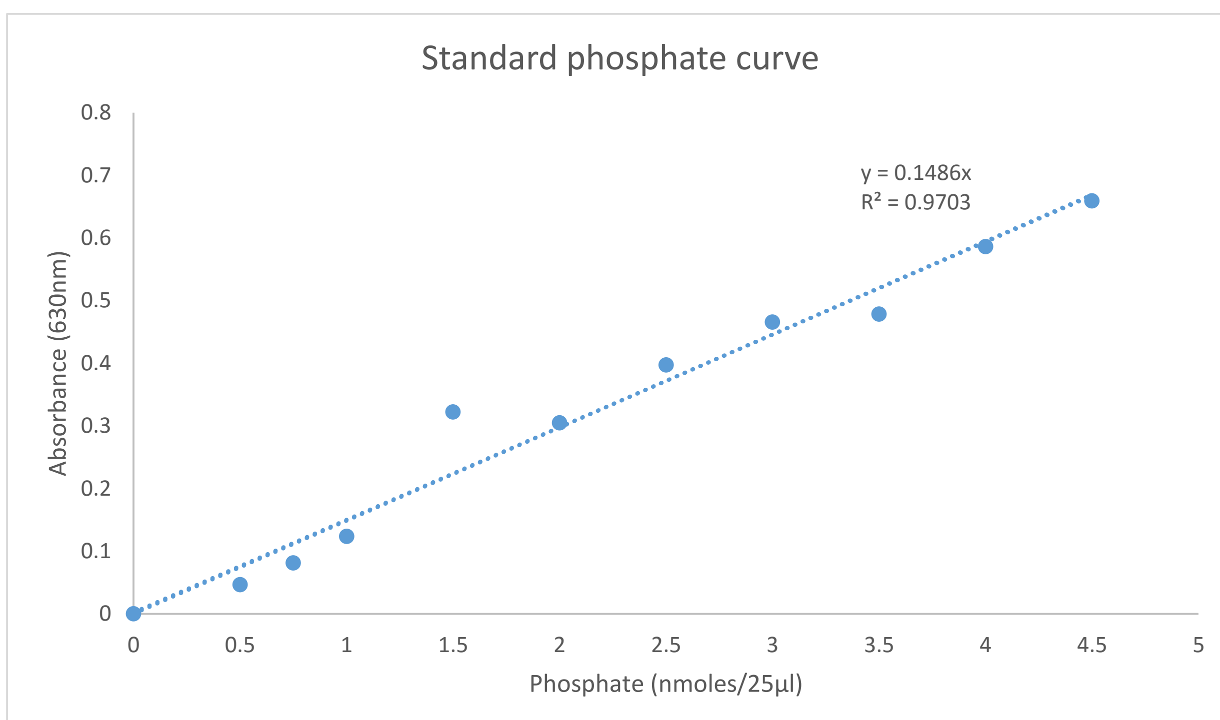


Figure 2.6: Standard curve for phosphate concentration estimation using Malachite green assay.

2.6. Crystallisation trials of SG-60

For crystallisation of protein, protein concentrations in the range of 5 – 10 mg/ml is required. At the same time, the quality of the protein is also important. The protein should be pure and homogenous. Crystal formation is facilitated usually when the protein sample is pure and homogeneous. In order to identify the combination of

salts, precipitating agent and pH at which crystallization occurs, initial screening was done using commercial available screens. Five different types of screens (each containing 96 different conditions) were set up using the sitting drop method for vapour diffusion. The screens are:

1. Molecular dimensions Structure Screen 1 & 2 HT-96 (MD1-30)
2. Molecular dimensions Crystal Strategy Screen I (MD1-31)
3. Molecular dimensions Crystal Strategy Screen II (MD1-32)
4. Jena Bioscience Screen Classic HTS II
5. Hampton Research PEG/Ion HT (HR2-139)

The concentration of protein used was 4.1 mg/ml. Screens were set up in 96 well-plates using a nano-litre pipetting robot known as Mosquito® (TTP Labtech). The drop ratio was 1:1 of 100 nl of protein and 100 nl of condition. The plates were covered using clear sealing tape and kept at 18°C. Screens were checked using stereo microscope (Leica S8 APO). The initial crystallization hits obtained in the screen were optimised further by varying conditions from hits during screening. For optimisation screens, 48 well-plate was used by varying drop ratio to 1:2 of 1 µl protein and 0.5 µl of condition. The concentration of protein used for optimisation screen was 4.5 mg/ml.

2.7. Binding assay using mant-labeled nucleotides

Affinity of nucleotides to the protein was performed to check the binding of mant-GDP and mant-GTP using Fluoromax-4 spectrofluorometer (Horiba Scientific). Mant is a fluorophore and its fluorescence intensity increases when bound to the protein. Changes in fluorescence signal can be measured and used to estimate exchange of nucleotide, serving as an indicator of binding affinity of guanine nucleotide to a protein (Remmers et al., 1990). The fluorescence of mant-GDP (0.4 µM) and mant-GTP (0.4 µM) was measured by exciting it at 360nm and measuring the emission at 440 nm wavelength. This assay was done using buffer A50 (50 mM Tris pH 8.0, 50mM NaCl) with and without addition of 2 mM Mg^{2+}. The volume of the reaction was 200 µl. The concentration of protein used was 4 µM, 8 µM for mant-GDP and 4 µM, 10 µM for mant-GTP.

2.8. Circular dichroism spectroscopy

Circular Dichroism (CD) spectroscopy is a method used for the analysis of secondary structure content of a protein. It measures differences in the absorption of left and right circularly polarized light which arise due to structural asymmetry. CD measurements were done using Jasco j-815 CD spectrometer. The buffer used contained 10 mM potassium phosphate and 100 mM potassium fluoride. The potassium phosphate buffer was prepared by mixing KH_2PO_4 and K_2HPO_4 at appropriate ratio. They were mixed in a ratio of 0.6:9.4 respectively so that exact pH 8.0 was obtained. The wavelength was varied between 185 nm to 260 nm. First the measurement for buffer was taken, then the spectrum of protein was measured.

2.9. High Performance Liquid Chromatography (HPLC)

HPLC is a type of liquid chromatography used to separate, identify and quantify compounds dissolved in a solution. It was done to identify whether the protein was bound with nucleotide or not and also to identify the type of nucleotide and quantify the fraction of molecules with bound nucleotide. Concentration of protein used was 15 µM. Protein was heated at 65 °C for 10 minutes, and the supernatant obtained on centrifuging was injected onto a DNAPac PA100 column (Dionex). Volume of samples loaded was 35 µl. DNAPac is a type of anion exchange column which can resolve nucleotides even at a resolution of one base. In this experiment, it was used to resolve GDP and GTP. Buffers used were: a) 2 mM Tris pH 8.0 and b) 2 mM Tris pH 8.0 with 1.25 M NaCl.

3. Results and Discussion

The chapter describes and discusses the results obtained including cloning, protein purification standardisation, biochemical characterisation and crystallization.

3.1. Cloning

All cloning experiments were done by using NEB® Turbo strain of *Escherichia coli* using pHis17 vector. Clones are verified by restriction digestion using NdeI and BamHI. Then confirmation was done by sequencing.

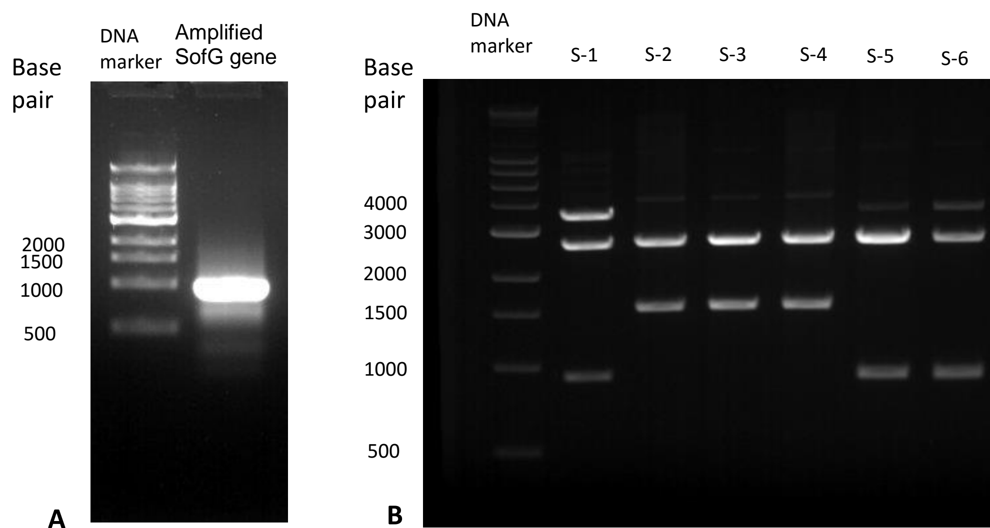


Figure 3.1: (A) Amplification of SofG by PCR. (B) Clone verification of SG-H6. Sample numbers S-1, S-5 and S-6 were positive clones.

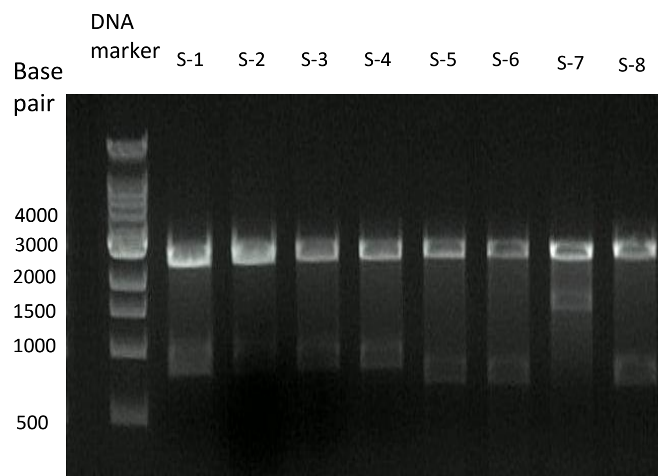


Figure 3.2: Clone verification of SG-18 and SG-60. S-1, S-2, S-3, S-4 were positive clones for SG-18 and S-5, S-6, S-8 were positive clones for SG-60.

The figures show PCR amplification of SofG (Figure 3.1 A) and clone confirmation by double digestion of NdeI and BamHI of SG-H6 (Figure 3.1 B; refer Table 2.1 for nomenclature of the constructs), SG-18 and SG-60 (Figure 3.2) and SG-277 and SG-60-277 (Figure 3.3). Cloning experiments of SG-258 and SG-60-258 were tried various times but were not successful. Subsequently, I went ahead with expression, standardisation and purification for the constructs, SG-H6, SG18, SG-60, SG-277 and SG-60-277.

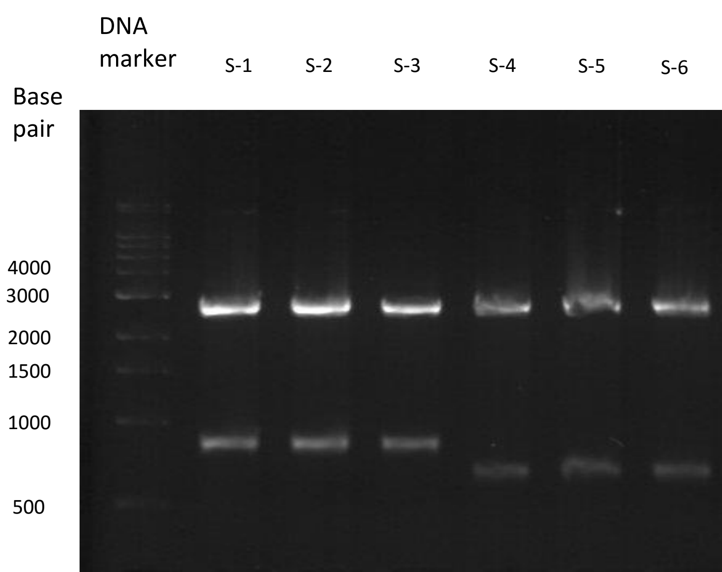


Figure 3.3: Clone verification of SG-277 and SG-60-277. S-1, S-2 and S-3 were positive clones for SG-277. S-4, S-5, S-6 were positive clones for SG-60-277.

3.2. Optimisation of expression and solubility of constructs

Different strains of *E. coli* and different conditions were tested for the optimal expression and solubility of the protein of interest. Conditions attempted for overexpression optimisation are listed below:

1. Different expression strains of *E.coli*: BL21AI, BL21DE3, Rosetta, C41 and C43.
2. Induction $O.D_{600}$ of culture were checked at 0.4, 0.6 and 0.8.
3. Temperatures of growing cultures after induction were checked at 30°C, 25°C and 18°C.

SG-H6, SG-U, SG-18 did not overexpress at all (data not shown). SG-60, SG-277 and SG-60-277 showed reasonable overexpression in BL21AI strain of *E.coli*, however the amount of protein in the soluble fraction was negligible even at an

induction temperature of 18°C. The best construct in terms of expression and solubility was observed to be SG-60, and further purification trials were carried out with this construct. Overexpression and solubility profiles of different constructs are shown in Figures 3.4, 3.5 and 3.6.

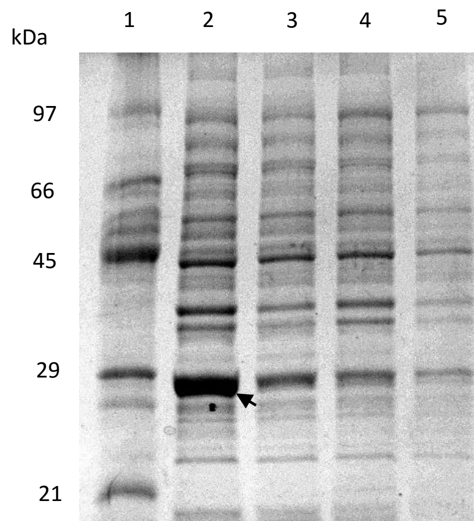


Figure 3.4: Expression check of SG-60. The arrow points to 27.5 kDa SG-60. Lanes: 1 – Marker; 2 - Total cell lysate (Induced); 3 - Soluble fraction of lysate (Induced); 4 - Total cell lysate (Uninduced); 5 - Soluble fraction of lysate (Uninduced).

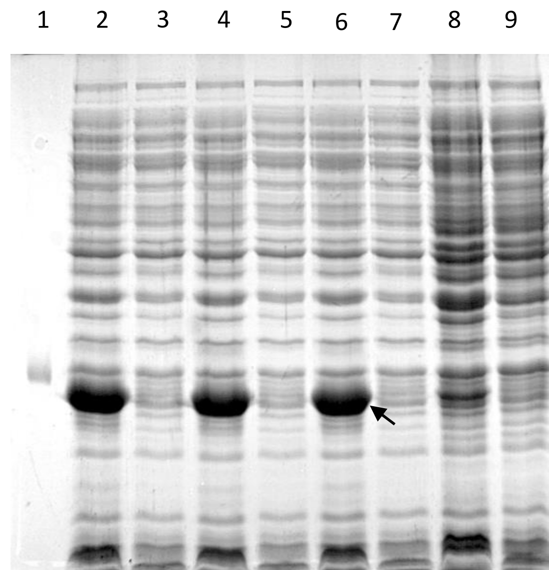


Figure 3.5: Expression check of SG-60-277. The arrow points to 25 kDa SG-60-277. Lanes: 1 - Marker (27.5 kDa); 2, 4, 6 - total cell lysate for samples induced at OD_{600} at 0.4, 0.6, 0.8; 3, 5, 7 - soluble fraction of the lysate for samples induced at OD_{600} at 0.4, 0.6, 0.8; 8 - Total cell lysate (Uninduced); 9 - Soluble fraction of lysate (Uninduced)

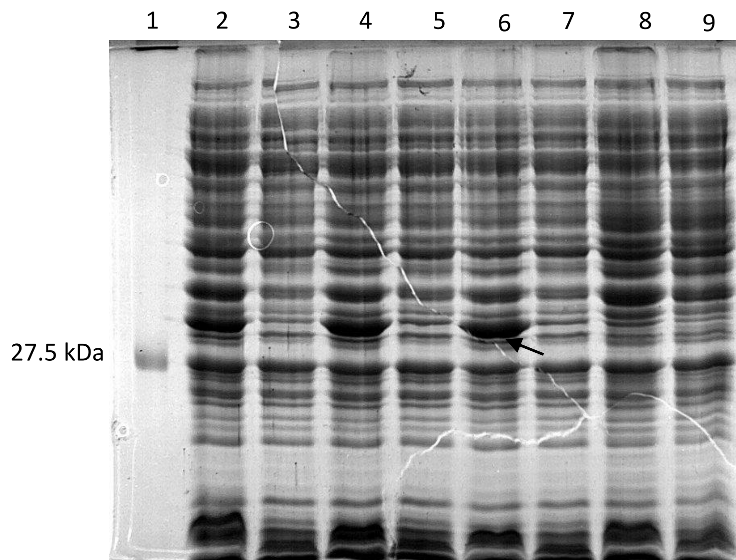


Figure 3.6: Expression check of SG-277. The arrow points to 31.5 kDa SG-60-277. Lanes: 1 - Marker (27.5 kDa); 2, 4, 6 - total cell lysate for samples induced at OD_{600} at 0.4, 0.6, 0.8; 3, 5, 7 - soluble fraction of the lysate for samples induced at OD_{600} at 0.4, 0.6, 0.8; 8 - Total cell lysate (Uninduced); 9 - Soluble fraction of lysate (Uninduced).

The overexpression temperature of SG-60 was further optimised and the highest percentage in soluble fraction was observed on growing at 15°C (Figure 3.7). So, the best expression and solubility condition was obtained for SG-60 in BL21AI strain with induction by 0.2% final arabinose and growing at 15°C for 16 hours. The protein expression and solubility status of various constructs are summarised below in Table 3.1.

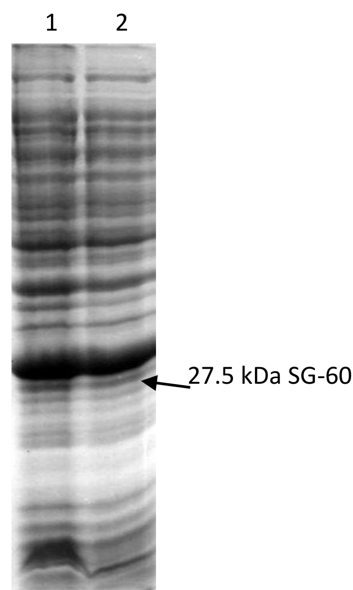


Figure 3.7: Expression check of SG-60 at 15°C. Lanes: 1 - Total cell lysate, 2 - Soluble fraction of lysate of sample incubated at 15 °C after induction.

Table 3.1: Expression and solubility status of various constructs of *M. xanthus* SofG.

Construct description	Construct name	Cloning status	Expression	Expression strain	Solubility
Full length untagged	SG-U	✓	No expression	-	-
C-terminal $(His)_6$ full length	SG-H6	✓	No expression	-	-
18 to 298 $(His)_6$	SG-18	✓	No expression	-	-
60 to 298 $(His)_6$	SG-60	✓	Expressed	BL21AI	Soluble
1 to 258 $(His)_6$	SG-258	×	-	-	-
60 to 258 $(His)_6$	SG-60-258	×	-	-	-
1 to 277 $(His)_6$	SG-277	✓	Expressed	BL21AI	Insoluble
60 to 277 $(His)_6$	SG-60-277	✓	Expressed	BL21AI	Insoluble

3.3. Standardization of protein purification

Crystallization of a protein for structure determination using X-ray crystallography requires high concentrations (about 10 mg/ml) of the protein and in a pure of homogenous state. Different purification trials were carried out to find out the suitable stable condition and optimising the yield of the protein. In order to standardise the purification, there is need of carry out the experiments by varying different conditions like pH, salt, additives etc.

As the construct SG-60 is C-terminal hexa-histidine tagged, it can be purified by using affinity chromatography using Ni-NTA column at pH 8.0. Figure 3.8 shows a representative protein purification profile of SG-60 from a six litre culture on SDS PAGE. The protein binds well to the 5-ml Ni-NTA column as observed from the absence of the protein band in the flow-through from the column (Figure 3.8, Lane 3). Initial rounds of wash using buffer with lower concentration of imidazole such as 10 mM, 25 mM and 50 mM removed many of the non-specifically bound proteins. SG-60 elutes at 250 mM of imidazole. Though the fractions eluted contained pure protein, the main challenge was precipitation of the protein. Precipitation starts immediately after elution and most of the protein got precipitated. So, in order to decrease the amount of precipitation and get more amount of protein, various other

strategies were tried by varying from standard condition of Tris pH 8.0 and 200 mM NaCl.

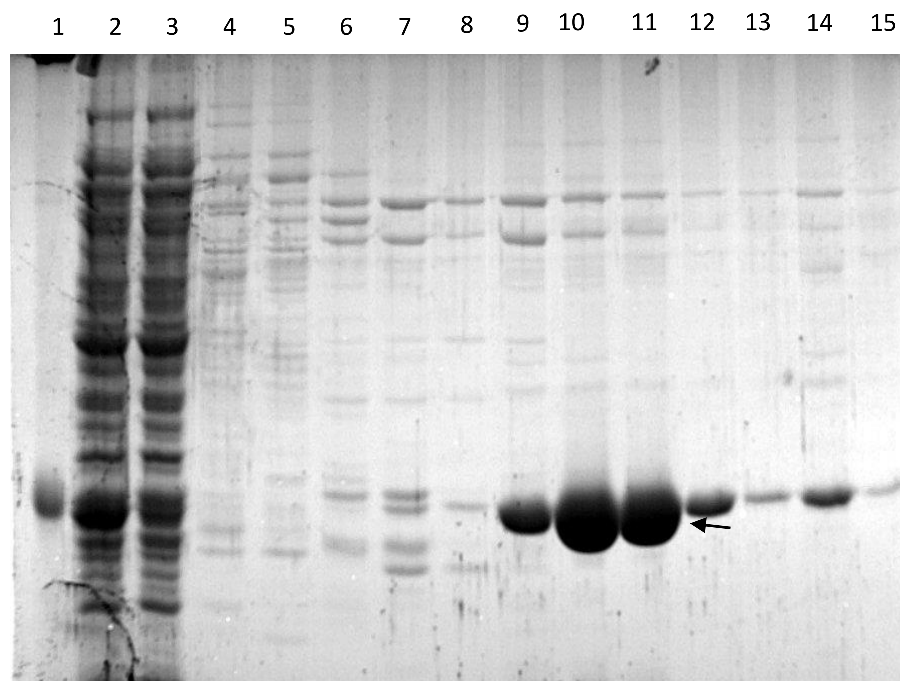


Figure 3.8: Purification profile for Ni-NTA affinity purification of SG-60. The arrow shows the 27.5 kDa band corresponding to purified SG-60. Lanes: 1 - Marker (27.5 kDa); 2 – Soluble fraction of cell lysate (load to Ni-NTA column); 3 – Flow-through of Ni-NTA column; 4, 5, 6 - 2% B, 5% B, 10% B washes respectively; 7, 8 - 20% B fractions; 9, 10, 11, 12 - 50% B fractions; 13, 14, 15 - 100% B fractions (refer Materials and Methods for constituents of Buffer B).

Different modifications in the basic purification protocol, which were attempted are listed below:

1. Increase in salt (NaCl) concentration in lysis buffer:
   Varying the salt concentration might help to increase the solubility of the protein and reduce precipitation. So, NaCl concentration in the lysis buffer was changed from 200 mM to 500 mM. But this does not give any positive result.
2. Addition of EDTA and DTT in fraction collection tube:
   Precipitation of protein after elution from Ni-NTA could be due to aggregation caused by multiple hexa-histidine tagged proteins binding to leached Ni^{2+} ions or due to disulfide bonds through exposed cysteines. Hence, addition of EDTA and DTT immediately after elution might help reduce precipitation.

However, addition of EDTA and DTT did not help in reducing the precipitation, in the case of SG-60.

3. Ni-NTA chromatography followed by gel filtration chromatography:
   Though the precipitation problem was not sorted out, I tried to proceed with the amount of protein left in solution after removing the precipitated protein. For reducing the volume of the sample for injection into a gel filtration column, concentration of the fractions from Ni-NTA was carried out using centricons (Millipore). But the precipitation during concentration blocked the centricon. Due to this problem, other fast concentrating methods were tried like transferring the protein to a dialysis bag and addition of sucrose or PEG 20K on top of dialysis bag. This helps in concentrating very fast but the protein still precipitated. Then the remaining amount of concentrated protein was loaded to gel filtration column i.e. Superdex 200 (GE Lifesciences). But protein was observed in void fraction. It might be due to aggregation of protein as it is not stable at that condition. So, other options such as ion exchange chromatography and ammonium sulphate precipitation were important to try out.
4. Ni-NTA chromatography followed by ion exchange chromatography:
   The observation of SG-60 in the void in gel filtration suggested that the protein could be an aggregate. Also, the peak had a higher absorbance at 260 nm than 280 nm, indicating that the protein could be non-specifically bound to DNA. Hence, passing the Ni-NTA fractions through an ion exchange column could help in removing the bound DNA. Also, use of specialised and highly efficient ion exchange columns such as MonoS results in the elution of the protein into a very small volume, thereby reducing the amount of time spent in concentrating the protein, and improving quality and quantity.

   The theoretical isoelectric point (pI) of SG-60 is 9.7 (as predicted by Protparam (http://web.expasy.org/protparam/) and if the pH of buffer is 8.0, then the protein should bind to cation exchange column and could be eluted by gradient of high salt. So, trying MonoS (GE LifeSciences) was a good option. The SG-60 fractions were loaded onto MonoS column after dialysing into low salt concentration (25 mM NaCl) buffer. However, the protein did not bind to MonoS (Figure 3.9). This could be because the exposed surface of the

protein might not be positively charged, and there could be negatively charged patches that allow binding to MonoQ, an anion exchange column. This was unsuccessful as the protein did not bind to MonoQ also.

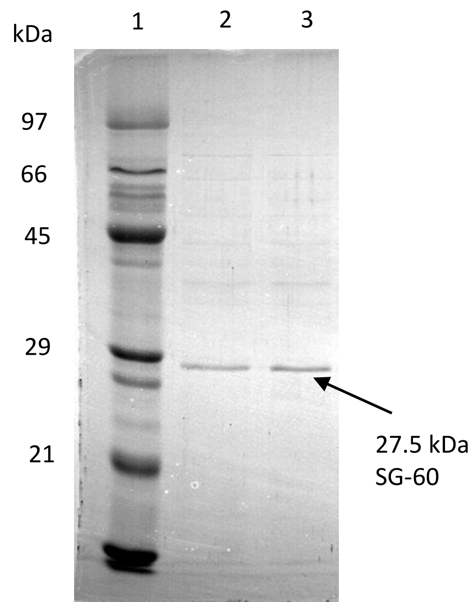


Figure 3.9: Run profile of ion exchange chromatography using MonoS. Lanes: 1- Marker; 2- Load of MonoS; 3- Flow-through

Varying pH of buffer might help to stabilise the protein and reduce precipitation. And also going far away from pI of protein might help in binding to MonoS column. So, instead of Tris pH 8.0, it was changed by HEPES pH 6.5. But still the protein did not bind to MonoS column and came out in flow-through.

5. Ni-NTA followed by Ammonium sulphate precipitation:

   Ammonium sulphate precipitation will serve to remove non-specifically bound DNA, and also facilitates concentration of the protein. Ammonium sulphate precipitated SG-60 at 50 % saturation, but the protein could not be resolubilised from the precipitate. So, ammonium sulphate precipitation did not help in purification.

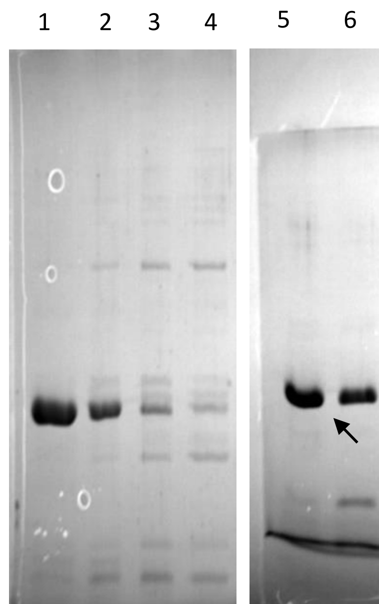


Figure 3.10: Run profile of size exclusion chromatography using Superdex 75. The arrow denotes band corresponding to 27.5 kDa SG-60. Lanes: 1- Marker (27.5 kDa); 2- Load of Superdex 75 (2 µl); 3, 4- Void fractions; 5, 6 - Monomeric fractions (10 µl)

6. Addition of CHAPS in lysis buffer:

   CHAPS is non denaturing zwitterionic detergent for solubilizing membrane proteins and also helps in breaking protein-protein interactions. And it generally does not inactivate protein function. Final amount of 0.04% CHAPS was added to lysis buffer. This improved the quality of protein, as judged by the presence of a monomeric peak of SofG in gel filtration. The gel filtration run results in two peaks, one in void and one closer to monomeric SofG (low peak height – Figure 3.10). The void peak had high A260/A280 ratio (4.5). So, addition of DNAse might help in reducing void.

7. CHAPS in all buffers and DNAse in lysis buffer, followed by Superdex 75:

   Elution profile of gel filtration trials following Ni-NTA affinity chromatography showed that the void peak vanished and only monomer peak was observed, but the peak height was still very low as the precipitation was huge. So, the protein yield (120 µl of 1 mg/ml concentration from 6-litre culture) was not enough to proceed for further crystallization experiments. Thus, there was still need of optimisation.

8. Addition of GDP in fraction collection tube of Ni-NTA followed by gel filtration: Instability of the protein might be due to nucleotide-free form after elution from Ni-NTA column. Many small Ras-like GTPases are known to be unstable in the absence of bound nucleotide (Balaji et al., 2000). So, if the protein does not come in nucleotide-bound state from the cell and if the purification procedure results in loss of the bound GDP, addition of extra GDP to the elution fractions should stabilise the protein. Thus, 0.5 mM final concentration of GDP was added to fraction collection tube. After addition of GDP, precipitation stopped. The sample was loaded to Superdex 75 gel filtration column and the elution profile showed a single peak of 190 mAu height, at an elution volume corresponding to monomeric molecular weight (refer section 3.5 for standard curve and details of size estimation). 0.1 mM final concentration of GDP was also added to fractions of Superdex 75. The quality of protein improved a lot as observed from the reduced precipitation. The requirement of CHAPS and DNAse for solubilising the GDP-bound SG-60 had to be confirmed.

9. Absence of DNAse and CHAPS in the lysis buffer and GDP addition to the Ni-NTA fractions:
   CHAPS was completely removed from the purification protocols. Equimolar amount of GDP and Mg^{2+} was added to fractions of Ni-NTA and Superdex 75 (0.5 mM in Ni-NTA fractions and 0.1 mM in Superdex 75 fractions). After removal of CHAPS, the quality of protein remains the same, so CHAPS was removed from the protocol. This protocol yielded about 120 μl of 4.1 mg/ml protein from six litre culture and was used for setting up of crystallisation trials. The SDS-PAGE profile of Superdex 75 fractions and input to Superdex 75 column is shown in Figure 3.11.

10. GDP and Mg^{2+} in Ni-NTA fractions but not in Superdex 75 fractions:
    Since there was excess GDP in Superdex 75 fractions, it might create problems in performing biochemical assays. So, there was need of purifying protein without excess GDP. So, excess GDP was not added in Superdex 75 fractions. This led to slight precipitation but yielded about 2.7 mg/ml protein of 100 μl from six-litre culture. This precipitation shows that the binding of GDP

might be of low affinity, leading to loss of bound GDP and precipitation. This protein without excess GDP was used for biochemical assays.

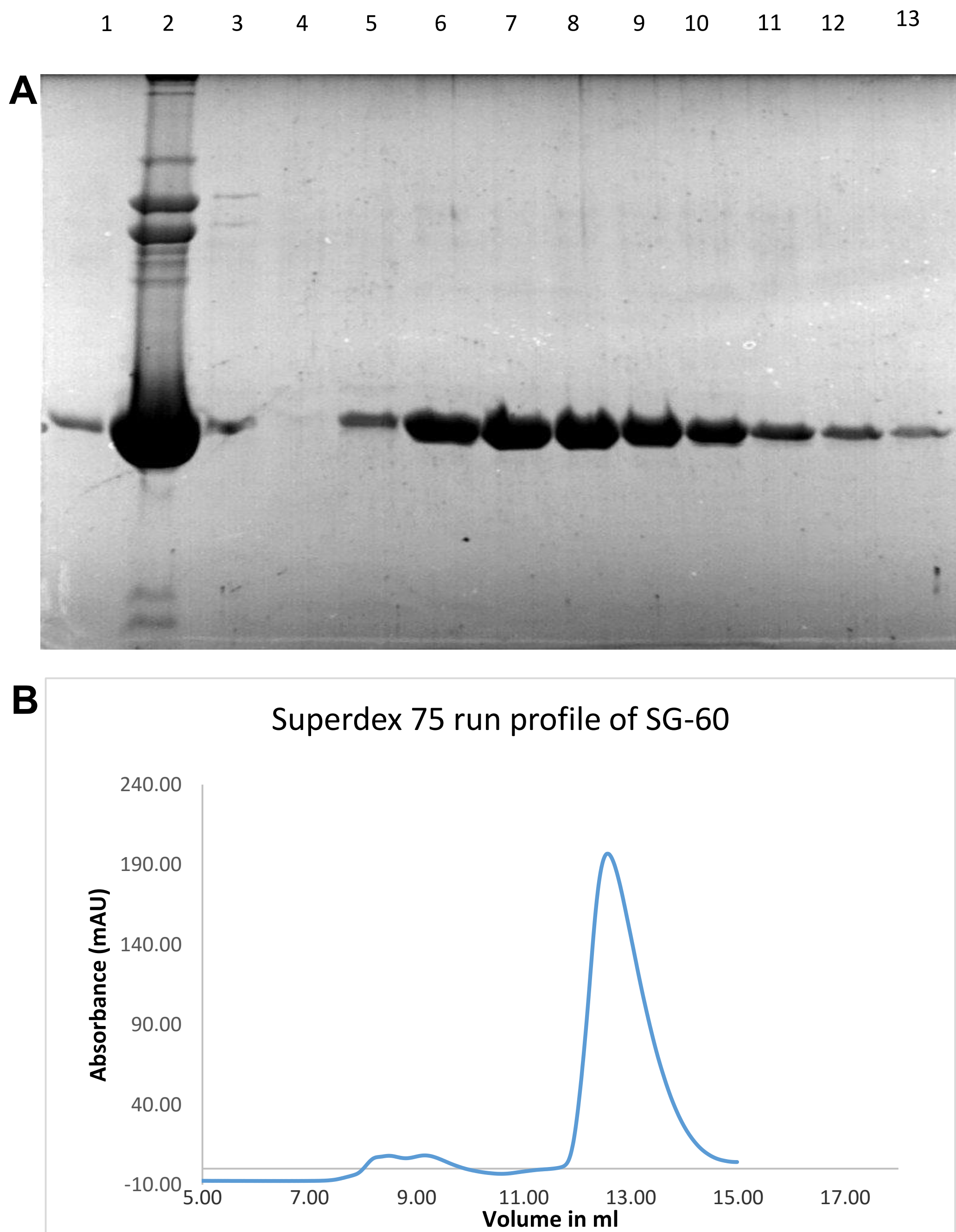


Figure 3.11: Superdex 75 run profile of SG-60 after addition of GDP. A) Fractions run on SDS-PAGE - Lanes: 1 - Marker (27.5 kDa); 2 - Load of Superdex 75 (2.5 µl); 3- Flow-through of centricon; 4-13 - Monomeric fractions (5 µl). B) Corresponding chromatography profile

3.4. Biochemical assays

GTPase assay: Malachite green assay

Malachite green assay is a method for measuring inorganic free phosphate (iP) released in a reaction. We can estimate if our protein is active and can hydrolyse GTP to GDP and iP from this assay. This assay was performed in two different concentrations of protein i.e. 2 μM and 10 μM, in which both were done in triplicates so that an accurate measurement can be obtained. Necessary controls were also included for ruling out phosphate contamination from the buffer, purified protein, and only GTP.

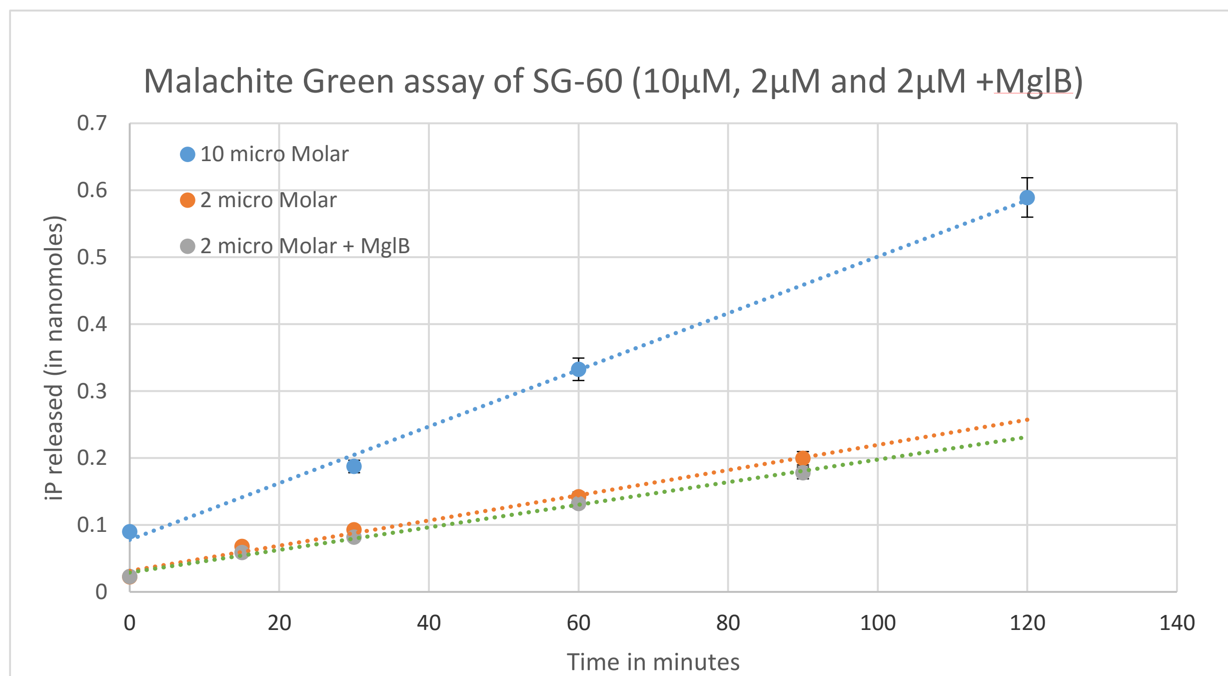


Figure 3.12: Comparison of iP released by Malachite green assay of SG-60 at 10 μM, 2 μM concentrations and 2 μM SofG + 4 μM MglB.

If we compare two plots of different protein concentrations i.e. 2 μM and 10 μM, we saw that the amount of iP released showed an increase in 10 μM compared to 2 μM (Figure 3.12). So, we can conclude that the protein that we purified is active and can act as a GTPase. The amount of iP released is in nanomoles. The low activity could be because the reaction conditions are not optimal and has to be standardised further. A fraction of the protein could also be misfolded. Low activity is also possible due to the truncated version of SofG. Also, SofG activity may be stimulated by a GTPase activating protein (GAP), which has not been identified yet.

SofG has around 45% sequence identity with MglA, another GTPase in *Myxococcus xanthus* motility. MglB has been shown to act as a GAP for MglA. The residues of MglA that help in MglB interaction are conserved in SofG (Figure 1.7). Hence it is possible that MglB could act as a GAP for MglA. GTPase assay was also done in the presence of wild type MglB. However, there was no increase in hydrolysis of GTP after addition of MglB (Figure 3.12).

Further experiments have to be carried out to confirm that SofG-60 is properly folded and is also capable of binding GTP efficiently.

3.5. Size exclusion chromatography: Monomeric state of SG-60

The elution state of SG-60 could be find out from size exclusion chromatography by comparing with markers of known molecular weights. The elution volume was compared to that of markers (Figure 3.13). The ratio of elution volume (V_e) to void volume (V_o) of different markers were calculated (refer Table 2.5). The volume of elution of blue dextran i.e. void was at 8.5 ml. Based on the equation derived from the curve, the oligomeric status of SG-60 in solution was estimated to be monomeric, corresponding to 27.6 kDa, for an elution volume and V_e/V_o for SG-60 12.7 ml and 1.4941 (=12.7/8.5), respectively.

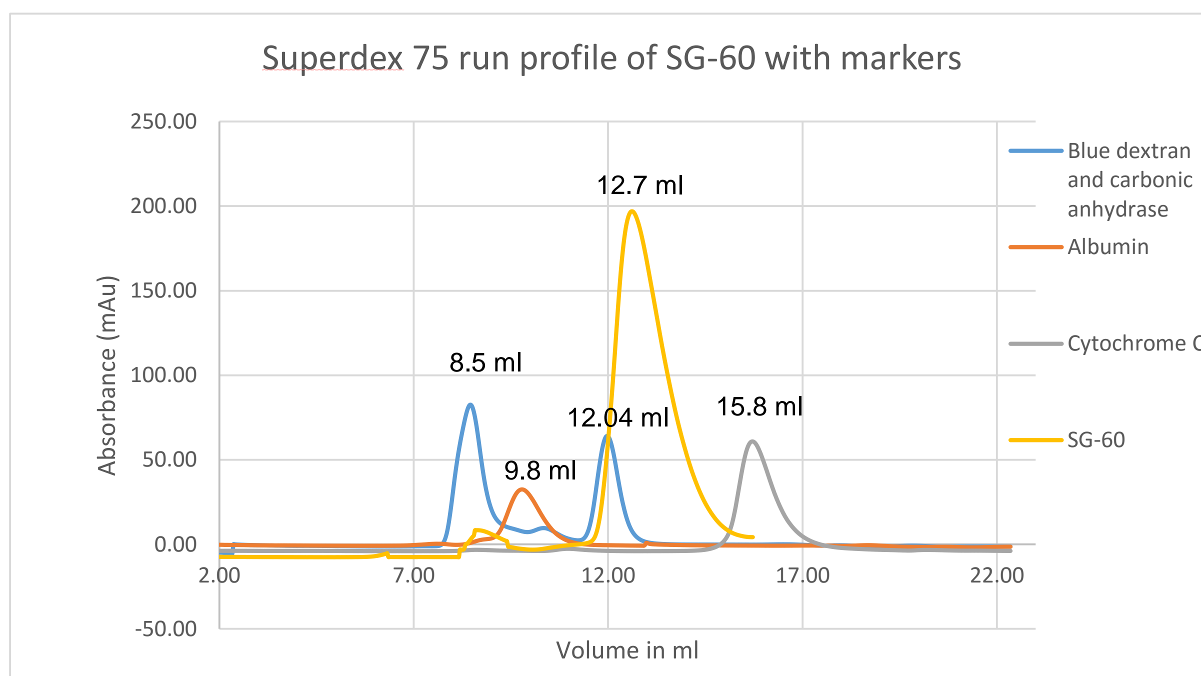


Figure 3.13: Superdex 75 run profile of SG-60 with plot of markers. The elution volumes are 8.5 ml (blue dextran), 9.8ml (albumin), 12.04 ml carbonic anhydrase), 15.8 ml (cytochrome C) and 12.7 ml SG-60.

3.6. Crystallisation trials

Crystallisation trials were set up using purified SG-60. Needle-like clusters of crystals were seen in different conditions, as shown in Figure 3.14. And also small crystals were seen in some conditions.

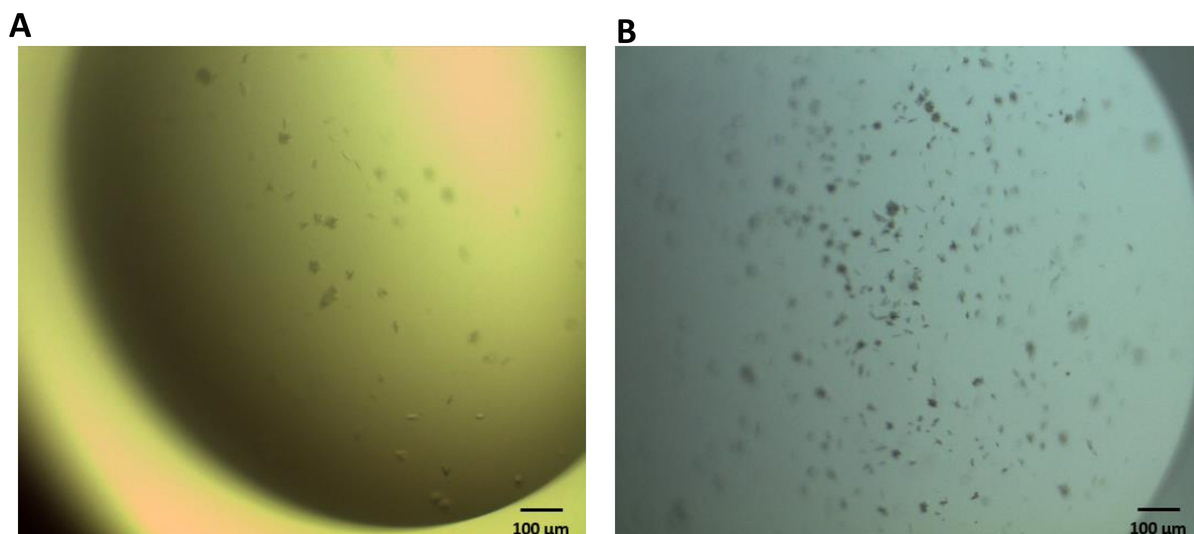


Figure 3.14: Needles and small crystals from screening. (A). Condition: 0.1 M MES pH 6.5, 12% w/v PEG 20k. (B) Condition: 1.6 M ammonium sulfate, 2% w/v PEG 1000, 100 mM HEPES sodium salt pH 7.5

Since hits were seen in different variety of conditions, it is possible that those hits were from SG-60, not from the crystallisation condition. So, these needles and small crystals can be improved to grow to bigger crystals. Thus, setting up of optimisation screen was needed. After examining screened conditions where hits were seen, conditions of optimisation screen were decided in such a way that compound that were found common in most of the hits were selected. Variations of concentration of those compounds were used in setting up of optimisation screen. Further optimisation with higher concentration of protein has to be set up.

3.7. Binding assay using mant-labeled nucleotides

Fluorescence-based assay for nucleotide binding was first done with mant-GDP addition to 4 µM of protein. However, no exchange was detected. After that, concentration of protein was increased to 8 µM. But, no exchange was still detected (Figure 3.15). Then the same was done with mant-GTP with and without Mg^{2+}. But in both cases there was still no conclusive raise in peak.

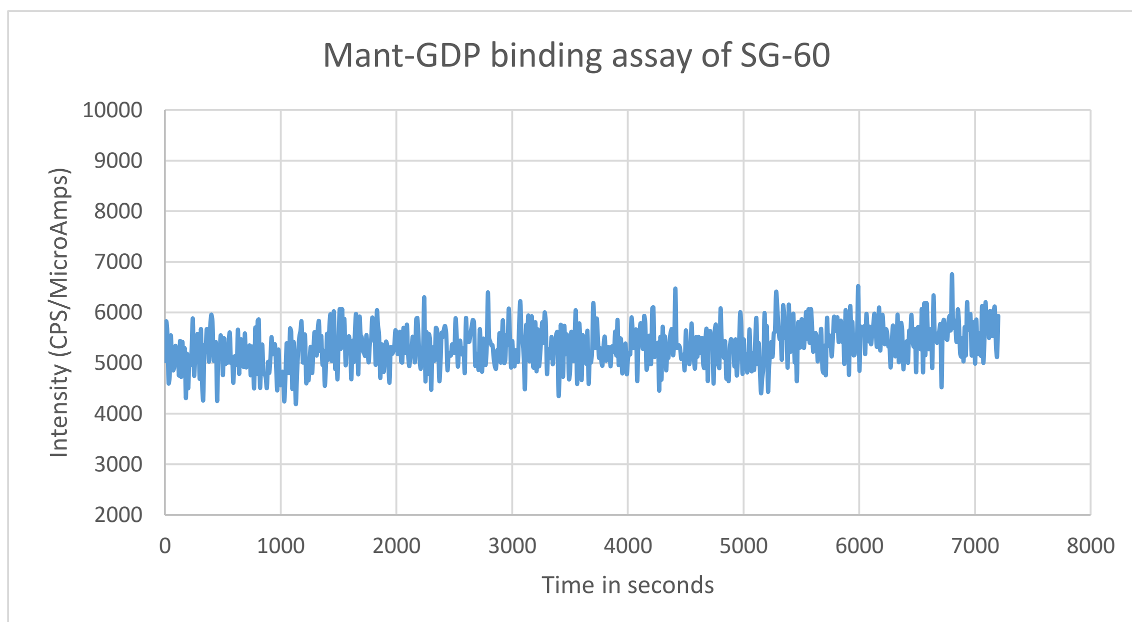


Figure 3.15: Mant-GDP binding assay of SG-60

From these results, there was a suspicion whether the protein was well-folded or not. But, the Malachite green GTPase assay showed activity of the protein. Activity won't be possible if the protein was denatured. The monomeric peak in gel filtration is also an indication of a folded protein. So, to confirm the formation of secondary structures, Circular Dichroism (CD) was an option. The binding experiments also need to be repeated by varying the concentration of protein and concentration of mant-GDP and mant-GTP.

3.8. Circular Dichroism (CD) Spectroscopy

Circular dichroism showed a mixed spectrum of different secondary structures. The main characteristic is the alpha helix. Prediction of secondary structure was done using online software K2D3 (k2d3.ogic.ca). It showed 40.35% as α helix and 11.87% as β strand. This showed that the purified protein i.e. SG-60 was not unfolded. It indeed formed secondary structure. The spectra showed characteristics of α helices, with minima at approximately 207 and 222 nm (Figure 3.16). This was actually a positive sign that we can carry on with the crystallisation trials and also with assays.

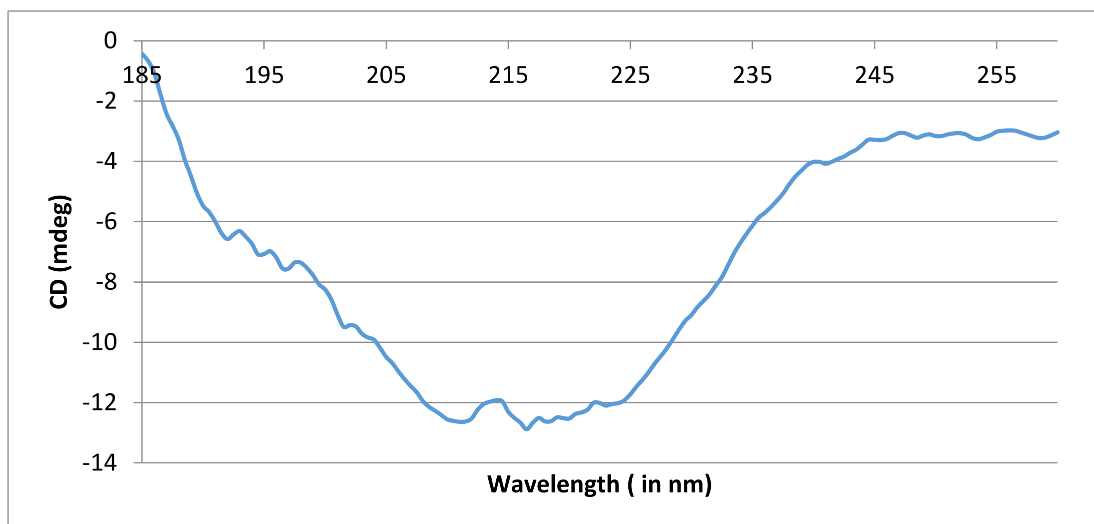


Figure 3.16: CD spectrum of SG-60

3.9. High Performance Liquid Chromatography (HPLC):

HPLC run of SG-60 in order to detect bound nucleotide species did not give highly conclusive result. Since the purification protocol was optimised by addition of excess GDP, it was expected that the protein should be in the GDP-bound state. The result of the HPLC showed neither GDP nor GTP bound with the protein, which meant that the protein was in nucleotide-free form. The run of buffer and the protein looked similar (Figure 3.18). This experiment need to be repeated to confirm that the protein was actually nucleotide-free form.

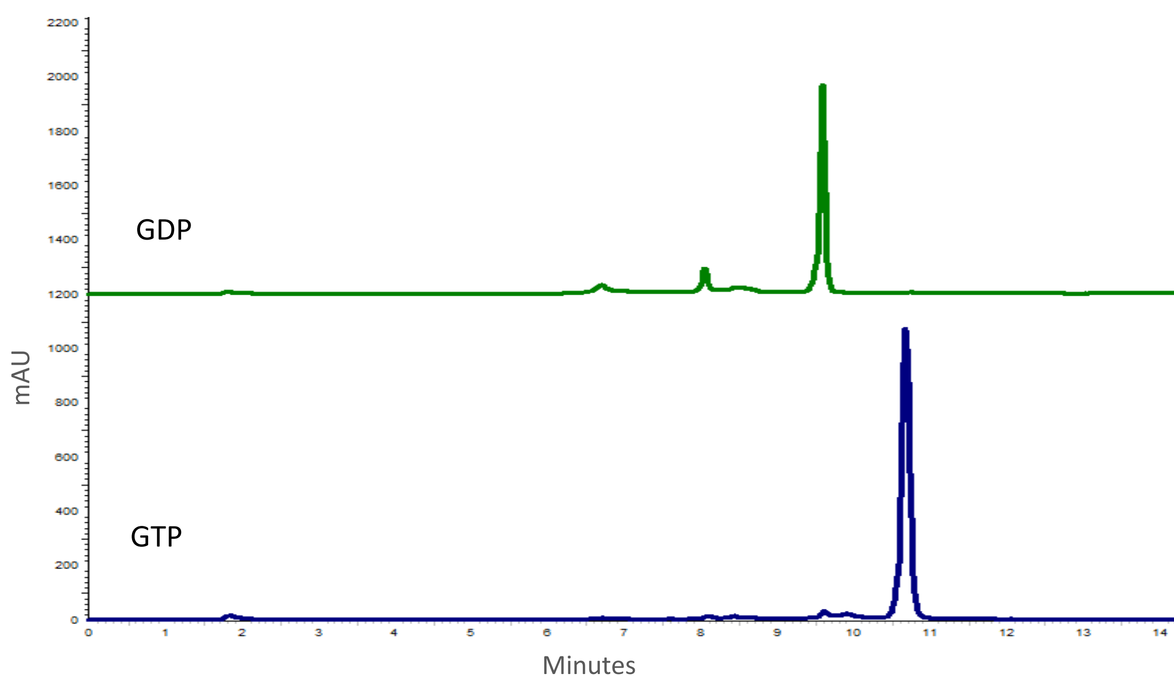


Figure 3.17: HPLC run of nucleotides. GDP (retention time of 9.6 minutes) and GTP (retention time of 10.7 minutes).

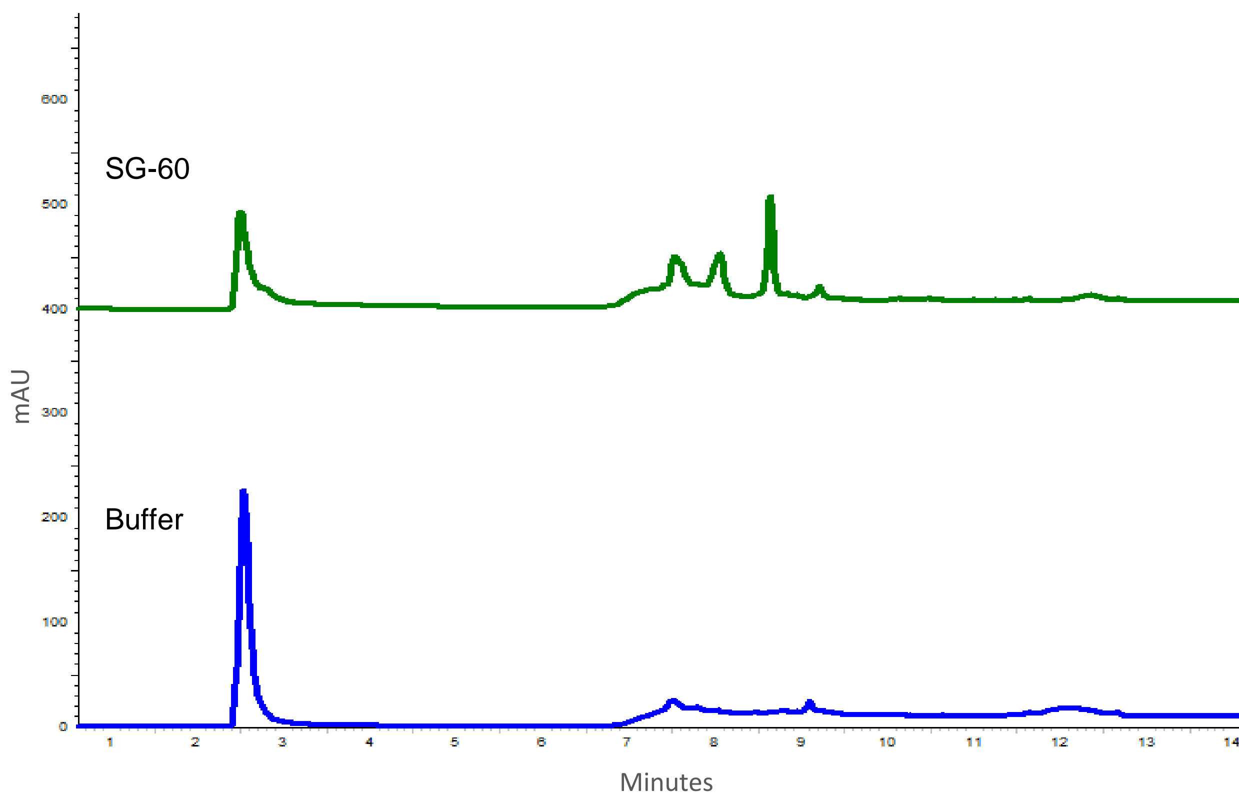


Figure 3.18: HPLC run of SG-60.

4. Conclusion and Future Prospects

Different experiments in this report covered some brief ideas about the *in vitro* behaviour of protein of interest i.e. SofG. The work reported has contributed to the successful overexpression and purification of *Myxococcus xanthus* SofG for biochemical and biophysical characterisation. The protein folded well as can be seen from CD and gel filtration experiments. The presence of secondary structure signatures in CD spectrum and elution of the protein as a monomer in the gel filtration run showed that the protein forms a globular fold. GTPase assays showed a weak activity for SofG. The experiments for measuring GTP/GDP binding affinity for SofG were inconclusive, and needs to be repeated. Crystallization attempts have been initiated and further trials have to be carried out with protein at higher concentration to improve the crystal quality.

The thesis work was initiated from scratch, beginning from amplification of the gene from the *Myxococcus xanthus* genomic DNA. SofG was a difficult protein to handle due to difficulties encountered in expression and solubility of various constructs. Out of the 8 constructs which were designed, 6 were successfully cloned, 3 expressed and only one was soluble and reasonable for proceeding with further purification. One of the major challenges encountered was the problem of precipitation during purification. Standardisation of growth conditions and standardisation of purification protocols had to be tried simultaneously, as the number of variables was large. There were also temporary problems due to loss of expression of competent cell stock, bacteriophage contamination during growth of large-scale cultures, etc., which made completion of the project within a short span of 10 months very challenging.

Results and information from this report could be used as the background for carrying out further investigation in this project. Experiments and results till date can be used for designing further experiments and trials. Some of the immediate future work will be further optimisation of purification to get high concentration of protein and setting up of crystallisation trials. This can yield crystals which could lead to solving the structure of the protein. If the protein is in nucleotide-free form, then structure could be observed at both GDP-bound form and GTP-bound form using non-hydrolysable analogue. This will help in understanding the conformational changes during GTP hydrolysis, and how it can function as a GTPase switch.

Regarding biochemical assays, the HPLC run does not show any nucleotide bound, though the protein gets stabilised (as observed by disappearance of visible precipitate) by addition of GDP. Hence, estimating the binding affinities for GDP and GTP is essential. Also, mutation studies can be done by mutating the active site residues responsible for hydrolysing GTP to confirm the observed GTP hydrolysis results. This can help in proving the role of active site residues, currently predicted based on residue conservation. Since the model suggests about its interaction with BacP, we could purify BacP also and see the interaction between SofG and BacP, to validate the current hypothesis. Other than BacP, we could also check its interaction with other motility proteins, to understand the molecular basis behind localization of PilB and PilT, the social motility proteins.

Long term plan also includes *in vivo* studies to validate the results obtained through the *in vitro* experiments. A combination of both *in vivo* and *in vitro* studies are essential to obtain mechanistic insights into motility of *Myxococcus xanthus*.

References

Balaji, P., Louis, R., Gerrit, J.K.P., Christian, H., and Alfred, W. (2000). Triphosphate structure of guanylate-binding protein 1 and implications for nucleotide binding and GTPase mechanism. EMBO J. 19(17): 4555–4564.

Berleman, JE., and Kirby, JR. (2009). Deciphering the hunting strategy of a bacterial wolfpack. FEMS Microbiol. Rev. 33, 942–57.

Blackhart, BD., and Zusman, DR. (1985). "Frizzy" genes of *Myxococcus xanthus* are involved in control of frequency of reversal of gliding motility. Proc Natl Acad Sci U.S.A. 82(24), 8767-70.

Bulyha, I., Schmidt, C., Lenz, P., Jakovljevic, V., Höne, A., Maier, B., Hoppert, M., and Søgaard-Andersen, L. (2009). Regulation of the type IV pili molecular machine by dynamic localization of two motor proteins. Mol Microbiol. 74(3), 691-706.

Bulyha, I., Lindow, S., Lin, L., Bolte, K., Wuichet, K., Kahnt, J., van der Does, C., Thanbichler, M., and Søgaard-Andersen, L. (2013). Two small GTPases act in concert with the bactofilin cytoskeleton to regulate dynamic bacterial cell polarity. Dev. Cell 25, 119−131.

Chen, H., Keseler, I M., and Shimkets, L J. (1990). Genome size of *Myxococcus xanthus* determined by pulsed-field gel electrophoresis. J Bacteriol. 172(8), 4206–4213.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, TJ., Higgins, DG., and Thompson, JD. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 31(13), 3497–3500.

Geladopoulos, TP., Sotiroudis, TG., and Evangelopoulos, AE. (1991). A malachite green colorimetric assay for protein phosphatase activity. Anal Biochem. 192(1), 112-6.

Heider, D., Hauke, S., Pyka, M., and Kessler, D. (2010). Insights into the classification of small GTPases. Adv Appl Bioinforma Chem. 3, 15–24.

Hwang, J., and Inouye, M. (2001). An essential GTPase, der, containing double GTP-binding domains from *Escherichia coli* and *Thermotoga maritima*. J Biol Chem. 276(33), 31415-21.

Jarrell, KF., and McBride, MJ. (2008). The surprisingly diverse ways that prokaryotes move. Nat Rev Microbiol. 6(6), 466-76.

Jakovljevic, V., Leonardy, S., Hoppert, M., and Søgaard-Andersen, L. (2008). PilB and PilT are ATPases acting antagonistically in type IV pilus function in *Myxococcus xanthus*. J. Bacteriol.190, 2411–2421.

Jones, DT. (1999). Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol. 292(2), 195-202.

Josenhans, C., and Suerbaum, S. (2002). The role of motility as a virulence factor in bacteria. Int. J. Med. Microbiol. 291(8), 605-14.

Konovalova, A., Petters, T., and Søgaard-Andersen, L. (2010). Extracellular biology of *Myxococcus xanthus*. FEMS Microbiol Rev. 34(2), 89-106.

Lauffenburger, D., Aris, R., and Keller, K. (1982). Effects of cell motility and chemotaxis on microbial population growth. Biophys J. 40(3), 209–219.

Leonardy, S., Miertzschke, M., Bulyha, I., Sperling, E., Wittinghofer, A., and Søgaard-Andersen, L. (2010). Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. EMBO J. 29, 2276–2289.

Li, Y., Sun, H., Ma, X., Lu, A., Lux, R, Zusman, D., and Shi, W. (2003). Extracellular polysaccharides mediate pilus retraction during social motility of *Myxococcus xanthus*. Proc Natl Acad Sci U S A. 100(9), 5443-8.

Łos, M., Czyz, A., Sell, E., Wegrzyn, A., Neubauer, P., Wegrzyn, G. (2004). Bacteriophage contamination: is there a simple method to reduce its deleterious effects in laboratory cultures and biotechnological factories? J Appl Genet. 45(1), 111-20.

Mauriello, E M F., and Zusman, D R. (2007). Polarity of motility systems in *Myxococcus xanthus*. Curr Opin Microbiol. 10(6), 624–629.

Mauriello, E M F., Mignot, T., Yang, Z., and Zusman, DR. (2010). Gliding Motility Revisited: How Do the *Myxobacteria* Move without Flagella? Microbiol Mol Biol Rev. 74(2), 229-49

Mignot, T., John, P. M Jr., and Zusman, R D. (2005). Regulated Pole-to-Pole Oscillations of a Bacterial Gliding Motility Protein. Science. 310, 855-857.

Nan, B., and Zusman, D. R. (2011). Uncovering the Mystery of Gliding Motility in the *Myxobacteria*. Annu Rev Genet. 45, 21–39.

Nielsen, E., Cheung, AY., And Ueda, T. (2008). The Regulatory RAB and ARF GTPases for Vesicular Trafficking. Plant Physiology. 147, 1516-1526.

Paduch, M., Jeleń, F., and Otlewski, J. (2001). Structure of small G proteins and their regulators. Acta Biochim Pol. 48(4), 829-50.

Remmers, AE., Posner, R., and Neubig, RR. (1990). Fluorescent guanine nucleotide analogs and G protein activation. J Biol Chem. 269(19), 13771-8.

Ruzheinikov, SN., Das, SK., Sedelnikova, SE., Baker, PJ., Artymiuk, PJ., García-Lara, J., Foster, SJ., and Rice, DW. (2004). Analysis of the open and closed conformations of the GTP-binding protein YsxC from *Bacillus subtilis*. J Mol Biol. 339(2), 265-78.

Shi, W., Köhler, T., and Zusman, DR. (1993). Chemotaxis plays a role in the social behaviour of *Myxococcus xanthus*. 9(3), 601-11.

Treuner-Lange, A., and Søgaard-Andersen, L. (2014). Regulation of cell polarity in bacteria. J Cell Biol. 206(1), 7-17.

Van den, E F., and Löwe, J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. J Biochem Biophys Methods. 67(1), 67-74.

Verstraeten, N., Fauvart, M., Versées, W., and Michiels, J. (2011). The universally conserved prokaryotic GTPases. Microbiol Mol Biol Rev. 75(3), 507-42.

Vetter, IR., and Wittinghofer, A., (2001). The guanine nucleotide-binding switch in three dimensions. Science. 294, 1299-1304.

Wennerberg, K., Rossman, KL., and Der, CJ. (2005). The Ras superfamily at a glance. J Cell Sci. 1;118, 843-6.

Zhang, Y., Franco, M., Ducret, A., and Mignot, T. (2010). 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, e1000430.