CRYSTALLIZATION TRIALS OF GLOBIN DOMAIN OF ENVIRONMENTAL STRESS PROTEIN RsbR FROM *Saprospira grandis*

SITI SURIAWATI BADAI

UNIVERSITY SAINS MALAYSIA
2011

CRYSTALLIZATION TRIALS OF GLOBIN DOMAIN OF ENVIRONMENTAL STRESS PROTEIN RsbR FROM *Saprospira grandis*

by

SITI SURIAWATI BADAI

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Science

February 2011

Acknowledgements

I am deeply grateful to my supervisor, Professor Maqsudul Alam, for his detailed and constructive comments, and for his important support throughout this work. I warmly thank Professor Nazalan Najimudin for his extensive discussions during the cluster meeting. I wish to express my sincere thanks to Dr. Jennifer Saito who gave me important guidance during my first steps into molecular work studies and for revising the English of my thesis. My warm thanks are due to Dr. Teh Aik Hong and Dr. Masaomi Kanbe who directed me in protein structure and protein purification studies. Their kind support and guidance have been of great value in this study. Many thanks to lab members Nurohaida Abd Aziz, Lee Ling Sze, Khoo Bee Feong, Ng Fui Ling, Ong Su Yean, Tan Hock Siew, Muhd Khairul Luqman Muhd Sakaff, and Teh Beng Soon for the stimulating discussions and for all the fun we have had in the last two years. I would like to thank my juniors, they were joyful companions, and in no less way did I also learn a lot from them by teaching them. I owe my loving thanks to my family. Without their encouragement and understanding it would have been impossible for me to finish this work. My special gratitude is due to Mohd Izar Omar for his loving support.

TABLE OF CONTENTS

CHAPTER 2 – MATERIALS AND METHODS

CHAPTER 3 – RESULTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF ABBREVIATIONS

The following abbreviations were used in the text:

3D	Three dimensional
Ald	Aldolase
Apr	Aprotinin
BD	Blue dextran 2000
BSA	Bovine Serum Albumin
C	Conalbumin
CA	Carbonic anhydrase
CBD	Chitin-binding domain
c-di-GMP	cyclic diguanosine monophosphate
CooA	CO-oxidation activator protein
CtrHb	*Chlamydomonas eugametos* truncated hemoglobin
CV	Column Volume
Cygb	Cytoglobin
DTT	Dithiothreitol
EAL	Referring to a conserved central sequence pattern
EM	Cryo–electron microscopy
ERERQR	Domain between the globin and GGDEF
F	Ferritin
GAF	cGMP-regulated cyclic nucleotide phosphodiesterases, adenylate cyclases, and bacterial transcription factor FhlA
GCS	Globin coupled sensor
GGDEF	Referring to a conserved central sequence pattern

HAMP	Histidine kinases, Adenylyl cyclases, Methyl-accepting chemotaxis proteins, and Phosphatases
HATPase_c	Histidine kinase-like ATPase
HemAT	Heme-based aerotactic transducers
HGFE	H; His-tag, G; Globin, FE; First and last amino acid of this sequence
HGME	H; His-tag, G; Globin, ME; First and last amino acid of this sequence
HGMN	H; His-tag, G; Globin, MN; First and last amino acid of this sequence
HisKA	Histidine kinase homologs
HNOB	Heme–NO-binding
IGME	I; Intein-tag, G; Globin, ME; First and last amino acid of this sequence
IGMN	I; Intein-tag, G; Globin, MN; First and last amino acid of this sequence
IMPACT	Intein Mediated Purification with an Affinity Chitin-binding Tag
IPTG	Isopropyl-β-D-thiogalactopyranoside
LBD	Nuclear-receptor class ligand-binding domain
MCP	Methyl-accepting chemotaxis protein
Ngb	Neuroglobin
NO	Nitric oxide
O	Ovalbumin
PAS	Per-Arnt-Sim
PCR	Polymerase Chain Reaction

PDB	Protein Data Bank
Pgb	Protoglobin
PMSF	Phenylmethylsulfonyl fluoride
PtrHb	*Paramecium caudatum* truncated hemoglobin
R	RNase A
REC	Receiver domain
Rsb	Regulator of sigma B
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
Ser	Serine
STAS	Sulfate Transporter and AntiSigma-factor antagonist
T	Thyroglobulin
Thr	Threonine

Percubaan pengkristalan domain globin daripada protein tekanan persekitaran RsbR dari *Saprospira grandis*

ABSTRAK

RsbR daripada bakteria perakus pantai *Saprospira grandis*, terdiri daripada dua domain struktur, domain globin N-terminus (residu 12-157) dan domain STAS C-terminus (residu 165-276). Pemerhatian struktur stressosom menunjukkan bahawa pelipatan globin daripada domain N-terminus RsbR mungkin memainkan peranan penting sebagai sensor tekanan dalam pengisyaratan persekitaran tekanan. Dalam kajian ini, cebisan gen yang mengekodkan domain globin daripada RsbR telah diklonkan dan protein rekombinan telah diekspreskan, ditulenkan, dan dikristalkan. Penyaringan pengkristalan awal domain globin daripada RsbR telah menghasilkan kristal berbentuk bulat. Morfologi lemah kristal ini seterusnya dioptimumkan dengan penyaringan pH, analisis kesan pilihan larutan penimbal terhadap kualiti kristal, percubaan pelbagai kepekatan mendakan, percubaan garam yang berlainan dari kelas bahan kimia yang sama, penetapan skrin pengkristalan pada suhu yang berbeza, penyemaian goresan, dan penyaringan dengan kit Screen Detergen HT. Namun, tiada peningkatan kualiti kristal diperhatikan. Empat variasi rekombinan tambahan daripada domain N-RsbR telah dibina dan percubaan pengkristalan masih sedang dilakukan.

Crystallization Trials of Globin Domain of Environmental Stress Protein RsbR from *Saprospira grandis*

ABSTRACT

RsbR from the coastal scavenger bacterium *Saprospira grandis* consists of two structural domains, the N-terminal globin domain (residues 12-157) and C-terminal STAS domain (residues 165-276). Observation of the stressosome structure indicates that the globin fold of the N-terminal domain of RsbR might play crucial role as stress sensor in the environmental stress signalling. In this study, we have cloned the gene fragment encoding the globin domain of RsbR and expressed, purified, and crystallized the recombinant protein. Initial crystallization screening of the globin domain of RsbR yielded round shaped crystals. The poor morphology of this crystal was further optimized by screening the pH, analyzing the effect of buffer choice on crystal quality, testing a reasonable range of precipitant concentration, trying different salts from the same chemical class, setting crystallization screens at different temperatures, streak seeding, and screening with the Detergent Screen HT kit. However, no improvement on crystal quality was observed. Four additional recombinant variants of N-RsbR domain were constructed and crystallization trials are currently being performed.

CHAPTER 1

LITERATURE REVIEW

1.1 Globin Superfamily

The discovery of hemoglobins in prokaryotes, fungi, plants and animals (Table 1.1) shows that the gene for hemoglobin is very ancient. They exhibit a wide variety of functions, including the reversible binding of oxygen for transport and storage, to cytoprotection against reactive oxygen species, nitric oxide (NO) scavenging, signaling of oxygen dependent metabolic pathways, and electron transfer (Wajcman *et al.*, 2009).

Table 1.1: Selected hemoglobins illustrate the diversity of proposed functions and regulation

Class	Exemplary genus	Hemoglobin	Regulation	Function (demonstrated and proposed)
Vertebrate	*Homo*	HbA	Hypoxia-induced increase in production of erythropoietin, which stimulates proliferation and differentiation of erythroid precursors, the progeny of which express Hb at a high level	Oxygen transport between tissues
Plant	*Glycine*	Lb	Nodulin-specific increase in transcription	May sequester oxygen away from nitrogenase May transport oxygen to electron transport chain in nodule
Plant	*Glycine*	Nonsymbiotic Hb	Induced by hypoxia?	Intracellular oxygen movement
Alga	*Chlamydomonas*	LI637 Hb	Light-inducible expression in chloroplast	Oxygen bound to LI637 Hb can be reduced. It may serve to accept electrons, sequester oxygen or deliver oxygen inside the organelle
Fungi	*Saccharomyces*	YHB (a flavo-hemoglobin)	Induced by high levels of oxygen or reactive oxygen species, or by blocking electron transport Repressed by hypoxia Induction is mediated by the transcription factors HAP1 and HAP2/3/4	Can transfer electrons from NADPH to heme iron May serve to protect from oxidative stress
Bacteria	*Alcaligenes*	FHP (a flavo-hemoglobin)	Induced anaerobically Promoter contains a potential binding site for NarL and FNR	Proposed electron transfer Possible role in anaerobic metabolism, perhaps gas metabolism during denitrification
Bacteria	*Vitreoscilla*	Hb	Induced by hypoxia Promoter contains binding sites for FNR	Can serve as terminal electron acceptor during respiration May scavenge oxygen

Table from HARDISON, R. 1998. Hemoglobins from bacteria to man: evolution of different patterns of gene expression. *J Exp Biol,* 201**,** 1099-117, Table 1, Page 1100.

Initially, our knowledge of globin proteins was limited to the α- and β-subunits of hemoglobins in vertebrates and myoglobins. The emergence of high throughput genome sequencing facilities and powerful high performance bioinformatic tools revealed a rapid growth of a "globin superfamily".

The first plant hemoglobins were discovered in the root nodules of legumes. These symbiotic hemoglobins (leghemoglobins) facilitate the diffusion of oxygen to the respiring bacteroids in the root nodule (Appleby, 1984). The ensuing years brought the discovery of hemoglobins distinct from leghemoglobin, called non-symbiotic hemoglobins, in root nodules of a nonleguminous plant (Bogusz *et al.*, 1988) and in the monocotyledon cereals (Taylor *et al.*, 1994).

Further studies have found hemoglobins in jawless vertebrates and in diverse invertebrates ranging from flies (arthropods) to earthworms (annelids) to nematodes (Riggs, 1991; Dixon and Pohajdak, 1992; Sherman *et al.*, 1992).

More biochemical reactivities have been implicated in a wider variety of potential functions than those traditionally associated with animal and plant hemoglobins by the discovery of hemoglobins in unicellular organisms. Hemoglobins that are found in the chloroplast of *Chlamydomonas* are light-inducible. This hemoglobin could serve as an electron acceptor in the electron transport system or another redox system based on the reduction of bound O_2 (Couture and Guertin, 1996).

Flavohemoglobins discovered in *Escherichia coli* (Vasudevan *et al.*, 1991), *Bacillus subtilis* (LaCelle *et al.*, 1996), *Alcaligenes eutrophus* (Cramm *et al.*, 1994), and yeast *Saccharomyces* (Zhu and Riggs, 1992; Iwaasa *et al.*, 1992) are involved in protection against NO (Poole *et al.*, 1996; de Jesus-Berrios *et al.*, 2003). Flavohemoglobins consist of two domains, one for binding heme and one for binding a flavin cofactor. The three-

dimensional (3D) structure of *Alcaligenes* flavohemoglobin (Figure 1.1) corresponds to the classical globin fold, demonstrating the homology between the bacterial and eukaryotic hemoglobins (Ermler *et al.*, 1995).

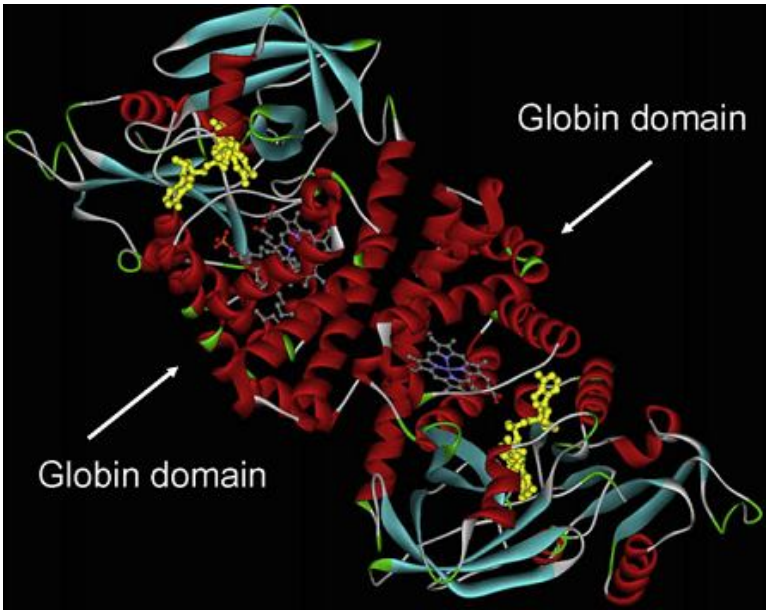

Figure 1.1: Structure of the flavohemoglobin from *Alcaligenes eutrophus*. This molecule has been crystallized as a dimer. The helical part of the molecule, which corresponds mostly to the globin domain surrounding the heme group, is shown in red. (PDB accession code 1CQX). Figure from WAJCMAN, H., KIGER, L. & MARDEN, M. C. 2009. Structure and function evolution in the superfamily of globins. *C R Biol*, 332**,** 273-82, Figure 2, Page 275.

Truncated globins are small hemoproteins with sequences shorter than mammalian hemoglobins by 20 to 40 residues (Wittenberg *et al.*, 2002). These molecules provide long-term ligand or substrate storage, NO detoxification, O_2/NO sensing, redox reactions, and O_2 delivery under hypoxic conditions (Wittenberg *et al.*, 2002; Vuletich and Lecomte, 2006). They were found in unicellular eukaryotes (Takagi, 1993), cyanobacteria (Potts *et al.*, 1992; Kaneko and Tabata, 1997; Scott *et al.*, 2002), a

nemertean (Vandergon *et al.*, 1998), and a large number of bacteria (Couture *et al.*, 1999b; Wittenberg *et al.*, 2002). Globins longer than normal (>160 aa), similar to the "truncated" Hbs, were observed in a green alga (Couture *et al.*, 1994; Couture *et al.*, 1999a) and in plants (Watts *et al.*, 2001).

Rapid accumulation of genomic information has resulted in the discoveries of new types of globins, including neuroglobins (Ngbs), cytoglobins (Cygbs), and globin coupled sensors (GCSs). Ngbs (Burmester *et al.*, 2000; Trent *et al.*, 2001) and Cygbs (Burmester *et al.*, 2002; Hankeln *et al.*, 2005) have been found in all vertebrates from humans to birds and fish (Hankeln *et al.*, 2005). There is also an eye-specific globin in the domestic chicken *Gallus gallus*, GbE found related to Cygb (Kugelstadt *et al.*, 2004), and a fifth type of globin gene in fish and amphibians, GbX which is thought to have been lost in the higher vertebrates (Roesner *et al.*, 2005).

GCSs comprise of an N-terminal globin domain that binds (‘‘senses’’) compatible ligands and a C-terminal variable signaling domain that relays (‘‘transduces’’) this ligand-bound state in a form that is compatible with the host cell. GCSs were discovered in several bacterial and archaeal groups (Hou *et al.*, 2000; Hou *et al.*, 2001; Freitas *et al.*, 2003).

A phylogenetic analysis of GCS globin domains suggest that these sensors descended from an ancient globin or protoglobin (Pgb) (Freitas *et al.*, 2005; Freitas *et al.*, 2004). Characterization of two Pgbs from the archaea, *Ap*Pgb from the obligately aerobic hyperthermophile *Aeropyrum pernix* and *Ma*Pgb from the strictly anaerobic methanogen *Methanosarcina acetivorans*, found that both *Ap*Pgb and *Ma*Pgb conform to the globin sequence motifs, contain heme, and demonstrably bind O_2, CO, and NO (Freitas *et al.*, 2004).

1.2 Globin fold

The classical globin molecule has been described fifty years ago by Kendrew (Kendrew *et al.*, 1958). This molecule is characterized by 8 helices designated A through H and arranged in a 3-over-3 helical sandwich structure. This arrangement forms a hydrophobic pocket which contains a heme prosthetic group. Helices A, B, C, and E are on the distal side of the heme and helices F, G, and H are on the proximal side (Figure 1.2). The proximal histidine in the F helix (His F8) binds the heme cofactor to the protein by coordinating to the heme iron's fifth coordinate. Gaseous ligands bind to the iron's sixth coordinate at the opposite side of the heme plane.

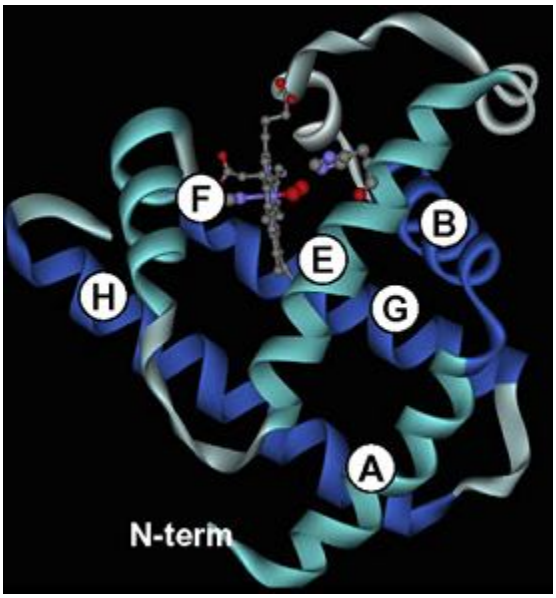

Figure 1.2: Structure of "classical" globin fold. Sperm whale myoglobin is the typical example of a 3/3 globin, the heme group is surrounded by 3 helices on its proximal site (F, G, and H) and 3 helices on its distal site (A, B, and E). Figure from WAJCMAN, H., KIGER, L. & MARDEN, M. C. 2009. Structure and function evolution in the superfamily of globins. *C R Biol,* 332**,** 273-82, Figure 1, Page 275.

Analysis of amino acid sequences and protein folds has recently unraveled the structural bases and details of several proteins from the “truncated hemoglobin” family. The crystal structure of the truncated hemoglobin from the ciliated protozoan *Paramecium caudatum* (PtrHb) shows that the tertiary structure of this protein is based on a 2-over-2 α-helical sandwich (Figure 1.3) (Pesce *et al.*, 2000; Milani *et al.*, 2001). The 2/2 globin fold results from extensive and complex modifications of the “classical” globin fold (Figure 1.4).

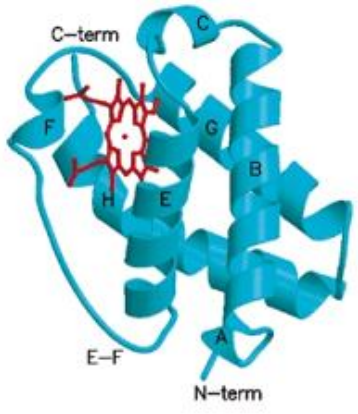

Figure 1.3: Structure of 2/2 globin. *P. caudatum* hemoglobin is an example of a 2/2 Hb. Here the heme group is sandwiched between helices B and E on the distal site and helices G and H on the proximal site (PDB accession code 1DLW). Figure from PESCE, A., COUTURE, M., DEWILDE, S., GUERTIN, M., YAMAUCHI, K., ASCENZI, P., MOENS, L. & BOLOGNESI, M. 2000. A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family. *EMBO J,* 19**,** 2424-34, Figure 2, Page 2428.

The A-helix and the CD-D region are drastically reduced or fully absent, with a short segment linking the C- and E-helices forcing the E-helix very close to the heme distal face. The proximal F-helix is almost completely replaced by an extended pre-F-helix (E-F region) polypeptide loop. The CD region and the E-, F-, and G-helices build the protein crevice where the heme is inserted, stabilized, and shielded from the solvent (Figure 1.5). Stabilization of the bound heme is achieved through direct Fe coordination

to the proximal HisF8 residue, whose conformation is mostly staggered relative to the pyrrole nitrogen atoms, thus supporting fast oxygen association and electronic donation to the bound distal ligand (Pesce *et al.*, 2000).

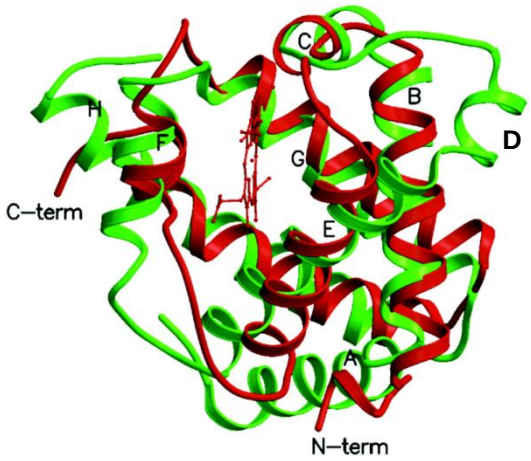

Figure 1.4: Comparative view of the 2/2 and "classical" globin folds. A ribbon diagram of the structural overlay between *Chlamydomonas eugametos*, CtrHb (red ribbon; PDB accession code 1DLY) and 3/3 sperm whale Mb (green ribbon; PDB accession code 1EBC). Figure from WITTENBERG, J. B., BOLOGNESI, M., WITTENBERG, B. A. & GUERTIN, M. 2002. Truncated hemoglobins: a new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes, and plants. *J Biol Chem*, 277**,** 871-4, Figure 2, Page 872.

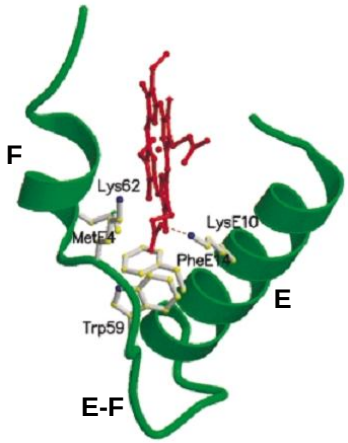

Figure 1.5: A hydrophobic cluster preventing solvent access to the heme pocket from the pre-F region, in CtrHb. The figure includes the heme group, the E-helix, the pre-F loop and the F α-helical turn. Residues LysE10 and Lys62, which are electrostatically linked to the heme propionates, are also displayed. Figure from PESCE, A., COUTURE, M., DEWILDE, S., GUERTIN, M., YAMAUCHI, K., ASCENZI, P., MOENS, L. & BOLOGNESI, M. 2000. A novel two-over-two alpha-helical sandwich fold is characteristic of the truncated hemoglobin family. *EMBO J,* 19**,** 2424-34, Figure 4, Page number 2430.

1.3 Heme-based sensors

Heme-based sensor proteins are the key regulators of adaptive responses to fluctuating O_2, CO, and NO levels. These sensors achieve their responses by coupling a heme-binding domain to a linked transmitter domain. There are six different types of heme-binding modules: the heme-binding PAS domain (Per-Arnt-Sim), GCS, CO-oxidation activator protein (CooA), and heme–NO-binding (HNOB), cGMP-regulated cyclic nucleotide phosphodiesterases, adenylate cyclases, and bacterial transcription factor FhlA (GAF) and nuclear-receptor class ligand-binding domain (LBD) (Gilles-Gonzalez and Gonzalez, 2008). The transmitters for coupling to such heme-binding domains include histidine protein kinases, cyclic nucleotide phosphodiesterases, chemotaxis methyl-carrier protein receptors, and transcription factors of the basic helix-loop-helix and helix-turn-helix classes (Gilles-Gonzalez and Gonzalez, 2005).

1.4 GCS

Heme-based sensors utilize either a globin domain or a PAS domain to detect O_2, CO, NO, and sulfides (S^{2-}). In contrast to PAS domains that can bind various cofactors, globins are known to bind only heme. GCSs are multidomain proteins, consisting of an N-terminal globin domain fused to a variety of C-terminal transmitter domains. Functional classification of the GCSs is based on the transmitter domain(s) they possess, broadly falling under either aerotaxis or gene regulation.

1.5 GCS Functional Classification

Heme-based sensors were classified as either aerotactic or gene regulating (Freitas *et al.*, 2003). The gene-regulating group was further subdivided into three

subgroups: protein-DNA (Gekakis *et al.*, 1998; Hogenesch *et al.*, 1998; Lanzilotta *et al.*, 2000; Reick *et al.*, 2001; Dioum *et al.*, 2002), protein–protein (David *et al.*, 1988; Hou *et al.*, 2000), and 2nd messenger pathways (Ross *et al.*, 1991; Gilles-Gonzalez *et al.*, 1991; Zhao *et al.*, 1999; Delgado-Nixon *et al.*, 2000; Chang *et al.*, 2001; Sasakura *et al.*, 2002). The resulting organization scheme is illustrated in Figure 1.6.

1.5.1 Aerotactic

In February of 2000, the first heme-based aerotactic transducers (HemAT) from *Halobacterium salinarum* and *B. subtilis* were discovered (Hou *et al.*, 2000). HemATs compose of an N-terminal globin domain and a C-terminal homologous to bacterial methyl-accepting chemotaxis proteins (MCP). Recently, three new archaeal HemATs were discovered in two organisms. First, HemAT from *Haloarcula marismortui*, a close relative of *H. salinarum* (Baliga *et al.*, 2004), exhibits HAMP (Histidine kinases, Adenylyl cyclases, Methyl-accepting chemotaxis proteins, and Phosphatases) :MCP domains. The other two HemATs were found in the archaeon *Natronomonas pharaonis* DSM 2160. One of these HemATs couples a sensor globin with a PAS:PAS domain combination (Freitas *et al.*, 2008). This extra sensing capability may aid this archaeon in dealing with both high osmolality (3.5 M NaCl) and high pH (lake pH ~11). In particular, the magnetotactic proteobacterium *Magnetospirillum magnetotacticum* possesses two aerotactic transducers, whereas *Magnetococcus* MC-1 cells possess only one (Freitas *et al.*, 2003). Magnetotaxis has been shown to work in conjunction with aerotaxis (Frankel *et al.*, 1997).

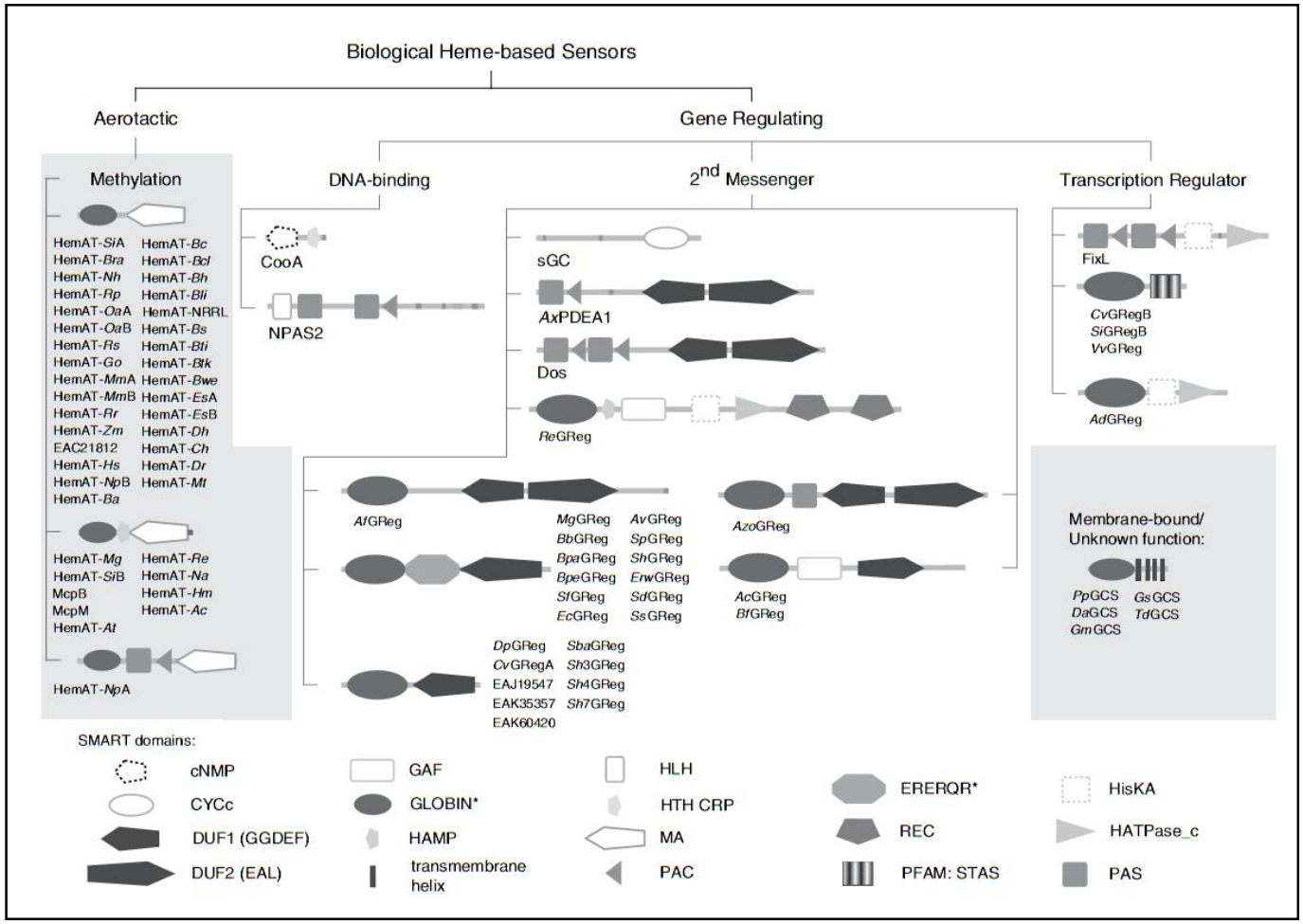

Figure 1.6: Classification scheme of biological heme-based sensors. Heme-based sensors and their domain organization are illustrated. Individual globin-coupled sensors are assigned to their respective class on the basis of the known/putative functions of their signaling domains. The name ERERQR is a name given to the domain between the globin and GGDEF (DUF1) domain and based on the ERERQR motif it contains (Freitas *et al.*, 2003). The source organisms of each GCS are listed in Appendix A. Figure from FREITAS, T. A., SAITO, J.A., WAN, X., HOU, S. & ALAM, M. 2008. Protoglobin and Globin-coupled Sensors. *In:* GHOSH, A. (ed.) *The Smallest Biomolecules Diatomics and their Interactions with Heme Proteins.* First Edition ed. Oxford: Elsevier, Figure 2, Page 186.

1.5.2 Modulation of a Second Messenger

Thirty GCSs were identified in this group, incorporating either the GGDEF domain or a GGDEF:EAL domain pair (Freitas *et al.*, 2008). The *Ax*PDEA1 protein (GGDEF:EAL domains) from *Gluconacetobacter xylinus* is a heme-binding protein that regulates cellulose production in response to cellular O_2 levels by linearizing cyclic diguanosine monophosphate (c-di-GMP). The GCS from *Acidithiobacillus ferrooxidans* has the same signaling domain organization as heme-PAS proteins *Ec*Dos (aka YddV) and *Ax*PDEA1 (Chang *et al.*, 2001) and may possess a similar function. The GCS from *Burkholderia fungorum* exhibits a GAF:EAL domain organization, where GAF domains are nucleotide-specific cAMP- and cGMP-regulating domains (Galperin *et al.*, 2001; Kanacher *et al.*, 2002; Hurley, 2003) with broad cellular roles as far reaching as in the human rod photoreceptors (Paglia *et al.*, 2002). *Azoarcus* sp. EbN1 possesses a GCS, like the HemAT from *N. pharaonis*, containing a PAS:PAS domain coupled with the sensor globin domain. More interestingly, the GCS from γ-Proteobacterium *Reinekea* sp. MED297 (*Re*GReg) has seven identifiable domains (sensor_globin:HAMP:GAF: HisKA:HATPase_c:REC:REC). HATPase_c is a histidine kinase-like ATPase, which binds and hydrolyzes ATP. Receiver domains (REC) contain a phosphoacceptor site that is phosphorylated by histidine kinase homologs (HisKA) and forms homodimers. Although the *Reinekea* GCS exhibits the HisKA and HATPase_c domains like FixL, it is classified in this category solely on the presence of its GAF domain (Freitas *et al.*, 2008).

1.5.3 Protein-Protein Interactions

Currently, four proteins populate this subdivision of the GCSs (Freitas *et al.*,

2008). The δ-Proteobacterium *Anaeromyxobacter dehalogenans* is a metabolically versatile facultative anaerobe with efficient mechanisms for bioremediation (Cole *et al.*, 1994; Loffler *et al.*, 1999; Sanford *et al.*, 2002; He and Sanford, 2002; He and Sanford, 2003). *A. dehalogenans* GCS called *Ad*GReg consists of a catalytic HATPase_c domain and a HisKA histidine kinase domain and probably functions as a sensory histidine kinase like FixL. Three GCSs with a C-terminal STAS domain (sulfate transporter and antisigma-factor antagonist) have been identified in *Vibrio vulnificus*, *Chromobacterium violaceum*, and *Silicibacter sp. TM1040*. These proteins may regulate gene expression by functioning as an anti-antisigma factor (antisigma-factor antagonist or ASA) similar to SpoIIAA in *B. subtilis* spore formation (Aravind and Koonin, 2000).

1.5.4 Unclassified GCSs

Membrane-spanning (~four transmembrane helices) GCSs found in the anaerobic δ-Proteobacteria *Geobacter sulfurreducens*, *Geobacter metallireducens*, *Pelobacter propionicus* DSM 2379, and *Desulfuromonas acetoxidans*, and the ε-Proteobacterium *Thiomicrospira denitrificans* are included in this small group of GCSs. Other GCSs have been labeled as unclassified because they have only partial C-terminal signaling domains or lengthy sequences with no recognizable domain (Freitas *et al.*, 2008).

1.6 Bacterial Stress Response

Most bacteria live in a dynamic environment including variations in temperature, ionic strength, pH, and depletion of nutrients. A common strategy used by microorganisms to survive multiple stresses involves a signal transduction cascade that controls the simultaneous expression of a number of stress-responsive genes. Such signalling cascade in *B. subtilis* leads to the activation of the general stress sigma factor, σ^B, and enhanced transcription of its large regulon to provide a global response to the imposed stress (Chen *et al.*, 2003) (Figure 1.7). The regulation of σ^B involves the product of nine genes in the σ^B operon, namely RsbR, RsbS, RsbT, RsbU, RsbV, RsbW, RsbX, RsbP, and RsbQ (Kalman *et al.*, 1990; Wise and Price, 1995; Vijay *et al.*, 2000; Brody *et al.*, 2001). These genes are termed rsb, for regulator of sigma B.

The *B. subtilis* stressosome consists of multiple copies of the regulators of σ^B: RsbS, RsbR, and four paralogs of RsbR (Dufour *et al.*, 1996; Chen *et al.*, 2003; Gaidenko *et al.*, 2006) which sequester a kinase, RsbT (Chen *et al.*, 2003; Murray *et al.*, 2005). During stress, RsbT is released from the RsbR-RsbS complex by phosphorylating the STAS domains of RsbR and RsbS (Murray *et al.*, 2005). RsbT then interacts with the serine/threonine (PP2C) phosphatase, RsbU (Chen *et al.*, 2003; Akbar *et al.*, 1997). RsbX, the phosphatase for RsbS-P, is responsible for the feedback mechanism by which the level of activity returns to the pre-stress levels (Smirnova *et al.*, 1998; Voelker *et al.*, 1997; Yang *et al.*, 1996)

Two PP2C phosphatases, RsbU and RsbP, catalyze the dephosphorylation of RsbV-P, and define which stress signals converge on RsbV-P (Vijay *et al.*, 2000; Voelker *et al.*, 1996; Yang *et al.*, 1996). Environmental stress signals lead to activation of RsbU whereas energy stress signals activate RsbP. Dephosphorylated RsbV then binds the

RsbW anti-σ factor, forcing it to release σ^B (Alper *et al.*, 1996; Benson and Haldenwang, 1993; Dufour and Haldenwang, 1994) to activate expression of the general stress regulon, comprising more than 150 genes (Delumeau *et al.*, 2006)

In addition to trapping RsbT, the stressosome plays a vital role as a signaling hub that integrates a diverse array of signals to activate the σ^B cascade and thus the stress response. The complex could act as a receptor of the environmental stress signal itself. Observation of the globin fold of the N-terminal domain of RsbR (Murray *et al.*, 2005) is evocative of the role that globins play in sensing gaseous ligands. Alternatively, the complex could act as an attenuator or amplifier of the stress signal by controlling the number of RsbT molecules that are released from the stressosomes as a function of the intensity of the imposed stress (Delumeau *et al.*, 2006). The question that remains, however, is how stress is signaled to the stressosome, resulting in the release of RsbT to activate RsbU and hence σ^B?

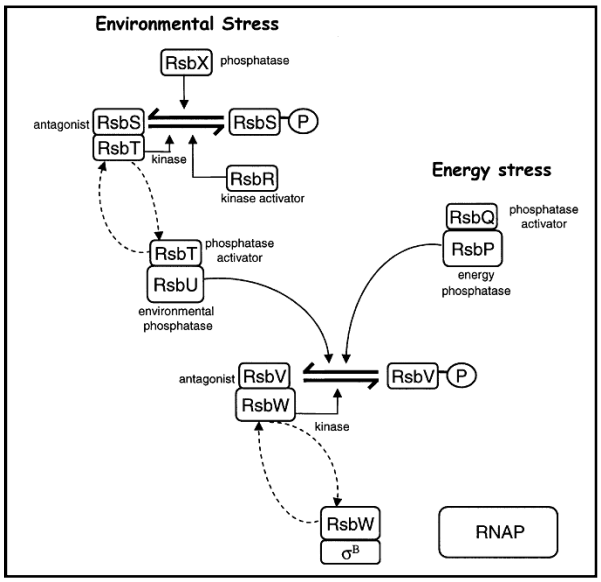

Figure 1.7: The σ^B signalling cascade of *B. subtilis*. Figure from CHEN, C. C., LEWIS, R. J., HARRIS, R., YUDKIN, M. D. & DELUMEAU, O. 2003. A supramolecular complex in the environmental stress signalling pathway of Bacillus subtilis. *Mol Microbiol,* 49**,** 1657-69, Figure 1, Page 1658.

1.7 *Saprospira grandis*

S. grandis is a gliding bacterium that can be found along sea shores all over the world. This Gram-negative species forms long helical filaments extending from each *S. grandis* cell as long as 500 μm (Lewin, 1997). When such cells are suspended in liquid, they easily tangle with one another to make big aggregates of twisted filaments. *S. grandis* plays an important role as a coastal scavenger. *S. grandis* also preys on enteric bacteria such as *Salmonella* and *E. coli*, both of which can survive in seawater (Aizawa, 2005). *S. grandis* traps prey cells passively, doing little until actively motile prey cells bump against its long filaments and then adhere to its cell surface (Figure 1.8).

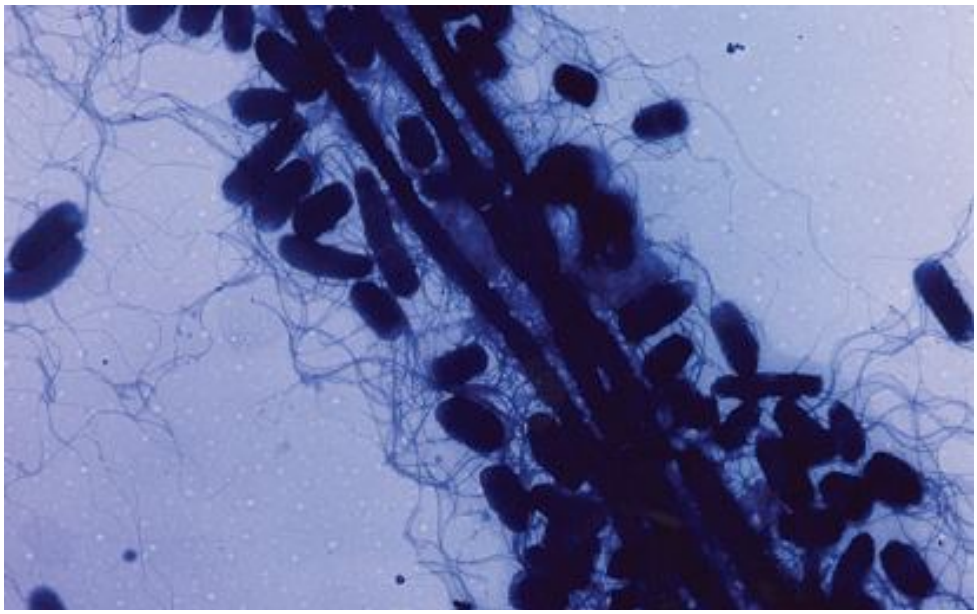

Figure 1.8: Wild-type *Salmonella typhimurium SJW1103* cells were trapped along the filamentous cell surface of *S. grandis*. Figure from AIZAWA, S. I. 2005. Bacterial Gliding Motility: Visualizing Invisible Machinery. *ASM News,* 71**,** 71-75, Figure 2, Page 74.

Lewin calls this kind of approach as “ixotrophy”, meaning the process of feeding on prey bacteria caught on sticky substance (Lewin, 1997). The final step in *S. grandis* ixotrophy is digestion of the prey bacteria and use of the organic matter as an energy source. The mode of killing is still poorly understood. It is speculated that *S. grandis* uses bacteriocins in killing prey cells. *S. grandis* produces rod-shaped particles called rhapidosomes (Figure 1.9). These rhapidosomes strikingly resemble the R-type pyocin of *Pseudomonas aeruginosa* that is a phage-tail-type bacteriocin.

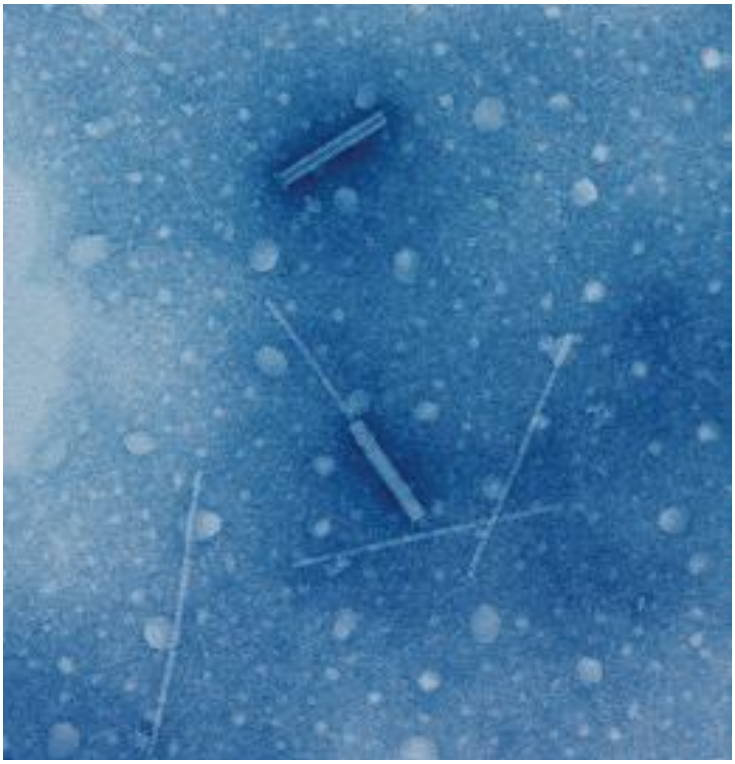

Figure 1.9: Electron micrographs of rhapidosome, a bacteriocin of *S. grandis*. Figure from AIZAWA, S. I. 2005. Bacterial Gliding Motility: Visualizing Invisible Machinery. *ASM News*, 71**,** 71-75, Figure 3, Page 74.

1.8 RsbR in *S. grandis*

The stressosome is the signaling hub that integrates a diverse array of signals to activate the σ^B cascade and thus the stress response. However, how is stress signalled to the stressosome, activate the σ^B cascade and thus the stress response? Unlike the two-component system in *B. subtilis*, there are no obvious sensory domains in the proteins that compose the environmental stress signalling pathway.

To better understand the assembly and function of the stressosome, we have identified the Rsb proteins in *S. grandis*. Annotation of the *S. grandis* genome found RsbR, RsbS, RsbT, and RsbW-U. RsbW-U contained parts of the domains from both RsbW and RsbU. There are five different globin-coupled RsbR in the *S. grandis* genome but only one is found neighbouring other stress related proteins. This stress-related protein has two structural domains (Figure 1.10). The N-terminal of RsbR (N-RsbR) is a myoglobin-like sensing domain (residues 12-157) and the C-terminal (residues 165-276) is homologous to sulfate transporter and anti-factor antagonist (STAS).

Observation of the globin fold of the N-terminal domain of RsbR in *B. subtilis* (Murray *et al.*, 2005) suggest that this domain plays a crucial role in sensing gaseous ligands, but there is no bound cofactor in this globin protein. Here we report the cloning, expression, purification, and crystallization trials of the globin domain of environmental stress protein RsbR from *S. grandis*.

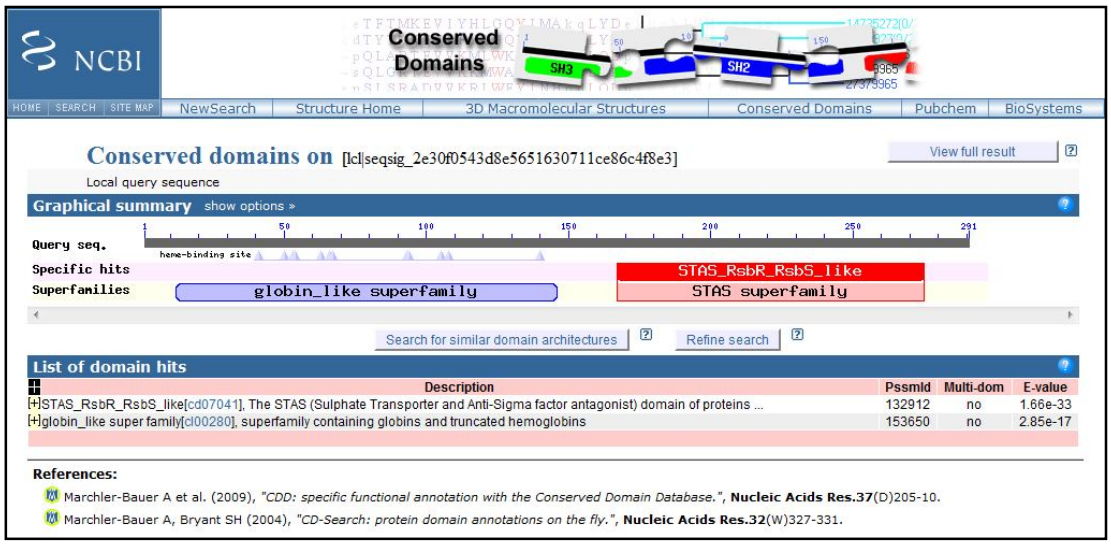

Figure 1.10: Conserved domains of RsbR protein. A search for conserved domains using NCBI's Conserved Domain software revealed that RsbR found in *S. grandis* contain globin domain at N-terminal and STAS domain at C-terminal

1.9 Objectives

The goal of this research is to elucidate the molecular mechanism of oxygen sensing for the sensor domain of RsbR from *S. grandis*.

Objective 1: Cloning, Expression, and Purification of RsbR globin domain

Objective 2: Crystallization of RsbR globin domain

Objective 3: Data collection and structure determination of RsbR globin domain

CHAPTER 2

MATERIALS AND METHODS

2.1 Overview of Protein Structure Determination

The flow chart of the processes that involved in the protein structure determination is illustrated in Figure 2.1.

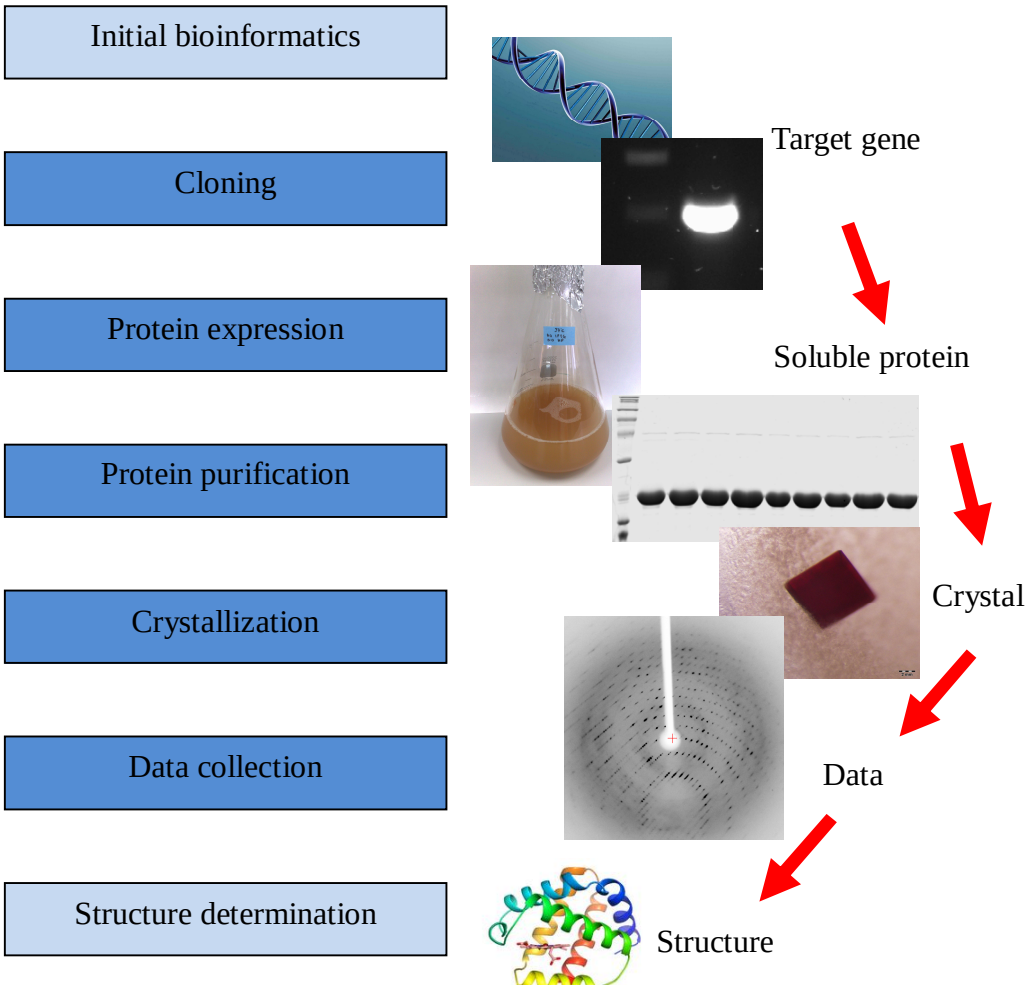

Figure 2.1: Overview of protein structure determination. The bar on the left side of the figure lists major stages of a crystal structure determination project. The dark blue shading indicates experimental procedures while the light shading indicates work performed *in silico*.

2.2 Genomic DNA Extraction

Genomic DNA was isolated using the GNome Kit. Overnight culture ~5 ml was centrifuged and the cell pellet was resuspended with 1.85 ml of Cell Suspension Solution. RNA was degraded by adding 50 µl of RNase Mixx. To lyse the cells, 100 µl of Cell Lysis/Denaturing Solution was added to the suspension and incubated at 55ºC for 1.5 hours. Protein was denatured by adding 25 µl of Protease Mixx and incubated at 55ºC for 3-4 hours. To precipitate the protein, 500 µl of "Salt-Out" Mixture was added and mixed thoroughly. The sample was then divided into 1.5 ml microcentrifuge tubes and incubated at 4ºC for 10 minutes. The tubes were centrifuged at 14,000 rpm for 10-15 minutes at 4ºC. The supernatant was transferred to a 15 ml falcon tube and mixed with 2 ml of TE buffer. Slowly, 8 ml of ice-cold 100% ethanol was added and the DNA at the interface was spooled with a glass rod with a small hook. The DNA was air-dried for 45-60 minutes and then dissolved in 30-40 µl of TE or HPLC water. Quality of genomic DNA was checked on an agarose gel.

2.3 Polymerase Chain Reaction (PCR)

High fidelity PCR was performed using *PfuTurbo* DNA polymerase (Stratagene). The reactions were set up according to Table 2.1, with *PfuTurbo* DNA polymerase being added after a 2 minute hot start at 95˚C. This was followed by 25 cycles of 95˚C for 30 seconds, 55˚C for 30 seconds, and 72˚C for 60 seconds, with a final extension at 72˚C for 7 minutes. PCR products were analyzed by agarose gel electrophoresis.

Table 2.1: PCR reaction setup

Reagent	Amount
DNA	0.5-1 µl
10X Pfu Buffer	5 µl
dNTPs (2.5 mM)	5 µl
DMSO	3 µl
Primer 1 (10 pmol/µl)	2 µl
Primer 2 (10 pmol/µl)	2 µl
HPLC water	31 µl
PfuTurbo DNA Polymerase (2.5 U/µl)	1 µl
Final volume	50 µl

2.4 PCR Purification

DNA fragments from PCR were purified from primers, nucleotides, polymerases, and salts using the QIAquick PCR Purification Kit (Qiagen). The PCR sample was mixed with 5 volumes of Buffer PB, applied to a QIAquick spin column, and centrifuged for 1 minute at 14,000 rpm. The flow-through was discarded and the bound DNA was washed twice with 750 µl Buffer PE. After the flow-through was discarded, the column was centrifuged for an additional 1 minute to remove residual ethanol. The QIAquick column was placed in a clean microcentrifuge tube and air-dried for 20 minutes. To elute, 30 µl of HPLC water was added to the center of the QIAquick membrane, allowed to stand for 1-2 minutes, and then centrifuged for 1 minute.

2.5 Restriction Enzyme Digestion

Blunt-end products with specific restriction sites at the C-terminal and N-terminal were generated using PCR. These products were digested with corresponding restriction enzyme to facilitate the cloning of the genes into the appropriate vector. Restriction enzymes and alkaline phosphatase were purchased from Promega. The preparative digestion reaction was set up as indicated in Table 2.2 and incubated in a 37ºC water bath for 1 hour. The vector was dephosphorylated by adding 1 µl of alkaline phosphatase (1 U/µl) and incubating at 37ºC for an additional 1 hour. This was done to prevent self-ligation of partially digested vector.

Table 2.2: Preparative digestion setup

Reagent	Amount
DNA	5 µl
10X Enzyme Buffer	3 µl
Restriction Enzyme 1	0.5 µl
Restriction Enzyme 2	0.5 µl
BSA	0.3 µl
HPLC water	20.7 µl
Final volume	30 µl

A mini-digestion was used to screen for insert-containing plasmids. The reaction was set up as in Table 2.3.

Table 2.3: Mini-digestion setup

Reagent	Amount
DNA	3 µl
10X Enzyme Buffer	1.5 µl
Restriction Enzyme 1	0.25 µl
Restriction Enzyme 2	0.25 µl
BSA	0.15 µl
HPLC water	9.85 µl
Final volume	15 µl