

**THE STRUCTURE OF GLYCEROL-3-PHOSPHATE
CYTIDYLYLTRANSFERASE FROM
STAPHYLOCOCCUS AUREUS**

**X-RAY CRYSTALLOGRAPHIC STUDIES OF GLYCEROL-3-
PHOSPHATE CYTIDYLYLTRANSFERASE FROM
*STAPHYLOCOCCUS AUREUS***

By

VERONICA CHIU NGA YIM, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

© Copyright by Veronica Chiu Nga Yim

Master of Science (2002)
(Biochemistry)

McMaster University
Hamilton, Ontario

TITLE: X-ray Crystallographic Studies of Glycerol-3-Phosphate Cytidylyltransferase
from *Staphylococcus Aureus*

AUTHOR: Veronica Chiu Nga Yim, B.Sc. (McMaster University)

SUPERVISOR: Professor A.M. Berghuis

NUMBER OF PAGES: 80

ABSTRACT

Glycerol-3-phosphate cytidyltransferase from *Staphylococcus aureus* complexed with CTP (TarD_{Sa}-CTP) was crystallized by the hanging drop-vapor diffusion method at 22°C. Determination of crystallization condition included examination of the amount of precipitant, investigation of the effects of small molecules, and alteration of the rate of diffusion. With these three optimization steps, crystals suitable for x-ray diffraction study were produced. During data processing, TarD_{Sa}-CTP was determined to belong to the space group P3₁21, with unit-cell dimensions a=b=92.2 and c=156.1 Å. The crystal structure of TarD_{Sa}-CTP was solved to 3.0 Å by molecular replacement, using TagD from *Bacillus subtilis* as a search model.

Unlike the search model, TarD_{Sa} appears as a tetramer in the asymmetric unit. This result also confirms the gel-filtration and ultracentrifugation studies that were done previously. Although TarD_{Sa} crystals were grown in the presence of CTP, the crystal structure does not reveal convincing data for the location and position of this co-factor. However, the data suggests a possible location for CTP in one of the four subunits in an orientation that differs from that of TagD_{Bs}. Unfortunately, the resolution of this data set at 3.0 Å is not high enough to corroborate this finding.

ACKNOWLEDGEMENTS

First I would like to say I am extremely thankful to my supervisor, Dr. Albert Berghuis, for his guidance and support in the past two academic years. Without his help, I would never have achieved as much accomplishment throughout the master's study, especially with the publication on the crystallization and preliminary x-ray diffraction studies of glycerol-3-phosphate cytidylyltransferase from *Staphylococcus aureus*. I am also appreciative for his patience whenever I asked questions concerning the x-ray crystallographic studies.

I would also like to acknowledge my co-supervisor, Dr. Eric Brown, for his thoughtful opinions on why the CTP binding sites were not filled in the tetramer structure of TarD_{Sa}. I am also grateful that his enthusiasm about this project inspired me throughout the study. Further thanks to David Badurina in Dr. Eric Brown's lab for keeping me updated with his kinetics results of TarD_{Sa}.

I would like to further thank my committee supervisory member, Dr. Lynne Howell from the University of Toronto, for kindly providing the references which supported the speculation about the CTP binding conformation. I would like to thank her for her helpful discussion regarding crystallization technique and structural analysis.

I would like to thank every member of Dr. Berghuis' laboratory for their assistance: in particular, Laura Rossi for showing me how to purify TarD_{Sa} and grow crystals, Ahmad Mirza, the "hugemycin" expert, for sharing his wonderful stories and daily adventures, David Burk and Jonathan Cechetto, for assisting me in solving the technical problems on the computer, and Noboru Ishiyama, for keeping me awake when driving to Brookhaven.

I am extremely thankful to my very good friend, Desiree Fong, for proofreading this thesis and further refining this structure for me. She also showed me how to mount crystals, make a better arrangement for my PowerPoint presentation, and generate the structural figures using the Molscrip software. I would like to extend my thanks to her for her friendship and support. I would also like to thank her for her wonderful advice and comments throughout this study.

I would like to acknowledge the staff members of the National Synchrotron Light Source at Beamline X8C at the Brookhaven National Laboratory for assistance with data collection.

I would like to express my final gratitude to my parents for being supportive and patient with me during my graduated studies. I would like to particularly thank my mom for her stay in Hamilton, and her wonderful cooking every day.

TABLE OF CONTENTS

Abstract.....	i
Acknowledgements.....	ii
Table of Contents.....	iv
List of Figures.....	vii
List of Tables.....	x
List of Abbreviations.....	xi
CHAPTER 1 INTRODUCTION.....	1
1.1 Multi-Antibiotics Resistance in <i>Staphylococcus aureus</i>	1
1.2 Introduction to Cell Wall Composition of Gram-Positive Bacteria and Glycerol-3-Phosphate Cytidylytransferase.....	6
1.3 Sequential and Functional Comparisons of TarD _{Sa} and Other Structural-Related Enzymes.....	11
1.4 Structural Analysis of TagD _{Bs} with bound CTP.....	16
1.5 Thesis Objective.....	20
CHAPTER 2 PURIFICATION AND CRYSTALLIZATION OF TarD_{Sa}.....	21
2.1 Over-expression and Purification.....	21
2.1.1 Materials.....	21
2.1.2 Over-experssion and Purification of TarD _{Sa}	22
2.2 Crystallization.....	24
2.2.1 Materials.....	24

2.2.2	General Crystallization Set Up.....	24
2.2.3	Optimization Tool I: Fine Screening Technique.....	25
2.2.4	Optimization Tool II: Crystallization Using Additives.....	27
2.2.5	Optimization Tool III: Crystallization Oils.....	27
2.3	Results and Discussion.....	29
2.3.1	Crystallization of TarD _{Sa}	29
2.3.2	Co-crystallization of TarD _{Sa} with CTP.....	30
2.3.3	Co-crystallization of TarD _{Sa} with Glycerol-3-Phosphate.....	36
CHAPTER 3 RESULTS AND DISCUSSION.....		40
3.1	Data Collection and Processing.....	40
3.2	Phase Determination.....	43
3.3	Model Building and Refinement.....	50
3.4	Results and Discussion.....	55
3.4.1	Overall Structure and Subunit Folding.....	55
3.4.2	Tetramerization of TarD _{Sa} vs. Dimerization of TagD _{Bs}	58
3.4.3	A Tentative CTP Binding Site in TarD _{Sa}	60
3.5	Summary and Future Work.....	63
3.5.1	Summary.....	63
3.5.2	Future Work and Concluding Remarks.....	64
REFERENCES.....		66
APPENDIXES.....		71
	Appendix I Molecular Replacement Method.....	71

AI.1	Introduction and Basic concept of Molecular Replacement.....	71
AI.2	Rotation Function.....	74
AI.3	CNS: Direct Rotation and PC Refinement.....	77
AI.4	Translation Function.....	79
AI.5	$R_{\text{conventional}}$ vs. R_{free}	80
Appendix II Procedure for Preparing Siliconized Cover-glass.....		81

LIST OF FIGURES

Figure 1.1	Gram Positive <i>S. aureus</i> under Microscopic Examination.....	1
Figure 1.2	The Similarities and Differences of Cell Wall Compositions in <i>S. aureus</i> and <i>B. subtilis</i>	7
Figure 1.3	Proposed Teichoic Acid Polymer Biosynthesis in <i>S. aureus</i>	8
Figure 1.4	Enzymatic Reaction Performed by TarD _{Sa} and TagD _{Bs}	9
Figure 1.5	Roles Performed by the Reactant GroP and Product CDP-glycerol During the Linkage Unit Assembly.....	9
Figure 1.6	Comparisons of the Enzymatic Reactions of PS, PPAT, QtRS, and TagD _{Bs}	14
Figure 1.7	Sequence Alignment of TarD _{Sa} with TagD _{Bs}	15
Figure 1.8	Crystal Structure of TagD _{Bs} Bound with Two CTP at 2.0Å.....	16
Figure 1.9	CTP Binding Site in the Crystal Structure of TagD _{Bs}	17
Figure 1.10	CTP Binding Interactions in the Crystal Structure of TagD _{Bs}	18
Figure 2.1	Purification Gel for TarD _{Sa}	23
Figure 2.2	Reference Point in the System of Fine Screen Setting.....	26
Figure 2.3	Quasi Crystals of Apo-TarD _{Sa}	29
Figure 2.4	Crystals of TarD _{Sa} -CTP Grown with a Combination of Three Optimization Strategies.....	32
Figure 2.5	Soaking Procedures.....	35
Figure 2.6	Crystals of TarD _{Sa} -GroP Obtained from the First Crystallization Condition and the Izit Test Result.....	36

Figure 2.7	Crystals of TarD _{Sa} -GroP Obtained from the Second Crystallization Condition and the Izit Test Result of the Fine Screen	37
Figure 2.8	Crystals of TarD _{Sa} -GroP Obtained from the Repetitive Fine Screen of the Second Crystallization Condition and the Optimal Additive and Oils Conditions.....	39
Figure 2.9	Examination of the Composition of the Non-stained Crystals Obtained from the Additive and Oils Conditions.....	39
Figure 3.1	Proposed Scheme Used to Explain How the First Answer Calculated from Previous Rotation and Translation is Fixed in the Translation Search.....	47
Figure 3.2	Proposed Scheme Used to Explain the Result of the Translation Search with the Second Rotation Solution	47
Figure 3.3	Proposed Scheme Used to Explain the Result of the Translation Search with the Third Rotation Solution	48
Figure 3.4	Proposed Scheme Used to Explain the Result of the Translation Search with the Fourth Rotation Solution	48
Figure 3.5	CTP in a Fo-Fc sa omit at 1 sigma contour.....	50
Figure 3.6	Ramachandran Plot of TarD _{Sa}	52
Figure 3.7	Overall Structure of TarD _{Sa}	55
Figure 3.8	Monomer folding of TarD _{Sa}	56
Figure 3.9	Overlay of TarD _{Sa} with Subunit TagD _{Bs}	56
Figure 3.10	Dimer Interface Region in TagD _{Bs} Vs Tetramer Interface Region in TarD _{Sa}	58
Figure 3.11	Overlay of CTP in TagD _{Bs} with the One in TarD _{Sa}	60
Figure 3.12	Location of the CTP observed in a Fo-Fc sa omit map at 1 sigma contour.....	60
Figure 3.13	TagD _{Bs} 's CTP Modeled in the Crystal Structure of TarD _{Sa}	61
Figure AI.1	Basic Concept of Rotation and Translation Function.....	71

Figure AI.2	Diagrams of Patterson Function and Intra- and Inter-molecular Vectors..	74
Figure AI.3	Maximum Overlap of Patterson Functions Resulted in the Rotation Search.....	75
Figure AI.4	Direct Rotation	76

LIST OF TABLES

Table 1.1	Most Common Syndromes Caused by <i>S.aureus</i>	2
Table 1.2	Historical Progression of Antibiotic Resistance of <i>S.aureus</i>	5
Table 1.3	Lists of All Enzymes Structurally Related to TagD _{Bs} and TarD _{Sa}	11
Table 2.1	Variables and Gradient Employed in the Fine Screen Setting.....	25
Table 2.2	Fine screen set up.....	25
Table 3.1	Space Groups Defined in the Primitive Hexagonal Family.....	41
Table 3.2	Data Processing Statistics for the Diffraction Studies of Crystals of TarD _{Sa}	42
Table 3.3	Lists of the Direct Cross Rotation Results.....	44
Table 3.4	Lists of Solutions Obtained from the Translation Search in Space Group P ₃ ₁ 21 Using the TagD _{Bs} Dimer as the Search Model.....	46
Table 3.5	Lists of Solutions Obtained from the Translation Search in Space Group P ₃ ₂ 21 Using the TagD _{Bs} Dimer as the Search Model.....	46
Table 3.6	Lists of Solutions Obtained from the Translation Search of an Asymmetric Tetramer in Space Group P ₃ ₁ 21 Using the Rotation-Searches dimers as the Search Model.....	48
Table 3.7	Statistics of the Final Refined Crystal Structure of TarD _{Sa} at 3.0Å.....	53

LIST OF ABBREVIATIONS

Ade	adenosine
Amp	ampicillin
ATP	adenosine triphosphates
Ca(OAc) ₂	calcium acetate
CCT	choline-phosphate cytidyltransferase
CDP	cytidine diphosphate
CMP	cytidine monophosphate
CTP	cytidine triphosphate
Cyt	cytidine
D-Ala	D-alanine
DTT	dithiothreitol
ECT	ethanolamine-phosphate cytidyltransferase
EDTA	ethylenediaminetetraacetic acid
EtRNA	glutaminyI-tRNA synthetase
GlcNAc	N-acetylglucosamine
GroP	glycerol-3-phosphate
HEPES	<i>N</i> -[2-hydroxyethyl]piperazine- <i>N'</i> -[2-ethanesulfonic acid]
KDa	kilo Dalton
MAD	multiple wavelength anomalous diffraction
ManNAc	N-acetylmuramic acid

MES	2-[<i>N</i> -morpholino]ethanesulfonic acid
MRSA	methicillin-resistant <i>Staphylococcus aureus</i>
Mg ²⁺	magnesium
MgCl ₂	magnesium chloride
NaCl	sodium chloride
NCBI	National Center for Biotechnology Information
NLS	National Synchrotron Light Source
PC	Patterson correlation
PDB	Protein Data Bank
PEG	polyethylene glycol
PPAT	phosphopantetheine adenytransferase
PPi	pyrophosphate
PS	pantotheate synthetase
QtRNA	glutamyl-tRNA synthetase
sa	simulated annealing
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
TagD _{Bs}	glycerol-3-phosphate cytidyltransferase from <i>Bacillus subtilis</i>
TarD _{Sa}	glycerol-3-phosphate cytidyltransferase from <i>Staphylococcus aureus</i>
VRE	vancomycin-resistant <i>enterococci</i>
YtRNA	tyrosyl-tRNA synthetase

CHAPTER 1

INTRODUCTION

1.1 Multi-Antibiotic Resistance in *Staphylococcus Aureus*

Staphylococci aureus (*S. aureus*), gram-positive bacteria, appear in pairs of spherical clusters (Figure 1.1) and grow selectively at temperatures from 30 to 37°C and pHs from 7 to 7.5. *S. aureus* is always found in nasal passages, throat, hair and skin of humans and animals in the community. However,

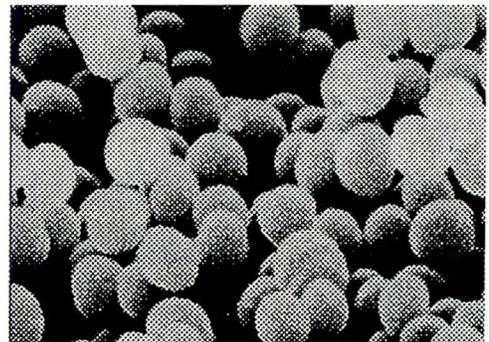


Figure 1.1 Gram positive *S. aureus* under microscopic examination (Picture is taken from Microbiology an Introduction, Fifth Edition, p.283)

S. aureus does not usually cause any diseases in a healthy individual unless the colonization is initiated by binding of a surface protein, protein A, to an extracellular matrix protein, fibronectin (Kobayash et al., 1999). *S. aureus* can be spread to patients simply from direct contact with health-care workers. If this organism is able to penetrate and survive in tissues, it will cause a spectrum of bacterial infections in patients. Surgical wound infections, endocarditis, and scalded skin and toxic shock syndromes are a few examples (Table 1.1). The bacterium's ability to release numerous toxins (e.g. enterotoxin, exfoliative toxin, cytolytic toxins, etc.) is associated with its powerful pathogenicity. The rapid development of resistance in *S. aureus* to the clinical antibiotics administered also makes the treatment of these infections more difficult (Lowy, 1998).

Table 1.1 Most Common Syndromes Caused by *S. aureus* (Lowy, 1998)

Furuncle or carbuncle
Impetigo bullosa
Surgical wound infection (e.g. osteomyelitis)
Pyomyositis
Botryomycosis
Acute or right-sided endocarditis
Epidural abscess
Toxic shock syndrome
Scalded skin syndrome

Whenever antibiotics are prescribed, there is a possibility for variant strains to survive. Antibiotic resistance is evolved from the addition of genetic information to the chromosome via transformation, or spontaneous DNA mutations in chromosomal genes themselves (Lewis, 1995). This mutant can then share the genetic material encoding for the resistance phenotype with other antibiotic-susceptible strains and species through extra-chromosomal elements. This is how antibiotic resistances develop and spread in bacteria.

Since the beginning of the antibiotic therapy era (~1940), *S. aureus* had quickly responded to many drugs through the genetic actions mentioned above. The chronology of antimicrobial resistance of *S. aureus* is listed in Table 1.2. During the Second World War, penicillin, a β -lactam antibiotic, was found to be the most powerful antimicrobial agent for curing infected wounds in wartime casualties as well as hospitalized patients. However, only several years after the introduction of penicillin, *S. aureus* was the first bacteria to acquire resistance against β -lactam antibiotics. Penicillin-resistance in *S.*

aureus is caused by the production of a β -lactamase that is encoded by the gene *bla_Z*, which catalyzes the breaking of the β -lactam ring, thus inactivating β -lactam antibiotics (Lowy, 1998). β -lactams are substrate analogs designed to inactivate the membrane-bound peptidases, PBP proteins, which links the nascent peptidoglycan chains to pre-existing peptidoglycan layers of the bacterial cells (Hiramatsu, 2001). Therefore, inhibition of PBP proteins that causes the discontinuation of peptidoglycan biosynthesis is found to impair bacterial growth.

Resistance to Methicillin was detected in *S. aureus* in the late 1960s, and the isolation of MRSA had been identified as multi-resistance to the β -lactam antibiotics. At Johns Hopkins University Hospital in the early 1960s, 27% of patient fatalities resulted from infections caused by these organisms. From 1975 to 1991, the percentage of MRSA increased from 2.4% to 29% in United States hospitals (Archer, 1998), and MRSA has now spread worldwide (Lowy, 1998). The evolution of MRSA is due to the acquisition of a unique penicillin-binding protein designated PBP2' or PBP2A that has a greatly reduced affinity for most β -lactam antibiotics (Zhang et al., 2001; Archer & Bosilevac, 2001).

The occurrence of MRSA is due to the vast production of PBP2A, which is synthesized to replace other PBP proteins in the bacteria. Although the exact mechanism for PBP2A is not known, methicillin is suspected to be the inducible factor for PBP2A protein synthesis.

Vancomycin became the next antibiotic for treating diseases caused by MRSA. Reported by the National Nosocomial Infections Surveillance system of the Centers for Disease Control and Prevention (CDC), the percentage of MRSA with sensitivity only to Vancomycin increased from 2.8% to 56.2% between 1987 and 1997 (Lowy, 1998). More importantly, *S. aureus* strains have been shown to develop resistance to the antibiotic vancomycin (Smith, 1999). In 1997, reports of reduced susceptibility to vancomycin in Japan, Michigan and New Jersey represented the most frightening incidence of drug resistant *S. aureus* (Marchese et al., 2000). It also illustrates that the outbreaks of vancomycin-resistant strains in *S. aureus* (VRSA) have already emerged in different nations. The thickening of the cell wall is a major issue of the development of VRSA. Because more peptidoglycan layers (~30-40 layers) are produced by VRSA strains, vancomycin will have difficulty in penetrating to the cytoplasmic membrane where the synthesis of peptidoglycan occurs. The trapping of vancomycin will also destroy the outer layer of thickened peptidoglycan, clogging the entry to the inner layer of the cell wall polymers (Hiramatsu, 2001).

Zyvox, released by New Jersey-based Pharmacia Corporation in the last year, was the latest antibiotic approved for treating infections including both MRSA and VRE¹. This antimicrobial agent, which is in a new class of antibiotics named oxazolidinones, is distinct from other drugs since it employs a different mode of inhibition. Zyvox is able to terminate the bacteria's multiplication early in the life cycle by stopping the production

¹ <http://id.medscape.com/reuter/prof/2000/04/04.19/20000419rglt007.html>

of proteins that are vital to their growth (Champney & Miller, 2002). Only one year later, however, *S. aureus* resistance to Zyvox has been reported in an 85-year-old patient undergoing peritoneal dialysis in Westport-CT². Fortunately, this resistant strain was still susceptible to other clinical antibiotics, while most other MRSA were not. These incidences strongly suggest that *S. aureus* strains can readily become fully resistant to all available antimicrobial agents in the near future.

Table 1.2 Historical progression of antibiotic resistance of *S. aureus* (Delage, 1999)

<u>Antibiotic</u>	<u>Year Introduced</u>	<u>Reports of resistance</u>
Penicillin	1941	1940s
Streptomycin	1944	mid-1940s
Tetracyclin	1948	1950s
Erythromycin	1952	1950s
Methicillin	1959	late 1960s
Gentamicin	1964	mid-1970s
Vancomycin	1958	1997
Zyvox	2000	2001

² <http://id.medscape.com/reuters/prof/2001/07/07.20/20010719clin004.html>

1.2 Introduction to Cell Wall Composition of Gram-Positive Bacteria and Glycerol-3-Phosphate Cytidylyltransferase

To date, cell wall biosynthesis is still categorized as the primary choice of therapeutic target against many gram-positive bacteria, due to its absence in eukaryotes. Prevention of extensive cell wall breakdown and protection from cell lysis are crucial factors for bacterial survival. Penicillin and methicillin are such antibiotics aimed at halting the formation of the cell wall polymers.

Cell wall composition of gram-positive bacteria contains three parts: peptidoglycan, linkage unit, and teichoic acid polymer, as illustrated in Figure 1.2. Teichoic acid polymer serves as the major constituent in the cell wall of gram-positive bacteria. In *S. aureus*, this polymer consists of a repeating unit of ribitol phosphate, with D-Ala and glucose molecules branched on the 2- and 4-positions respectively. Instead, this polymer in *Bacillus subtilis* is composed of a chain of glycerol-3-phosphate with a glucose unit attached to the hydroxyl group of the glycerol molecule. Teichoic acid's polymer is covalently bonded to the peptidoglycan via a phosphodiester bond between the linkage unit of teichoic acid and the hydroxyl of *N*-acetylmuramic acid of the peptidoglycan layer. The linkage unit in the cell wall composition of gram-positive microorganisms is relatively consistent and usually contains a disaccharide component of *N*-acetylglucosamine and *N*-acetylmuramic acid and three consecutive glycerol-3-phosphates (GroP).

peptidoglycan. Several mutation studies in the poly-glycerol phosphate teichoic acid gene cluster (*tag*) of *B. subtilis* have identified the roles of the *tag* gene cluster in the teichoic acid biosynthesis (Briehl et al., 1989; Mauel et al., 1991; Pooley et al., 1991).

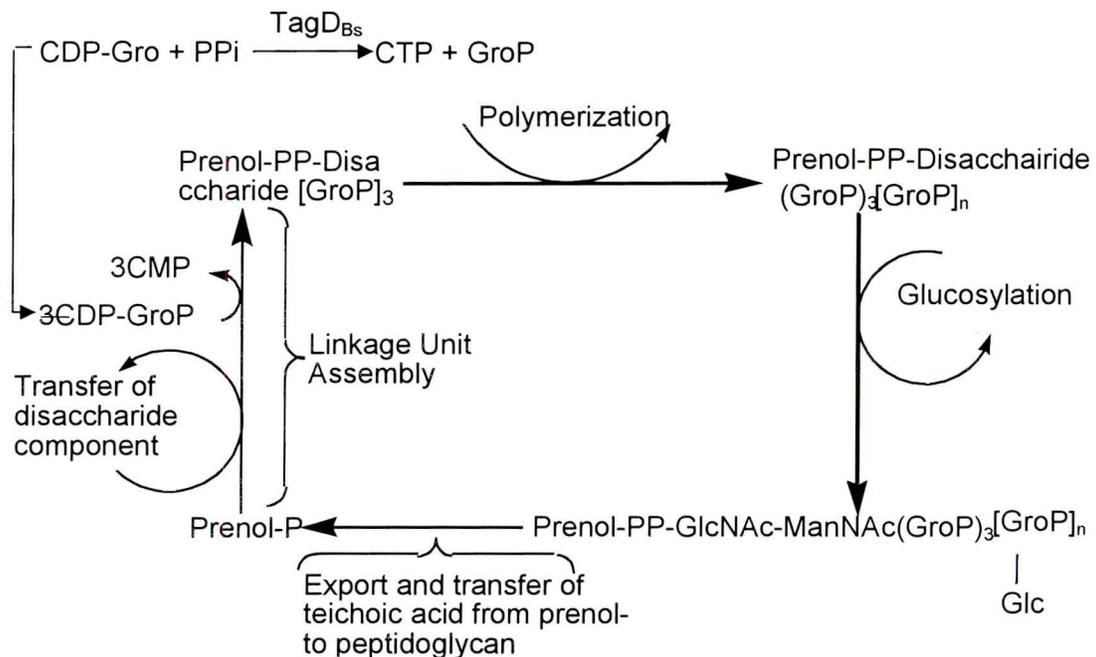


Figure 1.3 Proposed Teichoic acid polymer biosynthesis in *B. subtilis* 168. This pathway is shown to require three basic steps: 1) linkage unit assembly, 2) polymerization and glucosylation, and 3) export and transfer of teichoic acid from prenel-P to peptidoglycan. This pathway in *S. aureus* is proposed to generate and transfer ribitol phosphate instead of glycerol phosphate during the polymerization of teichoic acid unit.

A precise deletion of *tagD_{Bs}* (Bhavsar, 2001) has shown that the absence of glycerol-3-phosphate cytidyltransferase leads to irregular cell shape and division followed by cell lysis. These findings suggest that glycerol-3-phosphate cytidyltransferase is a potential drug target against gram-positive pathogens. Glycerol-3-phosphate cytidyltransferase from *B. subtilis* (*TagD_{Bs}*) is the enzyme that catalyzes

the formation of CDP-glycerol for the linkage unit assembly (Baddiley, 1970). Through a complementary mutation studies performed by David Badurina in Dr. Eric Brown's lab, at McMaster University, the gene *tarD_{Sa}* was shown to have the similar physiological role as the function of the gene *tagD_{Bs}* in the living cell (Figure 1.4). However, their enzymatic mechanisms were found to be very different from each other. This will be discussed in the following section. The fate of the enzymatic product of TarD_{Sa} and TagD_{Bs} is that three synthesized CDP-glycerol molecules will be converted to three CMP(s) followed by a release of three glycerol-3-phosphate(s). They will be then transferred to the prenol-P-disaccharide component in order to assemble the linkage unit of teichoic acid (Figure 1.5).

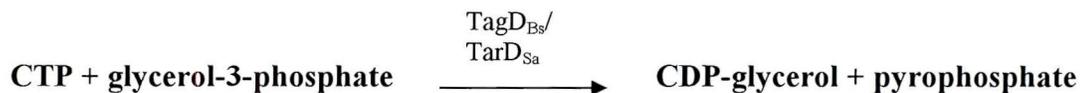


Figure 1.4 Enzymatic reaction performed by TarD_{Sa} and TagD_{Bs}. They are both responsible for the formation of CDP-glycerol and pyrophosphate from glycerol-3-phosphate and CTP

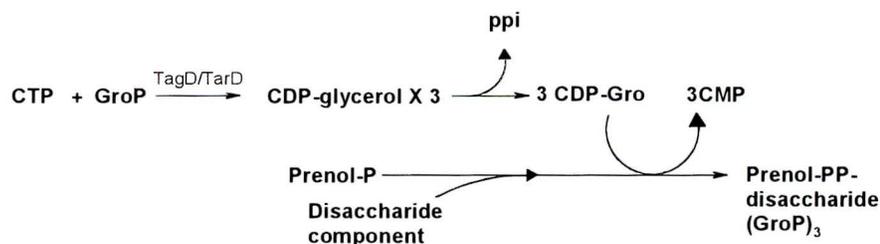


Figure 1.5 Roles played by the reactant GroP and product CDP-glycerol during the linkage unit assembly. Three products of CDP-glycerol will be transferred and converted to three CMP, and the products of this reaction GroP will be linked to the prenol-PP-disaccharide unit.

TarD_{Sa} is the subject of this study. David Badurina had been focused on the functional characterization of TarD_{Sa}. By using the analytical S-200 gel-filtration column, he determined the molecular weight of TarD_{Sa} in oligomeric state to be 64kDa. This finding was also confirmed by sedimentation analysis, done in the Princess Margaret Hospital at the University of Toronto. Subunit molecular weight of ~16kDa was also observed on the 15% SDS-PAGE gel, suggesting that TarD_{Sa} probably exists as a tetramer in solution. The K_m values of CTP and glycerol-3-phosphate for TarD_{Sa} are 36.1±5.8µM and 20.7±4.1µM respectively. He had also determined that the cytidyltransferase reaction catalyzed by TarD_{Sa} is an order bi-substrate reaction. The enzyme in which the glycerol-3-phosphate binds first and the pyrophosphate is released last. In contrast, TagD_{Bs} begins with a binding of CTP and ends with a formation of cytidylated reaction product.

To monitor the enzymatic reaction of TarD_{Sa}, David Badurina performed an enzymatic assay and used HPLC to purify the reaction products. The eluted species corresponding to the enzymatic product was collected and determined to be CDP-glycerol using the mass spectroscopy analysis. The presence of divalent cation, Mg²⁺, was essential for the catalytic binding of CTP and glycerol-3-phosphate. To select for the most optimal condition in which TarD_{Sa} remains monodisperse, dynamic light scattering experiments (Mikol, 1990; Zulauf & D'Arcy, 1992) were performed by Laura Rossi, a technician in Dr. Berghuis's lab at McMaster University. Ms. Rossi found that TarD_{Sa} in 25mM Hepes at pH 8.0 and 1mM DTT gave a monodispersity index of ~0.20.

1.3 Sequential and Functional Comparisons between TarD_{Sa} and Other Structural-related Enzymes

TarD_{Sa} is made up of 132 amino acids. The sequence is available from the National Centre for Biotechnology Information (NCBI), GI code CAA62900. The sequence of TarD_{Sa} is shown to bear two cytidylyltransferase signature motifs: HXGH and RTXGISTT. Cytidylyltransferase members that carry both of these conserved sequences include glycerol-3-phosphate cytidylyltransferase from *B. subtilis* (TagD_{Bs}), choline-phosphate cytidylyltransferase (CCT) from rat (Sweitzer & Kent, 1994), and ethanolamine-phosphate cytidylyltransferase (ECT) from human (Weber et al., 1999). The first conserved motif, HXGH, is also consistent in other cytidylyltransferase superfamily members (Von Delft et al., 2001) such as pantotheate synthetase (PS), phosphopantetheine adenylyltransferase (PPAT), and the class I aminoacyl-tRNA synthetases family: glutamyl-tRNA synthetase (QtRNA), glutaminyl-tRNA synthetase (EtRNA), and tyrosyl-tRNA synthetase (YtRNA), as listed in Table 1.3.

Table 1.3 List of all enzymes structurally related to TarD_{Sa} and TagD_{Bs}

1. Apo-Structure of Pantotheate Synthetase (PS) from *E. coli* (PDB code: 1iho)
2. Phosphopantetheine Adenylyltransferase (PPAT) in complex with 3'-Dephospho-Coa from *E. coli* (PDB code: 1b6t)
3. Class I aminoacyl-tRNA synthetases family:
 - Apo-Structure of Glutamyl-tRNA Synthetase (QtRNA) (PDB code: 1gln)
 - Glutaminyl-tRNA Synthetase (EtRNA) complexed with tRNA and ATP (PDB code: 1gtr)
 - Apo-Structure of Tyrosyl-tRNA Synthetase (YtRNA) (PDB code: 2ts1)

In class I aminoacyl tRNA synthetases, the two His residues in the HXGH motif have been shown to be involved in the catalytic binding to the phosphate of ATP, so that the pentacoordinate transition state is stabilized (Perona et al., 1993; Venkatachalam et

al., 1998). This is confirmed by the crystal structure of glutaminyl-tRNA synthetase complexed with tRNA and ATP (Arnez & Steitz, 1994). Likewise, the HXGH motif in TagD_{Bs} has been interpreted to be the functional conserved sequence for the cytidylyltransferase activity through the mutagenesis experiments. These studies have showed that the mutated His14 and His17 in TagD_{Bs} resulted in reduction in k_{cat} by several orders of magnitude, suggesting that these two residues are essential for activity of the enzyme (Young et al., 1997). The structure of TagD_{Bs} with bound CTP (Weber et al., 1999) also confirms the enzymatic roles played by the two His residues in the HXGH motif by the observation of the residues binding to the α - and β -phosphate of the substrate.

The enzymatic reactions performed by cytidylyltransferase and class I aminoacyl tRNA synthetases have shown that these two families are functionally related (Von Delft et al., 2001). This explains why the HXGH motif is present in their amino acid sequences. Figure 1.6 schematically depicts a comparison of the enzymatic mechanisms in PS, PPAT, EtRNA, and TagD_{Bs}. All of the compared enzymes including TagD_{Bs} begin with a nucleophilic attack on the α -phosphate atom of the substrate of either ATP or CTP, followed by a release of pyrophosphate and the formation of either cytidylylated/adenylated intermediate or reaction product.

RTXGISTT is the second dominant motif found in the following cytidylyltransferases: RTEGISTT in TagD_{Bs} and TarD_{Sa}, RTEGISTS in CCT, and RTQGVSTT in ECT. This sequence in TagD_{Bs} is found close enough to interact with the HXGH motif, and stabilizes the β - and γ -phosphate of CTP according to the structural analysis of TagD_{Bs} (Weber et al., 1999). For class I aminoacyl-tRNA synthetase, a different amino acid sequence (KMSKS in QtRNA, KFGKT in YtRNA synthetase, VMSKR in EtRNA synthetase) is utilized to perform the similar stabilization of the substrate ATP. Although RTXGISTT and VMSKR bear a different kind of binding conformation in these two classes of protein, the structural comparison of the crystal structures of TagD_{Bs} and QtRNA synthetase reveals that ¹¹⁷IST and ²³¹MSK are the structurally equivalent residues (Weber et al., 1999). Mutational analysis of YtRNA synthetase also shows that movement is occurring in the ²³⁰KFGKT region during the formation of tyrosyl adenylate (Fersht et al., 1988). This suggests that movement of the ²³⁰KFGKT sequence is required for substrate binding.

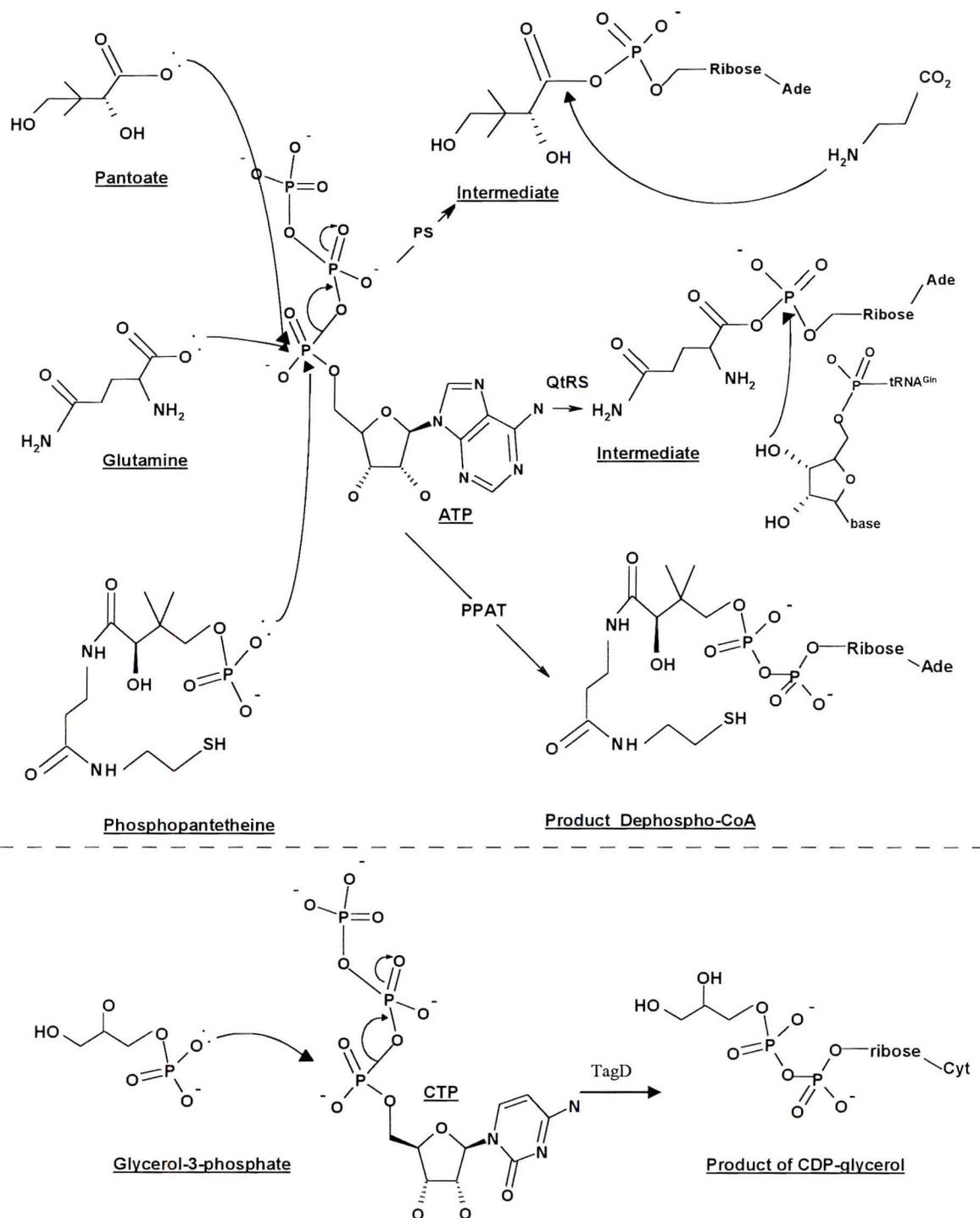


Figure 1.6 Comparisons of the enzymatic reactions of PS, PPAT, QtRS, and TagD_{Bs}. (Picture is adapted from reference of crystal structure of *E. coli* Pantothenate Synthetase (Von Delft, 2001)) PS, PPAT, QtRS, and TagD_{Bs} share a similar mechanism with a nucleophilic attack on the oxygen atom on the α -phosphate of the triphosphate of either ATP or CTP. PS and QtRS require two steps while PPAT and TagD_{Bs} have only one step.

Among the cytidylyltransferase family members, TagD_{Bs} shares the highest sequence identity (68%) with TarD_{Sa}. Alignment of TagD_{Bs} with TarD_{Sa} was performed using BLAST2 as provided on the NCBI³ web site (Figure 1.7). Although TarD_{Sa} and TagD_{Bs} have different kinetic properties, they have been shown to be functionally related (see also Section 1.2) by complementary mutation studies. Since the crystal structure of glycerol-3-phosphate cytidylyltransferase from *Bacillus subtilis* with bound CTP (PDB code: 1coz) has already been solved and published, TagD_{Bs} would be the best choice for a search model in the 3-D structure determination of TarD_{Sa} using molecular replacement methods.

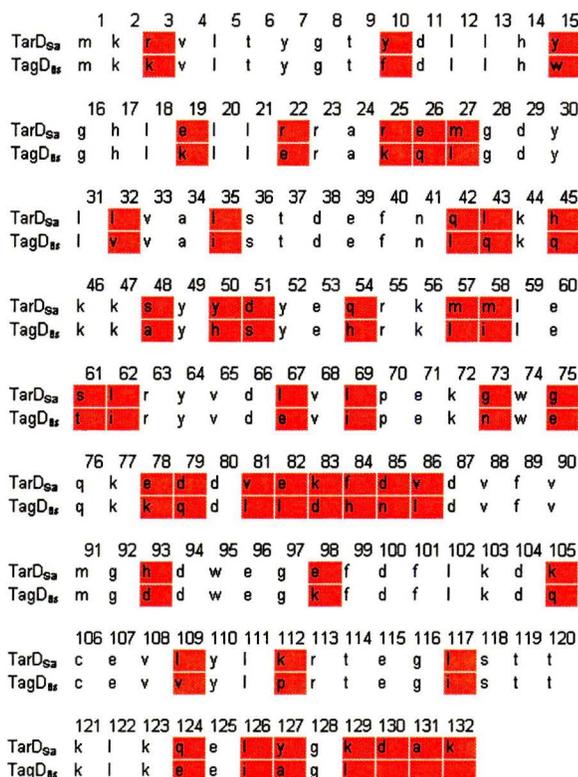


Figure 1.7 Sequence alignment of TarD_{Sa} with TagD_{Bs}. The result is shown with 68% sequence identity.

³<http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html>

1.4 Structural Analysis of TagD_{Bs} with Bound CTP

The three-dimensional structure of *B. subtilis* TagD_{Bs} had been determined by the Martha Ludwig Lab and published in August, 1999 (Weber et al., 1999). The crystal structure of TagD_{Bs} was determined by the MAD technique in space group R3 to 2.8Å, and then further refined in space group P2₁ to 2.0Å using molecular replacement methods. TagD_{Bs} contains two identical molecules in the asymmetric unit, forming a homodimer (Figure 1.8). Each monomer of TagD_{Bs} contains a Rossmann fold, as seen in other crystal structures of PS, PPAT, QtRNA, YtRNA and EtRNA (Von Delft et al., 2000). A Rossmann fold consists of five strands forming a parallel β sheet with a topology of “3-2-1-4-5” sandwiched between two layers of α-helices: (1) helices A, C, and D and (2) helix B on the other side. The referenced short 3₁₀ helix segment (helix E) acts as a cross-connector between strands 4 and 5 (Weber et al., 1999).

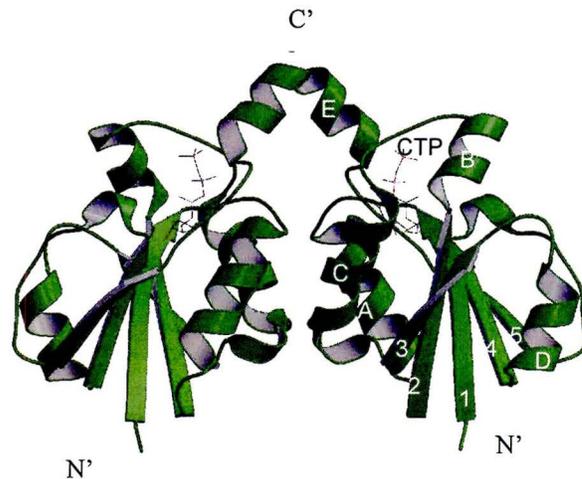


Figure 1.8 Crystal structure of TagD_{Bs} bound with two CTP at 2.0 Å. TagD_{Bs} packs as a dimer with the presence of Rossmann fold in each monomer. Five strands of parallel β sheets with a topology of “3-2-1-4-5” are arranged in the center and also surrounded by helices A-E. N' and C' represent the N- and C-terminus, respectively.

As shown in Figure 1.9, the substrate-binding site in each monomer of TagD_{Bs} is formed by amino acid sequences ⁸GTFDLLHWGH and ¹¹³RTEGISTT (Weber et al., 1999). These are the conserved residues in the superfamily of cytidylyltransferase (see also Section 1.3). Besides these two conserved regions, Lys44, Lys46, and Trp74 are also clustered on the other side of the binding pocket. According to the docking studies (Weber et al., 1999), these residues are predicted to be the binding site for glycerol-3-phosphate. The ₃₁₀ segment between strands 4 and 5 and Trp74 also contribute to this region, serving as the rim of the binding bowl. This study also states that one side of the binding bowl in TagD_{Bs} is open to the solvent. Thus, some residues may need to relocate for the entry of CTP.

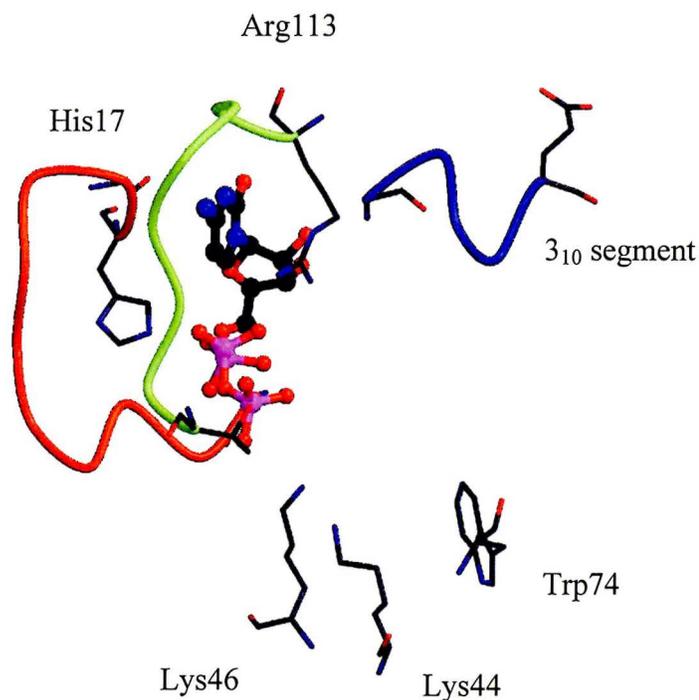


Figure 1.9 CTP binding site in the crystal structure of TagD_{Bs}. The binding bowl of TagD_{Bs} is made of two conserved motifs, ⁸GTFDLLHWGH and ¹¹³RTEGISTT, the ₃₁₀ segment between strands 4 and 5, and the residues, Lys44,46 and Trp 74, which form the putative glycerol binding site.

Interactions for CTP binding primarily take place in the conserved regions of residues ¹⁴HWGH and ¹¹³RTEGISTT. As depicted schematically in Figure 1.10 (left), the amino group inside the imidazole ring of His14 is interacting with the β-phosphate oxygen of CTP. Residues including His17, Thr9, and Phe10 all face toward an α-phosphate. The imidazole side chain of His17 and the backbone amide of Phe10 and Thr9 are also close enough to form interactions with the oxygens on the α-phosphate of CTP.

Figure 1.10 on the right shows that the ¹¹³RTEGISTT motif is sitting at the other side of the substrate-binding bowl, and is partly stabilizing the negative charges of CTP with the backbone amide of Thr119 and Thr120 and the NH group in the side chain of Arg113. The backbone carbonyl oxygen atom of Thr114 and Ile117, on the other hand, is forming an interaction with the exocyclic 4-NH₂ group. The backbone amide of Thr114 is also interacting with the N3 on the pyrimidine ring. However, Mg²⁺, the essential divalent cation for cytidylyltransferase activity, is not observed in this crystal structure.

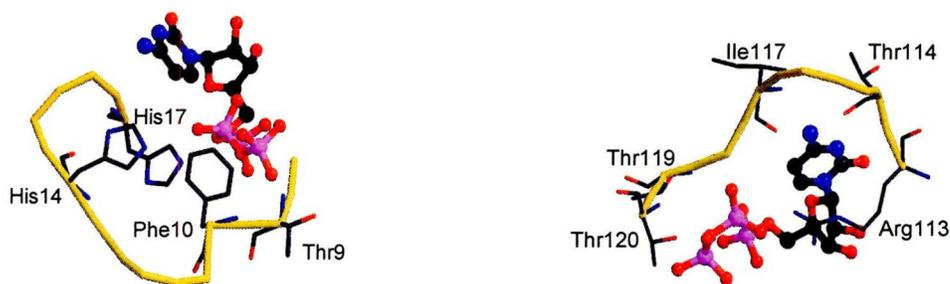


Figure 1.10 CTP binding interactions in the crystal structure of TagD_{Bs}. They are shown to be mediated by motifs ⁸GTFDLLHWGH (left) and ¹¹³RTEGISTT (right).

The crystal structure of TagD_{Bs}-CTP complex is the first member of the cytidyltransferase family to be determined, and thus provides significant information towards monomer structure and CTP binding interactions in the given enzyme family. For example, the structural and binding information obtained from the crystal structure of TagD_{Bs}-CTP has contributed to the study of eukaryotic cytidyltransferase (CCT) in Kent's group. Determination of other binding modes such as GroP in TagD_{Bs} and TarD_{Sa} and the essential divalent cation, Mg²⁺, can be further investigated by x-ray crystallographic studies. In this research, TarD_{Sa} is the subject for the 3D structural analysis, and the binding information obtained from this crystal structure will be used in rational drug design against *S. aureus*.

1.5 Thesis Objectives

S. aureus is the leading cause of infections acquired in hospitals and in the community. Throughout the 1940s to 1990s, various antibiotics were developed to combat *S. aureus* infections. However, resistance to all of these antimicrobial agents has developed in *S. aureus*. The incidence of multi-drug resistant *S. aureus* obviously endangers public health, and this significantly highlights the immediate need for antimicrobial research on a new or improved drug which can successfully combat these strains.

TarD from *S. aureus* catalyzes the conversion of CDP-glycerol and pyrophosphate from CTP and glycerol-3-phosphate during teichoic acid biosynthesis. The absence of glycerol-3-phosphate cytidyltransferase from *B. subtilis* (TagD_{Bs}) has been shown to be lethal to the growth of the cell, suggesting that TarD_{Sa}, a functionally related protein with TagD_{Bs}, is a potential target for the development of novel drugs against *S. aureus*.

The objective of this study is to determine the structures of TarD_{Sa} with its substrates by x-ray crystallographic technique. The structural information obtained from this study will be used for further understanding of the enzymatic mechanism and the design of inhibitors for this enzyme (Klebe, 2000; Amzel, 1998; Gane & Dean, 2000). These inhibitors may prove to be lead compounds for the development of novel antimicrobial agent against *S. aureus*.

CHAPTER 2

PURIFICATION AND CRYSTALLIZATION OF TarD_{Sa}

2.1 Over-expression and Purification of TarD_{Sa}

2.1.1 Materials

E. coli W3110/EB58, the cell strain containing the over-expression system for TarD_{Sa}, was kindly provided by Dr. Eric Brown, at McMaster University. The chemicals for preparation of buffer solutions and substrates including CTP and glycerol-3-phosphate were purchased from Sigma. Bradford Protein Assay reagents and SDS-PAGE low range markers were purchased from Bio-Rad. Buffer solutions employed are as follow:

Column I: Q-Sepharose ion-exchange column

Buffer A: 25mM HEPES pH8.0, 1mM DTT, 5mM EDTA

Buffer B: 25mM HEPES pH8.0, 1mM DTT, 5mM EDTA, 1M NaCl

Column II: Reactive Blue 4 affinity column

Buffer C: 25mM HEPES pH8.0, 1mM DTT, 10mM MgCl₂

Buffer D: 25mM HEPES pH8.0, 1mM DTT, 10mM MgCl₂, 1mM CTP

Column III: S-200 sizing column

Buffer E: 25mM HEPES pH8.0, 1mM DTT

2.1.2 Over-expression and Purification of TarD_{Sa}

The gene *TarD_{Sa}* had been previously cloned into the expression vector p*tarD_{Sa}* (EB58) and transformed into *E. coli* W3110 cells by Michela Zolli in Dr. Eric Brown's Lab, at McMaster University. She also established the purification protocol for TarD_{Sa} as the following. Single colony of *E. coli* W3110/EB58 was picked on the LB-Amp (50µg/mL) agar plate to inoculate 25mL of LB-Amp (50µg/mL) for overnight culture at 37°C. The overnight culture was then used to inoculate 1L LB-Amp (50µg/mL), which was grown at 37°C. When the OD₆₀₀ reached a value of ~0.8, the grown culture was induced with isopropyl-β-thiogalactopyranoside (IPTG) at a final concentration of 1mM at 37°C for 5 hours. The cultures were harvested by centrifugation at 5000g for 10 minutes, washed with 0.85% NaCl and pelleted by centrifugation at 5000g for 10 minutes. The cell pellet was resuspended with Buffer A and lysed by two passes through a French Pressure Cell Press (20,000 psi). Following this, the lysate was centrifuged for one hour at 15,000g, and the collected supernatant was applied to a Q-Sepharose column (~50mL). TarD_{Sa} protein was eluted with a linear gradient of 1M NaCl at ~30% of Buffer B. Fractions containing TarD_{Sa} were dialyzed overnight against Buffer C. After dialysis, the fractions were loaded onto a Reactive blue column (~30mL), and TarD_{Sa} was eluted with Buffer D which contains 1mM CTP. The pooled fractions were concentrated by an 80% ammonium sulfate cut, followed by a Sephadex 200 sizing column (~124mL). The progress of purification was monitored by 15% acrylamine-Coomassie blue and silver-stained SDS-PAGE (Figure 2.1). The concentration of pure protein was determined using the Bradford Assay (Bradford, 1976).

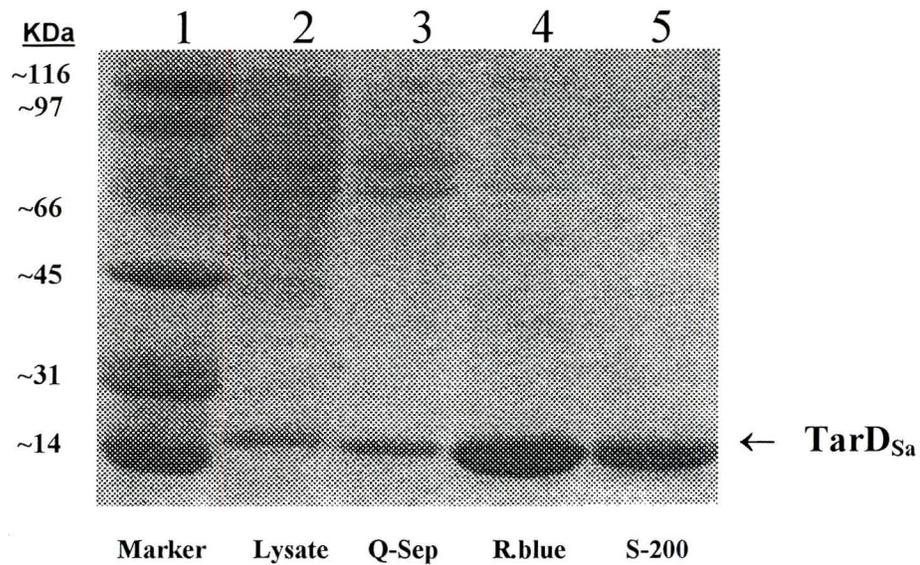


Figure 2.1 Purification gel for TarD_{Sa}. Each purification step is monitored by 15% acrylamine-Coomassie blue and silver-stained SDS-PAGE loaded with 1 μ l protein solution per lane; lane 1, low molecular weight markers in kDa; lane 2, the whole cell lysate; lane 3, TarD_{Sa} eluates from Q-Sepharose column; lane 4, TarD_{Sa} eluates from reactive blue affinity column; lane 5, TarD_{Sa} eluates from S-200 sizing column. TarD_{Sa} migrates with an apparent molecular weight of 16 kDa.

2.2 Crystallization

2.2.1 Materials

Linbro tissue culture multi-well plates were purchased from ICN (Aurora, Ohio). Crystal Screen I and II and Additive Screens I and II were obtained from Hampton Research (Laguna Niguel, CA). Cryo Screens I and II were supplied from Emerald BioStructures (Seattle, WA) as were Wizard Screens I and II. Coverslips were purchased from Fisher Scientific and dimethyldichlorosilane solution employed as the siliconization reagent was obtained from BDH. Silicone lubricant high vacuum grease was obtained from Dow Corning. IZIT ($C_{16}H_{18}ClN_3S$) dye purchased from Hampton Research was used to determine the presence of protein crystals.

2.2.2 General Crystallization Set Up

Prior to crystallization, the concentration of purified protein was adjusted to 10mg/mL in 25mM HEPES and 1mM DTT. Cover slips were prepared and siliconized using standard procedures (Appendix II). Crystallization was carried out using hanging drop-vapor diffusion method in Linbro plates (McPherson, 1982). 0.6mL of reservoir solution was pipetted into each well; 2 μ L of reservoir solution and 2 μ L of protein solution were placed onto a siliconized cover slip and sealed with the silicone lubricant high vacuum grease. Trays were stored at two different temperatures: 4°C and 22°C (Blow et al., 1994). Initial screening employed 290 conditions including Crystal Screens I and II, Cryo I and II screens, and Wizard I and II screens (Jancarik and Kim, 1991).

2.2.3 Optimization Tool I: Fine Screening Technique

Optimization of preliminary crystallization screen (Cudney, 1994) is designed by the trail and error method. A crystallization condition is defined by multiple variables e.g. precipitant concentration (Jovine, 2000), pH of buffer (McPherson, 1995), salt concentration, etc. Any one of these variables can be altered for the fine screen set up. Other variables and gradients employed are listed in Table 2.1. Each fine screen setting will be also examined at both 4°C and 22°C (Blow et al., 1994). Each fine screen manipulates two variables simultaneously (Table 2.2) and utilizes a four by six grid to screen the variables (Figure 2.2). The preliminary/initial condition found from the commercial screen is placed in 2nd row in column D as a starting point for the gradient setting. A more refined screen will be required to examine the neighboring promising conditions.

Table 2.1 Variables and gradients employed in the fine screen setting

<u>Variables</u>	<u>Gradient</u>
1. Precipitant concentration	± 5%
2. pHs of buffers	± 0.5pH
3. Protein concentration	± 1mg/mL
4. Salt concentration	± 0.5M
5. Ratios of protein to precipitant solution in the droplet	arbitrary
6. Amount of well solution	± 0.1mL

Table 2.2 Fine screen set up (Numbers listed in Variables I and II are according to Table 2.1)

Fine Screen	Variable I	Variable II
A	1	2
B	1	3
C	2	4
D	5	6
E	3	4

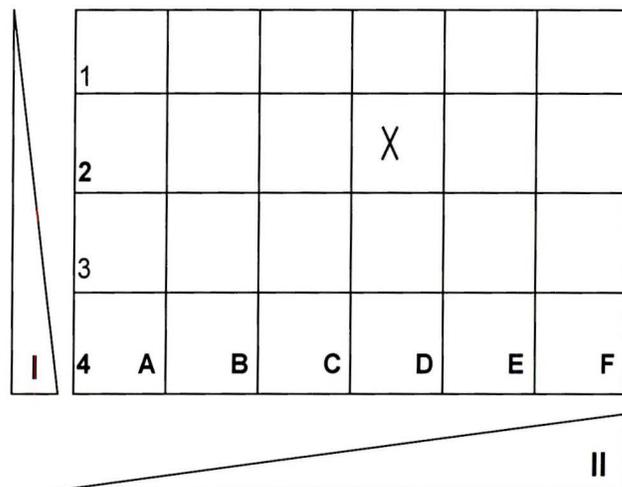


Figure 2.2 Reference point in the system of fine screen setting. This diagram represents the Linbro box containing 24 wells with 4 rows (1-4) and six columns (A-F). The starting position is placed in row 2 and column D contained the preliminary crystallization condition. Variables are examined either from top to bottom or from left to right across the Linbro plate. For example, the pH of the buffer in the preliminary condition is at 6.5. During fine screening, the pH of this buffer being variable I is designed to be varied in the order of 6.0, 6.5, 7.0, and 7.5.

2.2.4 Optimization Tool II: Crystallization using additives

Examination of the effects of small molecules on crystal growth were used as an approach for improving the quality of TarD_{Sa} crystals (Sousa, 1995; Trakhanov, 1995). Additive Screening Kits I, II and III are the optimization tools from Hampton Research⁴. They are designed for improving the preliminary crystallization condition such as the quality and size of the macromolecular crystals. Each kit contains 24 additives including salts and ions which may cause a significant effect on the crystallization. According to the user guide⁵, the effect of small “additives” molecules may have the ability to manipulate the sample/sample and sample-solvent interactions with the macromolecule. Additionally, they may also perturb the solvent in the structure in order to stabilize the protein crystal. Each optimization trial screened all three additive kits, totaling among 72 unique additive solutions at 4°C and 22°C. The instruction for the additive screen is to apply 10% of the total drop volume. For volatile additives, a mixture of 1:9 ratios of additive solution to well solution need to be employed as the proper way of usage.

2.2.5 Optimization Tool III: Crystallization using oils

During crystallization, rapid vapor diffusion may result in excess nucleation of small macromolecular crystals that are not useful for x-ray diffraction analysis. This problem, however, can be improved by applying a layer of oil over the reservoir solution which acts as a barrier between the drop and the reservoir solution such that the vapor

⁴ <http://www.hamptonresearch.com/hrproducts/2420.html>

⁵ <http://www.hamptonresearch.com/techcenter/pdf/2420G.pdf>

diffusion rate is reduced (Chayen, 1997). The viscosity and thickness of the oil layer is also proportional to the rate of vapor diffusion. Therefore, a mixture of silicon oil, the thinner oil, and paraffin oil, the thicker oil, in various ratios and volumes need to be examined for the best result⁶. Various ratios of silicon⁷ to paraffin⁸ oils were assigned for each column of the tray in the order of 4:1, 3:1, 2:1, 1:1, 1:2, and 1:3 from left to right. Volumes of the mixture of silicon and paraffin oils were increased for each row from 0.2mL to 0.35mL in 0.05mL increment.

⁶ <http://www.hamptonresearch.com/techcenter/pdf/3411G.pdf>

⁷ <http://www.hamptonresearch.com/hrproducts/3411.html>

⁸ <http://www.hamptonresearch.com/hrproducts/3413.html>

2.3 Results and Discussion

2.3.1 Crystallization of TarD_{Sa}

Crystallization of TarD_{Sa} in the first trial was performed at a concentration of 10mg/mL with 290 conditions from the initial screens. Some preliminary conditions produced rosette-shape crystals as shown in Figure 2.3. However, those stacking crystals in the figure on the right were unable to be broken into one individual piece for x-ray diffraction studies. Subsequent optimization screens were also set up and examined at both 4°C and 22°C. Unfortunately, no significant result was obtained from any fine-screening procedures. Crystallization of TarD_{Sa} in apo-form was unable to be performed successfully; therefore, co-crystallization of TarD_{Sa} with CTP was pursued.

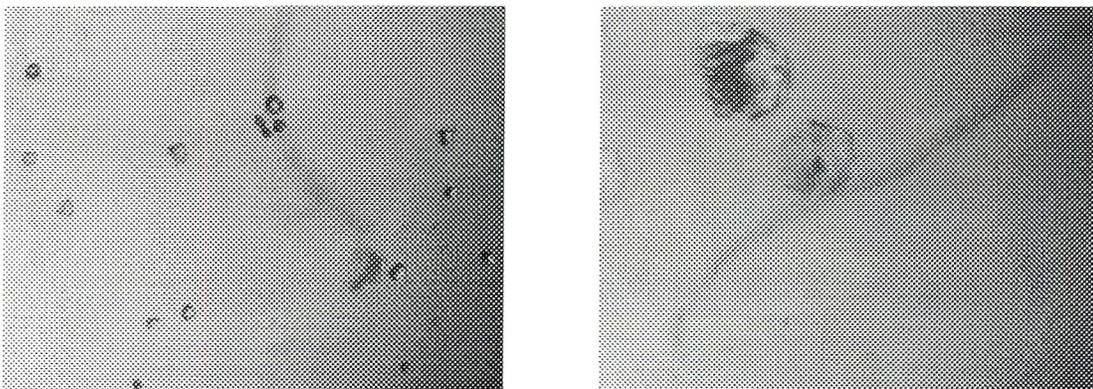


Figure 2.3 Quasi crystals of apo-TarD_{Sa}. Crystals of apo-TarD_{Sa} grown in condition of 1.6M sodium citrate pH6.5 at 22°C (left) and condition of 10% (w/v) PEG-3000, 0.1M imidazole pH8.0, 0.2M LiSO₄ at 4°C (right).

2.3.2 Co-crystallization of TarD_{Sa} with CTP

The purified TarD_{Sa} at 10mg/mL was mixed with 3.3mM CTP and MgCl₂ and stored at 22°C for two days prior to crystallization. Co-crystallization of TarD_{Sa} with CTP started with the use of the commercial screens. Long, thin, needle-like crystalline objects (Figure 2.4a) were initially found in 25% (v/v) ethylene glycol. However, the quality of crystals was not marginally improved with 8% (v/v) ethylene glycol at 22°C in this fine screen (Figure 2.4b). Through screening with small additive molecules, excess nucleation of rod-shape crystals (Figure 2.4c) were obtained from the addition of 0.5μL of 30% (w/v) D(+) glucose to the drop. The IZIT dye diluted 10X with double deionized water was then added to the drop. The IZIT dye test can distinguish protein crystals from salt crystals because biological macromolecule crystals generally have larger solvent channels allowing a small molecule dye to penetrate and stain them blue⁹. These needle-liked crystals with a maximal length of approximately 0.4mm did stain blue with the IZIT dye; however, they were still too small for x-ray diffraction analysis. In order to reduce nucleation and to improve the size of crystals, various ratios of silicon to paraffin oils in different volumes were employed to slow down the rate of vapour diffusion. In about two weeks, the best crystals were found in the well containing 250μL silicon and paraffin oils over the reservoir solution at 22°C in a ratio of 2:1. As shown in figure 2.4d, several protein crystals had approximate dimensions of 0.2 x 0.02 x 0.02 mm. Crystals were fished out from the droplet using Hampton Research cryoloops and dipped into a cryoprotectant agent before freezing. The cryoprotectant contained 25% (v/v) ethylene

⁹ <http://www.hamptonresearch.com/hrproducts/4710.html>

glycol saturated with D(+)-glucose. The cryoprotectant employed is very important for flash cooling of the crystals, because it serves as a stabilization agent which prevents ice-crystal formation when the crystal is frozen. Crystals were frozen in a suitable cryoprotectant by liquid nitrogen and stored at $\sim -180^{\circ}\text{C}$. Crystals diffracted to 3.0\AA at the National Synchrotron Light Source (NSLS). The same crystallization condition at 4°C did not produce any crystals, indicating the importance of temperature for crystallization. In conclusion, determination of the crystallization condition for TarD_{sa}-CTP involved three optimization procedures. Each of the steps displayed a different degree of improvement on the crystal quality and size, which ultimately resulted in crystals that diffracted to 3.0\AA resolution (Yim et al., 2001).

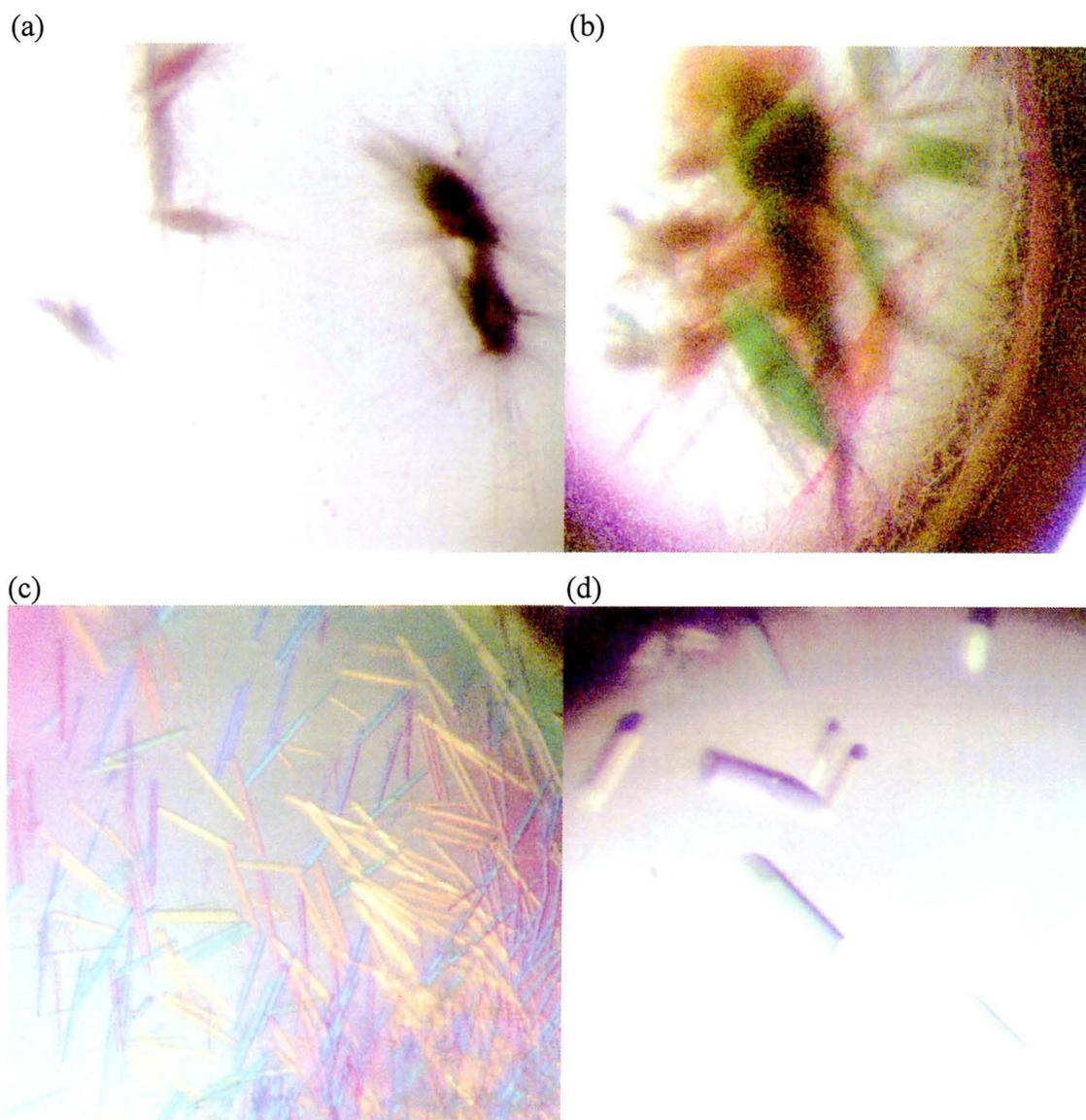


Figure 2.4 Crystals of TarDSa-CTP grown with a combination of three optimization strategies

- (a) Needle-like crystalline material with a maximal length of approximately 0.4mm was initially found to grow in 25% (v/v) ethylene glycol.
- (b) Crystals examined in condition of 8% (v/v) ethylene glycol did not show marginal improvement.
- (c) Shower of small crystals were produced by adding 0.5 μ l of D(+) glucose to the drop
- (d) The rate of diffusion was slowed down by employing a combination of silicon and paraffin oils over the well solution. Crystals (0.2 x 0.02 x 0.02 mm) grown for the diffraction study were equilibrated in conditions of 8% (v/v) ethylene glycol, D(+) glucose, and the mixture of silicon and paraffin oils in the ratio of 2:1.

Because the CTP binding sites were not filled in the tetramer structure of TarD_{sa} (see Chapter 3), co-crystallization of TarD_{sa} with higher concentrations of CTP were performed at both 22°C and 4°C using the commercial screens. Unfortunately, no crystals had been found yet. The original condition which previously produced TarD_{sa} with CTP crystals, did not produce any crystals even if CTP's concentration was increased. This indicates that CTP concentration has a significant influence on crystal growth.

Although the co-crystallization technique cannot be successfully employed, soaking CTP into the pre-formed crystals is another approach to saturate the substrate binding sites in the crystals of TarD_{sa}. Many macromolecular crystals contain high solvent content in the crystal lattice which can be as high as 95% of the total composition (Matthews, 1968; Cohen et al., 1971). Some solvent molecules will form interactions with the surface of protein. However, some solvent molecules may be loosely linked with the protein and be replaced by small molecules diffusing through the solvent channels of the crystal. Moreover, there are examples of macromolecules that maintain their activity in the crystalline state, for example, glycogen phosphorylase (Johnson et al., 1983). The only disadvantage of the soaking technique is that there is a high possibility of cracking the crystals due to the disruption of the crystal contacts and intermolecular interactions of the macromolecule. Figure 2.5 shows the soaking procedure (Petock et al., 2001). Using the Hampton Research cryoloop, crystals were transferred from the coverslip to the soaking condition which contained 100µL of the reservoir solution plus 15mM CTP. Crystals were soaked in the new solution for two days, then they were transferred to

another soaking condition containing a higher concentration of CTP (30mM CTP), and soaked for four additional days. Crystals were subsequently cryoprotected by transferring them from 8% to 25% (v/v) ethylene glycol saturated with D(+) glucose. However, crystals diffracted only to 5.0Å at NSLS. This might be due to the occurrence of crystal damage during the transferring and soaking procedures. The crystal lattice could be destroyed even when deterioration of crystals is unable to be observed.

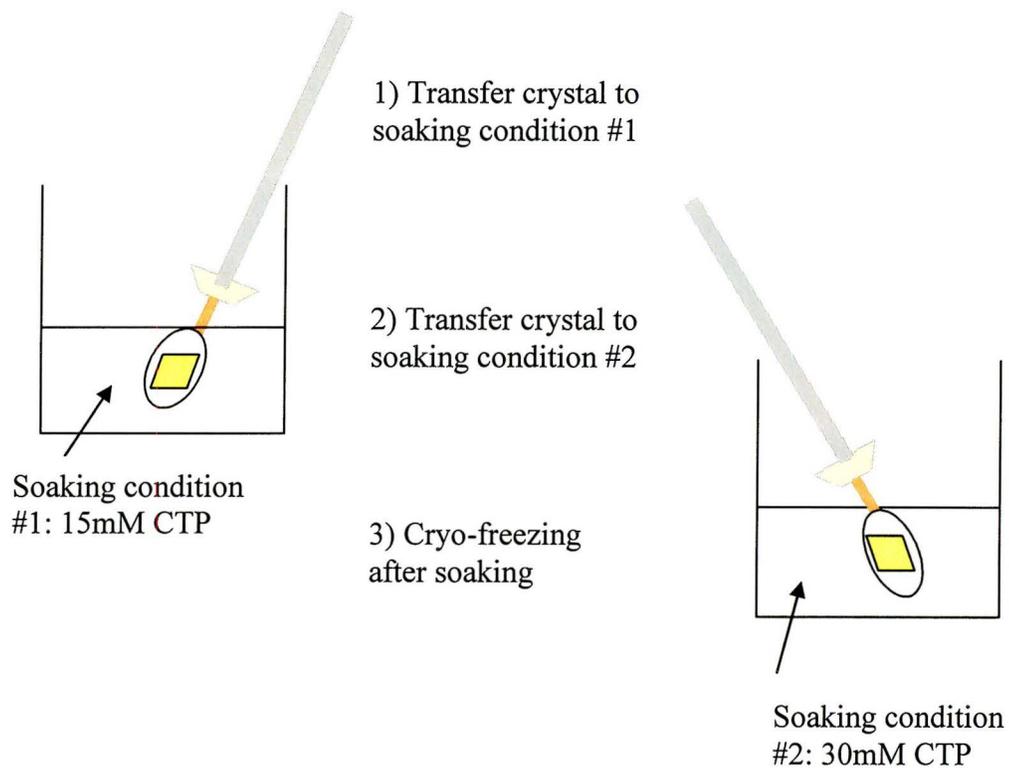
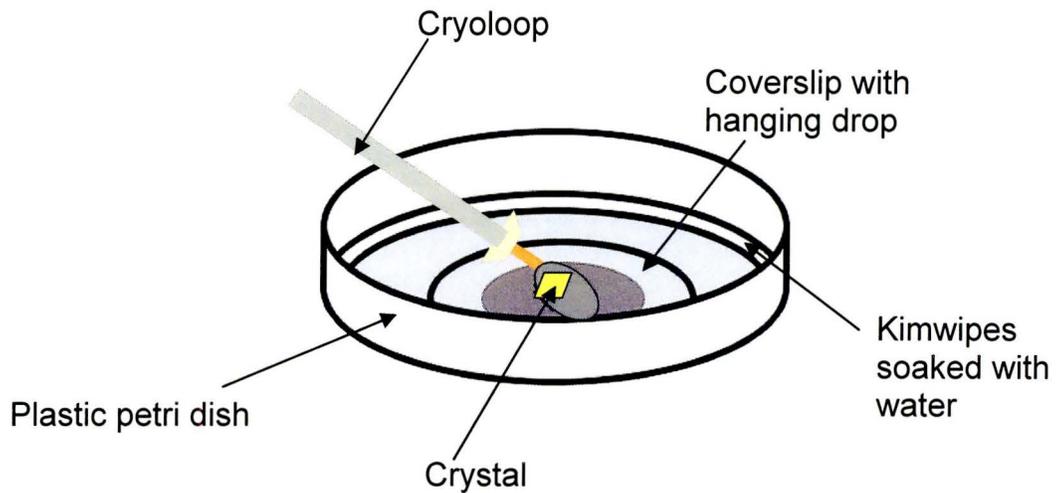


Figure 2.5 Soaking procedures. Soaking procedures were prepared by making two solutions containing 15mM and 30mM CTP. The crystals inside the hanging drop were placed onto the cover slip in the centre of Petri dish. The plastic petri dish was encircled with wet Kimwipes. This way, the rate of evaporation of the drop was reduced. The grown crystals were transferred and soaked in the first condition containing 15mM CTP in 100 μ L for two days, then soaked in the second condition (30mM CTP) for four additional days. The crystals after soaking were cryo-frozen and tested in NSLS.

2.3.3 Co-crystallization of TarD_{Sa} with Glycerol-3-Phosphate

As in the previous set up, TarD_{Sa} was mixed with 3.3mM of glycerol-3-phosphate (GroP) and MgCl₂ before co-crystallization was carried out. Crystals were firstly observed to grow from 12% (W/V) PEG-20000 and 0.1M MES pH6.5. Fine-screening of this condition and different kinds of optimization techniques such as the application of mixtures of silicon and paraffin oils and additive screens did not significantly improve the size and quality of the crystals. Although larger crystals were obtained at a higher protein concentration (14mg/mL), they were very fragile and easily dissolved in cryoprotectant in Figure 2.6 on the left. The IZIT dye was used to show that these crystals contain protein (Figure 2.6 right). However, they only diffracted very weakly at NSLS. For flash cooling of the protein crystals, 18% of (W/V) PEG-20000 and 0.1M MES pH6.5 was employed as the cryoprotectant reagent.



Figure 2.6 Crystals of TarD_{Sa}-CTP obtained from the first crystallization condition and the IZIT test result. Crystals of TarD_{Sa}-CTP were grown in conditions of 12% (W/V) PEG-20000 and 0.1M MES pH6.5 (left). IZIT dye was able to stain these crystals blue, indicating they contained protein (right).

Rosette shape crystals obtained from the second crystallization conditions are shown in Figure 2.7 (left). The refined conditions, 20% (v/v) (V/V) PEG-600, 0.1M cacodylate at pH5.5, 0.2M Ca(OAc)₂, were found to yield a stack of plate crystals. The

Izit test also showed a positive result on these crystals (right). Unfortunately, these crystals could not be readily reproduced. Stained crystals were still taken to NSLS, but they diffracted to only 5.0Å.

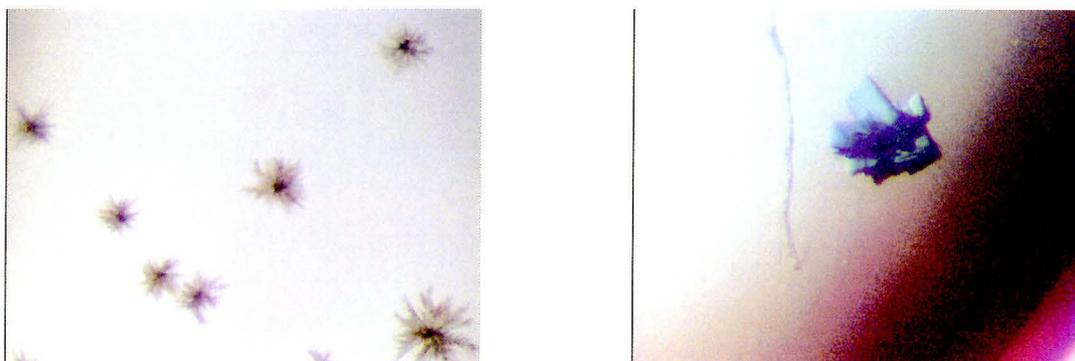


Figure 2.7 Crystals of TarD_{Sa}-CTP obtained from the second crystallization condition and the Izit test result of the fine screen. Crystals grown in 40% (v/v) (V/V) PEG-600, 0.1M cacodylate pH6.5, and 0.2M Ca(OAc)₂ (left). Optimal crystals were grown in 20% (v/v) (V/V) PEG-600, 0.1M cacodylate pH5.5, and 0.2M Ca(OAc)₂ and confirmed as protein by the Izit test (right).

After extensive repeated trials of the fine-screen condition described above, several crystals were obtained in one single well at 4°C (Figure 2.8a). Although the crystallization condition remained the same, the crystals were different from the initial ones and had the same crystal form as those of TarD_{Sa} with CTP. Crystals were brought to NSLS; however, the resolution limit observed was only 5.0Å. Optimization procedures were utilized to improve the reproducibility of the crystals grown in conditions containing 20%(w/v) (V/V) PEG-600, 0.1M cacodylate pH5.5 and 0.2M Ca(OAc)₂. Subsequently, showers of small crystals (0.1 x 0.1 x 0.03 mm) were easily reproduced through the addition of 0.5µL of 1.0M ammonium sulfate to the drop (Figure 2.8b). The oil technique was used to improve the crystal size. In about one week, the dimensions of the crystals increased to 0.2 x 0.1 x 0.05 mm (Figure 2.8c) with the

application of 250 μ l mixture of silicon and paraffin oil in 2:1 ratio over the well solution. A combination of the additive and oil techniques showed the best result for the improvement of crystal size (0.4 x 0.2 x 0.05 mm), however, these crystals did not stain blue with the Izit dye (Figure 2.9d). The Izit test is not always definitive in the determination of protein crystals. The Izit crystal dye did not stain macromolecular crystals in about 1-3 % of reported use¹⁰. These non-staining crystals were then collected and dissolved in 10 μ L deionized water, boiled in the SDS-PAGE running buffer for approximate 5-10 minutes, and run on the SDS-PAGE gel. On the silver-stained gel, a thick single band of 16KDa was present (Figure 2.9), thereby confirming that these crystals were indeed protein crystals. The failure of this staining process on these crystals may result from an unusual crystal packing under a specific crystallization condition. Crystals were taken to NSLS, but they diffracted poorly to \sim 5.0 \AA .

¹⁰ <http://www.hamptonresearch.com/techcenter/pdf/4710G.pdf>

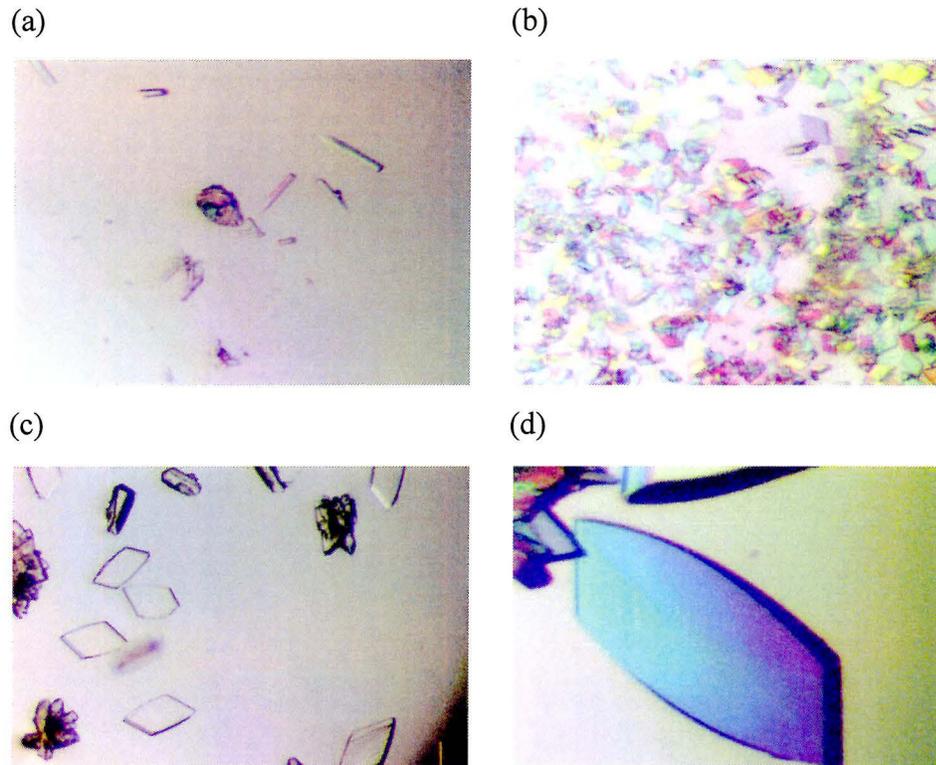


Figure 2.8 Crystals of TarD_{Sa}-CTP obtained from the repetitive fine screen of the second crystallization condition and the optimal additive and oils conditions.

- (a) Crystals (0.06 x 0.02 x 0.02 mm) were obtained from the repetitive fine screen of 20%(v/v) PEG- 600, 0.1M cacodylate pH5.5, and 0.2M Ca(OAc)₂, but resulted in a different morphology.
- (b) By adding 0.5 μ l ammonium sulfate to the drop containing the protein (TarD_{Sa}-CTP) and reservoir solutions of 20%(v/v) PEG- 600, 0.1M cacodylate pH5.5, and 0.2M Ca(OAc)₂, excess nucleation of crystals (0.1 x 0.1 x 0.03 mm) occurred.
- (c) The crystal size was optimized to the dimensions of 0.2 x 0.1 x 0.05 mm using the oils technique with a 2:1 ratio of silicon to paraffin oil.
- (d) Izit test result of the crystals (0.4 x 0.2 x 0.05 mm) grown in both additive and oils conditions.

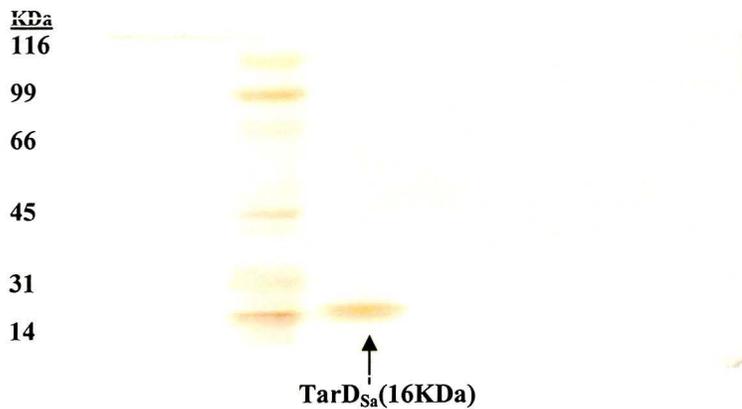


Figure 2.9 Examination of the composition of the non-stained crystals obtained from the fine additive and oils conditions. The dissolved, non-stained crystals of TarD_{Sa}-CTP with molecular weight of 16KDa were shown on the 15% SDS-PAGE silver-stained gel.

CHAPTER 3

RESULTS AND DISCUSSION

3.1 Data Collection and Processing

Crystals of TarD_{Sa} grown in 8% (v/v) ethylene glycol, 30% (w/v) D(+)-glucose, and 250 μ l of silicon and paraffin oils were first mounted on the home x-ray source (Rigaku RU200 rotating anode, equipped with osmic double focusing mirrors and an RaxisIIC area detector). The crystal diffracted very poorly and the resolution obtained was only to 4.0 \AA . These crystals were subsequently taken to the National Synchrotron Light Source and diffracted to 3.0 \AA at the beam line X8C using an ADSC Quantum-4 CCD. The data set was collected with an oscillation range of 1.5° and a crystal to detector distance at 280mm using a wavelength of 1.072 \AA . A second set of data was collected using the same method described above. However, this data set was only 50% completed, due to the occurrence of crystal damage after a long period of x-ray exposure.

The crystallographic data of TarD_{Sa} was processed using the programs Denzo and Scalepack (Otwinowski and Minor, 1997). Primitive hexagonal (Table 3.1) was found to be the cell lattice type that could fit the reflection peaks. Analysis of symmetry-related reflections and systematic absences indicated that the candidate for the space group of crystal TarD_{Sa} was either P₃₁21 or its enantiomorph, P₃₂21. The correct space group was subsequently established by the translation search in the process of the molecular

replacement method (see Section 3.2). The unit-cell dimensions of the crystals are $a=b=92.15\text{\AA}$, $c=156.12\text{\AA}$, $\alpha=\beta=90^\circ$, and $\gamma=120^\circ$. Table 3.2 lists the significant statistics for TarD_{Sa} in data processing.

The number of molecules per unit cell and volume of solvent in TarD_{Sa} were estimated by Matthews' average solvent content analysis (Drenth, 1994). In brief, the ratio of the unit cell volume and the molecular weight of the protein within the unit cell, which is called the Matthews volume (V_M), is generally found between 1.7 and $3.5\text{\AA}^3/\text{Da}$ (Equation 3.1). Due to restrictions imposed by the space group, the number of molecules in the unit cell for TarD_{Sa} is most likely 24, giving a V_M of $2.98\text{\AA}^3/\text{Da}$. Since there are six symmetry operators in space groups P3₁21 and P3₂21, TarD_{Sa} crystals may contain four molecules per asymmetric unit. Based on the equations $V_{\text{protein}}=1.23/V_M$ and $V_{\text{solvent}}=1-1.23/V_M$, they suggest that there is about 58% of the crystal volume occupied by solvent.

$$V_M = \frac{\text{Volume of the unit cell}}{\text{Number of molecules per unit cell} * \text{Molecular weight of the protein}} \quad \text{Equation 3.1}$$

Table 3.1 Space groups defined in the primitive hexagonal family

P3(n)
P3(n)12
P3(n)21
P6(n)
P6(n)22

Table 3.2 Data processing statistics for the diffraction studies of crystals TarD_{Sa}
(The following data were taken from the final scalepack output)

Highest resolution (Å):	3.0
Space group:	P3 ₁ 21 / P3 ₂ 21
Unit cell dimensions:	a=b=92.15, c=156.12, $\alpha=\beta=90^\circ$, $\gamma=120^\circ$
Molecules per unit cell ($V_M = 2.98$):	24
Molecules per asymmetric unit ($V_M = 2.98$):	4
Solvent content ($V_M = 2.98$):	59%
Number of observations:	106644
Number of unique reflections:	15217
Completeness (%):	95.2
Rsym (%):	13.5
Redundancy:	7.5

3.2 Phase Determination

The published structure of *B. subtilis* TagD (Weber et al., 1999) was chosen to be the search model for the phase determination of the present crystal structure (see section 1.3). Prior to building a search model, a sequence alignment of the TagD_{Bs} protein against TarD_{Sa} was performed and found to have 68% sequence homology, using the pairwise BLAST sequence alignment tools provided by the National Center for Biotechnology Information (see Section 1.3).

43 out of 128 residues in TagD_{Bs} are different from those in the amino acid sequence of TarD_{Sa}, these were altered to Ala. Because TagD_{Bs} is a dimer, we speculated that TarD_{Sa} crystals contained two dimers in the asymmetric unit. The first step of molecular replacement was to apply a fast direct rotation search in space group P3₁21, using the mutated TagD_{Bs} dimer as the search model. The resolution range was chosen from 15.0Å to 4.0Å. The initial coarse grid search was sampled on a 6° interval. This angular grid size should be small enough to ensure the detection of the solution (Brünger, 1997).

After the initial coarse grid search, a maximum of 20 peaks were chosen to be examined in subsequent fine grid searches to obtain a precise orientation (angular grid size = 2.0°). Four solutions were found in the cross rotation list (Table 3.3). They were determined to have a peak height of ~5.5σ above the mean value, and this suggested that the chosen peaks had good signal-to-noise ratios. A clear separation was also observed

between the top peaks (four) and subsequent answers, indicating the precision of the solutions. Among these four solutions, two are related to the other two by the inversion centre.

Table 3.3 List of the direct cross rotation results

Resolution: 15 – 4.0 Å

Angular grid: 2.0°

Mean of rotation function: 0.0279

Standard deviation around mean: 0.0119

Theta1,	Theta2,	Theta3,	$C(\Omega)_{\text{Direct}}$
247.521	75.123	149.000	0.0514
251.997	81.552	217.094	0.0485
112.479	75.123	91.000	0.0484
109.518	81.552	24.421	0.0471
0.042	-2.020	6.102	0.0377
145.856	-2.020	151.915	0.0363
114.720	2.020	120.780	0.0350

↓
20th peaks

The top solutions identified in the rotation search were subjected to a translation search. Before and after the translation search, a PC refinement of all the selected peaks was performed to generate a more accurate model to increase the success of molecular replacement. The first round of translation search was performed in space group $P3_121$ and utilized TagD_{Bs} dimer as the search model. Only the top four solutions are shown in Table 3.4 as the best answers for the first translation search. An analogous translation search was performed in space group $P3_221$, showed relatively lower translation correlation coefficients ($C(\Omega)_{\text{Trans}}$) (Table 3.5). It was apparent that $P3_121$ was the correct space group for the TarD_{Sa} crystal. Note that $C(\Omega)_{\text{Trans}}$ represents the agreement between

the structural factors of the observed and search models. Therefore, the higher the value $C(\Omega)_{\text{Trans}}$ is, the smaller the discrepancy between the calculated and actual model will be.

Based on the Matthew average solvent content (Matthew, 1968) analysis, there are probably four molecules present in the asymmetric unit. However, the packing of the solution calculated from the previous rotation and translation search represents only half of the contents of the asymmetric unit. Since the location of the first dimer is already known, its location will remain fixed while the location of second dimer is tested for each rotation peak. To do so, the orientation and position of the top ranking solution in Table 3.4 was treated as the first dimer group in the asymmetric unit (the orange dimer in Figure 3.1). After searching the optimal orientation and position for this selected dimer using PC refinement, a second translation search was started to locate the other dimer group. As a result, an asymmetric tetramer was created in space group $P3_121$ (Figure 3.2). The result (Table 3.6) was determined to comprise 40.45% protein content associated with the best $C(\Omega)_{\text{Trans}}$: 0.531. The second solution in the translation search was shown to embrace 24.14% protein packing instead of 40.45% given in the first solution. Because the location of the second translation solution, which was sampled on the third rotation peak, was equivalent to the fixed dimer, there would be no additional protein content adding to the asymmetric unit in the result (Figure 3.3). The solution sampled on the fourth rotation peak yielded a similar $C(\Omega)_{\text{Trans}}$ value to the first translation solution (Figure 3.4). Since the first solution displayed the highest $C(\Omega)_{\text{Trans}}$ value, it was employed during structure modeling and refinement.

Table 3.4 List of solutions obtained from the translation search in space group P3₁21 using the TagD_{Bs} dimer as the search model

Solution	thetal1	thetal2	thetal3	transX	transY	transZ	C(Ω) _{Trans}	packing
1	247.78	75.78	148.46	-15.06	78.81	-103.39	0.273	0.2144
2	251.79	81.72	216.30	52.03	56.60	13.31	0.289	0.2139
3	112.08	76.08	91.61	41.56	63.23	23.45	0.273	0.2146
4	108.12	82.16	23.76	-36.66	60.60	-96.51	0.286	0.2139
↓ 20								

Table 3.5 List of the solutions obtained from the translation search in space group P3₂21 using the TagD_{Bs} dimers as the search model

Solution	thetal1	thetal2	thetal3	transX	transY	transZ	C(Ω) _{Trans}	packing
1	250.30	76.38	148.28	-15.65	71.62	8.29	0.154	0.2148
2	250.18	80.39	217.31	40.92	29.90	-89.10	0.141	0.1647
3	109.55	76.69	91.71	35.30	59.64	-87.17	0.154	0.2145
4	107.96	81.27	23.37	19.13	-0.67	-96.04	0.152	0.2132
↓ 20								

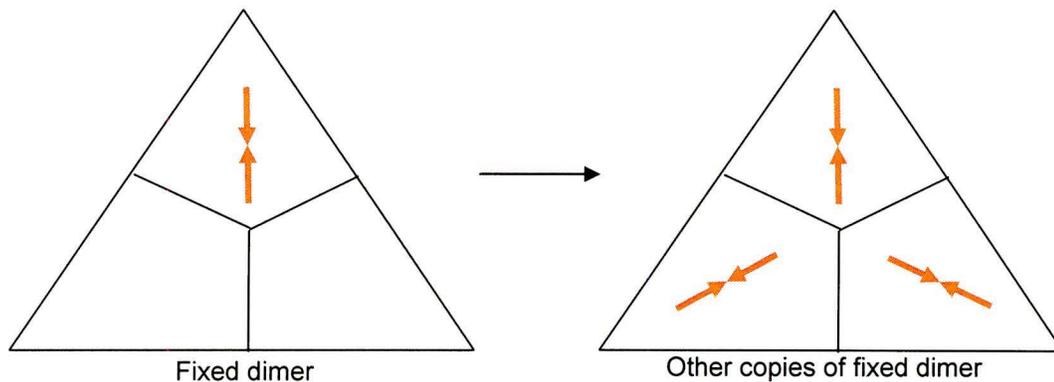


Figure 3.1 Proposed scheme used to explain how the first answer calculated from previous rotation and translation is fixed in the second translation search. Orientation and position of the first solution is fixed in the second translation search. Other copies of solution (1) are presumably present in the model (right).

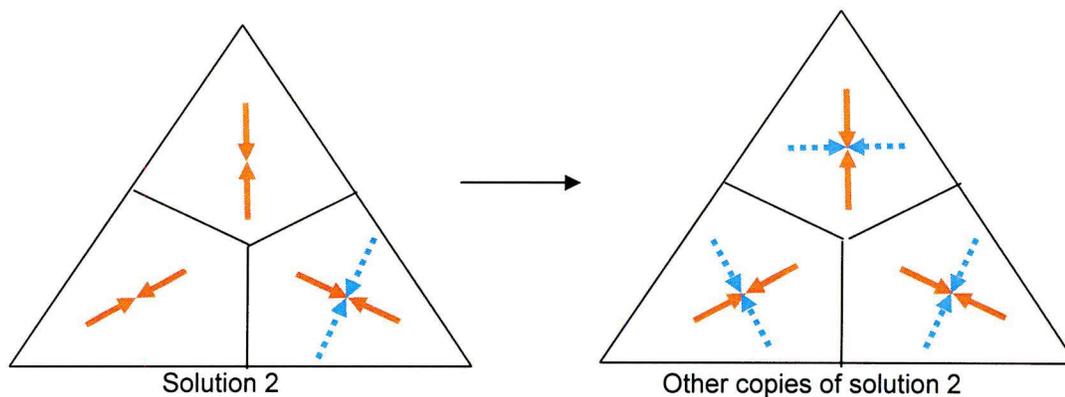


Figure 3.2 Proposed scheme used to explain the result of the translation search with the second rotation solution. When the proper position of solution (2) is located, other copies of molecules are subsequently filled, as shown in this figure. An asymmetric tetramer is then established, giving the highest correlation coefficient in the result, and solvent contents of ~ 60% as expected for protein crystals.

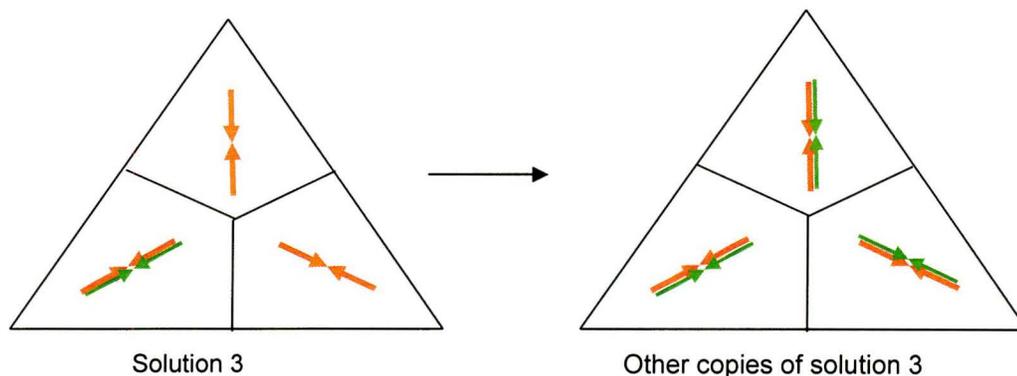


Figure 3.3 Proposed scheme used to explain the result of the translation search with the third rotation solution. Solution (3) of the translation search is symmetry related to the fixed dimer.

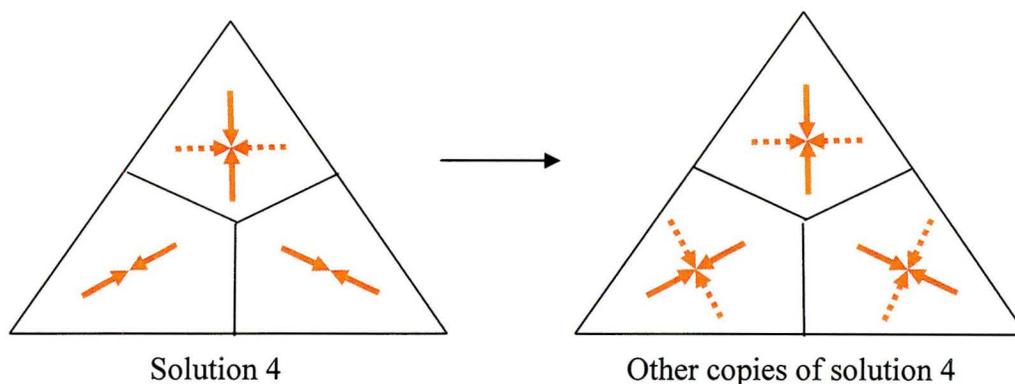


Figure 3.4 Proposed scheme used to explain the result of the translation search with the fourth rotation solution. Solution (4) of the translation search is symmetry related to solution (2).

Table 3.6 List of solutions obtained from the translation search of an asymmetric tetramer in space group $P3_121$ using the TagD_{Bs} dimer as the search model

Solution	thetal1	thetal2	thetal3	transX	transY	transZ	$C(\Omega)_{\text{Trans}}$	Packing
1	252.32	81.88	215.99	52.21	56.30	13.53	0.531	0.4045
2	112.70	76.06	91.02	40.88	63.62	-54.69	0.297	0.2414
3	107.62	82.23	24.12	-36.73	60.70	-18.42	0.529	0.4047

3.3 Model Rebuilding and Refinement

The phases determined from molecular replacement were used as the initial estimate of phases for building the first model of TarD_{Sa}. After calculating the $2F_o-F_c$ and F_o-F_c electron density maps, the model was interpreted by manual manipulations on a silicon graphics computer, with the program O (version 6.2) (Jones et al., 1991). The majority of the model displayed reliable agreement with the observed map, except for several residues, Asp11, Leu12, Leu13, His14, and Tyr15 in subunits I and III, which did not fit the electron density. The coordinates of these residues were then adjusted using the program O (Jones et al., 1991). Unfortunately, model fitting was unable to be performed after residue 112 in every subunit because the electron density is missing around the C-terminal regions.

During the course of model building, there was electron density showing a tendency to contain CTP substrates in subunits I and IV of TarD_{Sa}. However, the density in the subunit IV was found to be very weak, and thus we were justified in including CTP in the model of TarD_{Sa}. Although the CTP in subunit I did fit into the density (Figure 3.5), it was inaccurate, since the refinement of this CTP in the density was very poorly behaved. The refinement resulted in B-factors over 100, indicating that CTP was not present or only partly present. This CTP conformation was different from that seen in the crystal structure of TagD_{Bs} and this will be discussed in Section 3.4. Although there was CTP in the crystallization setup (see Section 2.3.2), which was also mandatory to obtain

the crystals of TarD_{Sa}, the poor diffraction data at 3.0Å could not confirm this speculative CTP conformation and also its difference from TagD_{Bs}'s CTP.

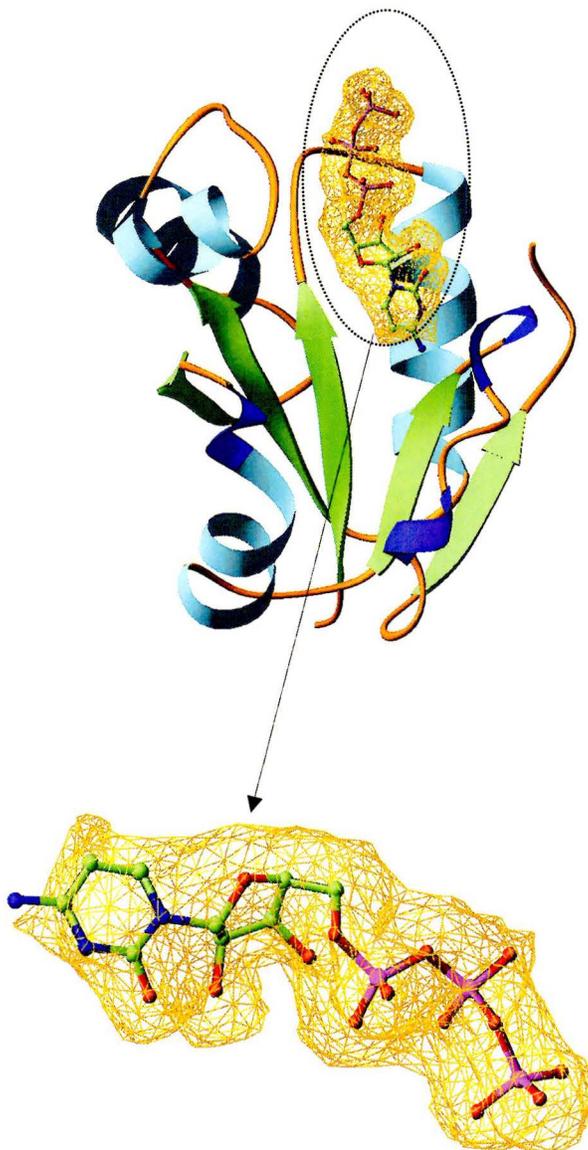


Figure 3.5 CTP in a Fo-Fc sa omit map at 1 sigma contour. Subunit I of TarD_{Sa} is shown in ribbon. CTP is shown in ball-and-stick. This figure is generated using the program “Ribbons” (Carson, 1997).

The initial calculated structure factors are usually revealed to have poor agreement with the observed structure factors. The purpose of the refinement process is to improve the discrepancy between the calculated and observed structure factors. Rigid body refinement, the same as the PC refinement discussed (see also Appendix in section AI.3 and AI.4), was first applied to refine the position and orientation of the model. Subsequent refinement was done using CNS (Brünger, 1992), and carried out by the method of maximum likelihood (Pannu & Read, 1996). Because of the limited resolution of TarD_{Sa}, which is to only 3.0Å, the geometric parameters such as bond lengths, bond angles, torsion angles, and van der Waals contacts, were treated as restrained. Simulated annealing refinement was also employed for searching the invisible C-terminal regions, but it still failed to identify convincing positions for residues beyond Lys112.

The refinement cycles were interspersed with manual interventions in which residues originally modeled as alanines (sequence differences between TagD_{Bs} and TarD_{Sa} were modeled as alanines for search model purposes) were replaced with appropriate residues. The course of model rebuilding and refinement was monitored by computing the $R_{\text{conventional}}$ and R_{free} values (see also Appendix I in section AI.5). The starting $R_{\text{conventional}}$ and R_{free} values were 0.39 and 0.50. After cycles of refinement, the final structure was refined to $R_{\text{conventional}}$: 0.24 and R_{free} : 0.27.

To reduce the risk of over-fitting the diffraction data, the refinement process was terminated when the R_{free} value started rising. The Ramachandran plot of the TarD_{Sa}

model (Figure 3.6) was computed using program PROCHECK (Laskowski R.A., 1993). CNS (Brünger, 1992) was also employed to validate the proper stereochemistry and torsion angles of the protein groups for the model. Table 3.7 lists the statistics of the final refined crystal structure of TarD_{Sa} at 3.0Å.

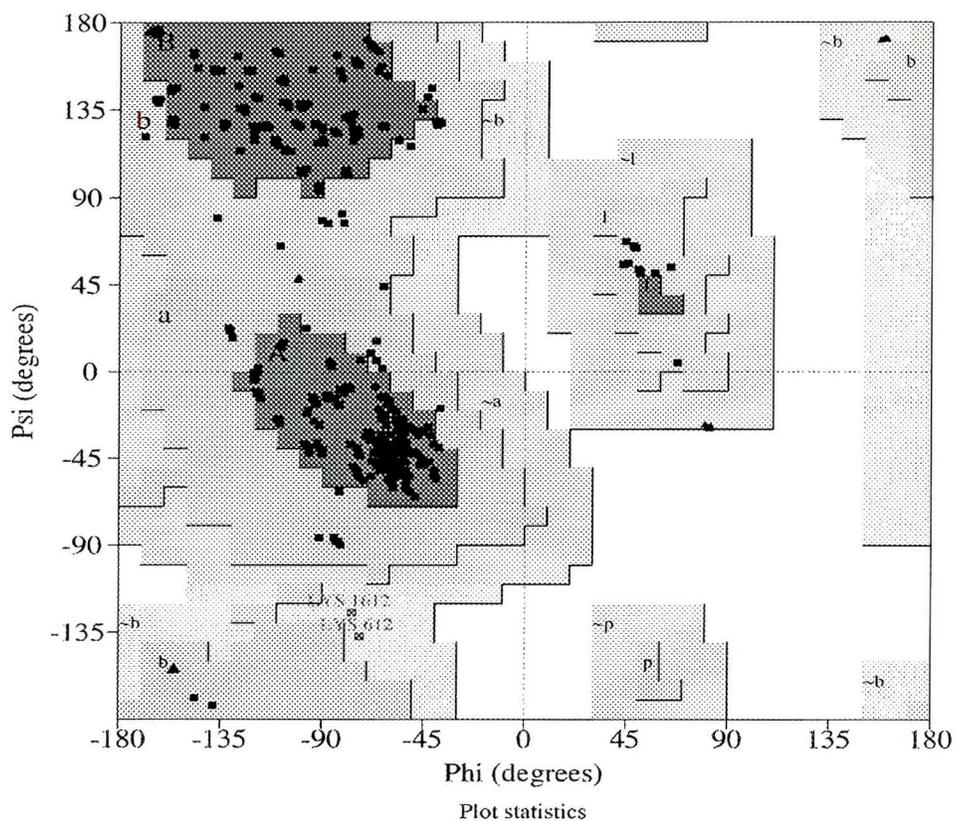


Figure 3.6 Ramachandran plot TarD_{Sa}. Ramachandran plot of TarD_{Sa} showing the sterically reasonable values of the angles ϕ and ψ . The shaded areas indicate particularly allowed regions of these angles. Darker shaded regions represent the more favorable values of these angles. Triangles represent the actual angles ϕ and ψ measured for Gly residues. All other residues are represented by squares.

Table 3.7 Statistics of the final refined crystal structure of TarD_{Sa} at 3.0 Å

Refinement Statistics	
Number of reflections used for R factor calculations	14712
Resolution Range (Å)	50-3
% of observation on resolution shell ¹	
3.38Å	98.9
3.23Å	83.2
3.00Å	70.2
R _{free} (%) ²	0.27
R _{conventional} (%) ³	0.24
Stereochemistry Statistics	
Number of glycine residues in protein	28
Number of proline residues in protein	4
Number of residues in protein	448
Number of atoms in protein	3881
Number of atoms with occupancy set to 0 ⁴	38
Residues in Ramachandran plot (%) ⁵	
Most favorable	87.0
Additionally allowed	12.6
Generously allowed	0.5
Disallowed	0
Deviation from Ideal Values	
Rmsd in bond lengths (Å)	0.008
Rmsd in bond angles (°)	1.2
Rmsd in dihedral angles (°)	22.2
Rmsd in improper angles (°)	0.7
Coordinate Errors⁶	
Sigma coordinate error	0.52(working set) 0.57(test set)
Thermal Factor⁷	
Overall B mean value (Å ²)	23.2
Overall anisotropic B value ⁸	
B11(Å ²)	-7.9
B22(Å ²)	-13.2
B23(Å ²)	0
B12(Å ²)	-13.2
B13(Å ²)	0
B33(Å ²)	15.8

¹The completeness (%) were reported from the final output of scalepack

^{2,3} See section 3.2 for definitions of R_{conventional} and R_{free}

⁴The occupancy of these atoms was set to zero due to the absence of electron density in these regions

⁵Values were analyzed using PROCHECK (Laskowski, 1993)

⁶These values represent the estimated coordinate errors calculated by the Sigma A method

⁷An overall thermal factor is used to describe the extent of movement generated for the whole model during the temperature change

⁸The anisotropic B-values describe the directional dependency of the overall average B-factor

3.4 Structural Discussion of TarD_{Sa}

3.4.1 Overall Structure and Subunit Folding

The crystal structure of TarD_{Sa} is depicted in Figure 3.7. The crystal complex of TarD_{Sa} packs as a tetramer; however, there is no four-fold symmetry axis found in this tetramer, suggesting that it is a dimer of dimers. As shown in Figure 3.8, the structure of each subunit includes residues from 1 to 112, and consists of four α -helices labeled from A to D, and five parallel β -strands numbered from I to V. Two short stretches of α -helices connecting strands IV and V of TarD_{Sa} are 3_{10} segments analogous to the crystal structure of TagD_{Bs} (Weber et al., 1999).

Like TagD_{Bs}, each subunit of TarD_{Sa} contains a Rossmann fold. The Rossmann fold consists of a five-stranded parallel β -sheet sandwiched between α -helices. Structural comparison of all the C α atoms in each subunit of TarD_{Sa} and TagD_{Bs} has been shown to yield a root mean square deviation of $\sim 0.6\text{\AA}$, and the overlay figure is shown in Figure 3.9. Although class I aminoacyl tRNA synthetase, PS, and PPAT share relatively lower sequence homology with TarD_{Sa} (see examples in section 1.3), they are also observed to contain a Rossmann fold and similar modes of substrate binding in their structures.

The C-terminal regions after residue 112 in every subunit of TarD_{Sa} are missing in the crystal structure (see also Section 3.3). However, both of the C-termini are visible in the dimer of TagD_{Bs}, and even shown to have binding interactions with the substrates

CTP. This observation suggests that TarD_{Sa} may be crystallized in a conformation in which the C-termini of the four subunits are disordered in the absence of CTP, since it lacks a series of interactions which are probably used to anchor this missing region. Thus, the uninterpretable density of the four C-termini in TarD_{Sa} is perhaps due to the low concentration of CTP used during crystallization (see also Section 2.3.2).

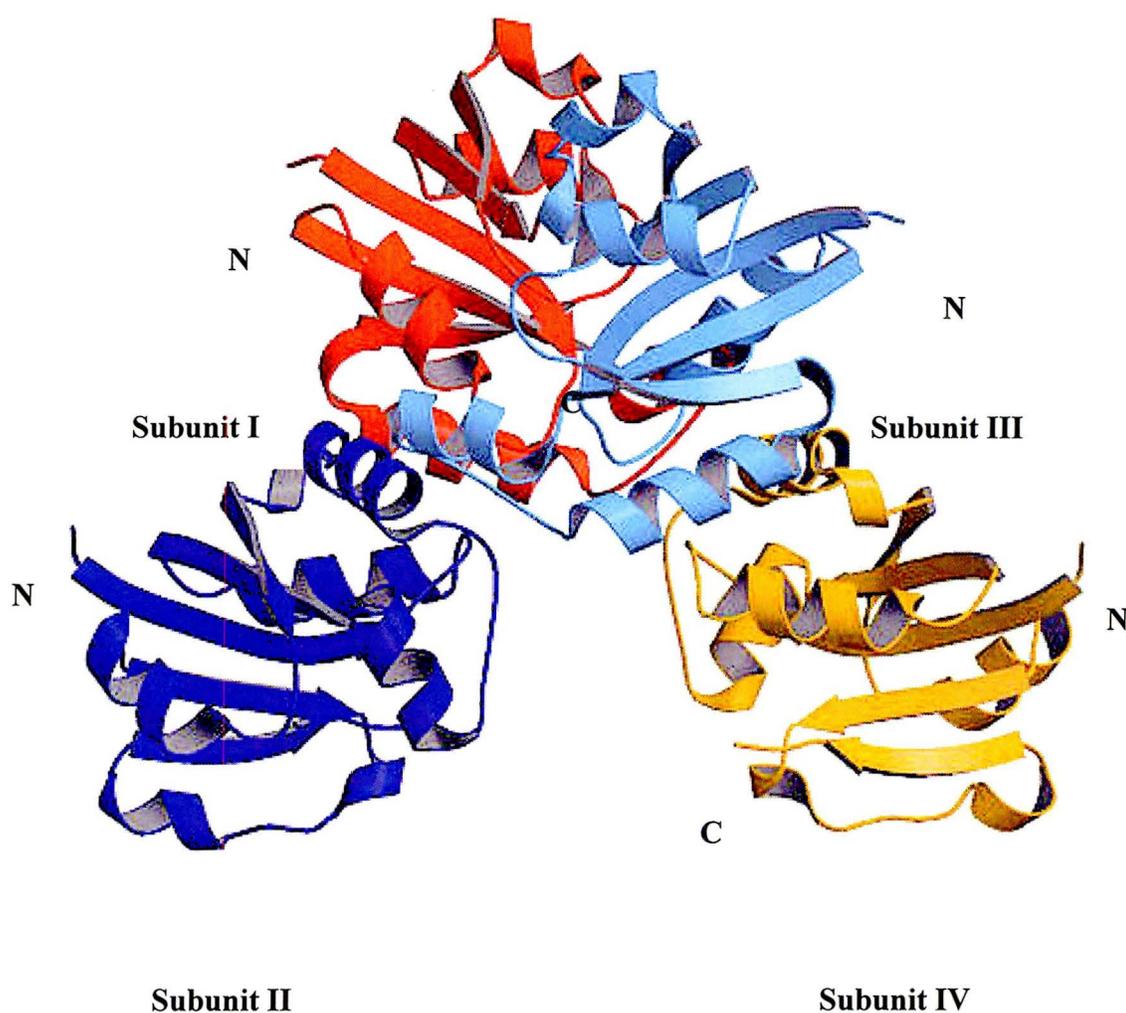


Figure 3.7 Overall Structure of TarD_{Sa}. Crystal structure of TarD_{Sa} diffracted at 3.0Å is shown to contain four molecules in the asymmetric unit. The N and C terminus are labelled, and the four subunits are in four different colours: red, I; blue, II; sky blue, III; gold, IV. 112 out of 132 residues in each subunit of TarD_{Sa} are defined in this structure. The remaining residues in the C-terminus of each subunit are missing in the electron density map.

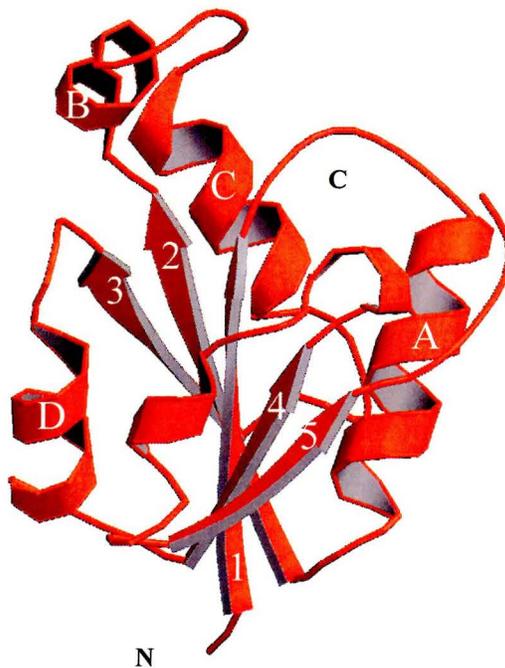


Figure 3.8 Monomer folding of TarD_{Sa}. A Rossmann fold is present in each subunit of TarD_{Sa}. As described in the crystal structure of TagDBs, the construction of each subunit structure of TarD_{Sa} is also based on the Rossmann fold pattern. As seen in this figure, five strands of parallel β sheets are packed inside and wound by four α -helices.

TagD_{Bs}

TarD_{Sa}

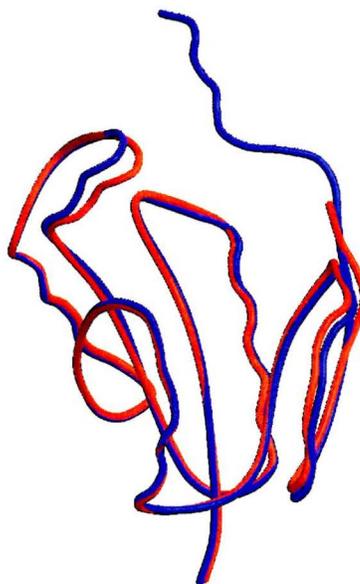


Figure 3.9 Overlay of TagD_{Bs} with subunit of TarD_{Sa}. TagD_{Bs} contains 128 residues, whereas TarD_{Sa} contains only 112 residues.

3.4.2 Tetramerization of TarD_{Sa} vs. Dimerization of TagD_{Bs}

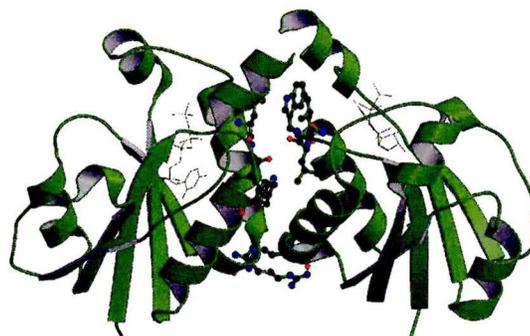
The crystal structure analysis reveals that TarD_{Sa} and TagD_{Bs} contain a different number of molecules in the asymmetric unit: dimerization of TagD_{Bs} vs. tetramerization of TarD_{Sa}. The gel-filtration and ultracentrifugation studies also indicate that TarD_{Sa} exists as a tetramer in solution (see also Section 1.3). The dimer interface of TagD_{Bs} is hydrophobic, involving the ⁶³RYVDEVI sequences and ¹⁴HWGH motifs, as shown in Figure 3.10a. Other residues participated in this region include Leu12 and Leu13 from each subunit (Weber et al., 1999).

This interface arrangement is similarly observed in the crystal structure of TarD_{Sa}, forming the twofold local symmetry between subunits I and II and between subunits III and IV of TarD_{Sa}. As shown in Figure 3.10b, there is an interface region between subunits I and III which involve residues in their first _{3₁₀} helix: His93(I) and His93(III) interact with a separation of 3.09Å. Subunits I and IV, as well as subunits II and III, do not seem to have any significant interactions between them, except for some residues that precede helix C (Figure 3.10c).

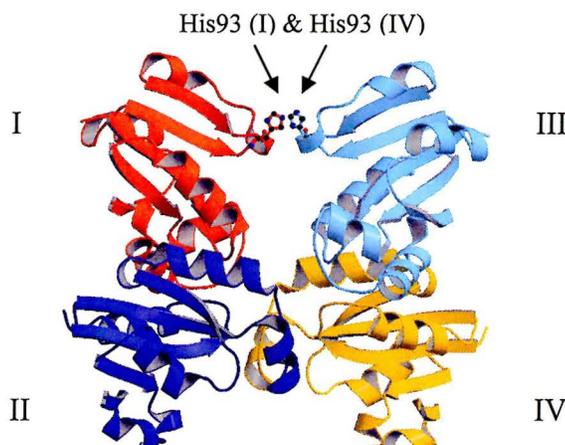
Looking at the structure of TarD_{Sa}, subunits I and III show that there is not enough room for both C-termini to be near the core. Along with this problem, one of the C-termini may not be able to serve as an optimal position for catalytically competent CTP binding. As a consequence, it is reasonable to have three out of four or even fewer catalytically active subunits in the crystal structure of TarD_{Sa}. The tetramer could lead to

a CTP conformation different from the one observed in TagD_{Bs}, and this difference could also occur between different subunits of TarD_{Sa}.

(a)



(b)



(c)

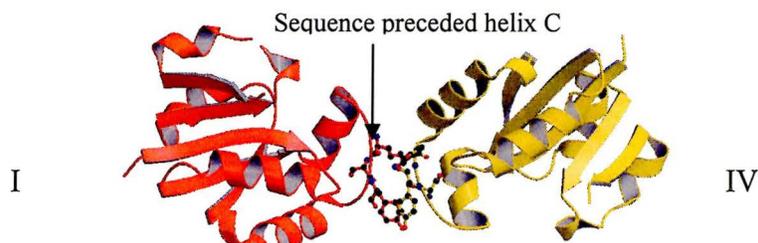


Figure 3.10 Dimer interface region in TagD_{Bs} vs. tetramer interface regions in TarD_{Sa}

- (a) Residues involved in the interface region of TagD_{Bs} are shown in ball-and-stick. Residues Arg63, Trp15, Leu12, and Leu13 are sitting in the interface region of TagD_{Bs}. This interface region is also found in TarD_{Sa} and located between subunits I and II, as well as between subunits III and IV.
- (b) A local twofold symmetry is formed between a pair of subunits I and II and a pair of subunits III and IV. The 3₁₀ segments of subunits I and III with the closest contacts of residues His93(I) and His93(III), shown in ball-and-stick, are found in this interface region.
- (c) Residues that precede helix C are shown in ball-and-stick, forming the closest interactions between subunits I and IV, as well as subunits II and III.

3.4.3 A Tentative CTP Binding Site in TarD_{Sa}

The binding pocket of the tentative CTP in the Fo-Fc *sa* omit map at 1 sigma contour in Section 3.3 (Figure 3.11) is found to be very similar to that of TagD_{Bs}. Owing to the absent electron density towards the C-terminus, this CTP binding site is lined by three clusters of residues in TarD_{Sa}, whereas there are four in the crystal structure of TagD_{Bs} (see Section 1.4). The ⁸GTYDLLHYGH sequence is shown to be responsible for part of the active site. Residues including Lys44, Lys46, and Trp74 are buried at the other side of the binding bowl, which is predicted to be the active site for glycerol-3-phosphate (Weber et al., 1999).

The first ₃₁₀ segment following the IV β -strand, ⁹¹MGHDW, also takes part in the formation of the substrate binding pocket. Like the CTP in TagD_{Bs}, the CTP is found very close to the motifs of ⁸GTYDLLHYGH. However, the conformation of CTP in subunit I of TarD_{Sa} is shown to be different from the one seen in the structure of TagD_{Bs} (Figure 3.11). The difference of CTP in the crystal structures of TagD_{Bs} and TarD_{Sa} is depicted in Figure 3.11, TarD_{Sa}'s CTP lies in a flat position instead of folding inward, leading to a different CTP binding mode.

In a different binding mode, the six-membered ring of CTP in TarD_{Sa} sits closer to the ₃₁₀ segment and binds to the residues Met91 and Trp95 (Figure 3.12). In TagD_{Bs}, this six-membered ring of CTP is interacting with residues Thr114 and Ile117 of the second conserved motif of ¹¹³RTEGISTT (Figure 3.13). The presence and difference of

the CTP observed in the crystal structure of TarD_{Sa} remain questionable, because low-resolution data cannot be used to strongly support the detailed molecular analysis.

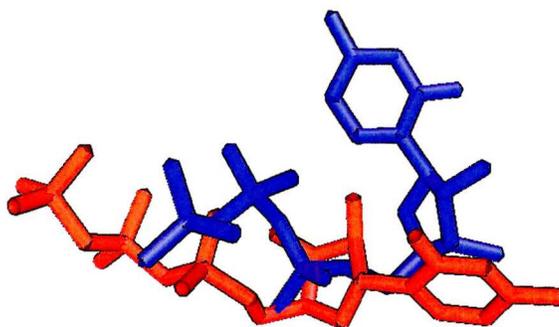


Figure 3.11 Overlay of CTP in TagD_{Bs} with the One in TarD_{Sa}. CTP in TagD_{Bs} is shown in blue. CTP in TarD_{Sa} is shown in red.

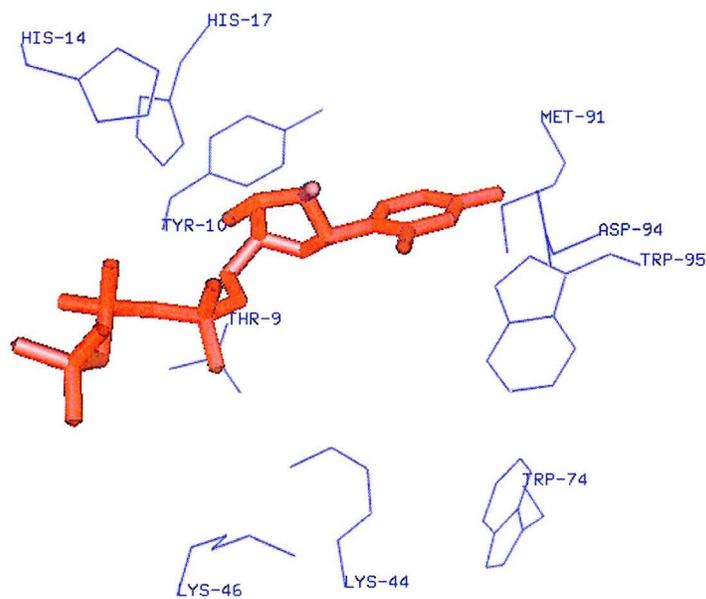


Figure 3.12 Location of the CTP observed in a Fo-Fc σ_a omit map at 1 sigma contour (see also Section 3.3) in subunit I of TarD_{Sa}. TarD_{Sa}'s CTP is shown in ball-and-stick (red). Side chains of TarD_{Sa} residues are shown by blue sticks.

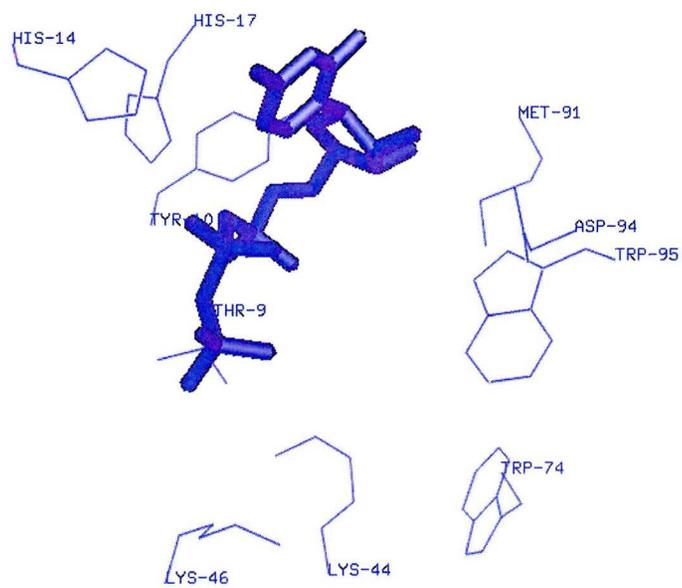


Figure 3.13 TagD_{Bs}'s CTP modeled in the crystal structure of TarD_{Sa}
TagD_{Bs}'s CTP is shown in ball-and-stick (blue). Side chains of TarD_{Sa} residues are shown by blue sticks.

3.5 Summary and Future Work

3.5.1 Summary

Determination of the crystallization condition for the crystal of TarD_{Sa} required a combination of three optimization methods. Although application of various amounts of precipitant did not significantly improve the crystal quality, examination of the effects of small molecules and the oils technique used for control of vapor diffusion rate greatly assisted the optimization of crystal growth. A single crystal was grown in equilibration with 8% (v/v) ethylene glycol, 30% (w/v) D(+) glucose, and 250 μ l of silicon oil and paraffin oil in a 2:1 ratio at 22°C. Crystals thus obtained diffracted to 3.0Å at the BNL-NSLS facilities where a complete data set was collected.

The data set was processed and analyzed by Denzo and Scalepack. Crystals of TarD_{Sa} was determined to be in space group P3₁21 or its enantiomorph P3₂21, with the unit cell dimensions a=92.15Å, b=92.15Å, c=156.12Å, $\alpha=90^\circ$, $\beta=90^\circ$, and $\gamma=120^\circ$. Based on the results obtained from the translation search, the correct space group for this crystal form was determined to be P3₁21. The crystal structure of TarD_{Sa} is shown to contain four molecules in the asymmetric unit, in agreement with the gel-filtration and ultracentrifugation analysis.

A tentative CTP substrate is seen in subunit I of the crystal structure of TarD_{Sa} and is in a different conformation from the one seen in TagD_{Bs}. However, this CTP conformation is not reliable, since the refinement of the CTP in this density shows the B-

factors to be over 100. Due to the low resolution crystal structure, this speculative CTP conformation could not be confirmed.

Co-crystallization of TarD_{Sa} with a higher concentration of CTP was also performed to fill up the empty CTP binding pockets of TarD_{Sa}. Unfortunately, soaking and repetitive co-crystallization of TarD_{Sa} with higher concentration of CTP failed. The overall structure and subunit fold of TarD_{Sa} is shown to be very similar to the TagD_{Bs} structure, with a root mean deviation of $\sim 0.6\text{\AA}$. Crystals of TarD_{Sa} with glycerol were obtained from several crystallization conditions; however, they diffracted only to 5.0\AA at NSLS.

3.5.2 Future Work and Concluding Remarks

Two crystallization conditions have been obtained recently and found to grow a single crystal, as shown in Figures 3.18. The protein condition contains 10mg/mL of TarD_{Sa}, with 10mm of glycerol and MgCl₂. These crystals are cryo-frozen in liquid nitrogen and will be tested on the next NSLS trip. If data collection of shown crystals is not successful, co-crystallization of TarD_{Sa} with glycerol-3-phosphate (TarD_{Sa}-GroP) could be continued. The obtained glycerol binding information will greatly support the mutational studies focused on this active site.

Additionally, the missing electron density in the core region of TarD_{Sa} could perhaps be resolved if crystal structures of TarD_{Sa} can be determined at higher resolution.

Co-crystallization of TarD_{Sa} with the enzymatic product, CDP-glycerol, will be also required to provide a complete understanding of CTP and GroP binding to the protein.



Figure 3.13 Crystals of TarD_{Sa}-GroP grown in the new crystallization conditions. A single crystal of TarD_{Sa}-GroP (left) was grown in cryo-conditions of 35% (v/v) 2-ethoxyethanol, Na/K phosphate at pH6.2, and 0.2M NaCl at 22°C. The baby crystal was broken off before cryo-freezing. Another crystallization condition of 35% (v/v) 2-propanol, citrate pH5.5, and 5% (w/v) PEG-1000 also yielded a couple crystals of TarD_{Sa}-GroP (right).

REFERENCES

- ¹ <http://id.medscape.com/reuter/prof/2000/04/04.19/20000419rglt007.html>
- ² <http://id.medscape.com/reuters/prof/2001/07/07.20/20010719clin004.html>
- ³ <http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html>
- ⁴ <http://www.hamptonresearch.com/hrproducts/2420.html>
- ⁵ <http://www.hamptonresearch.com/techcenter/pdf/2420G.pdf>
- ⁶ <http://www.hamptonresearch.com/techcenter/pdf/3411G.pdf>
- ⁷ <http://www.hamptonresearch.com/hrproducts/3411.html>
- ⁸ <http://www.hamptonresearch.com/hrproducts/3413.html>
- ⁹ <http://www.hamptonresearch.com/hrproducts/4710.html>
- ¹⁰ <http://www.hamptonresearch.com/techcenter/pdf/4710G.pdf>
- ¹¹ <http://www.structmed.cimr.cam.ac.uk/Course/MolRep/molrep.html#concepts>

Amzel, L.M. (1998) Structure-based drug design. *Current Opinion in Biotechnology*, **9**, 366-369.

Arnez, J.G. & Steitz, T.A. (1994) Crystal structures of unmodified tRNA (Gln) complexed with glutamyl-tRNA synthetase and ATP suggests a possible role for pseudo-uridines in stabilization of RNA structure. *Biochemistry*, **33(24)**, 7560-7567.

Archer, G.L. & Bosilevac, J.M. (2001) Signaling antibiotic resistance in *Staphylococci*. *Science*, **291**, 1915-1916.

Archer, G.L. (1998) *Staphylococcus aureus*: A well-armed pathogen. *CID*. **26**, 1179-1181

Baddiley J. (1970) Structure, biosynthesis, and function of teichoic acids. *Accounts. Chem. Res.* **3**, 98-105.

Bhavsar, A.P., Beveridge, T.J., & Brown, E.D. (2001) Precise deletion of *tagD* and controlled depletion of its product, glycerol-3-phosphate cytidyltransferase, leads to irregular morphology and lysis of *Bacillus subtilis* grown at physiological temperature. *Journal of Bacteriology*. **183**, 6688-6693.

- Blow, D.M., Chayen, N.E., Lloyd, L.F. & Saridaki, S.E. (1994) Control of nucleation of protein crystals. *Protein Science*, **3**, 1638-1643.
- Briehl, M., Pooley, H.M., & Karamata, D. (1989) Mutants of *Bacillus subtilis* 168 thermosensitive for growth and wall teichoic acid biosynthesis. *J. Gen. Microbiol.* **135**, 1325-1334.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantization of microgram quality of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.
- Brünger, A.T. (1992) X-PLOR version 3.1 A system for X-ray Crystallography and NMR, Yale University Press, New Haven, CT.
- Brünger, A.T. (1993) Assessment of phase accuracy by cross validation: the Free R value. Methods and applications. *Acta Cryst.* **D49**, 24-36.
- Brünger, A.T. (1997) Patterson Correlation Searches and Refinement, p. 558-579, *In Carter, C.W. Jr, Sweet, R. (ed.), Methods in Enzymology 276* Academic Press Inc.
- Carson, M. (1997) "Ribbons" *Methods in Enzymology*, 277, 493-505. R.M. Sweet and C.W. Carter, eds, Academic Press.
- Champney, W.S. & Miller, M. (2002) Linezolid is a specific inhibitor of 50S ribosomal subunit formation in *Staphylococcus aureus* cells. *Current microbiology.* **44**, 350-356.
- Chayen, N.E. (1997) A novel technique to control the rate of vapour diffusion, giving larger protein crystals. *Journal of Applied Crystallography.* **30**, 198-202.
- Cohen, S.L. (1996) Domain elucidation by mass spectrometry. *Structure*, **4**, 1013-1016.
- Crowther, R.A. & Blow, D.M. (1967) A method of positioning a known molecule in an unknown crystal structure. *Acta Cryst. A* **23**, 544-548.
- Cudney, B. et al. (1994) Screening and optimization strategies for macromolecular crystal growth. *Acta Cryst.* **D50**, 414-423.
- Delage, G. et al. (1999) Control of methicillin-resistant *S.aureus* in Canadian paediatric institutions is still a worthwhile goal. *Paediatrics & Child Health.* **4(5)**, 337-341.
- Drenth, J. (1999) Principles of Protein X-Ray Crystallography, Springer-Verlag New York, Inc.

- Engh, R.A. & Huber, R. (1991) Accurate bond and angle parameters for x-ray protein structure refinement. *Acta Cryst.* **A47**, 392-400.
- Gane, J.G. & Dean, P.M. (2000) Recent advances in structure-based rational drug design. *Current Opinion in Structural Biology*, **10**, 401-404.
- Hiramatsu, K. (2001) Vancomycin-resistant *Staphylococcus aureus*: a new model of antibiotic resistance. *Lancet Infectious Diseases*, **1**, 147-155.
- Jancarik, J. & Kim, S.H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.*, **24**, 409-411.
- Johnson, L.N., Stura, E.A., Sansom, M.S., & Babu, Y.S. (1983) Oligosaccharide binding to glycogen phosphorylase b. *Biochem Soc Trans*, **2**, 142-4.
- Jones, T.A. & Kjeldgaard, M. (1998) Essential O
- Jovine, L. (2000) A simple technique to control macromolecular crystal nucleation efficiently using a standard vapor-diffusion setup. *J. Appl. Cryst.* **33**, 988-989.
- Klebe, G. (2000) Recent developments in structure-based drug design. *J. Mol. Med.*, **78**, 269-281.
- Kobayashi, N., Urasawa, S., Uehara, N., & Watanabe, N. (1999) Analysis of Genomic Diversity Within the Xr-region of the Protein A Gene in Clinical Isolates of *Staphylococcus aureus*. *Epidemiol. Infect.* **122**, 241-249.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., & Thornton, J.M. (1993) PROCHECK: a program to check the stereochemistry of protein structures. *J. Appl. Cryst.* **26**, 283-291.
- Lewis, R. (1995) The risk of antibiotic-resistant infections. *FDA Consumer magazine*, <http://www.fda.gov/fdac/features/795.antibio.html>
- Lowy, F.D. (1998) Review Articles: *Staphylococcus aureus*. *The New England Journal of Medicine*. **339**, 520-532.
- Luft, J.R., Albright, D.T., Baird, J. K., & DeTitta, G.T. (1996) The Rate of Water Equilibration in Vapor-diffusion Crystallization: Dependence on the Distance from the Droplet to the Reservoir. *Acta Cryst.* **D52**, 1098-1106.
- Marchese, A., Schito, G.C., & Debbia, E.A. (2000) Evolution of antibiotic resistance in gram-positive pathogens. *J Chemother*, **12(6)**, 459-462.
- Mattews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.

- Mauel, C., Young, M., & Karamata, D. (1991) Genes concerned with synthesis of poly(glycerol phosphate), the essential teichoic acid in *Bacillus subtilis* strain 168, are organized in two divergent transcription units. *J. Gen. Microbiol.* **137(Pt4)**, 929-41.
- McPherson, A. (1982) Preparation and Analysis of Protein Crystals. J. Willey and Sons. New York.
- McPherson, A. (1995) Increasing the size of microcrystals by fine sampling of pH limits. *J. Appl. Cryst.* **28**, 362-365.
- Mikol, V., Hirsch, E., & Giege, R. (1990) Diagnostic of Precipitant for Biomacromolecule Crystallization by Quasi-elastic Light-scattering. *J. Mol. Biol.* **213**, 187-195.
- Otwinowski, Z. & Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, p. 307-325, *In Carter, C.W. Jr, Sweet, R. (ed.), Methods in Enzymology* **276** Academic Press Inc.
- Pannu, N.S. & Read, R.J. (1996) Improved structure refinement through maximum likelihood. *Acta Cryst.* **A52**, 659-668.
- Petock, J.H., Wang, Y, F., DuBois, G.C., Harrison, R.N., & Weber, I.T. (2001) Effects of different post-crystallization soaking conditions on the diffraction of *Mtcpl* crystals. *Acta Cryst.* **D57**, 763-765.
- Pooley, H.M., Abellan, F.X., & Karamata, D. (1991) A conditional-lethal mutant of *Bacillus subtilis* 168 with a thermosensitive glycerol-3-phosphate cytidyltransferase, an enzyme specific for the synthesis of the major wall teichoic acid. *J. Gen. Microbiol.* **137**, 921-928.
- Sousa, R. (1995) Use of glycerol, polyols and other protein structure stabilizing agents in protein crystallization. *Acta Cryst.* **D51**, 271-277.
- Rossmann, M.G. & Blow, D.M. (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Cryst.* **A 15**, 24-31
- Smith, T.L. et al. (1999) Emergence of vancomycin resistance in *Staphylococcus aureus*. *N. Engl. J. Med.* **340**, 493-501.
- Sweitzer, T.D. & Kent, C. (1994) Expression of wild type and mutant rat liver CTP: phosphocholine cytidyltransferase in a cytidyltransferase deficient Chinese hamster ovary cell line. *Arch. Biochem. Biophys.* **311**, 107-116.

Tortora, G.J., Funke, B.R., & Case, C.L. (1995) Microbiology an Introduction, Edition 5. P.283. The Benjamin/Cummings Publishing Company, Inc.

Trakhanov, S. & Quioco, F.A. (1995) Influence of divalent cations on protein crystallization. *Protein Science*, **4**, 9: 1914-1919.

Venkatachalam, K.V., Fuda, H., Koonin, E.V., & Strott, C.A. (1998) Site-selected Mutagenesis of a Conserved Nucleotide Binding HXGH Motif Located in the ATP Sulfurylase Domain of Human bifunctional 3'-Phosphoadenosine 5'-Phosphosulfate Synthase. *J. Biol. Chem.* **274**, 5: 2601-2604.

Weber, C.H., Young, S.P., Sanker, S., Kent, C., & Ludwig, M.L. (1999) A prototypical cytidylyltransferase: CTP: Glycerol-3-phosphate cytidylyltransferase from *Bacillus subtilis*. *Structure*, **7**, 1113-1124.

Yim, V.C.N., Zolli, M., Badurina, D.S., Rossi, L., Brown, E.D., & Berghuis, A.M. (2001) Crystallization and preliminary X-ray diffraction studies of glycerol-3-phosphate cytidylyltransferase from *Staphylococcus aureus*. *Acta Cryst. D* **57**, 918-920.

Von Delft, F. et al. (2001) The crystal structure of E.coli pantothenate synthetase confirms it as a member of the cytidylyltransferase superfamily. *Structure*, **9**, 439-450.

Young et al. (1997) Identification of functional conserved residues of CTP: glycerol-3-phosphate cytidylyltransferase. *J. Biol. Chem.* **272**(24), 15161-15166.

Zhang, H.Z., Hackbarth, J., Chansky, K.M., & Chambers, H.F. (2001) A proteolytic transmembrane signaling pathway and resistance to β -Lactams in *Staphylococci*. *Science*, **291**, 1962-1964.

Zulauf, M. & D'Arcy, A. (1992) Light scattering of proteins as a criterion for crystallization. *J. Crystal Growth*. **122**, 102-106.

APPENDIXES

APPENDIX I

Molecular Replacement Method

AI.1 Introduction and Basic Concept of Molecular Replacement

Phasing by molecular replacement (Rossmann and Blow, 1962) is a convenient method to determine the structure of a protein which has a similar or homologous structure to a protein that has already been solved. The choice of a protein for the solution of the unknown model can be a structure that is homologous in its amino acid sequence. The general rule of thumb for the homology of the selected model is to share more than 40% sequence identity with the target structure¹¹. In the Protein Data Bank, a number of known crystal structures with different folding patterns can be taken as the probe model for the unknown structures. Besides the sequence homology against the unknown protein, the completeness and quality of the experimental data are also important to the result of molecular replacement. Molecular replacement is practiced not only in determination of a known structure that exists in different crystal forms such as space group and unit cell parameters, but also in studies of substrate-binding and site-directed mutagenesis. The information obtained from the substrate/inhibitor binding studies can be established to enhance the process of rational drug design. This research is aimed at the application of molecular replacement method to solve the crystal structure of *S.aureus* TarD using TagD_{Bs}, as a search model.

¹¹<http://www..structmed.cimr.cam.ac.uk/Course/MolRep/molrep.html#concepts>

The objective of molecular replacement is to place and orient the search model in a crystal lattice such that it closely matches the experimentally obtained crystals. In this respect, the search model is defined as a model constructed based on homology with a known structure that matches the real structure of the protein as closely as possible. What is required is to determine for each search model present in the asymmetric unit, the position and orientation can be described by two sets of three parameters¹¹. Searching for these parameters, however, is too computationally intensive, inefficient, and prone to false solutions. In molecular replacement, this problem is subdivided into two functional problems: rotational and translational. A basic concept of rotational and translational search is illustrated schematically in Figure AI.1. The search and experimental model are symbolized as M_1 and M_2 , respectively. To superimpose M_1 and M_2 , first is to apply a rotation on the starting position M_1 , and then a translation is applied on the rotated version M_1' to the target position M_2 .

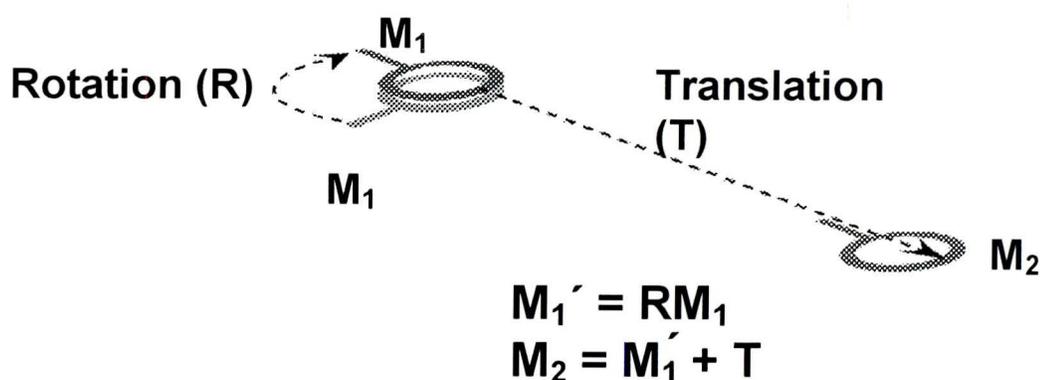


Figure AI.1 Basic concept of rotation and translation function. Superposition of the probe and target molecules, M_1 and M_2 , requires two steps: rotation and translation. M_1 is rotated to a proper orientation M_1' by applying the rotation function $M_1 = RM_1'$. The rotated molecule M_1' is subsequently translated to a proper position by the function of $M_2 = M_1' + T$.

If the rotation and translation of the probe model can be performed successfully in molecular replacement, the correct orientation and position of the target model in the crystal lattice can be defined. The calculated phases from the model structure factor may be then applied with the observed structure factor and amplitudes to build an initial electron density map for the target model. The quality of the fit to the calculated density can be examined using graphic software. However, a starting model usually contains errors so that the initial calculated data may not agree with the observed data. For example, there may be some false interpretation of breaks and connections observed in the electron density map. Therefore, the initial model will undergo cycles of model rebuilding and refinement in order to obtain a precise protein structure.

AI.2 Rotation Function

The phase of the structure factor is the key factor to determine the orientation and position of the unknown model; however, this piece of information is not available in the data set. Fortunately the Patterson function can generate vectors between atoms of the model molecules without phase information. Figure AI.2a&b illustrate how to make a Patterson map for two atoms in a unit cell. In this figure, there is a peak of electron density for the first atom at position 1 and a peak of electron density for the second atom at position 2. The corresponding Patterson map will have a vector from peak 2 to peak 1 and have another one from peak 1 to peak 2. Both vectors are of equal size but of opposite direction, thus the Patterson will display inversion symmetry. The unit cell in real space and Patterson space are identical. If there are N peaks in the real cell, the corresponding Patterson map will have $N^2 - N$ peaks excluding those overlapped at the origin (Drenth, 1999). A Patterson map also contains two set of vectors: intramolecular vectors (self-Patterson vectors) and intermolecular vectors (cross-Patterson vectors). Intramolecular vectors as shown in Figure AI.2c are marked from one atom to another within the same molecule. Intermolecular vectors (Figure AI.2d) are the ones that cross from one molecule to another, and they are related by the crystallographic symmetry. Note that the intramolecular vectors rely on the orientation of the model molecule in the unit cell while intermolecular vectors depend on both orientation and position of the model¹¹.

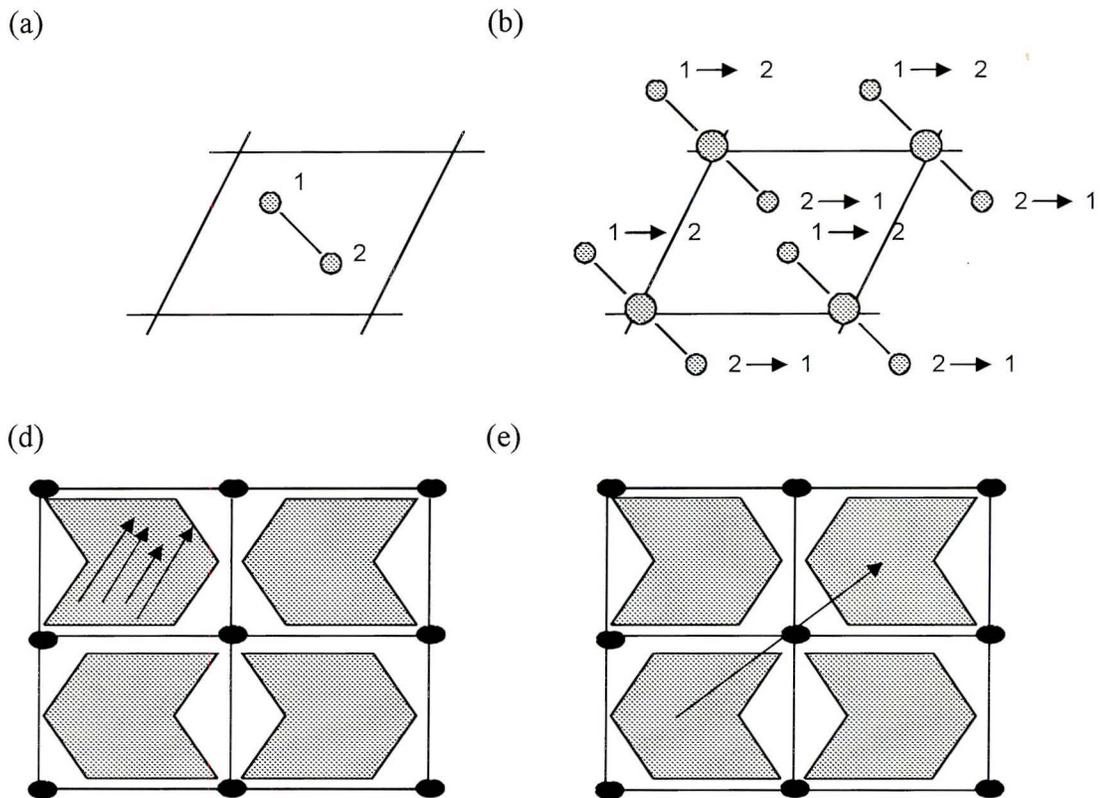


Figure A1.2 Diagrams of Patterson Function and Intra- and Inter-molecular vectors.

- (a) Two atoms are present in the unit cell
- (b) Patterson function yields a vector from peak 2 to peak 1 and another one from peak 1 to peak 2
- (c) Intermolecular vectors are joining the atoms within the same molecule
- (d) Intramolecular vectors are the vectors present between two atoms in two different molecules

The purpose of the rotation function in molecular replacement is to compare two Patterson maps in order to superimpose two molecules belonging to two different orientations. In other words, a correct orientation of the target model is based on the angular relationship between two identical molecules within one asymmetric unit (Drenth, 1999). The rotation function (Rossmann and Blow, 1962) shown below is a product function used in real space Patterson function rotation.

$$F_{\text{Rot}}(\Omega) = \int_U \mathbf{P}_1(\mathbf{u})\mathbf{P}_2(\mathbf{u}_r) d\mathbf{u} \quad (\text{Equation A.1})$$

In the above equation, F_r is the rotation function representing the overlap of the Patterson function of P_1 with P_2 . A maximum overlap of Patterson maps is obtained (right in Figure AI.3) if the molecule is correctly oriented in the target cell (middle in Figure AI.3).

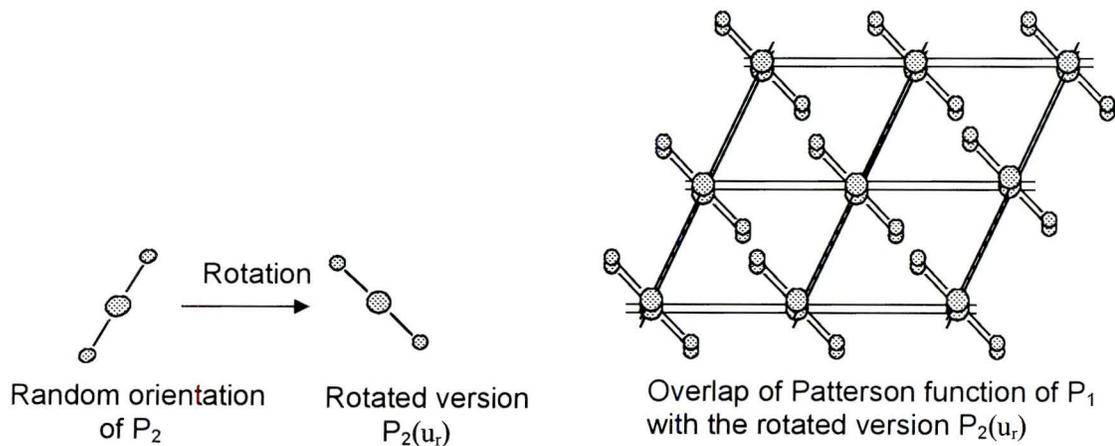


Figure AI.3 Maximum overlap of Patterson Functions resulted in the rotation search. By applying rotation operation (R), P_2 is rotated to a correctly orientation, $P_2(u_r)$, and contributes to a maximum overlap of Patterson maps between P_1 and P_2 .

AI.3 CNS: Direct Rotation and PC Refinement

The real-space intramolecular vectors are shorter than the intermolecular vectors on average. To exclude the integration of intermolecular vectors when using real space Patterson rotation search (see section AI.2), such computation will use vectors in the region near the origin. Due to this limitation, some peaks that contribute to the correct solution of the rotation search may not be included. Therefore, direct rotation has been introduced in the molecular replacement programs CNS. The difference of direct rotation from real space rotation is to place the search model in an artificial cell corresponding to the same dimensions of the observed unit cell (Figure AI.4). Instead of rotating the Patterson map of the search model, the whole model itself is integrated in the rotation function. Reciprocal-space structure factors are then recomputed for each tested orientation of the model and compared with the observed model. There are three cross rotation functions in CNS program: real space Patterson function (Rossmann and Blow, 1962), direct rotation, and fast direct rotation (Brünger, 1997). Fast direct rotation was employed for the phase determination of the crystal structure of TarD_{Sa} in this experiment.

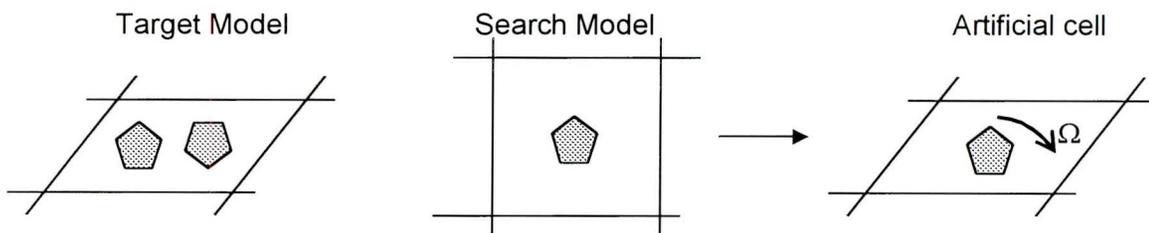


Figure AI.4 Direct Rotation. Direct rotation search is employed in the CNS program. The model is placed in an artificial unit with cell dimensions and angles equivalent to those in target protein. Instead of moving the Patterson map, this method will conduct a rotation of the whole model unit. Structure factors are recomputed for the obtained orientation of the model. The accuracy of the solutions is monitored by the correlation coefficient Ω .

However, a full direct rotation search must calculate FFT for each sampled orientation, and that will take tremendous amounts of time. In CNS, the computation of direct rotation can be accelerated by a fast direct rotation method that performs the finer searches just around the top solutions in the initial coarse grid search. Before the translation search, those rotation peaks identified in the finer searches will be optimized by least squares PC refinement (Brünger, 1997) to produce the best fit to the actual structural factor. The whole structure can be also divided into several parts treated as constraint for the PC refinement. On the other hand, it is possible to assign the whole model as a rigid body in the position and orientation PC refinement. The criterion of fit of the direct rotation peaks will be monitored by calculating a correlation coefficient, $C(\Omega)_{\text{Direct}}$ (Equation AI.2).

$$C(\Omega)_{\text{Direct}} = \frac{\sum_{\text{hkl}} [(|\mathbf{F}_{\text{obs}}|^2)(|\mathbf{F}_{\text{cal}}(\Omega)|^2) - (|\mathbf{F}_{\text{obs}}|^2)(|\mathbf{F}_{\text{cal}}(\Omega)|^2)]}{\left[\sum_{\text{hkl}} (|\mathbf{F}_{\text{obs}}|^4 - |\mathbf{F}_{\text{obs}}|^2)^2 (|\mathbf{F}_{\text{cal}}(\Omega)|^4 - |\mathbf{F}_{\text{cal}}(\Omega)|^2)^2 \right]^{1/2}} \quad \text{(Equation AI.2)}$$

Then, a coefficient correlation translation followed by a PC refinement will be carried out based on the selected rotation peaks.

AI.4 Translation Function

After the orientation of the model is defined, the next step in molecular replacement is to find its proper position in the unit cell. This is done by performing a translation search. The translation search can be accomplished using Patterson Function studied by (Crowther and Blow, 1967). This way, the correct position for the search model is achieved by the computation of a maximum overlap between the observed and calculated Patterson maps. In CNS, however, a reciprocal space correlation coefficient translation is actually employed (Brünger, 1992). The oriented model will be moved around in the unit cell to find the best position matching with the actual model. The structure factors are recomputed and compared to the observed structure factor by calculating the correlation coefficient shown below (Equation AI.3).

$$C(\mathbf{t}) = \frac{\sum_{\mathbf{hkl}} (|\mathbf{F}_{\text{obs}}|^2 - \overline{|\mathbf{F}_{\text{obs}}|^2}) \times (|\mathbf{F}_{\text{cal}}(\mathbf{t})|^2 - \overline{|\mathbf{F}_{\text{cal}}(\mathbf{t})|^2})}{\left[\sum_{\mathbf{hkl}} (|\mathbf{F}_{\text{obs}}|^4 - \overline{|\mathbf{F}_{\text{obs}}|^2})^2 \sum_{\mathbf{hkl}} (|\mathbf{F}_{\text{cal}}(\mathbf{t})|^4 - \overline{|\mathbf{F}_{\text{cal}}(\mathbf{t})|^2})^2 \right]^{1/2}} \quad \text{(Equation AI.3)}$$

AI.5 $R_{\text{conventional}}$ Vs R_{free}

The correctness of the model is accessed by the R factor. The R factor is used to monitor the fit of a model to the diffraction data by measuring the agreement between the observed (F_o) and calculated (F_c) structure-factor amplitudes as the following: $R_{\text{factor}} = \sum ||F_o| - |F_c|| / |F_o|$. Due to the relatively low resolution of protein structures (i.e. $>0.7\text{\AA}$), it is possible to introduce errors in the model without affecting the R_{factor} adversely. Therefore, a method of cross validation using the R_{free} value was proposed by Brünger (1992, 1993). Firstly, diffraction data are divided into two categories, the 90% 'working' set and the 10% 'test' set. Any changes or refinement to the model will only be guided by the 90% working set, and its correctness is subsequently evaluated by calculating an R_{factor} for this data, referred to as the $R_{\text{conventional}}$. The 10% 'test' set will be excluded during structure refinement, and a separate R_{factor} , R_{free} , will be calculated for assessing the fit of the model to the 'test' set at every stage of the refinement process. Since R_{free} is closely related to the error in the phase, it should detect the over-fitting problem more efficiently.

APPENDIX II

Procedure for Preparing Siliconized Cover-glass

- (1) Place the non-siliconized cover-slips onto the custom-made rack.
- (2) Fill the graduated cylinder with the dimethyldichlorosilane solution to about 300mL.
- (3) Immerse the rack placed with all the non-siliconized cover-slips into the dimethyldichlorosilane solution.
- (4) Rinse the cover-slips with dimethyldichlorosilane solution for 30 seconds.
- (5) Rinse the cover-slips with the tap water, then with deionized water for 1-2 minutes.
- (6) Dry cover-slips in the oven at 350°F for 90 minutes.