STRUCTURAL BASIS FOR THE REGULATION OF A CONSERVED TONB-DEPENDENT IRON TRANSPORT SYSTEM VIA CELL-SURFACE SIGNALING IN GRAM-NEGATIVE BACTERIA

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Jaime Lea Jensen

In Partial Fulfillment of the Requirements
for the Degree of
DOCTOR OF PHILOSOPHY

Major Department:
Chemistry and Biochemistry

June 2017

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Structural Basis for the Regulation of a Conserved TonB-dependent Iron Transport System via Cell-surface Signaling in Gram-negative Bacteria

By

Jaime Lea Jensen

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Christopher Colbert, Ph.D.
Chair

Kenton Rodgers, Ph.D.

Sangita Sinha, Ph.D.

Glenn Dorsam, Ph.D.

Approved:

July 7, 2017
Date

Gregory Cook, Ph.D.
Department Chair

ABSTRACT

Cell-surface signaling (CSS) pathways are highly conserved systems in Gram-negative bacteria that allow the cell to efficiently respond to environmental stimuli through transcriptional regulation. Three distinct proteins are involved in this process: an outer membrane (OM) protein that senses the extracellular signal, an inner membrane (IM) sigma regulator protein that transmits the signal from the OM protein to the cytoplasm, and an extracytoplasmic function (ECF) sigma factor that initiates transcription of stimulus response genes. One such CSS pathway regulates bacterial iron acquisition- an essential process for bacterial survival and pathogenesis. Under iron-limited conditions, expression of the OM transporter is upregulated by signal transduction through the IM protein to the sigma factor. The goal of this work is to provide a structural rationale for distinctive signal transduction through the CSS pathway that regulates ferric siderophore uptake in Gram-negative bacteria, by structurally characterizing these proteins from *Pseudomonas capeferrum*, with a focus on the IM protein, PupR.

The solution structures of an OM transporter, PupA, an OM transducer, PupB, a PupA$^{\text{PupB-NTSD}}$ chimera, and the OM proteins with the PupR C-terminal cell-surface signaling domain (CCSSD) were probed by SEC-SAXS to examine global architectural differences amongst the OM proteins.

The X-ray crystal structure of the PupB N-terminal signaling domain (NTSD):PupR-CCSSD complex was determined. The PupB-NTSD exhibits a conserved βαβ-repeat motif. Unexpectedly, the CCSSD subdomain contains the same fold, which is the first time this fold had been identified at a protein's C-terminus. The other subdomain of the CCSSD, designated the C-terminal juxtamembrane (CJM) subdomain, has a novel, β-solenoid-like motif. Analysis of the CCSSD by CD spectroscopy and SEC-SAXS indicated that the domain is highly flexible,

and is significantly stabilized by the PupB-NTSD. Concurrently, the PupB-NTSD structure was determined by NMR, and contrasted with published NTSDs to evaluate structural variation that may account for disparities in functionality.

The PupR N-terminal anti-sigma domain (ASD) was solved by X-ray crystallography and presents as a dimer in solution- the first description of a transmembrane ASD to assume an oligomeric form.

Structural characterization of these proteins suggests novel implications for CSS through the TonB-dependent ferric siderophore uptake pathway.

ACKNOWLEDGEMENTS

First and foremost, I must acknowledge my dissertation advisor, Dr. Chris Colbert. I am indebted to him for his continuous support, advice, generosity, and seemingly superhuman levels of patience. He graciously subsidized multiple travel opportunities to attend conferences and workshops to expand my training, many of which I think I would not have had the possibility to attend otherwise, and his open-door policy ensured that I and my labmates had research and intellectual assistance readily available. He also provided me with the chance to contribute to multiple projects beyond my primary thesis research, which considerably expanded my knowledge and technical abilities. I am proud to be Dr. Colbert's first graduate student trainee, and am extremely proud of the work that I was able to accomplish in his lab.

I am grateful for the support of Dr. Sangita Sinha. Taking her comprehensive biochemistry course renewed my interest in biochemical research and launched my curiosity and appreciation for structural biology. Her dedication to her work is inspiring, and I am thankful for her drive to instill critical thinking in all students that pass through her lab. She also provided significant advice and resources for many aspects of my projects.

Many thanks to my other committee members- Dr. Kent Rodgers, Dr. Glenn Dorsam, and Dr. Birgit Pruess. They have been incredibly supportive and encouraging since my early graduate student days, and provided suggestions and advice for my work throughout the years. I am especially grateful to Dr. Rodgers- you asked the tough questions and required perceptive answers, but I am a better scientist for it.

I am very grateful to have had the assistance of many collaborators on my projects: Dr. David Neau at NE-CAT at the Advanced Photon Source (APS); Dr. Srinivas Chakravarthy at Bio-CAT at the APS; Andrea Balbo and Dr. Huaying Zhao, both from the National Institute of

Biomedical Imaging and Bioengineering at the NIH; Dr. Qiong Wu in the NMR lab at UT Southwestern; Dr. Chad Brautigam at UTSW; and Drs. Angel Ugrinov, Ganesh Bala, and John Bagu from the Materials Characterization Lab, the Core Synthesis and Analytical Services facility, and the NMR lab, respectively, at NDSU. All provided their time and expertise to assist us, and in some cases, performed experiments when we did not have access to some instruments.

My appreciation goes to my labmates over the years, especially Dr. Minfei Su and Dr. Yang Mei, who were around when I started and provided a sounding board for ideas and recommendations when I had to troubleshoot various problems. Karen, Yue, Shane, Ben, Shreya, Srini, Beau, and Lizzy, as well as other students in the Chem and Biochem department and various undergrad assistants, past and present, were there through the ins and outs of the graduate school odyssey. We saw it all and lived to toil another day! Best of luck to Beau, who is continuing the cell-surface signaling work.

Extra special thanks to those who saw me through outside of the lab. To Jenny, Kip, Chris, Linnea, Meghan, Adrian, and my little buddy, Mia, your friendship has been invaluable. Much thanks to my fellow trainees at Red River Traditional TaeKwonDo for the sweat and stress relief, and to Grand Master Jim for his encouragement, his confidence in me, and his efforts to improve his students' form and character inside and outside the dojang.

Last, but certainly not least, thanks to my family, both immediate and extended, for their support, encouragement, understanding (as much as possible), and giving me the space to pursue my random intellectual interests. Much love and gratitude to all the Jensens, Saias, Frelichs, Andersons, Brodens, and McKewens.

DEDICATION

For my parents, Jon and Marion Jensen, for their wholehearted support in this endeavor-

Amare et sapere vix deo conceditur,

and

In memory of my mother, Suzanne Margaret Saia Jensen

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
APS	Advanced Photon Source
ASD	Anti-sigma domain
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
βME	β-mercaptoethanol
BMRB	Biological Magnetic Resonance Bank
CC	Correlation coefficient
$CC_{1/2}$	Pearson correlation coefficient between random half-datasets
CCD	Charge coupled device
CCSSD	C-terminal cell-surface signaling domain
CD	Circular dichroism
CHAPSO	3-([3-Cholamidopropyl]dimethylammonio)-2-hydroxy-1-propanesulfonate
CJM	C-terminal juxtamembrane
CMC	Critical micelle concentration
CRE	Carbapenem-resistant Enterobacteriaceae
CS	Chemical shifts
CSS	Cell-surface signaling
CTD	C-terminal domain
Da	Dalton
DDFA	Differential detergent filtration assay
DDM	n-Dodecyl β-D-maltoside

DMPC	1,2-Dimyristoyl-sn-glycero-3-phosphorylcholine
DNA	Deoxyribonucleic acid
DOPE	Discrete optimized protein energy
DTT	Dithiothreitol
ECF	Extracytoplasmic function
E. coli	*Escherichia coli*
EM	Electron microscopy
ESI-MS	Electrospray ionization mass spectrometry
FC-13	Fos-choline 13
GFP	Green fluorescent protein
GI	Gastrointestinal
GSH	Glutathione Sepharose
GST	Glutathione S-transferase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horse radish peroxidase
HSQC	Heteronuclear single quantum correlation/coherence
HT	High-throughput
ICU	Intensive care unit
IM	Inner membrane
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
JCSG	Joint Center for Structural Genomics

KBM	King’s B medium
kD	kilodalton
LB	Luria Bertani
LDAO	Lauryldimethylamine N-oxide
LGA	Local-global alignment
LN_2	Liquid nitrogen
LPS	Lipopolysaccharide
MBP	Maltose binding protein
MHz	Megahertz
MPQS	ModPipe Quality Score
MR	Molecular replacement
MW	Molecular weight
MWCO	Molecular weight cutoff
NCS	Non-crystallographic symmetry
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser spectroscopy
NSD	Normalized spatial discrepancy
NTD	N-terminal domain
NTSD	N-terminal signaling domain
OM	Outer membrane
OMPs	Outer membrane proteins
PDZ	PSD95, Dlg1, zo-1 domain
Pb	Pseudobactin

PBP	Periplasmic binding protein
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDC	Protein-detergent complex
PEG	Polyethylene glycol
PI	Propidium iodide
PMF	Proton motive force
PVDF	Polyvinylidenedifluorine
R_{free}	Free R-factor
RIP	Regulated intramembrane proteolysis
RMSD	Root-mean-square deviation
RNAP	RNA polymerase
ROSET	Rotational surveillance and energy transfer
RSM	Rhizosphere medium
R_{work}	Working R-factor
ROS	Reactive oxygen species
SAD	Single-wavelength anomalous dispersion
SANS	Small-angle neutron scattering
SAXS	Small-angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SeMet	Selenomethionine

STN	Secretin and TonB N-terminus short domain
SUMO	Small ubiquitin-like modifier
SV-AUC	Sedimentation velocity analytical ultracentrifugation
TB	Terrific broth
TBDT	TonB-dependent transporter
TBS-T	Tris buffered saline + Tween
TEV	Tobacco etch virus
Tet	Tetracycline
TFE	2,2,2-trifluoroethanol
TLS	Translation, libration, screw-motion
TM	Transmembrane
TOCSY	Total correlated spectroscopy
WT	Wild type

LIST OF SYMBOLS

a	height of the micelle
b	cross sectional axis of the detergent-micelle complex
$c(s)$	differential sedimentation coefficient distribution
D_{max}	maximum particle dimension
e	torus ellipticity
$\mathrm{Fe^{2+}/Fe^{3+}}$	ferrous/ferric iron
g	relative centrifugal force (RCF)
$\mathrm{I(0)}$	scattering intensity at zero-angle
I	scattering intensity
$I/\sigma I$	signal-to-noise ratio
$\mathrm{K_a}$	association constant
$\mathrm{K_d}$	dissociation constant
k_{ex}	rate exchange constant
mAU	milli-absorbance units
$P(r)$	pairwise-distance distribution function
q	momentum transfer
R_g	radius of gyration
S	s-value (Svedberg units)
s_W	weight average s-value
t	height of the hydrophilic region of a detergent belt
$\mathrm{T_m}$	melting temperature
ΔG	change in Gibbs free energy

ΔH	change in enthalpy
ΔS	change in entropy
$\Delta\omega$	difference in angular frequency
ε	extinction coefficient
θ	angle between incident X-rays, or degree ellipticity
ρ	Pearson correlation coefficient
σ	transcriptional regulator (sigma factor)
λ	wavelength
χ	chi value (goodness of fit)
ϕ	C(i-1),N(i), Ca(i),C(i) torsion angle
ψ	N(i),Ca(i),C(i),N(i+1) torsion angle
Å	angstrom

1. INTRODUCTION

1.1. The Post-antibiotic Era

The events of the 20th century realized many of humanity's most significant changes across all aspects of life, from the social and political to the technological and medicinal. Some of the greatest achievements of the era came with exponential developments in science and technology, resulting in space exploration, nuclear power, computational science, and the progression of modern medicine. The transition from the 19th to the 20th centuries, especially, saw a distinct transformation in quality of life through sanitation efforts, serological testing, and the implementation of routine vaccinations. The discovery and prescription of antibiotics, in particular, resulted in a sharp decline in infant and child mortality, as well as the extension of life expectancy by nearly three decades (1); this represented a marked change from deaths by infectious diseases to deaths from degenerative diseases, such as diabetes, heart disease, and cancer (Figure 1.1). Prior to the introduction of antibiotics into contemporary medical intervention, five out of every 1000 mothers died during childbirth, one-third of individuals who contracted pneumonia died, and one-tenth of those who presented with a skin infection required amputation of a limb (2). In addition to directly combating infections, antibiotics have been a critical player in other major medicinal advances. Without antibiotics, surgeries of any kind, organ transplants, or usage of medical devices such as catheters or pacemakers would be extraordinarily risky. In today's society, use of antibiotics extends beyond chemotherapy and prophylaxis; antimicrobials are now included in cleaners, soaps, lotions, textiles, and over 700 other household products (3). These biocides are incorporated to prevent transmission of pathogenic microorganisms to noninfected or immunocompromised individuals, instead of treating an existing infection. However, these agents still possess targeted antibacterial activity.

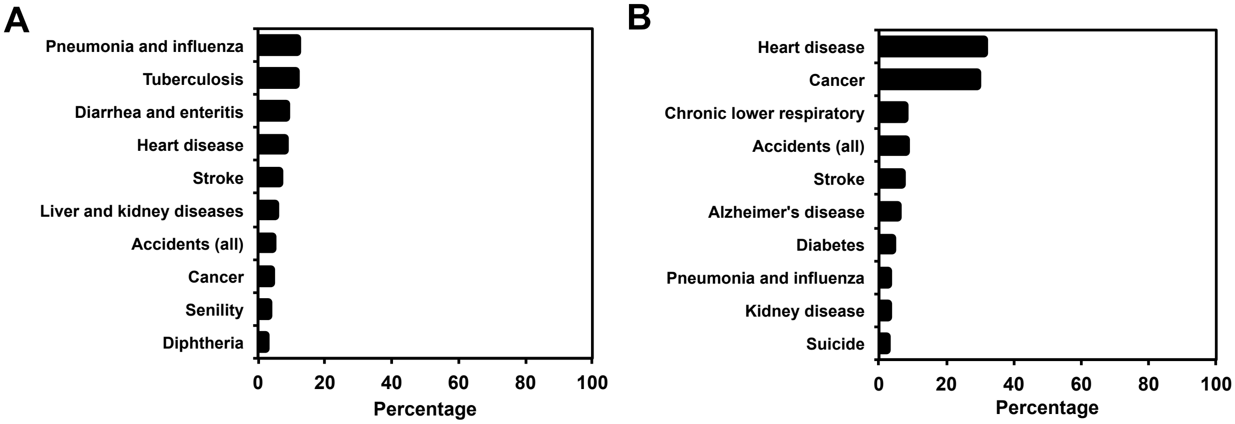


Figure 1.1. The ten leading causes of deaths as a percentage of all reported deaths for the years 1900 and 2015. (A) The top three causes of death for 1900 were infectious diseases. (B) By 2015, the leading causes of death had shifted to degenerative diseases. (Adapted from CDC data (4) and (5)).

Despite the enormous success of antibiotics on the reduction of infections and mortality, serious concerns plague continued usage of these molecules. These considerations were prominently noted by Alexander Fleming, the discoverer of penicillin, in his 1945 Nobel Prize in Medicine acceptance speech: “It is not difficult to make microbes resistant to penicillin in the laboratory by exposing them to concentrations not sufficient to kill them. There is the danger that the ignorant man may easily under-dose himself and by exposing his microbes to non-lethal quantities of the drug make them resistant” (6). Fleming’s conjecture was proven correct, as microbes resistant to penicillin were first detected even before widespread usage of the drug in 1943. As antibiotic utilization became commonplace throughout the 20th century, the timeline between development of new antibiotics and detection of resistance in microbial populations was significantly decreased (Figure 1.2). In one such example, resistance to Levofloxacin, a broad-spectrum fluoroquinolone typically prescribed for the treatment of respiratory tract infections, urinary tract infections, tuberculosis, and meningitis, among other infections, was first detected in 1996- the same year that the drug was approved for use in the United States (7).

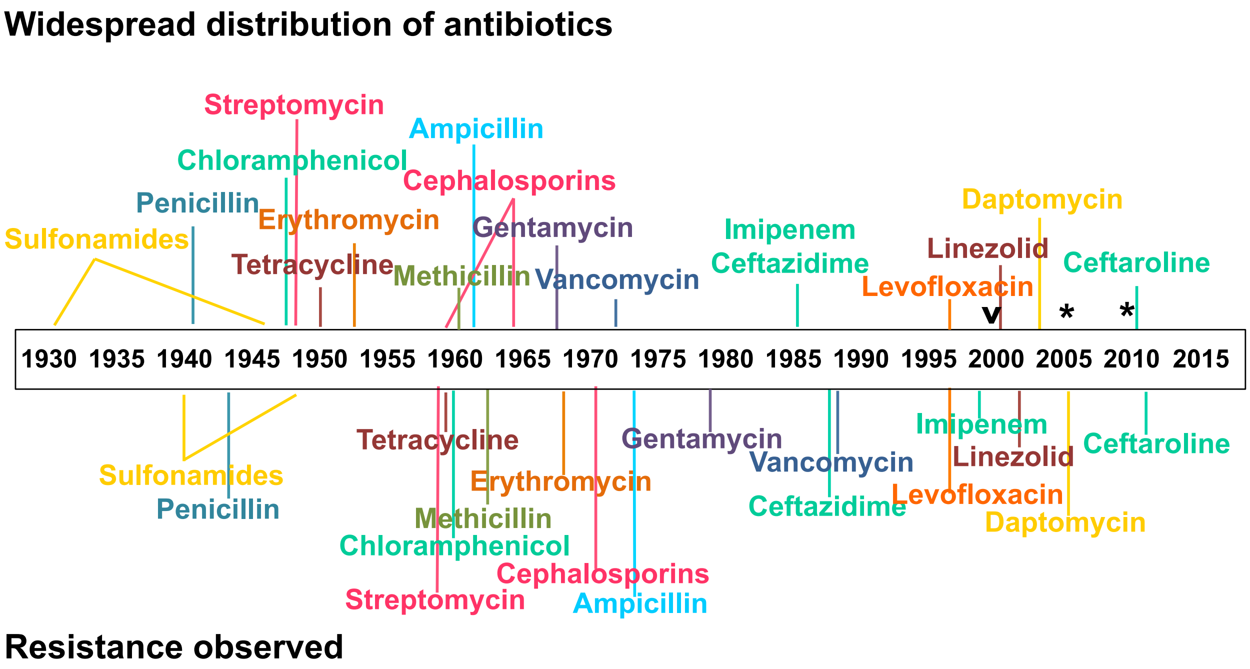


Figure 1.2. Timeline of initial antibiotic prescription vs. observations of resistance. (Adapted from (7)). $^{\mathrm{V}}$ indicates emergence of extensively drug-resistant organisms; * indicates observation of pan-drug-resistant organisms.

Resistance to antibiotics represents a grave health concern. Recent estimates indicate that drug resistant bacteria cause over two million infections in humans per year in the US (7). Additionally, infections by these microorganisms result in nearly 25000 deaths per year, and healthcare costs alone associated with these infections pose a significant economic burden, estimated to be around twenty billion dollars, not accounting for loss of productivity (7-9). Given the recession in novel antibiotic development, as well as increasing bacterial drug resistance, many healthcare professionals have designated humanity’s position within the chronology of modern medicine as the “post-antibiotic era” (10).

1.1.1. Causes of Antibiotic Resistance

Several factors have been implicated in the rapid escalation and emergence of microbial populations resistant to antibacterial therapeutics. Bacteria may possess inherent resistance through spontaneous genomic mutations, which are subsequently passed to daughter cells through binary fission. Thus, bacterial colonies adapt to antibiotic exposure through survival and proliferation of resistant individual organisms that retain selective mutations, allowing them to swiftly overwhelm susceptible populations as a result of vertical evolution by natural selection (11). Bacteria may also acquire resistance through horizontal gene transfer (horizontal evolution) via mobile genetic elements, such as plasmids, from related or unrelated species. Both of these methods require selective pressure to propagate resistant bacterial populations, and as such, overuse of antibiotics is the primary determinant of emergence of resistance (8). Several epidemiological studies have exposed a direct correlation between utilization of antibiotics and observation of resistant bacterial strains (12). In the US, over-prescription of antibiotics is endemic; incredibly, one study revealed that, on average in 2010, 22 doses of antibiotics were prescribed per person (13). In many countries, antibiotics are easily accessible as they are unregulated and inexpensive. Over-prescription also derives from a lack of appropriate diagnostics, so that broad spectrum antibiotics are used before the infectious agent has been identified; in many cases, antibiotic consumption is either unnecessary or substandard, which also contributes to the spread of resistance (14).

Overuse of antibiotics extends beyond the hospital or clinic to agriculture. Over 80% of antibiotics sold in the US are used in livestock production; antibiotics are fed to livestock to prevent infection and improve the health of the animal, producing greater growth and higher yields (15). Residual antibiotics may be present when the animal products are ingested by humans, and the livestock excrete a high percentage of ingested drugs, which are then disseminated into the soil and groundwater (7). Resistant bacteria may also contaminate meat products or improperly washed produce, and are subsequently transmitted directly to the consumer.

The expansion of antibiotic resistance is compounded by the decline of development of new antibiotics. Pharmaceutical companies increasingly invest limited resources in novel antibacterial drugs, as the sale of chemotherapeutics for chronic conditions, such as diabetes, is considerably more profitable (16). Given the cost of bringing a new drug to market, which may easily exceed several millions of dollars, the return on investment may be roughly 5% of a new drug for a degenerative disease (17), and since the rate at which bacteria may develop resistance to an antibiotic is unpredictable, a new antibacterial agent takes years to characterize and license, but may only be effective for a limited timespan.

1.2. Gram-negative Bacteria and Human Health

1.2.1. Features of the Gram-negative Bacterial Cell Wall

Gram-negative bacteria, classified based on their inability to retain crystal violet as part of the Gram stain method for bacterial differentiation, encompass a wide variety of ubiquitous bacteria- both commensal and pathogenic. In contrast to Gram-positive bacteria, Gram-negative microbes have two membranes- an outer membrane (OM) composed of phospholipids and glycolipids or lipopolysaccharides (LPS) at the inner and outer leaflets, respectively, and a cytoplasmic or inner membrane (IM) bilayer of diverse phospholipids, such as phosphatidylethanolamine, phosphatidylglycerol, phosphatidylcholine, and cardiolipin, although the lipid composition of the IM varies amongst bacterial species (18,19). Both membranes are highly complex and heterogeneous, and contain assorted transmembrane or membrane-anchored proteins, and often include lipid microdomains. The OM and IM are separated by the periplasm, an aqueous but dense compartment with a thin peptidoglycan layer, and varying proteins and ions (Figure 1.3) (20). By contrast, Gram-positive bacteria possess a single phospholipid bilayer covered with a thick peptidoglycan lamina (Figure 1.3).

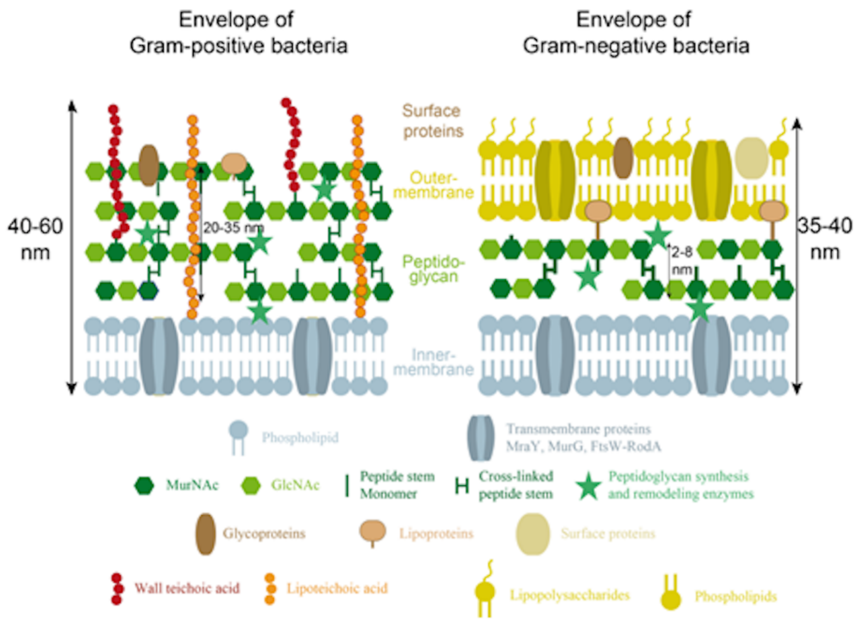


Figure 1.3. Schematic of the cell walls of Gram-positive and Gram-negative bacteria (21).

1.2.2. Considerations for Development of Antibiotics Targeting Gram-negative Bacteria

Generally speaking, Gram-negative bacteria are more resistant to antibiotics than Gram-positive bacteria (22). The intricacies of the Gram-negative bacterial cell wall present unique complications for antibiotic design. The existence and complexity of the OM provide a protective barrier for the bacterial cell against toxic compounds. Since the majority of antibiotics have intracellular targets, these compounds must be able to traverse both the OM and the IM. Permeation of the OM is complicated by the high hydrophobicity of the bilayer; passive diffusion through a porin (for small molecules) may occur (23,24), and active transport through a ligand-gated channel protein (for higher molecular weight molecules) is also exploited. Most of these OM proteins (OMPs) are β-barrels with variable numbers of β-strands. Available antibiotics that selectively target Gram-negative bacteria are summarized in Table 1.1, although the activity of many are limited to particular Gram-negative species, and resistance has been observed for all.

Table 1.1. Antibiotic targets and associated cell death mechanisms of Gram-negative bacteria (22,25).

Antibiotic class	Action	Bacterial cell target	Outcome	Example	[b]Activity
Quinolones	Interferes with DNA supercoiling	Topoisomerase II or IV	dsDNA breaks and arrest of DNA repair	Ciprofloxacin, levofloxacin	Bactericidal
Monobactams	Inhibits cell wall synthesis	Penicillin-binding protein (PBP)	Disruption of peptidoglycan synthesis	Aztreonam	Bactericidal
Aminoglycosides	Inhibits protein synthesis	30S ribosomal subunit	Mistranslation of mRNA	Streptomycin, kanamycin	Bactericidal
[a]Carbapenems	Inhibits cell wall synthesis	Penicillin-binding protein (PBP)	Disruption of peptidoglycan synthesis	Imipenem, Doripenem	Bactericidal
Cephalosporins (third generation)	Inhibits cell wall synthesis	Penicillin-binding protein (PBP)	Disruption of peptidoglycan synthesis	Cefotaxime, Ceftazidime	Bactericidal
Cephalosporins (fourth generation)	Inhibits cell wall synthesis	Penicillin-binding protein (PBP)	Disruption of peptidoglycan synthesis	Cefepime	Bactericidal
Cephalosporins (fifth generation)	Inhibits cell wall synthesis	Penicillin-binding protein (PBP)	Disruption of peptidoglycan synthesis	, Ceftobiprole	Bactericidal
Rifamycins	Inhibits DNA-dependent transcription	RNA polymerase	Inhibition of RNA synthesis during elongation	Rifampicin	Bactericidal
Macrolides	Inhibits protein synthesis	50S ribosomal subunit	Inhibition of translocation of tRNA	Erythromycin, azithromycin	Bacteriostatic
Chloramphenicol	Inhibits protein synthesis	50S ribosomal subunit	Prevention of protein chain elongation during translation	Chloromycetin	Bacteriostatic
Sulfonamides	Prevents growth and replication	Dihydropteroate synthetase	Inhibition of folate synthesis	Sulfanilamide, sulfadiazine	Bacteriostatic
Tetracyclines	Inhibits protein synthesis	30S ribosomal subunit	Inhibition of aminoacyl-tRNA binding to mRNA-ribosome	Tetracycline, doxycycline	Bacteriostatic
[a]Polypeptides	Disrupts cell membrane	Outer, inner membranes	Disruption of LPS, solubilization of membrane	Bacitracin, colistin, polymyxin B	Bactericidal

[a] "Last-resort" antibiotics

[b] Bactericidal antibiotics kill bacteria; bacteriostatic antibiotics inhibit bacterial growth and division.

1.2.3. Gram-negative Bacteria as Important Human Pathogens

Bacterial colonization of the human body is pervasive; the ratio of bacterial cells to human cells is approximately 1.3:1, according to a recently revised assessment from the conventional 10:1 estimate (26). While several Gram-negative species are commensal microbial flora, their localization is restricted to specific areas of the body, including the skin and mucosa, upper and lower gastrointestinal tracts, and respiratory tract (27). Certain microbes may become opportunistic pathogens when the normal microbiota is altered, the host has a weakened immune system, or integumentary barriers are breached, in the case of wounds or insertion of a medical device (28). Gram-negative bacteria are known to cause multiple types of infections via diverse mechanisms. Common species are *Escherichia coli*, a common cause of foodborne illness, gastrointestinal and urinary tract infections; *Neisseria gonorrhoeae* or *N. meningitides*, the causative agents of gonorrhea and meningitis, respectively; and *Pseudomonas aeruginosa*, which may elicit a variety of infections, including those of the skin, and GI, urinary, or respiratory tracts (29). *P. aeruginosa* respiratory infections are especially prevalent in individuals with cystic fibrosis (30-37). Gram-negative bacteria account for over 30% of nosocomial infections, and are the most common causal agents of infections in the ICU (9,38,39).

In 2013, the Centers for Disease Control and Prevention released a threat assessment of notable pathogens, categorized as “urgent,” “serious,” and “concerning.” Of the 18 total microorganisms described in the report, two Gram-negative bacteria are classified as “urgent,” and six as “serious.” All of the Gram-negative pathogens described are multi-drug resistant; in the case of carbapenem-resistant Enterobacteriaceae (CRE), resistance has also been detected for “last resort” antibiotics, leaving few or no treatment options, resulting in the deaths of half of patients who develop sepsis from CRE (7). The danger to human health from Gram-negative pathogens and multi-drug resistant Gram-negative pathogens, in particular, is undeniable. The combination of nosocomial infections from multi-drug resistant bacteria is especially troubling.

1.3. Bacterial Survival, Pathogenesis, and Dependence Upon Iron Availability

Efforts toward development of novel antimicrobials to treat multi-drug resistant Gram-negative bacteria target critical bacterial survival pathways. Survival of the microorganism, and thus, its ability to cause disease, is dependent on the microbe’s capacity to acquire nutrients from its environment. One such element, iron, is an essential cofactor and prosthetic group in many enzymes that regulate basic biological functions, such as respiration, DNA replication, and regulation of gene expression; over 100 enzymes involved in primary and secondary metabolism contain iron as part of iron-sulfur clusters or heme (40). The reversible redox properties of Fe^{2+}/Fe^{3+} at physiological conditions make it an ideal cofactor. Iron is pervasive in the environment, and ferrous iron is soluble in aqueous solutions at neutral pH, but its availability for biologics is low, as it has poor solubility in the ferric form, which is most prevalent in aerobic conditions. Additionally, ferrous iron readily reacts with hydrogen peroxide to form reactive oxygen species. In the mammalian host, iron sequestration by common proteins such as ferritin, lactoferrin, and transferrin, represent a critical innate immune response to infection. This leads to a blood serum iron concentration of 10^{-24} M (41), while many bacteria require concentrations around 10^{-7} M to survive (42).

Iron regulation has been implicated in Gram-negative bacterial pathogenesis, as overcoming iron limitation in the host results in virulence that promotes infection and survival (43). A mouse model of *P. aeruginosa* lung infectivity demonstrated a 10% survival rate after 14 days amongst mice infected with wild-type (WT) *P. aeruginosa* PAO1, while PAO1 strains lacking genes responsible for iron acquisition (both transport and regulation) resulted in an avirulent phenotype, suggesting that iron uptake and regulations contributes to virulence (44). In *E. coli* and *P. fluorescens*, multiple genes that abet iron homeostasis are also virulence factors associated with infection (45,46). Iron-controlled genes have been identified in *Vibrio cholerae* that, upon mutation, decrease virulence of the bacterium (47-49). In *P. aeruginosa*, the sigma factor PvdS regulates the production of the secreted virulence factors exotoxin A, PrpL endoprotease, and AprA metalloprotease (50,51). Interestingly, it has been demonstrated in *P. aeruginosa* and *P. putida* that antibiotic efficacy is influenced by iron concentration through oxidative stress management (52). Efflux of iron-containing molecules restricts bacterial growth but may also reduce susceptibility to oxidative and nitrosative stress, and to diverse antimicrobials (53). Mutations of *fur,* which encodes for the ferric uptake regulator, a genome-wide transcriptional repressor, exhibit reduced virulence in various animal models of infection, indicating a role for Fur in iron homeostasis and pathogenesis (54). Iron homeostasis must be precisely monitored, a process which is facilitated by the regulation of bacterial iron acquisition pathways.

1.4. Gram-negative Bacterial Iron Acquisition Pathways

Gram-negative bacteria acquire iron through several similar conduits. Although some abundant nutrients may passively diffuse into the periplasm by ion channels, metals and larger molecules must be actively transported into the periplasm, and subsequently, the cytoplasm. Bacteria must either scavenge iron cofactors from extracellular proteins, or release then transport self-synthesized iron-chelating molecules back into the cell. Import into the periplasm requires an electrochemical gradient, or proton motive force (PMF), to facilitate active transport, as ATP is not present in the periplasm (55).

1.5. The TonB-Dependent Iron Transport and Regulatory System

TonB-dependent transporters (TBDTs) are a class of bacterial iron acquisition proteins. Located in the outer membrane of Gram-negative bacteria, these high affinity active transporters are responsible for the binding and transfer of >600 Da ferric complexes (siderophores or heme), in addition to nickel complexes, vitamin B_{12}, and carbohydrates, into the periplasmic space (Figure 1.4). Since siderophore receptors concentrate iron against a gradient, TBDTs require energy via the proton motive force through coupling with the inner membrane TonB complex, TonB-ExbB-ExbD, for transport across the outer membrane. Expression of TBDTs is tightly regulated due to iron toxicity, as ferrous iron (Fe^{2+}) may readily react with cellular hydrogen peroxide, via the Fenton reaction, to form detrimental reactive oxygen species, such as hydroxyl radicals and superoxide (56).

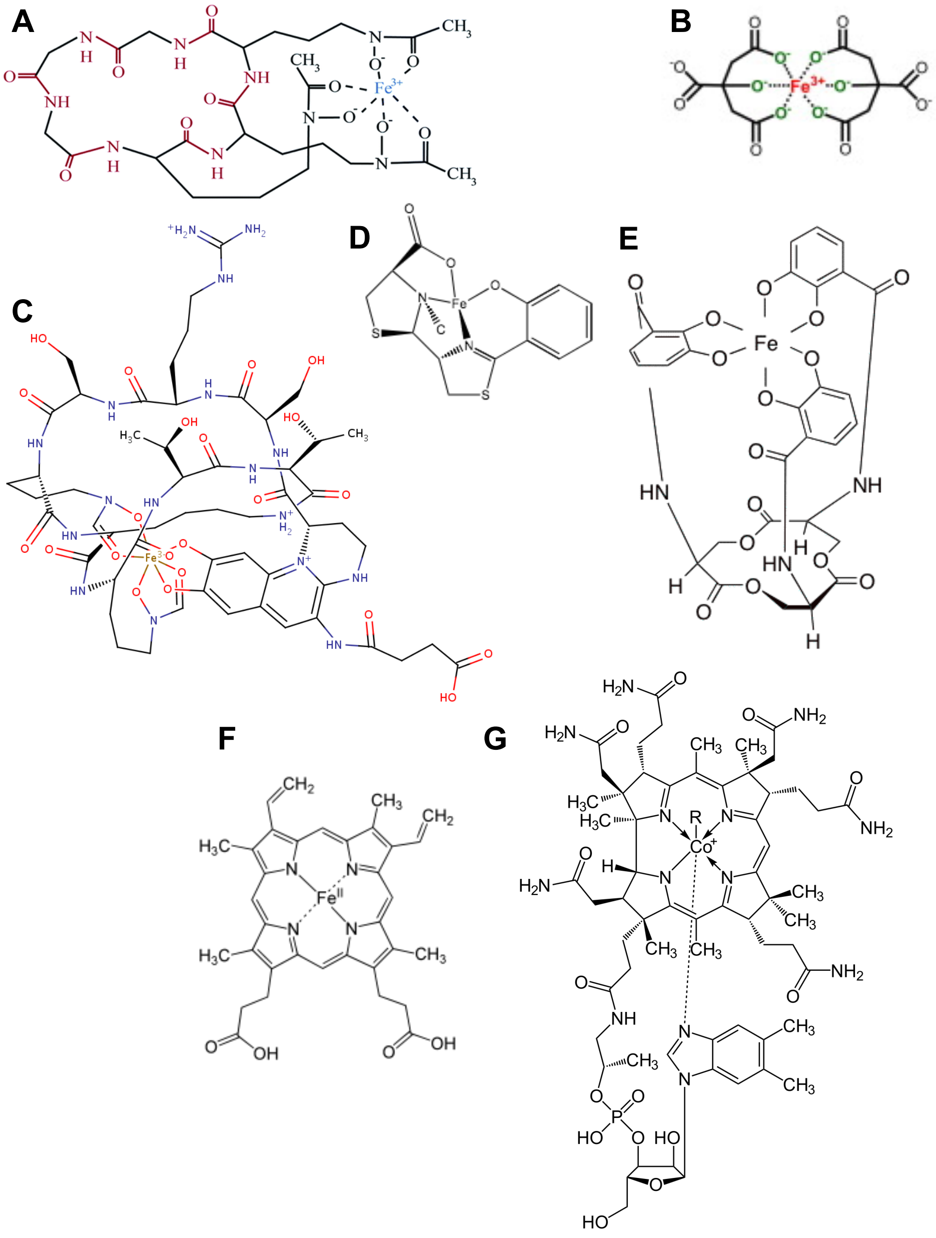


Figure 1.4. Various metallo-molecules transported by TBDTs. (A) Ferrichrome, (B) Ferric citrate, (C) Pyoverdine, (D) Pyochelin, (E) Enterobactin, (F) Heme B, (G) Cobalamin (vitamin B_{12}). Siderophores, high affinity metal chelators, are synthesized by the bacterium and released into the extracellular space to bind available iron. Siderophores may contain a variety of moieties for iron coordination, such as catecholates (enterobactin), phenolates (pyochelin), hydroxamates (desferrioxamine), carboxylates (achromobactin), or mixed types (citrate-hydroxamate, aerobactin) (40).

All characterized siderophores possess a higher affinity for ferric iron (Fe^{3+}) than ferrous iron (Fe^{2+}). Many microorganisms synthesize and secrete more than one siderophore; *P. capeferrum* may utilize pyoverdine, as well as at least ten pseudobactins (57). Additionally, some bacteria are able to use xenosiderophores- siderophores produced by other microbes present in their environment, as microbial competition is high for this limited nutrient (58). Each TBDT recognizes and transports a specific siderophore or structurally related siderophore. After uptake into the periplasm, the ferric siderophore may be reduced to release free iron, in which case the apo siderophore is recycled into the extracellular environment by an efflux pump and the reduced iron is shuttled into the cytoplasm by binding to a periplasmic binding protein (PBP) domain of an ATP-binding cassette (ABC) transporter (59-61). Similarly, ferric siderophore is translocated into the cytoplasm by interaction of the siderophore:PBP complex with the permease of the ABC transporter; this transport process is coupled to ATP hydrolysis (62). Bacterial ABC transporters typically consist of five domains: a PBP, a channel for siderophore translocation formed by two transmembrane polypeptides, and two ATPases (63). ABC transporters are able to bind a variety of ligands. Once the ferric siderophore enters the cytoplasm, iron may be released by a number of proposed mechanisms: siderophore hydrolysis, proton-assisted dissociation of the iron:siderophore complex, or reduction of the iron center (64,65). Several cytoplasmic enzymes have been identified that couple siderophore hydrolysis and iron reduction for dissociation of the iron from the siderophore (66-68). Following iron reduction and release, it may be incorporated directly into metallo-enzymes, or other metalloproteins, such as ferritin, for storage (63).

While the transport mechanism of TBDTs has yet to be fully elucidated, a topic of even greater ambiguity is the signal transduction pathway originating with ferrisiderophore uptake and concluding with the initiation of transcription of genes involved in iron transport; this pathway is one of the primary methods of control of TBDT expression, and subsequently, iron uptake under iron starvation conditions (Figure 1.5). This pathway typifies a critical survival mechanism for both benign and pathogenic bacteria, and as such, represents an attractive target for novel antimicrobial therapeutics, as was recently demonstrated through pharmacological disruption of the ferripyoverdine regulatory system in *Pseudomonas aeruginosa*, which reduced pathogenicity of the microbe *in vivo* (69).

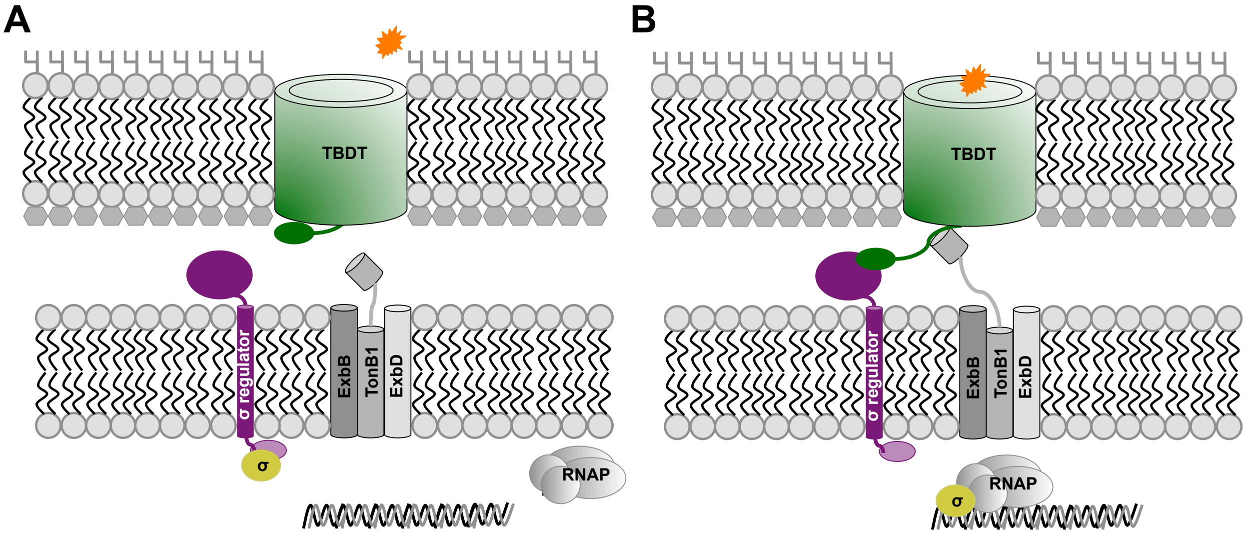


Figure 1.5. The canonical TBDT regulatory pathway. (A) The regulatory system at rest. The apo TBDT (green) does not bind to either the TonB complex (shades of grey) or the inner membrane sigma regulator protein (purple), and the sigma factor (yellow) is thus sequestered to the inner membrane by the sigma regulator. Transcription of target genes has not been initiated. (B) The regulatory pathway is activated by binding of ferric siderophore (orange star) to the TBDT. This results in association of the TonB box of the TBDT with the inner membrane TonB complex, which initiates transport of the siderophore. Concurrently, the NTSD of the TBDT contacts the periplasmic region of the sigma regulator, and the sigma factor is released. The sigma factor binds the promoter of the target response gene (the TBDT), and recruits RNA polymerase for initiation of transcription.

1.5.1. Structural Features of TBDTs are Conserved

Despite low primary sequence similarity, TBDTs of known structure possess conserved structural features (70); these include a 22-stranded antiparallel transmembrane β-barrel that is 55-70 Å high, with an inserted plug domain (which contains a variety of conserved motifs) (70), in addition to a semi conserved region near the N-terminus (71). This region, designated the TonB box (or Ton box), contacts the transperiplasmic TonB complex during ferrisiderophore transport. Residues of the extracellular loops of the TBDT involved in siderophore recognition and binding are specific for their target ligand and are poorly conserved, indicating evolutionary divergence (70,72). Conservation of residues of the TBDT β-barrel, plug domain, and TonB box suggest that the TBDTs integrate homologous transport mechanisms for ligand uptake, and in some cases, may bind and transport heterologous or xenosiderophores (73). Point mutations within the TonB box or deletion of this region abolish siderophore import (74-76).

1.5.2. Transport May Induce Conformational Changes to the TonB Box

Great variation is observed in substrate uptake by TBDTs, although specific siderophores or families of siderophores are typically transported exclusively by their target receptors. Extracellular loops function to trap and secure ferrisiderophore to the plug domain. Binding of ferric substrate involves hydrogen bonding of the iron-complexed portion of the siderophore to specific residues of the plug domain, while the remainder of the siderophore is solvent accessible, as observed with the ferrichrome-free and bound structures of FhuA. After siderophore binding, the most significant conformational changes manifest on the periplasmic side of the plug domain and in the TonB box, although these alterations are not homologous for TBDTs; few changes have been observed in the structural features of the β-barrel. In the holo-form of FhuA, one α helix of the plug domain is unfolded (77).

Experimental evidence suggests that interaction of the TonB box with the TonB complex permits movement of the plug domain, which releases the bound siderophore into the periplasmic space. The TonB box of BtuB, a transporter of vitamin b_{12}, undergoes a substrate-dependent “folding-to-unfolding transition” in the presence of ligand that is also observable in

FecA (78); the TonB box of FhuA appears to be a highly flexible and unstructured region both with and without substrate (77). In the case of BtuB, unfolding of the TonB box upon substrate binding has been shown via EPR to extend the TonB box region of the N-terminus an additional 20-30 Å into the periplasmic space (79). Given the structural homology of these proteins, the disparity of conformation of the TonB box is surprising and indicates that mechanisms other than conformational switches of the TonB box may ultimately be responsible for transmembrane transport and signaling in some TBDT systems (70). Additionally, substrate binding to specific receptors may signal through different periplasmic or inner membrane regulatory proteins.

Limited structural information exists to corroborate specific TonB interactions with active siderophore transport. The original structure solution of the FpvA-ferripyoverdine complex displayed an amino N-terminal signaling domain (NTSD) folded with the rest of the receptor; furthermore, the TonB box was not observed in the electron density (80). The fold of the FpvA NTSD displays structural similarity to the NTSD of FecA and PupA (which share 25% sequence identity) (81). Further structural studies revealed the TonB box is within a β-strand that is part of a mixed four-stranded β-sheet with the three-stranded β-sheet of the NTSD; however, in this conformation, the TonB box is buried within the receptor and cannot associate with TonB. Further conformational changes must occur for the NTSD to unfold into the periplasmic space. Binding of ferripyoverdine to FpvA induces a conformational change resulting in a flexible orientation of the TonB box, which allows for a "β lock exchange" between an extra β-strand of the NTSD, formed by the TonB box, and TonB, to configure the transient 4-stranded β-sheet observed in the TonB:TonB box interaction (82). This positioning of the NTSD and flexibility of the N-terminus of the transporter provides the necessary structural arrangement for TonB binding and active siderophore transport into the periplasm, without complete displacement of the plug domain. The NTSD is not conserved across TBDTs; in several TBDTs, the domain does not exist, so the TonB box is somewhat more accessible.

Structures of BtuB, FecA, and FhuA have been solved in complex with a C-terminal portion of TonB, and crystallization of the complexes reveals disulfide crosslinks and a salt-bridge coupling TonB with the TonB box (79). The BtuB plug domain unfolds in a stepwise and localized manner (83), and association of the TonB box of BtuB to the TonB complex across the periplasm reconfigures an extracellular loop of the transporter; thus, TonB box interaction with the TonB complex and substrate binding appear to be allosterically coupled (84).

The C-terminal domain (CTD), consisting of the C-terminal 144 residues of TonB, has been characterized by solution NMR, and a truncated version of this domain (76 residues) was solved by crystallography as a homodimer (79,85-90). The TonB CTD crystallized as a dimer, although further work demonstrated that dimerization is not essential for TonB function (91). Recent structural studies of the Ton components and the fully assembled complex revealed the stoichiometry of the Ton interaction (92). ExbB is a pentamer, with a centralized transmembrane pore for proton passage. ExbD forms a dimer, with part of one monomer fit within the ExbB pore, and the second monomer outside of the ExbB pentamer ring. The full Ton complex contains at least one TonB protein, which is located outside of the ExbB pentamer. TonB association does not structurally affect the ExbB:ExbD oligomer, but the interaction of TonB with ExbD is necessary for a functional complex. The complex is highly sensitive to changes in pH, and is cation selective. The structural and electrostatic properties of the Ton complex lead the authors to propose two primary mechanistic models for PMF mobilization for energy production (Figure 1.6). The “electrostatic piston” model relies on translational movement of the transmembrane helix of ExbD. In the “rotational” model, the ExbD helix rotates within the ExbB pore (92,93), similar to the bacterial flagellar motor. The actual movement of the Ton complex is likely a combination of these models.

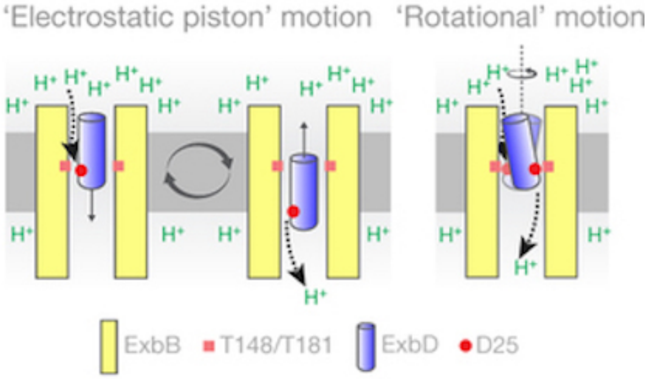


Figure 1.6. Proposed mechanisms for proton trafficking through the Ton complex. Critical residues for harnessing the PMF are highlighted in red. (From (92)).

1.5.3. Energy Requirements for Siderophore Transport

Common active transport mechanisms utilize ion gradients (Na^+ or K^+) or ATP hydrolysis across the cytoplasmic membrane to stimulate transport. Gram-negative bacteria do not possess these resources in the periplasmic space, and so must utilize the proton gradient harnessed by the TonB system to import nutrients that are too large to enter the cell via passive diffusion. In *E. coli*, FepA, the ferric enterobactin transporter, appears in the cellular envelope in much greater concentrations than TonB (~35:1). Passage of one molecule of ferric enterobactin transpires in 10-15 seconds, based upon post-uptake binding measurements, with a maximum turnover number of ~5/minute, although the turnover number of the receptor is low (3-5 minutes), due to the disparity between FepA and TonB concentrations (94). Some FepA are therefore inactive when bound with substrate while awaiting TonB coupling and energy input, estimated at 33-35 kcal/mol. These values represent the first attempt at approximating the energy requirements for transportation of a single ferrisiderophore across only the bacterial outer membrane, which delineates the complete transport pathway's rate-limiting step, based on estimations of total siderophore uptake and release into the cytoplasm (94). The cell's total energy expenditure for the entire uptake and signaling process is unknown. The proposed rotational surveillance and energy transfer (ROSET) model couples the rotational motion of the dimeric-monomeric TonB-CTD equilibrium with the electrochemical proton gradient of the proton motive force mediated by the ExbBD complex (93). When the monomeric form of the TonB-CTD encounters a TonB box from an outer membrane transporter rendered flexible and accessible by ligand binding, a 4-stranded β-sheet is formed between the TonB-CTD and the TonB box. The kinetic energy of the binding reaction and the rotation motion of the TonB-CTD, driven by the proton gradient, triggers iron uptake (95). ExbD experiences cycles of proton depletion and replenishment as TonB is activated (96). This is accomplished through the adjustment of the TonB-CTD by ExbD so that it dissociates from ExbD to interact with the TonB box of the TBDT (97,98).

1.5.4. Conservation of TBDT Signaling Genomic Architecture Across Gram-negative Species

Genome-wide analyses of cell surface-signaling (CSS) systems have demonstrated that these pathways are widespread amongst Gram-negative bacteria (99). A subset of the TBDT regulatory systems possesses both transport and signaling capabilities. One such example is the ferric citrate transport or *fec* system, originally discovered in *E. coli* K-12. The organization governing this network includes an outer membrane signal transducer (transducers contain an NTSD); an anti-sigma transmembrane protein that transmits the signal from the periplasmic space to the cytoplasm; and an extracytoplasmic function (ECF)-subfamily sigma factor, which initiates transcription of the TBDT through promoter recognition and RNA polymerase (RNAP) recruitment. In *E. coli*, these proteins are designated FecA, FecR, and FecI, respectively. The prototypical genetic arrangement of the Fec regulon is the tandem sigma factor, anti-sigma factor, and transducer (71). Several diverse proteobacteria, such as *Pseudomonas aeruginosa, Pseudomonas putida, Klebsiella pneumoniae,* and *Nitrosomonas europaea* possess one or more Fec-type regulatory systems and genes encoding for the TonB-ExbB-ExbD complex (71). Further bacterial species contain analogous Fec regulons. Additionally, sequence divergence allows for a Fec-type system to regulate transport of ferric complexes other than ferric citrate.

The *fec* locus is transcribed via promoters upstream of f*ecI* and *fecA*, and both promoters may be repressed by Fe^{2+}-Fur (ferric uptake regulator), which is the ultimate iron sensor and regulatory element of this system (100). *fecIR* are transcribed under iron-limited conditions, when Fur repression is relieved. Transcription of *fecA* occurs when ferric citrate is bound to FecA, which is typically minimally transcribed in the absence of ferric citrate to maintain a restricted supply of FecA in the outer membrane (100). Release of FecI and recruitment of RNAP to the *fecA* promoter results in transcription of only the transducer gene (101). There is no evidence to support the autoregulation of *fecIR* (102). Genes immediately downstream of the *fec* or *fec*-type locus often encode for ABC-type transporters and/or proteins involved in siderophore biosynthesis (103,104).

1.5.5. Coupling Transport with Signaling

The ROSET model of iron complex transport, while rationalizing the position of the TonB box in terms of localization of the region for TonB-CTD interaction and subsequent movement, does not account for the location or possible interactions of the TBDT-NTSD throughout this process. Although molecular mechanisms explicating the interactions between the outer membrane TBDT and the inner membrane regulatory protein were previously uncharacterized, FecA signaling sequence mutants generated to explore reduction in outer membrane transport substantially decrease rates of transcription (105). While complete elimination of the NTSD in FecA does not affect transport (106), specific mutations leading to structural changes in the signaling domain reduce interactions with the inner membrane σ regulatory protein, which decreases transcription initiation (107). Transcriptional activation is also thought to involve structural uncoupling of the NTSD with the β-barrel and plug domain of the TBDT once substrate has been bound, and result in further association of the NTSD with the inner membrane regulatory protein (70,108). Clearly, interactions between the primary domains of the TBDT- the NTSD, the plug domain, and the β-barrel- influence both transport and signaling events; conflicting reports suggest that transport of siderophore is not always necessary for signal transduction, since gene expression analysis suggests that transcription initiation of these particular pathways is lessened but not completely inhibited (109). Additionally, in the case of the Fec system, initiation of transcription is not dependent on ferric citrate, FecA, or TonB (110,111). Thus, it remains possible that stochastic co-localization of the TBDT-NTSD and the periplasmic domain of the anti-sigma factor is sufficient to initiate transcriptional regulation.

1.6. Features of the TonB-Dependent Regulatory Pathway

The characterization of the primary players in the Fec system of *E. coli* has dramatically broadened conventional knowledge concerning the cell-surface signaling process leading to initiation of transcription by the RNAP core enzyme of the TBDT gene as a consequence of iron uptake. Studies of the analogous Fpv system from *Pseudomonas aeruginosa* and Pup system from *P. capeferrum* have also yielded significant findings, enhancing the field's cognizance of the functionality of related TBDT regulatory proteins.

1.6.1. The TBDT

X-ray crystal structures of several TonB-dependent transporters have been solved (Table 1.2). The majority of these TBDTs transport ferric siderophore complexes; however, several facilitate uptake of other metallomolecules. FecA from *E. coli* binds and transfers ferric citrate; BtuB, from *E. coli,* cobalamin; Cir, from *E. coli,* colicin I; HasR and ShuA, from *S. marcescens* and *S. dysenteriae*, respectively*,* heme; TbpA, from *N. meningitides*, human transferrin; FrbA/FetA, from *N. meningitides*, Fe^{3+}; ZnuD, from *N. meningitides*, Zn^{2+}; and FusA, from *Pectobacterium atrosepticum*, transports plant ferredoxin (see Table 1.1 for references). TBDTs contain several conserved structural features; namely, 13-50 residue N-terminal signal peptide that is cleaved from the mature protein during protein secretion and translocation; a 5-8 residue region, the TonB box, that directly interacts with the TonB protein to facilitate active transport; and a 22-stranded transmembrane β-barrel, occluded by a 60-110 residue plug domain. A select few TBDTs also possess an N-terminal signaling domain (NTSD), a 70-90 residue globular domain located between the signal peptide and the TonB box. TBDTs that have this domain are classified as TonB-dependent transducers, as the NTSD allows for signal transduction across the periplasmic space.

Table 1.2. Available X-ray crystal structures of known TonB-dependent transporters. Residue numbers correspond to the full-length proteins, including the signal peptides, as annotated by UniProt (http://www.uniprot.org).

Protein	**PDB entry**	**Organism**	**Transport compound**	**Ligand**	**Residues in structure**	**Residues in sequence**	**Residues in SP**	**Residues in TonB box**	**Residues in NTSD**	**Citation**
FhuA	2FCP	*E. coli*	Ferrihydroxymate	--	52-747	747	1-33	40-47	--	(112)
	1FCP	*E. coli*	Ferrihydroxymate	Ferrichrome + LPS	52-747	747	1-33	40-47	--	(112)
	1QKC	*E. coli*	Ferrihydroxymate	Albomycin	52-747	747	1-33	40-47	--	(113)
	1QJQ	*E. coli*	Ferrihydroxymate	Phenylferricrocin	52-747	747	1-33	40-47	--	(113)
	1QFG	*E. coli*	Ferrihydroxymate	--	52-747	747	1-33	40-47	--	(114)
	1QFF	*E. coli*	Ferrihydroxymate	Ferrichrome	52-747	747	1-33	40-47	--	(114)
	1FI1	*E. coli*	Ferrihydroxymate	Rifamycin + LPS	52-747	747	1-33	40-47	--	(115)
	2GRX	*E. coli*	Ferrihydroxymate	Ferricrocin + TonB	41-747	747	1-33	40-47	--	(90)
	4CU4	*E. coli*	Ferrihydroxymate	Microcin J25	53-747	747	1-33	40-47	--	(116)
	1BY3	*E. coli*	Ferrihydroxymate	--	53-747	747	1-33	40-47	--	(77)
	1BY5	*E. coli*	Ferrihydroxymate	Ferrichrome	51-747	747	1-33	40-47	--	(77)
FepA	1FEP	*E. coli*	Ferric enterobactin	--	33-746	746	1-22	34-41	--	(117)
FecA	1KMP	*E. coli*	Ferric citrate	Ferric dicitrate	128-774	774	1-33	56-63	34-113 (Not in structure)	(118)

Table 1.2. Available X-ray crystal structures of known TonB-dependent transporters. Residue numbers correspond to the full-length proteins, including the signal peptides, as annotated by UniProt (http://www.uniprot.org) (continued).

Protein	**PDB entry**	**Organism**	**Transport compound**	**Ligand**	**Residues in structure**	**Residues in sequence**	**Residues in SP**	**Residues in TonB box**	**Residues in NTSD**	**Citation**
FecA	1KMO	*E. coli*	Ferric citrate	--	114-774	774	1-33	56-63	34-113 (Not in structure)	(118)
	1PO3	*E. coli*	Ferric citrate	Ferric dicitrate	128-774	774	1-33	56-63	34-113 (Not in structure)	(119)
	1PO0	*E. coli*	Ferric citrate	Dicitrate	114-774	774	1-33	56-63	34-113 (Not in structure)	(119)
	1PNZ	*E. coli*	Ferric citrate	--	114-774	774	1-33	56-63	34-113 (Not in structure)	(119)
FpvA	1XKH	*P. aeruginosa*	Ferripyoverdine	Ferripyoverdine	129-815	815	1-43	130-134	44-118 (Not in structure)	(120)
	2IAH	*P. aeruginosa*	Ferripyoverdine	Ferripyoverdine	44-815	815	1-43	130-134	44-118 (Not in structure)	Wirth, C., *et al.* Unpublished
	2O5P	*P. aeruginosa*	Ferripyoverdine	--	44-815	815	1-43	130-134	44-118	(82)
	2W78	*P. aeruginosa*	Ferripyoverdine	$\text{Ferripyoverdine}_{13525}$	44-815	815	1-43	130-134	44-118	(121)
	2W77	*P. aeruginosa*	Ferripyoverdine	$\text{Ferripyoverdine}_{18\text{-}1}$	44-815	815	1-43	130-134	44-118	(121)
	2W76	*P. aeruginosa*	Ferripyoverdine	Ferripyoverdine	44-815	815	1-43	130-134	44-118	(121)
	2W75	*P. aeruginosa*	Ferripyoverdine	--	44-815	815	1-43	130-134	44-118	(121)
	2W6U	*P. aeruginosa*	Ferripyoverdine	$\text{Ferripyoverdine}_{\text{G173}}$	44-815	815	1-43	130-134	44-118	(121)
	2W6T	*P. aeruginosa*	Ferripyoverdine	$\text{Ferripyoverdine}_{\text{synthetic}}$	44-815	815	1-43	130-134	44-118	(121)
	2W16	*P. aeruginosa*	Ferripyoverdine	$\text{Ferripyoverdine}_{\text{C-E}}$	44-815	815	1-43	130-134	44-118	(121)
FptA	1XKW	*P. aeruginosa*	Ferripyochelin	$\text{Ferripyoverdine}_{188}$	56-720	720	1-38	33-40	--	(122)
PirA	5FP2	*P. aeruginosa*	Ferric enterobactin	Random peptide	60-724	724	1-28	119-126	--	Moynie, L., *et al.* Unpublished
PiuA	5FOK	*P. aeruginosa*	Ferric enterobactin	--	62-753	753	1-35	53-60	--	Moynie, L., *et al.* Unpublished
PirA	5FR8	*A. baumannii*	Nonspecific ferrisiderophore	--	48-754	754	1-24	116-123	--	Moynie, L., *et al.* Unpublished
PiuA	5FP1	*A. baumannii*	Nonspecific ferrisiderophore	--	43-743	743	1-28	46-53	--	Moynie, L., *et al.* Unpublished
BtuB	2GSK	*E. coli*	Cobalamin (Vitamin $B_{12)}$	Cobalamin + TonB	25-614	614	1-20	26-33	--	(79)
	3M8D	*E. coli*	Cobalamin	Cobalamin	27-614	614	1-20	26-33	--	(123)
	3M8B	*E. coli*	Cobalamin	--	26-614	614	1-20	26-33	--	(123)
	3RGN	*E. coli*	Cobalamin	--	26-614	614	1-20	26-33	--	(124)
	3RGM	*E. coli*	Cobalamin	--	27-614	614	1-20	26-33	--	(124)
	2GUF	*E. coli*	Cobalamin	--	21-614	614	1-20	26-33	--	(125)
Cir	2HDI	*E. coli*	Colicin I	Colicin Ia - R domain	31-663	663	1-25	31-38	--	(126)
	2HDF	*E. coli*	Colicin I	--	30-663	663	1-25	31-38	--	(126)
HasR	3DDR	*S. marcescens*	Heme	Heme + HasA	147-899	899	1-34	134-141	70-121 (Not in structure)	(127)
	3CSN	*S. marcescens*	Heme	HasA	147-899	899	1-34	134-141	70-121 (Not in structure)	(127)
	3CSL	*S. marcescens*	Heme	Heme + HasA	147-899	899	1-34	134-141	70-121 (Not in structure)	(127)
FauA	3EFM	*B. pertussis*	Ferric alcaligin	--	56-734	734	1-35	48-55	--	(128)
ShuA	3FHH	*S. dysenteriae*	Heme/hemoglobin	--	29-660	660	1-28	30-37	--	Brillet, K., *et al.* Unpublished.
FetA	3QLB	*P. fluorescens*	Ferric enantiopyochelin	Ferric enantiopyochelin	73-734	734	1-50	63-70	--	(129)
TbpA	3V89	*N. meningitidis*	Transferrin	Serotransferrin	51-915	915	1-24	38-45	--	(130)
	3V8X	*N. meningitidis*	Transferrin	Serotransferrin	54-915	915	1-24	38-45	--	(130)
FrpB/ FetA	4B7O	*N. meningitidis*	Fe^{3+}	--	39-720	720	1-22	66-73	--	(131)
	4AIQ	*N. meningitides*	Fe^{3+}	Fe^{3+}	34-720	720	1-22	66-73	--	(131)
	4AIP	*N. meningitidis*	Fe^{3+F}	--	39-720	720	1-22	66-73	--	(131)
ZnuD	4RVW	*N. meningitidis*	Zn^{2+}	Zn^{2+}	46-758	758	1-24	34-41	--	(132)
	4RDT	*N. meningitides*	Zn^{2+}	Zn^{2+}	44-758	758	1-24	34-41	--	(132)
	4RDR	*N. meningitidis*	Zn^{2+}	Zn^{2+}	42-758	758	1-24	34-41	--	(132)
FyuA	4EPA	*Y. pestis*	Ferric yersiniabactin	--	35-673	673	1-22	30-37	--	(133)
FusA	4ZGV	*Pectobacterium atrosepticum*	Ferredoxin	--	52-860	860	1-13	46-53	--	(134)

1.6.1.1. The TBDT N-terminal Signaling Domain

The NTSD typically comprises the first 70-90 residues of the mature TBDT. The FpvA crystal structures, in addition to solution NMR structures of the NTSDs of FecA from *E. coli* (PDB 2D1U and 1ZZV) (81,135); PupA from *P. capeferrum* (formerly *P. putida* WCS358) (PDB 2A02) (81); and the HasR-NTSD X-ray crystal structure (PDB 2M5J) (136) reveal a conserved βαβ-βαββ motif. The hydrophobic core is stabilized by non-polar amino acids such as Leu, Ile, and Phe, as well as the aliphatic regions of polar sidechains. Therefore, it appears that the NTSD fold is not affected by either substrate binding, transport, or signaling. Clearly, the overall fold of the NTSD is maintained throughout signaling. Experimental evidence indicates that deletion of the NTSD in FecA results in a transport competent protein, but without transcriptional initiation, due to regulatory incompetence of the protein; hence, the NTSD plays a critical role in the regulatory process (137,138).

1.6.2. The ECF Anti-sigma/Sigma Regulator Protein

The mediating protein within the TBDT CSS cascade is the anti-sigma or sigma regulator protein, which transmits the signal from the TBDT-NTSD across the inner membrane to the cytoplasm. There are a variety of regulatory mechanisms that govern anti-sigma factor:sigma factor interaction and release. The ECF anti-sigma proteins differ from cytoplasmic anti-sigma proteins in that they transverse the bacterial inner membrane. A canonical ECF anti-sigma protein may be traditionally divided into three domains, categorized as the periplasmic signal sensor or CTD (defined in this work as the C-terminal cell-surface signaling domain, or CCSSD, because of its uniqueness to CSS); the transmembrane signal transmitter, which is a single transmembrane α-helix; and the cytoplasmic signal receiver or N-terminal domain (designated herein as the anti-sigma domain) (105), each of which will be discussed in turn.

1.6.2.1. The Anti-sigma/Sigma Regulator Protein: C-terminal Cell-surface Signaling Domain

The CCSSD of the sigma regulator is conventionally defined as the entire periplasmic domain, or the final ~200 residues of the sigma regulator protein. The CCSSD of FecR (residues 101 to 317) interacts both *in vivo* and *in vitro* with the N-terminal signaling domain (NTSD) of FecA (residues 34-113), although FecR residues 237-317 were shown to be necessary and sufficient for binding to the FecA-NTSD (105,139). Deletion of residues 48-102 of the FecA-NTSD abolish binding of full-length FecA to FecR, as analyzed by affinity pull-down; FecA transport activity remained unaffected (139), suggesting that the signaling component of FecA is confined to the NTSD. FecR and FecR-like proteins contain a conserved heptapeptide repeat motif flanked by three leucine residues and one (semi-conserved) valine residue ($Lx_6Lx_7Lx_7L/V$) resembling a leucine zipper motif or a leucine-rich repeat. Site-directed mutagenesis within the FecR leucine motif reduced binding to the FecA-NTSD, suggesting that the leucine motif plays a direct role in NTSD:anti-sigma CCSSD association. Specifically, FecR mutations L254E, L269G, and F284L, displayed a significant decrease in activity compared to WT FecR (105). A double mutation of FecR residues D138 and V197 specifically caused an increase in transcriptional activation activity and association to FecR, while, surprisingly, a FecA D43 mutant restored the activity of several FecR mutants (105). A binding pocket on the FecA-NTSD was proposed as part of an interaction interface with the FecR-CCSSD, as derived from statistical coupling analysis and site-directed mutagenesis; this includes D45, which is located just prior to helix 3 of the FecA-NTSD (81). Thus, the proposed interaction interface of the NTSD and anti-sigma factor CCSSD includes the leucine rich region of the CCSSD (105), and the CCSSD binding pocket on the NTSD, of which D45 appears to play a key role (81).

1.6.2.2. The Anti-sigma/Sigma Regulator Protein: Transmembrane Helix

The anti-sigma regulatory process of the TBDT system is unique in that the signal begins in the extracellular space and is transferred through three cellular compartments to the cytoplasm. Thus, a single monomeric integral-membrane sensor protein spans the bacterial inner membrane to convey this signal. An important aspect of the σ regulator is the transmembrane helix- a single helical region that is conserved amongst FecR homologs (140). Based on secondary structure and transmembrane helix predictions, the transmembrane region spans residues 85-100 of FecR and 94-114 of FpvR (141). These single-pass transmembrane regions are characterized by low sequence complexity and incorporation of non-polar residues, flanked on the cytoplasmic side with positively-charged residues (K or R), and on the periplasmic side with bulky, hydrophobic residues (W and/or Y). These serve to anchor the helix within the non-polar transmembrane environment, as well as interact with the charged cytoplasm:inner membrane interface.

1.6.2.3. The Anti-sigma/Sigma Regulator Protein: N-terminal Domain

A semi-conserved cytoplasmic N-terminal domain of 80-85 residues has been identified in multiple sigma regulators (142). The σ^{70} superfamily, including the group IV ECF sigma factors, to which PupI, FecI, and FpvI belong, is largely regulated by sigma regulators each containing a related N-terminal domain, despite the low sequence similarity among the N-terminal domain-containing sigma regulators (142-144). The primary function of the N-terminal domain is to sequester their cognate sigma factors by tethering them to the cytoplasm:inner membrane interface for release only when the appropriate signal is transmitted; thus, these domains have been defined (in this work) as anti-sigma domains (ASDs). FecR proteins possess three N-terminal tryptophan residues essential for FecR activity (105,139). Previously, mutational analyses of these particular tryptophan residues disrupted the FecR:FecI interaction, and thus, rendered them transcriptionally inactive (110). These residues are highly conserved across FecR-like proteins, and appear to be critical for maintaining the stability of the domain's hydrophobic core (145).

1.6.3. The ECF Sigma Factor

Prokaryotic transcription is principally regulated at the initiation step. Sigma factors play a crucial role in the initiation of transcription by mediating promoter recognition through RNA polymerase recruitment; in some cases, sigma factors regulate expression of virulence genes and virulence-associated genes in bacterial pathogens (146). To alter gene expression and adequately react to environmental or cellular stresses, sigma factor association with RNAP must be reversible. Gram-negative bacterial species typically produce multiple sigma factors; *E. coli* contains seven sigma factors, while *P. aeruginosa* harbors twenty-four (147,148). Of these sigma factors, the Group 4 sigmas are the largest and most diverse class; Group 4 includes the extracytoplasmic function (ECF) factors, which respond to environmental changes to modify the bacterial gene expression profile (149). Under iron-limiting conditions, a Group 4 ECF sigma factor (of the σ^{70} family, which bind to the -10 and -35 promoter regions for interaction with RNAP) (150) directs expression of a TBDT gene encoding for the suitable cell surface receptor for siderophore transport. Expression of both the sigma and anti-sigma factors are coupled, so that the anti-sigma factor regulates its cognate sigma factor in response to siderophore binding. In the absence of ferrisiderophore, the anti-sigma factor suppresses the sigma factor by blocking its interaction with RNAP; in the presence of siderophore, the anti-sigma factor releases the sigma factor. Presumably, this occurs through either coupled release of the anti-sigma domain:sigma factor complex, or proteolysis and degradation of the anti-sigma domain for release of the sigma factor alone. As has been demonstrated in the Fec system, the *fecABCDE* and f*ecIR* genes create distinct transcripts (151). Transcription of *fecIR* is regulated by the Fur repressor protein and not by siderophore transport; additionally, they are not autoregulated (102). The *fecABCDE* operon is co-regulated by iron concentration sensed by the Fur protein, and by the activity of FecI and FecR through transport and signaling of an extracellular iron source (100). Under iron-limiting conditions, transcriptional initiation of the Fec operon is dependent upon FecI (151). Analogous systems in other Gram-negative bacteria have been shown to function in a similar manner.

Nearly one-third of ECF sigma factors are controlled by anti-sigma proteins that interact via an N-terminal ASD (142). Approximately 85 residues at the FecR N-terminus interacts with FecI, and the initial 61 residues of the N-terminus of FecR were sufficient to interact with FecI. In the *fpv* system, the first 67 residues of FpvR are sufficient to associate with both FpvI and PvdS (152).

ECF sigma factors appear to be quite unstable in solution, as demonstrated for FecI, which required solubilization from inclusion bodies (139). This may indicate that the sigma:anti-sigma proteins are typically co-expressed. Even after renaturation, FecI activity was maintained, as it was able to bind the *fecA* promoter region and recruit RNAP (153). Additionally, the ASD of FecR was shown to be required for FecI activity, indicating that the ASD may preserve FecI in a stable conformation (101,139). It is unknown whether the FecR-ASD acts only as an anti-sigma regulator to release FecI once the signal from the ferrisiderophore-transporter binding initiates contact between the NTSD and the FecR-CCSSD, or if the FecR-ASD also serves as a chaperone for FecI and is essential to mediate FecI association with RNAP.

1.6.4. Regulatory Mechanisms of the TBDT System

1.6.4.1. Primary Transcriptional Regulation by Fur

The ferric uptake regulator (Fur) is a transcriptional repressor that serves as the global controller of cellular iron concentrations. Fur recognizes a 19-bp inverted repeat sequence, the "Fur box"- 5' - GATAATGATAATCATTATC – 3' (154). Fur, when bound to Fe^{2+}, undergoes a conformational change that induces dimerization and DNA binding through sequence recognition specificity (155). Under iron-limited conditions, Fur dissociates from DNA, and expression of genes encoding for iron transporters, iron transport regulators, siderophore biosynthesis, iron metabolism, and other cellular functions is initiated (156-158). DNA binding by Fur physically blocks RNAP from the promoter regions of target genes, thereby inhibiting transcription (54).

1.6.4.2. Regulated Intramembrane Proteolysis

A proposed mechanism of signal activation through the sigma regulator involves differential proteolysis as a method of control of anti-sigma and sigma factor functionality. Both FecR and FpvR regulated intramembrane proteolysis (RIP) data implies that a conformational change induced by siderophore uptake and rearrangement of the transporter NTSD also prompts conformational variation in the sigma regulator, causing it to become more sensitive to cleavage of the periplasmic domain by a membrane-anchored protease, initially by a site-1 protease, which allows for subsequent proteolysis by a site-2 protease (110,139,159-161).

Binding of ferripyoverdine to FpvA results in complete proteolysis of FpvR (160). Proteolysis of FpvR activates and releases sigma factors PvdS and FpvI to initiate genes required for pyoverdine synthesis and transport. In the absence of pyoverdine, differential proteolysis of FpvR occurs, but the FpvR subfragments produced inhibit PvdS and FpvI function. MucP, an ortholog of the site-2 protease RseP, along with other proteases (including a site 1-like protease), is involved in the FpvR degradative pathway in both the presence and absence of pyoverdine (160). Given that targeted expression of regulatory genes and continual proteolysis in the presence and absence of signal is relatively energetically costly for the cell, it is not well understood how this mechanism benefits the bacterium. The exact recognition sequence of RseP remains uncharacterized; however, the protease appears to have wide substrate specificity (162).

Recently, two specific proteases have been tied to CSS regulatory systems involving TonB-dependent transporters that possess dual siderophore transport and signal transduction through an inner membrane sigma regulator to release an ECF sigma factor, in both *Pseudomonas putida* KT2440 and *Pseudomonas aeruginosa* PAO1, under iron starvation conditions (161). The first is Prc (or Tsp), the tail-specific protease, which is a serine-type protease with endoproteolytic activity. Prc identifies a particular target sequence, usually A/L-A/Y-A, then cleaves the periplasmic protein some distance from the recognition sequence. Prc contains a PDZ domain (named for the first three proteins found with the domain- postsynaptic density protein, Drosophila disc large tumor suppressor, and zonula occludens-1 protein), which has been implicated in intracellular signaling and co-localization of signaling molecules to membrane interfaces (163). In *P. putida* lacking *prc*, aerobactin-mediated expression of the target transporter protein was abrogated; constitutive expression of Prc from a plasmid restored signaling activity, both in the presence and absence of aerobactin (161). Additional studies demonstrated that the site-2 protease RseP further proteolyzed the sigma regulator to release the anti-sigma domain:sigma factor complex, and was essential for transcriptional activation through the sigma regulator (160). Mutations of RseP, a site-2 intramembrane protease, reduced FecA concentrations in the outer membrane, which were restored by introduction of plasmid-encoded wild-type *rseP* (160).

RIP by RseP requires proteolysis of the periplasmic domain of the sigma regulator by a site-1-like protease (160), similar to the proteolysis of the RseA:σ^E complex by the membrane-anchored serine protease DegS (164-166). Sequential cleaveage of RseA by DegS is inhibited by binding of RseB to the periplasmic domain of RseA; the X-ray crystal structure of this complex reveals that, although the DegS cleavage site of RseA is disordered, part of its recognition sequence is buried within the hydrophobic pocket of RseB (165). In the absence of DegS, σ^E activity cannot be induced, and *ΔdegS* cells exhibit a slow growth phenotype (167).

Similarly, the sigma factor AlgU in *P. aeruginosa* is regulated by the anti-sigma factor MucA (168). Interestingly, mutations in the C-terminal region of MucA, which dramatically lower the stability of the anti-sigma factor and result in a mucoid phenotype, were a common features amongst *P. aeruginosa* isolates from cystic fibrosis patients (30,169). In this example, mutations in MucA rendered it more susceptible to proteolysis by Prc, which has no effect on wild-type (WT) MucA (170). This may trigger intramembrane proteolysis by an unidentified site-2 protease, resulting in the release of AlgU and the observed mucoidy phenotype. Thus, cooperativity between the site-1 and site-2 proteases is essential for signal transduction.

1.7. The TBDT System as a Target for Novel Antibiotics

The TBDT iron transport and regulatory pathway presents an attractive target for new or alternative antimicrobial therapeutics. The “Trojan horse” antibiotics exploit the molecular recognition and transport mechanisms of Gram-negative species for directed antimicrobial activity, via conjugation of an antibiotic compound to a siderophore, or to a molecule that mimics natural siderophores. Common strategies have included conjugation of a siderophore with a familiar β-lactam (penams, cephalosporins, monobactams, or carbapenems), which target synthesis of the peptidoglycan layer of the cell wall, and have low efficacy against Gram-negative pathogens under normal conditions. Several *in vitro* studies have demonstrated that these conjugates exhibits 10-1000 fold increase in activity, and discloses accelerated inhibition of bacterial growth compared to the parent antibiotics (171). Similar compounds analyzed *in vivo* (murine models) displayed an increase in antimicrobial efficacy, even under iron-depleted conditions (172). A monosulfactam coupled with an iron-chelating dihydroxypyridone moiety presented potent activity, even against multi-drug resistant *P. aeruginosa* and *Acintobacter* clinical isolates. These findings were verified in a mouse model of septicemia; however, these compounds have yet to undergo clinical trials to test human tolerability (173,174).

Some microbes synthesize similar siderophore-antibiotic compounds, called sideromycins, to target other bacteria in its environment, in order to compete for limited resources. Several of these sideromycins have been characterized- albomycins (175,176), ferrimycins (177), danomycins (178), salmycins (178), and microcins (179-181). Additionally, a number of synthetic sideromycins have been developed in recent years that have demonstrated antimicrobial activity (171,182-184).

This novel class of antibiotics addresses several complications associated with chemotherapeutic treatment of Gram-negative bacterial infections. First, by coupling a compound that the bacterium recognizes with an antimicrobial agent, the bacterium actively transports the compound into the cell, which overcomes the major hurdle of limited access through the complex Gram-negative bacterial cell wall. Second, recognition of the siderophore conjugate allows for evasion of some bacterial drug resistance mechanisms, such as elimination of undesirable molecules from the cell by efflux pumps. Third, by manipulating a bacterial survival pathway that recognizes and utilizes very distinct molecular signatures, application of antibiotics that are directed toward specific iron acquisition pathways allows for purposeful treatment of specific bacterial infections, and a move from usage of broad-spectrum antibiotics toward narrow-spectrum antimicrobials. Last, targeting of the regulatory pathway of the TonB-dependent transducers may allow for artificial upregulation of expression of the transducer, which would result in more transducers in the outer membrane (which are typically sparingly expressed) for transport of additional siderophore-antibiotic conjugate molecules.

1.8. Specific Aims of This Research

Given the importance of the control of iron transport through the TonB-dependent pathway for maintenance of iron homeostasis for Gram-negative bacterial survival, and the possibilities for targeting this system with novel antimicrobials, it is necessary to understand the fundamental regulatory components of this pathway. The goal of this work is to provide a structural rationale for a distinct signal transduction pathway that regulates ferric siderophore uptake in Gram-negative bacteria, by addressing the following specific aims:

1. To investigate structural variation between the transporter PupA and the transducer PupB that allows for association of the NTSD with the CCSSD, in the context of the full-length outer membrane protein.
2. To structurally and biochemically characterize the PupB-NTSD for comparison to the PupA-NTSD, and to identify residues within the PupB-NTSD that enable discrete signal transfer by protein:protein interactions with the PupR-CCSSD.
3. To study the structural details of the interaction of the PupB-NTSD with the PupR-CCSSD, clarify the association interface of the two proteins, and identify mutations that abrogate the interaction.

4. To examine the structural features of PupR, a representative of a class of anti-sigma inner membrane proteins with almost no available structural data.
5. To characterize the interaction of PupR with its cognate sigma factor PupI, and elucidate details that result in release of the protein from the inner membrane:cytoplasm interface for interaction with a target DNA promoter and recruitment of RNAP.

2. TOWARD A STRUCTURAL UNDERSTANDING OF DIFFERENTIAL SIGNALING THROUGH THE OUTER MEMBRANE IRON TRANSPORT PROTEIN PUPA AND TRANSDUCER PUPB

2.1. Introduction

The response of the OM TonB-dependent transporters to ferric siderophore is critical for CSS under iron-limited conditions. Although the primary protein players in this system have been identified, the signal transduction mechanism is still poorly understood. An additional unknown is the specificity by which homologous receptors signal through discrete inner membrane regulators to initiate transcriptional control. CSS TBDTs have an overall architecture consisting of an N-terminal signal peptide, which is cleaved during protein maturation, an NTSD, a plug domain, and a β-barrel domain (Figure 2.1) (140,185). The NTSD, the soluble periplasmic domain of the receptors, is critical for signaling. Of structurally characterized TBDTs, only FecA, FpvA, and HasR contain an NTSD, and NTSDs have been identified in the siderophore receptors PupA and PupB (Figure 2.1). As TBDT regulation via CSS is limited to transporters that also include an NTSD, these transporters have been categorized as TonB-dependent transducers (71). Interestingly, only the structures of FpvA include an ordered NTSD, albeit in multiple conformations, indicating that the NTSDs may exist in a disordered state relative to the barrel and plug domains.

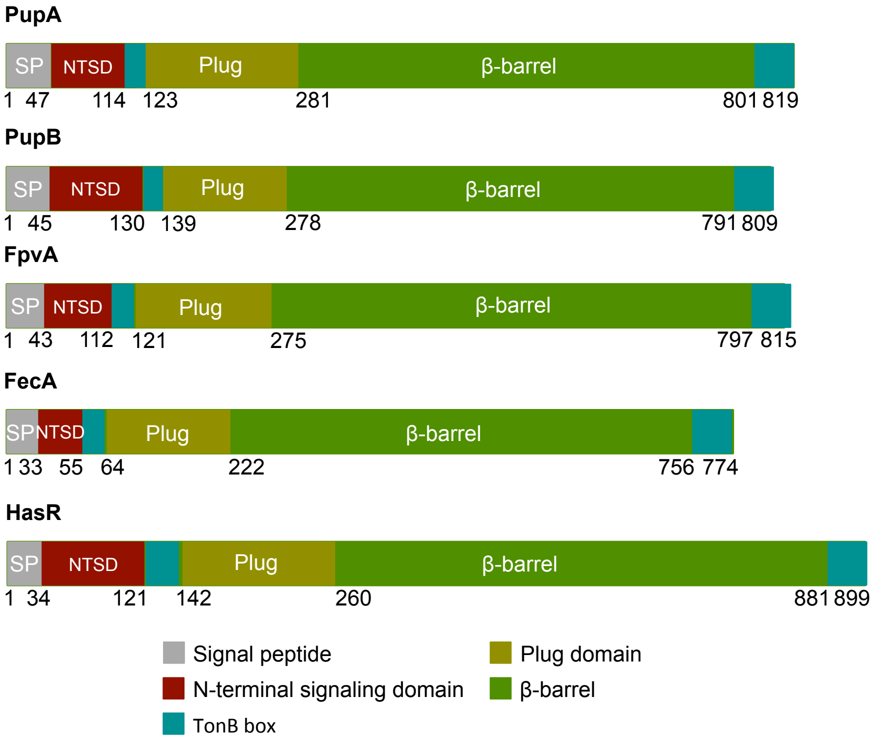
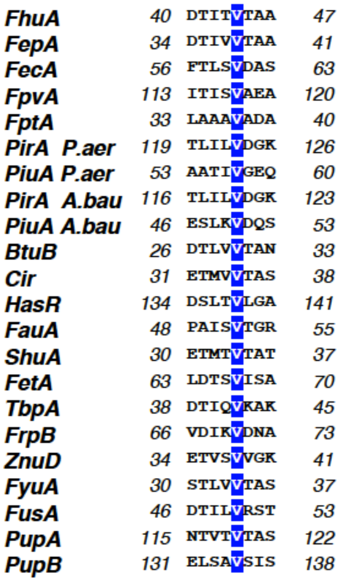


Figure 2.1. Conserved domain architecture of the homologs PupA and PupB from *Pseudomonas capeferrum*, and orthologs FpvA from *Pseudomonas aeruginosa*, FecA from *E. coli*, and HasR from *Serratia marcescens*.



FhuA	40	DTITVTAA	47
FepA	34	DTIVVTAA	41
FecA	56	FTLSVDAS	63
FpvA	113	ITISVAEA	120
FptA	33	LAAAVADA	40
PirA P.aer	119	TLILVDGK	126
PiuA P.aer	53	AATIVGEQ	60
PirA A.bau	116	TLILVDGK	123
PiuA A.bau	46	ESLKVDQS	53
BtuB	26	DTLVVTAN	33
Cir	31	ETMVVTAS	38
HasR	134	DSLTVLGA	141
FauA	48	PAISVTGR	55
ShuA	30	ETMTVTAT	37
FetA	63	LDTSVISA	70
TbpA	38	DTIQVKAK	45
FrpB	66	VDIKVDNA	73
ZnuD	34	ETVSVVGK	41
FyuA	30	STLVVTAS	37
FusA	46	DTILVRST	53
PupA	115	NTVTVTAS	122
PupB	131	ELSAVSIS	138

Figure 2.2. Alignment of the 8-residue TonB box sequences of TonB-dependent transporters of known structure, and PupA and PupB. The sequences are highly variable, with only a single invariant valine at the fifth position. Uniprot identification of the protein sequences from which the TonB box sequences were derived are as follows: FhuA, P06971; FepA, P05825; FecA, P13036; FpvA, P48632; FptA, P42512; PirA, *P. aeruginosa*, Q9I527; PiuA, *P. aeruginosa*, G3XCY8; PirA, *Acinetobacter baumannii*, D0C8V9; PiuA, *Acinetobacter baumannii*, D0CAH3; BtuB, P06129; Cir, P17315; HasR, Q79AD2; FauA, Q9X6A5; ShuA, P72412; FetA, C5I2D9; TbpA, Q9K0U9; FrpB, Q841A2; ZnuD, Q9JZN9; FyuA, P46359; FusA, Q6D8U4; PupA, P25184; PupB, P38047. Note: for FpvA, residues 113-120 have been used for alignment, while residues 130-134 are predicted to be part of the TonB box (120); however, this region does not contain a valine residue that is hypothesized to be conserved across TonB boxes.

Conformational flexibility during substrate binding and/or transport appears to be localized to the TonB box region (Figure 2.2), as was demonstrated by the reduction in HasR-NTSD dynamics in the presence of HasB (analogous to TonB) as measured by chemical shift perturbation (186). Mutation of part of the TonB box of FpvA (residues 129-135) abrogate ferripyoverdine transport and also significantly disrupt pyoverdine-mediated signaling (187). There is very little alteration between the NTSDs of apo-FvpA and ferripyoverdine-bound FpvA, suggesting that this domain undergoes minimal conformational change upon substrate binding to the TBDT and transport into the periplasm, and that substrate binding only alters the flexible TonB box region, and subsequently, the dynamics of the NTSD as a whole. This was confirmed through electron microscopy (EM) and small-angle X-ray scattering experiments (SAXS) of free HasR compared to the HasR:heme:HasA complex, which presented a protracted disordered region between the HasR-NTSD and the HasR β-barrel (186). This results in an NTSD that extends far into the periplasmic space, positioning the NTSD for interaction with its cognate anti-sigma factor. In the FpvA crystal structure (PDB 2O5P), the β-barrel and the NTSD are in close proximity, with the N-terminal TonB box hidden within the β-barrel in chain A of 2O5P (one of the FpvA structures in the asymmetric unit) (Figure 2.3). The absence of an NTSD in the other TBDT structures indicates that the linker region (including the TonB box) between the NTSD and the β-barrel is highly dynamic; it is unknown whether the interaction of the NTSD with the β-barrel and the compact linker observed in the FpvA structures is an accurate representation of one physiological state of the apo-transporter, or a result of crystal packing. Of the FpvA structures, PDBs 2IAH and both chain A and chainB of 2O5P display an alternative location of the NTSD (Figure 2.3B-C). In 2IAH, the NTSD is presented on the opposite edge of the periplasmic space:membrane interface, and is also rotated nearly 180 degrees about the x plane (Figure 2.3A). Residues 118 – 135, which include part of the TonB box, are missing from the crystal structures and thus are presumably disordered. Furthermore, there is no structural evidence to demonstrate association of TonB with the TonB box and concurrent binding of the CCSSD with the NTSD.

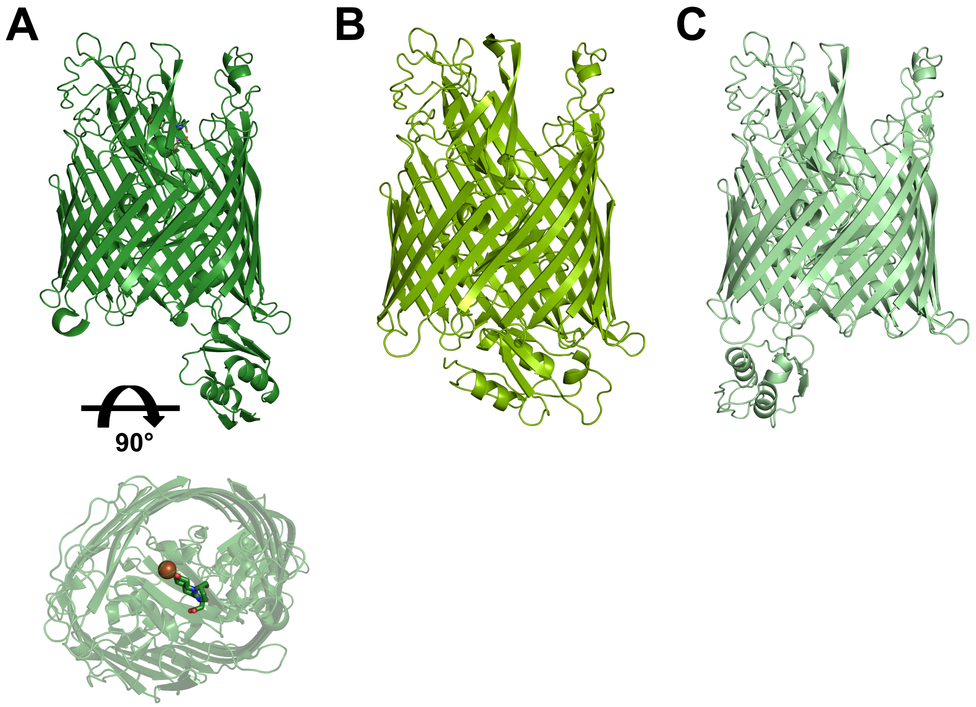


Figure 2.3. X-ray crystal structures of FpvA from *Pseudomonas aeruginosa*. (A) Ferripyoverdine-bound FpvA (PDB 2IAH) (Worth, C. *et al.* Unpublished). Pyoverdine is displayed as sticks, with Fe as a sphere. (B,C) Apo-FpvA (PDB 2O5P) (82). The two FpvA molecules within the asymmetric unit show the NTSD in varying orientations.

In order to examine structural variation between PupA and PupB that may allow for differential signaling to occur through one protein (PupB) but not the other (PupA), both transmembrane receptors were purified to homogeneity in the presence of detergent. Additionally, a chimeric protein was constructed, wherein the PupA-NTSD was replaced with the PupB-NTSD, within full-length PupA (designated PupA$^{\text{PupB-NTSD})}$. The stability and solubility of PupA in an array of diverse detergents was assessed, and crystallization of PupA for structure solution was pursued. The solution states of detergent-solubilized PupA, PupB, and PupA$^{\text{PupB-NTSD}}$ were analyzed by SEC-SAXS, and the interaction of the sigma regulator PupR’s CCSSD, characterized in Chapter 4, with the NTSDs was probed in the context of the full-length receptors. Several receptor and protein detergent complex (PDC) models were constructed to evaluate possible solution states of the receptors and receptor complexes, which were compared to the experimental SAXS data.

2.2. Materials and Methods

2.2.1. Cloning and Expression of the Pup Transporters

Several expression constructs were created for purification and interaction studies with the full-length Pup transporters. The *pupA* gene was amplified from *Pseudomonas capeferrum* genomic DNA and cloned into pET17b between NdeI and HindIII sites. A His_6 tag was inserted by site-directed mutagenesis between the signal peptide and the signaling domain of PupA, in order to facilitate purification by affinity chromatography. The *pup*A gene, including the His_6 tag, was transferred to pET24a, so that constructs containing Amp^R (pET17b) and Kan^R (pET24a) were available. Similarly, the *pupB* gene was amplified from *Pseudomonas capeferrum* genomic DNA and cloned into pET17b between NdeI and XhoI sites, and a His_6 tag inserted between the *pupB* signal peptide and signaling domain sequences. Following site-directed mutagenesis for addition of the affinity tag, the *pupB* gene was transferred to pET29b.

In order to structurally assess the differences between the signaling domains of PupA and PupB in the context of the full transporters, and to determine binding of the full length transporters to the PupR-CCSSD *in vitro*, a chimeric protein was created, which contained the NTSD of PupB, and the plug and barrel of PupA. This was accomplished by a combination of site-directed mutagenesis, digest with restriction endonucleases, and ligation. A BamHI site was inserted into the pET17b-PupA sequence between the His_6 and the signaling domain sequences. An NheI site was inserted into the same construct, between the signaling domain and the plug domain of PupA. This allowed for the PupA-NTSD encoding region to be excised from *pupA*, and replaced with the PupB-NTSD coding insert. The Quick Ligation protocol and reagents

(NEB) were utilized to ligate the PupB-NTSD into the pET17b-PupA construct. The *pupA*$^{PupB\text{-}NTSD}$ chimeric gene was then transferred to pET24a. The sequences of all expression constructs were verified through DNA sequencing by MCLab.

E. coli BL21(DE3)pLysS cells were transformed with ~100 ng of one of the expression constructs described above according to the standard heat shock method (188). Transformed cells were incubated at 37 °C in LB medium with 100 µg/mL ampicillin or 15 µg/mL kanamycin until OD_{600} ~ 1.0. Protein expression was induced with the addition of 1 M IPTG to a final concentration of 0.5 mM, and cultures were incubated at 20 °C for 18 hr. Cells were harvested by centrifugation at 4000 *g* for 30 min, resuspended in 25 mM HEPES pH 7.5, 150 mM NaCl, pelleted a second time, and the cell pellets stored at -80 °C until use.

2.2.2. Purification and Solubilization of Pup Transporters

The initial purification protocol for PupA, developed by Dr. Chris Colbert, was modified for isolation of PupB and PupA$^{\text{PupB-NTSD}}$.

Frozen pellets from cells expressing PupA were suspended in 25 mM HEPES pH 7.5, 150 mM NaCl, and lysed by emulsification. The crude extract was clarified by centrifugation at 183000 *g* for 45 min in a Beckman L8-70M ultracentrifuge, using a F40L Fiberlite rotor (Thermo Scientific). Pelleted cell membranes were homogenized in 25 mM HEPES pH 7.5, 150 mM NaCl, 25 mM imidazole, and solubilized with 1% w/v FC-13. After gentle stirring at 4 °C to solubilize the protein, a second centrifugation was performed at 183000 *g* for 30 min to remove insoluble proteins cellular debris. The supernatant from the second centrifugation was loaded onto a 5 mL HisTrap (GE Lifesciences) column equilibrated with 25 mM HEPES pH 7.5, 150 mM NaCl, 25 mM imidazole, 0.1% w/v LDAO. The column was washed with 100 mL of equilibration buffer, and eluted with 25 mL of elution buffer (equilibration buffer, with imidazole increased to 125 mM). IMAC elution fractions containing the protein of interest were pooled and concentrated with a 100 kD MWCO filtration device (Millipore) to ~ 10 mg/mL. Size exclusion chromatography was performed with either a Superdex XK 200 16/60 or Superdex 200 10/300 GL column (GE Lifesciences). Isocratic elution with 25 mM HEPES pH 7.5, 150 mM NaCl, 0.1% LDAO was performed at a flow rate of 0.3 mL/min. Fractions containing PupA were pooled, concentrated to 10-12 mg/mL, and stored at 4 °C or flash-frozen in LN_2 for storage at -80 °C. Typical yields were 4.0-5.5 mg/6 L of cells. Protein concentration was determined by absorbance at 280 nm using the molar extinction coefficient $\varepsilon_{280} = 138050$ L M^{-1} cm^{-1}, with the theoretical molecular weight derived from the protein sequence (without the signal peptide) of 86968.1 Da. This represents a roughly 25% recovery of protein from the initial purification stages.

To purify PupB and $PupA^{PupB\text{-}NTSD}$, the same protocol as outlined for PupA was followed, with the exception of maintaining FC-13 throughout purification. During IMAC, the concentration of FC-13 was decreased to 0.1% w/v. This concentration or 0.05% w/v FC-13 was used during SEC and SEC-SAXS. Typical yields for PupB and $PupA^{PupB\text{-}NTSD}$ were roughly 9.5 mg/6L cells and 2.3 mg/6L cells, respectively.

2.2.3. Differential Detergent Filtration Assay with PupA

To identify additional detergents that may improve PupA stability and solubility for single-crystal optimization, a differential filtration detergent assay (Memb-PASS, Omscientia) was utilized (189,190) to test the stability and obtain relative size information of the protein in new detergents. Approximately 500 µg of purified His_6-PupA was combined with 2.4 mL of Ni-NTA resin equilibrated in 25 mM HEPES pH 7.5, 150 mM NaCl, 25 mM imidazole, 0.1% LDAO. The total volume of the slurry was adjusted to 6 mL, and allowed to incubate overnight at 4 °C. Following the overnight incubation, 50 μL of the slurry was pipetted into each well of a 96-well filter microplate (0.2 μm filter), then washed with 150 μL of equilibration buffer and centrifuged at 2000 *g* for 2 min. The flow-through was collected in a 96-well microplate. The wash step was repeated 3 times. A second series of washes was performed with 30 μL of 25 mM HEPES pH 7.5, 150 mM NaCl, 25 mM imidazole that was previously combined with a new detergent from the 94 assay conditions, followed by centrifugation at 2000 *g* for 2 min. The second wash series was repeated 6 times. Seventy μL of elution buffer, consisting of 25 mM HEPES pH 7.5, 150 mM NaCl, 125 mM imidazole, and the new detergent, was added to each well, and the plate was centrifuged at 2000 *g* for 2 min. The eluate was collected in a new 96-well microplate. Half of the eluate was added to a 100 kD MWCO filtration plate, and the other half to a 300 kD MWCO filtration plate, followed by centrifugation at 2000 *g* for 4 min. The flow-through solution was collected from each plate.

A Minifold I Spot-blot 96-well dot-blot apparatus (Whatman) was assembled as specified in the Omscientia protocol with pre-wet 0.2 μm nitrocellulose membrane. Ten percent trichloroacetic acid was added to each well, and mixed with 10 μL of the 100 kD MWCO elutions. After removal of the protein solution through the vacuum manifold, the blot was washed 3 times with PBS pH 7.0. This protocol was repeated for the 300 kD MWCO elutions. The membranes were then washed with ddH_2O and blocked with TBST + 5% w/v milk powder. Each blot was incubated with 1:5000 dilution of Anti-His_6-HRP antibody (Clontech) for 1 hr at 4 °C, and washed 3 times with TBST. His_6-tagged PupA proteins were detected with the ECL Western Blotting Substrate (Pierce) and imaged with an Odyssey fluorescence plate imager (LI-COR Biosciences). Scan intensity was optimized so as to eliminate spot signal saturation, and the intensities quantified following average background signal correction.

2.2.4. Pull-down Analysis of the PupA Transporter and PupB Transducer with the PupR-CCSSD

Purified transporter/transducer was combined with purified PupR-CCSSD that had been exchanged into 25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol, 0.1% w/v FC-13, in a 1:1 molar ratio. Size exclusion chromatography was performed as previously described. SDS-PAGE with Coomassie blue staining was utilized to qualitatively analyze SEC elution fractions. SECs were also run in the same buffer conditions with each protein separately, in order to compare the chromatograms and elution volumes of each individual protein with potential protein complexes.

2.2.5. Crystallization and X-ray Diffraction of PupA

Crystal optimization of PupA was based upon nucleation conditions previously identified by C. Colbert. Hanging drop vapor diffusion experiments were manually constructed, and were comprised of 50 mM Tris pH 7.8-8.0 or 50 mM Tricine pH 8.0, 350 mM NaCl, and 0.1% w/v LDAO, and various concentrations of PEG 1k (20-30% w/v) and glycerol (0-20% v/v) as the reservoir solutions. PupA at ~11 mg/mL in 25 mM HEPES pH 7.0, 100 mM NaCl, 0.1% LDAO were combined 1:1 with the reservoir solution. Trays were incubated at 20 °C and monitored frequently.

Lipidic bicelles were utilized for crystallization experiments to mimic a native bilayer environment (191). A MemMagic bicelle screen kit (Molecular Dimensions), containing solutions with varying concentrations of dimyristoyl-phosphatidylcholine (DMPC) and 3-[{3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO), was obtained. Bicelles were combined with concentrated protein (~11 mg/mL) in a 1:4 bicelle:protein ratio; this mixture was incubated on ice for 30 min. Hanging drop vapor diffusion trays were assembled with the reservoir solutions described above, and protein:bicelle solution and reservoir solution were combined in a 1:1 ratio. Trays were incubated at 20 °C. Microcrystals typically appeared after approximately 1 week, and were harvested with micromeshes (10-25 μm openings) (MiTeGen) for diffraction screening, using the reservoir solution + 20% v/v glycerol as a cryoprotectant. Mounted microcrystals were flash-cooled in LN_2.

Several cryo conditions were tested for PupA crystals grown in LDAO. These included washing the crystals in MiTeGen cryo-oil and flash-cooling in LN_2; washing in MiTeGen cryo-oil, flash-freezing, followed by flash annealing (allowing the frozen crystal in cryoprotectant to thaw for a few seconds, then repeating the flash-freezing process); soaking the crystals in 45% PEG 1k, as a cryoprotectant; soaking in 45% PEG 1k, followed by flash annealing; soaking in 45% PEG 1K + 10% v/v glycerol; dehydration of the crystal droplet for 5-10 min, then transfer of crystal to MiTeGen oil prior to flash-freezing; dehydration, transfer to 20% v/v glycerol, followed by flash-annealing; dehydration, then crystal transfer to 30% sucrose; dehydration, transfer to 30% sucrose, then flash-annealing; and lastly, crystal drop dehydration, followed by rehydration with reservoir solution, the flash-freezing with 20% v/v glycerol as the cryoprotectant.

Diffraction data were recorded at APS NE-CAT, beamline 24-ID-E with a microfocus X-ray beam. In the case of the protein:bicelle crystals, rastering was necessary to identify the location of the microcrystals in the micromesh, as they were not readily visible. A single PupA:LDAO crystal initially diffracted to ~6 Å, and a partial data set to 7 Å was collected. Data processing was performed with either RAPD or autoPROC, and the space group determined.

2.2.6. Solution Characterization of Pup Transporters by SEC-SAXS

2.2.6.1. SAXS: the Theory

Small-angle X-ray scattering of biological macromolecules provides low-resolution (in the 10-50 Å range) structural information concerning the folding, flexibility, shape, and assembly state of a protein (or other biological macromolecule) in solution, and how the protein responds to changes in the solution environment (192). A solution containing pure and homogeneous protein is exposed to a monochromatic X-ray beam, and the intensities of the X-rays scattered at small angles (0.1-10 °) is recorded. The background-corrected intensities are proportional to the average scattering from a molecule in multiple orientations. The scattering intensity, I, is expressed as a function of the magnitude of the scattering vector (also called the momentum transfer), q:

$$q = \frac{4\pi Sin\theta}{\lambda} \qquad \text{(Eq. 2.1)}$$

where 2θ is the scattering angle and λ is the wavelength of the X-ray beam, and plotted as a one-dimentional, radially-averaged curve (Figure 2.4A).

Monodispersity of a protein in solution is assessed by the Guinier formula, the plot of which enables estimation of the radius of gyration, R_g, of the sample. If, at q x $R_g < 1.3$, the plot of log(intensity) vs. q^2 is linear, the sample is monodisperse. The slope of the fitted line yields the R_g, and the y-intercept is the forward scattering intensity, $I(0)$ (Figure 2.4B) (193). Thus, the R_g is the average distance from the center of density of the molecule, and $I(0)$ is proportional to the molecular weight of the protein.

The plot of $I(q) \times q^2$ vs q, or the Kratky plot, provides a qualitative estimate of the folded state or flexibility of the protein (194). A well-folded, globular protein will yield a symmetrical, bell-shaped curve that returns to baseline at high *q* (Figure 2.4C). The curve of an unfolded or disordered protein increases then plateaus at high *q* (Figure 2.4C).

The maximum dimension of the molecular in solution, D_{max}, is the value of the protein's radius when the pairwise distribution function, *P(r)*, equals zero (195). Additionally, the shape of the *P(r)* plot indicates the general shape of the protein (Figure 2.4D).

A low-resolution molecular envelope may be calculated from the one-dimensional scattering profile. Various programs, such as DAMMIN and DAMMIF in the ATSAS suite, create *ab initio* bead models by iteratively altering a random bead configuration with a theoretical scattering pattern that aligns with the experimental data (196).

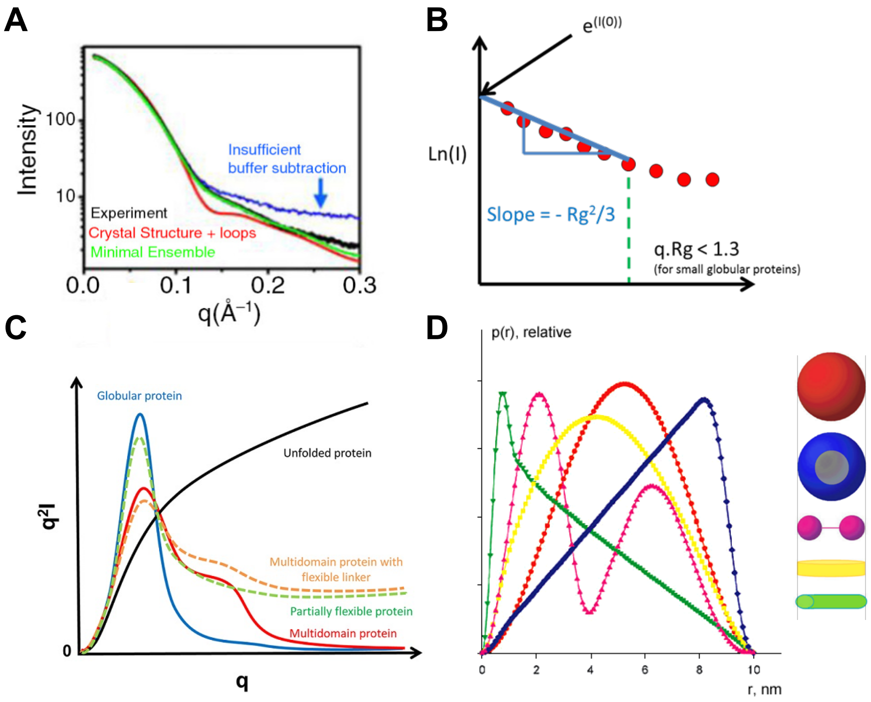


Figure 2.4. Basic example analyses of SAXS data for proteins of various conformations. (A) An experimental scattering profile, with the scattering intensity plotted vs momentum transfer (From (197). (B) A Guinier plot, where the R_g is calculated from the slope of the line, and the y-intercept yields the forward scattering intensity (From http://psldc.isis.rl.ac.uk/Guinier/Guinier%20plot.htm). (C) A Kratky plot, indicating proteins with various folding states (From https://www-ssrl.slac.stanford.edu/~saxs/analysis/assessment.htm). (D) *P(r)* curves corresponding to proteins of diverse shapes (From https://www.saxier.org/forum/viewtopic.php?t=663).

2.2.6.2. SAXS: the Experiment

SAXS data were recorded at BioCAT (beamline ID-18) at the Advanced Photon Source. Scattering data were collected using a Pilatus 3 1M detector at a distance of 3 m from the sample at a wavelength of 1.03 Å (~12 keV), covering a momentum transfer range (*q*) of 0.006-0.35 $Å^{-1}$. Prior to measurements, an inline Superdex 200 (10/300) Increase column was equilibrated with 25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol, 0.05% w/v FC-13. Protein samples comprising either 1.3 mg PupA, 4 mg PupB, 2.3 mg $\text{PupA}^{\text{PupB-NTSD}}$, 1:1 $\text{PupA}^{\text{PupB-NTSD}}$:PupR-CCSSD (3 mg total protein), or 1:1 PupB:PupR-CCSSD (4 mg total protein) were injected onto the SEC column with a flow rate of 0.6 mL/min, and the scattering recorded by exposing the column eluate to the X-ray beam for 1 sec with a 2 sec delay. Samples containing the transporter plus the PupR-CCSSD were combined and incubated for 30 min at 4 °C prior to data collection. Measurements were performed at ambient temperature. Scattering data were normalized to the incident X-ray beam, and scattering from the SEC buffer was subtracted with Igor Pro and BioCAT beamline in-house software.

Data analysis was performed utilizing the ATSAS suite (198). PRIMUS was used for data merging, calculating the radius of gyration with a Guinier approximation, and evaluating protein order by Kratky plot (193). Validation of the monodispersity of the protein samples was accomplished by examining the linearity of the Guinier region. The pair distribution function and maximum particle dimension were determined in GNOM (199). For each protein sample, 10-20 independent *ab initio* bead models that agreed with the experimental scattering data were calculated in DAMMIF (196). The resulting molecular envelope was fit with the available FpvA X-ray crystal structures (PDB entry 2O5P) with either SUPCOMB (200) or SITUS (201,202).

To evaluate the conformational plasticity of the region between the NTSD and the plug domain of the transporter, which includes the TonB box, residues 118-135 or 75-135 of FpvA were specified as flexible, and used as the input for EOM 2.0 in ATSAS. EOM was utilized to generate flexible conformers of transporters that align with the SAXS profile, assigning the NTSD and plug-barrel domains as rigid bodies, specifying of the linker between the two domains (including the TonB box) as a flexible region, and inputting the full FpvA sequence.

Memprot (April 2017 release, (203)) was employed to model the detergent corona around the transmembrane surface of the transporters, using the FpvA crystal structure as the input model (PDB 2O5P chainA), following the general methodology for modeling of aquaporin-0 in DDM, and incorporating local modifications from the modeling of HasR in DDM (186,204,205). A hollow toroidal detergent corona was built around the transmembrane portion of FpvA, with the inner hydrophobic and outer hydrophilic phases distinctly modeled. Input values included the number of electrons in the hydrophilic head and hydrophobic tail of FC-13 (97 and 105, respectively), D_{max} determined from the SAXS data, an estimation of the radius of the transmembrane region of the protein, the q range to be used in the calculations (0-0.3 $Å^{-1}$), and the densities of the hydrophilic and hydrophobic regions of the detergent molecule (206). Parameters describing the detergent belt that fit the experimental SAXS data were identified with Memprot, including the height of the micelle, a; the height of the hydrophilic region of the detergent belt, t; the cross-sectional axis of the detergent-micelle complex, b; and the ellipticity of the torus, e. For this analysis, b was locally modified to reflect the local protein radius, which is not a rigid parameter for β-barrels, as they are inherently asymmetric. Output from Memprot was automatically fed into CRYSOL to fit the theoretical scattering calculated from the generated PDC model to the experimental scattering, to identify parameter values that accurately represented the experimental scattering.

2.3. Results

2.3.1. Description of Transporters/Transducers of *Pseudomonas capeferrum*

A manual keyword search of the annotated genome of *Pseudomonas capeferrum* was conducted using the NCBI Genome Browser (https://www.ncbi.nlm.nih.gov/genome/) to identify and further classify TonB-dependent receptors as transporters or transducers. Twenty-two distinct TBDTs were found; all but one of the TBDTs are hypothesized to transport iron complexes. The ligand for this last transporter is currently unknown. Characterization of the TBDTs as transporters or transducers was based on the identification of conserved *IRB* architecture (Table 2.1). The secretin and TonB N-terminus short (STN) domain (the Pfam classification for the NTSD) was absent in all transporters, with the exception of two proteins, one of which is PupA. All of the identified transducers are A-type TBDTs, in which the sigma and anti-sigma factors are encoded by an operon directly upstream of the transducer (71).

Table 2.1. Genes in *Pseudomonas capeferrum* annotated as TonB-dependent receptors.

Uniprot ID	Protein name	Classification	Number of residues	Ligand	STN?	Regulator + sigma factor?
A0A084CJ44	PupA	Transporter	819	Pseudobactin 358	Yes	No
A0A084CH10	PupB	Transducer	809	Pseudobactin BN7/BN8	Yes	Yes
A0A084C9T9	--	Transducer	811	*	Yes	Yes
A0A084CAI5	--	Transporter	713	*	No	No
A0A084CHK2	--	Transporter	716	*	No	No
A0A084CE89	--	Transporter	721	*	No	No
A0A084C4R7	--	Transducer	802	Ferrioxamine	Yes	Yes
A0A084CBV3	--	Transducer	809	*	Yes	Yes
A0A084C5R2	--	Transporter	765	*	No	No
A0A084CCF8	--	Transducer	814	*	Yes	Yes
A0A084CBV9	--	Transporter	770	*	No	No
A0A084CJ99	--	Transducer	813	*	Yes	Yes
A0A084CI55	--	Transducer	806	*	Yes	Yes
A0A084CI43	--	Transducer	825	*	Yes	Yes
A0A084CD49	--	Transporter	684	#	No	No
A0A084CDS8	--	Transporter	858	*	No	No
A0A084C901	--	Transducer	810	Ferripyoverdine	Yes	Yes
A0A084CF11	--	Transporter	802	*	Yes	No
A0A084CHI4	--	Transducer	809	*	Yes	Yes
A0A084CGZ1	--	Transporter	712	*	No	No
A0A084C6Z8	--	Transducer	796	*	Yes	Yes
A0A084C5R7	--	Transporter	695	*	No	No

* Unspecified siderophore. # Unknown ligand

2.3.2. Expression and Purification of PupA, PupB, and PupA$^{\text{PupB-NTSD}}$

A chimeric transducer was successfully cloned by replacement of the *pupA-NTSD* sequence with the *pupB-NTSD* within the full-length *pupA* gene. Similarly, a pET17b-based expression vector was constructed to express of full-length PupB, with a His_6 tag inserted by site-directed mutagenesis between the signal peptide and N-terminal signaling domain.

Full-length PupA, PupB, and the PupA$^{\text{PupB-NTSD}}$ chimera were expressed in BL21(DE3)pLysS, solubilized with 1% w/v FC-13, and purified by IMAC and SEC in the presence of 0.1 or 0.05% w/v FC-13 to maintain protein solubility. Relatively homogeneous samples were obtain from this purification strategy, as qualitatively assessed by SDS-PAGE (Figure 2.5), although the final yields were quite low, as expected for purification of an integral membrane protein. PupA could be exchanged into 0.01% w/v LDAO (a significantly less expensive detergent) during IMAC; PupB precipitated heavily in 0.01-0.02% LDAO. Stability of the chimera in LDAO was not explored. Precipitate of PupB was also observed in 0.05% w/v FC-13, but the majority of protein aggregate could be separated by SEC. Protein was concentrated using 100 kD MWCO filtration devices, to minimize concentration of free detergent micelles.

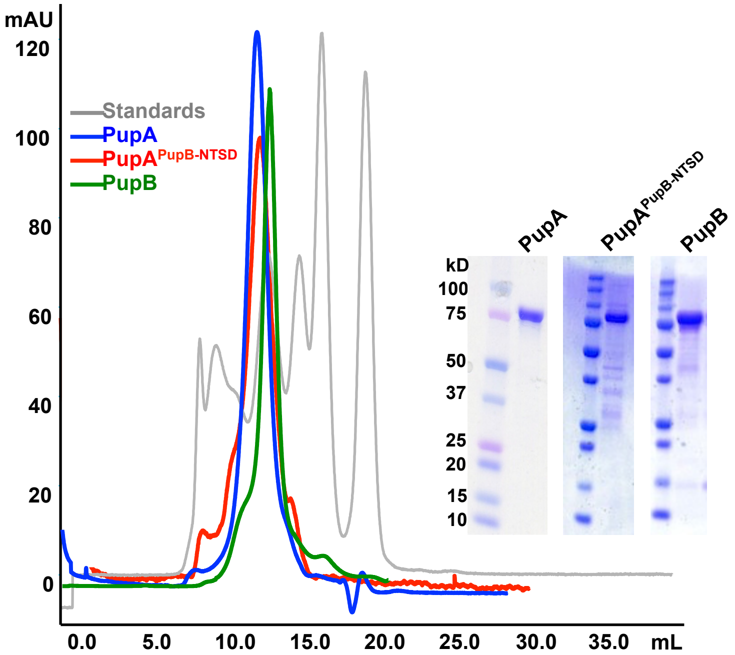


Figure 2.5. Superimposed SEC profiles of PupA, PupA$^{PupB\text{-}NTSD}$, and PupB solubilized in 0.05% w/v FC-13, with SDS-PAGE results. Apparent molecular mass of the transporters are ~150 kD. Theoretical molecular weight: PupA, 87 kD; PupA$^{PupB\text{-}NTSD}$, 86 kD, PupB, 84 kD (inset).

2.3.3. Initial Crystallization and X-ray Diffraction of PupA

Crystallization conditions for PupA previously identified by C. Colbert were optimized to improve crystal quality and diffraction properties, as prior experiments had indicated a poorly ordered crystal lattice. PupA crystals grown in 0.1% or 0.2% LDAO were improved to >100 µm in one dimension (Figure 2.6A). These conditions also yielded showers of 30-50 µm crystals (Figure 2.6B), and crystals with visible defects near the edges (Figure 2.6C). Bicelles, which have led to successful crystallization and structure solution of other β-barrel membrane proteins (207,208), were incorporated to explore an additional medium for crystallization. Protein stocks were combined in a 1:4 bicelle:protein ratio, and used for crystallization experiments, following the reservoir conditions for crystallization with LDAO. However, all crystals obtained typically had one dimension smaller than ~50 µm (Figure 2.7), and it was unclear whether the microcrystals grown under these conditions were protein or detergent/phospholipid. Screening of the crystals at NE-CAT beamline 24-ID-E did not clarify this issue, as no diffraction was detected.

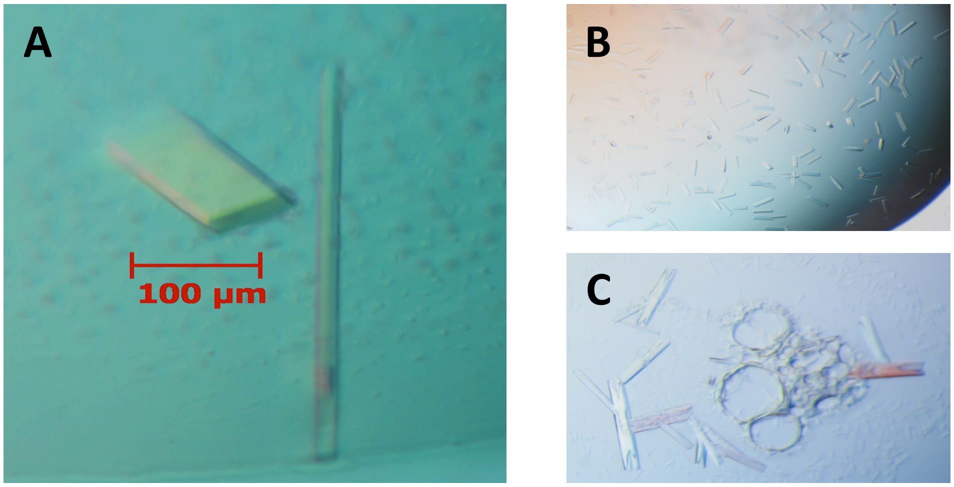


Figure 2.6. Characteristic PupA crystals. Crystallization conditions: (A,B) PupA, 11 mg/mL; Reservoir, 50 mM Tris pH 7.8, 350 mM NaCl, 5% v/v glycerol, 26% PEG 1k, 0.2% w/v LDAO. (C) PupA, 11 mg/mL; Reservoir, 50 mM Tris pH 7.8, 350 mM NaCl, 28% PEG 1k, 0.2% w/v LDAO. Trays were incubated at 20 °C; crystals were visible after one week.

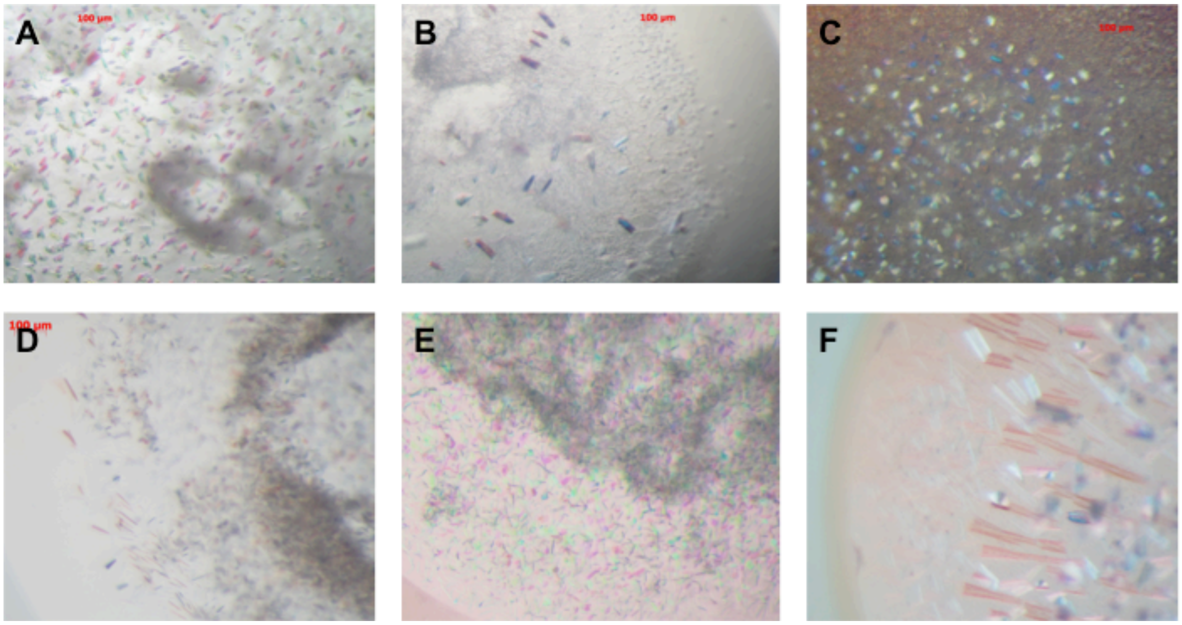


Figure 2.7. PupA crystals grown in DMPC:CHAPSO bicelles. Reservoir solution: 50 mM Tris pH 7.8, 350 mM NaCl; Protein: 6.3 or 4 mg/mL in 25 mM HEPES pH 7.5, 150 mM NaCl, 0.2% LDAO. (A) 14% PEG 1k, 2% v/v glycerol, 2.5% w/v DMPC:CHAPSO. (B) 14% PEG 1k, 2% v/v glycerol, 3.5% DMPC:CHAPSO. (C) 14% PEG 1k, 1% v/v glycerol, 4% DMPC:CHAPSO. (D) 18% PEG 1k, 8% v/v glycerol, 2.5% DMPC:CHAPSO. (E) 18% PEG 1k, 8% v/v glycerol, 3% DMPC:CHAPSO. (F) 18% PEG 1k, 4% DMPC:CHAPSO. Trays were incubated at 20 °C; images were acquired after one week.

To further improve protein packing and crystal contacts, multiple cryoprotectant solutions were analyzed, and crystal dehydration immediately prior to flash-freezing was assessed. As availability of reasonably-sized crystals for diffraction experiments was limited, several dehydration and cryoprotectant conditions were explored, but not duplicated. Of the approaches defined in Chapter 2.2.5, limited crystal dehydration, followed by washing of the crystal in MiTeGen cryo oil, yielded the highest resolution during diffraction experiments to date. A single lattice with reasonable reflection intensities was observed (Figure 2.8), and one data set was collected, although completeness to ~6 Å was quite poor (Table 2.2). The data set was processed with RAPD and autoPROC, to acquire initial collection statistics (Table 2.2).

Unfortunately, due to low resolution and a minimal number of unique reflections, the only useful information derived from the data set was determination of the space group, *$P2_12_12$*.

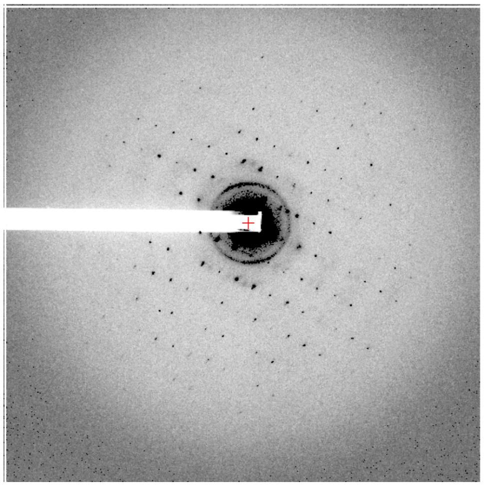


Figure 2.8. Diffraction pattern from the PupA crystal with highest resolution (6.9 Å). The PupA crystal was grown in 50 mM Tris pH 7.8, 350 mM NaCl, 28% PEG 1k, 0.2% w/v LDAO at 20 °C with 11 mg/mL protein solution. The crystal was briefly dehydrated, as described above, then washed with MiTeGen cryo oil and flash-cooled in LN_2.

Table 2.2. X-ray data collection statistics for PupA.

	His_6PupA (RAPD)	**His_6PupA (autoPROC)**
Data collection		
Beamline	24-ID-E	24-ID-E
Wavelength (Å)	0.9792	0.9792
Space group	*P222*	*P222*
		(PHENIX: *$P2_12_12$*)
Unit-cell parameters (Å, deg)	79.2, 83.8, 315.7	79.5, 84.1, 317.1
	α, β, γ = 90	α, β, γ = 90
Resolution range (Å)	157.86-6.92 (7.29-6.92)	63.43-8.32 (8.4-8.32)
Unique observations	3679	2162
Multiplicity	3.4 (3.4)	3.3 (3.6)
Completeness (%)	97.6 (96.4)	95.7 (100)
Mean *I/σI*	10.8 (0.7)	10.0 (2.9)

Values in parentheses pertain to the highest resolution shell. The same data set was processed with either RAPD or autoPROC.

2.3.4. Differential Detergent Filtration Assay with PupA

A differential detergent filtration assay was utilized to identify detergents in which PupA was stable and soluble, with a critical micelle concentration (CMC) compatible with forming a low MW protein-detergent complex (PDC). This latter factor is especially desirable, as the space occupied by the PDC is often quite large, which impacts protein stability, and how contacts are formed during crystallization (189). This assay provided a high-throughput method to evaluate 94 chemically-diverse detergents. For this assay, protein stability is defined as the quantity that is inversely proportional to the fraction of aggregate/large particle PDCs. This method employs a preliminary affinity chromatography step, followed by elution of the bound protein, and separation by filter plates of two different molecular weight cut off (MWCO) values: 100 kD and 300 kD. Protein that is unstable upon buffer exchange into the new detergent solution precipitates on the affinity resin. Protein that elutes via low MWCO filtration is considered stable. Stable protein may also elute via high MWCO filtration, but in this case, it would form a large PDC. Protein eluted by both filtration devices was quantified by Western dot blot, with an anti-His_6-HRP antibody for detection of His_6-PupA (Figure 2.9).

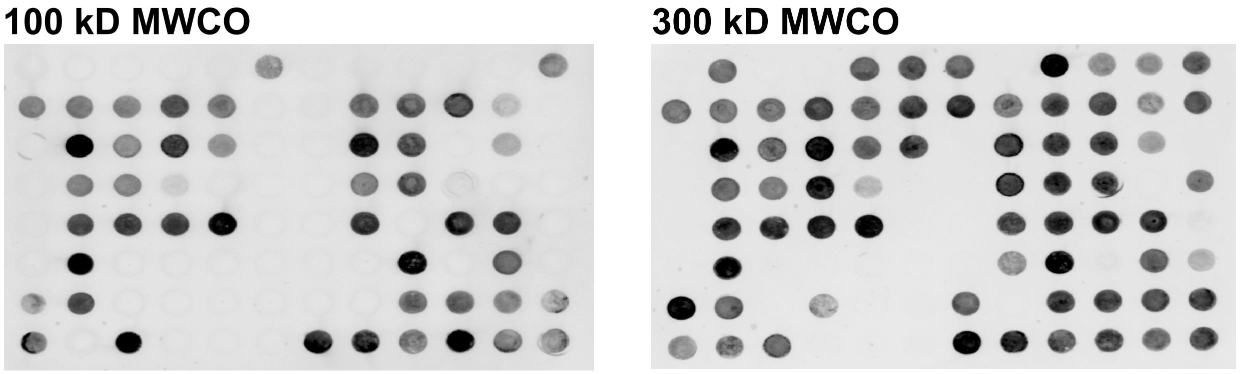


Figure 2.9. Dot blots of eluted PupA after exchange into new detergent solutions. Ten μL of the elutions from each filtration plate was spotted onto nitrocellulose membrane, and visualized by Western blot with anti-His_6-HRP. His_6-tagged PupA was detected with ECL Western blotting substrate; the blots were imaged with an Odyssey fluorescence plate imager.

The normalized intensities derived from the dot blots were plotted with as a quartile grid, with the protein filtrate from the high MWCO plate plotted on the x-axis, and the ratio of the low/high MWCO filtrates on the y-axis (Figure 2.10). This data display allowed for rapid identification of detergents with a low PDC that were stable in solution. Twenty-four detergents with various chemical properties were detected (Table 2.3).

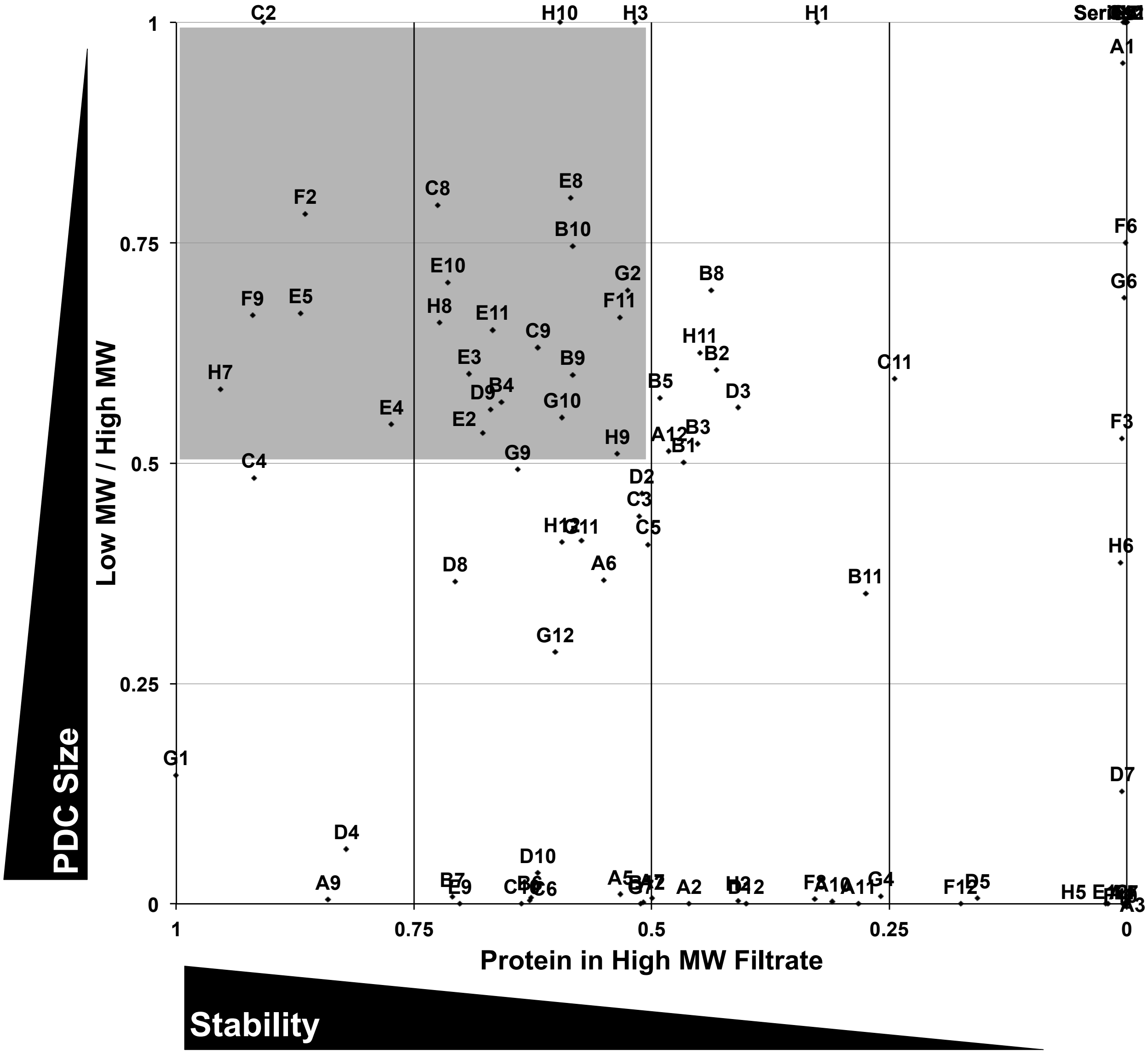


Figure 2.10. Quartile grid plot of normalized intensities derived from Western blot analysis of the filtrates from the DDFA. The well numbers from the detergent solution plate are indicated next to their corresponding point. The region in grey represents the condition with the most stable protein with the smallest PDC size.

Table 2.3. Detergents compatible with PupA stability and solubility identified with the DDFA.

Position in Assay	**Detergent**	**Abbreviation**	**CMC (mM)**
B4	Fos-choline 12	FC-12	1.5
B9	1,2-Dihexanoyl-sn-glycero-3-phosphocholine	6-DHPC	15.0
B10	1,2-Diheptanoyl-sn-glycero-3-phosphocholine	7-DHPC	1.4
C2	14:0 LysoPC (1-tetradecanoyl-2-hydroxy-sn-glycero-3-phosphocholine)	LPC-14	0.06
C8	N-(1,3-bis(Glucopyranoside)propan-2-yl)-3-Butyl-3-Cyclohexylheptanamide	CY-TRIPGLU	1.8
C9	N-(1,3-bis(Glucopyranoside)propan-2-yl)-3-Butyl-3-Phenylheptanamide	Ph-TRIPGLU	3.6
D9	Tetraethylene glycol monooctyl ether	C8E4	8.0
E2	Octaethylene glycol monodecyl ether (Anapoe C12E8)	C12E8	0.09
E3	Polyoxyethylene(9)dodecyl ether (Anapoe C12E9)	C12E9	0.05
E4	Polyoxyethylene(10)dodecyl ether (Anapoe C12E10)	C12E10	0.2
E5	Polyoxyethylene(8)tridecyl ether (Anapoe C13E8)	C13E8	0.1
E8	Octyl-2-hydroxyethyl-sulfoxide	OHES	24.2
E10	n-Heptyl-β-D-thioglucopyranoside	HTG	29.0
E11	n-Octyl-β-D glucopyranoside	OG	18.0
F2	3-Cyclohexyl-1-propyl-β-D-glucoside	CYGLU-3	28.0
F9	Trans-4-(trans-4’-propylcyclohexyl)cyclohexyl-α-D-maltoside	PCC-a-M	0.036
F11	3-Cyclohexyl-1-propyl-β-D-maltoside	CYMAL-3	30.0
G2	6-Cyclohexylhexyl β-D-maltoside	CYMAL-6	0.56
G10	n-Undecyl-β-D-thiomaltopyranoside	UDM	0.59
H3	n-Octyl-β-D-thiomaltopyranoside	OTM	8.5
H7	n-Dodecyl-β-D-thiomaltopyranoside	DDTM	0.05
H8	Octyl glucose neopentyl glycol	MNG-8	1.02
H9	Decyl maltose neopentyl glycol	MNG-10	0.036
H10	Lauryl maltose neopentyl glycol	MNG-12	0.01

2.3.5. SEC-based Pull-down Assay of PupR-CCSSD with PupA and PupB

Pull-down analysis was conducted with purified PupR-CCSSD, full-length PupA, and full-length PupB, to assess complex formation of the full-length transmembrane proteins with the PupR-CCSSD. Both PupA and PupB were purified individually, with the final SEC step utilizing 25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol, 0.05% w/v FC-13 as the elution buffer. Prior pull-down trials had indicated that the buffer typically used for SEC of PupA or PupB caused PupR-CCSSD to precipitate; thus, all proteins were exchanged into the buffer in which PupR-CCSSD was stable, with the addition of 0.05% w/v FC-13 to maintain solubility of the membrane proteins. PupA or PupB was combined with PupR-CCSSD in a 1:1 molar ratio, and incubated at 4 °C for ~30 min prior to SEC. Concentrations and availability of all proteins necessitated the usage of low quantities of protein, which resulted in weak absorbance signals during SEC elution and also in SDS-PAGE (Figure 2.11A-B). However, a small leftward shift in the elution peak of PupB:PupR-CCSSD was discernible, as compared to the SEC profiles of the individual proteins, suggested complex formation. Conversely, no shift was observed in the SEC profile of the PupA:PupR-CCSSD mixture, implying that a complex was not constituted.

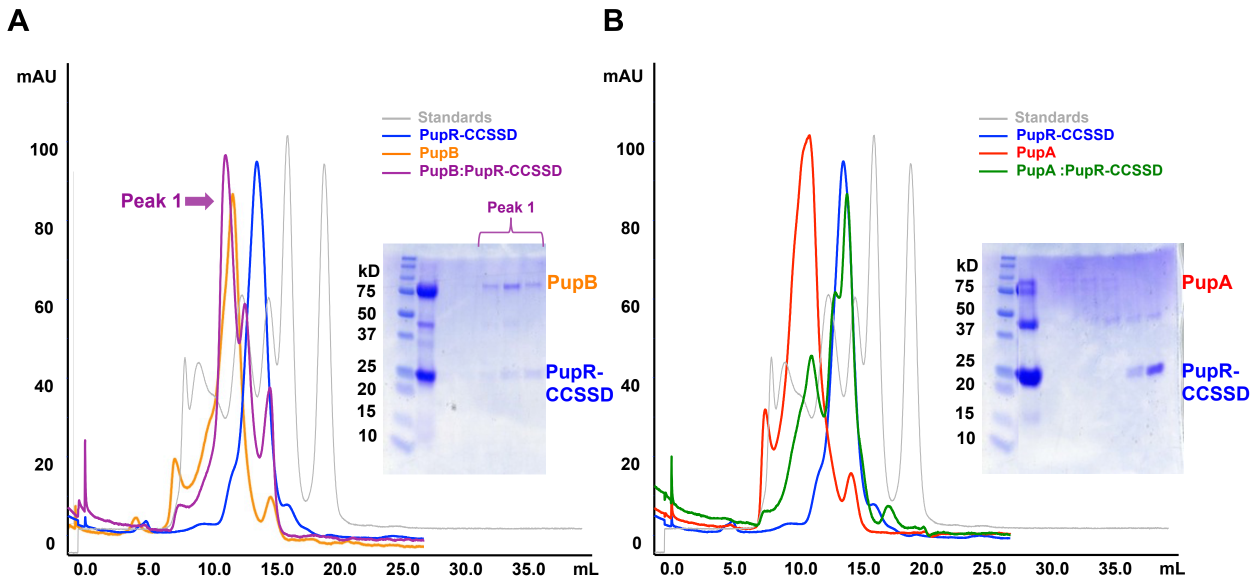


Figure 2.11. SEC-based pull-down analysis of PupR-CCSSD with PupA and PupB. (A) SEC and SDS-PAGE profiles of PupR-CCSSD, PupB, and a 1:1 molar ratio of PupB:PupR-CCSSD. (B) SEC and SDS-PAGE profiles of PupR-CCSSD, PupA, and a 1:1 molar ratio of PupA:PupR-CCSSD.

2.3.6. SEC-SAXS Analysis of the PupA Transporter, PupB Transducer, and PupA$^{\text{PupB-NTSD}}$ Transducer, and Complexes of the Transducers with the PupR-CCSSD

SEC-SAXS experiments were performed to investigate the size, shape, and solution structure of the PupA transporter, the PupB transducer, the PupA$^{\text{PupB-NTSD}}$ transducer, and complexes of PupB:PupR-CCSSD and PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD, and detergent molecules forming a toroid around the transmembrane region of the transporter/transducers were modeled, as these contributed to the SAXS signal. SAXS measurements were recorded concurrent with elution of the protein:detergent complex from a size exclusion column (Figures 2.12A, 2.14A 2.15A, 2.16A, 2.17A, 2.18A, 2.19A). The incorporation of inline SEC enabled examination of homogeneous protein:detergent samples, without aggregated protein or free detergent micelles. The scattering profiles for all data sets displayed a two-component ellipsoid curve, indicative of a protein:detergent complex (209,210) (Figures 2.12B, 2.14B 2.15B, 2.16B, 2.17B, 2.18B, 2.19B). Similarly, Kratky plots displayed a sharp rise at $q > 0.2$, which is also characteristic of a PDC. All data examined in the low q region in the Guinier plots confirmed linearity, indicating the samples were monodisperse (Figure insets). All molecular weight estimations from the SAXS data were significantly different than expected (>10%) (see Table 2.4), which has previously been observed for SEC-SAXS of other β-barrel PDCs (186). Therefore, estimation of MW must proceed with caution for the transducer:CCSSD complexes, as the MW difference of the complex may be masked by errors in MW estimates of only the transducers. A data set was collected from two different protein preparations of both PupB and the PupB:PupR-CCSSD complex.

The same methodology was followed for data analysis of the PupA transporter, PupB transducer, and PupA$^{\text{PupB-NTSD}}$ transducer. Following buffer subtraction, EOM in ATSAS was utilized for SAXS-based refinement of the NTSD orientation and length and flexibility of the linker between the NTSD and the plug domain, using the X-ray crystal structure of the *P. aeruginosa* transducer FpvA as the input molecule. The resulting models were visually evaluated; candidate structures wherein the NTSD was located in the same plane as the β-barrel were eliminated from further refinement, as downstream optimization would have resulted in steric clashes between the NTSD and the detergent micelle.

Models derived from EOM which best fit the experimental data were advanced for further processing in Memprot. A hollow toroid detergent band was constructed around the transmembrane region of FpvA. Optimization of the geometric parameters of the course-grain representation of the detergent corona and adaption to FpvA culminated in identification of parameter values for models that most accurately fit the experimental data. Theoretical scattering of these models was compared to the experimental scattering using CRYSOL (211). These results are summarized in Tables 2.4 and 2.5. The final models depict the average spatial and temporal distance distributions of the full PDCs in solution.

Treatment of the SAXS data of the PupB:PupR-CCSSD and PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD complexes likewise involved modeling with EOM and recognition of acceptable candidate structures. The FpvA-NTSD was superimposed with the PupB-NTSD:PupR-CCSSD crystal structure (described in Chapter 4) to engineer the transducer:CCSSD complex. These models were then optimized with Memprot, for composition of the detergent belt, and subsequent evaluation for correlation of the output models with the experimental data.

2.3.6.1. Solution State Characterization of PupA

Optimization of the PDC model for fitting to the PupA data set yielded a compact structure, with a minimally flexible TonB box, similar to the FpvA crystal structures (120) (Figure 2.12E-F). The low q region in the Guinier plot was linear, indicating that the sample was monodisperse (Figure 2.12B, inset). The Kratky plot revealed minimal flexibility within the complex, as demonstrated by the near-convergence of the peak to the x-axis at higher q (Figure 2.12C). The R_g was 33 Å, and the D_{max} was 128 Å, as estimated from the $P(r)$ distance distribution plot (Figure 2.12D). The final PupA:FC-13 corona model (Figure 2.12E) exhibited an excellent fit to the experimental data, with a final χ = 1.74, which is quite comparable to PDC model fitting to SAXS data as reported for the HasR structure (186).

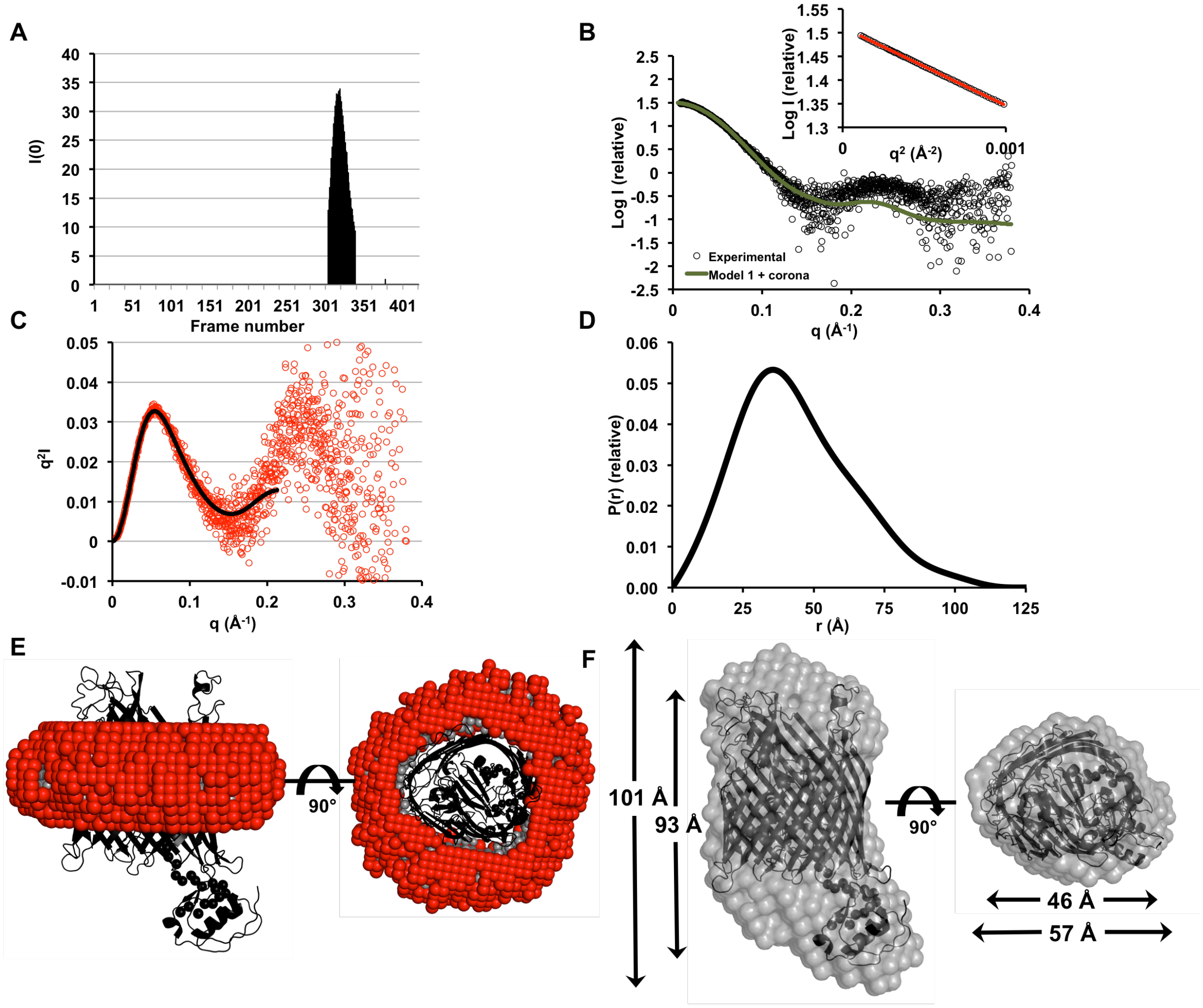


Figure 2.12. SEC-SAXS analysis of the PupA transporter. (A) Recorded scattering intensity vs. frame number, after buffer subtraction. (B) Experimental scattering profile, fit with the theoretical scattering profile of the final PDC model. The theoretical scattering of the final PupA PDC model fit the experimental scattering with a final $\chi = 1.74$. (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log (intensity). (C) Kratky profile. (D) Distance distribution function (*P(r)*) plot. (E) Overall view of the PupA:FC-13 detergent complex. Red spheres indicate the hydrophilic region of the detergent corona; grey spheres signify the hydrophobic interior. The FpvA-NTSD and plug/barrel domains are shown in black ribbon, with the flexible linker displayed as black spheres. (F) Fit of the EOM-refined FpvA structure with the *ab initio* bead models calculated from the experimental SAXS data. Bead models are shown as transparent grey surface representations.

The relative compactness of PupA in solution allowed for generation of a series of structures in EOM with reasonable χ values (Figure 2.13).

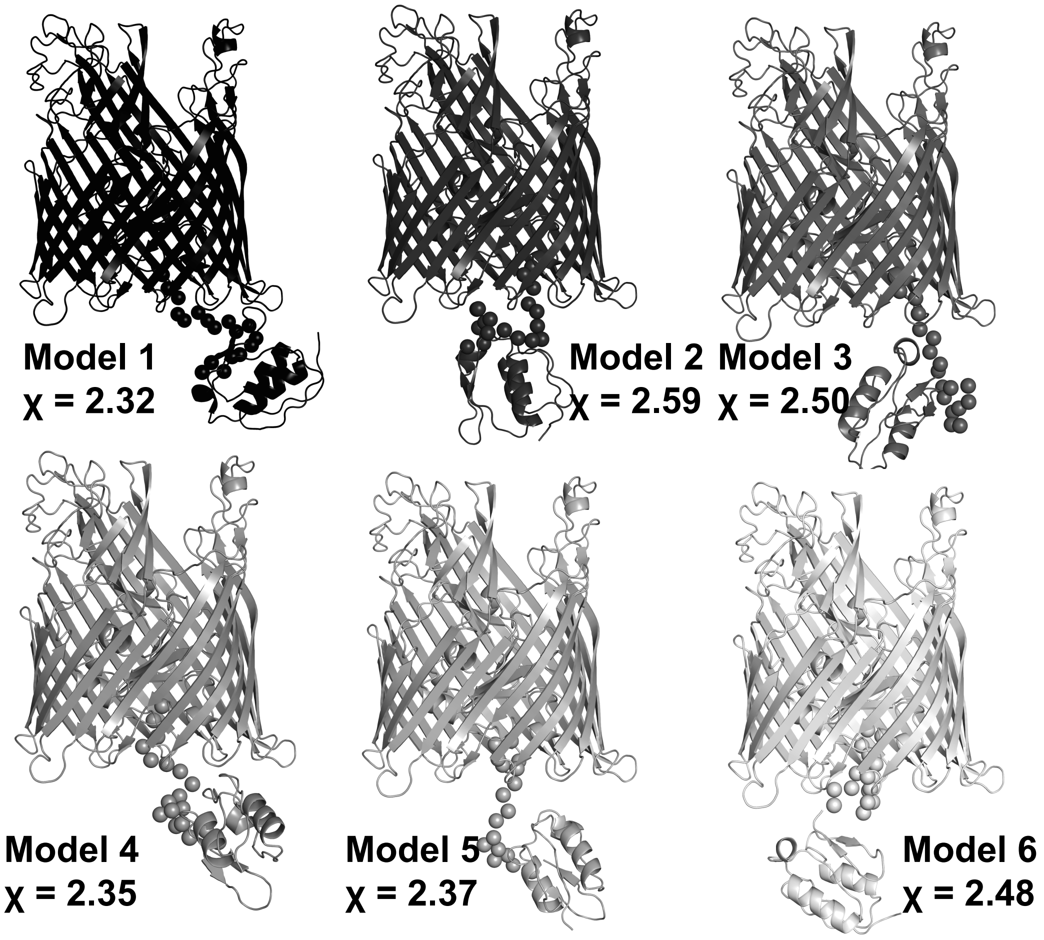


Figure 2.13. Models derived from EOM for the PupA SAXS data, using FpvA as the input structure.

2.3.6.2. Solution State Characterization of PupB

For the 1st PupB sample, the R_g was 49 Å, and the D_{max} was 238 Å, as estimated from the *P(r)* distance distribution plot (Figure 2.14D). For the 2nd PupB sample, the R_g was 45 Å, and the D_{max} was 198 Å (Figure 2.15D). The theoretical scattering of the final PupB:FC-13 corona models (Figure 2.14E) fit reasonably well to the experimental data for sample 1 and quite poorly for sample 2 (Table 2.4). The Kratky plot revealed significant flexibility within the PupB PDCs (Figure 2.14C, 2.15C). Optimization of the PDC models for fitting to the PupB data sets yielded extended structures with a protracted linker/TonB box, compared to that of PupA (Figure 2.14F, 2.15F).

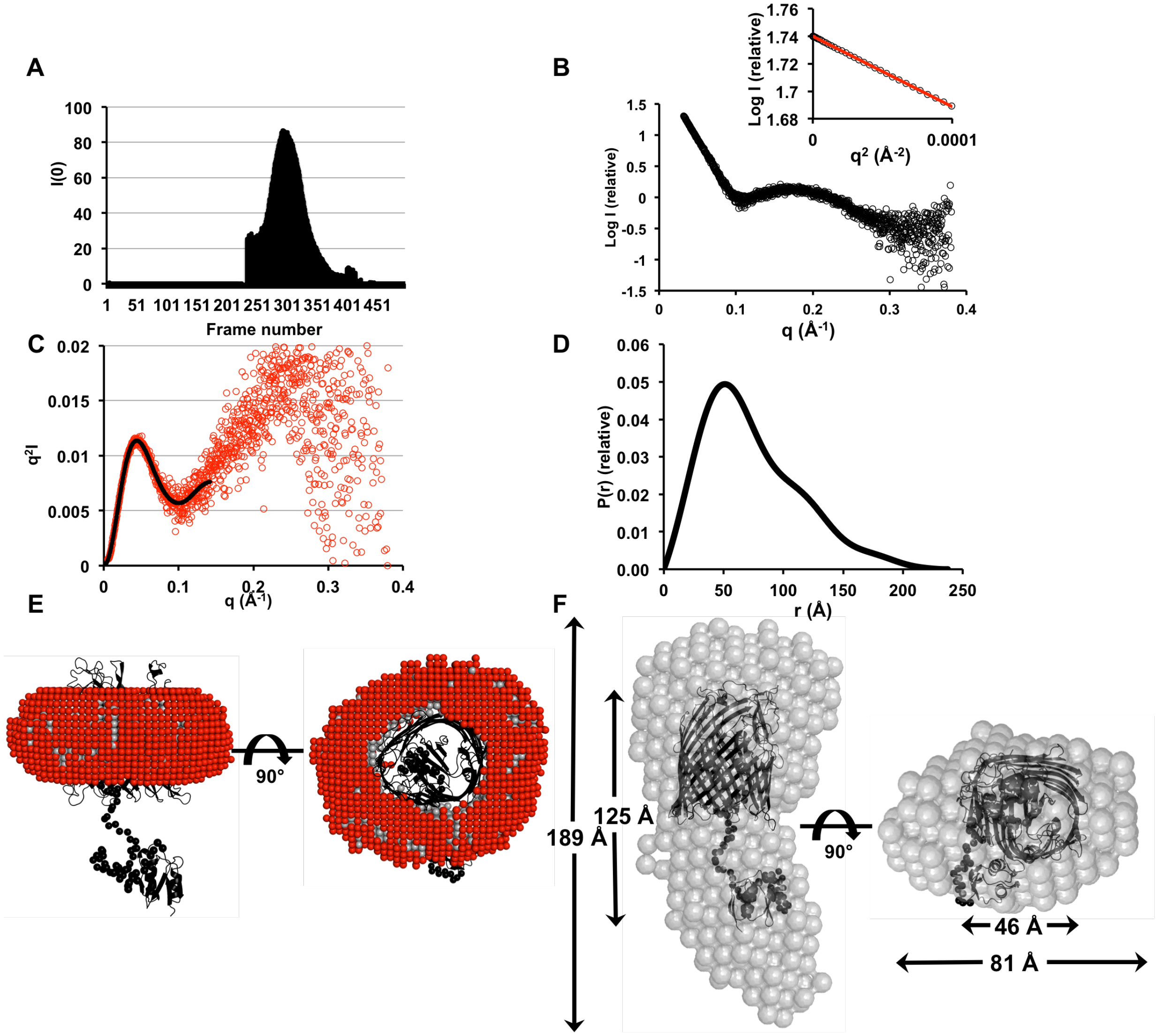


Figure 2.14. SEC-SAXS analysis of the 1st PupB sample. (A) Recorded scattering intensity vs. frame number, after buffer subtraction. (B) Experimental scattering profile. (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log (intensity). (C) Kratky profile. (D) Distance distribution function (*P(r)*) plot. (E) Overall view of the PupB:FC-13 detergent complex. (F) Fit of the EOM-refined FpvA structure with the *ab initio* bead models calculated from the experimental SAXS data.

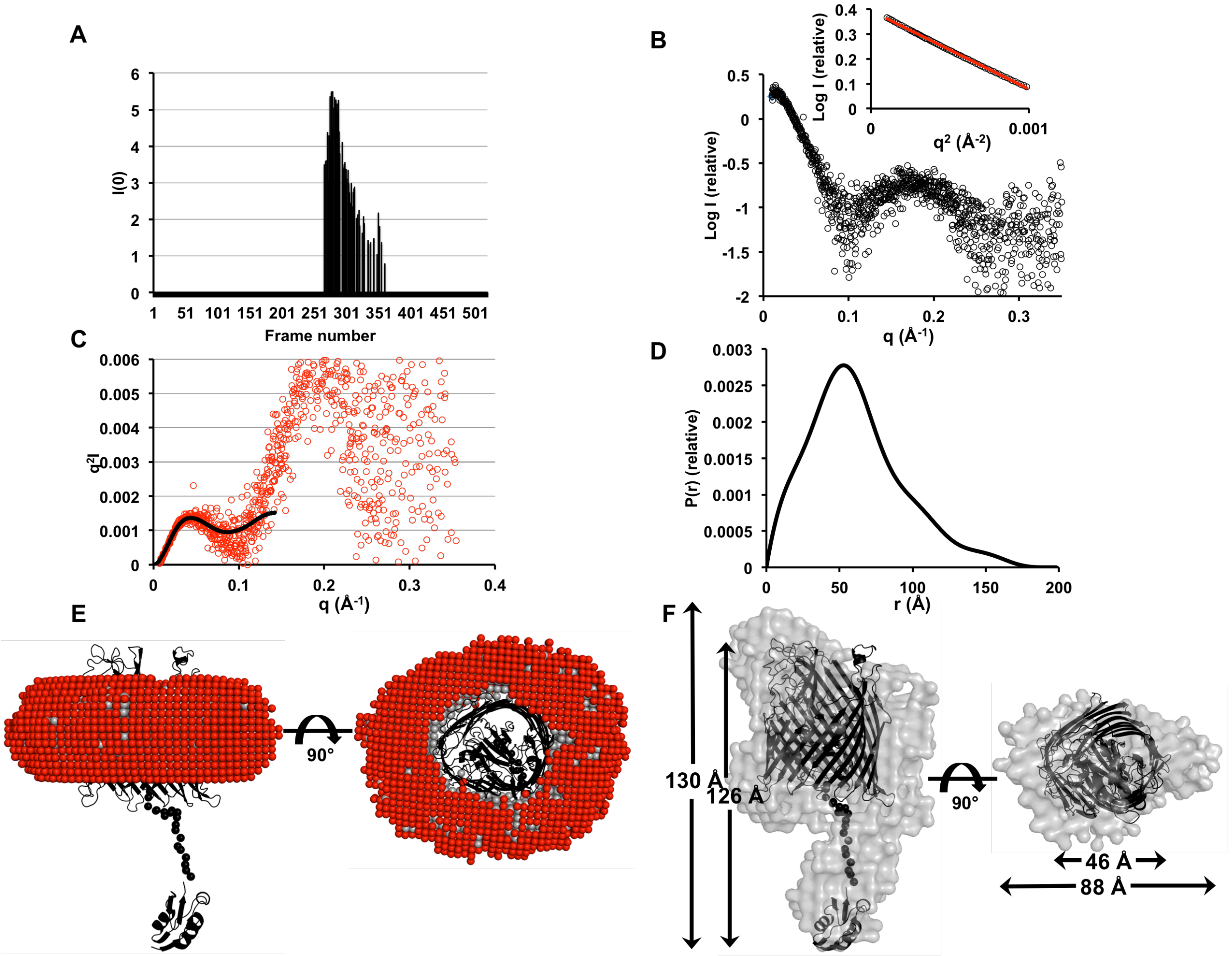


Figure 2.15. SEC-SAXS analysis of the 2nd PupB sample. (A) Recorded scattering intensity vs. frame number, after buffer subtraction. (B) Experimental scattering profile. (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log (intensity). (C) Kratky profile. (D) Distance distribution function (*P(r)*) plot. (E) Overall view of the PupB:FC-13 detergent complex. (F) Fit of the EOM-refined FpvA structure with the *ab initio* bead models calculated from the experimental SAXS data.

2.3.6.3. Solution State Characterization of PupA$^{PupB\text{-}NTSD}$

For the PupA$^{\text{PupB-NTSD}}$ sample, the R_g was 37 Å, and the D_{max} was 193 Å, as estimated from the *P(r)* distance distribution plot (Figure 2.16D). The theoretical scattering of final PupA$^{\text{PupB-NTSD}}$:FC-13 corona models (Figure 2.16E) fit reasonably well to the experimental data (Table 2.4). The Kratky plot revealed moderate flexibility within the PupA$^{\text{PupB-NTSD}}$ PDC (Figure

2.16C). Optimization of the PDC model for fitting to the PupA$^{\mathrm{PupB\text{-}NTSD}}$ data set yielded a somewhat extended structure, in between those of PupA and PupB (Figure 2.16F).

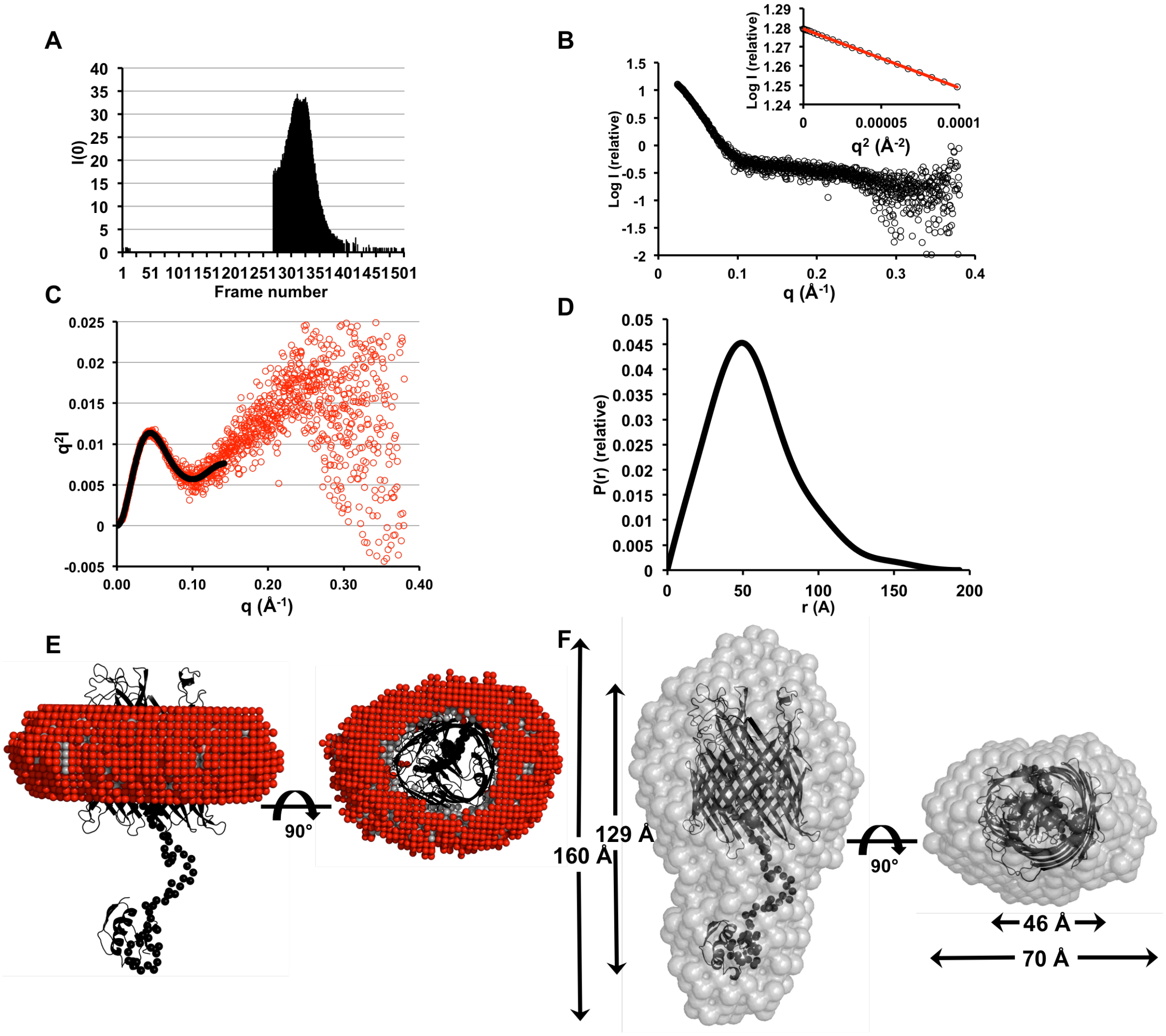


Figure 2.16. SEC-SAXS analysis of the PupA$^{\mathrm{PupB\text{-}NTSD}}$ transducer. (A) Recorded scattering intensity vs. frame number, after buffer subtraction. (B) Experimental scattering profile. (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log (intensity). (C) Kratky profile. (D) Distance distribution function (*P(r)*) plot. (E) Overall view of the PupA$^{\mathrm{PupB\text{-}NTSD}}$:FC-13 detergent complex. (F) Fit of the EOM-refined FpvA structure with the *ab initio* bead models calculated from the experimental SAXS data.

2.3.6.4. Solution State Characterization of the PupA$^{PupB\text{-}NTSD}$:PupR-CCSSD Complex

For the PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD sample, the R_g was 43 Å, and the D_{max} was 156 Å, as estimated from the *P(r)* distance distribution plot (Figure 2.17D). The theoretical scattering of the final PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD:FC-13 corona models (Figure 2.17E) fit moderately well to the experimental data (Table 2.4). The Kratky plot revealed significant flexibility within the PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD:PDC (Figure 2.17C). Optimization of the PDC model for fitting to the PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD data set yielded a lopsided structure with, presumably, the NTSD:CCSSD favored on one side of the bead model (Figure 2.17F).

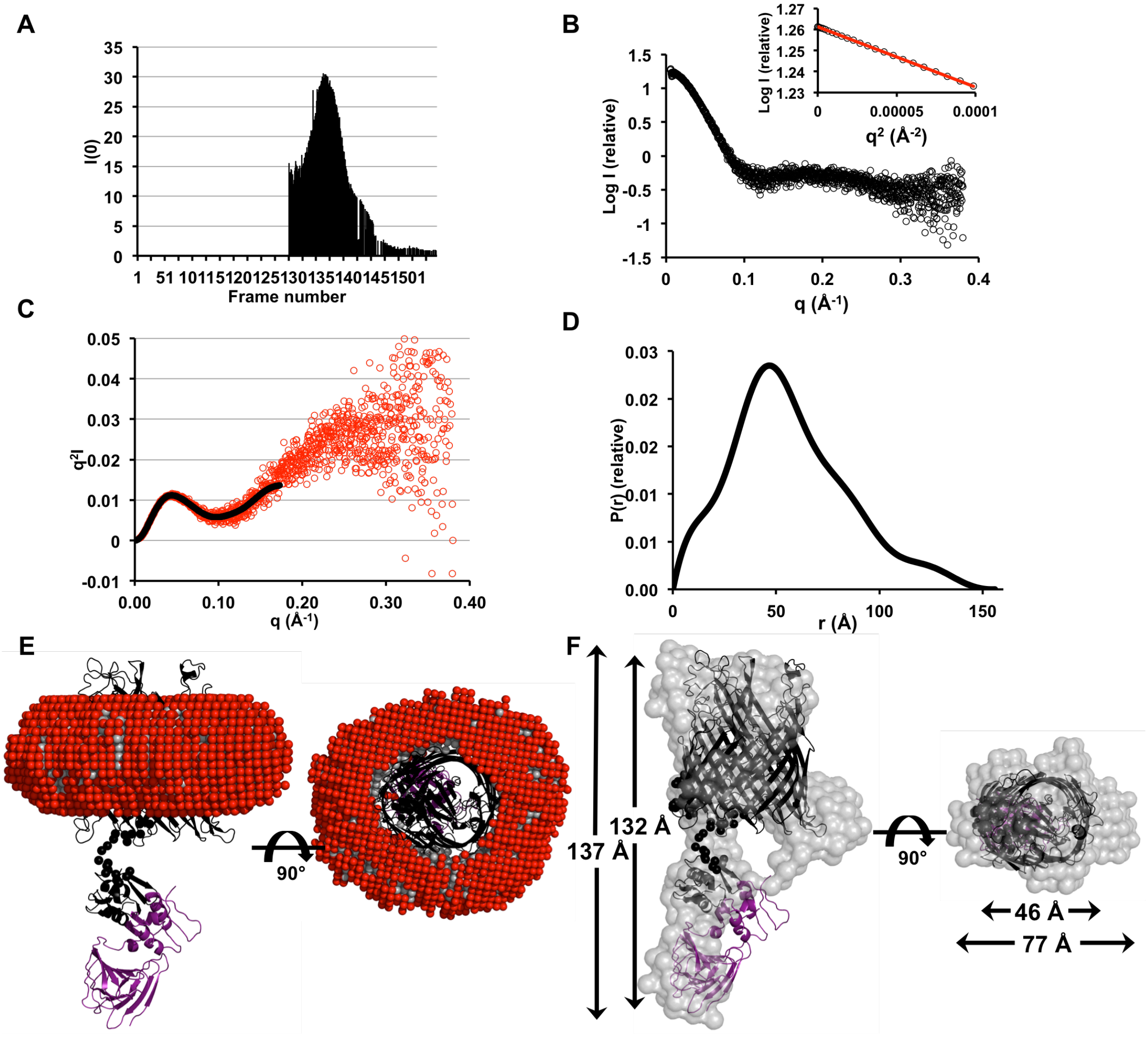


Figure 2.17. SEC-SAXS analysis of the PupA$^{PupB\text{-}NTSD}$:PupR-CCSSD complex. (A) Recorded scattering intensity vs. frame number, after buffer subtraction. (B) Experimental scattering profile. (Inset) Guinier plot of the low q region of the X-ray scattering data, as a function of log (intensity). (C) Kratky profile. (D) Distance distribution function ($P(r)$) plot. (E) Overall view of the PupA$^{PupB\text{-}NTSD}$:PupR-CCSSD:FC-13 detergent complex. (F) Fit of the EOM-refined FpvA structure and the PupR-CCSSD with the *ab initio* bead models calculated from the experimental SAXS data. The PupR-CCSSD is illustrated in purple.

2.3.6.5. Solution State Characterization of the PupB:PupR-CCSSD Complex

For the PupB:PupR-CCSSD samples, the R_g was 47 Å, and the D_{max} was 197 Å for the 1st sample; the R_g was 48 Å, and the D_{max} was 213 Å for the 2nd sample. The theoretical scattering of the final PupB:PupR-CCSSD:FC-13 corona models (Figure 2.18E, 2.19E) the experimental data fit quite poorly for the 1st sample and moderately well for the 2nd sample (Table 2.4). The Kratky plot revealed significant flexibility within the PupB:PupR-CCSSD:PDCs (Figure 2.18C, 2.19C). Optimization of the PDC models for fitting to the PupB:PupR-CCSSD data sets again revealed somewhat lopsided bead models (Figure 2.18F, 2.19F). Visually, these models agree with the calculated molecular envelopes quite well.

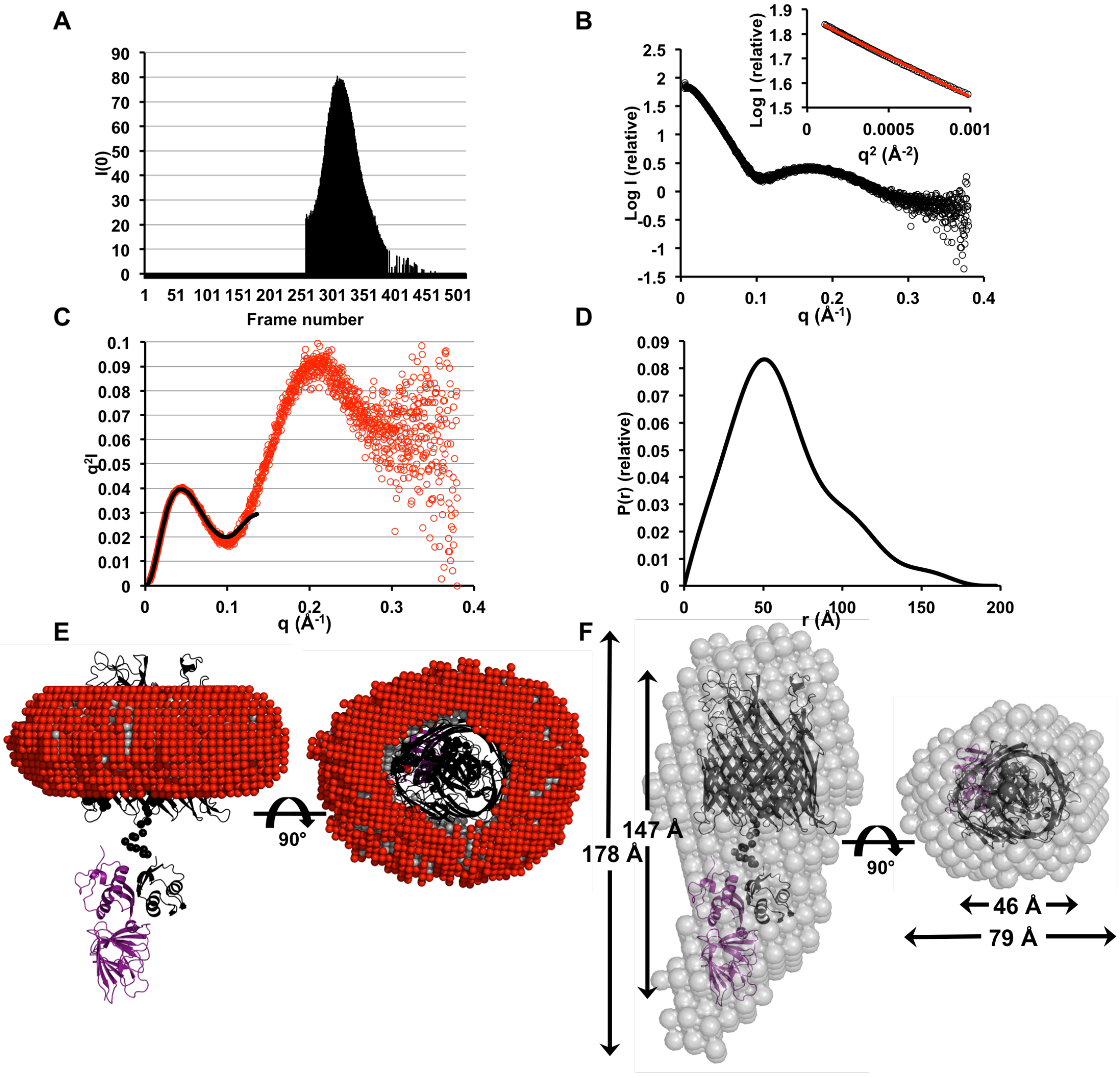


Figure 2.18. SEC-SAXS analysis of the 1st PupB:PupR-CCSSD sample. (A) Recorded scattering intensity vs. frame number, after buffer subtraction. (B) Experimental scattering profile. (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log (intensity). (C) Kratky profile. (D) Distance distribution function (*P(r)*) plot. (E) Overall view of the PupB:PupR-CCSSD:FC-13 detergent complex. (F) Fit of the EOM-refined FpvA structure and the PupR-CCSSD with the *ab initio* bead models calculated from the experimental SAXS data.

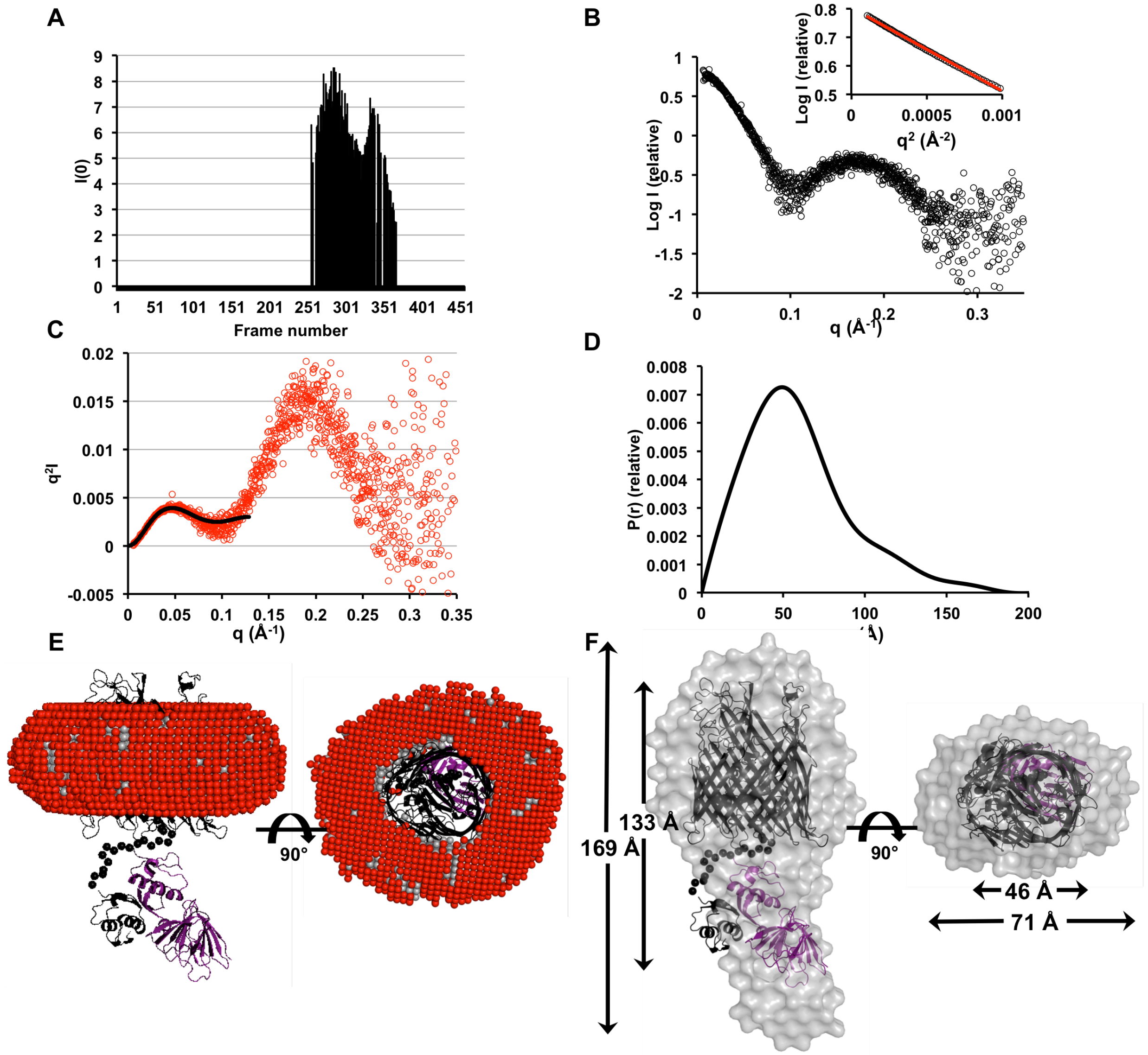


Figure 2.19. SEC-SAXS analysis of the 2nd PupB:PupR-CCSSD sample. (A) Recorded scattering intensity vs. frame number, after buffer subtraction. (B) Experimental scattering profile. (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log (intensity). (C) Kratky profile. (D) Distance distribution function (*P(r)*) plot. (E) Overall view of the PupB:PupR-CCSSD:FC-13 detergent complex. (F) Fit of the EOM-refined FpvA structure and the PupR-CCSSD with the *ab initio* bead models calculated from the experimental SAXS data.

2.3.6.6. Summary of SEC-SAXS Studies of the Pup Transporter and Transducers

SAXS analyses of the PupA transporter, the PupB transducer, the PupA$^{\text{PupB-NTSD}}$ chimeric transducer, and the PupB:PupR-CCSSD and PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD complexes provides a low-resolution overview of the heterogeneous solution state of these molecules and their complexes. Additional geometric parameters determined by these computations, such as volumes and surface areas of the models described above, are outlined in Table 2.5. Comparison of the Kratky profiles and *P(r)* plots of the PupA transporter to those of the PupB and PupA$^{\text{PupB-NTSD}}$ transducers suggests greater flexibility of the transducers, as well as more extended solution structures, as demonstrated by the greater D_{max} values calculated from the transducer samples.

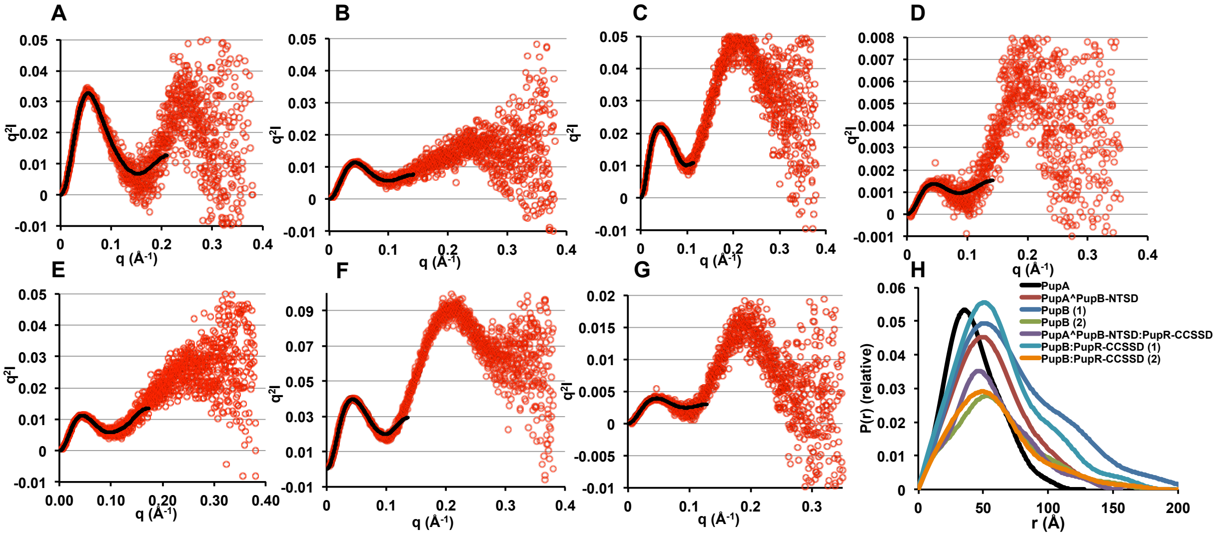


Figure 2.20. Comparison of the Kratky profiles and distance distribution function plots of all transporter/transducer samples. (A) PupA Kratky plot. (B) PupA$^{\text{PupB-NTSD}}$ Kratky plot. (C) PupB sample 1 Kratky plot. (D) PupB sample 2 Kratky plot. (E) PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD Kratky plot. (F) PupB:PupR-CCSSD sample 1 Kratky plot. (G) PupB:PupR-CCSSD sample 2 Kratky plot. (H) *P(r)* plot of all samples.

Table 2.4. Summary of parameters determined from SAXS data for PupA, PupB, PupA$^{\text{PupB-NTSD}}$, and the transducer:PupR-CCSSD complexes, and theoretical scattering of relevant X-ray crystal structures.

						Optimized detergent toroid parameters						
Sample	**Total protein (mg)**	**R_g (Å)**	**D_{max} (Å)**	**Estimated MW (kD)**	**Theoretical MW (kD)**	***a***	***b***	***t***	***e***	**Detergent aggregation number**[a]	χ^2	χ
FpvA crystal structure[b]	--	27	91	68.8	86.5							
PupA	1.3	33	128	76.1	87.0	28	2.4	3.3	1.1	158 ± 15	2.51	1.74
PupA$^{\text{PupB-NTSD}}$	2.3	37	193	71.6	86.2	32	3.0	4.2	0.85	242 ± 28	2.84	2.5
PupB (1)	2.2	49	238	75.3	84.5	34.4	3.1	4.0	1.0	276 ± 29	12.2	8.8
PupB (2)	4.0	45	198	107	84.5	35	4.8	4.6	0.9	332 ± 37	2.9	2.6
PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD	3.0	43	156	85.5	110.2	34	3.2	4.1	0.95	279 ± 36	6.5	4.6
PupB:PupR-CCSSD (1)	4.0	47	197	243	108.5	34	4.2	4.5	0.9	314 ± 48	48.02	16.0
PupB:PupR-CCSSD (2)	5.0	48	213	207	108.5	34.5	6.1	4.1	0.95	341 ± 27	7.82	4.28
PupR-CCSSD crystal structure[c]	--	19	70	21.7	24.1							

[a] Detergent aggregation number = average number of calculated heads and tails in detergent corona (average number of detergent molecules). Aggregation number (in H_2O) of FC-13 = 87; formula weight = 365.5 Da (Anatrace, Inc.). Micelle molecular weight of FC-13 = 31.8 kD (aggregation number multiplied by molecular weight of detergent monomer).

[b] PDB 2O5P (120).

[c] This work, chapter 4.

Table 2.5. Comparison of geometric parameters calculated from SAXS data to the final protein or PDC models.

Model	R_g (Å)	D_{max} (Å)	Porod volume (kÅ^3)	Calculated volume (kÅ^3)	Calculated surface area (kÅ^2)
PupA experimental atomic bead model	33	128	162.4	180.0	31.1
PupA protein model	29	106	132.9	129.1	22.3
PupA PDC model	32	112	172.0	246.7	30.5
PupA corona only	36	97	87.2	121.9	22.2
PupA$^{\text{PupB-NTSD}}$ experimental atomic bead model	44	193	363.1	430.8	53.1
PupA$^{\text{PupB-NTSD}}$ protein model	34	136	143.6	129.7	26.1
PupA$^{\text{PupB-NTSD}}$ PDC model	37	138	208.4	303.7	37.4
PupA$^{\text{PupB-NTSD}}$ corona only	39	113	123.6	179.0	27.9
PupB experimental (1) atomic bead model	53	205	499.1	596.5	68.9
PupB protein model (1)	34	132	137.2	129.5	25.6
PupB PDC model (1)	37	136	215.9	323.9	37.4
PupB corona only (1)	39	108	141.9	201.3	29.9
PupB experimental (2) atomic bead model	41	151	194.7	251.4	26.3
PupB protein model (2)	34	144	131.7	131.6	23.5
PupB PDC model (2)	38	142	228.5	364.6	38.2
PupB corona only (2)	41	114	169.8	238.4	32.3
PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD experimental atomic bead model	42	148	267.1	166.2	29.9
PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD protein model	41	155	188.5	167.9	30.2
PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD PDC model	41	152	219.6	363.1	41.8
PupA$^{\text{PupB-NTSD}}$:PupR-CCSSD corona only	39	106	141.8	202.2	29.9
PupB:PupR-CCSSD experimental (1) atomic bead model	47	187	421.4	493.8	54.9
PupB:PupR-CCSSD protein model (1)	41	157	187.1	167.5	29.7
PupB:PupR-CCSSD PDC model (1)	42	160	227.2	383.4	42.7
PupB:PupR-CCSSD corona only (1)	40	114	160.2	201.3	29.9
PupB:PupR-CCSSD experimental (2) atomic bead model	45	180	351.7	411.0	45.1
PupB:PupR-CCSSD protein model (2)	39	137	162.3	168.3	30.3
PupB:PupR-CCSSD PDC model (2)	41	151	231.2	405.5	44.5
PupB:PupR-CCSSD corona only (2)	41	114	176.5	243.9	32.6

Volumes of experimental atomic bead models were determined by DAMMIF, which calculates the volume as (volume of a single dummy atom)*(number of dummy atoms)/0.74, where 0.74 is the value of close-packing of equal spheres (196). Volume and surface area values for the protein, PDC, and corona models were obtained from 3V (www.3vee.molmovdb.org/volumeCalc.php) (212).

2.4. Summary and Conclusions

This chapter summarizes considerable efforts made toward structurally characterizing the differential outer membrane proteins PupA and PupB, and how their structural variation, particularly in the NTSD and flexible linker region, may promote signal transduction via interaction with the inner membrane sigma regulator. This represents an innovative study utilizing two homologous TBDTs from the same organism to understand iron transport and regulatory discrimination amongst various receptors and receptor complexes.

An inspection of the *P. capeferrum* genome divulged several various TonB-dependent receptors with common global architectural features. These receptors may be classified as either transporters or transducers depending on the arrangement of related genes within the same operon. Transducers are a subset of transporters, as they possess both metabolite transport activity, as well as the ability to autoregulate their expression in metabolite-deplete conditions. Transducers exhibit similar genomic arrangements, in that a sigma factor and sigma regulator that nominally control transducer expression are located immediately upstream of the transducer gene. Transporters appear to lack this operon design. Additionally, transducers include an N-terminal signaling domain, which is essential for downstream transcriptional activation upon receipt of signal. TBD receptors that are strictly transporters do not retain this domain, with the exception of two TBDTs distinguished in the *P. capeferrum* genome- *pupA* and an unidentified transporter. Since deletion of the NTSD does not affect transport, as demonstrated for FecA (138), it is unknown what function this domain retains. Previous reports claim that transporters that lack a regulatory function do not possess this domain (105,106,138); clearly, PupA is an exception.

Initially, efforts were focused on the crystallization of the receptors, in order to elucidate the structural location and orientation of the NTSD, in relation to the receptor, as the NTSD was not resolved in the X-ray crystal structures of FecA, whereas crystal structures of FpvA revealed the NTSD in three different orientations (118,120). SAXS analysis of the PupA transporter, the PupB transducer, and the PupA$^{\text{PupB-NTSD}}$ chimera sheds some light on this puzzle.

Comparison of the SAXS scattering profiles and *ab initio* bead model calculated from experimental data indicate that the linker region between the NTSD and the plug and barrel domains, containing the TonB box, extends an additional 20-30 Å in PupB and the chimera beyond that of PupA. The greater R_g, D_{max}, and volume values of the transducers suggest a more extended structure. The values presented here may even be conservative estimates, as examination of activated HasR (bound to HasB- the TonB-like protein of *S. marcescens)* appeared to extend 70-90 Å into the periplasm (186). Additionally, qualitative analysis of the flexibility of the transporter and transducers indicate greater plasticity of the transducers, most assuredly in the linker region, but possibly in other areas of the transmembrane protein, as well.

Analysis of the transducer:CCSSD complexes also provide new insight into this interaction. This work presents the first interaction to be observed *in vitro* between the full-length transducer and the periplasmic domain of the sigma regulator. Previous studies have focused on the interaction between the NTSD and the sigma regulator periplasmic domain, but none in the context of the full transducer in an environment resembling the native phospholipid bilayer. Given the considerable difference in molecular mass between the transducers and the CCSSD, and the dubious molecular mass estimates derived from the SAXS data, it is not certain that the complex was the primary species in solution. However, preliminary separation by SEC intimates that complex formation does indeed occur. Initial fitting of the SAXS experimental scattering of the transducer:CCSSD complexes with the theoretical scattering of only the transducer affirmed that scattering by the transducer alone or the transducer:detergent complex alone did not account for all experimental scattering. Likewise, the model of the transducer oligomer fit rather poorly to the experimental data. Thus, the transducer:CCSSD complex is likely observed in the SAXS data.

The final agreement of the constructed models with the experimental SAXS data was, as a whole, reasonable; serious improvements could be made, although this is currently severely limited by the lack of availability of algorithms to handle complex molecules that contain additional scattering components (detergent), multiple domains, multiple flexible regions, and protein binding partners. The programs utilized in this work incorporate one or more of these features, but none appear to handle all factors in a single calculation.

Further optimization of the solution state of PupA, PupB, and PupA$^{\text{PupB-NTSD}}$ may yet result in atomic-level details of these receptors. The results of the differential detergent filtration assay will no doubt inform future efforts to clarify structural differences between PupA and PupB that account for variability in signal transduction, in a membrane-mimicking environment.

3. COMPARATIVE ANALYSIS OF THE N-TERMINAL SIGNALING DOMAINS OF THE OUTER MEMBRANE PROTEINS PUPA AND PUPB

3.1. Introduction

The amino-terminal signaling domain (NTSD) is an essential component of the TonB-dependent iron transport CSS regulatory pathway. Deletion of this domain from the outer membrane receptor results in loss of signal transduction, although transport activity is maintained (185).

Structure solution of the FecA-NTSD, the PupA-NTSD, and the HasR-NTSD by solution state nuclear magnetic resonance (NMR), as well as the X-ray crystal structure determination of the full-length FpvA receptor (PDB 2IAH and 2O5P), reveal that the NTSD maintains a conserved, globular fold (81). This fold has been described as a βαβ-repeat or a βαβ-βαββ motif (81). An NTSD has also been identified in the pseudobactin BN7/BN8 transducer from *P. capeferrum*. However, despite sharing 33% sequence identity, 39% sequence similarity, the conserved βαβ-repeat structural motif, the PupB-NTSD is signaling competent, while the PupA-NTSD is signaling incompetent through the PupR-PupI signaling pathway (140). Substitution of the PupB-NTSD sequence with the PupA-NTSD in the full-length PupB receptor did not alter transport capacity, but rendered the protein signaling inactive (140). Conversely, replacement of the PupA-NTSD with the PupB-NTSD in the full length PupA transporter enabled signaling. This suggests that PupA either possesses its own anti-sigma factor:sigma factor regulatory network, or transcriptional regulation of PupA is maintained by a distinct control system. Either of these options are viable, as *P. putida* possesses 19 ECF sigma factors, most of which are clustered with an anti-sigma regulator (213,214). Since PupA and PupB facilitate transport of specific ferrisiderophores- PupA for pseudobactin 358 and PupB for pseudobactin BN7 and

BN8- this may reflect the adaptation of pseudobactin receptors for diverse siderophores based on microbial habitat and iron competition.

To explore structural differences between the PupB-NTSD and PupA-NTSD that allow for the former to signal through the *pupIRB* CSS regulatory pathway, but not the PupA-NTSD, both domains were expressed and purified. Isotopically-labeled PupB-NTSD was utilized for 2D and 3D NMR spectroscopy, to determine its high-resolution solution structure. The solution state of the PupB-NTSD was also probed by SEC-SAXS, and multiple mutants of the PupB-NTSD and PupA-NTSD were created to analyze binding capabilities to the PupR-CCSSD.

3.2. Materials and Methods

3.2.1. Expression and Purification of the PupB-NTSD

The N-terminal signaling domain of PupB from *Pseudomonas capeferrum*, amino acids 49-128, was originally amplified and cloned into a pGEX4T1 vector (GE Lifesciences). The PupB-NTSD fragment was excised from pGEX4T1 via *BamHI* and *XhoI* restriction sites and cloned into a pGEX$_r$ vector, which included a TEV protease site (ENLYFQG) between the glutathione S-transferase (GST) tag and the PupB-NTSD sequence.

E. coli BL21(DE3)pLysS cells were transformed with the pGEX$_r$-PupB-NTSD vector, and the transformants used for expression. Cells were grown in LB medium with 100 μg/mL ampicillin and expression induced by addition of 0.5 mM IPTG at OD_{600} = 0.6-0.8 as measured using an Implen Nano-Photometer Classic (Implen, Inc.). Following induction, the cells were incubated overnight at 20°C, collected via centrifugation at 4000 *g*, and the cell pellets stored at -80°C until use. For uniform isotope labeling, protein was expressed in cells grown in M9 minimal media containing 3 g/L [$^{13}C_6$]-D-glucose and 1 g/L $^{15}NH_4Cl$ (Cambridge Isotope Laboratories, Inc.).

To purify the PupB-NTSD, cell pellets were suspended in lysis buffer (25 mM Tris pH 7.8, 150 mM NaCl, 2 mM DTT) and disrupted in an emulsifier at 15000 psi. Cellular debris was separated by centrifugation at 20,000 *g* for 45 min. Clarified lysate was applied to glutathione Sepharose 4 Fast Flow resin (GE Lifesciences) and washed with lysis buffer. Recombinant TEV protease was applied to the column in an approximate TEV protease:protein mass ratio of 1:10, then incubated for 16 h at 4°C to remove the GST tag. Cleaved PupB-NTSD was washed off the column with lysis buffer, and yielded an 82-residue product; 80 residues comprising the signaling domain (residues 49-128), and 2 N-terminal residues as a cloning artifact (GS). The NTSD was further purified by gel filtration on a Superdex 200 column (GE Lifesciences) in lysis buffer without DTT, then concentrated to 15-20 mg/mL (1.8-2.5 mM). Concentrated protein was flash-frozen in LN_2 and stored at -80°C. Typical average protein yields were 7 mg/L (from expression in LB media) and 4 mg/L (from expression in M9 minimal media). The purity of the collected protein was estimated to be >95% by SDS-PAGE stained with Coomassie Blue (215). Protein concentration was determined by measuring absorbance at 280 nm and using the molar extinction coefficient $\varepsilon_{280} = 2980\ M^{-1}\ cm^{-1}$ and a theoretical molecular weight of 8101 g/mol.

To determine the interaction of the PupB-NTSD and PupA-NTSD with the PupR-CCSSD, the *gst-pupB-NTSD* region of the primary expression vector was cloned into pET41. Similarly, *pupA-NTSD* was cloned into pET41, with an N-terminal GST tag. Site-directed mutagenesis via the QuikChange II protocol was performed using these constructs, for downstream qualitative estimation of the binding capacity of PupR-CCSSD with PupB-NTSD or PupA-NTSD mutants. Co-expressions of MBP-PupR-CCSSD with either GST-PupB-NTSD or GST-PupA-NTSD or mutants thereof, as well as pull-down analysis of GST-NTSD mutants with MBP-PupR-CCSSD, were carried out as described in chapter 4.2.6.

3.2.2. NMR Data Collection

NMR data collection and data processing were performed by Dr. Qiong Wu at the University of Texas Southwestern Medical Center Biomolecular NMR facility. Some in-house trials were performed to characterize the samples prior to high-field experiments.

Protein samples used for data collection and chemical shift assignment (NTSD = 82 residues) contained 500 µL of 50-100 µM U-^{15}N-labeled PupB-NTSD, or 500 µM- 1 mM U-^{15}N/^{13}C-labeled PupB-NTSD that was dialyzed into 50 mM Na phosphate pH 6.0, 100 mM NaCl overnight at 4 °C. Prior to data collection, D_2O was added to a final concentration of 10% v/v. NMR data were recorded at 25 °C on Varian Inova 600 and 800 MHz spectrometers equipped with triple-resonance cryogenic probes. The following double and triple resonance experiments were performed: ^{1}H-^{15}N HSQC, ^{1}H-^{13}C HSQC (all carbon), ^{1}H-^{13}C HSQC (aliphatic carbon only), HNCO, HNCACB, CBCA(CO)NH, C(CO)NH, H(CCO)NH, HCCH TOCSY, HNHA, HBCB(CGCD)HD, HBCB(CGCDCE)HE, 2D ^{1}H-^{1}H NOESY, ^{1}H-^{15}N HSQC NOESY, and ^{1}H-^{13}C HSQC NOESY. Spectra were processed with NMRPipe (216).

3.2.3. NMR Chemical Shift Assignment

Backbone chemical shift assignments were initially performed via semi-automation with RunAbout in NMRViewJ (217,218), then transferred for continued manual assignment in CCPNmr (219). Sequential assignment of backbone chemical shifts were accomplished with through-bond experiments: 3D HNCO, HNHA, HNCACB, and CBCA(CO)NH (220,221). Side chain resonances were assigned from CBCA(CO)NH, H(CCO)NH, C(CO)NH, HCCH TOCSY, HBCB(CGCD)HD, and HBCB(CGCDCE)HE spectra (222). Aromatic side chain assignments were obtained from HBCB(CGCD)HD, HBCB(CGCDCE)HE, and 2D ^{1}H-^{1}H NOESY.

3.2.4. Structure Determination

Secondary structure prediction based upon H, N, CO, Cα, Cβ, and Hα chemical shifts relative to average characteristic chemical shift perturbations derived from known protein structures, was acquired from TALOS+ (223,224). This method also yielded ϕ and ψ torsion angle (dihedral) restraints, which were compared to the dihedral restraints procured from DANGLE in CCPNmr (225). ^{1}H-^{15}N HSQC NOESY were manually assigned to roughly 20% unambiguity, then fed into Ambiguous Restraints for Iterative Assignment (ARIA) for repeated structure calculation coupled with CNS for refinement (226-229). Additional initial dihedral restraints from DANGLE and the HNHA experiment were also used in structure calculation. The final refinement cycle parameters in CNS included energy minimization in explicit solvent (water) (230). Further ARIA runs incorporated chemical shifts from the ^{1}H-^{13}C HSQC NOESY experiment, and involved manual and automated NOESY assignments.

In addition to the traditional structure calculations approach via iterative NOESY assignment and refinement cycles, CS-Rosetta was utilized to obtain an auxiliary PupB-NTSD structure. Rosetta performs a Monte Carlo search through a specified number of conformations to find a set of minimal energy conformations for *de novo* structure prediction; it does not rely upon available structural templates, as in homology modeling (231-234). Instead, Rosetta samples backbone fragments (3-9 residues) from the PDB with similar ϕ and ψ angles, which are replaced in an extended peptide chain (the given protein sequence), and "folded" into a series of low energy conformations. CS-Rosetta combines the Rosetta framework with backbone chemical shifts for the fragment generation step (235).

3.2.5. Circular Dichroism Spectroscopy of PupB-NTSD and PupA-NTSD Mutants

The WT and mutant NTSDs were dialyzed overnight in10 mM potassium phosphate, pH 6.8, 100 mM $(NH_4)_2SO_4$, at 4°C. Continuous scanning CD spectra were measured at 4 °C between 180 and 250 nm using a Jasco J-815 spectrometer with a PFD-425S Peltier cell holder and a 1 mm quartz cell. The spectra were buffer subtracted, and the secondary structure content estimated using CONTIN and CDSSTR, incorporated in the CDPro suite (236).

3.2.6. Solution Characterization of PupB-NTSD by SEC-SAXS

SAXS data collection and analysis on a 6 mM sample of PupB-NTSD were determined as described in Chapter 4.2.10. The resulting models had an average χ^2 of 1.1 ± 0.002, with an averaged normalized spatial discrepancy (NSD) for the 10 DAMMIF calculations of 0.48 ± 0.01.

3.2.7. Structural Comparison of the PupA- and PupB-NTSDs

Structure solution of the PupB-NTSD by NMR and X-ray crystallography (as part of the PupB-NTSD:PupR-CCSSD complex, discussed in Chapter 4) allowed for distinct atomic-level variations of the static (X-ray crystal structure) and dynamic (NMR) protein state to be observed and contrasted with other available NTSD structures. All structural comparisons and RMSD calculations were performed in PyMOL (Schrödinger, LLC).

3.3. Results

3.3.1. Purification of the PupB-NTSD and PupA-NTSD

Constructs containing the gene fragments of 48-128 of PupB and 48-128 of PupA were successfully cloned into two different expression constructs for expression and purification of the individual NTSD, and for co-expression and pull-down analysis of the NTSDs with the PupR-CCSSD. Both NTSDs could be purified to homogeneity, as detected by SDS-PAGE (Figure 3.1, inset), and SEC yielded apparent MWs nearly identical to the theoretical MW of each NTSD (Figure 3.1).

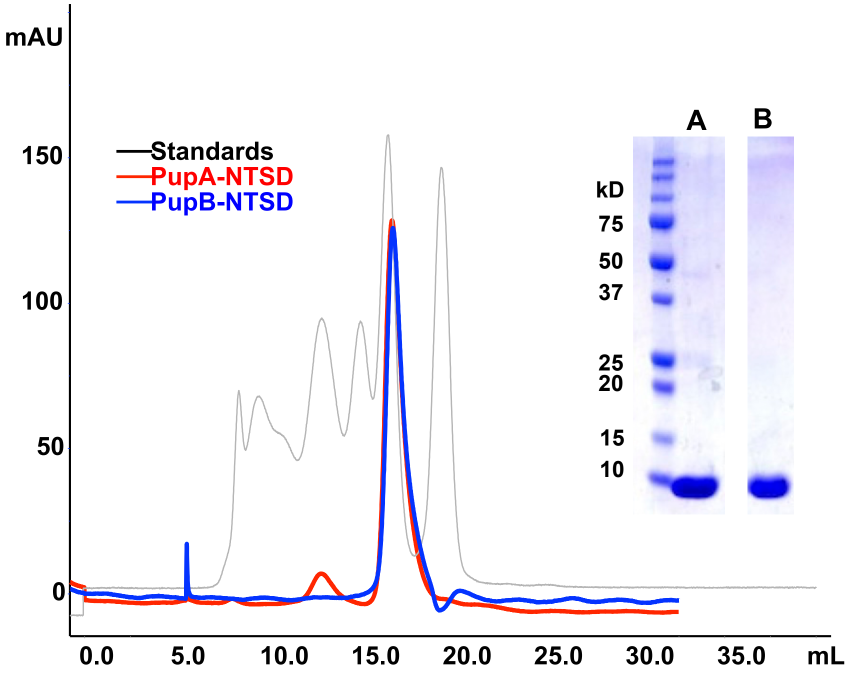


Figure 3.1. SEC and SDS-PAGE profiles of PupA-NTSD and PupB-NTSD. The SDS-PAGE indicates that both samples are homogeneous. The theoretical MW of the PupA-NTSD and PupB-NTSD are 8.6 kD and 8.2 kD, respectively. The apparent MW of the NTSDs are 8.5 kD and 8.4 kD.

3.3.2. Chemical Shift Assignments

Resonance assignments of the PupB-NTSD were obtained with standard strategies based on double- and triple-resonance NMR experiments. The protein was single and double uniformly labeled with ^{15}N and ^{15}N-^{13}C, respectively. Backbone resonances in the 2D ^{1}H-^{15}N HSQC spectrum (Figure 3.2), and C_α and C_β atoms were sequentially assigned to all non-proline residues based on correlations observed in HN(CA)CB and CBCA(CO)NH spectra. Backbone carbonyl resonance assignments were made from the HNCO experiment. H_α and H_β assignments were obtained from C(CO)NH and H(CCO)NH experiments. Initial backbone assignments were made using the automated RunAbout program in NMRviewJ, then checked manually in

CCPNmr. Remaining sidechain resonance assignments were made in CCPNmr through analysis of the 3D HCCH-TOCSY, ^{15}N-edited NOESY, ^{13}C-edited NOESY, and 2D ^{1}H-^{1}H NOESY experiments.

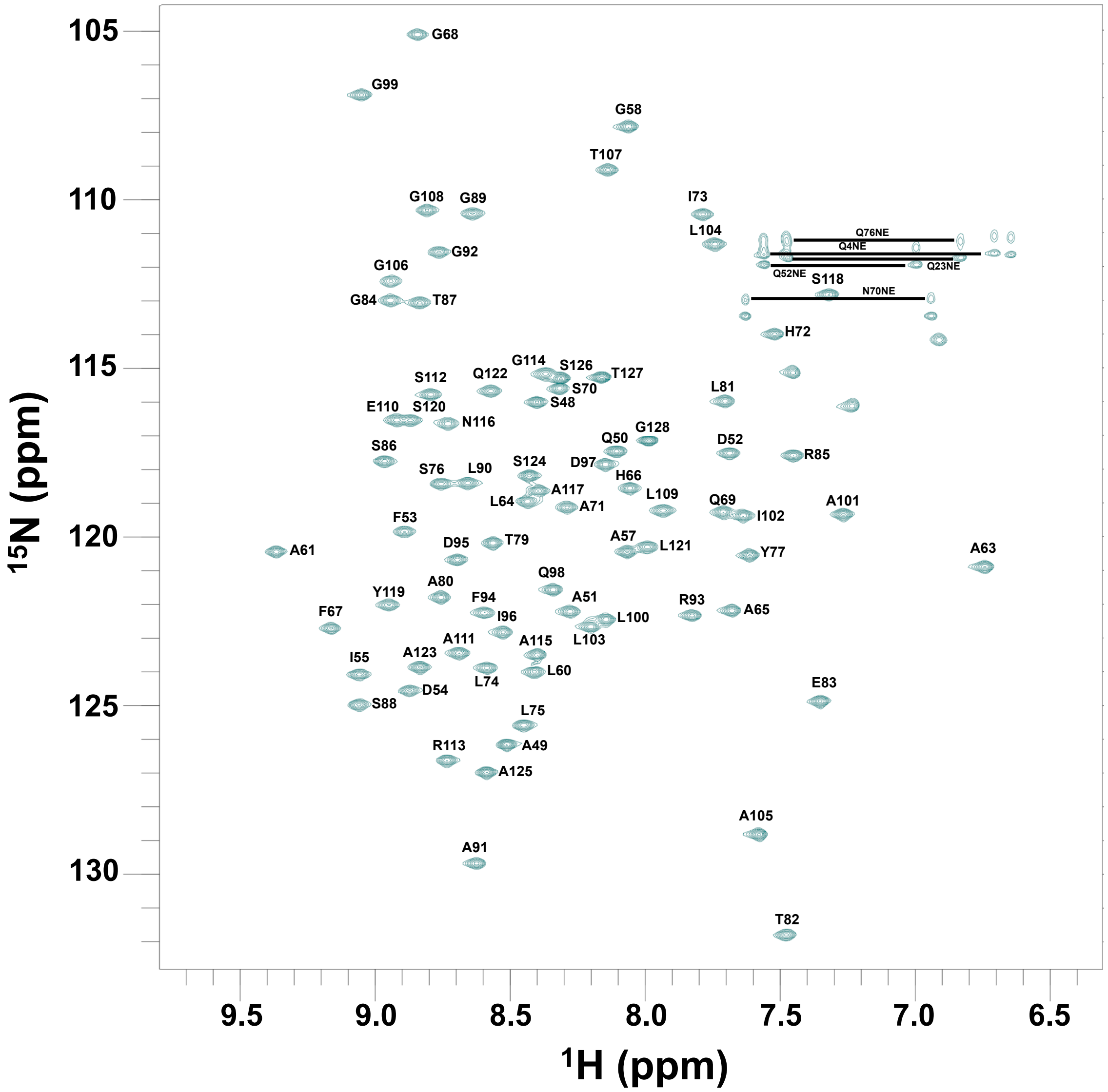


Figure 3.2. 2D ^{1}H-^{15}N HSQC spectrum for the PupB-NTSD, labeled with the residue assignments. The spectrum was acquired at 600 MHz on a 1 mM protein sample in 50 mM Na phosphate pH 6.0, 100 mM NaCl, 10% v/v D_2O at ambient temperature. Assigned cross-peaks for the NH_2 sidechains of Asn and Gln are connected by horizontal lines. The residue numbers correspond to that of the full-length PupB (including signal peptide). Peak separation observed in the spectrum indicated that the domain was well-folded.

All backbone atoms were assigned, except for proline amides, which lack an amide proton. The PupB-NTSD contains 4 proline residues, which created some difficulties in assigning chemical shifts. Overall, 97.3% of all hydrogens, 89.6% of all carbons, and 85.0% of all nitrogens were assigned. The remaining unassigned atoms are primarily those that are not typically assigned, such as Asp C_{Υ}, Arg guanidine groups, and some Tyr or Phe sidechain protons. These assignments have been deposited into the Biological Magnetic Resonance Bank (BMRB) (http://www.bmrb.wisc.edu) under accession number 27141. CheckShift (http://www.checkshift.services.came.sbg.ac.at) was utilized to assess the quality of chemical shift assignments, by comparison of the target chemical shifts to a set of reference chemical shifts (237,238). CheckShift output includes the target shifts with reference correction. Proposed chemical shift offset values were -0.02, +0.02, -0.29, +0.35, and 0 for N, C, C_{β}, C_{α}, and H, respectively, indicating that referencing was consistent and accurate, with respect to the reference chemical shifts.

3.3.3. Secondary Structure Conservation of the PupB-NTSD

Secondary structure prediction of the PupB-NTSD suggests that the domain comprises the conserved βαβ-repeat motif that is observed in other NTSDs (Figure 3.3).

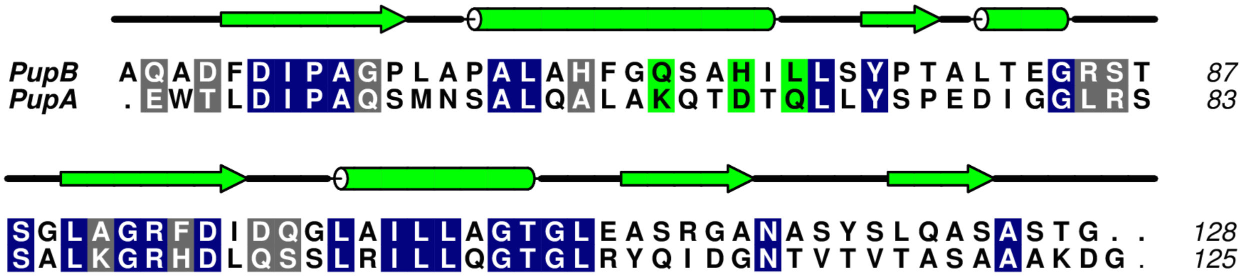


Figure 3.3. Secondary structure and sequence conservation of the PupB- and PupA-NTSDs. Residues highlighted in navy are invariant between the two domains. Residues in grey did not significantly alter interaction with the PupR-CCSSD, when mutated to the corresponding residue of the other NTSD. Residues in green partially or completely abrogated interaction of the NTSD with the PupR-CCSSD, when mutated to the corresponding residue of the PupA-NTSD.

Secondary structure elements for the PupB-NTSD were predicted using the Φ and ψ dihedral angles generated by TALOS based on chemical shift assignments (Figure 3.4) (223). TALOS classifies its dihedral angle constraints as "good," "bad," "ambiguous," "dynamic," or "unclassified," through comparison of acceptable Φ and ψ distributions to the input chemical shift values, and graphing the distributions onto a Ramachandran plot. Empirical prediction of the Φ/ψ distributions of the PupB-NTSD resulted in the following designations: 5 dynamic residues (1 on the N-terminus and 4 on the C-terminus), 17 ambiguous residues, and 58 good residues. The N- and C-terminal residues are unclassified. The 17 ambiguous residues primarily map to regions within the domain that are predicted to be disordered.

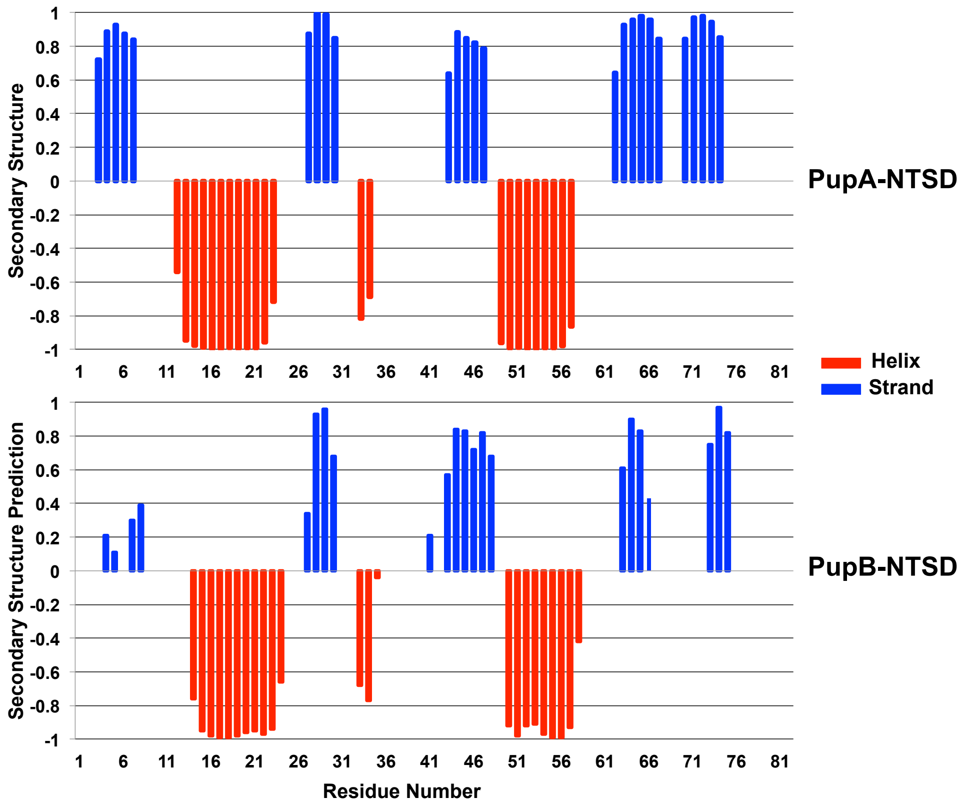


Figure 3.4. Comparison of secondary structure elements from the PupA-NTSD NMR structure (PDB 2A02) and the TALOS output for the PupB-NTSD. The blue and red bars indicate consensus β-strands and α-helices, respectively. Residue numbering is from the TALOS input, which requires that the sequence begins with number 1. To convert to the full-length NTSD sequence numbering, add 47 to the listed numbers.

3.3.4. Structure Solution of the PupB-NTSD

De novo structure solution of the PupB-NTSD was attempted through semi-automated NOESY cross-peak assignments implemented in the program ARIA (226). Dihedral angle restraints from DANGLE in CCPNmr and derived from the HNHA experiment were examined; 189 values were considered good and introduced as constraints into the structure calculations. A total of 607 unambiguous constraints from the NOESY experiments were used as initial input constraints. However, structure solution could not be completed. Incorporation of additional restraints, such as hydrogen bond constraints from deuterium exchange experiments and/or residual dipolar coupling analysis would facilitate structure solution.

To investigate structure solution by a more automated method, CS-ROSETTA was utilized for solution with the automated ROSETTA toolbox and the assigned chemical shifts. The CS-ROSETTA server trims the flexible N- and C-termini of protein structures during structure calculation. Five thousand structures were calculated, and the lowest energy structural ensemble extracted. Ten structures that converged to the lowest relative energy and also were structurally most similar as indicated by low RMSD values amongst them, were selected (Figure 3.5, 3.6A-C). RMSD values for the structured domain (residues 49-121) were 0.654 ± 0.221 Å.

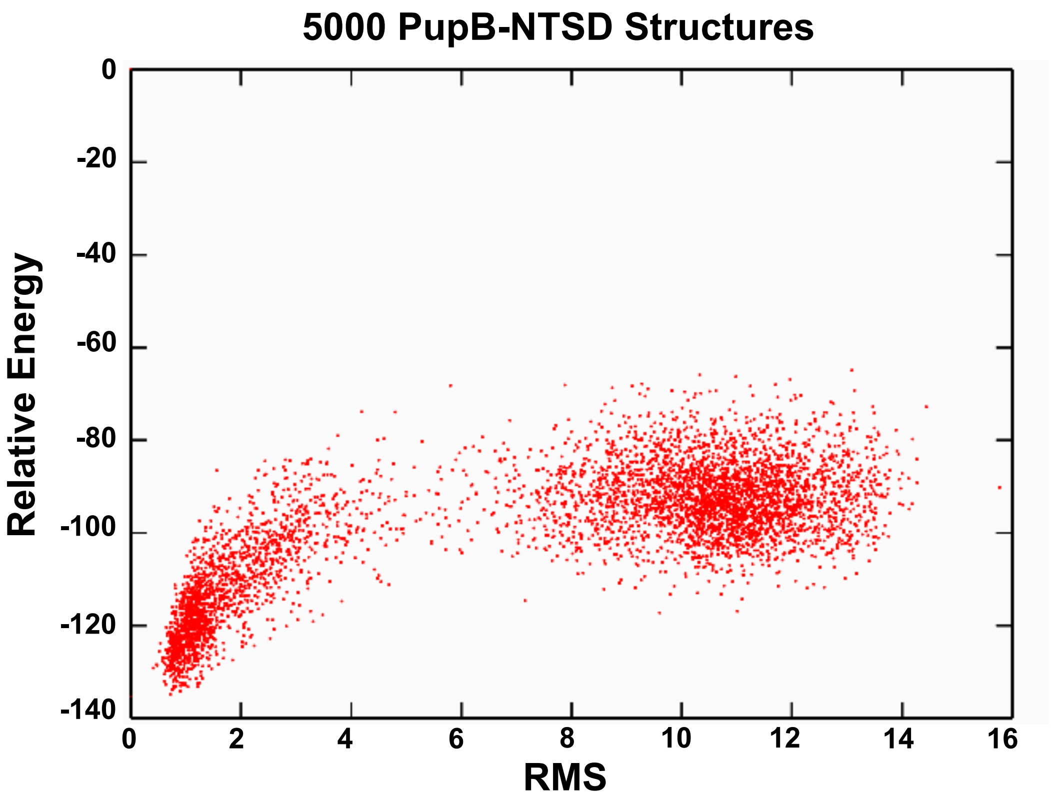


Figure 3.5. PupB-NTSD structures generated by CS-ROSETTA, with the relative energy of each structure plotted as a function of their RMS. Shown are the results from the creation of 5000 structures; results displayed in the lower left-hand corner of the plot are considered ideal.

Although CS-ROSETTA is considered a robust, *de novo* structure generator, it does rely on selection of protein fragments from the PDB based on similar chemical shifts. However, assignment of isotropic chemical shifts is generally based on empirical correlations derived from average chemical shifts of assignments deposited into the BMRB. The convergence of several low-energy, well-aligned structures demonstrates the validity of the final solution structures (Figure 3.6). In this case, additional data, such as the crystal structure, is available to assess the reliability of the final CS-ROSETTA structure. The PupB-NTSD crystal structure (discussed further in chapter 4) superimposes well on the lowest energy NMR structure, with an RMSD of 1.26 Å (Figure 3.6D).

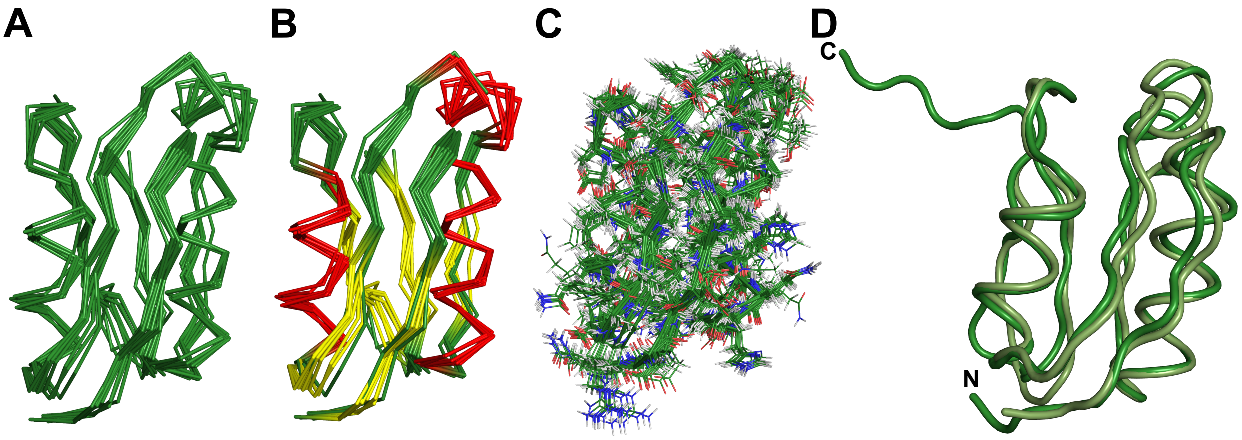


Figure 3.6. Representations of the 10 PupB-NTSD NMR conformers. (A) Ribbon diagram. (B) Ribbon diagram, with helices shown in red, and β-strands in yellow. (C) Stick representation. (D) Alignment of the lowest energy NMR structure with the PupB-NTSD crystal structure, with an RMSD of 1.26 Å.

3.3.5. SEC-SAXS Analysis of the PupB-NTSD

SEC-SAXS was performed to probe the low-resolution structure of the PupB-NTSD for comparison with the SEC-SAXS data of the PupR-CCSSD:PupB-NTSD (Chapter 4). The Guinier plot was linear in the $0 < q < 0.01$ Å, verifying the monodispersity of the sample. The radius of gyration (R_g), calculated from the Guinier region, was 13 Å, and the D_{max} was 36 Å, as determined by *P(r)* analysis (Figure 3.7A-B). Theoretical R_g and D_{max} calculated from the PupA-NTSD and FecA-NTSD NMR structures were 14 Å and 43 Å for the PupA-NTSD, and 14 Å and 45 Å for the FecA-NTSD, respectively. The lowest energy PupB-NTSD NMR and the PupB-NTSD crystal structure were fit with the experimental SAXS data; the R_g of the crystal structure was 12 Å, while the R_g of the NMR structure was 11 Å. The D_{max} values were 37 and 38 Å for the PupB-NTSD crystal structure and NMR structure, respectively. The Kratky plot indicated that the PupB-NTSD was reasonably well-folded (Figure 3.7C). Estimation of the molecular weight from the SAXS data was 8.2 kD, with the maximum q set to 0.35 Å^{-1}. The theoretical molecular mass of the PupB-NTSD is 8.1 kD; thus, the SAXS data confirmed that the PupB-

NTSD is monomeric in solution. The theoretical SAXS curve was calculated for the PupB-NTSD crystal structure, and fit to the experimental PupB-NTSD scattering profile (Figure 3.7D), yielding a χ value of 1.31. A low-resolution *ab initio* bead model was computed from the SAXS data, and fit with the X-ray crystal structure of PupB-NTSD, the NMR structure of PupA-NTSD, and the FecA-NTSD NMR structure (Figure 3.8).

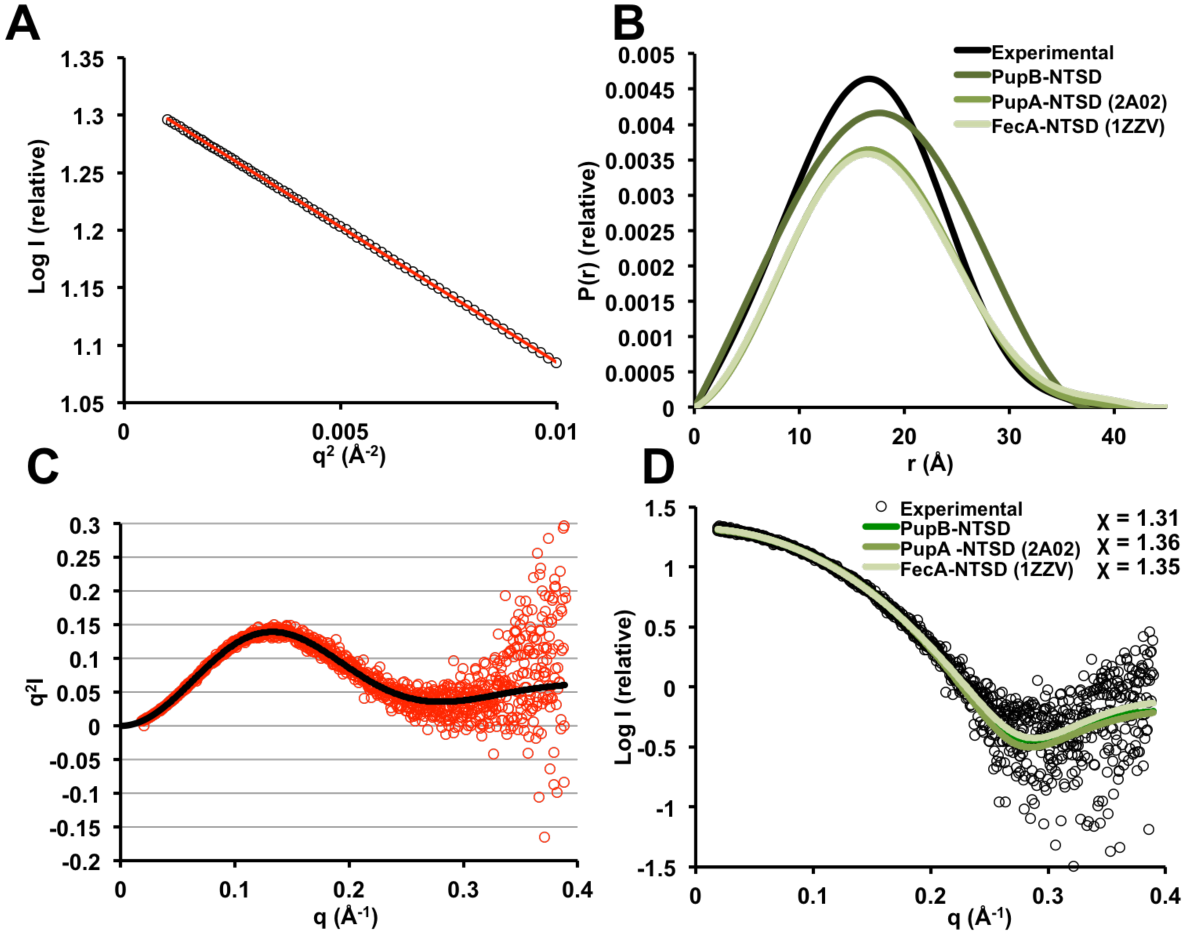


Figure 3.7. SEC-SAXS analysis of the PupB-NTSD. (A) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log(intensity). (B) Distance distribution *P(r)* for the experimental data (black), the theoretical curve calculated from the PupB-NTSD crystal structure (dark green), and the theoretical curves calculated from the PupA-NTSD and the FecA-NTSD (lighter shades of green). (C) Kratky plot calculated from the experimental scattering profile. (D) Experimental scattering profile, fit with the theoretical scattering profiles of the PupB-NTSD crystal structure, PupA-NTSD NMR structure, and FecA-NTSD NMR structure.

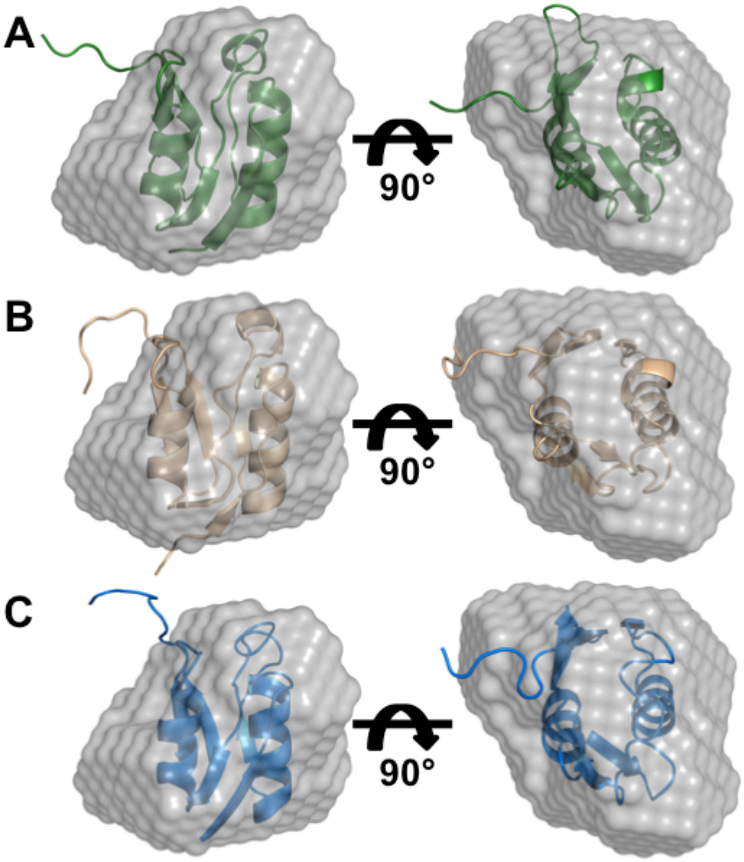


Figure 3.8. Low resolution *ab initio* bead model from the PupB-NTSD SAXS data, fit with the PupB-NTSD crystal structure, the PupA-NTSD, and the FecA-NTSD. The bead model derived from the SAXS data is displayed as a grey surface. (A) PupB-NTSD crystal structure (green). (B) PupA-NTSD NMR structure (tan). (C) FecA-NTSD NMR structure (blue). The correlation coefficients (CC) of the fit of the crystal or NMR structures to the molecular envelope calculated from the PupB-NTSD data is 0.9.

3.3.6. Assessment of Mutations on Structure and Function of the PupB-NTSD and the PupA-NTSD

To assess the ability of the PupB-NTSD and the PupA-NTSD to bind the PupR-CCSSD *in vitro*, pull-down assays were performed with GST-PupB-NTSD or GST-PupA-NTSD with MBP-PupR-CCSSD. The ability of the NTSDs to bind various truncations of the PupR-CCSSD was also analyzed (see Chapter 4). As expected, only the PupB-NTSD associated with the PupR-CCSSD (Figure 3.9).

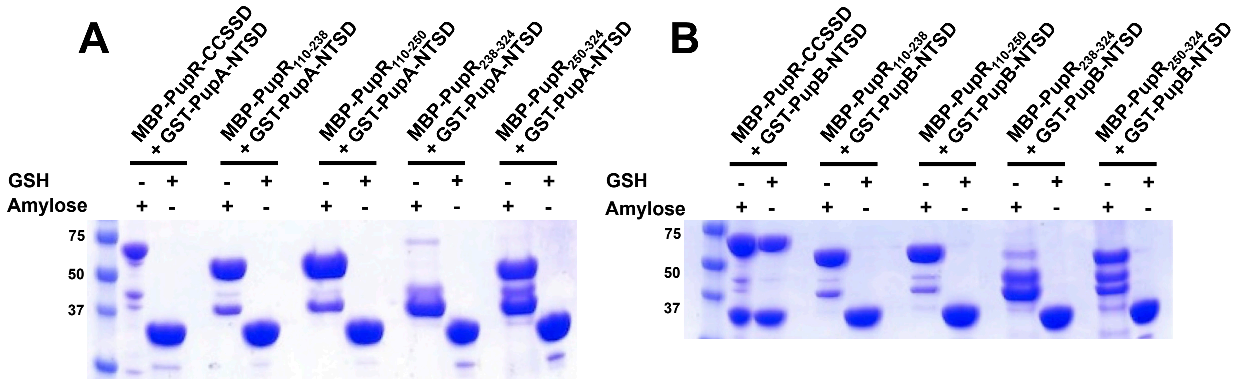


Figure 3.9. SDS-PAGEs of affinity pull-downs with GST-PupB-NTSD or GST-PupA-NTSD and MBP-tagged PupR-CCSSD truncations. PupR-CCSSD constructs used: full-length PupR-CCSSD (residues 110-324), residues 110-238, residues 110-250, residues 238-324, and residues 250-324. (A) Affinity pull-downs over either GSH or amylose resin indicate no interaction with PupA-NTSD; (B) interaction is observed between GST-PupB-NTSD and MBP-PupR-CCSSD (full-length domain).

To evaluate the association of various PupB-NTSD and PupA-NTSD mutants with the PupR-CCSSD, site-directed mutagenesis was performed to replace target residues of one NTSD with the corresponding residue of the other NTSD. Initially, residues identified for mutation were within or near the proposed NTSD:anti-sigma factor interaction interface, as determined by *in vivo* transcription initiation assays of the FecA-NTSD (81,105,138,151). Each single mutant was purified, as described for WT PupB-NTSD, and analyzed by CD spectroscopy for estimation of secondary structural elements (Figure 3.10 and Table 3.1). All mutants yielded spectra characteristic of folded protein; the helix, strand, and coil content of each mutant was estimated by CONTIN in CDpro (239). The only mutant identified by this method that abrogated binding of the PupB-NTSD to the PupR-CCSSD was H72D. However, the inverse mutation, PupA-NTSD D71H, did not permit PupA-NTSD:PupR-CCSSD complex formation. None of the PupA-

NTSD mutants analyzed, including several double, triple, or quadruple mutants, enabled association of the PupA-NTSD with the PupR-CCSSD.

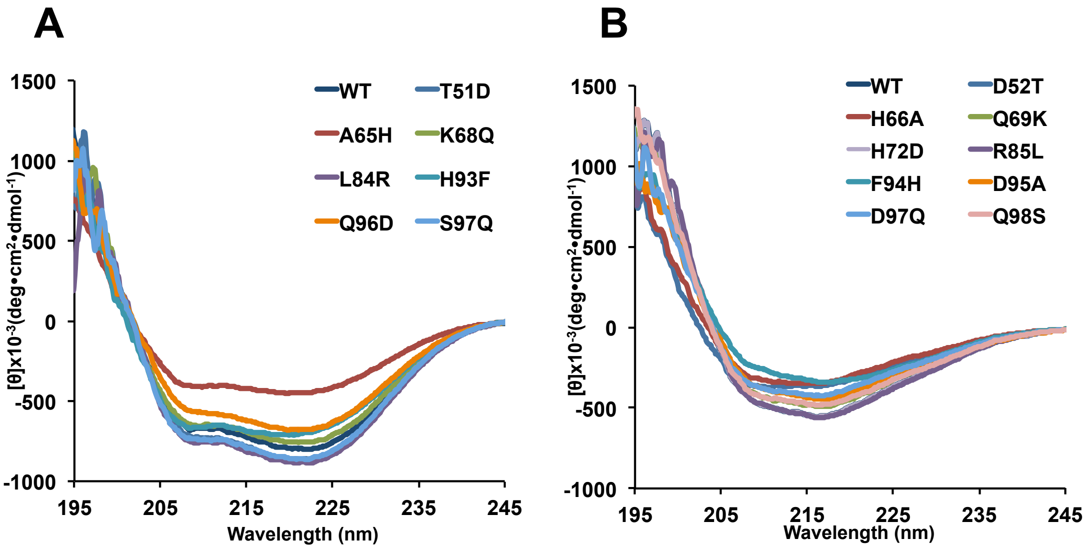


Figure 3.10. Circular dichroism spectra of WT and mutant PupA-NTSD, and WT and mutant PupB-NTSD. (A) PupA-NTSDs. (B) PupB-NTSDs. CD data was collected with 50 µM protein, with absorption measured from 195-245 nm.

Table 3.1. Estimated secondary structure content of PupA-NTSD vs. PupB-NTSD.

Sample	Helix	Strand	Coil	Total
PupA-NTSD NMR Structure	**0.300**	**0.210**	**0.490**	**1.000**
WT	0.307	0.176	0.517	1.000
T51D	0.285	0.202	0.513	1.000
A65H	0.185	0.298	0.517	1.000
K68Q	0.271	0.209	0.521	1.001
L84R	0.333	0.172	0.496	1.001
H93F	0.275	0.216	0.510	1.001
Q96D	0.264	0.232	0.504	1.000
S97Q	0.277	0.215	0.508	1.000
PupB-NTSD Crystal Structure	**0.280**	**0.280**	**0.439**	**0.999**
WT	0.115	0.361	0.524	1.000
D52T	0.172	0.311	0.518	1.001
H66A	0.141	0.340	0.518	0.999
Q69K	0.209	0.281	0.511	1.001
H72D	0.166	0.378	0.455	0.999
R85L	0.205	0.302	0.493	1.000
F94H	0.181	0.351	0.468	1.000
D95A	0.164	0.316	0.519	0.999
D97Q	0.194	0.305	0.502	1.001
Q98S	0.222	0.304	0.474	1.000

3.4. Summary and Conclusions

Both the PupB-NTSD and PupA-NTSD were expressed and purified to homogeneity, and the PupB-NTSD used for 2D and 3D NMR experiments for high-resolution structure solution. The solution structure of the PupB-NTSD was determined by assignment of the PupB-NTSD chemical shifts, although an alternative method for structure calculation was be employed for successful solution.

The PupB-NTSD is a small, globular domain that shares ~33% sequence identity and 39% sequence similarity with the PupA-NTSD. Additionally, as determined from these studies, the secondary and tertiary structural elements are conserved across NTSDs. Therefore, specific residues must mediate the interaction of the PupB-NTSD with the PupR-CCSSD that allows for signal transfer, and these distinct differences also render the PupA-NTSD signaling inactive through this pathway. Despite site-directed mutagenesis and isolation of multiple mutants of both domains, no single mutants were identified that granted PupA-NTSD:PupR-CCSSD complex formation. Intricate secondary structure rearrangements away from the interaction interface may account for further conformational changes that contribute to domain interaction. Further analysis is necessary to establish binding specificity, and thus, signal transfer.

4. STRUCTURAL BASIS FOR CELL-SURFACE SIGNALING THROUGH THE INNER MEMBRANE SIGMA REGULATOR PROTEIN, PUPR

4.1. Introduction

Structural evidence for the interaction between an NTSD and sigma regulator CCSSD has thus far been limited to proposed interaction sites mapped onto NTSD structures, and NMR titrations involving the HasR-NTSD and the HasS-CCSSD. These latter experiments suggested that the HasS-CCSSD is structurally unstable, as the domain had to be isolated from inclusion bodies followed by denaturation and renaturation with buffer containing 0.05% w/v Zwittergent 3-14 (136). Multiple residues throughout the HasS-CCSSD demonstrated chemical shifts upon interaction with the HasR-NTSD. Since two regions of the NTSD were found to interact with the CCSSD, a “wrapping mode” was proposed in which the CCSSD in its partially disordered state envelops the NTSD. Additionally, the region of the CCSSD identified as the H2 helix (T_{257}-$QALAAQLNRYR_{267}$) was shown to interact with detergent micelles, which the authors postulate mimic the periplasmic space:inner membrane interface environment and suggest that this helix contacts the inner membrane (136). An X-ray crystal structure of a putative anti-sigma CCSSD from *Parabacteroides distasonis* ATCC 8503 (PDB 4M0N and 4M0H; unpublished), which shares secondary structure similarities with that predicted for the TBDT sigma regulator CCSSDs, shows a well-folded structure containing two subdomains. These two subdomains appear to stabilize each other through extensive interactions between the N-terminal and C-terminal regions of the CCSSD.

Initial cleavage of the sigma regulator CTD has also been proposed to include an autoproteolytic event via N-O acyl rearrangement through the conserved residues G191 and T192 of FoxR from *P. aeruginosa* (161,240,241). However, these residues are not conserved in many sigma regulators, so the self-cleavage mechanism may not be required for complete proteolysis and transcriptional activation. Analogous sigma regulators FoxR and FiuR of *P. aeruginosa* are anti-sigma factors also possessing “pro-sigma” activity, although FpvR of the same organism is a purely anti-sigma regulator. This clarification in function illustrates the findings that FpvR only serves to repress FpvI by tethering it to the inner membrane, whereas the ASD of the anti-sigma factor remains in complex with the sigma factor, and the complex is necessary for full functionality of the sigma factor (152,160,242).

To structurally assess the CCSSD:NTSD interface, the X-ray crystal structure of the periplasmic domain of PupR in complex with the PupB-NTSD was determined to 1.6 Å, which represents the first such cell-surface signaling NTSD:CCSSD structure. To distinguish the periplasmic domain of the regulatory protein from other periplasmic regulatory proteins, the CTD has here been renamed to the C-terminal cell-surface signaling domain (CCSSD). The full-length PupR-CCSSD is composed of two subdomains; the N-terminal 128 residues form a 14-stranded mixed-parallel/anti-parallel β-sheet, with a novel β-solenoid-like fold, which has been designated the C-terminal juxtamembrane (CJM) subdomain. A 12-residue flexible linker connects the CJM to the remaining 74 residues of the PupR-CCSSD. Interestingly, these final residues form a subdomain of the same βαβ-repeat motif (STN domain) similar to that found in the PupB-NTSD, which is conserved across TonB-dependent transporter protein signaling domains of known structure. Additionally, the two subdomains of the CCSSD appear to stabilize each other, such that the full-length CCSSD is required for interaction with the NTSD, as demonstrated by affinity pull-downs with the PupB-NTSD and various truncations of the PupR-CCSSD. Analysis of the solution states by small-angle X-ray scattering of the PupR-CCSSD alone and in complex with PupB-NTSD suggests that the PupR-CCSSD possesses a significant degree of flexibility and instability, which is reduced when the domain is associated with the NTSD.

4.2. Materials and Methods

4.2.1. Cloning of PupR-CCSSD Constructs

Secondary and tertiary structure predictions of PupR were used to estimate domain boundaries of the proposed periplasmic region of the protein. Five expression constructs were developed, containing PupR residues 110-324, 110-238, 110-250, 238-324, or 250-324. Each region was amplified by PCR, using pUC19-PupR as a template. Each region was separately cloned between NcoI and XhoI sites of the pMBP-Parallel1 vector (243).

4.2.2. Protein Expression and Purification of FL PupR-CCSSD

Chemically competent *E. coli* C41(DE3) cells (Lucigen) were transformed by this pMBP-Parallel1-PupR-CCSSD (PupR residues 110-324) plasmid by standard protocol. Transformed cells were grown at 37°C in LB medium supplemented with 100 μg/mL ampicillin to an optical density $OD_{600\ nm}$ of 0.7-0.9, and PupR-CCSSD expression was induced with 0.5 mM IPTG. After induction, the culture temperature was maintained at 20°C for 20 h, and cells were harvested by centrifugation at 4000*g* for 30 min. Harvested cells were stored at -80 °C until use, to assist cell lysis during purification via a freeze-thaw cycle. For protein purification, 30 g of cells were suspended in lysis buffer (25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol, 2 mM DTT) and then lysed with a Nano DeBEE homogenizer, with 2-3 passes at 15000 psi. The crude extract was clarified by centrifugation at 20000*g* for 45 min.

PupR-CCSSD was purified by amylose affinity chromatography coupled with on-column cleavage by recombinant TEV protease. The supernatant after lysis clarification was loaded onto an amylose gravity column and washed with lysis buffer. TEV protease was added to the column in an approximate 1:10 mass ratio, based on the protein binding capacity of the amylose resin, and incubated for 16 h at 4°C to remove the MBP tag. TEV proteolysis yielded a 219-residue product, containing PupR residues 110-324, with an N-terminal 4-residue cloning artifact (GAMG). The column was washed with lysis buffer to elute the cleaved PupR-CCSSD. Contamination by MBP necessitated a second pass of eluted PupR-CCSSD over equilibrated amylose resin. The PupR-CCSSD protein was pooled, concentrated with a Millipore centrifugal filter unit (MWCO = 10000 kD), and loaded onto a 16/60 Superdex 200 (GE Lifesciences) size exclusion column. SEC was performed with 25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol at a flow rate of 0.4 mL/min. Fractions containing PupR-CCSSD were pooled and concentrated to 10 mg/mL and stored at 4°C or flash-frozen as pellets in LN_2 and stored at -80°C until use. Protein yield was approximately 4 mg per 1 L of culture.

4.2.3. Determination of Protein Purity and Concentration

The purity of the collected protein was estimated to be ~80% by SDS-PAGE stained with Coomassie Blue (215). Contamination by MBP was observed in the final protein samples. Protein concentration was determined by absorbance at 280 nm using the molar extinction coefficient $\varepsilon_{280} = 29450\ M^{-1}\ cm^{-1}$ and a theoretical molecular weight of 24067 g/mol. Incorporation of SeMet and MW of WT protein was confirmed by electrospray ionization mass spectrometry (ESI-MS).

4.2.4. Preparation of Selenomethionine-derivatized PupR-CCSSD

Selenomethionine (SeMet; Acros Organics) derivatized PupR-CCSSD was prepared following a modified protocol involving methionine synthesis suppression (244,245). *E. coli* C41(DE3) cells transformed with pMBP-Parallel1-PupR-CCSSD were grown at 37°C to saturation in 3 mL LB medium with 100 μg/mL ampicillin, then transferred to pre-warmed M9 minimal media containing 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% w/v glucose, and 100 μg/mL ampicillin and incubated at 37°C. Once the $OD_{600\ nm}$ reached 1.0, the media was supplemented with SeMet, Lys, Thr, Phe, Leu, Ile, and Val, and the temperature lowered to 20°C. IPTG was added to 0.5 mM, and the cultures grown for an additional 18 h. Purification of SeMet PupR-CCSSD was performed as described for the native protein. Approximately 2 mg of SeMet protein was obtained per 1 L of cells.

4.2.5. Site-Directed Mutagenesis of PupR-CCSSD and PupB-NTSD

Point mutations of WT PupR-CCSSD or WT PupB-NTSD were created using the QuikChange II protocol (Stratagene), with the pMBP-Parallel1-PupR-CCSSD vector or the pGEXr-PupB-NTSD vector as a template, and primers containing the desired mutation, using the most common codons. The mutant gene sequences were verified by DNA sequencing (performed by MCLab), and the plasmids used to transform *E. coli* BL21(DE3)pLysS competent cells for protein expression. The mutant proteins were utilized for co-expression and pull-down assays.

4.2.6. Co-expression and Pull-down Assays of PupR-CCSSD:PupB-NTSD Complexes

E. coli BL21(DE3)pLysS cells were co-transformed with pMBP-Parallel1-PupR-CCSSD and pET41-GST-PupB-NTSD by standard protocol. Co-transformants were selected with LB medium containing 100 μg/mL ampicillin and 15 μg/mL kanamycin. MBP-PupR-CCSSD and GST-PupB-NTSD were co-expressed following the same procedure as for the individually expressed proteins. Harvested cells were lysed in 25 mM HEPES pH 7.5, 400 mM LiCl, 10% glycerol via emulsification. Lysate was clarified by centrifugation at 20000 *g* for 45 min, and the resulting supernatant divided into two equal aliquots, and combined with either 5 mL of amylose resin, or 5 mL glutathione sepharose (GSH) resin. The gravity-flow columns were gently rocked for 30 min at 4°C to facilitate protein binding. Each column was washed with 10 CV of lysis buffer, then eluted with 5 CV of elution buffer (for amylose affinity chromatography: lysis buffer + 20 mM maltose; for GSH affinity chromatography: lysis buffer + 15 mM glutathione). Total protein content of each 5 mL elution fraction was determined by Bradford assay, and 20 μg of total protein loaded onto a 4-20% TGX SDS-PAGE (Bio-rad). Gels were stained with Coomassie Blue and qualitatively analyzed for protein association. This protocol was repeated for all pull-down analyses.

4.2.7. Western Blot Analysis of MBP-PupR-CCSSD:GST-PupB-NTSD Complexes

Twenty μg of total protein from pull-down assay elution fractions were resolved on SDS-PAGE and transferred to nitrocellulose membranes. Membranes were incubated with either Anti-MBP-HRP (NEB) or Anti-GST-HRP (GE Lifesciences) antibodies at a 1:10000 dilution in TBS-T with 5% w/v milk powder. MBP- or GST-tagged proteins were detected using ECL Western Blotting Substrate (Pierce) and documented with a Storm 865 Imaging Station (GE Lifesciences).

4.2.8. Binding Affinity Determination of PupR-CCSSD with PupB-NTSD by Isothermal Titration Calorimetry

ITC was performed with a Low Volume Nano ITC (TA Instruments). Both proteins were loaded into separate dialysis cassettes, and co-dialyzed into 25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol. All experiments were performed at 15°C, with 25 injections of 2 μL each, with stirring at 250 rpm. PupB-NTSD, concentrated to 2 mM, was titrated into PupR-CCSSD, with a concentration of 200 μM. Titrations were repeated in triplicate, with the values from a buffer-into-buffer titration subtracted from the values derived from the protein-to-protein titration. Data were analyzed with either NanoAnalyze (TA Instruments) with an independent, single-site model, or NITPIC (246) for data integration, followed by data processing with SEDPHAT (247-250) and plotting of isotherms in GUSSI (248). Processing in SEDPHAT included data refinement incorporating the local incompetent fraction as a function of the concentration compensation factor (251).

4.2.9. NMR Titration Experiments of PupR-CCSSD into ^{15}N PupB-NTSD

NMR titration experiments were used to probe the structural dynamics of the PupB-NTSD upon binding to the PupR-CCSSD. ^{15}N isotopically labeled PupB-NTSD was purified as described in chapter 3.2.1. 2D ^{1}H-^{15}N HSQC spectra were collected on a Bruker 400 MHz instrument (in-house), as well as an 800 MHz instrument with a room temperature probe (UTSW). Due to the instability of the PupR-CCSSD, low protein concentration was necessary in order to explore a broader titration range. ^{15}N-labeled PupB-NTSD at 250 µM, 100 µM, 80 µM, and 55 µM, was combined with 0 µM, 250 µM, 50 µM, and 75 µM PupR-CCSSD, respectively, for molar ratios of 1:0, 1:2.5, 1:0.625, and 1:1.36 PupB-NTSD:PupR-CCSSD in 25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol, 10% v/v D_2O. A second set of experiments on the same instrument and in the same buffer conditions were tested with PupB-NTSD to PupR-CCSSD ratios of 1:0, 1:1, 1:2, and 1:3. Data were processed with NMRPipe, and shift intensities differences assessed in NMRviewJ and CCPNmr.

4.2.10. Circular Dichroism Spectroscopy and Thermal Denaturation of PupB-NTSD, PupR-CCSSD, and PupR-CCSSD:PupB-NTSD

The PupB-NTSD was diluted to 50 µM (0.204 mg/mL) in 10 mM potassium phosphate, pH 6.8, 100 mM $(NH_4)_2SO_4$, after dialysis in the same buffer overnight at 4°C. Continuous scanning CD spectra were measured at 4 °C between 180 and 250 nm using a Jasco J-815 spectrometer with a PFD-425S Peltier cell holder and a 1 mm quartz cell. The spectra were buffer subtracted, and the secondary structure content estimated using CONTIN and CDSSTR, provided in CDPro (236). This procedure was repeated for the PupR-CCSSD alone, and the PupR-CCSSD:PupB-NTSD complex, generated by combining 50 μM each of purified PupR-CCSSD and PupB-NTSD (for a final volume of 300 μL), and incubating on ice for 30 min.

CD melting and re-folding curves were recorded at 216 nm (based on the observed minima during CD scanning) with 50 μM PupB-NTSD, 50 μM PupB-NTSD, or 50 μM each of PupR-CCSSD:PupB-NTSD protein by increasing the temperature from 10 to 85 °C in 1 °C increments with a slope of 1 °C/min. Protein unfolding was monitored during both heating and cooling. Complete CD spectra at 4 °C were recorded just prior to each experiment. Melting temperatures were determined by fitting a standard Boltzmann sigmoidal curve to the molar ellipticity in Origin 8 (OriginLab Corp., Northampton, MA). The final melting temperature was defined as the inflection point after fitting.

4.2.11. Crystallization and Data Collection of PupR-CCSSD

PupR-CCSSD in SEC buffer was concentrated to ~300 μM using a Millipore centrifugal filter unit (MWCO = 10k). Screening for potential crystallization conditions was performed with the MCSG crystallization suite (Anatrace). Hits were identified in 100 mM MES:NaOH pH 6.5, 400-700 mM NaCl, 18-23 % PEG 4000. Optimization via manual grid screening was carried out holding the MES:NaOH pH 6.5 buffer constant, while varying concentrations of NaCl and PEG 4000. The sitting drop vapor diffusion method was employed using a 1:1 v/v mixture of protein solution to reservoir solution (1 μL each). Protein crystals were grown at 20°C and appeared to mature after 5 days (Figure 4.1). Crystals in a single well were stained with Izit Crystal Dye (Hampton Research), and crushed with a cat whisker to verify that they were protein, and not salt, crystals. Crystals that grew to at least 100 μm in length were cryoprotected with MiTeGen LV CryoOil and flash-cooled in liquid nitrogen.

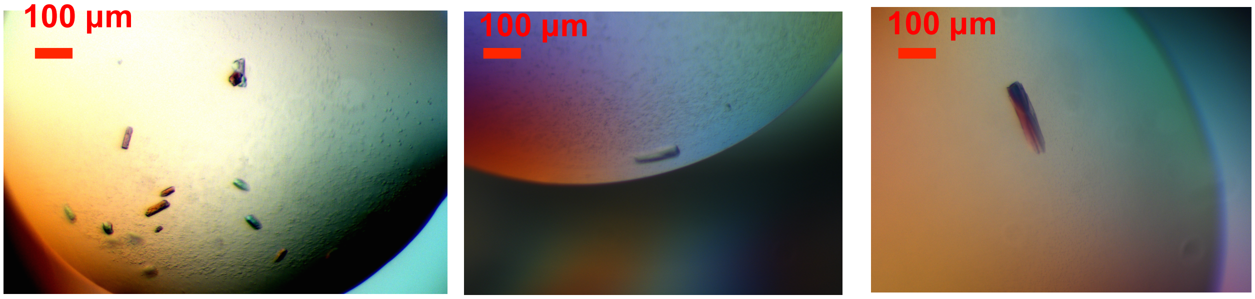


Figure 4.1. Single crystals and crystal clusters of native PupR-CCSSD.

Crystals were screened for diffraction at NE-CAT 24-ID-E at the APS. One crystal out of ~20 crystals screened diffracted to 2.6 Å. The Labelit autoindexing function of RAPD (https://rapd.nec.aps.anl.gov/rapd) initially suggested the space group *C2*, with mosaicity at 1.5°. BEST, the data collection strategy program within RAPD, recommended the strategy outlined in Table 4.1. One data set based on this collection strategy was gathered; however, reflection intensities were very weak and the crystal succumbed to radiation damage before a complete data could be collected (Table 4.2). The data was processed with XDS through RAPD, and separately with autoPROC. Unfortunately, the number of unique reflections was too low to be used for structure solution (Table 4.2).

Table 4.1. X-ray data collection strategy for PupR-CCSSD, as indicated by BEST.

Omega Start	**Omega End**	**Range of rotation**	**Number of images**	**Delta Omega**	**Exposure Time (sec)**	**Detector Distance (mm)**	**% Beam Transmission**
16.0	166.0	150.0	300	0.5	1.0	500	7.0

Table 4.2. X-ray data collection statistics for the PupR-CCSSD.

	PupR-CCSSD
Data collection	
Beamline	24-ID-E
Wavelength (Å)	0.9792
Space group	*I222*
Mosaicity (deg)	0.47
Unit-cell parameters (Å, deg)	51.8, 53.8, 69.2 α,β,γ = 90
Resolution range (Å)	41.5-3.81 (4.26-3.81)
Unique observations	2423 (622)
Multiplicity	3.0 (2.9)
Completeness (%)	80.1 (77.1)
CC(1/2)	0.971 (0.793)
R_{merge} (%)	31.5 (103.2)
Mean $I/\sigma I$	2.7 (1.4)
Data processing program	autoPROC

4.2.12. Crystallization, Data Collection, and Structure Solution of the PupR-CCSSD:PupB-NTSD Complex

Purified PupR-CCSSD (native or SeMet) and PupB-NTSD were combined in a 1:1 molar ratio and incubated on ice for approximately 30 min. The MCSG crystallization suite (Anatrace) was used for crystallization condition screening. A number of conditions with crystal hits were identified (Table 4.3). Attempts at manually optimizing crystallization conditions via sitting drop vapor diffusion resulted in reproducible crystals in 200 mM Na tartrate or Na K tartrate, 20-25% PEG 3350 (Figure 4.2). Single crystals were cryoprotected with MiTeGen CryoOil and flash-frozen in liquid nitrogen.

Table 4.3. Favorable crystallization conditions for PupR-CCSSD:PupB-NTSD identified from various MCSG protein crystallization screens.

Screen ID	Well	Salt	Buffer and pH	Precipitant	Protein concentration (each)	Ratio (protein:reservoir)
MCSG-1	A8	None	100 mM Tris:HCl pH 7.0	20% PEG 2k MME	200	3:1, 1:1, 1:3
MCSG-1	D12	None	200 mM NH_4Cl pH 6.3	20% PEG 3350	200	1:1
MCSG-1	B12	None	100 mM Bis-Tris:HCl pH 6.5	28% PEG 2k MME	400	1:1, 1:3
MCSG-1	E10	200 mM NH_4 tartrate dibasic	None	20% PEG 3350	400	3:1, 1:1, 1:3
MCSG-1	C12	None	100 mM Bis-Tris:HCl pH 6.5	25% PEG 3350	400	3:1, 1:1, 1:3
MCSG-1	E6	200 mM K_2SO_4	None	20% PEG 3350	400	1:3
MCSG-1	G4	200mM K Na tartrate	None	20% PEG 3350	200, 400	3:1, 1:1, 1:3
MCSG-2	A2	150 mM KBr	None	30% PEG 2k MME	200, 400	1:3
MCSG-2	C5	200 mM Na citrate	None	20% PEG 3350	200	1:1
MCSG-2	F11	200 mM Na tartrate dibasic	None	20% PEG 3350	400	3:1, 1:1, 1:3
MCSG-2	G6	200 mM NaCl 1 M $(NH_4)_2SO_4$	100 mM Na citrate:citric acid pH 5.5	None	200	1:1
MCSG-2	F12	200 mM $LiNO_3$	None	20% PEG 3350	400	1:1
MCSG-2	H10	200 mM NaCl	None	20% PEG 3350	200, 400	3:1
MCSG-3	E8	None	100 mM HEPES:NaOH pH 7.5	10% PEG 8k, 10% ethylene glycol	400	1:1
MCSG-3	H1	Other: 200 mM proline	100 mM HEPES:NaOH pH 7.5	10% PEG 3350	400	1:1
MCSG-4	E7	None	100 mM HEPES:NaOH pH 7.5	25% PEG 1k	200	1:3
MCSG-4	F1	200 mM Na acetate	100 mM HEPES:NaOH pH 7.5	20% PEG 3k	200	1:1
MCSG-4	F5	200 mM $LiSO_4$	100 mM Tris:HCl pH 8.5	20% PEG 4k	400	1:1
MCSG-4	H2	100 mM Na acetate	None	25% PEG 4k, 8% isopropanol	400	1:1
MCSG-4	H6	10 mM Na citrate	None	33% PEG 6k	400	1:1
MCSG-4	F8	None	100 mM MES:NaOH pH 7.5	30% PEG 4k	200	3:1, 1:1, 1:3
MCSG-4	F6	100 mM Na acetate	100 mM HEPES:NaOH pH 7.5	22% PEG 4k	200	1:1
MCSG-4	G9	100 mM Na acetate	None	20% PEG 4k, 5% isopropanol	200	1:1
MCSG-4	H5	100 mM LiCl	100 mM HEPES:NaOH pH 7.5	25% PEG 6k	200	1:1

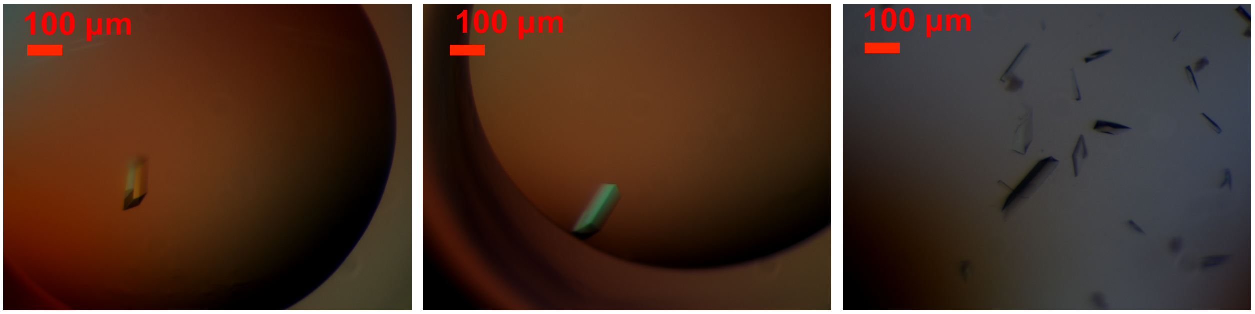


Figure 4.2. Single crystals and crystal clusters of native or SeMet PupR-CCSSD:PupB-NTSD.

Diffraction data were collected at NE-CAT beamlines 24-ID-E and 24-ID-C at the APS under cryogenic conditions (~100 K). The data set collected from the native PupR-CCSSD:PupB-NTSD crystals was processed using autoPROC (252) with components MOSFLM (253) for indexing, XDS (254) and SCALA (255) for processing and scaling, and POINTLESS (256) for space group determination. Initially, structure solution was attempted by molecular replacement (MR) with PHASER MR in PHENIX, using either the PupA-NTSD or a putative anti-sigma factor from *Parabacteroides distasonis* (PDB ID 4M0N, Joint Center for Structural Genomics, unpublished), which was predicted to have the same secondary structural elements as the PupR-CCSSD, although it shares only 22% sequence identity with the PupR-CCSSD. MR using either structure as a starting model failed, as well as structures truncated to either Cβ or Cα. Thus, experimental phases were required for structure solution.

Diffraction data from a single, orthorhombic crystal of SeMet PupR-CCSSD:PupB-NTSD was processed with HKL2000 (257), and the structure determined by single-wavelength anomalous dispersion (SAD) phasing using intensities measured at 12.66 keV (0.9792 Å), with a high resolution limit of 1.5 Å. Three of the four selenium sites per PupR-CCSSD monomer (three in the PupR-CCSSD sequence, and one from the N-terminal cloning artifacts) were located, and initial phasing performed with AutoSol in PHENIX (258). No methionine residues are present in the PupB-NTSD. These phases were used to calculate initial electron density maps for automated model building of the PupR-CCSSD with AutoBuild (259). A polyalanine chain was fit into excess electron density, followed by *in silico* mutation of residues with strong $2F_o$-F_c electron density. Refinement was carried out in PHENIX (260) and iterative manual modeling building in COOT (261). Ten percent of the data were excluded from refinement and used as the test set to calculate R_{free} (PHENIX Refine default). The SeMet PupR-CCSSD structure was subsequently used as a molecular replacement (MR) model to phase a 1.76 Å data set of native PupR-CCSSD:PupB-NTSD. The original data set was collected to 1.60 Å; however, a lower resolution cut-off was required during data processing to improve the average $I/\sigma I$ and CC(1/2) to within an acceptable range. MR was performed with Phaser-MR (262), followed by AutoBuild (259) and refinement in PHENIX. Automated TLS group determination (263) and individual atomic B-factors were used during refinement for both the SeMet and native structures. Excess density was observed in two regions around the PupR-CCSSD; two tartrate molecules were positioned within these regions, based on the contents of the reservoir solution. Water oxygen atoms were positioned, with subsequent visual verification. Data collection and refinement statistics are summarized in Table 4. All figures were prepared using PyMOL v.1.5.0.4 (Schrödinger, LLC). Analysis of surface areas, protein interfaces, assemblies, and interactions were determined using the PISA server (http://www.ebi.ac.uk/pdbe/pisa) (264). RMSD comparisons were carried out in PyMOL. Model validation was performed using MolProbity (265) and the PDB Validation Server (https://validate-rcsb-1.wwpdb.org/). Atomic models and structure factors have been deposited into the Protein Data Bank (http://www.rcsb.org) under PDB entries 5URL and 5UU0. The DaliLite v.3 webserver (http://ekhidna.biocenter.helsinki.fi/dali_server/start) was utilized to identify structural homologs available in the Protein Data Bank (266).

Table 4.4. X-ray data collection, phasing, and refinement statistics for the PupR-CCSSD:PupB-NTSD complex.[a]

	Native	**SeMet derivative**
Beamline	24-ID-E	24-ID-C
Wavelength (Å)	0.9792	0.9792
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å, deg)	43.4, 44.6, 141.0 $\alpha,\beta,\gamma = 90$	43.6, 44.7, 141.3 $\alpha,\beta,\gamma = 90$
Resolution range (Å)	42.5-1.76 (1.767-1.761)	141.34-1.51(1.53-1.51)
Total observations	190024 (1895)	258089 (2816)
Unique observations	27078 (2741)	43910 (1511)
Multiplicity	7.0 (6.9)	5.9 (1.9)
Completeness (%)	96.9 (99.3)	98.0 (70.2)
CC(1/2)	0.999 (0.966)	0.999 (0.765)
R_{merge}[b] (%)	5.6 (26.8)	5.2 (40.1)
R_{merge} (anom, %)	--	4.5 (42.9)
Mean $I/\sigma I$	25.2 (6.8)	18.9 (1.6)
Data processing program	autoPROC	HKL2000
Refinement		
Refinement program	PHENIX	PHENIX
Resolution range (Å)	42.5-1.76 (1.82-1.76)	42.6-1.60 (1.614-1.558)
Molecules per asymmetric unit	2	2
R_{work} (%)	14.9	15.3
R_{free} (%)	20.0	18.4
RMSD stereochemistry		
Bond lengths (Å)	0.014	0.018
Bond angles (deg)	1.48	1.96
No. of atoms	2661	2787
PupR-CCSSD:PupB-NTSD	2325	2412
Ligands (tartrate)	20	20
Waters	316	355
Total average B (Å^2)	21.7	18.1
PupR-CCSSD	19.5	16.1
PupB-NTSD	23.9	18.6
Tartrate	25.3	20.0
Waters	29.7	27.4
Ramachandran plot (%)		
Preferred	98	98
Allowed	2	2
Outliers	0	0
PDB code	5UU0	5URL

[a] Values in parentheses pertain to the highest resolution shell.

[b] $R_{merge} = \frac{\sum_{hkl}\sum_j \left|I_{hkl,j} - \langle I_{hkl}\rangle\right|}{\sum_{hkl}\sum_j I_{hkl,j}}$

4.2.13. SEC-SAXS Measurements and Analysis

SAXS data were recorded at BioCAT (beamline ID-18) at the Advanced Photon Source. Scattering data were collected using a Pilatus 3 1M detector at a distance of 3 m from the sample at a wavelength of 1.03 Å (~12 keV), covering a momentum transfer range (q) of 0.006-0.35 Å^{-1}. Prior to measurements, an inline Superdex 200 (10/300) Increase column was equilibrated with 25 mM HEPES pH 7.5, 400 mM LiCl, 10% v/v glycerol. Protein samples at 400 μM (PupB-NTSD:PupR-CCSSD), or 800 μM (PupB-NTSD:PupR-CCSSD) were injected onto the SEC column with a flow rate of 0.6 mL/min, and the scattering recorded by exposing the column eluate to the X-ray beam for 1 sec with a 2 sec delay. Measurements were performed at ambient temperature. Data was also collected on 200 or 400 μM PupR-CCSSD, following the protocol specified above.

Scattering data were normalized to the incident X-ray beam, and scattering from the SEC buffer was subtracted with Igor Pro and BioCAT beamline extensions. Data analysis was performed utilizing the ATSAS suite (198). PRIMUS was used for data merging, calculating the radius of gyration from a Guinier approximation, and evaluating protein order by Kratky plot (193). Linearity of the Guinier region confirmed monodispersity of the protein samples. The pairwise distribution function and maximum particle dimension were determined in GNOM (199). For each protein sample, 10-20 independent *ab initio* bead models that agreed with the experimental scattering data were calculated in DAMMIF (196). The models had an average χ^2 of 1.1 ± 0.002 and χ^2 of 1.0 ± 0.0004, with an averaged normalized spatial discrepancy (NSD) for the DAMMIF calculations of 0.62 ± 0.03 and 0.89 ± 0.05 for the PupR-CCSSD and the PupR-CCSSD:PupB-NTSD complex, respectively. The resulting molecular envelope was fit with the available PupR-CCSSD X-ray crystal structure using either SUPCOMB (200) or SITUS (201,202).

To evaluate whether the observed domain instability was a result of conformational plasticity of the PupR-CCSSD linker region, the two subdomains of PupR were specified as rigid bodies, with the linker between the two domains defined as flexible, and used as the input for multi-state modeling with MultiFoXS (267). MultiFoXS combines conformational sampling via a rapidly exploring random tree search to produce 10,000 conformations based on the provided rigid body structure and the flexible region(s) with multi-state model enumeration and scoring by minimizing the chi score for optimal fit to the experimental SAXS data. The models with the lowest chi values and deviations were identified. Additionally, EOM 2.0 in ATSAS (268) was utilized to generate PupR-CCSSD flexible conformers that align with the SAXS profile, using the two subdomains and full PupR-CCSSD sequence as the input files.

4.3. Results

4.3.1. The PupR-CCSSD is Unstable and Monomeric in Solution

Although expression and purification of a soluble MBP-tagged PupR-CCSSD was successful, purification of the stable domain with proteolysis of the MBP tag by TEV protease required considerable effort. Several truncations of the PupR-CCSSD (residues 110-324) were constructed: PupR$^{110-238}$, PupR$^{110-250}$, PupR$^{238-324}$, and PupR$^{250-324}$, all of which proved to be quite unstable once the MBP tag had been cleaved from the fusion protein. Instability was partially alleviated for the full PupR-CCSSD by addition of 10% v/v glycerol during cell lysis and purification, as well as adjustment of salts in the purification buffers, based on the Hofmeister series. Screening of KCl, NaCl, LiCl, and $MgCl_2$ at various concentrations lead to the identification of 400 mM LiCl as the optimal ionic condition for maintaining soluble PupR-

CCSSD. However, despite these optimizations, MBP still featured as a considerable contaminant in the final stages of purification (Figure 4.3 inset), and significant precipitate was observed during protein concentration at various stages of purification.

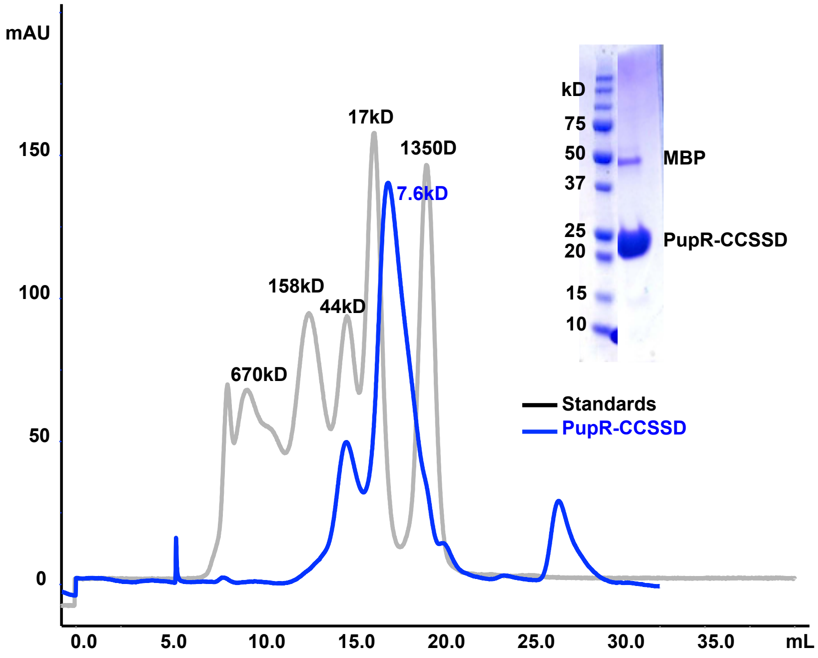


Figure 4.3. SEC chromatogram of the PupR-CCSSD. The domain exhibits delayed elution over a Superdex 200 (10/300), indicating that the domain is likely interacting with the resin. Thus, the apparent molecular weight of PupR-CCSSD is ~8 kD, while the theoretical molecular weight is 24.1 kD, as approximated from sample migration in SDS-PAGE. Contaminating MBP is also observable in the final sample after purification; the molecular weight of the MBP after TEV proteolysis is approximately 43 kD.

Results from pull-down analysis of GST-PupB-NTSD with MBP-tagged truncations of the PupR-CCSSD suggested that only the full-length CCSSD interacted with the PupB-NTSD (Figure 4.4). Since the PupR-CCSSD:PupB-NTSD X-ray crystal structure (described later) indicates a limited interaction interface with the PupB-NTSD, the working hypothesis was that the full-length domain is conformationally unstable when not associated with the PupB-NTSD, but that the N-terminal region of the domain stabilizes the C-terminal region.

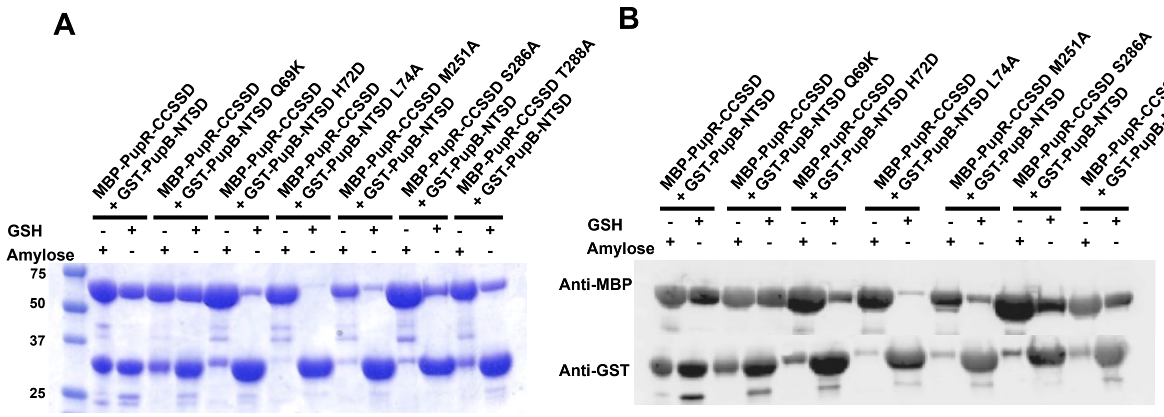


Figure 4.4. SDS-PAGE and Western blotting of affinity pull-downs with GST-PupB-NTSD and MBP-tagged PupR-CCSSD truncations. PupR-CCSSD constructs used: full-length PupR-CCSSD (residues 110-324), residues 110-238, residues 110-250, residues 238-324, and residues 250-324. (A) Affinity pull-downs over either GSH or amylose resin indicate an interaction of the two domains only for the full-length PupR-CCSSD, as confirmed by Western blotting (B).

Circular dichroism spectroscopy and CD spectroscopy thermal denaturation also revealed significant flexibility in the PupR-CCSSD. Thermal denaturation of the PupR-CCSSD yielded a melting temperature of 40.2 °C, determined by fitting the standard Boltzmann sigmoidal curve to the T_m profile recorded while heating the domain from 20 to 60 °C. Protein renaturation was not observed during cooling, as the PupR-CCSSD precipitated at high temperatures (Figure 4.5B).

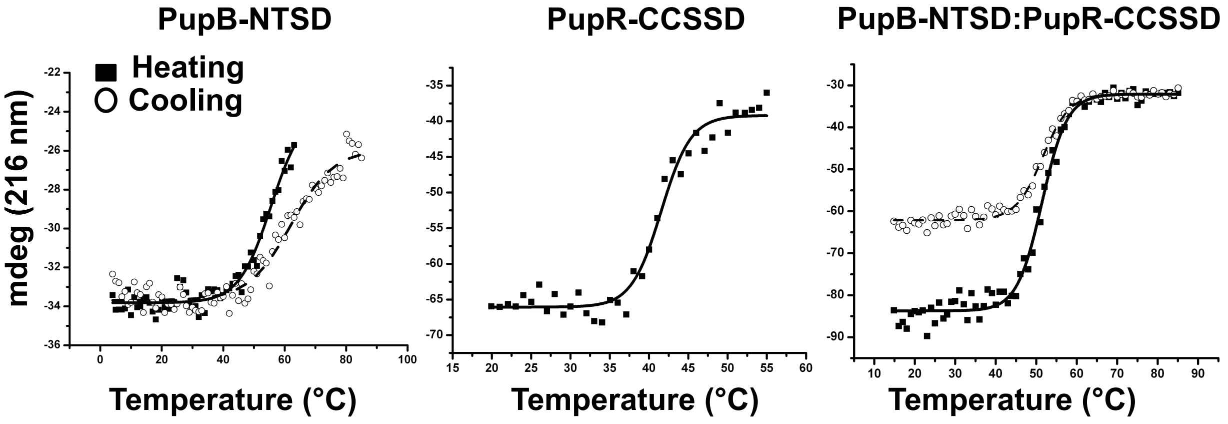


Figure 4.5. Thermal denaturation CD profiles of PupB-NTSD, PupR-CCSSD, and PupR-CCSSD:PupB-NTSD. Unfolding (■ *heating*) and refolding (**O** *cooling*) was monitored at 216 nm.

SEC-SAXS was performed to probe the low-resolution structure and solution properties of the PupR-CCSSD. This technique also allowed for a more accurate determination of the molecular mass and possible oligomeric state of the PupR-CCSSD, as this information could not be unambiguously determined from SEC alone. The linearity of the Guinier plot in the $0 < q < 0.003$ Å verified the monodispersity of the sample. The radius of gyration (*Rg*), calculated from the Guinier region, was 23 Å, and the D_{max} was 75 Å, as determined by *P(r)* analysis (Figure 4.6A-B). Theoretical R_g and D_{max} calculated from the PupR-CCSSD crystal structure were 23 Å and 70 Å, respectively. The Kratky plot indicated that the PupR-CCSSD was partially unfolded or flexible in solution (Figure 4.6C); this is consistent with the circular dichroism results (Figure 4.7), which suggested that nearly 50% of the domain was random coil or disordered. The sample molecular weight estimated from the SAXS data was 26 kD, with the maximum q set to 0.35 Å^{-1}. The theoretical molecular mass of the PupR-CCSSD is 24.1 kD; thus, the SAXS data support the finding that the PupR-CCSSD is monomeric in solution. The theoretical SAXS curve was calculated for the PupR-CCSSD crystal structure (discussed in detail in Chapter 4.3.3), and fit to the experimental PupR-CCSSD scattering profile (Figure 4.6D and Figure 4.8A), yielding a chi value of 2.35. The higher chi value is due to the poor fit around $q = 0.2$ Å^{-1}.

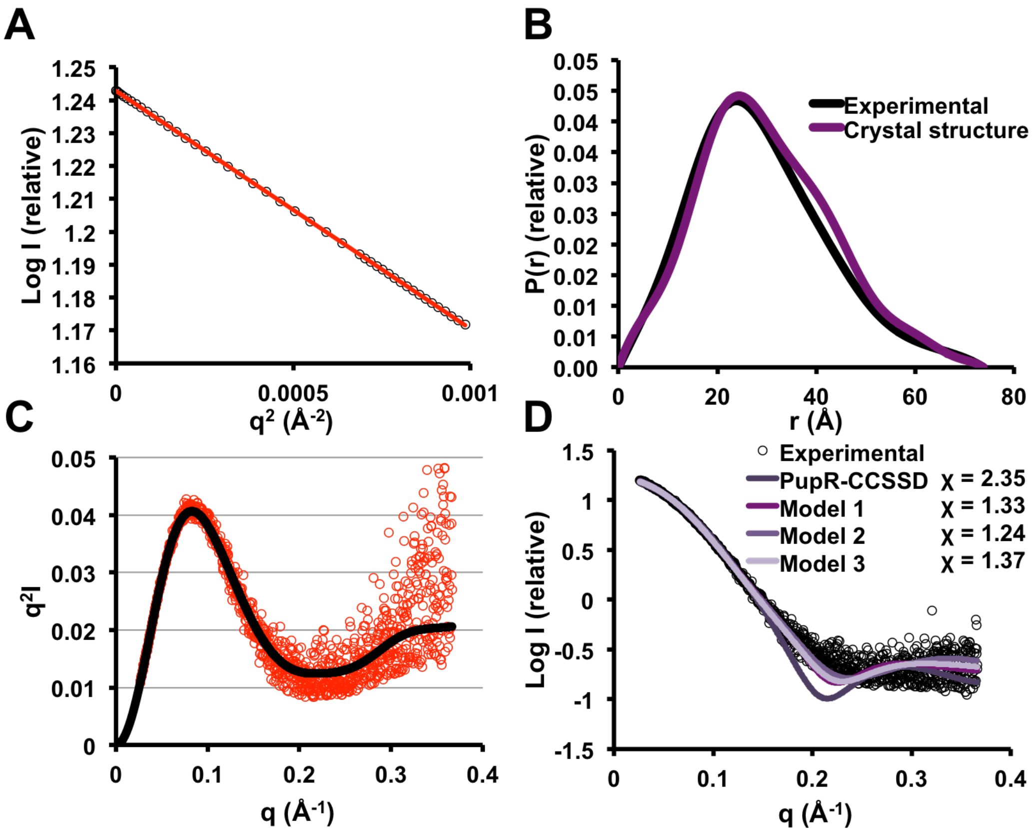


Figure 4.6. SEC-SAXS analysis of the PupR-CCSSD. (A) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log(intensity). The linearity of the region indicates that the sample is monodisperse (non-aggregated), with the slope of the line equal to the radius of gyration. (B) Distance distribution *P(r)* for the experimental data (black) and the theoretical curve calculated from the PupR-CCSSD crystal structure (purple). (C) Kratky plot. (D) Experimental scattering profile, fit with the theoretical scattering profiles of the PupR-CCSSD crystal structure, and Models 1-3, generated by structural conformation sampling. The chi values (χ) for each structure are indicated.

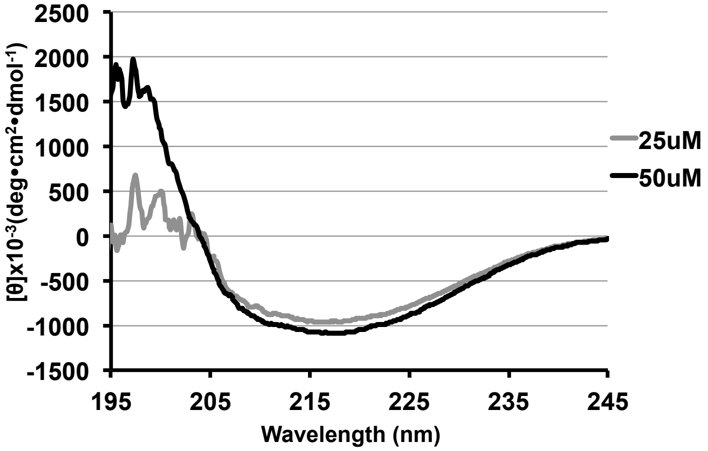


Figure 4.7. Circular dichroism spectra of the PupR-CCSSD at two concentrations. The noise at the low wavelengths likely indicates that the buffer exchange prior to CD data collection (from SEC buffer to CD buffer) was incomplete, and some residual glycerol remained in solution. Estimations of secondary structure content with CONTIN and CDSSTR yielded 13.7%, 35.6%, and 50.6% for helix, strand, and coil, respectively. These values closely align with the secondary structure content of the PupR-CCSSD structural homolog (PDB entry 4M0N), which has helix = 12.9%, strand = 27.2%, and coil = 59.9%.

To examine the possibility that the PupR-CCSSD is conformationally heterogeneous in solution, MultiFoxS was used to generate ~10000 conformers, based on the likelihood of the plasticity of the linker region. The two PupR-CCSSD subdomains, comprising residues 110-238 and 250-324, were maintained as rigid bodies, while the residues of the linker region were defined as flexible, so their backbone phi and psi angles were sampled by the MultiFoxS sampling algorithm. To create Model 1 (Figure 4.6D and Figure 4.8B), the linker region was defined as residues 239-250. With these parameters, it was determined that either a two- or three-state model best fit the experimental data. This indicates that the PupR-CCSSD likely assumes two or three predominant species in solution. For the two-state model, the predominant conformation (Figure 4.8B) has an R_g = 20.2 Å and comprises 89% of the solution state, while the secondary species has an R_g = 22.4 Å and is sampled 11% of the time. For the three-state model, the primary conformation (also Figure 4.8B) is present as 75.8% of the weighted population, while the secondary and tertiary conformations have R_g = 20.4 Å and 22.4 Å, sampled at 17.1% and 7.1%, respectively. Theoretical scattering calculated from the predominant conformer identified by this method fit well to the experimental SAXS data, with an improved chi value of 1.33.

A similar method was attempted to create Model 2 (Figure 4.5D and Figure 4.8C), with the flexible region defined as residues 232-250. With this extended flexible linker, a two- or three-state model was again suggested. The two-state model resulted in R_g = 21.7 Å and 18.7 Å, sampled at 87% and 13%. The three-state model yielded R_g = 19.8 Å, R_g = 18.7 Å, and R_g = 24.9 Å, sampled at 23.1%, 48.5%, and 28.4%, respectively. The predominant species of the best-fitting conformers from each model (Figure 4.8C) improved the chi value to 1.24.

EOM was also used to generate models with a flexible linker region that fit the experimental scattering data (Model 3, Figure 4.6D). Residues 239-249 were defined as the flexible region. The best EOM model resulted in a chi value of 1.37 (Figure 4.8D). EOM assumes the protein is homogeneous in solution.

Low-resolution *ab initio* bead models were calculated from the experimental data, with no symmetry operators enforced. The PupR-CCSSD X-ray crystal structure and each predominant conformer generated from EOM or MultiFoxS were docked into the resulting SAXS envelope using SUPCOMB in the ATSAS suite (Figure 4.8).

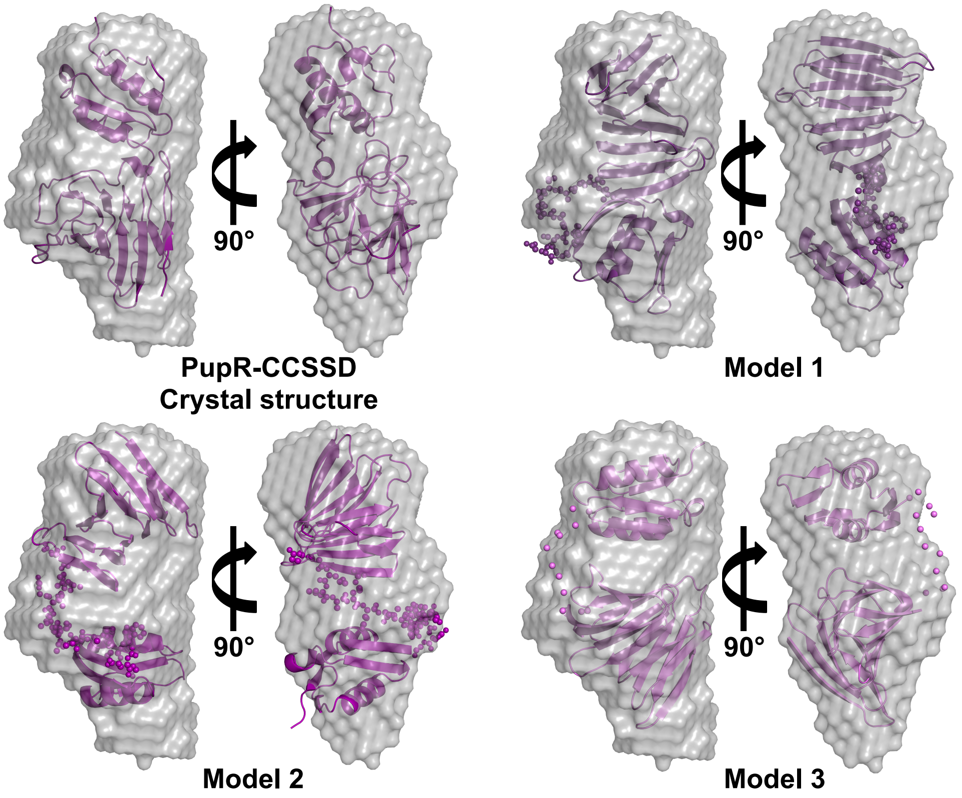


Figure 4.8. *Ab initio* bead model of PupR-CCSSD with the calculated surface superimposed with the PupR-CCSSD X-ray crystal structure and PupR-CCSSD models. (A) PupR-CCSSD X-ray crystal structure. (B) Model 1, which defined the linker region as residues 239-250; (C) Model 2, which defined the linker region as residues 232-250; and (D) Model 3, the EOM output. The two subdomains of the PupR-CCSSD are shown in shades of purple ribbon, with the linker regions of Models 1, 2, and 3 displayed as non-bonded spheres.

4.3.2. The PupR-CCSSD is Stabilized via Interaction with the PupB-NTSD

Thermal denaturation coupled to circular dichroism spectroscopy indicated that, when the PupB-NTSD (residues 49-128 of PupB) is combined with the PupR-CCSSD, the PupR-CCSSD is stabilized (4.4C). The PupR-CCSSD alone denatured at 40.2 °C, and could not be refolded upon cooling. When combined with the PupB-NTSD, however, the T_m of the complex was 51.4 °C and 52.9 °C for heating and cooling, respectively. Comparison to the T_m profile of the PupB-NTSD alone (Figure 4.4A) suggests that the values are a product of complex denaturation, and not simply an average of the two proteins denaturing independently.

The interaction of the PupR-CCSSD and PupB-NTSD was quantified using isothermal titration calorimetry, by titrating a 10-fold excess of PupB-NTSD into PupR-CCSSD. The PupB-NTSD binds to the PupR-CCSSD with a moderate binding affinity of 1.5 ± 0.38 μM. was determined, with favorable free energy of association (*ΔG*) (Figure 4.9). Data analysis with NanoAnalyze reflected a 1:1 stoichiometry. In order to compensate for the presence of MBP in the final PupR-CCSSD sample, and to account for PupR-CCSSD that precipitated during titration, an incompetent fraction was applied as a local parameter per experiment, and globally fit to a one-site model using SEDPHAT (248,250). Concentration errors were also incorporated into local incompetent fractions for each individual titration, and the thermodynamic binding parameters considered global parameters for all titrations. Thus, the local incompetent fraction range was found to be 5.1-16.7% amongst the three experiments.

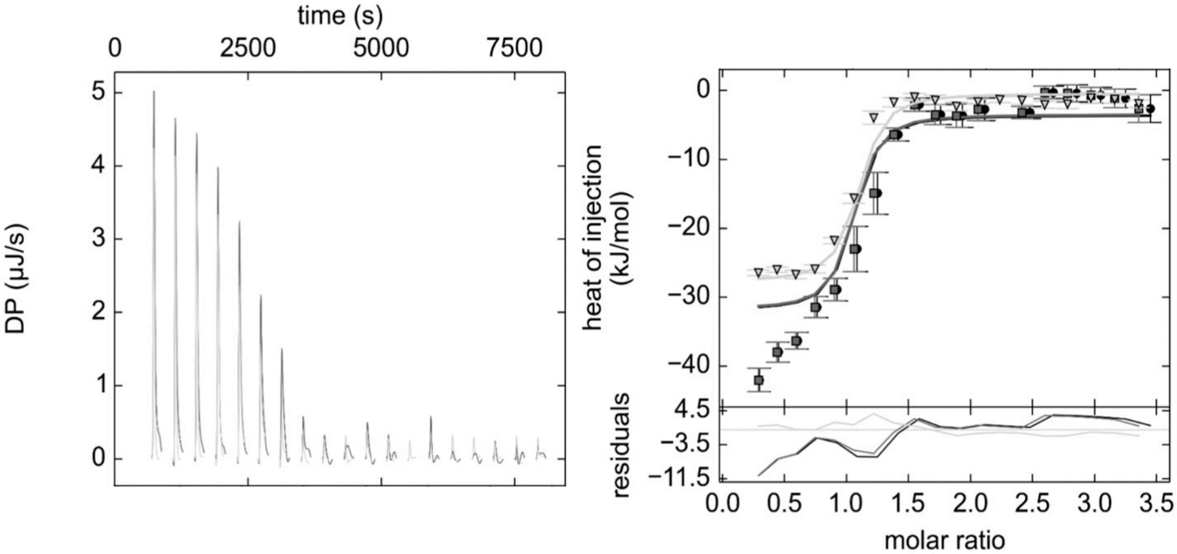


Figure 4.9. ITC profile of PupB-NTSD binding to PupR-CCSSD. Three such separate titrations were fit with a one-site binding model.

NMR titration was employed to validate the binding affinity of PupB-NTSD with PupR-CCSSD, and also to monitor subtle conformational changes induced in the PupB-NTSD upon PupR-CCSSD binding. 2D ^{1}H-^{15}N correlation HSQC experiments were initiated, and ^{15}N PupB-

NTSD was combined with PupR-CCSSD in various molar ratios: 1:0.625, 1:1, 1:1.36, 1:2, 1:2.5, and 1:3. Titration spectra were compared to the spectra of apo PupB-NTSD. The standard 2D ^{1}H-^{15}N HSQC spectra of PupB-NTSD exhibits well-resolved NH signals with ^{1}H and ^{15}N resonance widths characteristic of a well-folded monomer (Figures 3.2 and 4.10A). Addition of PupR-CCSSD up to a 1:2 PupB-NTSD:PupR-CCSSD ratio elicited no response by PupB-NTSD (Figure 4.10B-C), and thus, no significant changes in chemical shifts or peak broadening were observed. Cursory examination of changes in signal intensity indicated minimal drops in peak intensities, but residues mediating the interaction could not be identified. Notably, addition of PupR-CCSSD in 2.5 times molar excess of PupB-NTSD resulted in significant loss of spectral intensity for nearly all NH resonances, with no peak broadening detected for the remaining resonances (Figure 4.10D). The observation reduction in tumbling rate of the molecules may be due to a couple of different phenomena: 1) the formation of larger non-specific aggregates, 2) broad-scale conformational changes, 3) intermediate exchange, when the off-rate is approximately equal to the chemical shift difference between the free and bound state (269). At high concentrations, PupR-CCSSD may precipitate; however, addition of PupB-NTSD considerably improves stability of the CCSSD, as demonstrating by CD-monitored thermal denaturation experiments. Assuming a simple two-state equilibrium with second-order binding kinetics (for 1:1 complex formation), the rate exchange constant, k_{ex}, is approximately equal to the difference in angular frequency, Δω:

$$k_{ex} = k_{on}[L] + k_{off} \qquad \text{(Eq. 4.1)}$$

$$\Delta\omega = \omega_P - \omega_{PL} \qquad \text{(Eq. 4.2)}$$

Where k_{on} and k_{off} are the association and dissociation rate constants, and the difference in angular frequency, Δω, equals the difference between the unbound resonance, ω_P, and the bound resonance, ω_{PL} (270). Resonances in intermediate exchange have been shown to disappear, and sometimes, not reappear, due to limitations of the ligand (namely, concentration and solubility). Thus, it seems that titration of PupB-NTSD with PupR-CCSSD results in intermediate exchange on the NMR time scale, which correlates with the moderate binding affinity calculated by ITC (1-10 μM). Due to the limitation of inability to concentrate PupR-CCSSD beyond 400 μM, ratios greater than 1:3 were not attempted, so it is unknown if signal intensity may be recovered.

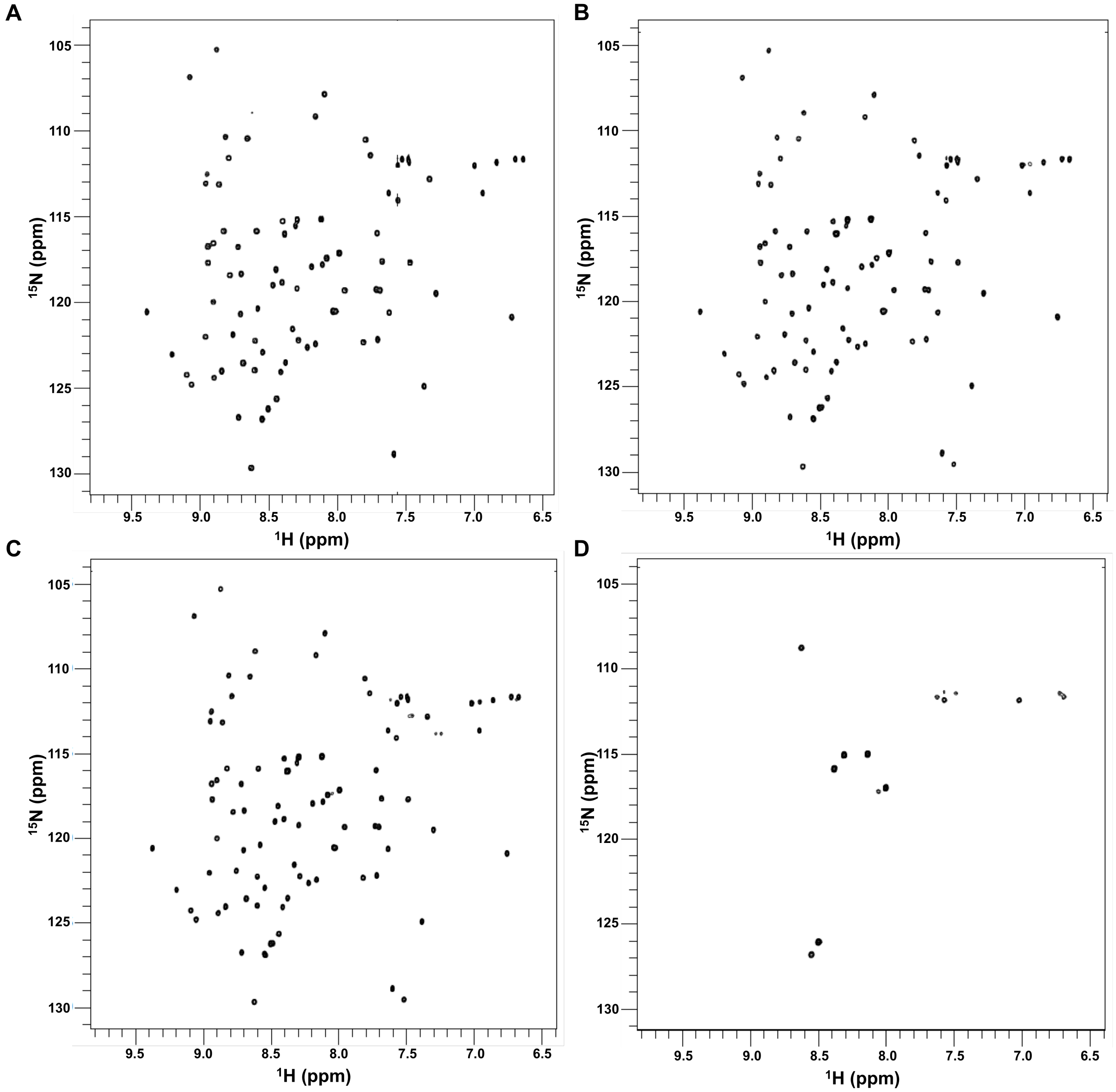


Figure 4.10. 2D ^{1}H-^{15}N HSQC spectra upon titration of PupB-NTSD with the PupR-CCSSD. The panels correspond to varying ratios of ^{15}N-labeled PupB-NTSD and unlabeled PupR-CCSSD. (A) Apo ^{15}N PupB-NTSD. (B) 1:0.625 ^{15}N PupB-NTSD:PupR-CCSSD. (C) 1:1.4 ^{15}N PupB-NTSD:PupR-CCSSD. (D) 1:2.5 ^{15}N PupB-NTSD:PupR-CCSSD. All spectra are shown at the same contour level.

4.3.3. The PupR-CCSSD:PupB-NTSD Complex

SEC-SAXS data were recorded from a 1:1 molar ratio of the PupR-CCSSD and the PupB-NTSD, and the experiment performed as described for the PupR-CCSSD alone. The

Guinier plot (Figure 4.11A) indicates that the sample is monodisperse, and the R_g = 26 Å. The distance distribution yields a D_{max} = 86.7 Å (Figure 4.11B). Examination of the Kratky plot (Figure 4.11C) suggests some flexibility in the complex. Fitting of the experimental scattering data to the theoretical scattering of the PupR-CCSSD alone, or the PupR-CCSSD:PupB-NTSD complex from the X-ray crystal structure, implies that the complex was the likely species in solution (χ = 4.35 for the CCSSD, and χ = 1.92 for the complex); plasticity of the complex that is not accounted for in the crystal structure may explain the high chi value. The molecular weight of the complex estimated from the SAXS data is 35.5 kD, which deviates by less than 10% from the theoretical molecular weight of the complex (32.3 kD). This suggests that the primary species observed in the SEC-SAXS scattering profile is a 1:1 PupR-CCSSD:PupB-NTSD complex.

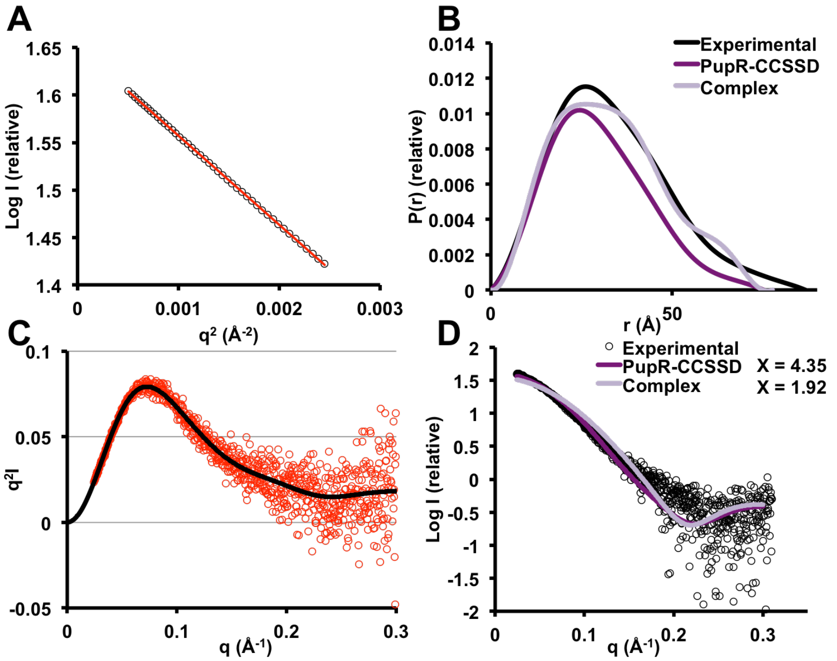


Figure 4.11. SEC-SAXS analysis of the PupR-CCSSD:PupB-NTSD complex. (A) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log(intensity). (B) Distance distribution *P(r)* for the experimental data (black), the theoretical curve calculated from the PupR-CCSSD crystal structure (dark purple), and the theoretical curve calculated from the PupR-CCSSD:PupB-NTSD crystal structure (light purple). (C) Kratky plot. (D) Experimental scattering profile, fit with the theoretical scattering profiles of the PupR-CCSSD crystal structure and the PupR-CCSSD:PupB-NTSD complex.

The X-ray crystal structure of the PupR-CCSSD:PupB-NTSD complex was solved by single-wavelength anomalous diffraction, with a selenomethionine PupR-CCSSD derivative used for recording anomalous diffraction. Crystals belonging to space group $P2_12_12_1$ of native and SeMet-containing PupR-CCSSD complex diffracted to 1.76 and 1.50 Å, respectively. Three of the four possible selenium sites of the PupR-CCSSD were identified, and this substructure enabled calculation of phases and automated building of one PupR-CCSSD per crystal asymmetric unit. Excess density was observed within the asymmetric unit after selenium site location and construction of one PupR-CCSSD chain. Identification of PupB-NTSD aromatic residue sidechains enabled a poly-alanine chain corresponding to the PupB-NTSD fold to be positioned within this density, and the domain was subsequently manually built in COOT (261). The completed 1.6 Å model was used as a search model for MR to determine the structure from the native data set at 1.76 Å resolution. R_{work} and R_{free} converged to 15.3% and 18.4%, respectively, for the SeMet model. No Ramachandran outliers were present in the final model (Table 4.4). The final SeMet SAD model had 213 of 219 residues of the PupR-CCSSD (chain R), and 80 of the 82 residues of the PupB-NTSD (chain B). Excess electron density at the N-terminus of the PupB-NTSD enabled an additional residue to be built in the final native model. Two molecules of tartrate from the crystallization reservoir solution were also placed in the asymmetric unit of each model as indicated by the clear electron density.

The structure of the PupB-NTSD from the PupR-CCSSD:PupB-NTSD complex contains the canonical βαβ-repeat motif observed in the PupA-NTSD and FecA-NTSD NMR structures (Figure 4.12). The PupB-NTSD shares 37.1% sequence identity with the PupA-NTSD, and 30.5% sequence identity with the FecA-NTSD. The secondary and tertiary structure elements appear to be conserved amongst TonB-dependent transporter proteins. The PupB-NTSD, as seen with the PupA-NTSD and FecA-NTSD, contains an internal pseudo- two-fold symmetry axis, with two α-helices sandwiched between two anti-parallel β-sheets. As this fold is conserved amongst TBDT signaling domains, as well as the N-terminal region of secretins of the bacterial type II and III secretory system, this domain has been designated an STN domain (SMART accession number SM00965) for "secretin and TonB N-terminus short domain."

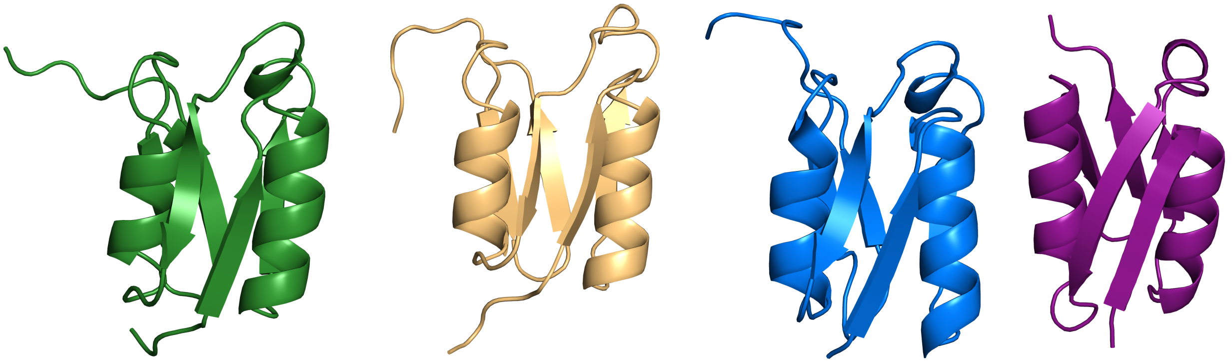


Figure 4.12. X-ray crystal structure of the PupB-NTSD, PupA-NTSD NMR structure, FecA-NTSD NMR structure, and PupR-CCSSD STN X-ray crystal structure. The PupB-NTSD is shown in green, the PupA-NTSD (PDB 2A02) in tan, the FecA-NTSD (PDB 1ZZV) in blue, and the PupR-CCSSD STN in purple; all are displayed in superimposable orientations.

The X-ray crystal structure of the PupR-CCSSD:PupB-NTSD complex includes residues 111-323 of the periplasmic C-terminal region of PupR, previously described in the literature as the C-terminal domain. To clarify this domain type in functional terms, it has been renamed the C-terminal cell-surface signaling domain (CCSSD). The structure of this domain reveals that it is composed of two subdomains of markedly different folds. Residues 110-238 compose the N-terminal subdomain, the structure of which is two 7-stranded β-sheets. One β-sheet is a mixed-parallel/anti-parallel β-sheet, where β-strands 5 and 7 form a small parallel sheet, sandwiched between the rest of the β-strands, each linked by loops or β-arcs. The other β-sheet is an anti-parallel β-sheet. The two β-sheets form a twisted shape reminiscent of a β-solenoid, wherein the core of the twisted sheets is stabilized by the primarily hydrophobic or nonpolar sidechains of the residues of the β-strands. The novel fold of the N-terminal region of the CCSSD appears to be unique for periplasmic anti-sigma proteins, as the only structural homolog identified containing the same fold is the putative anti-sigma protein (PDB entry 4M0N). As this N-terminal subdomain of the CCSSD is adjacent to the inner membrane, it is designated here as the C-terminal juxtamembrane (CJM) subdomain.

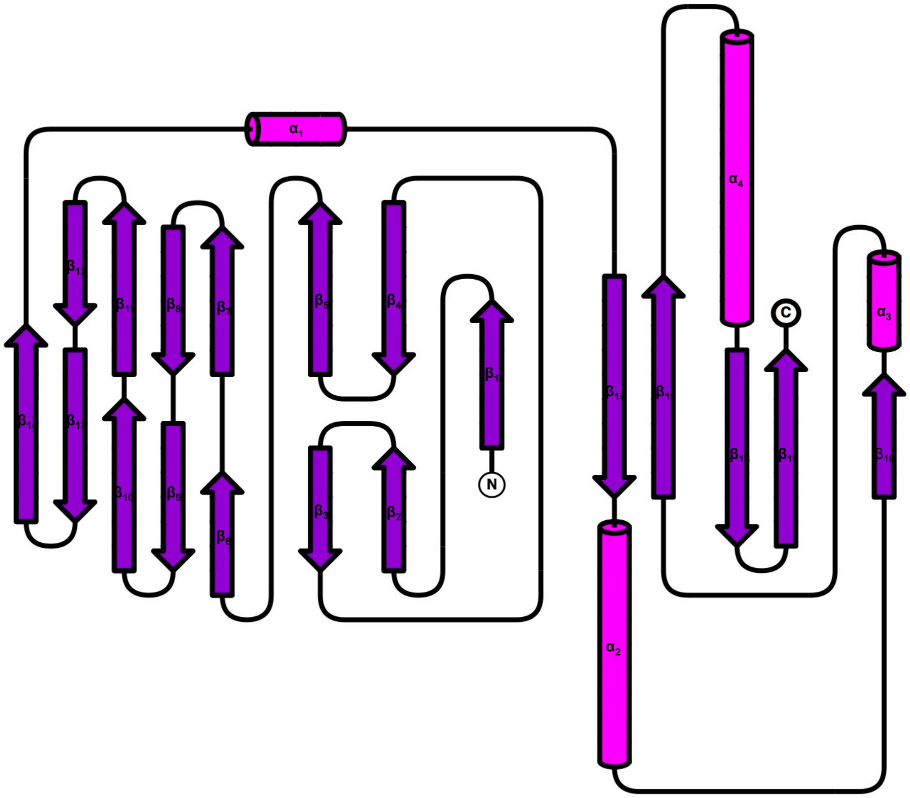
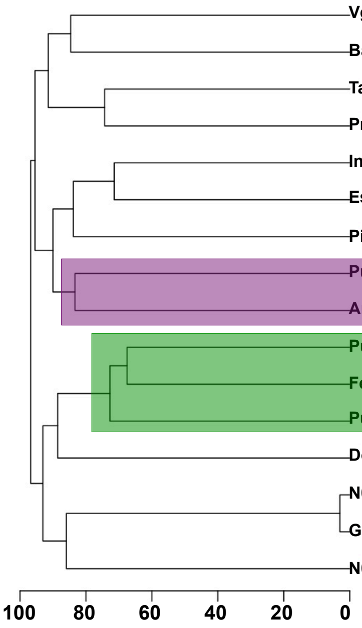


Figure 4.13. Topology of the PupR-CCSSD. β-strands are displayed as purple arrows, and α-helices as pink cylinders.

Surprisingly, the C-terminal subdomain of the CCSSD, residues 250-324, exhibit the same βαβ-repeat motif observed in the PupB-NTSD (Figure 4.12). This subdomain connects to the CJM subdomain via an 11-residue linker that has a primarily extended structure, apart from a single helical turn of residues 246-249. The core of this subdomain is stabilized by conserved nonpolar or hydrophobic residues, similar to the PupB-NTSD (Figures 4.15 and 4.16). The βαβ-repeat motif is not unique to the N-terminal signaling domains of TBDTs (Figure 4.12). Structural homologs have been identified in diverse proteins such as the periplasmic domain of InvG the *Salmonella enterica* type III secretion system (PDB 4G08) (271); the periplasmic domain of the *N. meningitidis* PilQ protein of the Type IV pili (PDB 4AR0) (272); the periplasmic domain of the outer membrane secretin EscC from enteropathogenic *E. coli* (EPEC) (PDB 3GR5) (273); the periplasmic domain of the secretin GspD from EPEC (PDB 3EZJ) (274); the periplasmic domain of the type II secretin GspD from EPEC (PDB 4JTM) (275); the periplasmic domain of DotD of the Type IVB secretion system of *Legionella pneumophila* (PDB 3ADY) (276); residues 97-176 of the tail protein of prophage MuSo2 from *Shewanella oneidensis* (PDB 3CDD; Chang, C., *et al*., unpublished); the C-terminal domain of VirB7 of the Type IV secretion system from *Xanthomonas citri* (PDB 3OV5 and 2L4W) (277); residues 93-177 of the tail protein from *N. meningitidis* (PDB 3D37; Zhang, R. *et al.*, unpublished); residues 116-199 of the baseplate hub protein gp27 of the Bacteriophage T4 cell-puncturing device (PDB 1K28 and 2Z6B) (278,279); residues 102-184 of Vgrg1 from the *P. aeruginosa* needle tip of the Type VI secretion system (PDB 4UHV (280) and 4MTK; Sycheva, L., *et al.*, unpublished); and the periplasmic domain of a putative anti-sigma factor from *Parabacteroides distasonis* ATCC 8503 (PDB 4M0N and 4M0H; JCSG, unpublished) . Interestingly, with the exception of the bacteriophage T4 gp27 domain, all of the aforementioned domains are located in the Gram-negative bacterial periplasmic space. Many of the NTSD-like domains of the secretins are located at the N-terminal end of the full-length protein (273,274); however, the “STN” designation is outdated, as this domain is clearly 1) not always involved in TBDT signaling, 2) not always located in a secretin complex, and 3) not always located at the N-terminus. Thus, the presence of an STN at the C-terminus of a cell-surface signaling anti-sigma regulator protein represents a new sub-category of this domain type.

Structure-based sequence alignment followed by structural clustering using Structure Alignment-based Clustering of Proteins (STRALCP) (281) was performed to predict the "relatedness" of the structural homologs determined by searching the PDB via DALI. Unlike DALI, which utilizes a single input structure to identify structural homologs and score them based on RMSD to the target structure, STRALCP incorporates a Local-Global Alignment (LGA) algorithm, which performs an "all-against-all" comparison, so each structure in a given set is used as a reference for comparison to all other structures within the set. Interestingly, despite no reliance on primary sequence alignment or consideration of protein function, STRALCP clustered all of the NTSDs together, and the PupR-CCSSD with the *Parabacteroides distasonis* anti-sigma factor structure (4M0N). The structural homologs examined have 6-18% sequence identity to PupB-NTSD (excluding the other NTSD), and 5-18% identity to PupR-CCSSD STN.



	DALI Z-score (PupR-CCSSD STN)	RMSD	DALI Z-score (PupB-NTSD)	RMSD
VgrG1 Type VI secretion protein (*Pseudomonas aeruginosa*)	4.3	3.2	3.6	3.4
Baseplate structural protein GP27 (*Bacteriophage T4)*	5.1	2.6	3.7	4.0
Tail protein (*Neisseria meningitidis* MC58)	7.2	2.6	5.1	3.9
Prophage MuSo2 tail protein (*Shewanella oneidensis*)	7.1	2.6	5.7	3.9
InvG periplasmic domain (*Salmonella enterica*)	11.5	1.7	5.9	3.3
EscC periplasmic domain (*E. coli*)	8.9	2.1	7.2	3.5
PilQ N0 domain (*Neisseria meningitidis*)	6.7	2.0	5.1	3.5
PupR-CCSSD (*Pseudomonas capeferrum*)			9.0	2.6
Anti-sigma factor (*Parabacteroides distasonis*)	22.2	1.7	8.4	2.2
PupB-NTSD (*Pseudomonas capeferrum*)	7.5	2.6		
FecA-NTSD (*E. coli*)	7.2	2.4	10.8	2.3
PupA-NTSD (*Pseudomonas capeferrum*)	8.1	2.2	10.6	2.2
DotD Type IV secretion protein (*Legionella pneumophila*)	6.1	2.6	6.0	3.4
N0 domain Type II secretin (*E. coli*)	9.0	1.9	6.8	2.5
GspD N-terminal domain (*E. coli*)	8.6	1.9	6.9	2.4
N0 domain VirB7 Type IV secretion protein (*Xanthomonas citri*)	6.8	2.4	5.4	2.8

Figure 4.14. Dendrogram of the structure-based sequence similarity of available structures with the βαβ-repeat motif, or STN domain. Clustering was accomplished with STRALCP (281), a multi-criteria-based clustering approach utilizing the LGA algorithm, with an LGA cutoff of 40% (the minimum structural similarity score for clustering). DALI (266) Z-scores and RMSDs are indicated for structural alignment with the PupR-CCSSD STN domain or the PupB-NTSD.

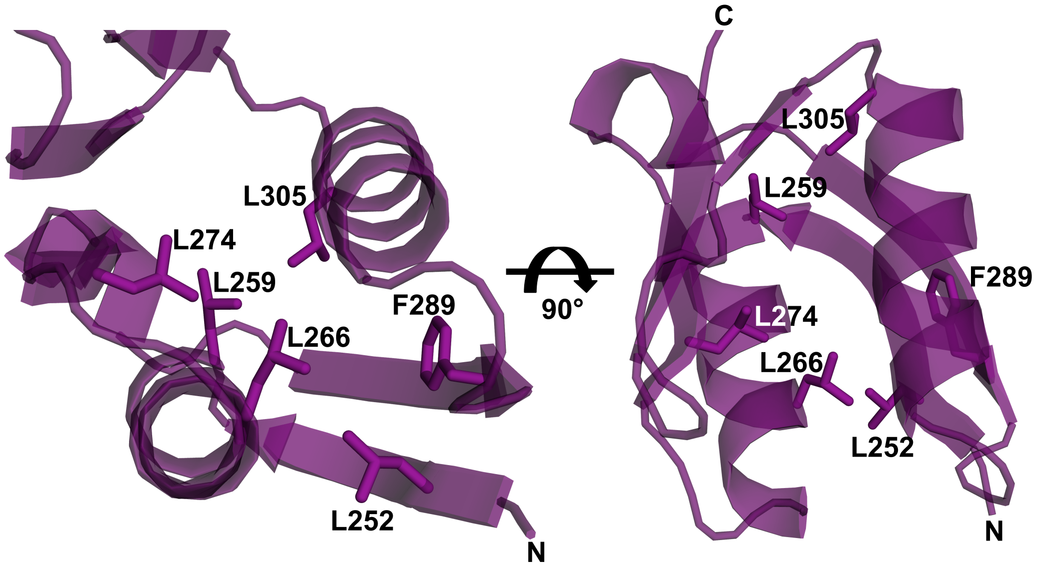


Figure 4.15. Conserved hydrophobic and nonpolar residues stabilize the core of the PupR-CCSSD C-terminal STN. Residues L252, L259, L274 are invariant amongst putative anti-sigma regulator proteins (Figure 1.14); residues L266, L305, F289 are semi-conserved.

The interaction interface observed in the crystal structure includes α_1 and β_2 of the PupB-NTSD and β_{17} and α_2 of the PupR-CCSSD (Figure 4.17). The total surface area buried upon the association of the two domains is 1438.6 Å^2, as calculated with PISA. The interface is roughly 6.3% of the total surface of the PupR-CCSSD, and 15.3% of the total surface of the PupB-NTSD. Several critical residues mediate this interaction, as shown in Figure 4.18.

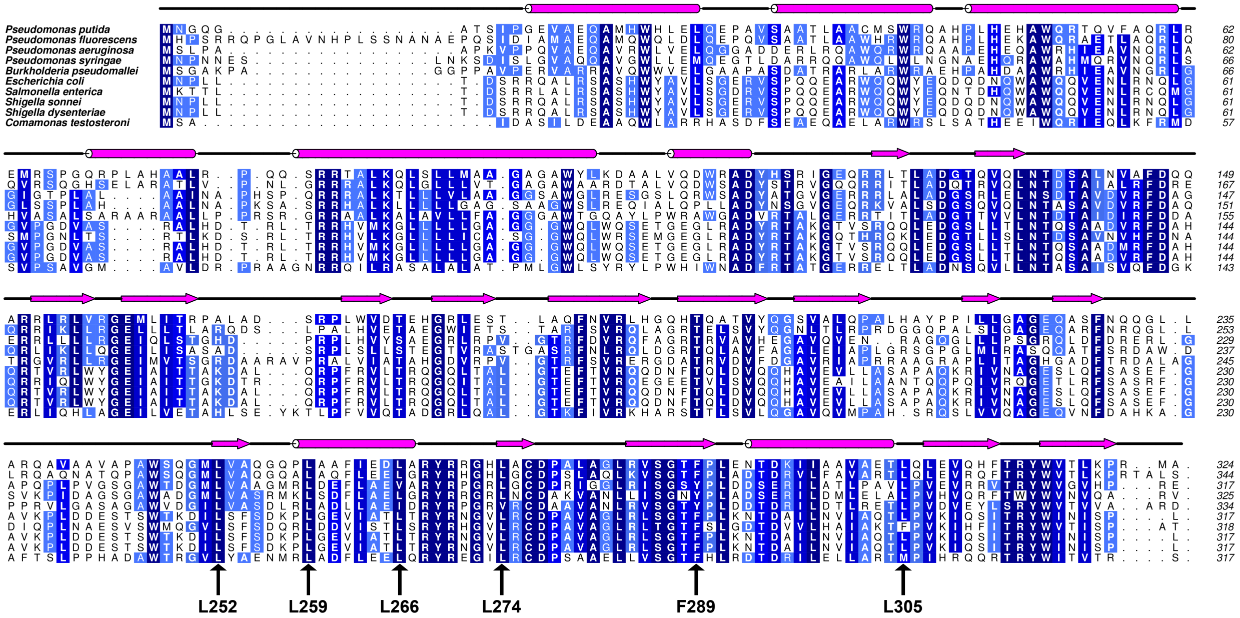


Figure 4.16. Sequence alignment and secondary structure predictions of anti-sigma regulator proteins from various proteobacteria with putative iron transport TBDTs. The inner membrane regulator proteins of PupR (Q52209, *Pseudomonas capeferrum*) were aligned with PupR orthologs (Q4KDP8, *Pseudomonas fluorescens*; Q48ML1, *Pseudomonas syringae*), FpvR (Q91192, *Pseudomonas aeruginosa*), FecR (P23485, *Escherichia coli*), and FecR orthologs (Q63LI7, *Burkholderia pseudomallei*; V2K8B2, *Salmonella enterica*; Q3YU71, *Shigella sonnei*; B3WZ95, *Shigella dysenteriae*; H1RKR8, *Comamonas testosteroni*). Sequence alignment was performed using T-COFFEE and edited with Aline. Invariant residues are highlighted in navy blue, and conserved residues in varying lighter shades of blue, with the lightest shade corresponding to the least conserved residues. Secondary structure prediction was obtained using Psipred, and is displayed above the primary sequence, with purple cylinders denoting helices, purple arrows denoting β-strands, and black lines denoting coils.

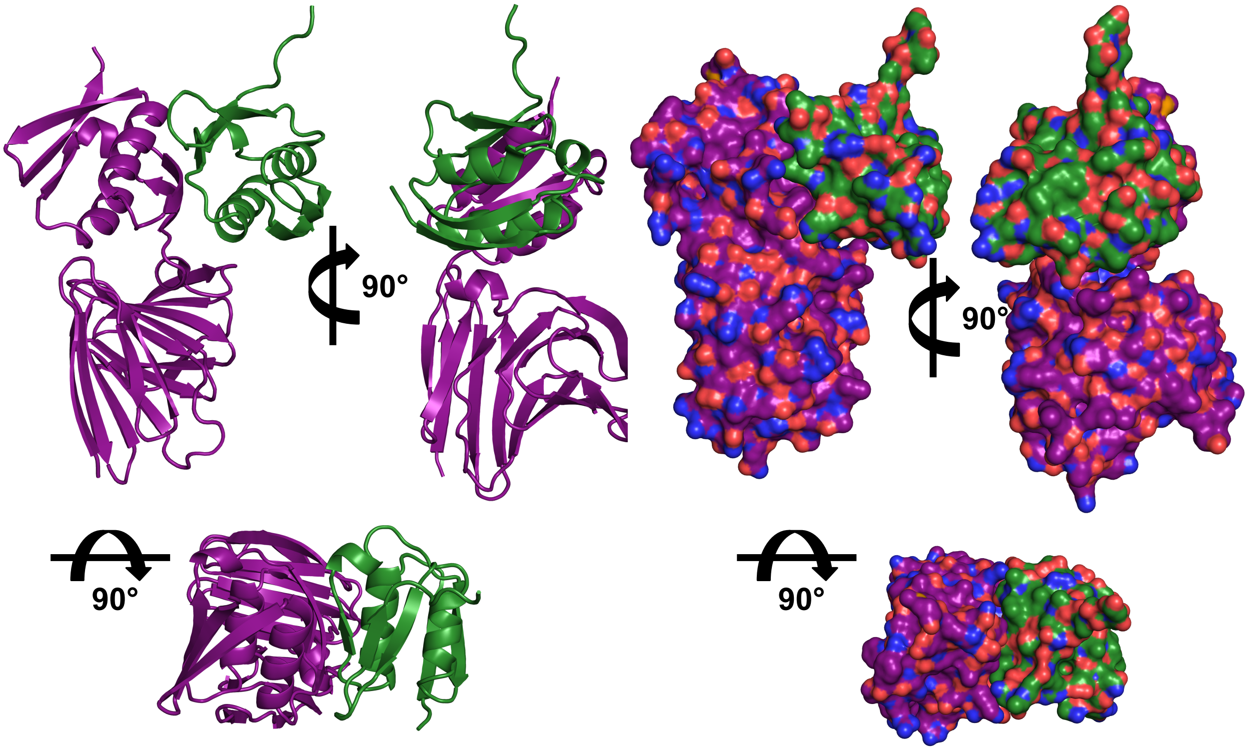


Figure 4.17. The X-ray crystal structure of the PupR-CCSSD:PupB-NTSD complex. The PupR-CCSSD is displayed as purple ribbon or surface representation; the PupB-NTSD is shown as green ribbon or surface. The surface representations are colored according to atom type: carbon, purple or green; oxygen, red; nitrogen, blue; sulfur, yellow.

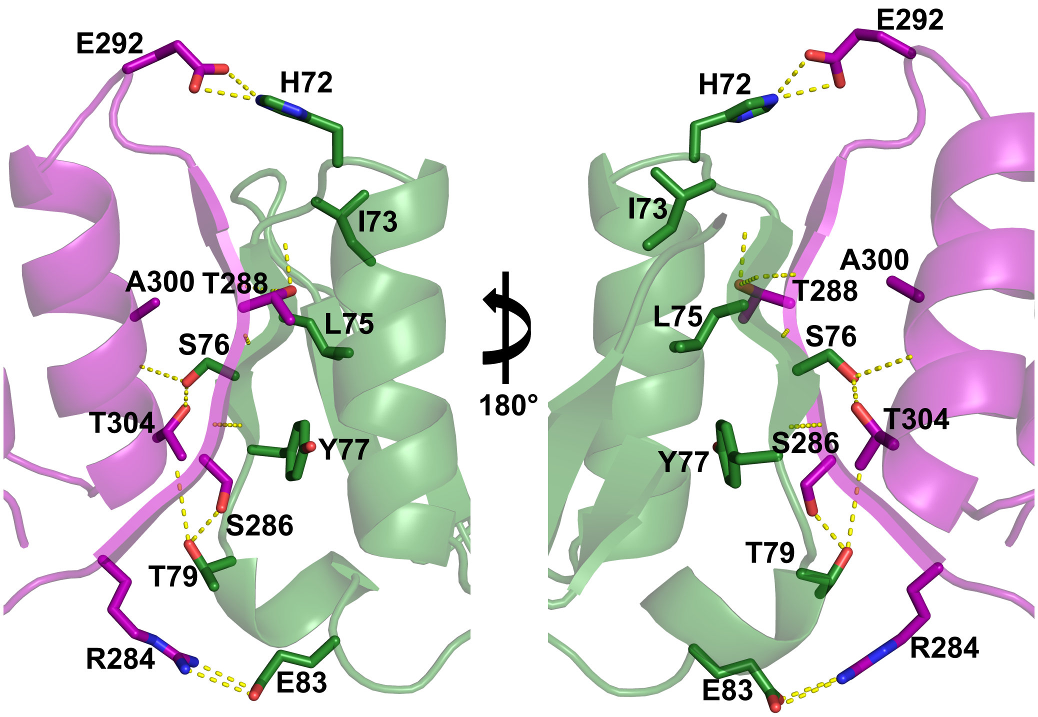


Figure 4.18. Residues at the interface of the PupR-CCSSD and PupB-NTSD complex that stabilize the interaction of the two domains. The PupR-CCSSD is displayed in purple, and the PupB-NTSD in green. Yellow dashes indicate hydrogen bonds or salt bridges.

H72 of PupB-NTSD and E292 of PupR-CCSSD form a salt bridge (Figure 4.19A), as does E83 of PupB-NTSD and R284 of PupR-CCSSD (Figure 4.19D) (values <100 indicate residues on the PupB-NTSD; values >200 indicate residues on the PupR-CCSSD). Although the interface is fairly minimal, an extensive hydrogen-bonding network stabilizes the interface. The backbone amide of L75 H-bonds the backbone carboxyl and sidechain hydroxyl of T288, and the backbone carboxyl of L75 H-bonds the backbone amide of T288; additionally, this interaction is stabilized by an H-bond between the backbone carbonyl of I73 and the sidechain hydroxyl of T288 (Figure 4.19G). The sidechain hydroxyl of S76 H-bonds the backbone carboxyl of A300 and the sidechain hydroxyl of T304 (Figure 4.19C). The backbone amide of T77 H-bonds the backbone carboxyl of S286, the sidechain hydroxyl of T79 H-bonds the sidechain hydroxyl of S286, and the sidechain hydroxyl of T79 coordinates with the backbone amide of S286 (Figure

4.19F). Additionally, L74 is 84% buried at the interface (Figure 4.19B), and M251 is 98% buried at the interface (Figure 4.19E). All hydrogen bonds are in the range of 2.9 – 3.8 Å, so these are mostly weak, low energy bonds.

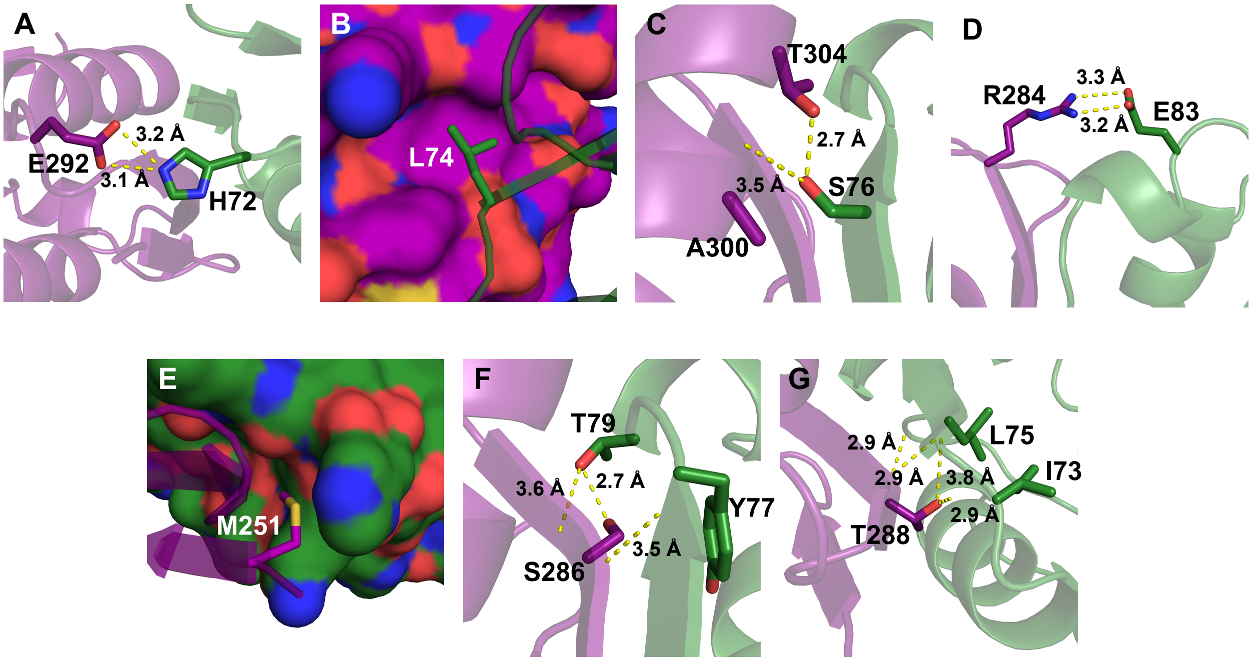


Figure 4.19. PupR-CCSSD:PupB-NTSD interface residues identified for site-directed mutagenesis. The following mutations were assessed: (A) E292H, H72D; (B) L74A; (C) S76L, T304A; (D) R284E, E83R; (E) M251A; (F) S286A, T79P; (G) T288A. Residues that only contributed backbone atoms to hydrogen bonding were not mutated. Not shown: Q69K.

4.3.4. Confirmation of the PupR-CCSSD:PupB-NTSD Interaction Interface

In order to verify that the crystallographic interface displayed in the asymmetric unit was the true interaction interface, several of the residues whose sidechains mediated salt bridges, H-bonds, or hydrophobic interactions were altered by site-directed mutagenesis (Figure 4.19). These residues and their corresponding mutations included PupB-NTSD residues Q69K, H72D, and L74A, as well as PupR-CCSSD residues M251A, S286A, and T288A. Residues Q69 and H72 were mutated to their corresponding residues of the PupA-NTSD. Alanine mutations were incorporated for the other residues identified at the PupR-CCSSD:PupB-NTSD interface. The association of PupB-NTSD and PupR-CCSSD was assessed by affinity pull-down, as previously described (Chapter 3.3.6), and verified by Western blotting (Figure 4.20). Of these mutations, H72D, L74A, and M251A clearly disrupt the interaction between the two domains (Figure 4.20). Mutations of S286A and T288A appear to limit, but not completely abrogate, the interaction (Figure 4.20).

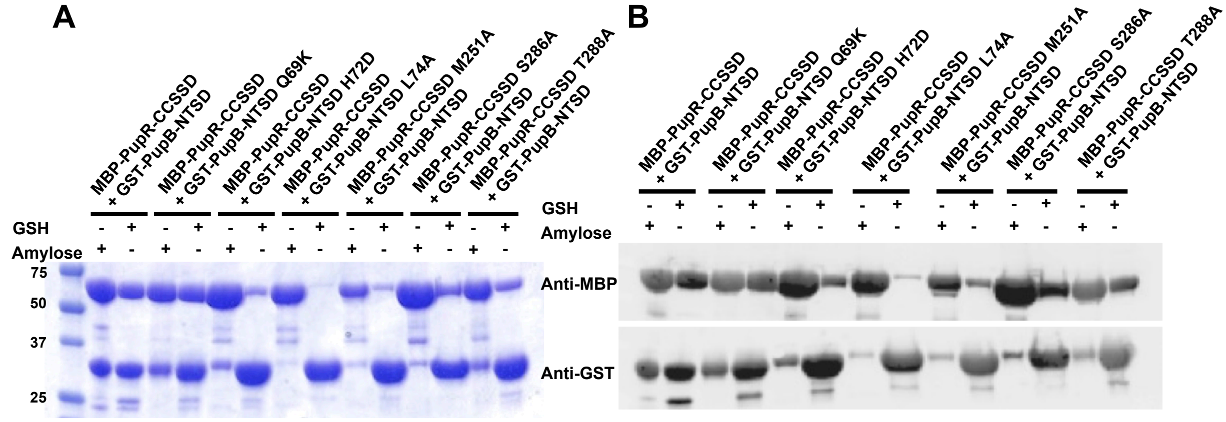


Figure 4.20. SDS-PAGE and Western blotting of affinity pull-downs with mutants of GST-PupB-NTSD and MBP-PupR-CCSSD. (A) Affinity pull-downs over either GSH or amylose resin analyzed by SDS-PAGE and confirmed by Western blotting (B).

In the structure solved by SAD, the E292 sidechain is replaced with a water molecule, which is within hydrogen bonding distance of the N1 of the H72 imidazole ring, the backbone amide of L291, the backbone amide of E292, and the backbone carboxyl of G250. The closest oxygen of the glutamate sidechain, OE2, is 4.7 Å from N1 of H72. The orientation of the E292 sidechain was modeled according to best-fit rotamer, as the sidechain atoms were not well defined in the electron density and low σ. The placement of the E292 sidechain and water molecule in the structure solved from the native data set correlate well with electron density when contoured at <1 σ. Thus, the H72-E292 interaction is not maintained between the native and the SeMet-derived X-ray crystal structures, suggesting some flexibility of the E292 sidechain.

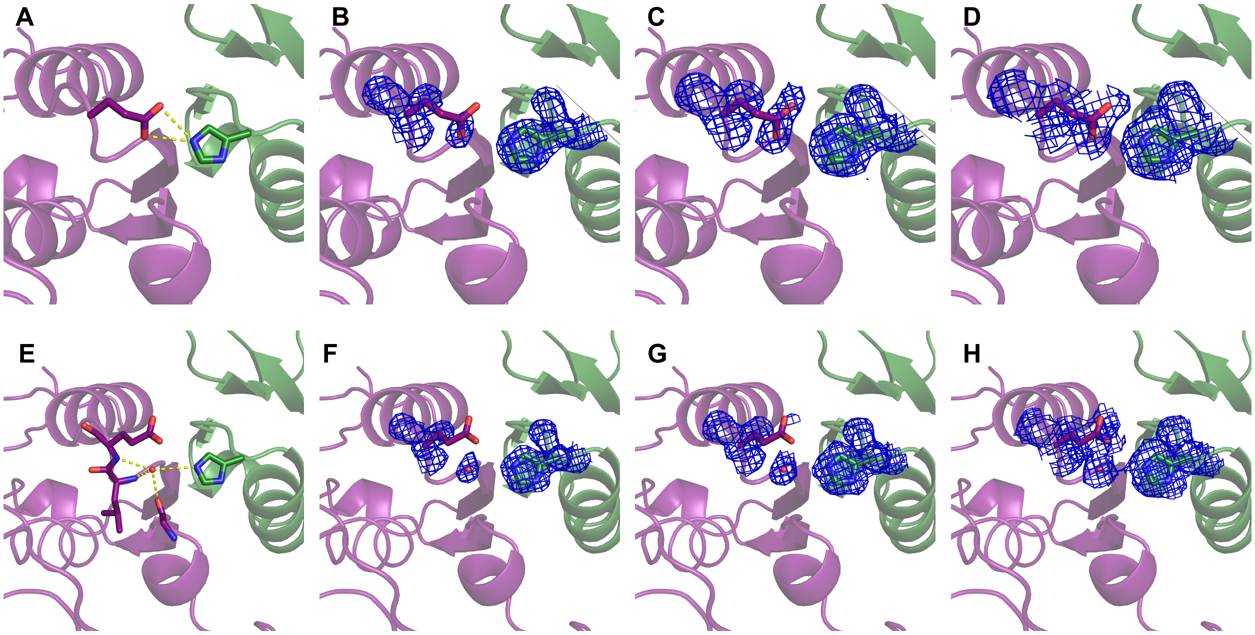


Figure 4.21. Differences in residue coordination of PupR-CCSSD E292 between the PupR-CCSSD:PupB-NTSD complex structures solved from the native or SeMet data. (A) Interaction of H72 and E292 from the structure solved from the native data set, contoured at (B) 1 σ, (C) 0.5 σ, and (D) 0 σ. (E) Interaction of H72 with a water molecule in place of the E292 sidechain carboxyl from the structure solved from the SeMet data set, contoured at (F) 1 σ, (G) 0.5 σ, and (H) 0 σ.

4.4. Summary and Conclusions

The structure of the PupB-NTSD:PupR-CCSSD complex resolves many questions concerning the interaction of these two domains, and refutes much of the canonical data and proposed mechanisms for interaction. While only the STN subdomain of the CCSSD is required for interaction with the NTSD, the CJM subdomain of the CCSSD is essential for structural stabilization of the entire periplasmic region of PupR. The complete periplasmic region is well-folded and incorporates a common signaling motif for the STN subdomain, while the CJM domain represents a previously uncharacterized fold. The interaction between the NTSD and CCSSD is relatively minimal; nearly all of the residues previously identified as critical for interaction with homologous proteins, such as with the Fec proteins, do not map to the

NTSD:CCSSD interface of the crystal structure (81,139,282). Site-directed mutagenesis of several of the interface residues revealed that hydrophobic packing drives the formation of this interface, as mutation of L74 of the PupB-NTSD and M251 of the PupR-CCSSD abolished complex formation. Additionally, H72 of the PupB-NTSD, and S286 and T288 form weak electrostatic or hydrogen bonds to further stabilize the complex. FecR residues 237-317 were previously shown to be necessary and sufficient for binding to the FecA-NTSD (105,139). This is contrary to pull-down analysis of various PupR-CCSSD truncations with the PupB-NTSD, which demonstrated that only the full-length CCSSD interacted with the PupB-NTSD *in vitro*.

Based on the limited global conformational changes between the STN subdomain of the free structure of the putative anti-sigma factor (PDB 4M0N and 4M0H) and the PupB-NTSD:PupR-CCSSD complex, the LLLV motif serves only to stabilize the hydrophobic core of the STN subdomain; it does not directly mediate the interaction with the NTSD, as previously proposed (105). Helix 2 of the STN subdomain of the CCSSD partially contacts the NTSD; thus, it is unlikely that this region interacts with the inner membrane (136). The PupR-CCSSD, while highly flexible in solution, remains folded, which is unmistakable from the presented SEC-SAXS data. Interaction between the CCSSD and NTSD further stabilizes the CCSSD, and both domains are clearly well-folded, which excludes the notion of the wrapping mode for the association of the complex (186). Thermodynamic and kinetic experiments indicate that PupB-NTSD and PupR-CCSSD interact with moderate binding affinity, and intermediate exchange between the bound and unbound states.

Based on this novel structural information, a new model of signal transduction through the inner membrane sigma regulator emerges. The PupR-CCSSD, when not associated with the NTSD, is highly flexible, resulting in non-specific proteolysis by a site-1 protease, which does not initiate transcriptional activation. When the PupR-CCSSD:PupB-NTSD complex is formed, the stabilized PupR-CCSSD is primed for site-specific proteolysis by a site-1 protease such as Prc. This then initiates signal transfer by conformational rearrangement such that proteolysis by a site-2 protease, like RseP, may occur. This final proteolytic event releases the anti-sigma domain:sigma factor complex, thus initiating transcription of the *pupB*.

5. STRUCTURAL CHARACTERIZATION OF THE PUPR-ASD AND ITS IMPLICATIONS FOR A NOVEL REGULATORY MECHANISM

5.1. Introduction

The sigma regulators, including PupR, are predicted to consist of an N-terminal cytoplasmic region, a transmembrane helix, and a C-terminal periplasmic region (283). Involvement of a proteolytic cascade in signaling through the sigma regulator to release the anti-sigma domain (ASD) has been demonstrated with FecR and FpvR (109). The canonical regulated intramembrane proteolysis (RIP) model comprises sequential proteolysis of the intramembrane anti-sigma protein by an external site-1 protease, followed by cleavage of the anti-sigma protein within the membrane by a site-2 protease, which releases the anti-sigma:sigma factor complex (284). The RseA:σ^E RIP cascade from *E. coli* is the most well-studied representative of this system. Initially, the site-1 protease DegS catalyzes cleavage of the RseA periplasmic domain. This proteolysis step provides the substrate for the site-2 protease RseP, which proteolyzes the cytoplasmic domain of RseA to release the RseA-ASD:σ^E complex (285,286). Conversely, other anti-sigma-ASDs release their cognate sigma factors through conformational change in response to oxidative stress (RsrA), binding of cobalt or nickel (CnrY), or loss of zinc (ChrR, RslA) (142,287-290). In certain anti-sigma factors, such as FecR, FpvR, and FoxR, cytoplasmic or periplasmic proteolysis appears to be preceded by a conserved autoproteolytic event induced via an enzyme-independent N-O acyl rearrangement (161,240,241). However, the residues responsible for this occurrence are not conserved in PupR.

FecR is cleaved at residue 85 within or near the cytoplasmic membrane to release FecI (109). Binding of ferripyoverdine to FpvA results in complete proteolysis of FpvR. Proteolysis of FpvR activates and releases sigma factors PvdS and FpvI to initiate genes required for pyoverdine synthesis and transport. In the absence of pyoverdine, differential proteolysis of FpvR occurs, but the FpvR subfragments produced inhibit PvdS and FpvI function (160). MucP, an ortholog of the site-2 protease RseP, along with other uncharacterized proteases (including a site 1-like protease), is involved in the FpvR degradative pathway in both the presence and absence of pyoverdine (160). Given that targeted expression of regulatory genes and continual proteolysis in the presence and absence of signal is relatively energetically costly for the cell, it is not well understood how this mechanism benefits the bacterium, and is one of the fundamental questions that this work seeks to address. The exact recognition sequence of RseP remains uncharacterized; however, the protease appears to have wide substrate specificity (162). Analogous σ regulators FoxR and FiuR of *P. aeruginosa* are anti-sigma factors also possessing "pro-sigma" activity, meaning that their continued association with the cognate sigma factor is required for transcriptional activation, although FpvR of the same organism is a purely anti-sigma regulator. Similarly, in the Pup system of *P. capeferrum*, a PupR knockout decreased *pupB* expression through PupI activation of transcription by only 50% of that observed in the wild-type strain (283).

Release of sigma factors by structurally characterized ASDs has previously been shown to occur by two distinct mechanisms: either degradation of the ASD, or loss of affinity between the ASD and the sigma factor. The ASD of *E. coli*, RseA, is proteolytically degraded to release σ^E (167,291). The ASDs of *Rhodobacter sphaeroides*, ChrR, and *Mycobacterium tuberculosis*, RslA, coordinate Zn^{2+}, and loss of the metal ion induces a loss of affinity for the cognate sigma factors (142,289,292). RskA of *M. tuberculosis* loses affinity between the ASD and the sigma factor, σ^K, upon reduction of a disulfide bond in σ^K (293). However, in contrast to these sigma regulators, FecR, FpvR, and likely PupR, the ASDs remain associated with their sigma factor during activation of transcription. This last group of sigma regulators has not previously been structurally characterized.

Here, the X-ray crystal structure of the cytoplasmic anti-sigma domain of PupR was determined to 2.0 Å resolution. Size exclusion chromatography, small-angle X-ray scattering, and sedimentation velocity analytical ultracentrifugation all indicate that, in contrast to other ASDs, the PupR-ASD exists as a dimer in solution. Mutagenesis of residues at the dimer interface identified from the crystal structure disrupts dimerization and protein stability, as determined by SV-AUC and thermal denaturation circular dichroism spectroscopy. The X-ray crystal structure of the PupR cytoplasmic domain revealed a conserved domain architecture related to the anti-sigma domains (145). Several structures of anti-sigma domains in complex with their affiliated sigma factors have been solved (PDBs 2Z2S, 4NQW, 3HUG, 1OR7) (142,289,291,293), but the PupR-ASD structure is the first for an integral membrane sigma regulator in the absence of a sigma factor. Interestingly, despite the low sequence identity amongst the ASDs of known structure, each contains a conserved core 3-helix bundle, followed by a 4-10 residue loop. Additionally, the PupR-ASD structure displayed a dimeric interface, which was confirmed by small-angle X-ray scattering, analytical ultracentrifugation, and site-directed mutagenesis. An ASD dimer had previously never been reported.

5.2. Materials and Methods

5.2.1. Protein Expression and Purification of PupR-ASD

Bioinformatics analyses were used to identify residues 1-82 as the cytoplasmic domain of *Pseudomonas capeferrum* PupR (PupR-ASD). These residues were cloned between NcoI and XhoI sites of the pMBP-Parallel1 vector (243). Chemically competent *E. coli* BL21(DE3)pLysS cells were transformed by this pMBP-Parallel1-PupR-ASD plasmid via standard heat shock protocol. Transformed cells were grown at 37°C in LB medium supplemented with 100 μg/mL ampicillin to an optical density $OD_{600\ nm}$ of ~0.8, as measured using an Implen Nano-Photometer Classic (Implen, Inc.), and PupR-ASD expression was induced by the addition of 1 M isopropyl thio-β-D-galactoside (IPTG) to a final concentration of 0.5 mM. After induction, the culture temperature was maintained at 20°C for 18 h, and cells were harvested by centrifugation at 4000*g* for 30 min. Harvested cells were stored at -80 °C until use, to assist cell lysis during purification via a freeze-thaw cycle. For protein purification, 30 g of cells were suspended in lysis buffer (25 mM HEPES:HCl, pH 7.5, 150 mM NaCl) and then lysed with a Nano DeBEE homogenizer (BEE International), with 2-3 passes at 15000 psi. The crude extract was clarified by centrifugation at 20000 *g* for 45 min.

MBP-PupR-ASD was purified by amylose affinity chromatography coupled with on-column cleavage by TEV protease. The supernatant after lysis clarification was loaded onto the gravity column containing ~15 mL of amylose resin, and washed with wash buffer: 25 mM HEPES:HCl pH 7.5, 150 mM NaCl, 2 mM DTT. Recombinant TEV protease was added to the column in a 1:10 mass ratio, and incubated for 16 h at 4°C to remove the MBP tag. TEV proteolysis yielded an 86-residue product, containing PupR residues 1-82, with an N-terminal 4-residue cloning artifact (GAMG). The column was washed with wash buffer to elute the liberated PupR-ASD. These fractions were diluted with 25 mM Tris:HCl pH 7.8 to an NaCl concentration of ~50 mM, then loaded onto a MonoQ anion exchange column (10/100GL, GE) equilibrated with 25 mM Tris:HCl pH 7.8. Contaminants after amylose affinity bound to the MonoQ column, but due to the high pI of the PupR-ASD (pI = 8.3), PupR-ASD was eluted in the wash fractions. The PupR-ASD protein was pooled, concentrated with a Millipore centrifugal filter unit (MWCO = 3000 kD), and loaded onto a tandem HR 10/300 Superdex 200-Superdex

75 (GE Lifesciences) size exclusion columns. Isocratic elution was performed with 25 mM HEPES:HCl pH 7.5, 150 mM NaCl at a flow rate of 0.4 mL/min. Fractions containing PupR-ASD were pooled and concentrated to 15 mg/mL and stored at 4°C or flash-frozen as pellets in LN_2 and stored at -80°C until use. Protein yield was approximately 4 mg per 1 L of culture.

5.2.2. Preparation of Selenomethionine-derivatized PupR-ASD

Selenomethionine (SeMet) derivatized PupR-ASD was prepared following a modified protocol involving methionine synthesis suppression (244,245). *E. coli* BL21(DE3)pLysS cells transformed with pMBP-Parallel1-PupR-ASD were grown at 37°C to saturation in 3 mL LB medium with 100 µg/mL ampicillin, then transferred to pre-warmed M9 minimal media containing 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% w/v glucose, and 100 µg/mL ampicillin and incubated at 37°C. Once the $OD_{600\ nm}$ reached 1.0, the media was supplemented with SeMet (Acros Organics), Lys, Thr, Phe, Leu, Ile, and Val (RPI), and the temperature lowered to 20°C. IPTG was added to 0.5 mM, and the cultures grown for an additional 18 h. Purification of SeMet PupR-ASD was performed as described for the native protein. Approximately 3 mg of SeMet protein was obtained per 1 L of cells.

5.2.3. Site-Directed Mutagenesis of PupR-ASD

Aspartate-encoding point mutations of WT PupR-ASD were created using the QuikChange II protocol (Stratagene), with the pMBP-Parallel1-PupR-ASD vector as a template, and primers containing the desired aspartate mutation, using either of the two most common codons. The mutant gene sequences were verified by DNA sequencing (MCLab), and the plasmids used to transform *E. coli* Arctic Express (DE3) competent cells (Stratagene). The mutant proteins were prepared in the same manner as the WT, with the pH of the buffers maintained at 8.0 throughout purification.

5.2.4. Determination of Protein Purity and Concentration

The purity of the collected protein was estimated to be >90% by SDS-PAGE stained with Coomassie Blue (215). Protein concentration was determined by absorbance at 280 nm using the molar extinction coefficient ε_{280} = 16500 M^{-1} cm^{-1} and a theoretical molecular weight of 9502 g/mol. Incorporation of SeMet was confirmed by ESI mass spectrometry.

5.2.5. Crystallization of PupR-ASD

The PupR-ASD was concentrated to 12 mg/mL in SEC buffer prior to crystallization experiments. Crystallization screening was performed with the SaltRx HT, Crystal Screen HT, Index HT, PEGRx HT, and Additive Screen HT (Hampton Research), using an ARI Crystal Gryphon liquid handling robot (Art Robbins Instruments). Each screen contained 96 conditions, deposited into 3-well Intelli-plates (ARI); the three wells had ratios of protein:reservoir solution- 3:1, 1:1, and 1:3, respectively, for a total of 384 conditions. Each well was equilibrated against 50 μL of reservoir solution. Conditions conducive to initial crystal growth were identified in multiple wells (Table 5.1).

Table 5.1. Favorable crystallization conditions for PupR-ASD identified from various Hampton protein crystallization screens.

Screen ID	Salt	Buffer and pH	Precipitant	Ratio (protein:reservoir)
Index HT	3.0 M NaCl	0.1 M Bis-Tris pH 5.5	None	1:1
Index HT	3.0 M NaCl	0.1 M HEPES pH 7.5	None	1:1
Index HT	3.0 M NaCl	0.1 M Tris pH 8.5	None	1:1
SaltRx HT	2.2 M NaCl	0.1 M Na acetate trihydrate pH 4.6	None	3:1
SaltRx HT	3.0 M NaCl	0.1 M HEPES pH 7.5	None	3:1, 1:1, 1:3
SaltRx HT	3.0 M NaCl	0.1 M Tris pH 8.5	None	3:1, 1:1, 1:3

Subsequent manual screening yielded crystals from conditions containing high salt and buffers with a pH around 6.0. Reproducible needle-like clusters were obtained via sitting-drop vapor diffusion by combining 1 μL protein solution with 1 μL reservoir solution (100 mM Bis-Tris pH 6.0, 3 M NaCl) and equilibrating against 500 μL reservoir solution at 20°C. These crystals were used for microseeding fresh droplets of 2 μL protein:reservoir by serial dilution and resulted in the growth of larger crystal clusters (Figure 5.1). Crystals initially appeared 2-3 days after microseeding, and reached final dimensions after approximately 2 weeks. Larger clusters (Figure 5.1) were broken into ~200 μm x 20 μm x 20 μm pieces, and cryo-protected by repeated passing through MiTeGen LV CryoOil (MiTeGen) prior to manual mounting with 0.1-0.2 mm MiTeGen MicroLoops™ or Hampton Mounted CryoLoops™ with stainless-steel rods and reusable goniometer bases, and flash-frozen in LN_2. Crystals of SeMet PupR-ASD were grown and harvested in a similar manner.

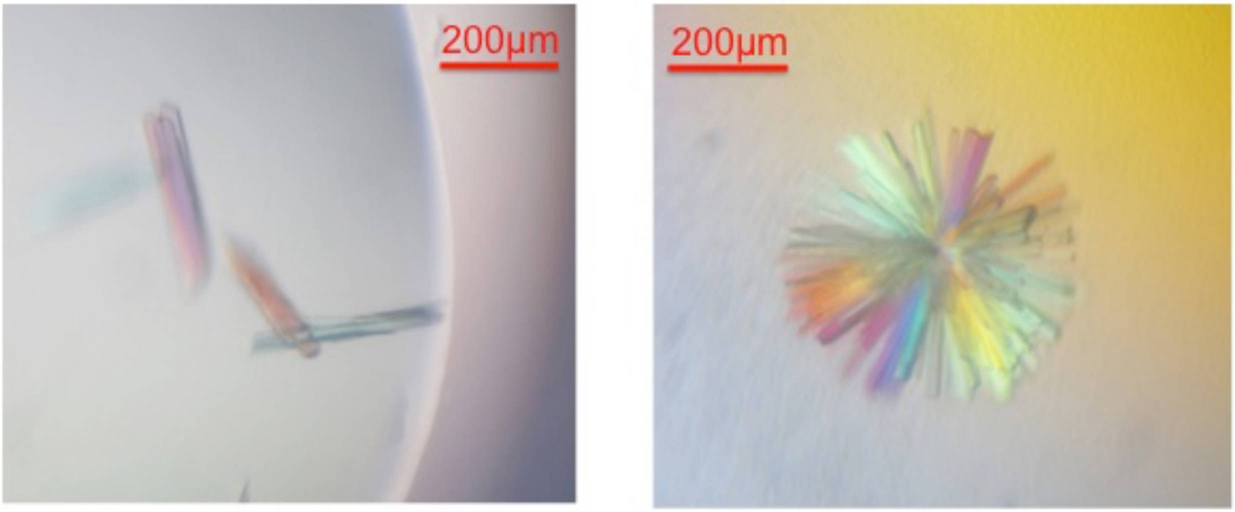


Figure 5.1. Single crystals and crystal clusters of selenomethionine-derived PupR-ASD.

5.2.6. Diffraction Measurements and X-ray Structure Determination

Initial diffraction experiments were performed in the NDSU Materials Characterization Laboratory using Cu-Kα radiation from a Bruker Kappa Apex II Duo sealed tube X-ray generator with a low-powered Cu IµS microfocus source and an Apex II CCD detector. Crystals were maintained at cryogenic conditions (~100 K) during diffraction screening with the use of an LN_2 cryostream. Diffraction were collected with a crystal-to-detector distance of 75 mm, with 1° oscillation per 5 min exposure. Intensities indicative of diffraction by the protein crystal were detected, although the observed lattice was very weak, with the maximum resolution at ~5 Å.

As the fold of the PupR-ASD was unknown, and a search of the Protein Data Bank did not reveal any structures with sequence identity greater than 20%, initial phasing was attempted by soaking crystals in Bis-Tris buffer with 3.0 M BrCl prior to mounting (294,295). Although the integrity of the crystals was maintained and a highly complete data set was collected at 0.92 Å (the K absorption edge for bromine) to a resolution of 2.0 Å, the anomalous signal could not be distinguished from the noise of the disordered solvent region (Table 5.2); these crystals were screened at GMCA-CAT, beamline 23-ID-D, APS. To overcome this complication, selenomethionine-containing PupR-ASD was expressed in selenomethionine-supplemented minimal media, and purified as previously described.

Table 5.2. X-ray data collection from the PupR-ASD Br soak experiment.

	Br Soak
Data collection	
Beamline	APS 23-ID-D
Wavelength (Å)	0.9792
Space group	$P2_1$
Unit-cell parameters (Å, deg)	33.5, 135.3, 34.9
	90, 113.5, 90
Resolution range (Å)	32.0-2.17 (2.25-2.17)
Unique observations	14850
Multiplicity	6.1 (2.1)
Completeness (%)	98.7 (94.6)
Mean *I*/σ*I*	16.1 (2.8)
Data processing program	HKL2000

Values in parentheses pertain to the highest resolution shell.

Diffraction data were collected at NE-CAT beamlines 24-ID-E and 24-ID-C at the Advanced Photon Source (APS), Argonne National Laboratory, Argonne, IL. All data sets were processed using the RAPD automated processing suite (https://rapd.nec.aps.anl.gov/rapd), which incorporates XDS for integration and scaling (254). The SeMet PupR-ASD structure was determined by single-wavelength anomalous dispersion (SAD) phasing using intensities measured at 12.66 keV. Three of the five selenium sites per monomer (four in the PupR-ASD sequence, and one from the N-terminal cloning artifacts) were located, and initial phasing performed with AutoSol in PHENIX (258). These phases were used to calculate initial electron density maps for automated model building with AutoBuild (259). Refinement was carried out in PHENIX (260) and iterative manual modeling building in COOT (261). The SeMet PupR-ASD structure was subsequently used as a molecular replacement (MR) model to phase a 2.0 Å data set of native PupR-ASD. MR was performed with Phaser-MR (262), followed by AutoBuild (259) and refinement in PHENIX, using automated TLS group determination (263). Water oxygen atoms were positioned, with subsequent visual verification. Data collection and refinement statistics are summarized in Table 2. All figures were prepared using PyMOL v.1.5.0.4 (Schrödinger). Analysis of surface areas, protein interfaces, assemblies, and interactions were determined using the PISA server (http://www.ebi.ac.uk/pdbe/pisa) (264). RMSD comparisons were carried out in PyMOL. Atomic models and structure factors have been deposited into the Protein Data Bank (http://www.rcsb.org) under PDB entries 5CAM and 5COS.

Table 5.3. X-ray data collection, phasing, and refinement statistics for the PupR-ASD.[a]

	Native	**SeMet derivative**
Data collection		
Beamline	APS 24-ID-C	APS 24-ID-E
Wavelength (Å)	0.9792	0.9792
Space group	$P2_1$	$P2_1$
Unit-cell parameters (Å, deg)	33.5, 134.6, 34.9 90, 113.7, 90	33.5, 134.7, 34.9 90, 113.8, 90
Resolution range (Å)	44.87-2.02 (2.13-2.02)	134.7-2.17 (2.29-2.17)
Total observations	62331 (9246)	99992 (5749)
Unique observations	18320 (2685)	14546 (1796)
Multiplicity	3.4 (3.4)	6.9 (3.2)
Completeness (%)	98.8 (99.4)	97.2 (82.0)
CC(1/2)	0.997 (0.913)	0.997 (0.879)
R_{merge}[b] (%)	4.5 (44.2)	11.3 (71.1)
R_{merge} (anom, %)		6.6 (59.4)
Mean $I/\sigma I$	13.3 (2.4)	18.2 (1.7)
Data processing program	RAPD	RAPD
Refinement		
Refinement program	PHENIX	PHENIX
Resolution range (Å)	33.66-2.02 (2.09-2.02)	67.35-2.17 (2.25-2.17)
Molecules per asymmetric unit	4	4
R_{work} (%)	21.6	17.8
R_{free} (%) (5%)	25.7	21.7
RMSD stereochemistry		
Bond lengths (Å)	0.010	0.018
Bond angles (deg)	1.19	1.51
No. of atoms	2023	2002
PupR-ASD	1920	1906
Waters	103	96
Average *B* (Å^2)		
PupR-ASD	51.9	55.0
Waters	49.7	45.6
Ramachandran plot (%)		
Preferred	99	99
Allowed	1	1
Outliers	0	0
PDB code	5COS	5CAM

[a] Values in parentheses pertain to the highest resolution shell.

[b] $R_{merge} = \frac{\sum_{hkl}\sum_j \left| I_{hkl,j} - \langle I_{hkl} \rangle \right|}{\sum_{hkl}\sum_j I_{hkl,j}}$

5.2.7. Size Exclusion Chromatography SAXS Data Collection and Analysis

SEC-SAXS was performed on the WT PupR-ASD at Bio-CAT beamline 18-ID of the APS. An inline Superdex 75 10/300 column (GE Lifesciences) was equilibrated with 25 mM Tris pH 7.8, 150 mM NaCl at a flow rate of 0.7 mL/min. SEC-SAXS data were recorded at a wavelength of 1.033 Å on a Pilatus 3 1 M detector at a sample-to-detector distance of 3.0 m, covering a momentum transfer range of 0.0055 Å $< q <$ 0.3316 Å^{-1}. Data were normalized to the incident X-ray beam intensity, and buffer-subtracted prior to analysis using the ATSAS program suite (198). Linearity of the Guinier region was examined using PRIMUS (193) and the Kratky plot calculated to assess disorder of the molecule in solution. GNOM (199,296) was used to compute the radius of gyration, R_g; pairwise distribution function, $P(r)$; and the maximum particle dimension, D_{max} (199,296). Ten independent *ab initio* bead model reconstructions were calculated using DAMMIF (196) and enforcing *P2* symmetry. The resultant dummy atom models were averaged and filtered with DAMAVER (297) with default parameters. The models had an average χ^2 of 1.2 ± 0.014, with an averaged normalized spatial discrepancy (NSD) for the 10 DAMMIF calculations of 0.85 ± 0.19. The resultant three-dimensional averaged envelope was superimposed upon the PupR-ASD crystal structure using SITUS and the UCSF Chimera package (201,202,298). CRYSOL (211) was used to compare theoretical scattering curves calculated from monomer and dimer atomic models against the experimental SAXS data. PyMOL was used to visually evaluate fits of the PupR-ASD and for manual docking. The molecular weight of the PupR-ASD was determined with the SAXS MoW server (http://www.if.sc.usp.br/~saxs/) (299). Theoretical scattering profiles calculated for the PupR-ASD monomer and dimer X-ray structures were fit to the experimental scattering profiles using the FoxS server (http://www.modbase.compbio.ucsf.edu/foxs/) (300,301). EOM 2.0 (268,302) in the ATSAS suite was employed to model the flexible N- and C-termini of the PupR-ASD. The full PupR-ASD sequence, along with either the monomer or dimer crystal structure, was used to generate 10,000 random conformers, which were subsequently pooled by EOM to generate ensembles that best fit the experimental data. The theoretical scattering profiles of the best models, as identified by the program, were fit to the experimental scattering intensities with the FoxS server.

5.2.8. Analytical Ultracentrifugation of WT and Mutant PupR-ASDs

All sedimentation velocity analytical ultracentrifugation (SV-AUC) experiments and data analysis were conducted by Andrea Balbo and Huaying (Joy) Zhao of the Biomedical Engineering and Physical Science Shared Resource and the Dynamics of Macromolecular Assembly Section of the Laboratory of Cellular Imaging and Macromolecular Biophysics, respectively, of the National Institute of Biomedical Imaging and Bioengineering at the National Institutes of Health, Bethesda, Maryland.

The SV-AUC experiments were performed in a Beckman ProteomLab XL-I analytical ultracentrifuge (Beckman Coulter) at 20 °C following standard protocol (303) and boundary profile analysis (304). PupR-ASD samples in a buffer of 25 mM HEPES pH 8.0, 150 mM NaCl were loaded into a sample cell assembly containing a standard double-sector charcoal-filled Epon centerpiece with a 12 mm or 3 mm path length and sapphire windows. Sample cell assemblies loaded in a rotor were equilibrated at 20 °C for at least 2 h prior to sedimentation. Sedimentation profiles were acquired using absorbance optics at 280 nm in combination with interference optics at 20 °C at 201,240 *g* (rotor speed of 50,000 rpm at a radius of 7.2 cm).

Various concentrations were explored for each protein construct to populate monomer and dimer species. For WT PupR-ASD, seven samples were prepared with concentrations that ranged from 1.1 to 102 μM. Mutant PupR-ASD proteins were prepared at a few concentrations in this range. The monomer-dimer self-association of each sample was characterized using previously reported procedures (303). Briefly, a standard c*(s)* model (304) was used to analyze the sedimentation profiles. Precise concentrations were determined using refractive index increments, *dn/dc*, predicted from the amino acid sequence of PupR-ASD in SEDFIT (305) and subsequent integration of the *c(s)* peaks from the interference data, resulting in the total protein concentration in mg/mL corresponding to the integrated peaks in the *c(s)* distributions. Signal weighted-average *s*-value (s_w) isotherms (306) were created by integration of the c*(s)* peaks in the distribution between 0 and 4 S for the data and plotted as a function of protein concentration in SEDPHAT (306). A monomer-dimer self-association model was applied to the s_w isotherms in SEDPHAT (307). For WT PupR-ASD, the *s*-value of the monomer was fixed at 1.06 S, based on the value determined from the mutants, which do not form dimers under the concentration range tested; *s*-value of the dimer was fitted along with the binding constant (K_d). For the mutants, *s*-value of monomer and dimer were fixed at 1.06 and 1.71 S (determined from the WT s_w isotherm) respectively; thus, the only fitted parameter is the K_d. All experimental SV data and best-fit values are presented in units of experimental *s*-values. Error analysis of K_d for WT was performed using the error surface projection method and F-statistics (308) and the 95% confidence intervals were reported.

5.2.9. Circular Dichroism Spectroscopy and Thermal Denaturation

The PupR-ASD was diluted to 10 μM (0.095 mg/mL) in 10 mM potassium phosphate, pH 8.0, 100 mM $(NH_4)_2SO_4$. Continuous scanning CD spectra were measured at 4 °C between 180 and 250 nm using a Jasco J-815 spectrometer with a PFD-425S Peltier cell holder and a 1 mm quartz cell. The spectra were buffer subtracted, and the secondary structure content estimated using K2D2 (309); as well as CONTIN, CDSSTR, and SELCON, provided in CDPro (236). Additionally, WT PupR-ASD was diluted to 10 μM with CD buffer and 0%, 10%, 25%, and 40% 2,2,2-trifluoroethanol (TFE), and incubated on ice for 30 min. CD spectra were measured and analyzed as previously described.

CD melting and re-folding curves were recorded at 222 nm with 10 μM WT or mutant PupR-ASD protein in CD buffer by increasing the temperature from 4 to 85 °C in 1 °C increments with a slope of 1 °C/min. Continuous scanning CD indicated that the PupR-ASD was primarily helical in solution; thus, heating and cooling were monitored at 222 nm. Protein unfolding was monitored during both heating and cooling. Complete CD spectra at 4 °C were recorded just prior to each experiment. Melting temperatures were determined by fitting a standard Boltzmann sigmoidal curve to the molar ellipticity in Origin 8 (OriginLab Corp., Northampton, MA). The final melting temperature was defined as the inflection point after fitting.

5.3. Results

5.3.1. The Tertiary Structure of the PupR-ASD is Conserved amongst ASDs

The PupR-ASD crystals belonged to space group $P2_1$ with unit cell parameters of a = 33.5 Å, b = 113.8 Å, c = 34.0 Å, and β = 113.7 ° (Table 2). Crystals of native PupR-ASD and the SeMet derivative diffracted to 2.0 and 2.2 Å, respectively. SeMet SAD data were used to locate 12 of the possible 20 selenium sites, enabling the placement of four monomers per crystal asymmetric unit (Figure 5.2A) This initial PupR-ASD model was refined to 2.2 Å model. A monomer from the SeMet structure was used as an MR search model to determine the structure from the native data set at 2.0 Å resolution. R_{work} converged to 17.8% and R_{free} to 21.7%. No Ramachandran outliers are present in the structure (Table 5.3).

The final model at 2.0 Å resolution is comprised of residues 10-65 of chain A, 11-64 of chain B, 8-66 of chain C, and 5-65 of chain D; electron density corresponding to the N- and C-termini of each monomer was not resolved. The structure of the PupR-ASD monomer is a tightly packed, three-helix bundle (Figure 5.2B). The three-helix core is stabilized by hydrophobic interactions, primarily between helices 2 and 3, including π-π stacking of W40 and H47 (plane-plane distance = 4.0 Å) (Figure 5.2B). Residues analogous to W40, W20, and W51 are invariant in homologous putative iron transport anti-sigma regulator protein N-terminal domains across bacterial species (Figure 5.3).

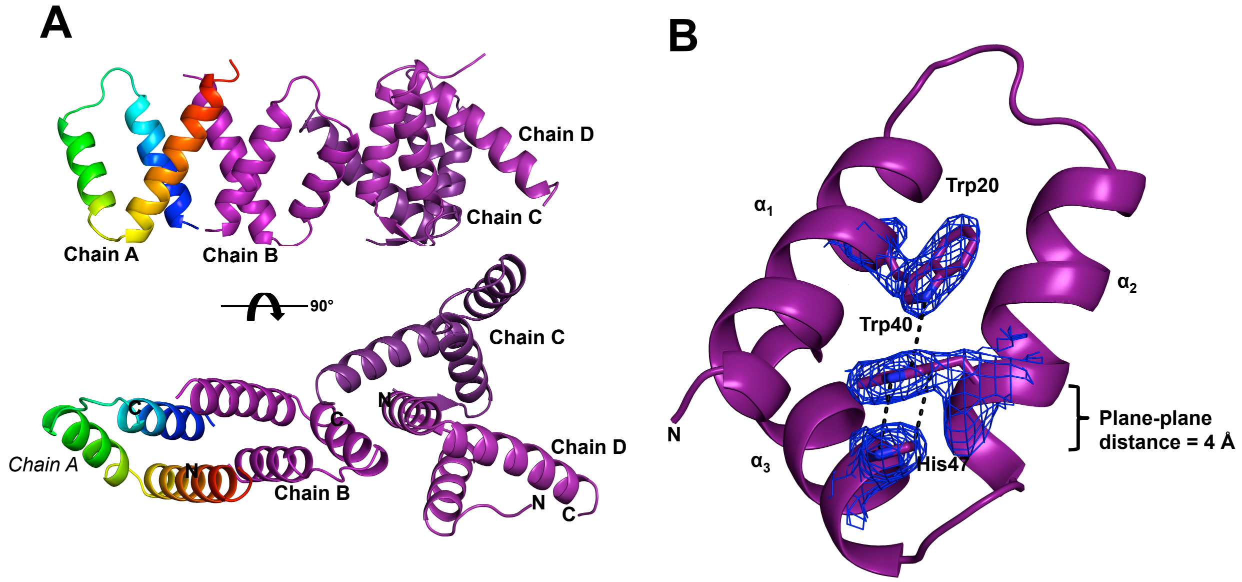


Figure 5.2. Asymmetric unit of the PupR-ASD crystal. (A) The four PupR-ASD monomers in the asymmetric unit are displayed in ribbon. Monomer A is rendered in rainbow colors with blue at the N-terminus and red at the C-terminus, with monomer B in purple. Monomers C and D are displayed in shades of grey. The symmetric dimer comprises chains A and B; chains C and D are related by noncrystallographic symmetry to each other and to chains A and B. (B) Electron density (blue mesh) for the conserved aromatic residues involved in stabilizing the PupR-ASD core. Helices 1-3 of the PupR-ASD monomer form the core of the ASD, with off-centered parallel π-π stacking of His47 and Trp40, which is further stabilized by Trp20. The ASD backbone is displayed as a purple ribbon, with aromatic residues in stick, color-coded by atom type: N, blue; O, red; S, yellow; and C, magenta.

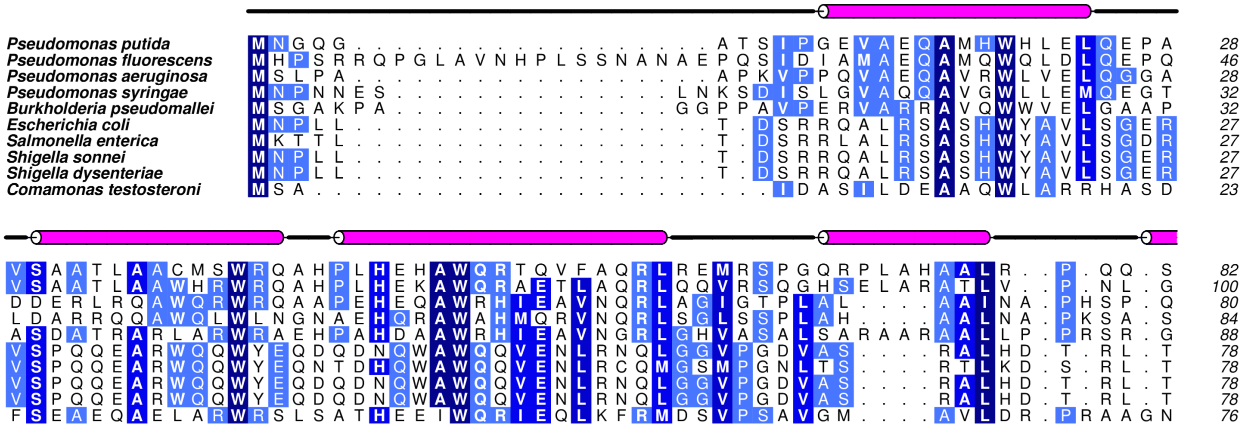


Figure 5.3. Sequence alignment and secondary structure predictions of ASDs from various proteobacteria with putative iron transport TBDTs with cognate anti-sigma regulators. The ASDs of PupR (Q52209; *Pseudomonas capeferrum*) were aligned with PupR homologs (Q4KDP8; *Pseudomonas fluorescens*; Q48ML1, *Pseudomonas syringae*), FpvR (Q91192; *Pseudomonas aeruginosa*), FecR (P23485; *Escherichia coli*), and FecR homologs (Q63LI7, *Burkholderia pseudomallei*; V2K8B2, *Salmonella enterica*; Q3YU71, *Shigella sonnei*; B3WZ95, *Shigella dysenteriae*; H1RKR8, *Comamonas testosterone*). Sequence alignment was performed using T-COFFEE and edited with Aline. Invariant residues are highlighted with navy blue, and conserved residues in differing shades of blue, with the lightest shade corresponding to the least conserved residues. Secondary structure prediction obtained with Psipred is displayed above the primary sequence, with red cylinders denoting helices and black lines denoting coils.

Despite sharing only 7-16% sequence identity, the PupR-ASD and other known ASDs share a common structural fold comprising a three-helix core bundle (Figure 5.4). PupR-ASD helices 1-3 superimpose with root-mean-square deviations (RMSD) of 2.52-3.18 Å over 40-48 residues of the ChrR-ASD from *R. sphaeroides* (142) (PDB ID 2Z2S), RseA from *E. coli* (291) (PDB ID 1OR7), RskA from *M. tuberculosis* (293) (PDB ID 4NQW), and RslA from *M. tuberculosis* (289) (PDB ID 3HUG). ChrR and RslA are both zinc-binding anti-sigma factors, with their tertiary structure stabilized by a Zn^{2+}, which is coordinated by a conserved His-X_3-Cys-X_2-Cys motif. Loss of the Zn^{2+} results in loss of structure and anti-sigma function. In contrast, the PupR-ASD, RseA-ASD, and RskA-ASD comprise another group of ASDs that are stabilized by hydrophobic cores, although the PupR-ASD is the only ASD of known structure that stabilizes its hydrophobic core via π-π stacking of aromatic residue side chains- two tryptophans and one histidine (Figure 5.2B). The four previously determined ASD structures also include a 4 to 10 residue loop preceding a fourth, C-terminal helix, whose location varies among the assorted ASDs (Figure 5.4). Interestingly, PupR residues 68-82, which are analogous to this region, are disordered in the PupR-ASD structure. CD spectra collected with increasing amounts of TFE added to the WT PupR-ASD demonstrated an increase in helical content, indicating that the region containing the fourth helix may only be ordered in the presence of a binding partner (Table 5.4, Figure 5.5) (310-312).

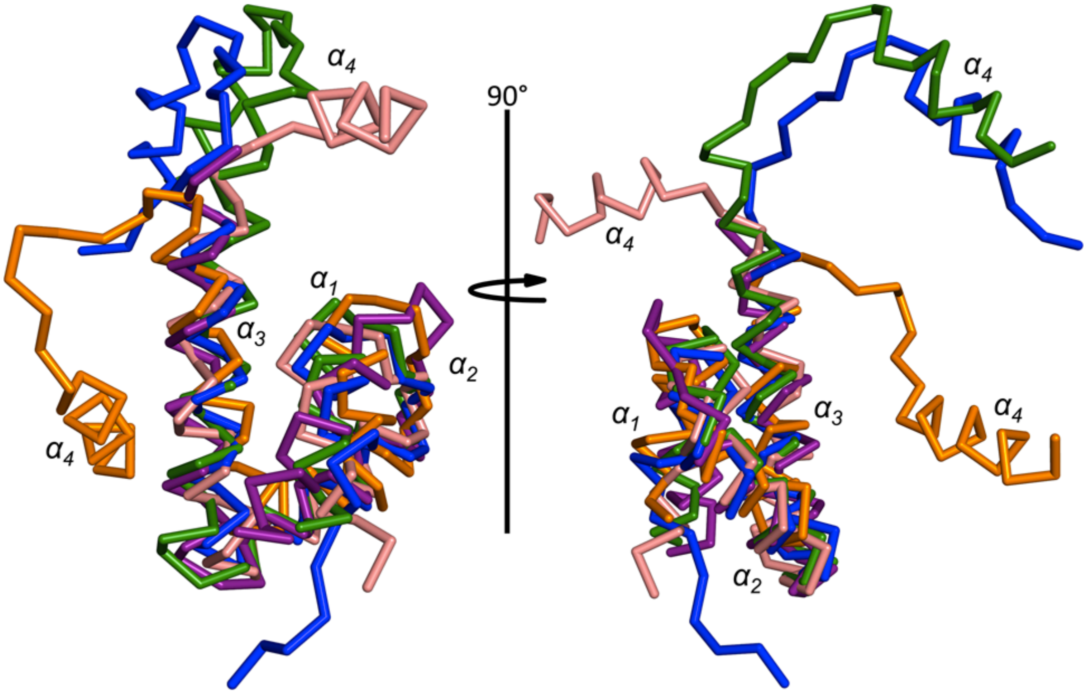


Figure 5.4. Known ASD structures contain a conserved core helical bundle. Helices 1-3 of *R. sphaeroides* ChrR (blue; PDB entry 2Z2S; Campbell, *et al.,* 2007), *M. tuberculosis* RskA (green; PDB entry 4NQW; Shukla, e*t al.*, 2014), *M. tuberculosis* RslA (salmon; PDB entry 3HUG; Thakur, e*t al.*, 2010), and *E. coli* RseA (orange; PDB entry 1OR7; Campbell, *et al.*, 2003) were superimposed upon the PupR-ASD (purple). The orientation of the fourth helix varies amongst ChrR, RskA, RslA, and RseA.

Table 5.4. Estimated secondary structure content of PupR-ASD WT with TFE. The helical content of the ASD based on the X-ray crystal structure is 56%; assuming an order 4th helix, the helical content would be 67%.

Sample	Helix	Strand	Coil	Total
0% TFE	0.54	0.09	0.36	0.99
10% TFE	0.60	0.07	0.32	0.99
25% TFE	0.56	0.09	0.35	1.00
40% TFE	0.67	0.04	0.28	0.99

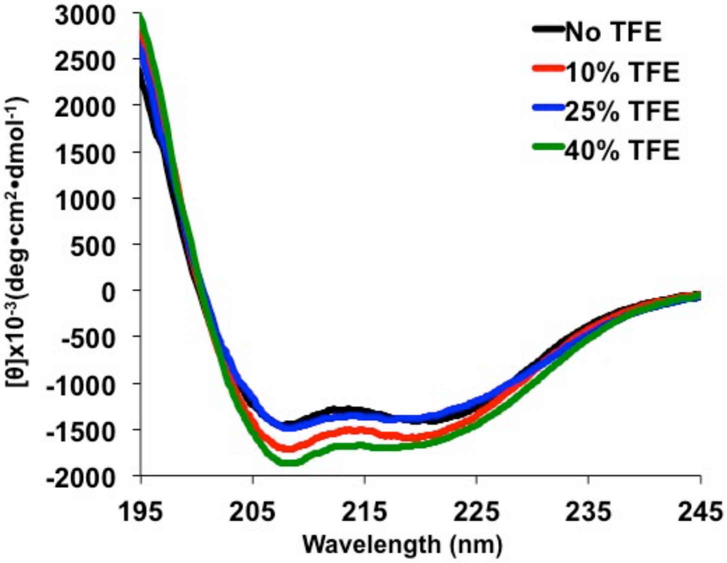


Figure 5.5. CD spectra of PupR-ASD WT with increasing concentrations of TFE.

5.3.2. The PupR-ASD Forms a Dimer in Solution

The relative molecular weight and oligomeric state of the native PupR-ASD was determined by analytical size-exclusion chromatography during purification of the domain. The relative molecular weight of the PupR-ASD subunit as estimated by SDS-PAGE is approximately 9000 kD, consistent with the expected molecular weight of 9502 Da. However, the PupR-ASD eluted from the size-exclusion column as a single Gaussian-shaped peak with an apparent molecular mass of 19 kD, suggesting that the PupR-ASD purified as a homodimer (Figure 5.6A-C).

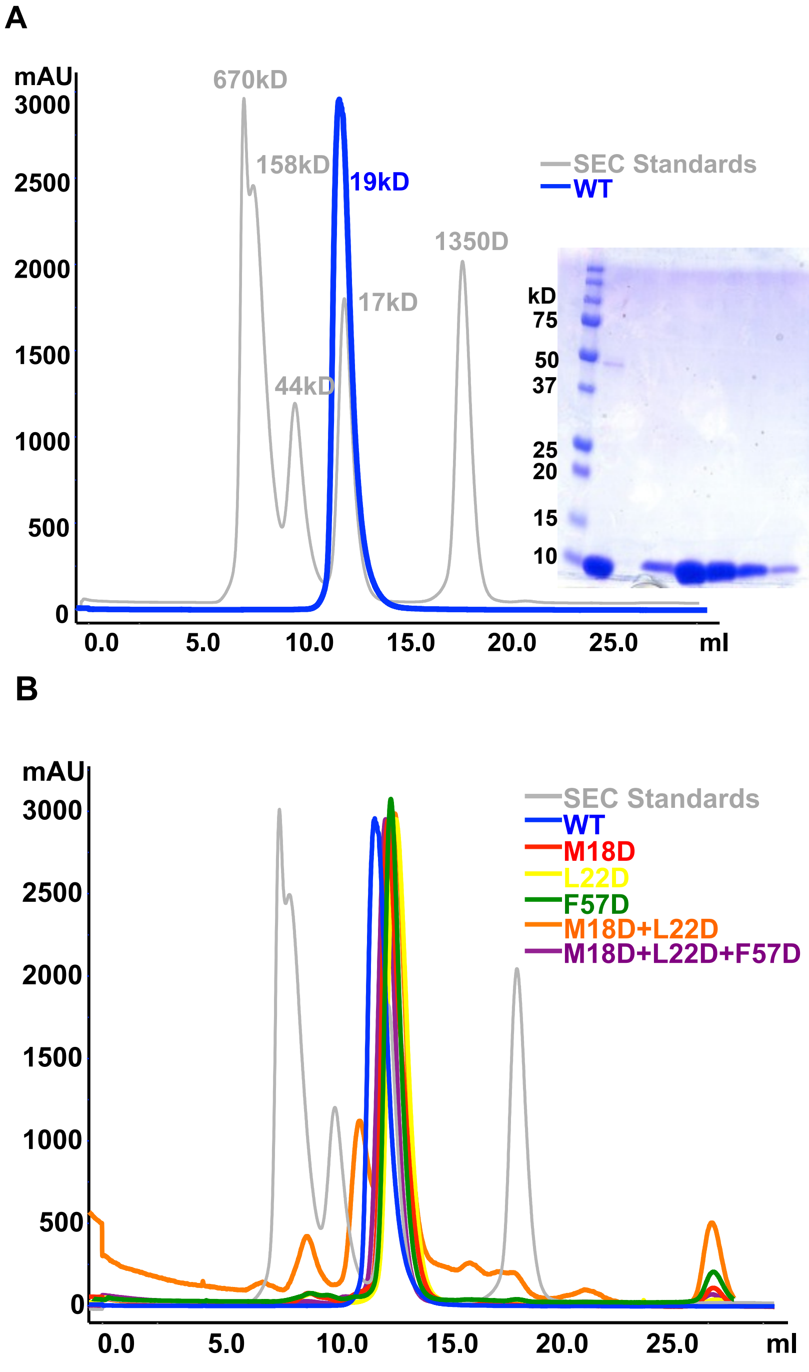


Figure 5.6. Size exclusion chromatograms of WT and mutant PupR-ASD. (A) SEC chromatogram and corresponding SDS-PAGE of WT PupR-ASD. (B-C) Comparison of SEC chromatograms of WT PupR-ASD with interface mutants.

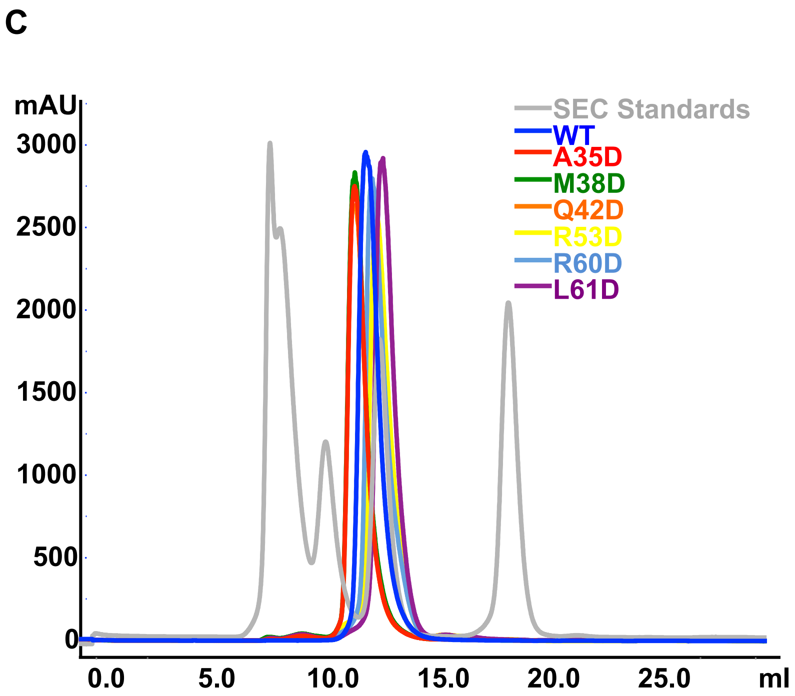


Figure 5.6. Size exclusion chromatograms of WT and mutant PupR-ASD (continued). (A) SEC chromatogram and corresponding SDS-PAGE of WT PupR-ASD. (B-C) Comparison of SEC chromatograms of WT PupR-ASD with interface mutants.

Small-angle X-ray scattering in conjunction with size-exclusion chromatography (SEC-SAXS) was performed on the purified PupR-ASD to characterize its solution state (Figure 5.8a). The Guinier plot was linear in the $0 < q < 0.1$ Å^2 range (Figure 5.8b), verifying that the sample was monodisperse. The Guinier plot indicated that the PupR-ASD has an $R_g = 18.7 \pm 0.02$ Å, with the *P(r)* analysis yielding a D_{max} corresponding to 64.5 Å (Figure 5.8c). The Kratky plot analysis (Figure 5.8d) revealed that the PupR-ASD is partially unfolded, consistent with the X-ray structure, wherein the 11 N-terminal and 18 C-terminal residues are not observed in the electron density map, and are therefore presumed to be disordered. An estimate of the PupR-ASD molecular weight was calculated from the scattering data with the maximum *q* set to 0.15 Å^{-1}. This calculation indicated the species in solution had a molecular weight of 19.8 kD, demonstrating that the dimer was the dominant species in the SEC-SAXS sample. Thus, the SEC and the SAXS results are consistent with, and indicative of, the PupR-ASD existing as a homodimer in solution.

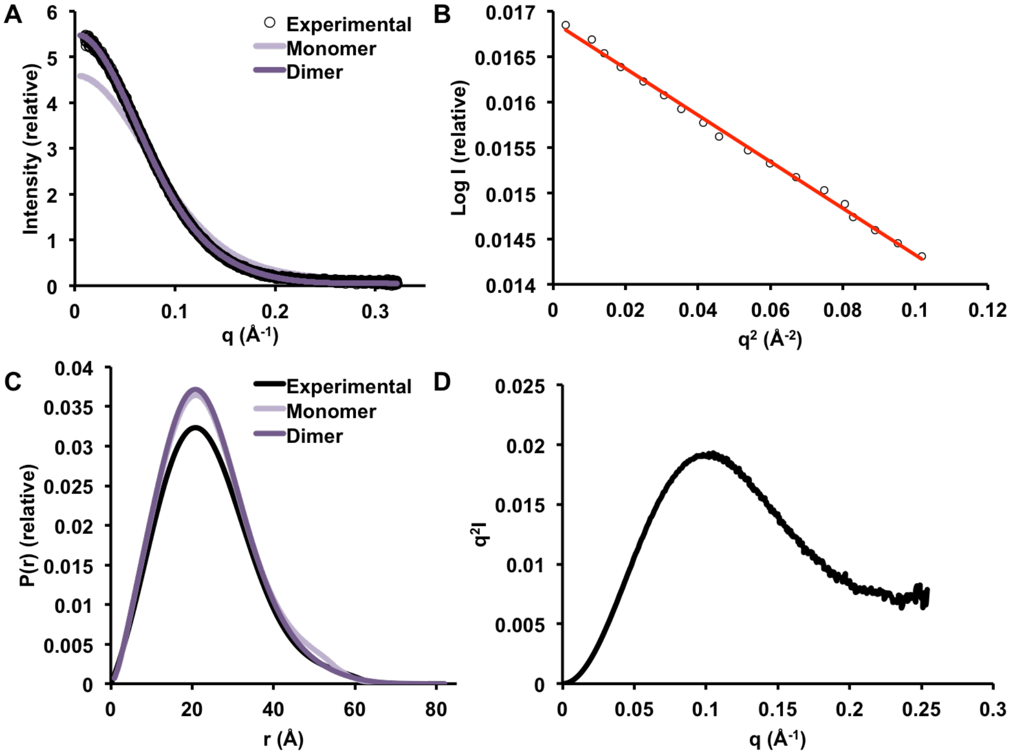


Figure 5.7. SEC-SAXS analysis of the PupR-ASD. (A) Experimental scattering profile of the PupR-ASD (black) and the theoretical scattering profiles of the PupR-ASD dimer (green) and monomer (red) calculated from the X-ray crystal structure. (B) Guinier plot of the low *q* region of the X-ray scattering data. (C) Distance distribution *P(r)* plot for the experimental data (black) and the theoretical curves calculated for the PupR-ASD monomer (red) and dimer (green). (D) Kratky plot calculated from the experimental scattering profile.

SV-AUC was utilized to quantify PupR-ASD self-association. Sedimentation coefficient distributions, *c(s)* of WT were obtained for SV data at seven different concentrations, ranging from 1.1 µM to 101.9 µM, of WT PupR-ASD. The monomer had a sedimentation coefficient of 1.06S, and the dimer, ~1.71 S (Figure 5.8A). An s_w isotherm was then calculated from the dilution series and fit with a monomer-dimer equilibrium binding model (Figure 5.8B), to obtain a K_d = 10.2 µM (95% confidence interval [7.58, 13.79 µM]) for the WT PupR-ASD (Table 5.5).

Table 5.5. Melting temperatures determined by CD and estimated K_d's by SV-AUC for PupR-ASD constructs.

Protein	T_m (°C)	K_d (μM)	AUC Concentrations (μM)
WT	53.3	10.2	1.1 – 102
M18D	37.3	226	5.0, 63.1
L22D	43.3	>200	9.45
A35D	44.0	13.5	6.36, 50.0
M38D	44.7	11.6	4.39, 51.21
Q42D	43.7	18.2	6.84, 45.67
R53D	47.7	562	5.27, 41.92
F57D	48.4	>200	12.3
L61D	48.8	>500	103.7

Interestingly, the SEC and SEC-SAXS samples were loaded at a range of concentrations from 100 μM (1.8 mg/mL) to 1.5 mM (28 mg/mL): that is, at concentrations where formation of the dimer was favored. Thus, the combined SEC, SEC-SAXS, and SV-AUC experiments indicate that an equilibrium exists between the monomeric and dimeric states of the ASD, and suggest that, at normal physiological concentrations at the bacterial inner membrane, the ASD may transition between these states.

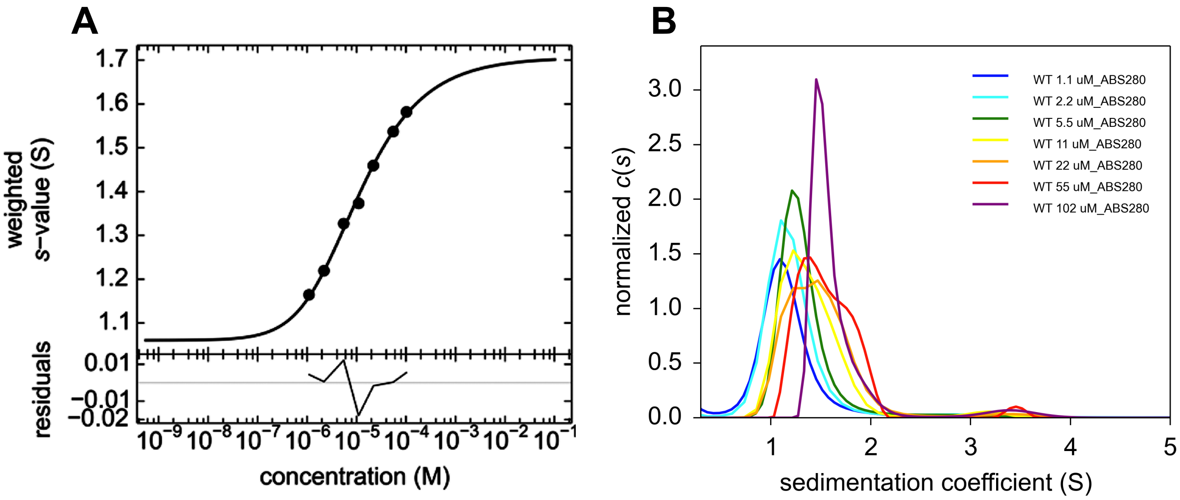


Figure 5.8. SV-AUC analysis of WT PupR-ASD dimerization. (A) s_w isotherm for WT PupR-ASD created by integration of the *c(s)* peaks in the distribution between 0 and 4 S for the data recorded at 20 °C, with error estimates calculated using SEDFIT. (B) SV-AUC *c(s)* distributions for the WT PupR-ASD at seven different concentrations.

5.3.3. The PupR-ASD Forms Symmetric and Asymmetric Dimers in the Asymmetric Unit of the Crystal

The X-ray crystal structure was analyzed for potential dimer interfaces. Two distinct sets of interactions were observed between different monomers in the asymmetric unit. One set consisted of asymmetric interactions stabilized by hydrogen bonding between the N-terminus of chain C and the loop between helix 1 and helix 2 of chain D (Figure 5.2). The other set comprised symmetric interactions between chain A and chain B, with substantial stabilization by hydrophobic interactions between residues on helices 1 and 3 of each monomer (Figure 5.2). The potential symmetric dimer created by the latter set was related by 2-fold, rotational, noncrystallographic symmetry. In contrast, a closed dimer is not created by the asymmetric interactions, and it is unclear why such interactions would create only a dimer, rather than a linear chain of subunits. This suggests that the symmetric dimer observed in the crystal structure is more likely to correspond to the solution dimer.

The PISA server (264) was utilized to analyze the two potential interfaces identified in the crystal structure. For the symmetric dimer, the total surface area buried upon the association of two subunits is 1740 $Å^2$, with a complexation significance score of 1.0. The complexation significance score assesses the maximal fraction of the total free energy of binding for a given interface, and a score of 1.0 denotes a clear significance of this interface for dimer assembly. Thus, this interaction would bury approximately 33% of the total surface area of each subunit upon formation of a symmetrical dimer. In contrast, the asymmetric interaction buries only 1395 $Å^2$ of total surface area, with a complexation significance score of 0.0. This provides additional support that the symmetric homodimer is likely the assembly observed in solution (Figure 5.9).

The symmetric homodimer is stabilized by hydrophobic packing between helices 1 and 3 of each monomer (Figure 5.9). Overall, helix 1 of monomer A (Helix 1^A – chain ID in superscript) runs parallel to helix 3 in monomer B (Helix 3^B). This interface is stabilized by a "knobs-in-grooves" hydrophobic packing between residues M18 and L22 on Helix 1^A and residues R53 (the aliphatic region), F57, and L61 on Helix 3^B (Figure 5.9A). This unit has approximately a 45° tilt to the equivalent Helix 1^B:Helix 3^A unit to create a four-helix bundle. The bundle is further stabilized by interactions between F57 and L61 of Helix 3^A with F57 and L61 of Helix 3^B. The asymmetric interaction involves an association between monomers C and D (Figure 5.2). In contrast to the symmetric interaction, the interface is a three-helix bundle consisting of Helix 1^C, Helix 3^C, and Helix 2^D. This interaction involves $A14^D$, $M18^D$, and $L22^D$ packing against $L34^C$ and $M38^C$, and the aliphatic region of $R41^C$. Additionally, there is hydrogen bonding between the main chain of residues $I9^C$ and $G11^C$ and residues $P27^D$, $V29^D$, and $A31^D$, similar to an antiparallel β-sheet (Figure 5.9B).

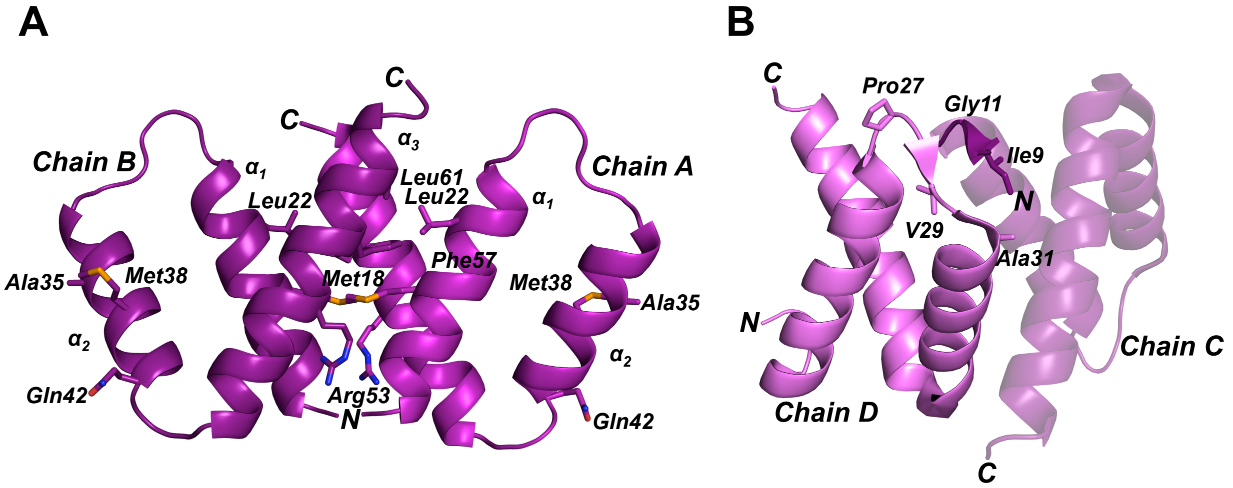


Figure 5.9. PupR-ASD symmetric and asymmetric dimer interfaces. Residues selected for mutation to aspartic acid are shown in stick and colored by atom type as in Figure 5.3. The N- and C-termini are also labeled. (A) The symmetric dimer. (B) An asymmetric dimer with hydrogen bonding between chain C residues I9 and G11, and chain D P27, V29, and A31.

Full-length PupR consists of an N-terminal cytoplasmic domain, a transmembrane domain, and a C-terminal periplasmic domain. In the symmetric homodimer, the C-termini of the two monomers are positioned within 11.2 Å of each other on the same face of the dimer (Figure 5.9), suggesting that this side faces the inner membrane, facilitating the simultaneous insertion of the two transmembrane domains from homodimer into the inner membrane. Conversely, the hypothetical dimer based on the asymmetric contacts would position the two C-termini 41.8 Å apart on opposite faces, which, in the full-length protein, would complicate the simultaneous insertion of the transmembrane helices into the inner membrane since the chains would run in opposite directions. Thus, it is concluded that the symmetrical interactions represent the homodimer seen in solution, while the asymmetric interactions merely represent crystal contacts.

The PupR-ASD SAXS data was analyzed to further explore ASD dimerization. Theoretical SAXS curves were calculated for the crystal structures of the ASD monomer and the symmetric homodimer, and these curves were fit to the experimental PupR-ASD scattering profile. The experimental data fits the calculated scattering curve for the homodimer ($\chi = 8.4$) much better than for the monomer ($\chi = 23.3$). Since both the X-ray structure and the Kratky plot indicated that the ASD was partially disordered, fitting of the theoretical scattering curves of the symmetric dimer and the monomer were then performed using the program EOM 2.0 (268,302). The inclusion of the disordered regions of the PupR-ASD in the models dramatically improved the fits of the calculated scattering curve for the homodimer (best $\chi = 2.2$) relative to that for the monomer ($\chi = 13.5$). This clearly indicates that the scattering curve calculated from the model comprising the symmetric PupR-ASD homodimer as well as the disordered regions fits best to the experimental scattering data (Figure 5.7A).

A low-resolution molecular bead model was also calculated from the experimental SAXS data. The initial model showed an envelope for the PupR-ASD that was C2 symmetric in nature. Manual fitting of the symmetric dimer found in the crystal structure into the averaged a*b initio* SAXS envelope further supported this as the plausible assembly. An improved bead model was calculated by enforcing 2-fold rotational symmetry (Figure 5.10). The symmetric homodimer was docked into this SAXS envelope using SITUS, with a resulting correlation coefficient of 0.87, demonstrating a reasonable fit to the calculated *ab initio* bead model. Thus, the symmetric PupR-ASD homodimer represents the structure of this domain at the concentrations explored in this work.

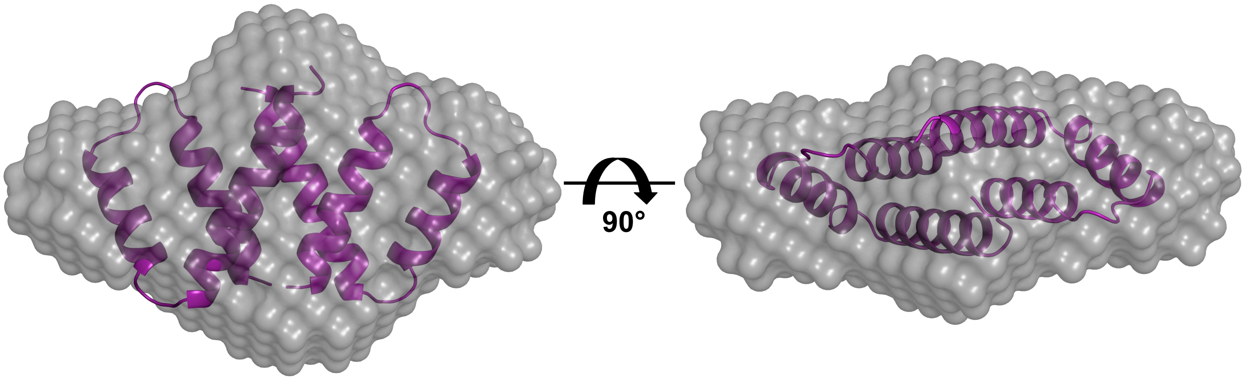


Figure 5.10. *Ab initio* bead model of the PupR-ASD dimer with the calculated surface superimposed with the dimer crystal structure. The bead model is displayed as a grey surface, and the PupR-ASD dimer in purple.

5.3.4. Residues at the Dimer Interface are Essential for ASD Stability

To investigate the impact of residues at the subunit symmetric and asymmetric interfaces, residues were identified for mutation to Asp, to assess quaternary, tertiary, and secondary structure of the ASD. Among the residues M18, L22, R53, F57, and L61 (Figure 5.10) found at the PupR-ASD dimer interface, the three residues with the greatest buried surface area (BSA) were presumed to be of critical importance and were selected for mutagenesis to aspartic acid to force repulsion at the interface. The three residues selected were M18 (BSA = 103 $Å^2$), L22 (BSA = 120 $Å^2$), and F57 (BSA = 109 $Å^2$). In addition to single mutants of each of these residues, an M18D and L22D double mutant, and a triple Asp mutant were created. M38, a residue involved only at an asymmetric interface was mutated to Asp. As controls, several other residues, involved in the asymmetric interactions that were deemed to be crystal contacts, were also mutated to aspartic acid. Residues R53 (BSA = 92 $Å^2$) and L61 (BSA = 88 $Å^2$), two residues that participate in both the dimer interface and the asymmetric crystal contacts, were mutated to Asp. Lastly, Asp mutants of A35 and Q42 were developed; these two residues are not involved in symmetric or asymmetric dimer formation.

The mutant PupR-ASD proteins were analyzed by SV-AUC to assess whether dimerization was altered compared to WT. As expected, the symmetric dimer interface mutations dramatically impacted quaternary structure. Mutation of M18, L22, or F57 to Asp substantially weakened the dimer, resulting in an increase of the dimer dissociation constant to over 200 μM (Table 5.4). Even more strikingly, mutation of R53 or L61 resulted in K_d's of > 500 μM. In contrast, and consistent with our expectations, mutation of A35, M38, and Q42, the control residues and the residue involved only in asymmetric contacts, only marginally impacted K_d's. Thus, mutation of any of the residues involved in the symmetric dimer interface resulted in significant disruption of the PupR-ASD homodimer.

Next, CD spectroscopy (Figure 5.11) was employed in conjunction with thermal denaturation (Figure 5.12) to assess how these mutations influence secondary structure and the stability of the tertiary structure. WT PupR-ASD has a T_m of 53.3 °C. Relative to the WT, mutation of the three residues that contribute only to the symmetric dimer interface also significantly destabilized the ASD tertiary structure (Table 5.4). The L22D and F57D mutations decreased the T_m by 10 and ~5 °C, respectively, while the M18D mutation had the greatest destabilizing effect, reducing the T_m by 16 °C. The double and triple mutants did not maintain the ASD secondary structure and were largely unfolded.

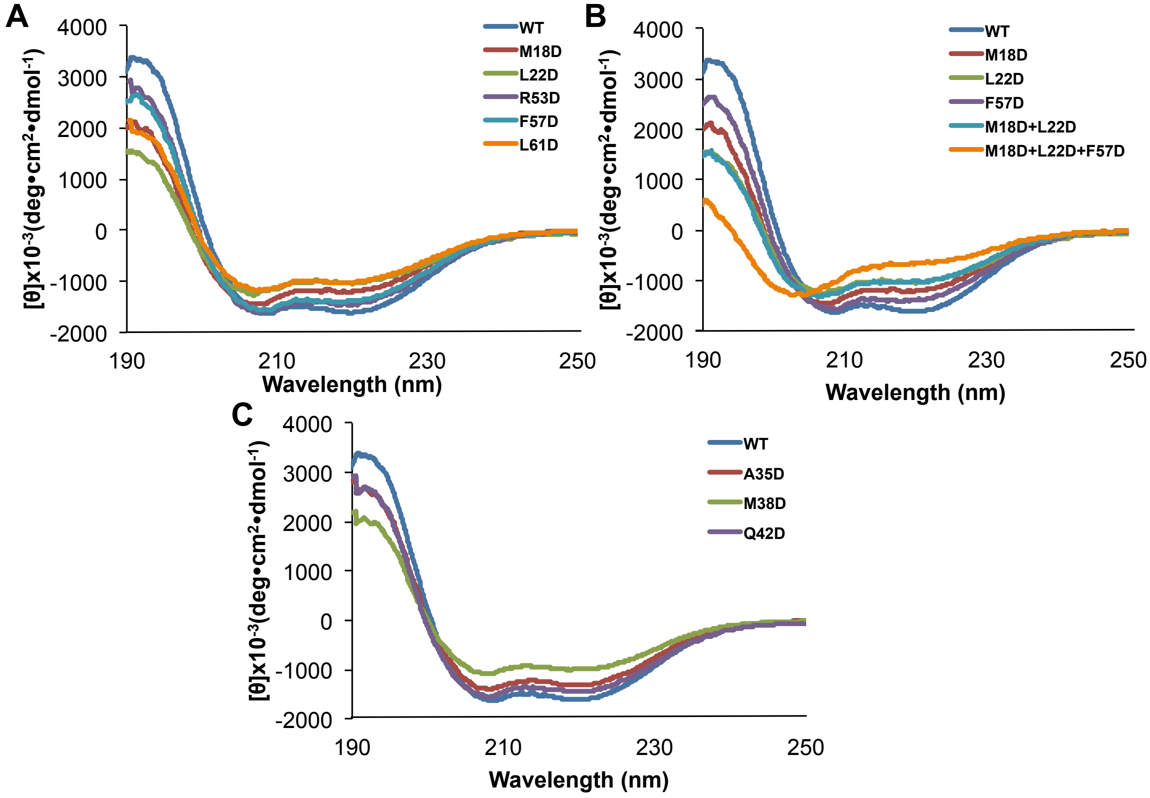


Figure 5.11. CD spectra of WT and mutant PupR-ASD. (A) WT PupR-ASD and Asp mutants of residues at only the symmetric dimer interface (M18D, L22D, and F57D), and residues found at both the symmetric and asymmetric dimer interfaces (R53D and L61D). (B) WT PupR-ASD and Asp mutants of symmetric dimer interface residues, including an M18D+L22D double mutant, and an M18D+L22D+F57D triple mutant. (C) WT PupR-ASD and Asp mutants of control residues.

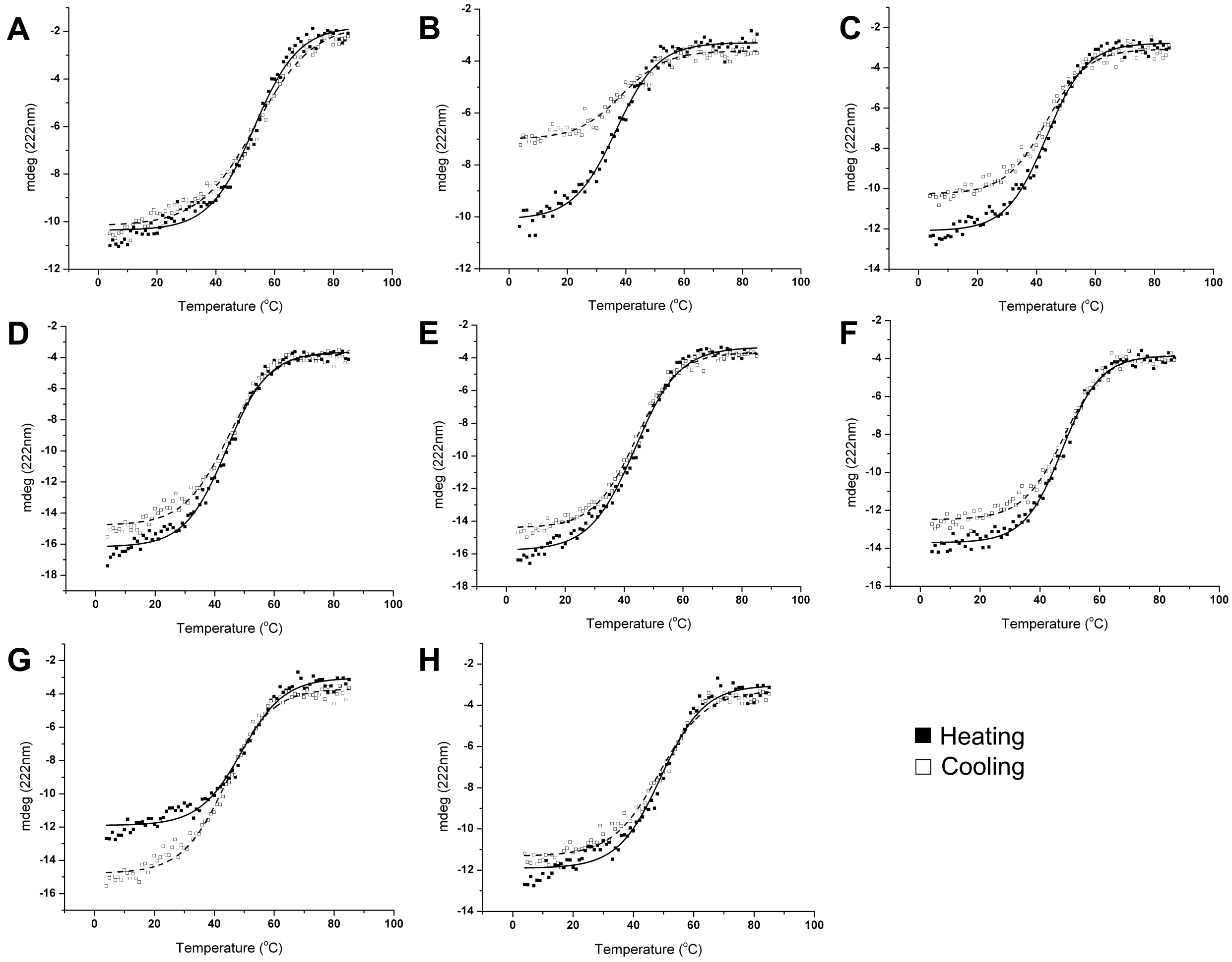


Figure 5.12. Thermal denaturation profiles of PupR-ASD WT and mutants. Unfolding (heating) and refolding (cooling) were monitored at 222 nm from 4 to 85 °C. T_m was determined by fitting a standard Boltzmann sigmoidal curve to the ellipticity (in mdeg). (A) WT. (B) M18D. (C) L22D. (D) A35D. (E) Q42D. (F) R53D. (G) F57D. (H) L61D.

Lastly, some of the symmetric homodimer interface mutants also disrupted secondary structure content of the PupR-ASD (Figure 5.11). Even more strikingly, the secondary structure of a double mutant (M18D + L22D) was almost entirely disrupted, while the secondary structure of a triple mutant (M18D + L22D + F57D) was completely disrupted. Conversely, the secondary structure of the PupR-ASD was not disrupted by mutations of either the residues present at the asymmetric crystal contacts (A35D, M38D, or Q42D), or the residues involved in both sets of contacts (R53D and L61D). Thus, all the symmetric dimer interface mutations negatively impact dimer stability. Further, this combined SV-AUC and CD analysis indicates that M18 is critical for maintaining all levels of ASD structure.

5.4. Summary and Conclusions

The PupR-ASD core fold comprises a three-helix bundle, similar to the fold found in other recently-determined ASD structures. Unlike the zinc-binding ASDs, ChrR and RslA, the PupR-ASD has a relatively extensive hydrophobic core stabilized by W40, H47, W20, and W51. This hydrophobic stabilization is more extensive than that found in the ASDs of RseA and RskA. This stability may be of physiological importance to the function of the iron transport sigma regulators, PupR, FecR, and FpvR, since it has been shown that these ASDs remain attached to their cognate sigma factor after activation of the sigma factor, and are required for sigma factor interaction with RNAP (109,153,242,283). Indeed, mutation and subsequent bacterial 2-hybrid analysis of the three conserved N-terminal tryptophan residues of FecR were found to be essential for FecR activity (105,139), as mutation of these residues disrupted the FecR:FecI interaction, rendering the complex transcriptionally inactive (110). These residues are highly conserved amongst FecR-like proteins. In contrast, ChrR and RslA are degraded and release the sigma factor during activation. Therefore, the stability of the ASD, and its independence from external stabilizing elements, such as a metal ion, is likely central to the mechanism of regulation by the iron-transport ASDs.

The fourth helix found in the ASDs of ChrR, RseA, RskA, and RslA serves to block the interaction of the sigma factor with the RNAP core. Removal of this fourth helix from the σ regulator then becomes an important step for activation of these sigma factors. The lack of an ordered fourth helix in the structure of PupR-ASD may indicate that this class of ASDs regulates its sigma factor solely by tethering it to the membrane. Since the fourth helix is C-terminal to the core helical bundle of the ASD, a model wherein the sigma factor is sandwiched between the larger ASD dimer and the inner membrane would limit access of RNAP or DNA to the sigma factor. Recent evidence shows that intramembrane proteolysis of FpvR, and presumably subsequent release from the membrane, results in activation of PvdS and FpvI (160). Currently, however, the possibility cannot be ruled out that the fourth helix in the PupR-ASD is ordered only in the presence of its sigma factor, PupI. Subsequent CD measurements of the PupR-ASD in the presence of TFE demonstrated an increase in helical content, which also suggests that the fourth helix of the PupR-ASD may only become ordered when associated with PupI. In this scenario, the fourth helix might block RNAP binding until proteolysis of the ASD:sigma factor complex from the transmembrane region, upon which it would revert to a disordered state and allow association to the DNA promoter. Further studies are required to resolve these varying hypotheses.

Isolation of the PupR-ASD as a dimer was unexpected. Identification of a symmetric dimer from the X-ray crystal structure necessitated an in-depth biophysical investigation of the role of the residues at the dimer interface. These studies confirmed that the residues at this interface are critical components of dimer packing and stability. The presence of the dimer does not preclude interaction with a sigma factor. The dimeric interface does not interfere with regions of the ASD responsible for recognizing the σ_2 region of the sigma factor, as found in the ChrR and RseA structures. Superposition of the ChrR:σ^E complex and the RseA:σ^E complex upon the PupR-ASD homodimer reveal that the surface analogous to the surface of these σ regulators which interacts with the σ_2 region of σ^E (σ_2^E) is localized to the exposed face of the PupR-ASD dimer. The C-terminal σ_4 region of σ^E (σ_4^E) does sterically clash with the second monomer of the PupR-ASD homodimer. However, the flexible linker between σ_2^E and σ_4^E

should enable the rearrangements that would allow interaction of the σ_4 region with the second PupR-ASD subunit of the dimer. Alternatively, the σ_4 region may adopt an unknown alternate conformation that does not interact with the ASD allowing interaction of two sigma factors with the dimer. Comparison of the σ_4^E structure from the ChrR and RseA complexes with that observed in the complex of σ_4^E with AsiA (313) demonstrates the flexibility of the σ_4 region. In either case, an ASD dimer could allow binding of the sigma factor in the orientation previously discussed. This is examined further in Chapter 6.

Dimerization has at least two potential mechanistic implications. First, the dimer is important prior to association with the σ factor, but is disrupted concomitant with σ factor association, allowing each monomer to interact with a separate sigma factor. Thus, the dimer may stabilize the ASD and protect against degradation, thereby preventing loss of sigma factor inhibition and unintended activation. However, given that some of our mutant proteins remained folded even when dimer assembly was disrupted, in this mechanism, it is unclear why a dimer is necessary prior to interacting with the sigma factor. Although an *in vivo* LexA-based repression assay demonstrated heterodimer formation between the FecR-ASD and FecI (139), it remains unclear what the true stoichiometry is of this interaction.

The second mechanistic possibility is that since the ASD tethers the sigma factor to the inner membrane, the PupR-ASD dimer is necessary for tethering a dimeric state of PupI. It is currently unknown whether the PupR-ASD dimer is a relevant physiological state.

6. EXPLORING THE INTERACTION OF THE SIGMA FACTOR PUPI WITH THE PUPR ANTI-SIGMA DOMAIN AND *IN VIVO* TRANSCRIPTIONAL ACTIVATION BY PUPI

6.1. Introduction

Group 4 sigma factors are the largest and most diverse class of sigma factors; this class includes the ECF sigma factors, which respond to environmental stimuli to modify the bacteria gene expression profile (284). ECF sigmas retain two primary conserved regions, of the four that are found in all proteins of the σ^{70} family (Figure 6.1). Region 2 associates with the -10 region and promotes DNA melting, and region 4 recognizes the -35 promoter sequence (149). Later structural work identified σ^{70} regions 2-4 as separate domains joined by flexible linkers (314).

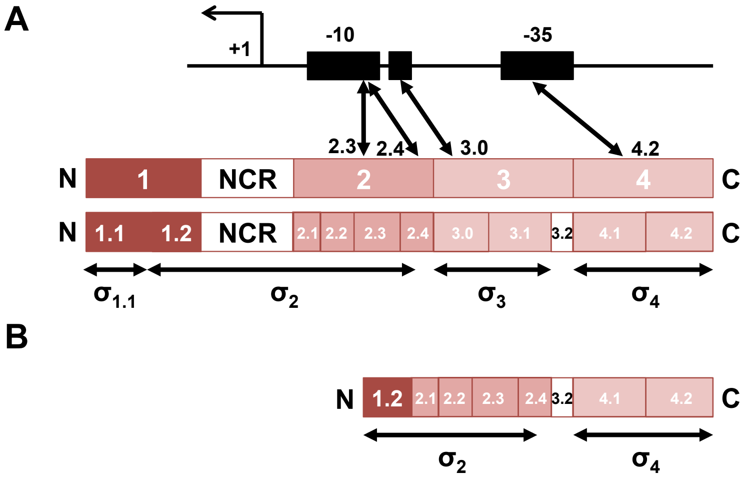


Figure 6.1. Domain organization of the σ^{70} family. (A) Group 1 σ^{70} factors contain four structural domains- $\sigma_{1.1}$, σ_2, σ_3, and σ_4. Within each domain, conserved σ regions are defined: region 1, divided into regions 1.1 and 1.2; region 2, containing regions 2.1-2.4; region 3, divided into regions 3.0-3.2; and region 4, containing 4.1 and 4.2. A non-conserved region (NCR) separates region 1.2 from region 2.1. The regions of the Group 1 σ^{70} factor that associate with various promoter elements (-10 and -35 elements) are delineated. (B) Group 4 σ^{70} factors (ECF sigma factors) are the most minimal sigma factors, lacking both $\sigma_{1.1}$ and σ_3. Structurally, the ECF sigma factors are predicted to only retain the σ_2 and σ_4 domains. (Adapted from (149,284)).

A prime example of the Group 4 ECF sigmas are those that control TBDT expression during iron acquisition. Bacterial two-hybrid assays demonstrated that, specifically, regions 4.1 and 4.2 of FecI, which include a predicted helix-turn-helix motif (residues 139-158), interact with the FecR-ASD (101,315). X-ray crystal structures of the RseA-ASD from *E. coli* in complex with σ^E and the ChrR-ASD from *R. sphaeroides* in complex with σ^E both reveal a similar binding mode, despite the low sequence similarity between the anti-sigma-ASDs (Figure 5.5 and 6.13). Both exhibit the established 3-helix bundle of the ASD that lies between the σ_2 (region 2) and σ_4 (region 4) domains of σ^E. This orientation of the complex occludes the β-flap-tip-helix binding determinant which further blocks interactions with the RNAP β' coiled-coil (284,291). Amongst ECF sigma factors, region 2 displays the greatest sequence variation, while the primary sequence of region 4 is comparatively well-conserved (316). The sequence divergence of region 2, which recognizes the -10 promoter element, likely indicates variation in promoter-binding specificity. Regions 2 and 4 of the sigma factor bind the -10 and -35 promoter recognition sequences, respectively. However, there is evidence to suggest that FecI also interacts with a region near the +13 position, as substitutions close to +13 of the f*ecA* promoter disrupts association by the FecI:RNAP complex (153). This finding indicates that the FecI:RNAP may display a novel interaction with the *fecA* promoter region (317). The promoter sequences recognized by ECF sigma factors have been identified for multiple promoters (318).

The RNA polymerase core enzyme is composed of five subunits: β', β, α^I, α^{II}, and ω (Figure 6.2). The RNAP holoenzyme is formed by association of a sigma factor with the core enzyme to initiate transcription at specific promoters. Several X-ray crystal structures of the RNAP holoenzymes from *Thermus aquaticus* and *Thermus thermophilus* reveal a primary interaction interface between the RNAP β' subunit and σ^{70}; in particular, the β' coiled-coil and the helical region 2.2 of the sigma factor (PDBs 1KU7, 1KU2, 1IW7, 4Q5S, 4Q4Z, 4XLQ, 4XLP, 4XLN) (314,319-324).

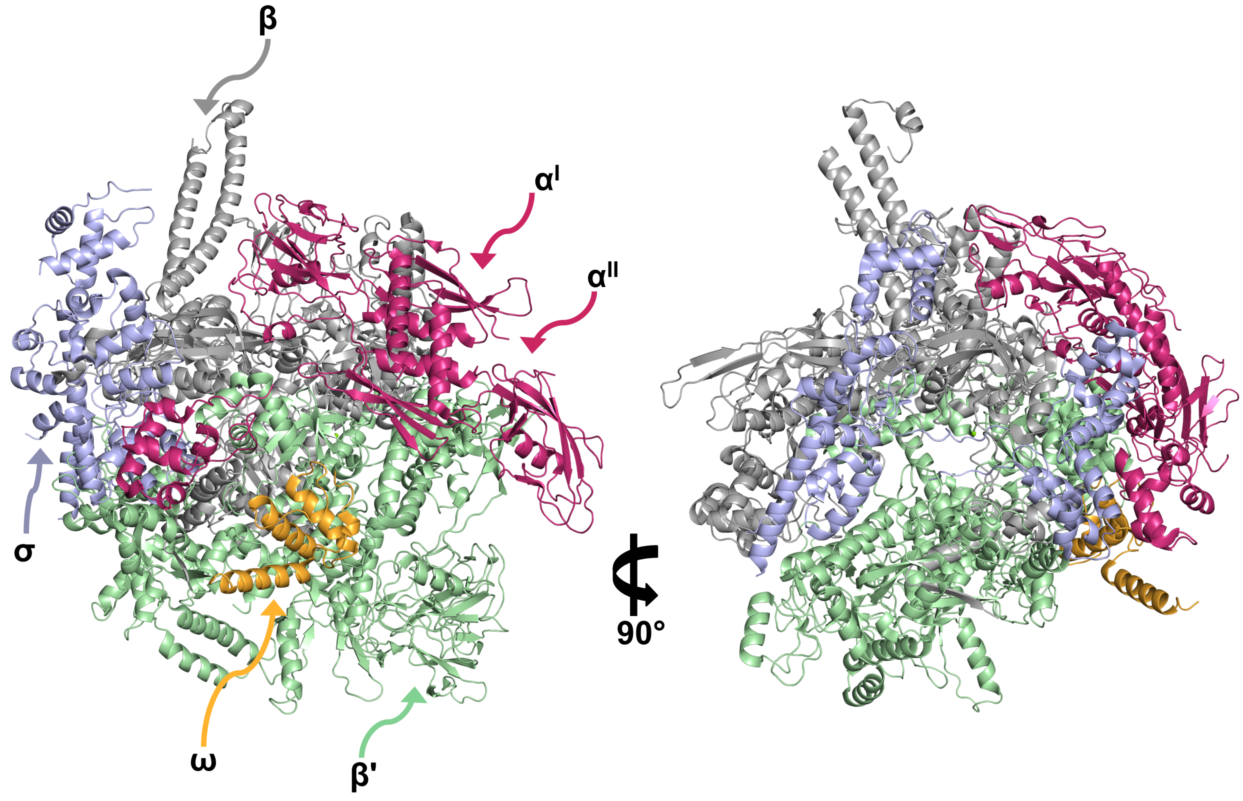


Figure 6.2. X-ray crystal structure of the *E. coli* RNA polymerase σ^{70} holoenzyme (325). The core enzyme in bacteria consists of 5 subunits: α^{I} and α^{II} (magenta), β (grey), β' (green), and ω (orange) (PDB 4YG2). Several sigma factors may recruit the core RNAP to a DNA promoter to form the holoenzyme; this structure contains a housekeeping sigma factor σ^{70}, RpoD (blue). RpoD is missing the first 93 residues ($\sigma_{1.1}$), but retains the remainder of the protein (613 residues), which includes σ^{70} factor domains σ_2, σ_3, and σ_4.

FecI interacts with the β' fragment residues 1-313 by affinity pull-down; interestingly, $\text{FecR}_{1\text{-}85}$ also co-eluted with FecI and $\beta'_{1\text{-}313}$, suggesting that the FecR-ASD remains bound to FecI even during association with RNAP (318). FecI is also more susceptible to trypsin degradation when not bound to the FecR-ASD. This supports the model wherein the FecR-ASD:FecI complex is released from the inner membrane via RIP and regions of FecI necessary for promoter and RNAP interaction are exposed. Point mutants in region 2 of FecI reduces transcription of *fecA*, in addition to decreasing binding of FecI to β' (106). A similar system is found in *P. aeruginosa*, although the FecR-like protein FpvR regulates two sigma factors- FpvI and PvdS; additionally, FpvR and FpvI/PvdS are not co-transcribed, as is observed for other Fec-like regulatory CSS systems. Although FpvR controls PvdS availability, PvdS remains active in mutants lacking FpvR. This may relate to PvdS' enhanced role as a sigma factor that initiates the transcription of multiple genes, while FpvI only regulates the transcription of *fpvA* (109). Mutations in region 2 of PvdS also reduced its affinity for RNAP (326). Consequently, two subgroups of transmembrane signaling regulators have been experimentally identified- in one group, the activity of the sigma factor relies upon the anti-sigma factor (as in the *fec* system), and in the second group, the anti-sigma domain functions solely as an anti-sigma factor, and the sigma activity does not depend on the anti-sigma-ASD, once cleaved from the transmembrane anti-sigma region (as in the *fpv* system) (111). Thus, at least in the case of the *fec* system, the anti-sigma-ASD presents both "pro-sigma" and "anti-sigma" activity. In the *fpv* system, degradation of the FpvR-ASD may lead to active sigma factor, as in the RseA-ASD:σ^E complex, which requires degradation by the ClpXP protease to release active σ^E (327).

To investigate the structural features of the sigma factor PupI and its interaction with the PupR-ASD, large quantities of purified, homogenous PupI were required for biophysical analysis. Given the instability of PupI in solution, the only developed strategy that enabled purification of PupI was via co-expression and purification with the PupR-ASD, or solubilization of PupI from inclusion bodies with low concentrations of detergent. Preliminary efforts toward understanding the solution characteristics and PupI:PupR-ASD complex interaction are described. Analysis by SEC-SAXS suggests that the PupI:PupR-ASD complex is highly flexible, and molecular weight estimates of the complex in solution suggest a 1:1 stoichiometry.

6.2. Materials and Methods

6.2.1. Secondary Structure Prediction and Homology Modeling of PupI

The PupI primary sequence was aligned with the sequences of FecI and FpvI with ClustalW (328). Sequence alignment of putative iron regulatory ECF sigma factors from multiple proteobacteria was accomplished with T-COFFEE (329) and manually edited in Aline (330). Psipred (331) and Jpred4 (332) were used to predict secondary structure.

Homology modeling and tertiary structure prediction of PupI was performed with MODELLER (333,334) via ModWeb (https://modbase.compbio.ucsf.edu/modweb/). A search of the Protein Data Bank for available structures of sigma factors yielded several structures, as outlined in Table 6.1 and Figure 6.3. MODELLER uses several criteria for selecting structural homologs from ModBase with which to construct a model from the input primary sequence: the ModPipe Protein Quality Score (MPQS), a composite score including sequence identity of the target to the template, the coverage of the modeled sequence, the e-value, as defined by NCBI's PSI-BLAST (335); the z-Discrete Optimized Protein Energy (DOPE) score, an improved reference state corresponding to non-interacting atoms in a homogenous sphere; and GA341, a reliability score for the model derived from statistical potentials (336). Additional monitored statistics include the TSVMod NO35, or the estimated native overlap of the model to the template(s) at 3.5 Å; the TSVMod RMSD of the predicted model to the template(s); and MTALL vs. MSALL, wherein MTALL relies on a training set based on template structures(s) (which is considered more reliable), whereas MSALL considers a training set based on similarities in predicted or experimentally-derived secondary structure elements. Acceptable generated models fall within or near the range of values for these criteria, as listed in Table 6.2.

Table 6.1. σ^{70} (E, C, H, R, W) and σ^{54} (N) structures compared to the σ^{70} factor PupI.

PDB ID	Protein	In Complex?	Residues in Structure	Total Residues	Residues Aligned to PupI	Region Sequence ID to PupI	Region Used for Template
1OR7	σ^{E}	RseA-ASD	1-187	191	15-181	19%	15-181
2O8X	σ^{C}	--	117-176	185	118-170	28%	118-170
2O7G	σ^{C}	--	3-91	185	20-91	25%	--
3MZY	σ^{H}	--	45-204	204	46-188	15%	46-188
1H3L	σ^{R}	--	25-102	227	39-102	16%	39-102
5FGM	σ^{R}	--	143-207	227	143-207	18%	--
2K9L	σ^{N}	--	60-135	398	60-135	15%	--
5WUR	σ^{W}	RsiW-ASD	2-187	187	125-187	20%	--

Table 6.2. Selection criteria statistics for MODELLER model reliability of PupI based on available structures.

PDB ID	MPQS	TSVMod Method	TSVMod RMSD	TSVMod NO35	E-value	GA341	z-DOPE
1OR7	1.07608	MTALL	27.13	0.043	3.3e-8	0.54	0.67
2O8X	0.658058	MTALL	1.518	0.961	0	0.85	-0.12
3MZY	0.678122	MSALL	12.648	0.175	0	0.10	-0.24
1H3L	0.679162	MTALL	2.636	0.855	0	0.18	-1.53
Reliable values	≥ 1.1	MTALL > MSALL	< 1.0	< 3.5 Å	< 0.0001	≥ 0.7	< 0

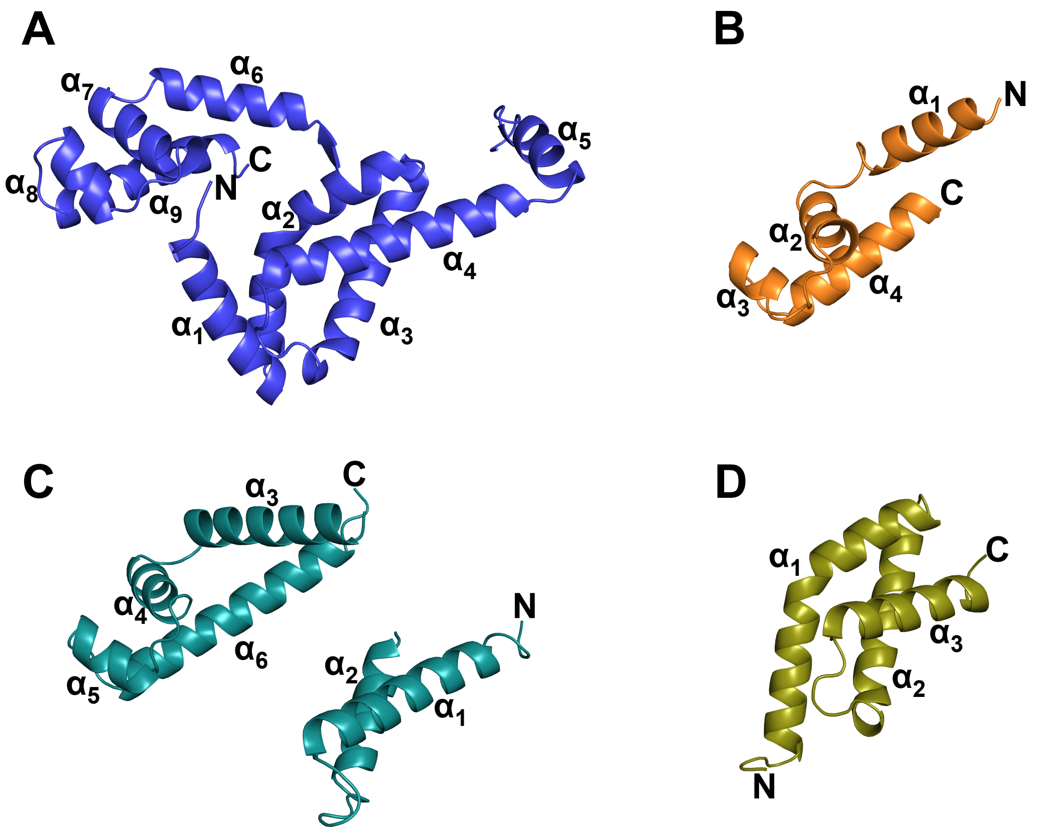


Figure 6.3. Selected crystal structures used for PupI homology modeling. (A) σ^W (PDB 1OR7) (291), (B) σ^C (PDB 2O8X) (337), (C) σ^H (PDB 3MZY) (Stein, A. J., *et al.* Unpublished), and (D) σ^R (PDB 1H3L) (338).

All tertiary and quaternary X-ray crystal structures, or those derived from homology modeling, were visualized and compared in PyMOL (Schrödinger, LLC).

6.2.2. Cloning of PupI Expression Constructs

Several expression constructs of PupI with various affinity tags were created (Table 6.3). Briefly, *pupI* was amplified from *Pseudomonas capeferrum* WCS358 by PCR, to include a 5' BamHI or NdeI restriction endonuclease site and 3' XhoI or HindIII site. PCR products and vectors were digested with the appropriate restriction endonucleases, and Antarctic phosphatase added to the vectors, followed by separation on a 1% agarose gel and extraction of the digested DNA. Ligation of the *pupI* insert with the vector was accomplished with the Quick Ligation protocol and reagents (NEB). The ligation products were transformed into *E. coli* Top10 cells, and proper ligation of the insert was verified by DNA sequencing by MCLab.

Table 6.3 Summary of cloned PupI expression constructs.

Expression Vector	Restriction Endonuclease Sites	Affinity Tag	Final Construct ID	Selectable Marker
pMBP-Parallel-1	BamHI + XhoI	MBP + long linker	pMBP-PupI	Amp^R
pET17b	NdeI + XhoI	No tag	pET17b-PupI	Amp^R
pET28a	NdeI + XhoI	N-terminal His_6 + Thrombin cleavage site	pET28a-PupI	Kan^R
pET28a	BamHI + XhoI	N-terminal StrepII + SUMO	pET28a-SUMO-PupI	Kan^R
pET28a	BamHI + XhoI	N-terminal StrepII + SUMO + TEV cleavage site	pET28a-SUMO-TEV-PupI	Kan^R
pET28a	BamHI + XhoI	N-terminal StrepII + SUMO + C-terminal His_6	pET28a-SUMO-PupI-His_6	Kan^R
pET28a	BamHI + XhoI	N-terminal StrepII + SUMO + TEV cleavage site + C-terminal His_6	pET28a-SUMO-TEV-PupI-His_6	Kan^R
pET29b	NdeI + XhoI	C-terminal His_6	pET29b-PupI	Kan^R
pET-Duet	BamHI + HindIII	No tag	pET-Duet-PupI	Amp^R
pET-Duet	PupI = BamHI + HindIII PupR-ASD = NdeI + XhoI	PupI = no tag PupR-ASD = N-terminal His_6	pET-Duet-PupI+PupR-ASD	Amp^R
pSV	BamHI + XhoI	N-terminal His_6 + MBP + TEV cleavage site	pSV-PupI	Kan^R

6.2.3. Expression Trials of PupI

To determine ideal expression conditions for PupI, the following variables were tested for all the constructs described above: cell lines, growth media, induction temperature, expression time, and concentration of IPTG. BL21(DE3)pLysS, Arctic Express, C43, or C41 chemically-competent cells were transformed with a PupI expression construct, and incubated in LB, 2x LB, TB, or M9 minimal medium supplemented with the appropriate antibiotic at 37 °C, until an optical density of ~0.6-0.8 at 600 nm was obtained. After induction of protein expression with 0.5 or 1 mM IPTG, the cells were incubated for either 4 hrs at 37 °C, or 16-18 hrs at 20 °C, 16 °C, or 10 °C.

6.2.4. Strategies for Purifying PupI

In addition to optimization of expression conditions, numerous purification strategies were evaluated to identify methods that produced stable, homogeneous, and pure PupI. Determination of purification conditions was complicated by initial observations that PupI or His_6-tagged PupI was quite unstable in solution, and often expressed as inclusion bodies. Methods tested to minimize or overcome these obstacles included protein denaturation in 6 M urea, followed by dialysis to remove the urea and refold the protein (339); solubilization of insoluble protein with 0.1% w/v FC-13 or 2 mg/mL Anzergent 3-10 (Anatrace) (153); solubilization of inclusion bodies in PBS pH 8.0 + 8 M urea, dialysis into FoldIt buffer (340); purification of an MBP-PupI fusion protein; and purification of a SUMO-PupI fusion protein.

The only method tested that resulted in homogenous and soluble PupI utilized solubilization of His_6-PupI or PupI-His_6 in FC-13. Chemically competent BL21(DE3)pLysS cells were transformed with either pET28a-PupI (N-terminal His_6) or pET29b-PupI (C-terminal His_6), and transformed cells were grown in LB medium with 15 μg/mL kanamycin at 37 °C to an optical density of 0.8 at 600 nm. IPTG was added to a working concentration of 0.5 mM, and the culture temperature maintained at 20 °C for 16-18 hrs. Cells were harvested and lysed in 25 mM Tris:HCl pH 7.0, 300mM NaCl via emulsification. After centrifugation at 20000 *g* for 40 min, the pellets were homogenized in 25 mM Tris:HCl pH 7.0, 300 mM NaCl, 25 mM imidazole; the volume was measured and FC-13 was added to 0.1% w/v. Protein was solubilized by gentle stirring at 4 °C for 30 min. A second centrifugation was performed, with the same settings as the first centrifugation step. Affinity chromatography was performed using a 5 mL HisTrap column (GE Lifesciences), with wash buffer: 25 mM Tris:HCl pH 7.0, 300 mM NaCl, 25 mM imidazole, 0.1% w/v FC-13, and elution buffer: 25 mM Tris:HCl pH 7.0, 300 mM NaCl, 125 mM imidazole, 0.1% w/v FC-13. Elution fractions containing PupI were pooled, concentrated, and loaded onto a Superdex 75 (10/300) size exclusion column. SEC was performed at a flow rate of 0.3 mL/min, with buffer 25 mM Tris:HCl pH 7.0, 300 mM NaCl, 0.1% w/v FC-13. A second SEC step was incorporated to further purify the PupI, and fractions containing PupI were concentrated to 5-10 mg/mL, and stored at -80 °C after flash-freezing protein aliquots in LN_2. Protein homogeneity was monitored throughout purification by SDS-PAGE with Coomassie Blue staining, and sample concentration determined by absorbance at 280 nm using the molar extinction coefficient ε_{280} = 16960 M^{-1} cm^{-1} and a theoretical molecular weight of 21276 g/mol.

6.2.5. Western Blotting and N-terminal Sequencing of PupI

Twenty μg of total protein from various stages of the purification of the PupR-ASD:PupI-His_6 complex were resolved on SDS-PAGE and transferred to nitrocellulose membranes under 25 V overnight at 4 °C. Membranes were blocked with TBS-T with 5% w/v milk powder, followed by incubation with either Anti-His_6-HRP (Clontech) or Anti-StrepII-HRP (NEB) antibodies at a 1:5000 dilution. His_6- or StrepII-tagged proteins were detected with ECL Western Blotting Substrate (Pierce) and imaged with a Storm 865 Imaging Station (GE Lifesciences).

To verify separation of PupI-His_6 from StrepII-SUMO, N-terminal sequencing by Edman degradation was performed. For sample preparation, 10 μg of protein were separated by SDS-PAGE and transferred to polyvinylidenedifluoride (PVDF) via a semi-dry electroblotter (Owl) after soaking the membrane in 10 mM CAPS pH 11, 10% methanol. Transfer was complete after 1 hr at 20 V. The membrane was stained with amido black for ~2 min, then destained 3 times with 5% acetic acid. Stained species thought to correspond to StrepII-SUMO and PupI-His_6 were cut from the membrane, and sent to the Tufts University Core Facility for Edman sequencing. Five cycles of peptide degradation were performed.

6.2.6. Co-expression and Purification of PupI with the PupR-ASD

Purification of PupI:PupR-ASD from cells co-transformed with MBP-PupI:MBP-PupR-ASD, MBP-PupI:His$_6$-PupR, or His$_6$-PupI:MBP-PupR-ASD was unsuccessful, as the result was either non-expression of one of the proteins, or precipitation of PupI during the procedure. However, co-expression was accomplished using the combination of pHis$_6$-PupR-ASD and pET28-SUMO-tev-PupI-His$_6$ vectors.

Chemically competent *E. coli* BL21(DE3)pLysS cells were transformed with pHis$_6$-PupR-ASD and pET28-SUMO-TEV-PupI-His$_6$. Transformed cells were grown at 37 °C in LB medium supplemented with 100 μg/mL ampicillin and 15 μg/mL kanamycin until an optical density of ~1.0 at 600 nm was achieved. The temperature of the culture was decreased to 20 °C, and protein expression was induced with the addition of 1 M IPTG to a final concentration of 0.5 mM. The culture temperature was maintained at 20 °C for 18 hrs; the cells were harvested by centrifugation at 4000*g* for 30 min. Cells were resuspended in 25 mM Tris:HCl pH 8.0, 150 mM NaCl, centrifuged a second time, and the cell pellets stored at -80 °C until use.

For protein purification, approximately 20 g of cells (3L of expression culture) were lysed in 25 mM Tris:HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol, 2 mM DTT or 5 mM β-mercaptoethanol (βME) by passage through an emulsifier at 17000 psi. Centrifugation at 20000*g* was performed for 45 min to remove cellular debris, and the supernatant applied to 10 mL of Strep-Tactin resin (IBA Life Sciences). After washing the column with lysis buffer, TEV was added to the column in a 1:10 mass ratio, and incubated for 16-18 hrs at 4 °C to remove the Strep-SUMO tag from the PupI construct, and the His$_6$ tag from the PupR-ASD. The proteolyzed complex was eluted from the column with the addition of lysis buffer. Wash fractions containing the complex were pooled, concentrated, and loaded onto either a Superdex 200 (10/300 GL) or

Superdex 75 (10/300 GL) size exclusion column. Isocratic elution at a flow rate of 0.3 mL/min with 25 mM Tris:HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol was performed. SEC fractions containing the complex were pooled, concentrated, and rerun over the size exclusion column to better separate contaminants and uncleaved fusion protein. Fractions from the second SEC step were concentrated to 5-10 mg/mL and stored at 4 °C or flash-frozen in LN_2. Protein purity was assessed throughout purification by SDS-PAGE stained with Coomassie Blue, and protein concentration was determined by Bradford assay. A standard preparation yielded approximately 3 mg of total protein per 1 L of expression culture. One explanation for the low yield is the very limited binding capacity of the Strep-Tactin resin, which only binds 3 mg of protein per 1 ml of resin.

6.2.7. Crystallization Trials with PupI in Detergent and the PupI:PupR-ASD Complex

His_6-PupI was concentrated to 6 mg/mL in 25 mM Tris:HCl pH 8.0, 300mM NaCl, 0.05% w/v FC-13, and the PupI-His_6:PupR-ASD complex was concentrated to 5-10 mg/mL in 25 mM Tris:HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol. Each sample was utilized for the Pre-crystallization Test (PCT, Hampton Research) to assess optimal protein concentration for crystallization screening. PCT results yielded no precipitate from reagents A1 and B1, and very light precipitate from A2 and B2, indicating that the proteins were sufficiently concentrated to perform screens. MCSG-1 and MCSG-2 (Microlytic, Anatrace) were utilized to identify protein crystallization conditions. Each screen contained 96 reagents, and reservoir solution was combined with protein sample at 3:1, 1:1, and 1:3 v/v ratios, for a total of 576 conditions. Crystallization trays were incubated at 20 °C and monitored daily.

6.2.8. PupI:PupR-ASD SEC-SAXS

SEC-SAXS was performed on samples from two different preparations of the PupR-ASD:PupI-His$_6$ complex at beamline 18-ID (Bio-CAT) of the APS. An inline Superdex 200 10/300 GL Increase column (GE Lifesciences) was equilibrated with 25 mM Tris:HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol at a flow rate of 0.7 mL/min. For the first sample, 5mg of total protein was loaded; for the second, 8mg of the complex was used; the final volume of both samples was 500 µL. SEC-SAXS data were recorded at a wavelength of 1.033 Å on a Pilatus 3 1 M detector at a sample-to-detector distance of 3.0 m, covering a momentum transfer range of 0.0055 Å $< q <$ 0.388 Å^{-1}. Data were normalized to the incident X-ray beam intensity, and buffer-subtracted using Bio-CAT beamline software prior to analysis with the ATSAS program suite (198). Linearity of the Guinier region was examined using PRIMUS (193) and the Kratky plot calculated to assess disorder of the complex in solution. GNOM (199,296) was used to compute the radius of gyration, R_g; pairwise distribution function, $P(r)$; and the maximum particle dimension, D_{max} (199,296). Ten independent *ab initio* bead model reconstructions were calculated using DAMMIF (196). The resultant dummy atom models were averaged and filtered with DAMAVER (297) using default parameters. The models had an average χ^2 of 2.31 ± 0.004, with an averaged normalized spatial discrepancy (NSD) for the 10 DAMMIF calculations of 0.51 ± 0.04. The resultant three-dimensional averaged envelope was superimposed with the PupI homology model or the PupI model:PupR-ASD crystal structure with SITUS and the UCSF Chimera package (201,202,298). CRYSOL (211) and FoxS were used to compare theoretical scattering curves calculated from PupI monomer models and PupR-ASD:PupI-His$_6$ heterodimer atomic models against the experimental SAXS data. PyMOL was used to visually evaluate fits of the crystal structures and derived models. The molecular weight of the complex was determined with the SAXS MoW server (http://www.if.sc.usp.br/~saxs/) (299). EOM 2.0 (268,302) in the ATSAS suite was employed to model the flexible linker region between the two σ^{70} domains of PupI (σ_2 and σ_4). MultiFoxS was also used to model multi-conformational states that PupI may sample in solution, and predict how often those states are sampled.

6.2.9. *In vivo* Analysis of Transcriptional Activation by PupI

6.2.9.1. Cloning of Reporter Vectors

Two reporter vectors were constructed to assess the ability of PupI to transcriptionally activate PupB production. Two low copy number, broad host range, shuttle vectors were obtained from the Standard European Vector Architecture 2.0 (SEVA) repository (http://www.seva.cnb.csic.es) (341-343). The plasmids, designated by SEVA as pSEVA-231 and pSEVA-531, included the same backbone architecture, with the exception of the selectable marker. pSEVA-231 contained a kanamycin resistance gene, and pSEVA-531 contained a tetracycline resistance marker. Additionally, two vectors containing the desired insert for the reporter vector were acquired from Aldevron. The two custom vectors had a pUC57 backbone, with either an ampicillin or kanamycin selectable marker. The insert included the features displayed in Figure 6.4, constituting the coding region for PupR residues 182-324 and PupB residues 1-128 (the signal peptide and N-terminal signaling domain) as a green fluorescent protein (GFP) fusion.

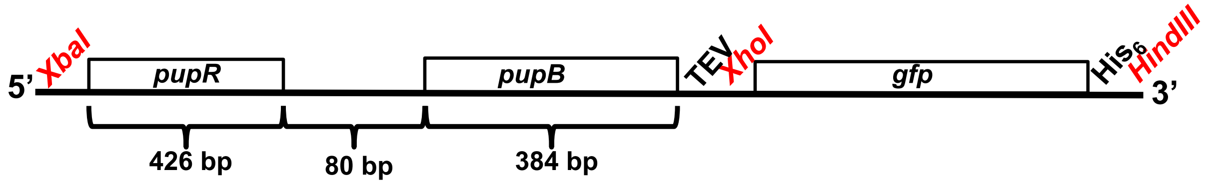


Figure 6.4. Schematic representation of the reporter vector insert. The insert contains the coding regions for PupR residues 182-324 and PupB residues 1-128, separated by 80 bp, which is half of the genomic region between the *pupR* and *pupB*, including the proposed PupI binding site. Other features of the vector are a 5' XbaI and 3' HindIII restriction endonuclease recognition sequences, a TEV protease recognition sequence following the *pupB* truncation, an XhoI site immediately upstream of the GFP-encoding region, and a His_6 affinity tag immediately following the *gfp* gene.

The pUC57 and pSEVA vectors were digested with XbaI and HindIII, and the pSEVA digests incubated with Antarctic phosphatase prior to purification of the digest fragments by gel extraction. The insert and backbone were combined in a 1:1 molar ratio, and ligated via the Quick Ligation protocol (NEB), and the ligation product transformed into *E. coli* Top10 cells by the heat shock method. Cells containing the ligated vector were isolated in LB medium with either 10 μg/mL tetracycline or 15 μg/mL kanamycin, and DNA sequencing (MCLab) confirmed proper ligation of the insert into the vector backbone. The resulting reporter vectors were designated pSEVA-231-RBSD-GFP or pSEVA-531-RBSD-GFP.

6.2.9.2. Triparental Mating of Pseudomonas capeferrum

The WT *P. capeferrum* used in this work was a gift from Dr. Peter Bakker at Utrecht University, Utrecht, Netherlands, and the three mutant *P. capeferrum* strains were a gift from Dr. Margot Koster, also at Utrecht University. The three mutants were created by Koster, *et al.* via gene replacement of a target gene with a chromosomal insertion element (Ω interposon or Tn5) containing an antibiotic resistance marker (344,345). *P. capeferrum* KV51 and KV52 were developed by Tn5 insertion into *pupI* and *pupB*, respectively, with a kanamycin resistance gene (57). *P. capeferrum* BWV29 was created by gene replacement of *pupR* with the Ω interposon carrying a streptomycin resistance marker (140). All *P. capeferrum* possess intrinsic resistance to nalidixic acid (57,140,346).

All attempts to transform *P. capeferrum* with the reporter vector failed, including via the standard heat shock method with chemically competent cells and electroporation. Triparental mating (conjugation) was used to transfer the reporter construct into four *P. capeferrum* strains (347). Briefly, *E. coli* DH5α cells were transformed with the reporter vector by following the heat shock method. *E. coli* containing the helper plasmid pRK2013 were acquired from the American Type Culture Collection (ATCC). DH5α cells harboring the reporter vector (pSEVA-531-RBSD-GFP), *E.coli* carrying pRK2013, and each *P. capeferrum* strain were grown to late log phase (approximately 16 hrs incubation at 37 °C for *E. coli* and 30 °C for *P. capeferrum*). The two *E.coli* strains and one *P. capeferrum* strain were combined in a 1:1:1 ratio, such that each strain was added to a concentration of 1.10^9 cells per mL. Following a short centrifugation step to pellet the cells, the mating mixture was suspended in 100 μL of King's B medium (KBM) and plated on a KBM agar plate. After incubation for 16 hrs at 30 °C, cells were suspended in 1 mL of KBM, diluted 1:10 in KBM, and spread onto KB agar with 40 μg/mL tetracycline and 25 μg/mL nalidixic acid. For the *P. capeferrum* strains KV51 and KV51, 50 μg/mL kanamycin was also used; for the BWV29 strain, 50 μg/mL streptomycin was included. Plates were incubated for 16 hrs at 30 °C.

As a control, *E. coli* DH5α cells were transformed with the reporter construct by electroporation. LB medium with 10 μg/mL tetracycline was used to select for cells that retained the plasmid.

6.2.9.3. Flow Cytometry Analysis of Transcriptional Activation and Cell Viability

To determine transcriptional activation via PupI binding to the reporter vector, a variety of growth conditions were explored to determine cellular response through the Pup regulatory system. Untransformed cells and cells carrying the pSEVA-531-RBSD-GFP reporter vector were grown in LB broth overnight at 37 °C (DH5α) or 30 °C (*P. capeferrum*). Five mL cultures were supplemented with the appropriate antibiotics for selection, as described above. After approximately 16 hrs of growth, cultures were centrifuged at 4000 *g* to pellet the cells, then suspended in 1 mL PBS pH 7.0. Fifty μL of the cell suspension was utilized to inoculate a second set of cultures, with the conditions outlined in Table 6.4.

Table 6.4 Bacterial growth conditions evaluated for transcriptional activation by PupI *in vivo*.

Condition #[a]	**Cells**	**Mutation**	**Reporter present?**	**Media**	**Supplement**
1	DH5α	None	No	LB	None
2	DH5α	None	No	LB	100 μM $FeCl_3$
3	DH5α	None	No	LB	20 uM Pseudobactin[b]
4	DH5α	None	Yes	LB	None
5	DH5α	None	Yes	LB	100 μM $FeCl_3$
6	DH5α	None	Yes	LB	20 uM Pseudobactin
7	*P. capeferrum* KV51	*pupI*::Tn5	No	KBM	None
8	*P. capeferrum* KV51	*pupI*::Tn5	No	KBM	100 μM $FeCl_3$
9	*P. capeferrum* KV51	*pupI*::Tn5	No	KBM	20 uM Pseudobactin
10	*P. capeferrum* KV51	*pupI*::Tn5	No	RSM[c]	None
11	*P. capeferrum* KV51	*pupI*::Tn5	No	RSM	100 μM $FeCl_3$
12	*P. capeferrum* KV51	*pupI*::Tn5	No	RSM	20 uM Pseudobactin
13	*P. capeferrum* KV51	*pupI*::Tn5	Yes	KBM	None
14	*P. capeferrum* KV51	*pupI*::Tn5	Yes	KBM	100 μM $FeCl_3$
15	*P. capeferrum* KV51	*pupI*::Tn5	Yes	KBM	20 uM Pseudobactin
16	*P. capeferrum* KV51	*pupI*::Tn5	Yes	RSM	None
17	*P. capeferrum* KV51	*pupI*::Tn5	Yes	RSM	100 μM $FeCl_3$
18	*P. capeferrum* KV51	*pupI*::Tn5	Yes	RSM	20 uM Pseudobactin

[a] Growth conditions 7-18 were repeated for each *P. capeferrum* strain: WT (no mutation), KV52 (*pupB*::Tn5), and BWV29 (*pupR*::Ω).

[b] Pseudobactin was prepared by Dr. Ganesh Bala, Montana State University, following standard siderophore isolation methods (348,349); the compound was suspended in PBS 7.0, and the concentration determined using $\varepsilon_{400} = 1.4 \times 10^4\ M^{-1}\ cm^{-1}$ (346).

[c] Rhizosphere medium (350).

All cultures were incubated for 12 hrs at 37 °C (DH5α) or 30 °C (*P. capeferrum* strains). After 12 hrs, cultures were centrifuged at 4000 *g*, and suspended in 500 µL PBS pH 7.0; 100 µL was set aside for cell viability analysis. Flow cytometry was performed on an Accuri C6 flow cytometer (BD Biosciences), by analyzing the percentage of GFP fluorescing cells vs. non-fluorescing cells (cells lacking the reporter vector). The following settings were applied: 10,000 events collected, with threshold settings FSC-H = 10,000 and SSC-H = 8,000.

Cell viability was also assessed by adding 5 µg/mL propidium iodide (PI) (BD Biosciences) solution to each 100 µL cell aliquot, then incubating the solution for 5 min at room temperature. Percentage of non-PI dyed (total cells) vs. PI-dyed cells (dead cells) were calculated. All flow cytometry data was analyzed with Accuri C6 software. Each set of conditions were assessed in triplicate, using different cell cultures for each repeat, and the results averaged amongst the three experiments. Significance of transcriptional activation as indicated by GFP expression was determined using a two-tailed, heteroscedastic student's t-test, wherein $p \leq 0.05$ is considered significant, when compared to unactivated (non-fluorescing) cells.

6.3. Results

6.3.1. Cloning and Purification of PupI and the PupI:PupR-ASD Complex

PupI was cloned into twelve different expression constructs by standard amplification of the PupI gene from *Pseudomonas capeferrum*, with addition of 5' and 3' restriction sites, for digestion and ligation into the vector of choice (Table 6.3). Successful clones were verified by sequencing through MCLab. Extensive cloning of PupI into multiple vectors allowed for a variety of conditions to be tested for expression of PupI as a fusion protein; constructs containing MBP, StrepII-SUMO, and N- or C-terminal His_6 tagged PupI were developed. Furthermore, numerous expression and purification conditions were assessed in an effort to produce stable

PupI for further analysis (summarized in 6.2.3 and 6.2.4). Only addition of detergent (FC-13, specifically) to solubilize PupI from inclusion bodies yielded soluble PupI, although efforts to remove the detergent by dialysis lead to pervasive protein precipitation. Thus, a low concentration of FC-13 (0.05% w/v) was maintained during purification of PupI.

Since purification of stable PupI alone was inadequate, as structural characterization in the presence of a detergent was undesirable, attention was shifted toward isolation of the PupI:PupR-ASD complex. As prior experiments demonstrated that the PupR-ASD was quite stable in solution, the working hypothesis was that co-expression and purification of PupI with the PupR-ASD would positively influence PupI stability. Co-expression trials of MBP-PupR-ASD or His$_6$-PupR-ASD with the PupI constructs listed in Table 6.3 led to identification of pHis$_6$-PupR-ASD with pET28a-StrepII-SUMO-TEV-PupI-His$_6$ as targets for co-purification of the complex. Affinity tags larger than His$_6$ or StrepII-SUMO appeared to prevent complex formation. Additionally, the N-terminal affinity tags could be easily removed from both proteins in the first purification step without disrupting the PupI:PupR-ASD complex (Figure 6.5).

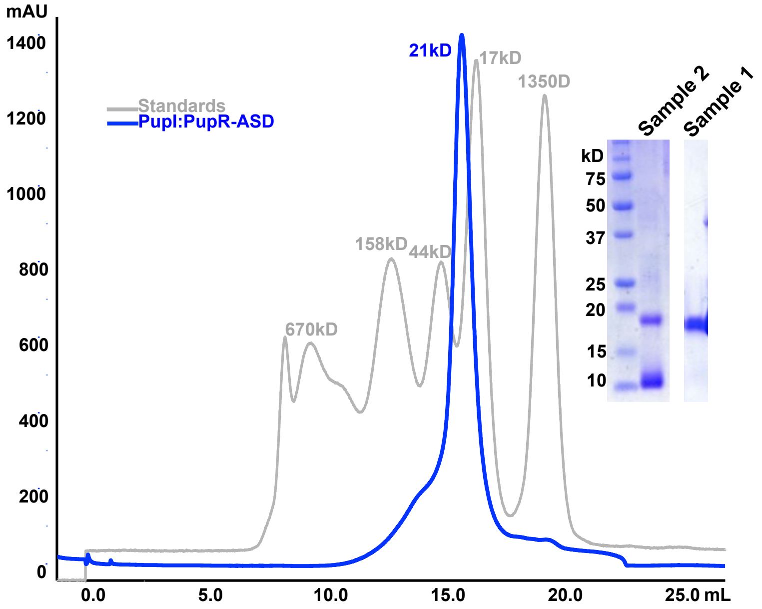


Figure 6.5. SEC chromatogram of the PupI:PupR-ASD complex. The complex exhibits delayed elution over a Superdex 200 (10/300), which was unexpected; the expected MW of the complex is ~31 kD. The theoretical MW of the PupR-ASD is 9.5 kD, and the MW of PupI-His$_6$ is 20.7 kD. The three lanes of the SDS-PAGE contain (L to R) the molecular weight marker, SEC-SAXS sample 2, and SEC-SAXS sample 1.

The results from N-terminal sequencing did not ambiguously demonstrate that StrepII-SUMO was being separated from PupI-His$_6$ during purification. The samples excised from the amido black-stained PVDF membrane were from the lower part of the band in lane 5 (Figure 6.6) for the StrepII-SUMO sample, and from the primary band around 20 kD in lane 7. Both samples presented with a weak signal and high background during the Edman degradation reaction, and the signal from cycle 4 could not be identified and subsequent reactions resulted in multiple possibilities for residue identification. Sequential degradation of the sample thought to contain StrepII-SUMO yielded the following sequence of amino acids from the N-terminus: glycine, serine, alanine (or methionine), unknown, and serine (or phenylalanine or leucine). The N-terminal residues of the StrepII-SUMO tag are methionine, glycine, serine, alanine, tryptophan, and serine, indicating that the StrepII-SUMO tag was most likely present in the sample. Sequential degradation of the second sample containing PupI-His$_6$ gave the following calls from the N-terminus: glycine or alanine, serine, alanine (or methionine or proline), unknown, and proline or leucine. The N-terminal residues of the PupI-His$_6$, after TEV proteolysis during purification, were glycine, serine, methionine, leucine, and proline. Again, although there is some ambiguity in determination of the residues due to weak signal, the second sample appears to be PupI-His$_6$. Thus, while StrepII-SUMO is a considerable contaminant during purification of PupI-His$_6$, it is mostly removed by SEC.

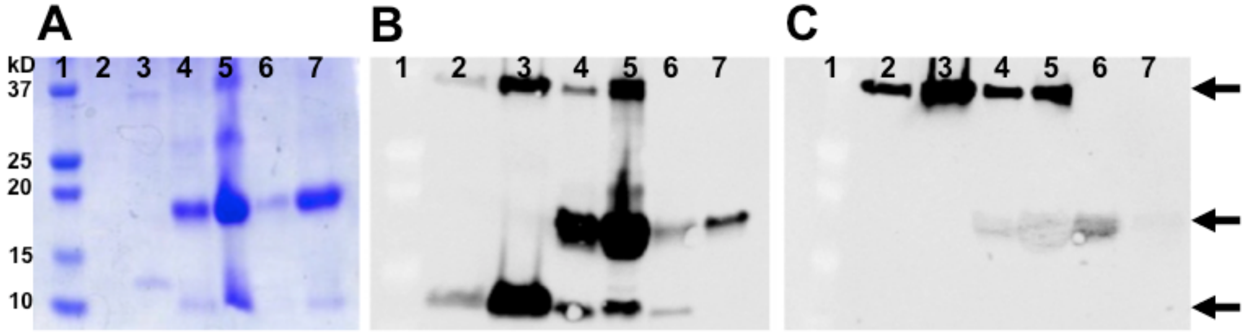


Figure 6.6. SDS-PAGE and Western blotting of protein samples collected throughout purification of the PupI:PupR-ASD complex. Lanes: 1. Molecular weight marker (All-Blue Precision Plus, Bio-rad), 2. Pre-induction sample, 3. After-induction sample, 4. After addition of TEV, 5. SEC load, 6. 1st SEC elution fraction, 7. 2nd SEC pool. (A) SDS-PAGE stained with Coomassie Blue. (B) Western blot probed with Anti-His$_6$-HRP. (C) Western blot probed with Anti-StrepII-HRP. The arrow at 37 kD indicates StrepII-SUMO-TEV-PupI-His$_6$ (~35 kD); StrepII-SUMO and PupI-His$_6$ both have an apparent molecular weight <20 kD (theoretical molecular weight of StrepII-SUMO = ~15 kD; PupI-His$_6$ = ~21 kD); His$_6$-PupR-ASD is ~ 11 kD.

6.3.2. Crystallization Trials of PupI and PupI:PupR-ASD

To investigate the structural details of PupI and the PupI:PupR-ASD complex, crystallization trials were initiated. Crystallization conditions for both PupI alone and PupI:PupR-ASD were screened using two commercially-available sparse matrix screens, along with three different ratios of protein to reservoir solution, for a total of 576 conditions. However, no crystals were observed in any of the crystallization conditions.

6.3.3. PupI Domain Architecture and Homology Model

Similar to other ECF sigma factors, PupI contains two conserved domains, which incorporate two regions involved in DNA binding and RNA recruitment (Figure 6.7). ECF sigma factors retain σ_2 and σ_4 domains; σ_2 contains the polymerase core binding region, and σ_4 harbors the DNA binding domain, which includes the predicted helix-turn-helix motif. Secondary structure predictions of the sigma factors of the TonB-dependent transducer-type suggest that they contain eight helices.

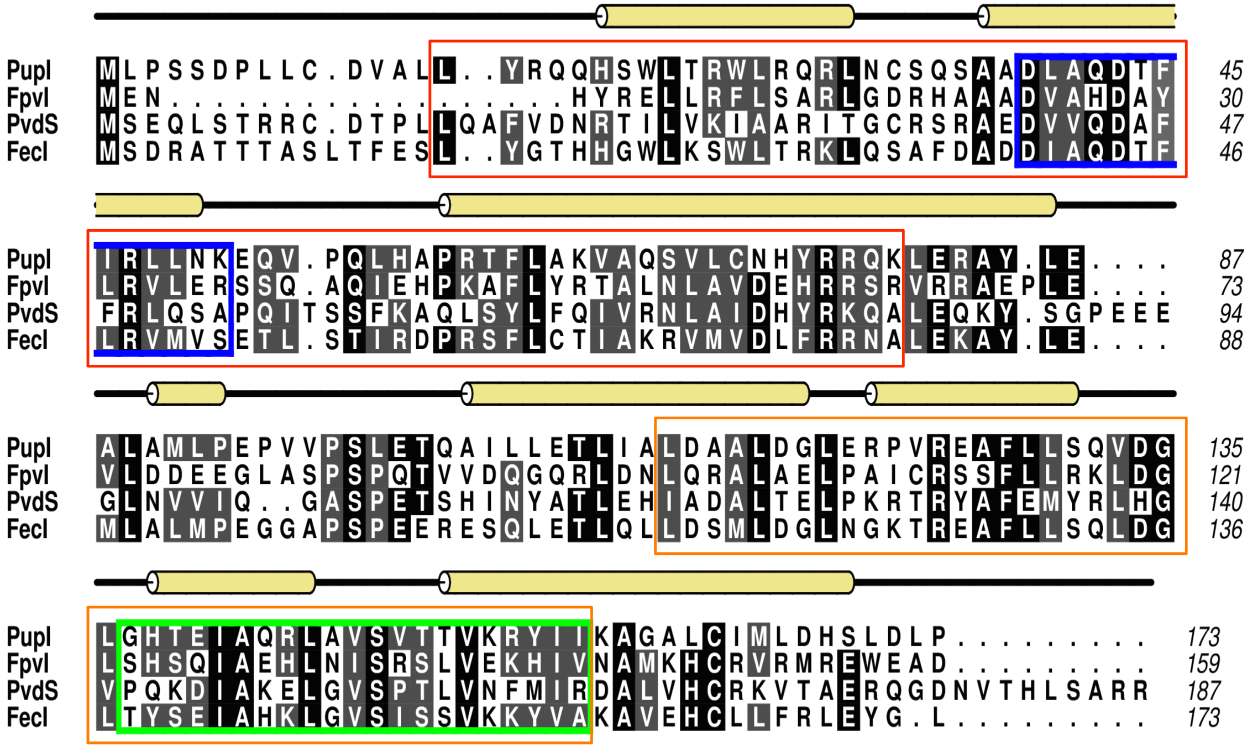


Figure 6.7. Sequence alignment of PupI with orthologs FpvI and PvdS from *P. aeruginosa* and FecI from *E. coli*. The alignment was performed with T-COFFEE, and the secondary structure prediction with Jnet. The gold cylinders indicate predicted helical regions. The region outlined in blue is the polymerase core binding region, and the area outlined in green is the DNA binding domain (helix-turn-helix motif). PupI shares 30.2% identity and 54.9% similarity with FpvI, 27.2% identity and 47.8% similarity with PvdS, and 42.4% identity and 67.5% similarity with FecI. The red outline indicates the predicted σ_2 domain of PupI, and the orange outline is the predicted σ_4 domain of PupI.

Secondary structure prediction of PupI based on the primary sequence yielded a principally helical model, consistent with secondary structure predictions of other ECF sigma factors, and existing X-ray crystal structures of σ^{70} factors (Figure 6.3). Comparison of the PupI secondary structure prediction to the secondary structure of σ^{E} (291), which is the most complete crystal structure of a σ^{70} factor available, demonstrated reasonable agreement between PupI and σ^{E}.

To create a homology model of PupI, MODELLER through ModWeb was utilized. Four structures were identified by ModBase as possible structural homologs, and these served as the basis for the construction of the PupI model (Figure 6.3). Although several of the criteria for formulation of a reliable 3D model were not met (Table 6.2), MODELLER produced two homology models for PupI, which were defined by the program as reliable, despite the template structures having <30% sequence identity to the target. The first PupI model (Figure 6.8A) covers residues 5-163 of 173, and presents as an all-α structure, with two three-helix bundles (helices 1-3 and helices 6-8) linked by an extended coil, flanked by two more helices (helix 4 and helix 5). The second PupI model (Figure 6.8B) displays a more compact structure overall, but lacks residues 1-38 and 159-173. A total of six helices are exhibited, with a flexible linker between helices 2 and 3.

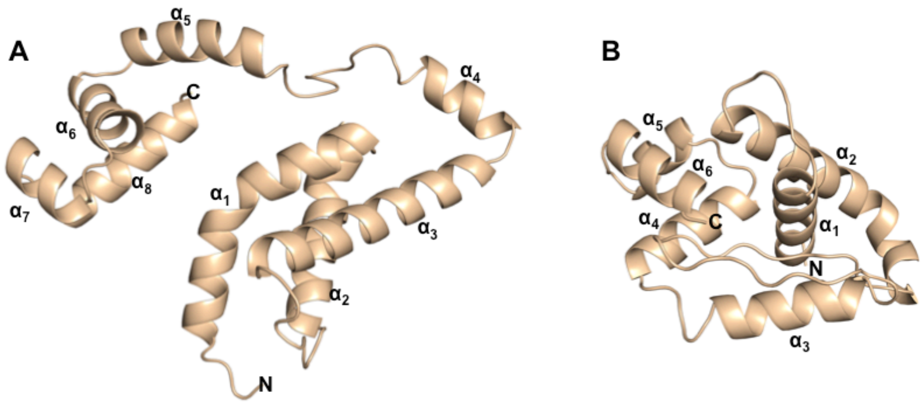


Figure 6.8. PupI homology models generated by MODELLER. (A) First PupI model (MODELLER #1) is a more elongated structure, and consists of PupI residues 5-163, with an extended loop between helices 4 and 5. (B) Second PupI model (MODELLER #2) comprises a more compact, globular structure of only 6 helices, with a loop between helices 2 and 3.

6.3.4. Solution Characteristics of the PupI:PupR-ASD Complex and Comparison to PupI Models

SEC-SAXS was employed to probe the solution properties of the PupI:PupR-ASD complex, purified from co-expression of His_6-PupR-ASD and StrepII-SUMO-TEV-PupI-His_6. Two separate samples of the complex were examined; in the first, only PupI-His_6 was distinguishable in the SDS-PAGE prior to analysis. It was unknown if this was a result of degradation of the PupR-ASD, minimal staining of the species after SDS-PAGE, or suboptimal binding of the PupR-ASD to PupI. However, it was anticipated that molecular weight estimations by SEC-SAXS would help to resolve this issue. The second SEC-SAXS sample contained both PupI and the PupR-ASD, as verified by SDS-PAGE.

Analysis of the SAXS data from both samples indicated that the Guinier plot was linear in the low q range (Figures 6.9C and 6.10C, inset), so both samples were monodisperse. The R_g values derived from the Guinier plot were 23 Å and 25 Å for samples one and two, respectively. D_{max} values derived from the $P(r)$ plot were 91.7 Å and 87.5 Å (Figures 6.9D and 6.10D). Scrutiny of the Kratky plot (Figures 6.9B and 6.10B) revealed that the samples are partially flexible and may contain multiple domains with flexible linkers, consistent with the PupI homology models. This phenomenon is also reflected in the *P(r)* function (Figures 6.9D and 6.10D), where the extended tail at high *r* is indicative of an elongated molecule.

Most interestingly, the molecular weight estimations from DAMMIF were 22 kD for sample one, and 32 kD for sample two. The expected molecular mass of PupI alone is ~21 kD, while the theoretical molecular mass of the PupR-ASD is ~10 kD. Taken with the apparent molecular mass of the species present in the samples by SDS-PAGE (Figure 6.5), these results suggest that the primary species in sample one is PupI, while sample two contains the PupI:PupR-ASD complex.

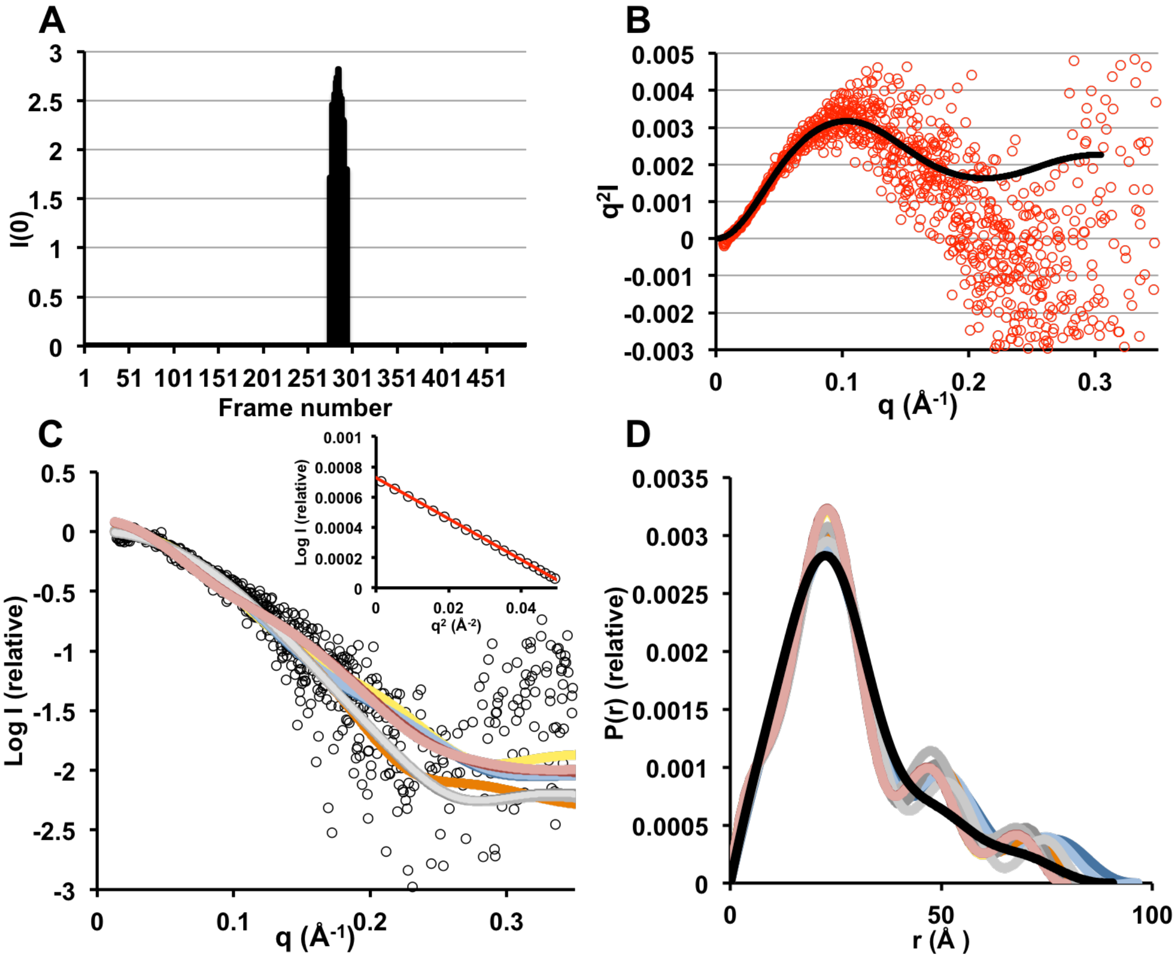


Figure 6.9. SEC-SAXS analysis of the first PupI sample. (A) Trace of scattering intensity, I(0) vs. frame number, following buffer subtraction. (B) Kratky plot calculated from the experimental scattering profile. (C) Experimental scattering profile (open circles), fit with the theoretical scattering profiles (lines) of MODELLER PupI model #1 (yellow), MODELLER PupI model #2 (green), MultiFoxS PupI single-state model (navy), MultiFoxS PupI two-state model (light blue), σ^E:RseA (PDB 1OR7) (dark orange), σ^W:RsiW (PDB 5WUR) (light orange), MODELLER PupI model #1:PupR-ASD crystal structure (dark grey), MODELLER PupI model #1:PupR-ASD EOM model 1 (medium grey), MODELLER PupI model #1:PupR-ASD EOM model 2 (light grey), PupI single-state model:PupR-ASD crystal structure (red), and PupI two-state model:PupR-ASD crystal structure (pink). (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log(intensity). (D) Distance distribution *P(r)* for the experimental data (black), and the theoretical scattering of the eleven models described above.

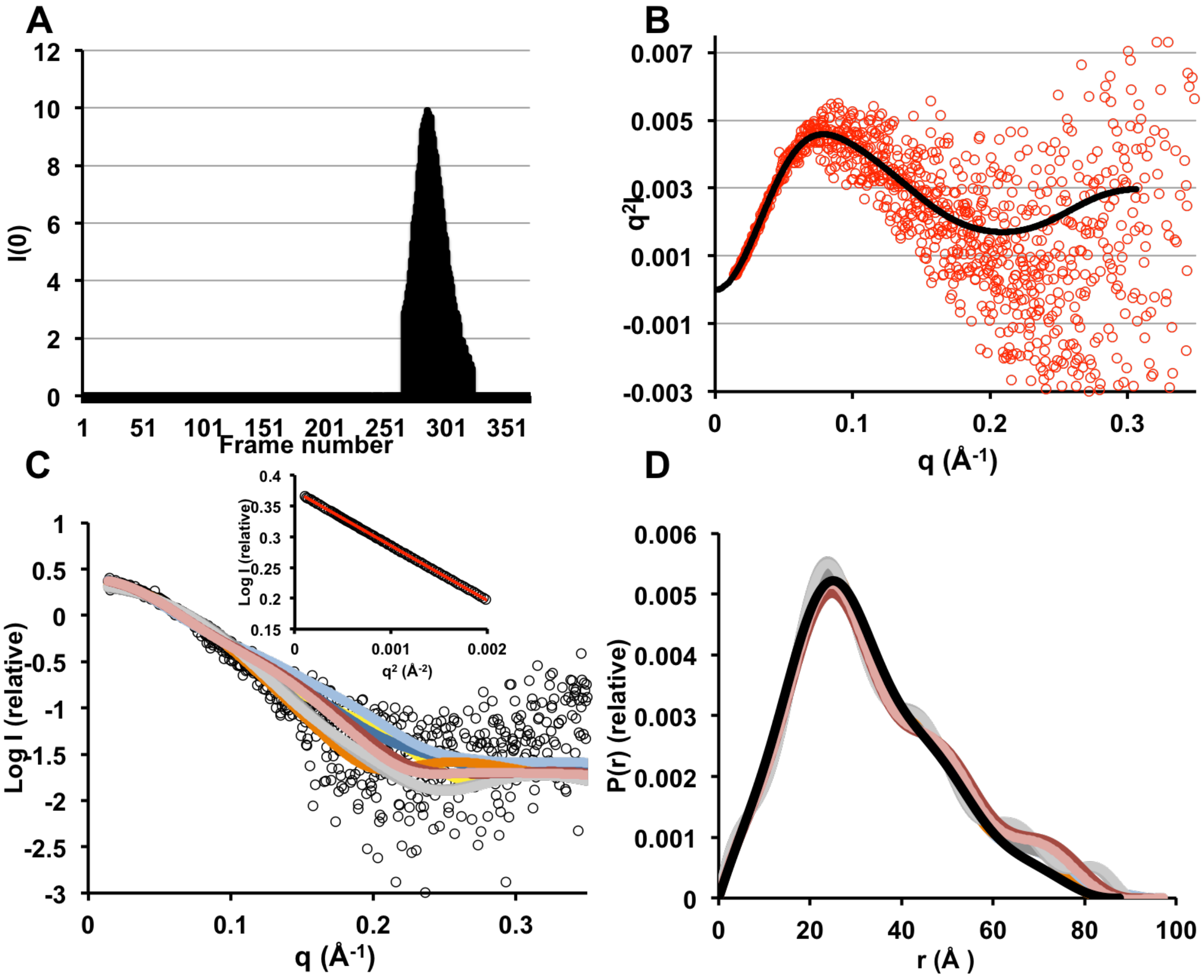


Figure 6.10. SEC-SAXS analysis of the second PupI sample, containing the PupI:PupR-ASD complex. (A) Trace of scattering intensity, I(0) vs. frame number, following buffer subtraction. (B) Kratky plot calculated from the experimental scattering profile. (C) Experimental scattering profile (open circles), fit with the theoretical scattering profiles (lines) of MODELLER PupI model #1 (yellow), MODELLER PupI model #2 (green), MultiFoxS PupI single-state model (navy), MultiFoxS PupI two-state model (light blue), σ^E:RseA (PDB 1OR7) (dark orange), σ^W:RsiW (PDB 5WUR) (light orange), MODELLER PupI model #1:PupR-ASD crystal structure (dark grey), MODELLER PupI model #1:PupR-ASD EOM model 1 (medium grey), MODELLER PupI model #1:PupR-ASD EOM model 2 (light grey), PupI single-state model:PupR-ASD crystal structure (red), and PupI two-state model:PupR-ASD crystal structure (pink). (Inset) Guinier plot of the low *q* region of the X-ray scattering data, as a function of log(intensity). (D) Distance distribution *P(r)* for the experimental data (black), and the theoretical scattering of the eleven models described above.

Several structural models of PupI and the PupI:PupR-ASD complex were constructed for comparison to the collected SAXS data. These included two homology models (MODELLER model #1 and #2, Figure 6.8A-B), a one-state or two-state model from MultiFoxS (Figure 6.11A), MODELLER PupI model #1 with the PupR-ASD crystal structure (Figure 6.11B), MODELLER PupI model #1 with PupR-ASD EOM model 1 (Figure 6.11C), MODELLER PupI model #1 with PupR-ASD EOM model 2 (Figure 6.11D), PupI one-state model with PupR-ASD crystal structure (Figure 6.11E), and PupI two-state model with PupR-ASD crystal structure (same as Figure 6.11B). Multi-state modeling of PupI with MultiFoxS indicated that the extended conformation of PupI, shown in Figure 6.11A, is only sampled 5.3% of the time, whereas the more globular PupI model (Figure 6.8A and Figure 6.11B- shown with PupR-ASD) is sampled 94.7%, and is therefore the predominant species in solution. The R_g, D_{max}, and fit of the theoretical scattering from the model to the experimental scattering were calculated for each model (Table 6.5). These values were also determined for the X-ray crystal structures of the σ^E:RseA complex (PDB 1OR7) and σ^W:RsiW complex (PDB 5WUR). Additionally, representative PupI or PupI:PupR-ASD complex models were fit with SUPCOMB to the pseudo-atomic bead models generated by DAMMIF (Figure 6.12). Comparison of R_g, D_{max}, and fits (χ or χ^2 values) of the theoretical scattering of the models listed above to the experimental scattering did not expose significant differences between the various models. Further analysis is required to fully corroborate the SEC and SDS-PAGE output with the SEC-SAXS results.

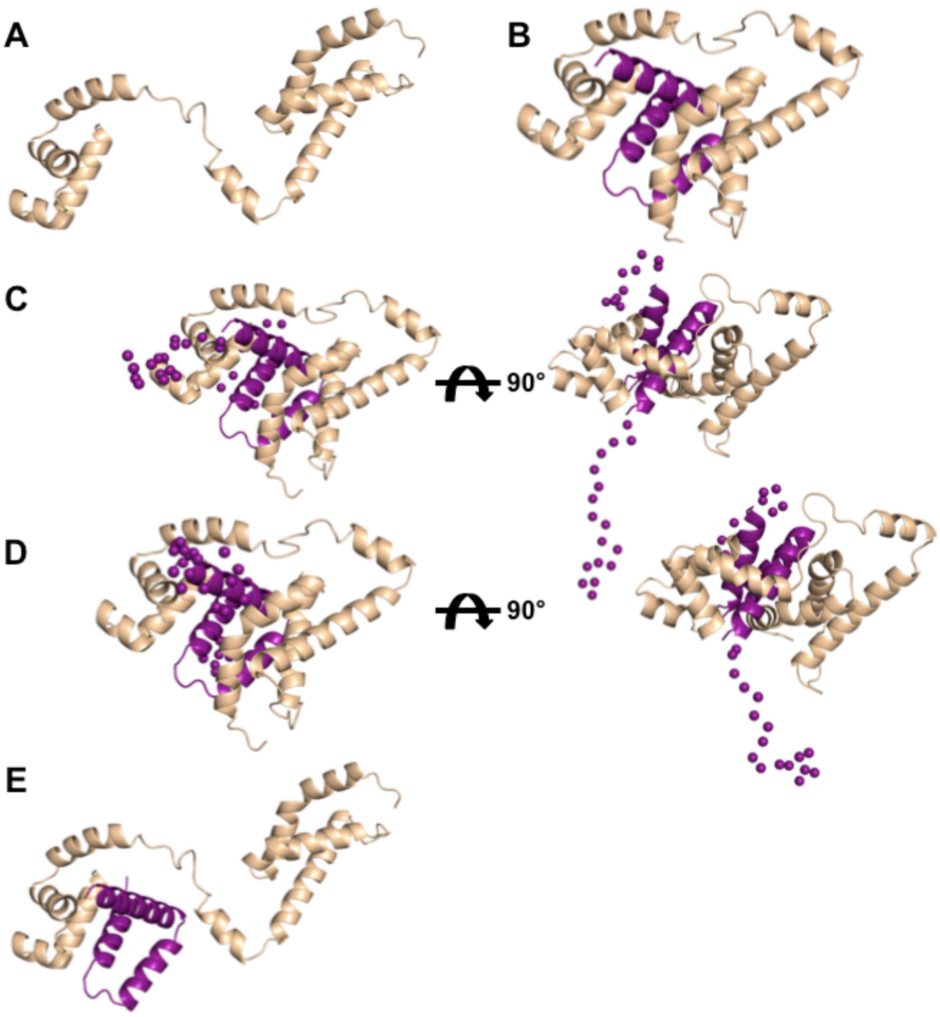


Figure 6.11. PupI or PupI:PupR-ASD complex models created for comparison to SAXS data. The PupI is shown in pale yellow, and the PupR-ASD in purple. (A) The dominant PupI species from multi-state modeling, based on MODELLER PupI model #1, and assuming flexibility of the linker between the two domains. (B) MODELLER PupI model #1 with the PupR-ASD crystal structure (PDB 5COS chainA). (C) MODELLER PupI model #1 with PupR-ASD EOM model #1, which creates an ensemble of atomic models to fit the SAXS data, including modeling of flexible regions not included in the crystal structure (in this case, the N- and C-termini, shown as purple spheres). (D) MODELLER PupI model #1 with PupR-ASD EOM model #2. (E) PupI one-state model with PupR-ASD crystal structure.

Table 6.5. Comparison of model sizes and fits of theoretical scattering of models to experimental scattering.

Model	R_g (Å)	D_{max} (Å)	Fit to scattering profile from Sample 1 (χ)	Fit to scattering profile from Sample 2 (χ)
Sample #1 (Experimental)	23	91.7	--	--
Sample #2 (Experimental)	25	87.5	--	--
MODELLER #1	23	80.6	1.78	1.60
MODELLER #2	23	83.5	1.77	1.59
One-state Model PupI	26	96.3	1.64	1.55
Two-state Model PupI	25	96.5	1.73	1.87
MODELLER #1:PupR-ASD	26	85.7	1.65	1.67
MODELLER PupI model #1:PupR-ASD EOM model 1	27	90.7	1.65	1.52
MODELLER PupI model #1:PupR-ASD EOM model 2	27	90.6	1.65	1.53
PupI one-state model:PupR-ASD	27	94.9	1.87	1.48
PupI two-state model:PupR-ASD	27	95.6	1.88	1.42
σE:RseA	26	93.7	1.73	1.43
σW:RsiW	22	84.5	1.67	1.39

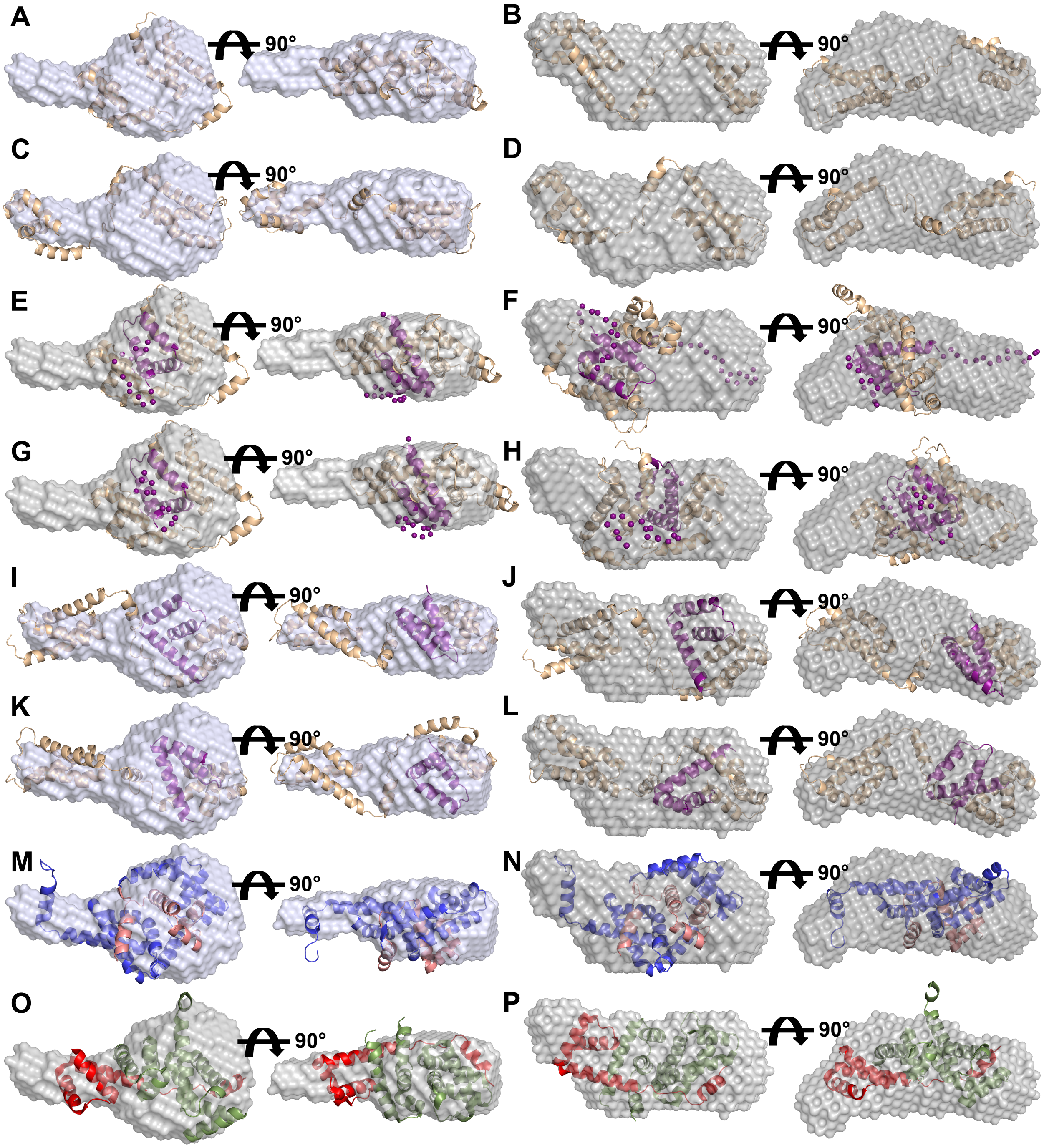


Figure 6.12. Pseudo-atomic models of experimental SAXS data fit with PupI or PupI:PupR-ASD structures. (A, C, E, G, I, K, M, O) Bead model derived from SAXS sample #1. (B, D, F, H, J, L, N, P) Bead model derived from sample #2. (A-B) Fit to MODELLLER PupI model #1. (C-D) Fit to one-state PupI model. (E-F) Fit to MODELLER PupI model #1 with PupR-ASD EOM model #1. (G-H) Fit to MODELLER PupI model #1 with PupR-ASD EOM model #2. (I-J) Fit to MODELLER PupI model #1 with PupR-ASD crystal structure. (K-L) Fit to PupI one-state model with PupR-ASD crystal structure. (M-N) Fit to σ^E:RseA complex (PDB 1OR7). (O-P) Fit to σ^W:RsiW complex (PDB 5WUR).

6.3.5. Probing the Response by PupI in *P. capeferrum* to Various Growth Conditions

Four strains of *P. capeferrum*, plus *E. coli* DH5α cells, were monitored for transcriptional activation by PupI via flow cytometry. For each strain, twelve conditions were assessed, including non-transformed and transformed cells containing the reporter vector, two types of growth media, and addition of either 100 μM $FeCl_3$ or 20 μM pseudobactin. These experiments were carried out to analyze the function of PupI in its native environment, under iron-rich and iron-limiting conditions. It was hypothesized that PupI would activate expression of $PupB_{128}$-GFP under iron-deplete conditions, through association of PupI with the *pupB* promoter. KBM and RSM are two different *Pseudomonas* growth media, typically utilized for siderophore binding assays (57,140,346), and were incorporated in this analysis to examine two different growth media; iron is a component of neither media. Both KBM and RSM were supplemented with either $FeCl_3$ or isolated pseudobactin. Following incubation of cells in the specific growth conditions for 12 hrs, cell populations were analyzed for relative GFP fluorescence of reporter-containing cells, compared to non-transformed cells. Cell viability was also determined as the percentage of dead vs total cells.

The three mutant *P. capeferrum* strains contained replacements for genes *pupI*, *pupB*, and *pupR*, and thus, were presumed to function similarly to genomic deletions, wherein the gene of interest is deleted, and is not expressed. Transcriptional activation by PupI, indicated by GFP expression was significant for all four strains of *P. capeferrum* grown in KBM, when GFP expression of cells transformed with the reporter vector is compared to GFP expression of non-transformed cells (Figure 6.13). Interestingly, cells with *pupI* and *pupR* replacements exhibited similar GFP expression levels. This was unexpected for the *pupI* replacement, wherein the absence of PupI should have rendered the cells transcriptionally inactive. Results for the *pupR* replacement were as expected, given that the absence of PupR would lead to unregulated PupI. The *pupB* replacement still displayed GFP expression, although expression levels were lower than observed for WT *P. capeferrum*.

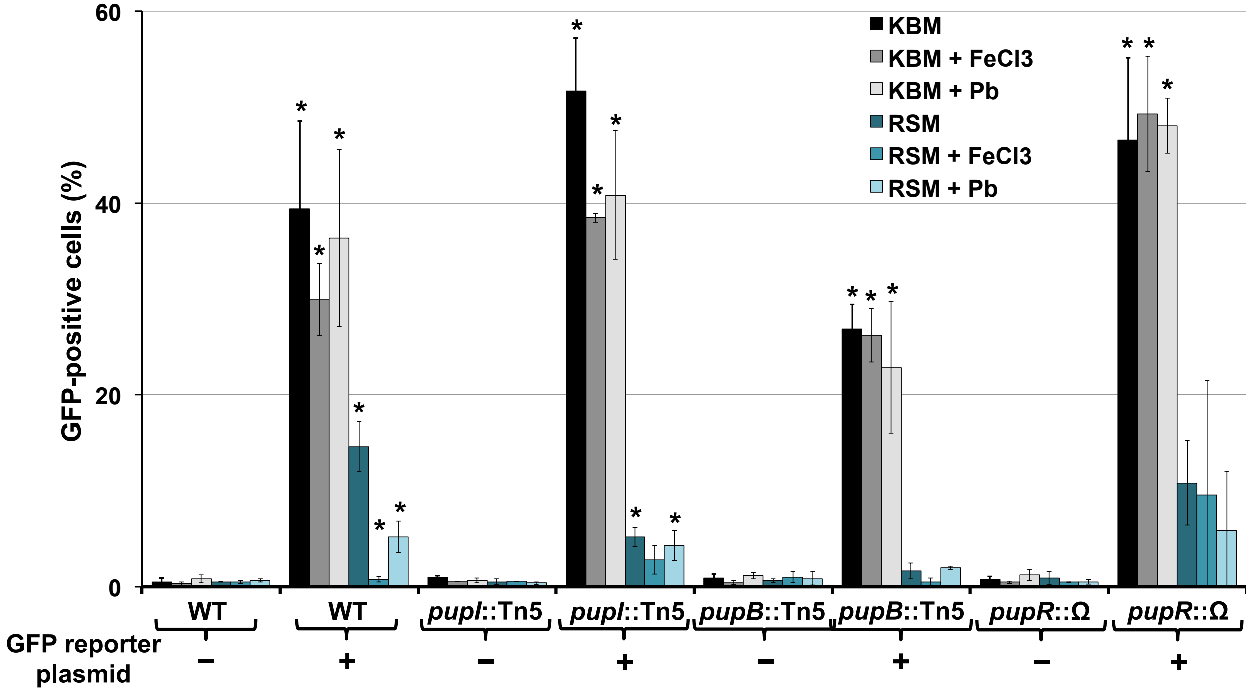


Figure 6.13. PupI-induced expression of the PupB promoter-GFP reporter gene in *P. capeferrum*. WT *P. capeferrum* and three genomic mutants were analyzed, with cells lacking the reporter vector used as controls. Each strain was grown in either KBM or RSM, with addition of $FeCl_3$ or pseudobactin. Series labeled "+GFP" indicate the cells were transformed with the reporter vector. Each column is the average of three experiments. *$p \leq 0.05$ was considered significant.

Unexpectedly, cells grown in RSM displayed much lower levels of GFP expression than cells grown in KBM. RSM was developed to imitate the environment of soil-dwelling bacteria, so it is formulated with standard macroelement composition of soil at field capacity water content. The medium includes zinc at a final concentration of 7.0 x 10^{-4} mM. Fur, which regulates expression of PupI and PupR, is a zinc metalloprotein, and Zn^{2+} is essential for Fur to bind DNA (351,352). Under iron-limiting conditions, it is plausible that $Zn^{2}+$ may occupy the site typically populated by Fe^{2+}, which would repress expression of PupI and PupR. It is unknown what cellular concentrations of iron are necessary to replace Fur-bound zinc, and at what concentration zinc serves to repress iron acquisition genes, as low intracellular iron levels typically cause Fur dissociation from DNA promoters.

Viability of cells was demonstrated by staining cell populations with PI, which binds exposed DNA from dead or dying cells. All samples analyzed exhibited >80% viability (Figure 6.14).

Transformation of *E. coli* DH5α cells with the pSEVA-531-RBSD-GFP reporter vector was verified by cell growth in 10 μg/mL tetracycline. Conditions outlined in Table 6.4 were examined for transcriptional activation. As expected, no GFP fluorescence was detected (data not shown), as *E. coli* lacks PupI or a close ortholog that could induce gene expression via association with the reporter vector.

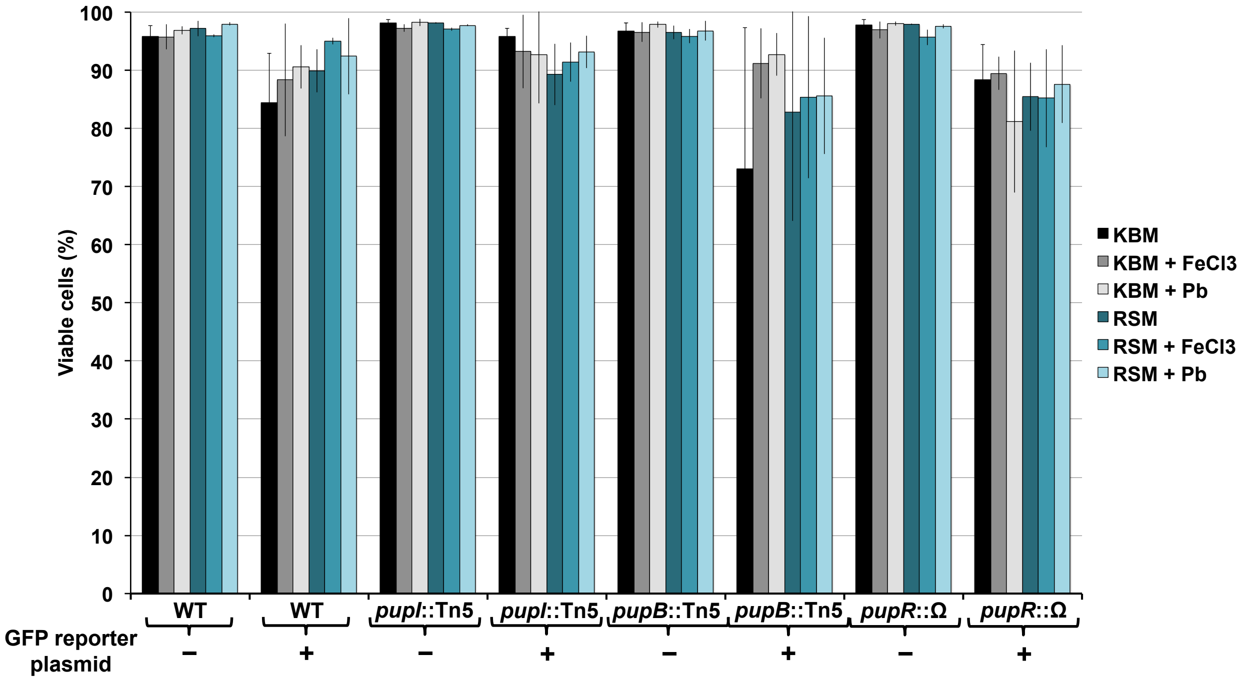


Figure 6.14. Viability of cells utilized for the reporter assay. Cell viability (average of 3 separate experiments) was calculated as the difference between total viable cells and the number of dead cells vs. total cells.

6.4. Summary and Conclusions

After numerous trials and errors, protocols to over-express and purify homogeneous PupI and PupI:PupR-ASD complex have been developed. Although PupI alone appears quite unstable in solution, association with the PupR-ASD significantly improves PupI solubility. These efforts allowed for preliminary solution characterization of the PupI:PupR-ASD interaction by SEC-SAXS.

One of the outstanding questions of this work is the stoichiometry of the sigma factor:ASD complex. The possible stoichiometry observed from SEC-SAXS of PupI:PupR-ASD does not correlate to data from the free PupR-ASD dimer. While this was unexpected, given the extensive characterization of the PupR-ASD dimeric interface, it is not entirely without precedent. The dimeric form of PupR-ASD was the first report of a transmembrane anti-sigma factor with a dimeric anti-sigma domain. As discussed in Chapter 5, all known ASD:sigma factor structures exhibit a heterodimeric complex. However, ASiA, encoded by bacteriophage T4, is one example of a dimeric anti-sigma domain dissociating to form a 1:1 complex with its cognate sigma factor (353).

The anti-sigma factor AsiA interacts with the σ^{70} subunit of RNAP, which is a critical step in transcriptional activation during T4 infection (353). The structure of AsiA is a small, six-helix bundle (or five-helix bundle with an additional single-turn helix) of only 90 residues (354). It is localized to the cytoplasm, and only shares sequence similarity with other phage proteins or structural similarity with other DNA-binding proteins. AsiA has been shown to interact with the C-terminal region of σ^{70}, which contains the σ_4 domain. By binding to the σ_4 domain, AsiA weakens the association of σ^{70} to RNAP. In solution, AsiA forms a stable dimer; upon titration of σ^{70}-derived peptides into AsiA, the dimer dissociates to accommodate σ^{70} at the dimeric interface

(355). Thus, while dimer reorganization is energetically unfavorable, the high affinity of σ^{70} binding drives dimer dissociation (353,354). This is in agreement with prior functional data, which depicted a heterodimeric AsiA: σ^{70} complex (356). The hydrophobic core present within the AsiA dimer is exchanged with the sigma factor for minimal solvent exposure.

Thus far, there is limited experimental evidence to suggest that PupI:PupR-ASD association mimics that of AsiA:σ^{70}; however, a couple of factors advance this hypothesis. Dimerization of ECF sigma factors has not been observed, and, in the case of the *pupIRB* operon, there is not a repetitive sequence in the PupB promoter that would be indicative of a recognition sequence for dimeric PupI, as observed for the Fur box- a known dimeric (or higher oligomer) sigma factor.

The question becomes whether PupI binds to a PupR-ASD monomer or dimer. Sequence alignment and secondary structure prediction of PupI with similar ECF sigma factors and sigma factors of known structure demonstrate conservation of secondary structure. Subsequent homology modeling provided a reasonable tertiary structure prediction for PupI. Comparison of the two available X-ray crystal structures of a σ^{70} factor:anti-sigma domain complex reveal two binding modes for the complex (Figure 6.15). In the first mode (Figure 6.15A), the sigma factor wraps around the ASD: in the second mode (Figure 6.15B), the reverse is observed- the ASD wraps around the sigma factor. In both of these cases, the region of the sigma factor corresponding to the σ_4 domain, interacts with the ASD, and has also been demonstrated to interact with the -35 promoter sequence, as shown for the Fec system (315). Superposition of the PupR-ASD dimer with the core ASD in both Figure 6.15A and 6.15B reveal that the dimeric interface is the same interface that supports sigma factor:ASD association. Therefore, the ASD dimer would need to undergo significant rearrangement to accommodate a similar association.

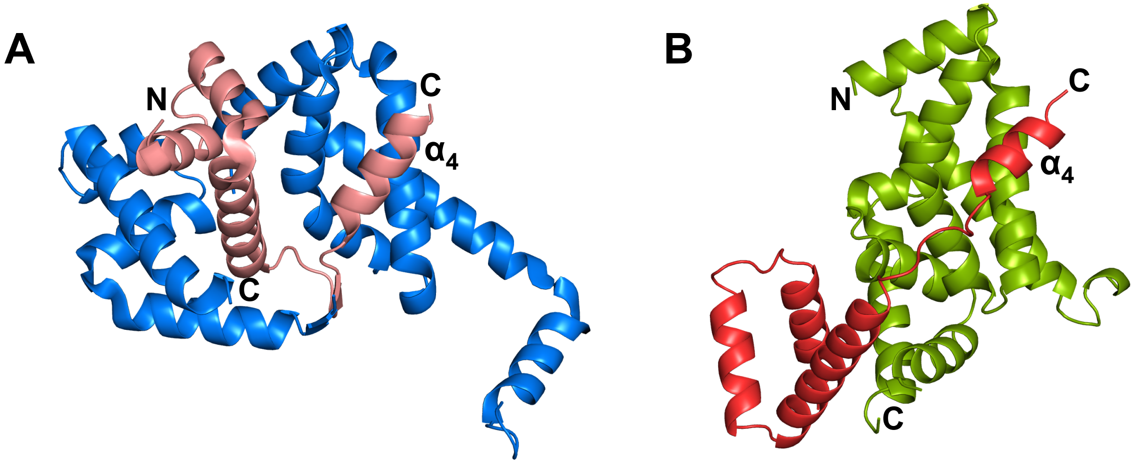


Figure 6.15. Available X-ray crystal structures of sigma factor:anti-sigma domain complexes. (A) σ^E:RseA (PDB 1OR7) (291); RseA is displayed in blue, and σ^E in salmon. (B) σ^W:RsiW (PDB 5WUR) (357); σ^W is shown in green, and RsiW in red. Both sigma factors are missing the flexible linker between the two sigma domains. Note: structures are not superimposed, due to signification structural differences between the two sigma factor.

Initial (unpublished) work by Stiefel, *et al.* suggested that FecR may form a dimer or oligomer, though whether this was a result of aggregation, what domains may commence oligomerization, or how interconversion proceeds between a monomer and oligomer were not proposed (358). Dimerization does not occur through the transmembrane region, as revealed by cysteine scanning (359). Although dimerization of free PupR-ASD was observed *in vitro* in the absence of PupI, the physiologically-relevant oligomerization state of this domain is unknown; whether PupR and PupI (as monomers or oligomers) exist independently in the cell, or if their translation is coupled is also presently ambiguous. As expression of both *pupI* and *pupR* are controlled by Fur repression via the same promoter, transcriptional coupling seems likely.

To extrapolate to the PupI:PupR-ASD interaction, SEC-SAXS data suggests a 1:1 association. This would correlate with all available examples of ECF:ASD interactions. Additionally, either wrapping modes 1 or 2 described above could be implemented in the

PupI:PupR-ASD interaction, as these regions on PupI are not highly charged, as presented in the PupI homology models, so the hydrophobic core exposed upon PupR-ASD dissociation would remain buried in the heterodimer.

The sequence alignments and homology modeling of PupI will facilitate future efforts to obtain soluble domains of PupI and PupI:PupR-ASD complexes for further structure and functional analyses. Continued optimization of purification conditions are required in order to obtain enough protein for crystallization trials, as the concentrations of the protein stocks used for crystal screening were not sufficient to initiate crystal nucleation.

Further analysis is required to discern the mechanism of transcriptional activation by PupI *in vivo.* Although some general trends observed in GFP expression levels by PupI activation were as expected, several findings are, as of yet, unexplainable by current data. Decreased transcriptional activation by PupI under iron-limited conditions was unanticipated, and does not align with the model of signaling through PupI that is prevalent in the literature (101,102,140,150,154,185,282,284,317,360). Previous work by Koster, *et al*. that utilized the *P. capeferrum* genomic mutants reported that KV51, the *pupI*::Tn5 mutant, displayed limited *pupB* promoter activity. Additionally, the two papers that cite usage of KV51 report the mutation as *pupI*::Tn5 in one publication, and *pupR:*:Tn5 in another (57,140); modern high-throughput DNA sequencing may resolve the exact gene replacement region. Lastly, gene replacement of any of the *pup* transport or regulatory genes severely limited, but did not completely eliminate, *pupB* promoter activity. *P. capeferrum* and other microorganisms have many ECF sigma factors, and corresponding regulators; they may also have with analogous regulatory pathways that perform redundant functions.

7. CONCLUSIONS, PERSPECTIVES, AND FUTURE DIRECTIONS

CSS pathways in Gram-negative bacteria are highly conserved, and allow the cell to respond rapidly and efficiently to extracellular stimuli. These pathways involve three distinct proteins: the OM transducer, the IM sigma regulator, and the ECF sigma factor. Interaction of the proteins in a semi-conserved, semi-sequential manner results in transcription of stimulus response genes, namely, the OM transducer. In the model studied here, the *P. capeferrum* ferric BN7/BN8 signaling axis includes the OM transducer PupB, the IM sigma regulator PupR, and the sigma factor PupI. The ultimate objective of the work described in this dissertation was to obtain structural details that account for the regulation of this conserved bacterial iron acquisition pathway, a critical bacterial survival process and target for innovative antimicrobial compounds.

A search of the annotated *P. capeferrum* genome revealed that, in general, transducers harbor an NTSD (STN domain), with a cognate sigma regulator and sigma factor for transcriptional activation of the transducer through signal transfer; transporters appear to lack these regulatory elements. Since transporters lack the sigma regulator and sigma factor required for transcriptional activation of the transporter gene, it is unclear why a transporter might possess an NTSD when it is transcriptionally inactive. Two TBDTs were identified which retained an STN, but were deficient of any apparent associated regulatory proteins. One of these proteins, PupA, was purified to homogeneity, and crystallization of the transporter achieved. Although well-diffracting crystals were not procured, several detergents were identified with a differential detergent filtration assay, to be used for maintaining protein stability and solubility, and to improve protein packing to facilitate crystallization.

Solution characterization by SEC-SAXS of the OM transporter PupA and the OM transducer PupB was performed to delineate low-resolution structural features that contribute to selectivity of PupB for signaling through PupR. Previously, minimal structural information was available to corroborate signal specificity through explicit transducers and regulators in the context of the full-length OM protein, and how structural variation may allow specialized transducers to signal through distinct sigma regulators. This work is the first to demonstrate that a full-length, detergent solubilized OM transducer interacts *in vitro* with the C-terminal region of a sigma regulator. SAXS analysis also established that the linker region between the NTSD and the transporter domains containing the TonB box exhibits greater flexibility for the transducers, such as PupB, compared to the PupA transporter. Specifically, the linker region of the PupB transducer and PupA$^{\text{PupB-NTSD}}$ chimeric transducer extended an additional 20-30 Å beyond that observed for PupA. Qualitative analysis of the flexibility of the transporter and transducers suggested that the transducers display greater conformational flexibility than the transporter. Higher resolution structural studies are necessary to further describe global changes that occur to the OM proteins during signaling. Optimization of crystallographic conditions, as well as solution experiments such as small-angle neutron scattering (SANS) would provide contrast variation for differentiation between detergent and protein. As demonstrated for related β-barrel proteins, SANS measurements can provide more detailed information with respect to orientation of flexible protein components, while masking the signal from one solution component (such as the detergent) to obtain a more accurate solution profile with minimal scattering contributions from the detergent (361).

The NTSD is essential for signal transduction through the TBDT regulatory pathway. A comparative analysis of the high-resolution structures of the PupB-NTSD, determined by NMR and X-ray crystallography, with existing NMR structures of the PupA-NTSD and FecA-NTSD, indicated that the NTSD from PupB also retains the βαβ-repeat motif observed in the other

NTSD structures. The PupB-NTSD and PupA-NTSD exhibit very similar secondary structure elements, and share 33% sequence identity; thus, differences in primary sequence between the two NTSDs must account for the signaling competency of the PupB-NTSD. Over fifteen single mutations of PupA-NTSD were examined, in which the residue of interest was mutated to the corresponding residue of the PupB-NTSD. Despite these extensive analyses, no PupA-NTSD mutations were identified that enabled PupA-NTSD interaction with the PupR-CCSSD, thereby facilitating signal transduction. Further mutational studies are necessary to elucidate the exact residues that account for signal specificity through the NTSD.

The X-ray crystal structure of the PupB-NTSD:PupR-CCSSD complex was determined to 1.6 Å resolution, and resolved several key concerns of the interaction of these two domains. Conventionally, the CCSSD was designated the CTD based on its position in the protein sequence; this domain was renamed to establish the domain in terms of its involvement in CSS. The structure of the complex revealed that the CCSSD comprises two subdomains: the C-terminal juxtamembrane (CJM) subdomain, residues 110-250, and the STN subdomain, residues 251-324. The CJM subdomain structure disclosed a previously uncharacterized, novel all-β, solenoid-like fold. The identification of the STN domain at the C-terminus of the regulator protein was surprising, as STN domains had previously not been located in the C-terminal region of any structurally characterized proteins. Site-directed mutagenesis of the residues at the interface of the complex resulted in several mutants that completely or nearly completed abrogated binding of the PupB-NTSD to the PupR-CCSSD. Of note, the interface identified in the crystal structure and confirmed by mutagenesis, is not the interface previously proposed to mediate the interaction between an NTSD and sigma regulator (81,139,159). Investigation of the

CCSSD by CD spectroscopy and SEC-SAXS revealed that it is metastable and highly flexible in solution, and is significantly stabilized upon interaction with the PupB-NTSD.

The X-ray crystal structure of the PupR N-terminal domain, classified here as an anti-sigma domain, was determined to 2.0 Å. The PupR-ASD exhibits a core composed of a three-helix bundle, similar to other ASD structures. Contrary to other ASDs of known structure, the PupR-ASD contains an extensive hydrophobic core, stabilized by three conserved tryptophans and one semi-conserved histidine. Since anti-sigma domains from transmembrane sigma regulators, such as FpvR, remain in complex with their cognate sigma factors after activation, the hydrophobic core may be essential to maintain stability of the sigma factor. In contrast to the other structurally characterized ASDs, the predicted fourth helix is disordered in the PupR-ASD. Since all other ASD structures include the sigma factor the ASD regulates, the fourth helix may be ordered when bound to PupI.

A surprising feature of the isolated PupR-ASD was the formation of a dimer; this represents the first description of a dimeric ASD for a transmembrane sigma regulator. The physiological relevance of this observation has yet to be demonstrated, although this presents new and interesting possibilities for signal transduction through the ASD (Figure 7.1). The stoichiometry of the PupR-ASD: PupI complex has not been validated, although preliminary SAXS analysis suggests that a heterodimer may be the primary species in solution. As discussed in chapter 6, a 1:1 PupR-ASD:PupI complex could accommodate either wrapping mode revealed by the σ^E:RseA and σ^W:RsiW structures; either of these binding modes would retain the hydrophobic core buried within the PupR-ASD dimer interface, as exposure of this interface upon dimer dissociation would allow burial of the same interface by PupI. Continued efforts to delineate the interaction interface of PupR-ASD:PupI, as well as optimize *in vivo* transcriptional activation, are required to solve the next step in the workings of the TBDT iron acquisition regulatory system.

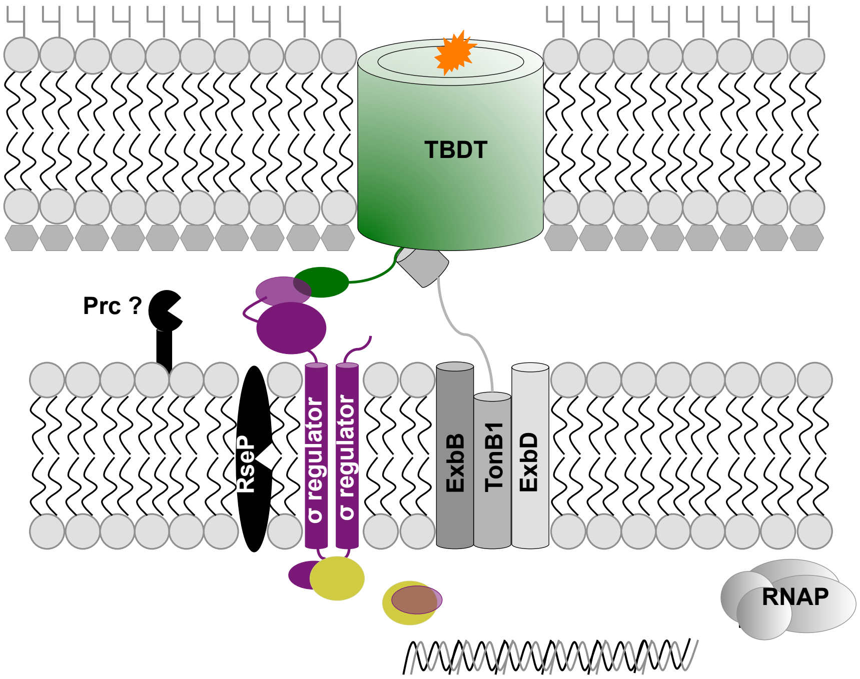


Figure 7.1. A revised model of the regulation of siderophore uptake. Binding of ferric siderophore to the TBD transducer results in TonB association with the TonB box for active transport aided by the PMF, and extension of the transducer NTSD farther into the periplasm. Contact of the NTSD with the CCSSD stabilizes the CCSSD, such that it becomes resistant to non-specific proteolytic degradation, and/or allows for site-specific proteolysis by a site-1 protease, such as Prc. Formation of a dimer by the ASD may also serve to stabilize the entire protein or set of proteins, and anchor the hyper-flexible CCSSD. Proteolysis by the site-1 protease provides the substrate for the site-2 protease, such as RseP, resulting in release of at least one ASD:sigma factor complex from the inner membrane. The complex is targeted to the transducer promoter, the sigma factor binds the DNA promoter, RNAP is recruited, and expression of the transducer occurs. Once Fur senses acceptable intracellular iron levels, Fur represses expression of the sigma factor and anti-sigma factor.

REFERENCES

1. CDC. (1999) Achievements in Public Health, 1900-1999: Control of Infectious Diseases. 48(29) Ed., Morbidity and Mortality Weekly Report.
2. McKenna, M. (2013) Imagining the Post-Antibiotics Future. http://www.Medium.com.
3. Levy, S. B. (2001) Antibacterial household products: cause for concern. *Emerg Infect Dis* **7**, 512-515.
4. CDC. Leading Causes of Death, 1900-1998.
5. Jiaquan, X., Murphy, S. L., Kochanek, K. D., and Arias, E. (2016) Mortality in the United States, 2015. *NCHS Data Brief* **267**.
6. Fleming, A. (1945) Penicillin. *Nobel Lecture.*
7. CDC. (2013) Antibiotic resistance threats in the United States.
8. Ventola, C. L. (2015) The antibiotic resistance crisis: part 1: causes and threats. *P T* **40**, 277-283.
9. Mehrad, B., Clark, N. M., Zhanel, G. G., and Lynch, J. P., 3rd. (2015) Antimicrobial resistance in hospital-acquired gram-negative bacterial infections. *Chest* **147**, 1413-1421.
10. Zucca, M., and Savoia, D. (2010) The post-antibiotic era: promising developments in the therapy of infectious diseases. *Int J Biomed Sci* **6**, 77-86.
11. Read, A. F., and Woods, R. J. (2014) Antibiotic resistance management. *Evol Med Public Health* **2014**, 147.
12. (2013) The antibiotic alarm. *Nature* **495**, 141.
13. Van Boeckel, T. P., Gandra, S., Ashok, A., Caudron, Q., Grenfell, B. T., Levin, S. A., and Laxminarayan, R. (2014) Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data. *Lancet Infect Dis* **14**, 742-750.
14. Lushniak, B. D. (2014) Antibiotic resistance: a public health crisis. *Public Health Rep* **129**, 314-316.
15. Michael, C. A., Dominey-Howes, D., and Labbate, M. (2014) The antibiotic resistance crisis: causes, consequences, and management. *Front Public Health* **2**.
16. Gould, I. M., and Bal, A. M. (2013) New antibiotic agents in the pipeline and how they can help overcome microbial resistance. *Virulence* **4**, 185-191.
17. Bartlett, J. G., Gilbert, D. N., and Spellberg, B. (2013) Seven ways to preserve the miracle of antibiotics. *Clin Infect Dis* **56**, 1445-1450.
18. Grimard, V., Lensink, M., Debailleul, F., Ruysschaert, J. M., and Govaerts, C. (2014) The role of lipid composition on bacterial membrane protein conformation and function. in *Bacterial membranes: structural and molecular biology* (Remaut, H., and Fronzes, R. eds.), Horizon Scientific Press.
19. Chatterjee, S. N., and Chaudhuri, K. (2012) Gram-negative bacteria: the cell membranes. in *Outer Membrane Vesicles of Bacteria*, Springer Briefs in Microbiology.
20. Silhavy, T. J., Kahne, D., and Walker, S. (2010) The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**, a000414.
21. Jean, N., Bougault, C., and Simorre, J. P. The structure of the bacterial cell wall. http://www.glycopedia.eu/e-chapters/the-structure-of-bacterial-cell/article/the-bacterial-cell-wall.
22. Miller, S. I. (2016) Antibiotic Resistance and Regulation of the Gram-Negative Bacterial Outer Membrane Barrier by Host Innate Immune Molecules. *mBio* **7**.

23. Delcour, A. H. (2009) Outer membrane permeability and antibiotic resistance. *Biochim Biophys Acta* **1794**, 808-816.
24. Beveridge, T. J. (1999) Structures of Gram-negative cell walls and their derived membrane vesicles. *J Bacteriol* **181**, 4725-4733.
25. Kohanski, M. A., Dwyer, D. J., and Collins, J. J. (2010) How antibiotics kill bacteria: from targets to networks. *Nature Rev Microbiol* **8**, 423-435.
26. Sender, R., Fuchs, S., and Milo, R. (2016) Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol* **14**, e1002533.
27. Human Microbiome Project. (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214.
28. Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., and Knight, R. (2009) Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694-1697.
29. National Institutes of Allergy and Infectious Diseases. (2016) Gram-negative bacteria. NIH, https://www.niaid.nih.gov/research/gram-negative-bacteria.
30. Davies, J. C. (2002) *Pseudomonas aeruginosa* in cystic fibrosis: pathogenesis and persistence. *Paediatr Respir Rev* **3**, 128-134.
31. Davies, J. C. (2002) New therapeutic approaches for cystic fibrosis lung disease. *J R Soc Med* **95 Suppl 41**, 58-67.
32. Pritt, B., O'Brien, L., and Winn, W. (2007) Mucoid *Pseudomonas* in cystic fibrosis. *Am J Clin Pathol* **128**, 32-34.
33. Mossialos, D., and Amoutzias, G. D. (2009) Role of siderophores in cystic fibrosis pathogenesis: foes or friends? *Int J Med Microbiol* **299**, 87-98.
34. Reid, D. W., Anderson, G. J., and Lamont, I. L. (2009) Role of lung iron in determining the bacterial and host struggle in cystic fibrosis. *Am J Physiol Lung Cell Mol Physiol* **297**, L795-802.
35. Eber, E., and Zach, M. S. (2010) *Pseudomonas aeruginosa* infection in cystic fibrosis: prevent, eradicate or both? *Thorax* **65**, 849-851.
36. Martin, L. W., Reid, D. W., Sharples, K. J., and Lamont, I. L. (2011) *Pseudomonas* siderophores in the sputum of patients with cystic fibrosis. *Biometals* **24**, 1059-1067.
37. Winstanley, C., O'Brien, S., and Brockhurst, M. A. (2016) *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. *Curr Trends Microbiol* **24**, 327-337.
38. Vincent, J. L., Rello, J., Marshall, J., Silva, E., Anzueto, A., Martin, C. D., Moreno, R., Lipman, J., Gomersall, C., Sakr, Y., Reinhart, K., and Investigators. (2009) International study of the prevalence and outcomes of infection in intensive care units. *JAMA* **302**, 2323-2329.
39. Peleg, A. Y., and Hooper, D. C. (2010) Hospital-acquired infections due to gram-negative bacteria. *N Engl J Med* **362**, 1804-1813.
40. Miethke, M., and Marahiel, M. A. (2007) Siderophore-based iron acquisition and pathogen control. *Microbiol Mol Biol Rev* **71**, 413-451.
41. Raymond, K. N., Dertz, E. A., and Kim, S. S. (2003) Enterobactin: an archetype for microbial iron transport. *PNAS* **100**, 3584-3588.
42. Andrews, S. C., Robinson, A. K., and Rodriguez-Quinones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**, 215-237.
43. Hantke, K. (2001) Iron and metal regulation in bacteria. *Curr Opin Microbiol* **4**, 172-177.

44. Minandri, F., Imperi, F., Frangipani, E., Bonchi, C., Visaggio, D., Facchini, M., Pasquali, P., Bragonzi, A., and Visca, P. (2016) Role of iron uptake systems in *Pseudomonas aeruginosa* virulence and airway infection. *Infect Immun* **84**, 2324-2335.
45. Schaible, U. E., and Kaufmann, S. H. (2004) Iron and microbial infection. *Nature Rev Microbiol* **2**, 946-953.
46. Sun, Y. Y., and Sun, L. (2015) *Pseudomonas fluorescens*: iron-responsive proteins and their involvement in host infection. *Vet Microbiol* **176**, 309-320.
47. Goldberg, M. B., Boyko, S. A., and Calderwood, S. B. (1990) Transcriptional regulation by iron of a *Vibrio cholerae* virulence gene and homology of the gene to the *Escherichia coli* fur system. *J Bacteriol* **172**, 6863-6870.
48. Goldberg, M. B., DiRita, V. J., and Calderwood, S. B. (1990) Identification of an iron-regulated virulence determinant in *Vibrio cholerae*, using TnphoA mutagenesis. *Infect Immun* **58**, 55-60.
49. Goldberg, M. B., Boyko, S. A., Butterton, J. R., Stoebner, J. A., Payne, S. M., and Calderwood, S. B. (1992) Characterization of a *Vibrio cholerae* virulence factor homologous to the family of TonB-dependent proteins. *Mol Microbiol* **6**, 2407-2418.
50. Xiong, Y. Q., Vasil, M. L., Johnson, Z., Ochsner, U. A., and Bayer, A. S. (2000) The oxygen- and iron-dependent sigma factor pvdS of *Pseudomonas aeruginosa* is an important virulence factor in experimental infective endocarditis. *J Infect Dis* **181**, 1020-1026.
51. Llamas, M. A., van der Sar, A., Chu, B. C., Sparrius, M., Vogel, H. J., and Bitter, W. (2009) A Novel extracytoplasmic function (ECF) sigma factor regulates virulence in *Pseudomonas aeruginosa*. *PLoS Pathog* **5**, e1000572.
52. Yeom, J., Imlay, J. A., and Park, W. (2010) Iron homeostasis affects antibiotic-mediated cell death in *Pseudomonas* species. *J Biol Chem* **285**, 22689-22695.
53. Frawley, E. R., Crouch, M. L., Bingham-Ramos, L. K., Robbins, H. F., Wang, W., Wright, G. D., and Fang, F. C. (2013) Iron and citrate export by a major facilitator superfamily pump regulates metabolism and stress resistance in *Salmonella typhimurium*. *PNAS* **110**, 12054-12059.
54. Troxell, B., and Hassan, H. M. (2013) Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Front Cell Infect Microbiol* **3**, 59.
55. Sen, K., Hellman, J., and Nikaido, H. (1988) Porin channels in intact cells of *Escherichia coli* are not affected by Donnan potentials across the outer membrane. *J Biol Chem* **263**, 1182-1187.
56. Touati, D. (2000) Iron and oxidative stress in bacteria. *Arch Biochem* **373**, 1-6.
57. Koster, M., van de Vossenberg, J., Leong, J., and Weisbeek, P. J. (1993) Identification and characterization of the pupB gene encoding an inducible ferric-pseudobactin receptor of *Pseudomonas putida* WCS358. *Mol Microbiol* **8**, 591-601.
58. Llamas, M. A., Sparrius, M., Kloet, R., Jimenez, C. R., Vandenbroucke-Grauls, C., and Bitter, W. (2006) The heterologous siderophores ferrioxamine B and ferrichrome activate signaling pathways in *Pseudomonas aeruginosa*. *J Bacteriol* **188**, 1882-1891.
59. Schalk, I. J., Abdallah, M. A., and Pattus, F. (2002) Recycling of pyoverdin on the FpvA receptor after ferric pyoverdin uptake and dissociation in *Pseudomonas aeruginosa*. *Biochem* **41**, 1663-1671.
60. Greenwald, J., Hoegy, F., Nader, M., Journet, L., Mislin, G. L., Graumann, P. L., and Schalk, I. J. (2007) Real time fluorescent resonance energy transfer visualization of ferric

pyoverdine uptake in *Pseudomonas aeruginosa*: A role for ferrous iron. *J Biol Chem* **282**, 2987-2995.

61. Yeterian, E., Martin, L. W., Lamont, I. L., and Schalk, I. J. (2010) An efflux pump is required for siderophore recycling by *Pseudomonas aeruginosa*. *Environ Microbiol Rep* **2**, 412-418.
62. Hvorup, R. N., Goetz, B. A., Niederer, M., Hollenstein, K., Perozo, E., and Locher, K. P. (2007) Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science* **317**, 1387-1390.
63. Schalk, I. J., and Guillon, L. (2013) Fate of ferrisiderophores after import across bacterial outer membranes: different iron release strategies are observed in the cytoplasm or periplasm depending on the siderophore pathways. *Amino Acids* **44**, 1267-1277.
64. Brickman, T. J., and McIntosh, M. A. (1992) Overexpression and purification of ferric enterobactin esterase from *Escherichia coli*: Demonstration of enzymatic hydrolysis of enterobactin and its iron complex. *J Biol Chem* **267**, 12350-12355.
65. Mies, K. A., Wirgau, J. I., and Crumbliss, A. L. (2006) Ternary complex formation facilitates a redox mechanism for iron release from a siderophore. *Biometals* **19**, 115-126.
66. Miethke, M., Hou, J., and Marahiel, M. A. (2011) The siderophore-interacting protein YqjH acts as a ferric reductase in different iron assimilation pathways of *Escherichia coli*. *Biochem* **50**, 10951-10964.
67. Matzanke, B. F., Anemuller, S., Schunemann, V., Trautwein, A. X., and Hantke, K. (2004) FhuF, part of a siderophore-reductase system. *Biochem* **43**, 1386-1392.
68. Hannauer, M., Barda, Y., Mislin, G. L., Shanzer, A., and Schalk, I. J. (2010) The ferrichrome uptake pathway in *Pseudomonas aeruginosa* involves an iron release mechanism with acylation of the siderophore and recycling of the modified desferrichrome. *J Bacteriol* **192**, 1212-1220.
69. Imperi, F., Massai, F., Facchini, M., Frangipani, E., Visaggio, D., Leoni, L., Bragonzi, A., and Visca, P. (2013) Repurposing the antimycotic drug flucytosine for suppression of *Pseudomonas aeruginosa* pathogenicity. *PNAS* **110**, 7458-7463.
70. Noinaj, N., Guillier, M., Barnard, T. J., and Buchanan, S. K. (2010) TonB-dependent transporters: regulation, structure, and function. *Annu Rev Microbiol* **64**, 43-60.
71. Koebnik, R. (2005) TonB-dependent trans-envelope signalling: the exception or the rule? *Curr Trends Microbiol* **13**, 343-347.
72. Kim, M., Fanucci, G. E., and Cafiso, D. S. (2007) Substrate-dependent transmembrane signaling in TonB-dependent transporters is not conserved. *PNAS* **104**, 11975-11980.
73. Wiener, M. C. (2005) TonB-dependent outer membrane transport: going for Baroque? *Curr Opin Struct Biol* **15**, 394-400.
74. Gudmundsdottir, A., Bell, P. E., Lundrigan, M. D., Bradbeer, C., and Kadner, R. J. (1989) Point mutations in a conserved region (TonB box) of *Escherichia coli* outer membrane protein BtuB affect vitamin B12 transport. *J Bacteriol* **171**, 6526-6533.
75. Cadieux, N., Bradbeer, C., and Kadner, R. J. (2000) Sequence changes in the ton box region of BtuB affect its transport activities and interaction with TonB protein. *J Bacteriol* **182**, 5954-5961.
76. Endriss, F., Braun, M., Killmann, H., and Braun, V. (2003) Mutant analysis of the *Escherichia coli* FhuA protein reveals sites of FhuA activity. *J Bacteriol* **185**, 4683-4692.
77. Locher, K. P., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J. P., and Moras, D. (1998) Transmembrane signaling across the ligand-gated FhuA receptor:

Crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell* **95**, 771-778.

78. Lukasik, S. M., Ho, K. W., and Cafiso, D. S. (2007) Molecular basis for substrate-dependent transmembrane signaling in an outer-membrane transporter. *J Mol Biol* **370**, 807-811.
79. Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Outer membrane active transport: structure of the BtuB:TonB complex. *Science* **312**, 1396-1399.
80. Wirth, C., Meyer-Klaucke, W., Pattus, F., and Cobessi, D. (2007) From the periplasmic signaling domain to the extracellular face of an outer membrane signal transducer of *Pseudomonas aeruginosa*: crystal structure of the ferric pyoverdine outer membrane receptor. *J Mol Biol* **368**, 398-406.
81. Ferguson, A. D., Amezcua, C. A., Halabi, N. M., Chelliah, Y., Rosen, M. K., Ranganathan, R., and Deisenhofer, J. (2007) Signal transduction pathway of TonB-dependent transporters. *PNAS* **104**, 513-518.
82. Brillet, K., Journet, L., Celia, H., Paulus, L., Stahl, A., Pattus, F., and Cobessi, D. (2007) A beta strand lock exchange for signal transduction in TonB-dependent transducers on the basis of a common structural motif. *Structure* **15**, 1383-1391.
83. Flores Jimenez, R. H., and Cafiso, D. S. (2012) The N-terminal domain of a TonB-dependent transporter undergoes a reversible stepwise denaturation. *Biochem* **51**, 3642-3650.
84. Sikora, A., Joseph, B., Matson, M., Staley, J. R., and Cafiso, D. S. (2016) Allosteric signaling is bidirectional in an outer-membrane transport protein. *Biophys J* **111**, 1908-1918.
85. Chang, C., Mooser, A., Pluckthun, A., and Wlodawer, A. (2001) Crystal structure of the dimeric C-terminal domain of TonB reveals a novel fold. *J Biol Chem* **276**, 27535-27540.
86. Koedding, J., Polzer, P., Killig, F., Howard, S. P., Gerber, K., Seige, P., Diederichs, K., and Welte, W. (2004) Crystallization and preliminary X-ray analysis of a C-terminal TonB fragment from *Escherichia coli*. *Acta Crystallogr D Biol Crystallogr* **60**, 1281-1283.
87. Kodding, J., Killig, F., Polzer, P., Howard, S. P., Diederichs, K., and Welte, W. (2005) Crystal structure of a 92-residue C-terminal fragment of TonB from *Escherichia coli* reveals significant conformational changes compared to structures of smaller TonB fragments. *J Biol Chem* **280**, 3022-3028.
88. Sean Peacock, R., Weljie, A. M., Peter Howard, S., Price, F. D., and Vogel, H. J. (2005) The solution structure of the C-terminal domain of TonB and interaction studies with TonB box peptides. *J Mol Biol* **345**, 1185-1197.
89. Shultis, D. D., Purdy, M. D., Banchs, C. N., and Wiener, M. C. (2006) Crystallization and preliminary X-ray crystallographic analysis of the *Escherichia coli* outer membrane cobalamin transporter BtuB in complex with the carboxy-terminal domain of TonB. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 638-641.
90. Pawelek, P. D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C. M., Moiseeva, N., Allaire, M., and Coulton, J. W. (2006) Structure of TonB in complex with FhuA, *E. coli* outer membrane receptor. *Science* **312**, 1399-1402.
91. Koedding, J., Howard, P., Kaufmann, L., Polzer, P., Lustig, A., and Welte, W. (2004) Dimerization of TonB is not essential for its binding to the outer membrane siderophore receptor FhuA of *Escherichia coli*. *J Biol Chem* **279**, 9978-9986.

92. Celia, H., Noinaj, N., Zakharov, S. D., Bordignon, E., Botos, I., Santamaria, M., Barnard, T. J., Cramer, W. A., Lloubes, R., and Buchanan, S. K. (2016) Structural insight into the role of the Ton complex in energy transduction. *Nature* **538**, 60-65.
93. Jordan, L. D., Zhou, Y., Smallwood, C. R., Lill, Y., Ritchie, K., Yip, W. T., Newton, S. M., and Klebba, P. E. (2013) Energy-dependent motion of TonB in the Gram-negative bacterial inner membrane. *PNAS* **110**, 11553-11558.
94. Newton, S. M., Trinh, V., Pi, H., and Klebba, P. E. (2010) Direct measurements of the outer membrane stage of ferric enterobactin transport: postuptake binding. *J Biol Chem* **285**, 17488-17497.
95. Klebba, P. E. (2016) ROSET model of TonB action in Gram-negative bacterial iron acquisition. *J Bacteriol* **198**, 1013-1021.
96. Gresock, M. G., Kastead, K. A., and Postle, K. (2015) From homodimer to heterodimer and back: Elucidating the TonB energy transduction cycle. *J Bacteriol* **197**, 3433-3445.
97. Ollis, A. A., Kumar, A., and Postle, K. (2012) The ExbD periplasmic domain contains distinct functional regions for two stages in TonB energization. *J Bacteriol* **194**, 3069-3077.
98. Ollis, A. A., and Postle, K. (2012) Identification of functionally important TonB-ExbD periplasmic domain interactions *in vivo*. *J Bacteriol* **194**, 3078-3087.
99. Van Hove, B., Staudenmaier, H., and Braun, V. (1990) Novel two-component transmembrane transcription control: regulation of iron dicitrate transport in *Escherichia coli* K-12. *J Bacteriol* **172**, 6749-6758.
100. Angerer, A., and Braun, V. (1998) Iron regulates transcription of the *Escherichia coli* ferric citrate transport genes directly and through the transcription initiation proteins. *Arch Microbiol* **169**, 483-490.
101. Braun, V., Mahren, S., and Ogierman, M. (2003) Regulation of the FecI-type ECF sigma factor by transmembrane signalling. *Curr Opin Microbiol* **6**, 173-180.
102. Ochs, M., Angerer, A., Enz, S., and Braun, V. (1996) Surface signaling in transcriptional regulation of the ferric citrate transport system of *Escherichia coli*: mutational analysis of the alternative sigma factor FecI supports its essential role in fec transport gene transcription. *Mol Gen Genet* **250**, 455-465.
103. Crosa, J. H., and Walsh, C. T. (2002) Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev* **66**, 223-249.
104. Davidson, A. L., and Chen, J. (2004) ATP-binding cassette transporters in bacteria. *Annu Rev Biochem* **73**, 241-268.
105. Enz, S., Brand, H., Orellana, C., Mahren, S., and Braun, V. (2003) Sites of interaction between the FecA and FecR signal transduction proteins of ferric citrate transport in *Escherichia coli* K-12. *J Bacteriol* **185**, 3745-3752.
106. Braun, V., and Mahren, S. (2005) Transmembrane transcriptional control (surface signalling) of the *Escherichia coli* Fec type. *FEMS Microbiol Rev* **29**, 673-684.
107. Braun, V., and Endriss, F. (2007) Energy-coupled outer membrane transport proteins and regulatory proteins. *Biometals* **20**, 219-231.
108. Sen, T. Z., Kloster, M., Jernigan, R. L., Kolinski, A., Bujnicki, J. M., and Kloczkowski, A. (2008) Predicting the complex structure and functional motions of the outer membrane transporter and signal transducer FecA. *Biophys J* **94**, 2482-2491.
109. Braun, V., Mahren, S., and Sauter, A. (2005) Gene regulation by transmembrane signaling. *Biometals* **18**, 507-517.

110. Stiefel, A., Mahren, S., Ochs, M., Schindler, P. T., Enz, S., and Braun, V. (2001) Control of the ferric citrate transport system of *Escherichia coli*: mutations in region 2.1 of the FecI extracytoplasmic-function sigma factor suppress mutations in the FecR transmembrane regulatory protein. *J Bacteriol* **183**, 162-170.
111. Breidenstein, E., Mahren, S., and Braun, V. (2006) Residues involved in FecR binding are localized on one side of the FecA signaling domain in *Escherichia coli*. *J Bacteriol* **188**, 6440-6442.
112. Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K., and Welte, W. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science* **282**, 2215-2220.
113. Ferguson, A. D., Braun, V., Fiedler, H. P., Coulton, J. W., Diederichs, K., and Welte, W. (2000) Crystal structure of the antibiotic albomycin in complex with the outer membrane transporter FhuA. *Protein Sci* **9**, 956-963.
114. Ferguson, A. D., Welte, W., Hofmann, E., Lindner, B., Holst, O., Coulton, J. W., and Diederichs, K. (2000) A conserved structural motif for lipopolysaccharide recognition by procaryotic and eukaryotic proteins. *Structure* **8**, 585-592.
115. Ferguson, A. D., Kodding, J., Walker, G., Bos, C., Coulton, J. W., Diederichs, K., Braun, V., and Welte, W. (2001) Active transport of an antibiotic rifamycin derivative by the outer-membrane protein FhuA. *Structure* **9**, 707-716.
116. Mathavan, I., Zirah, S., Mehmood, S., Choudhury, H. G., Goulard, C., Li, Y., Robinson, C. V., Rebuffat, S., and Beis, K. (2014) Structural basis for hijacking siderophore receptors by antimicrobial lasso peptides. *Nat Chem Biol* **10**, 340-342.
117. Buchanan, S. K., Smith, B. S., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D., and Deisenhofer, J. (1999) Crystal structure of the outer membrane active transporter FepA from *Escherichia coli*. *Nat Struct Biol* **6**, 56-63.
118. Ferguson, A. D., Chakraborty, R., Smith, B. S., Esser, L., van der Helm, D., and Deisenhofer, J. (2002) Structural basis of gating by the outer membrane transporter FecA. *Science* **295**, 1715-1719.
119. Yue, W. W., Grizot, S., and Buchanan, S. K. (2003) Structural evidence for iron-free citrate and ferric citrate binding to the TonB-dependent outer membrane transporter FecA. *J Mol Biol* **332**, 353-368.
120. Cobessi, D., Celia, H., Folschweiller, N., Schalk, I. J., Abdallah, M. A., and Pattus, F. (2005) The crystal structure of the pyoverdine outer membrane receptor FpvA from *Pseudomonas aeruginosa* at 3.6 angstroms resolution. *J Mol Biol* **347**, 121-134.
121. Greenwald, J., Nader, M., Celia, H., Gruffaz, C., Geoffroy, V., Meyer, J. M., Schalk, I. J., and Pattus, F. (2009) FpvA bound to non-cognate pyoverdines: molecular basis of siderophore recognition by an iron transporter. *Mol Microbiol* **72**, 1246-1259.
122. Cobessi, D., Celia, H., and Pattus, F. (2005) Crystal structure at high resolution of ferric-pyochelin and its membrane receptor FptA from *Pseudomonas aeruginosa*. *J Mol Biol* **352**, 893-904.
123. Freed, D. M., Horanyi, P. S., Wiener, M. C., and Cafiso, D. S. (2010) Conformational exchange in a membrane transport protein is altered in protein crystals. *Biophys J* **99**, 1604-1610.
124. Freed, D. M., Khan, A. K., Horanyi, P. S., and Cafiso, D. S. (2011) Molecular origin of electron paramagnetic resonance line shapes on beta-barrel membrane proteins: the local solvation environment modulates spin-label configuration. *Biochem* **50**, 8792-8803.

125. Cherezov, V., Yamashita, E., Liu, W., Zhalnina, M., Cramer, W. A., and Caffrey, M. (2006) *In meso* structure of the cobalamin transporter, BtuB, at 1.95 A resolution. *J Mol Biol* **364**, 716-734.

126. Buchanan, S. K., Lukacik, P., Grizot, S., Ghirlando, R., Ali, M. M., Barnard, T. J., Jakes, K. S., Kienker, P. K., and Esser, L. (2007) Structure of colicin I receptor bound to the R-domain of colicin Ia: implications for protein import. *EMBO J* **26**, 2594-2604.

127. Krieg, S., Huche, F., Diederichs, K., Izadi-Pruneyre, N., Lecroisey, A., Wandersman, C., Delepelaire, P., and Welte, W. (2009) Heme uptake across the outer membrane as revealed by crystal structures of the receptor-hemophore complex. *PNAS* **106**, 1045-1050.

128. Brillet, K., Meksem, A., Lauber, E., Reimmann, C., and Cobessi, D. (2009) Use of an in-house approach to study the three-dimensional structures of various outer membrane proteins: structure of the alcaligin outer membrane transporter FauA from *Bordetella pertussis*. *Acta Crystallogr D Biol Crystallogr* **65**, 326-331.

129. Brillet, K., Reimmann, C., Mislin, G. L., Noel, S., Rognan, D., Schalk, I. J., and Cobessi, D. (2011) Pyochelin enantiomers and their outer-membrane siderophore transporters in fluorescent pseudomonads: structural bases for unique enantiospecific recognition. *JACS* **133**, 16503-16509.

130. Noinaj, N., Easley, N. C., Oke, M., Mizuno, N., Gumbart, J., Boura, E., Steere, A. N., Zak, O., Aisen, P., Tajkhorshid, E., Evans, R. W., Gorringe, A. R., Mason, A. B., Steven, A. C., and Buchanan, S. K. (2012) Structural basis for iron piracy by pathogenic *Neisseria*. *Nature* **483**, 53-58.

131. Saleem, M., Prince, S. M., Rigby, S. E., Imran, M., Patel, H., Chan, H., Sanders, H., Maiden, M. C., Feavers, I. M., and Derrick, J. P. (2013) Use of a molecular decoy to segregate transport from antigenicity in the FrpB iron transporter from *Neisseria meningitidis*. *PloS One* **8**, e56746.

132. Calmettes, C., Ing, C., Buckwalter, C. M., El Bakkouri, M., Chieh-Lin Lai, C., Pogoutse, A., Gray-Owen, S. D., Pomes, R., and Moraes, T. F. (2015) The molecular mechanism of Zinc acquisition by the neisserial outer-membrane transporter ZnuD. *Nat Commun* **6**, 7996.

133. Lukacik, P., Barnard, T. J., Keller, P. W., Chaturvedi, K. S., Seddiki, N., Fairman, J. W., Noinaj, N., Kirby, T. L., Henderson, J. P., Steven, A. C., Hinnebusch, B. J., and Buchanan, S. K. (2012) Structural engineering of a phage lysin that targets Gram-negative pathogens. *PNAS* **109**, 9857-9862.

134. Grinter, R., Josts, I., Mosbahi, K., Roszak, A. W., Cogdell, R. J., Bonvin, A. M., Milner, J. J., Kelly, S. M., Byron, O., Smith, B. O., and Walker, D. (2016) Structure of the bacterial plant-ferredoxin receptor FusA. *Nat Commun* **7**, 13308.

135. Garcia-Herrero, A., and Vogel, H. J. (2005) Nuclear magnetic resonance solution structure of the periplasmic signalling domain of the TonB-dependent outer membrane transporter FecA from *Escherichia coli*. *Mol Microbiol* **58**, 1226-1237.

136. Malki, I., Simenel, C., Wojtowicz, H., de Amorim, G. C., Prochnicka-Chalufour, A., Hoos, S., Raynal, B., England, P., Chaffotte, A., Delepierre, M., Delepelaire, P., and Izadi-Pruneyre, N. (2014) Interaction of a partially disordered antisigma factor with its partner, the signaling domain of the TonB-dependent transporter HasR. *PloS One* **9**, e89502.

137. Mahren, S., Schnell, H., and Braun, V. (2005) Occurrence and regulation of the ferric citrate transport system in *Escherichia coli* B, *Klebsiella pneumoniae*, *Enterobacter aerogenes*, and *Photorhabdus luminescens*. *Arch Microbiol* **184**, 175-186.
138. Kim, I., Stiefel, A., Plantor, S., Angerer, A., and Braun, V. (1997) Transcription induction of the ferric citrate transport genes via the N-terminus of the FecA outer membrane protein, the Ton system and the electrochemical potential of the cytoplasmic membrane. *Mol Microbiol* **23**, 333-344.
139. Enz, S., Mahren, S., Stroeher, U. H., and Braun, V. (2000) Surface signaling in ferric citrate transport gene induction: interaction of the FecA, FecR, and FecI regulatory proteins. *J Bacteriol* **182**, 637-646.
140. Koster, M., van Klompenburg, W., Bitter, W., Leong, J., and Weisbeek, P. (1994) Role for the outer membrane ferric siderophore receptor PupB in signal transduction across the bacterial cell envelope. *EMBO J* **13**, 2805-2813.
141. Lamont, I. L., Beare, P. A., Ochsner, U., Vasil, A. I., and Vasil, M. L. (2002) Siderophore-mediated signaling regulates virulence factor production in *Pseudomonas aeruginosa. PNAS* **99**, 7072-7077.
142. Campbell, E. A., Greenwell, R., Anthony, J. R., Wang, S., Lim, L., Das, K., Sofia, H. J., Donohue, T. J., and Darst, S. A. (2007) A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Mol Cell* **27**, 793-805.
143. Greenwell, R., Nam, T. W., and Donohue, T. J. (2011) Features of *Rhodobacter sphaeroides* ChrR required for stimuli to promote the dissociation of sigma(E)/ChrR complexes. *J Mol Biol* **407**, 477-491.
144. Maillard, A. P., Girard, E., Ziani, W., Petit-Hartlein, I., Kahn, R., and Coves, J. (2014) The crystal structure of the anti-sigma factor CnrY in complex with the sigma factor CnrH shows a new structural class of anti-sigma factors targeting extracytoplasmic function sigma factors. *J Mol Biol* **426**, 2313-2327.
145. Jensen, J. L., Balbo, A., Neau, D. B., Chakravarthy, S., Zhao, H., Sinha, S. C., and Colbert, C. L. (2015) Mechanistic Implications of the Unique Structural Features and Dimerization of the Cytoplasmic Domain of the *Pseudomonas* Sigma Regulator, PupR. *Biochem* **54**, 5867-5877.
146. Kazmierczak, M. J., Wiedmann, M., and Boor, K. J. (2005) Alternative sigma factors and their roles in bacterial virulence. *Microbiol Mol Biol Rev* **69**, 527-543.
147. Sharma, U. K., and Chatterji, D. (2010) Transcriptional switching in *Escherichia coli* during stress and starvation by modulation of sigma activity. *FEMS Microbiol Rev* **34**, 646-657.
148. Potvin, E., Sanschagrin, F., and Levesque, R. C. (2008) Sigma factors in *Pseudomonas aeruginosa*. *FEMS Microbiol Rev* **32**, 38-55.
149. Paget, M. S., and Helmann, J. D. (2003) The sigma70 family of sigma factors. *Genome Biol* **4**, 203.
150. Staron, A., Sofia, H. J., Dietrich, S., Ulrich, L. E., Liesegang, H., and Mascher, T. (2009) The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol Microbiol* **74**, 557-581.
151. Enz, S., Braun, V., and Crosa, J. H. (1995) Transcription of the region encoding the ferric dicitrate-transport system in *Escherichia coli*: similarity between promoters for fecA and for extracytoplasmic function sigma factors. *Gene* **163**, 13-18.

152. Redly, G. A., and Poole, K. (2005) FpvIR control of fpvA ferric pyoverdine receptor gene expression in *Pseudomonas aeruginosa*: demonstration of an interaction between FpvI and FpvR and identification of mutations in each compromising this interaction. *J Bacteriol* **187**, 5648-5657.
153. Angerer, A., Enz, S., Ochs, M., and Braun, V. (1995) Transcriptional regulation of ferric citrate transport in *Escherichia coli* K-12. Fecl belongs to a new subfamily of sigma 70-type factors that respond to extracytoplasmic stimuli. *Mol Microbiol* **18**, 163-174.
154. Crosa, J. H. (1997) Signal transduction and transcriptional and posttranscriptional control of iron-regulated genes in bacteria. *Microbiol Mol Biol Rev* **61**, 319-336.
155. Hantke, K. (1981) Regulation of ferric iron transport in *Escherichia coli* K12: isolation of a constitutive mutant. *Mol Gen Genet* **182**, 288-292.
156. Bagg, A., and Neilands, J. B. (1987) Molecular mechanism of regulation of siderophore-mediated iron assimilation. *Microbiol Rev* **51**, 509-518.
157. Bagg, A., and Neilands, J. B. (1987) Ferric uptake regulation protein acts as a repressor, employing iron (II) as a cofactor to bind the operator of an iron transport operon in *Escherichia coli*. *Biochem* **26**, 5471-5477.
158. Stojiljkovic, I., Baumler, A. J., and Hantke, K. (1994) Fur regulon in Gram-negative bacteria. Identification and characterization of new iron-regulated *Escherichia coli* genes by a fur titration assay. *J Mol Biol* **236**, 531-545.
159. Wriedt, K., Angerer, A., and Braun, V. (1995) Transcriptional regulation from the cell surface: conformational changes in the transmembrane protein FecR lead to altered transcription of the ferric citrate transport genes in *Escherichia coli*. *J Bacteriol* **177**, 3320-3322.
160. Draper, R. C., Martin, L. W., Beare, P. A., and Lamont, I. L. (2011) Differential proteolysis of sigma regulators controls cell-surface signalling in *Pseudomonas aeruginosa*. *Mol Microbiol* **82**, 1444-1453.
161. Bastiaansen, K. C., Ibanez, A., Ramos, J. L., Bitter, W., and Llamas, M. A. (2014) The Prc and RseP proteases control bacterial cell-surface signalling activity. *Environ Microbiol* **16**, 2433-2443.
162. Koide, K., Ito, K., and Akiyama, Y. (2008) Substrate recognition and binding by RseP, an *Escherichia coli* intramembrane protease. *J Biol Chem* **283**, 9562-9570.
163. Ponting, C. P., Phillips, C., Davies, K. E., and Blake, D. J. (1997) PDZ domains: targeting signalling molecules to sub-membranous sites. *Bioessays* **19**, 469-479.
164. Kim, D. Y., Jin, K. S., Kwon, E., Ree, M., and Kim, K. K. (2007) Crystal structure of RseB and a model of its binding mode to RseA. *PNAS* **104**, 8779-8784.
165. Kim, D. Y., Kwon, E., Choi, J., Hwang, H. Y., and Kim, K. K. (2010) Structural basis for the negative regulation of bacterial stress response by RseB. *Protein Sci* **19**, 1258-1263.
166. Kim, D. Y. (2015) Two stress sensor proteins for the expression of sigmaE regulon: DegS and RseB. *J Microbiol* **53**, 306-310.
167. Ades, S. E., Connolly, L. E., Alba, B. M., and Gross, C. A. (1999) The *Escherichia coli* sigma(E)-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor. *Genes Dev* **13**, 2449-2461.
168. Schurr, M. J., and Deretic, V. (1997) Microbial pathogenesis in cystic fibrosis: co-ordinate regulation of heat-shock response and conversion to mucoidy in *Pseudomonas aeruginosa*. *Mol Microbiol* **24**, 411-420.

169. Rowen, D. W., and Deretic, V. (2000) Membrane-to-cytosol redistribution of ECF sigma factor AlgU and conversion to mucoidy in *Pseudomonas aeruginosa* isolates from cystic fibrosis patients. *Mol Microbiol* **36**, 314-327.

170. Reiling, S. A., Jansen, J. A., Henley, B. J., Singh, S., Chattin, C., Chandler, M., and Rowen, D. W. (2005) Prc protease promotes mucoidy in mucA mutants of *Pseudomonas aeruginosa*. *Microbiol* **151**, 2251-2261.

171. Zheng, T., and Nolan, E. M. (2014) Enterobactin-mediated delivery of beta-lactam antibiotics enhances antibacterial activity against pathogenic *Escherichia coli*. *JACS* **136**, 9677-9691.

172. McPherson, C. J., Aschenbrenner, L. M., Lacey, B. M., Fahnoe, K. C., Lemmon, M. M., Finegan, S. M., Tadakamalla, B., O'Donnell, J. P., Mueller, J. P., and Tomaras, A. P. (2012) Clinically relevant Gram-negative resistance mechanisms have no effect on the efficacy of MC-1, a novel siderophore-conjugated monocarbam. *Antimicrob Agents Chemother* **56**, 6334-6342.

173. Page, M. G., Dantier, C., and Desarbre, E. (2010) In vitro properties of BAL30072, a novel siderophore sulfactam with activity against multiresistant gram-negative bacilli. *Antimicrob Agents Chemother* **54**, 2291-2302.

174. Tillotson, G. S. (2016) Trojan horse antibiotics-a novel way to circumvent Gram-negative bacterial resistance? *Infect Dis* **9**, 45-52.

175. Braun, V., and Braun, M. (2002) Active transport of iron and siderophore antibiotics. *Curr Opin Microbiol* **5**, 194-201.

176. Mislin, G. L., and Schalk, I. J. (2014) Siderophore-dependent iron uptake systems as gates for antibiotic Trojan horse strategies against *Pseudomonas aeruginosa*. *Metallomics* **6**, 408-420.

177. Bickel, H., Mertens, P., Prelog, V., Seibl, J., and Walser, A. (1965) Constitution of ferrimycin A1. *Antimicrob Agents Chemother* **5**, 951-957.

178. Tsukiura, H., Okanishi, M., Ohmori, T., Koshiyama, H., Miyaki, T., Kitazima, H., and Kawaguchi, H. (1964) Danomycin, a New Antibiotic. *J Antibiot* **17**, 39-47.

179. de Lorenzo, V. (1984) Isolation and characterization of microcin E492 from *Klebsiella pneumoniae*. *Arch Microbiol* **139**, 72-75.

180. de Lorenzo, V., Martinez, J. L., and Asensio, C. (1984) Microcin-mediated interactions between *Klebsiella pneumoniae* and *Escherichia coli* strains. *J Gen Microbiol* **130**, 391-400.

181. Nolan, E. M., and Walsh, C. T. (2008) Investigations of the MceIJ-catalyzed posttranslational modification of the microcin E492 C-terminus: linkage of ribosomal and nonribosomal peptides to form "trojan horse" antibiotics. *Biochem* **47**, 9289-9299.

182. Ji, C., Miller, P. A., and Miller, M. J. (2012) Iron transport-mediated drug delivery: Practical syntheses and *in vitro* antibacterial studies of tris-catecholate siderophore-aminopenicillin conjugates reveals selectively potent anti-pseudomonal activity. *JACS* **134**, 9898-9901.

183. Fardeau, S., Dassonville-Klimpt, A., Audic, N., Sasaki, A., Pillon, M., Baudrin, E., Mullie, C., and Sonnet, P. (2014) Synthesis and antibacterial activity of catecholate-ciprofloxacin conjugates. *Bioorg Med Chem* **22**, 4049-4060.

184. Ji, C., and Miller, M. J. (2015) Siderophore-fluoroquinolone conjugates containing potential reduction-triggered linkers for drug release: synthesis and antibacterial activity. *Biometals* **28**, 541-551.

185. Braun, V. (1997) Surface signaling: novel transcription initiation mechanism starting from the cell surface. *Arch Microbiol* **167**, 325-331.
186. Wojtowicz, H., Prochnicka-Chalufour, A., Cardoso de Amorim, G., Roudenko, O., Simenel, C., Malki, I., Pehau-Arnaudet, G., Gubellini, F., Koutsioubas, A., Perez, J., Delepelaire, P., Delepierre, M., Fronzes, R., and Izadi-Pruneyre, N. (2016) Structural basis of the signaling through a bacterial membrane receptor HasR deciphered by an integrative approach. *Biochem J* **473**, 2239-2248.
187. Shirley, M., and Lamont, I. L. (2009) Role of TonB1 in pyoverdine-mediated signaling in *Pseudomonas aeruginosa*. *J Bacteriol* **191**, 5634-5640.
188. Sambrook, J., Fritsche, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, New York.
189. Vergis, J. M., Purdy, M. D., and Wiener, M. C. (2010) A high-throughput differential filtration assay to screen and select detergents for membrane proteins. *Anal Biochem* **407**, 1-11.
190. Vergis, J. M., and Wiener, M. C. (2011) The variable detergent sensitivity of proteases that are utilized for recombinant protein affinity tag removal. *Protein Expr Purif* **78**, 139-142.
191. Ujwal, R., and Bowie, J. U. (2011) Crystallizing membrane proteins using lipidic bicelles. *Methods* **55**, 337-341.
192. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40**, 191-285.
193. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering analysis. *J. Appl. Crystallogr.* **36**, 1277-1282.
194. Kikhney, A. G., and Svergun, D. I. (2015) A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* **589**, 2570-2577.
195. Grant, T. D., Luft, J. R., Carter, L. G., Matsui, T., Weiss, T. M., Martel, A., and Snell, E. H. (2015) The accurate assessment of small-angle X-ray scattering data. *Acta Crystallogr D Biol Crystallogr* **71**, 45-56.
196. Franke, D., and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342-346.
197. Boldon, L., Laliberte, F., and Liu, L. (2015) Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application. *Nano Rev* **6**, 25661.
198. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D., Konarev, P. V., and Svergun, D. I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* **45**, 342-350.
199. Semenyuk, A. V., and Svergun, D. I. (1991) Gnom- a program package for small-angle scattering data-processing. *J. Appl. Crystallogr.* **24**, 537-540.
200. Kozin, M., and Svergun, D. (2001) Automated matching of high- and low-resolution structural models. *J Appl Cryst.* **34**, 33-41.

201. Wriggers, W., Milligan, R. A., and McCammon, J. A. (1999) Situs: a package for docking crystal structures into low-resolution maps from electron microscopy. *J. Struct. Biol.* **125**, 185-195.
202. Wriggers, W., and Birmanns, S. (2001) Using situs for flexible and rigid-body fitting of multiresolution single-molecule data. *J Struct Biol* **133**, 193-202.
203. Perez, J., and Koutsioubas, A. (2015) Memprot: a program to model the detergent corona around a membrane protein based on SEC-SAXS data. *Acta Crystallogr D Biol Crystallogr* **71**, 86-93.
204. Berthaud, A., Manzi, J., Perez, J., and Mangenot, S. (2012) Modeling detergent organization around Aquaporin-0 using small-angle X-ray scattering. *JACS* **134**, 10080-10088.
205. Koutsioubas, A., Berthaud, A., Mangenot, S., and Perez, J. (2013) Ab initio and all-atom modeling of detergent organization around Aquaporin-0 based on SAXS data. *J Phys Chem B* **117**, 13588-13594.
206. Lipfert, J., Columbus, L., Chu, V. B., Lesley, S. A., and Doniach, S. (2007) Size and shape of detergent micelles determined by small-angle X-ray scattering. *J Phys Chem B* **111**, 12427-12438.
207. Ujwal, R., Cascio, D., Colletier, J. P., Faham, S., Zhang, J., Toro, L., Ping, P., and Abramson, J. (2008) The crystal structure of mouse VDAC1 at 2.3 A resolution reveals mechanistic insights into metabolite gating. *PNAS* **105**, 17742-17747.
208. Faham, S., Ujwal, R., Abramson, J., and Bowie, J. U. (2009) Practical aspects of membrane proteins crystallization in bicelles. in *Curr Topics Memb* 109-125.
209. Lipfert, J., and Doniach, S. (2007) Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu Rev Biophys Biomol Struct* **36**, 307-327.
210. Doniach, S., and Lipfert, J. (2012) Small and wide angle x-ray scattering from biological macromolecules and their complexes in solution. in *Comprehensive Biophysics*, Elsevier. 376-397.
211. Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) CRYSOL- a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768-773.
212. Voss, N. R., and Gerstein, M. (2010) 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Res* **38**, W555-562.
213. Martinez-Bueno, M. A., Tobes, R., Rey, M., and Ramos, J. L. (2002) Detection of multiple extracytoplasmic function (ECF) sigma factors in the genome of *Pseudomonas putida* KT2440 and their counterparts in *Pseudomonas aeruginosa* PA01. *Environ Microbiol* **4**, 842-855.
214. Thakur, P. B., Vaughn-Diaz, V. L., Greenwald, J. W., and Gross, D. C. (2013) Characterization of five ECF sigma factors in the genome of *Pseudomonas syringae* pv. syringae B728a. *PloS One* **8**, e58846.
215. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, H. A., and Struhl, K. (1991) *Curr Prot in Mol Biol*, Greene Publishing Associates, New York.
216. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe - a multidimensional spectral processing system based on Unix pipes. *J Biomol Nmr* **6**, 277-293.
217. Johnson, B. A., and Blevins, R. A. (1994) NMR View: A computer program for the visualization and analysis of NMR data. *J Biomol Nmr* **4**, 603-614.

218. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol Biol* **278**, 313-352.
219. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59**, 687-696.
220. Bax, A., and Ikura, M. (1991) An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. *J Biomol Nmr* **1**, 99-104.
221. Constantine, K. L., Goldfarb, V., Wittekind, M., Friedrichs, M. S., Anthony, J., Ng, S. C., and Mueller, L. (1993) Aliphatic 1H and 13C resonance assignments for the 26-10 antibody VL domain derived from heteronuclear multidimensional NMR spectroscopy. *J Biomol Nmr* **3**, 41-54.
222. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., Rance, M., and Skelton, N. J. (2007) *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed., Academic Press.
223. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol Nmr* **13**, 289-302.
224. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol Nmr* **44**, 213-223.
225. Cheung, M. S., Maguire, M. L., Stevens, T. J., and Broadhurst, R. W. (2010) DANGLE: A Bayesian inferential method for predicting protein backbone dihedral angles and secondary structure. *J Magn Reson* **202**, 223-233.
226. Linge, J. P., Habeck, M., Rieping, W., and Nilges, M. (2003) ARIA: Automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315-316.
227. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921.
228. Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T. E., and Nilges, M. (2007) ARIA2: Automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **23**, 381-382.
229. Bardiaux, B., Malliavin, T., and Nilges, M. (2012) ARIA for solution and solid-state NMR. *Methods Mol Biol* **831**, 453-483.
230. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) Refinement of protein structures in explicit solvent. *Proteins* **50**, 496-506.
231. Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data. *PNAS* **105**, 4685-4690.
232. Shen, Y., Vernon, R., Baker, D., and Bax, A. (2009) *De novo* protein structure generation from incomplete chemical shift assignments. *J Biomol Nmr* **43**, 63-78.
233. Shen, Y., Bryan, P. N., He, Y., Orban, J., Baker, D., and Bax, A. (2010) *De novo* structure generation using chemical shifts for proteins with high-sequence identity but different folds. *Protein Sci* **19**, 349-356.

234. Lange, O. F., Rossi, P., Sgourakis, N. G., Song, Y., Lee, H. W., Aramini, J. M., Ertekin, A., Xiao, R., Acton, T. B., Montelione, G. T., and Baker, D. (2012) Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *PNAS* **109**, 10873-10878.
235. Kaufmann, K. W., Lemmon, G. H., Deluca, S. L., Sheehan, J. H., and Meiler, J. (2010) Practically useful: What the Rosetta protein modeling suite can do for you. *Biochem* **49**, 2987-2998.
236. Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (2001) Analysis of protein circular dichroism spectra based on the tertiary structure classification. *Anal Biochem* **299**, 271-274.
237. Ginzinger, S. W., Gerick, F., Coles, M., and Heun, V. (2007) CheckShift: Automatic correction of inconsistent chemical shift referencing. *J Biomol Nmr* **39**, 223-227.
238. Ginzinger, S. W., Skocibusic, M., and Heun, V. (2009) CheckShift improved: Fast chemical shift reference correction with high accuracy. *J Biomol Nmr* **44**, 207-211.
239. Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**, 2876-2890.
240. Bastiaansen, K. C., Otero-Asman, J. R., Luirink, J., Bitter, W., and Llamas, M. A. (2015) Processing of cell-surface signalling anti-sigma factors prior to signal recognition is a conserved autoproteolytic mechanism that produces two functional domains. *Environ Microbiol* **17**, 3263-3277.
241. Bastiaansen, K. C., van Ulsen, P., Wijtmans, M., Bitter, W., and Llamas, M. A. (2015) Self-cleavage of the *Pseudomonas aeruginosa* cell-surface signaling anti-sigma factor FoxR occurs through an N-O acyl rearrangement. *J Biol Chem* **290**, 12237-12246.
242. Edgar, R. J., Xu, X., Shirley, M., Konings, A. F., Martin, L. W., Ackerley, D. F., and Lamont, I. L. (2014) Interactions between an anti-sigma protein and two sigma factors that regulate the pyoverdine signaling pathway in *Pseudomonas aeruginosa*. *BMC Microbiol* **14**, 287.
243. Sheffield, P., Garrard, S., and Derewenda, Z. (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expr Purif* **15**, 34-39.
244. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**, 105-124.
245. Doublie, S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol* **276**, 523-530.
246. Scheuermann, T. H., and Brautigam, C. A. (2015) High-precision, automated integration of multiple isothermal titration calorimetric thermograms: new features of NITPIC. *Methods* **76**, 87-98.
247. Brautigam, C. A. (2015) Fitting two- and three-site binding models to isothermal titration calorimetric data. *Methods* **76**, 124-136.
248. Brautigam, C. A., Zhao, H., Vargas, C., Keller, S., and Schuck, P. (2016) Integration and global analysis of isothermal titration calorimetry data for studying macromolecular interactions. *Nat Protoc* **11**, 882-894.
249. Zhao, H., and Schuck, P. (2015) Combining biophysical methods for the analysis of protein complex stoichiometry and affinity in SEDPHAT. *Acta Crystallogr D Biol Crystallogr* **71**, 3-14.

250. Zhao, H., Piszczek, G., and Schuck, P. (2015) SEDPHAT--a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **76**, 137-148.
251. Houtman, J. C., Brown, P. H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L. E., and Schuck, P. (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci* **16**, 30-42.
252. Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* **67**, 293-302.
253. Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
254. Kabsch, W. (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132.
255. Evans, P. R. (1997). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **33**, 22-24.
256. Evans, P. (2006). *Acta Cryst.* **D62**, 72-82.
257. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326.
258. Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H., and Hung, L. W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582-601.
259. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., and Adams, P. D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**, 61-69.
260. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221.
261. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.
262. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.
263. Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62**, 439-450.
264. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797.
265. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21.
266. Holm, L., and Rosenstrom, P. (2010) Dali server: Conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-549.

267. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., and Sali, A. (2016) FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res* **44**, W424-429.
268. Tria, G., Mertens, H. D., Kachala, M., and Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207-217.
269. Bieri, M., Kwan, A. H., Mobli, M., King, G. F., Mackay, J. P., and Gooley, P. R. (2011) Macromolecular NMR spectroscopy for the non-spectroscopist: beyond macromolecular solution structure determination. *FEBS J* **278**, 704-715.
270. Ziarek, J. J., Peterson, F. C., Lytle, B. L., and Volkman, B. F. (2011) Binding site identification and structure determination of protein-ligand complexes by NMR a semiautomated approach. *Methods Enzymol* **493**, 241-275.
271. Bergeron, J. R., Worrall, L. J., Sgourakis, N. G., DiMaio, F., Pfuetzner, R. A., Felise, H. B., Vuckovic, M., Yu, A. C., Miller, S. I., Baker, D., and Strynadka, N. C. (2013) A refined model of the prototypical *Salmonella* SPI-1 T3SS basal body reveals the molecular basis for its assembly. *PLoS Pathog* **9**, e1003307.
272. Berry, J. L., Phelan, M. M., Collins, R. F., Adomavicius, T., Tonjum, T., Frye, S. A., Bird, L., Owens, R., Ford, R. C., Lian, L. Y., and Derrick, J. P. (2012) Structure and assembly of a trans-periplasmic channel for type IV pili in *Neisseria meningitidis*. *PLoS Pathog* **8**, e1002923.
273. Spreter, T., Yip, C. K., Sanowar, S., Andre, I., Kimbrough, T. G., Vuckovic, M., Pfuetzner, R. A., Deng, W., Yu, A. C., Finlay, B. B., Baker, D., Miller, S. I., and Strynadka, N. C. (2009) A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nature Struct Mol Biol* **16**, 468-476.
274. Korotkov, K. V., Pardon, E., Steyaert, J., and Hol, W. G. (2009) Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. *Structure* **17**, 255-265.
275. Korotkov, K. V., Delarosa, J. R., and Hol, W. G. (2013) A dodecameric ring-like structure of the N0 domain of the type II secretin from enterotoxigenic *Escherichia coli*. *J Struct Biol* **183**, 354-362.
276. Nakano, N., Kubori, T., Kinoshita, M., Imada, K., and Nagai, H. (2010) Crystal structure of *Legionella* DotD: insights into the relationship between type IVB and type II/III secretion systems. *PLoS Pathog* **6**, e1001129.
277. Souza, D. P., Andrade, M. O., Alvarez-Martinez, C. E., Arantes, G. M., Farah, C. S., and Salinas, R. K. (2011) A component of the Xanthomonadaceae type IV secretion system combines a VirB7 motif with a N0 domain found in outer membrane transport proteins. *PLoS Pathog* **7**, e1002031.
278. Kanamaru, S., Leiman, P. G., Kostyuchenko, V. A., Chipman, P. R., Mesyanzhinov, V. V., Arisaka, F., and Rossmann, M. G. (2002) Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**, 553-557.
279. Koshiyama, T., Yokoi, N., Ueno, T., Kanamaru, S., Nagano, S., Shiro, Y., Arisaka, F., and Watanabe, Y. (2008) Molecular design of heteroprotein assemblies providing a bionanocup as a chemical reactor. *Small* **4**, 50-54.
280. Spinola-Amilibia, M., Davo-Siguero, I., Ruiz, F. M., Santillana, E., Medrano, F. J., and Romero, A. (2016) The structure of VgrG1 from *Pseudomonas aeruginosa*, the needle tip of the bacterial type VI secretion system. *Acta Crystallogr D Struct Biol* **72**, 22-33.

281. Zemla, A., Geisbrecht, B., Smith, J., Lam, M., Kirkpatrick, B., Wagner, M., Slezak, T., and Zhou, C. E. (2007) STRALCP--structure alignment-based clustering of proteins. *Nucleic Acids Res* **35**, e150.

282. Harle, C., Kim, I., Angerer, A., and Braun, V. (1995) Signal transfer through three compartments: transcription initiation of the *Escherichia coli* ferric citrate transport system from the cell surface. *EMBO J* **14**, 1430-1438.

283. Mettrick, K. A., and Lamont, I. L. (2009) Different roles for anti-sigma factors in siderophore signalling pathways of *Pseudomonas aeruginosa*. *Mol Microbiol* **74**, 1257-1271.

284. Paget, M. S. (2015) Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. *Biomol* **5**, 1245-1265.

285. Alba, B. M., Leeds, J. A., Onufryk, C., Lu, C. Z., and Gross, C. A. (2002) DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma(E)-dependent extracytoplasmic stress response. *Genes Dev* **16**, 2156-2168.

286. Kanehara, K., Ito, K., and Akiyama, Y. (2002) YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA. *Genes Dev* **16**, 2147-2155.

287. Kang, J. G., Paget, M. S., Seok, Y. J., Hahn, M. Y., Bae, J. B., Hahn, J. S., Kleanthous, C., Buttner, M. J., and Roe, J. H. (1999) RsrA, an anti-sigma factor regulated by redox change. *EMBO J* **18**, 4292-4298.

288. Trepreau, J., Girard, E., Maillard, A. P., de Rosny, E., Petit-Haertlein, I., Kahn, R., and Coves, J. (2011) Structural basis for metal sensing by CnrX. *J Mol Biol* **408**, 766-779.

289. Thakur, K. G., Praveena, T., and Gopal, B. (2010) Structural and biochemical bases for the redox sensitivity of *Mycobacterium tuberculosis* RslA. *J Mol Biol* **397**, 1199-1208.

290. Rajasekar, K. V., Zdanowski, K., Yan, J., Hopper, J. T., Francis, M. L., Seepersad, C., Sharp, C., Pecqueur, L., Werner, J. M., Robinson, C. V., Mohammed, S., Potts, J. R., and Kleanthous, C. (2016) The anti-sigma factor RsrA responds to oxidative stress by reburying its hydrophobic core. *Nat Commun* **7**, 12194.

291. Campbell, E. A., Tupy, J. L., Gruber, T. M., Wang, S., Sharp, M. M., Gross, C. A., and Darst, S. A. (2003) Crystal structure of *Escherichia coli* sigmaE with the cytoplasmic domain of its anti-sigma RseA. *Mol Cell* **11**, 1067-1078.

292. Newman, J. D., Anthony, J. R., and Donohue, T. J. (2001) The importance of zinc-binding to the function of *Rhodobacter sphaeroides* ChrR as an anti-sigma factor. *J Mol Biol* **313**, 485-499.

293. Shukla, J., Gupta, R., Thakur, K. G., Gokhale, R., and Gopal, B. (2014) Structural basis for the redox sensitivity of the *Mycobacterium tuberculosis* sigK-RskA sigma-anti-sigma complex. *Acta Crystallogr D Biol Crystallogr* **70**, 1026-1036.

294. Dauter, Z., and Dauter, M. (1999) Anomalous signal of solvent bromides used for phasing of lysozyme. *J Mol Biol* **289**, 93-101.

295. Dauter, Z., Dauter, M., and Rajashankar, K. R. (2000) Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. *Acta Crystallogr D Biol Crystallogr* **56**, 232-237.

296. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495-503.

297. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860-864.

298. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612.
299. Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I., and Craievich, A. F. (2010) Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* **43**.
300. Schneidman-Duhovny, D., Hammel, M., and Sali, A. (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res* **38**, W540-544.
301. Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., and Sali, A. (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys J* **105**, 962-974.
302. Bernado, P., Mylonas, E., Petoukhov, M. V., Blackledge, M., and Svergun, D. I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *JACS* **129**, 5656-5664.
303. Zhao, H., Berger, A. J., Brown, P. H., Kumar, J., Balbo, A., May, C. A., Casillas, E., Jr., Laue, T. M., Patterson, G. H., Mayer, M. L., and Schuck, P. (2013) Analysis of high-affinity assembly for AMPA receptor amino-terminal domains. *J Gen Physiol* **141**, 747-749.
304. Brown, P. H., Balbo, A., and Schuck, P. (2008) Characterizing protein-protein interactions by sedimentation velocity analytical ultracentrifugation. *Curr Protoc Immunol* **Chapter 18**, Unit 18:15.
305. Zhao, H., Brown, P. H., and Schuck, P. (2011) On the distribution of protein refractive index increments. *Biophys J* **100**, 2309-2317.
306. Schuck, P. (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* **320**, 104-124.
307. Zhao, H., and Schuck, P. (2012) Global multi-method analysis of affinities and cooperativity in complex systems of macromolecular interactions. *Anal Chem* **84**, 9513-9519.
308. Johnson, M. L. (1992) Why, when, and how biochemists should use least squares. *Anal Biochem.* **206**, 215-225
309. Perez-Iratxeta, C., and Andrade-Navarro, M. A. (2008) K2D2: Estimation of protein secondary structure from circular dichroism spectra. *BMC Struct Biol* **8**, 25.
310. Sonnichsen, F. D., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochem* **31**, 8790-8798.
311. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1998) Trifluoroethanol effects on helix propensity and electrostatic interactions in the helical peptide from ribonuclease T1. *Protein Sci* **7**, 383-388.
312. Glover, K., Mei, Y., and Sinha, S. C. (2016) Identifying intrinsically disordered protein regions likely to undergo binding-induced helical transitions. *Biochim Biophys Acta* **1864**, 1455-1463.
313. Lambert, L. J., Wei, Y., Schirf, V., Demeler, B., and Werner, M. H. (2004) T4 AsiA blocks DNA recognition by remodeling sigma70 region 4. *EMBO J* **23**, 2952-2962.
314. Campbell, E. A., Muzzin, O., Chlenov, M., Sun, J. L., Olson, C. A., Weinman, O., Trester-Zedlitz, M. L., and Darst, S. A. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* **9**, 527-539.

315. Mahren, S., Enz, S., and Braun, V. (2002) Functional interaction of region 4 of the extracytoplasmic function sigma factor FecI with the cytoplasmic portion of the FecR transmembrane protein of the *Escherichia coli* ferric citrate transport system. *J Bacteriol* **184**, 3704-3711.
316. Lonetto, M. A., Brown, K. L., Rudd, K. E., and Buttner, M. J. (1994) Analysis of the *Streptomyces coelicolor* sigE gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. *PNAS* **91**, 7573-7577.
317. Brooks, B. E., and Buchanan, S. K. (2008) Signaling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. *Bioch Biophys Acta* **1778**, 1930-1945.
318. Mahren, S., and Braun, V. (2003) The FecI extracytoplasmic-function sigma factor of *Escherichia coli* interacts with the beta' subunit of RNA polymerase. *J Bacteriol* **185**, 1796-1802.
319. Burgess, R. R., Arthur, T. M., and Pietz, B. C. (1998) Interaction of *Escherichia coli* sigma 70 with core RNA polymerase. *Cold Spring Harb Symp Quant Biol* **63**, 277-287.
320. Arthur, T. M., and Burgess, R. R. (1998) Localization of a sigma70 binding site on the N terminus of the *Escherichia coli* RNA polymerase beta' subunit. *J Biol Chem* **273**, 31381-31387.
321. Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., and Darst, S. A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285-1290.
322. Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* **417**, 712-719.
323. Basu, R. S., Warner, B. A., Molodtsov, V., Pupov, D., Esyunina, D., Fernandez-Tornero, C., Kulbachinskiy, A., and Murakami, K. S. (2014) Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme. *J Biol Chem* **289**, 24549-24559.
324. Bae, B., Feklistov, A., Lass-Napiorkowska, A., Landick, R., and Darst, S. A. (2015) Structure of a bacterial RNA polymerase holoenzyme open promoter complex. *Elife* **4.**
325. Murakami, K. S. (2013) X-ray crystal structure of *Escherichia coli* RNA polymerase sigma70 holoenzyme. *J Biol Chem* **288**, 9126-9134.
326. Wilson, M. J., and Lamont, I. L. (2006) Mutational analysis of an extracytoplasmic-function sigma factor to investigate its interactions with RNA polymerase and DNA. *J Bacteriol* **188**, 1935-1942.
327. Flynn, J. M., Levchenko, I., Sauer, R. T., and Baker, T. A. (2004) Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes Dev* **18**, 2292-2301.
328. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-4680.
329. Di Tommaso, P., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J. M., Taly, J. F., and Notredame, C. (2011) T-Coffee: A web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* **39**, W13-17.

330. Bond, C. S., and Schuttelkopf, A. W. (2009) ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr* **65**, 510-512.
331. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics* **16**, 404-405.
332. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G. J. (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-394.
333. Webb, B., and Sali, A. (2014) Comparative protein structure modeling using MODELLER. *Curr Protoc Bioinformatics* **47**, 5 6 1-32.
334. Webb, B., and Sali, A. (2014) Protein structure modeling with MODELLER. *Methods Mol Biol* **1137**, 1-15.
335. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402.
336. Melo, F., Sanchez, R., and Sali, A. (2002) Statistical potentials for fold assessment. *Protein Sci* **11**, 430-448.
337. Thakur, K. G., Joshi, A. M., and Gopal, B. (2007) Structural and biophysical studies on two promoter recognition domains of the extra-cytoplasmic function sigma factor sigma(C) from *Mycobacterium tuberculosis*. *J Biol Chem* **282**, 4711-4718.
338. Li, W., Stevenson, C. E., Burton, N., Jakimowicz, P., Paget, M. S., Buttner, M. J., Lawson, D. M., and Kleanthous, C. (2002) Identification and structure of the anti-sigma factor-binding domain of the disulphide-stress regulated sigma factor sigma(R) from *Streptomyces coelicolor*. *J Mol Biol* **323**, 225-236.
339. Sambrook, J., and Russell, D. W. (2006) Purification of expressed proteins from inclusion bodies. *CSH Protoc* **2006.**
340. Anthony, J. R., Green, H. A., and Donohue, T. J. (2003) Purification of *Rhodobacter sphaeroides* RNA polymerase and its sigma factors. *Methods Enzymol* **370**, 54-65.
341. Silva-Rocha, R., Martinez-Garcia, E., Calles, B., Chavarria, M., Arce-Rodriguez, A., de Las Heras, A., Paez-Espino, A. D., Durante-Rodriguez, G., Kim, J., Nikel, P. I., Platero, R., and de Lorenzo, V. (2013) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* **41**, D666-675.
342. Durante-Rodriguez, G., de Lorenzo, V., and Martinez-Garcia, E. (2014) The Standard European Vector Architecture (SEVA) plasmid toolkit. *Methods Mol Biol* **1149**, 469-478.
343. Martinez-Garcia, E., Aparicio, T., Goni-Moreno, A., Fraile, S., and de Lorenzo, V. (2015) SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. *Nucleic Acids Res* **43**, D1183-1189.
344. Ruvkun, G. B., and Ausubel, F. M. (1981) A general method for site-directed mutagenesis in prokaryotes. *Nature* **289**, 85-88.
345. Prentki, P., and Krisch, H. M. (1984) In vitro insertional mutagenesis with a selectable DNA fragment. *Gene* **29**, 303-313.
346. Bitter, W., Marugg, J. D., de Weger, L. A., Tommassen, J., and Weisbeek, P. J. (1991) The ferric-pseudobactin receptor PupA of *Pseudomonas putida* WCS358: Homology to TonB-dependent *Escherichia coli* receptors and specificity of the protein. *Mol Microbiol* **5**, 647-655.

347. Marugg, J. D., Nielander, H. B., Horrevoets, A. J., van Megen, I., van Genderen, I., and Weisbeek, P. J. (1988) Genetic organization and transcriptional analysis of a major gene cluster involved in siderophore biosynthesis in *Pseudomonas putida* WCS358. *J Bacteriol* **170**, 1812-1819.
348. Meyer, J. M., and Abdallah, M. A. (1978) The fluorescent pigment of *Pseudomonas fluorescens*: biosynthesis, purification and physiochemical properties. *J Gen Microbiol* **107**, 319-328.
349. Yang, C. C., and Leong, J. (1984) Structure of pseudobactin 7SR1, a siderophore from a plant-deleterious *Pseudomonas*. *Biochem* **23**, 3534-3540.
350. Buyer, J. S., Sikora, L. J., and Chaney, R. L. (1989) A new growth medium for the study of siderophore-mediated interactions. *Biol. Fertil. Soils* **8**, 98-101.
351. Jacquamet, L., Aberdam, D., Adrait, A., Hazemann, J. L., Latour, J. M., and Michaud-Soret, I. (1998) X-ray absorption spectroscopy of a new zinc site in the fur protein. *Biochem* **37**, 2564-2571.
352. Althaus, E. W., Outten, C. E., Olson, K. E., Cao, H., and O'Halloran, T. V. (1999) The ferric uptake regulation (Fur) repressor is a zinc metalloprotein. *Biochem* **38**, 6559-6569.
353. Urbauer, J. L., Adelman, K., Urbauer, R. J., Simeonov, M. F., Gilmore, J. M., Zolkiewski, M., and Brody, E. N. (2001) Conserved regions 4.1 and 4.2 of sigma(70) constitute the recognition sites for the anti-sigma factor AsiA, and AsiA is a dimer free in solution. *J Biol Chem* **276**, 41128-41132.
354. Urbauer, J. L., Simeonov, M. F., Urbauer, R. J., Adelman, K., Gilmore, J. M., and Brody, E. N. (2002) Solution structure and stability of the anti-sigma factor AsiA: Implications for novel functions. *PNAS* **99**, 1831-1835.
355. Gilmore, J. M., Bieber Urbauer, R. J., Minakhin, L., Akoyev, V., Zolkiewski, M., Severinov, K., and Urbauer, J. L. (2010) Determinants of affinity and activity of the anti-sigma factor AsiA. *Biochem* **49**, 6143-6154.
356. Adelman, K., Orsini, G., Kolb, A., Graziani, L., and Brody, E. N. (1997) The interaction between the AsiA protein of bacteriophage T4 and the sigma70 subunit of *Escherichia coli* RNA polymerase. *J Biol Chem* **272**, 27435-27443.
357. Devkota, S. R., Kwon, E., Ha, S. C., Chang, H. W., and Kim, D. Y. (2017) Structural insights into the regulation of *Bacillus subtilis* sigW activity by anti-sigma RsiW. *PloS One* **12**, e0174284.
358. Templeton, D. (2002) *Molecular and Cellular Iron Transport*, CRC Press.
359. Kramer, R., and Jung, K. (2009) *Bacterial Signaling*, John Wiley & Sons.
360. Visca, P., Leoni, L., Wilson, M. J., and Lamont, I. L. (2002) Iron transport and regulation, cell signalling and genomics: lessons from *Escherichia coli* and *Pseudomonas*. *Mol Microbiol* **45**, 1177-1190.
361. Gabel, F., Lensink, M. F., Clantin, B., Jacob-Dubuisson, F., Villeret, V., and Ebel, C. (2014) Probing the conformation of FhaC with small-angle neutron scattering and molecular modeling. *Biophys J* **107**, 185-196.