STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE *Bordetella pertussis* CyaA-RTX SUBDOMAIN FRAGMENT

KITTIPONG SUVARNAPUNYA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE
(MOLECULAR GENETICS AND GENETIC ENGINEERING)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2012

Thesis
entitled

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE *Bordetella pertussis* CyaA-RTX SUBDOMAIN FRAGMENT

……
Mr. Kittipong Suvarnapunya
Candidate

……....
Assoc. Prof. Chanan Angsuthanasombat, Ph.D.
Major advisor

……
Assoc. Prof. Albert Ketterman, Ph.D.
Co-advisor

……
Assist. Prof. Gerd Katzenmeier, Ph.D.
Co-advisor

..……..
Prof. Banchong Mahaisavariya,
M.D., Dip Thai Board of Orthopedics
Dean
Faculty of Graduate Studies
Mahidol University

..……..
Assoc. Prof. Apinunt Udomkit, Ph.D.
Program Director
Master of Science Program in
Molecular Genetics and Genetic Engineering
Institute of Molecular Biosciences
Mahidol University

Thesis
entitled

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE *Bordetella pertussis* CyaA-RTX SUBDOMAIN FRAGMENT

was submitted to the Faculty of Graduate Studies, Mahidol University
for the Master degree of Sciences
(Molecular Genetics and Genetic Engineering)
on
November 16, 2012

……
Mr. Kittipong Suvarnapunya
Candidate

..……..
Assoc. Prof. Apinunt Udomkit, Ph.D.
Chair

……....
Assoc. Prof. Chanan Angsuthanasombat, Ph.D.
Member

……
Assist. Prof. Gerd Katzenmeier, Ph.D.
Member

……
Lect. Chalermpol Kanchanawarin, Ph.D.
Member

..……..
Prof. Banchong Mahaisavariya,
M.D., Dip Thai Board of Orthopedics
Dean
Faculty of Graduate Studies
Mahidol University

..……..
Prof. Prasert Auewarakul,
M.D. Dr. med
Director
Institute of Molecular Biosciences
Mahidol University

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere appreciation and gratitude to my advisor, Associate Professor Dr. Chanan Angsuthanasombat for his helpful guidance, valuable assistance and generous encouragement throughout the course of this research. I would like to extend my sincere thanks to Associate Professor Dr. Apinunt Udomkit, Assistant Professor Dr. Gerd Katzenmeier, Dr, Chalermpol Kanchanawarin for their time and comment and suggestions.

Gratefully thanks are extended to the Institute of molecular bioscience and Faculty of graduate studies, Mahidol University for providing me great opportunity to do this project

Finally, I would like to specially thank my family and research group members for their advice and encouragement throughout my entire education.

Kittipong Suvarnapunya

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE *Bordetella pertussis* CyaA-RTX SUBDOMAIN FRAGMENT

KITTIPONG SUVARNAPUNYA 5036581 MBMG/M

M.Sc. (MOLECULAR GENETICS AND GENETIC ENGINEERING)

THESIS ADVISORY COMMITTEE: CHANAN ANGSUTHANASOMBAT, Ph.D., ALBERT KETTERMAN, Ph.D., GERD KATZENMEIER, Ph.D.,

ABSTRACT

The bi-functional adenylate cyclase-hemolysin toxin (CyaA) is secreted by *Bordetella pertussis*, whooping cough pathogen, as a major toxin during its colonization in human upper respiratory tract. As a member of repeat-in-toxin (RTX), CyaA contains five blocks of consecutive nona-peptide repeat in the C-terminal. Additionally, to exert its activity, the toxin needed to be acylated at lysine-983 by acyltransferase, CyaC. This study is aimed to elucidate the structure of the RTX region using molecular modeling technique. CyaA-RTX (residues 931-1603) was put into molecular modeling using I.3 lipase from *pseudomonas sp.* as model template. The molecular parameters were all in acceptable range with 3.8% amino acid in not allowed region of Ramachandran plot. The main chain and side chain parameters were equivalent or better compared to the standard value. The model suggested that the regions were stabilized by four major forces, network of triggerable calcium binding, core hydrophobic interactions, beta-roll structures and inter-block interactions. In addition, *in vivo* studies reveal that the CyaA-RTX (residues 751-1706) co-expressed with the active CyaC, was unacylated. This infers the importance of the PF region (residues 482-750) in the activation mechanism of the toxin. Understanding the folding and mechanism of the toxin pave the way for better treatment and prevention of whooping cough and others RTX toxin related illness.

KEY WORDS: *Bordetella pertussis*/ ADENYLATE CYCLASE-HEMOLYSIN TOXIN/ REPEAT-IN-TOXIN/ RTX

131 pages

การศึกษาโครงสร้างและหน้าที่การทำงานของโปรตีนสารพิษ CyaA ในส่วน RTX จากเชื้อแบคทีเรีย Bordetella pertussis

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE Bordetella pertussis CyaA-RTX SUBDOMAIN FRAGMENT

กิตติพงศ์ สุวรรณะบุณย์ 5036581 MBMG/M

วท.ม. (อณูพันธุศาสตร์และพันธุวิศวกรรมศาสตร์)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์: ชนันท์ อังศุธนสมบัติ, Ph.D., GERD KATZENMEIER, Ph.D., ALBERT KETTERMAN, Ph.D.

บทคัดย่อ

โรคไอกรนเกิดจากเชื้อแบคทีเรีย *Bordetella pertussis* ซึ่งระหว่างการก่อโรคที่บริเวณทางเดินหายใจส่วนบนจะหลั่งโปรตีนสารพิษหลายชนิดออกมา หนึ่งในโปรตีนสารพิษที่สำคัญคือ CyaA โปรตีนสารพิษนี้จัดอยู่ในกลุ่ม repeat-in-toxin (RTX) ซึ่งมีลักษณะพิเศษคือ มีชุดของลำดับกรดอะมิโนจำนวน 9 ตำแหน่ง ทั้งหมด 5 ชุด ที่ส่วน C-terminal นอกจากนี้โปรตีนสารพิษ CyaA ต้องการการกระตุ้นโดยการเติมหมู่สายโซ่ acyl โดยเอนไซม์ CyaC ที่ตำแหน่ง lysine-983 ในการศึกษานี้ ต้องการที่จะสร้างโครงสร้างจำลองระดับโมเลกุลของโปรตีนสารพิษ CyaA ในส่วน RTX (ลำดับกรดอะมิโนที่ 931-1603) ด้วยเทคนิค molecular modelling โดยมีโปรตีน I.3 lipase จาก pseudomonas sp. เป็นต้นแบบ ค่ามาตรฐานระดับโมเลกุลต่างๆอยู่ในช่วงที่ยอมรับได้ ค่ามุมของพันธะเพปไทด์ส่วนที่ยอมรับไม่ได้ใน Ramachandran plot อยู่ที่ 3.8% ค่า main chain และ side chain parameters อยู่ในช่วงมาตรฐานหรือดีกว่ามาตรฐาน โครงสร้างจำลองที่ได้แสดงให้เห็นปฏิสัมพันธ์สี่กลุ่มในการก่อโครงร่างของ CyaA-RTX อันได้แก่ แรงจากการจับกับไอออนแคลเซียม, hydrophobic interaction, beta-roll structures และแรงจากปฏิสัมพันธ์ระหว่างบล็อก นอกจากนี้ การศึกษาการผลิตโปรตีนสารพิษ CyaA ส่วน RTX (ลำดับกรดอะมิโนที่ 751-1706) ควบคู่กับเอนไซม์ CyaC ภายในเซลเดียวกัน พบว่า CyaA ที่ผลิตได้ไม่ถูกเติมหมู่สายโซ่ acyl แสดงให้เห็นว่าโปรตีน CyaA ส่วน PF (ลำดับกรดอะมิโนที่ 482-750) มีความสำคัญต่อกลไกการเติมหมู่สายโซ่ acyl ของ CyaC การศึกษาการก่อโครงสร้างและกลไกการกระตุ้นโปรตีนสารพิษ CyaAนี้ เป็นพื้นฐานเพื่อพัฒนาการป้องกันรักษาโรคไอกรน และโรคที่เกี่ยวข้องกับโปรตีนสารพิษในกลุ่มนี้ต่อไป

131 หน้า

CONTENTS

CONTENTS (cont.)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (cont.)

LIST OF ABBREVIATIONS

% (v/v)	Percent volume by volume
% (w/v)	Percent weight by volume
sec	Second (s)
min	Minute (s)
h	Hour (s)
mm	Millimeter (s)
μm	Micrometer (s)
nm	Nanometer (s)
g	Gram (s)
mg	Milligram (s)
μg	Microgram (s)
ng	Nanogram (s)
ml	Milliliter (s)
μl	Microliter (s)
M	Molar (s)
mM	Millimolar (s)
μM	Micromolar (s)
pmol	Picomole (s)
psi	Pound-force per square inch
°C	Degrees Celsius
bp	Base pair (s)
kb	Kilobase (s)
kDa	Kilodalton (s)
U	Unit (s)

LIST OF ABBREVIATIONS (cont.)

AC	Adenylate cyclase domain
ACT	Adenylate cyclase-hemolysin toxin
ATP	Adenosine triphosphate
Amp	Ampicillin
BSA	Bovine serum albumin
B. pertussis	*Bordetella pertussis*
cAMP	Cyclic adenosine monophosphate
CyaA	Adenylate cyclase-hemolysin toxin
CyaA-PF	Pore-forming domain of adenylate cyclase-hemolysin toxin
CyaC	Acyltransferase
CD	Circular dichroism
CTAB	Cetyl trimethylammonium bromide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTPs	Deoxynucleotide triphosphates (dATP, dGTP, dTTP, dCTP)
DTT	1,4-Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
E. coli	*Escherichia coli*
et al.	And others
Hly	Hemolysin
IPTG	Isopropyl-β-D-thiogalactopyranoside
L	Liter (s)
LB	Luria-Bertani medium

LIST OF ABBREVIATIONS (cont.)

MS	Mass spectrometry
OD	Optical density (-ies)
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PF	Pore-forming fragment
PMSF	Phenylmethylsulphonylfluoride
ProCyaA-PF	Pore-forming domain of inactive adenylate cyclase-hemolysin toxin
RNA	Ribonucleic acid
RNase	Ribonuclease
RTX	Repeat-in-Toxin
rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
T_a	Annealing temperature
TAE	Tris acetate EDTA
Tris	Tris (hydroxymethyl)-aminomethane
UV	Ultraviolet
V	Volt (s)

CHAPTER I
INTRODUCTION

Bordetella pertussis is a Gram-negative, aerobic cocco-bacillus usually colonized on ciliate cells of human respiratory tract and causes a highly contagious disease, whooping cough. It produces many virulent factors including the bi-functional protein adenylate cyclase-haemolysin toxin (CyaA). CyaA contains two domains, AC domain which is responsible for adenylate cyclase activity and pore-forming domain (PF). This PF domain is further divided into four functional regions, which are, hydrophobic segments (residues 400-750), acylation region (residues 751-1006), repeat-in-toxin (RTX) moiety (residues 1007-1657) and secretion signal (residues 1658-1706) (Figure 1.1). The hydrophobic part was proved to account for pore-forming activity while the RTX portion was served as a specific receptor-binding domain. The RTX subdomain, ranging from residue 750 to almost C-terminus of the protein, contains over 43 nona-peptide repeats of a consensus sequence GGXGXDXUX (U is hydrophobic amino acid, X can be any amino acid). The RTX repeats are separated into five blocks divided by 40-50 amino acid linkers. In *B. pertussis*, CyaA was translates along with an acyltransferase, namely, CyaC. The protein was accounted for acylation of CyaA at the acylation site, lysine-983. Once acylated, CyaA was secreted out of the cell through type I secretion system (T1SS) comprised of protein in the same operon; CyaB, CyaD and CyaE. Outside the cell, in the present of mM range calcium ion, the tandem repeats in the RTX subdomain associated with Ca^{2+} to fold into β-roll structure, this formation is understood to be an active receptor-binding structure.

Recently, the CyaA-RTX subdomain (residues 751-1706) was cloned into pCyaA-RTX plasmid and over expressed in *E.coli* BL21 as a soluble protein. However, though the CyaA-PF fragment (residues 482-1706) was expressed and purified in similar manner can be *in vitro* acylated by CyaC while CyaA-RTX was fail to be acylated in the same condition. Since *in vivo* acylation may give more suitable

condition compared to our earlier *in vitro* reaction, co-expression of CyaA-RTX and CyaC was considered as an alternative for this enzymatic acyl-transferring reaction.

Moreover, molecular modeling is a power tool for constructing the three-dimensional structure for the protein that cannot be obtained by conventional methods. However, the accurate molecular model can be obtained only from particularly high homology template. To make it harder, the fact that among the RTX protein family, there are only few proteins have a linker region. Nevertheless, extracellular lipase from *pseudomonas* sp. MIS38 fit both criteria and is the best candidate for determining the structure of CyaA-RTX *via in silico* molecular modeling technique.

This study combines the molecular biology methods with the bioinformatics technique to contribute the structural importance and structure-function relationships of the RTX and the adjacent regions of the *Bordetella pertussis* adenylate cyclase-haemolysin toxin.

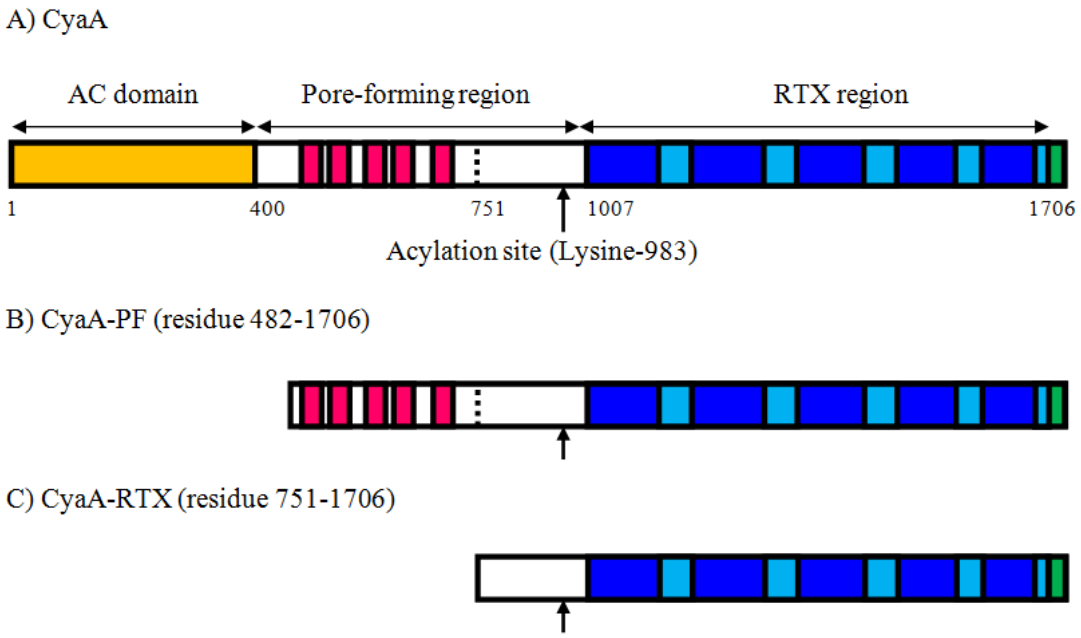


Figure 1.1 Structural organization of the CyaA toxin

(A) Diagram of the CyaA toxin-domain arrangement is shown in color code. The AC domain, hydrophobic pore-forming region, RTX repeat, linker region and secretion signal are shown in yellow, magenta, blue, light blue and green, respectively. The two CyaA clones that used in this experiment are (B) the CyaA-PF containing PF region and RTX region and (C) CyaA-RTX ranging from residue 751 (dash line) to the C-terminal. Noted that the two clones still retain the acylation site (arrow).

CHAPTER II
OBJECTIVES

The function of a protein is dependent on and related to its structure either; at global structure, such as, present or absent of a whole protein fragment, or at local conformation such as interactions between several small motifs. The information of three- dimensional structures greatly enhances the knowledge on the mechanism of the protein action. In order to understand the structure and function relationships of adenylate cyclase-hemolysin toxin (CyaA) especially in the RTX region, several approaches were selected;

1. to study the importance of CyaA-PF and CyaA-RTX subdomain on CyaC- dependent *in vivo* acylation activation.
2. to obtain the molecular model of the RTX region using molecular modeling method.
3. to use the attained information for describing the early stage of protein folding after being secreted.

CHAPTER III
LITERATURE REVIEW

Nowadays, there were about 30-50 million pertussis cases and more than 300,000 deaths each year. Even in one of the most concrete medical care system country like United States of America (USA) where nearly all infants got their DTAP vaccine, still, an incident is still high. The study in 2008 obviously shows the increase in pertussis infection in US citizens (**Figure 3.1**). From January to June 2010, there was an outbreak in California, 1337 cases with 5 infants' deaths. Moreover, there were total of 27,550 cases report that year in USA, increased from 16,858 cases (5.54 cases per 100,000 populations) in 2009. It was estimated that there was total of 300,000-500,000 case world-wide and 30,000 deaths each year (CDC-NCIRD 1).

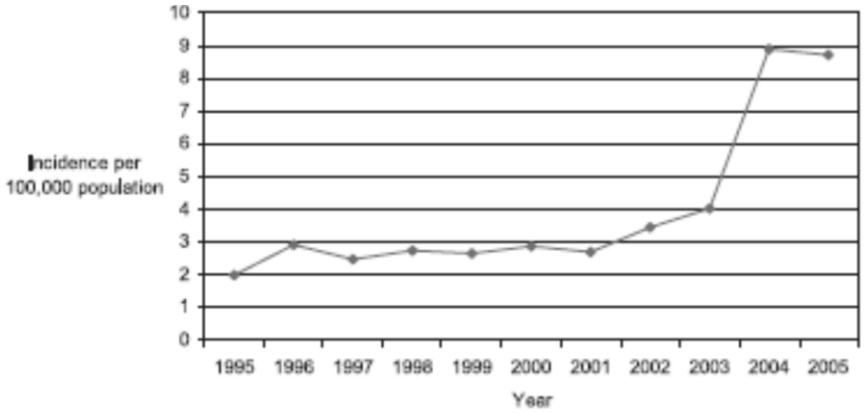


Figure 3.1 Pertussis incidents in US citizens during 1995-2005

World health organization (WHO) statistic in 2010 (2) reveal interesting point of view (**Table 3.1**). The reported case for pertussis were not co-related to vaccination coverage, on the other hand, the higher income country tend to have more reported cases. This indicated the lack of pertussis surveillance in developing country which leads to lower number of reported cases.

Table 3.1 Incident and DTP3 vaccination coverage based on country income group (2)

Country (by income group)	Selected example	Pertussis incident	DTP3 coverage (%)
Lower income	Laos, Cambodia, Ethiopia	7,314	80
Lower middle income	Thailand, Philippines, China	26,770	79
Higher middle income	Brazil, Russia federation, Turkey	10,913	93
High income	Singapore, Japan, South Korea	61,210	95

For Thailand, according to WHO (2) in 2010, the coverage of pertussis vaccination (DTP1/DTP3) is more than 99% of the population. In 2006, report by bureau of epidemiology revealed that the reported cases per 100,000 populations were 0.12. However, this number is much underestimated dues to many factors, such as, most Thais physician presumed that all the infants get the DTP vaccination and thus this disease was hardly found. In fact, effective DTP vaccination needs total of 5 dosages. The first administration is during 2 month-old period, then 4 more boosters until the child reach 4 years old. The vaccination program thus easily misunderstood and/or forgets by many parents. Moreover, the actual geological vaccination coverage especially in the rural province of Thailand is often overlooked. The migration of labor force from neighboring country is also increase the risk of pertussis spreading. There is no clear evidence or survey that indicates the concrete success of the vaccination program.

The recurrence rates of pertussis in many countries are gradually increasing (2,3). Therefore understanding the infection mechanism of the disease from the beginning is very important.

As an air-born disease, during the early infection, *B. pertussis* colonized in human upper respiratory tract and secretes many toxins. One of the important toxins is

adenylate-cyclase hemolysin toxin (CyaA). This toxin was discovered in 1982 by Confer D. and Eaton J. (4) and known for its high activity in converting adenosine triphosphate (ATP) into cyclic adenosine monophosphate (cAMP). Later in 1993 and 1994, Hewlett E. et al. (5) and Hackett M. et al. (6) reveals the importance of palmitoylation on lysine-983 of CyaA by an acyltransferase, CyaC, produced from CyaC gene in Cya Operon. As described by Barry et. al. in 1991 (7), CyaC is a protein in CyaA operon upstream of CyaA with reversed open reading frame. The other 3 protein in this operon are, CyaB, CyaD and CyaE, located downstream on the CyaA gene. These 3 proteins are responsible to the translocation of CyaA by forming type 1 secretion system complex (T1SS) which will be discuss later on.

The CyaA is 1706 amino acids protein contains 2 domains, adenylatecyclase (AC) domain and pore-forming (PF) domain. The first 400 amino acid built up AC domain which responsible for adenylatecyclase activity of the toxin. While the PF domain is further divided into 4 regions which are PF region (residues 401-750), acylation region (residues 751-1006), repeat-in-toxin (RTX) moiety (residues 1007-1657) and secretion signal (residues 1658-1706) (Figure 1.1).

In past 2 decades many researcher studies the structure, function and importance of *B. pertussis* CyaA toxin and other related RTX protein such as *E. coli* Hemolysin A (HlyA).

In 1995, Rose et.al. (8) reveal the calcium binding activity of the RTX region of the CyaA by calcium ion titration with different CyaA mutant. The deconvolution far UV CD spectra of mutant that contain RTX region were shifted between 0.1-2 mM range calcium concentrations. Menestrina G. *et al.* (9) studied the pore formation of *E.coli* HlyA, the smaller CyaA toxin homolog using PC:PS planer lipid bilayer system. The result indicated formation of cation selective channel which prefer K^+ over Na^+. Ludwig *et al.* (10) used 2-D polyacrylamide gel electrophoresis to studied the modification on certain lysine-564 and lysine-690 (equivalent to Lysine-860 and Lysine-983 of CyaA). The results indicated that the acylation on the lysine residue required adjacent flanking region or structure and was required for membrane binding activity.

In 1999, Ladant D. and Ullmann A. (11) proposed mode of action of CyaA toxin during the host cell invasion. The AC domain was translocated across the host cell membrane and interacted with calmodulin to produce overabundant cyclic AMP. Bejerano M. *et al.* (12) studied the AC activity of CyaA mutants and found the residue 1636-1651 at the C-terminal of the toxin were important for the AC activity and can be activated by calcium ion. Lin *et al.* (13) compared the gene organization of 4 different RTX toxin; RtxA operon from *V. cholera*, HlyA operon from *E. coli*, LktA operon from *Pesteurella haemolytica* and CyaA operon from *B. pertussis*. Except the difference it the gene length, the organization of all 4 operon were very similar and suggested the relation of the RTX toxin were not only the protein level but also th gene and organism level.

In 2001, Rhodes C. *et al.* (14) used EPR assay to quantitate the affinity of divalent cation bound to CyaA. The results suggested that the calium binding site in the RTX region were very selective and can't be replaced by outher divalent cation in the normal ion concentration range. Locht *et al.* (15) discussed many virulence factors secreted from B. pertussis. These virulence factors were related through BvgAS regulon which response to the iron concentration. (**Figure 3.2**)

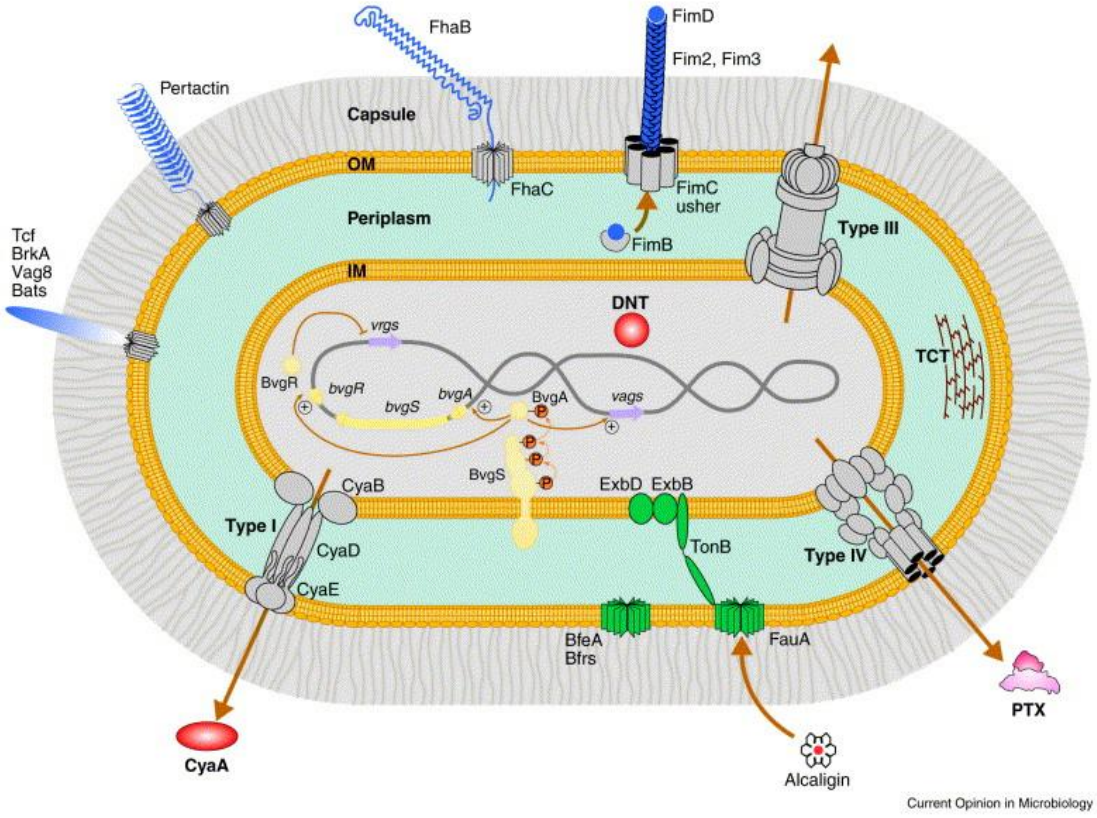


Figure 3.2 CyaA and other virulence factors secreted from *B. pertussis* (15)

CyaA secreted through CyaBDE complex during the colonization of *B. pertussis* along with other toxin.

In 2001 Guermonprez P. *et al.* (16) firstly reported the important of CD11b/CD18 integrin as the CyaA receptor on monocyte. This was confirmed by El-azami-el-idrisso M. *et al.* in 2003 (17), using cell binding assay by many CyaA-FLAG mutants and anti-CD11b antibody. It was concluded that the CyaA used the RTX region to bound to CD11b integrin on target cells. Masin J. *et al.* (18) used the sheep red blood cells and liposome system to characterize the membrane binding activity of CyaA. The results suggested the mode of action of CyaA toward sheep red blood cell was difference from that of liposome. The results further indicated that the acylation is only required for the toxin to interact with the cellular component other than the normal lipid bilayer. Moreover, Knapp O. *et al.* (19) showed that the channel formation of CyaA in liposome was depended on calcium concentration. Benabdelhak H. *et al*. (20) used surface plasmon resonance technique to demonstrate that the C-terminal region of *E.coli* HlyA was responsible for HlyB binding during the membrane translocation thorough type one secretion system complex. In 2004 Gray M. and co-worker (21) show that the binding of CyaA to target cell required newly secreted toxin and close contact.

In 2005 Donato G. *et al.* (22) studied CyaA and CyaC from *B. Hinzii.* The result indicated that AC activity independent from CyaC and may related to the differences between host immune system.

In 2007 Basler M. *et al.* (23) Show the important of helical hairpin structure and three glutamic acid residues; E-516, E-570, E-581. Substitution of the E-570 residue by site-directed mutagenesis abolished the hemolytic activity while the other two mutations become super hemolytic. This indicated the location and ion-selective ability of the two residues.

CHAPTER IV
MATERIAL AND METHODS

4.1 Materials

4.1.1 Chemicals and reagents

Ampicillin	Sigma
Chloramphenicol	Sigma
Cetyl trimethyl ammonium bromide (CTAB)	Sigma
RNaseA	Sigma
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Sigma
1, 4-Dithiothreitol (DTT)	Sigma
Coomassie brilliant blue R-250	Sigma
Phenylmethylsulphonylfluoride (PMSF)	Sigma
p-Nitrophenyl palmitate (pNPP)	Sigma
Deoxyribonucleotide triphosphates (dNTPs)	Promega
Standard DNA markers	Gibco BRL, New England Biolabs
SDS-PAGE molecular mass standards	Bio-RAD
Bradford protein assay reagent	Bio-RAD

All other unlisted chemicals and reagents were analytical grade purchased from various suppliers.

4.1.2 Enzymes and accessory buffers

Pfu DNA polymerase (cloned)	Promega
T_4 DNA Ligase	Promega
Restriction endonucleases	New England Biolabs/Promega

All restriction endonuclease enzymes and their buffers were commercially provided by specified companies.

4.1.3 Plasmid vector

The recombinant plasmids, pCyaAC-PF expressing CyaA-PF and CyaC and pCyaA-PF (without CyaC), which have been cloned from the *B. pertussis* DMST15659 Thai isolate (24) were used as a DNA template for site-directed mutagenesis and protein expression.

4.1.4 Bacterial strains and

E. coli, strain JM109 [*endA1*, *recA1*, *gyrA96*, *thi*, *hsdR17*(r_k^-, m_k^+), *relA1*, *supE44*, Δ*(lac-pro AB)*, (F' *traD36 proAB lacI*^{q}Δ(lacZ)M15)] was purchased from Promega.

E. coli, strain BL21(DE3)pLysS [F^-, *ompT*, *hsdS*$_B$(r_B^-, m_B^-), *dcm*, *gal*, λ(DE3), pLysS, Cm^r] was purchased from Promega.

E. coli strain JM109 was used for general DNA manipulation whereas *E. coli* strain BL21(DE3)pLysS was used for expression of the recombinant plasmids.

4.1.5 Sheep red blood cells

Sheep erythrocytes drawn in Alsever's solution were purchased from National Laboratory Animal Center, Mahidol University.

4.1.6 Synthetic oligonucleotide primer

Mutagenic and sequencing oligonucleotide primers were purchased from Sigma PROLIGO, Singapore. The mutagenic primer sequences showed below.

Introduction of restriction site, *Nde*I (underlined), by point mutation (bold.)

```
Amino acid              -  -  -  -  M  A  N  S
Forward primer             5′-GAACATATGGCCAATTCG-3′
Reverse primer      3′-CGTCACGAACATATGGC-5′
```

4.1.7 Culture media

For each litre of Luria-Bertani (LB) medium, dissolved 10 g peptone, 5 g yeast extract and 10 g NaCl in distill water. For LB agar, additional 10 g of bacto-agar was added. After the content is completely dissolved adjust the volume to 1 litre. pH of the medium tested by universal pH paper was around 6-7. The medium was sterilized by autoclaving at 121˚C, 15 psi for 15 min. After cooled down,

ampicilin100µg/ml final concentration was aseptically added to media used for expression. For selective media used to culture stock *E.coli* BL21 DE3 pLysS, ampicilin 100 µg/ml and choramphenicol 34 µg/ml final concentration were supplemented.

4.2 Methods

Construction of pCyaAC-RTX plasmid co-expressing CyaA-RTX (residue 751-1706) and CyaC, acyltransferase, was done using pCyaAC-PF as a template. Then *Nde*I site was introduced to the plasmid by site-directed mutagenesis. Plasmid was verified by restriction endonuclease pattern and DNA sequencing. *E.coli* BL21(DE3)pLysS was selected as expression system. The recombinant protein, CyaA-RTX, was characterized by various biochemical methods including SDS-PAGE, LC/MS/MS and hemolysis assay. *In silico* study was also done to improve understanding on the structure and functions basis of the protein.

4.2.1 Plasmid DNA preparation by alkaline lysis method

Desired clone either *E. coli* JM109 or *E. coli* BL21(DE3)pLysS was culture at 37°C overnight. The cells were harvested by centrifugation at 5000 rpm for 1 min. Supernatant was removed. The cell pellet was resuspended in 100µl cold solution I (25 mM Tris-HCl pH 8.0, 10mM EDTA). Then 200µl of solution II (0.2M NaOH, 1%(v/v) SDS) was add into the cell suspension. The tube was inverted a few time and leaved in ice bath for 2 minute. Then 150µl of solution III (5M potassium acetate-acetic acid pH 5.5) was added. The tube was inverted several times before centrifugation at 12,000 rpm. for 5 minutes. The supernatant was transfer into the new tube. Plasmid DNA was precipitated with cold isopropanol for 15 minutes followed by centrifugation at 12,000 rpm. for 15 minutes. The DNA pellet was washed several times with cold 70% ethanol before allowed to dry for 10 minutes. The pellet was resuspended in double distilled water.

4.2.2 Agarose gel electrophoresis of DNA

Agarose gel electrophoresis was prepared by dissolved desired amount of agarose powder corresponding to 0.8% or 1% gel with 1x TAE buffer (40 mM Tris-Cl pH 8.0, 40 mM acetic acid and 2.5 mM EDTA) then heating in microwave oven with maximum power for 30-60 sec. The gel solution was left for cooling down before pour into the tray. The solidified gel was submerged in TAE buffer in agarose gel running chamber. Diluted DNA sample was mixed with loading dye (15% (w/v) ficoll 400, 0.1% bromophenol blue, 5 mM EDTA) at 5:1 ratio and load in to a well in agarose gel. λ/*Hind*III and/or λ/*Bst*EII molecular standard with known concentration were used as internal base pair marker. After electrophoresis, the gel was soaked in 2µg/ml ethidium bromide solution for 3 min and de-stained in distill water for 15-30 min. The DNA pattern was visualized under UV light using Geldoc® system. Amount of DNA was estimated by comparing the band with the internal base pair marker.

4.2.3 Restriction endonuclase digestion of plasmid DNA

The restriction endonuclease digestion of plasmid DNA was performed by mixing 2 µl of 10x restriction endonuclease buffer, 5-10 U of restriction endonuclease, plasmid DNA and segmentally added the double distilled water to a final volume of 20 µl. The reaction was carried out at the recommended temperature and time for each endonuclease. The digested DNA patterns were analyzed by agarose gel electrophoresis.

4.2.4 Construction of recombinant pCyaAC-RTX

The retraction endonuclease site was introduced into the pCyaAC-PF using Stratagene's QuickChange™ Site-directed mutagenesis kit (**Figure 4.1**). The mutagenic primer (from section 4.1.5) was used as primer in PCR reaction. The PCR reaction contains; 100 ng od DNA template (pCyaAC-PF), 200 µM of dNTPs, 10 pmol of each forward and reverse primer, 5 µl 10x *Phusion* buffer and 0.5 µl *Phusion* DNA polymerase. The volume was adjusted to 50 µl with double distilled water.

The temperature cycle for PCR reaction is showed in Table 4.1. After the PCR reaction, the plasmid product was digested with *Dpn*I to remove the DNA template by adding 1 µl of *Dpn*I then incubate at 37°C for 2 h. The digested products

were analyzed by 0.8% agarose gel electrophoresis before being transformed into *E. coli* competent cells.

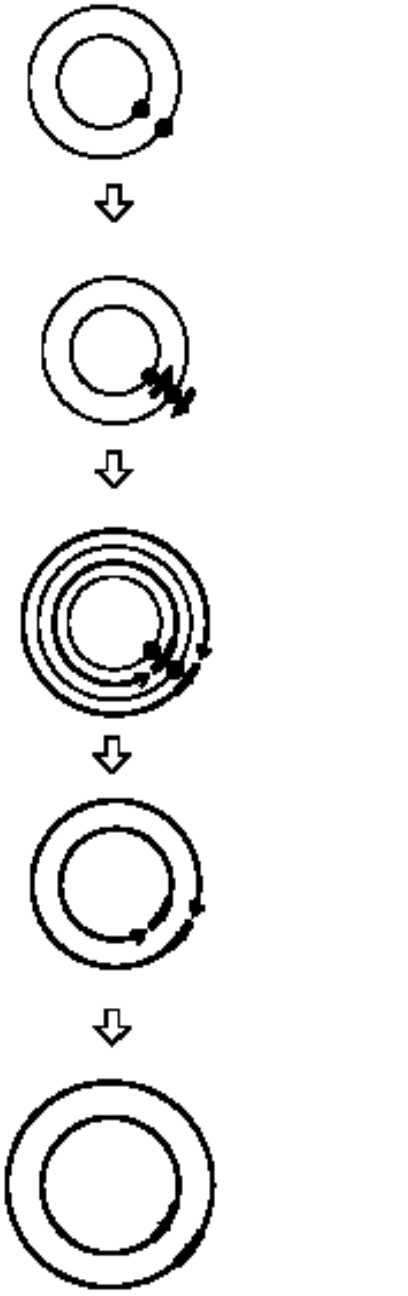


Figure 4.1 QuickChange site-directed mutagenesis technique

Overview of QuickChange site-directed mutagenesis method (adopted from Stratagene's QuickChange instruction manual).

Table 4.1 Temperature cycling parameters for site-directed mutagenesis

	Temperature (°C)	Time (min:sec)	Cycles
Pre-heat	98	0:30	1
Denaturing	98	0:10	25
Annealing	56	0:30	
Extension	72	2:30	
Final extension	72	7:00	1

4.2.5 Preparation of competent cells by $CaCl_2$ method

Desired strain of *E.coli* cells was cultured overnight at 37°C in LB broth then inoculated at 100x dilution into fresh LB broth and cultured for 2 more h. The culture was put into ice bath for 10 min and harvested buy centrifugation at 3000 rpm, 4°C for 10 min. The cell pellet was resuspended in 100 mM cold $CaCl_2$ solution. After left in ice bath for 5 min, the cells suspension was centrifugation at 3000 rpm, 4°C for 10 min. The cell pellet was resuspended again in 100 mM cold $CaCl_2$ solution and left in ice bath for 1 h. Glycerol was added to the cell suspension to 20% (v/v) final concentration. The cell suspension was aliquoted (200 µl each) and kept at -80°C.

4.2.6 Transformation of plasmid into competent cells

One ng of plasmid DNA or 50 ng of PCR product was added into an aliquot of competent cells. The suspension was mix and left on-ice for 30min. The cell suspension was put into heat shock in 42°C water bath for 5 min then immediately adds 800 µl of LB broth and incubates at 37°C for with shaking. After 1 h, the transformed cell was spred on LB agar plate containing 100 µg/ml ampicillin (for *E. coli* JM109) or 100 µg/ml ampicillin and 34 µg/ml chloramphenicol (for *E. coli* BL21(DE3)pLysS) and then incubated at 37°C overnight. The recombinant clones were verified by restriction analysis

4.2.7 Protein expression

Starter plate was created by immediately streak from -80˚C glycerol stock onto LB agar supplement with ampiclin and choramphenicol. The plate was incubated at 37˚C overnight (16-24 h) then a single colony was selected and picked into 3 ml LB broth in 15 ml centrifugal tube. The starter was incubated for 12 h at 37˚C before inoculated at ratio 1:100 into expression media containing only ampicilin. The culture was incubate at 25˚C or 30˚C until OD_{600} reach 0.5-0.6, IPTG was added to the final concentration of 0.1 mM then the culture was leaved at same condition for 6 h. before harvested by centrifugation at 6000xg, 4˚C for 15 min. Pellet was washed twice with either RBC buffer (20 mM Tris-Cl pH 7.4, 5 mM $CaCl_2$ and 125 mM NaCl) or Tris-Ca buffer (20 mM Tris-Cl pH 8.0 and 5 mM $CaCl_2$). The supernatant was discarded; pellet was kept at 4˚C for further experiment.

4.2.8 Sodium dodecyl sulfate-polyacrylamides gel-electrophoresis analysis (SDS-PAGE)

The protein sample was prepared by mixing the sample (protein solution or cell pellets) with 4× loading buffer (60 mM Tris-Cl pH 7.5, 2% SDS, 10% glycerol, 0.025% bromophenol blue, 100 mM DTT) double distilled water was added to the total volume of 80µl if need and boiled at 100°C for 5 min. The sample was vigorously vortex until homogeneous and centrifuge at 12,000×*g* for 10 min. Supernatant was used to load into the SDS-PAGE gel.

Electrophoresis of protein was performed by method described by Laemmli. SDS-PAGE was carried out using the Bio-Rad Mini Protein II system. The separating gel contain 10% or 12% acrylamide with 3% cross-linker (bis-acrylamide), 375 mM Tris-Cl pH 8.8, 0.1% SDS, 0.1% ammonium persulfate and 0.04% TEMED. The stacking gel composed of 5% acrylamide with 3% cross-linker, 25 mM Tris-Cl pH 6.8, 0.1% SDS, 0.1% ammonium persulfate and 0.04% TEMED. The gel was run in Tris-glycine buffer (25mM Tris, 192 mM Glycine and 0.1%SDS). Board range protein marker (Bio-Rad) was used as an internal protein molecular weight standard. Electrophoresis was performed at room temperature with constant current at 26 mA or 52 mA for 1 gel or 2 gels, respectively. To visualize the protein band, the gel was soaked in staining solution (0.1% Coomassie brilight blue R-250, 50% ethanol and 10% glacial acetic acid in distill water) After 1 h, the was transferred into destaining solution (10% ethanol and 10% acetic acid in distill water) and left rotated until background was cleared.

4.2.9 Protein concentration assay by Bradfords method

Protein concentration was determined by Bradford method using Bio-Rad protein assay reagent kit (Bio-Rad, USA). The calibration curve was prepared from a standard protein, BSA. The BSA standards were prepared by diluted the standard protein to obtain 0, 1, 2, 3, 4, 5 mg/ml standard protein solutions. The 50 µl of each dilution of standard protein solution was mixed with 200 µl of Bradford reagent dyes. The solutions were mixed in a 96-well flat bottom microtiter plate. The absorbance of the resulting solutions was measured at 545 nm (Spectromax 250 ELISA plate reader with internal software; Molecular Devices, Sunnyvale, CA). The 10-50 µl of sample

protein were measured in the same manner and calculated protein concentration compared with standard protein.

4.2.10 Protein preparation from *E.coli* culture

4.2.10.1 Protein was harvested from packed *E.coli* by French press. For large scale protein expression, >100 OD_{600} of packed cell, French press approach was selected to disturb bacterial cell because the method have high yield and suitable for large volume of the sample. Firstly, packed cell was resuspended in 20 ml of RBC buffer (20 mM Tris-Cl pH 7.4, 5 mM $CaCl_2$ and 125 mM NaCl). The protease inhibitor, PMSF was supplement in the buffer. Cell suspension was subjected into cold French press cell chamber and operate at 10,000 psi for at least three times. The whole cell lysate was then centrifuged at 8,000 rpm, 4˚C for 15 min. The supernatant was carefully transferred to new tube and refer as crude lysate. The crude lysate was kept at 4˚C for next experiment.

4.2.10.2 Protein was harvested from packed *E.coli* by sonication. For small scale protein expression, <100 OD_{600} of packed cell, sonication technique was selected to break bacterial cell because the approach is quick, simple and suitable for small volume of the sample. Initially, packed cell was resuspended in 5 ml of RBC buffer (20 mM Tris-Cl pH 7.4, 5 mM $CaCl_2$ and 125 mM NaCl). The protease inhibitor, PMSF was supplement in the buffer. Cell suspension was retained in ice bath and sonicated using manually config profile. The profile was set to 1 sec on, 2 sec off for 3 min with 15% power. The resulting solution should be clear with trace amount of foam. The solution was then centrifuged at 8,000 rpm, 4˚C for 15 min. The supernatant was carefully transferred to new tube and refer as crude lysate. The crude lysate was kept at 4˚C for further experiment.

4.2.11 Sheep red blood cell preparation

Sheep red blood cell (sRBC) was maintained in 1:1 (V/V) Alsever's solution at 4˚C prior to use. The sRBC was prepared by adding RBC buffer (20mMTris-Cl pH 7.4, 5 mM CaCl2, 125 mMNaCl) to the total volume of 1 ml. The washing step was done by gently inverts the tube several times. The sRBC was centrifuge at 2000×g for 5 min at room temperature (25-28˚C). Supernatant was

removed by pipetting and new RBC buffer was added. The sRBC was washed for 3-5 times until the supernatant was clear. The sRBC was resuspended in the same volume as drawn from the bottle. The resulting sRBC suspension was referred as stock sRBC.

To determine amount of sRBC used for the assay; 10, 20, 30, 40, 50 µl stock sRBC was diluted in RBC buffer. Triton X-100 was added to each suspension to the final concentration of 0.1% (V/V). Then 200 µl of each dilution was subjected to microtiter plate and the absorbance at 540 nm was read. Standard curve was ploted and the amount of sRBC suspension used was the volume that gives OD_{540} around 0.8.

4.2.12 Hemolytic activity assay

Hemolytic assay was done based on method modified from Busaba (25). In brief, soluble protein from *E.coli* lysate from section 4.2.10 was diluted to the desired concentration with RBC buffer. Then the volume was adjusted to 980 µl before addition of 20 µl sRBC. The reactions were gently mixed by inverting several times and incubate at 37˚C for 6 h. At the end of the incubation, the reaction was centrifuged at 12,000 rpm for 2 min to remove unlysed RBC and cell debris. 200 µl of supernatant was subject to flat bottom 96 wells microtiter plate and measure OD_{540}

4.2.12.1 Enhance hemolysis test was performed based on hypothesis that the toxin form a pre-pore oligomerization before attacked the sRBC cell membrane. Hence, although CyaA-RTX is lack of the PF sub-domain, the C-terminal part may be, as minimal requirement, enough for the oligomerization of the toxin. These oligomerized toxins might have ability to hemolyzed sRBC comparable to that of CyaA-PF homo-oligomer. Thus incubation of the active CyaA-PF toxin with CyaA-RTX may enhance the hemolysis activity of the toxin in dose dependent manner. On the other hand, if the CyaA-RTX can oligomerized with the active CyaA-PF, the oligomerized product may malfunction due to lack of PF-sub domain of the CyaA-RTX. So incubating the active CyaA-PF toxin with CyaA-RTX may abolish the hemolysis activity of the toxin in dose dependent manner. Protein solution was prepared by mixed either acylated CyaA-PF or non-acylated CyaA-PF with CyaA-RTX and adjust the volumn to 980 µl. The protein solution had final concentration of 1, 2 or 3 mg/ml **(Table 4.2)**. The mixture was left at room temperature for 18 min.

Subsequently 20 µl sRBC was added to the solution. The solution was kept at 37°C for 6 h, before measure the OD_{540} as describe above.

4.2.12.2 In case of competitive hemolysis assay, it is hypothesized that the truncated toxin may inhibit toxin binding to target cell membrane by prior bound to the target molecule on sRBC. Competitive protein solution (CyaA-RTX, CyaAC-RTX) was incubated with 20 µl of sRBC for 18 min at room temperature before addition of active toxin (CyaAC-PF) **(Table 4.3)**. The solution was kept at 37°C for 6 h, before measure the OD_{540} as describe above.

4.2.13 Mass spectrometric analysis

SDS-PAGE of the soluble cell lysate was performed according to section 4.2.8. The band at corresponds molecular weight was excised and analyzed by LC/MS/MS.

Table 4.2 Enhanced hemolysis assay

	Acylated CyaA-PF	Enhancing protein	Remarks
(1) 100% hemolysis	-	-	Add 1µl of triton X-100 before measure OD_{540}
(2) Negative control (0% hemolysis)	-	-	
(3) Positive control	1 mg/ml	-	Final protein concentration is 1 mg/ml
(4) Positive control	3 mg/ml	-	Final protein concentration is 3 mg/ml
(5) Negative control (for enhanced hemolysis)	-	2 mg/ml CyaA-RTX	Final protein concentration is 2 mg/ml
(6) Negative control (for enhanced hemolysis)	-	2 mg/ml CyaA-RTX co-expressed with CyaC	Final protein concentration is 2 mg/ml
(7) Enhanced hemolysis	1 mg/ml	2 mg/ml CyaA-RTX	Final protein concentration is 3 mg/ml
(8) Enhanced hemolysis	1 mg/ml	2 mg/ml CyaA-RTX co-expressed with CyaC	Final protein concentration is 3 mg/ml

Table 4.3 Competitive hemolysis assay

	Acylated CyaA-PF	Competitive protein	Remarks
(1) 100% hemolysis	-	-	Add 1µl of 10% (V/V) triton X-100 before measure OD_{540}
(2) Negative control (0% hemolysis)	-	-	
(3) Positive control	1 mg/ml	-	Final protein concentration is 1 mg/ml
(4) Positive control	3 mg/ml	-	Final protein concentration is 3 mg/ml
(5) Negative control (for competitive hemolysis)	-	2 mg/ml CyaA-RTX	Final protein concentration is 2 mg/ml
(6) Negative control (for competitive hemolysis)	-	2 mg/ml CyaA-RTX co-expressed with CyaC	Final protein concentration is 2 mg/ml
(7) Competitive hemolysis	1 mg/ml	2 mg/ml CyaA-RTX	Final protein concentration is 3 mg/ml
(8) Competitive hemolysis	1 mg/ml	2 mg/ml CyaA-RTX co-expressed with CyaC	Final protein concentration is 3 mg/ml

4.2.14 Molecular modeling of CyaA-RTX fragment

Several short (150 - 300 residues) CyaA-RTX fragments (residues 751-1706) were subjected to server-based template identifying tools (http://swissmodel.expasy.org). The several candidates were selected subjected to pairwise alignments. Among the templates, I.3lipase from *pseudomonas sp.* (LipA) gave highest score and was selected as a model template. Several pairwise alignments which have overlapped region of around 9-27 amino acid were performed. All of them were subjected into SWISS-MODEL web-based molecular modeling server. The overlapped models that have best molecular parameter were joined together and energy minimized. The final model was evaluated by PRO-CHECK software. Summary of CyaA-RTX molecular model construction is showed in **Figure 4.2**.

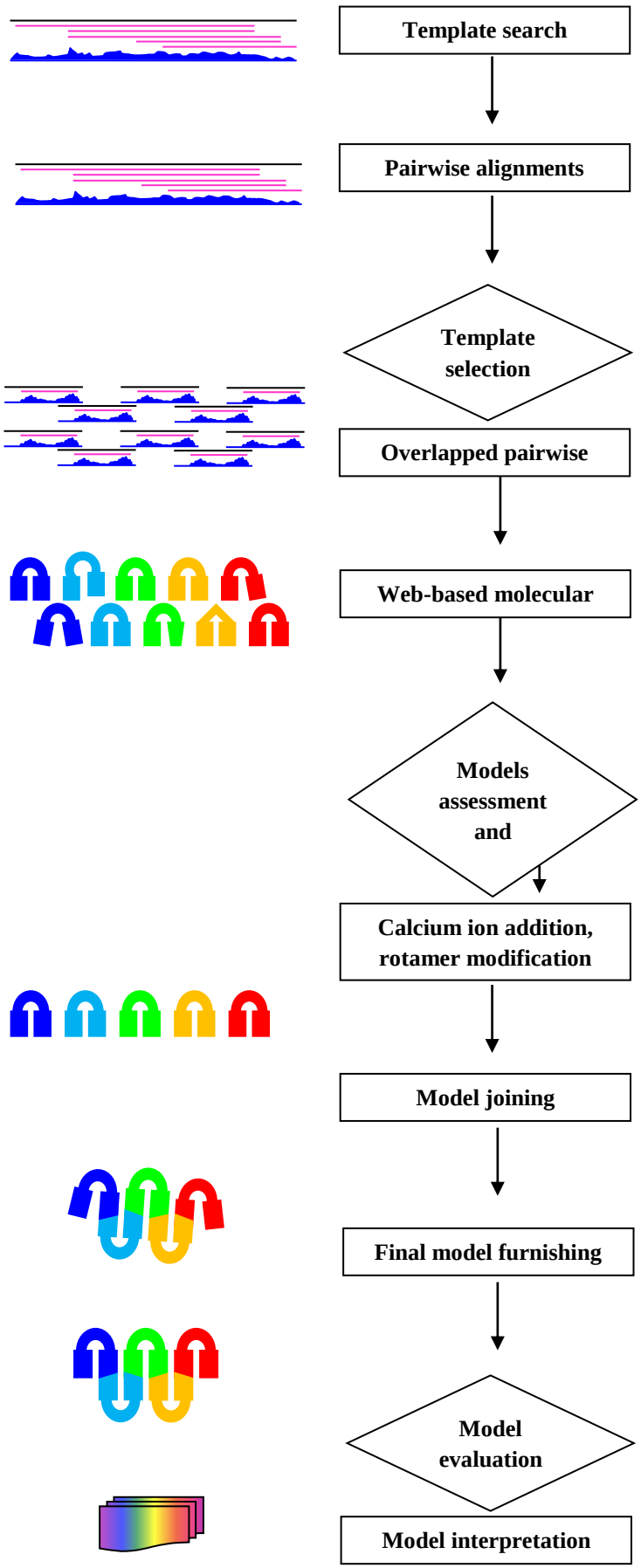


Figure 4.2 Molecular modeling protocols

CHAPTER V
In vitro RESULTS:
EXPRESSION AND FUNCTIONAL CHARACTERIZATION OF THE CyaA-RTX SUBDOMAIN FRAGMENT

Recent study has shown that CyaA-PF can be palmitoylated by CyaC (22). Here, the pCyaAC-RTX co-expressing CyaA-RTX and CyaC was constructed to study the function of CyaA-RTX compared to the CyaA-PF

5.1 Construction of pCyaAC-RTX plasmid

Contruction of pCyaAC-RTX plasmid co-expressing CyaA-RTX (residue 751-1706) and CyaC, acyltransferase, was done using pCyaAC-PF (24) as a template. *Nde*I site was introduced into the plasmid by site-directed mutagenesis. The pCyaAC-RTX was obtained by partial digestion with *Nde*I, selection of a derived band and re-ligation.

5.1.1 Plasmid preparation of pCyaAC-PF template

Plasmid DNA (pCyaAC-PF) was extracted and purified. The purifies plasmid shows three bands with size corresponding to linear (≈7.5 kb), super-coiled (≈5.0 kb) and circular conformation (≈9.4 kb). The plasmid was verified by double digestion with two restriction emzymes. The band patterns are found to correspond with expected sizes of the digested fragments. The verified plasmid was further used as template for PCR based mutagenesis.

5.1.2 Generation of an added *Nde*I site in pCyaAC-PF

Purified pCyaAC-PF was used as template for mutation by quick-change mutagenesis method. The resulting mutant PCR product which was digested with *Dpn*I shows a single band with size corresponding to 7.5 kb. The treated plasmid was

transformed into *E. coli* JM109. The extracted plasmids from resulting colonies were checked by restriction enzyme digestion, showing the correct band pattern.

5.1.3 Partial digestion of pCyaAC-PF with *Nde*I

pCyaAC-PF was partially digested with *Nde*I. The band pattern was found to correspond with each size of the digested products **(Figure 5.1)**. The desired band (≈6.8 kp) was cut from the gel for DNA purification.

5.1.5 Ligation of 6.8 kb fragment

The purified 6.8 kb fragment was subjected to several different ligation reactions. The ligated product was transformed into *E. coli* B21(DE3) pLysS and selected on selection media, LB-ampicillin. Three clones were obtained that show corrected size plasmid on agarose gel.

5.1.6 DNA sequencing of pCyaAC-RTX

Plasmid DNA from transformed *E. coli* B21(DE3) pLysS was sequenced for verifying correct sequence of CyaA-RTX (residues 751-1706) and full-length CyaC genes.

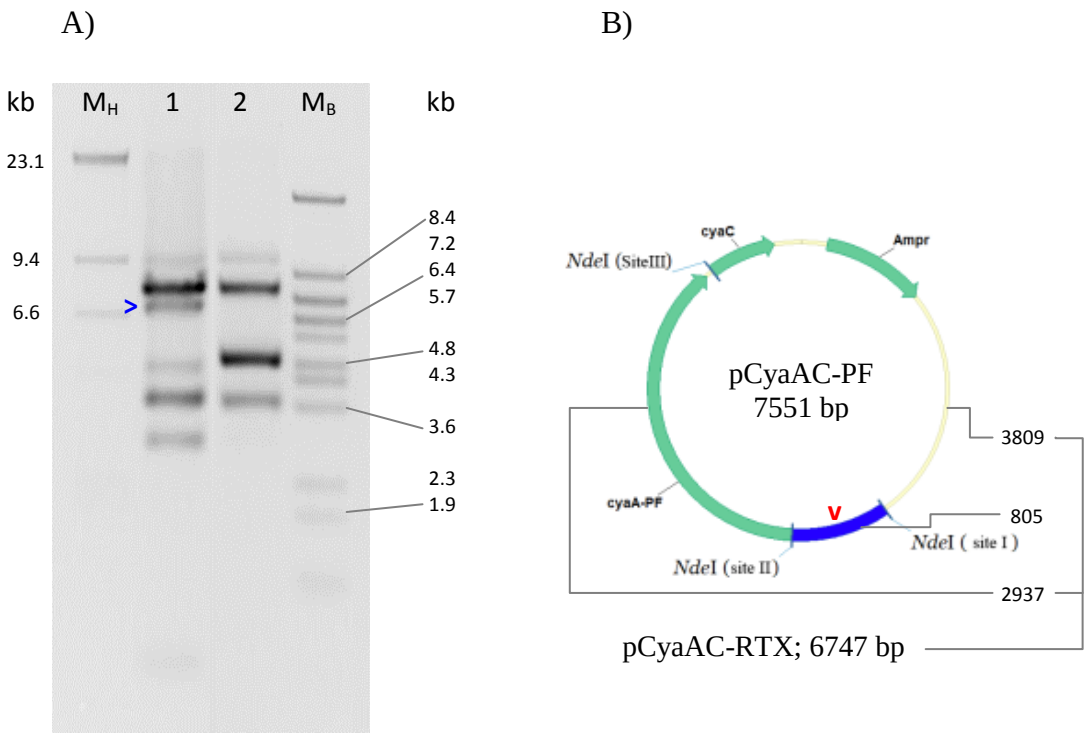


Figure 5.1 Partially *Nde*I digested pCyaAC-PF for pCyaAC-RTX construction

(A) PCR product of pCyaAC-PF with added *Nde*I before (lane 2) and after (lane 1) partially digested with *Nde*I. The fragment with 6.8 kb size (>) was excised and used for purification and transformation. M_H is λ/*Hind*III molecular mass standard. M_B is λ/*Bst*EII molecular mass standard.

(B) Plasmid map of pCyaAC-PF with an added *Nde*I site (v). The PF region is highlighted in blue

5.2 Co-expression of CyaA-RTX with CyaC from pCyaAC-RTX

Recombinant *E.coli* BL21(DE3)pLysS cells harboring pCyaAC-PF, pCyaA-RTX or pCyaAC-RTX were used to produce each protein of interest under IPTG induction. SDS-PAGE gel of IPTG-induced total cell lysate shows bands coresponding to CyaA-PF, CyaA-RTX and CyaC **(Figure 5.2)**. Band of CyaA-PF shows molecular mass around ~150 kDa while that of CyaA-RTX is around ~125 kDa. Both recombinant cells that carrying CyaC gene (pCyaAC-PF and pCyaAC-RTX) have a ~25 kDa band which corresponds to CyaC. All three recombinant (CyaA-PF, CyaA-RTX and CyaC) bands were absent from total cell lysate of *E.coli* carrying only pET17b.

5.2.1 Effect of temperature on CyaA-RTX and CyaC production

Productions of both CyaA-RTX and CyaC from pCyaAC-RTX at 30˚C and 37˚C show comparable amounts of protein. Protein quantities at 0, 2, and 4 hr after induction show increasing trend and become more steady around 6 hr of induction **(Figure. 5.3)**.

5.2.2 Solubility of CyaA-PF, CyaA-RTX and CyaC

All three proteins were detectable in the soluble fraction after cell disruption by sonication **(Figure 5.4)**.

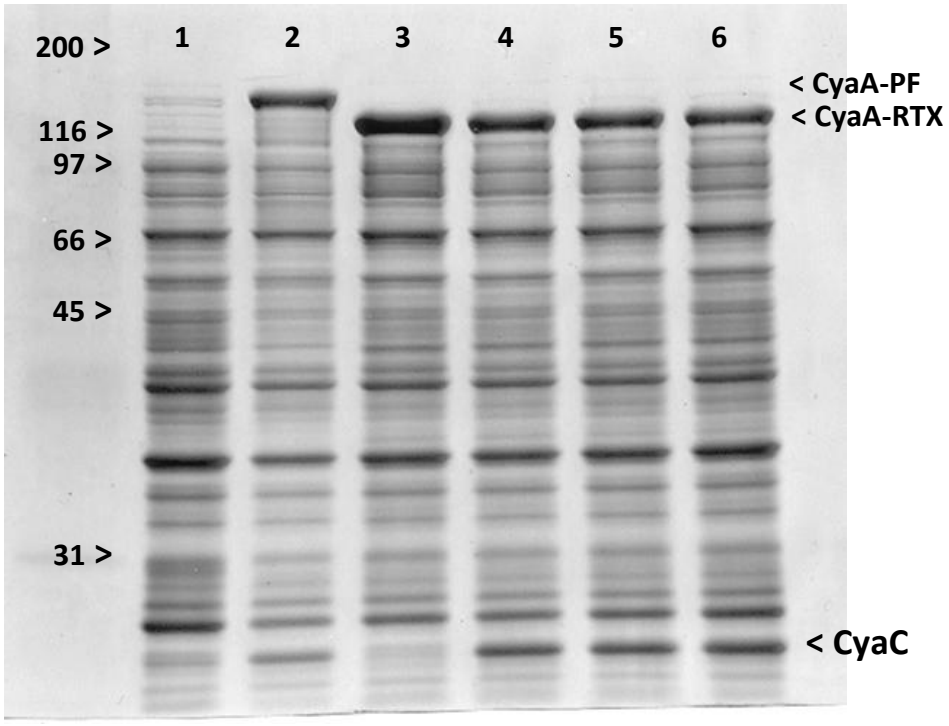


Figure 5.2 Expression profile of pCyaAC-RTX

Coomasie brilliant blue-stained SDS-PAGE (10% gel) of *E.coli* cell lysates expressing proteins from the indicated plasmids.

Lane 1: pET17b

Lane 2: pCyaAC-PF, co-expressed CyaA-PF with CyaC

Lane 3: pCyaA-RTX, produced only CyaA-RTX

Lane 4-6: different clones of pCyaAC-RTX, expressing both CyaA-RTX and CyaC

Numbers on the left indicate molecular mass standard (kDa).

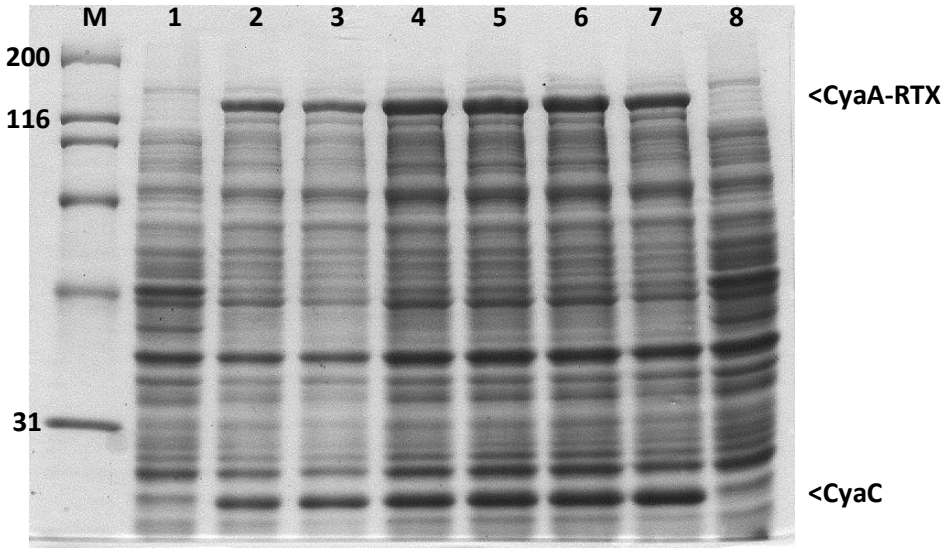


Figure 5.3 Effect on culturing temperature and time of induction

Coomasie brilliant blue-stained SDS-PAGE (10% gel) of total *E.coli* lysate containing pET17b or pCyaAC-RTX growth at different temperatures with various times of induction.

Lane 1: total cell lysate of *E.coli* containing pET17b, cultured at 30˚C with 6-hr IPTG induction

Lane 2: total cell lysate of *E.coli* containing pCyaAC-RTX, cultured at 30˚C for 2 hr after IPTG induction

Lane 3: total cell lysate of *E.coli* containing pCyaAC-RTX, cultured at 37˚C for 2 hr after IPTG induction

Lane 4: total cell lysate of *E.coli* containing pCyaAC-RTX, cultured at 30˚C for 4 hr after IPTG induction

Lane 5: total cell lysate of *E.coli* containing pCyaAC-RTX, cultured at 37˚C for 4 hr after IPTG induction

Lane 6: total cell lysate of *E.coli* containing pCyaAC-RTX, cultured at 30˚C for 6 hr after IPTG induction

Lane 7: total cell lysate of *E.coli* containing pCyaAC-RTX, cultured at 37˚C for 6 hr after IPTG induction

Lane 8: total cell lysate of *E.coli* containing pCyaAC-RTX, cultured at 30˚C for 6 hr without IPTG induction

Lane M: Board range molecular mass standards.

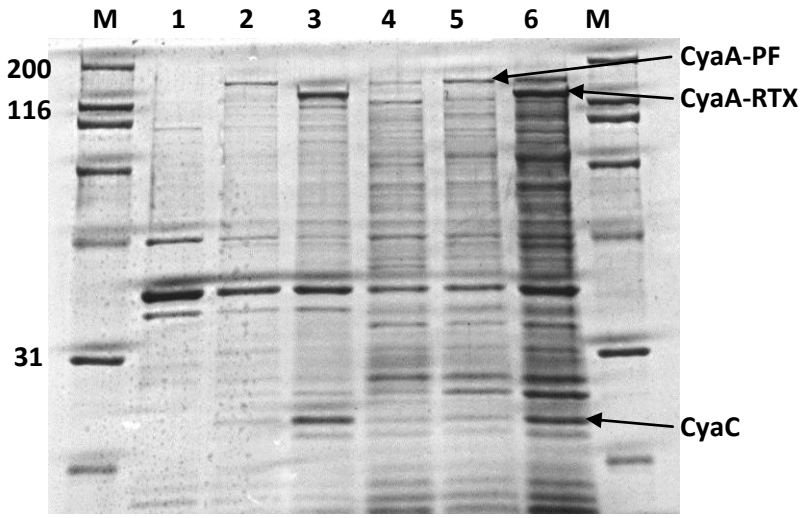


Figure 5.4 Solubility of CyaA-PF, CyaA-RTX and CyaC

Coomassie brilliant blue-stained SDS-PAGE gel of insoluble pellet and soluble fraction from *E.coli* cell lysates carrying pET17b, pCyaAC-PF or pCyaAC-RTX.

Lane 1: insoluble pellet of *E.coli* harboring pET17b

Lane 2: insoluble pellet of *E.coli* harboring pCyaAC-PF

Lane 3: insoluble pellet of *E.coli* harboring pCyaAC-RTX

Lane 4: soluble fraction of *E.coli* harboring pET17b

Lane 5: soluble fraction of *E.coli* harboring pCyaAC-PF

Lane 6: soluble fraction of *E.coli* harboring pCyaAC-RTX

Lane M: Board range molecular mass standards (kDa)

5.3 *In vitro* activity of CyaC from pCyaAC-RTX

Hydrolysis activity of *para*-nitrophenylpalmitate (pNPP) by CyaC from pCyaAC-PF and pCyaAC-RTX was tested using spectroscopic method. The efficiency of CyaC was monitored for at least 300 sec. The graphs were nearly linear with little decrease in slope after 60 sec. The slope of the graphs were calculated at 0-60 sec **(Figure 5.5)**. At this high substrate concentration, the initial velocity was near the maximum velocity. Specific activities were computed as 1 mg of enzyme that liberated 1 μmole of *p*-nitrophenol in 1 min at 25˚C. The soluble CyaC enzyme from both clones showed similar specific activity that is comparable to the refolded wild-type CyaC **(Table 5.1)**.

5.4 Acylation at K983 on CyaA protein from pCyaAC-RTX

The coomasie blue staining shows a thick band of CyaA-RTX overexpressed from plasmid CyaAC-RTX. This protein bands was cut for LC/MS/MS and MALDI-TOF MS analysis along with the CyaAC-PF band as positive control for *in vivo* acylation at K^{983}. Results from mass-spectrometry show non-acylated peak with m/z ratio of 613.31 **(Figure 5.6)**.

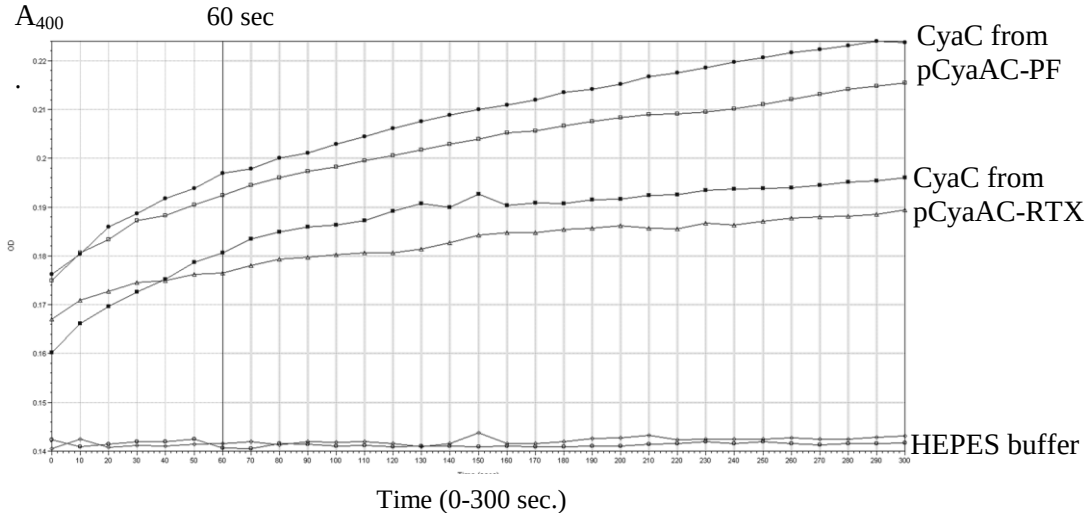


Figure 5.5 Activity of CyaC from pCyaAC-PF and pCyaAC-RTX

pNPP hydrolysis activity of CyaC from pCyaAC-PF and pCyaAC-RTX. Slope of the first min reaction was used for specific activity calculation. However, the reactions were very fast which might give underestimated results. The figure shows one example of the three duplicates.

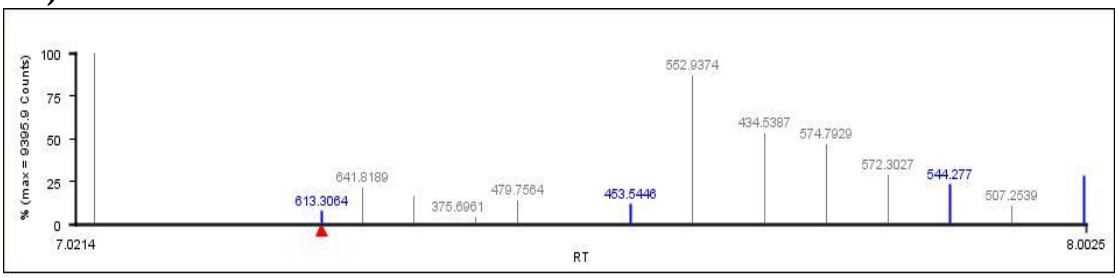
Table 5.1 Specific activity of CyaC for hydrolysis of pNPP

Enzyme source	**Specific activity*** (U/mg protein ± SEM)
Soluble CyaC from pCyaAC-PF	248.0 ± 42.2
Soluble CyaC from pCyaAC-RTX	223.6 ± 15.1
Refolded CyaC from pCyaC**	289.4 ± 1.2
Bovine serum albumin**	4.2 ± 1.1

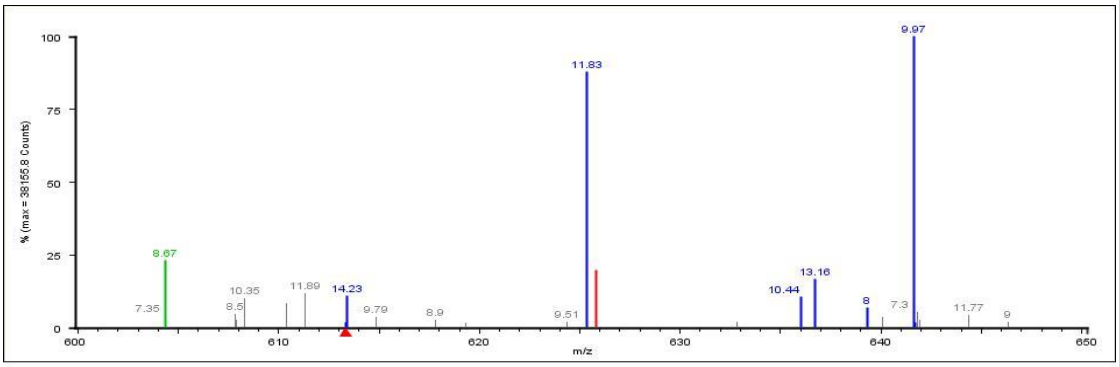
* One unit (U) of enzyme was defined as amount of CyaC liberating 1μmol of *p*-nitrophenol per min at 25˚C. SEM represents standard errors of the mean from three independent wxperiments.

** From Thamwiriyasati *et.al.* (26)

A) MALDI-TOF MS



B) LC/MS/MS (precursor ion)



C) LC/MS/MS (product ion)

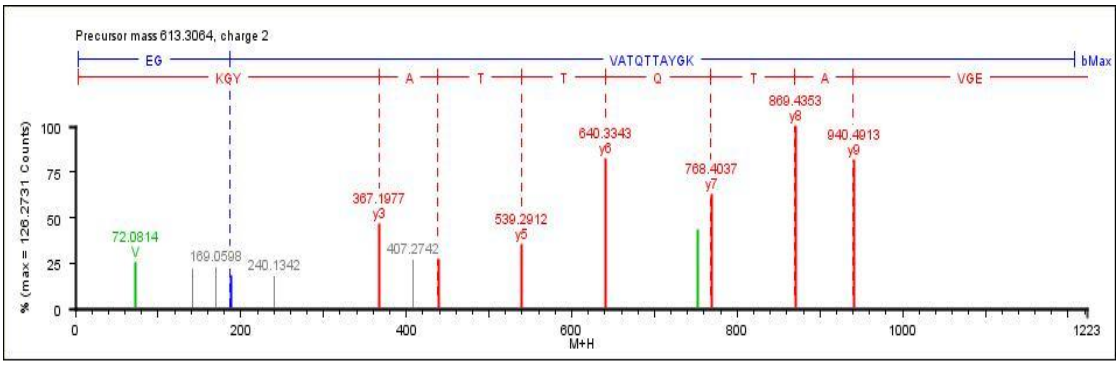


Figure 5.6 Mass spectrometry spectra of the non-acylated acylation site K^{983} of CyaA-RTX

Spectrum representing non-acylated fragment from two methods, matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) and liquid chromatography-tandem mass spectrometry (LC/MS/MS). (A) MALDI-TOF MS and (B) LC/MS/MS spectra showing a peptide *m/z* of 613.31 (red triangle) that is equivalent to the non-acylated peptide, E^{972}GVATQTTAYGK^{983}. (C) LC/MS/MS second dimension spectra of the 613.31 *m/z* fragment representing amino acid sequence of the peptide.

5.5 Hemolysis activity of CyaA-PF and CyaA-RTX

Effect of CyaA on sheep red blood cells (sRBC) was studied by two approaches, by visual observation and spectroscopic technique.

5.5.1 Sheep red blood cells morphology observation

After centrifugation the sRBC was pelleted at the bottom of microcentrifuge tubes. Pellets of the different reactions found were divided into two groups, first, the unchanged in color. The blood pellet of the first group still showed vivid red and easily to be resuspended **(Figure 5.7, B, C)**. The first group includes a negative control, CyaAC-RTX reaction and CyaA-RTX reaction. The pellet from second group showed darker red color **(Figure 5.7, B, C)**. The sRBC seems to clamp together, ably but harder to be resuspended. This group was includes CyaAC-PF reaction and all reaction that CyaA-PF proteins **(Figure 5.7 A)**. Supernatant of these second group show clear pale red color different from the first group that have clear colorless supernatant.

The observation was also done under light microscope. sRBC incubated with CyaAC-PF showed partially hemolyzed population while most of the sRBC incubated with CyaAC-RTX seemed to be normal **(Figure 5.8)**.

5.5.2 Hemoglobin release assays

sRBC incubated with crude *E.coli* lysate harboring pET17-b was used as a negative control (0% hemolysis) and was used to subtract all values from others reaction. The same condition was used for the 100% hemolysis control but in final step, 0.1% final concentration of Triton X-100 was added to make complete hemolysis of sRBC. CyaAC-PF (final concentration of 1 mg/ml), as a positive control, shows hemolysis activity around 18.3±9.0%. However, the same protein with increase final concentration to 3 mg./ml shows hemolysis activity around 38.3±2.8%, which is about 2.1 times more than that of 1 mg control. CyaAC-RTX and CyaA-RTX controls have hemolysis activity less than 5% (3.2±2.4% and 2.9+3.6%, respectively). Mixtures of the two protein solutions CyaAC-PF and CyaAC-RTX or CyaAC-PF and CyaA-RTX, incubated together before being added into the reaction solutions given hemolysis activity around 18.9±6.1% and 18.0±5.2%, respectively. The CyaA-RTX and CyaAC-RTX preincubated with sRBC for 18 min (5% of 6 hr) before adding CyaAC-PF give % hemolysis equal to 18.9±10.6% and 18.2±9.6%, respectively **(Figure 5.9)**.

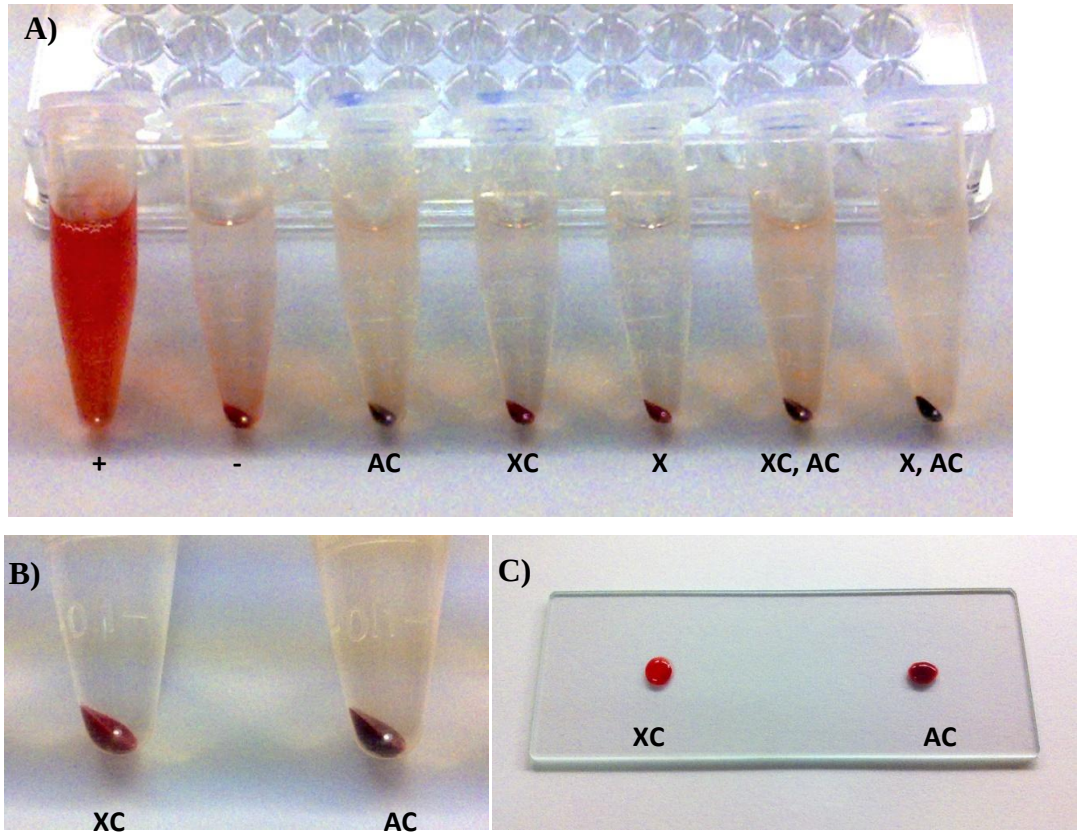


Figure 5.7 Hemolysis assays of CyaA-PF and CyaA-RTX

(A) Hemolysis assay after 6 h of incubation at 30˚C. From left, positive control (+), negative control (-), blood incubated with total lysate from *E.coli* expressing protein from pCyaAC-PF (AC), blood incubated with total lysate from *E.coli* expressing protein from pCyaAC-RTX (XC), blood incubated with total lysate from *E.coli* expressing protein from pCyaA-RTX (X), blood incubated with total lysate from two batches of *E.coli* expressing protein from pCyaAC-PF and pCyaAC-RTX (XC, AC), blood incubated with total lysate from two batches of *E.coli* expressing protein from pCyaAC-PF and pCyaA-RTX(X, AC)

(B) Color of blood and hemolized supernatant from hemolysis assay incubated with CyaA-RTX co-expressed with CyaC (XC) or CyaA-PF co-expressed with CyaC (AC)

(C) Color of resuspened blood from hemolysis assay incubated with CyaA-RTX co-expressed with CyaC (XC) or CyaA-PF co-expressed with CyaC (AC)

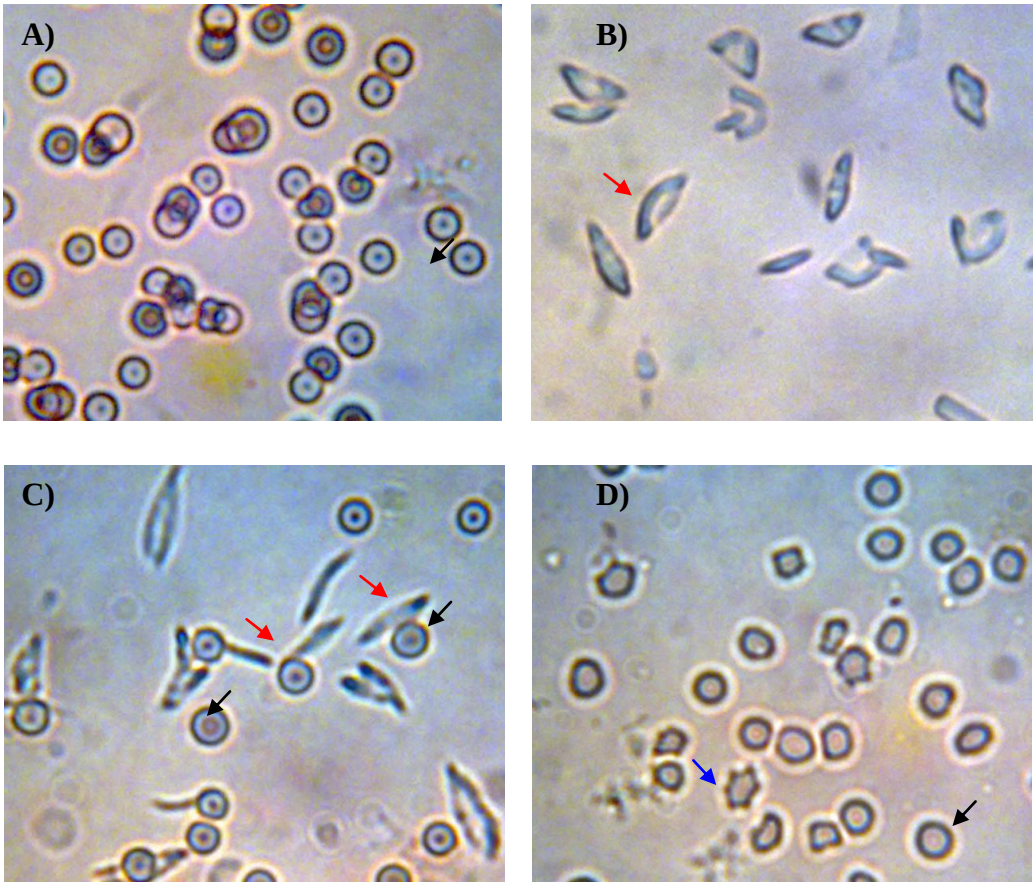


Figure 5.8 sRBC after 6 h of incubation with CyaAC-PF or CyaAC-RTX under light microscope

(A) Healthy sRBC from negative control

(B) Hemolysed sRBC

(C) sRBC incubated with crude extract of *E.coli* carrying pCyaAC-PF. Two types of cells were observed, healthy sRBC (black arrow) and hemolyzed sickle-shape sRBC (red arrow)

(D) sRBC incubated with crude extract of *E.coli* carrying pCyaAC-RTX. Mostly, the sRBC seems to be normal (black arrow), however, some shrink or star-like (crenate) sRBC was seldom observed (blue arrow).

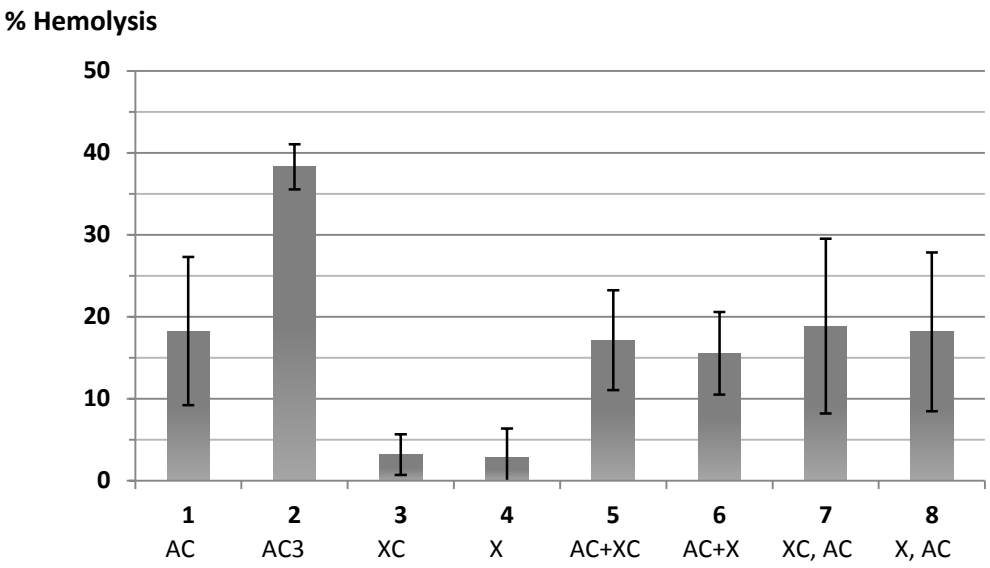


Figure 5.9 Hemolysis activity of CyaA-PF and truncated, CyaA-RTX.

Hemolysis activity of the indicated proteins incubated with sRBC for 6 h. Total cell lysate of *E.coli* harboring pET17b was used as a negative control (0% hemolysis). 1 µl of triton-X100 was added to negative control to give 100% hemolysis value. Error bars indicate standard deviation. Results in columns 1, 3 and 4 were averaged from six independent experiments, whereas, columns 2, 5, 6, 7 and 8 were from three independent experiments and each experiment was performed in duplicate.

Column 1: 1 mg of total *E.coli* cell lysate expressed CyaA-PF and CyaC

Column 2: 3 mg of total *E.coli* cell lysate co-expressed CyaA-PF and CyaC

Column 3: 2 mg of total *E.coli* cell lysate co- expressed CyaA-RTX and CyaC

Column 4: 2 mg of total *E.coli* cell lysate expressed CyaA-RTX

Column 5: 1 mg of total *E.coli* cell lysate co-expressing CyaA-PF and CyaC preincubate with 2 mg of total E.coli cell lysate co-expressed CyaA-RTX and CyaC for 18 min before being added into hemolysis reaction

Column 6: 1 mg of total *E.coli* cell lysate co-expressing CyaA-PF and CyaC preincubate with 2 mg of total *E.coli* cell lysate co-expressed CyaA-RTX for 18 min before being added into hemolysis reaction

Column 7: 2 mg of total *E.coli* cell lysate co-expressing CyaA-RTX and CyaC preincubate with sRBC for 18 min before adding 1 mg of total *E.coli* cell lysate co-expressed CyaA-PF and CyaC

Column 8: 2 mg of total *E.coli* cell lysate expressing CyaA-RTX preincubate with sRBC for 18 min before adding 1 mg of total *E.coli* cell lysate co-expressed CyaA-PF and CyaC

CHAPTER VI

In silico RESULTS: COMPUTATIONAL ANALYSIS OF CyaA TOXIN AND MOLECULAR MODELING OF THE CyaA-RTX SUBDOMAIN FRAGMENT

6.1 Primary Structure based analysis

6.1.1 General protein characters

CyaA, CyaA-PF and CyaA-RTX sequences wer submitted to ProtParam tool on ExPASy server. Some hypothetical information was listed below. It is notably that the protein contains no cysteine residue (*).

CyaA
Amino acid composition:

Ala (A)	226	13.2%	Leu (L)	147	8.6%
Arg (R)	96	5.6%	Lys (K)	43	2.5%
Asn (N)	68	4.0%	Met (M)	22	1.3%
Asp (D)	160	9.4%	Phe (F)	42	2.5%
Cys (C)	0	0.0% *	Pro (P)	37	2.2%
Gln (Q)	72	4.2%	Ser (S)	98	5.7%
Glu (E)	84	4.9%	Thr (T)	79	4.6%
Gly (G)	239	14.0%	Trp (W)	15	0.9%
His (H)	34	2.0%	Tyr (Y)	41	2.4%
Ile (I)	69	4.0%	Val (V)	134	7.9%

The instability index (II) is computed to be 24.96 (Stable)
Molecular weight: 177 521.4 (without calcium ions)
Molecular weight: 178 564.1 (with 35 calcium ions)

CyaA-PF
Amino acid composition:

Ala (A)	176	13.5%	Gly (G)	198	15.2%
Arg (R)	64	4.9%	His (H)	25	1.9%
Asn (N)	52	4.0%	Ile (I)	54	4.1%
Asp (D)	134	10.3%	Leu (L)	118	9.0%
Cys (C)	0	0.0%	Lys (K)	28	2.1%
Gln (Q)	56	4.3%	Met (M)	15	1.1%
Glu (E)	61	4.7%	Phe (F)	28	2.1%

Pro (P)	22	1.7%	Trp (W)	13	1.0%
Ser (S)	70	5.4%	Tyr (Y)	29	2.2%
Thr (T)	61	4.7%	Val (V)	102	7.8%

Molecular weight: 134 460.3 (without calcium ions)
Molecular weight: 135 503.0 (with 35 calcium ions)
CyaA-RTX
Amino acid composition:

Ala (A)	97	10.1%	Leu (L)	84	8.8%
Arg (R)	51	5.3%	Lys (K)	22	2.3%
Asn (N)	48	5.0%	Met (M)	9	0.9%
Asp (D)	120	12.6%	Phe (F)	21	2.2%
Cys (C)	0	0.0%	Pro (P)	15	1.6%
Gln (Q)	35	3.7%	Ser (S)	40	4.2%
Glu (E)	43	4.5%	Thr (T)	48	5.0%
Gly (G)	152	15.9%	Trp (W)	11	1.2%
His (H)	20	2.1%	Tyr (Y)	21	2.2%
Ile (I)	44	4.6%	Val (V)	75	7.8%

Molecular weight: 99 895.6 (without calcium ions)
Molecular weight: 100 938.3 (with 35 calcium ions)

6.1.2 Domain organization of the *Bordetella pertussis* CyaA toxin

On the following page, the figure shows a domain organization of the CyaA toxin which can be divided into 2 domains, adenylate cyclase domain (AC) and pore-forming domain (PF) **(Figure 6.1)**. The PF domain is further separated in to 4 regions, namely, the hydrophobic pore-forming region, acylation sites, RTX moiety and secretion signal. While hydrophobic region and acylation site are not distinctly divided, the RTX region is more structural distinct as region ranging from approximately residues 1015 to1657. Due to their structure similarity (as discussed further), a region in front of repeat block I including acylation site, and a region between block V and C-terminal secretion signal were named as linker 0 and linker V, respectively.

6.1.3 Domain boundary prediction

The results from 11 disorder prediction method were aligned and selected for consensus domain linker region **(Figure 6.2)**. There are 254 residues (14.89%) which appear to be highly disordered and 105 residues (6.15%) possibly highly flexible. Since the RTX region was highly charged with many aspartic acid residues, it is presumably that the region would give a disorder signal. There were only 3 predicted these order regions that were not in the RTX region. The first disorder

region was 49 residues long (residues 366-415). This region was previously proved to be the link between AC domain and PF domain. The second and third disorder regions are found within the hydrophobic-acylation region, residues 441 to 457 (17 residues) and residues 846 to 860 (15 residues). Furthermore, the RTX region also shows disordered characteristic due to its high loop content nature.

6.1.4 Secondary structure prediction

CyaA sequence was subjected to 11 primary structure based-secondary structure prediction methods. The consensus prediction is shown in **Figure 6.3**. The predicted membrane interact helices were found to be similar to those studied by Powthongchin *et.al.* The slight difference in range of each helix was arisen from different prediction methods, in which, here was a general structure prediction whereas the previously report used several trans-membrane prediction approaches.

To further affirm the prediction, the crystal structure of the AC domain solved by Gou *et. al.* (PDBID: 1YRU) was compared with the prediction in **Table 6.1**. The correct prediction in this region was around 65%. The actual accuracy was significantly higher than this since the missed calculations were usually come from the different lengths of the secondary structure.

In 2006, Bauche *et.al.* studied the secondary structure of 701-amino acid truncated CyaA (residues 1006-1706) by deconvolution-curve fitting of far UV-CD spectra. Four years later, Chenal *et. al.* revised the secondary structure of the same construct by deconvolution-curve fitting of FTIR amide I band technique. The secondary structures from these studies were compared to the prediction in **Table 6.2**. The secondary structure of the molecular model from this study was also compared. Results from all four methods were agreed very well.

```
  1 MQQSHQAGYANAADRESGIPAAVLDGIKAVAKEKNATLMFRLVNPHSTSLIAEGVATKGLGVHAK
    SSDWGLQAGYIPVNPNLSKLFGRAPEVIARADNDVNSSLAHGHTAVDLTLSKERLDYLRQAGLVT
    GMADGVVASNHAGYEQFEFRVKETSDGRYAVQYRRKGGDDFEAVKVIGNAAGIPLTADIDMFAIM
    PHLSNFRDSARSSVTSGDSVTDYLARTRRAASEATGGLDRERIDLLWKIARAGARSAVGTEARRQ
    FRYDGDMNIGVITDFELEVRNALNRRAHAVGAQDVVQHGTEQNNPFPEADEKIFVVSATGESQML
    TRGQLKEYIGQQRGEGYVFYENRAYGVAGKSLFDDGLGAAPGVPSGRSKFSPDVLETVPASPGLR
    RPSLGAVERQ
401 DSGYDSLDGVGSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHWGQRAL
    QGAQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAAS
500 LSAAVFGLGEASSAVAETVSGFF  RGSSRW
529 AGGFGVAGGAMALGGGIAAAVGA  GMSLTDDAPAGQKAAAGAE
571 IALQLTGGTVELASSIALALAAA  RGVTSGLQ
602 VAGASAGAAAGALAAALSPMEIYGLV  QQSHYADQLDKLAQESSAYGYEGDALLAQLYRDKT
                                AAEGAVAGVSAVLST
678 VGAAVSIAAAASVVGAPVAVV    TSLLTGALNGILRGVQQPIIEKLANDYARKIDELG
                              GPQAYFEKNLQARHEQL
751 NSDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSVDVFVDRFVQGERVAG
    QPVVLDVAAGGIDIASRKGERPALTFITPLAAPGEEQRRRTKTGKSEFTTFVEIVGKQDRWRIR
    DGAADTTIDLAKVVSQLVDANGVLKHSIKLDVI
931     GGDGDDVVL   ANASRIHYD
        GGAGTNTVS
958 YAALGRQDSITVSADGERFNVRKQLNNANVYREGVATQTTAYGKRTENVQYRHVELARVGQL
    VEVDTLEHVQHII                                                  Linker 0
1015    GGAGNDSIT   GNAHDNFLA
        GGSGDDRLD   GGAGNDTLV
        GGEGQNTVI   GGAGDDVFL          Repeat block I
      QDLGVWSNQLD   GGAGVDTVK
1089 YNVHQPSEERLERMGDTGIHADLQKGTVEKWPALNLFSVDHVKNIENLH                 Linker I
1138    GSRLNDRIA   GDDQDNELW
        GHDGNDTIR   GRGGDDILR
        GGLGLDTLY   GEDGNDIFL          Repeat block II
      QDDETVSDDID   GGAGLDTVD
1212 YSAMIHPGRIVAPHEYGFGIEADLSREWVRKASALGVDYYDNVRNVENVI                 Linker II
1262    GTSMKDVLI   GDAQANTLM
        GQGGDDTVR   GGDGDDLLF
        GGDGNDMLY   GDAGNDTLY          Repeat block III
        GGLGDDTLE   GGAGNDWFG
      QTQAREHDVLR   GGDGVDTVD
1354 YSQTGAHAGIAAGRIGLGILADLGAGRVDKLGEAGSSAYDTVSGIENVV                 Linker III
1403    GTELADRIT
        GDAQANVLR   GAGGADVLA
        GGEGDDVLL   GGDGDDQLS          Repeat block IV
        GDAGRDRLY   GEAGDDWFF
      QDAANAGNLLD   GGDGRDTVD
1486 FSGPGRGLDAGAKGVFLSLGKGFASLMDEPETSNVLRNIENAV                       Linker IV
1529    GSARDDVLI   GDAGANVLN
        GLAGNDVLS   GGAGDDVLL
        GDEGSDLLS   GDAGNDDLF          Repeat block V
        GGQGDDTYL FGVGYGHDTIY
       ESGGGHDTIR   INAGADQLW
1622 FARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIH                              Linker V
1658 AANQAVDQAGIEKLVEAMAQYPDPGAAAAAPPAARVPDTLMQSLAVNWR  1706
```

Figure 6.1 (previous page) Domain organization of the CyaA toxin
Complete sequence, 1706 amino acids, of the CyaA toxin showing five distinctive elements of the protein. First 400 amino acids of the adenylatecyclase domain (AC) are shown in orange. Hydrophobic segment is shown in magenta. Five putative membrane reacting amphiphatic helices are underlined and separated from adjacent regions. Acylation region (residues 751 to 1014) is in black with the acylation sites K^{860} and K^{983} hi-lighted. RTX moiety is shown in blue. Each GGXGXDXUX nona-peptide repeat is separated and shown in bold. Two sets of these repeat would form one β-roll layer. Approximately eight to ten adjacent repeat united to form a β-roll repeat block. There are five blocks of repeats (I to V) and each is divided by different linker sequences (Linkers I to IV). Additionally, linker 0 region (residues 985 to 1014) and linker V region (residues 1622 to 1657) were defined by sequence homology. Around fifty amino acids in the C-terminal (green) were proved to be secretion signal. Numbers on the left indicate amino acid positions.

```
1 MQQSHQAGYANAADRESGIPAAVLDGIKAVAKEKNATLMFRLVNPHSTSLIAEGVATKGLGVHAK
  SSDWGLQAGYIPVNPNLSKLFGRAPEVIARADNDVNSSLAHGHTAVDLTLSKERLDYLRQAGLVT
  GMADGVVASNHAGYEQFEFRVKETSDGRYAVQYRRKGGDDFEAVKVIGNAAGIPLTADIDMFAIM
  PHLSNFRDSARSSVTSGDSVTDYLARTRRAASEATGGLDRERIDLLWKIARAGARSAVGTEARRQ
  FRYDGDMNIGVITDFELEVRNALNRRAHAVGAQDVVQHGTEQNNPFPEADEKIFVVSATGESQML
  TRGQLKEYIGQQRGEGYVFYENRAYGVAGKSLFDDGLGAAPGVPSGRSKFSPDVLETVPASPGLR
  RPSLGAVERQ
401 DSGYDSLDGVGSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHWGQRALQG
     AQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAAS
500 LSAAVFGLGEASSAVAETVSGFF    RGSSRW
529 AGGFGVAGGAMALGGGIAAAVGA    GMSLTDDAPAGQKAAAGAE
571 IALQLTGGTVELASSIALALAAA    RGVTSGLQ
602 VAGASAGAAAGALAAALSPMEIYGLV QQSHYADQLDKLAQESSAYGYEGDALLAQLYRDKT
                                 AAEGAVAGVSAVLST
678 VGAAVSIAAAASVVGAPVAVV TSLLTGALNGILRGVQQPIIEKLANDYARKIDELG
                            GPQAYFEKNLQARHEQL
751 ANSDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSVDVFVDRFVQGERVAG
    QPVVLDVAAGGIDIASRKGERPALTFITPLAAPGEEQRRRTKTGKSEFTTFVEIVGKQDRWRIRD
    GAADTTIDLAKVVSQLVDANGVLKHSIKLDVI
931 GGDGDDVVL   ANASRIHYD
    GGAGTNTVS
958 YAALGRQDSITVSADGERFNVRKQLNNANVYREGVATQTTAYGKRTENVQYRHVELARVGQLVEV
    DTLEHVQHII
1015      GGAGNDSIT   GNAHDNFLA
          GGSGDDRLD   GGAGNDTLV
          GGEGQNTVI   GGAGDDVFL
        QDLGVWSNQLD   GGAGVDTVK
1089  YNVHQPSEERLERMGDTGIHADLQKGTVEKWPALNLFSVDHVKNIENLH
1138      GSRLNDRIA   GDDQDNELW
          GHDGNDTIR   GRGGDDILR
          GGLGLDTLY   GEDGNDIFL
        QDDETVSDDID   GGAGLDTVD
1212  YSAMIHPGRIVAPHEYGFGIEADLSREWVRKASALGVDYYDNVRNVENVI
1262      GTSMKDVLI   GDAQANTLM
          GQGGDDTVR   GGDGDDLLF
          GGDGNDMLY   GDAGNDTLY
          GGLGDDTLE   GGAGNDWFG
        QTQAREHDVLR   GGDGVDTVD
1354  YSQTGAHAGIAAGRIGLGILADLGAGRVDKLGEAGSSAYDTVSGIENVV
1403      GTELADRIT
          GDAQANVLR   GAGGADVLA
          GGEGDDVLL   GGDGDDQLS
          GDAGRDRLY   GEAGDDWFF
        QDAANAGNLLD   GGDGRDTVD
1486  FSGPGRGLDAGAKGVFLSLGKGFASLMDEPETSNVLRNIENAV
1529      GSARDDVLI   GDAGANVLN
          GLAGNDVLS   GGAGDDVLL
          GDEGSDLLS   GDAGNDDLF
          GGQGDDTYL FGVGYGHDTIY
         ESGGGHDTIR   INAGADQLW
1622  FARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIH
1658  AANQAVDQAGIEKLVEAMAQYPDPGAAAAAPPAARVPDTLMQSLAVNWR  1706
```

Figure 6.2 (previous page) Putative disorder region in the CyaA toxin

Complete sequences of the CyaA toxin representing consensus predicted disorder regions within the protein. Red letter codes are potentially disordered regions (at least 6 out of 11 predictions). Browns are possibly disordered regions (at least 4 out of 11 predictions). Green are consensus order regions. These disordered regions are likely to be domain linkers or intrinsically disordered structure. Numbers on the left indicate amino acid position. Three long disordered non-RTX regions are underlined.

```
1 MQQSHQAGYANAADRESGIPAAVLDGIKAVAKEKNATLMFRLVNPHSTSLIAEGVATKGLGVHAK
  SSDWGLQAGYIPVNPNLSKLFGRAPEVIARADNDVNSSLAHGHTAVDLTLSKERLDYLRQAGLVT
  GMADGVVASNHAGYEQFEFRVKETSDGRYAVQYRRKGGDDFEAVKVIGNAAGIPLTADIDMFAIM
  PHLSNFRDSARSSVTSGDSVTDYLARTRRAASEATGGLDRERIDLLWKIARAGARSAVGTEARRQ
  FRYDGDMNIGVITDFELEVRNALNRRAHAVGAQDVVQHGTEQNNPFPEADEKIFVVSATGESQML
  TRGQLKEYIGQQRGEGYVFYENRAYGVAGKSLFDDGLGAAPGVPSGRSKFSPDVLETVPASPGLR
  RPSLGAVERQ
401 DSGYDSLDGVGSRSFSLGEVSDMAAVEAAELEMTRQVLHAGARQDDAEPGVSGASAHWGQRALQG
    AQAVAAAQRLVHAIALMTQFGRAGSTNTPQEAAS
500 LSAAVFGLGEASSAVAETVSGFF      RGSSRW
529 AGGFGVAGGAMALGGGIAAAVGA      GMSLTDDAPAGQKAAAGAE
571 IALQLTGGTVELASSIALALAAA      RGVTSGLQ
602 VAGASAGAAAGALAAALSPMEIYGLV  QQSHYADQLDKLAQESSAYGYEGDALLAQLYRDKT
                                  AAEGAVAGVSAVLST
678 VGAAVSIAAAASVVGAPVAVV       TSLLTGALNGILRGVQQPIIEKLANDYARKIDELG
                                  GPQAYFEKNLQARHEQL
751 ANSDGLRKMLADLQAGWNASSVIGVQTTEISKSALELAAITGNADNLKSVDVFVDRFVQGERVAG
    QPVVLDVAAGGIDIASRKGERPALTFITPLAAPGEEQRRRTKTGKSEFTTFVEIVGKQDRWRIRD
    GAADTTIDLAKVVSQLVDANGVLKHSIKLDVI
931 GGDGDDVVL   ANASRIHYD
    GGAGTNTVS
958 YAALGRQDSITVSADGERFNVRKQLNNANVYREGVATQTTAYGKRTENVQYRHVELARVGQLVEV
    DTLEHVQHII
1015      GGAGNDSIT   GNAHDNFLA
          GGSGDDRLD   GGAGNDTLV
          GGEGQNTVI   GGAGDDVFL
          QDLGVWSNQLD   GGAGVDTVK
1089  YNVHQPSEERLERMGDTGIHADLQKGTVEKWPALNLFSVDHVKNIENLH
1138      GSRLNDRIA   GDDQDNELW
          GHDGNDTIR   GRGGDDILR
          GGLGLDTLY   GEDGNDIFL
        QDDETVSDDID   GGAGLDTVD
1212  YSAMIHPGRIVAPHEYGFGIEADLSREWVRKASALGVDYYDNVRNVENVI
1262      GTSMKDVLI   GDAQANTLM
          GQGGDDTVR   GGDGDDLLF
          GGDGNDMLY   GDAGNDTLY
          GGLGDDTLE   GGAGNDWFG
        QTQAREHDVLR   GGDGVDTVD
1354  YSQTGAHAGIAAGRIGLGILADLGAGRVDKLGEAGSSAYDTVSGIENVV
1403      GTELADRIT
          GDAQANVLR   GAGGADVLA
          GGEGDDVLL   GGDGDDQLS
          GDAGRDRLY   GEAGDDWFF
        QDAANAGNLLD   GGDGRDTVD
1486  FSGPGRGLDAGAKGVFLSLGKGFASLMDEPETSNVLRNIENAV
1529      GSARDDVLI   GDAGANVLN
          GLAGNDVLS   GGAGDDVLL
          GDEGSDLLS   GDAGNDDLF
          GGQGDDTYL FGVGYGHDTIY
        ESGGGHDTIR   INAGADQLW
1622  FARQGNDLEIRILGTDDALTVHDWYRDADHRVEIIH
1658  AANQAVDQAGIEKLVEAMAQYPDPGAAAAAPPAARVPDTLMQSLAVNWR  1706
```

Figure 6.3 (previous page) Predicted secondary structure of CyaA

Complete sequences of the CyaA toxin representing consensus predicted secondary structure. Helices, β-sheet and random coil are colored in blue red and black, respectively. The repeat blocks are in bold. Remark that 3-4 residues at the end of the nona-peptide repeats were usually in β structure. The predicted membrane interact helices are solid underlined. C-terminal region of linkers II and III (underlined) were α-helices predicted by 4 out of 11 methods.

Table 6.1 Secondary structure of the AC domain from two different methods

Method		**Total (residues)**	Coil	β-sheet	α-helix (and 3_{10}-helix)
X-ray crystallography* (PDBID: 1YRU, residues 7-225 and 233- 365)		351 *100%*	152	63	136 (19 as 3_{10}helix)
Amino acid sequences-based prediction (Residue 7-225 and 233- 365)	Total Prediction+	351	162	60	129
	Correct prediction++	**230** ***65.5%***	102 *67.1%*	34 *54.0%*	94 *69.1%*
	False prediction	121	60	26	35

* from Q. Guo *et.al.* (2005), residues 1-6 and 226-232 were absent from the crystal structure.

+ Including false positive

++ Compared with the crystal structure.

Table 6.2 Secondary structure of the RTX region from four different methods

Method	Total (residues)	Coil	β-sheet	α-helix
Amino acid sequences-based prediction (Residues 931-1603)	673 *100%*	401 *59.6%*	162 *24.1%*	110 *16.3%*
Homology modeling (Residues 931-1603)	673 *100%*	384 *57.1%*	277 *41.2%*	12 *1.8%*
Deconvolution-curve fitting of FTIR amide I' band* (Residues 1006-1706)	701 *100%*	*61%* (Coil + β-turn)	*39%*	-
Deconvolution-curve fitting of far UV-CD spectra** (Residues 1006-1706)	701 *100%*	*60%* (turn + others)	*28%*	*12%*

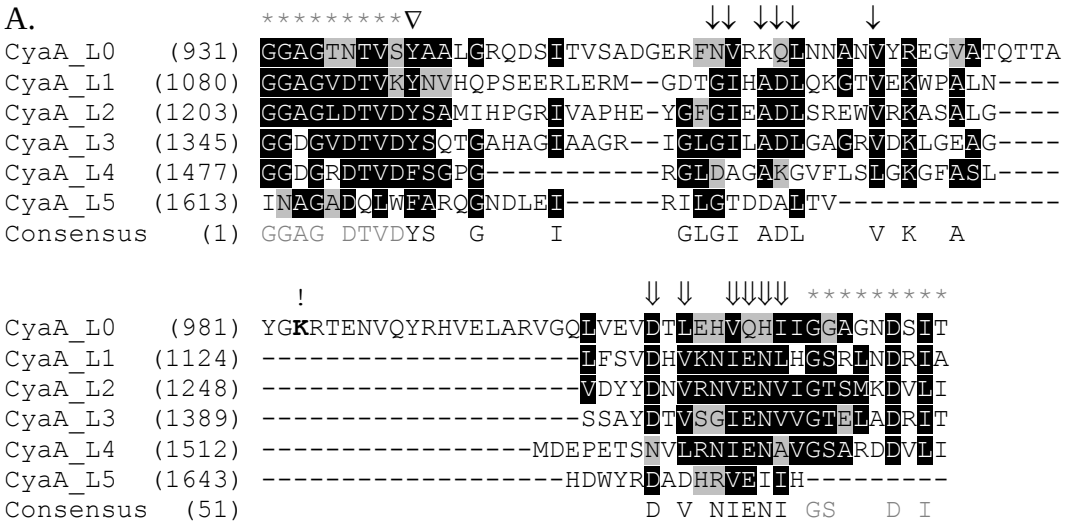
* from A. Chenal *et.al.* (2010)
** from C. Bauche *et.al.* (2006)

6.1.4 Studies on linker fragment in the RTX region

The RTX region of the CyaA protein was divided in to 5 blocks and each block contains 8-10 nona-peptide repeats. In between each block of the repeat, there are 43-50 non-repeat sequences inserted. With this criteria, there are 4 linker regions; linker I (residues 1087-1135), linker II (residues 1210-1259), linker III (residues 1353-1401) and linker IV (residues 1485-1527). Additional sequence between residues 940-1013 were also considered and named linker 0. These sequences fulfill only part of the criteria since they lies between 2 distinct nona-peptide repeat and the block 1 repeat. Each linker seems to have unique sequences and properties so studies on these fragment were performed based on amino acid composition.

6.1.4.1 Alignments of the linker fragments Six linker sequences (Linkers 0, I, II, III, IV and V) were extracted from the complete sequences of CyaA. The excised sequences contain the linker region plus 1 flanking block of the repeat. Multiple alignment was performed using AlignX software. The software was calculated based on ClustalW algorithm. Blossom62 was used as scoring matrix with gap opening penalty of -10 and gap extension penalty of -0.1. The output data were manually adjusted to provide the best result. Some consensus sequences were identified (Figure 5.4A) and used for further studies. Predicted secondary structures of the linker region were also aligned **(Figure 6.4B)**.

A.


CyaA_L0 (931) GGAGTNTVSYAALGRQDSITVSADGERFNVRKQLNNANVYREGVATQTTA
CyaA_L1 (1080) GGAGVDTVKYNVHQPSEERLERM--GDTGIHADLQKGTVEKWPALN----
CyaA_L2 (1203) GGAGLDTVDYSAMIHPGRIVAPHE-YGFGIEADLSREWVRKASALG----
CyaA_L3 (1345) GGDGVDTVDYSQTGAHAGIAAGR--IGLGILADLGAGRVDKLGEAG----
CyaA_L4 (1477) GGDGRDTVDFSGPG----------RGLDAGAKGVFLSLGKGFASL----
CyaA_L5 (1613) INAGADQLWFARQGNDLEI------RILGTDDALTV--------------
Consensus (1) GGAG DTVDYS G I GLGI ADL V K A
CyaA_L0 (981) YGKRTENVQYRHVELARVGQLVEVDTLEHVQHIIGGAGNDSIT
CyaA_L1 (1124) ----------IFSVDHVKNIENLHGSRLNDRIA
CyaA_L2 (1248) ----------VDYYDNVRNVENVIGTSMKDVLI
CyaA_L3 (1389) ----------SSAYDTVSGIENVVGTELADRIT
CyaA_L4 (1512) ----------MDEPETSNVLRNIENAVGSARDDVLI
CyaA_L5 (1643) ----------HDWYRDADHRVEITH---------
Consensus (51) D V NIENI GS D I

B.

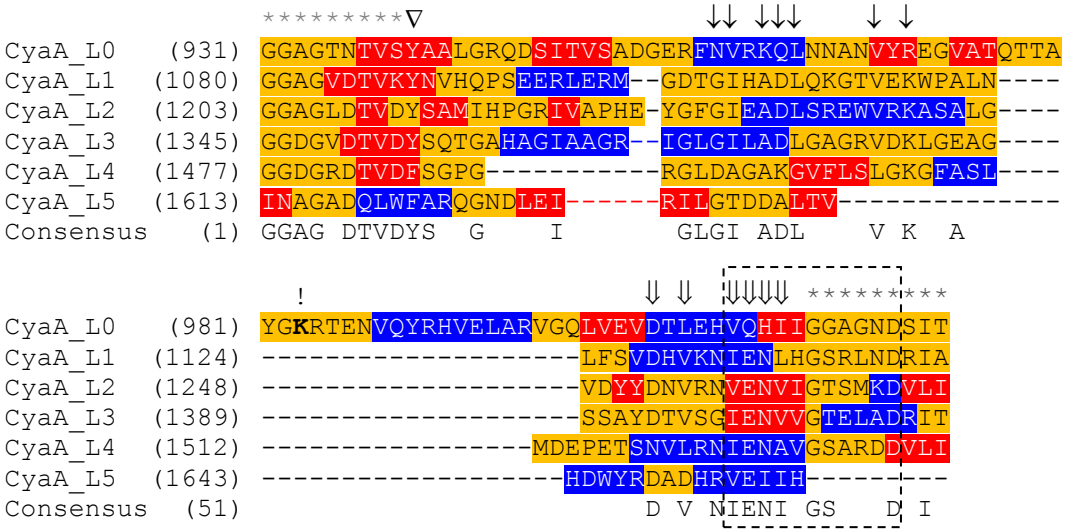

CyaA_L0 (931) GGAGTNTVSYAALGRQDSITVSADGERFNVRKQLNNANVYREGVATQTTA
CyaA_L1 (1080) GGAGVDTVKYNVHQPSEERLERM--GDTGIHADLQKGTVEKWPALN----
CyaA_L2 (1203) GGAGLDTVDYSAMIHPGRIVAPHE-YGFGIEADLSREWVRKASALG----
CyaA_L3 (1345) GGDGVDTVDYSQTGAHAGIAAGR--IGLGILADLGAGRVDKLGEAG----
CyaA_L4 (1477) GGDGRDTVDFSGPG----------RGLDAGAKGVFLSLGKGFASL----
CyaA_L5 (1613) INAGADQLWFARQGNDLEI------RILGTDDALTV--------------
Consensus (1) GGAG DTVDYS G I GLGI ADL V K A
CyaA_L0 (981) YGKRTENVQYRHVELARVGQLVEVDTLEHVQHIIGGAGNDSIT
CyaA_L1 (1124) ----------LFSVDHVKNIENLHGSRLNDRIA
CyaA_L2 (1248) ----------VDYYDNVRNVENVIGTSMKDVLI
CyaA_L3 (1389) ----------SSAYDTVSGIENVVGTELADRIT
CyaA_L4 (1512) ----------MDEPETSNVLRNIENAVGSARDDVLI
CyaA_L5 (1643) ----------HDWYRDADHRVEIIH---------
Consensus (51) D V NIENI GS D I

Figure 6.4 (previous page) Multiple alignments of the four linkers of CyaA and adjacent sequences

Six linkers with a block of each flanking repeat are labeled as CyaA_L1 (linker I), CyaA_L2 (linker II), CyaA_L3 (linker III), CyaA_L4 (linker IV) and CyaA_L5 (linker V). Residues which are marked with (*) are the GGXDXGXUX repeats adjacent to the linkers. Numbers in brackets indicate positions of the first residue. CyaA_L0 is the sequence prior to block 1 of the repeat containing an acylation site, K^{983}, which is marked with exclamation mark (!). Aromatic residues following each block of repeats are marked by triangle (∇). The conserved GIXADL and VXK are marked by arrows (↓). The other conserved predicted amphiphatic helix, DXVXXUENU, is marked by block arrows (⇓). (A) white letter amino acids highlighted in black (x) are closely related residues and highlighted in gray (x) are similar amino acids. To compare the predicted secondary structure, the aligmnet are re-colored in (B) White letter with blue background (x) are α-helix, red background (x) are β-sheet and black letters with yellow background (x) are random coil. The putative aliphatic helices are marked by dash box. Notably that 4 out of 11 methods predict these regions in linker II and III as helices.

6.1.4.2 Studies on consensus GIXADL sequence From the alignment, linkers I, II and III share an analogous GIXADL sequence (X is any amino acid). However, predicted secondary structure of these GIXADL sequences shares no comparable organization. To rule out chance of occurring as a random sequence, 22 random sequences were generated from RandSeq on ExPASy website. The amino acid propensities of the random peptides were based on standard amino acid abundant table. All 22 sequences were checked against PROSITE database. Statistic data are shown in **Figure 6.5**. The GIXADL sequence was found to be more than 2sd above the mean and fall in higher than 90^{th} percentile in Weibull distribution.

Moreover, some crystal structures of the proteins that have the GIXADL sequence were investigated. The lists of these proteins are shown in **Table 6.3**. Nonetheless, there was no conserved or significant three-dimensional structure among these proteins.

6.1.4.3 Studies on consensus DXVXXUENU sequence Unlike GIXADL sequences, DXVXXUENU sequences (U is hydrophobic residue) tend to have a conserved helix structure. These sequences have at least one hydrophobic amino acid for every three residues, suggesting that they would form an amphipathic helix. So helical projections of these residues were draw. The hydrodynamic were written in the middle of each helix. (**Figure 6.6**) Of the 6 possible sequences, all of them show 11 potential amino acids for amphipathic helices with at least 4 residues on the hydrophilic surface.

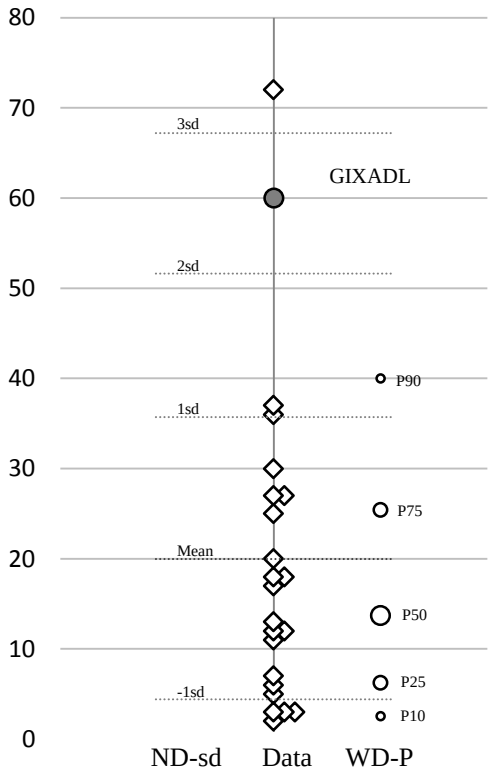


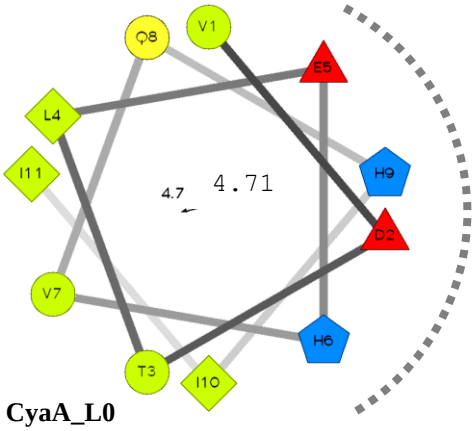
Figure 6.5 Statistical data of GIXADL and 22 random sequences

Number of hits using GIXADL and 22 random sequences searched in PROSITE database using advance search function. Diamond dots in the data column represent numbers of hit for random sequences and gray dot (60 hits) for GIXADL sequence. For evaluation, twenty-two random sequences were used for generating statistical data. ND-sd and WD-P were normal distribution standard deviation and Weibull distribution percentile, respectively.

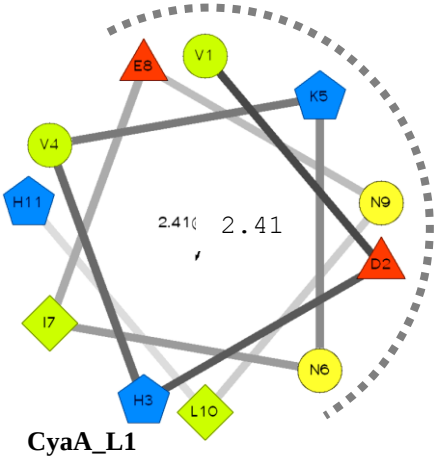
Table 6.3 PDB ID of proteins that contain GIXADL sequence

Consensus sequences	Protein	#AA[a]	PDB entry (Chain-ID)	Source organism[b]
GILADL	S-adenosyl-L-methionine-dependent methyltransferase (mraW)	285	1WG8-A, 1WG8-B	*Thermus thermophilus* and many bacteria
GILADL	putative GTP pyrophosphokinase	739	1VJ7-A, 1VJ7-B	*Streptococcus euismilis*
GIQADL	Hydroxymethylpyrimidine / phosphomethylpyrimidine kinase	266	1JXH-A, 1JXI-A, 1LLJ-A	*Salmonella typhimurium* and many bacteria
GIQADL	Pyridoxine kinase	271	2I5B-A, 2I5B-B, 2I5B-C, 2I5B-D, 2I5B-E	*Bacillus subtilis* and *Treponema pallidum*
GIVADL	Electron transfer flavoprotein subunit alpha	333	1EFV-A	*homo sapiens*' mitochondrial protein, in mitochondria of rice and many mammals
GIVADL	Arginyl-tRNA synthase	577	1LLE-A	*Salmonella typhimurium* and many bacteria
GIPADL	Leucoanthocyanidin dioxygenase	339	1GP4-A, 1GP5-A, 1GP6-A, 2BRT-A	*Arabidopsis thaliana*
GIPADL	Malate synthase A	533	3CUZ-A, 3CV1-A, 3CV2-A	*Escherichia coli*

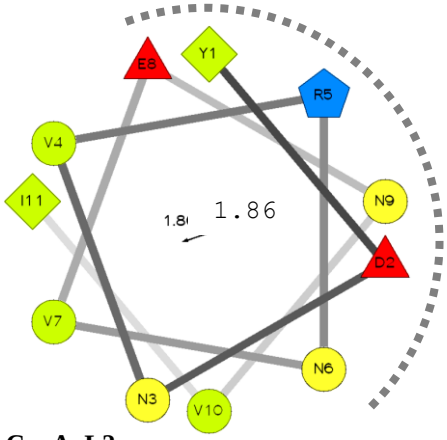
#AA[a] : Number of total amino acids in the protein.
Source organism[b] : The first organism mentioned produce the protein for 3D structure determination. The others are organisms that proved to produce the same protein.



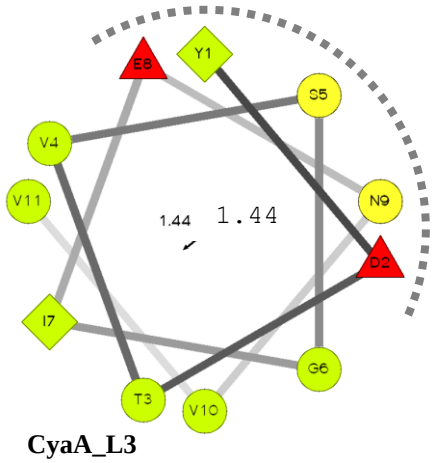
CyaA_L0
AA# 1003-1015



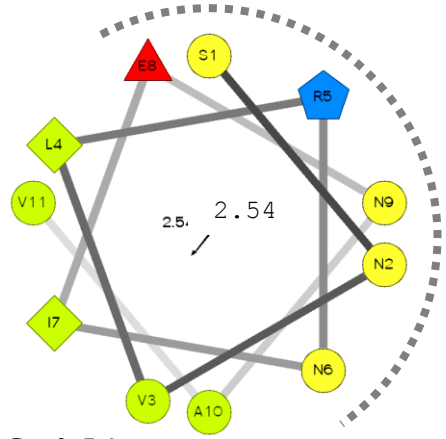
CyaA_L1
AA# 1127-1137



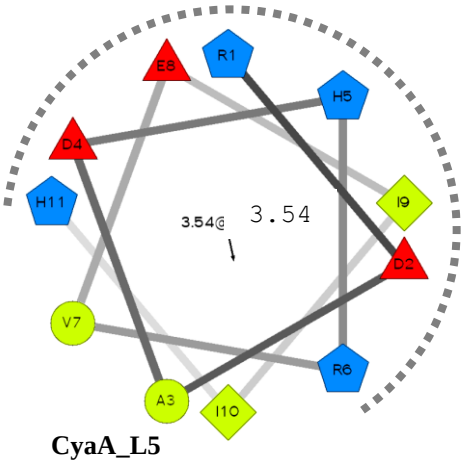
CyaA_L2
AA# 1251-1261



CyaA_L3
AA# 1392-1402



CyaA_L4
AA# 1518-1528



CyaA_L5
AA# 1647-1657

Figure 6.6 (previous page) Amphiphilic helices at the end of each linker

Helical projections illustrate 11 amino acid predicted helices at the end of each linker. Hydrophobic or aromatic residues, un-charged hydrophilic residues, positively charged residues and negatively charged residues are shown in green, yellow, blue and red, respectively. Hydrophilic surfaces are emphasized with dash curves. The hydrodynamic are written in the middle of each helix.

6.1.5 Amino acid composition of the RTX nona-peptide repeat

Amino acids abundant in 45 repeats are shown in **Table 6.4** according to their repeat positions. First G position (**G**GXGXDXUX) was the most conserved rsidues having 38 glycine out of 45 positions. The next G position (G**G**XGXDXUX) was less conserve with 21 glycine and 8 aspartic acid residues. The X position (GG**X**GXDXUX) mostly appears as alanine and aspartic acid (18 and 8 positions, respectively) the third G position (GGX**G**XDXUX) was somewhat conserved with 33 glycine out of 45 positions. The next X position (GGXGXDXUX) mostly contains hydrophilic residue, 14 aspartic acid residues and 10 asparagine residues. The conserve D position located only aspartic acid and asparagine residues (37 and 8 positions, respectively). The third X position lays Y shape side chain amino acids i.e. 15 threonine and 10 valine residues. The conserved U position comprised of 27 leucine residues, 7 isoleucine residues, 6 valine residues and 4 phenylalanine residues. The last X position covers many amino acids being nonspecific side-chain properties. From 405 amino acid positions of the nona-peptide repeat, there were 23 positively charged R groups which mostly the R residues (16 positions). On the other hand, the negatively charged R groups dominate region through 76 aspartic acid residues and 11 glutamic acid residues. There were total of 64 hydrophilic uncharged R groups in the region (22, 10, 12 and 20 residues of asparagine, glutamine, serine and threonine, respectively). The 93 of 98 glycine residues were apportion in the conserved G position, whereas the 18 of 32 alanine residues were located in the first X position. More than half of the isoleucine and leucine residues (7 out of 13 and 27 out of 43, respectively) were at the U position while the less were uniformly distributed over the less of the position. There were 8, 5 and 7 residues of aromatic R group, phenylalanine, tryptophan and tyrosine, respectively. There was no proline and cysteine amino acid in this region.

Table 6.4 Amino acids with frequency of each position in the repeat position

	G_1	G_2	X_3	G_4	X_5	D_6	X_7	U_8	X_9	**Total**
Positively charged R groups										
K	*				1				1	2
R		1	2	1	2		4		6	16
Negatively charged R groups										
D	1	8	8		14	**37**	1		7	76
E		3	4	1			2		1	11
Polar uncharged R groups										
N		2	1		10	8			1	22
Q	1	1	1	3	1		3			10
S	1	2	2		3		1		3	12
T		2	1				15		2	20
Aliphatic R groups										
A	1	3	18	1	6				3	32
G	**38**	**21**	4	**33**	1				1	98
H		1		1	3					5
I	1						2	**7**	3	13
L	1	1	2	2	2		3	**27**	5	43
P										**0**
V	1		1	1	2		10	**6**	1	22
Aromatic R groups										
F							1	**4**	3	8
W				1			2		2	5
Y			1					1	5	7
Sulfur containing R groups										
C										**0**
M				1			1		1	3

* Gray numbers : Amino acids that occurred less than 5% at the position (< 3 residues)
Dark gray numbers: Amino acids that occurred less than 15% at the position (< 7 residues)
Black numbers : Amino acids that occurred more than 15% at the position (≥ 7 residues)
In bold : Consensus sequence at the position

6.2 Molecular modeling of CyaA-RTX region.

To have better understanding of the RTX subdomain, models of the region were generated via computational molecular modeling.

6.2.1 Alignments of the RTX subdomain with template structure for local molecular modeling

Sequence of the RTX subdomain of the CyaA protein was subjected to pairwise alignments with other known structure RTX subdomain from other proteins. Until now, there are only two solved crystal structures of the wild-type RTX subdomain that contain linker sequence, *pseudomonas* sp. MIS38 Lipase (I3 lipase, PDBID: 2z8x) and *serratia marcescens* lipase (LipA, PDBID: 1qub). The template RTX sequence was divided into 5 fragments. Each portion has a linker and 5-6 repeats of each flanking repeats. The first fragment (CyaA_L0) has acylation site instead of a linker. Each fragment was subjected to pairwise alignments with both known structures for 4 criteria; 1) ClustalW auto alignment based on primary structure, 2) manually modified primary structure alignment, 3) alignment based on secondary structure and 4) manually modified alignment based on both primary and secondary structures. So maximum of 8 alignments were performed for each query. All sequence alignments were submitted to Swiss-PDB modeling server using alignment mode. The resulting models were evaluated based on;

- Ramachandran plot (ϕ/ψ-angle)
- QMEAN Z-score (overall structure assessment; C-beta interaction, atom pairwise energy, Solvation energy, Torsion angle energy, Secondary structure agreement, Solvent accessibility agreement)
- Residue geometry (Chi1-Chi2 plot, planar group)
- Main-chain parameter (peptide planar (ω-rotation), bond length, bond angle)
- Side-chain parameter (bond length, bond angle)

The best model of each block was selected for further manual adjusting. The alignment of the each selected model is shown in **Figure. 6.7**. Calcium ions were added to each model by 3-dimentional fitting with the template molecules (2z8x). It should be noted that not all of the calcium binding sites were filled since some sites

were vacant in the template molecule. To give best inter-residue contact and the calcium ion contact, the rotamer of some residues were manually adjusted. Structure of the nona-peptide repeat were checked against template where consensus GGXGXD should form a short turn while consensus XUX forms a short β-strand which hydrophobicly interacted with the XUX sequence of the next repeat. Energy minimization were carried out for each model using steepest descent method from DeepView/Swiss-PdbViewer software until the energy dropped less than 0.1 kjoles per step.

CyaA_L0 (residues 949-1042)

```
CYAA_L0   1    ---------G GAGTNTVSYA ALGRQDSITV SADGERFNVR KQLNNANVYR
2Z8XA     400  DGGGYNVILG GAGNNTLDLQ KSVNTFDFAN DGAGNLYVRD ANGGISITRD

CYAA_L0   42   EGVATQTTAY GKRTENVQYR HVELARVGQL VEVDTLEHVQ HIIGGAGNDS
2Z8XA     450  IGSIVTKEPG FLWGLFKDDV THSVTASGLK VGSNVTQYDA SVKGTNGADT

CYAA_L0   92   ITGNAHDNFL AGGAGNDSIT GNAHDNFLAG GSGDDRLDGG AGDNTLV---
2Z8XA     500  LKAHAGGDWL FGLDGNDHLI GGVGNDVFVG GAGNDLMESG GGADTFLFNG
```

CyaA_L1 (residues 1034-1211)

```
CYAA_L1   1    ---------- ---------- ---GGAGNDT LVGGEGQNTV IGGAGDDVFL
2Z8XA     352  RANTWVQDLN RNAETHKGST FIIGSDSNDL IQGGSGNDYL EGRAGNDTFR

CYAA_L1   28   QDLGVWSNQL DGGAGVDTVK YNVHQPSEER LERMGDTGIH A-----DLQK
2Z8XA     400  --DGGGYNVI LGGAGNNTLD LQKSVNTFDF ANDGAGNLYV RDANGGISIT

CYAA_L1   73   GTVE------ ---------- -----KWPAL NLFSVDHVKN IENLHGSRLN
2Z8XA     448  RDIGSIVTKE PGFLWGLFKD DVTHSVTASG LKVGSNVTQY DASVKGTNGA

CYAA_L1   102  DRIAGDDQDN ELWGHDGNDT IRGRGGDDIL RGGLGLDTLY GEDGNDIFLQ
2Z8XA     498  DTLKAHAGGD WLFGLDGNDH LIGGVGNDVF VGGAGNDLME SGGGADTFLF

CYAA_L1   152  DDETVSDDID GGAGLDTVD- ----------
2Z8XA     548  NGAFGQDRVV GFTSNDKLVF LGVQGVLPND
```

CyaA_L2 (residues 1165-1333)

```
CYAA_L2   1    ---------- ---------- ---GRGGDDI LRGGLGLDTL YGEDGNDIFL
2Z8XA     352  RANTWVQDLN RNAETHKGST FIIGSDSNDL IQGGSGNDYL EGRAGNDTFR

CYAA_L2   28   QDDETVSDDI DGGAGLDTVD YSAMIHPGRI VA-----PHE YGFGIEADLS
2Z8XA     400  --DGGGYNVI LGGAGNNTLD LQKSVNTFDF ANDGAGNLYV RDANGGISIT

CYAA_L2   73   REWVRKASAL GVDYY----- ---------- -----DNVRN VENVIGTSMK
2Z8XA     448  RDIGSIVTKE PGFLWGLFKD DVTHSVTASG LKVGSNVTQY DASVKGTNGA

CYAA_L2   103  DVLIGDAQAN TLMGQGGDDT VRGGDGDDLL FGGDGNDMLY GDAGNDTLYG
2Z8XA     498  DTLKAHAGGD WLFGLDGNDH LIGGVGNDVF VGGAGNDLME SGGGADTFLF

CYAA_L2   153  G--LGDDTLE GGAGNDWFG- ----------
2Z8XA     548  NGAFGQDRVV GFTSNDKLVF LGVQGVLPND
```

CyaA_L3 (residues 1307-1476)

```
CYAA_L3   1    ---------- ---------- ---GDAGNDT LYGGLGDDTL EGGAGNDWFG
2Z8XA     352  RANTWVQDLN RNAETHKGST FIIGSDSNDL IQGGSGNDYL EGRAGNDTFR

CYAA_L3   28   QTQAREHDVL RGGDGVDTVD YSQTGAHAGI AA-----GRI GLGILADLGA
2Z8XA     400  --DGGGYNVI LGGAGNNTLD LQKSVNTFDF ANDGAGNLYV RDANGGISIT

CYAA_L3   73   GRVD------ ---------- -----KLGEA GSSAYDTVSG IENVVGTELA
2Z8XA     448  RDIGSIVTKE PGFLWGLFKD DVTHSVTASG LKVGSNVTQY DASVKGTNGA

CYAA_L3   102  DRITGDAQAN VLRGAGGADV LAGGEGDDVL LGGDGDDQLS GDAGRDRLYG
2Z8XA     498  DTLKAHAGGD WLFGLDGNDH LIGGVGNDVF VGGAGNDLME SGGGADTFLF
```

```
CYAA_L3   152  E--AGDDWFF QDAANAGNLL D---------
2Z8XA     548  NGAFGQDRVV --GFTSNDKL VFLGVQGVLP
```

CyaA_L4 (residues 1438-1602)

```
CYAA_L4   1    ---------- ---------- ---GGDGDDQ LSGDAGRDRL YGEAGDDWFF
2Z8XA     352  RANTWVQDLN RNAETHKGST FIIGSDSNDL IQGGSGNDYL EGRAGNDTFR

CYAA_L4   28   QDAANAGNLL DGGDGRDTVD FSGPGRGLDA GA-----KGV F-----LSLG
2Z8XA     400  --DGGGYNVI LGGAGNNTLD LQKSVNTFDF ANDGAGNLYV RDANGGISIT

CYAA_L4   68   KGFASL-MDE PETS------ ---------- -----NVLRN IENAVGSARD
2Z8XA     448  RDIGSIVTKE PGFLWGLFKD DVTHSVTASG LKVGSNVTQY DASVKGTNGA

CYAA_L4   96   DVLIGDAGAN VLNGLAGNDV LSGGAGDDVL LGDEGSDLLS GDAGNDDLFG
2Z8XA     498  DTLKAHAGGD WLFGLDGNDH LIGGVGNDVF VGGAGNDLME SGGGADTFLF

CYAA_L4   146  G--QGDDTYL FGVGYGHDTI Y---------
2Z8XA     548  NGAFGQDRVV --GFTSNDKL VFLGVQGVLP
```

Figure 6.7 Pairwise alignment for molecular modeling
The alignment of the models those were selected for the final modeling. The low sequence similarity was due to taking consideration of the secondary structure during adjusting process.

6.2.2 Model joining and furnishing and evaluation of technical statistic

As the models were separated into fragments, joining them together was a crucial step to obtain the whole protein domain. Pair fitting function in Pymol software was used to align the overlapped repeats together. To do this the last and the first repeats of each model were ignored because they are likely to be misfolded. The α-carbon positions of at least nine amino acids of the second repeat were marked and pair fitted to the α-carbon of the second last repeat of the prior repeat. The bond that was selected to join the model were selected from the XUX sequence, for the reason that they both form a β-strand and have less movement. The selected bonds were cut and rejoined using build function. The same procedures were performed throughout all five linker models. The final model was then again energy minimized using DeepView/Swiss-PdbViewer software. The free energy was minimized using steepest descent method until, each step, the energy dropped less than 0.1 kJ/mole and then applied the conjugate gradient method until the energy declined less than 0.01 kJ/mole per step.

Some technical statistical data of the finish model were evaluated.

Ramachandran plot (ϕ/ψ-angle) is display in **Figure 6.8**. From 690 residues, there are 547 non-G and non-P residues. Within this group, there are 523 residues (95.6%) in favoured regions, additional allowed regions or generously allowed regions.

The QMEAN score is a composite score consisting of a linear combination of 6 terms. The pseudo-energies of the contributing terms are given below together with their Z-scores with respect to scores obtained for high-resolution experimental structures of similar size solved by X-ray crystallography.

		Z-score;
β-C interaction energy:	1.00	(2.35)
All-atom pairwise energy:	2.37	(2.49)
Solvation energy:	-4073.13	(5.62)
Torsion angle energy:	54.12	(4.30)
Secondary structure agreement:	0.0%	(-1.87)
Solvent accessibility agreement:	69.6%	(-3.21)

Total QMEAN score: 0.499 (Z-score: -3.90)

The score expected to be 1.0. The acceptable range of Z-values is ±4.0.*

* The QMEAN score was ranged between 0 and 1, which 1 is the ideal. Z-score is the number of standard deviations that the score deviated from the expected value.

Additional information (not included in QMEAN scoring)

	Z-score;
Average bond length	0.73
Average bond angle	1.18
Ramachandran plot (ϕ/ψ-angle)	-0.56
Omega (ω) angle	7.96
Chi1-Chi2 (χ_1-χ_2) angle	-0.40
Outside/inside residues distribution	1.15
Structure average packing	-3.40

Main-chain parameter and side chain parameter were showed in **Table 6.4**. Main-chain parameters were within acceptable range, as the side-chain parameters were better than expected value. Complete directory of all parameters are shown in appendix A.

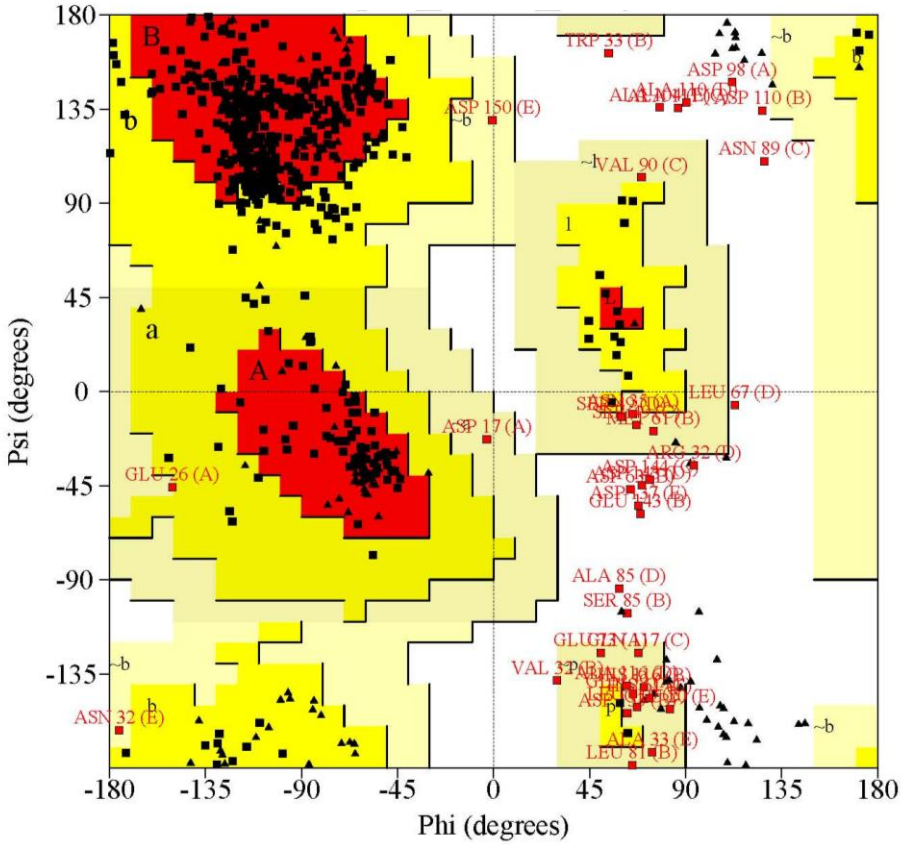


Figure 6.8 Ramachandran plot of the RTX region molecular model of CyaA toxin

The Psi/Phi angles of all residues of the model were plotted as black triangle for glycine and square for the others. Except glycine that has no constrain from its side chain, the others residues that fell in the generously allowed regions (pale yellow) and disallowed regions (white) are labeled (by name, position and chain) and colored in red. More than 93.3% of the residues stay in the most favoured regions (red) and additional allowed regions (yellow). Raw data and some other parameters are shown in appendix A. Plot statistics are shown below;

Resides in most favoured regions (red)	397	72.6 %
Resides in additional allowed regions (yellow)	113	20.7 %
Resides in generously allowed regions (pale yellow)	20	3.7 %
Resides in disallowed regions (white)	17	3.1 %
Number of non-glycine and non-proline residues	547	100 %
Number of glycine residues (triangle)	131	
Numder of proline residues	6	
Total	690	

Table 6.5 Main-chain and side-chain parameters of CyaA-RTX molecular model

Stereochemical parameter	Model values		Comparison values	
	No. of data points	Parameter value	Typical value	Remark
Main-chain parameter				
Ramachandran plot quality assessment	547	72.6	76.6 ±10.0	Inside
Peptide bond planarity (omega angle sd)	679	8.0	6.0 ±3.0	Inside
bad non-bonded interactions	1	0.1	10.5 ±10.0	Better
Alpha carbon tetrahedral distortion	559	2.4	3.1 ±1.6	Inside
Hydrogen bond energies	388	0.9	0.9 ±0.2	Inside
Overall G-factor	690	-0.4	-0.6 ±0.3	Inside
Side-chain parameter				
Chi-1 gauche minus sd	121	9.4	22.7 ±6.5	Better
Chi-1 trans sd	130	9.0	22.7 ±5.3	Better
Chi-1 gauche plus sd	242	9.0	21.3 ±4.9	Better
Chi-1 pooled sd	493	10.3	22.0 ±4.8	Better
Chi-2 trans sd	121	11.6	23.1 ±5.0	Better

6.2.3 CyaA-RTX molecular model

Overall structure of CyaA-RTX from G^{931} to Y^{1603} is illustrated in **Figure 6.9** representing 5 blocks of the repeat joining together side by side in head-to-tail fashion. The K^{938} residue is on the surface of the modeled molecule, exposed its amine side chain to the aqueous environment (**Figure 5.10B**). The surface of the region is shown to have unequally distribution of the electrostatic charge (**Figure 5.10A**). Each repeat block structure was similar however the linkers are relatively diverse (**Figure 5.12**). Dimension of the condensed repeat sequence without linker is around 25 × 19 × 10 angstrom for a 10 repeats block (Figure 5.13).

The insight repeat organization shown in **Figure 5.14A** appears as a stack of repeat each comprised of 2 consecutive nona-peptide sequences. The stacks finally formed a β-roll block. Every position of the repeat has its own function i.e. turn-forming structure (Figure 5.14B), β-sheet forming structure (Figure 5.15), conserve calcium-binding residues (Figure 5.16), conserved hydrophobic core scaffold (Figure 5.17) and cross-block interaction surface (Figure 5.18).

Furthermore the specially consensus QD sequence stereo positioning is shown in **Figure 5.19**. The predicted α-helices interacting with repeat block is shown in **Figure 5.20**. Summary of the interaction stabilizing each block of repeat is shown in **Figure 5.21**.

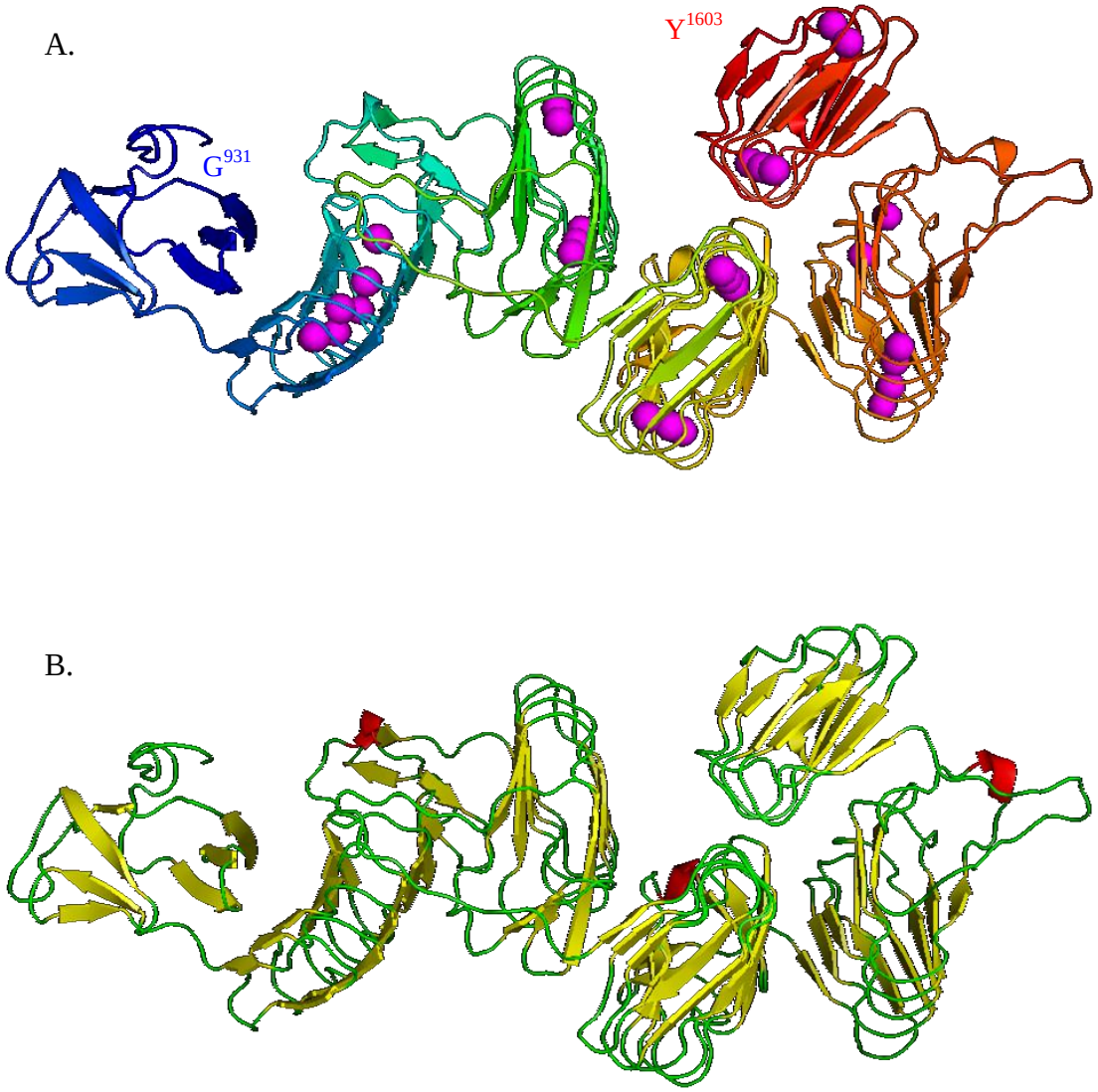


Figure 6.9 Overall structure of the RTX region molecular model

The whole structure of the predicted molecular model of the RTX region starting from G^{931} to Y^{1603} presented in cartoon format. The C-terminal of the structure is shortened due to lack of homology between CyaA and the template. (A) N-terminal to C-terminal end of the protein is colored in blue to red spectrum. Calcium ions are colored in magenta. (B) Secondary structure organization of the RTX region in which β-sheet, α-helix and random coil are shown in yellow red and green, respectively.

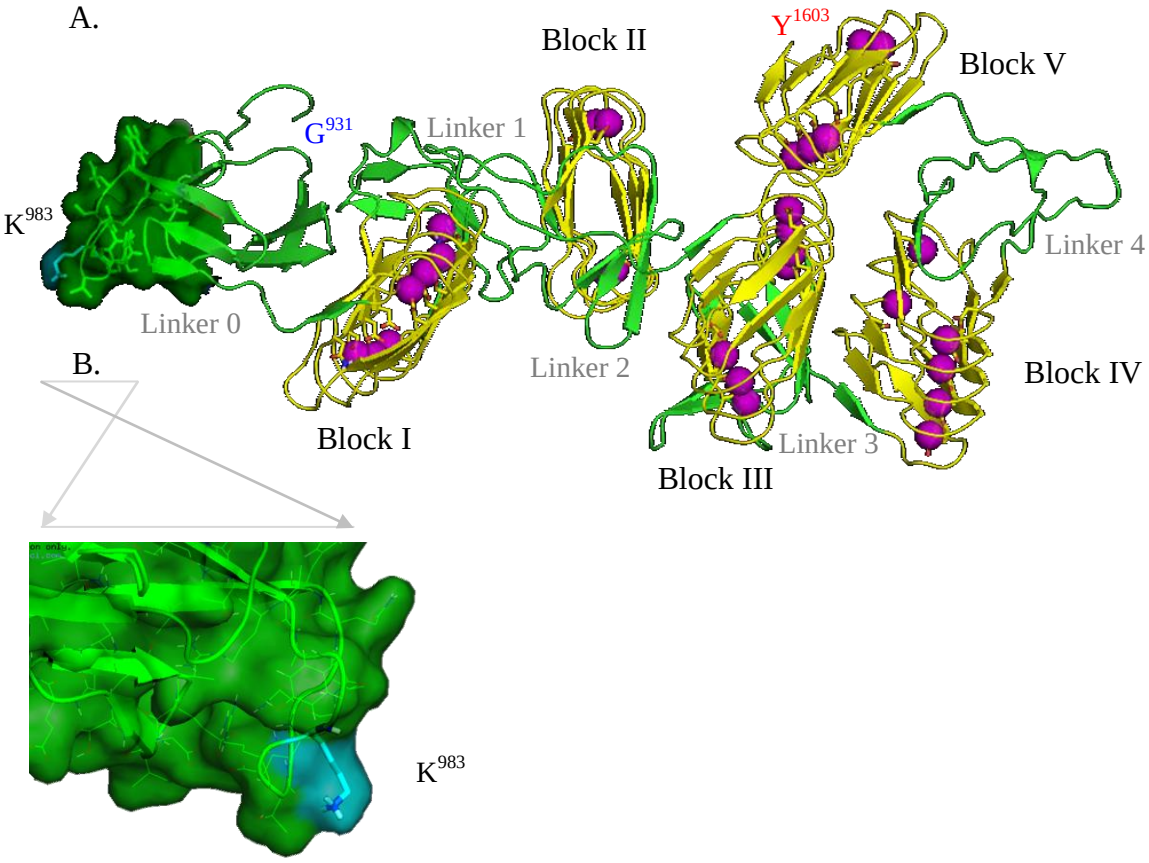


Figure 6.10 Five repeat blocks of the modeled CyaA-RTX

(A) Five yellow repeating blocks of the nona-peptide repeats using its unique β-roll structure to bind to the magenta calcium ions. Linker regions are coloured in green. The repeat blocks are arranged in linear order with block V turned around to interact with block III. Part of the linker 0 surface is shown with acylation site, K^{983}, colored in cyan. (B.) The flexible solvent- exposed loop of the acylation site is magnified showing amine group of the K^{983} residue pointing towards the protein exteria.

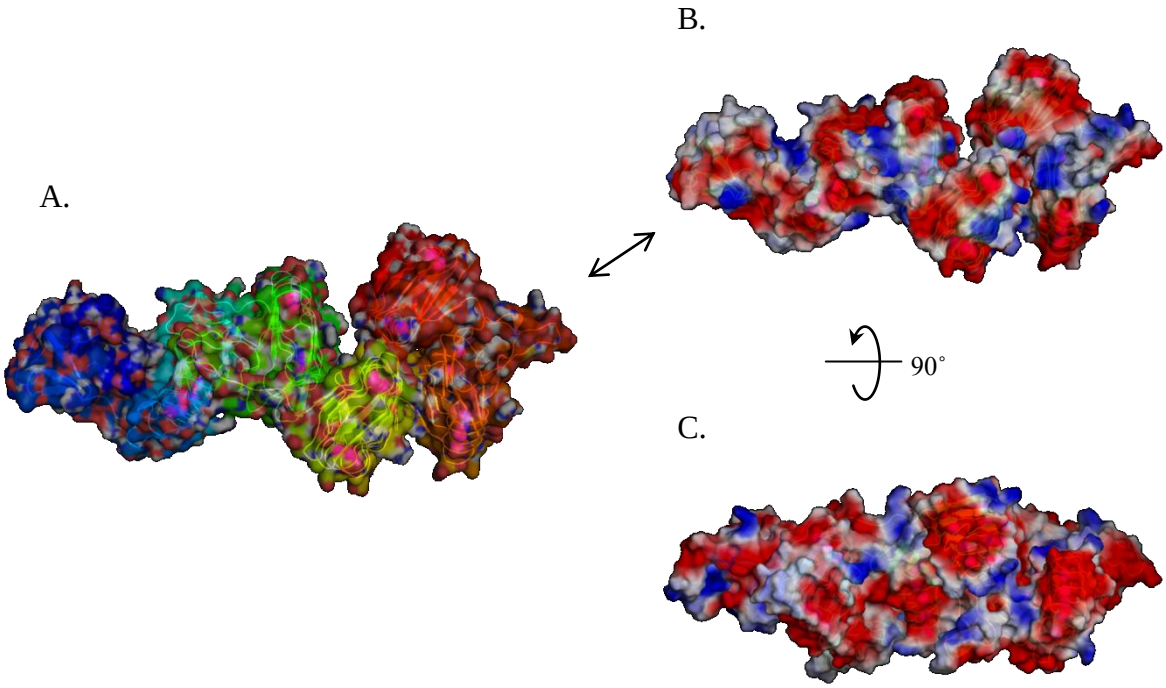


Figure 6.11 Surface of the RTX region

The protein surfaces are shown (A) colored by amino acid position, (B, C) by charge of the local area. Absolute red and blue represent negatively and positively charged regions, respectively. The calculation was performed assuming in vacuum without any solvent interaction. The calcium ions were not included in the computation thus the actual electrostatic charge should be less electronegative. (C) represent 90˚ bottom-up rotation along the X-axis from (B)

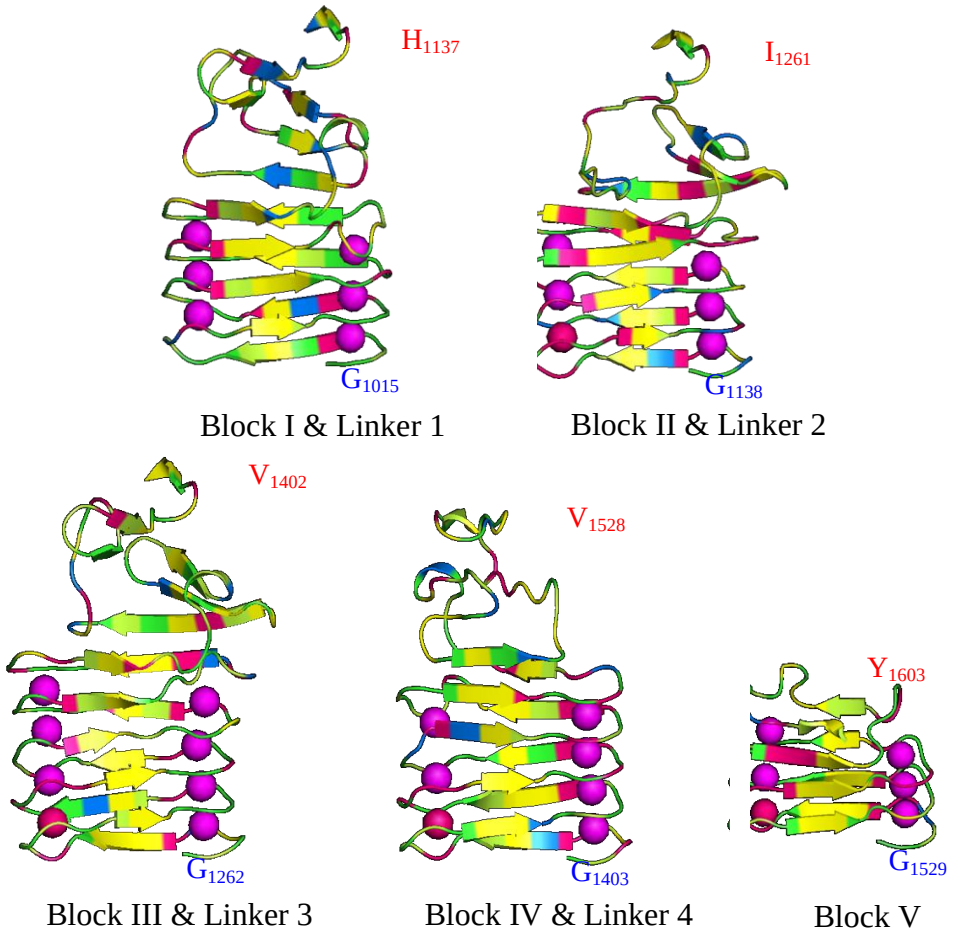


Figure 6.12 Structure of the repeat blocks and linkers.

Main-chain traces showing 5 blocks of repeats from CyaA. The repeat blocks are organized into a right-handed β-roll structure, whereas the linkers tend to form several flexible β-hairpins with some short α-helices. The first residue of the block and the last residue of the linker are indicated in blue and red, respectively. Amino acids are colored in the following color code based on Kyte and Doolittle hydropathy index.

Blue : Positively charged ; K and R
Red : Negatively charged; D and E
Green : Small or hydrophilic ; G, H, N, Q, S and T
Yellow-green : Slightly hydrophobic ; A, M, P, W and Y
Yellow: Hydrophobic : F, I, L and V

Magenta calcium ions were added into the molecule based on their coordination in crystal structure template. Deep pink calcium ions were supplementary added based on their range to the adjacent side chains.

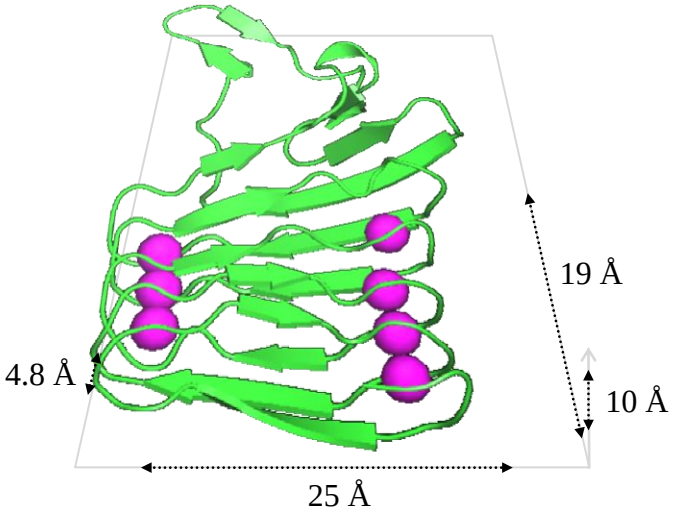
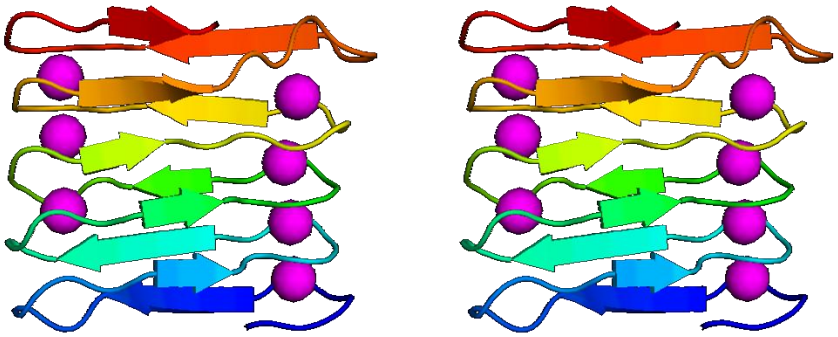


Figure 6.13 Repeat block dimension

The distance between main-chain atoms of each repeat was measured. The results were very similar to one another, thus, block III was selected and shown as an example. Two consecutive nona-peptide repeats would have about 25 Å wide and 10 Å thick. When shaped into β-roll, each stack of repeats are about 4.8 Å away from each other. For block III that has 10 repeats, total length of the stacking repeat is around 19 Å.

A.



B.

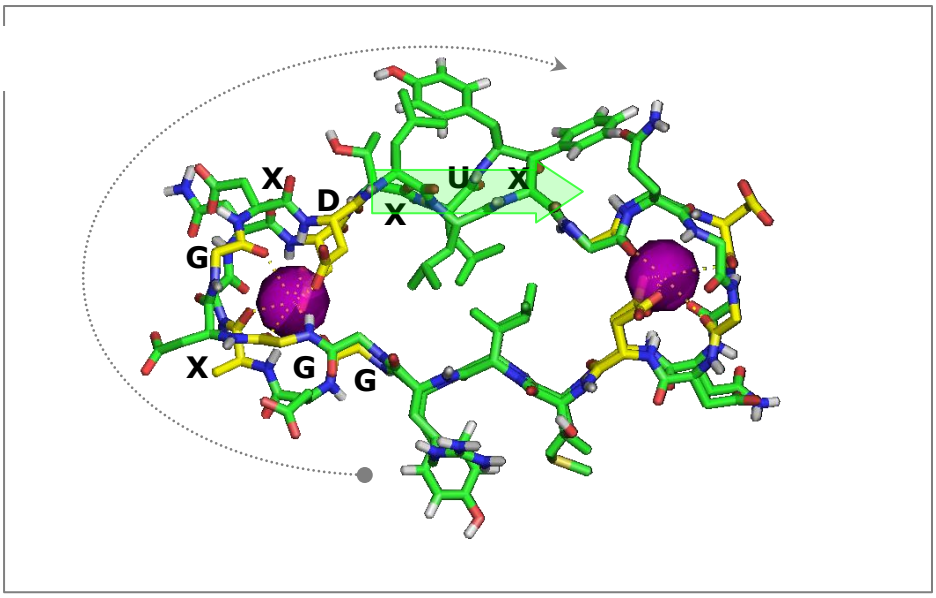


Figure 6.14 The repeat structure

(A) β-roll structure of the block III repeats displayed as cross-eye stereo view.

(B) Cross section of the repeats showing 4 repeats bound to 2 calcium ions. Consensus primary structure, GGXGXDXUX repeat, is indicated, starting from gray dot and going around the repeat clockwise. Typically, the XUX sequence would form a short beta strand (green arrow) that was a little tilted and made the next repeat shifted (in this case, into the paper). The third repeat positioned behind the first is at virtually the same organization as well as the fourth behind the second and so on. These foldings of the repeat finally form a β-roll structure. Oxygen, nitrogen and hydrogen atoms are displayed in red, blue and white, respectively. Residues that interact with calcium ions are marked in yellow.

Table 6.6 Amino acid frequency of each position of the repeat position

	G_1	G_2	X_3	G_4	X_5	D_6	X_7	U_8	X_9	Total
Positively charged R groups										
K	*				1				1	2
R		1	2	1	2		4		6	16
Negatively charged R groups										
D	1	8	8		14	**37**	1		7	76
E		3	4	1			2		1	11
Polar uncharged R groups										
H		1		1	3					5
N		2	1		10	8			1	22
Q	1	1	1	3	1		3			10
S	1	2	2		3		1		3	12
T		2	1				15		2	20
Aliphatic R groups										
A	1	3	18	1	6				3	32
G	**38**	**21**	4	**33**	1				1	98
I	1						2	**7**	3	13
L	1	1	2	2	2		3	**27**	5	43
P										**0**
V	1		1	1	2		10	**6**	1	22
Aromatic R groups										
F							1	**4**	3	8
W				1			2		2	5
Y			1					1	5	7
Sulfur containing R groups										
C										**0**
M				1			1		1	3

* Gray numbers : Amino acids that occur less than 5% at the position (< 3 residues)
Dark gray numbers : Amino acids that occur less than 15% at the position (< 7 residues)
Black numbers : Amino acids that occur more than 15% at the position (≥ 7 residues)
In bold : Consensus sequence at the position

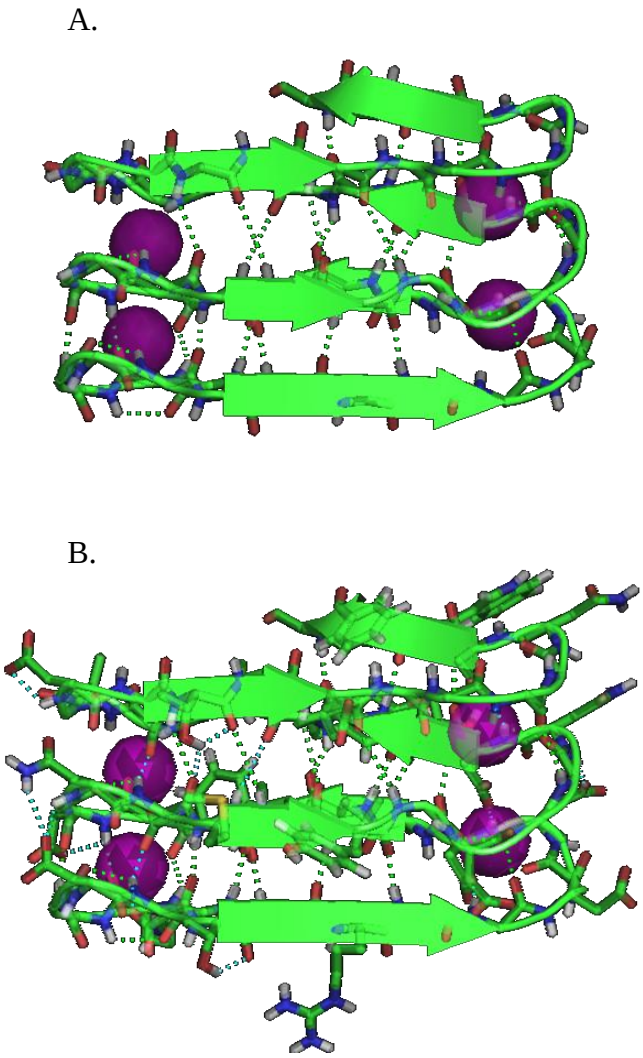
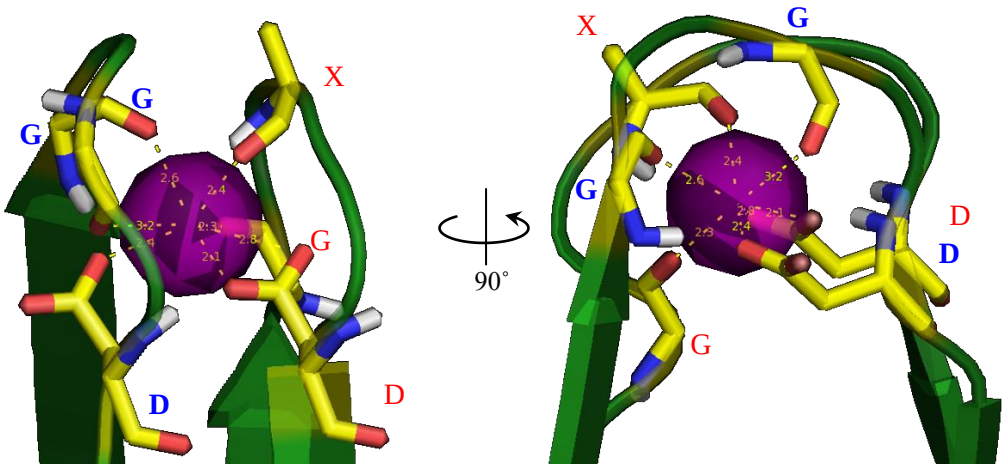


Figure 6.15 Hydrogen bonding in the β-roll structure

As the β-roll formation occurs in each block of repeat, two parallel β-sheets are also formed. (A) Hydrogen bonding of the main-chain interaction was mainly provided from the parallel β-sheets (green dash). (B) Moreover, the hydrogen bonding is also come from main chain to side chain and side chain to side chain interactions (cyan dash). Besides stabilizing the β-sheets, the bonding also built up less flexible loops suitable for calcium ion docking. Supposedly, this hydrogen bonding should contribute significant strength in stabilizing the entire structure, especially at the early step of the protein folding.

A.



Repeat 1 Repeat 3 N-terminal C-terminal

B.

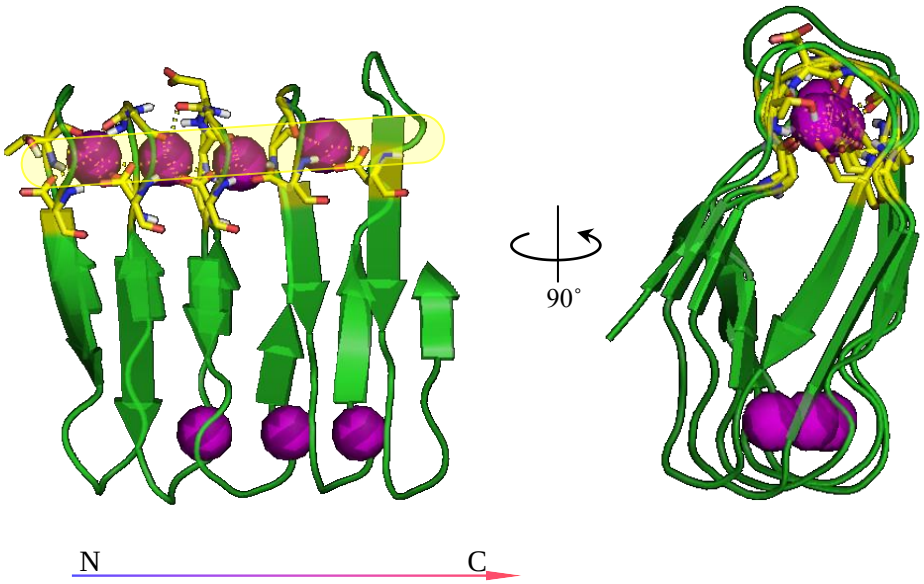


Figure 6.16 Calcium ion interactions

The GGXGXD sequence form a stable loop pointing the aspartic side chain toward interior of the protein provided an electronegative environment, readily for calcium ion docking. (A) In local structure, each nona-peptide repeat could bind to 1 calcium ion using the carboxylic group on the consensus aspartic acid side chain and stabilized with carbonyl group from the peptide bond on the main chain of the G residues (G**G**X**G**X**D**XUX: blue) and the other G and X residues of the second next repeat (**G**G**X**GX**D**XUX: red). All residues that involved in calcium interaction were colored in yellow. (B.) Model of block III repeat acquired 7 calcium ions out of theoretically 8 binding position. The ions are arranged in nearly linear line to each other. The adjacent ions share the same aspartic residue which the two oxygen atoms settled in conjugated form created a resonance π-electron bridge (yellow cylinder).

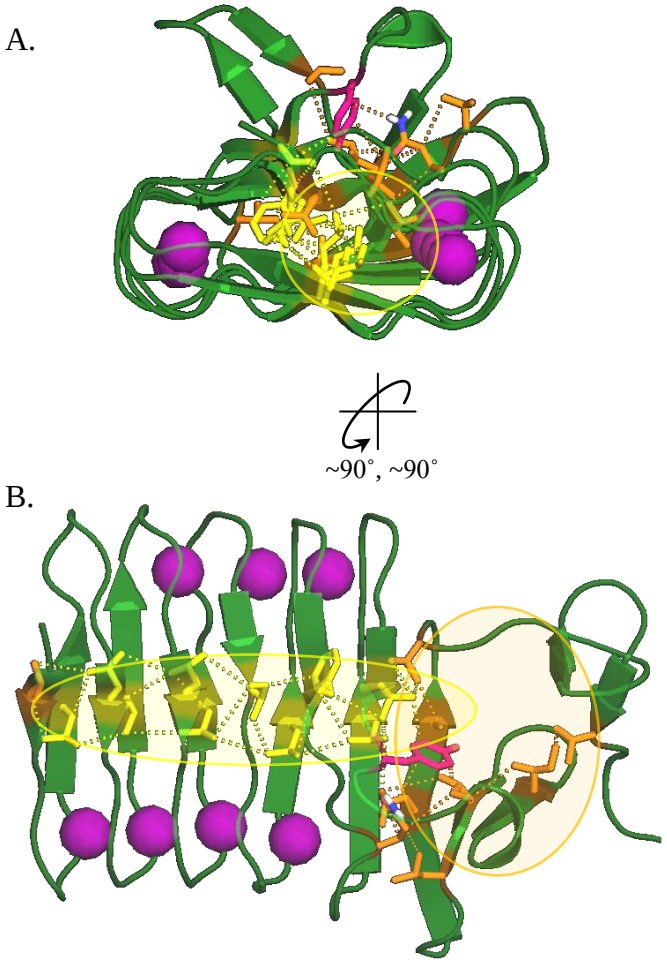


Figure 6.17 Hydrophobic interactions within the block

In the middle of β-roll structure laid a highly hydrophobic region. This region was assembled from U residue of the repeat. (GGXGXDX**U**X) Front view (A) and top view (B) of the repeat block showing yellow intra-repeat hydrophobic amino acids and orange residues of the linker regions that are involved in hydrophobic interaction. Some possible interactions are shown in yellow dash, involved with repeat U residues, or in orange dash for the others. Conserved aromatic residues are colored in red and apparently act as the core residues.

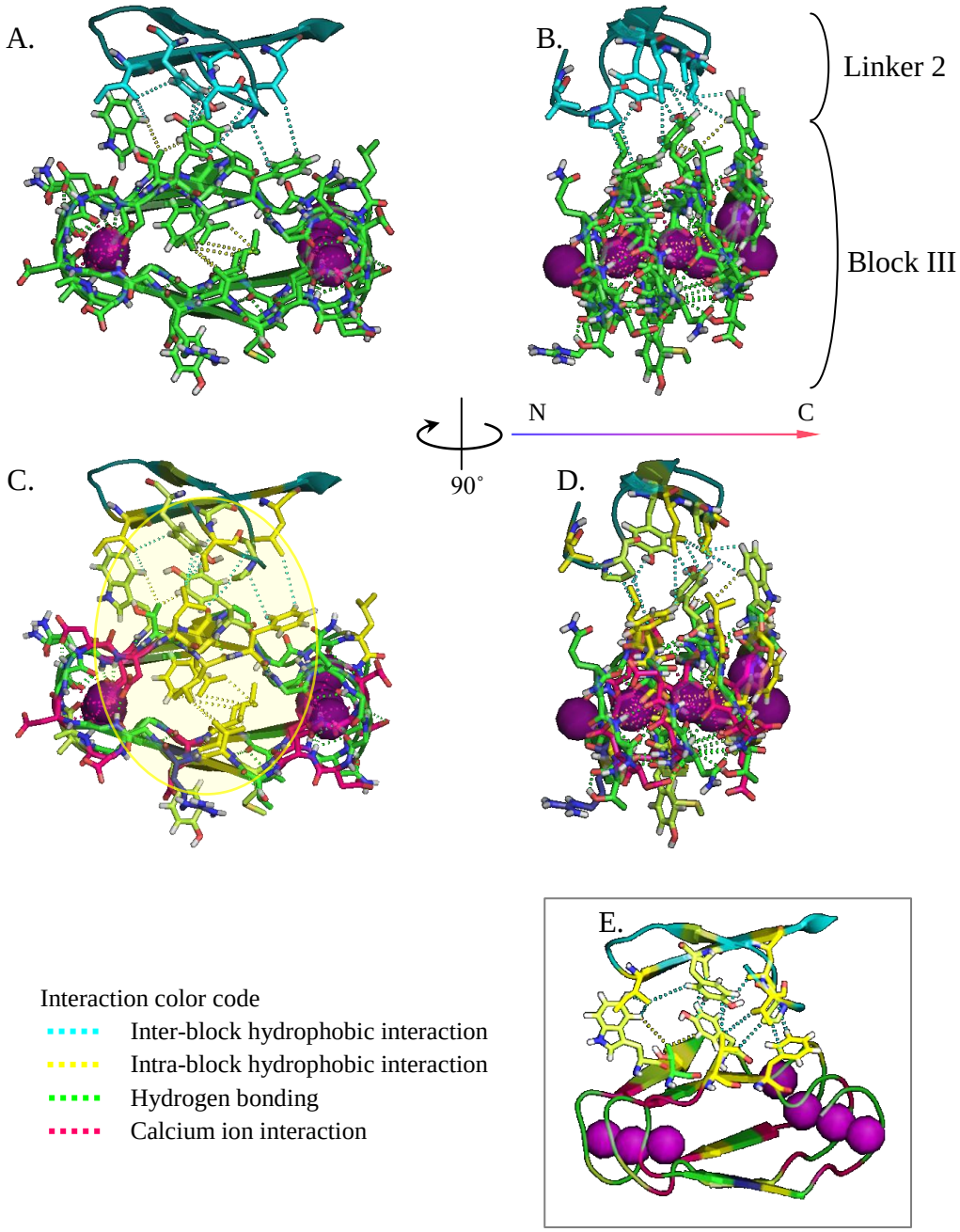
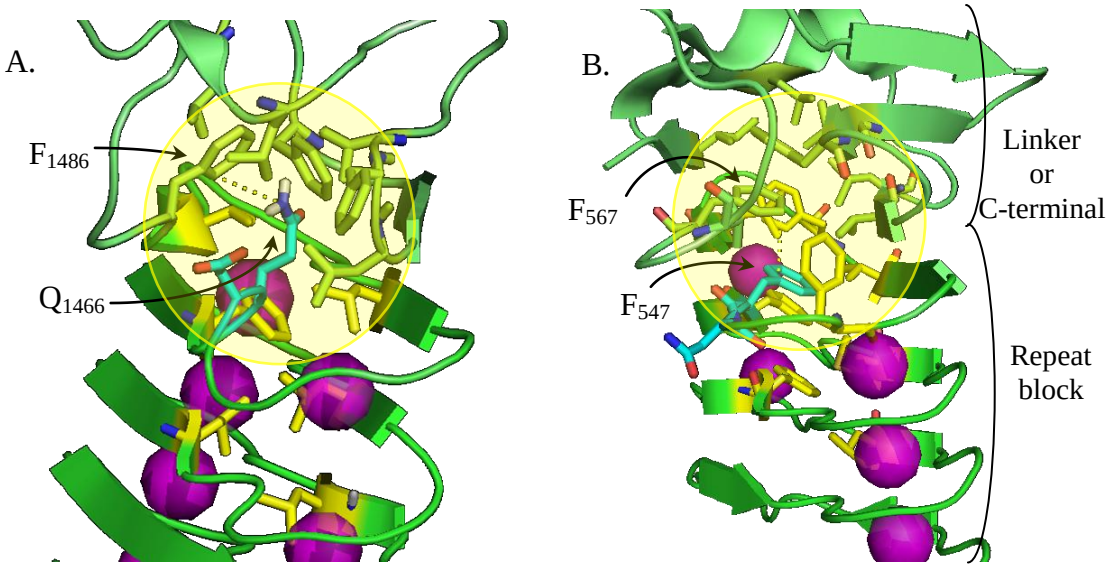
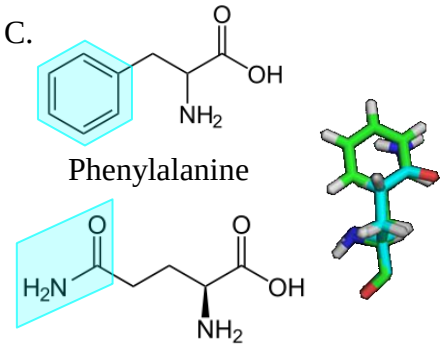


Figure 6.18 Inter-block interactions

The repeat blocks not only form internal block interactions, they also form block to block (linker) contacts where the X residues (GGXGXD**X**U**X**) interact with hydrophobic or, sometimes, hydrophilic residues. (A) and (B) Front and side views of block III repeat (green) forming hydrophobic interactions with the previous linker 2 (cyan). (C) and (D) Front and side views showing hydrophobic core in yellow circle. (E) Perspective view showing only inter block interactions.



CyaA Model LipA (PDBID: 2Z8X)



Phenylalanine

Glutamine

Figure 6.19 Consensus QD sequences
The second last repeat (or third for black V) of every block laid additional QD sequence in the loop region. While aspartic acid residue faces its side chain toward exterior of the protein, glutamine residues point theurs side chain inward the hydrophobic core (yellow circle). (A.) Example of interactions between block V, Q_{1466} with the hydrophobic interior, especially with F_{1486} which is the consensus aromatic residues at the end of each repeat. (B) Crystal structure of *pseudomonas sp*. LipA protein showing the similar region. In this case, the $F_{547}N$ sequence which asparagine side chain points outward and phenylalanine turns toward the hydrophobic core and interacts with the consensus aromatic F_{567}. Hydrophobic amino acids of the repeat, hydrophobic amino acid of the linker and the QD or FN sequence are shown in yellow, yellow-green and cyan, respectively. (C.) Chemical structures of phenylalanine and glutamine. Possible delocalized π-electron plane is highlighted in cyan. The three-dimension superimposed structure is shown to the right with phenylalanine in green and glutamine in cyan.

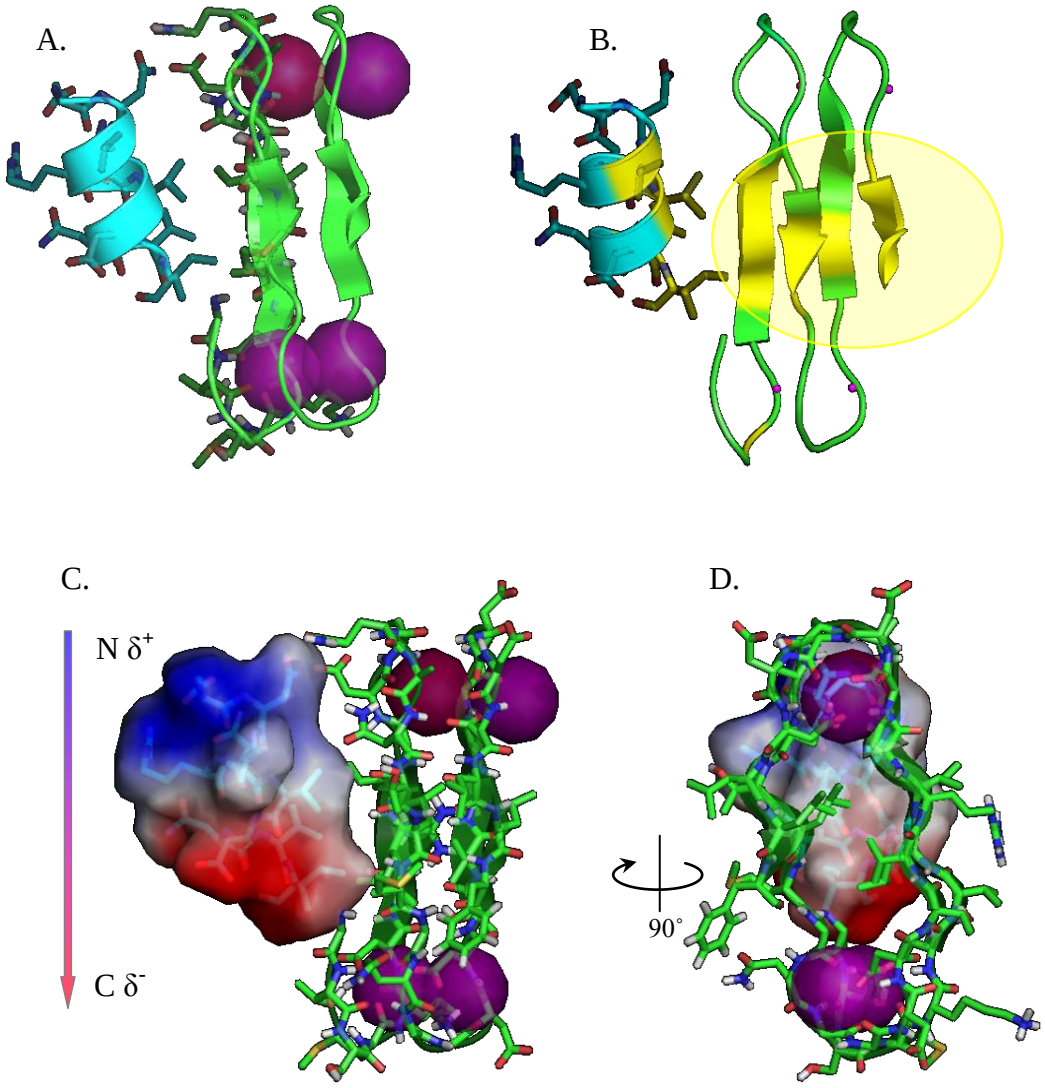


Figure 6.20 Repeat capping by consensus amphipathic helix

The model was generated from block III repeat and amphipathic helix from linker 2 as an example of the putative interaction between the repeat block and consensus amphipathic helix at the end of every linker. (A) and (B) display 11-rsidue amphipathic helix vertically positioned from N-end to C-end. Turn the hydrophobic surface toward the repeat block, the helix using its hydrophobic residues (yellow) to form hydrophobic core (yellow circle). (C) and (D) Side and back views representing electro static potential surface of the helix. Both side chain electrostatic and helix dipole moment were used for computation. The N-terminal and of the helix is electropositive whereas C-end is electronegative. On the other hand, the contact site is nearly neutral (white).

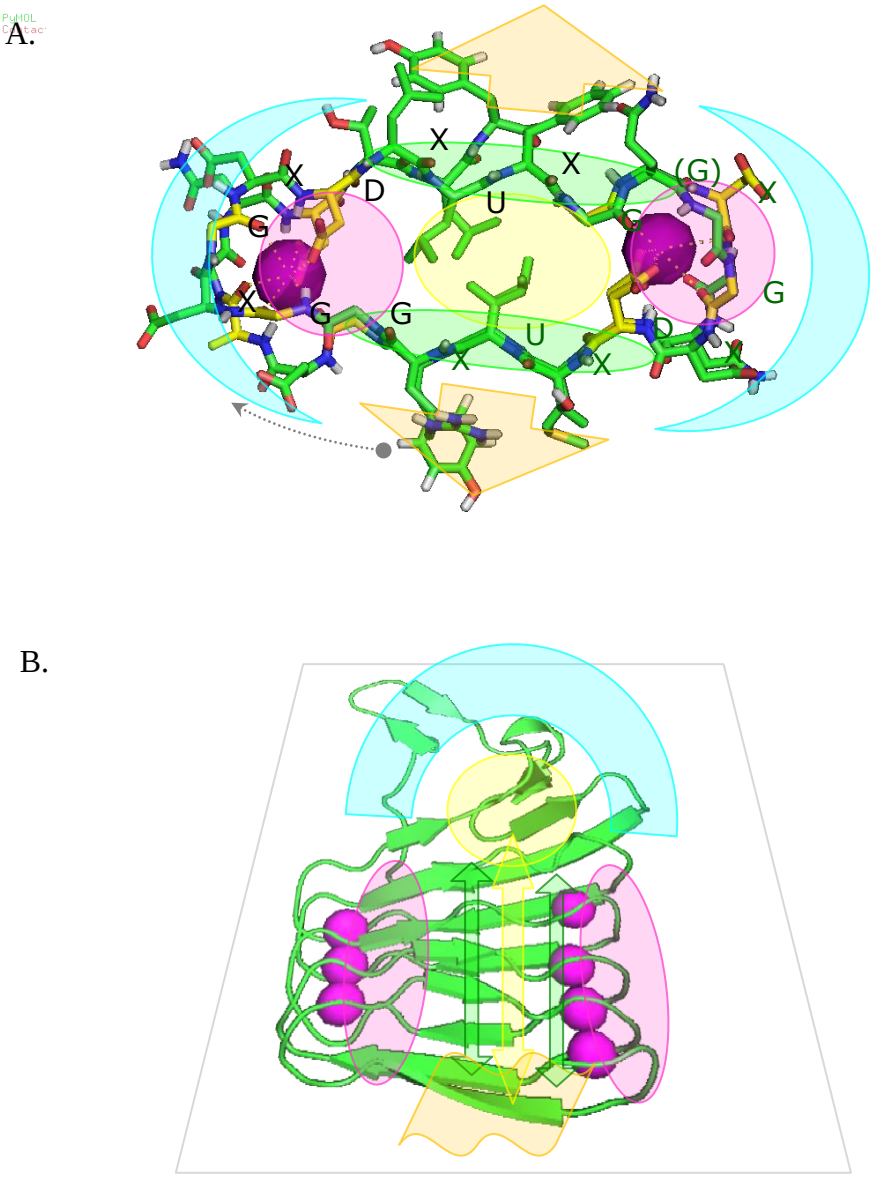


Figure 6.21 Summary of the forces that stabilize the β-roll structure

Force recapitulation of one repeat block and linker of the RTX. (A) Two stacks, four repeats, are shown for simplicity. Sequence starts from gray dot and rotate clockwise. First force that contributed to the structure is hydrophobic interaction. The hydrophobic core (yellow oval) is formed in the middle of the repeat, stabilizing an internal part of the protein, while the second force, β-sheet hydrogen bonding (green oval), is built up flanking the hydrophobic center. Utilizing its negatively charged amino acid, the positively charged calcium ions are drawn to the site and bind next to the hydrophobic core with the less flexible loop and consensus aspartic side chain (magenta circle). This third force makes the structure more strong, rigid and, the most

important, triggerable. Fourthly, the exterior part of the loop which is rich in hydrophilic or small residues (cyan curve) can easily solvate the molecule in the water based environment. With all these interactions the repeat block should be stable enough to accommodate the interactions with next adjacent block of repeat (orange arrow). (B) Perspective of the repeat with the linker is drawn. Including the neighboring linkers, additional forces are shown. Alongside of the hydrophobic interaction (yellow arrow), β-sheet hydrogen bonding (green arrows) and calcium interaction (magenta oval) previously shown, hydrophobic capping at the N- and C-terminal end is clearly demonstrated. The C-end of the repeat block is laid with an additional hydrophobic region (yellow oval) provided from some residues from the linker that obviously interact with the hydrophobic core of the repeat. The linker of outer surface was covered with hydrophilic residues to promote a stable solvent interaction (cyan curve). At the same time, N-terminal end of the repeat is stabilized by the hydrophobic contact with the amphipatic helix from the previous linker (orange wave).

6.3 Molecular modeling of outer membrane secretion machinery, CyaE.

CyaE modeled structure was constructed using *E.coli* TolC as a template. The modeling parameters were showed in appendix B. Molecular model of CyaE is shown in **Figure6.22.** CyaE is a homotrimer outer membrane integrated protein. Internal structure of one CyaE monomer reveals an electropositive channel (**Figure 6.23**). The cell exterior channel, exit of CyaE, has opening size approximately 20 angstrom, on the other hand, the periplasmic channel entrance has pore size less than 9 angstrom (**Figure 6.24**).

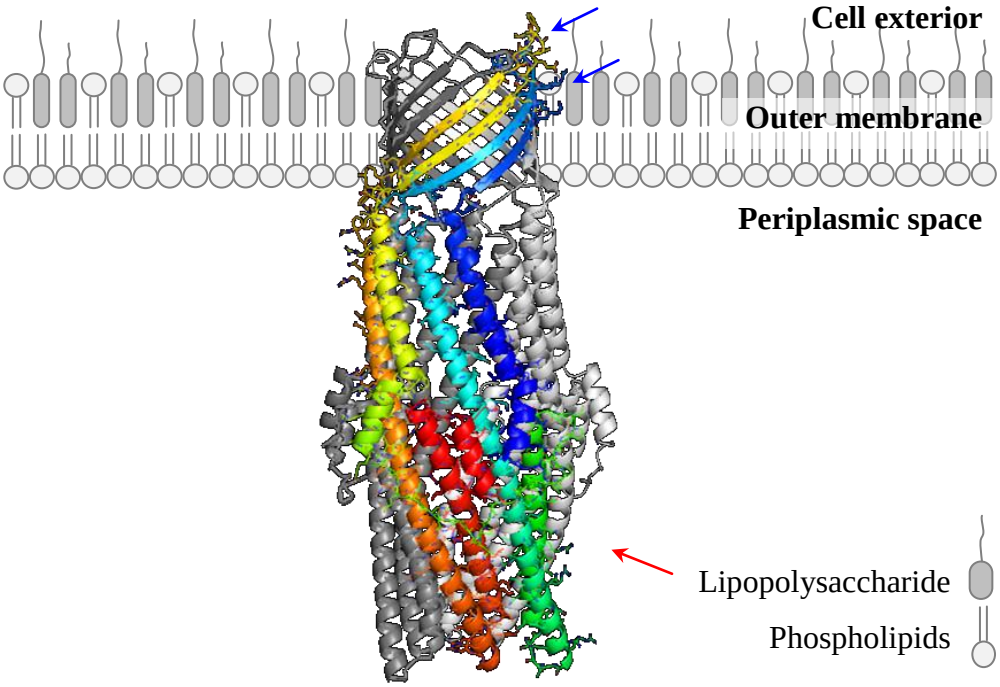


Figure 6.22 Side view of the trimeric CyaE transporter

A subunit of CyaE is shown as rainbow color with the N-terminal end in blue and the C-terminal end in red. The two other subunits are colored in light gray and dark gray. The three most diverse regions are shown by arrow; two exterior loops (blue) and equatorial region (red). The two exterior constructed loops would depend on their RTX protein product. The equatorial region would has the configuration appropriate for theirs inner membrane transporter, CyaBD complex, docking procedure.

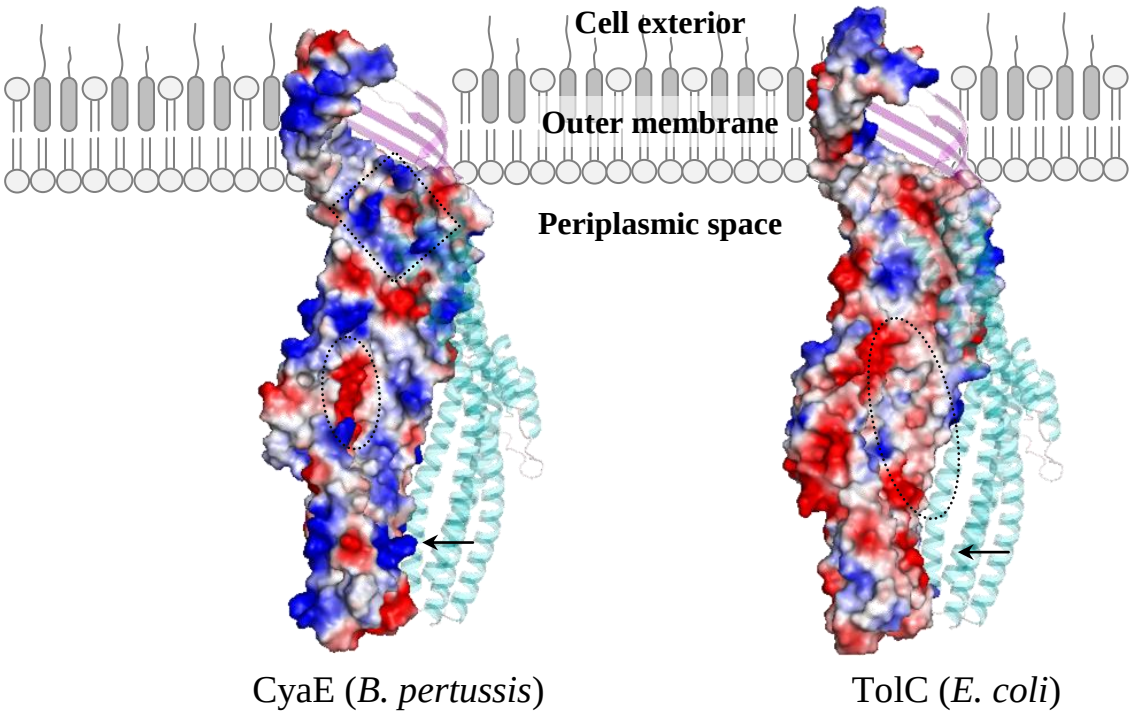


Figure 6.23 Vacuum electrostatic of a subunit of CyaE compared with TolC

A subunit of each protein was used for surface electrostatic computation while the second subunit is shown in cartoon representation with helical structure in cyan and sheet in magenta. The third subunit in front is omitted, thus, this figure illustrats the internal morphology of the tunnel. The potential positive electrostatic and negative electrostatic areas are shown in blue and red, respectively. A smaller electronegative surface is illustrated for CyaE where the larger area is shown for TolC (circle). This provides CyaE a large electropositive area. The four unique arginine residues in CyaE (R^{136}, R^{143}, R^{344} and R^{356}) are shown in diamond. These residues form up an electropositive pocket near the outer membrane pore opening. The arginine and glutamic acid periplasmic gate keeper residues of CyaE and TolC, respectively, are marked by arrows.

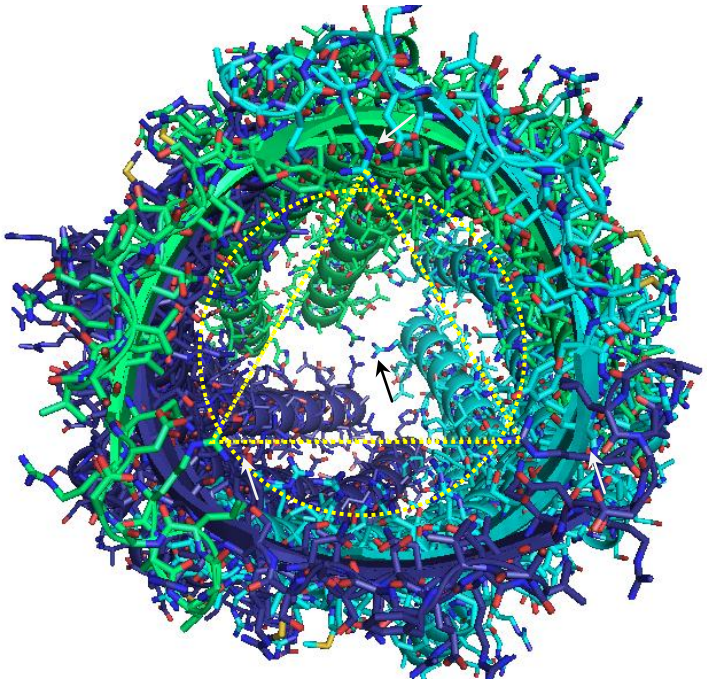


Figure 6.24 Model of *B. pertussis* CyaE outer membrane transporter

View from outside of the outer membrane of trimeric CyaE. Three subunits are represented in green, cyan and blue. The distance between each pairs of the three guanidinium carbons of arginine residues at the exterior pore opening is about 19 angstroms (white arrows, yellow triangle). Consequently, the diameter of the pore opening (yellow circle) would be approximately 20 angstroms. In the middle, on the periplasmic pore opening side, the three gate keeper arginine residues are shown (black arrow). The distance between these residues at the open stage should be somewhat less than 9 angstroms. With these structure, CyaA would enter the transporter as a low complex peptide and depart with more complex, bigger structure but not as large as β-roll bundle.

CHAPTER VII
DISCUSSION

Adenylate cyclase-hemolysis toxin (CyaA) was secreted from *Bordetella pertussis* as one of the key virulence factor. However the insight understanding of the toxicity mechanism of this toxin is still needed to be elucidated. CyaA is firstly translated inside the bacterial cell as non-acylated pro-toxin. Then undergo post translation modification by palmitoylation at Lys^{983} by CyaC, an acyltransferase. The protein is subsequently transported outside the cells by ATP-dependent type I secretion system through CyaB, CyaD and CyaE complex. Outside the cells, CyaA toxin binds to calcium ions in order to exert its cell invasive activity.

The organization of the domain in CyaA is still elusive. Nonetheless, the disorder prediction (**Figure. 6.2**) and secondary structure prediction (**Figure. 6.3**) gave hints toward what domain organization of CyaA should be. As consistent with previous studies, AC domain is a single standalone domain connected by a lengthy disorder region. The other disorder region at position 846-860 was a putative domain connector region. This agrees well with the secondary structure prediction that designates the helices rich region ranging approximately from position 420 to 810. While the subsequent sequence tend to have more beta structure. This suggested that the PF domain is actually divided into two sub-domains, the hydrophobic pore-forming sub domain, approximately residues 400-850 and RTX sub-domain, approximately residues 850-1706. The RTX sub-domain contains acylation site (K^{983}) and secretion signal (residues 1658-1706).

The first hallmark of the toxin is the needs of acylation activation at certain lysine residue (K^{983}). Without this activation the toxin is fail to invade the host cells.

Previously, CyaA-PF was expressed in *E.coli* system as a soluble protein which can be acylated by CyaC in both *in vitro* and *in vivo*. In this work, pCyaAC-RTX was successfully constructed using site-directed mutagenesis technique. Both CyaA-RTX and CyaC were expressed as soluble proteins (**Figure 5.4**) comparable

with CyaAC-PF and CyaA-RTX from pCyaA-PF and pCyaA-RTX, respectively (**Figure 5.2**). These suggested that both proteins were in native folded conformation. In addition, *in vitro* pNPP hydrolysis activity assay of CyaC from pCyaAC-RTX showed that the protein was active and have specific activity comparable to soluble CyaC from pCyaAC-PF (**Figure 5.5, Table 5.1**). The yield of the soluble CyaA-RTX was marginally higher than those of CyaA-PF and was unaffected by culturing temperature (**Figure 5.2, 5.3**). However, LC/MS/MS results (**Figure 5.6**) showed that the CyaA-RTX was not acylated even co-expressed with CyaC.

It appeared that even CyaA-RTX was co-expressed as a soluble protein with active CyaC, it still cannot be acylated. These can be implied that the truncated region, residues 400-750 in the pore-forming region of CyaA, was important for CyaC in order to be able to acylate CyaA. This can give rise to 2 hypotheses; Firstly, CyaC directly recognized the truncated region as binding site. Secondly, the hydrophobic pore-forming region plays important role in providing the correct folding for the other region that is the CyaC recognition site.

The hemolysis assay reveal the even the CyaA-RTX was un-acylated. It was slightly effect the morphology of the sRBC (**Figure 5.7, 5.8**). The star-like (crenated) sRBC seldom observed. These suggested some sRBC were dehydration shock or the protein may somehow interfere with the lipid membrane. On this basis, the hemolysis assays using two type of CyaA were performed (**Figure 5.9**). The reaction 5 and 6 assumed that once CyaA absorbed on the target cell, they formed oligomer complex in order to penetrate to the cell membrane. Present of CyaA-RTX may interfere either in oligomerization step or penetration step and reduce the hemolysis activity. For reaction 7 and 8, they were assumed that CyaA-RTX can occupied the target molecule on the call membrane inhibit CyaA-PF activity. However the results from all 4 reactions show that CyaA-RTX cannot competes CyaA-PF in neither way.

The second hallmark of CyaA is the RTX region. Among RTX protein family, there are only few proteins that contain more than one repeat block. However, up till now, CyaA is the only identified toxin that contains 5 repeats block divided by 4 linkers. Different levels of analysis ranging from primary structure study to tertiary structure sturdy of the RTX region are discussed.

From linkers alignment (**Figure 6.4**), the results reveal some interesting sequences. Firstly the GIXADL sequence in the middle of linker I, II and III these sequences including the VXK 4 amino acid away from the first sequence. However, predicted secondary structure of these GIXADL sequence share no comparable organization. To rule out chance of occur as random sequence, 22 random sequences were generated from RandSeq software (amino acid propensity were based on standard table) and check on PROSITE database. Statistic data were shown in Figure 6.5. It seem that GIXADL have significant change of occurrence. The protein crystal structure that contains this sequence were reviews. (**Table 6.3**) Nonetheless, there was no conserved 3 dimensional structure among these proteins.

Unlike GIXADL sequences, DXVXXUENU sequences, where U was any hydrophobic residues, lay at the end of each linker tend to had helix structure from the prediction. (**Figure 6.3**) These sequences have at least one hydrophobic amino acid divided by two other residues, which may be suggested that they would form amphipathic helix. Nevertheless the molecular modeling fail to create these structure since the helices was absent from the template protein. Helical projections of these predicted helices reveal amphipathic helices for every sequence. Although linker I had some histidine spanning the sequence, it still had hydrophaty of 2.4 which is more than those of linker II and III which had 1.86 and 1.44, respectively. The amphipathic character may implied that the helices play a role in structural integrity by act as capping structure using its hydrophobic surface to interact with hydrophobic core of the repeat while the hydrophilic surface pointes toward protein exterior. Recent studies by show that lack of this sequence makes the repeat block failed to fold properly.

An additional noticeable point is that every repeat block ends with aromatic amino acid, Y_{941}, Y_{1090}, Y_{1213}, Y_{1355}, F_{1487} and F_{1623}, respectively. Comparing with crystal structure from *pseudomonas sp.* MIS38 I.3 lipase (PDBID: 2z8x) and *serratia marcescens* LipA (PDBID: 2qub) reveals that the aromatic residues in this position are responsible for stabilizing the hydrophobic core of the repeat. In these two proteins the aromatic residues at the end of the repeat form a stack of hydrophobic aromatic interaction at the end of the structure. The stacking contains 3-4 of the stacking unit which compromised the larger globular structure of the proteins. CyaA

resolves this problem in the other way, by comprise of a single aromatic residues surrounded with many small aliphatic hydrophobic residues. (**Figure 6.17**)

Molecular model of the RTX region from CyaA displayed linear fashion arrangement (**Figure 6.9**). The model parameters (**Section 6.2.2**) are in acceptable range. Although, the model has fairly low structure average packing thus slightly high solvent accessibility and several disfavored residues in the ramachadran plot but all of these are still in conventional range (**Figure 6.8**). On the other hand, the overall interaction energy, atom pairwise energy, bond length and bond angle are in standard value with mostly above average stereochemical parameter (**Table 6.4**). These mean that the model is good enough to represent the configuration of the RTX region.

Molecular model start from two discrete repeats at position G^{931}, through all four and a half repeat and four linkers which finally end at Y^{1603} in the middle of repeat block V (**Figure 6.9**). The linker 0 region forms up a globular domain with some flexible solvent exposed loop. This allows the acylation site, K^{983}, to be exposed to the environment readily to be palmitoylated by CyaC (**Figure. 6.10B**). The following repeat blocks were joining together in an I, II, III and IV order. The last block of repeat was bended back to in between blocks II and IV.

Without calcium ions, the region is highly electronegative (**Figure 6.11**). This would make the newly synthesized peptide very flexible and soluble. While upon calcium binding, the positively charged calcium ions would neutralized the charge, therefore, encourage more compact protein folding.

The approximated dimension for 10 repeats blocks (block III and V) are 25 angstrom wide, 19 angstrom long and 10 angstrom high (**Figure 6.13**). The smaller repeat blocks i.e. block I and II, are 25 x 15 x 10 angstrom. The linker regions are diverse so that they were not included in the measurement.

As previously mention, CyaA contains 5 blocks of repeats divided by short linkers. Each repeat block can bind up to 6-8 calcium ions (**Figure 6.12**). Each position of consensus nona-peptide repeat had its own structural roles (**Figure 6.14B**). The first G (**G**GXGXDXUX) position is the most conserved residue among 9 positions (**Table 6.4**). The residue sit aside the β-strand of the previous repeat and because of the highly flexible nature, it would abolish the β-sheet structure and twist the main chain into flexible loop arrangement, facilitating the turn like-loop structure.

The side chain of this residue was pointed toward the calcium binding pocket; thus, the non-glycine residue at the position would abolish the calcium binding activity (**Figure 6.16**). Obviously, the non-glycine at the position are only arise at the second last repeat that are facilitated with consensus extended QD sequence and tilted the loop so that the side chain of the non-glycine bended away from the ion binding pocket. The important of the consensus QD will be discussed later on.

The second G (G**G**XGXDXUX) was somewhat less stringent. Several aspartic acid and some polar residues were allowed at the sites which would help recruiting and stabilizing the positively charged calcium ions. The X residue (GG**X**GXDXUX) was mostly taken up by A (40%) and D (17.8%). The position lies at the edge of the structure and exposed to the environment. This position provides bending in the main chain to build up a tight and selective calcium binding loop. And also crucial for providing degree of firmness to the flexible loop hence the GGGG sequence was never observed in the repeat region. If the previous G position was not occupied by D or E residues, the acid residues sometimes take this position instead. The third G (GGX**G**XDXUX) bends its carboxyl group inside the molecule and delivers a short stretch to the loop that propels the next residue away from the first G position. Consequently, they creates a negatively charge pocket instead of β-turn (N to N+3) hydrogen bonding. The second X residue (GGXG**X**DXUX) was mostly D or N residues. This solvent exposed position would facilitate in calcium recruiting to the binding site and assist the region during the pre-molten globule state. In that condition, the hydrophobic core may repulse the consensus D resides from the center of the molecule which the molecules may fail to build up the D facilitated calcium binding pocket

The consensus D position (GGXGX**D**XUX) was occupied only by D (82.2%) or N (17.8%) residues. Which the N residues hardly ever appear in the middle of the block, instead, they normally appear at the first or last stack of the repeat. The residues side chain are pointed toward the interior of the protein and served as a key calcium binding residues. The next X residues (GGXGXD**X**UX) often be T (33.3%) or V (22.2%) amino acids. The side chains of this position turn toward protein exterior. The compact structure of T and V would provide tight interaction to the adjacent flanking repeat block. The hydrophobic U residues (GGXGXDX**U**X)

including L (60.0%), I (15.6%), V (13.3%) and F (8.9%), used to form up a hydrophobic cluster in the middle of the repeat block. And the last X residues (GGXGXDXU**X**) offer interaction to the nearby repeat block. No wonder that of the entire repeat region, there were no P and C residue found. The reason for P was that it would probably destroy β- sheet structure and made a 'too tight' turn for the G-rich loops. The C residue was absent for the other reason. Actually, the entire CyaA toxin does not contain any C residue. As if the toxin exposed to the oxidative environment outside the cells, it would be burden to have uncontrollable disulfide bonding of the C residues.

For one set of repeat, the GGXGXD sequences **(GGXGXD**XUX) form a hollow turn-like loop for calcium ion binding, while the XUX sequences (GGXGXD**XUX**) built up a short tilted β-strand that helps in structural integrity (**Figure 6.14B**). Although, the loop region consisted of many glycine residues, the loops are quite stable and unmovable. The loops stringent come from the tightly bound calcium ions and the compact structure of stacking repeat. A stack of repeat consists of two consecutive nona-peptide repeats arranged themselves as a rectangle block. Each pair stacked on the prior repeats and shape up a tower of repeats. Each stack was strengthening by hydrogen bonding typically from parallel β-sheet (**Figure 5.15**).

Hypothetically, RTX region of the CyaA can binds about 35 calcium ions from 45 repeats, however, the model present 29 calcium ions because the position of the calcium ion came from the template crystal structure which the repeat block is a bit shorter than those of CyaA. In local configuration, each nona-peptide repeat could binds to 1 calcium ion using the carboxylic group on the consensus aspartic acid side chain and stabilized with carbonyl group from the peptide bond on the main chain of the G and X residue (**G**G**X**GXDXUX) and the two other G from the previous repeat (G**G**X**G**XDXUX) (Figure. 5.16A). These stabilizing interactions are very important since calcium ion has 2 positive charges whilst carboxylic group has only one negative charge. The calcium binding sites ae linked together by sharing the same D residue. The electron bridge formed this way enable electron delocalization which compact the loop and repeat structure (**Figure 6.16B**).

Hydrophobic interaction is formed between U residues (GGXGXDX**U**X) of the adjacent repeat. Unlike other interaction, hydrophobic interaction is settled at the early phase of the protein folding and form up the molten globule. Start by the hydrophobic interaction inter linker region (**Figure. 6.17B orange circle**). Then followed by the repeat region (**Figure 6.17A and B yellow circle**). The key residues that link the interaction together are the aromatic residues at the end of the repeat block.

The consensus QD sequences were suspected to be the loop stabilizing sequence. These additional sequences were found at every second last repeat of the block. Superimpose structure between CyaA and LipA reveal some hydrophobic interaction between aromatic residues at the end of the repeat and linker region which the Q residues of CyaA (superiposed by F residues in LipA) were in the middle of the interaction. These two residues can substitute each other because they both have delocalized electron plane in their structure.

The final force that involved in the RTX folding was inter-repeat block interaction. The X and X residues (GGXGXD**X**U**X**) would interact with the residues from adjacent block, mostly, hydrophobic interaction (**Figure 6.18**).

To sum up, the force that built up the RTX structure are, firstly, hydrophobic interaction formed in the middle of the repeats structure. The interactions provide fundamental structure of the region. This is later capped by amphipathic

helices at the N-terminal and linker hydrophobic core at the C-terminal. Secondly, the ion bridges network at the calcium ion binding pockets. Third, Hydrogen bonding provided from the β-sheet structure covering the hydrophobic core. The forth interactions are provided by inter repeat block contact (**Figure 6.21**).

In order to study the folding pathway of the CyaA involving RTX domain, the structure of outer membrane protein transporter, CyaE, was modeled. The CyaA peptide was translocated from bacterial cytoplasm by bound to an ATP-dependent CyaBD complex. The peptide passes through the periplasmic space inside the channel that linked to the outer membrane protein, CyaE. Using crystal structure of CyaE homologue, *E. coli* TolC, as a template, molecular modeling reveals significant features of CyaE that indicate some key point of CyaA folding.

Same as TolC, CyaE is a hollow cylinder homotrimer. The tunnel in the periplasmic space is consisted of helices structure, whereas the membrane spanning region is built up from β-barrel structure (**Figure 6.22**). The diverse equatorial region and two exterior loops region are presumed to be the region specific to CyaD and CyaA, respectively. Inside the channel, CyaE provides adequate electropositive environment (**Figure 6.23**) which help stabilizing the high negatively charged RTX domain of CyaA while prevent folding of the domain. Near the cell exterior exit, a group of 4 arginine residues, R^{136}, R^{143}, R^{344} and R^{356}, formed up a positively charged pocket. This pocket would help attracted the RTX region toward cell exterior, while, repulsed the large positively charged ions from entering the channel.

The opening of the periplasmic gate of CyaE was somewhat less than 9 angstroms guarded by R^{429} gate keeper residue (Compared with *E.coli* TolC). Therefore, the peptide that can pass through the gate should be in a simple structure, for example, tight helix and random coiled. Note that the helix consisted of only alanine would have diameter around 6 angstroms, while the helix consisted of only arginine or tryptophan would have diameter approximately 16 angstroms. The more complex structure or the structure that required inter-strand interactions such as parallel and anti-parallel β-sheet would be difficult to pass through the gate. This means that CyaA folding is very simple during entering into CyaE.

On the other side, cell exterior pore opening is about 20 angstrom in diameter (**Figure 6.24**). This allows more complex construction or small globular

structure to pass. Nevertheless, β-roll structures are still too large to get through. This suggested that the folding state of CyaA at CyaE pore opening is not yet reach molten globule. For RTX-region, the expected structures are partially unfolded structure. The repeat region is unfolded with very small amount of β-sheet observed. The calcium binding pockets are unorganized since aspartic acid residues turn away from each other. The Linker region may be partially folded. Some region may turn toward the repeat region to help stabilized the highly positively charged region. Conversely, the amphipathic helices at the end of each linker are folded.

Protein folding is composed of different step and level of interaction, the local adjacent residues interaction and longer inter-region interaction. Folding mechanism of the CyaA toxin was proposed based on molecular modeling. The two long α-helices secreting signal was the first to be secreted however, this region is not significant for protein folding. The linker region especially linker 5 were among the first to fold since the driven force on these region come from solvation and hydrophobic interaction. The folding is start in C to N-terminal direction. The small globular linker region served as platform for folding of the triggerable repeat structure. Because of very tight calcium binding pocket with delocalized electron network and the fact that there are two calcium binding network in one β-roll structure, the calcium binding would occurred in a nearly instantly. When the block was complete, it was capped by an aliphatic helix at the N-terminal. This is also draw the linker close to the folded repeats. The folded repeat blocks served as a scaffold for the next repeat block through cross block interactions. The folding would continue until the block I was finished then the rest of the protein would build up using the RTX region as nuclei.

In conclusion, CyaA were transcripted along with CyaC, CyaB, CyaD and CyaE. After translated, CyaA was palmitoylated by CyaC and secreted out of *B. pertussis* cell through CyaBDE complex as a low complex protein. In calcium rich environment such as human trachea, the RTX region bound to calcium ions and undergoes conformational change. The formed β-roll structure were rigid and served as platform for other region of the toxin. The reorganization should be fast. The folding direction was C-terminal to N-terminal. The active CyaA now can bind to the receptor on target cell and can promote cell invasion activity (**Figure 7.1**).

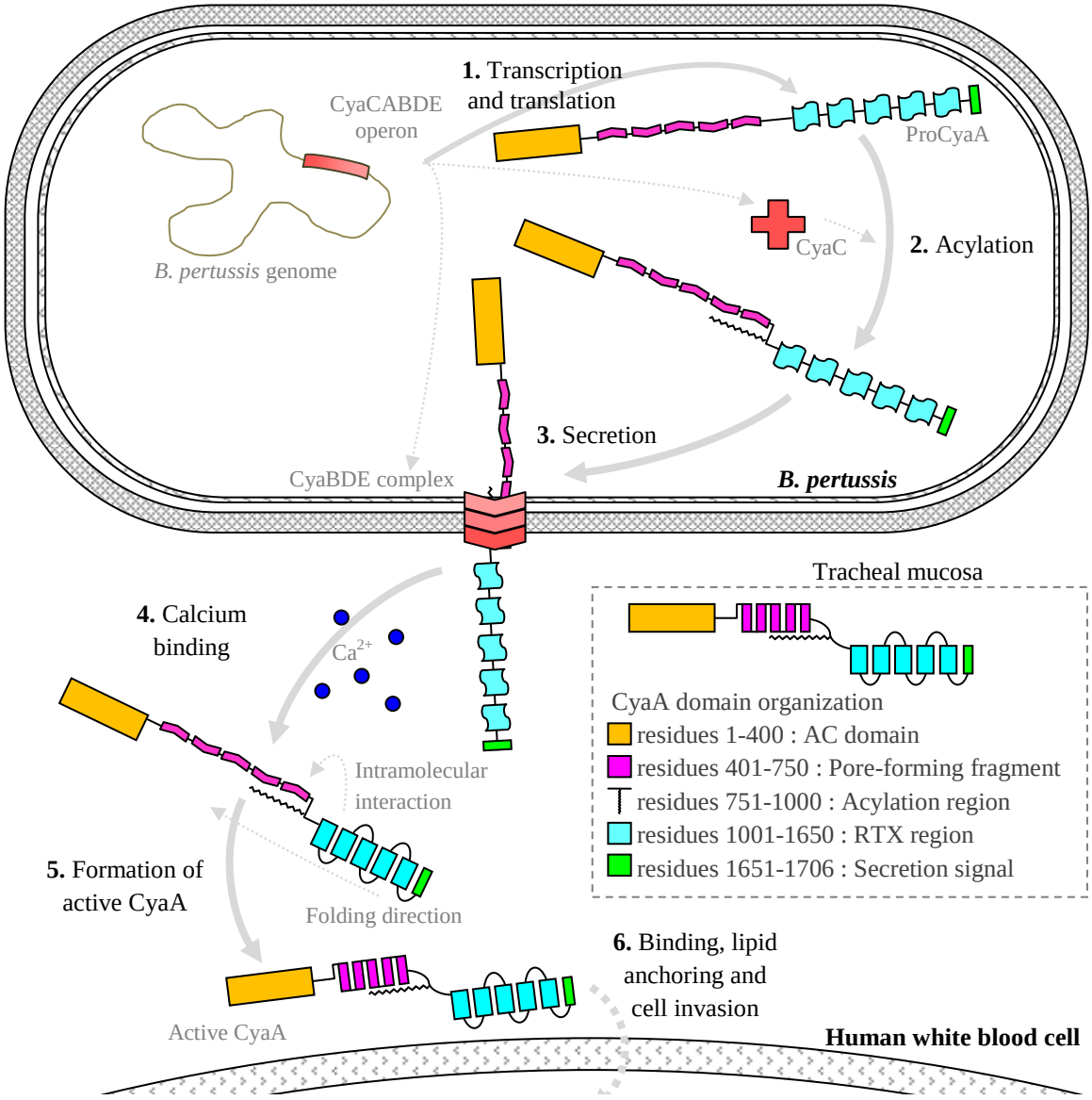


Figure 7.1 Schematic diagram represented mechanism of native CyaA

In nature CyaA was transcribed and translated from CyaCABDE operon (1). Once CyaA was translated as proCyaA, it undergoes post translational modification by being acylated at K_{983} by CyaA specific acyltransferase, CyaC (2). This inactive form has a secretion signal at the last 50 amino acids in the C-terminal. The secretion signal was recognized by type I secretion system CyaBDE complex. The toxin was secreted through 20 angstroms pore of CyaE as a partially unfolded protein (3). The toxin then exposed to the high calcium concentration in the host respiratory tract. Here, the RTX region binds to the calcium ion and promotes formation of β-roll structure (4). The maturation of all five block should be a fast process since inter block interactiob also play crutial role in the structure formation. The RTX region served as folding platform for other region through intramolecular interaction (5). The toxins subsequently bind to specific binding molecule(s), invade the host cells and exert its activities (6).

CHAPTER VIII
CONCLUSION

1. A recombinant plasmid encoding CyaAC-RTX was constructed by site-directed mutagenesis, restriction enzyme assisted-DNA cleavage and ligation. The plasmid was successfully transformed into *E.coli* BL21 (DE3) pLysS. Both CyaA-RTX and CyaC were over-expressed in high amount as soluble proteins.
2. The co-expressed CyaC was active with catalytic activity comparable to that expressed individually.
3. The co-expressed CyaA-RTX was unable to be *in vivo* acyalted by CyaC, and thus, appeared in a soluble unacylated form. This suggested that the pore-forming region (residues 400-750) was required for acylation reaction by CyaC.
4. At twenty times molar ratio, the unacylated CyaA-RTX toxin was unable to inhibit or compete with acylated CyaA-PF in hemolysis assay against sheep red blood cell.
5. Primary structure studies of CyaA reveals the long disorder region at the end of AC domain and RTX region, suggesting that the RTX region adopts intrinsically disorder structure in the absent of calcium ions.
6. Molecular modeling of the RTX region was performed in SWISS-MODEL server using *pseudomonas* sp. MIS38 Lipase (I3 lipase, PDBID: 2z8x) as template. The model had acceptable molecular parameters.
7. The CyaA-RTX model reveals a linear domain arrangement from acylation site through repeat block I-IV with the repeat block V bended toward block III. Linkers II and IV were on the same side as acylation region and might serve as membrane interacting region.
8. The GGXGXDXUX nona-repeat resemble the β-roll motif in which GGXGXD is built up with a calcium binding turn with XUX forming up a β-strand similar

to those of other RTX toxin. The tight and consolidate calcium-binding pockets imply a fast structure forming pathway.

9. The CyaA-RTX model suggests that the aromatic residues at the end of each repeat block play an important role in connecting the repeats and linker hydrophobic cores.
10. Pairwise alignment of CyaE and *E. coli* TolC, was subjected to SWISS-MODEL server. The CyaE trimer model reveals an electropositive tunnel with approximately 20 angstrom cell exterior pore opening. This can be implied that the CyaA could be secreted as a low complex peptide with lack of bound calcium ions.
11. The folding pathway of CyaA was proposed. After CyaA is translated and acylated by CyaC, the partially folded peptide passes through the membrane transporter, CyaBDE complex. Once outside of the cell, the disorder RTX region would bind to calcium ions and fold into a β-roll structure. This structure could then serve as a folding platform for the rest on the toxin. On the other hand, the main purpose of acylation is for target cell adhesion and cell invasion.

REFERENCES

1. Centers for Disease Control and Prevention. Guidelines for the Control of Pertussis Outbreaks. Centers for Disease Control and Prevention: Atlanta, GA, USA, 2012.
2. National Center for Health Statistics. Health, United States, 2011: With Special Feature on Socioeconomic Status and Health. Hyattsville, MD. 2012
3. World health statistics 2012 , Geneva, Switzerland: WHO Press 2012
4. Confer, D. L. and Eaton, J. W. (1982). Phagocyte impotence caused by an invasive bacterial adenylate cyclase. Science, 217(4563), 948-950
5. Hewlett, E. L., Ehrmann, I. E., Maloney, N. J., Fremgen, P. R., Barry, E. M., Weiss, A. A., & Gray, M. C. (1993). Adenylate-Cyclase Toxin from Bordetella-Pertussis - Characterization of Toxin-Catalyzed Intoxication and Hemolysis. Bacterial Protein Toxins, 23, 241-248.
6. Hackett, M., Guo, L., Shabanowitz, J., Hunt, D. F., & Hewlett, E. L. (1994). Internal Lysine Palmitoylation in Adenylate-Cyclase Toxin from Bordetella-Pertussis. Science, 266(5184), 433-435.
7. Barry E.M., Weiss, A. A., Ehrmann, I. E., Gray, M. C., Hewlett, E. L., Goodwin, M. M. (1991) Bordetella pertussis Adenylate-Cyclase Toxin and Hemolytic Activities Require a Second Gene, cyaC, for Activation.
8. Rose, T., Sebo, P., Bellalou, J., & Ladant, D. (1995). Interaction of Calcium with Bordetella-Pertussis Adenylate-Cyclase Toxin. Journal of Biological Chemistry, 270(44), 26370-26376.
9. Menestrina, G., Serra, M. D., Pederzolli, C., Bregante, M., & Gambale, F. (1995). Bacterial hemolysins and leukotoxins affect target cells by forming large exogenous pores into their plasma membrane. Escherichia coli hemolysin A as a case example. Bioscience Reports, 15(6), 543-551
10. Ludwig, A., Tengel, C., Bauer, S., Bubert, A., Benz, R., Mollenkopf, H. J., & Goebel, W. (1995). SlyA, a regulatory protein from Salmonella

typhimurium, induces a haemolytic and pore-forming protein in Escherichia coli. Molecular & General Genetics, 249(5), 474-486

11. Ladant, D., & Ullmann, A. (1999). Bordetella pertussis adenylate cyclase: a toxin with multiple talents. Trends in Microbiology, 7(4), 172-176.
12. Bejerano, M., Nisan, I., Ludwig, A., Goebel, W., & Hanski, E. (1999). Characterization of the C-terminal domain essential for toxic activity of adenylate cyclase toxin. Molecular Microbiology, 31(1), 381-392
13. Lin, W., Fullner, K. J., Clayton, R., Sexton, J. A., Rogers, M. B., Calia, K. E., . . . Mekalanos, J. J. (1999). Identification of a Vibrio cholerae RTX toxin gene cluster that is tightly linked to the cholera toxin prophage. Proceedings of the National Academy of Sciences of the United States of America, 96(3), 1071-1076.
14. Rhodes, C. R., Gray, M. C., Watson, J. M., Muratore, T. L., Kim, S. B., Hewlettt, E. L., & Grisham, C. M. (2001). Structural consequences of divalent metal binding by the adenylyl cyclase toxin of Bordetella pertussis. Archives of Biochemistry and Biophysics, 395(2), 169-176.
15. Locht, C., Antoine, R., & Jacob-Dubuisson, F. (2001). Bordetella pertussis, molecular pathogenesis under multiple aspects. Current Opinion in Microbiology, 4(1), 82-89.
16. Guermonprez, P., Khelef, N., Blouin, E., Rieu, P., Ricciardi-Castagnoli, P., Guiso, N. Leclerc, C. (2001). The adenylate cyclase toxin of Bordetella pertussis binds to target cells via the alpha(M)beta(2) integrin (CD11b/CD18). Journal of Experimental Medicine, 193(9), 1035-1044.
17. El-Azami-El-Idrissi, M., Bauche, C., Loucka, J., Osicka, R., Sebo, P., Ladant, D., & Leclerc, C. (2003). Interaction of Bordetella pertussis adenylate cyclase with CD11b/CD18 - Role of toxin acylation and identification of the main integrin interaction domain. Journal of Biological Chemistry, 278(40), 38514-38521.
18. Masin, J., Konopasek, I., Svobodova, J. & Sebo, P. (2004) Different structureal requiremetns for adenylate cyclase toxin interations with erythrocyte and liposome membranes. Biochim. Biophys Acta 1660, 144-154.

19. Knapp, O., Maier, E., Polleichtner, G., Masin, J., Sebo, P., & Benz, R. (2003). Channel formation in model membranes by the adenylate cyclase toxin of Bordetella pertussis: Effect of calcium. Biochemistry, 42(26), 8077-8084.

20. Benabdelhak, H., Kiontke, S., Horn, C., Ernst, R., Blight, M. A., Holland, I. B., & Schmitt, L. (2003). A specific interaction between the NBD of the ABC-transporter HlyB and a C-terminal fragment of its transport substrate haemolysin A. Journal of Molecular Biology, 327(5), 1169-1179.

21. Gray, M. C., Donato, G. M., Jones, F. R., Kim, T., & Hewlett, E. L. (2004). Newly secreted adenylate cyclase toxin is responsible for intoxication of target cells by Bordetella pertussis. Molecular Microbiology, 53(6), 1709-1719.

22. Donato, G. M., Hsia, H. L. J., Green, C. S., & Hewlett, E. L. (2005). Adenylate cyclase toxin (ACT) from Bordetella hinzii: characterization and differences from ACT of Bordetella pertussis. Journal of Bacteriology, 187(22), 7579-7588.

23. Basler, M., Knapp, O., Masin, J., Fiser, R., Maier, E., Benz, R. Osicka, R. (2007). Segments crucial for membrane translocation and pore-forming activity of Bordetella adenylate cyclase toxin. Journal of Biological Chemistry, 282(17), 12419-12429.

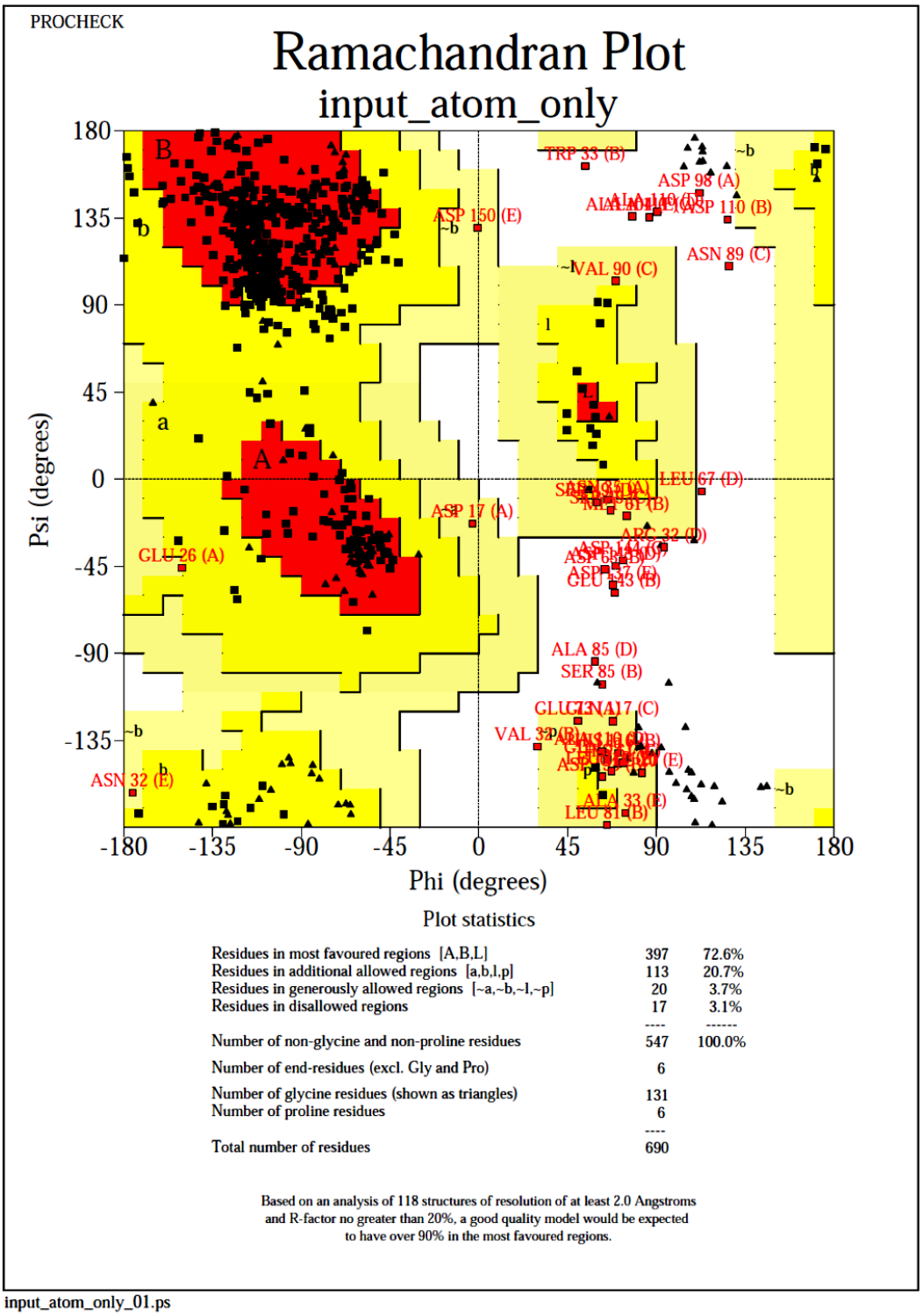
24. Powtongchin, B., Angsuthanasombat, C., (2008). High level of soluble expression in Escherichia coli and characterization of the CyaA pore forming fragment from bordetella pertussis Thai clinical isolate. Arch microbio, 189, 169-174.

25. Powtongchin, B., Angsuthanasombat, C., (2009). Effect on heamolytic activity of single proline substitutions in the bordetella pertussis CyaA pore-forming fragment. Arch microbio, 191, 1-9.

APPENDICES

APPENDIX A
CyaA-RTX MOLECULAR MODEL

The CyaA-RTX molecular model was summited to the web base molecular parameter verifying software. The complete results were showed here, including, Ramachandran plot, Detail ramachandran plot for each residues type, Chi1/Chi plot, Main chain parameter, Side-chain parameter, complete residue properties, Main chain bond length/angle, Planarity of certain residues and distorted geometry.

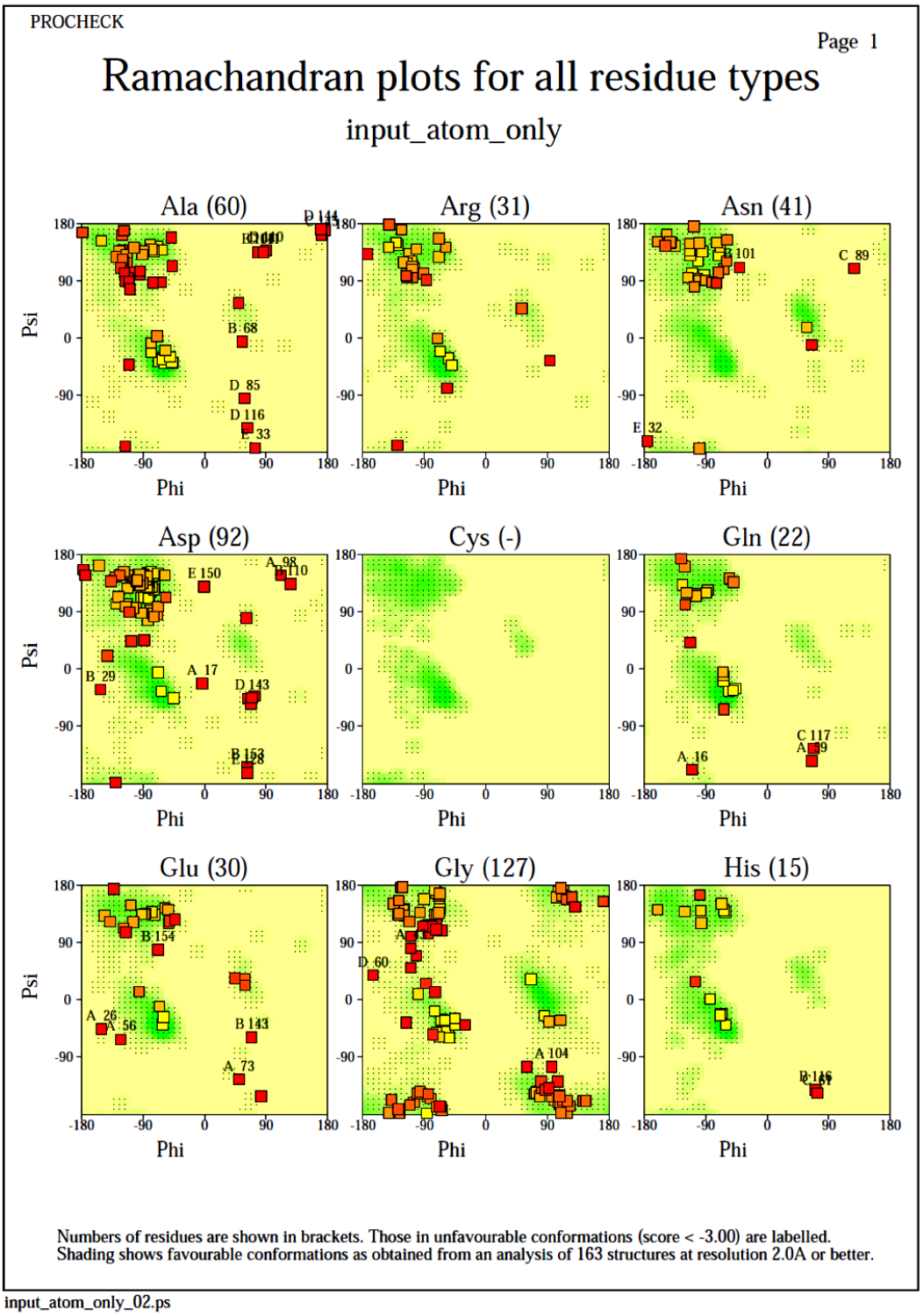


Plot statistics

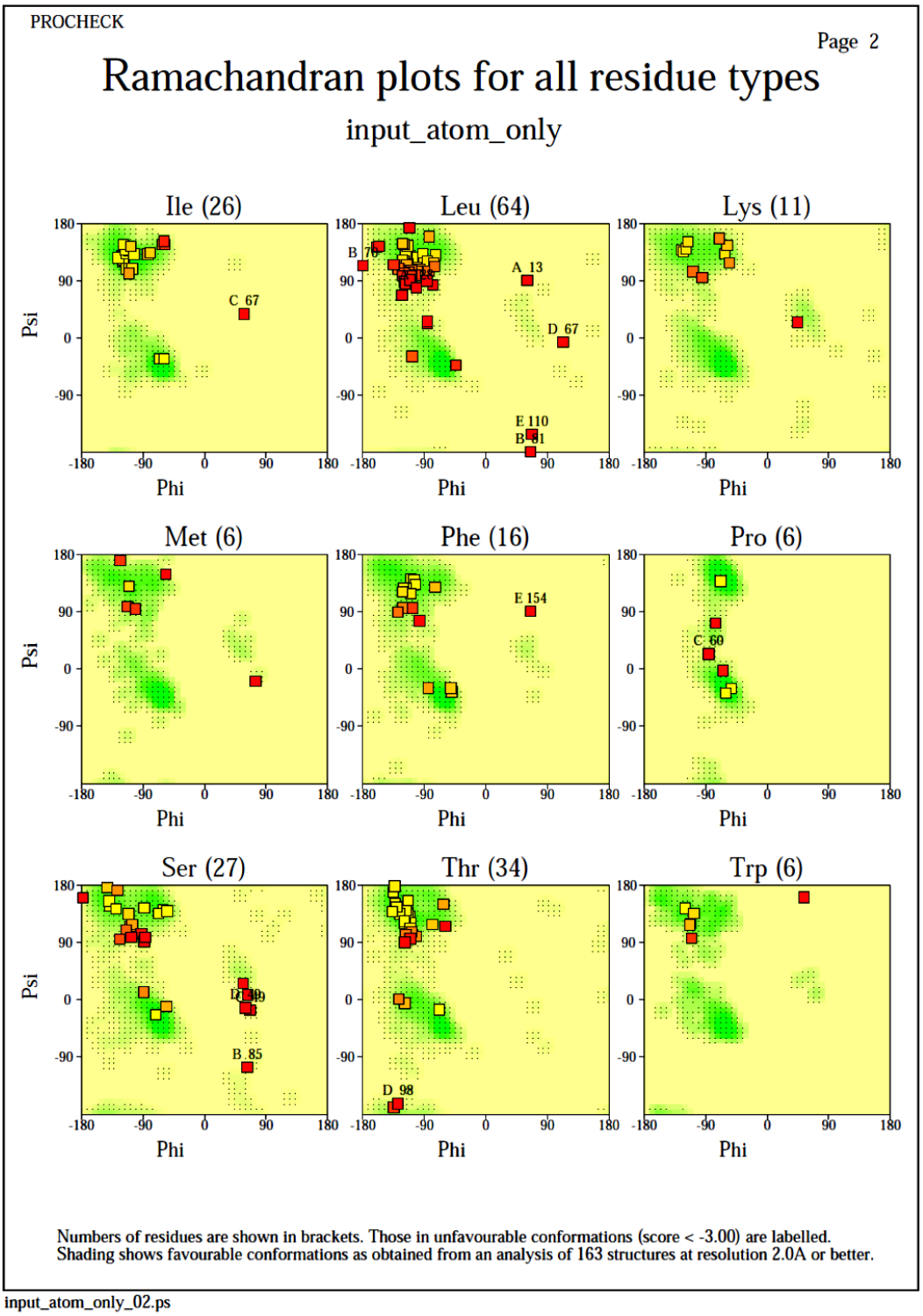
Residues in most favoured regions [A,B,L]	397	72.6%
Residues in additional allowed regions [a,b,l,p]	113	20.7%
Residues in generously allowed regions [~a,~b,~l,~p]	20	3.7%
Residues in disallowed regions	17	3.1%
Number of non-glycine and non-proline residues	547	100.0%
Number of end-residues (excl. Gly and Pro)	6	
Number of glycine residues (shown as triangles)	131	
Number of proline residues	6	
Total number of residues	690	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

input_atom_only_01.ps



input_atom_only_02.ps

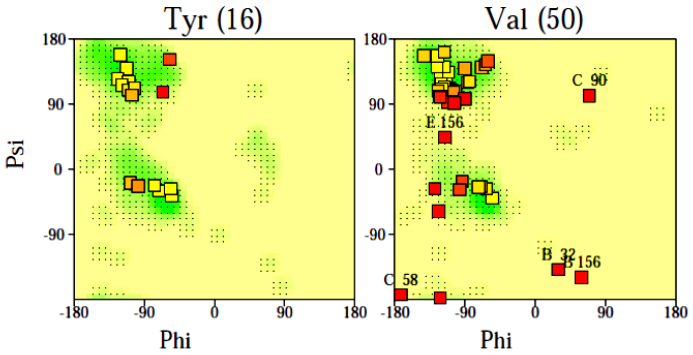


input_atom_only_02.ps

PROCHECK

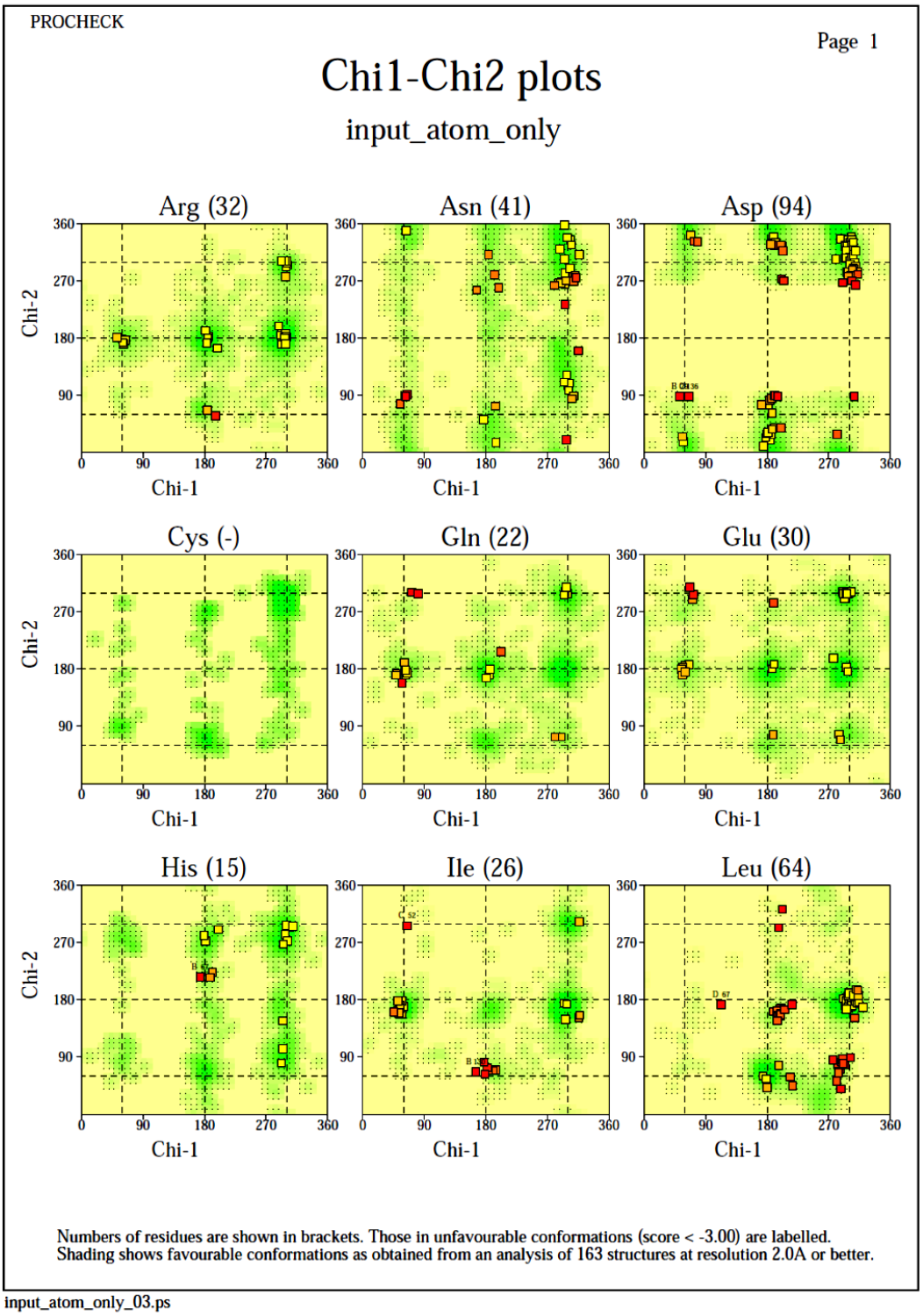
Ramachandran plots for all residue types

input_atom_only



Numbers of residues are shown in brackets. Those in unfavourable conformations (score < -3.00) are labelled.
Shading shows favourable conformations as obtained from an analysis of 163 structures at resolution 2.0A or better.

input_atom_only_02.ps



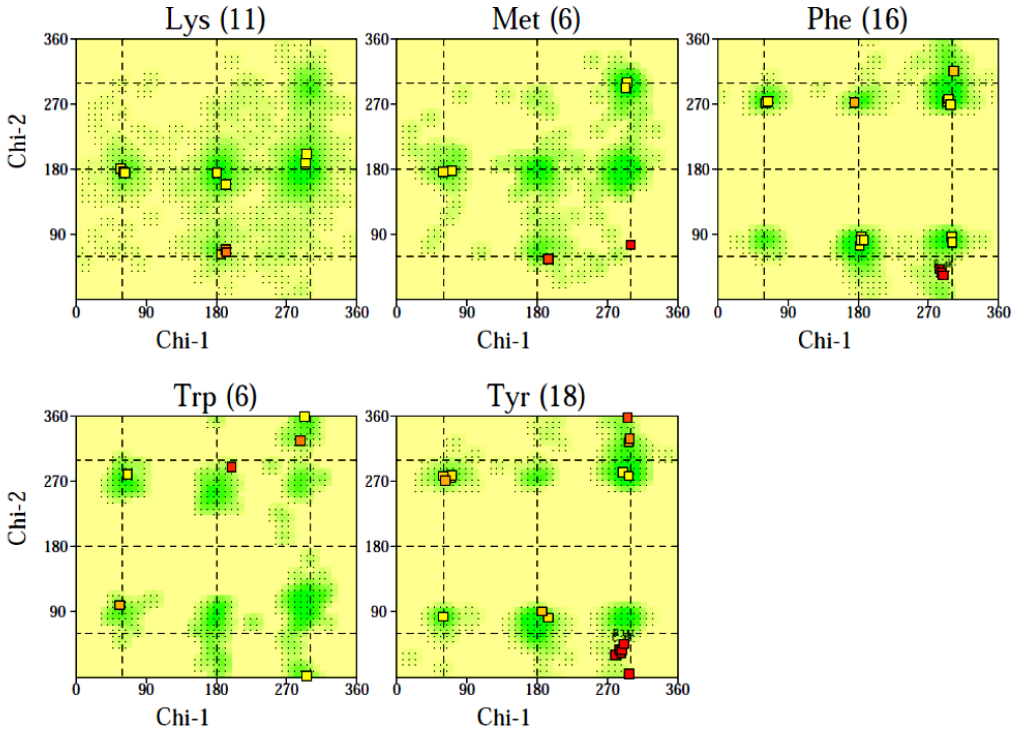
input_atom_only_03.ps

PROCHECK

Page 2

Chi1-Chi2 plots

input_atom_only



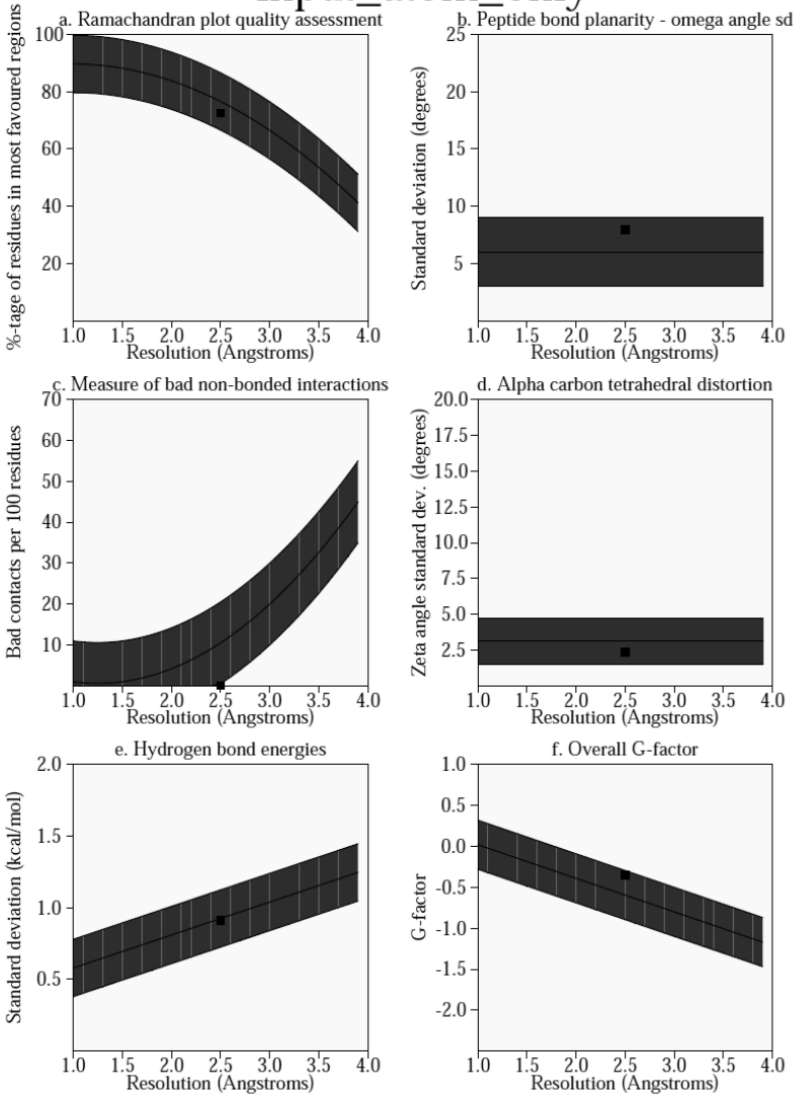
Numbers of residues are shown in brackets. Those in unfavourable conformations (score < -3.00) are labelled.
Shading shows favourable conformations as obtained from an analysis of 163 structures at resolution 2.0A or better.

input_atom_only_03.ps

PROCHECK

Main-chain parameters

input_atom_only



Plot statistics

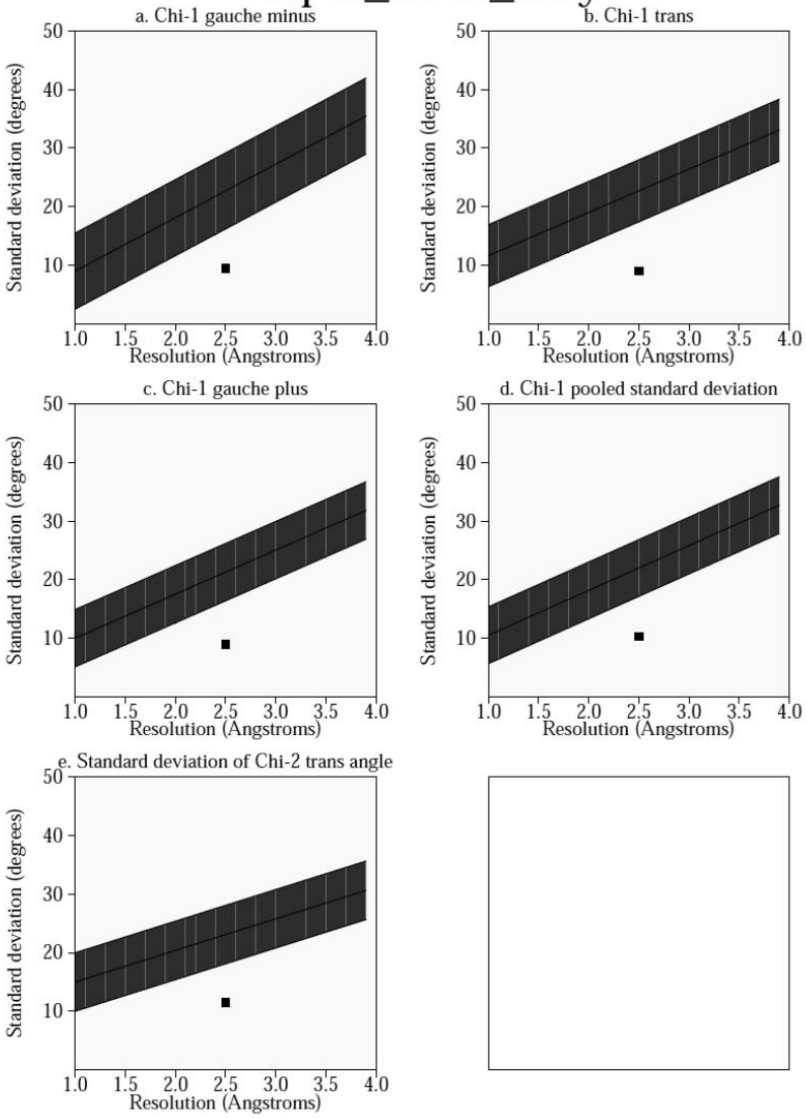
Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. %-tage residues in A, B, L	547	72.6	76.6	10.0	-0.4	Inside
b. Omega angle st dev	679	8.0	6.0	3.0	0.7	Inside
c. Bad contacts / 100 residues	1	0.1	10.5	10.0	-1.0	BETTER
d. Zeta angle st dev	559	2.4	3.1	1.6	-0.5	Inside
e. H-bond energy st dev	388	0.9	0.9	0.2	-0.1	Inside
f. Overall G-factor	690	-0.4	-0.6	0.3	0.8	Inside

input_atom_only_04.ps

PROCHECK

Side-chain parameters

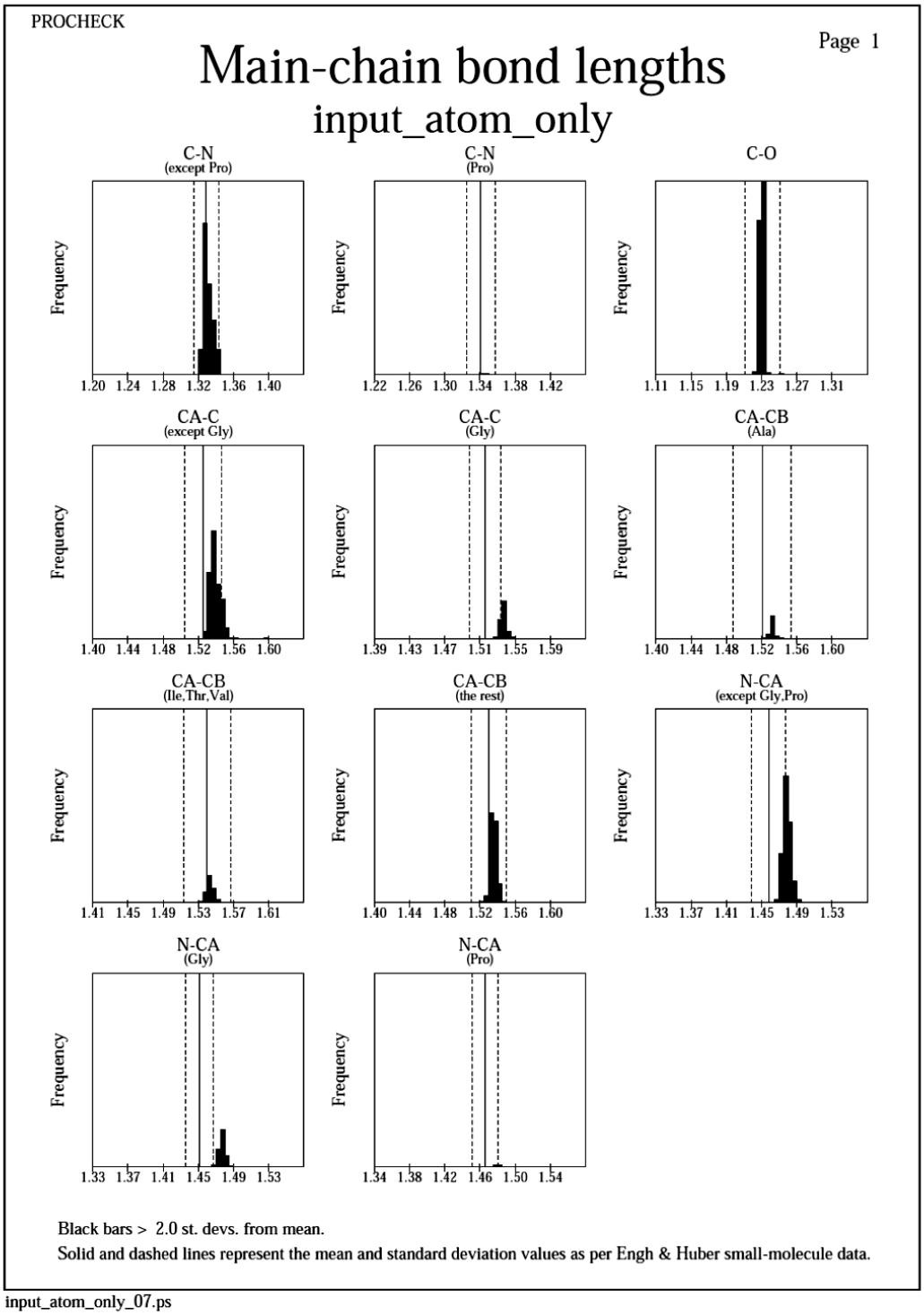
input_atom_only



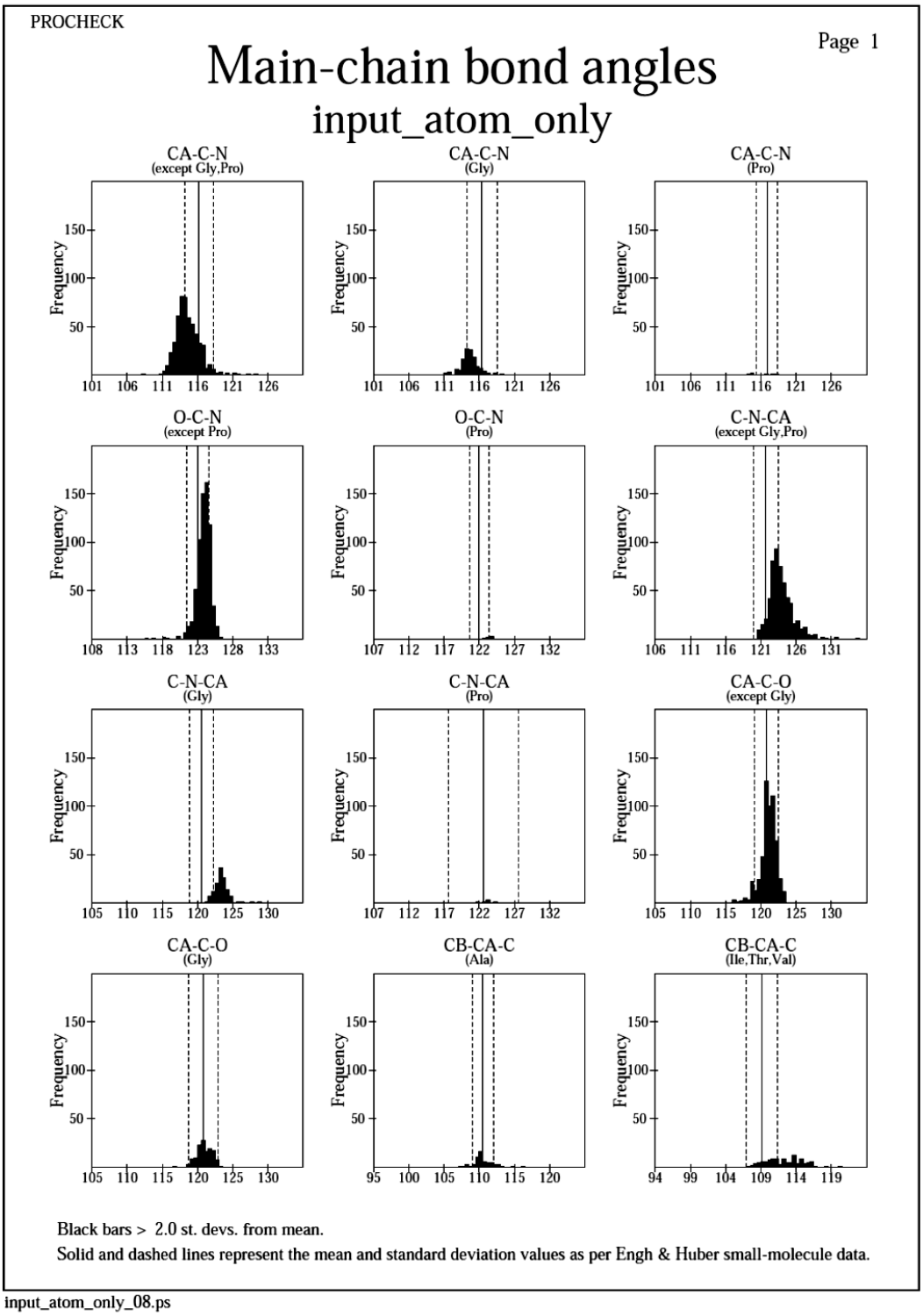
Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. Chi-1 gauche minus st dev	121	9.4	22.7	6.5	-2.0	BETTER
b. Chi-1 trans st dev	130	9.0	22.7	5.3	-2.6	BETTER
c. Chi-1 gauche plus st dev	242	9.0	21.3	4.9	-2.5	BETTER
d. Chi-1 pooled st dev	493	10.3	22.0	4.8	-2.4	BETTER
e. Chi-2 trans st dev	121	11.6	23.1	5.0	-2.3	BETTER

input_atom_only_05.ps



input_atom_only_07.ps

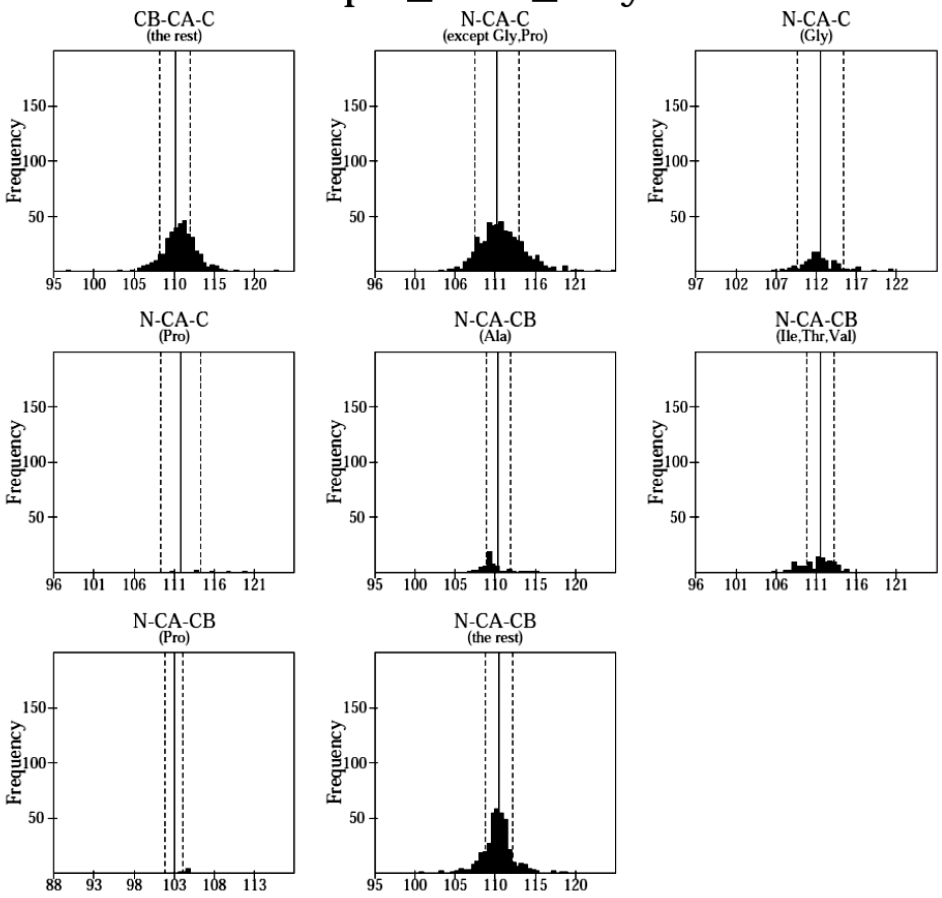

PROCHECK
Page 1
Main-chain bond angles
input_atom_only
CA-C-N
(except Gly,Pro)
CA-C-N
(Gly)
CA-C-N
(Pro)
O-C-N
(except Pro)
O-C-N
(Pro)
C-N-CA
(except Gly,Pro)
C-N-CA
(Gly)
C-N-CA
(Pro)
CA-C-O
(except Gly)
CA-C-O
(Gly)
CB-CA-C
(Ala)
CB-CA-C
(Ile,Thr,Val)
Frequency
Black bars > 2.0 st. devs. from mean.
Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.
input_atom_only_08.ps

PROCHECK

Page 2

Main-chain bond angles

input_atom_only



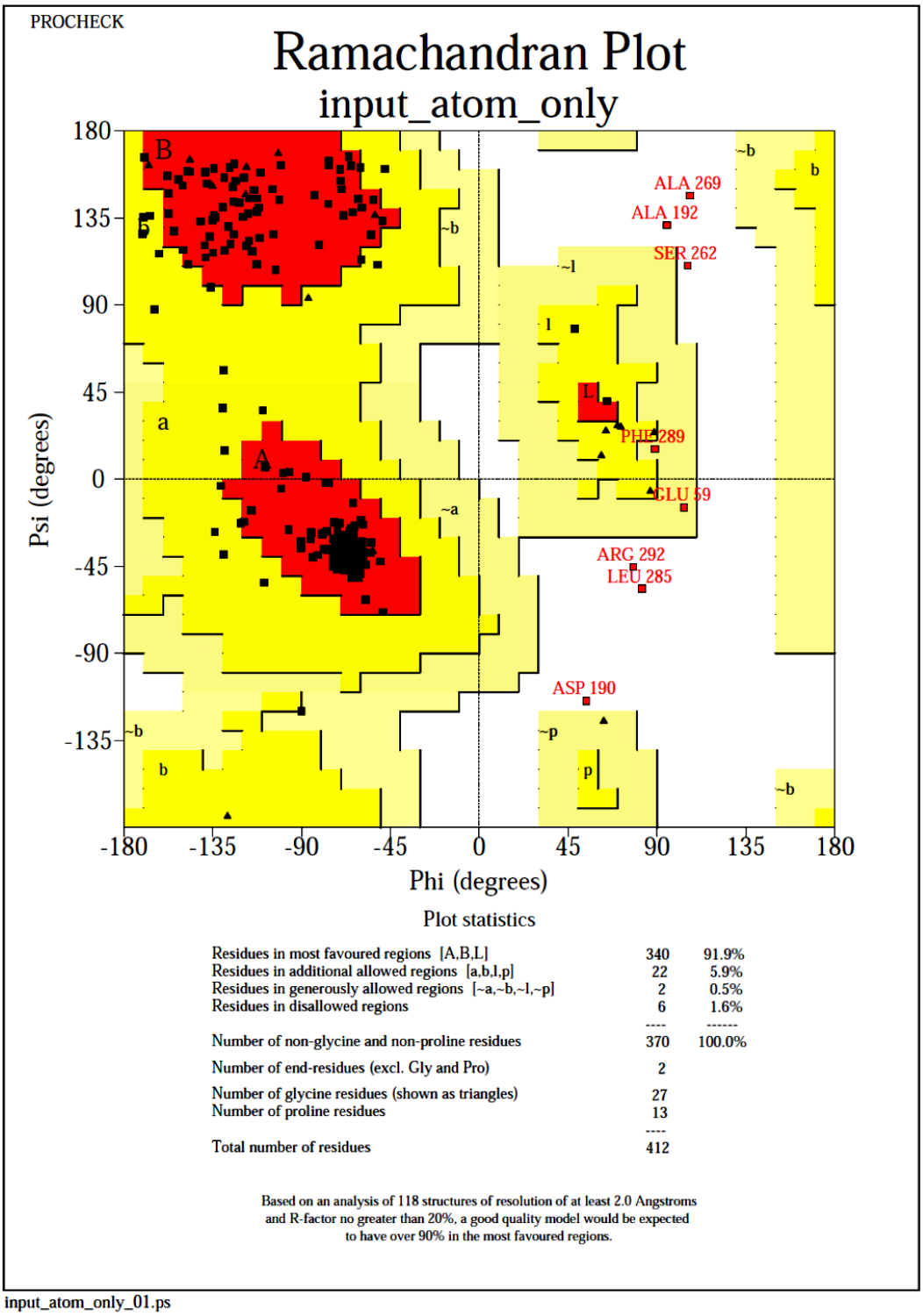
Black bars > 2.0 st. devs. from mean.

Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

input_atom_only_08.ps

APPENDIX B
CyaE MOLECULAR MODEL

The CyaE molecular model was summited to the web base molecular parameter verifying software. The complete results were showed here, including, Ramachandran plot, Detail ramachandran plot for each residues type, Chi1/Chi plot, Main chain parameter, Side-chain parameter, complete residue properties, Main chain bond length/angle, Planarity of certain residues and distorted geometry.

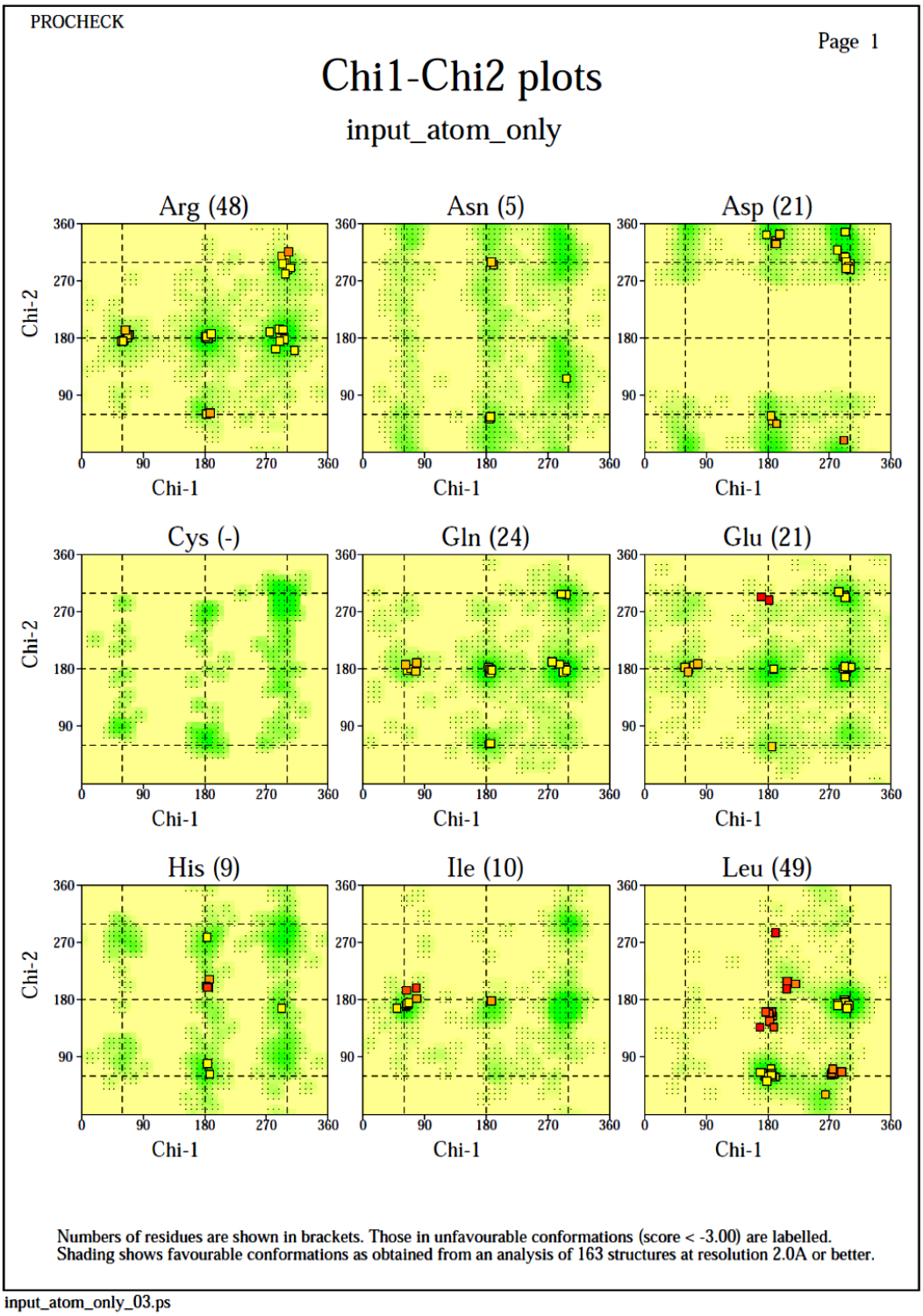


Plot statistics

Residues in most favoured regions [A,B,L]	340	91.9%
Residues in additional allowed regions [a,b,l,p]	22	5.9%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.5%
Residues in disallowed regions	6	1.6%
Number of non-glycine and non-proline residues	370	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	27	
Number of proline residues	13	
Total number of residues	412	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

input_atom_only_01.ps



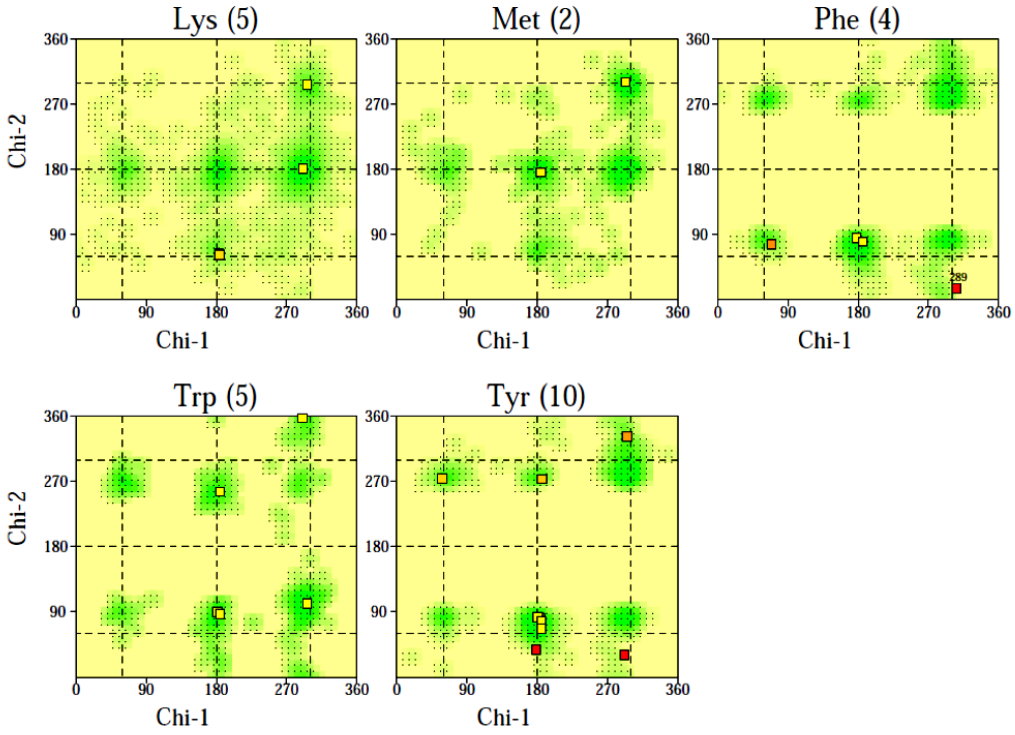
input_atom_only_03.ps

PROCHECK

Page 2

Chi1-Chi2 plots

input_atom_only



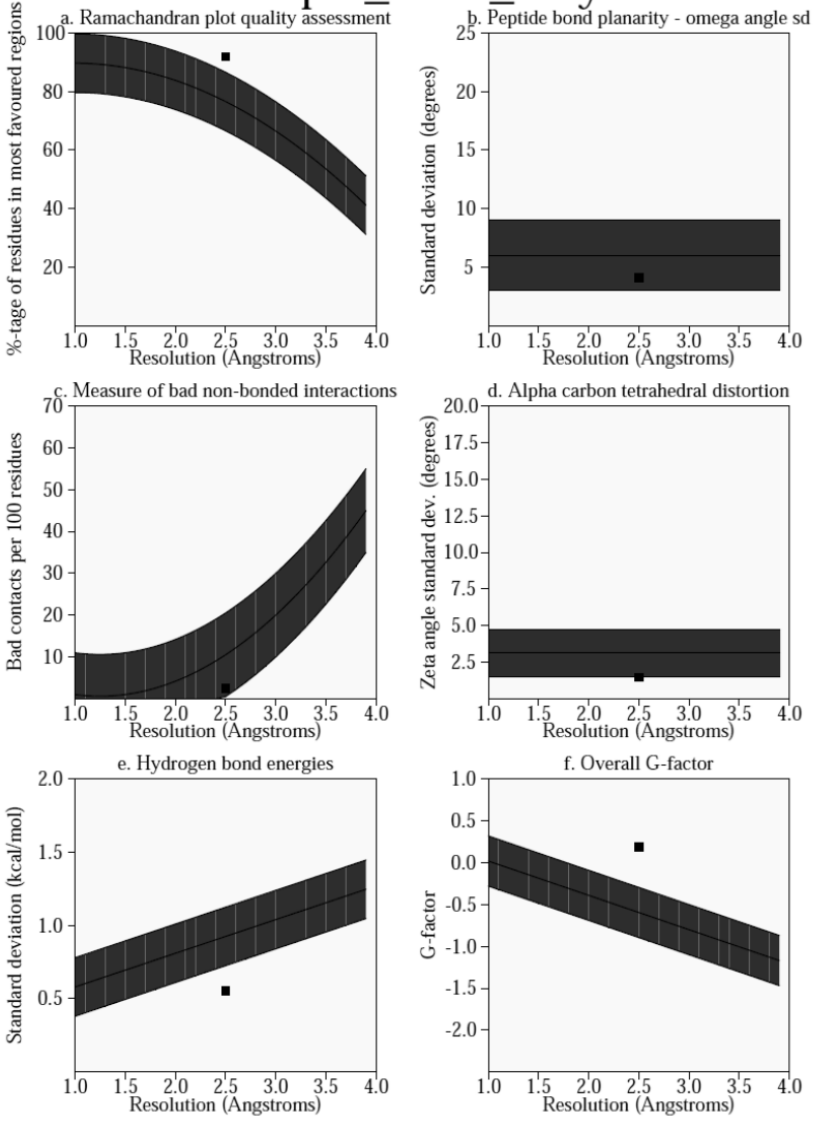
Numbers of residues are shown in brackets. Those in unfavourable conformations (score < -3.00) are labelled.
Shading shows favourable conformations as obtained from an analysis of 163 structures at resolution 2.0A or better.

input_atom_only_03.ps

PROCHECK

Main-chain parameters

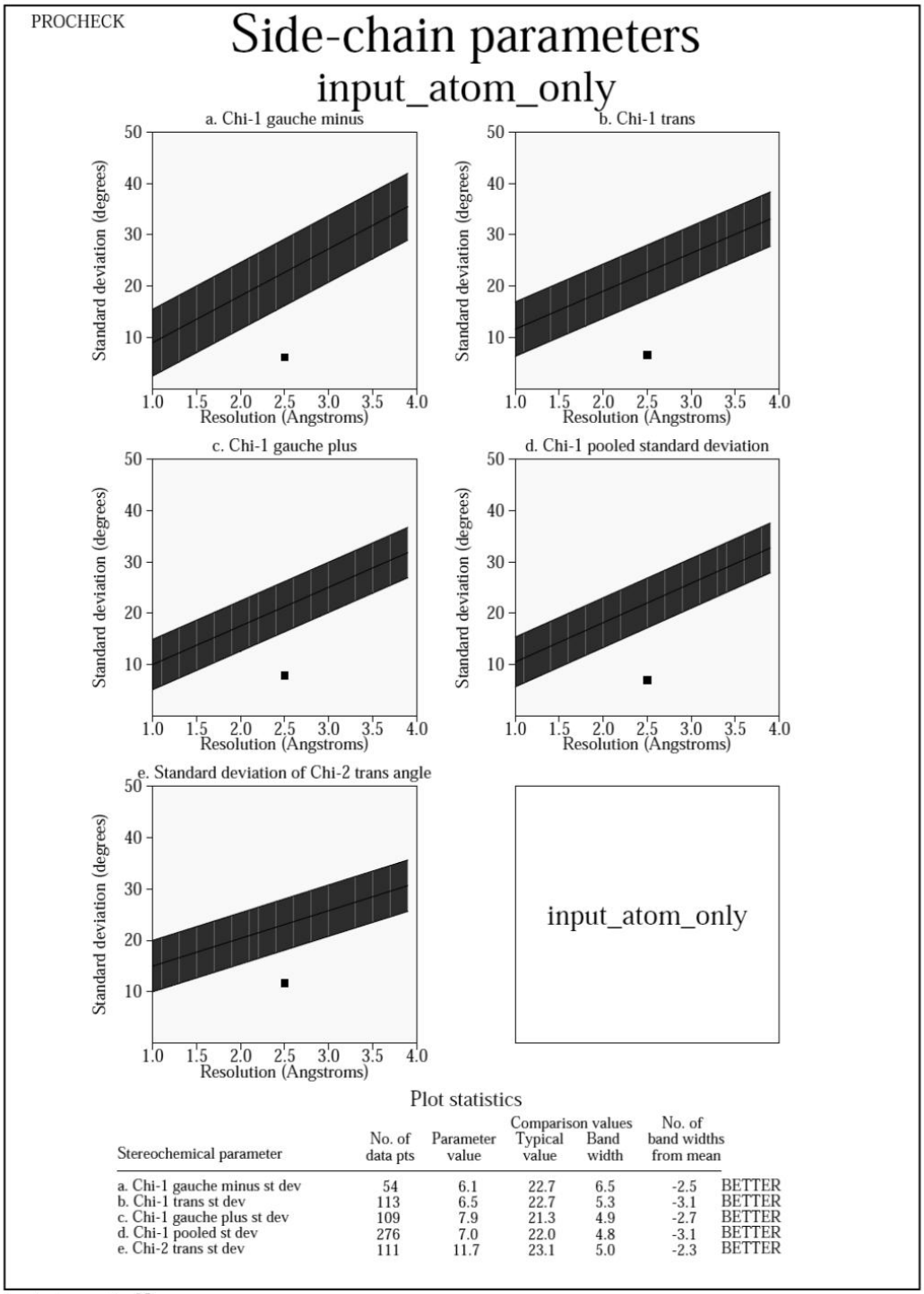
input_atom_only



Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. %-tage residues in A, B, L	370	91.9	76.6	10.0	1.5	BETTER
b. Omega angle st dev	410	4.1	6.0	3.0	-0.6	Inside
c. Bad contacts / 100 residues	10	2.4	10.5	10.0	-0.8	Inside
d. Zeta angle st dev	385	1.5	3.1	1.6	-1.0	BETTER
e. H-bond energy st dev	324	0.6	0.9	0.2	-1.8	BETTER
f. Overall G-factor	412	0.2	-0.6	0.3	2.6	BETTER

input_atom_only_04.ps



Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. Chi-1 gauche minus st dev	54	6.1	22.7	6.5	-2.5	BETTER
b. Chi-1 trans st dev	113	6.5	22.7	5.3	-3.1	BETTER
c. Chi-1 gauche plus st dev	109	7.9	21.3	4.9	-2.7	BETTER
d. Chi-1 pooled st dev	276	7.0	22.0	4.8	-3.1	BETTER
e. Chi-2 trans st dev	111	11.7	23.1	5.0	-2.3	BETTER

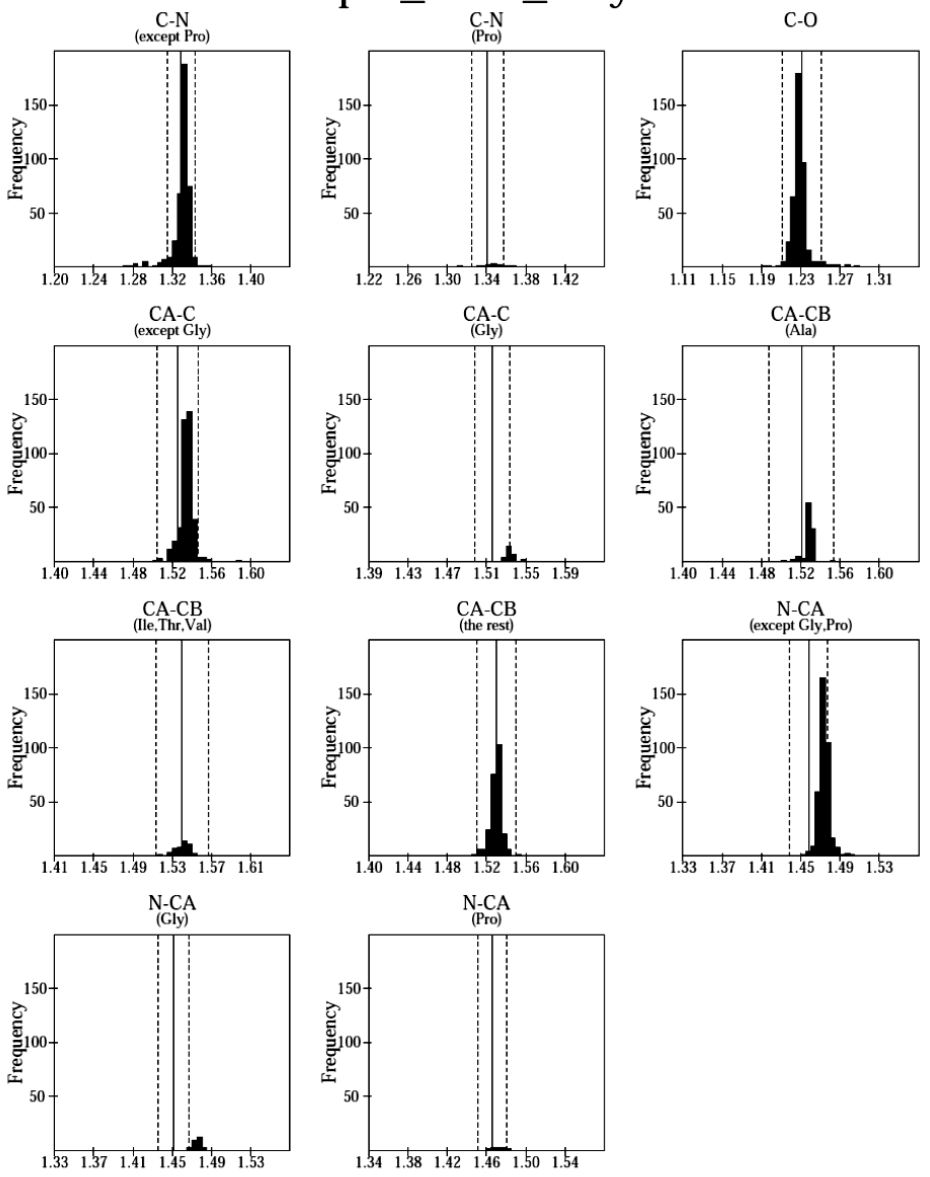
input_atom_only_05.ps

PROCHECK

Page 1

Main-chain bond lengths

input_atom_only



Black bars > 2.0 st. devs. from mean.

Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

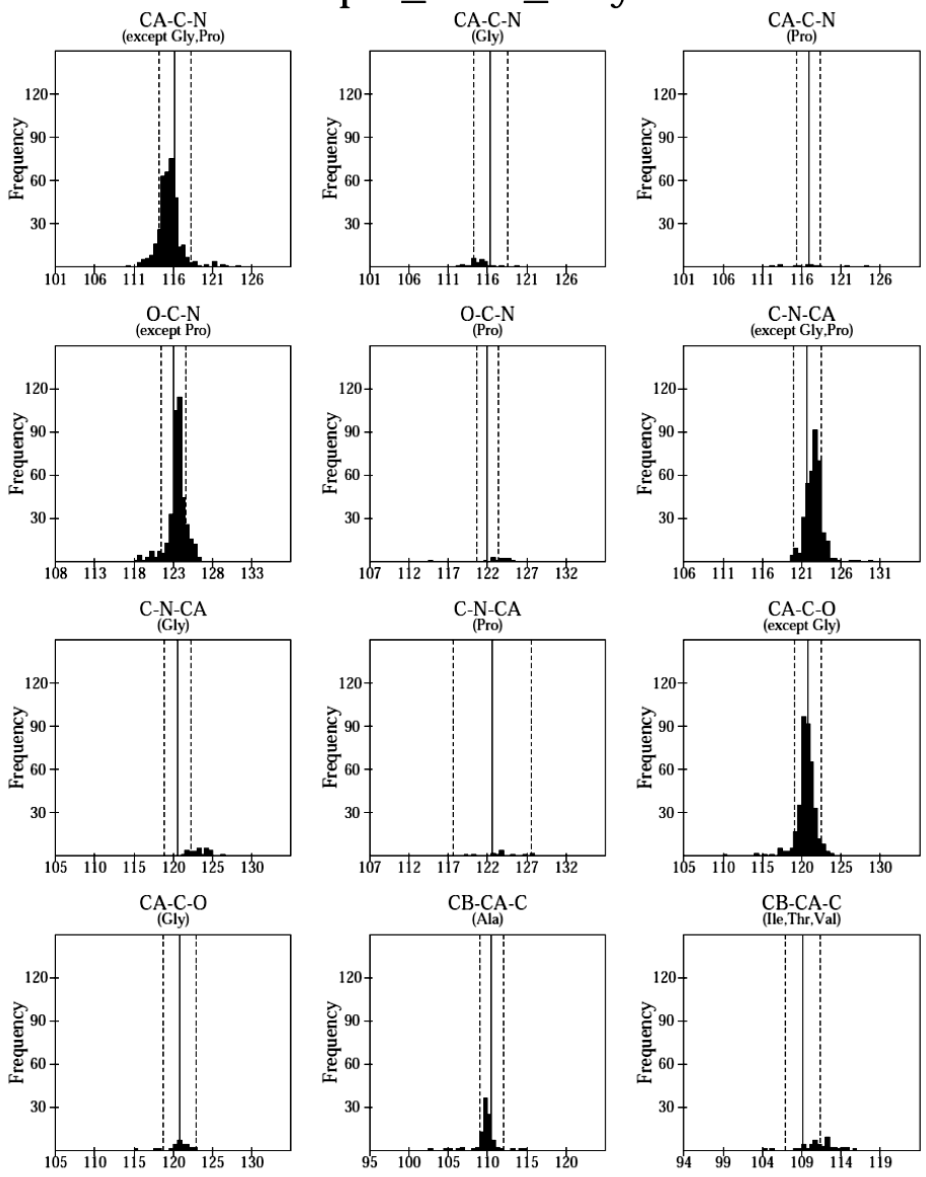
input_atom_only_07.ps

PROCHECK

Page 1

Main-chain bond angles

input_atom_only



Black bars > 2.0 st. devs. from mean.

Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

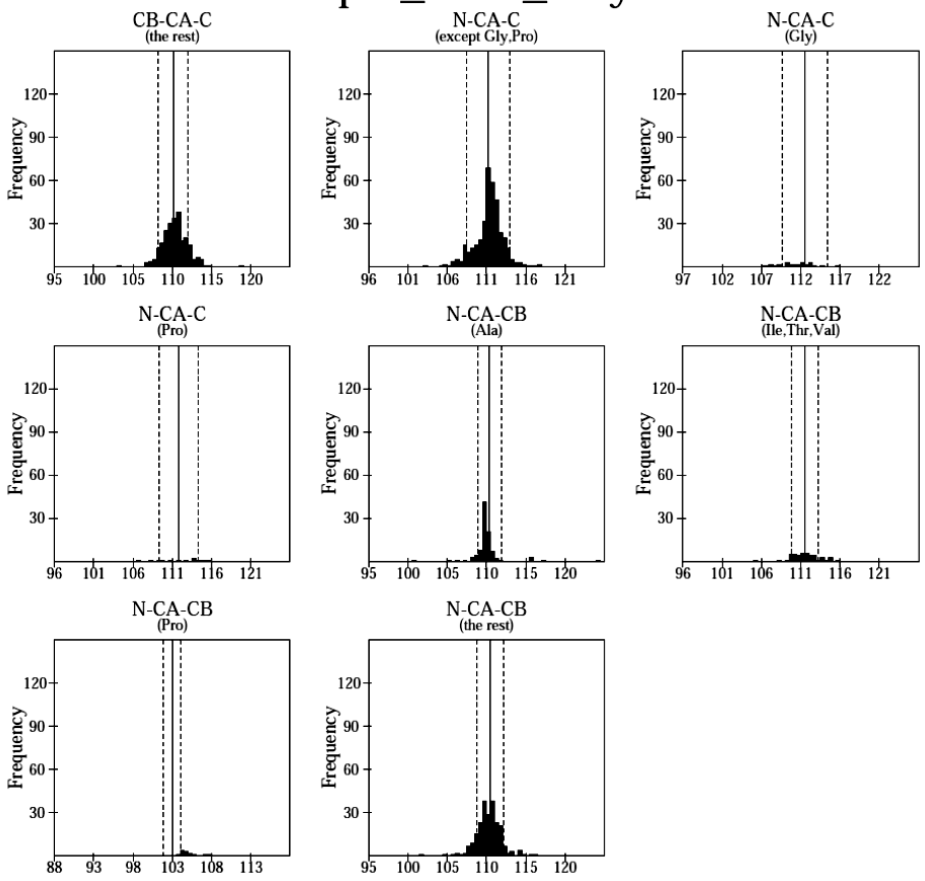
input_atom_only_08.ps

PROCHECK

Page 2

Main-chain bond angles

input_atom_only



Black bars > 2.0 st. devs. from mean.

Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

input_atom_only_08.ps

BIOGRAPHY

NAME	Mr. Kittipong Suvarnapunya
DATE OF BIRTH	May 10, 1985
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	B.Sc. (Biochemistry), Chulalongkorn University M.Sc. (Molecular Genetics and Genetic Engineering), Mahidol University
HOME ADDRESS	364 Phaholyothin 30, Jankasem, Jatujak, Bangkok, 10900, Thailand Email: tonmai_kub@hotmail.com