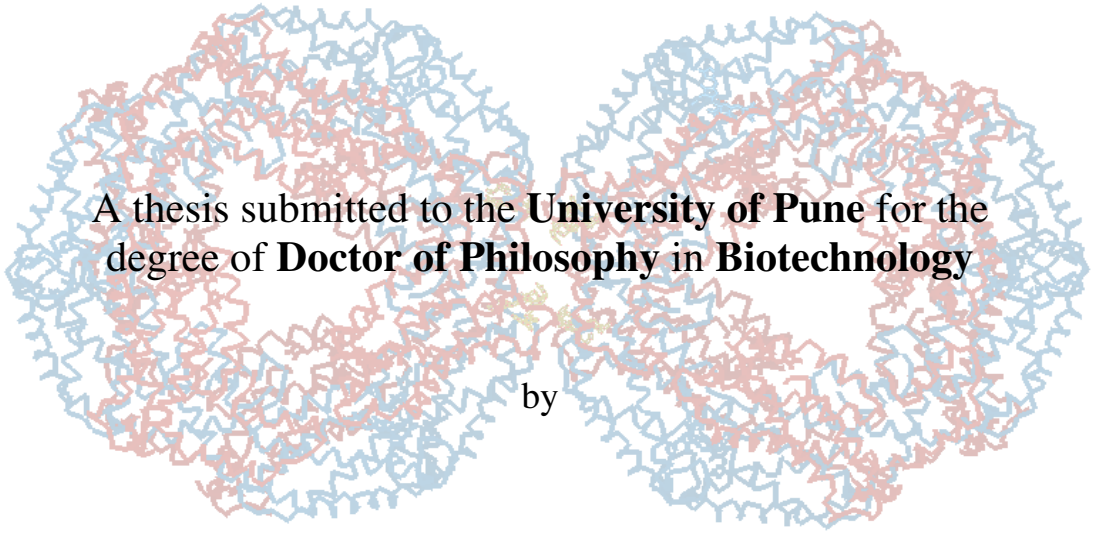
STRUCTURAL STUDIES ON

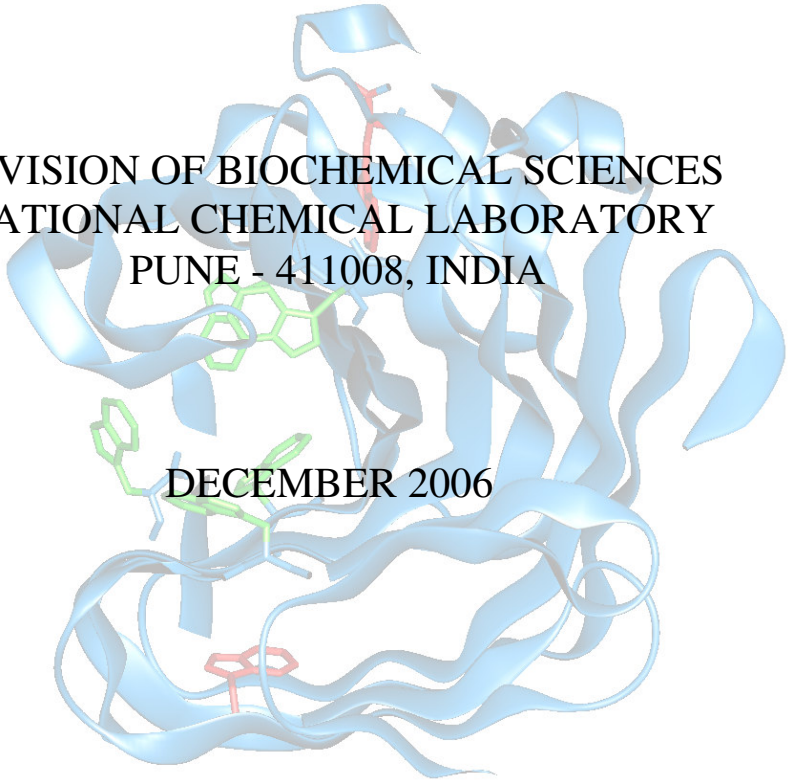
PHYCOCYANINS FROM THREE CYANOBACTERIAL SPP. and A XYLANASE FROM AN ALKALOPHILIC *BACILLUS SP.*



A thesis submitted to the **University of Pune** for the degree of **Doctor of Philosophy** in **Biotechnology**

by

L.SATYANARAYANA



DIVISION OF BIOCHEMICAL SCIENCES
NATIONAL CHEMICAL LABORATORY
PUNE - 411008, INDIA

DECEMBER 2006

DECLARATION

I hereby declare that the thesis entitled **'Structural studies on phycocyanins from three cyanobacterial spp. and a xylanase from an alkalophilic *Bacillus sp*'**, submitted for Ph.D degree to the University of Pune has not been submitted by me to any other university for a degree or diploma.

L. Satyanarayana
Senior Research Fellow
Div. of Biochemical Sciences
National Chemical Laboratory
Pune – 411 008, India

Date: 6.12.2006

Dedicated
… My Parents

AKNOWELEDGEMENTS

I express my reverence and gratitude for my guide Dr. C.G. Suresh, who was continuous source of inspiration throughout the course of my work. I do not have words to express my feelings about his dedication, hard work and quest for the research. His deep involvement, logical and analytical thinking has made a big difference towards the research. I consider myself to be fortunate to work under his guidance and I am indebted for shaping my research life. I am grateful to him for being an extraordinary teacher and seeking support to improve by all means.

It is my honor to thank Dr. Sushma Gaikwad, for exploring ideas in research and being with me. I am grateful to Dr. Balakrishnan for his help.

I am grateful to Dr. M.I Khan and Dr. Asmitha Prabunea for their continuous encouragement, motivation and support to seek my capabilities. My sincere thanks to Dr. Sandhya Misra and Dr. P.K.Gosh, (Director) for allowing me to at CSMCRI, Bhavnagar.

Thanks are also due to my past and present labmates K.N.Roa, Anu, Manish Chandra, Suresh Kumar, Uma, Poorva, Raamesh, Anitha, Sindu, Nitin, Nishant and Urvashi for the cordial and friendly atmosphere in the lab.

My gratitude to Dr. Mahan badbade, Dr. Puranik, Dr. Smitha, Dr. Rajesh and Manoj for there suggestions and encouragement.

I fall short of words to express my feelings towards my all time friends, Sharath, Srikanth, and Suchitra. They have made an outstanding impact on my outlook towards life and science. They have been there all the time with me in my ups and downs.

I would like to express my regards to all my friends and colleagues as their contribution has been stimulating over a period especially, Sarvanan, Anish. Nitin, Rohini, Anamika, Richa, Ravathi, Manoj, Bushan, Ramchander, Elan, Feroz, Anil, Atul Thakur, Raju, Sathish, Nagaraj, Sachin, Salva, Harish, Mukund, Mahesh Takker, Sambaji, Sathish,

Suhas, Ajay, Sagar, Ajit, Ambirsh, Rothas, Shabab, Atul, J.P, Sarvash, Santhose, Chatan, Afzal, Srinivas, Sushim, Noor, Arun. Sameer, Manish, Shashidher, Atul, Raman, Prakesh, Srinu, Shiva, Raghu.

I am indeed thankful to my Aunt 'Nanda', for providing me the nice lunch and encouragement and help during the last part of the work.

I owe my deepest gratitude to my parents, brother, sister, niece, nephew, and sister-in-law without whose enduring support and encouragement, it would not have been possible to embark upon this journey in life.

This acknowledgement would not be complete without thanking all my cousins and college friends; Sridher akkula, Aravinds, Rajinikanth, Ramu M.L, Santhose, Sridher, Ramakrishna, Rajesh, Uday for their moral support.

I am thankful to our office staffs Indira, Satyali, and Mari for their timely help rendered during the course of this work.

I also express my thanks to our efficient instrumentation staffs, Mr. Kamthe, Mr. Karanjkar, Mr. Trehan, Mr. Dev, and Mr. Prasad for the help and maintenance of the instruments.

I acknowledge Dr. Paul Ratnaswamy, ex, Director and Dr. S. Shivram, Director NCL for permitting to present this work in the form of thesis. I appreciate financial support from CSIR, in the form of Junior and Senior Research fellowship.

In the end I am glad to thank all my school, college teachers and friends.

L.Satyanarayana

December 2006.

Contents

ABBREVIATIONS

(List of abbreviations and standard notations used in the text)

A	: Absorbance
AMoRe	: Automated Molecular Replacement
ANS	: 1-Anilino 8-Sulphonic Acid
APC	: Allophycocyanin
AQE	: 6-aminoquinolyl-N-hydroxysuccinimiyl carbamate
AS	: Ammonium Sulphate
AT*B*XYL-C	: Alkalophilic Thermophilic Bacillus Xylanase-type C
AU	: Absorption Unit
BPB	: Bromo-Phenol Blue
Cc	: correlation coefficients
CCD	: Charge Coupled Device
CCP4	: Collaborative Computational Project No.4
CD	: Circular Dichroism
CM	: Carboxy Methyl
C-PC	: Cyanobacterial Phycocyanin
CsCl	: Cesium Chloride
DLS	: Dynamic Light Scattering
DMF	: Dimethyl formamide
DMSO	: Dimethyl Sulfoxide
DNSA	: Dinitrosalicylicacid
EDTA	: Ethylenediamietetraaceticacid
Gdn-HCl	:Guanidium Hydrochloride
His	: Histidine
HNBB	: 2-hydoxy-5-nitrobenzylbromide
HXT	: 1,2,6- Hexanetriol
IU	: International Unit
kcal/mol	: kilo Calories per mole
KI	: Potassium Iodide
L	: Liter
M.W/Mr.	: Molecular weight
MAD	: Multiwavelenght Anamolous Dispoersion
Min	: Minute
MIRAS	: Multiple isomorphous replacement with anomalous scattering
MPD	: 2-Methyl-2,4-Pentanediol
MR	: Molecular replacement
NaCl	: Sodium chloride
NBS	: N-bromo succinimide
NCS	: Non-crystallographic symmetry
NMR	: Nuclear Magnetic Resonance
OD	: Optical Density

PAGE	: Poly Acryamide Gel eletrophoresis
PDB	: Protein Data bank
PE	: Phycoerythrin
PEC	: Phycoerythrocyanin
PEG	: Poly Ethylene Glycol
PG	: Phenyl Glyoxal
PS II	: Photosystem II
RF	:Rotation Function
r.m.s.d	: Root Mean Square Deviarion
rpm	: Revolutions per Minute
SAD	: Single Wavelength Anomalous Diffraction
SCB	: Sodium Cacodylate buffer
SDS	: Sodium Dodecyl Sulphate
Sec or s	: Second
SIRAS	: Single Isomorphous Replacement with Anomalous Scattering
Sol	: Solution
TEA	:Terminal Electron Acceptor
TF	:Translation Function
Temp	: Temperature
Tris	: tris-hydroxylmethyl amino methane
Tyr	: Tyrosine
UV	: Ultra Violet
v	: volume
w	: weight
WRK	: Woodward's reagent K
X1	: Xylose
X2	: Xylobiose
X3	: Xylotriose
X4	: Xylotetrose
X5	: Xylopentose

α	: alpha
β	: beta
γ	: gamma
µmol	: Micro mole
∑	: summation
µg	: microgram
µl	: microliter
µm	: micrometer
µM	: micromolar
Å	: Angstrom
C	: Degree centrigtrade
Da	: Dalton
g	: Gram

h	: hour
K	: degree Kelvin
Kcat	: Turn Over number
Km	: Michaelis-Menten constant
Ksv	: Stern-Volmer constant
m	: meter
M	: Molar
mm	:milli meter
nm	: nanometer
U	: Unit
Vm	: Mattew's number
λ	: Lambda/Wave length
Σ	: sigma/Sum
χ	: Chi
ψ	: Psi
Φ	: Phi

ABSTRACT

This thesis describes the research on C-phycocyanins (C-PC) from three cyanobacteria and an alkaline thermoactive xylanase (AT*B*XYL-C) from a *Bacillus* species. The first chapter involves a general introduction to the proteins under study, based on extensive review of available reports, publications and communications. The second chapter incorporates all the materials and methods employed in the course of the work. Next two chapters (3 & 4) record the results of various experiments on cyanobaterial phycocyanins, their analysis and comparison to existing knowledge. Similarly, last two chapters (5 & 6) involve documentation of results and relevant discussion about the xylanase.

The major light-harvesting capacity of prokaryotic cyanobacteria and eukaryotic red algae is associated with large antennae complexes called phycobilisomes located on the surface of the photosynthetic thylakoid membranes. The phycobilisomes are composed of rods and a core, consisting of various phycobiliproteins and linker polypeptides. The phycobiliproteins are divided into three major classes: phycoerythrins, phycocyanins and allophycocyanins. The rods in phycobilisome normally have phycocyanin and the core has allophycocyanins. C-phycocyanins are composed of α and β subunits exhibiting high mutual affinity to form ($\alpha\beta$) monomers, which in turn aggregate into $(\alpha\beta)_3$ trimers and $(\alpha\beta)_6$ hexamers. Phycocyanins not only absorb light energy but also transfer the absorbed energy from phycoerythrins to allophycocyanins in the core and finally to the photosynthetic reaction center. In this thesis, we report the purification, crystallization and crystal structure analysis of C-PCs from the Indian cyanobacteria *Phormidium* and *Lyngbya spp.* of marine habitat and *Spirulina sp.* of freshwater habitat. The crystal structure analysis of C-PCs has thrown light upon how the organization of C-PC units as seen in crystals helps in energy transfer.

Xylanases produced by extremophiles are important due to their biotechnological applications and as model systems in structure-function studies. The paper and pulp industries use xylan-degrading enzymes for the pretreatment of paper pulp to enhance the bleaching effects. By using xylanases the quantity of bleaching chemicals can be reduced,

thereby reducing the amount of toxic by-products and results in cost benefits also. Many studies on xylanases belonging to family G/11 have been reported. However, information, especially structural characterization, about very few alkaline xylanases has been reported. The active site characterization, fluorescence studies and determination of three-dimensional structure are of great interest, since they can lead to a better understanding of the catalytic mechanism and structural stability which in turn helps in the rational design of enzyme with selected specificity and enhanced stability. In the second part of the thesis detailed structural and active site characterization and fluorescence studies on an alkaline thermoactive xylanase from an alkalophilic *Bacillus* sp (NCL 87-6-10) is presented.

Chapter 1

General introduction

This chapter deals with general introduction, history, classification, role and application of C-phycocyanins (C-PCs) and xylanases. A detailed survey of literature on different algal C-phycocyanins and bacterial family G/11 xylanases is presented in this chapter.

C-phycocyanins

Phycobiliproteins are major photosynthetic accessory pigments, found primarily in prokaryotic cyanobacteria, and two groups of eukaryotic algae, the red algae and cryptomonads. The red algae and cyanobacteria also contain chlorophyll-a, and the cryptomonads contain both chlorophyll-a and c. The chlorophylls are located in the thylakoid membranes, whereas the phycobiliproteins are found on the exterior of these membranes. The phycobiliproteins have capacity to harvest light energy at wavelengths at which chlorophyll-a shows less absorption. This property of phycobiliproteins is especially useful in underwater habitats, where the light is both attenuated and altered in spectral distribution by passage through water. In particular, light tends to become blue green as it proceeds down the water column, and this light is better absorbed by phycobiliproteins than by chlorophyll-a. The rods in phycobilisome normally include phycocyanin; they sometimes also contain phycoerythrin or phycoerythrocyanin at the tips of the rods in some species. The phycobilisome core is composed of

allophycocyanins and linker polypeptides. The hexameric CPCs, the major component of the rods, not only absorb light energy but also transfer the absorbed energy from phycoerythrins to allophycocyanins in the core. The energy is finally transferred to the photosynthetic reaction center.

Xylanases

The increased interest in the study of extremophilic microorganisms can be attributed mainly to their importance as biotechnological resources and their specific properties are expected to find applications in novel industrial processes. Useful insights gained from the study of enzyme systems in extremophiles can extend the understanding of protein chemistry along with expanding the horizon of application of biocatalysts. Xylan is the major component of hemicellulose, which accounts for 30% of the dry weight of some plant tissues. Microbial xylanases are the preferred catalysts for xylan hydrolysis due to their high specificity, mild reaction conditions, and negligible substrate loss by the efficient generation of product. Cellulase-free xylanases have attracted much interest, especially in paper and pulp industry. Xylanases play an important role in the development of eco-friendly processes by reducing the use of hazardous chlorine chemicals in bleaching processes. They are also important from the point of view of fundamental research, since multiple xylanases produced by an organism have diverse structure-function correlation. The commercial and academic importance of extremophilic microorganisms, prompted us to carry out biochemical and structural studies of a xylanase from an alkalophilic *Bacillus* sp.

Chapter 2

Materials and methods

This chapter describes the details of materials and methods used for the preparation of C-phycocyanin from *Spirulina sp., Phormidium sp., Lyngbya sp.* and alkaline thermoactive xylanase from alkalophilic *Bacillus sp.,* their fermentation, purification, crystallization, X-ray diffraction data collection, data processing, structure determination, structure refinement, analysis of the refined structure and biochemical/biophysical characterizations.

The proteins were purified by ion-exchange chromatography using DEAE-sepharose and CM-sepharose and by Gel-filtration using G-100. Hanging-drop vapour-diffusion method was used for crystallization. The X-ray data were collected using Raxis IV^{++} detector mounted on a rotating anode at NCL, Pune. India. The X-ray images were processed using DENZO and SCALEPACK program in HKL suite. The crystal structures were determined using molecular replacement technique implemented in AMoRe and PHASER. The REFMAC5 program was used for structure refinement in cycles along with the program QUANTA for display and model fitting. The programmes of CCP4 suit were used for most of the calculations.

Chapter 3

Purification, crystallization and preliminary X-ray characterization of C-phycocyanins from *Phormidium* and *Lyngbya spp.* (marine) and *Spirulina sp.* (fresh water).

C-phycocyanins from three cyanobacterial cultures of *Phormidium, Lyngbia spp*. (marine) and *Spirulina. sp* (freshwater) were purified to homogeneity and crystallized by hanging-drop vapour-diffusion method. Blue-coloured crystals in different crystal forms, monoclinic for all three species and hexagonal for two, were obtained. The crystals took 1–12 weeks to grow to full size using polyethylene glycols of different molecular weights as precipitants.

The biochemical and biophysical properties, i.e. molecular weight, stability etc. of C-PCs from fresh water and marine forms showed differences. The C-PCs were crystallized to investigate their structural differences in marine and fresh water forms. All of them crystallized in monoclinic space group $P2_1$ with unit-cell parameters a = 107.33, b = 115.64, c = 183.26 Å, β= 90.03° for C-PC from *Spirulina sp.* and had similar parameters for CPC crystals from *Phormidium and Lyngbya spp.* also. Crystals belonging to the hexagonal space group $P6_3$, with unit-cell parameters a = b = 154.97, c = 40.35 Å and a = b = 151.96, c = 39.06 Å, were also obtained for C-PCs from S*pirulina and Lyngbya spp.* X-ray diffraction data was collected at resolutions of 3.0 Å for the monoclinic and 3.2 and 3.6 Å for the hexagonal forms. The estimated solvent content was

around 50% for the monoclinic crystals assuming the presence of two hexamers, $(\alpha\beta)_6$, per asymmetric unit. By assuming the presence of one $(\alpha\beta)$ monomer per asymmetric unit the solvent content estimated for the hexagonal crystals was 66.5 and 64.1% respectively for *Spirulina and Lyngbya spp.*.

Chapter 4

Crystal structure analysis of C-phycocyanins from *Phormidium, Lyngbia* and *Spirulina spp.*

The crystal structures of C-phycocyanins from cyanobacteria (blue-green algae) *Phormidium, Lyngbya* and *Spirulina spp*. have been determined using molecular replacement technique. After refinement the final crystallographic R_{factor} and R_{free} were 19.9 and 24.0 (*Spirulina sp.* C-PC), 19.68 and 24.6 (*Phormidium sp.* C-PC) and 21.5 and 25.7 (*Lyngbya sp* C-PC*),* respectively, for monoclinic crystals of data at 3.0 Å resolution. The asymmetric unit contained two $(\alpha\beta)_6$ hexamers, each hexamer may be considered the functional unit in the native antenna rod of cyanobacteria.

The molecular structure resembles that of other reported C-PCs. The aggregation of two $(\alpha\beta)_6$-hexamers in the asymmetric unit suggests pathways of lateral energy transfer between adjacent hexamers involving phycocyanobilin chromophores. The chromophores involved, distance between two adjacent hexamers in the asymmetric units and conformation were different in C-PC crystals of marine (*Phormidium and Lyngbia spp.*) and freshwater (*Spirulina sp*) forms. In the asymmetric units monoclinic crystals of *Spirulina* C-PC the chromophores are closer to each other. Whereas the chromospheres between two hexamers in the asymmetric units of the crystals of *Phormidium* and *Lyngbya* C-PCs are not close. Conformation and closeness of chromophores between two hexamers in the asymmetric units of C-PCs from *Spirulina sp.* can facilitate efficient energy transfer. The growth of *Spirulina sp.* is faster compared to *Phormidium* and *Lyngbia spp.*. However, a careful analysis it is found that the difference in asymmetric

units, are a results of the choices of origin in the asymmetric unit. Thus both type of interactions are found in all C-PC crystals.

Chapter 5

Crystallization, active site characterization and stability studies of xylanase from alkalophilic *Bacillus sp.*

This chapter discusses the crystallization, chemical modification of active site residues and fluorometric studies on the stability of xylanase from *Bacillus sp.*

Initial screening provided conditions that gave thin crystals which were subsequently improved to get better crystals and higher resolutions of diffraction data. Enzyme inactivation using the residue or group specific reagents such as N-bromosuccinimide, Woodward's reagent K and Phenyl glyoxal combined with substrate protection studies identified Trp and Arg in substrate binding while carboxyl group in catalysis. Two population of tryptophans, one in polar and the other in hydrophobic environment, were detected by the intrinsic fluorescence spectrum of native xylanase. Thermal behaviour showed an increase in activity of the enzyme till 60°C and then a sudden drop at 70°C, associated with 2 to 3 times increase in fluorescence ascertaining role for Trp in activity.

1-Anilino-8-sulfonic acid (ANS) binding data showed presence of exposed hydrophobic patches at extreme acidic pH of the enzyme. Although, enzyme is inactivated by 1M guanidium hydrochloride (Gdn-HCl), substantial increase in fluorescence occurred only at 3M, further increase in Gdn-HCl resulted in a red shift in λ_{max} which could be an indication of unfolding. Fluorescence quenching studies using acrylamide, potassium iodide (KI) and cesium chloride (CsCl) showed maximum quenching due to acrylamide and no quenching due to CsCl. Quenching of fluorescence by KI is interpreted as indication of positively charged Trp environment. Maximum intrinsic fluorescence quenching by simple xylo-oligosaccharides such as xylobiose (X2)

and xylotriose (X3) were only 4%, 9%, respectively, indicating that the affinity of xylanase for latter is higher. All these observations have been correlated with the structure analysis in the next chapter.

Chapter 6

The three-dimensional crystal structure analysis of xylanase in two crystal forms, orthorhombic and tetragonal.

This chapter presents the complete three-dimensional structure analysis of the alkaline thermoactive xylanase from alkalophilic *Bacillus* sp. The structure solution was obtained by molecular replacement methods using PDB code 1H4G as model. The structures are refined at resolutions 2.8 (space group $P2_12_12_1$) and 2.4 Å (space group $P4_32_12$) to R_{factor}, R_{free} 16.9, 23.0 and 21.0, 24.0, respectively. Coordinates of space group $P2_12_12_1$ and $P4_32_12$ were deposited in PDB with PDB ID 2F6B and 2NQY respectevely. Comparison between the crystal structures of monoclinic and tetragonal forms and with other reported structures of family G/11 xylanases are discussed.

Chapter-1

INTRODUCTION

1.0 Introduction : Phycocyanins

Photosynthesis is the primary biochemical process by which solar energy is trapped by plants and is made available to the rest of the living kingdom. This is the primary method of harvesting solar energy by the biosphere and the major source of energy on earth.

The essential part of photosynthetic system is a set of pigments that can absorb and trap the available light radiation. 'Chlorophyll-a' is the common pigment found in all photosynthetic organisms. It transforms radiation energy into chemical energy. There are also accessory pigments that serve to absorb the radiation in those regions that are unabsorbed by chlorophyll-a. These include chlorophylls-b, c, d & e, carotenoids, xanthophylls and phycobilins. These pigments basically vary in the nature of absorption spectrum. This variation in the absorptive capability serves the organisms to colonize habitats with distinct illumination profile.

First part of this thesis concerns the study of the crystal structures of phycobiliproteins called C-phycocyanins (C-PCs). These proteins function as accessory light harvesting molecules in cyanobacteria. The representative forms chosen for the studies include *Phormidum, Lyngbya spp.* that are from marine habitat and *Spirulina sp.* of a fresh water origin. The structures are analyzed to understand the nature of multimerisation and its implications to energy transfer in these organisms.

1.1 Phycobiliproteins as light-harvesting apparatus in cyanobacteria and red algae

Phycobiliproteins are multimeric and chromogenic protein complexes involved in light absorption in cyanobacteria, red algae and cryptomonads. They are covalently linked to chromophoric prosthetic groups called bilins. Therefore, these proteins are also known as phycobiliproteins (PBP) or simply biliproteins (Lemberg, l928). The bilin prosthetic group is related to the bile pigment biliveridin, which serves as its biosynthetic

precursor (Glazer, 1985). Its special absorption spectrum allows it to absorb radiation in regions where pigments such as chlorophyl-a, b and carotenoids have reduced absorptivity (Fig.1.4). Phycobiliproteins are tools for physiological adaptation of life in habitats that experience reduced irradiation (subaquatic zones).

Fig. 1.1 Filamentous single cells of *Spirulina sp.*

(taken from Bidr.bau.ac.il/…/biotech/algal/va22.html)

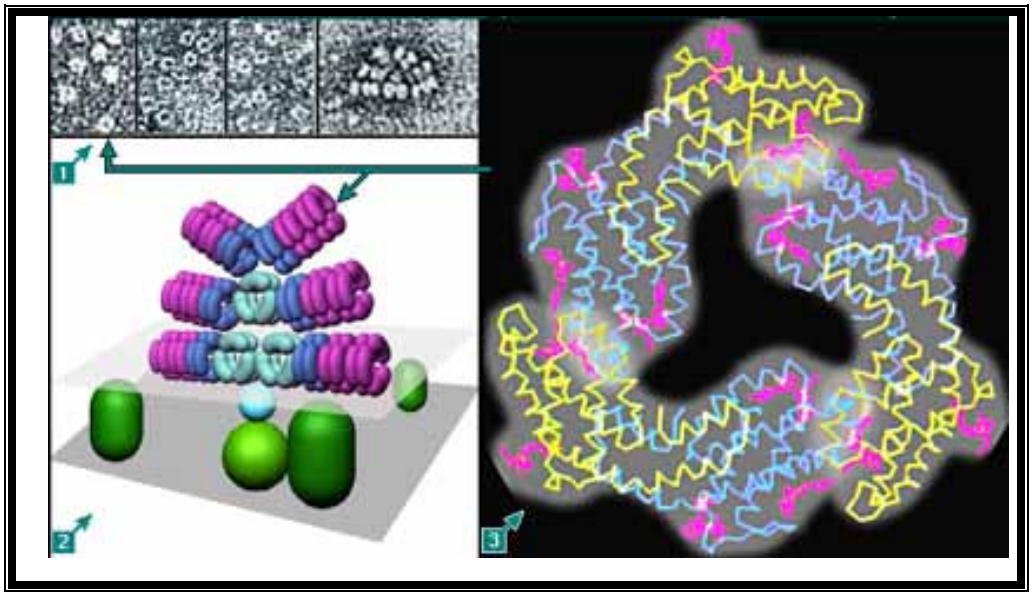


Fig. 1.2

Fig. 1.2

(1) Components of phycobilisomes in negatively stained preparations. (i) phycoerythrin aggregates, probably present as double disks with 3α and 3β subunits in each disc. (ii) C-PC aggregates showing the central cavity of the discs more clearly. (iii) allophycocyanin aggregates from the core of the phycobilisomes. (iv) a complete phycobilisome, showing 3 allophycocyanin aggregates in the core, with radiating rods each consisting of one phycocyanin (nearest to the core) and 2 distal phycoerythrin aggregates.

(2) A reconstruction of a phycobilisome, showing allo- phycocyanin cores (cyan colour) and 6 rods each consisting of 1 phycocyanin hexamer (blue) and 2 phycoerythrin (purple) hexamers. The phycobilisome sits on the outer face of a thylakoid membrane (translucent rectangle), complexed to a Photosystem II unit (pale blue and green spheres). Unattached Photosystem I (green cylinders) is also seen.

(3) Molecular organisation of a phycoerythrin trimer, consisting of 3 units each composed of an α (blue) and a β (yellow) protein. Chromophore molecules, which are covalently linked to cysteine residues, are shown in purple.

(Figure is taken from:
http://www.rsbs.anu.edu.au/profiles/Brian_Gunning/Web%20PCB/Ch%2010%20Plastids/Topic%2007%20Chloroplasts-Rhodophyta/10%2007%2016.htm)

1.2 Bilin chromophores

Bilins are the light absorbing chromophores in phycobiliproteins. They are open chain tetrapyrroles, covalently linked to specific cysteine residues of the apoprotein through thioester bonds. Bilins occur in four different forms in cyanobacteria. They are phycocyanobilin (PCB), phycoviolobilin (PXB) phycoerythrobilin (PEB) and phycourobilin (PUB). They are attached through their A rings or they have joint

attachments through their A and D rings. Figure 1.3 shows the chemical structures of four types of bilins and the position of their linkage to the apoprotein.

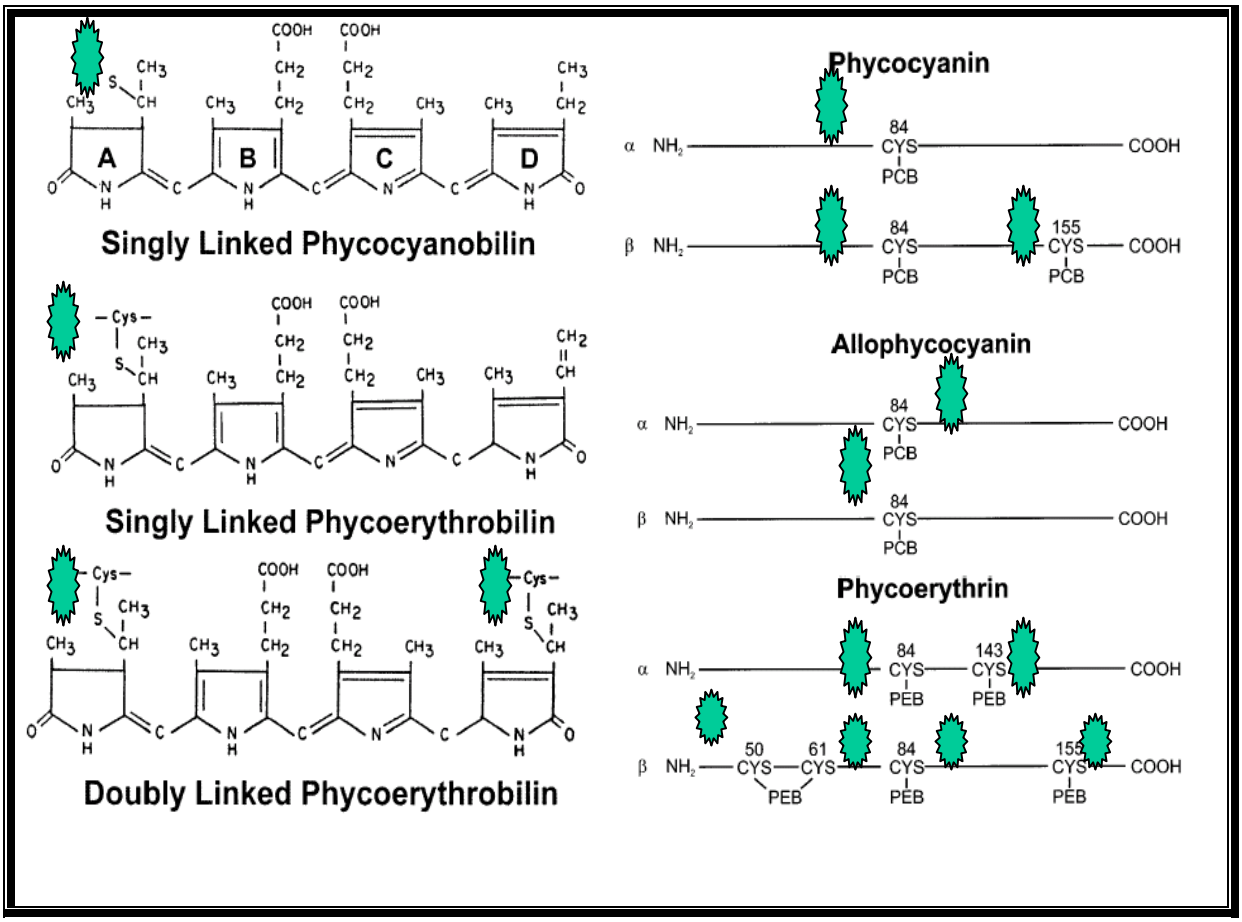


Fig.1.3 The chemical structures of chromophores: Singly linked phycocyanobilin (PCB) and singly and doubly linked phycoerythrobilin (PEB). Symbol indicates the thio-ether bond of chromophore to PBPs. (* modified from MacColl 1998.).

The variation in bilin structure is due to quantitative variation of conjugation. Phycocyanobilin has maximum conjugation and hence absorb low energy photons (λ_{max} = 610-620 nm). Phycoerythrobilin has less conjugation than PCB and hence absorbs slightly higher energy photon (λ_{max} = 545-565 nm) and phycourobilin being still less conjugated absorbs the highest range of absorption at around 495 nm. The spectroscopic

properties of each bilin within a PBP are strongly influenced by the conformation and environment imposed on them by the respective apoprotein.

The role of protein follows a hierarchical pattern in determining the properties of PBP complexes. Firstly, it influences the configuration and conformation of chromophores by providing a specific geometry. Secondly the protein serves as a scaffold in determining the special orientation of arrays of chromophores in a complex. Thirdly, the protein serves as a dielectric medium for energy transfer. And lastly, it mediates the interaction with other components, such as linkers, for the formation of supramolecular assemblies (phycobilisomes) (Huber, l99l). Further, the interaction of apo-protein with the bilins is also responsible for overlap of fluorescence and absorption spectrums of respective bilins, which is necessary for the efficient transfer of excitation energy. Figure 1.4 shows the spectral absorption range of each biliprotein and the overlap between them.

The variation in the spectral absorptivity of the bilins provides physiological diversity for utilization of different qualities of radiation.

1.3 Phycobiliproteins: Classification based on the nature of chromophore

On the basis of spectral absorption properties, phycobiliproteins are classified into four categories: Phycoerythrin (PE), Phycoerythrocyanin (PEC), C-Phycocyanin (C-PC) and Allophycocyanin (APC). The absorptive differences are due to the presence of different forms of bilins.

Table 1.1 Chromophoric content of the different phycobiliproteins.

Phycobiliprotein	Bilin
C-Phycoerythrin (PE)	Phycoerythrobilins (PEB)
C-Phycocyanin (C-PC)	Phycocyanobilins (PCB)
Allophycocyanin (APC)	Phycocyanobilins (PCB)
Phycoerythrocyanin (PEC)	Phycocyanobilin & Phycoviolobilin

The spectroscopic properties of individual phycobiliproteins depend in large measure on the chemical nature of the cognate bilin and effect of the microenvironment and geometry imposed by the apoprotein.

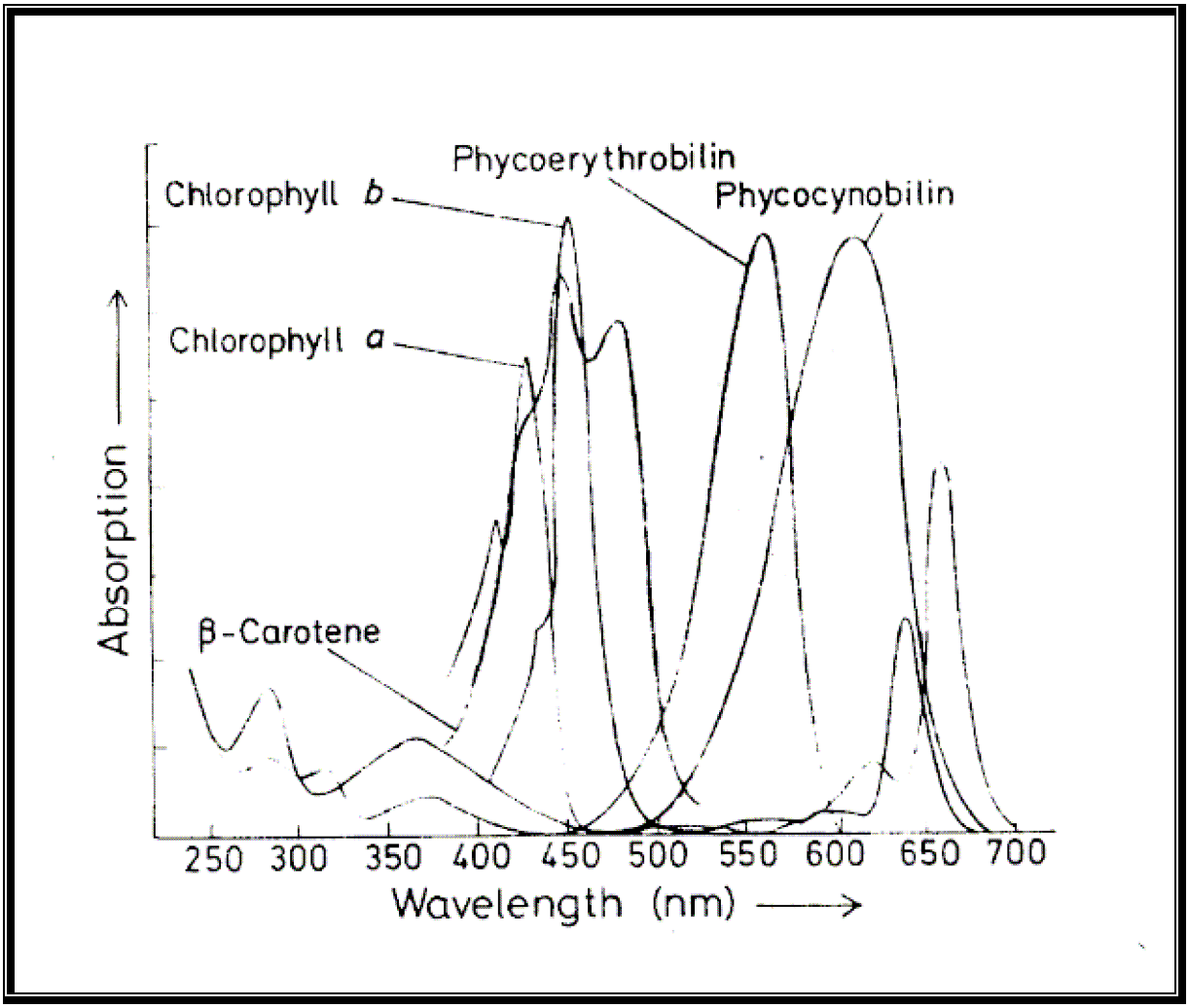


Fig. 1.4 Absorption spectra of light harvesting pigments.
(Glazer *et al.*, 1982, l983, 1985).

1.4 Phycobiliproteins: Structural organization

Evidence from electron microscopy and spectroscopic studies indicate that phycobilisome complexes share common structural features. Each phycobiliprotein contains two dissimilar polypeptide chains, α and β, of approximate molecular mass of 18 kDa and 19 kDa respectively. Each polypeptide chain carries one or more covalently attached bilins. The amino acid sequences of a number of PBPs have been determined. The sequences of α and β units are related to each other and to those of corresponding subunits of other PBPs. All these proteins are considered to have evolved from a common ancestral protein. The α- and β-subunit polypeptides have affinity to each other and form (αβ)-monomers which in turn aggregate into $(\alpha\beta)_3$ trimers and $(\alpha\beta)_6$ hexamers. The trimers are hollow cylindrical disc-shaped assemblage with a thickness of 30Å and a diameter of 110Å. The hexameric complexes are face-to-face dimers of trimeric assemblies.

Purification of native PBPs by conventional methods frequently leads to isolation of monomeric (αβ), trimeric $(\alpha\beta)_3$, or hexameric $(\alpha\beta)_6$ complexes (Berns & MacColl, 1989). In a few cases they are isolated in the form of $(\alpha\beta)_2$ dimers or as equilibrium mixtures of varying aggregates. The outcome depends on the purification protocol and on the source organism (Berns & MacColl, l989). Certain PEs are isolated as hexamer assemblies $(\alpha\beta)_6$ with a γ chain sitting in the middle of the hexamer.

Phycoerythrin (PE) has five phycoerythrobilins (PEB), two on the α-polypeptide and three on the β-chain. In the α-chain the chromophores are bound to cysteine at α-84 and α-143, and in the β-chain one chromophore doubly bound to β-50 and β-61 and the remaining two are attached at β-84 and β-155. C-Phycocyanin (C-PC) contains three phycocyanobilins (PCB), one on α-polypeptide chain (α-84) and two on β-chain (β-84 and β-155) (Table. 1.1). Allophycocyanin (APC) has two phycocyanobilins (PCB) as prosthetic groups, one on α-84 and another on β-84.

1.5 Structural organization of phycobilisomes

Phycobiliproteins occur as discreet organelle-like assemblies called phycobilisomes. Electron microscopic observations have indicated that phycobilisomes assume a periodic arrangement on the outer surface of thylakoids (Gantt et al., 1965, 1976; Bryant et al., 1991; Giddings et a1., 1983; MacColl, l998).

The external morphology of phycobilisomes varies with the source organism (Gantt et al., 1965, 1969, 1977: Glazer et al., l979; Bryant et al., 1979). The particles may be ellipsoidal, hemidiscoidal or a bundle of rod shaped elements. These differences in gross morphology do not reflect fundamental differences in the placement of major phycobiliproteins or in the functional properties of the particles (Glazer, l982). Phycobilisomes range in size from 7-15 million daltons. Hemidiscoidal phycobilisomes are most common among cyanobacteria.

Phycobilisomes contain two morphologically distinct domains - a 'core' made up of cylindrical elements laid horizontally above the thylakoid membrane to form the core and 'rods' that emanate in a centrifugal fashion. Figure 1.7 shows a schematic representation of phycobilisome light-harvesting assembly. It shows a tri-cylindrical core element with six rods emanating out of the core. In general, the number of cylinders that compose the rod or core also vary depending on the organism possessing them. Further, the lengths of each cylinder (containing stacked disks of PBPs) can vary depending on the physiological conditions and the quality of light available for a particular organism. Figures 1.5 & 1.7 show the most widely seen 'tri-cylindrical, six rods model' of phycobilisomes. The core is composed of APC and cognate linker polypeptides. The rods are composed of C-PCs and PE/PECs along with their respective linkers. While C-PCs invariably occupy the base of the rod, PE/PECs occupy the tip.

The structural organization of various PBP elements into rod and core envisages the formation of a spectroscopic ladder. The emission range of PE/PEC overlaps with the absorption range of C-PC. Further, the emission range of C-PC overlaps with the

absorption range of APC. Hence the absorbed energy at the tip of the rod migrates in a downhill fashion towards the core of phycobilisome assembly (Gantt, 1973; Glazer, 1989; MacColl, l998). The spectroscopic overlap between subsequent biliproteins in the path of energy transfer is further enhanced by the presence of different linkers within the biliproteins.

The interaction of the linkers with the biliprotein assembly and the chromophores therein is believed to be responsible for the convenient modification of spectral properties (Yu & Glazer, 1982). The spectroscopic studies with and without the presence of linkers have confirmed this view (Maxson et al., 1989; Pizarro & Sauer, 2001). These studies have revealed that interactions with linker shift the overall spectra towards red (longer wavelength) region. The modified spectral response due to the presence of the linkers increases the rate of energy transfer. The directional energy transduction occurs from PE/PEC → C-PC → APC → PSII (reaction center). The optimal spectral overlap between biliproteins ensures a highly efficient transfer of energy from one protein to another in a non-radiative fashion. The overall efficiency of excitation transfer within the phycobilisome assembly is estimated to be greater than 95% using a variety of kinetic and spectroscopic studies (Glazer et a1, 1985).

Linker polypeptides form crucial building blocks of phycobilisomes. They assist in binding the PBP disks together to form ‘rods’ and ‘core’. They are also known to interact with the chromophores of PBPs and modify their spectral properties favorably for energy transfer. Some of the linkers, which carry chromophores, are named γ subunits to distinguish from α and β polypeptides of PBPs.

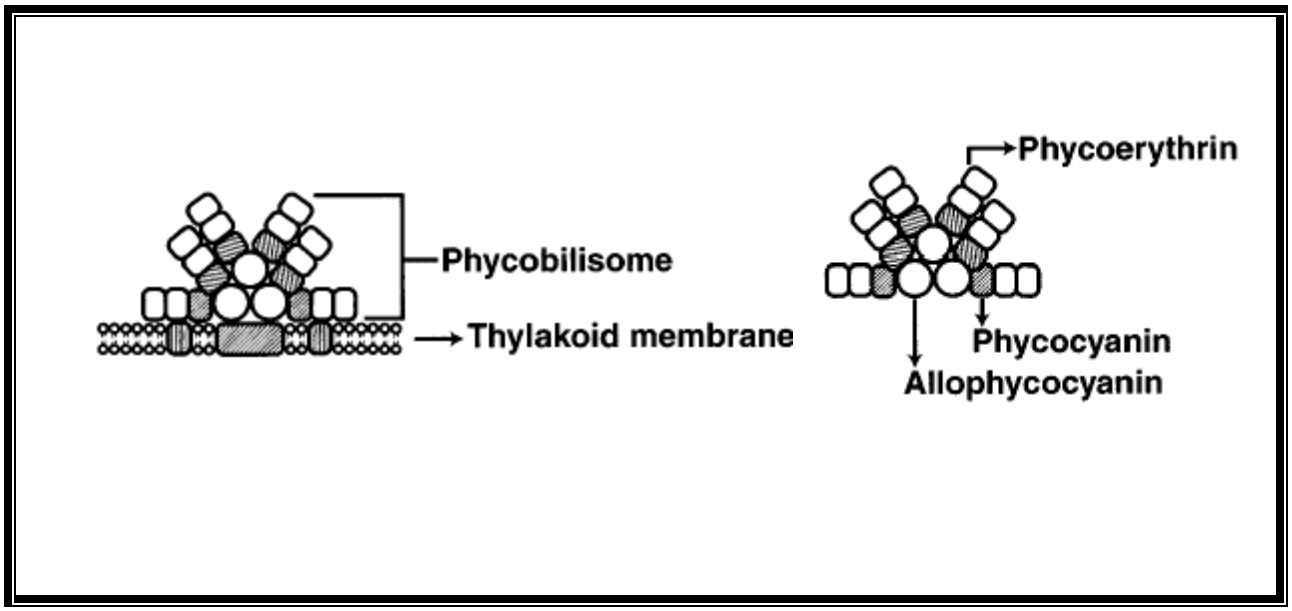


Fig. 1.5 Models of tricylindrical & hemidiscoidal phycobilisomes. (adapted from MacColl 1998.)

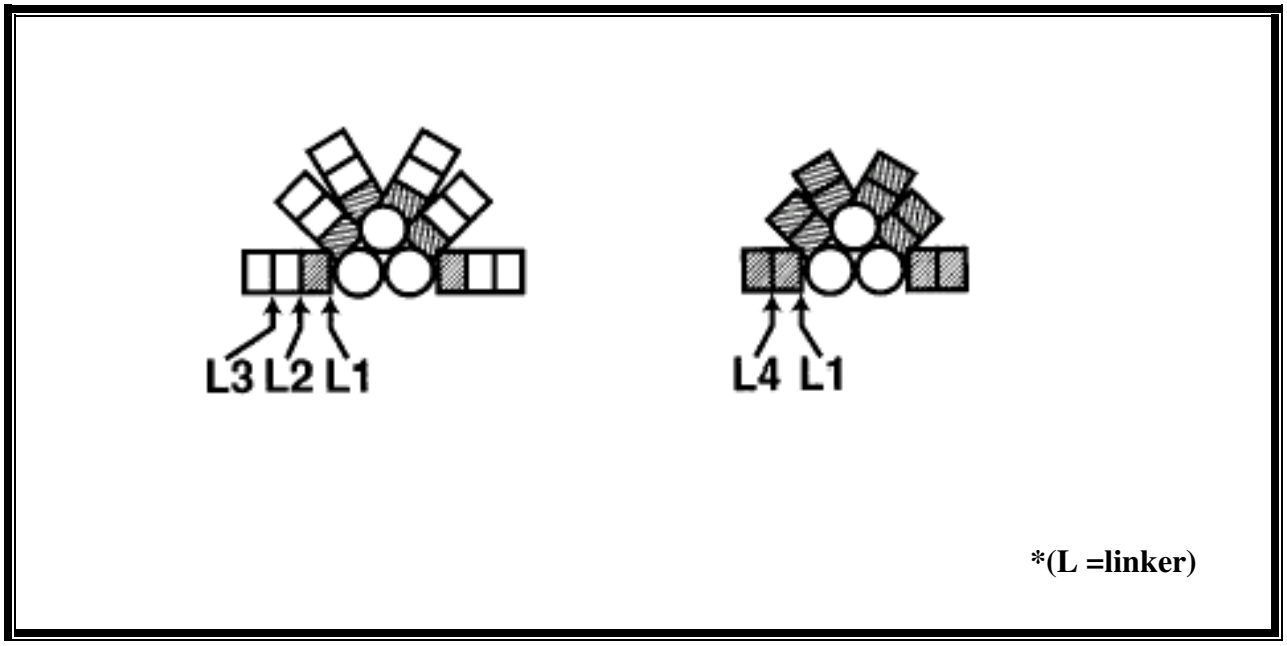


Fig. 1.6 Distribution of linkers in phycobilisomes (adapted from MacColl 1998.).

Table 1.2 Types of linker polypeptides present in phycobilisome

Linker size	Function
27000 (L1)	joins C-PC disk to core
33500 (L4)	joins second C-PC disk to first C-PC disk
31500 (L2)	joins first C-PE disk to C-PC disk
30500 (L3)	joins peripheral C-PE disk to end of rod

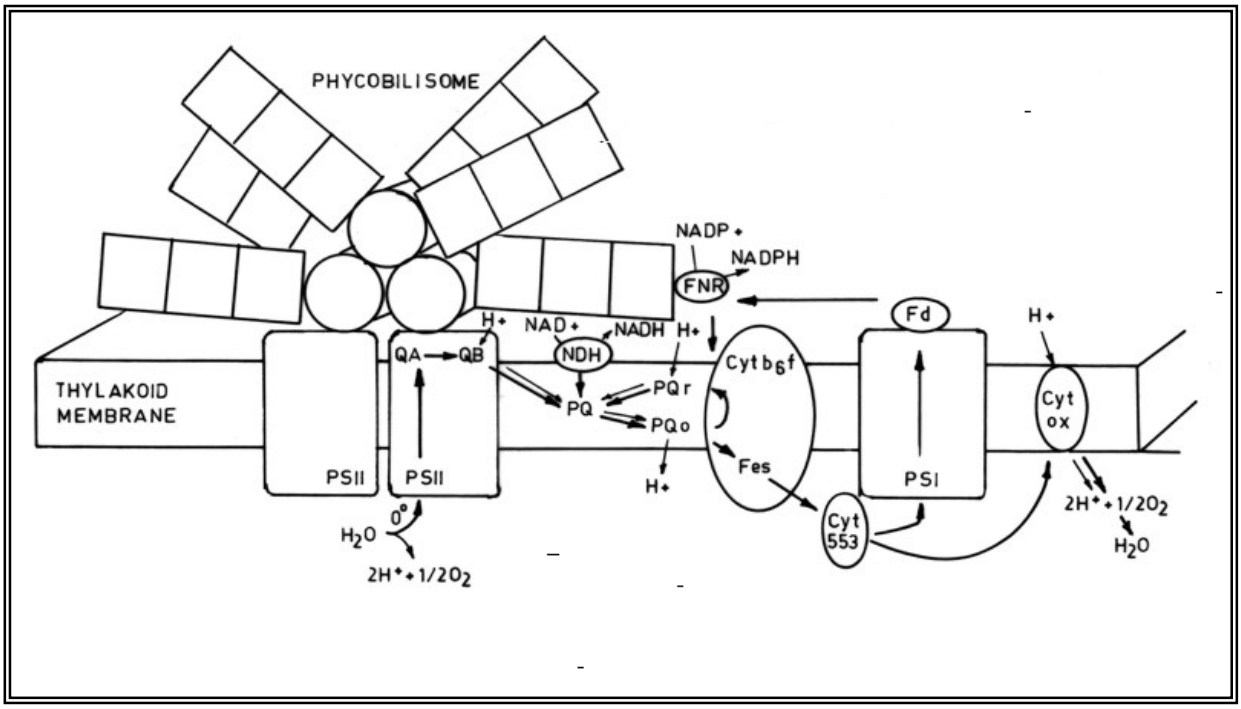


Fig. 1.7 Model of a thylakoid membrane showing photosynthetic electron transport chain (Glazer *et al.,* 1979, l983).

1.6 Crystal structures of phycobiliproteins

The first PBP to be solved by X-ray crystallography at high-resolution was C-PC from *M. laminosus* in 1985 (Schirmer *et a1*., 1985) followed by C-PC from *Agmenellum quadruplicatum* resolved at 3.0 Å, (Schirmer *et al*., 1986). Subsequently, both the structures were refined at higher resolution (Schirmer *et al*., 1987). In subsequent years C-PC from *Fremyella diplosiphon* (Duening *et al*., l99l) and structures of other phycobiliproteins namely PE from *Porphyridium sordidum*, PEC from *M. laminosus* and APC from *S. platensis* were reported, all from Prof. Huber's group at the Swiss Institute of Technology, Zurich, Switzerland (see Table 1.3 for respective references and details). Figure 1.8 shows the representative hexamer structure of linker free form of C-PC from *Spirulina platensis*. It may be noted that Fisher *et al*. (1980) published low-resolution crystal structures of C-PC and PE. However, they could drive very little information about the atomic level details aspects of the structures.

Table 1.3 List of the reported crystal structures of PBPs

PDB ID Year	PBP	Organism	Resolution (Å)	Space group	Asymmetric Unit	R_{cryst} & R_{free}	Reference
- 1980	B-PE	*Porphyridium cruentum*	5.0	R3	$(\alpha\beta)_2\gamma_{1/3}$	-	Fisher *et al*,. 1980
- 1992	R-PE	*Porphyridium sordidum*	2.2	R3	$(\alpha\beta)_2$	18.9% -	Fincer *et al*., 1992
- 1993	B-PE	*Porphyridium cruentum*	2.3	R3	$(\alpha\beta)_2$	18.4% -	Fincer & Huber, 1993.
1lia 1996	R-PE	*Porphyridium urceolata*	2.8	R3	$(\alpha\beta)_2$	18.0% -	Chang *et al*., 1996.
- 1990	PEC	*Mastigocladus laminosus*	2.7	$P6_3$	$(\alpha\beta)$ mono	19.2% -	Duerring *et al*., 1980.
- 1980	C-PC	*Anabaena variabilis*	5.0	$P6_3$	$(\alpha\beta)_2$	-	Fisher *et al*., 1980
- 1985	C-PC	*Mastigocladus laminosus*	2.1	$P6_3$	$(\alpha\beta)$ mono	21.7%	Schirmer *et al*., 1985. Schirmer *et al*., 1987.
- 1986	C-PC	*Agmenellum quadruplicatum*	2.5	R32	$(\alpha\beta)$ mono	18.4%	Schirmer *et al*., 1986.. Schirmer *et al*., 1987.
1cpc 1994	C-PC	*Fremyella diplosiphon*	1.66	R3	$(\alpha\beta)_2$	18.1%	Duerring *et al*., 1991.
1all 1995	ACP	*Spirulina platensis*	2.3	$P6_322$	$(\alpha\beta)$ mono	19.6%	Brejc *et al*., 1995.
1phn 1999	C-PC	*Cyanidium caladarium*	1.6	R32	$(\alpha\beta)$ mono	18.4 & 27.2	Stec *et al*., 1999
1b33 1999	APC	*Masatigocladus laminosus*	2.3	$P2_12_12_1$	2[$(\alpha\beta)$3+Lc7.8]	21.11 &	Reuter *et al*., 1999.

- 1999	APC	*Porphyra yezoensis*	2.2	R32	(αβ) mono	25.68 19.3 & (26.9)	Liu *et al.*, 1999.
- 1999	RPE	*Polysiphonia uceolata*	1.9	R32	(αβ) mono		Jiang *et al.*, 1999
1qgw 1999	RPE	*Rhodomonas CS24*	1.63	$P2_12_12_1$	α1α2ββ	14.9 & (18.8)	Wilk *et al.*, 1999
1b8d 1999	RPE	*Griffithsia monilis*	1.9	R3	(αβ)2 γ	17.5	Retter *et al.*, 1999.
1eyx 2001	RPE	*Gracilaria chilensis*	2.7	R3	(αβ)2 γ2	16 & 25	Contreras-Martel *et al.*, 2001
1ha7 2001	C-PC	*Spirulina platensis*	2.2	$P2_1$	2(αβ)6	19.2 & 23.8	Padyana *et al*, 2001
1gh0 2001	C-PC	*Spirulina platensis*	2.2	$P2_1$	2(αβ)6	18.9 & 23.7	Wang, *et al.*, 2001
1ktp 2002	C-PC	*Synechococcus vulcanus*	1.6	R32	(αβ) mono	21.6& 24.7	Adir, *et al.*, 2002.
1on7 2003	C-PC	*Thermosynechoc occus vulcanus*	2.7	$P6_3$	(αβ) mono	20.7 & 27.8	Adir, *et al.*, 2003.
1jbo 2003	C-PC	*Synechococcus elongates.*	1.45	R32	(αβ) mono	14.6 & 18.8	Nield, *et al.* 2003.

Comparison of these structures reveals that the tertiary fold and quaternary associations of macromolecular assemblies are remarkably conserved across organisms. All the PBP crystal systems invariably belong to rhombohedral or hexagonal systems except in the case of linker bound APC from *M. laminosus* and C-PC from *Spirulina platensis*. This is because of the coincidence of molecular symmetry with the crystallographic symmetry. In several cases, although the crystals contained linker polypeptides in the central channel, the structure of the linkers could not be resolved because of the three-way disorder. The only example of a biliprotein with a linker polypeptide in its central cavity determined so far by crystallography is that of APC from

M. laminosus with a L7.8 KDal from Huber's Group (Reuter *et al*., 1999). The structure determination was made possible because of the crystallization of the complex in an orthorhombic crystal system (space group $P2_12_12_1$). The three-way disorder of the linker was lifted by non-coincidence of molecular symmetry of APC with crystal symmetry. The C-PC structures presented in this thesis crystallized in monoclinic space group which are unique examples among biliproteins. Subsequent to the publication of *S. platensis* C-PC at 2.2 Å (Padyana *et a1,* 2001; PDB ID: 1ha7), another report of the same crystal structure also appeared (Wang *et a1*., 2001; PDB ID: lgho).

Despite the low sequence homology, the molecular structures of α- and β-polypeptides are found to be remarkably similar even among various classes of phycobiliproteins. Further, the comparison of PBP structures from different classes indicates conservation of multimeric arrangement. The formation of trimer from monomer involves 3-fold symmetry. The trimers aggregate in a face-to-face fashion to form hexamers. The face-to-face association involves a two-fold symmetry between monomers of adjacent trimers. Hence the total hexamer formation involves a 32 (D3) point group symmetry. The dimension of linker free C-PC trimer is 30Å x 110Å with a central cavity of approximately 35Å diameter. Hexameric and trimeric phycobiliproteins stack one above the other forming the rod and core structures of the phycobilisome. The central channel formed by rods and core of *in vivo* cylindrical organization is occupied by linker polypeptides. Linkers together with the PBPs constitute the phycobilisome light-harvesting assembly.

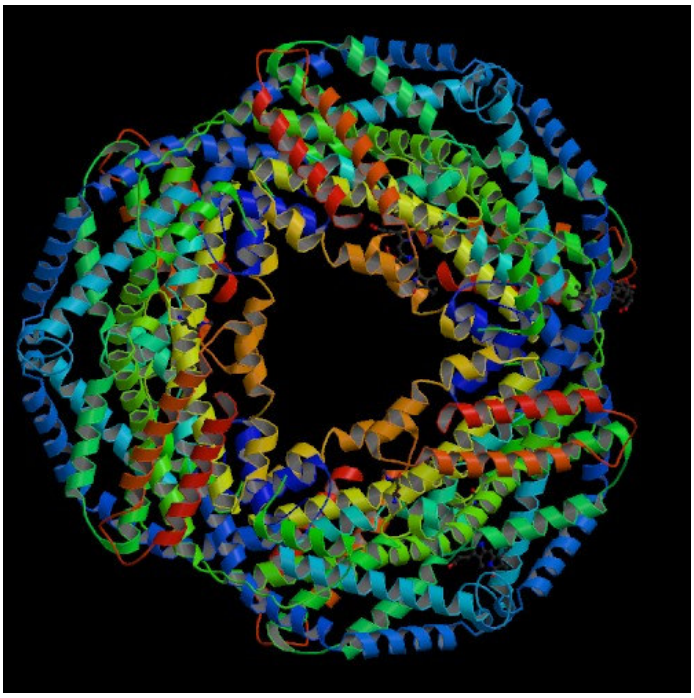


Fig. 1.8 Aggregation of C-PC monomers into trimers by three fold symmetry.

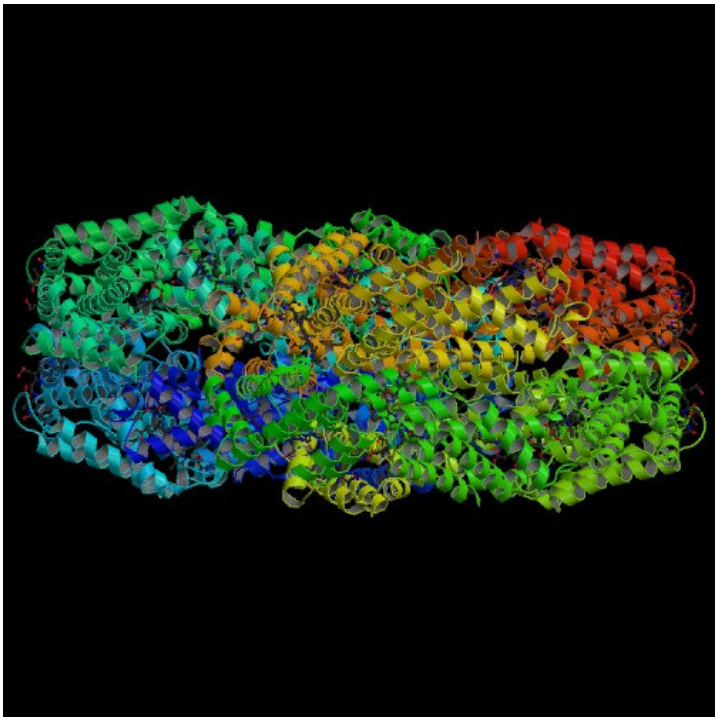


Fig. 1.9 Lateral view of the arrangement of hexamers in the monoclinic unit cell.

1.7 Function of phycobilisome in photosynthesis

Phycobilisome contain chromophoric phycobiliproteins that function as molecules that can absorb light energy at specific wavelengths. The energy is absorbed by open chain tetrapyrrole chromophores. The absorbed energy is then transferred to a photosynthetic reaction centre where it is converted into chemical energy. The energy migrates through the rods into the core and then to chlorophyll-a in the thylakoid membrane. These events are non-radiative, efficient and directional. Excitation energy transfer over several hundred chromophores typically occurs in 100 to 200 ps with 90% or higher efficiency. Phycobilisomes primarily transfer excitation energy to photosystem-II receptor, but under certain light conditions direct transfer to photosystem-I may also occur. Phycobilisomes occur in close proximity to the photosystem-II on the surface of the thylakoid membrane. There are indications that the two complexes might be docked through the ($\alpha\beta$) containing core trimer, by ferredoxin-NADP oxidoreductase or by some other mechanism (Ducret *et al.,* 1994)

1.8 Energy transfer between chromophores

Energy migration in phycobilisome is polarized, the flow being from the rods into the core and then into chlorophyll-a located in the thylakoid membrane. Fast laser kinetics, crystal structure analysis and various spectroscopic methods have been employed to study the mode of energy transfer.

Structures of various cyanobacterial phycobiliproteins has been reported by Huber *et al*. through X-ray diffraction analysis (Schirmer *et al*., 1985, 1986, 1987; Duerring *et al*., 1990, 1991; Brejc *et al*., 1995). All the trimeric and hexameric biliprotein assemblies have very similar general structure. Each polypeptide has nine α-helices connected by irregular turns and have similar tertiary structure. Each C-PC monomer contains three tetrapyrrole chromophores covalently attached to cysteine residues. Although these

tetrapyrroles are chemically identical, interactions with the protein make each of the three chromophores spectroscopically different.

The mechanism of energy transfer within multichromophore system is explained using the Forster's model of resonance energy transfer in the weak-coupling limit (Forster, 1965). Forster's theory of energy transfer for the weakly coupled case predicts that the rate of energy transfer depends on the overlap of the donor chromophore's emission spectra with the acceptor chromophore's absorption spectra.

Dale and Teale (1970) were the first to analyze the energy migration process for the isolated C-PC using fluorescence polarization. Since then many investigations have been carried out in this field. Debreczeny *et al*. (1993, 1995a, 1995b) have studied the energy transfer properties of C-PC monomers and trimers and established that Forster resonance energy transfer is the method used for the radiationless transfer of energy between the various pairs of chromophores on these two oligomers. From the results of the absorption and fluorescence spectra of the individual chromophores and the orientation of chromophores determined through X-ray crystallography (Schirmer *et al*., 1987; Duerring *et al*., 1990, 1991.), the rate constants of energy transfer for the various chromophore pairs were determined using Forster theory. In monomers, energy transfer is observed from β-155 to β-84 and α-84 to β-84, and slow energy transfer is found from β-155 to α-84. The β-84 chromophore is the lowest energy chromophore and serves as the chromophore from which energy is transferred out of the protein. The other two higher energy chromophores, β-155 and α-84, extend the range of light harvesting. The β-155 chromophore absorbs the highest energy radiation.

There are more possibilities of energy transfer in trimers, especially between closely spaced α-84 and β-84 chromophores on contiguous C-PC monomers in the trimer. The energy migration within hexamers and between different hexamers of C-PC is less known. The crystal structures reported here suggest that the β-84 and β-155 chromophores provide the pathway for energy migration between hexamers through arrangment seen in the crystals.

1.9 Economic significance of phycobiliproteins

Spirulina platensis is a unicellular filamentous blue - green algae. The genus '*Spirulina'* has eight species of which *Spirulina maxima* and *Spirulina platensis* are the widely found members. They form massive populations in tropical and subtropical waters characterized by high levels of carbonate and bicarbonate content and also high pH (Clement, 1975; Guerin-Dumartrait, & Moyse, 1975; Vonshak, 1997). *S. platensis* is the widely distributed species, found mainly in Africa, Asia and South America. Occurrence of *S. maxima* is confined to Central America. *S. platensis* is the most extensively studied and commercialized species among the *Spirulina* genus.

S. platensis contains two phycobiliproteins, APC and C-PC. C-PC is the major light-harvesting pigment protein (constituting up to 14% of biomass) present in the antenna rods whereas APC is a minor component present only at the core (Boussiba & Richmond, l979).

Recent studies have shown that the photodynamic property of *S. platensis* derived phycocyanin can be used in cancer treatment. It was found that phycocyanin specifically binds to cancer cells, thereby allowing anatomical imaging of tumors under *in vivo* conditions (Morcos, l998).

Immunomodulatory and hepato-protective activities have also been attributed to .C-PC from *S. platensis*. C-PC's fluorescence properties have been made use of in the development of phycofluoro probes for immunodiagnostics (Glazer & Stryer, l984; Siiman *et al*., 1999; William *et a1*., 2001). Recent reports indicate the participation of C-PCs in anti-inflammatory reactions (Romay *et al*., l998; Bhat *et al*., 1998).

Strong colour and non-carcinogenicity of C-PCs has made them an alternate choice in food and cosmetic industries. C-PC along with other natural colorants such as

carotenoids and chlorophylls are extensively used in food (chewing gums, dairy products, ice sherbets, jelly etc) and cosmetics (such as lipstick and eyeliners).

1.10 Scope of the work

The understanding of the crystal structure and the basis of light harvesting in a phycobilisome assembly protein (C-PCs) from marine and fresh water algal species forms one of the main themes of this thesis. The purified proteins are crystallized in a monoclinic form that helped us to look for any differences between the hexamer-haxamer association of C-PC from marine and fresh water cyanobacteria. The X-ray crystallographic structure investigations have led us to solve and refine the structure of light-harvesting phycobiliprotein C-PCs from marine and fresh water forms. The crystal structure provides a picture of variety of interactions involved in the formation of higher assemblies such as trimers, hexamers and lateral association of hexamers.

The monoclinic form of the crystal provides an exciting possibility of crystallizing a structurally identifiable linker protein along with the C-PC. The structure forms a unique example among PBPs in terms of the occurrence of whole of functional hexamer assembly and two such hexamers closely associated (> 30,000 atoms) in the crystal asymmetric unit. The refinement of the structure with the application of non-crystallographic restraints has allowed the accommodation of subtle asymmetry involved in highly symmetrical molecular assemblies. The crystal structure provides accurate structural basis for the understanding of comparable efficiencies between assemblies of proteins involving the physical mechanism of light absorption and energy transfer. The structure has also allowed determination of exact conformation of all the chromophores involved in light absorption and excitation transfer thereby providing correlation between structural differences and energy transfer efficiencies in marine and fresh water forms. The proposed model of lateral energy transfer is motivated by the occurrence in close proximity of two C-PC hexamers in the crystal asymmetric unit whose mutual association is not dictated by crystal symmetry. This is significant in the context of close association of C-PCs in antennae rods near the core region of phycobilisome light-harvesting

complex as seen by earlier electron microscopic studies (Gantt *et al*., 1976; Bryant *et al*., 1979; Glazer *et al*., 1979: Ducret *et al*., 1994). We also observe that the coupling interaction of laterally associated hexamers is through α-84 in marine whereas it is through β-l55 in fresh water forms. We also propose that this association *in vivo* could explain the difference in growth times of these organisms.

1.10 Introduction: Xylanases

Photosynthesis is one of the major reaction by which a large amount of energy is trapped as chemical energy. Lignocellulosic material is the largest byproduct of this process and as of now has had limited exploitation due to the recalcitrance of their chemical structures. The fast rate of depletion of non-renewable energy repositories has turned the focus on agricultural and forestry byproducts as a futuristic source. Abundance and renewable nature is at the root of the popularity.

The efficient enzymatic degradation of lignocellulosic materials requires a combination of hydrolyzing enzymes such as cellulases and hemicellulases. Complete enzymic hydrolysis of hemicellulose involves action of a battery of enzymes, of which endo-(1→4)-β-xylanase is one of the most crucial enzymes. Also, heightened interest has been shown in cellulase free xylanolytic enzymes from microbial sources for their potential application in the paper and pulp industry. For a successful integration of xylanases in biotechnological applications, a detailed understanding of the mechanism of enzyme action is desired.

1.11 Microbial sources of xylanases

The microorganisms that produce xylanases and other glycosidases have been found in extremely diverse habitats. Under mesophilic growth conditions, xylanolytic activity has been reported in a wide variety of different genera and species of bacteria, fungi, and yeasts (Gilkes *et al*., 1991). For example, xylan degradation occurs in certain strains of *Bacillus polymyxa*, *Bacillus pumilus, Bacillus subtilis, Cellulomonas fimi*, *Clostridium acetobutylicum, Streptomyces lividans, Streptomyces flavogriseus,*

Aspergillus fumigatus, Neurospora crassa, Trichoderma viride, Pichia stipitis, and *Candida shehatae* (Gosalbes *et al*., 1991; Wong *et al*., 1988). In addition to conditions of *meso*-temperature, a host of microorganisms inhabit extreme environmental conditions where they thrive and grow at temperatures above 50 °C, pH values euqalling or exceeding (Uffen, 1997) and/or in high ionic concentration of aqueous systems containing salt approaching saturating conditions (Wong *et al*., 1988). Use of thermo-tolerant and alkalophilic xylanases is especially desirable for kraft pulp treatment in the paper industry (Farrell & Skerker, 1992). Hyperthermophilic bacteria have been isolated that grow at temperatures above 80 °C. These microbes include *Thermotoga maritima* (Winterhalter & Liebl, 1995), *Thermotoga* sp. (Saul *et al*., 1995), *Caldocellum saccharolyticum* (Luthi *et al*., 1990), *Dictyoglomus* sp. and *Rhodothermus marinus.*

1.12 Structure of xylans, the substrate of xylanases

In plant cells, xylan is one of the three major structural polysaccharides and is localized in the cell wall matrix. The relative distribution of lignocellulosic components in the cell wall is dependent on the plant species and on the stage of growth and development. Xylans are typically polydispersed heteropolysaccharides with a homopolymeric backbone chain of 1,4-linked β-D-xylopyranosyl residues. The xylopyranosyl backbone is substituted at positions C-2, C-3 and C-5 to varying degrees depending upon the plant and the stage of development of the plant when the polymer is obtained. In monocots, at the C-2 positions 1→3-linked α-D-glucronic acid or 4-O-methyl-α-D-glucuronic acid might occur, while at C-3 of xylopyranose, one frequently finds 1→3 linked α-L arabinofuranose. In some xylans, particularly in hardwoods, xylopyranose residues may be O-acetylated at the C-2 or (more commonly) at the C-3 positions. Additionally, small amounts of phenolic components, such as ferulic and p-coumaric acids (associated with lignin) are esterified to xylan *via* their carboxyl groups to C-5 of xylose ring.

1.13 Xylanolytic enzyme system

Xylan is a complex polysaccharide containing different substituent groups in the backbone and side chains. Due to its heterogeneity, a complex enzyme system is required for the hydrolysis. Some major structural features are summarized in Fig.1.10 All the enzymes act cooperatively to convert xylan to its constituent sugar. The xylanolytic enzyme system is comprised of the following enzymes:[Endo-β-(1,4)-D-xylanase [β-(1→4)-D-xylan xylanohydrolase, E.C.3.2.1.8]: These enzymes act randomly on xylan to produce large amounts of xylo-oligosaccharides of various chain lengths. They are grouped into four different categories:

a) Non-arabinose liberating endoxylanases: This class of enzymes cannot act on arabinosyl initiated branch points at β- (1- 4) linkages and produce only xylobiose and xylose as the major end products. These enzymes can break down oligosaccharides as small as xylobiose.

b) Non-arabinose liberating endoxylanases: These are unable to cleave at α -(1-2) and α-(1-3) branch points produce mainly xylooligosaccharides larger than xylobiose. These xylanases are unable to cleave xylotriose and xylobiose.

c) Arabinose liberating endoxylanases: These enzymes can cleave the xylan chain at the branch points and produce mainly xylobiose, xylose and arabinose.

d) Arabinose liberating endoxylanases: These enzymes can hydrolyze the branch points and produce intermediate size xylooligosaccharides and arabinose.

Exo-β-(1-4) -D-xylanase [β-(1→4)-D-xylan xylanohydrolase]: These enzymes remove the single xylose units from the non-reducing end of the xylan chain.

β-xylosidase or xylobiase: These enzymes hydrolyze disaccharides like xylobiose and the higher xylooligosaccharides with decreasing specificity.

α-L-arabinofuranosidase (E.C. 3.2.1.55): These enzymes cleave side chain α-1,3 arabinofuranose from the xylan main chain. *α-(4-O-methyl)-D-glucuronidase* (E.C.3.2.1.): These can remove D-glucuronosyl or 4-Omethyl glucuronosyl residues from xylan.

Acetyl xylan esterase (E.C. 3.1.1.6)*:* These enzymes liberate acetyl groups from the xylan backbone.

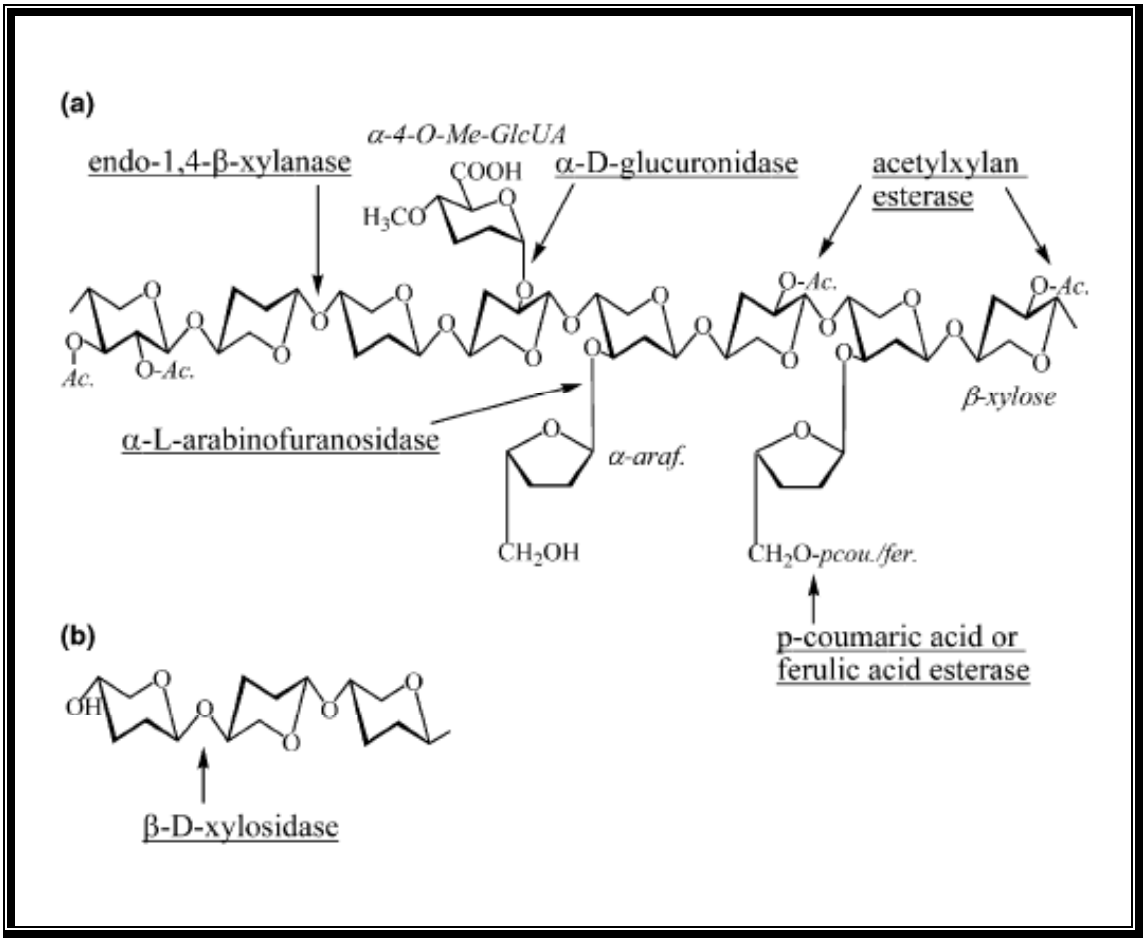


Fig.1.10 Structure of xylan and the sites of attack by xylanolytic enzymes. The backbone of the substrate is composed of 1,4-β-linked xylose residues. Ac: Acetyl group; α-araf: α-arabinofuranose; α-4-O-Me-GlcUA: α-4-O-methylglucuronic acid; pcou: p-coumaric acid; fer: ferulic acid. b Hydrolysis of xylooligosaccharide by xylosidase. Figure adapted from Collins *et al.,* 2005.

1.14 Xylanase induction and production

Under natural conditions, xylanolytic enzymes appear to be inducible by the products of their own action and are subject to catabolite repression by carbon sources such as glucose or xylose. Xylan, being a high molecular mass polymer, cannot enter the cell. Low molecular fragments of xylan namely, xylose, xylobiose, xylooligosaccharides, heterodisaccharides of xylose and glucose and their positional isomers, which are produced in the medium by small amounts of constitutively produced enzyme, stimulates induction of xylanases. Xylan has been shown to be the best inducer of xylanase production in many cases (Kelly *et al*., 1989; Nakamura *et al*., 1992). However, a few organisms show constitutive production of the enzyme (Debeire *et al*., 1990). Cellulose has also been observed to act as an inducer in a few cases (Deshpande *et al*., 1986; Morosoli *et al*., 1986; Stutzenberger & Bodine, 1992).

Induction can also be achieved by various synthetic alkyl, aryl β-D-xylosides and methyl β-D-xyloside (Nakanishi & Yasui, 1980; Marui *et al*., 1985). These compounds enable the production of xylanolytic enzymes in the absence of xylan and xylooligosaccharides. Cheaper hemicellulosic substrates like corncob, wheat bran, rice bran, rice straw, corn stalk and bagasse have been found to be most suitable for the production of xylanase in certain microbes. Wheat bran was found to be the best substrate for xylanase production by alkalophilic *Streptomyces* VP5, *Streptomyces* T-7 (Keskar *et al*., 1992) and *Penicillum funiculosum* (Mishra *et al*., 1985). Highest levels of xylanase were produced when *Trichoderma brachiatum* was grown on wood pulp. Corncob was the most suitable substrate for the production of xylanase by an alkalothermophilic *Thermomonospora* sp. (George *et al*., 2001).

Fungi produce higher levels of xylanase than bacteria and yeasts. However, fungal xylanases are generally associated with cellulases. Among fungi, the maximum activity reported is 3350 IU ml^{-1} from *Trichoderma reesei* (Haapala *et al*., 1994). However, Haltrich *et al*. (1992) reported maximum xylanase activity (27,000 IU g^{-1}) produced by solid-state fermentation of the fungus *Schizophyllum commune*. An increase

in xylanase production under solid-state fermentation has also been reported from a bacterial strain *B. licheniformis* A99.

Cellulase-free xylanases are desirable in the paper and pulp industry for effective bleaching of paper pulp without adversely affecting the quality of the pulp. Another potential application of cellulase-free xylanolytic systems is in the processing of plant fiber sources such as flax and hemp. Earlier attempts to obtain cellulase-free xylanase were made by treatment of the culture filtrate with mercurial compounds, bulk scale purification (Tan *et al*., 1987) or cloning of xylanase genes in heterologous non-cellulolytic hosts (Paice *et al*., 1988). These efforts have failed to achieve commercial success for practical applications in paper technology. However, search for naturally occurring microrganisms capable of selectively secreting high levels of xylanase have yielded promising results. Cellulase-free xylanase have been isolated from *Chainia, Saccharomonospora viridis, Streptomyces roseiscleroticus* (Grabski & Jeffries, 1991) and *Streptomyces* T-7 (Keskar *et al*., 1992).

1.15 Xylanase expression and secretion

It is important to understand the mechanism of release of xylan-degrading enzyme into the system for developing effective xylan-degrading technology. At present, much of the understanding of xylanolytic enzyme action comes from studies of microorganisms where 7 xylanases are secreted by the cells. However, in bacteria some cells appear to release the enzymes in the form of 'protein complexes' or 'xylanosomes' (Thomson, 1993). The suggestion that certain bacteria produce structured enzyme aggregates or xylanosomes (Thomson, 1993) is analogous to the formation of cellulosomes in some cellulose-degrading organisms. The first report of xylanosome from *B. fibrisolvens* (Lin & Thomson, 1991) appeared to consist of at least 11 xylanolytic active proteins ranging in size from 45 to 180 kDa. In a differential bacterial system, Shao *et al*. (1995) reported the localization of xylanase in the S-layer fraction of the thermophilic organism, *Thermoanaerobacterium sp*. strain JW/SL-YS 485. In *C. xylanolytica* a cell wall, associated xylan degrading system has been reported.

1.16 Purification of Xylanases

Xylanase purification schemes have generally utilized standard chromatographic techniques, mainly ion exchange and gel filtration, also using hydrophobic matrix. A xylanase from fungal maize pathogen *Cochiobolus carbonum* has been purified by hydrophobic interaction chromatography. Paul and Varma (1992) used Concavalin A-Sepharose chromatography to purify a xylanase which probably have a carbohydrate moiety. Immuno-affinity chromatography has been used to purify xylanase from *Trichoderma reesei.* Xylanases have also been purified using various techniques like isoelectric focussing, chromatofocussing, PAGE (Dey *et al*., 1992; Pereira *et al*., 2000) and FPLC (Simpson *et al*., 1991; Wong & Saddler, 1988). Purification of xylanases from crude culture filtrate using affinity precipitation with a commercial polymer Eudragit S-100 has also been reported (Gawande & Kamat, 1999). Some xylanases from an alkalophilic *Bacillus* sp. strain K-1 have been purified to homogeneity by affinity adsorption-desorption on insoluble xylan (Ratankhanokchai *et al*., 1999).

1.17 Properties of xylanases

Recent comprehensive reviews (Warren, 1996; Antranikian *et al*., 1997; Tony Collins et al., 2005) have described characterization of xylanases from microbial systems. The present section is limited to the properties of xylanases from extremophilic organisms. Studies of xylanases from alkalophilic and/or thermophilic organisms have led to the discovery of enzymes, which exhibit some unique properties. The molecular weights of the xylanases vary from 8-145 kDa (Sunna & Antranikian, 1997). The properties of some of the purified xylanases from extremophiles are described in Table 1.4 and 1.5. However, xylanases from *Thermoanaerobacterium sp*. (Shao *et al*., 1995) and *Thermotoga thermarum* (Sunna & Antranikian, 1996) showed higher molecular weights of 350 and 266 kDa, respectively. Horikoshi and Atsukawa (1973) were the first to report xylanase production from alkalophilic bacteria. The *Bacillus* sp. C-59-2 secreted two xylanases of molecular weight 43 and 17 kDa, respectively.

The purified xylanases exhibited a pH optimum of 6.0-8.0. Many of the xylanases produced by alkalophilic microorganisms such as *Bacillus* sp. (Okazaki *et al.,* 1984) and

Aeromonas sp. (Ohkushi *et al.,* 1985) showed remarkable stability at pH 9-10 but were not active above pH 8.0. An alkalophilic fungus having activity at pH 6.0-9.0 has been reported. Recently an alkali tolerant xylanase from *Aspergillus fischeri* was reported to exhibit remarkable stability at pH 9.0.

A small number of bacterial and fungal xylanases show maximal activities at temperatures 60-80 °C (McCarthy *et al*., 1985; Gruninger & Feichter, 1986; Khasin *et al*., 1993). The purified endoxylanases from various species belonging to the genus *Thermotoga* are optimally active at temperatures between 80 and 105 °C (Simpson *et al.,* 1991; Sunna *et al*., 1996; Winterhalter & Liebl, 1995). Xylanase from *Dictyoglomus* sp. exhibited a half-life of 80 min at 90 °C. (Mathrani & Ahring, 1991). *C. stercorarium* xylanase exhibited a temperature optimum of 70 °C and a half-life of 90 min at 80 °C. The thermophilic fungi include *Thermoascus aurantiacus* (Yu *et al*., 1987) which produces a thermostable xylanase reported to be stable at 70 °C for 24 h, *P. variota* (Krishnamurthy & Vithayathil, 1989) and *T. byssochlamydoides* with temperature optimum of 65-75 °C at pH 5-6.5.

1.18 Mode of action of xylanases

There are two types of xylanases: debranching (arabinose liberating) and non-debranching. The former catalyzes the removal of arabinose side-chain substituents in addition to cleaving main-chain linkages, while the latter cleaves only the main chain (Dekker & Richards, 1976).

Xylanases from *N. crassa* and *A. niger* are found to liberate arabinose from arabinoxylan (Mishra *et al.,* 1984; Takenishi & Tsujisaka, 1973). Majority of xylanases known are endotype enzymes which preferentially cleave internal glycosidic linkages in xylans and xylooligosaccharides and act by a random attack mechanism (Bérenger *et al*., 1985; Panbangred *et al*., 1983). However a few xylanases, such as those from *B. polymyxa* and *Chaetomium thermophile* are known to hydrolyze xylans to produce mainly xylobiose and traces of xylose and/or xylotriose by an exotype mechanism.

Xylanase V from *Aeromonas caviae* ME-1 is reported to be an unusual xylanase, which produced xylobiose as the only low molecular weight oligosaccharide from xylan by an exotype hydrolysis

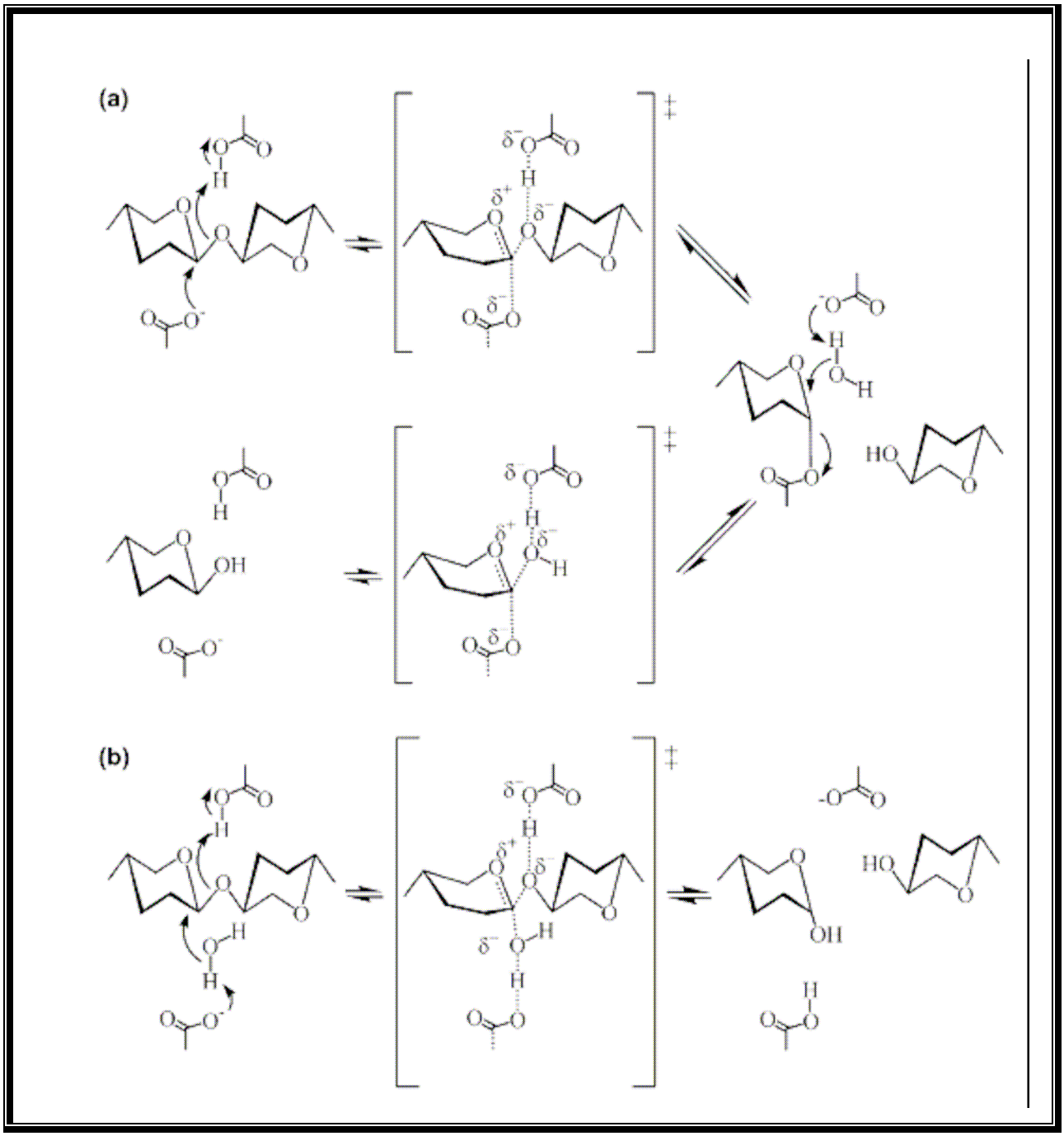


Fig. 1.11 General mechanisms for (a) retaining and (b) inverting glycosidases. Figure adapted from (Collins *et al.,* 2005)

1.19 Multiplicity of xylanases

It has become apparent that both fungal and bacterial cells produce a multiplicity of enzymes that belong to the same functional class and which sometimes exhibit broad plant polymer specificity (Gilkes *et al.,* 1991; Wong *et al*., 1988). These enzymes can be grouped into families based on conserved amino acid sequences in the catalytic domains and by hydrophobic cluster analysis. Thus, all high molecular weight xylanases belong to the family F/10, whereas low molecular weight xylanases belong to the family G/11. (Wong, *et al*., 1988; Gilkes *et al*., 1991; Henrissat, 1989). Enzyme multiplicity has been studied in the fungus, *T. reesei* and *T. harzianum* (Wong *et al.*, 1988). Results suggested that *T. reesei* produced four xylanases, each with different MW and p*I* values. Törönnen *et al*, (1992) cloned two *T. reesei* genes, *xyn1* and *xyn2* that appeared to encode separate products, XYL1 and XYL2 exhibiting similar molecular weights 19 and 21 kDa respectively but had p*I* values of 5.2 and 9.0. Similarly, *T. harzianum* produced three distinct xylanases. In analogous studies, reported that *Streptomyces lividans,* produces three xylanases encoded by three different genes, *xlnA* belongs to the family F, while *xlnB* and *xlnC* are members of the family 11.

Thomson (1993) suggested various mechanisms that could account for the multiplicity of function and specificity of the xylan degrading enzymes. Electrophoretically distinct xylanases could arise from posttranslational modification of a gene product such as differential glycosylation or proteolysis. The detection of minor xylanases may also be an artifact of the growth and /or purification conditions or these enzymes may have functions, which are not required in large amounts, e.g. hydrolysis of linkages not found frequently (Wong & Saddler, 1988). Multiple xylanases can also be produced from different alleles of the same gene (Wong *et al*., 1988) or may be a result of independent genes (Hazlewood & Gilbert, 1993).

1.20 Stereochemistry of product in xylanase action

Glycosyl hydrolases are classified as retaining or inverting enzymes depending on the stereochemistry of the released product. Retaining enzymes liberate products with the same anomeric configuration as the substrate by a double-displacement mechanism involving a glycosyl enzyme intermediate (Fig. 1.11a). Inverting enzymes form products with inverted configuration mediated by a single displacement mechanism (Fig. 1.11b).

The stereochemistry of the reaction products released as shown by NMR studies for both family 10 and 11 indicated that catalysis occurs through double displacement mechanism with the retention of the anomeric configuration in both the families (Biely *et al*., 1994). Several glycosyl hydrolases have been characterized, and their mechanism of hydrolysis is found to resemble that of lysozyme through an acid-base mechanism involving two residues (Sinnot *et al*., 1990). The first residue acts as a general catalyst and protonates the oxygen of the glycosidic bond. The second residue acts as a nucleophile, which in the case of retaining enzymes, interacts with the oxocarbonium intermediate or in case of inverting enzymes promotes the formation of an OH^- ion from a water molecule. In retaining glycosidases, distances between the nucleophile and the acid base catalyst are 5.4-5.5 Å (McCarter and Withers, 1994) whereas in inverting glycosidases, the corresponding distances are greater (9-9.5 Å) because for inversion to take place, water molecule has to be accommodated between the nucleophile and the enzyme.

1.21 Active site studies of xylanases

Useful information on the nature of groups at the active site can be obtained by investigating the effects of chemical modification on binding and /or catalysis. As substrates and competitive inhibitors bind to the active site, they frequently protect against modification/inactivation and can provide confirmatory evidence of involvement of particular residues in activity. The reports on the involvement of Trp residues in the reactions catalyzed by xylanases have all been based on inactivation by N-

bromosuccinimide (NBS). The participation of Trp in the active site of xylanases from various organisms has been reported (Kubackova *et al*., 1978; Keskar *et al*., 1989; Deshpande *et al*., 1990; Khasin *et al*., 1993). The role of Trp residues in substrate binding to catalytic domains of xylanase C from Fibrobacter succinogenes S85 has been shown (McAllister *et* al., 2000).

The only report on the identification of an essential Tyr residue in a xylanase from *Schizophyllum commune* xylanase. Chemical modification of xylanases from the same fungus indicated the involvement of carboxyl groups in the catalysis (Bray & Clarke, 1990). The involvement of Cys residue in the active site of a few bacterial xylanases has also been reported (Deshpande *et al.,* 1990; Keskar *et al*., 1989). Several specific reagents were used to identify reactive residues by competitive labelling in the presence or absence of substrate, in conjunction with kinetic analysis. Bray and Clarke (1994) using [^{14}C] EAC (1-(4-azonia-4,4-dimethyl-pentyl)-3- ethylcarbodiimide iodide) labelling of the enzyme followed by proteolytic cleavage, showed that the labeled reagent interacts with one Glu residue. The characterization and sequencing of the Cys containing active site peptide of the xylanase from *Streptomyces* T-7 (Keskar *et al*., 1989) and *Chainia* (Rao *et al*., 1996; Hegde *et al*., 1998) have been reported. The peptides showed the presence of a conserved Asp residue. The Cys residue in some xylanases may participate in covalent glycosyl-enzyme intermediate formation as has been proposed for the mutant T4 lysozyme in which Asp20 was replaced by a Cys residue (Hardy & Poteete, 1991).

1.22 Genetic engineering of xylanases

In recent years, several xylanases were subjected to site directed mutagenesis either in an attempt to improve their biochemical properties or to gain insight into structure-function relationships. Mutational studies were carried out on two *Bacillus* xylanases belonging to family 11. In xylanase from *Bacillus pumilus,* mutation of Glu93 and Glu182 resulted in decreased enzymatic activity, indicating that these residues were important for catalytic activity (Ko *et al*., 1992). Similarly, in *Bacillus circulans* xylanase Glu78 and Glu172 were identified as catalytic residues (Wakarchuk *et al*., 1994).

Mutational studies on xylanases belonging to family 10 included those of *T. saccharolyticum* at positions Asp 537, Asp602 and Glu600 (Lee *et al*., 1993), of *Streptomyces lividans* (XynA) at positions Glu128 and Glu236 (Moreau *et al.*, 1994) and of *C. fimi* (Cex) at position Glu127 (Macleod *et al*., 1994), all residues being replaced by their corresponding 'isoteric form'. Kinetic studies of all these mutants showed decreased activities towards the substrate consistent with the replacement of a catalytic residue. Roberge *et al*. (1997) showed that two His residues (His81 and His207) out of three were present in the active site of xylanase A from *S. lividans* and were found to be completely conserved in family 10 glycanases. The structural analysis revealed that they were involved in a network of hydrogen bonds which were responsible in maintaining the ionization states of the two catalytic residues (Glu128 and Glu236). Recently, it has been shown in xylanase from *B. circulans* that the substitution of an Asn residue with an Asp residue (N35D BCX) shifts its pH optimum from 5.7 to 4.6, with ~20% increase in activity.

Engineering of proteins by *in vitro* mutagenesis has become a process that allows almost any desired modification to be constructed in the laboratory. For example, introduction of several Cys residues into a xylanase from *Bacillus circulans* and the spontaneous formation of disulfide bridges resulted in increased thermostability of the enzyme by 15 °C (Wakarchuk *et al*., 1994). In the case of xylanase from *S. lividans* (Moreau *et al*., 1994) replacement of residue Arg156 with a Glu increased the optimum temperature of activity by 5 °C and half-life by 6 min. Similarly, site directed mutagenesis of a xylanase gene from *C. saccharolyticum* (Luthi *et al*., 1992) also yielded mutants with altered temperatures for stability and optimum, but with no change in optimum pH. Arase *et al.* (1993) reported stabilization of xylanases by random mutagenesis of the cloned gene fragment. However, reports on xylanase to date suggest that it has not been fully exploited to improve different properties of the enzyme.

1.23 Three-dimensional structures of xylanases

The three-dimensional structures of the enzymes at high resolution are indispensable for understanding the catalytic mechanism, the difference in substrate specificities among the enzymes belonging to the same family and for further improving their functions and stability through protein engineering. In this context, the number of xylanases whose three-dimensional structure solved is increasing rapidly. X-ray crystallographic studies on members of family 10 and family 11 have been reported. While xylanases of both families act upon the same substrate, there are significant differences in their structures. The three-dimensional structures of the family 10 catalytic domains are reported for *Streptomyces lividans* Xln A (Derewenda *et al*., 1994), *C. fimi* Cex (White *et al*., 1994), *Pseudomonas fluorescens* Xyn A (Harris *et al*., 1994), *C. thermocellum* Xyn Z (Domínguez, *et al*., 1995) and *T. auranticus* xylanase I. They all exhibit a tertiary fold of typical $(\alpha/\beta)_8$ barrel motif. Seen from the side, the molecule has a general 'salad bowl' shape. The active site is formed by an acidic cleft on the carboxyl-terminal side of the β-barrel and this cleft is well exposed to solvents (Ohmiya *et al*., 1997; Dupont *et al*., 1996). A long loop between β-strands 7 and 8 was observed for *P. fluorescens* xylanase (Harris *et al*., 1994) while the corresponding loops in *S. lividans* XlnA and *C. fimi* Cex were significantly shorter. This loop in Xyn A of *P. fluorescens* has been recently identified as the binding site for the calcium ion playing a role in stabilization the enzyme. It has been suggested that the occupation of Ca^{2+} binding loop with its ligand protects the enzyme from thermal inactivation, thermal unfolding and proteolysis. Site directed mutagenesis studies revealed that Asp residues at positions 256, 261 and 262 were pivotal for calcium binding (Spurway *et al*., 1997).

Among the family 10 xylanases, *T. auranticus* and *C. thermocellum* are from thermophilic organisms. A comparison with other mesophilic organisms from family 10 suggests that thermostability is affected mainly by improvement of the hydrophobic packing, favourable interactions of charged side-chains with the helix dipoles and introduction of Pro residues at the N-terminus of helices. *P. simplicissimum* xylanase is similar to other family 10 xylanase, but its active site cleft is much shallower and wider.

This probably accounts for the difference in catalysis and mode of action of the enzyme (Schmidt *et al*., 1998).

The three-dimensional structures of family 11 catalytic domains are reported for *Bacillus pumilus* XynA (Katsube *et al*., 1990), *B. circulans* xylanase (Wakarchuk *et al*., 1994) *Trichoderma harzianum* xylanase (Campbell *et al*., 1993), thermophilic *Bacillus* sp. xylanase (Pickersgill *et al*., 1993), *B. stearothermophilus* xylanase (Anna *et al*., 1997), *T. reesei* XynI (Törönnen & Rouvinen, 1994) and XynII (Törönnen *et al*., 1994), *Aspergillus kawachii* xylanase C (Fushinobu *et al*., 1998) and *Bacillus agaradhaerens* xylanase (Sabini et al., 1999 & 2001). Family 11 xylanases fold into a 'β sandwich' consisting of two-pleated antiparallel β-sheets, which are folded against each other, in a parallel manner, forming a cleft on one side of the protein structure. For *B. circulans* xylanase, an enzyme tetrasaccaharide complex was crystallized and the xylotetrose was found in the cleft, confirming the cleft's role as the active site of this xylanase (Wakarchuk, *et al*., 1994). The crystal structure of *B. pumilus* (Moriyama *et al*., 1987) was of ellipsoidal shape while two major xylanases from *T. reesei* (Törönnen *et al*., 1994) were reported to be monoclinic and those of *T. harzianum* were orthorhombic. In case of *B. pumilus* Xyn A (Katsube *et al*, 1990), although the structure was predominantly characterized as three large β- sheets, it was actually similar to the structures of *T. reesei* xylanases. Recently the crystal structure of xylanase (PVX) from a thermophilic fungus *Paecilomyces varioti Bainer* has been solved. This fungus has been attracting attention as a pathogen causing post-surgical infections (Kumar *et al*., 2000).

Table 1.4 Family 10 xylanases for which three-dimensional structures are available in Protein Data Bank

Xylanases	Organism	PDB code(s)	References
Xylanase (Xyn 10A)	*Cellulomonas fimi*	1EXP, 1FH7, 1FH8, 1FH9, 1FHD, 1J01, 2EXO,2HIS, 2XYL	White, *et al.,* (1994).
Xylanase A (Xyn 10A)	*Cellvibrio japonicus*	1CLX, 1E5N, 1XYS	Harris, *et al.,* (1994).
Xylanase F (Xyn 10C)	*Cellvibrio japonicus*	1US2, 1US3	Pell, *et al.,* (2004).
Xylanase C (Xyn 10B)	*Cellvibrio mixtus*	1UQY, 1UQZ, 1UR1, 1UR2	Pell, *et al.,* (2004).
Xylanase Z	*Clostridium thermocellum*	1XYZ	Dominguez, *et al.,* (1995).
Xylanase T-6	*Geobacillus stearothermophilus T-6*	1HIZ, 1R85, 1R86, 1R87	Coutinho, *et al.,* (1999), Mechaly, *et al.,* (2000)
Xylanase (Xyn A2)	*Geobacillus stearothermophilus T-6*	1N82	Teplitsky, *et al.,* (2000)
Xylanase A (Xyn A)	*Penicillium simplicissimumBT2246*	1B30, 1B31, 1B3V, 1B3W, 1B3X, 1B3Y, 1B3Z, 1BG4	Schmidt, *et al.,* (1998)
Xys 1	*Streptomyces halstedii*	1NQ6	Canals, *et al.,* (2003)
Xylanase A	*Streptomyces lividans*	1E0V, 1E0W, 1E0X, 1XAS, 1OD8	Derewenda, *et al.,* 1994
β-1,4-Xylanase	*Streptomyces olivaceoviridis E-86*	1ISV, 1ISW, 1ISX, 1ISY, 1ISZ, 1ITO, 1XYF	Fujimoto, *et al.,* (2000)
Xylanase	*Thermoascus aurantiacus*	1FXM, 1GOK, 1GOM, 1GOO, 1GOQ,	Lo Leggio, *et al.,* (1999) Natesh, *et al.,* 1999,
Xylanase B	*Thermotoga maritime*	1GOR, 1I1W,1I1X, 1K6A, 1TAX, 1TIX, 1TUX, 1VBR, 1VBU	Coutinho, *et al.,* (1999)

Table 1.5 Family 11 xylanases for which three-dimensional structures are available in Protein Data Bank

Xylanase	Organism	PDB code(s)	References
Xylanase C	*Aspergillus kawachii*	1BK1	Fushinobu, *et al.,* (1998)
Xylanase 1	*Aspergillus niger*	1UKR	Krengel, *et al.,* (1996)
Xylanase	*Bacillus agaradhaerens AC13*	1H4G, 1H4H, 1QH6, 1QH7	Sabini, *et al.,* (2001)
Xylanase A	*Bacillus circulans*	1BCX, 1BVV, 1C5H, 1C5I, 1HV0, 1HV11XNB, 1XNC, 2BVV,	Wakarchuk, *et al.,* (1994)
Xylanase	*Bacillus subtilis B230*	1IGO	Oakley, *et al.,* (2003)
Xylanase A	*Bacillus subtilis subsp. subtilis str. 168*	1AXK	Ay, *et al.,* (1998)
Xyn 11A	*Chaetomium thermophilum*	1H1A	Hakulinen, *et al.,* (2003)
Xylanase XynB	*Dictyoglomus thermophilum Rt46B.1*	1F5J	McCarthy, *et al.,* (2000)
Xyn 11A	*Nonomuraea flexuosa*	1M4W	Hakulinen, *et al.,* (2003)
Xylanase	*Paecilomyces varioti Bainier*	1PVX	Kumar, *et al.,* (2000)
Xylanase	*Streptomyces sp. S38*	1HIX	Wouters, *et al.,* (2001)
Xylanase	*Thermomyces lanuginosus*	1YNA	Gruber, (1998)
Xylanase	*Trichoderma harzianum E58 (Hypocrea lixii E58)*	1XND	Campbell, *et al.,* 1993, 125
Xylanase 1	*Trichoderma reesei (Hypocrea jecorina*	1XYN	Torronen, *et al.,* (1995)
Xylanase 2	*Trichoderma reesei (Hypocrea jecorina)*	1ENX, 1RED, 1REE, 1REF, 1XYO, 1XYP	Torronen, *et al.,* (1994)

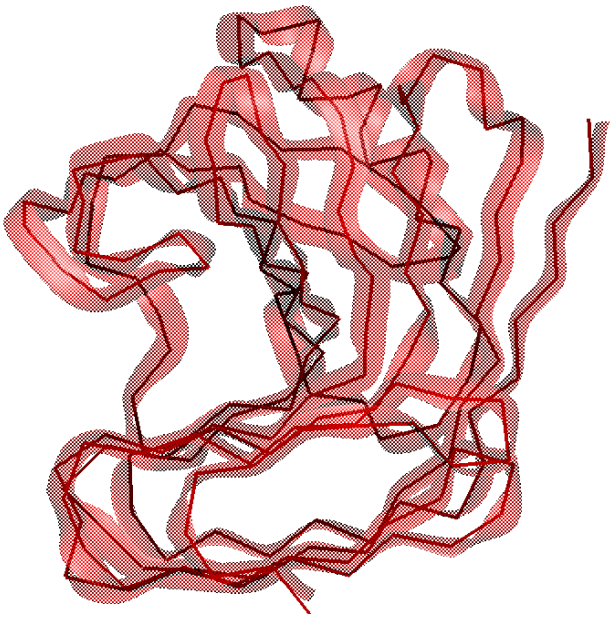


Fig. a.

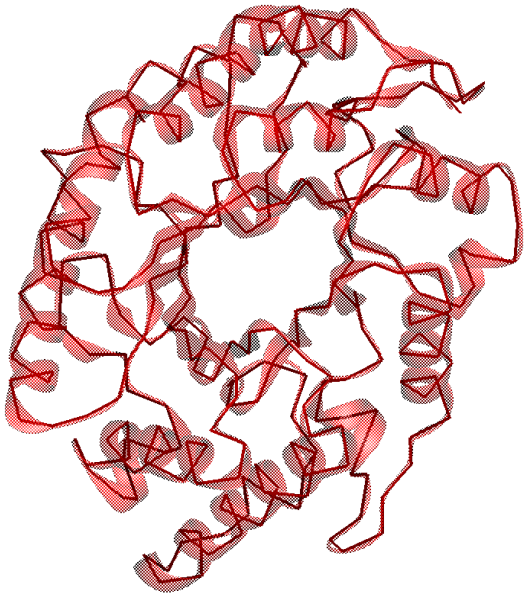


Fig. b.

Fig. 1.12 a) Crystal structure of family 11/G xylanase from *Dictyoglomus thermophilum*.

b) Crystal structure of family 10/F xylanase from *Thermoascus aurantiacus*.

1.24 Thermal denaturation/renaturation studies

Thermal denaturation studies of xylanase from *Streptomyces halstedii* by using differential scanning calorimetric (DSC) studies has been reported (Arribas *et al.,* 1994). *S. halstedii* produced two xylanases Xys1L (45 kDa) and Xys1S (35 kDa). Thermal denaturation of Xys1L revealed three thermodynamically independent domains, and that of Xys1S, which is a proteolytic fragment of Xys1L (without a C-terminal part), revealed two thermodynamically independent domains, each of which follows a two-state transition. The thermodynamic parameters of unfolding for each domain did not fit some of the correlations obtained for most compact globular proteins. In another study, catalytic activity measurements as a function of temperature, complemented with DSC data, were used to characterize the thermostability of the exoglucanase/xylanase Cex from C*ellulomonas fimi* (Nikolova *et al*., 1997). Xylanase (XynA) from the thermophilic bacterium *Thermotoga maritime* (Wassenberg *et al*., 1997) was studied to characterize the domain organization and stability of the recombinant enzyme and its isolated cellulose-binding domain (CBD). XynA and CBD were monomers with 116 and 22 kDa molecular masses. Denaturation/renaturation was used to gain insight into the folding mechanism of the complex multidomain protein. Guanidine hydrochloride (Gdn-HCl) induced unfolding of XynA leads to biphasic transitions. The shift of the transition of unfolding to higher guanidine hydrochloride concentration at acid pH was attributed to the CBD.

1.25 Biotechnological potential of xylanases

The potential biotechnological application of xylan and xylanases has been of tremendous importance to researchers. Commercial applications of xylanases involve conversion of xylan, which is present in wastes from agricultural and food industry, to xylose (Biely, 1985). Xylose and xylooligosaccharides have possible applications in the food industry as thickeners or as fat substitutes and as anti-freeze food additives. The hydrolytic products of xylan can be subsequently converted to liquid fuel, single cell proteins, solvents and artificial low calorie sweetners (Wong & Saddler, 1992).

Xylanases are of great importance in the paper and pulp industry as they replace toxic chemicals such as elemental chlorine and chlorine-dioxide for developing environment friendly processes. They play an important role in the debarking, deinking of recycled fibers and in the purification of cellulose for the preparation of dissolving pulps. Enzyme aided prebleaching was found to increase the brightness of the paper. Various organisms have been explored for producing xylanase enzymes for treatment of kraft pulps. Novo Nordisk marketed the first commercial xylanase preparation available for pulp bleaching under the name 'Pulpzyme HA' which was produced by a strain of *Trichoderma reesei*, subsequently Pulpzyme HB and HC from bacterial sources were also marketed (Pederson, 1989). Xylanases can be used in the production of dissolving pulp (purified cellulose) for making viscose rayons, cellulose esters and cellulose ethers and to remove undesired hemicellulose content (Viikari *et al*., 1993)

The use of xylanases has also been proposed in clarification of juices and wines (Biely, 1985), maceration of vegetable matters, liquefaction of coffee mucilage for making liquid coffee, recovery of oil from subterranean mines, extraction of flavors and pigments, plant oils and starch (McCleary, 1986) and to improve the efficiency of agricultural silage production (Wong and Saddler, 1992). The use of xylanases in bakery products has been suggested (Maat *et al*., 1992) and they were found to increase the specific volume, textural properties and shelf life of the bread. The application of xylanases in poultry diets has also been investigated (Bedford & Classen, 1992). Poultry are incapable of efficiently digesting cereals, the major ingredient in their diets. The addition of xylanases to the feed results in an improved nutritive value of wheat based diets for broilers. Xylanases with low pH optimum and broad pH stability would be most suitable for application in animal feed, where activity and stability at low pH is crucial. These biotechnological potentials of xylanases have prompted the search for suitable enzymes and technologies for large-scale economic production. The present investigation relates to the study of a cellulase-free xylanase (AT*B*XYL-C) from an alkalophilic thermophilic *Bacillus* sp. *(NCL 86-6-10).*

Chapter-2

MATERIALS AND METHODS

2.0 INTRODUCTION

In this chapter detail of the materials, methods and strategies used for purification, biochemical characterization, fluorescence studies, crystallization, X-ray data collection, data processing, structure solution and refinement of C-phycocyanins (C-PC) from cyanobacteria and xylanase from alkalophilic thermophilic bacteria (AT*B*XYL-C) are described. The programs used for analysis and comparison of protein structures have been detailed. Purification was carried out using different biochemical techniques such as chromatography to obtain pure samples of C-PCs and AT*B*XYL-C and the analysis of samples was using spectroscopic and other techniques such as electrophoresis. For crystallizing C-PCs and AT*B*XYL-C, mainly hanging-drop vapor-diffusion method was used as described in the following chapters. X-ray diffraction data were collected on Raxis IV^{++} image plate detector mounted on Rigaku rotating anode X-ray generator and processed using DENZO and SCALEPACK programs (Otwinowski, 1993; Otwinowski & Minor, 1997). Molecular Replacement (MR) method implemented in AMoRe program (Navaza, 1994; Navaza & Saludijan, 1997; Navaza, 2001) and PHASER were used to solve the structures of C-PCs and AT*B*XYL-C. The program REFMAC (Murshudov *et al*., 1997) was used for refinement and QUANTA (Accelrys) for graphics display and model fitting of the structures. Programs from CCP4 suite (Collaborative Computational Project, Number 4, 1994) were used for other calculations.

2.1 MATERIALS

Fermentation and purification of C-PCs and AT*B*XYL-C from the crude extract to the homogeneous protein solution involved chemicals such as Trizma, di-potassium hydrogen phosphate, potassium hydrogen phosphate, ethylenediaminetetraacetic acid (EDTA), ammonium sulfate (AS), sodium acetate, glycerol, β-mercapto-ethanol, bromo-phenol-blue (BPB), Acrylamide, N, N'-methylene bisacrylamide, sodium dodecyl sulfate (SDS), acetic acid, methanol, TEMED (N, N, N', N'- Tetramethylethylenediamine), Ammoniumpersulfate (APS), CM-sephadex, DEAE–sepharose CL-6B, molecular weight marker kits for SDS-PAGE etc, were purchased from Sigma Chemicals, St. Louis, USA.

The slab gel electrophoresis unit was purchased from Tarsons, India. The following chemicals were used extensively in the crystallization trials: sodium cacodylate buffer, 3-(N-morpholino) propanesulfonic acid MOPS buffer, N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid (HEPES) buffer, tri-sodium citrate, lithium sulphate, polyethylene glycol (PEG 20K, 8K, 4K, 1K), 2-methylpentane-2, 4-diol (MPD) and isopropanal. The additives used in crystallization trials to improve diffraction quality of crystals were ethelene glycol, glycerol, sucrose and PEG-400. Multi-well trays used in crystallization experiments were purchased from Becton Dickson and Company or Falcon. All chemicals used were of analytical grade. Silicon graphics Octane workstation was used for graphics display and computations.

2.2 Fermentation of Cyanobacteria

The marine cyanobacteria *Phormidium* and *Lyngbya spp.* were isolated from the rocky surface near the sea coast of Gujarat (latitude 21° 38' N and longitude 69°37'), in the west coast of India. These organisms were grown in batch cultures in the standard artificial seawater medium ASN III at pH 7.5 (Rippka *et al.,* 1979) and temperature 20 ± 2 °C with optimum light intensity of 60 lE m^{-2} s^{-1} provided by cool-white fluorescent tubes with a dark: light cycle of 12:12 h. The freshwater cyanobacterium *Spirulina sp.* was grown under same conditions except that the medium used was Zarouk medium (Chen *et at.,* 1996) at pH 10.

2.2.1 Extraction of phycobiliproteins: Cyanobacterial cells were harvested after 7-10 days (fresh water) and 15-20 days (marine) of incubation under laboratory-controlled conditions (temperature, pH, and light) by centrifugation at 8,000g for 30 min. The harvested cell mass was washed twice with distilled water and freeze-dried. One gram of freeze-dried cell mass was suspended in 100 ml of sodium- phosphate buffer (0.1 M, pH 7.0, containing 1 mM sodium azide). Further extraction procedure of phycobiliproteins consisted of repeatedly freezing the sample at -20°C and thawing to room temperature in the dark. The mixture was subsequently centrifuged at 10,000g for 30 min at 4 °C and phycobiliprotein containing clear supernatant was collected.

2.2.2 Estimation of phycobiliproteins: The absorbance of the supernatant was measured in a VARIAN CARY 50 BIO Scan UV–Visible NIR spectrophotometer at wavelengths 620, 652, and 562 nm. Concentration of C-PCs, A-PC and PE was calculated using the following equations (Bennett *et al.*, 1973):

$$\text{C-PCs(mg/ml)} = [A620 - 0.474\ (A652)]\ /5.34.$$

$$\text{A-PC (mg/ml)} = [A652 - 0.208\ (A620\}]\ /5.09.$$

$$\text{PE (mg/ml)} = [A562 - 2.41(PC) - 0.849\ (APC)]/9.62$$

2.2.3 Purification of C-PCs: The cyanobacteria *Phormidium, Lyngbya.* and *Spirulina spp.,* produce mainly C-PCs (with absorption maxima at 620 nm) and APC (with shoulder at 652 nm). The data of the purification of C-PCs from *Phormidium, Lyngbya and Spirulina spp.* are summarized in Table 3.3 (Chapter 3). All purification steps were performed in the dark at 15–20ºC. The clear supernatant of C-PCs was fractionated by precipitation with solid ASfirst at 25% and then at 50% saturation. The precipitate formed at 25% saturated AS was discarded and the supernatant was then brought to 50% AS saturation and allowed to stand for 4 h at 4°C. The precipitated protein containing mainly C-PCs was collected by centrifugation at 10,000g for 30min at 4°C, resuspended in acetate buffer (0.1 M, pH 4.5) and again centrifuged at 10,000g for 30 min at 4°C to precipitate out the basic linker polypetides. The precipitate was discarded and the supernatant was again brought to 50% AS saturation and allowed to stand for 4 h at 4°C prior to centrifugation at 10,000g for 30 min at 4°C. The precipitated C-PC was dissolved in 5 ml of sodium phosphate buffer (0.005 M, pH 7.0) and dialyzed against the same buffer overnight at 4 °C. This dialyzed solution was chromatographed on a DEAE– sepharose CL-6B column (1.5x15 cm.) equilibrated with the same buffer. The column was eluted with a linear increasing ionic concentration gradient using NaCl solution (0–0.25 M) at a flow rate of 0.5 ml/min. The C-PC eluted between NaCl concentrations of 0.10 and 0.20 M and was collected as 2 ml fractions. The absorption spectrum from 250 to 800 nm was recorded to monitor the purity of the fractions. Those fractions having ratio of A620/A280 > 4.0 were pooled together and brought to 50% saturation of AS. The

resultant precipitate of C-PC was dissolved in minimum volume of sodium-phosphate buffer (0.005 M, pH 7.0), dialyzed against water at 4°C, and freeze-dried into powder for storage.

2.2.4 Spectroscopic measurements: Absorption spectra of C-PC samples using UV-Visible radiation were recorded on a VARIAN CARY 50 BIO Scan UV–Vis, NIR spectrophotometer. The protein concentration for these experiments was in the range 0.1–0.3 mg/ml. In the region of visible radiation, C-PC and APC show maximum absorption at wavelengths of 620 and 652 nm, respectively. The purity was evaluated based on the two absorption ratios A620/A280 and A620/A652.

2.2.5 SDS–PAGE: Electrophoresis was carried out according to Laemmli (1990) using 15% gel containing 0.1% (w/v) SDS. Samples were pre-incubated with 2% (w/v) SDS, 10% (v/v) glycerol, 4.5% (v/v) β-mercaptoethanol, 0.025% (w/v) bromophenol blue and 60mM Tris buffer (pH 6.8), for about 4-10 min at 95ºC. Gels were run at room temperature and developed using Coomassie Blue R-250. Electrophore sample showed only two bands corresponding to α and β subunits of C-PCs (Fig. 3.1 Chapter 3), indicating that the linker polypeptdes were effectively removed during the purification procedure. Thus, the SDS-PAGE analysis clearly indicated that the C-PC preparations were pure and free of linker peptides. The molecular weights of subunits were determined by running the low molecular weight markers (Sigma SDS) along with the sample.

2.3 Purification of xylanase C from *Bacillus sp.* (NCL 86-6-10)

2.3.1 Enzyme production: The alkalophilic thermophilic *Bacillus* (NCL 87-6-10) was isolated from soil samples of Calicut in South India rich in decomposing coconut detritus. Fermentation of *Bacillus sp.* was carried out in wheat bran-yeast extract medium containing wheat bran (3% w/v), tryptone (1 to 3% w/v) and tween-80 (0.5 to 1%). The xylanase activity ranged from 130 to 170 IU/ml after 48 hours of growth (Balakrishnan. *et al.*, 1992, 2001).

2.3.2 Enzyme purification: The culture filtrate (500ml) was brought to 60% AS saturation by the addition of 215 g of the salt and the resssultant solution was stirred at 4 ºC for 3-4 h. The precipitate was collected by centrifugation at 10,000 x g for 30 min and dissolved in minimum volume of 20 mM potassium phosphate buffer pH 7.0. The extract was extensively dialyzed against the same buffer. The dialyzed enzyme solution was concentrated to ~5ml in Amicon ultra filtration unit using 10,000 cut off pore size membrane.The solution was loaded on to a CM-sephadex column (28x4 cm) pre-equilibrated with 20 mM potassium phosphate buffer pH 7.0. The bound enzyme was eluted with 0-0.8 M KCl gradient. The protein eluted at around 0.6 M KCl. The fractions showing AT*B*XYL-C activity were pooled and concentrated by Amicon ultra filtration. The purity of the enzyme was analysed by SDS-PAGE and the molecular weight was determined by mass spectrometry.

2.3.3 Enzyme Assay: AT*B*XYL-C activity was assayed by incubating 0.5 ml of enzyme with 0.5 ml xylan solution (substrate) and at 60 °C for 30 min. The reaction was stopped by adding 1 ml DNSA. The arrested solution was boiled for 5 min, diluted with 10 ml distilled water and the absorbance read at 540 nm. One enzyme unit is defined as the amount of enzyme required to liberate 1 µmol of xylose equivalent/ min under the assay conditions.

2.4 Chemical modification of active site residues

2.4.1 Modification of tryptophan: 1.0 ml of the enzyme solution (8-10 µM) was titrated with freshly prepared N-bromosuccinimide (NBS) (5 mM) at pH 6.0. The reaction was monitored using absorbance at 280 nm. The reagent was added till the absorbance at 280 nm started increasing. After each addition an aliquot 10 µl of the enzyme was removed and diluted to 200 µl with 50 mM tris-HCl buffer at pH 8.0 and then checked for residual enzyme activity under standard assay conditions. The number of Trp residues modified was quantified as described by Spande and Witkof (1967).

2.4.2 Modification of carboxyl groups: 1 ml of 4.0 µM AT*B*XYL-C solution was treated with Woodward's reagent K (15-30 mM) at pH 7.0. Aliquots of the enzyme were removed at specific time intervals and assayed for residual enzyme activity. Number of modified carboxylate residues was determined spectrophotometrically after removing the excess reagent by passing it through sephadex G-25 as described by Sinha and Brewer (1985).

2.4.3 Modification of arginine: 5 µM enzyme was treated with phenylglyoxal (15-60 mM) in 50 mM tris-HCl buffer pH 8.0. Aliquots of the modified enzyme were removed after 5, 10, 15 and 20 min and assayed for residual enzyme activity.

2.4.4 Substrate protection studies: The effect of the presence of substrate on inactivation during chemical modification reactions was studied. The enzyme was incubated in the presence of xylan and modifier together and then the residual activity was determined. The enzyme was incubated with same amount of xylan alone as control.

2.5 Fluorescence spectroscopic studies

Intrinsic fluorescence was measured using a Perkin-Elmer LS50B fluorescence spectrophotometer connected to a Julabo F25 water bath. Background emission due to buffer solution and denaturants was measured and subtracted from the protein spectra. The protein solution (4.1 µM) was excited at 280 nm and the emission recorded in the range of wavelengths 300-400 nm at 30 ºC. The slit-width on both the excitation and emission were set at 7 nm and the spectra were obtained at 100 nm/min.

2.5.1 Thermal denaturation: Effect of temperature on the fluorescence behavior of AT*B*XYL-C was studied using a thermostatic cuvette holder connected to an external constant temperature circulation water bath. The protein sample was incubated for 15 min at specified temperature before measuring the effect. The aliquots were removed at each specified condition for measuring the activity under standard assay conditions. For

renaturation experiments the samples were cooled back to 30 °C and left for one hour before recording the spectra. Fluorescence spectra were recorded as described above.

2.5.2 Light scattering: Fluorescence scattering at 400 nm was measured using a fluorimeter to detect and study protein aggregation under thermal denaturation conditions. The same experimental setup used for fluorescence studies was used for this purpose also. Both the excitation and the emission wavelengths were set at 400 nm. The excitation and emission slit widths were set at 5 and 2.5 nm, respectively. Scattering was observed for 120 sec.

2.5.3 Effect of pH: AT*B*XYL-C samples (4.1 µM) were incubated in an appropriate buffer for 4 h at 30 °C over a pH range of 1-12. The following buffers were used for these studies: 100 mM Glycine-HCl buffer for pH 1-3 and glycine-NaOH for pH 10-12. For other pH ranges, 100 mM acetate buffer (pH 4-5), 100 mM phosphate buffer (pH 6-7) and 100 mM Tris-HCl buffer (pH 8-9) were used. Fluorescence spectra were recorded as described above.

2.6.4 ANS-binding assay: The intermediate state of denatured and native AT*B*XYL-C at different denaturation conditions was analysed by hydrophobic dye (ANS) binding method. The final ANS concentration used was 40 µM, excitation wavelength (λ) 375 nm and total fluorescence emission was monitored between 400-550 nm. Reference spectrum with ANS at each buffer of respective pH and Gdn-HCl was subtracted from the spectrum of the sample.

2.6.5 Guanidium hydrochloide mediated unfolding: Protein samples (4.0 µM) were incubated with 1-6 M denaturant solution at pH 7 for 4 h to attain the equilibrium. After recording the scans, aliquots were removed from the sample to check for activity. Refolding experiments were conducted on ten fold diluted samples to reduce the Gdn-HCl concentration and incubated at 30 °C for 1h before recording the spectra.

2.6 Titration of xylo-oligosaccharides against xylanase

The fluorescence quenching measurements were made by titrating 2 ml of the enzyme (4.0 µM) solution with the saccharide solution (100 mM, 5 µl each time) followed by monitoring the change in fluorescence at 338 nm. The fluorescence of the buffer and saccharide solution were measured at appropriate wavelengths and these values were used to correct the observed fluorescence. Corrections were also made to compensate for the dilution upon addition of sugar and at the highest concentration of the saccharide, the volume change was less than 0.1% of the enzyme solution.The temperature of the enzyme and saccharide solutions was maintained at 30 ºC. The relative fluorescence intensity of the AT*B*XYL-C saturated with saccharide (F_∞) was obtained from the experimental data by plotting Fo/(Fo-F) versus 1/[S] and extrapolating to y-axis, where Fo is the fluorescence intensity of the enzyme alone and F is the fluorescence intensity of the enzyme at a saccharide concentration [S]. log {(Fo-F)/(F-F_∞)} was plotted against log[S] and the association constant (Ka) was determined from the plot by assuming the relation that the pKa of the complex equals the value of [S] when log {(Fo-F)/(F-F_∞)} = 0.

2.7 Circular dichroism (CD) spectroscopy

CD spectra were recorded on a J-715 spectro-polarimeter with a PTC343 Peltier unit (Jasco, Tokyo, Japan) at 25 ºC in quartz cuvette. All spectra were corrected for buffer contributions and converted to mean residue weight ellipticity. Each CD spectrum was accumulated from six scans at 100 nm/min with a 1 nm slit width and a time constant of 1s for a nominal resolution of 1 nm. Far UV CD spectra were collected in the range of wavelengths 200 - 250 nm using a cell of path length 0.1 cm for monitoring secondary structure. The tertiary structure of the enzyme was monitored with near UV CD spectra in the wavelength 250 - 300 nm using path length 1 cm. Protein concentrations were 5 µM

40 µM for far-UV and near-UV CD spectra collection. The protein was equilibrated in buffers of pH 1.0 and 7.0 six hours before data collection. The CD spectrum of the native AT*B*XYL-C was analyzed using the algorithm in CDPro program in order to determine the percentage of secondary structures.

2.8 Amino acid composition of xylanase

Amino acid composition of AT*B*XYL-C was analysed with AccQ-Fluor (Waters Corporation). Salt-free lyophilized AT*B*XYL-C (50µg) was hydrolyzed by using 6N HCl, in vacuum-sealed tubes at 110 °C for 24 h. Excess acid was removed by evaporation in vacuum at room temperature. The sample was then derivatized with 6-aminoquinolyl-N-hydroxysuccinimiyl carbamate (AQC) in borate buffer at pH 9. The hydrolysate (10 picomoles) was loaded on AccQ-Tag column equipped with a fluorescence detector.

2.9 CRYSTALLOGRAPHIC ANALYSIS

2.9.1 Crystallization: The first step considered crucial in protein structure determination is the growth of diffraction quality single crystals. In the absence of any single concrete theory behind the mechanism of crystallization, we have treated the protein crystallization as a trial and error procedure invoking experience and crystallization reports as guiding principles. It is accepted that the presence of impurities, ionic strength, pH, temperature, precipitating agent and several unspecified factors play role in crystallization process.

Crystallization is known to lower the free energy of proteins by ~3-6 kcal/mole relative to the solution state (Drenth & Haas, 1998). The general processes by which substances crystallize are similar for molecules of both microscopic (salts and small organics) and macroscopic (proteins, DNA, RNA) dimensions. Crystallization is one of several means by which a metastable supersaturated solution can reach a stable lower energy state by a reduction of solute concentration. There are three stages of crystallization common to all systems: nucleation, growth, and cessation of growth.

Nucleation is the process by which molecules or noncrystalline aggregates (dimers, trimers, etc.), which are free in solution, come together in such a way as to produce a thermodynamically stable aggregate with a repeating lattice. The formation of crystalline aggregates from supersaturated solutions does not however necessitate the formation of macroscopic crystals. Instead, the aggregate must first exceed a specific size (the critical size) defined by the competition of the ratio of the surface area of the aggregate to its volume (Feher & Kam, 1985; Boistelle & Astier, 1988). Once the critical size is exceeded, the aggregate becomes a supercritical nucleus capable of further growth. If the nucleus decreases in size so that it is smaller than the critical size, spontaneous dissolution will occur. The process of formation of nonspecific aggregates and noncrystalline precipitation from a supersaturated solution does not involve the competition between surface area and volume (*n*-mers add to the aggregate chain in a head-to-tail fashion forming a linear arrangement), and thus generally occurs on a much faster time scale than crystallization. The degree to which nucleation occurs is determined by the degree of supersaturation of the solutes in the solution. The extent of supersaturation is in turn related to the overall solubility of the potentially crystallizing molecule. Higher solubility allows for a greater number of diffusional collisions. Thus, higher degrees of supersaturation produce more stable aggregates (due to higher probability of collision of diffusing molecules) and therefore increase the likelihood of the formation of stable nuclei. In the case of a finite number of solute molecules, this condition generally results in the production of a large number of small crystals. At lower solute concentrations the formation of individual stable nuclei increases in rarity, thus favoring the formation of single crystals. Crystal growth generally starts at solute concentrations sufficient for nucleation to occur, and continues at concentrations beneath the nucleation threshold. The rate of growth is determined by a combination of the nature of the growing crystal surface and the diffusional rate.

As noted earlier, molecules crystallize from metastable supersaturated solutions as a means of lowering the overall solution free energy. Chemical precipitants are by and far the most widely used method of achieving supersaturation of macromolecules in order to induce crystallization. In general, the main influence of these compounds is on the

solvent (e.g. bulk water) rather than on the solute (the protein), with the notable exception of dye precipitants. For crystallization of proteins, the major classes of precipitants may be divided into six categories: salts, high molecular weight straight chain polymers (e.g. PEG), MPD, organic solvents, sulfonic dyes, and deionized water (Arakawa & Timasheff, 1985). Although the following discussions of the individual traits of these six categories use proteins as examples, most of these precipitants are applicable to other macromolecules including RNA, DNA, and polysaccharides.

Among various crystallization techniques known, hanging-drop vapor-diffusion method is widely adopted and has produced more crystallized proteins than all other methods combined. This method is simple, consumes less protein and it is easy to monitor the progress of crystallization. In a typical experimental set up using 24-well multiwell trays, 1-2 µl of protein solution was placed on a siliconized cover slip, mixed with 1-2 µl of the precipitant solution and allowed to slowly equilibrate against 500-1000 µl of reservoir solution of the precipitant. The concentrated solutions of pure C-PCs and AT*B*XYL-C were prepared in milli-Q water or buffers before setting up crystallization. The concentrations of proteins were estimated as described by Lowry *et al*., (1951). During initial crystallization trials it was found that in some cases the presence of certain salts (sodium formate in case monoclinic crystals of C-PCs and tri-sodium citrate in case of xylanase) was essential for growing crystals. Different precipitants such as ammonium sulfate, PEGs of different molecular weights, 2-methylpentane-2, 4-diol (MPD) and isopropanol were tried in the crystallization experiments and some were found very effective.

2. 9.2 Crystallizing proteins with PEG

Polymers: The use of high molecular weight linear polymers as precipitating agents was pioneered by Polson and coworkers who tried a variety of polymers including polyethylene glycol, dextran, polyvinyl alcohol, and polyvinyl pyrrolidone (Polson *et al*., 1964). Of these, polyethylene glycol (PEG) was found to be the most effective both by way of precipitating ability and cost effectiveness. PEGs are produced in a variety of

molecular weights, ranging from 200 (~3 monomers) to in excess of 1 million (~15000 monomers), and as mono- and di- methyl ethers. Like salts, PEGs compete with protein solutes for water and exert excluded volume effects (which vary according to the length of the polymer). However, unlike salts, PEGs decrease the effective dielectric constant of the solution, which increases the effective distance over which protein electrostatic effects occur. Solutions of polyethylene glycols have mean electron densities roughly equivalent to water and do not generally interact in a deleterious manner with heavy atom compounds, thus making them particularly well suited for macromolecular crystallization. PEGs with molecular weights less than 1000 are typically liquids and are generally used at concentrations above 40% v/v. PEGs with molecular weights above 1000 are generally solids and are used in the 5-50% w/v concentration range. All PEG solutions should be made with the inclusion of ~0.1% Na azide (toxic substance) to prevent bacterial growth. Also, buffering of high concentration (40%) of PEG solution with Na citrate at concentrations above 100 mM tends to cause the formation of phase transitions and color changes in the PEG solution, which suggests some form of reaction (probably cross linking by the citrate), and thus should be avoided.

2.9.3 Crystallization trials using (2-methyl-2, 4 pentane diol) (MPD) and Organic solvents (Isopropanol)

MPD is a small polyalcohol (2-methyl-2,4 pentane diol) which has properties midway between those of low molecular weight PEGs and organic solvents. MPD functions as a precipitant by a combination of activities, including competition for water, hydrophobic exclusion of protein solutes, lowering the dielectric constant of solution, and detergent-like effects. It is generally used in concentrations in excess of 40% v/v with water/buffer, and tends to cause phase transitions in the form of coacervate droplets, which are enriched in protein concentration (synonymous with those in Ray and Bracker, 1986).

Historically, organic solvents have typically been used as precipitants for protein crystallization as much by chance as by design. Crystallization by exposure to organic solvents is occasionally seen during protein purification, typically in the presence of

common solvents such as ethanol, methanol, acetone, isopropanol, DMSO, or tert-butanol (McPherson, 1990). Due to their hydrophobic nature, organic solvents cause phase transitions similar to those formed in the presence of MPD and lower the bulk dielectric of the solvent.

2.9.4 X-ray diffraction data collection

After growing protein crystals of suitable size and quality the immediate step in an X-ray diffraction experiment is to characterize the crystals and measure intensities of Bragg reflections from the crystals. Data collection is best performed as a highly interactive process. Protein crystals diffract X-rays with lesser intensity and lower resolution than the small molecule crystals. Thus, the protein data collection requires a high intensity X-ray source and a high sensitivity X-ray detector. In the last decade the combination of powerful tools such as synchrotron sources, image plate and CCD detectors and various softwares developed for collecting and processing diffraction data, increased computational and graphics power of workstations, all these have transformed protein crystallography into a powerful tool for structural study in the era of structural genomics (Beauchamp & Isaacs, 1999; Blundell *et al*., 2002). The X-ray storage-phosphor image plate (IP) is considered to be the most suitable detector for acquiring protein data with a home source, as well at synchrotron radiation facility. IP is a very sensitive detector, has wider dynamic range, and has high spatial resolution and high count rate capacity, which are the fundamental requirements of an X-ray area detector (Amemiya, 1997).

The IP detectors have enabled protein crystallographers to obtain very accurate data sets with laboratory X-ray sources. The plate can be erased by exposure to intense white radiation and can be used repeatedly. The size of commonly used IP is 300 mm and can be controlled by software running on a computer through an interface. Each diffraction image has a size of 18 MB, which can be stored in the computer. The crystal alignment was done through a CCD camera - TV monitor assembly. The processes of exposure, data collection, readout and storage of data are carried out automatically. All

the diffraction data of crystals described in the thesis have been collected at NCL, Pune, India. High intensity X-ray radiation usually damages protein crystals during longer exposure time for data collection. Now it is routine to collect the macromolecular data at cryogenic temperatures. This technique of flash cooling protein crystals at liquid nitrogen temperatures and collecting data offers several benefits. Some of the advantages include reducing radiation damage of the crystal on exposure to X-rays, improving the limits of resolution, decreasing thermal vibrations of the atoms, allowing storage and reuse of crystals, and helping to overcome the scaling problem by enabling collection and scaling of entire data from one crystal only (Garman & Schneider, 1997). In all the experiments described here, crystals were frozen under the cryostream of liquid nitrogen (~120 K) and different cryoprotectants like glycerol, PEG 400, MPD, 1, 2, 6-Hexanetri-ol (HXT) in the range 20 to 30% were used. Incase of C-phycocyanin PEG 400 has worked as cryoprotectant for crystals of this protein from different species in all its crystal forms (monoclinic and hexagonal).

During data acquisition the crystals were oscillated about an axis perpendicular to the X-ray beam, with a chosen, relatively small angle of oscillation, usually 0.5 to 1° per frame. Crystal to detector distance has been chosen based on the longest unit cell dimension, mosaic spread, etc., so that the diffraction spots are well resolved. The exposure time depends on the quality of crystal and oscillation range, larger the oscillation range, longer the exposure time required. However, the situation where the intensities crossing the limit of the dynamic range of image plate has been avoided. In our data collection, exposure times less than 600s per frame only was used. The range of rotation angle for data collection in order to acquire complete data set was chosen according to the symmetry of the crystal.

2.9.5 Data Processing: The analysis and reduction of single crystal diffraction data on an imaging plate detector consists of seven major steps:

l. Visualization and preliminary analysis of the original, unprocessed detector data.
2. Indexing of the diffraction pattern

3. Refinement of the crystal and detector parameters,
4. Integration of the diffraction maxima.
5. Finding the relative scale factors between frames.
6. Precise refinement of the crystal parameters using the whole data set.
7. Merging and carrying out statistical analysis of the symmetry related measurements depending upon the space group.

The first four steps of data processing were carried out by DENZO and XDISPLAYF and the steps five through seven were carried out using the program SCALEPACK (Otwinowski,1993). The auto indexing routine deduces crystal unit cell parameters and crystal orientation parameters from a single oscillation image. The parameters that specify the orientation of the crystal relative to the X-ray beam are the vertical axis, the spindle axis and the crystal rotz, roty, and rotx values. Distortion index and unit cell parameters consistent with all possible Bravais lattices are listed by the program. The lattice with the highest symmetry that fits the data with minimal distortion is chosen. The auto indexing program also gives the crystal orientation parameters. After a successful autoindexing step, the following parameters are refined in the given order:

l. CRYSTAL rotx, roty, rotz.
2. Y BEAM, X BEAM
3. UNIT CELL
4. CROSSFIRE x, y, xy
5. CASSETTE rotx, roty
6. RADIAL OFFSET and ANGULAR OFFSET
7. DISTANCE

Some of these parameters are highly correlated. In particular, crystal to film DISTANCE and RADIAL OFFSET are correlated with the unit cell parameters. So, they can be simultaneously refined only in cases where the quality of data is high and data at high-resolution are available. The input values for the program include the value for mosaicity,

spot shape, background shape, data collection parameters like the crystal to detector distance, the kind of detector used during data collection, oscillation etc.

DENZO allows for interactive visualization and modification of parameters such as the shape, size and profile-fitting radius of the spots. The user has the choice of visually selecting reflections for the auto indexing routine. After each cycle of refinement of the parameters, DENZO updates the display and prints out the numerical summary of the refinement cycle. The output lists the new values for the refined parameters and the shift in their values during the refinement cycle. The output also gives the χ^2 values for the X and Y positions of the predicted spots. The χ^2 values represent the average ratio of the squared error in fitting and the expected error. A good refinement will have χ^2 values near 1.0. The magnitude of the values is not a very critical indicator, as these represent only the comparison of the spatial differences between the observed and the predicted reflections to an error model that might be overly biased. χ^2 values of even 2.0 or 3.0 are accepted, because the position of the predicted reflection, and hence the intensity, is still very accurate. Very large values for the χ^2 parameters indicate that something is seriously wrong with indexing, refinement or the detector. DENZO finally writes out a list of hkl's and unscaled intensities for each image (Otwinowski & Minor, 1997).

The program SCALEPACK is used to scale the raw intensities output by DENZO. This program calculates single isotropic scale and B factors for each of the “films" or "batches" of the processed data that are input. The output is a scaled, merged data. The multiplicative correction factor applied to intensities (I) and σ(I) is given as

$$S = \frac{e^{2B\,(\mathrm{Sin}\theta/\lambda)^2}}{\mathrm{scale}}$$

Merging of the symmetry related measurements is done as follows:

The multiplicative factor that is applied to each measurement is calculated from the scale and B factor of the corresponding frame and applied as follows:

$$W = \frac{1}{(\sigma . S . E_1)^2 + (\{Icorr\} . E_2)^2}$$

Where E1 and E2 are the input variables 'error scale factors" and 'estimated error' respectively and

$$\{Icorr\} = \frac{\Sigma \, \mathbf{Icorr} \, . W}{\Sigma W}$$

The output $\sigma(I)$ is $\frac{1}{\mathbf{square\ root}\ \Sigma W}$

During the refinement, the scale (and B if requested) for all frames are shifted simultaneously to minimize the difference between the <Icorr>'s and the Icorr is for individual measurements, summed over all reflections on all images. <Icorr> are re-determined in each cycle as described earlier. One or more 'films' or 'batches' that is

input is designated as the reference, and its scale and B factors for the reference frame are not refined.

The assessment of the high-resolution limit of the diffraction pattern is done in two ways: The first is to calculate the mean ratio of the intensity to the error; $I/\sigma(I)$. The second is the agreement between symmetry related reflections, i.e. R_{merge}. The first parameter is an indication of the quality of data. Whereas $R_{merge,}$ although describes quality and reliability of data, depends on overcoming conditions of low redundancy of data, omission of weak or partial reflections, use of sigma cut-offs in the data set all of which can lead to artificially low R_{merge}.

The quality of the scaled and merged output data is also assessed using intensity statistics.

The normalized χ^2 given by the formula:

$$\chi^2 = \frac{\Sigma(I - \{I\})^2}{\frac{\text{Error}^2 * N}{N - 1}}$$

χ^2 is a good indicator of the data quality.

The R_{linear} and R_{square} are defined as

$$R_{linear} = \frac{\Sigma(I - \{I\})}{\Sigma I}$$

$$R_{square} = \frac{\Sigma(I - \{I\})^2}{\Sigma I^2}$$

A good data set should show χ^2 values close to 1.0. The total error and the statistical error should also match closely. The R_{linear} and R_{square} values should also be close to each other. If R_{square} is larger than R_{linear}, then the distribution of the deviation ($I-I_{ave}$) is skewed. This means that there are too many outliers in the data. The data collection statistics for the data set are given in a tabular form later in this chapter where the details are presented.

Mathew's number: Once the space group symmetry and unit cell dimensions of the crystal are known, it is possible to estimate the number of molecules in the crystallographic asymmetric unit and the solvent content of the protein crystals with the knowledge of the molecular weight of protein. The following equations are used (Matthews, 1968).

$$V_m = (\text{Unit cell volume}^* z) / (MW^* n)$$
$$V_{solv} = 1 - (1.23/ V_m)$$

Where V_m is the Mathew's number, n is the number of molecules in the unit and z is the Avogadro's number; (V_{solv}) is the solvent content of protein crystals. The Matthew's

number and the solvent content were calculated for all the crystal forms of C-PCs and xylanase. The solvent content of the crystal was calculated by assuming two hexamers (monoclinic system) and monomer (hexagonal system) molecule in case of C-PCs and dimer in both case of orthorhombic and monoclinic forms of the AT*B*XYL-C in the asymmetric unit.

2.9.6 Sequence alignment: Sequence alignments of C-PCs from *Phormidium, Lyngbya* and *Spirulina spp.* and AT*B*XYL-C from alkalophilic *Bacillus sp.* were carried using available amino acid sequence or from crystal structure determined. The pair wise alignment was carried out using the program CLAWSTAL W available at European Bioinformatics Institute (http://www.ebi.ac.uk) and **EsPript** (Gasteiger *et al.,* 2003.).

2.9.7 Molecular replacement: Crystallographic technique of elucidating the molecular image as in the crystal unit cell using X-ray diffraction data involves the computation of electron density ρ(x, y, z) through out the volume V of the unit cell and is,

$$\rho(x, y, z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} F_{hkl} e^{-2\pi i(hx+ky+lz)}$$

where F_{hkl} is the structure factor defined by the equation

$$F_{hkl} = |F_{hkl}| \, e^{i\alpha_{hkl}}$$

The amplitude $|F_{hkl}|$ which is proportional to the square-root of intensity I_{hkl} can be obtained from intensity of diffraction spot. However, any attempt to determine the associated phases, α_{hkl}, from experimental intensities is not straightforward. Direct methods derive the phases from amplitudes through statistical relationships satisfied by

structure invariants. It has been the traditional method for solving small molecule structures and recently being applied to determine the structures of high-resolution (~1.0 Å or better) macromolecules containing about 2000 atoms (Burla et a., 2000; Weeks& Miller, 1999). In general, macromolecular structures are solved by one or more of the three popular methods *viz.,* Multiple Isomorphism Replacement (MIR), Multiwavelength Anomalous Diffraction (MAD) and Molecular Replacement (MR). Other hybrid methods such as Single Isomorphous Replacement with Anomalous Scattering (SIRAS), Multiple isomorphous replacement with anomalous scattering (MIRAS) and Single Wavelength Anomalous Diffraction (SAD) involve the combination of MIR and MAD. MIR, first introduced in 1954 (Green *et al*.,), has been used widely for macromolecular structure determination. Structure solution by anomalous dispersion, first introduced in 1956 (Ramachandran & Raman) has caught the imagination of crystallographers more recently (Hendrickson *et al*., 1985; Hendrickson, 1991) with the development of tunable synchrotron radiation sources and recombinant incorporation of selenomethynyl derivative into proteins (Hendrickson *et al*., 1990). The recent developments in preparation of halide derivatives by a quick soak and using them for phasing of macromolecule is gaining importance in the light of high-through put structure determination efforts (Dauter et a1., 2000; Dauter & Dauter, 2001). MAD method exploits the inequalities of symmetry related reflections (Friedel pairs) due to a particular scattering species to absorb X-rays of specified wavelength while MIR operates on the differences in the intensities duly generated from isomorphous derivative crystals incorporating heavy atom scatterers. The difficulty in preparing heavy atom derivatives is overcome by introducing a new method called Molecular Replacement (MR) method (Rossmann, 1972). This method requires the availability of a search molecule whose structure is similar to the structure of the target molecule. The MR method has become very popular due to availability of many accurately determined protein structures in PDB.

In this chapter, the structure determination of hexameric C–PC from *Phormidium, Lyngbya and Spirulina spp.* and a xylanase from alkalophilic thermophilic *Bacillus sp.* using the molecular replacement method, its refinement and structure validation are

discussed. A brief description of the MR method follows as it was used for the structure determination presented in this thesis.

When the structure of a similar or homologous molecule is known, MR is the simplest technique to determine the target structure. In essence, it involves generating a preliminary model of the target crystal structure by orienting and positioning the search molecule within the unit cell of the target crystal so as to account for the diffraction pattern. The problem in MR is to find in general the six parameters, three rotational and three translational. Which would describe how the search molecule is placed in the unit cell? Rossmann & Blow (1962) had shown that this six-dimensional search could be reduced to a sequence of two three-dimensional searches in which first the orientation and then the position of the search molecule is determined. The rotation function formulation developed by Rossmann & Blow (1961) involves looking for agreement between the Patterson functions of the search and target structures as a function of their relative orientation. If the two Patterson's are calculated for the same structure the search is a self-rotation function and conveys information on the rotational transformations relating to the subunits of oligomeric proteins. If the Patterson function refers to different crystals, it is a cross-rotation function. The cross-rotation function provides information on the orientation of the known molecule in the unknown cell. First a rotational search is carried out, looking for agreement between positive peaks of the first Patterson representing mainly intra-atomic vectors of a protein subunit, with those of a rotated version of the second Patterson. A function to evaluate this agreement index is defined as

$$R = \int P2(x_2)\, P_1(x_1)\, dx_1$$

where Patterson function P_2 is rotated w. r. t. P_1. Any point x_1 in P_1 is related to any other point x_2 in P_2 by a rotation matrix C as $x_2 = [C]x_1$. The integral is over a spherical volume U centered at the Patterson origin. A maxima in the rotation function indicates a potential orientation of the search molecule in the target crystal.

The two Patterson functions expressed in reciprocal space are:

$$P_1(x_1) = l/V_1 \, \Sigma_h \, l \, F_h \, l^2 \cos (2\pi h.x_1)$$
$$P_2(x_2) = l/V_2 \, \Sigma_p \, l \, F_p \, l^2 \cos (2\pi h.x_2)$$

where h and p are reciprocal lattice vectors and V_1 and V_2 are the volumes of the unit cells of crystals 1 and 2, respectively. The rotation function is derived by expanding the Patterson function as a Fourier series (Rossmann & Blow, 1962). The resulting expression is given as

$$R = U/V^3 \, \Sigma_h \Sigma_p \, l \, F_p \, l^2 \, l \, F_h \, l^2 \, Gh.h'$$

where V is the volume of the crystal unit cell, F_h, F_p are structure factor amplitudes corresponding to Patterson's P_1 and P_2, respectively, and Gh,h' is a spherical interference function whose magnitude depends on the reciprocal lattice vectors h and h' as well as the volume U within which the integral is evaluated. The non-integral reciprocal lattice vector h' is given by

$$h' = p \, [C \,]$$

$$\mathbf{Gh.h' = \frac{3\,[\sin(2\pi H.r) - (2\pi H.r)\cos(2\pi H.r)]}{(2\pi H.r)^3}}$$

where r is the radius of the spherical volume used for integration and H = h + h'. The value of G will be significant only for those integral reciprocal lattice points h which are close to -h'. The number of points, which need to be included in the summation, depends on r and the reciprocal cell dimensions. Mostly, 27 points centered around -h' have been shown to be sufficient (Tollin & Rossmann, 1966).

The matrix algebra to relate these coordinate systems has been detailed by Rossmann& Blow (1962). Rotation searches are usually performed using Eulerian angles (α. β, γ) or spherical polar angles (Φ, ω, κ).Crowther (1972) developed a procedure for computing rotation functions, which is much faster than the method of Rossmann & Blow (1962). His method is based on the expansion of the Patterson function as sum of spherical harmonics. If a molecule possesses more than one non-crystallographic symmetry axis, then searching simultaneously for all the axes enhances the power of the rotation function peaks (Rossmann, 1972; Tong & Rossmann, 1990). Such a search is called locked rotation search.

Once the orientation is determined, next step is to find the translation vector (defined by translation function *T*) that defines the location of the molecule within the unit cell. The translation function (*T*) was first proposed by Crowther & Blow (1967). It is derived from the general form of the translation function given by the equation

$$T(t) = \int_{V} Po\{u \cdot t) \cdot P(u)\, du$$

Where t is the intermolecular vector, u is vector in Patterson space, V is primitive unit cell volume, P(u) represents the observed Patterson function of the crystal and P_o(u, t) represents the cross Patterson vectors (calculated intermolecular Patterson vectors). When t becomes t_o, the computed cross Patterson vectors fit correctly to the observed Patterson function P, and T(t_o) will have a large positive value. There are several forms of translation function and various available packages use one or more of these formalisms. The success of translation search depends on the accurate orientation of the model during rotation, completeness and accuracy of the model, and the resolution range and selection of diffraction data.

The success of the MR method in general relies on (Navaza & Saludjian, 1997)

1. The completeness and quality of the data.
2. The extent of structural similarity between the molecular models and the actual molecules that constitute the crystal.
3. The size of each molecular model with respect to the content of the crystal cell.
4. The excellence of the criterion used as an indicator of the quality of agreement.

Several MR packages are available. AMoRe, X-PLOR/CNS, Merlot, PATSEE, GLRF are some of the packages. The principal difference between these packages is the way in which the calculations can be carried out, whether in real space or in reciprocal space (Turkenburg & Dodson, 1996). AMoRe that was designed with motto to deal with difficult problems has many novel functions incorporated in it. We have used AMoRe implemented in CCP4 suite (Collaborative Computational Project, Number 4., 1994) in the present study.

2.9.8 Structure refinement: Crystallographic refinement aims at optimizing the agreement between atomic model and observed diffraction data by making use of chemical restraints in the case of low resolution data. The structures reported in this thesis were refined using the program in CCP4. Molecular dynamics methods are exploited by CCP4i to probe conformational space of the molecule while minimizing the difference between the observed and calculated structure factors (Brunger *et al*, 1987, 1990, 1997). There are options for rigid body refinement, positional refinement, restrained and unrestrained individual B-factor refinement, group B-factor refinement, occupancy refinement and electron density map calculations. Features of the program pertinent to the thesis are discussed briefly in subsequent sections. Model building and map fitting were done manually using the Quanta software (Accelrys).

Refinement techniques generally fall into two categories, depending upon whether the calculations are performed either in the real space or in the reciprocal space. Refinement methods based on reciprocal space is preferred over real space because the former ones are computationally less intensive. To prevent the model from going into a

local minimum, interactive graphics was used for checking the fit of the model to the electron density. To monitor the progress of refinement the crystallographic parameter called R factor is used. This is defined as,

$$R = \frac{\Sigma_{hkl} \left| \left| Fo(hkl) \right| - \left| Fc(hkl) \right| \right|}{\Sigma_{hkl} \left| Fo(hkl) \right|}$$

Where h, k, l, are the miller indices of Bragg reflections, and the summation is over all the reflections.

2.9.9 Constraints and restraints: During the course of refinement in macromolecules, some groups of atoms may have to be constrained or restrained to improve the ratio of observables to parameters. CCP4 suit has options to group atoms so that they move as rigid bodies, or, restrain or constrain the bond lengths, bond angles, non-crystallographic symmetry (NCS) and atomic positions to desired values by use of appropriate force constants. Restraints are used when limited freedom can be allowed for a parameter. Constraints are used to hold a parameter to an ideal value. A constraint is equivalent to a restraint having the value of force constant infinity. In NCS restraints, the molecules in the asymmetric unit are superposed by least squares superposition and the average co-ordinates (x_{av}) of individual atoms are computed. If x represents the co-ordinates of individual atoms then each atom can be restrained according to the mathematical term:

$$E_{NCS} = w(x - x_{av})^2$$

And the corresponding B-factor restraints are given by:

$$B_{NCS} = (b-b_{av})^2 / \boldsymbol{\sigma}^2_{NCS}$$

Where w is a weight function, b and b_{av} are the respective individual and average temperature factors of NCS related atoms and σ_{NCS} the target deviation for B-factor restraints. Each of the hexamers in C-PC asymmetric unit has molecular point group symmetry 32 (D3). However, the application of NCS constraints imposes a strict restriction on the refinement of related subunits or parts, which is usually violated for macromolecular structure data with resolution better than 3.5 Å (Kleywegt, 1996). Hence in the initial stages, refinement was carried out with NCS restrains with different weights applied between either all l2 monomers of hexamer $(\alpha\beta)_6$ or between the two $(\alpha\beta)_6$ hexamers of the asymmetric unit. However, inspection of relative change in R_{free} factor with and without NCS restraints (tight) (Kleywegt & Briinger, 1996) led us to keep the NCS restraints (tight).

2.9.10 Electron density maps and Model building: After every refinement cycle the model was inspected and manual rebuilding done by inspecting (2Fo-Fc), (Fo-Fc) and negative (Fo- Fc) electron density maps. CCP4 has options to calculate σ_A-weighted maps where the structure-factor amplitudes are weighted in order to reduce the model bias of an incomplete or partially incorrect structure. The Fourier coefficients calculated are given by:

$$F_{map} = (m\ Fo - D\ Fc) \exp (i\alpha_c)$$

Where m is the figure of merit and D is a measure of the error in the co-ordinates of the model, is as defined by Luzzati (1952). Electron density maps using the Fourier coefficients derived here have been observed to be superior to the conventional 'unweighted' maps (Main, 1979; Read, 1986). The (2Fo-Fc) and (Fo-Fc) maps were usually contoured at 1.0 σ and 2.0 σ, respectively, where σ refers to the r.m.s. deviation of the mean density in units of electrons/$Å^3$ in the maps.

Model building was done using the program Quanta (Accelrys). Manual rebuilding using quanta is less subjective because of the incorporation of information

from databases of known structures. Information from the main chain database (Jones & Thirup, 1986), side chain rotamer database (Ponder & Richards, 1987), the peptide orientations database (Zou & Mowbray, 1994) are available for model rebuilding and the results have been analyzed through real space electron-density fit.

Omit maps were computed very frequently during the process of model building and refinement. After omitting residues in a region in question and map calculation after refinement, the model was built into the omit density, thus overcoming any model bias. The model bias, if any, would clearly appear as –ve density in (Fo-Fc) maps and such regions could be built successfully by omit maps. The model that emerged after each cycle of model building and refinement was checked for unusual geometry at each residue, the peptide-flip values, and high temperature factors. The other parameters checked include values of torsion angles, including (Ramachandran angles φ and ψ), side chain torsion angles such as χl, χ2 etc. were verified between refinement cycles using the program PROCHECK. Special care was taken to resolve the side-chain flip conflicts for all Asn, Gln and His residues by analyzing their environment and interactions.

2.9.11 Refinement by maximum-likelihood method: The initial models of C-PCs and AT*B*XYL-C obtained from MR calculations were refined using the program REFMAC (implemented in CCP4) which makes use of maximum-likelihood equations. Bricogne (1991) have suggested a maximum-likelihood target that should be based on various probability distributions. One of the expected advantages of maximum likelihood refinement is a decrease in refinement bias, as the calculated structure-factor amplitudes will not be forced to match the observed amplitudes (Read, 1997). Use of appropriate likelihood targets through the incorporation of the effect of measurement of error and the use of cross-validation data to estimate the σ values (error) are the key ingredients in the likelihood refinement. Verification tests have shown that for protein structure refinement, maximum likelihood method is more than twice as effective compared to the least-squares method, in improving the model (Pannu & Read, 1996). REFMAC program can carry out rigid body restrained or unrestrained refinement using X-ray data (Murshudov *et al*., 1997).

The program minimizes the coordinate parameters to satisfy a maximum-likelihood or least squares residual. Before running the REFMAC, PROTIN program was run to invoke geometric restraints. PROTIN analyses the protein geometry and produces an output file containing restraint information. REFMAC also produces an output file with extension MTZ (named after three of its progenitors, McLaughlin, Terry and Zelinka) containing weighted coefficients for weighted mFo-DFc and 2mFo-DFc maps. About 5% of the reflections were kept aside during refinement to calculate R_{free} for cross validation (Brunger, 1992). NCS averaging is effectively done for an asymmetric unit composed of N similar objects related by non-crystallographic symmetry (NCS). Rossman and Blow (1963) proposed this method, by which the current phases of reflections can be improved by averaging over the electron densities of NCS related objects. NCS averaging requires an accurate estimate of NCS operators and exact information on the position and shape of the objects whose density has to be averaged (Vellieux & Read, 1997). Since both C-PCs and AT*B*XYL-C structures had more than one molecule in their asymmetric units, their initial models were refined taking advantage of the presence of non-crystallographic symmetry.

2.9.12 Analysis and validation of structures: The program PROCHECK was used for checking the stereochemistry and quality of the model. This program is written by Laskowski *et al*., (1993) and forms part of the CCP4 suite of programs. Once the model has improved, after every cycle of model fitting and run of REFMAC it was subjected to stereo chemical quality check. The program compares and assesses the quality of the model with the available structures of similar or better resolution than the reference structure. The output contains a comprehensive residue by residue listing of the parameters and plots of their values. The program highlights the regions of the structure where the conformations deviate considerably from ideal model. These can either be due to interesting properties of the structure or possible errors in interpretation, that require further investigation during next rebuilding step or analysis.

Chapter-3

Purification, Crystallization and Preliminary X-ray characterization of C-phycocyanin from *Phormidium, Lyngbya spp.* (marine) and *Spirulina sp.* (fresh water).

3.0 INTRODUCTION

In order to determine the three-dimensional structure of a biological macromolecule by X-ray crystallography, well-ordered single crystals of the molecule is necessary. It is usually possible to grow single crystals only from a sample that has been purified to homogeneity. Crystallization of macromolecules still remains one of the major rate limiting steps in single crystal diffraction work. There have been many attempts to understand the process of crystallization to arrive at a condition that will yield crystals of biological macromolecule (Carter *et al*., 1979; McPherson, 1976; Jancarik & Kim, l99l; Ducruix & Giege, 1992; McPherson, 1999). Recent attempts toward rational understanding of the process of crystallization involve study of crystal growth in microgravity conditions, *in situ* growth studies with high-resolution atomic force microscopy (Malkin *et al*., l995a, b) and dynamic light-scattering experiments (Thiubault *et al*., 1992). The systematic screening of all the parameters involved in the crystallization of macromolecules is still time consuming. To quickly reach the crystallization condition with less labour, several automated systems have been introduced. In an approach towards reducing the task of laborious screening and to reduce the time involved in identifying the appropriate condition, a method known as sparse matrix screening (Jancarik & Kim, l99l) and Crystal Statagy Screen I & II has been very effective. The method is based on a statistical survey of known crystallization conditions as deposited at Biological Macromolecule crystallization Database IBMCDI (Gilliland *et al*., l994) and identifying the unique variety of diverse conditions at which the macromolecules have been crystallized. We have achieved the crystallization of CPC from *Phormidium, Lyngbya and Spirulina spp.* by manual screening of a subset of the sparse matrix conditions. This chapter presents the purification, crystallization, X–ray diffraction data collection and data processing of the protein C-phycocyanin from *Phormidium* and *Lyngbya spp*. (Marine) and *Spirulina sp.* (Fresh water),

3.1 Cultivation of *Phormidium* and *Lyngbya spp.* (marine), and *Spirulina sp.* (fresh water).

3.1.1 Isolation: The marine cyanobacteria *Phormidium* and *Lyngbya spp.* were isolated from the rocky surface near the sea coast of Gujarat (latitude 21º 38' N and longitude 69º 37'), in the west coast of India. These organisms were grown in batch cultures in the standard artificial seawater medium ASN III (Rippka *et al*., 1979) at pH 7.5 and temperature 20 ± 2 °C with optimum light intensity of 60 lE m^{-2} s^{-1} provided by cool-white fluorescent tubes with a dark : light cycle of 12:12 h. The freshwater cyanobacterium *Spirulina sp*. was grown in batch cultures in Zarouk medium (Chen *et al*.,1996) at pH 10 and temperature 20 ± 2 °C with optimum light intensity of 60 lE $m^{-2}s^{-1}$ provided by cool-white fluorescent tubes with a dark light cycle of 12:12 h. This work has been done in our collaborator's laboratory at CSMCRI, Bhavnagar Gujarat, India.

3.1.2 Extraction of C-phycocyanin: The fresh cyanobacterial cells were harvested after 7-10 days (fresh water) and 15-20 days (marine) of incubation under laboratory-controlled conditions (temperature, pH, and light) by centrifugation at 8,000xg for 30 min. The harvested cell mass was washed twice with distilled water and freeze-dried. One gram of freeze-dried cell mass was suspended in 100 ml of sodium- phosphate buffer (0.1 M, pH 7.0, containing 1 mM sodium azide). Repeated freezing at -20°C and thawing at room temperature in the dark followed the further extraction of phycobiliproteins. The mixture was subsequently centrifuged at 10,000xg for 30 min at 4 °C and a phycobiliprotein containing clear supernatant was collected.

Table 3.1a Composition of ASN-III Medium

Compound	g l^{-1}	mM
NaCl	25.0	427
$MgCl_2.6H_2O$	2.0	9.8
KCl	0.5	6.7
$NaNO_3$	0.75	8.8
$K_2HPO_4.3H_2O$	0.02	0.09
$MgSO_4.7H_2O$	3.5	14.2
$CaCl_2.2H_2O$	0.5	3.4
Citric acid	0.003	0.015
Ferric ammonium citrate	0.003	0.015
EDTA(disodium magnesium)	0.0005	0.0015
Na_2CO_3	0.02	0.19
Trace metal mix A5+Co	5 ml	-
Deionized water	1 L	-
pH after autoclaving: 7.5		

Table 3.1b Trace Metals (A5+Co)

Compound	g /200ml
H_3BO_3	0.6
$MnCl_2.4H_2O$	0.4
$ZnSO_4.7H_2O$	0.044
$Na_2MoO_4.2H_2O$	0.06
$CuSO_4.5H_2O$	0.016
$Co(NO_3)_2.6H_2O$	0.017

Table 3.2 Zarouk medium Composition

Ingredient	g L^{-1}
NaHCO3	16.8
K2HPO4	0.5
NaNO3	2.5
K2SO4	1.0
NaCl,	1.0
MgSO4 z 7H2O	0.2
CaCl2,	0.04
FeSO4 z 7H2O	0.01
EDTA	0.08
H3BO3	2.86
MnCl2 z 4H2O	1.81
ZnSO4 z 7H2O	220
CuSO4 z 5H2O	79
MoO3	15
Na2MoO4	21
pH was maintained at 10.5	

3.1.3 Purification of C-PC: Purification details for the three C-PCs are summarized in Table 3.3. All steps in the purification were performed in the dark at 15–20ºC. The clear supernatant of C-PC was fractionated by precipitation with solid ammonium sulfate (AS) first at 25% and then at 50% saturation. The precipitate from 25% saturated AS was discarded. The supernatant was further brought to 50% AS saturation and allowed to stand for 4 h at 4°C. The precipitate proteins was collected by centrifugation at 10,000xg for 30min at 4°C and resuspended in acetate buffer (0.1 M, pH 4.5) to precipitate out the basic linker polypetides. The precipiatate was separated by centrifugation at 10,000xg for 30 min at 4°C and discarded. The supernatant was again brought to 50% saturation with solid AS and allowed to stand for 4 h at 4°C prior to centrifugation at 10,000xg for 30 min at 4°C. The pelleted C-PC was dissolved in 5 ml of sodium phosphate buffer (0.005 M, pH 7.0) and dialyzed extensively at 4°C against the same buffer. The dialyzed solution of C-PC was chromatographed on a DEAE– Sepharose CL-6B column (1.5x15 cm.). The column was developed with a linear gradient of NaCl (0–0.25 M) at a flow rate of 0.5 ml/min. C-PC eluted between 0.10 and 0.20 M NaCl concentrations and was collected in 2 ml fractions. Absorption spectrum from 250 to 800 nm was recorded to monitor the purity of fractions. All the fractions having purity ratio of A620/A280 > 4.0 were pooled together and brought to 50% saturation with solid AS. Separation factor (A620/A652) indicates the separation of C-PC and APC (Baussiba *et al*., 1979; MacColl *et al*., 1971). The precipitated C-PC was dissolved in a small volume of sodium-phosphate buffer (0.005 M, pH 7.0), dialyzed against water at 4°Cand freeze-dried for storage.

3.1.4 SDS PAGE: Purified C-PC sample was electrophorosed according to the method of Laemmli (1990) using 15% gel containing 0.1% (w/v) SDS. Samples were preincubated with 2% (w/v) SDS, 10% (v/v) glycerol, 4.5% (v/v) β-mercaptoethanol, 0.025% (w/v) bromophenol blue and 60mM Tris (pH 6.8), for about 4 min at 95ºC. Gels were run at room temperature and developed with coomassie blue R-250. Electrophorosed gel

revealed 2 bands corresponding to α and β subunits (Fig.3.1), indicating that the linker polypeptides were effectively removed in the purification procedure.

3.1.5 Ultraviolet-visible spectrometry: The purity of C-PC was determined by the ratio of absorbance at 620 nm (A620) to 280 nm (A280). The A620/A280 ratio is a measure of the purity of the folded C-PC. The value of the ratio of A620/A280 above 4.0 corresponds to pure C-PC (MacColl *et al*., 1994). The ratio A620/A280 is also a critical measure of the stability of the protein. The A620 peak is mainly due to the linear conformation of PCB chromophore in the folded protein (Glazer & Hixson, 1977). The unfolding of protein exposes the bilin molecules to the solvent (O'hEocha, 1965). It is reported that the PCB chromophores assume a cyclic conformation in solution with subsequent loss of absorbance and fluorescence (Riidiger, 1992) in the characteristic bands in the visible region of the spectrum. Purified C-PC in phosphate buffer (0.1 M, pH 7.0) was used for recording the spectra. The absorption spectra showed two peaks, at 342 and 620 nm (Fig.3.2). The absorption spectra of C-phycocyanin purified here are similar to the earlier reported spectra of phycocyanin from *S. platensis* (Boussiba & Richmond, 1979). The spectroscopic properties of phycobiliproteins are critically dependent on its state of assembly like monomer, trimers and hexmares.

3.1.5 Estimation of phycobiliproteins: The absorbances of phycobiliprotein containing supernatant were measured on a VARIAN CARY 50 BIO Scan UV–Vis, NIR spectrophotometer at wavelengths 620, 652, and 562 nm for calculating the concentrations of C-PC, APC, and PE, respectively, using the equations described in the Chapter 2, Section 2.2.2. (Bennett *et al.,* 1973).

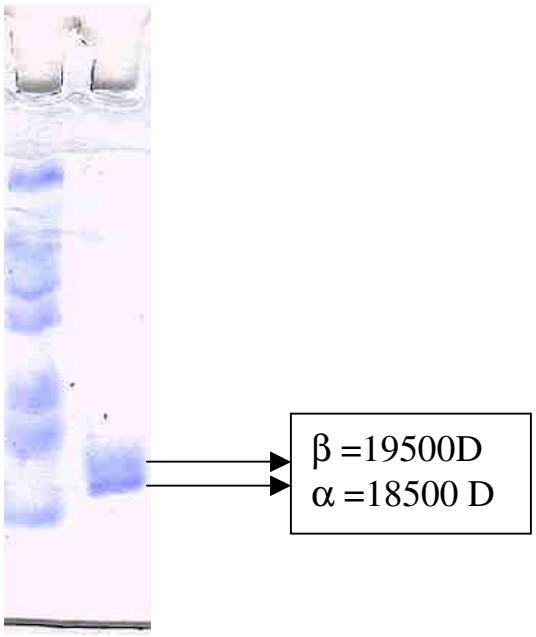


Fig. 3.1 SDS-PAGE gel showing two pure bands of α and β subunits

Lane1: Molecular Weight Markers & Lane2: C-CPC

Table 3.3 Purification of C-PCs

	Spirulina sp.			*Phormidium sp.*			*Lyngbya sp.*		
Purification step	**Purity ratio A620/A280**	**Separa-tion factor A620/A652**	**Recovery (%)**	**Purity ratio A620/A280**	**Separa-tion factor A620/A652**	**Recovery (%)**	**Purity ratio A620/A280**	**Separa-tion factor A620/A652**	**Recovery (%)**
Crude extract	**0.80**	**2.54**	**100**	**0.69**	**2.42**	**100**	**0.67**	**2.43**	**100**
Fractional precipitation with $(NH4)_2SO_4$ (25% saturation)	**0.82**	**2.64**	**95**	**0.73**	**2.50**	**98**	**0.91**	**2.80**	**92.5**
Fractional precipitation with $(NH4)_2SO_4$(50% saturation)	**2.66**	**3.14**	**81.5**	**1.62**	**2.90**	**84.6**	**1.46**	**3.19**	**67.0**
DEAE–Sepharose CL-6B	**4.42**	**5.82**	**45.6**	**4.43**	**7.97**	**35.2**	**4.59**	**5.07**	**36.8**

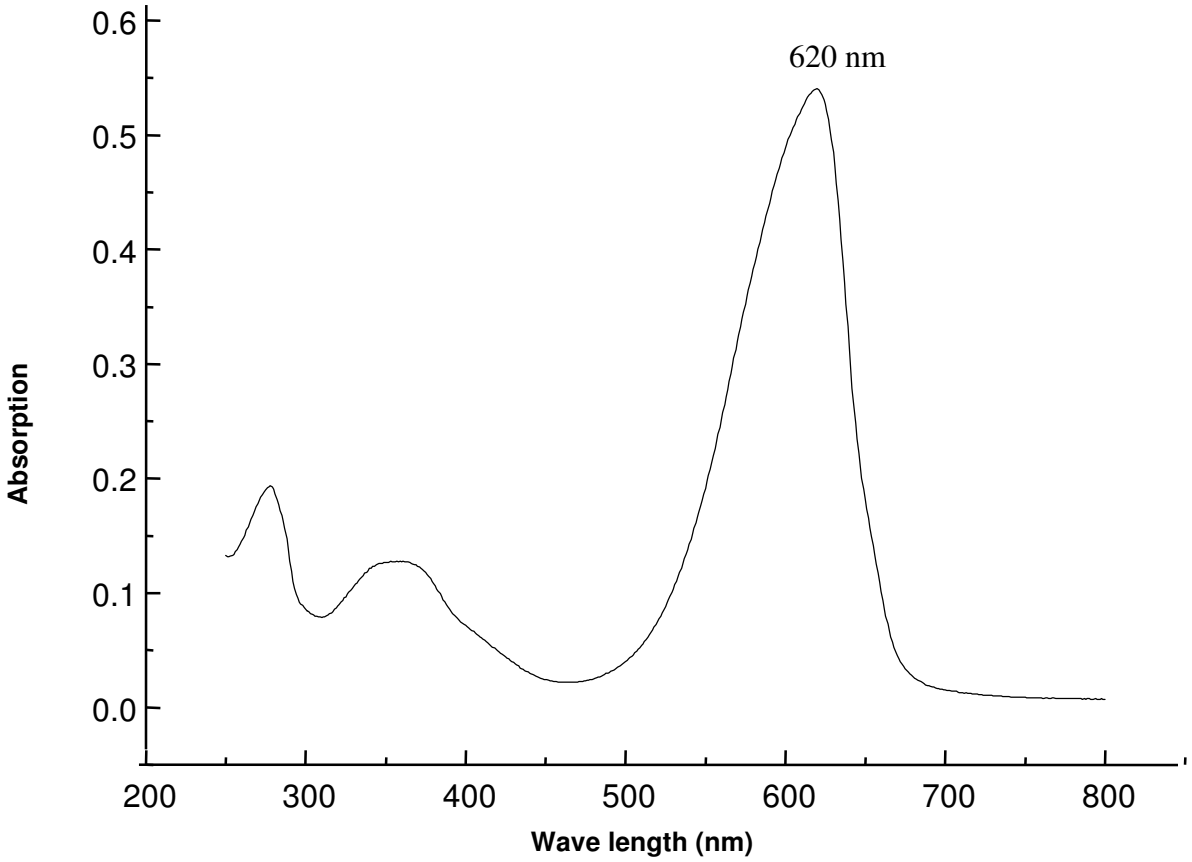


Fig. 3.2 Absorption spectra of C-PC from *Phormidium sp*.

3.2 Crystallization of C-PCs

3.2.1. Historical Perspective: The first ever report of crystallization of phycobiliproteins dates back to more than a century. Molisch in 1891 was able to crystallize phycoerythrin from living cells by adding a few drops of thiocarbon to whole cells from a Mediterranean algae and transferring them into 10% NaCl solution. After a few hours he observed under the microscope the appearance of PE crystals. Fascinated by the strong fluorescence of these crystals he further investigated this phenomenon and a year later reported the crystallization of C-PC upon addition of ammonium sulfate to a C-PC solution.Since then C-PC has been crystallized by Kylin (1910), Lemberg (1928), Sevedberg *et al*. (1929) and Hattori and Fujita (1958). Hackert *et al*. (1977) crystallized PC from *Agmenellum quadruplicatum*, Sweet *et al*. (1977) and Fisher *et al*. (1980) from *Anabaena variabilis*, Morisset *et al*. (1984) from *Chroomonas*, Mikir *et al* (1990) from the red alga *Porphyra tenera* and Duerring *et al*. (1991) from *Fremyella diplosiphon*. At the Swiss Federal Institute of Technology in Zurich, Professor Zuber's group determined the complete amino acid sequences of all *Mastigocladus laminosus* phycobiliproteins by Edman degradation method (Frank *et al*., 1978, Sidler *et al*., 1981; Fuglistaller *et al*, 1983, Suter *et a1*, 1987, Rumbeli *et al*., 1987a, 1987b). Using their isolation protocol for C-PC they routinely purified C-PC by dialysis against 100mM sodium phosphate and centrifuged to isolate C-PC-microcrystals. In order to obtain the crystal structure of a phycobiliprotein, attempts were made at crystallization by several groups (Dobler *et al*, 1972; Bryant *et al*., l976; Hacker *et al*, l977; Abad-Zapateuro *et al*, 1977; Fisher et a1, l980), but very often the crystals were either twinned or failed to diffract adequately.

3.2.2 Crystallization of C-PCs from *Phormidum, Lyngbia* and *Sprirulina spp.*: A survey of crystallization conditions of all the published structures of phycobiliproteins was conducted prior to the commencement of crystallization of C-PCs. This was done in order to identify the common factors, if any, among the various parameters of crystallization. However, it was clear that the conditions of crystallization of different phycobiliproteins were not similar to each other. A rigorous screening approach consisting of a variety of conditions with varying parameters involving different pH, buffers, salt and precipitant

concentrations were adopted. The initial screening for crystallization, which is essentially a derived subset of Hampton Crystal Screen kit I & II (Hampton Research, California, USA) and Crystal Strategy Screen I & II, consist of various molecular weights (in a range of 400 - 20,000) of polyethylene glycol (PEG), 2-methyl-2, 4-pentandiol (MPD), ammonium sulfate as precipitants and sodium chloride, magnesium chloride as additive at five different pH ranging from 4.0 to 10.0 and temperatures 295 K. Blue colored lyophilized protein was dissolved in sodium phosphate buffer pH 7 (0.05 and 0.01 mM) to a desired concentration of protein solution. The protein solution was centrifuged to remove particulate matter and the supernatant was used for all crystallization trials. The concentration of the protein used during screening trials was 20 mg/ml. The screening trials were performed using hanging-drop vapour-diffusion technique. In a typical hanging-drop crystallization setup, protein drop with lower precipitant concentration (1-2µl protein + 1-2µl reservoir) is laid in a pre-siliconized microscopic glass cover slide and hung from the top and is equilibrated over a well containing higher concentration of precipitant. Vacuum grease was used for sealing between cover glass and the plastic well. Plastic tissue culture plates with 4x6 wells (CORINING), which ideally suited the purpose were used. The large-scale purification of the C-PC sample was carried out to obtain ample quantity of the protein for use in crystallization trials. This helped to undertake as many experimental trials as possible until good quality crystals were obtained.

Initial screening yielded very small micro-crystals in conditions containing only PEG of molecular weight 4,000, 8,000 and 20,000 as precipitants and within a pH range of 6.0 to 7.5 at room temperature (295 K) in the dark over a period varying between 1- 4 weeks. The subsequent standardization of the condition was achieved by fine-tuning of the concentrations of various molecular weight PEGs within a pH range of 6.5 and 7 .5. Good quality crystals of reasonable dimensions were obtained of the protein in two crystal forms from three species of cyanobacteria within a period of 1-12 weeks, from conditions containing PEG 4000, PEG 8000 and PEG 20,000 as precipitants within the pH range of 6.5 to 7.0. Further refinement of the crystallization conditions include using a combination of additives (in small percentages of 5 – 10 % of w/v) such as sodium

formate with different molecular weight PEGs. Blue colored, good quality crystals of C-PC from *Spirulina, Phormidium, and Lyngbya spp.* could be reproducibly grown with 15-25% of PEG 4,000, PEG 8,000 and PEG 20,000 in reservoir solution with or without 0.72% (w/v) sodium formate at 295 K in two crystal forms: hexagonal and monoclinic. The crystal morphology was plate-like and large in two dimensions. The crystal morphology didn't indicate clear-cut faces of ideal crystallinity. However, careful examination under polarized microscope reveals a clear birefringence and extinction of polarized light, which is characteristic of anisotropic crystalline substances. C-PC had an intrinsic property to absorb light except the blue color, and the crystals exhibited fluorescence in the red region. The typical sizes of the crystals obtained had larger dimensions exceeding 0.5 - 0.2 mm and a thin side of about 0.1 -.02 mm which was reasonable for in-house data collection. Fig. 3.3, 3.4, 3.5, 3.6, & 3.7 show the images of blue colored CPC crystals obtained from three cyanobacterial species *Phormidium, Lyngbya spp* ,(marine) and *Spirulina sp*. (fresh water) and Table 3.4 lists the different crystallization conditions of C-PCs from all three species. It is a fascinating sight to observe the blue colored crystals glow red when viewed through polarizer.

.

Table 3.4 Optimized crystallization conditions for growing single crystals of various C-PCs from different organisms.

Source	Crystal Form & Space Group	Reservoir Solution	Growth Period	Crystal Dimension (mm)
Spirulina. sp.	Hexagonal, $P6_3$	0.05M sodium phosphate pH 7.0 and 20% (w/v) PEG 4K	3-4 weeks	0.4 x 0.2 x 0.1
Spirulina. sp.	Monoclinic, $P2_1$	0.01M sodium phosphate pH 6.5, 0.72 M sod. formate, and 13.5% PEG 4K	1week	0.3x0.15x0.1
Phormidium sp.	Monoclinic, $P2_1$	0.01M sodium phosphate pH 6.5, 0.72M sod. formate, 9%PEG 1K and 9% PEG 8K	1week	0.3x0.17x0.1
Lyngbya sp.	Hexagonal, $P6_3$	0.05M sodium phosphate pH 6.0 and 20% PEG 4K	10-12 weeks	0.15x0.15x0.1
Lyngbya sp.	Monoclinic, $P2_1$	0.01M sodium cacodylate pH 6.5. and 0.72M sodium formate, 7.2% PEG 20K	1 week	0.4x0.2x0.1

3. 3 Crystallization of C-phycocyanin from *Spirulina sp.* in monoclinic and hexagonal forms

The protein sample was prepared at 20 mg/ml concentration in 0.01 M phosphate buffer pH 7.0 and used for crystallization experiments by hanging-drop vapour-diffusion method at 295 K. 24 well format Linbro trays were used in all the experiments.

3.3.1. Hexagonal crystal form: Approximately 150 conditions were screened and multiple hits were obtained. Promising conditions were with pH 6.0 to 7.5 and 10-30% PEG 4000. The hit was refined to obtain an optimal condition involving 0.05 M phosphate buffer pH 6.5 and 20% PEG 4000. 1µl of protein solution was mixed with 1µl of mother liquor and equilibritated against 1ml of the same liquor.Rectangular plate like crystals grew to a final size of about 0.4 x 0.2 x 0.1 mm in about 3-4 weeks time and displayed bright blue colour (Fig.3.3).

X-ray diffraction data was collected from these hexagonal crystals at room temperature (295K) using R-AXIS IV^{++} image-plate area detector (size 300x 300 mm) and Cu Kα radiation (λ= 1.54189) generated by a Rigaku rotating-anode operating at 50 kV and 100 mA. The crystal-to-detector distance was kept at 250 mm. A total of 180 frames with 0.5º oscillation were collected. Of these frames 1-83 were used to get 98.5% completeness. The data was initially processed using CrystalClear software (Rigaku/MSC; Pflugrath, 1999) to verify the quality of data. Subsequently, the intensities between 30.0 - 3.2 Å were integrated and equivalent reflections were merged using the programs DENZO and SCALEPACK (Otwinowski & Minor, 1997).

3.3.2. Monoclinic form: Screening with Crystal Sceeen kits (Hampton Research) resulted in small crystals in a condition involving PEG 4000 and phosphate buffer pH 7.0. This condition was further refined to produce larger C-PC in crystals using 2 µl hanging drops containing equal volumes of protein solution (1 µl) and reservoir solution (1 µl) equilibrated against 500 ml reservoir resolution containing 0.72 M sodium formate, 13.5% PEG 4,000 (450 µl) and 50µl 0.1M sodium phosphate buffer 6.5 pH (Fig.3.4).

Here sodium formate and low ionic strength buffer helps in obtaining monoclinic crystals after 4 days. The crystal was transferred from crystallization drop into 0.1ml cryoprotectant solution containing reservoir solution supplemented with 30% PEG 400 for a few seconds, mounted in a nylon loop (0.05–0.1 mm, Hampton Research) and finally on a goniometer head under liquid nitrogen stream from a cryo-system at 110 K (X-Stream, Rigaku/MSC). The crystal-to-detector distance was kept at 250 mm. The crystal diffracted to a resolution of 3.0Å and was found to belong to space group $P2_1$, with unit-cell parameters a =107.33, b=115.64, c=183.26 Å, β= 90.03º. The value of R_{merge} calculated for reflection data was 9.5%. The completeness of the data set was 99.8% in 30.0-3.0 Å resolution range. Estimated solvent content and Matthews' coefficient showed that two $(\alpha\beta)_6$ hexamers occupied the asymmetric unit of the unit cell. Details of the crystal parameters and data-processing statistics are provided in Table 3.5.

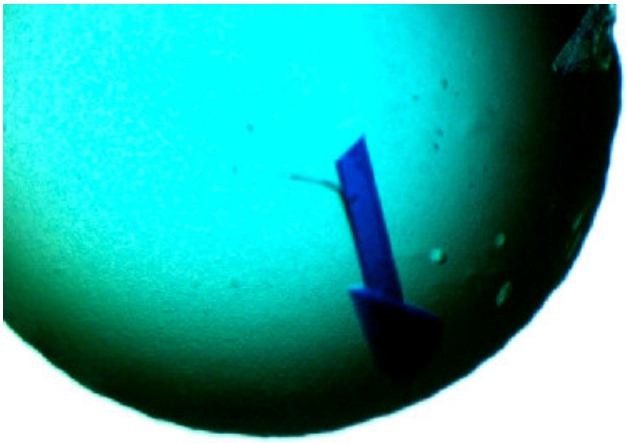


Fig. 3.3 Crystals of CPC from *Spirulina sp.* (Hexagonal)

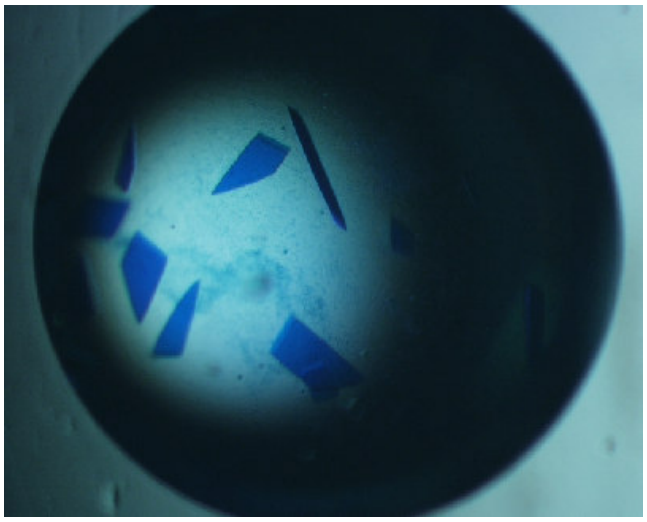


Fig. 3.4 Crystals of CPC from *Spirulina sp.* (Monoclinic).

3. 4 Crystallization of C-phycocyanin from *Phormidium sp.*

3.4.1 Monoclinc form: The crystallization condition for C-phycocyanin from *Phormidia sp*. was first screened using Hampton Research screen kit I and II. Three of the 96 conditions gave aggregates of thin plate-shaped crystals, all using either PEG 4K or PEG 8K as precipitant. Various concentrations of the precipitants and different buffers with various pH values were further tested using the hanging-drop vapor-diffusion method. More than 150 drops were set up, but little improvement was achieved. The conditions reported for the crystallization of C-phycocyanin from other species were also tested, but no diffraction quality crystals could be obtained. Finally, use of 0.72 M sodium formate as additive and 9% PEG 1K and 9% PEG 8K as the major precipitant yielded good crystals (Fig.3.5). The C-PC crystals used for X-ray diffraction data collection was grown at 295 K in a droplet composed of 2µl of protein solution and 2µl of reservoir solution. The protein was dissolved in 0.01 M sodium phosphate pH 7.0 to get 20mg/ml concentration. The reservoir solution (1ml) consisted of 0.01 M sodium phosphate buffer pH 6.5 and 9% PEG 1K, 9% PEG 8K and 0.72M sodium formate. Crystals of C-phycocyanin having dimensions 0.3 x 0.17 x 0.1 mm grew within a week. Just as in the

previous case 30% PEG 400 worked as cryoprotectant. The data was processed using DENZO and HKL2000 (Otwinowski & Minor, 1997). The X-ray data was 93% complete up to 3.0 Å resolution. The crystal parameters and data-collection statistics are shown in Table 3.5.

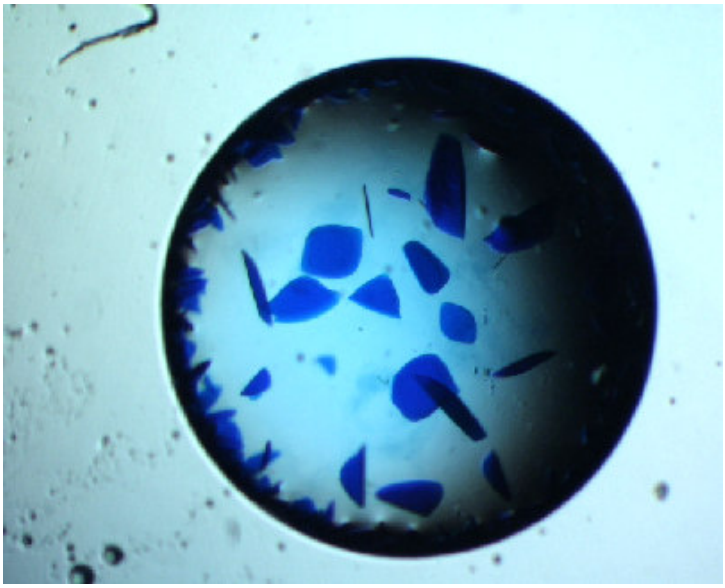


Fig. 3.5 Crystals of CPC from *Phormdium sp.* (Monoclinic)

3.5 Crystallization of C-phycocyanin from *Lyngbya sp.* in monoclinic and hexagonal forms

3.5.1 Hexagonal form: Crystallization was performed using the hanging-drop vapour-diffusion method. Conditions for crystallization were initially investigated using different molecular weights of PEG. Each drop was prepared by mixing 1µl of protein solution with 1µl of reservoir solution and equilibrated against 1ml of reservoir solution. Crystals grew in drops containing 20% PEG 4K and 0.05M sodium phosphate buffer pH 6.0 in 10-12 weeks time (Fig.3.6). For X-ray data collection, crystals were transferred into reservoir solution supplemented with 20% (w/v) PEG 400 and flash-frozen in a liquid-nitrogen stream at 100 K. A total of 115 images were recorded with an exposure time of 10 min per image and an oscillation angle of 0.5º. The data frames were indexed, integrated and scaled using CrystalClear (Rigaku/MSC). The data were separately

processed using DENZO and HKL2000 (Otwinowski & Minor, 1997). The crystal parameters and data-collection statistics are shown in Table 3.5.

3.5.2. Monoclinic Form: Crystallization was performed using the hanging-drop vapour-diffusion method. Conditions for crystallization were initially investigated using polyethylene glycols (PEGs) of different molecular weights but low molecular weight PEGs did not give good quality crystals. Each drop was prepared by mixing 1 µl of protein solution with 1 µl of reservoir solution and was equilibrated against 500 µl reservoir solutions. Crystals were obtained in condition containing 0.72 M sodium formate, 7.2% w/v PEG 20,000, 7.2% w/v PEG 550 MME and 0.01 M sodium cacodylate buffer of pH 6.5 after one week (Fig.3.7).

For X-ray data collection, crystals were transferred into reservoir solution supplemented with 20% (w/v) PEG 400 and flash frozen in a liquid-nitrogen gas stream at 100 K. The data wAS processed using DENZO and HKL2000 (Otwinowski & Minor, 1997).

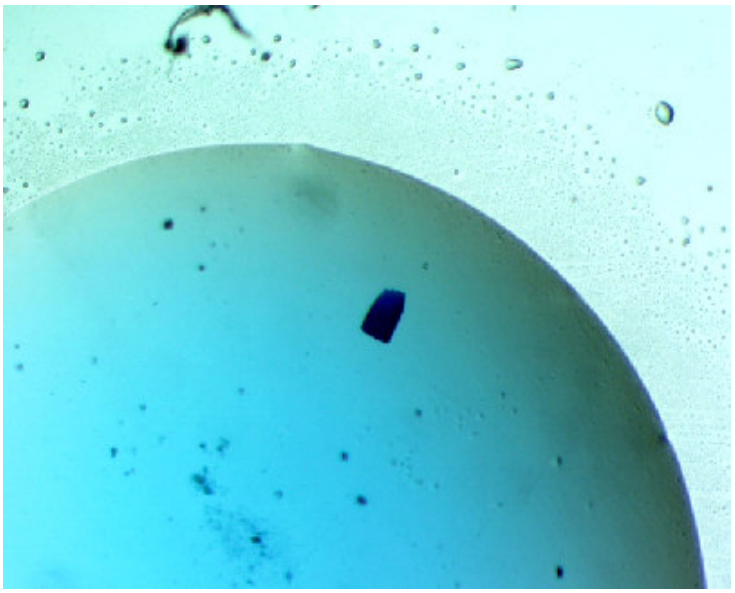


Fig. 3.6 Crystals of CPC from *Lyngbya sp.* (Hexagonal)

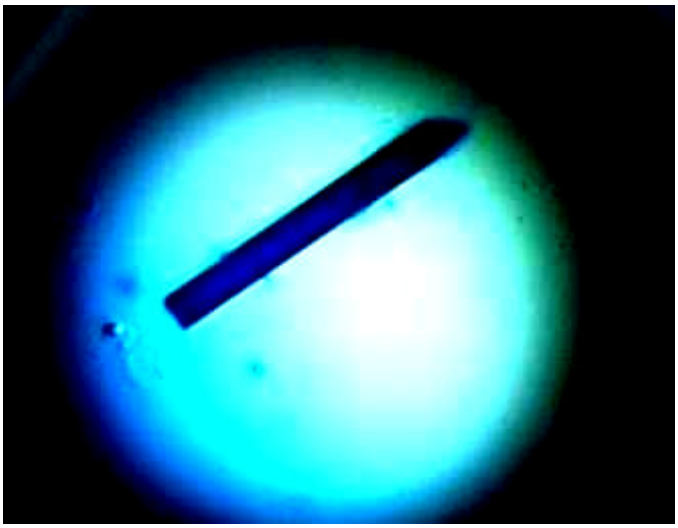


Fig. 3.7 Crystals of C-PC from *Lyngbya sp.* (Monoclinic)

3.6 X-ray diffraction data collection and processing

Since diffraction decays slowly on exposure of the crystals to X-rays at room temperature, data were collected from cryo-cooled C-PC crystals. One of the main advantages of the cryo-cooling is that it reduces the radiation damage by X-rays. The other advantages include easier crystal mounting which permits data collection from even very fragile crystals. Cryo-cooled crystals of C-PC were indeed stable against radiation damage and diffracted up to 3.0 Å. They were mounted in cryo-loop after transferring and soaking in appropriate cryo-protectant. The cryo-protectant consisted of 15-30% PEG 400 (w/v) supplemented in native well solution.Crystals diffracted weakly, hence two initial frames with different exposure times were collected to optimize the exposure time (10 minutes/frame). The crystals diffracted to better than 3.0Å resolution. The crystal-to-detector distance was adjusted to 200 to 250 mm. The oscillation range was selected in the range 0.5 -1° with six oscillations per min. The first frame was auto-indexed and checked to make sure that the data collection parameters set on the system are consistent. The auto-indexing routines yielded a slightly confusing index between primitive monoclinic and primitive orthorhombic systems. Auto indexing of different frames couldn't resolve this problem. Hence, data collection range was chosen to extend

the completeness in lower symmetry monoclinic system. The coefficient and cell parameters obtained from auto-indexing routines for the first frame were a=b=154.97, c=40.35 Å (hexagonal) and a=107.33, b=115.64, c=183.26 Å, β=90.03º (monoclinic). Once the full data was collected, the first frame was indexed and the parameters were refined. The refined parameters were then used to process all the frames in batch mode with DENZO to output " *.x " files containing raw intensity data, hkls and other details. The unit cell parameters on refinement were indicative of a monoclinic cell with β nearly (but not exactly) equal to 90º. The raw data were then scaled and merged using SCALEPACK into one single ' *.sca' file, which contained all the unique hkls and their corresponding scaled intensities and their estimated standard deviations.

The assignement of a monoclinic cell and exact space group was determined by scaling all the reflections with different space groups in SCALEPACK corresponding to monoclinic P2 and orthorhombic P222 system. SCALEPACK writes out the details of scaling and refinement into a log file along with statistics of the processed and merged data. The summary of scaling was verified for the two crystal systems (monoclinic and orthorhombic). The scaling parameters such as χ^2 and R_{merge} for P222 were 10.3 and 35.6% whereas χ^2 and R_{merge} for P2 were around 1.0 and 12.0%, respectively. This has clearly indicated that the crystal system is monoclinic. Further, verification of systematic absences led us to unambiguous assignment of the monoclinic space group $P2_1$. The statistics as given by SCALEPACK and summary of reflection intensities and R_{factor} by shells are given in Table 3.5. The total error and statistical error in individual shells and overall values are quite same, indicating good fit of error model to the data set. The corresponding linear R_{factor} and R^2_{factor} values are also close to each other indicating that everything is normal with detector and the data. R_{merge} of 20.0 to 35.0% in the highest resolution shell is well below the accepted range for a data set with I/σ(I) more than 2 in the highest resolution shell.

Table 3.5 Data collection statistics

Sources of C-PC	*Spirulina* Hexagonal	*Spirulina* Monoclinic	*Phormidium* Monoclinic	*Lyngbya* Hexagonal	*Lyngbya* Monoclinic
Space group	$P6_3$	$P2_1$	$P2_1$	$P6_3$	$P2_1$
Unit cell parameters	a=b=154.97 c=40.35Å	a=107.33 b=115.64 c=183.26Å β=90.03º	a=107.87 b=115.76 c=183.54 Å β=90.3º	a=b=151.96 c=39.06Å	a=107.45 b=115.33 c=183.36 Å β=90.08º
Resolution range in Å	30.0 - 3.2 (3.63 - 3.5)	40.0 - 3.0 (3.11 - 3.0)	40.0 - 3.0 (3.11 – 3.0)	50.0 - 3.6 (3.73 - 3.6)	25.0 - 3.0 (3.05 -3.0)
Total no. of reflections	37447	277752	235483	201951	217963
Unique reflections	7269	86909	82067	10380	86811
Completeness (%)	99.5 (99.0)	96.6 (93.8)	96.6(93.4)	95.1(90.2)	96.4(92.7)
Average I/σ(I)	14.35(8.40)	9.19(3.90)	5.27(2.46)	4.85(3.66)	6.78(3.65)
Rmerge (%)	9.1(20.7)	9.2(20.6)	12.7(29.7)	13.1(23.4)	9.1(20)
Unit cell Volume ($Å^3$/Da)	839347.6	2273079.5	2292597.9	781006.2	2272526.5
Matthews' coefficient V_m ($Å^3$/Da)	3.68	2.50	2.51	3.42	2.49
Solvent content (%)	66.5	50.6	50.9	64.1	50.6

3.7 Standardization of cryo-protectant

It was observed that during the data processing, the mosaicity of the crystal was very high in the order of 1.2º. It has been reported in the literature that the transfer of crystals from the mother liquor to the cryo-protectant marginally increases the mosaicity. Hence it was decided to determine the native mosaicity. A crystal of dimension 0.4 x 0.3 x 0.1 mm was mounted in a thin glass capillary and was exposed to the X-ray beam. A few frames were collected and processed using DENZO and SCALEPACK. The native mosaicity was determined to be as low as 0.4º indicating high crystallinity of the sample. Hence it was concluded that the high mosaicity seen in the cryo-protected crystal was due to incompatible cryocondition. Classically it is suggested that for mother liquor containing PEGs with molecular weights higher than 4K the best cryo-substitution is the low molecular weight PEGs, namely PEG 400 to PEG 1000 (Garman & Schneider, 1997). Other suggested common cryo-protectants are glycerol and MPD. However, it was found that use of 20-30% MPD or 30% glycerol as cryoprotectants either increases the mosaicity further or diminishes the quality of diffraction. The mosaicity with PEG 400 as cryoprotectant was around 0.4- 0.9º, which is higher than the native mosaicity of 0.4º. Data was collected using a crystal cryo-cooled to 110 K employing a solution of 20-30% PEG 400 supplemented in the well solution as cryo-protectant. The data was processed using HKL 2000 suite (Otwinowski & Minor, 1997). Auto-indexing routines gave solutions consistent with a monoclinic in one case and hexagonal space groups in the other. In monoclinic system the space group $P2_1$ was confirmed by looking at the systematic absences. The statistics of the processed data output by SCALEPACK are presented in Table 3.5.

3.8 Conclusions from crystallization experiments

C-PC from all three species of cyanobacteria, *Phormidium, Lyngbya* and *Spirulina spp.* were crystallized using different molecular weights PEGs, with and without sodium formate. C-PCs of marine origin crystallized using bulkier PEGs like PEG 8K and PEG 20K whereas those from fresh water grew using lower molecular weight PEG such as PEG 4K. Use of low ionic strength buffer and presence of sodium formate resulted in the growth of monoclinic crystals. This is the first structure of a C-PC of marine origin to be solved from a monoclinic unit cell. Sodium formate and low ionic strength buffers may have enhanced the interaction of trimers so as to form hexamers and thereby resulting in a monoclinic unit cell with two heaxameric molecules in the asymmetric unit. Absence of sodium formate in the crystallization condition resulted in hexagonal form as in the case of *Spirulina* and *Lyngbya spp*.

Chapter – 4

Crystal structure analysis of C-phycocyanins from *Phormidium, Lyngbya* and *Spirulina spp.*

4.0 Structure of C-PCs from marine and fresh water cyanobacteria.

Efficient absorption and transmission of solar energy by light-harvesting antennae drives the process of photosynthesis. Studies on light-harvesting antennae proteins have resulted in an understanding of diverse phenomena such as evolution of photosynthetic pigment systems (Apt *et al.,* 1995; Grossman *et a1.,* 1995; Tomitani *et al.,* 1999; Green, 2001) and bio-energetics (Sauer, 1975; Knox, 1975; van Grondelle, 1985; Huber, 1989; van Grondelle *et al.,* 1995.). Elucidation of energy transfer pathways has been crucial in understanding the highly efficient energetics of the system. A combination of approaches involving crystal structure analysis and a variety of spectroscopic studies has been instrumental in explaining the energetics of these natural systems. The models studied include light-harvesting systems from purple bacteria and cyanobacteria (Hu & Schulten, 1997; MacColl, 1998). These studies have assumed additional significance in view of the recent attempts to fabricate biomimetic, artificial nano-scale photodevices (Nguyen *et al.,* 2000; Gust *et al.,* 2001).

This structural construction results in a spectroscopic ladder of energy gradient with 'downhill' transfer of the absorbed light energy from PE/PECs at the antennae tip to C-PCs and thereafter to PS II through APC and TEA. In antenna rods, the efficiency of transfer is enhanced by the presence of mostly colorless linker polypeptides.

The chromophore in C-PC is called phycocyanobilin (PCB). Phycocyanobilins are covalently attached to the apoprotein at three different locations in the (αβ)-monomer. They are named α-84, β-84 and β-l55 indicating the position of cysteine residues along the respective polypeptide chains to which they are bound. Though the chemical structure of the chromophores is identical

their spectroscopic properties are modulated and modified by specific electrostatic interactions with the surrounding protein scaffold. The β-155 chromophore is a high-energy absorbing bilin (λ_{max} = 596-600 nm), α-84 being intermediate absorber (λ_{max} = 618-624 nm) and β-84 is the lowest (λ_{max} = 622-628 nm) energy absorber (MacColl, 1998). These differential spectroscopic responses of chromophores result in different spectroscopic parameters for each type of chromophore. Chromophores are divided into two functional spectroscopic classes. The sensitizing (s-) class absorbs at the blue edge (short wave length) of the absorption band and transfers the energy in a non-radiative fashion to the fluorescing (f-) chromophores (long wavelength) (Dale & Teale, 1979). The excitation at the red absorption edge by the f-chromophores causes depolarization resulting in energy hopping as a random walk (trap or limited diffusion) along the line of f-chromophores. The α-84 and β-l55 chromophores belong to the s-class and β-84 chromophore forms the f-class. The functional assignments of chromophores have been verified by extensive studies involving kinetic, spectroscopic and crystal structure analyses (Glazer *et al.,* 1985; Mimuro *et at.,* 1986; Sauer *et al.,* l987; Sauer & Scheer, 1988).

Intra-rod energy transfer pathways between the C-PC hexamers present in the antenna rod has been proposed based on analysis of C-PC crystal structures from *Synechococcus sp*. PCC (Syn.) 7002 (formerly *Agmenellum quadruplicatum*) (Schirmer *et al.,* 1987) *Fremyella displosiphon* (Duerring *et al.,* l99l) and *Spirulina platensis* (Padyan *et al.,* 2001). This chapter presents a possible mechanism for inter-rod energy transfer between the laterally associated C-PC hexamers based on the analysis of crystal structures of C-PCs from *Phormidium, Lyngya spp.* (marine forms) and *Spirulina sp.* (fresh water). This model is suggested based on the observation of two C-PC hexamers in close proximity present in the asymmetric unit of the crystals. This is significant in the context of close association of C-PCs near the core region of phycobilisome light-harvesting complex as seen by using electron microscopic studies reported previously (Gantt *et al.,* 1976: Glazer *et al.,* 1979: Bryant *et al.,* 1979).

4.1 Molecular replacement

The asymmetric unit of a phycocyanin in PDB entry 1GH0 contains 4 trimers corresponding to two $(\alpha\beta)_6$-hexamers. The coordinates of $(\alpha\beta)_3$-trimer was used for the model search using the AMoRe package (Navaza, 1994; Navaza & Saludjian, 1997) for molecular replacement. The molecular replacement calculation was carried out using diffraction data in the resolution range 20.0 - 4.0 Å.

The rotation function calculations (using the program 'roting' of AMoRe) resulted in small number of peaks with nearly similar correlation coefficient (Cc) and R-factors. The significant solutions indicated the orientations of four trimers in the asymmetric unit. The appearance of trimers and uniqueness of selected solutions were evident in further translational search calculations. The translation-function calculations (using program 'traing' of AMoRe) output correct solutions. The selection of solutions was by monitoring the high values of correlation factors Cc and low values of R_{factor}.

Inputting the correct solutions to 'fitting' program of AMoRe and performing a fast least-squares refinement using full data yielded further increase in Cc value and lowering of R-value. A high Cc and a low R-value are indications of correct solution (Navaza, 1994; Navaza & Saludjian, 1997). Table 4.1 lists the MR solutions with final orthogonal rotations (rx, ry, rz) and translations (tx, ty, tz), which were applied to the MR model to generate the initial model of C-PCs from the three cyanobacterial species. The MR solutions when viewed using molecular graphics made it clear that the association of the C-PC were as trimers forming hexamers. These hexameric C-PCs are more relevant form of aggregation from the functional considerations also.

Table 4.1 Molecular Replacement data.

Crystal	α	β	γ	Tx	Ty	Tz	Cc	R-factor
Spirulina **hexagonal**	50.1 4.12 18.60 12.26	162.9 125.63 12.35 26.69	234.5 184.60 35.42 275.08	0.514 0.0961 0.7671 0.5652	0.867 0.0604 0.9210 0.1417	0.000 0.000 0.000 0.000	76.3	34.6
Spirulina **monoclinic**	267.4 87.6 87.3 267.6	96.9 88.3 82.9 92.4	252.2 67.3 306.8 130.9	0.792 0.196 0.781 0.186	0.500 0.258 0.299 0.471	0.0593 0.086 0.580 0.082	67.9	39.3
Phormidium **monoclinic**	273.0 87.0 92.7 267.4	85.1 82.9 91.9 92.1	192.1 306.9 7.6 131.4	0.793 0.781 0.196 0.186	0.501 0.290 0.259 0.472	0.501 0.290 0.259 0.472	67.2	38.0
Lyngbya **hexagonal**	1 9.6	16.1	54.2	0.851	0.317	0.000	74.5	35.7
Lyngbya **monoclinic**	272.7 87.7 87.2 272.4	83.0 88.2 83.0 88.3	312.2 188.0 306.9 71.1	0.792 0.195 0.781 0.187	0.500 0.257 0.289 0.471	0.500 0.257 0.289 0.471	67.9	38.5

4.2 Analysis using sequence alignment of C-phycocyanins

The amino acid sequence of C-PC from *Phormidium* (PHOR), *Lyngbya* (LYNG*),* and *Spirulina spp.* (SPIR) were aligned with the sequences of other C-PCs whose crystal structures have previously been reported. The pair-wise alignment was carried out using the program **EsPript** (Gasteiger *et al.,* 2003.). Figures 4.1 and 4.2 show the alignment of sequences of α and β subunits of C-PCs along with secondary structural elements on top and accessible surface area at bottom.

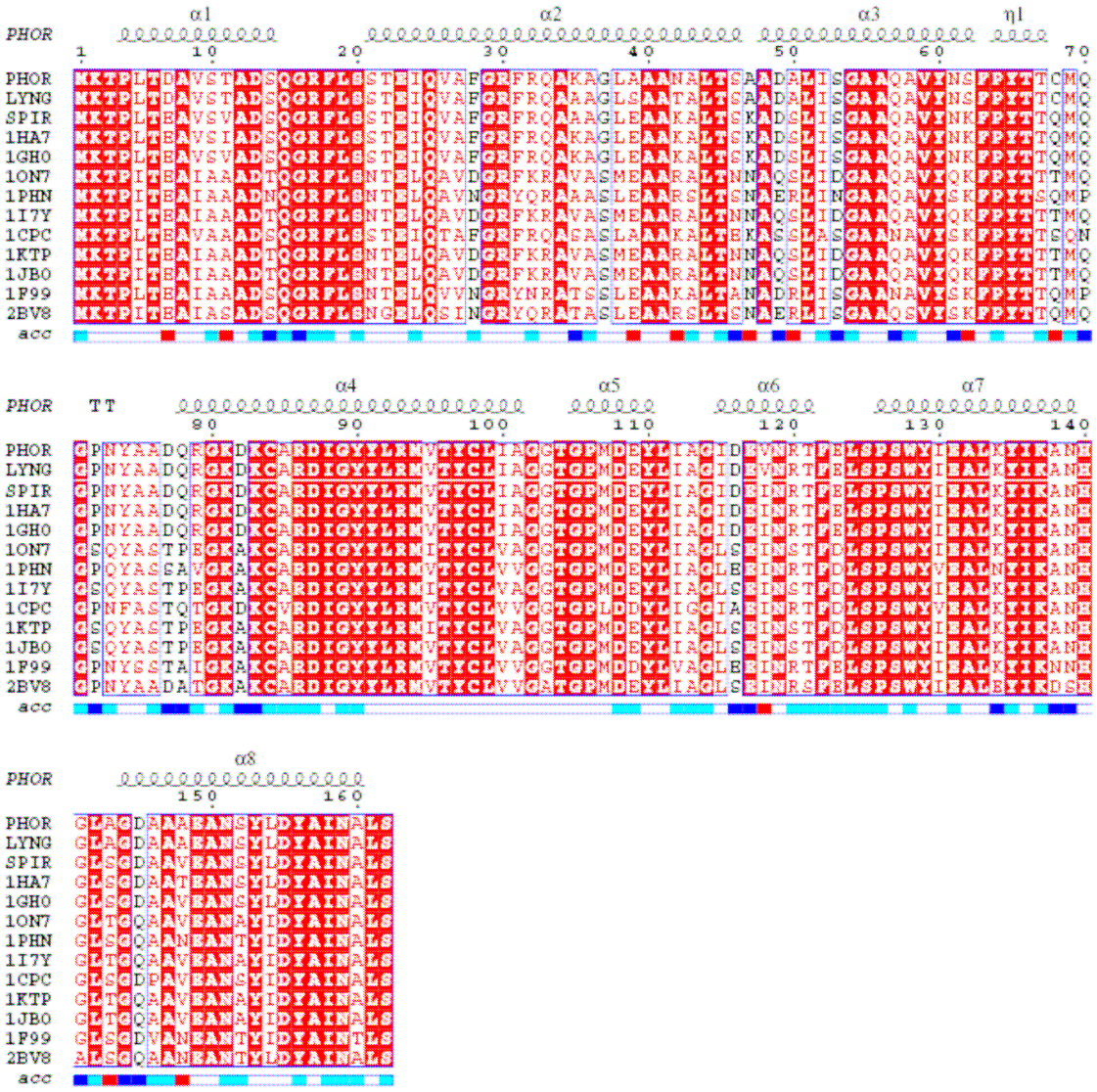


Fig. 4.1 Multiple sequence alignment of α-subunit of C-PCs from *Phormidium* (PHOR), *Lyngbya* (LYNG*), and Spiruluna spp.* (SPIR). Sequences were aligned with EsPript. The sequence titles are: 1GHO (*Spirulina platensis*), **1ON7** (*Thermosynechococcus vulcanus*), **1PHN** (*Cyanidium caldarium*), **1CPC** (*Fremyella diplosiphon*), **1KTP** (*Synechococcus vulcanus*), **1JBO** (*Synechococcus elongatus*), **1F99** (*Polysiphonia*), **2BV8** (*Gracilaria chilensis*).

Fig. 4.2 Multiple sequence alignment of C-PCs β-subunit of *Phormidium* (PHOR), *Lyngbya* (LYNG*), and Spiruluna spp.* (SPIR). Sequences were aligned with EsPript. The sequence titles are:1GHO (*Spirulina platensis* **1ON7** (*ThermoSynechococcus vulcanus*), **1PHN** (*Cyanidium Caldarium*), **1CPC** (*Fremyella diplosiphon*), **1KTP** (*Synechococcus vulcanus*), **1JBO** (*Synechococcus elongatus*), **1F99** ,(*Polysiphonia*), **2BV8** (*Gracilaria chilensis*).

4.3 Nomenclature of α and β subunits.

The 12 (αβ)-monomers of C-PC totaling 24 chains were named in alphabetical order A, B, C, . . ., X with alternate alphabets corresponding to α and β-chains, respectively. The first 12 chains A to L correspond to 'hexamer 1' and chains M to X correspond to 'hexamer 2' of the asymmetric unit. Table 4.2 shows different chain IDs used and the corresponding nomenclature followed for the 'double hexamer' structure of C-PCs.

Table 4.2 Designation of chain IDs

	Hexamer 1				Hexamer 2				
	A-subunit		β-subunit		α-subunit		β-subunit		
	Chain	Chain	Chain	Chain	Chain	Chain	Chain	Chain	
Trimer 1	A	1α	B	1β	M	1'α	N	1'β	Trimer 3
	C	2α	D	2β	O	2'α	P	2'β	
	E	3α	F	3β	Q	3'α	R	3'β	
Trimer 2	G	4^{α}	H	4β	S	4'α	T	4'β	Trimer 4
	I	5α	J	5β	U	5'α	V	5'β	
	K	6α	L	6β	X	6'α	W	6'β	

4.4 Structure refinement

Crystallographic refinement aims at optimizing the agreement of an atomic model with both observed diffraction data and chemical restraints, in the absence of data at atomic resolution. The structures reported in this thesis were

refined using programs incorporated in the CCP4 suite. The programs are available for structure solution, refinement, map calculations and analysis of the structures. Molecular dynamics methods are exploited by CCP4 to probe conformational space of the molecule while minimizing the difference between the observed and calculated structure factors (CCP4i suit). There are options for rigid body refinement, positional refinement, restrained and unrestrained individual B-factor refinement, grouped B-factor refinement, occupancy refinement and electron density map calculations. Model building and map fitting were done manually using Quanta software (Accelrys).

4.4.1 Model building and refinement: The model building and positional refinements were iteratively carried out until all the residues in the model were mutated to the residues compatible with the electron density map (residue information were also obtained from sequence alignment). The density for the backbone of all the 24 chains (12 each of α and β chains) were continuous with all the N and C-terminal residues clearly defined in the map except for a jump in the numbering of 2 residues (73 and 74) in α chain and 10 residues (145 to 154) in β subunit with respect to model. Although visual examination of a molecule of this size is very daunting, electron density for the whole set of molecules were examined and found that density was well defined for a majority of the side-chains.

4.4.2 Modeling the chromophores: The other milestones of model building include the modeling of the covalently attached chromophores into the apoprotein. The phycoyanobilin chromophore is an open tetrapyrrole moiety covalently attached to cysteinyl residues through thioether linkage. The molecular mass of phycocyanobilin chromophore is 5l8 Da. α-chain has one chromophore covalently attached to the cysteine located at position 84, hence the chromophore is also called α-84. In β-chain, phycocyanobilin chromophore is attached to cysteines at two different positions. One at position 84 and another at position 155, accordingly they are named β-84 and β-155. Hence an (αβ)-monomer has a

total of three chromophores. The crystallographic asymmetric unit contained two $(\alpha\beta)_6$ hexamers, therefore altogether a total of thirty six chromophores (eighteen in each of the hexamers) were present. The quality of the electron density map has allowed us to model all the 36 chromophores. The chromophores are designated by three-letter symbol CYC, following the convention of other two deposited structures (l CPC, Duerring *et al.,* l99l; l PHN, Stec *et a1.,* 1999). The topology and parameters for the refinement were obtained from the HICUP server/database (Kleywegt, & Jones, 1998).

4.4.3 B-factor refinement: Once the model building was in the final stages (one indication was that no further significant additional features could be seen in the residual (Fo-Fc) difference Fourier map), the individual B-factors, for all the atoms, were refined isotropically (using individual B-factor refinement protocol in CCP4 suit). The basic assumption made here is that all atomic vibrations have spherical symmetry. The low data to parameter ratio and no high resolution data permit no anisotropic refinement. Individual B-factor refinement was carried out for all three C-PCs alternatively along with positional refinement until R_{values} converged to below R_{crys} = 25.0% and R_{free} =28.0 % for the protein model. Adding of solvent molecules began at this stage.

4.4.4 Adding solvent molecules: Water molecules were added using (Fo-Fc) map as reference. Initially, about 20 water molecules were added based on (Fo-Fc) > 5σ using Quanta (X-solvent) program and subsequently refined using REFMAC5 in CCP4 suite. Finally waters were added at (Fo-Fc) > 3σ level.

The following criteria were used for selecting water molecules in the structure.

(i) When electron density is observed in both (2Fo-Fc) at lσ level and (Fo-Fc) at 5σ level initially and extended up to 3σ level in the final stages.

(ii) The distance between selected water molecule and atoms of protein or other water molecules has to be between 2.4 to 3.6 Å or around these limits for hydrogen bonding or van der Waals contact. At least one hydrogen bond must be

satisfied for selecting a water molecule. Geometry around water molecule for hydrogen bond formation also has been considered. After selecting waters, iterative cycles of B-factor refinement, positional refinement and model building using map from the newly improved phases were carried out to lead to a final refined model. Non-crystallographic symmetry was applied throughout during the refinement, although it was relaxed in the final stages. The final positional refinement of C-PCs converged to parameters as shown in Table 4.3.

Table 4.3. Refinement statistics of C-PCs from *Phormidium*, *Lyngbya* and *Spirulina spp*.

	Phormidium	***Lyngbya***	***Spirulina***
R value (%)	**20.1**	**21.8**	**20.9**
R_{free} (%)	**24.5**	**25.4**	**24.0**
R.m.s. deviation from target bond lengths	**0.021**	**0.022**	**0.020**
R.m.s. deviation from target bond angles	**2.51**	**2.28**	**2.37**
Residues in most favoured/ additionally allowed/ generously allowed / disallowed regions	**3330/194/4 /12**	**3379/150/0 /12**	**3360/161/7 /12**

4.4.5 Quality of the final refined model: The fully refined model consisted of 31,602 (*Spirulina*), 31,137 (*Phormidium*) and 31,130 (*Lyngbya*) non-hydrogen atoms of protein belonging to ~ 4207 amino acid residues in the asymmetric unit and also 1,548 atoms belonging to 36 chromophores. The crystallographic R_{factor} / R_{free}, RMSD from ideal geometry for bond length, bond angle, dihedral and impropers with estimated standard deviation of atomic coordinates for the final protein model are shown in Table 4.3.

The quality of the model was assessed at various stages of refinement for any potential outliers of bond lengths, bond angles, torsion angles etc., using the

program PROCHECK (Laskowski *et al*., 1993). The Ramachandran plots for the C-PCs model is given in Figures 4.3, 4.4 & 4.5. The plot shows ~95.2 % of the residues in most favoured regions of Ramachandran plot and only 12 residues are in disallowed regions. The (phi, psi) values of nearly all non-glycine residues are within the allowed regions of Ramachandran Plot. One exception is noticed for the position of Thr β-77 in the Ramachandran Map, the residue interacts through a hydrogen bond between its amide nitrogen and oxygen of chromophore α-84. The strained geometry of this residue is observed in other known phycobiliprotein structures also. The secondary structural features as observed in structures reported here are shown against the amino acid sequence of α and β- subunits in Figures 4.1 and 4.2.

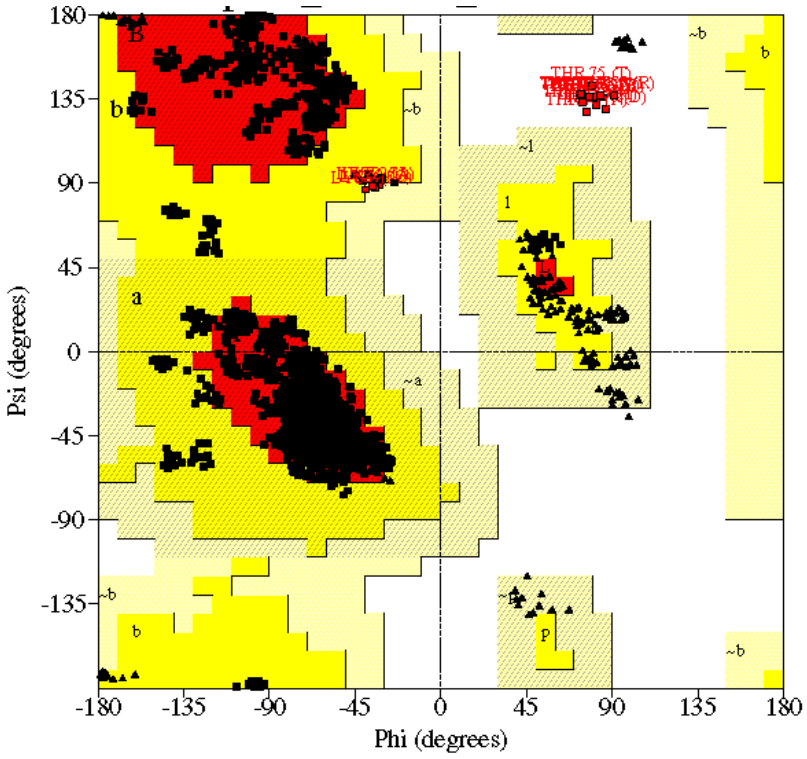


Fig. 4.3 Ramachandran Plot of C-PC from *Phormidium* sp.

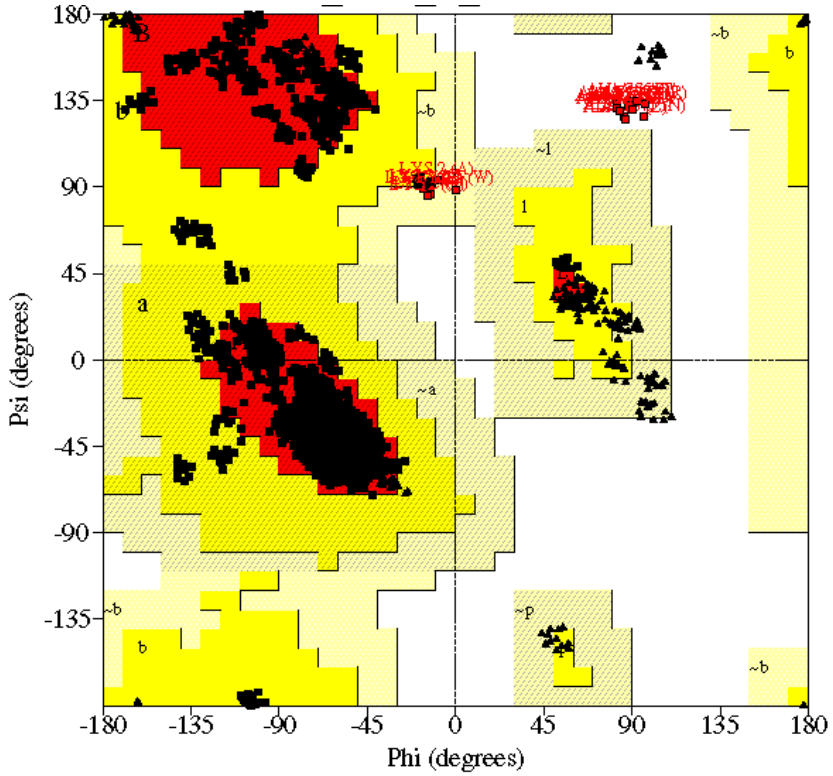


Fig. 4.4 Ramachandran Plot of C-PC from *Lyngbya* sp.

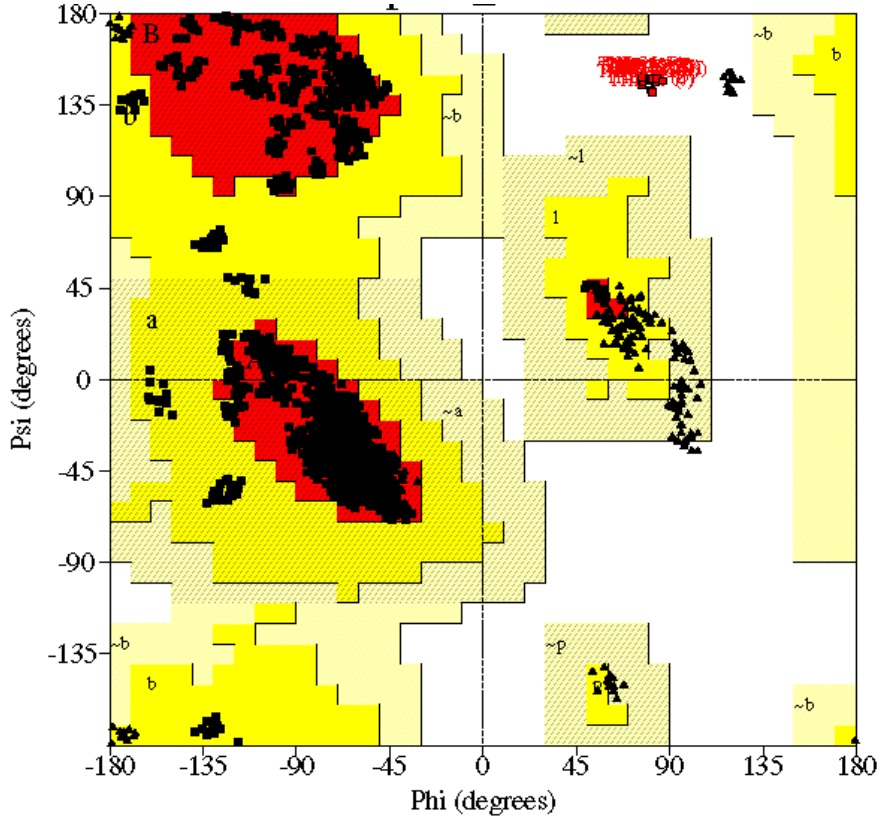


Fig. 4.5 Ramachandran Plot of C-PC from *Spirulina* sp.

4.5 Structure of α- and β-chains

The chain fold of α and β subunits is predominantly helical with a high proportion of about 60 to 65% amino acid residues located in the helix regions. The overall fold of each subunit is very similar to that in the previously reported crystal structures of C-PCs (Schirmer *et al.,* 1997; Duering *et al.,* l99l; Stec *et al.,*1999). This is expected due to high sequence homology of C-PCs from different organisms.

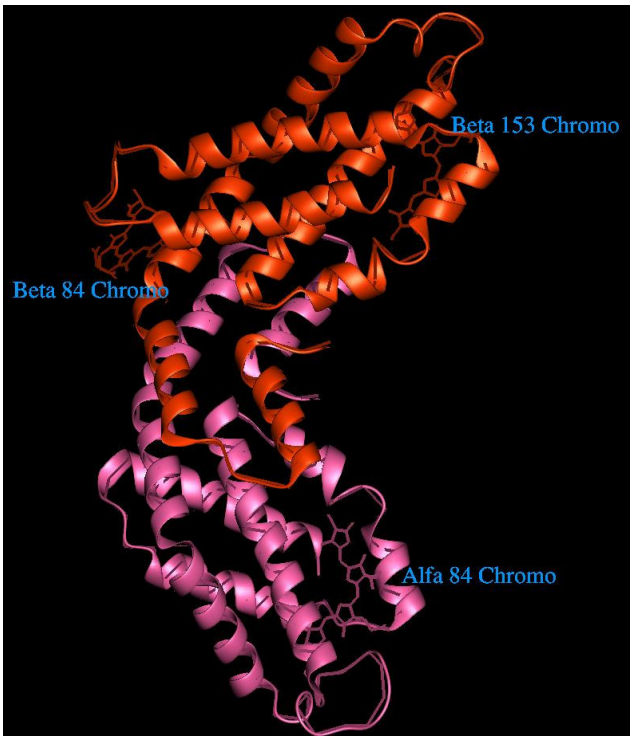


Fig. 4.6 Ribbon representation of the structure of α- and β-chains of C-PC is shown

The structure of both α and β subunits consists of nine helices, namely X, Y, A, B, E, F', F, G and H. The naming convention followed is as used in the first reported C-PC structure (C-PC from *M. laminosus*) (Schirmer *et al.,* 1985). The protein is organized into two distinct domains, a helical hairpin domain and a globin domain. The hairpin domain consists of two helices, X and Y, while helices A to H form the globin domain. The naming of the helices in the globin domain is derived from that of myoglobins due to its close similarity (Schirmer *et a1,* 1985). The helices Y, A and B can be considered as an elongated semicircular

helix. Only one helical H-bond is broken at the junction of these helices. Pairs of helices X/Y, A/E and G/H are anti-parallel and nearly coplanar. The α-chain has one phycocyanobilin (PCB) chromophore covalently attached to the cysteine residue at position 84, whereas the β-chain has two phycocyanobilin chromophores covalently attached to the cysteine residues at positions 84 and 155. Accordingly, the chromophores are termed as α-84, β-84 and β-155. α-84 and β-84 are located in topologically equivalent positions. Their covalent attachment is to the respective helices E. The chromophore β-l55 is attached at the end of the loop connecting G and H helices. The only residue in disallowed region of Ramachandran plot, Thr77, is of the β-chain. Thr77 occurs at the beginning of E helix, and its side chain is near to the α-84 chromophore. Its backbone amide nitrogen interacts with carbonyl oxygen of α-84 chromophore from the adjoining monomer through a hydrogen bond. The strained geometry of this residue is conserved in all known phycobiliprotein structures.

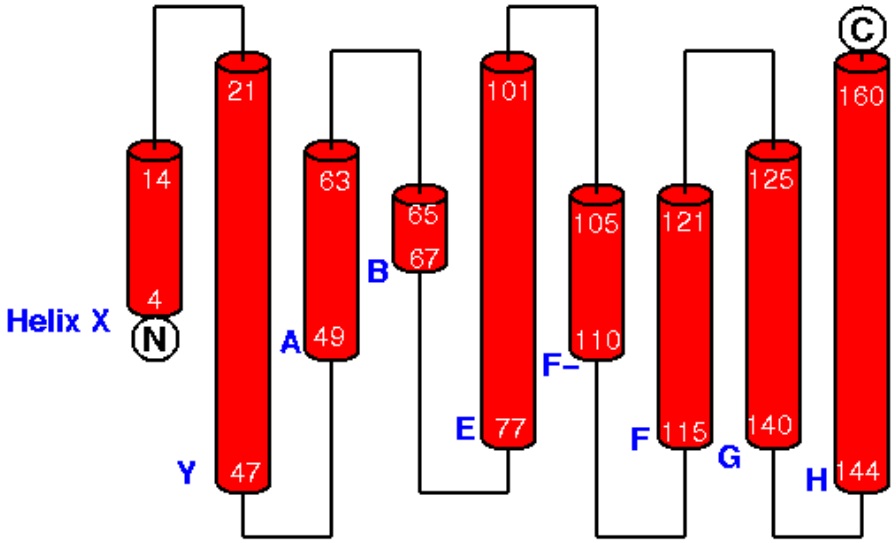


Fig. 4.7 Topological diagram of α-subunit

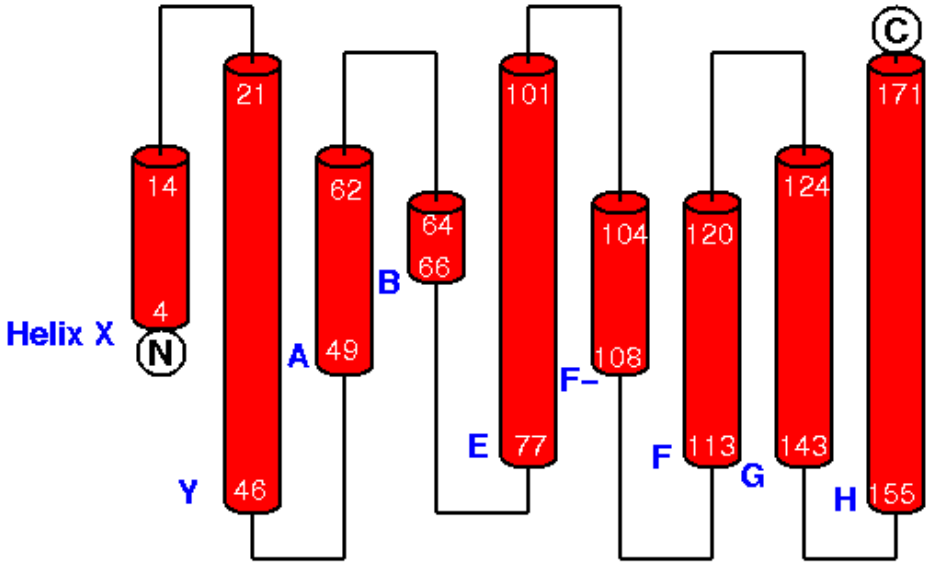


Fig. 4.8 Topological diagram of β-subunit

4.6 Comparison of α- and β-chains

Even though α- and β-subunits differ in their primary structure, their tertiary structures exhibit extensive similarity. The sequence identity between α- and β-subunits is only about 25% which is not very significant. In one attempt at prediction of protein structure the low sequence homology among the α- and β-chains of C-PCs from other organisms had led to the wrong prediction of very different structures for the subunits (Sidler *et al.,* 1981). Later Sweet and Eisenberg (1983) correctly predicted the structural relationships between α and β polypeptides of C-PCs based on the analysis of hydrophobicities of amino acid sequences. The first crystal structure of C-PC from *M. laminosus* (Schirmer *et al.,* 1985) showed a clear structural homology indicating the evolutionary relationship between the two polypeptides. The remarkable similarity of the tertiary fold of α and β chains in C-PCs is in agreement with the reported C-PC crystal structures (Schirmer *et al.,* 1987; Duerring *et al.,* l99l; Stec *et al.,* 1999). The structural superposition also corresponds to a rotation of 177º (for lα with respect to lβ) which gives rise to a pseudo 2–fold axis within the trimer, perpendicular to the 3-fold axis of assembly generating a pseudo 32 (or D3) point group symmetry within the trimer itself. This structural superposition also indicates two regions of residue insertions in α- and β-chains with respect to each other. The first one corresponds to a two-residue insertion in α-chain at position 73-74 extending the loop region between B and E helices. The second insertion corresponds to a twelve-residue stretch in β-chain. This occurs between the residues 144-155 in the loop region between the helices G and H, corresponding to a functional accommodation of an additional chromophore at β-l55. Hence the backbone around β-155 has no structural counterpart in the α-chain. As per the structural superposition of α- and β-chains, numbering gaps have been introduced in each chain so that the residue numbers indicate the equivalent structural positions between the chains.

4.7 Evolutionary relationship between C-PCs and Globins

Globins and C-PCs share neither sequence homology nor functional similarity. Yet, the first C-PC crystal structure *M. laminosus* (Schirmer *et al*., 1985) revealed that both the α- and β-chains show the globin fold between A to H helices. This feature is also seen in the crystal structure of C-PCs under study, as expected from sequence similarity to the models. This is in agreement with the earlier reports of distant evolutionary relationship between globins and C-PCs (Schirmer *et al*., 1985; Pastore & Lisk, 1990). The helices C and D of the globins (PDBID: lmbd; Phillips & Schoenbornl, 198l) are absent in C-PCs. The evidence for the evolutionary relationship arises from the topological similarity and the identical locations of phycocyanobilin (PCB) chromophores and heme binding pockets in the ridge between B and E helices of C-PCs and globins. Further, both the phycocyanobilin chromophores and heme groups are tetrapyrrole moieties. Pastore and Lesk (1990) made extensive comparison of C-PC and globin structures and concluded that the origin of C-PC and globin families may be a consequence of divergent evolution. This hypothesis is in agreement with the details observed in the structures under study.

4.8 Hexamer-Hexamer assembly of C-PC in monoclinic crystals

The major difference between the C-PC structures under study here and those C-PC structures already reported lie in the novel mode of molecular associations in the asymmetric unit of the crystal. The molecular associations observed in the asymmetric unit are not subjected to any of the symmetry restrictions imposed by the crystal symmetry. The near perfect molecular symmetries observed in most of the earlier C-PC structures and other phycobiliproteins were responsible for coincidence of molecular 3-fold axis with the crystallographic 3-fold axis resulting in only part of the hexamer unit occupying the asymmetric unit. The present C-PC structures form a unique

example of phycobiliproteins with the two hexamers closely associated with each other occupying the asymmetric unit.

4.8.1 Comparison of chromophore interactions in C-PCs from marine and fresh water organisms: Chromophores are the light absorbing centers of the C-PC. The chromophores of phycobiliproteins are open chain tetrapyrrole moieties. The phycocyanobilin chromophores are covalently attached to the respective cysteine residues in protein by thioether linkages. All together, thirty-six phycocyanobilin chromophores were located in the protein assembly, three chromophores for each (αβ)-monomer at α-84, β-84 and β-l55 positions.

Although the chemical structure of phycocyanobilins attached at α-84, β-84 and β-l55 is the same, their light absorption responses are remarkably different (Scheer, 198l; Glazer. 1985; Berns & MacColl, 1989). The spectral characteristics of biliproteins, such as C-PC, are dependent not only on the chemical nature of the chromophore but also on its conformation and interaction with the surrounding protein matrix. The fine-tuning of the spectral property of chromophores is achieved by suitable interactions with neighboring amino acids. The interactions are also responsible for restricting the chromophore to a rigid planar conformation, while the free phycocyanobilin chromophore assumes a cyclic conformation as characterized in solution (Rudiger, 1992). In this regard, we have carried out a detailed analysis on the intrinsic conformation of all the chromophores of *Phormidium and Lyngbya spp.* (marine water) and compared it with *Spirulina sp*. (fresh water form). This is important in estimating energy transfer efficiencies which is significantly influenced by the conformation and configuration of the chromophores.

The conformation of phycocyanobilin is similar to that seen in the model structure of C-PC from *Fremyella diplosiphon* (Duerring *et al*., l99l). The conformational variations are mostly responsible for the gross variations seen in the spectral absorption properties of α-84, β-84 and β-l55 chromophores. Spectroscopic studies carried out on the monomeric C-PCs at high-resolution has led to the determination of

absorption properties of each individual chromophore types by deconvoluting the complex spectrum into its components (Mimuro *et al*., 1986; Suer *et al*., l987; Holzwarth *et al*., 1987; Debreczeny *et al*., 1993; Demidov & Mimuro, 1995). The absorption maxima are closely placed around 620 nm for α-84 and β-84 while β-l55 absorbs at relatively lower wavelength of 600 nm. Accordingly, β-l55 absorbs at higher energies, α-84 absorbs at intermediate energy and β-84 at the lowest energy. The polar and ionic interactions of C-PCs also depict highly conserved environments of the chromophores. The network of interactions is responsible for maintaining the proper orientation of chromophores and for providing a suitable dielectric medium. This in turn influences the light absorption properties and energy transfer rates of each of the chromophores.

The energy of photon absorbed by any one of the chromophores is transferred to other chromophores within the C-PC in a non-radiative form. The arrangement of chromophores within a hexamer and in the supramolecular light-harvesting association is to set up a 'downhill' energy gradient such that the excitation is transferred in a directional fashion towards the photosynthetic reaction center (*viz.* PS II). The transfer of excitation energy at various subunit levels is modeled using the theoretical mechanism given by Forster (1948, 1967). The chromophores are treated as dipoles, which interact with the electromagnetic radiation and the transfer of energy between pairs of chromophores is treated as dipolar transitions. The Forster mechanism has been adopted for such calculations in a wide variety of systems. Models of energy transfer pathways can be constructed based on the elucidated pathways of energy transfer. The energy transfer rate between two chromophores is strongly dependent on the relative distance and orientations of the chromophores.

As mentioned before, the unique feature of the present set of structures is the occurrence of two hexamers in close association in the asymmetric unit. It is very interesting to probe the consequence of such an association in relation to the observations from *in vivo* light-harvesting assembly. Comparisons with the earlier electron microscopic studies (Gantt *et al.*, 1976, 1977; Glazer *et al*., 1979; Bryant *et*

al., 1979) of phycobilisome assemblies *in vivo* hinted that the observed association of C-PC hexamers assumes greater importance in terms of providing a quantitative model for excitation transfer of energy in lateral direction. The elucidation of the exact conformation of chromophores in the present structures based on the experimental X–ray diffraction data allowed us to see that there are no differences in chromophores' orientation and distance in marine and fresh water forms. However, the association shows the role of lateral association of hexamer assembly and its functional implication towards assessing the efficiency of mechanism of light harvesting in cyanobacterial system.

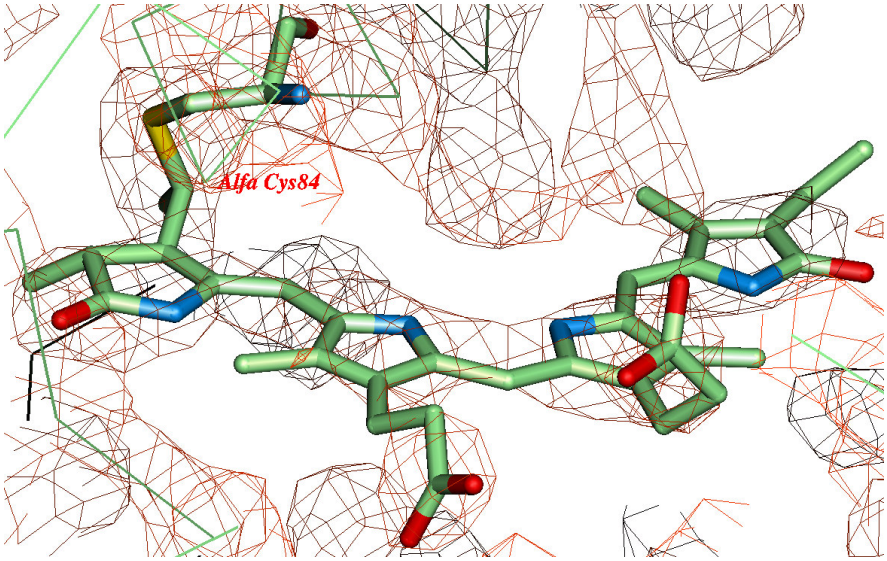


Fig. 4.9 Chromophore attached to a cysteine residue with (2Fo-Fc) map at 2σ.

4.8.2 The structure and interactions of C-PCs of marine and fresh water origin: There are two types interactions of hexamers seen in crystals of marine and fresh water C-PCs that can cause α-84 and β-155 to act as important chromophores for energy transfer. In C-PCs from *Phormidium* and *Lyngbya spp.* (marine forms) position α-68 in the amino acid sequence is occupied by a cysteine residue while it is glutamine in case of *Spirulina* C-PC (fresh water form). Multiple sequence alignment (Fig.4.1) reveals that this position is not conserved, other residues present are tyrosine

(*Synechococcus sp.*) and serine (*Mastigocladus laminosus*). This cysteine residue can play a critical role in multimer arrangement of C-PCs.

The region occupied by residues α-68Cys from two hexamers forms the interface between the hexamers in one side which is seen in the asymmetric unit of C-PCs from marine organisms (Fig. 4.12a). These residues (Fig. 4.10 & 4.11a) demonstrate a lateral approach towards each other of the hexamers. However, the distance of separation between these residues is 4.0Å (Fig. 4.11b) which is higher than 2.3Å required for the formation of a disulphide linkage. Formation of a disulphide bridge under *in vivo* conditions cannot be ruled out. This interaction is absent due to absence of Cys at this position in the fresh water C-PCs in their polymeric arrangement.

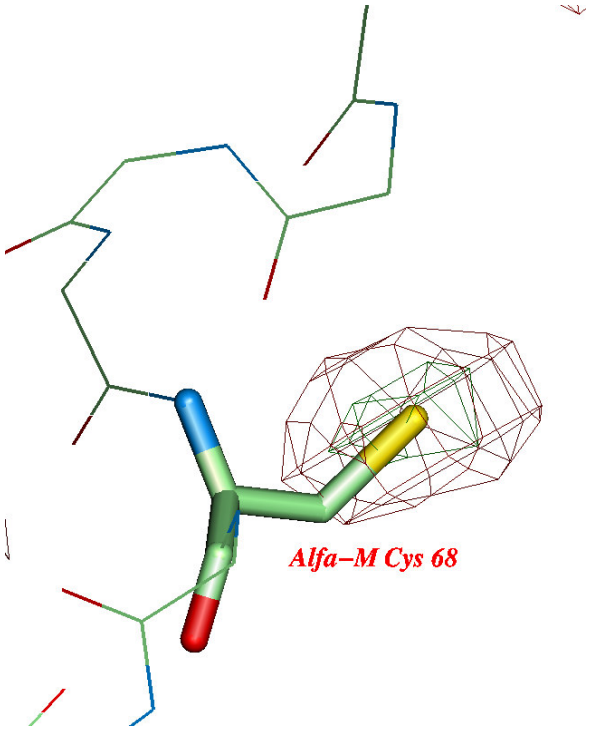


Fig.4.10 Residue α-M 68Cys with (Fo-Fc) map contoured at 5σ (green) and 3σ (red) in *Lyngbya sp.*

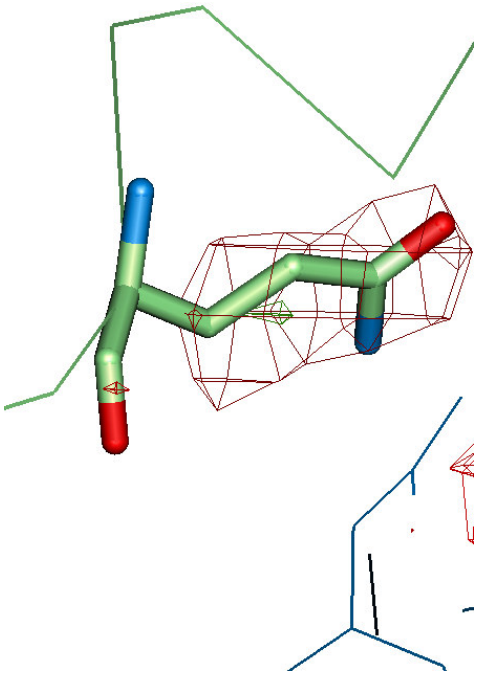


Fig. 4.11a Residue α-M Glu 68 with (Fo-Fc) map contoured at 5σ (green) and 3σ (red) in *Spirulina sp.*

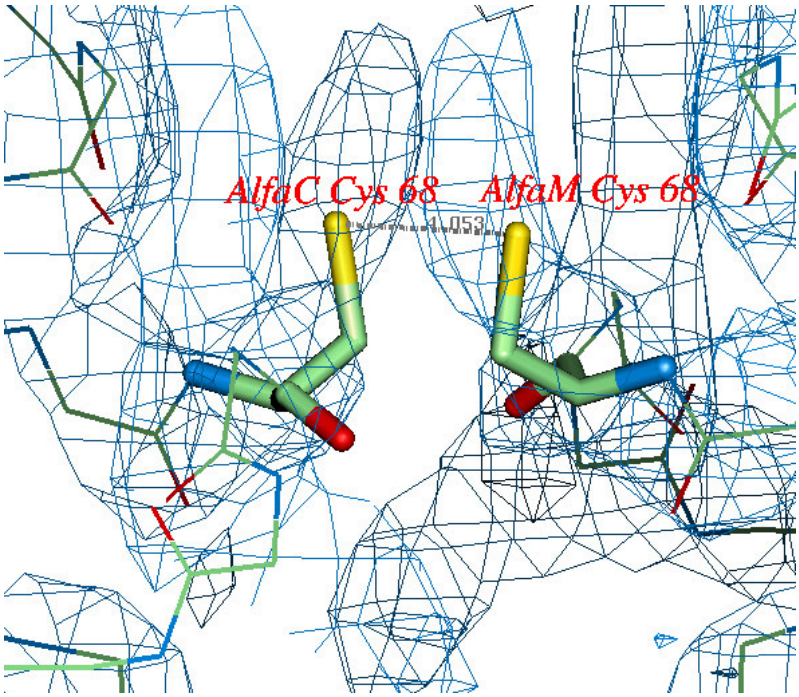


Fig. 4.11b The proximities of two sulfurs for potential disulfide bond formation between α-C 68Cys and α-M 68Cys in C-PC of marine species with (2Fo-Fc) map contoured at 1σ.

4.8.3 Transfer of energy between C-PC hexamers in cyanobacteria: The distance and conformation of the chromophores present at the interface between the hexamers of C-PCs from *Phormidium* and *Lyngbya spp.* in their structures were measured using the program Quanta (Acceryles). In the asymmetric unit of C-PCs from marine source the hexamer-hexamer interaction is due to residues belonging to α-subunits C & I of hexamer-I and residues of α-subunits M & W of hexamer-II. This brings α-84 chromophores into close proximity (C α-84, I α-84 of hexamer-I and M α-84, W α-84 of hexamer-II). The chromophore α-84 absorbs energy at 618-622 nm wave lengths (β-155 absorbs at high energy 600nm and β-84 absorbs at 624-628 nm) which is between absorption wavelengths of β-155 and β-84 chromophores (Scheer, 198l; Glazer. 1985; Berns & MacColl, 1989). The distance of separation between chromophores present in adjacent hexamers was measured from the central atom of the chromophores on hexamer-I and II. The least distance found was between chromophores of α-subunits C-84, I-84 of hexamer-I and M-84, W-84 of hexamer-II (Fig. 4.12a, b & 4.13). No previous report exists on the crystallization of C-PCs from marine cyanobacteria in monoclinic cell with two hexameric molecules in the asymenteric unit. The arrangement of hexamers in the unit cell may be similar to the association of hexamers under *in vivo* conditions.

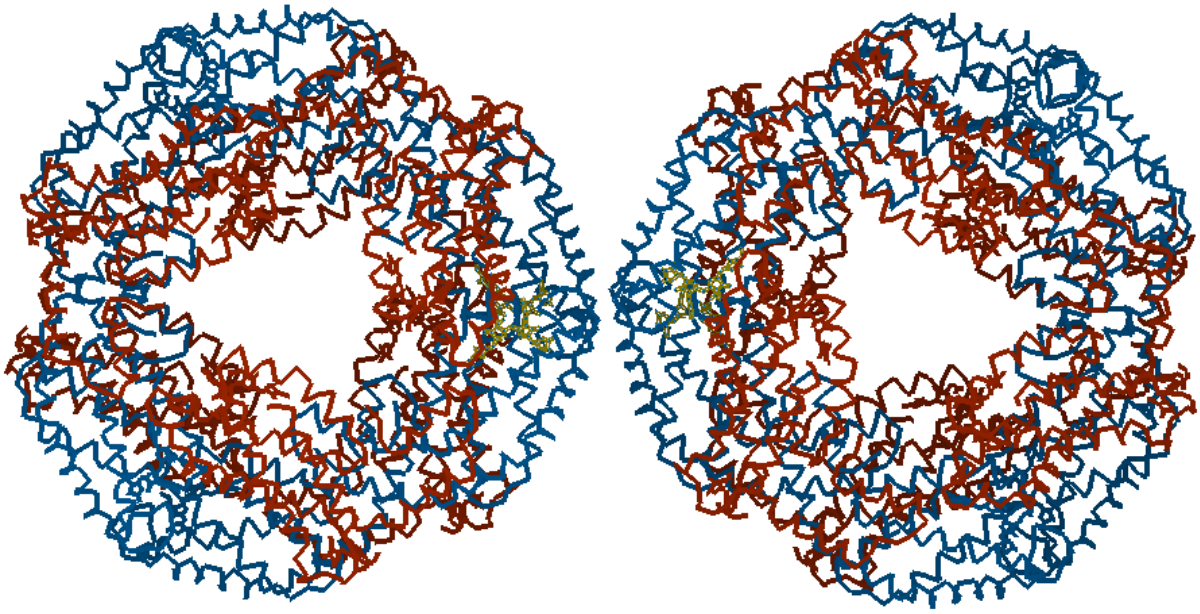


Fig. 4.12a

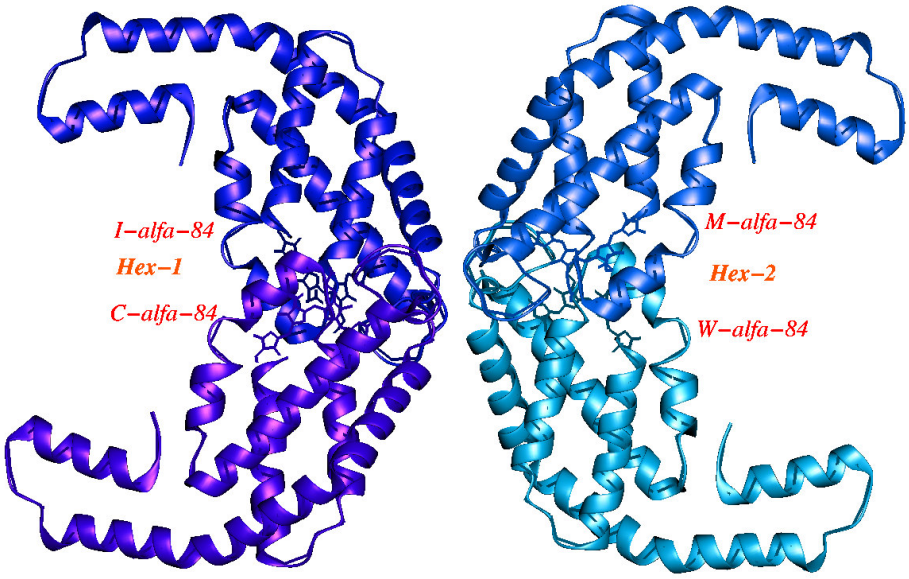


Fig. 4.12b

Fig 12: (a) The arrangement of two hexamers in the asymmetric unit of marine cyanobacteria. (b) Chromophore C-84, I-84 (Hex-1) and M-84, W-84 (Hex-2) present at the interface of hexamers in *Phormidium* C-PC (Marine)

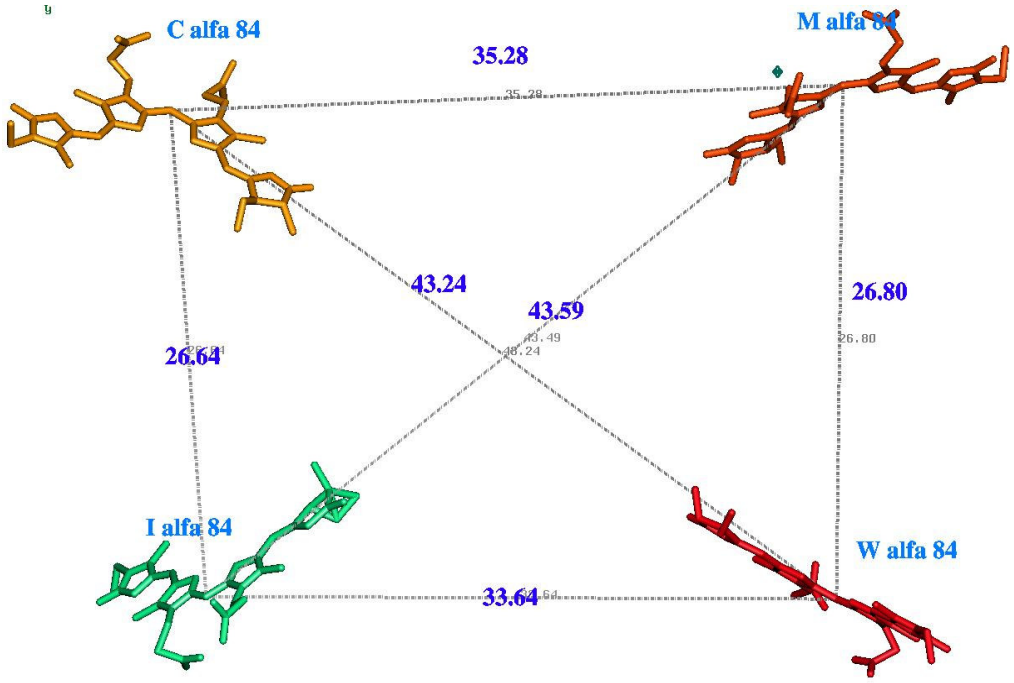


Fig. 4.13 Arrangement of chromophores C-84, I-84 (Hex-1) and M-84, W-84 (Hex-2) present at the interface of hexamers in *Phormidium* C-PC (Marine) are shown

In the asymmetric unit of C-PC structure from fresh water form (*Sprulina sp.*) the chromophores coming closer are β-155. The subunits and chromophores involved are D β-155, J β-155 of hexamer-I and N β-155, X β-155 of hexamer-II (Fig.4.14a ,b & 4.15). The distances between these chromophores are shown in Table 4.4. The shortest distance between D β-155 of hexamer–I and N β-155 of hexamer-II is 16.67Å which is almost half the shortest distance in the asymmetric unit of marine C-PCs (33.64 Å). Thus we observed apparently two types of associations in the crystal structures of C-PCs from marine and fresh water forms of cyanobacteria. The chromophore involved in the inter hexamer energy transfer in *Phormidium* and *Lyngbya spp.* is α-84where as from fresh water form (*Sprulina sp.*) is β-155

Table 4.4. Distances between chromophores of adjacent hexamers in the asymmetric unit of C-PCs from *Phormidium sp.* (Marine) and *Spirulina sp.* (Fresh water)

Marine forms			Fresh Water form	
α-Subunits involved between hexamers		Distance (Å)	β-Subunits involved between hexamers	Distance (Å)
1	C-84 and M-84	35.28	D-155 and N-155	16.67
2	I-84 and W-84	33.64	J-155 and X-155	25.85
3	C-84 and W-84	43.24	D-155 and X-155	19.31
4	I-84 and M-84	43.59	J-155 and N-155	20.27

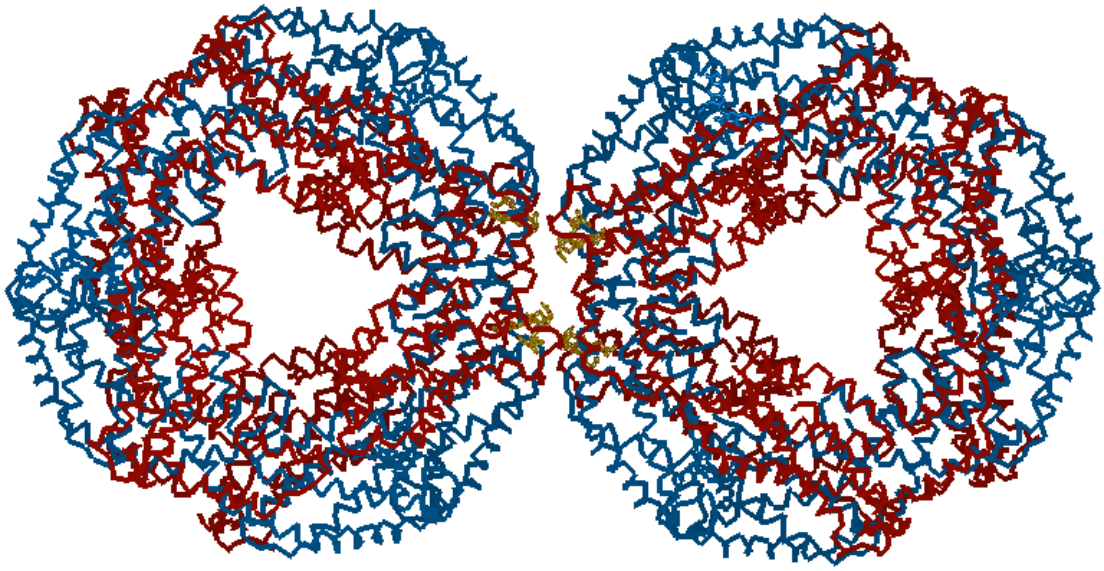


Fig. 4.14a

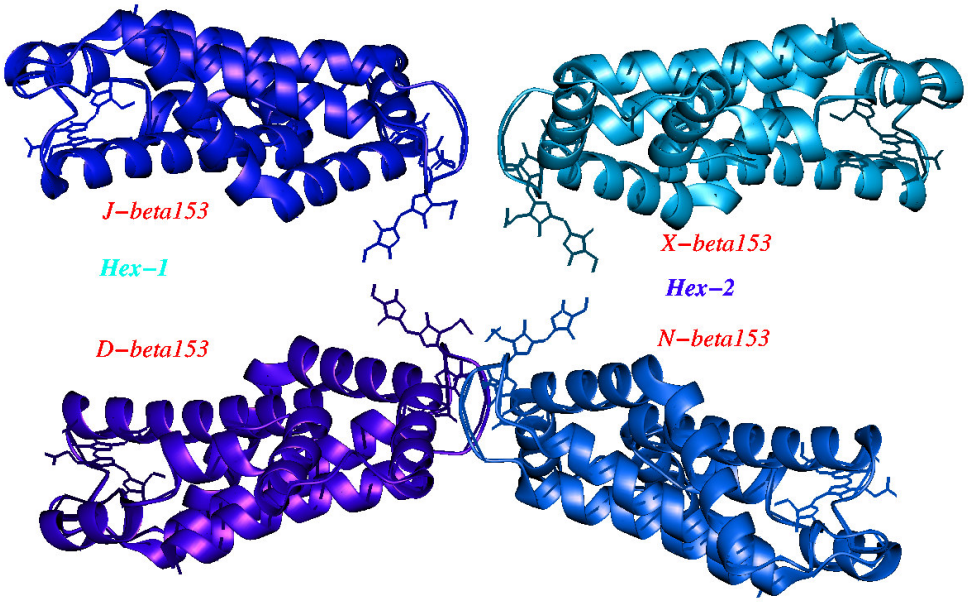


Fig. 4.14b

Fig. 14(a) The association of the two hexamers in the asymmetric unit of C-PC from fresh water form of cyanobacteria **(b) Chromophores D-155, J-155 of Hexamer-I and N-155, X-155 of Hexamer-II present at the interface of C-PC hexamers in *Spirulina sp.* (Fresh water form)**

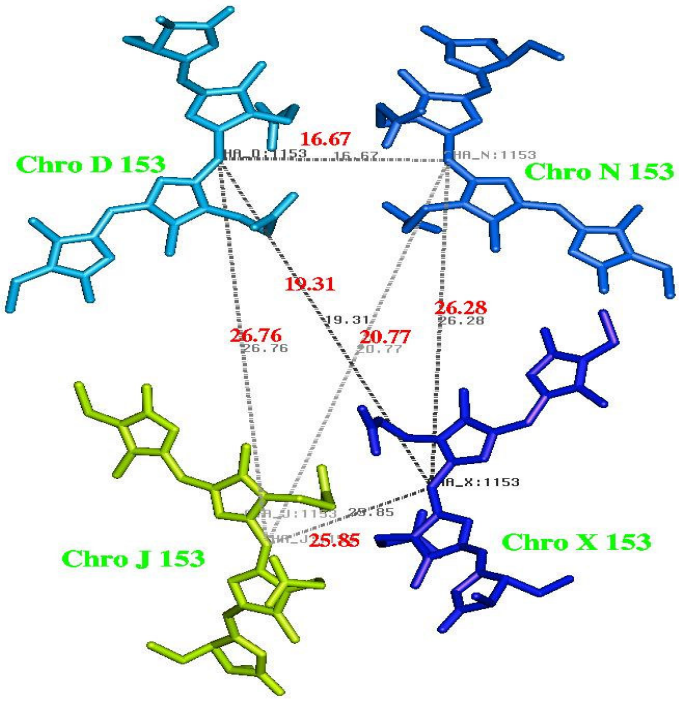


Fig. 4.15 Arrangement of chromophores D-155, J-155 (Hex-1) and N-155, X-155 (Hex-2) present at the interface of C-PC hexamers in *Spirulina sp*. (Fresh water form)

4.8.5 Comparison of the efficiency of energy harvesting reflected in the association of C-PCs from marine and fresh water forms:

Inter-chromophoric energy transfer is chiefly influenced by its conformation and distance between the interacting chromophores. The interacting chromophores in the asymmetric unit of C-PC structure from marine forms are α-84 while that in C-PCs from fresh water form is β-155. Previous studies have shown that β-155 chromophores absorb higher energy (λ_{max} = 596-600 nm) than α-84 (λ_{max} = 618-624 nm) while β-84 is the lowest energy absorber (λ_{max} = 622-628 nm) (MacColl, 1998). Energy migration through the chromophores is polarized with the flow direction being from high to low energy absorbing forms (β-155 → α-84 → β-84).

The interacting chromophores responsible for energy transfer between the hexamers in the asymmetric unit of C-PC from *Phormidium* and *Lyngbya spp. (*marine forms) are C-84, I-84 (Hex-I) and M-84, W-84 (Hex-II); their distances of separation are given in Table 4.4. The chromophores involved in energy transfer between the hexamers in the asymmetric unit of C-PC from *Spirulina* (fresh-water form) are D-155, J-155 (Hex-I) and N-155, X-155 (Hex-II) whose distances of separation are as given in Table 4.4. Clearly the separation between the critical chromophores in asymmetric unit of C-PC from fresh water form is almost half that of the marine water form. The proximity of the interacting chromophores should contribute towards efficiency of energy transfer as postulated by Foster theory (Forster, 1948, 1965, 1967). Apparently, we have observed a difference in growth periods of these cyanobacteria, shorter growth periods seen for *Spirulina sp.* (7-10 days) as compared to extended growth periods (15-20days) for *Phormidium sp. and Lyngbia sp*. If we consider the asymmetric units of C-PC from these organisms it will look as though C-PCs in freshwater forms like *Spirulina sp.* have better energy transfer efficiency than their homologues in marine forms like *Phormidium sp.* and *Lyngbia sp.* However, a careful analysis of the packing of these proteins revealed that, the apparent differences we observed in the asymmetric units were a mere coincidence and could be attributed to the selection of origin of the unit cell and asymmetric unit. Otherwise, both type of interactions, found in fresh water form and marine forms, were present in all the crystal structures as could be observed from the representative packing diagram shown in Fig. 4.16. Thus the assembly of C-PC observed in crystallographic studies cannot explain the growth difference between these organisms. However, the *in vivo* effect of a possible disulphide bridge formation due to mutation of interface residue α-68 in marine forms needs further investigation through mutational studies.

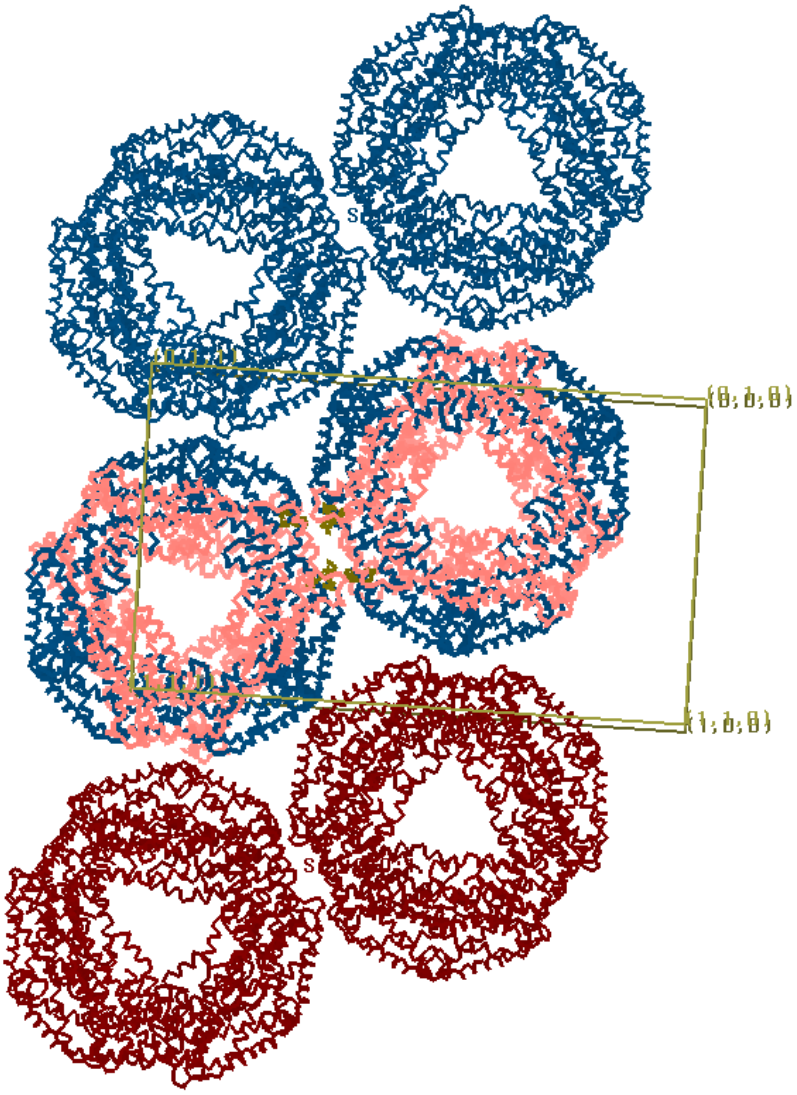


Fig. 4.16 Packing of C-PC molecules in monoclinic crystal forms. The two types contacts represented by asymmetric units of marine forms and fresh water form of cyanobacteria are marked.

.

Chapter-5

Crystallization, active site characterization and stability studies of AT*B*XYL-C from alkalophilic *Bacillus* sp.

5.0. INTRODUCTION

Xylanase, the major component of an enzymatic consortium, acts in nature by depolymerizing xylan molecules into monomeric pentosan units that are used by bacterial and fungal populations as a primary carbon source. Cellulase–free xylanases are a topic of considerable interest worldwide, following the realization of their potential impact in paper industry applications (Viikari *et al*., 1994; Bajpai *et al*., 1994). Search for xylanases that are stable and active at highly alkaline pH conditions have intensified in keeping with the requirements of the pulping operations, which are carried out at relatively high pH and temperatures (Zamost *et al*., 1991). Thermostable xylanases could be useful in such processes, in order to obtain higher reaction rates and greater solubility of the reactants. One of the ways to identify the industrially suitable enzymes stable and active at high temperature preparations is to look for the enzymes from extremophilic microorganisms (Gross, 1996). *Bacilli* are used in the industry for the production of enzymes of commercial significance. The enzymes are normally secreted into the extracellular medium facilitating their recovery. Alkalophilic and thermophilic *Bacillus* sp. (NCL 86-6-10) has been isolated in our laboratory and the potential application of the enzyme in the paper and pulp industry has been reported (Balakrishnan. *et al*., 1992). The organism produces two xylanases and the purification and characterization of one of them, we have termed it as alkaline thermophilic *Bacillus* xylanase-C (AT*B*XYL-C), has been reported (Balakrishnan *et al.*, 2001).

Studies on the structure-function relationship have always been one of the central issues in the investigations of biological macromolecules. Attempts to correlate structural changes in proteins to alterations in their biological activities have been hitherto largely confined to the modification of the side-chain functional groups. Chemical modification using group specific reagents is one of the approaches to study the role of amino acid residues present at the active site (Eyzaguirre *et al*., 1986; Means & Feeney, 1971). The utility of chemical modification is greatly extended by its use in conjunction with X-ray crystallographic studies, which mutually supplement one another. Significant progress has been made in the recent years in the understanding of the sequence of events that take place at the catalytic center of xylanases. The initial drive originated from the observation

that xylanases not only share amino acid similarities with lysozyme (Morosoli *et al.,* 1986) but also show a similar pattern of action (Biely *et al*., 1981). The studies on sequence similarities of xylanase, cellulase and lysozyme have supported the possibility that catalytic mechanism of these enzymes follows the same pathway. The difference in function among the glycosidases reflects the great diversity of molecular architecture of these enzymes and reinforces the need for simpler unambiguous methods to define functional residues in enzyme active sites. Despite the biotechnological importance of xylanases from extremophilic organisms, very few reports characterizing the structure-function relationship of these enzymes are documented. In the present chapter, we have attempted to correlate the kinetic and chemical modification studies delineating the presence of Arg, Trp and carboxylate containing residues in the catalytic site are described. Unfolding, inactivation and fluorescence quenching of AT*B*XYL-C were studied by thermal, chemical (guanidine hydrochloride) and pH denaturation. CD spectra were analysed at neutral and acidic pH (7.0 & 1.0).

The microenvironments of the Trp residues are also assessed by using various solute quenchers and by titration with oligosaccharides like xylobiose and xylotriose. Conformational transitions were monitored by intrinsic fluorescence and circular dichroism. The results from these experiments were analyzed with reference to the crystal structure of AT*B*XYL-C described in the next chapter.

5.1 Purification of AT*B*XYL-C

The crude preparation of alkalophilic *Bacillus sp.* (NCL-87-6-10) was reported to have two xylanases of differing molecular weights and p*I* in proportions of 1:9. The two xylanases were designated as 'A' and 'C', where 'A' represents anionic and 'C' denotes the cationic polypeptide. The molecular weight of AT*B*XYL-C was determined using mass spectrometry as 23,000 Daltons. The two xylanases were resolved by cation exchange chromatography with CM-Sephadex. The 'C' form was bound to the matrix while the 'A' form remained in the unbound fraction. The bound enzyme subsequently eluted in a KCl gradient at 0.6M concentration. The protein preparation was

homogeneous as shown by mass spectra. AT*B*XYL-C was purified approximately 12 fold with a yield of 39% and specific activity of 3030 IU/ml (Table 5.1).

Table 5.1 Purification of AT*B*XYL-C

Purification step	Xylanase form	Activity (IU)	Protein (mg)	Specific activity (IU/mg)	Recovery (%)	Fold Purification
Culture broth	A & C	46,000	192	239	100	-
AS fractions	A & C	40,800	91.0	448	88.7	1.9
CM Sephadex	C	18,100	6.06	3030.8	39.3	12.6

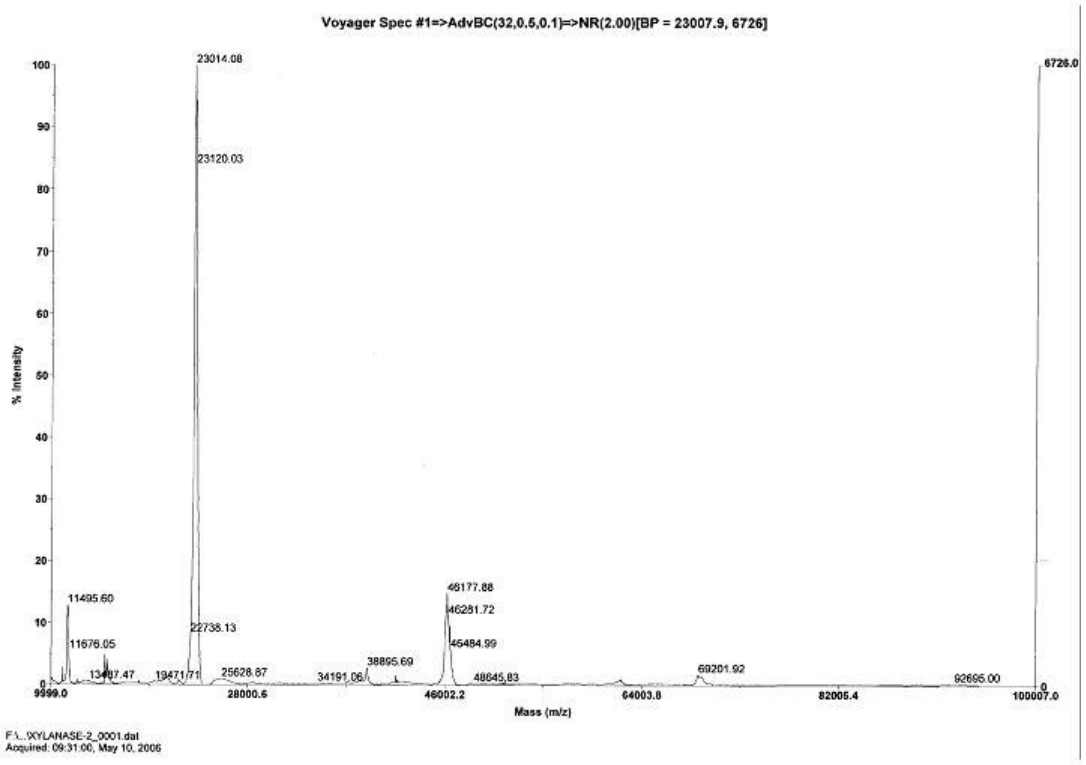


Fig. 5.1 Mass spectrum analysis of AT*B*XYL-C. Mass was determined by MALDI-TOF method. The spectra showed a signal corresponding to 23014 Da and at 46177 Da corresponding to a dimer.

5.2. Effect of various chemical modifiers on AT*B*XYL-C

The essentiality of specific functional groups for the activity of AT*B*XYL-C was investigated by the use of various chemical modifiers with narrow amino acid specificity.

5.2.1. Chemical modification of tryptophan with NBS: Treatment of AT*B*XYL-C with NBS resulted in fast inactivation of the enzyme with modification of one essential residue (Fig. 5.2). The effect of NBS on the enzyme activity of AT*B*XYL-C showed that after addition of each aliquot of NBS, there was a progressive decrease in the absorption at 280 nm with corresponding loss of activity. Modification of one Trp residue reduced the activity by 85%. The Km of the modified AT*B*XYL-C increased from 17.8 to 23.0 mg/ml indicating decrease in affinity of the modified enzyme towards the substrate presumably due to involvement of Trp residue in substrate binding (Table 5.4). About 1–2 mg of xylan protected the enzyme against inactivation (Table 5.2) thereby confirming the presence of the modified Trp residue at or near the active site. The importance of Trp residue in catalysis was further confirmed by crystal structure analysis (chapter 6).

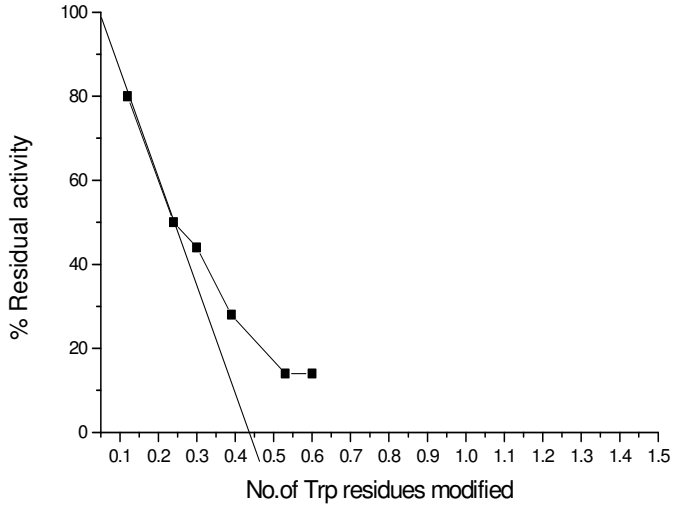


Fig. 5.2 Plot of percentage residual activity against number of Trp residues modified as determined by spectroscopic method.

Table 5.2 Substrate protection of AT*B*XYL-C against modification by NBS.

Components	Residual activity %
Enzyme + None	100
Enzyme + NBS	10
Enzyme + NBS + Xylan (1mg)	50
Enzyme + NBS + Xylan(2mg)	74

5.2.2. Chemical modification of carboxylate containing residues with Woodward's reagent K: The participation of carboxyl groups in the xylan hydrolysis was investigated by using Woodward's reagent K (WRK) at pH 6.0. The rate of inactivation of AT*B*XYL-C by WRK was biphasic with an initial fast phase and a subsequent slow phase (Fig. 5.3). The WRK treated enzyme showed almost 85% inactivation with modification of two carboxylate groups. Xylan protected the enzyme from inactivation confirming that the reactive carboxylate containing residues were at or near the active site (Table 5.3). Km of the modified protein remained unaltered, while V_{max} decreased indicating the participation of carboxylate containing residues in the catalysis (Table 5.4). Xylan (0.7mg) was able to provide 77% protection against inactivation, whereas 1mg of xylan provide 90% protection against inactivation Table 5.3. The hydrolytic activity of AT*B*XYL-C was unaffected by treatment with 5,5,-dithiobis nitrobenzoic acid (DTNB) indicating non-participation of cysteine residues in catalysis. This rules out the possibility of inactivation by WRK as due to cysteine modification.

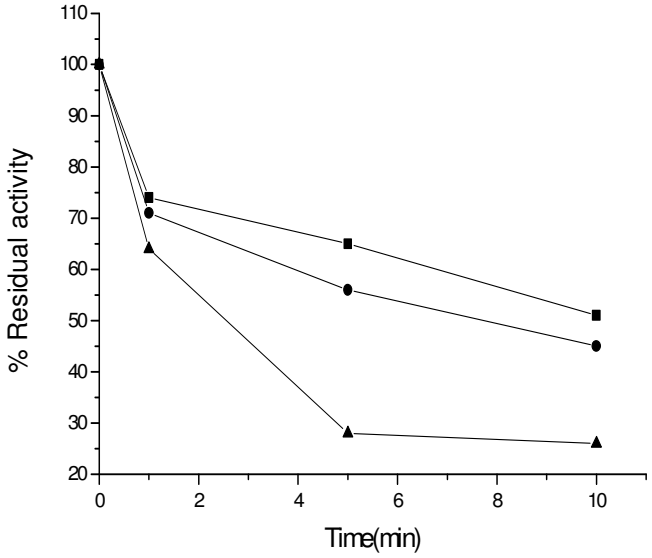


Fig. 5.3 Time course of AT*B*XYL-C inactivation with Woodward's reagent K. ■-15mM, ●-20mM & ▲-30mM WRK treated enzyme.

Table 5.3 Substrate protection of AT*B*XYL-C against modification by Woodward's reagent K.

Components	Residual activity %
Enzyme + None	100
Enzyme + WRK	60
Enzyme + WRK +Xylan(0.7mg)	77
Enzyme + WRK +Xylan(1 mg)	90

5.2.3. Chemical modification of arginine with phenylglyoxal: The purified AT*B*XYL-C on incubation with phenylglyoxal (60mM) at pH 8 for 30 min lost 40% of its activity (Fig. 5.4). No loss of activity was observed in the control samples. The log of residual activity plotted as a function of time at various phenylgyloxal concentrations was linear and showed time dependent inactivation (Fig. 5.4). The phenylgyloxal mediated inactivation followed pseudo first order kinetics at any fixed concentration of the reagent. The pseudo first order rate constant were calculated from the slopes plots of log (percent residual activity) versus reaction time (Fig. 5.4) and the order was determined from the plots of the first order rate constant against phenylglyoxal concentration. The double logarithmic

plot of slope against concentration of the reagent showed a straight line with slope one. Thus, it is concluded that one arginine residue is present at the active site. These plots (inset Fig.5.4) indicate that the loss of activity occurred as a result of modification of a single Arg residue. Km of the modified protein increased from 17.8 to 40 mg/ml indicating the role of arginine in substrate binding (Table 5.4). These results are further confirmed by the X-ray crystallographic studies.

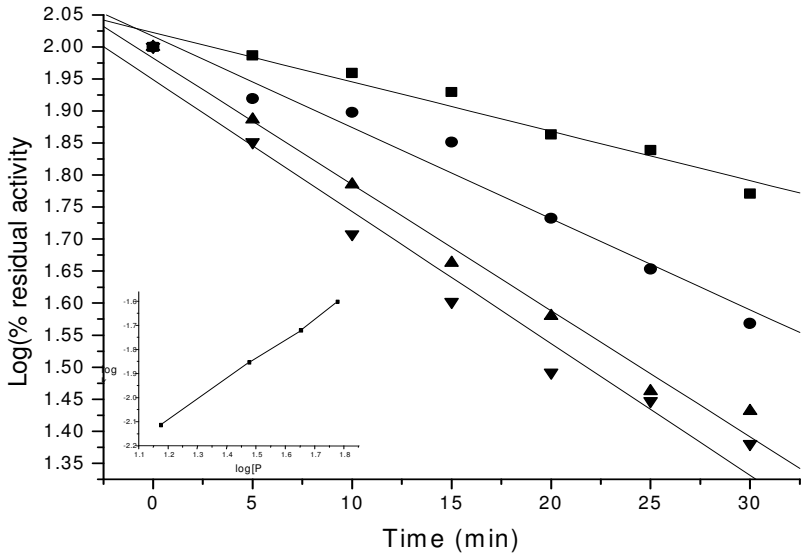


Fig. 5.4 Kinetics of inactivation of AT*B*XYL-C by phenylglyoxal. Enzyme was treated with phenylglyoxal of concentrations ■ –15, ● -30, ▲-45, and ▼-60mM. Inset shows double logarithmic plot of the first order rate constant against phenylglyoxal concentration.

Table 5.4 Kinetic constants of chemically modified AT*B*XYL-C

Percentage activity	Residue modified	Reagent used	K_m (mg/ml)	V_{max}	K_{cat}
100	None	None	17.8	1049	26225
51	Trp	NBS	23	606	15150
40	Carboxylate	WRK	17.54	874	21850
50	Arg	Phenylglyoxal	40	640	16000

5.3. Fluorescence Studies

Protein fluorescence is chiefly due to tryptophan residues. The fluorescence emission spectrum of AT*B*XYL-C showed maxima at 346nm and a shoulder at 330nm. This indicates the presence of two populations of tryptophans in AT*B*XYL-C. Tryptophan residues that are completely exposed to the solvent emit at 356nm. The signal at 330nm corresponds to tryptophan residues that are completely buried in the hydrophobic interior of the protein. The peak at 346nm is due to partially exposed tryptophans. Crystal structure of AT*B*XYL-C shows the presence of seven tryptophans, of which five (Trp 19, 87, 101, 104 and 140) are on the surface and two (Trp 45 and 166) in the interior.

5.3.1. Thermal denaturation: Fluorescence intensity of the native protein gradually decreases with increase in temperature from 30 - 60°C but there is no change in λ_{max}. However at 70°C the fluorescence intensity suddenly increases by 2.5-3 times (Fig. 5.5). Activity is observed to gradually increase as temperature increases from 30-60°C (thermal optima for activity is between 40-60°C) and starts to decrease at 70°C. The loss in activity at 70ºC and sudden increase in Trp emission indicates a significant change in conformation of the protein (Fig. 5.6). AT*B*XYL-C loses activity completely and precipitates at 80°C. Raleigh light scattering was used to monitor the formation of protein aggregate. On thermal denaturation, the protein first forms insoluble aggregate before total unfolding occurs with simultaneous loss of activity. Lowering of temperature from 80°C to 40 °C did not result in proper refolding of the polypeptide or restoration of hydrolytic activity. This indicates that the thermal denaturation is irreversible

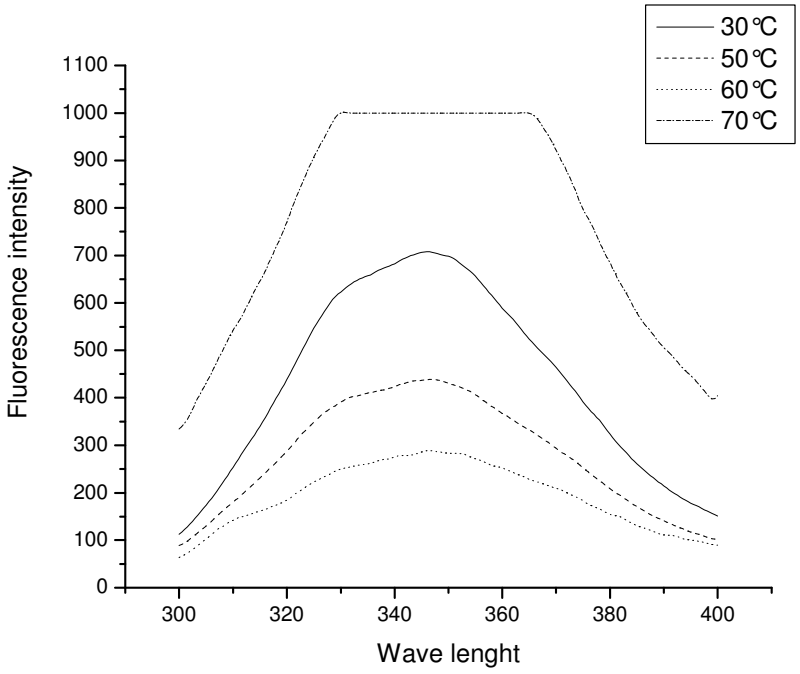


Fig. 5.5 Effect of temperature on the emission fluorescence of AT*B*XYL-C

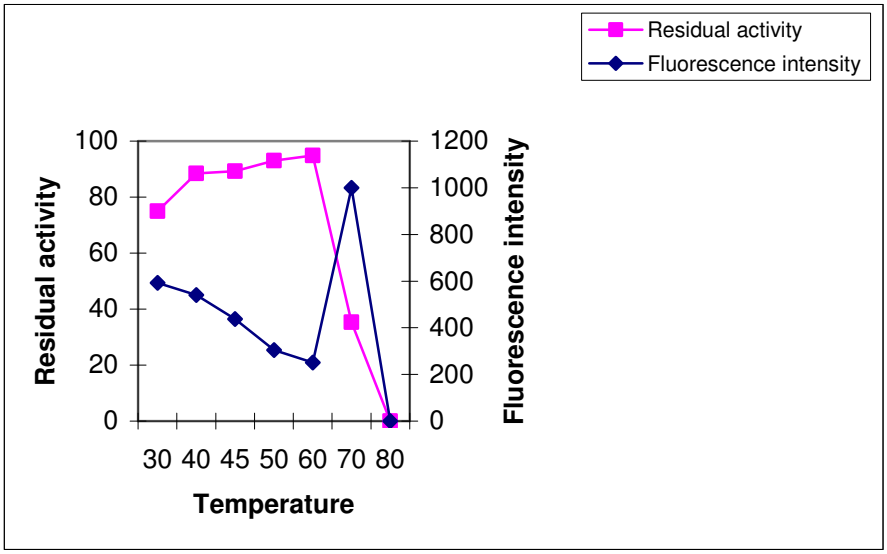


Fig. 5.6 Effect of temperature on hydrolytic activity and fluorescence emission of AT*B*XYL-C.

5.3.2 Effect of pH on activity and conformation: At extremely acidic and alkaline pHs, the fluorescence emission maxima of AT*B*XYL-C at 346nm becomes conspicuous. The shoulder at 330nm diminishes indicating reduction in the emission from buried trptophans (Fig.5.7). Maximum activity was at pH 8.0 with progressive reduction reaching negligible activity at extreme acidic and alkaline pHs.

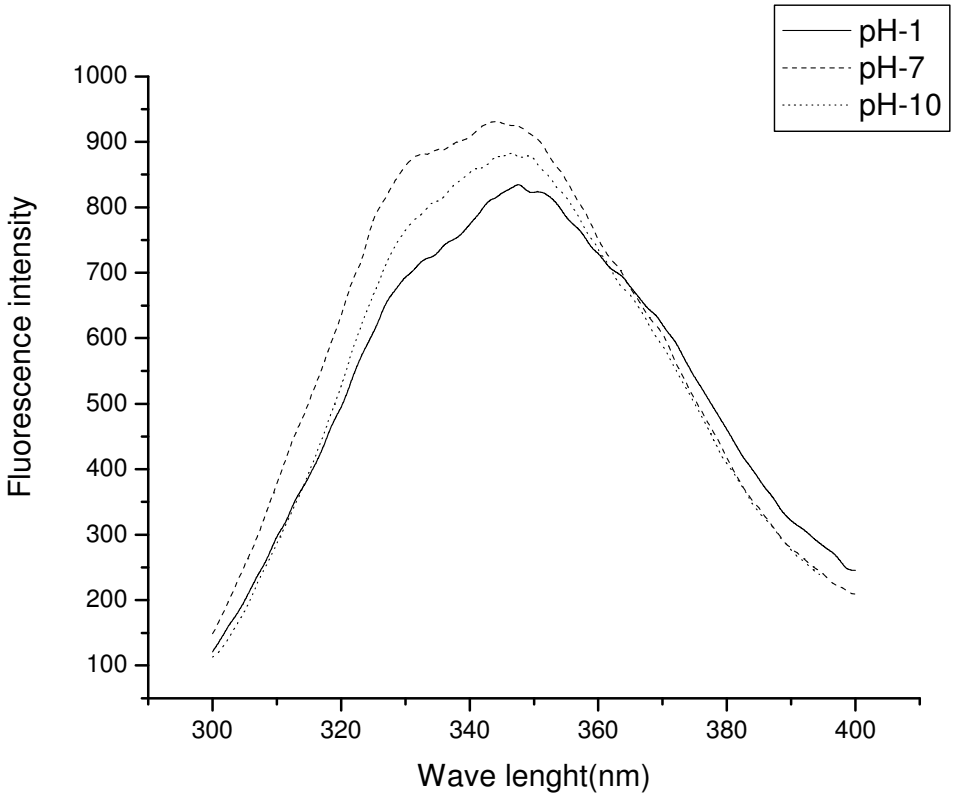


Fig. 5.7 pH dependence of conformational changes of AT*B*XYL-C

5.3.3. ANS binding studies: The hydrophobic dye ANS is widely used as a tool to identify and characterize partly folded states of proteins, including molten globule states. ANS binds to the hydrophobic regions of proteins with a dramatic increase of the fluorescence intensity and a blue shift of λ_{max} reflecting the binding of ANS molecule in a hydrophobic environment. The fluorescence intensity of ANS in presence of AT*B*XYL-C at pH 1.0 is much higher than that observed at pH 7 or pH 2.0 (Fig. 5.8). However, the ANS

fluorescence does not change in neutral and alkaline pH. At acidic pH, the λ_{max} of the fluorescence emission is shifted from 510 to 485 nm. pH dependence of ANS binding indicates that AT*B*XYL-C partially unfolds at acid pH indicating that the conformational change is mediated by the titration of Glu and/or Asp residue(s). It can be proposed that upon acid titration, intramolecular charge repulsion predominates over other intramolecular interactions thereby driving partial unfolding of the protein molecule. This induces a looser side-chain packing and causes exposure of hydrophobic groups to the solvent. There are several instances of proteins acquiring partially folded state in acidic pH.

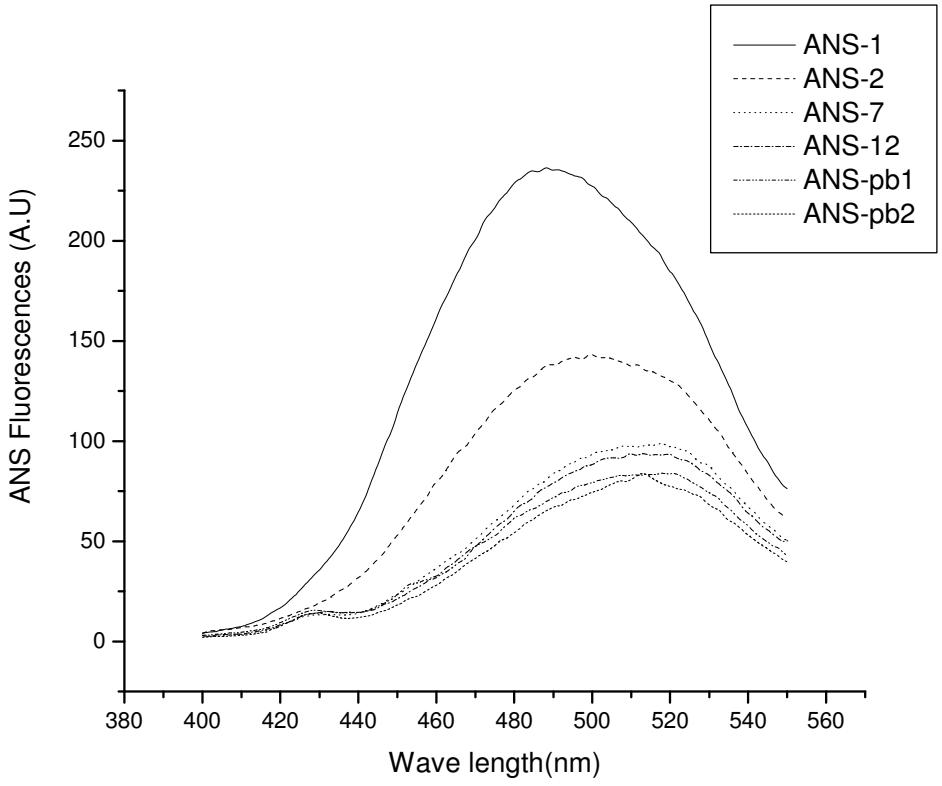


Fig. 5.8 ANS emission spectra were recorded with excitation at 375 nm and various pH. Protein concentration was 4 μM. (pb1 & 2 are phasphate buffer 1 & 2).

5.4. Circular Dichroism Spectrum

The CDPro analysis of the CD spectrum of AT*B*XYL-C showed that its secondary structure is comparable with those of previously studied family 11 xylanases. Based on crystallographic studies, we have shown (Balakrishnan *et al*., 2005) that the AT*B*XYL-C has antiparallel β-sheets, containing two domains of β sheets sandwiched to form a partially folded right hand structure with active site in the groove. The similarity in the contents of α and β structures in xylanases with other family 11 xylanases suggests similarity of its tertiary structure to that of other family 11 xylanases in the solution also. Family 10 xylanases exist as multiple domain proteins that contain eight-fold α/β barrel structure or TIM barrel structure, (Derewenda *et al*., 1994).

The far-UV CD spectrum (250–200 nm) of the native enzyme at pH 7.0 (Fig. 5.9) showed minima around 216 and shoulder at 224 nm, which is typical of a protein exhibiting β-sheets (approximately 35% for AT*B*XYL-C). CD analysis of the enzyme by CDPro at pH 1.0 and 7.0 indicated a loss in 'β-sheets and turns' and an increase in 'others'. This suggests opening of hydrophobic patches in AT*B*XYL-C at acidic pH. This is in corroboration with the enhanced ANS binding at pH 1.0. Change in the α-helical content does not seem to be associated with the change in enzyme activity. Conversely, a decrease in the 'β sheets and turns' can be correlated with decrease in AT*B*XYL-C activity noted at pH 1.0. These results suggest that in acidic pH a conformational change of the protein occurs, leading to a decrease in 'β sheets and turns' and an increase in 'other structures', resulting in the opening of the hydrophobic patches. The spectrum also shows a prominent shoulder at 224 nm revealing the α helical content (≈6%) of the enzyme. The mean molar residue ellipticities, [θ] plotted against wavelength (250-200 nm) at pH 1.0 and 7.0 are shown in figure 5.9.

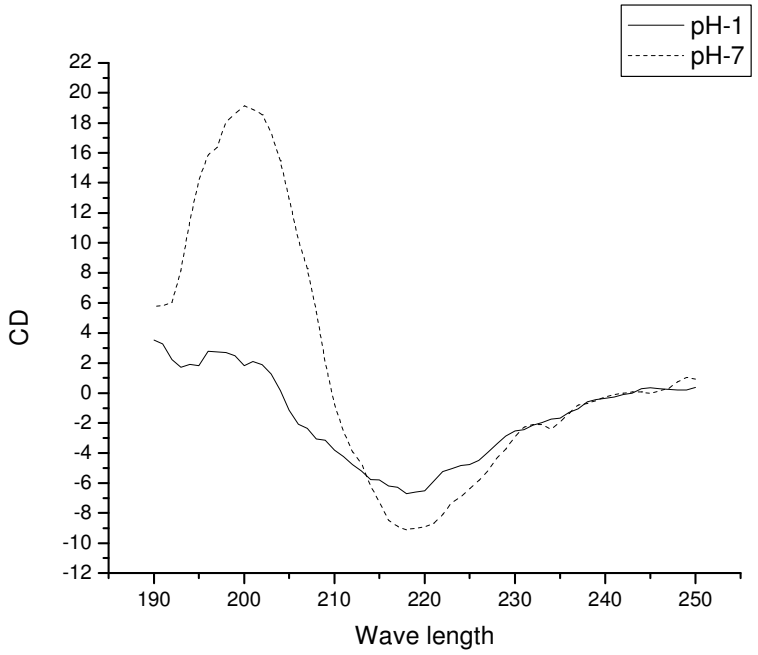


Fig. 5.9 CD spectra of AT*B*XYL-C. Far-UV CD spectra at pH 1 and 7 were recorded in the wavelengths 250-200 nm at 25 °C. The spectrum represents the average of six scans.

In the near-UV CD spectrum of the enzyme the negative band in the 270–290 nm range at pH 1.0 is far more prominent than that at pH 7.0 (Fig. 5.10). This is indicative of the totally different environments of aromatic residues at pH 1.0 and 7.0. At pH 1.0 a small change in the ellipticity occurs for the negative peak at 278 nm, suggesting an alteration in the microenvironment of aromatic residues. Finally there is a change in the tertiary structure of AT*B*XYL-C when pH decreases from 7.0 to 1.0.

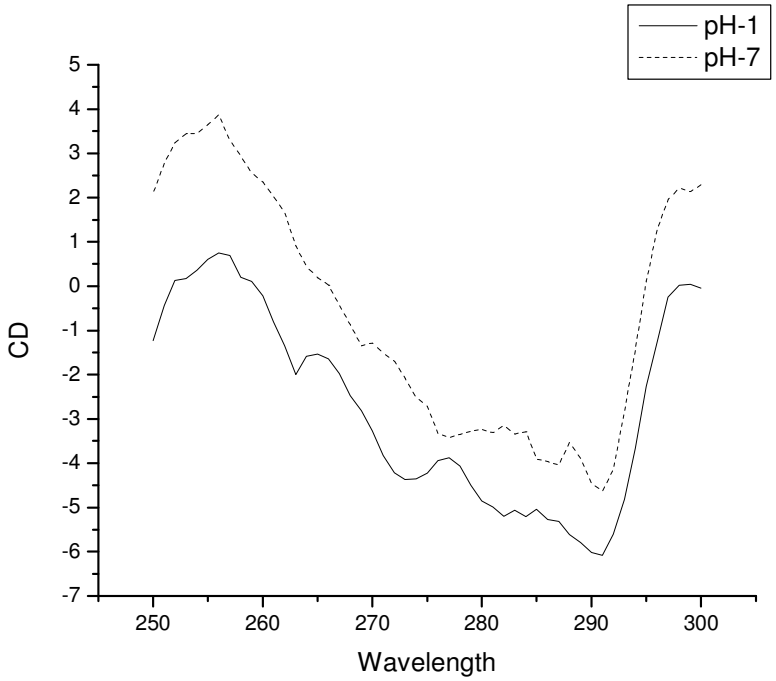


Fig. 5.10 CD spectra of native AT*B*XYL-C. Near-UV CD spectra at pH 1 and 7 were recorded in the range 300-250 nm at 25 ℃. Each spectrum represents the average of five scans.

5.5. Effect of Guanidine hydrochloride

The fluorescence intensity increases in the presence of Gdn-HCl up to 2 M concentrations and thereafter decreases sharply by 2.5-3 times at 6 M. No shift in λ_{max} was observed at concentrations 1, 2 and 3 M. However, decrease in intensity was observed at 4, 5, and 6 M with a red shift of λ_{max} by 3 to 4 nm indicating unfolding of protein at higher Gdn-HCl concentration. There was no residual activity at 1 M concentration of Gdn-HCl. The spectra of AT*B*XYL-C in the absence and presence of Gdn-HCl are shown in the Fig 5.11.

Inactivation of the enzyme at lower concentration of Gdn-HCl could be due to the binding of electropositive guanidinium to Asp and Glu residues at the active site. Previously described chemical modification studies have revealed the presence of carboxylate containing residues at the active site of AT*B*XYL-C.

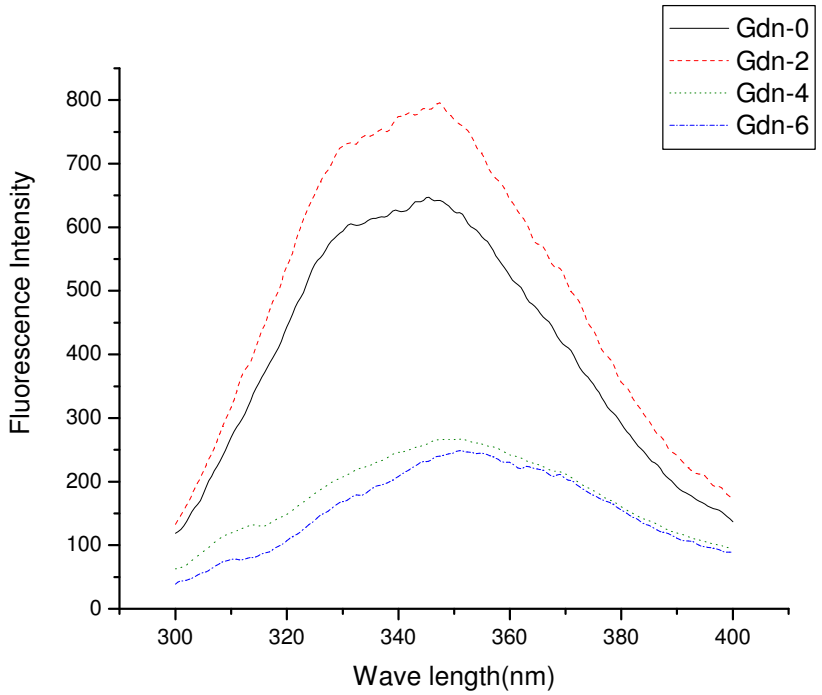


Fig. 5.11 Effect of Gdn-HCl on the tryptophanyl fluorescence of AT*B*XYL-C

5.6. Fluorescence quenching by acrylamide, KI and CsCl

The fluorescence characteristics of tryptophan residues depend strongly on the microenvironment and thus provide a sensitive probe for investigating the conformational state of the protein. Fluorescence quenching of the proteins in solutions have been widely used for studying the degree of exposure and electronic environment of aromatic amino acid residues. Acrylamide is an efficient quencher of tryptophan fluorescence and can distinguish between buried and exposed side chains. In contrast, KI and CsCl are highly hydrated and charged molecules, whose quenching ability are limited to surface exposed tryptophans and also depends upon the neighboring charged groups. The titration of AT*B*XYL-C with increasing concentration of acrylamide resulted in 100% quenching of fluorescence. CsCl showed no effect while KI partially quenched the fluorescence indicating electropositive environment of Trp. The f_a values were extrapolated from a replot of the quenching data according to modified Stern-Volmer equation. The K_{sv} value for acrylamide was much higher than that for KI and CsCl since it can penetrate into the interior and is a more efficient quencher.

Table 5.6 Fluorescence quenching of AT*B*XYL-C by acrylamide, KI and CsCl.

	Acrylamide	KI
K_{sv}	14.02	4
f_a	1	0.25
Quenching efficiency	100%	25%

KI and CsCl have been suggested as companion probes since they bear opposite charges and their relative quenching efficiencies should be related not only to the accessibility of the protein fluorophore but also to the net charge in the vicinity of the fluorophore. The fraction of tryptophans accessible to KI was 0.25 indicating that 25% of the tryptophans were in an electropositive environment. This may be due to the presence of Arg residues close to tryptophans.

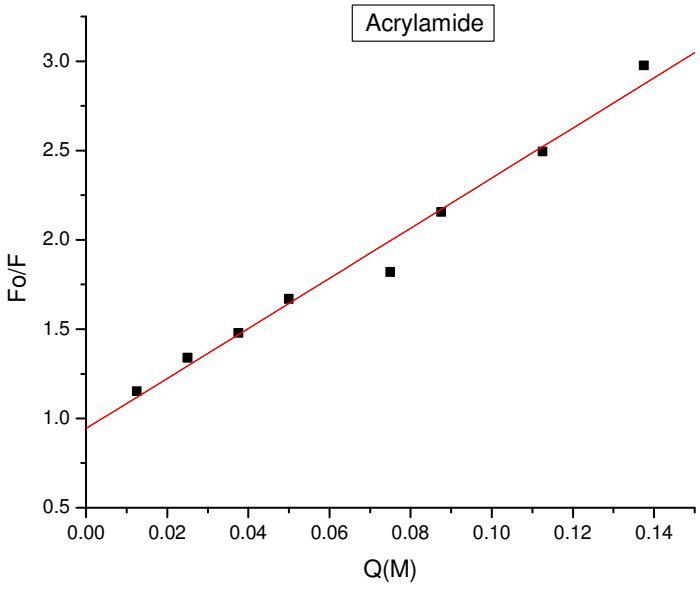


Fig. 5.12 Stern-Volmer plots for quenching of tryptophanyl fluorescence by acrylamide.

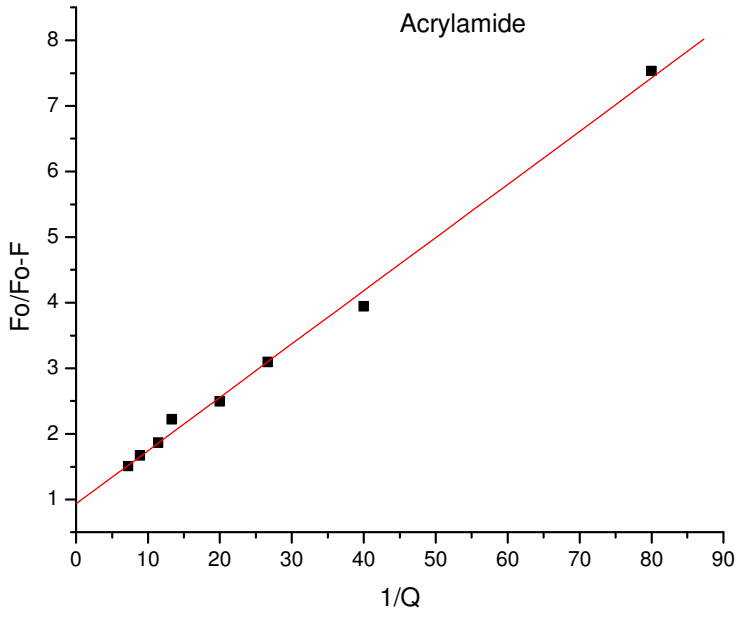


Fig. 5.13. Modified Stern-Volmer plot of AT*B*XYL-C due to acrylamide quenching

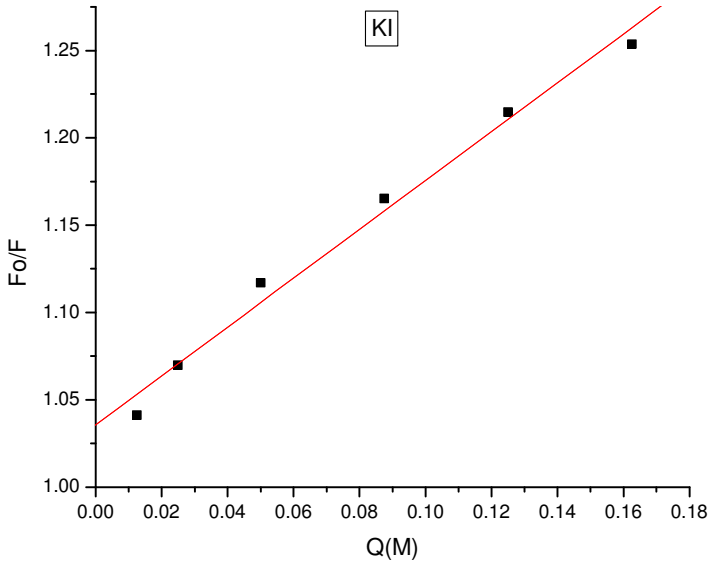


Fig. 5.14 Stern-Volmer plots of quenching of tryptophanyl fluorescence by KI

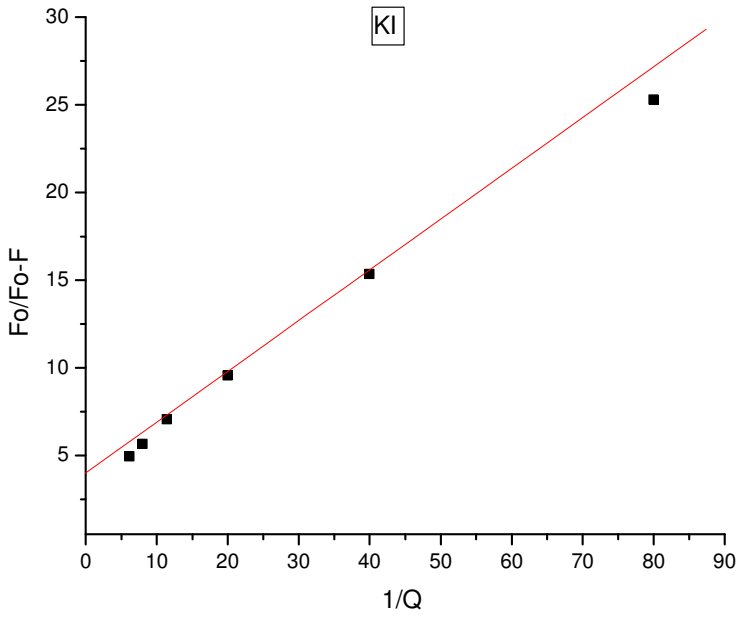


Fig. 5.15 Modified Stern-Volmer plot of AT*B*XYL-C due to KI quenching

5.7. Titration of AT*B*XYL-C with xylo oligosaccharides

5.7.1 Xylobiose and xylotriose: Addition of xylo-oligosaccharides to AT*B*XYL-C resulted in quenching of fluorescence. The maximum quenching of intrinsic fluorescence of the AT*B*XYL-C was 4% and 9% with xylobiose and xylotriose, respectively. The slope of the plot of log $\{(F_o\text{-}F)/(F\text{-}F_\infty)\}$ versus log [S] was unity for the above two saccharides used, indicating the formation of 1:1 complex between enzyme and ligand. The association constants of binding (K_a) calculated were **2.529×10^3** and **3.467×10^3** for xylobiose and xylotriose at 25ºC, respectively. This indicated that the affinity for xylo-oligosaccharides increased 1.4 times from X2 to X3. The difference in the affinities with differing chain lengths reflects the size of the binding site in AT*B*XYL-C.

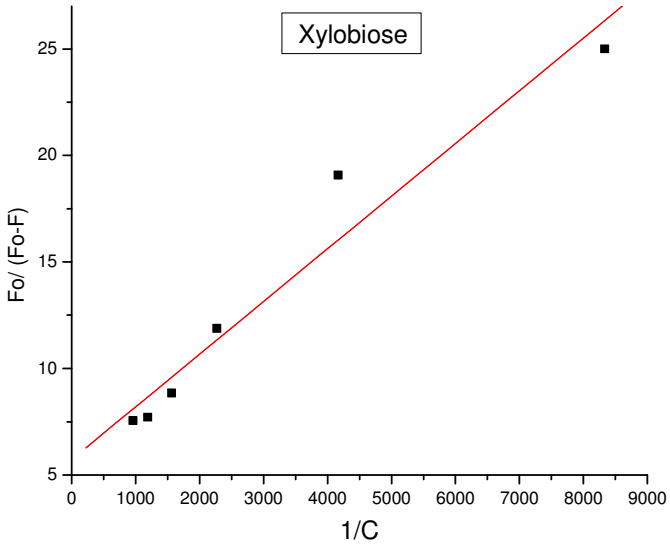


Fig. 5.16 $F_o/(F_o-F)$ vs 1/C graph for determination of F_∞ (xylobiose).

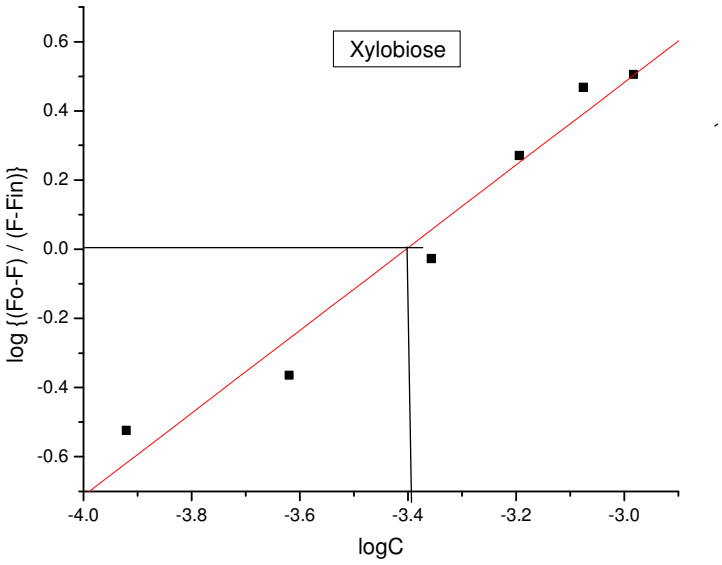


Fig. 5.17 log $\{(F_o-F)/(F_o-F_\infty)\}$ vs log C graph for determination of K_a (xylobiose).

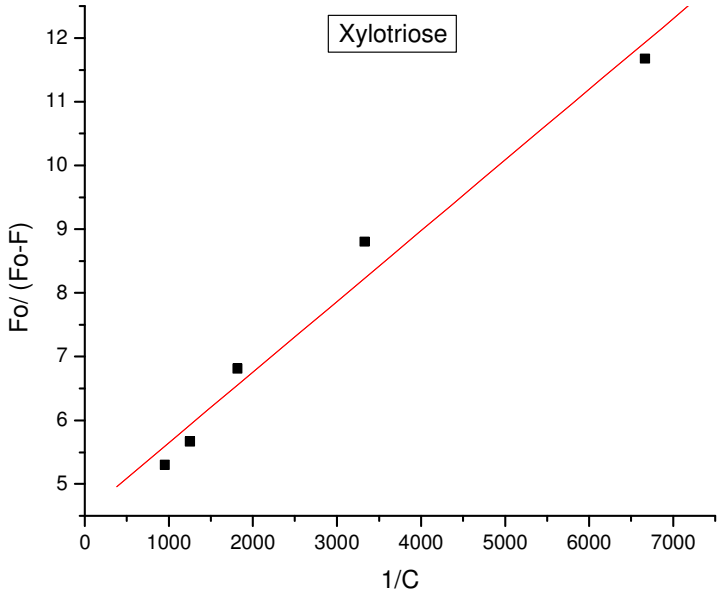


Fig. 5.18 $F_o/(F_o-F)$ vs 1/C graph for determination of F_∞ (xylotriose).

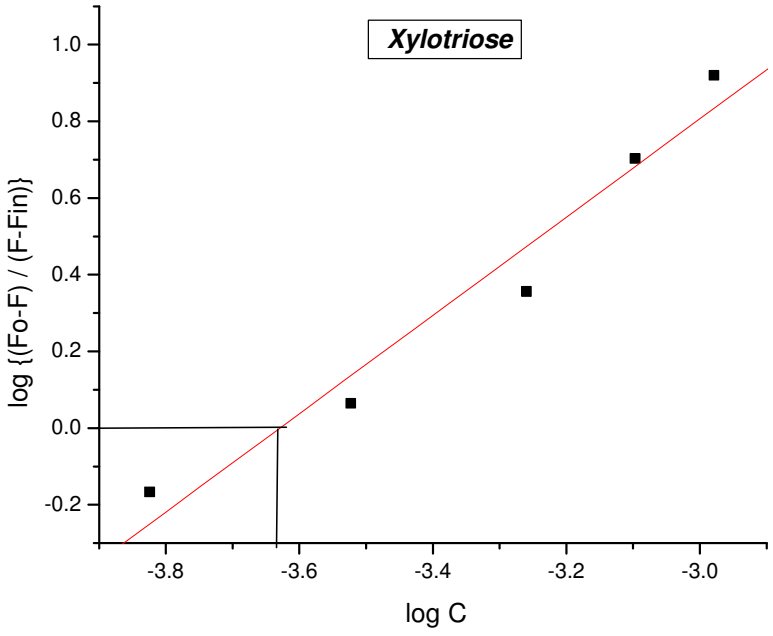


Fig. 5.19 log $\{(F_o-F)/(F_o-F_\infty)\}$ vs log C graph for determination of K_a (xylotriose).

5.8 Amino acid composition of AT*B*XYL-C

Amino acid analysis indicated the absence of cysteine residues in AT*B*XYL-C. However, one cysteine residue was detected during the crystal structure analysis of AT*B*XYL-C. The amino acid composition of AT*B*XYL-C independently determined was comparable to that of sequence obtained from crystal structure analysis except for a few missing residues in the electron density map(Table 5.5). It is also comparable to other family 11 xylanases whose sequences are known. AT*B*XYL-C has a higher proportion of positively charged residues (20 residues).

Table 5. 5 Amino acid composition of AT*B*XYL-C from crystal structure

Amino acid	No. of residues
Aspartic acid	8
Glutamic acid	7
Alanine	8
Valine	14
Leucine	9
Isoleucine	11
Proline	7
Phenylalanine	7
Metheonine	5
Glycine	24
Serine	18
Threonine	18
Tyrosine	13
Arginine	8
Histidine	3
Lysine	9
Tryptophan	7

5.9 Crystallization of Xylanase

5.9.1 Orthorhombic form: All crystallization experiments were performed using hanging-drop vapour diffusion method. Initial conditions were identified using Crystal Screen (Hampton Research, USA) at 295 K. The hanging drops set up on siliconized cover slips over 0.5 ml reservoir solutions in multi-well trays contained 1µl protein solution of 15 mg/ml in potassium phosphate buffer pH 7 (0.01M) mixed with 1µl reservoir solution. A range of conditions and concentrations of buffers and precipitants such as MPD (25–40%), sodium citrate (0.2–0.4 M) and sodium phosphate buffer (0.1–0.5 M), pH 6.0–7.5 yielded thin, long, needle like crystals (Fig. 5.20) that diffracted upto 3.5 Å resolution. Another set of conditions involving 25–40% isopropanol, 0.1–0.2M sodium citrate and 0.1M sodium cacodylate buffer, pH 6.5–7.5 were also tried. Sodium citrate was found to be an invariant factor in the crystallization process. Replacement of sodium citrate buffer by citrate phosphate buffer resulted in formation of needle like crystals. Best crystals were obtained from reservoir solutions containing isopropanol (30%), sodium citrate (0.2 M) and sodium cacodylate buffer (0.1 M) of pH 6.5 (Fig. 5.21). Crystal grew in about 15 days time and diffracted up to 2.8 Å resolution.

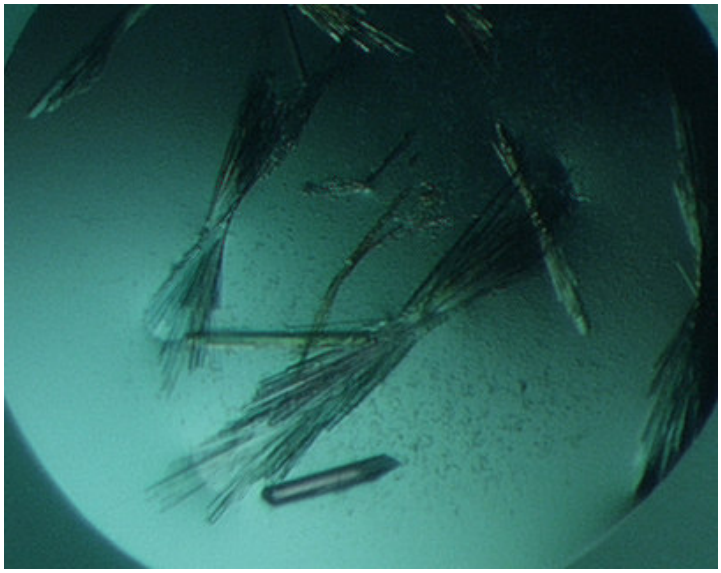


Fig. 5.20 Crystals from initial screening of AT*B*XYL-C for crystallization using MPD and sodium citrate. These crystals diffracted till 3.5Å.

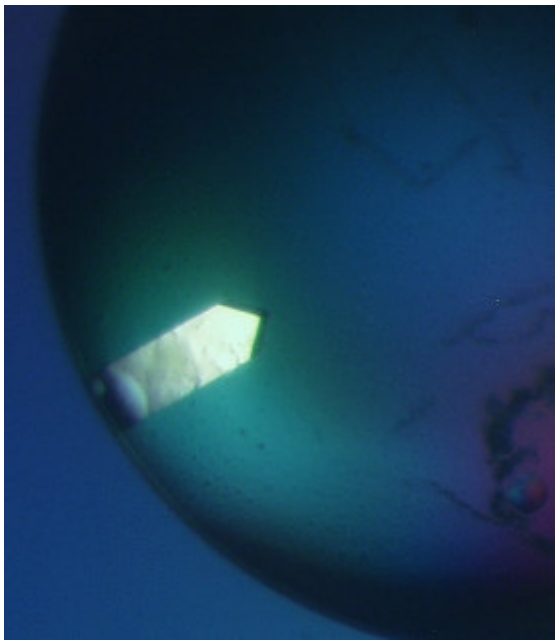


Fig. 5.21 Crystal shown here diffracted till 2.8 after improving the crystal quality with Isopropanal and sodium citrate

5.9.2 Effect of sodium and potassium salts on crystal quality: Improvement of the crystal quality and size was achieved by refining the initial crystallization condition. To obtain the xylo-oligosaccharide enzyme complex and also better resolution, the crystallization plates were setup with various precipitants, additives and different pH conditions.

Hanging drop were set-up by mixing 2µl protein solution (16mg/ml in 0.01M potassium phosphate buffer pH 7) and 2µl of reservoir solution. The composition of the reservoir solution was: 0.4M Na_2HP0_4, 10% MPD, 0.1M MES buffer pH 5.5. The resultant crystals diffracted upto 2.8Å (Fig. 5.22). Reduction of Na2HP04 from 0.4M to 0.2M and inclusion of 0.2M K_2HP0_4 in the above condition improved the diffraction quality upto 2.5Å (Fig. 5.23). The data was collected at room temperature (295K). The crystals belonged to orthorhombic space group.

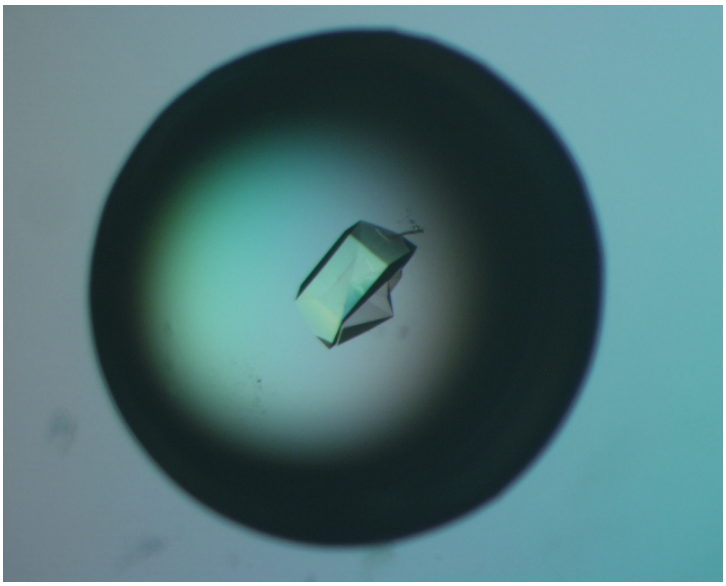


Fig. 5.22 Crystal grown using Na_2HP0_4 (0.4M), MPD 10%, MES buffer pH 5.5 (0.1M).

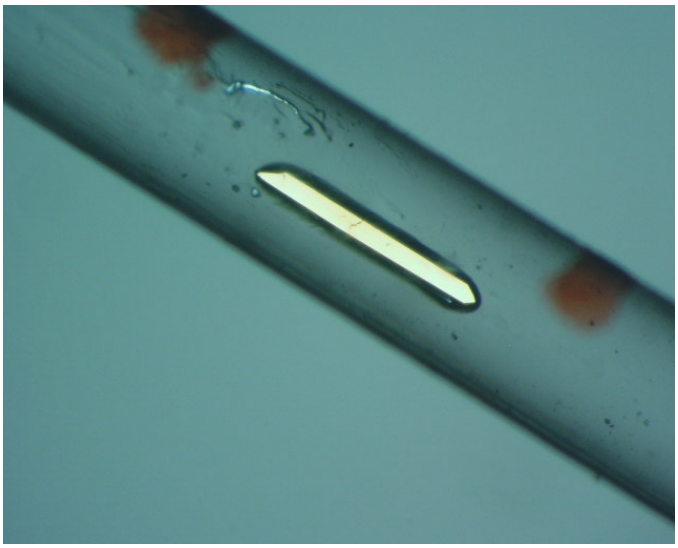


Fig. 5.23 Crystal diffracted at 2.5Å after addition Ka_2HP0_4 and halving the concentration of Na_2HP0_4.

5.9.3 Tetragonal crystals and crystalline complex with xylotriose: Crystallization plates were setup by varying the concentrations of sodium citrate and isopropanal. Crystals of two different morphologies that belonged to two different crystal systems appeared in the same hanging-drop i.e. long needle like crystals of monoclinic system and rectangular plate like crystals of tetragonal form. This was a result of increasing the concentrations of sodium citrate and isopropanol by two to three fold from the earlier condition (30% isopropanol and 0.2M sodium citrate in cacodylate buffer, 0.1 M pH 6.5) and slight decrease in pH. Tetragonal crystals were grown with isopropanol (40%), sodium citrate (0.4M) and sodium cacodylate buffer pH 5.5 (0.1M). Crystal grew in 5 days and diffracted up to 2.4 Å after addition of powdered xylotriose.

The tetragonal crystals were soaked in a reservoir solution consisting of mother liquor supplemented with xylotriose incubated for 4 days. The ligand was added so as to achieve 1:2 molar ratio of protein: ligand. Initially, the ligand was introduced as solution in milli-Q water. This resulted in crystal damage perhaps due to changes in the osmolarity of the components of the crystallization solution. Consequently, the solid ligand was directly added to the drop. The complex crystals diffracted to better than 2.4 Å resolution.

5.9.4 Data collection: X-ray diffraction data was collected at 295K on an R-AXIS IV^{++} image plate (Rigaku, Japan) using Cu Kα radiation generated by a RU-H2R (Rigaku) rotating-anode X-ray generator operated at 50 kV and 100mA and equipped with a confocal mirror focusing system (Osmic-MSI, USA).

The crystal-to-detector distance was kept at 175mm and each frame was recorded giving 1º oscillation and 180 s exposure time. The data processing was carried out on first best 110 images using DENZO and SCALEPACK (Otwinowski, 1997). Data collection statistics are listed in Table 5.7. The structure was determined using molecular replacement method (Navaza, 1994) placing 1H4G (Sabini *et al*., 1999 & 2001) as the model. There were two monomers in the asymmetric unit.

Diffraction data of tetragonal crystal complex was collected under cryogenic conditions at 100K using 20% PEG 400 as cryoprotectant .The data were processed using DENZO and SCALEPACK programs. Space group assignment was initially tried with

DENZO and SCALEPACK, but the outcome was ambiguous. It was found to be $P4_32_12$ by PHASER and AMoRe programs from CCP4 suite.

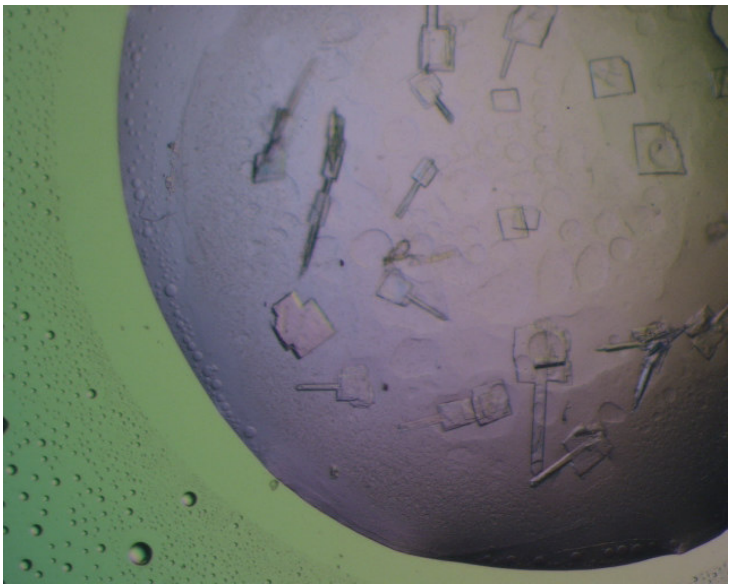


Fig. 5.24 Tetragonal crystals are formed when the concentration of isopropanal and sodium citrate of initial condition is increased.

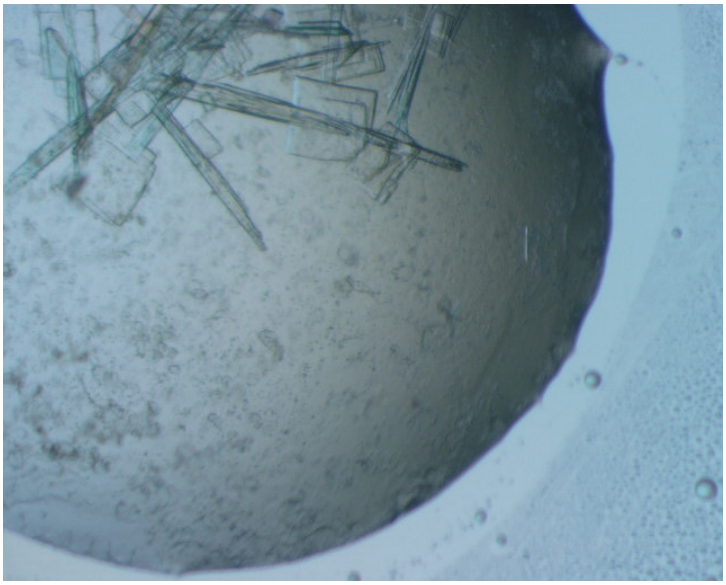


Fig. 5.25 Tetragonal crystal soaked in xylotriose diffracted upto 2.4Å resolution.

Table 5.7 Data collection statistics of orthorhombic and tetragonal forms of crystals

Temperature (K)	Orthorombic (2.8Å) 295	Orthorombic (2.5Å) 100	Tetragonal (2.4Å) 100
X-ray source	Cu Kα	Cu Kα	Cu Kα
Wave length (Å)	1.5418	1.5418	1.5418
Resolution range (Å)	20-2.8	30-2.5	30-2.4
Total No. Reflections	81516	93023	190598
No. unique reflections	10868	15981	17873
Data completeness (%)	93.9(95.9)	97.0(95.4)	99.8(98.6)
R_{merge}(%)	7.1 (23.0)	9.4 (40.1)	5.2 (16.6)
Mean I/ σ (I)	12.6	7.9	32.1
Subunits per asymmetric unit	2	2	2
Mosaicity (º)	0.3	0.4	0.3
Space group	$P2_12_12_1$	$P2_12_12_1$	$P4_32_12$
Unit Cell Dimensions (Å)	a= 73.39, b=77.28, c=79.22	a= 74.35, b=77.90, c=79.39	a=b= 78.36, c=140.83
Unit cell volume ($Å^3$)	449571	459867	864866
Matthews coefficient (V_m)($Å^3$ Da^{-1})	2.45	2.5	2.25
Solvent content (%)	49.64	50	45

Chapter 6

The three-dimensional crystal structure analysis of xylanase in two crystal forms, orthorhombic and tetragonal

6.0 INTRODUCTION

Xylanase (1, 4-β-D-xylan xylanohydrolase, E.C. 3.2.1.8) acts by depolymerizing xylan molecules into xylose units that are used by bacterial and fungal populations as primary carbon source. Xylanases produced by extremophiles are gaining importance due to their biotechnological applications and also for use as model system for structure–function studies. The paper and pulp industries make use of the substrate specificity of xylan degrading enzymes for pretreatment of paper pulp so as to enhance the bleaching effects. This reduces the quantity of bleaching chemicals required, thereby reducing toxic by-products and economic cost. The structural and functional aspects of family G/11 xylanases have been extensively studied (Torronen *et al.*, 1997). However structural characterization of very few alkaline xylanases has been reported. The three-dimensional structures of xylanases from different organisms have been determined (Katusbe, 1990; Campbell *et al*., 1993; Pickersgill, *et al*., 1993; Anna *et al*., 1997). These enzymes are shown to operate *via* a double displacement mechanism in which the anomeric configuration of the sugar molecule is retained. The family 11 xylanases are single domain proteins composed of three antiparallel β-sheets and one α-helix, with the active site between the second and third sheet. The three-dimensional structure of some of these enzymes has been described as 'partially closed right hand'. (Fig.6.7). Involvement of two Glu residues located 6–7 Å apart has been shown to participate in catalysis in xylanases from *Bacillus circulans* (Wakarchak *et al.*, 1994) and *Bacillus agaradhaerens* (Sabini *et al*., 2001). Similarly, two Tyr residues are shown to be involved in substrate binding while a Trp and an Arg are shown to participate either in stacking interaction or hydrogen bonding with the xylose ring of the substrate (Sabini *et al*., 2001; Wauters *et al*., 2001; Torronen *et al*., 1995; Krengel, 1996; Withers *et al*., 1995; Mais *et al.,* 1994). It has also been reported that family 11 xylanases with pH optima <5.0 have an aspartic acid adjacent to their general acid/base catalytic residue, while those with optima >5.0 have an asparagine in the same place. The structural characterization of xylanase active site is of great interest since it can lead to a better understanding of their catalytic mechanism and contribute significantly towards the rational design of specific oligosaccharide-binding sites. Identification and preliminary characterization of two

alkaline xylanases, xylanase 'A' and 'C', from an alkalophilic *Bacillus* (NCL 87-6-10) has been already reported [Balakrishnan *et al.*, 1992, 2002]. Of the two xylanases reported from the organism, xylanase A belonged to high molecular weight xylanase family (45,000 Da) with p*I* 5.3, and xylanase C to low molecular weight xylanase family (23,000 Da) with p*I* 8.9 (Balakrishnan *et al.*, 2002). Similarly the anionic enzyme xylanase A was reported to show a broad pH optimum for activity (6.0–10.0), while the pH optimum of the cationic enzyme xylanase C discussed here was reported as pH 8.0 at 40–60 ºC. The predominant hydrolytic product of this AT*B*XYL-C using xylan as substrate is X2 oligosaccharide. The present chapter describes the structural studies on xylanase 'C'. This enzyme is referred to as 'Alkaline Thermophilic *Bacillus* Xylanase-C' (AT*B*XYL-C) here. The structure of the enzyme was determined by x-ray crystallographic method. Details from chemical modification experiments and fluorescence spectroscopic studies support observations made from crystal structure studies. These studies provide an insight into the structural basis for the origin of specificity, mechanism of activity, alkalophilicity and thermostability.

6. 1 Structure determination using MR method

The structure of AT*B*XYL-C in both the crystal forms was determined by using a monomer as the search model (PDB: 1H4G). The data in the resolution range, 20.0-4.0Å for orthorhombic form and 20.0-3.0 Å for tetragonal form were used for MR calculations. The solutions with highest correlation coefficients (Cc) in the rotation function calculation turned out to be the correct ones in two crystal forms (Table 6.1 & 6.2). Both the orthorhombic and tetragonal forms had two unique solutions in the rotation function calculation. The translation function calculated using the solutions of rotation function, placed the molecule in correct positions in their respective unit cells. The application of rigid body fit in AMoRe program improved the solutions in all the cases (Table 6.1 & 6.2). The processed data of the tetragonal form did not have the necessary axial reflections for determining the correct space group. The correct space group including the chirality of the screw axis was confirmed through the translation function calculations using programme PHASER in CCP4 suit. The best Cc and R_{factor} were obtained when the

data were assigned the space group $P4_32_12$. The space group was independently checked using PDB:1H4G also (Table 6.3). This space group was used to solve the structure AMoRe in (CCP4i).

Table 6.1 Determination of rotation function, translation function solutions and rigid-body fit for orthorhombic crystals of AT*B*XYL-C

	α(°)	β(°)	γ(°)	Tx	T	Tz	Cc(%)*	R(%)*
RF. (Res. 20.0 4.0 Å)								
Sol-1	**90.0**	**60.9**	**7.3**	**----**	**----**	**----**	**26.0**	**--**
Sol-2	**91.2**	**123.9**	**184.8**	**----**	**----**	**----**	**21.4**	**---**
(Res.20-2.8) TF.								
Sol-1	**99.0**	**60.9**	**7.3**	**0.0232**	**0.2385**	**0.3312**	**49.4**	**48.4**
Sol-2	**91.2**	**123.9**	**184.8**	**0.0841**	**0.1767**	**0.1737**	**37.2**	**53.4**
2nd Sol.								
Sol-1	**99.0**	**61.9**	**7.3**	**0.0232**	**0.2385**	**0.3312**	**49.4**	**48.4**
Sol-2	**91.2**	**122.9**	**184.8**	**0.5825**	**0.1751**	**0.1751**	**60.8**	**43.9**
Rigid-body fit								
Sol-1	**99.7**	**61.9**	**5.3**	**0.0229**	**0.2396**	**0.3309**	**73.3**	**36.1**
Sol-2	**95.5**	**121.4**	**186.4**	**0.5865**	**0.1765**	**0.1724**		

***Cc is the correlation coefficient,**
***α, β, γ, are the Eulerian angles**

Table 6.2 Determination of rotation function, translation function solutions and rigid-body fit for tetragonal crystals of AT*B*XYL-C

RF (Res. 20 –3.0Å)	α(°)	β(°)	γ(°)	Tx	T	Tz	Cc(%)*	R(%)*
Sol-1	37.83	109.10	292.59	-----	-----	------	18.2	
Sol-2	32.35	101.12	282.74	-----	-----	------	16.0	_
Sol-3	34.00	38.86	296.49	-----	-----	------	11.6	
TF (Res. 20 –2.4 Å)								
Sol-1	37.83	109.10	292.59	0.1994	0.6828	0.3711	49.8	49.9
Sol-2	32.35	101.12	282.74	0.7041	0.4974	0.3625	47.7	49.8
Sol-3	34.00	38.86	296.49	0.459	0.0477	0.4635	31.7	55.9
2nd solution								
Sol-1	37.83	109.10	292.59	0.1994	0.6828	0.3711	49.8	49.9
Sol-2	32.35	101.12	282.74	0.2038	0.9969	0.8623	69.7	39.3
Fitting								
Sol-1	38.21	110.30	293.79	0.1989	0.6831	0.3709	76.6	35.7
Sol-2	31.65	99.51	281.55	0.2042	0.9971	0.8620		

*Cc is the correlation coefficient,
*α, β, γ, are the Eulerian angles

Table 6.3. Checking for correct space group using translation function in the case two best solutions from rotation function. (Model used was 1H4G)

	α(°) β(°) γ(°)	Tx T Y Tz	Cc(%)	R_{factor}(%)
P422				
Sol-1	37.83 109.10 292.59	0.0079 0.4879 0.3709	34.8	54.5
Sol-2	32.35 101.12 282.74	0.4148 0.5746 0.4736	34.1	55.1
$P42_12$				
Sol-1	37.83 109.10 292.59	0.6984 0.3379 0.4948	36.3	54.6
Sol-2	32.35 101.12 282.74	0.8695 0.6654 0.3625	36.1	54.5
$P4_122$				
Sol-1	37.83 109.10 292.59	0.1994 0.6820 0.2454	39.7	53.0
Sol-2	32.35 101.12 282.74	0.7042 0.4959 0.2379	38.2	53.8
$P4_12_12$				
Sol-1	37.83 109.10 292.59	0.1988 0.6830 0.1214	40.5	52.6
Sol-2	32.35 101.12 282.74	0.7047 0.4981 0.1138	39.4	53.1
$P4_222$				
Sol-1	37.83 109.10 292.59	0.0291 0.5133 0.1193	35.0	54.7
Sol-2	32.35 101.12 282.74	0.9397 0.5214 0.0200	34.5	55.1
$P4_22_12$				
Sol-1	37.83 109.10 292.59	0.1975 0.3369 0.2446	36.1	54.7
Sol-2	32.35 101.12 282.74	0.2052 0.3417 0.2368	36.0	54.8
$P4_322$				
Sol-1	37.83 109.10 292.59	0.7003 0.6825 0.4521	40.6	52.5
Sol-2	32.35 101.12 282.74	0.2035 0.4998 0.2811	39.7	53.3
$P4_32_12$				
Sol-1	37.84 109.10 292.59	0.1994 0.6828 0.3711	49.8	48.9
Sol-2	32.35 101.12 282.74	0.7041 0.4974 0.3625	47.7	49.8

6. 2 Refinement of the structures in orthorhombic and tetragonal forms

The model was determined by inputting coordinates of pdb file 1H4G and calculating the rotation matrices and translational vectors using molecular replacement calculation program AMoRe implemented in CCP4 the initial model for the structure in orthorhombic cell was determined. The refined coordinates in orthorhombic form were used for solving the structure in tetragonal form, again by molecular replacement method. During refinement of structures the diffraction data set was divided into two parts: a working set and a test set for cross validation. 5% of the data was set aside for calculation of R_{free}. The initial phases obtained from molecular replacement using the search model were improved by subsequent rigid body refinement retaining the amino acid sequence of input model. This was followed by several cycles of positional refinement using the data in the resolution shell of 20.0-2.8 (later improved to 2.5 Å when new data was collected) for orthorhombic form and 20-2.4 Å for tetragonal form. Refinement helped in improving the values of both R_{factor} and R_{free}. Overall temperature factor was refined initially and B-factors for individual atoms were refined in the final cycles. Electron density maps were calculated and side chain conformations were progressively modified and fitted into the observed density and their positions refined. The non-crystallographic symmetry (NCS) present in both the structures was used to improve the map quality. In the final refinement cycles NCS was released completely. The asymmetric unit has two identical subunits in both orthorhombic and tetragonal forms. The two polypeptide chains in the asymmetric unit were labeled A and B. The output of REFMAC refinement contained 1) the transformation matrices and the relative rotation angles for each NCS related pair of chains. 2) the number of positions and the number of atoms which differ by more than 2σ from the average position. The module X-SOLVATE of QUANTA was used interactively to add solvent molecules. The acceptance of solvent molecules was decided carefully as mentioned previously. After a few cycles of refinement those water molecules, which tend to reach the chosen limiting value 100 of B-factor, were deleted from the coordinate file. Omit maps were calculated for regions where the map was not well defined or the residues were not fitting into the density and then the correct residues

were chosen based on the shape of the calculated difference density. As the refinement progressed, difference Fourier maps (Fo-Fc) clearly revealed the position of xylo-oligosaccharide (xylobiose and xylose) bound to Arg 49 and other residues in the active site of all two subunits in the asymmetric unit of tetragonal form. After several cycles of refinement of the xylanase structures in the two crystal forms, the Cc and R_{factor} values improved, and remained with insignificant changes in the final cycles (Table 6.4). The final refined structure in orthorhombic form had 3340 protein atoms and 59 solvent atoms, whereas the tetragonal form had 3242 protein atoms, 174 solvent atoms and bound ligands.

Despite close sequence and structural similarity between AT*B*XYL-C and xylanase from *B. agaradhearens,* differences in sequence and side chain conformation were apparent in the final stages of refinement at 2.4 Å in tetragonal. The residues were eventually selected, replaced and fitted according to the observed electron density. Assuming same sequence length for xylanases from *B. agaradhearens* and *Bacillus sp. (NCL-87-6-10)*, 206 residues could be traced in the electron density map and no electron density was observed for the last 207^{th} residue (serine). In case of AT*B*XYL-C crystallized in tetragonal unit cell which is complexed with oligo-saccharide (xylotriose) electron density shows presence of xylobiose and xylose which are cleavage products of xylotriose.

During final cycles of refinement the stereochemistry and the geometry of the models were checked using PROCHECK. Ramachandran plot (Ramachandran *et al.*, 1968) showed that the residues were placed in the most favored region or allowed region of the map (Figs. 6.1 & 6.2). The statistics of the refinement is shown in Table 6.4.

Table 6.4 Refinement statistics of AT*B*XYL-C in orthorhombic and tetragonal forms

Crystal form/ structure	orthorhombic (2.5Å resolution)	orthorhombic (2.8Å resolution)	Tetragonal (2.4Å resolution)
Matthews coefficient (V_m)($Å^3$ Da^{-1})	2.5	2.3	2.25
Solvent content (%)	51	47	45.42
R value (%)	20.4	16.9	21.5
Rfree value (%)	25.4	23.0	27.5
No. water molecules	61	58	109
R.m.s. deviations from target bond lengths	0.015	0.02	0.014
R.m.s. deviations from target bond angles in distances	1.536	1.77	1.47
Residues in most favoured/ additionally allowed/ generously allowed regions (%)	86.1/13.9/0	87/ 12.4/ 0.1	87/ 13/0

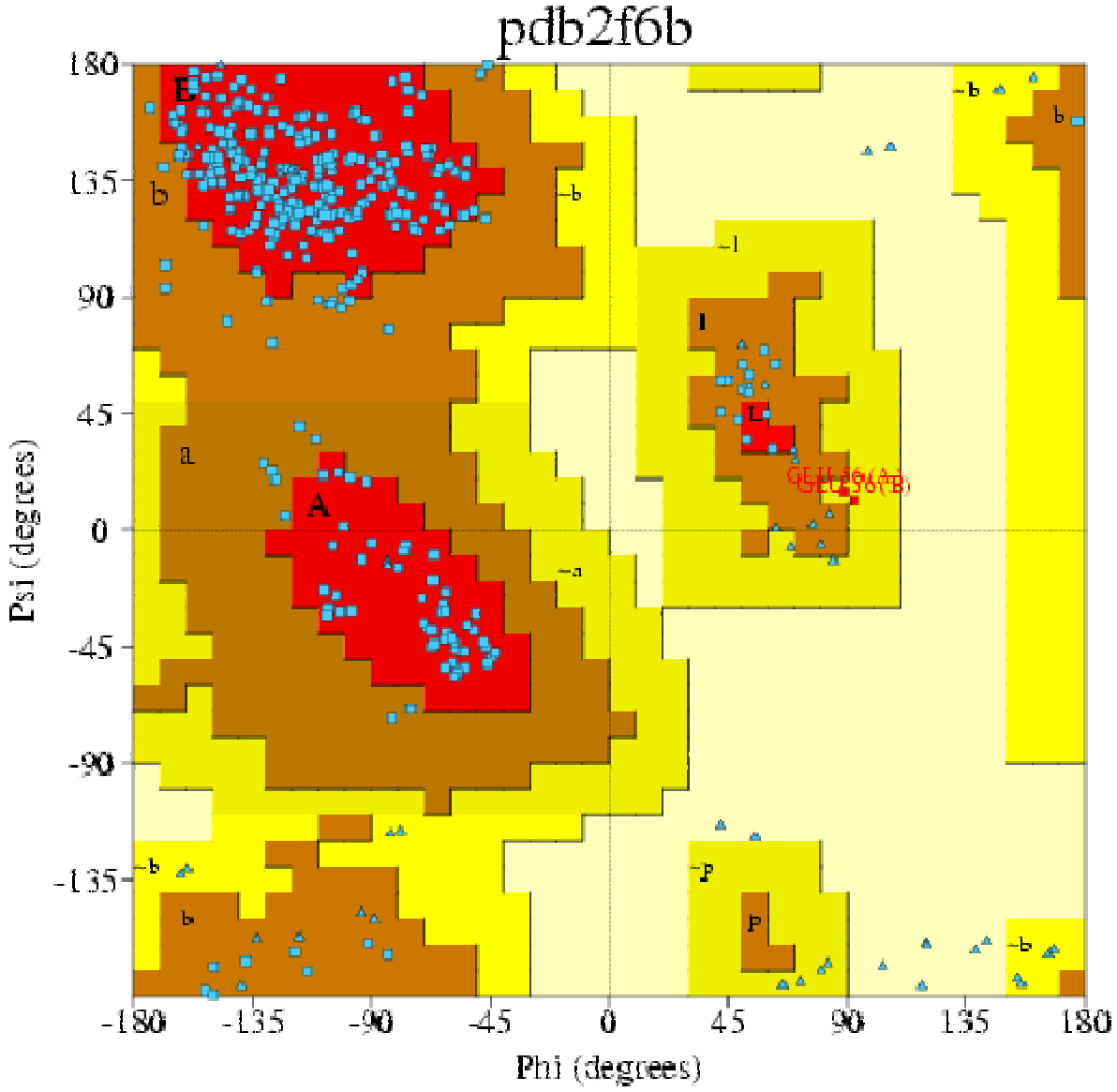


Fig. 6.1 Ramachandran plot of AT*B*XYL-C in orthorhombic form. The map was generated in PROCHECK. Triangles and squares represent glycine and non-glycine residues, respectively.

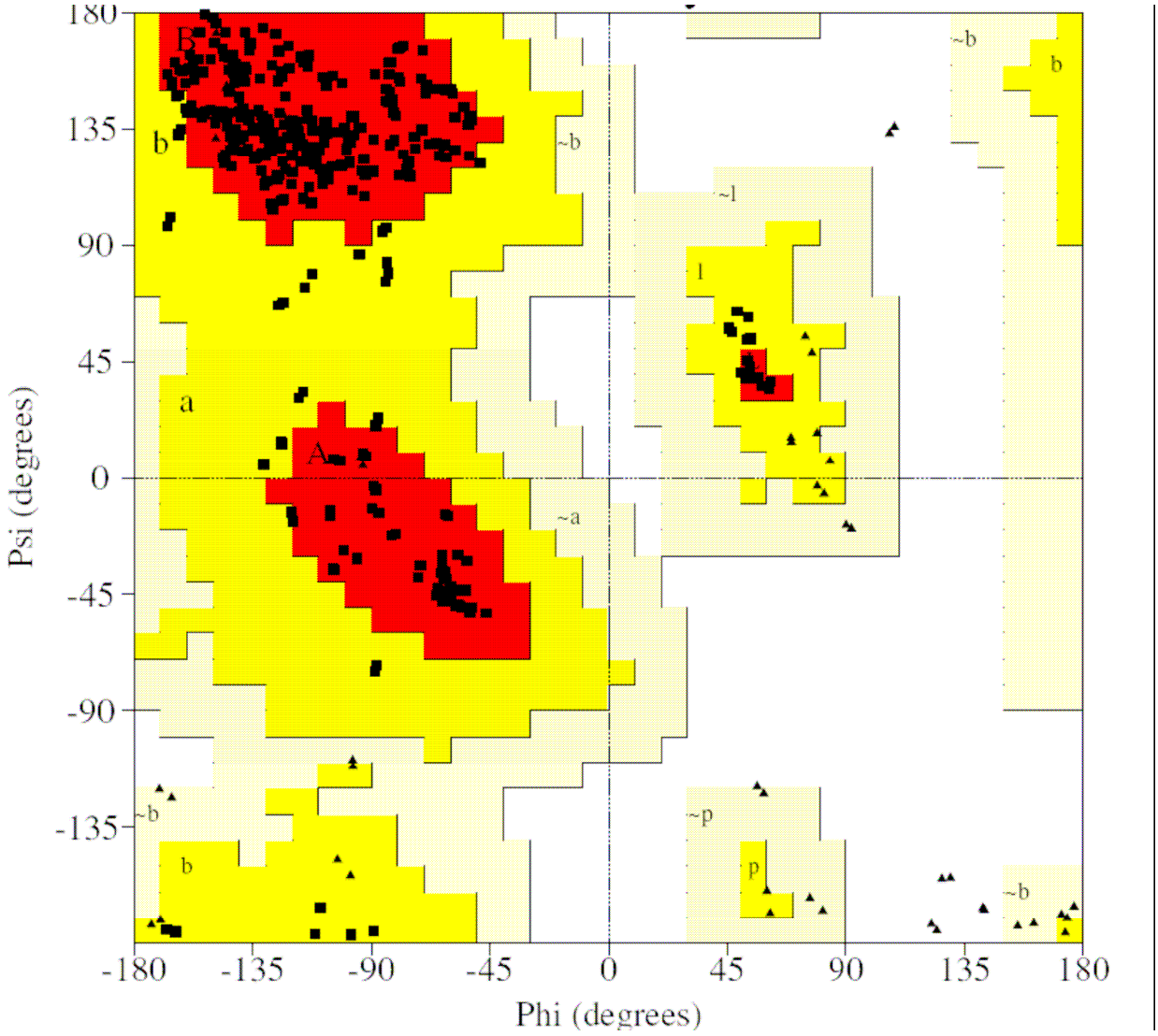


Fig. 6.2 Ramachandran plot of AT*B*XYL-C in tetragonal form. The map was generated in PROCHECK. Triangles and squares represent glycine and non-glycine residues, respectively.

6.3 Crystal structure of AT*B*XYL-C

6.3.1 Topology of secondary structure: The linear amino acid sequence of AT*B*XYL-C derived from the electron density map was aligned against the sequence of related xylanases from *T. reesei, A. niger, A. kawachii, N. flexuosa, C. thermophilum, D. thermophilum, T. lanuginosus, P. varioti, T. reesei –II, B.circulans, T. harzianum* (Fig.6.12 & Table 6.6), for which the three-dimensional structures are available. The glycine content of these proteins is slightly higher than usual and is distributed throughout the sequence rather than forming a collagen-like or any unusual structure. AT*B*XYL-C has a higher composition of positively charged residues (20 residues) than negatively charged ones (15 residues). Sequence similarity among these xylanases varies between 32 and 91% with maximum similarity existing between AT*B*XYL-C and *Bacillus agaradhaerens* xylanase.

The overall structure of xylanase has a fold (Fig.6.7) which is characteristic of family 11 endo-xylanases (Exs11) and a shape of partially folded right hand (as first described by Torronen & Rouvinen (1995) for the structure of XYNII from *T. reesei*). The two molecules in the asymmetric unit adopt similar conformation. The C^{α} atoms of the two chains superpose with an r.m.s deviation of 0.107 Å (A & B subunits in orthogonal form) and 0.073Å (A & B subunits in tetragonal form).

AT*B*XYL-C is a single-domain polypeptide containing two β-sheets (A and B). The structure contains a total of 15 β-strands, an α-helix and a 3_{10}-helix (Fig. 6.3).

The β-strands are arranged in two mostly antiparallel β-sheets designated as A and B. All β-strands have antiparallel hydrogen bonding. Both β-sheets A and B twist to form a cleft on one side of the protein. The hydrophobic faces of these β-sheets are packed against each other in a sandwich-like manner, forming the hydrophobic core of the protein. The hydrophilic face of β-sheet A forms a flat surface with many serine and threonine residues and is accessible to the solvent. The hydrophilic face of β-sheet B makes a surface for the cleft. β-sheet A is composed of six antiparallel β-strands (A1-A6). The β-strand A1 is very small being just two residues long. The short strands are

quite planar, whereas the long strands are twisted. First six strands are connected continuously from β-sheet A to β-sheet B. The β-sheet B contains β-strands B1-B9. Strands B1, B2 and B9 are short (five, seven and five residues), and B3, B4, B5 and B8 are longer. There is a twist of - 90° in the middle of the long β-strand B3, B4 B5 and B6. The twisted part of β-sheet B has been described as a separate sheet in xylanase from *B. pumilus* (Arase *et al*., 1993). The connection between strands B6 and A6 is interrupted by a long insertion. The insertion contains three β-strands: B7, B8 and B9. Strands A2, A3 and A6 are shorter (five or seven residues) than the middle strands A4 and A5 (nine to twelve residues).

α-helix is located between β-strands A6 and B4 and is packed against the hydrophobic face of β-sheet B. AT*B*XYL-C is clearly different from the other endo-xylanases in the nature of the loop-connecting strands B3 and B5 (residues 59-63). An insertion of five extra residues in this loop induces a different conformation resulting in a 3_{10}-helix. The 3_{10}-helix folds back against the body of the molecule, burying several hydrophobic residues.

The coiled regions are in general short and just provide short links (reverse turns and type II turns) between β-strands. Residue Pro91, which adopts a *cis* conformation, is involved in one of those short connections (between B5 and B6). As seen from Fig.6.12, this proline residue is conserved among the family 11 endo-xylanases. Although most of the loop regions are rather short, some longer coiled stretches are also found, in particular the so-called cord region (residues 103-112). The loop regions and turns of AT*B*XYL-C are characterized by higher temperature factors, while low B-factor regions correlate well with the secondary-structure elements, sheets and helices.

The overall structure has the shape of a 'right hand' (Fig.6.7), where the two β-sheets (A and B) form 'fingers' and the twisted part of β-sheet B and the α-helix form 'palm'. The long loop of the eight residues between β-strands B7 and B8 makes a 'thumb', which partly closes the cleft. A striking feature is a 10 residue long irregular loop between β-strands B6 and B9 forms a 'cord' 103-112 (GNWRPPGATPK), which partly closes the cleft on one side (Torronen & Rouvinen, 1995). The residues in the cord have

a clear electron density, indicating that this unit has a well-defined structure. The differences in conformations of the cord region can be attributed to the absence of hydrogen bonds between the cord region and the rest of the molecule, this short segment of the protein is stabilized by hydrophobic contacts alone.

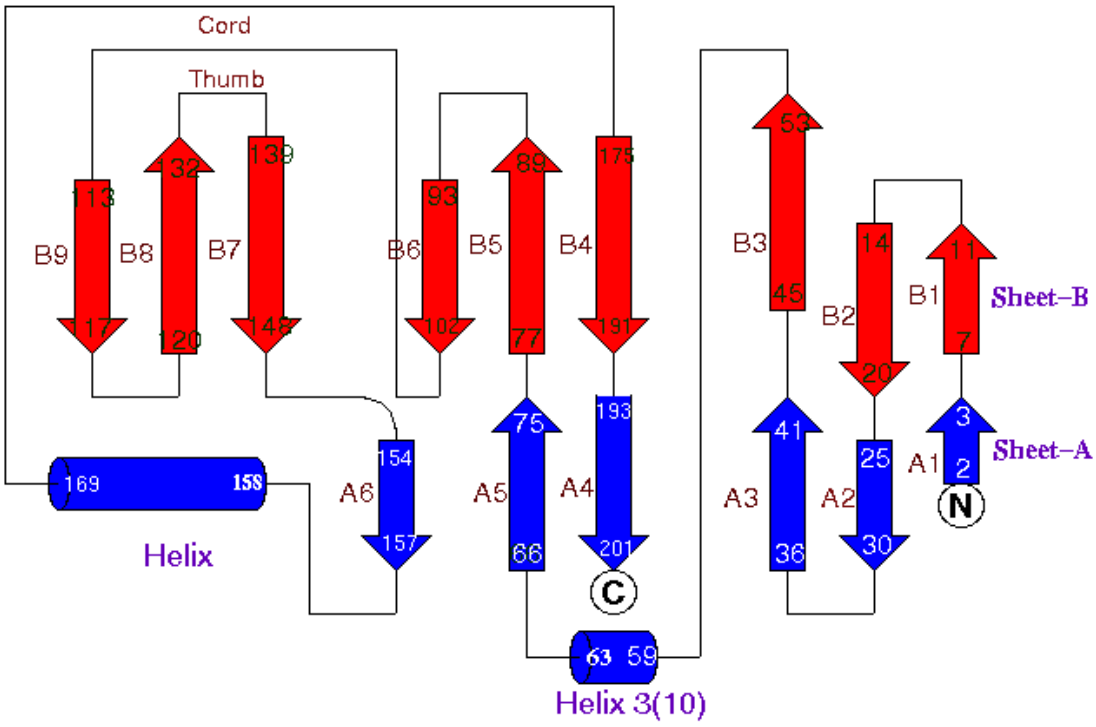


Fig. 6.3 Topological diagram of ATBXYL-C

6.3.2 N-terminal region: The topology of the elements of secondary structure of AT*B*XYL-C is compared with those of related family 11 endo-xylanases in Fig. 6.12. The major differences between these proteins is the presence (or absence) of B1 β-strand. Endo-xylanases lacking this first β-strand [e.g. proteins from *T. reseei* XYNI (1xyn), *A. niger* (1ukr), *A. kawachii* (1bk1) and *B. circulans* (1bcx)] position their N-terminal residues close to the C-terminal residue (type I; Fig.6.12), in contrast to those containing this B1 β-strand [e.g. proteins from *T. reseei* XYNII (1xyp), *T. harzianum* (1xnd), *T.*

lanuginosa (1yna), *P. varioti* (1pvx), *B. agaradhaerens* (1qh6) and *D. thermophilum* XynB (1f5j)] (type II; Fig. 6.12). Structurally, AT*B*XYL-C belongs to the type II family 11 endo-xylanases, as it contains B1. These differences contribute to structural stabilization of the family 11 xylanases. Indeed, several independent studies have shown that mutations in the N-terminal sequence increase the thermostability and thermophilicity of these proteins (Sung *et al*., 1998; Shibuya *et al*., 2000; Georis *et al*., 2000). Intramolecular hydrogen bonds stabilize the N-terminal region of AT*B*XYL-C. The pattern of hydrogen bonds involving the main-chain NH and CO groups in that region is characteristic of those observed in antiparallel β-sheets. All residues at the N-terminal have been traced in the electron-density maps. Hydrogen bonds involving side chains provide additional stabilization of the loop conformations.

6.3.3 Microenvironment of catalytic glutamates: As in other family 11 endo-xylanases, the two catalytic carboxylates in AT*B*XYL-C are provided by glutamic acid residues (Glu94 and Glu184; Fig. 6.4 & 6.6) situated at the bottom of the cleft, the distance of separation between them being 6.49 and 6.6 Å in chains A and B respectively in orthorhombic form. The distances between the catalytic acid groups is close to the value observed in the potentially active conformation of the xylanases from *B. circulans* (5.9 Å), *T. reesei* I (6.7 Å) and *T. harzianum* (6.4 Å). In *T. lanuginosus* and *T. reesei* II enzymes, where this distance is longer (10.7 and 10.9 Å, respectively), a conformational change is expected to generate a catalytically active species. We conclude therefore that the residues Glu94 and Glu184 in AT*B*XYL-C assume active conformation. These catalytic residues are also hydrogen bonded to a water molecule. In general, the structure and interactions of AT*B*XYL-C active site match that of other family 11 xylanases closely. Interactions involve residues that are conserved in other endo-xylanases, with exception that in enzymes with a low pH optimum, residue Asn45 is substituted by an aspartic acid residue (Krengel & Dijkstra, 1996). In AT*B*XYL-C, Asn45 points in the direction of active site and interacts with Glu184. It has been proposed that the guanidinium group of Arg129 is close enough to Glu94 to maintain a low pKa for this acid group. In AT*B*XYL-C, the distance between the side chain of Arg129 and Glu94 is 5.89 Å in molecule A and 5.7 Å in molecule B (orthorhombic form). This distance varies

in other members of family 11 endo-xylanases, underlining the flexibility of Arg129 side chain. Therefore, it is postulated that the binding of a substrate may induce repositioning of the Arg129 guanidinium group and thus modulate its involvement in the catalytic mechanism. A series of conserved aromatic residues line the base of the cleft. The position and orientation of these conserved residues in xylanase and in other Exs11 are quite similar. They define the substrate binding site, confirming the hypothesis that the substrate binding site of family 11 endo-xylanases is essentially preformed (Sabini *et al*., 1999; McCarthy *et al*., 2000).

6. 4 Active site exploration by chemical modification method and crystal structure analysis

Sequence alignment of various family 11 xylanases shows that residues involved in the formation of active site are conserved (Maios *et al.*, 1994). These conserved residues are identifiable in the active site of AT*B*XYL-C. Based on previously reported investigations it can be inferred that Glu94 in AT*B*XYL-C assumes the role of nucleophile while Glu184 functions as acid/base residue. Since there are no significant residue differences observed in the active site, the remarkable shift in optimum pH to alkaline range cannot be explained at this moment. The amino acid sequence of AT*B*XYL-C reported here is deduced from electron density map. The overall structure of the molecule and the distribution of amino acids in the structure targeted for chemical modification are shown in Fig.6.5 & 6.6.

It has been suggested that interaction between substrate and Trp stabilizes the transition state intermediate during hydrolysis of glycosidic bond. The stacking interaction is supposed to be the common binding interaction in glycosidases. The presence of a conserved Trp residue AT*B*XYL-C is confirmed by X-ray diffraction studies. Analysis of the three-dimensional structure of *B. circulans* xylanase has shown the presence of Trp at the active site. Several other reports also have identified Trp residue at the active site of xylanases. In AT*B*XYL-C the corresponding residue is identified as Trp19.

In *Bacillus circulans* xylanase, residues Glu78 and Glu172 have been shown to function as nucleophile and general acid/base catalyst respectively. Glu94 and Glu184 act in the same capacity in *Bacillus agaradhaerans*. In our chemical modification studies Glu94 and Glu184 would have been affected.

In the case of Arg modification by phenylglyoxal, the substrate xylan failed to protect the enzyme from inactivation. This could be due to steric hindrance towards accessing the active site by xylan once the arginine residue is modified. From the structural analysis of xylanase from *Bacillus agaradhaerens* and *B. circulans*, it is known that Arg forms hydrogen bonds with the substrate. One of the conserved arginines, Arg49 in this structure, would have reacted with the reagent.

Fig.6.6 shows some of the putative active site residues along with their corresponding electron densities. All the arginine, tryptophan, glutamic acid and aspartic acid residues present in AT*B*XYL-C are highlighted in Fig.6.5. The two active site glutamic acid residues (Glu 94 and 184), presumably involved in catalysis as nucleophile and acid/base groups, were potentially modified and responded to substrate protection by xylan, are shown in Fig. 6.6. Similarly, Trp 19 may be crucial for substrate binding. However, we do not have clear evidence on the degree of involvement of arginine residues (Arg49 or Arg129) on catalytic mechanism. An essential Arg is confirmed from kinetics of the chemically modified enzyme with phenylgyloxal. The other Arg could be the slow reacting one. Chemical modification studies could tackle neither the role of Gln143 that is hydrogen bonded to nucleophile Glu94, nor of Asn145 hydrogen bonded to acid/base residue Glu184, however, could ascertain their preponderance in deciding pH optimum of activity from studies on closely related structures and from their corresponding disposition in the present structure. As reasoned in another report (Poon David *et al.*, 2003) the higher thermostability of alkalophilic *Bacillus* xylanase may be indirectly responsible for higher pH optimum reported by Balakrishnan *et al*., (1992).

In conclusion, chemical modification studies in conjunction with three-dimensional structure determination X-ray diffraction method have helped to delineate the chemical architecture of the active site.

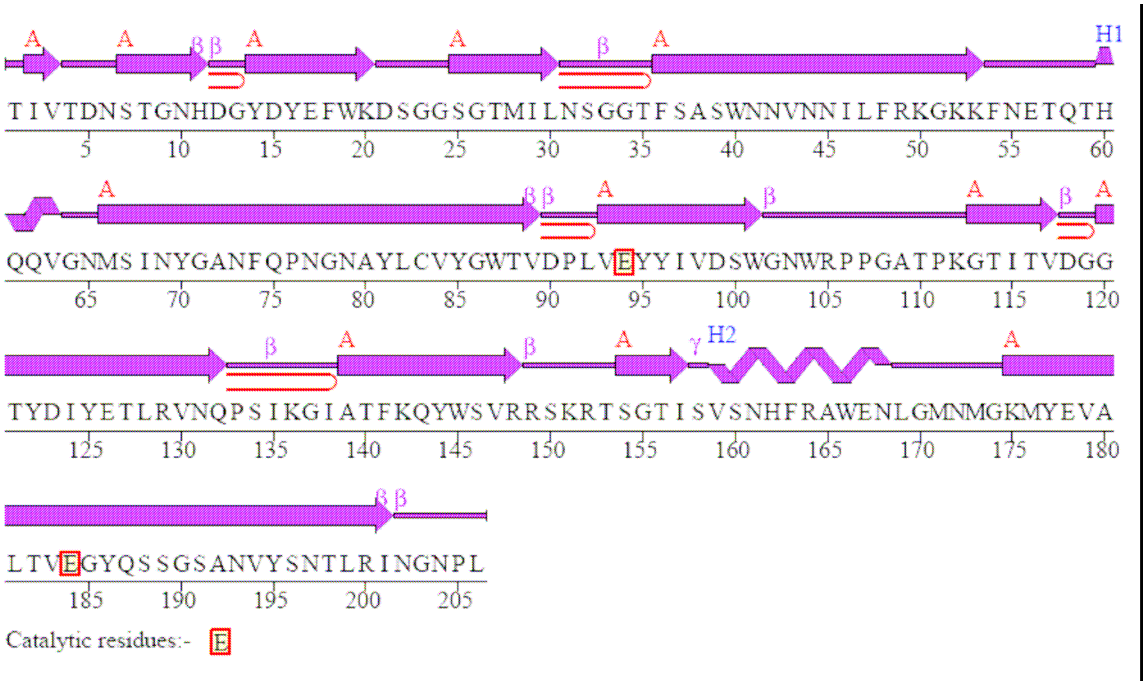


Fig. 6.4 Sequence of AT*B*XYL-C showing secondary structural elements in orthorhombic form

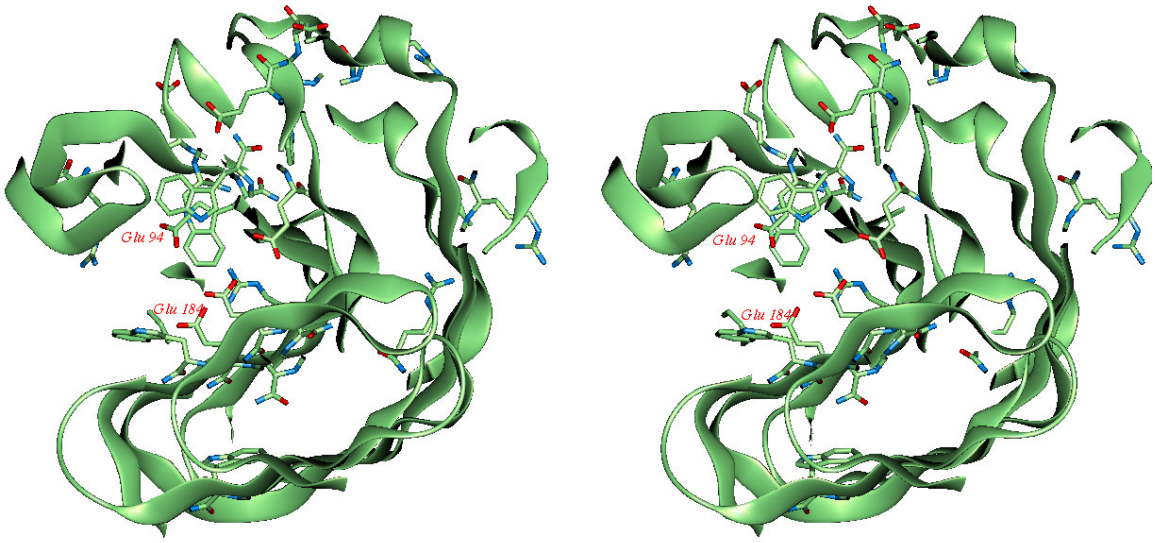


Fig. 6.5 Overall structure of monomer molecule is shown with the distribution of residues studied (Glu, Asp, Arg, Trp) by chemical modification method.

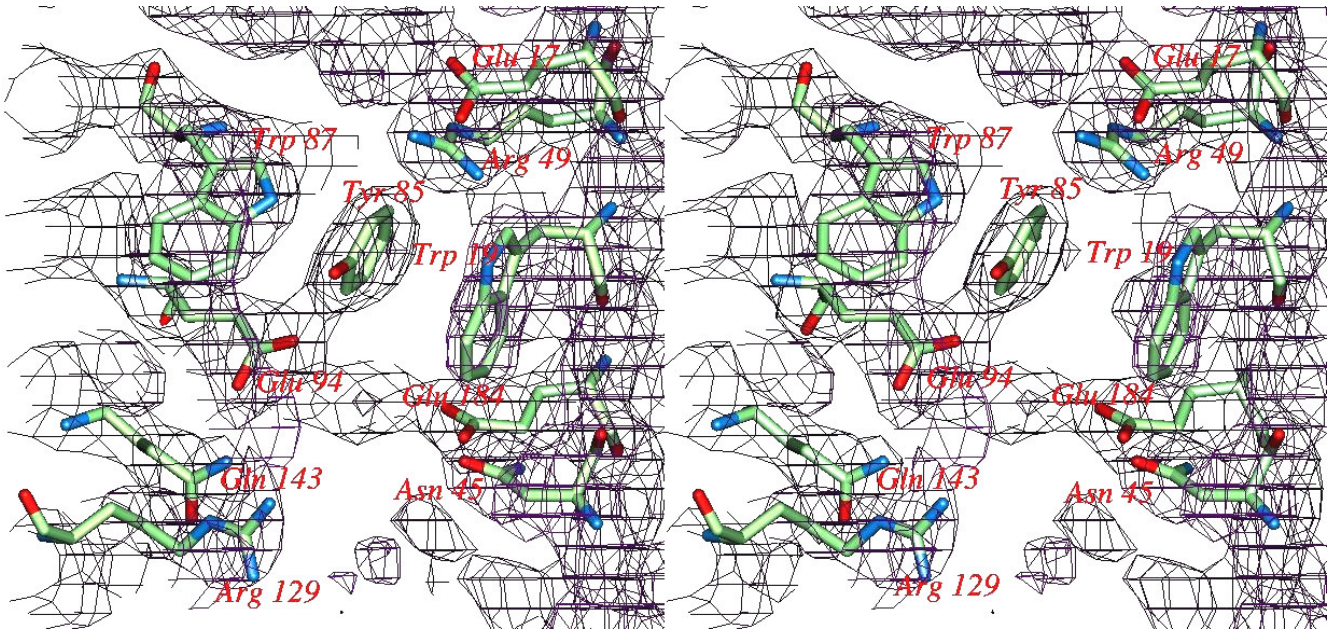


Fig. 6.6 Active-site of AT*B*XYL-C shown with (2Fo-Fc) map contoured at 1σ. Residues targeted by chemical modification are labelled.

6. 5 Analysis of fluorescence spectra of AT*B*XYL-C

6. 5.1 Two populations of tryptophan residues: Amino acid sequence deduced from X-ray crystallographic electron density map shows the presence of seven tryptophan residues in AT*B*XYL-C (Fig.6.7). Chemical modification of AT*B*XYL-C by tryptophan-specific reagents (NBS) and crystal structure of AT*B*XYL-C complexed with xylo-oligosaccharide (xylo-triose) indicate participation of a tryptophan residue in the active site of the enzyme (Balakrishnan *et al*., 2006). In the present work the main fluorescence emission at 346 nm and a shoulder at 330 nm have been attributed to two discreet Trp populations.

The tryptophan residues in AT*B*XYL-C exhibit unique spectroscopic features. The fluorescence spectrum of native AT*B*XYL-C shows an emission maxima at 346 nm

and a shoulder at 330nm when excited at 280 nm (Fig.5.7, Chapter-5), which could correspond to the partially exposed and completely buried populations of trptophans. It has been shown that tryptophan is essential for substrate binding in glycanases (Clark, 1987). Buried tryptophan residues in proteins usually show emission maxima in the range of 325 to 335 nm (Giraldi, *et al*., 1994). The emission maxima of L-tryptophan in water is 340 nm and completely exposed tryptophan side chain in proteins emit fluorescence at 350-356nm. The maxima of tryptophan emission in a protein is highly influenced by its conformation and the amino acids present in the vicinity (Brand *et al*., 1967). Interaction between the ground or excited states of aromatic groups with side chains of neighboring amino acids in the protein are known to occur. A mutant E246C of xylanase from *Pseudomonas fluorescens* (Szabo *et al*., 1985; Fftink *et al*., 1985) has shown that X5 is bound in the cleft in such a way that some or all of a chain of 20 water molecules that runs along the substrate binding cleft is expelled upon binding of X5. The three-dimensional structures of xylanases reveal that the active site clefts are dominated by highly conserved aromatic residues. This is represented by Trp19, Tyr85, Trp87, Tyr96, Tyr126, Trp145, Tyr186, and Tyr195 in the structure of AT*B*XYL-C. These residues are strictly conserved in other 11/G family xylanases. These hydrophobic residues are perfectly positioned to stabilize at least three to four xylose residues (Giraldi *et al.*, 1994; Brand *et al.*, 1967) and form stacking interaction and hydrogen bonds with xylose residues (Ghatge *et al*., 1993). Stacking interactions appear to be common among various glycosidases (Ghatge *et al*., 1993), and the pyranose rings of sugars with axial hydroxyls presents a hydrophobic surface that interacts with the aromatic side chains. The entropically driven xylo-oligosaccharide binding to *Chainia* endoxylanase indicates that hydrophobic stacking of xylose residues with aromatic residues within its active site cleft is the predominant driving force in the process of binding.

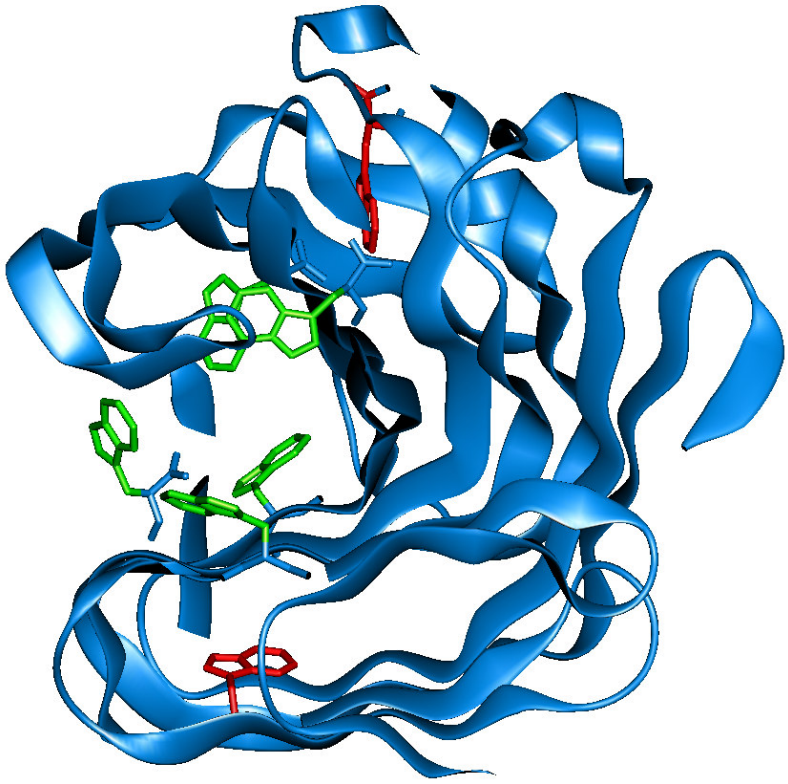


Fig. 6.7 Partial right hand folded AT*B*XYL-C showing two populations of tryptophan.

a. Partially exposed to hydrophilic environment (Green),

b. Hydrophobically buried (Red)

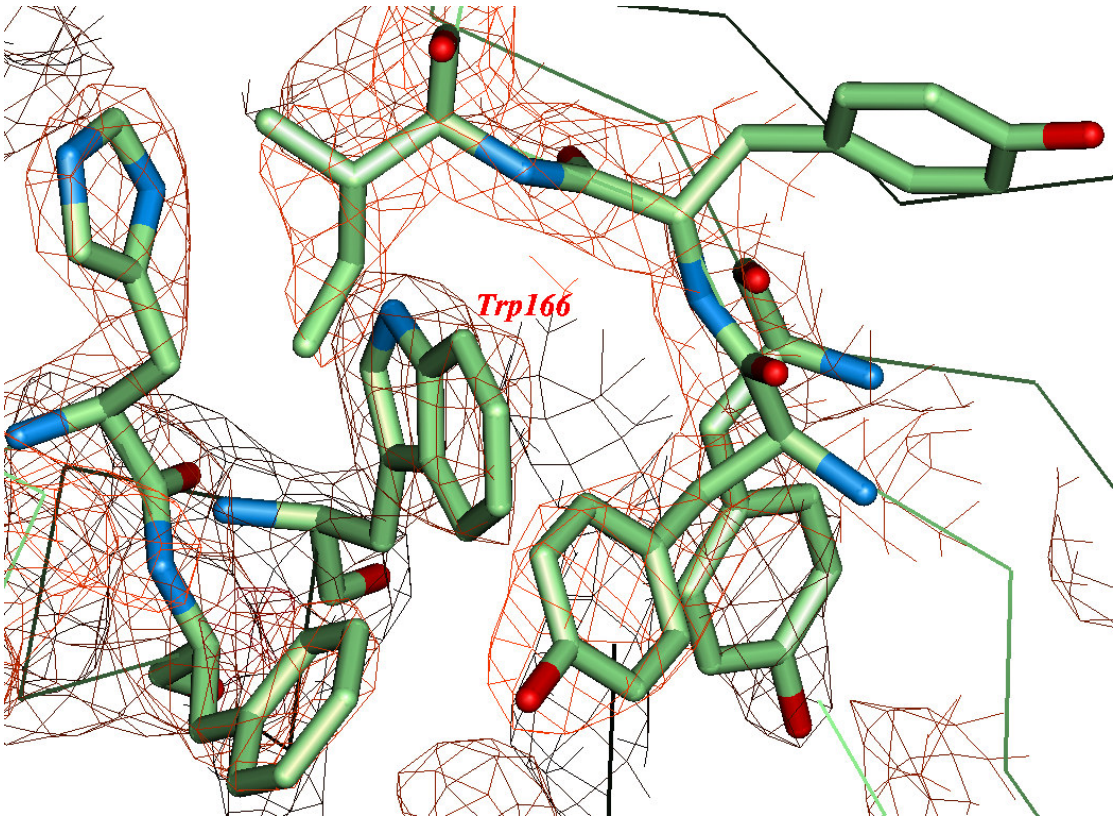


Fig. 6.8 Trp-166 which is completely buried in hydrophobic environment surrounding Phy, Tyr, and Ile (along with (2Fo-Fc) map contoured at 2σ). No interacting water molecules are seen around Trp 166.

6.5.2 Electropositive microenvironment of tryptophan 19: The fluorescence characteristics of tryptophan residues depend strongly on the microenvironment and thus provide a sensitive probe to study the conformational state of the protein. Fluorescence quenching of proteins in solutions have been widely used for studying the degree of exposure and electronic environment of aromatic amino acid residues. Acrylamide is an efficient quencher of tryptophan fluorescence and can distinguish between buried and exposed side chains. In contrast, KI and CsCl are highly hydrated, charged molecules and their quenching ability is limited to surface exposed tryptophans and also depends upon the neighboring charged groups. Titration of AT*B*XYL-C with increasing concentration of acrylamide resulted in 100% quenching of the fluorescence. CsCl showed no effect while KI partially quenched the fluorescence indicating electropositivity of Trp environment. The Ksv and fa values were extrapolated from a replot of the quenching data according to modified Stern-Volmer equation. The Ksv value for acrylamide was much higher than that of KI and CsCl since it can penetrate into the interior and is an efficient quencher.

The fluorescence quenchers KI and CsCl have been considered as companion probes since they bear opposite charges and their relative quenching efficiencies is dependent on accessibility and net charge in the vicinity of the fluorophore. The fraction of tryptophans accessible to KI was found to be 0.25, indicating that 25% tryptophans are accessible to iodide ions. There are a total of seven tryptophan residues in AT*B*XYL-C and 25% of seven tryptophans (1.75) is taken as one tryptophan. Trp19, which is located in the active site, may be accessible to iodide ions. Crystal structure of AT*B*XYL-C reveals that the microenvironment of Trp19 is electropositive due to its close proximity to residues Asn6 and Arg49 and N atoms of main chain amide groups. The electropositive environment of Trp19 seen in the structure is shown in Fig. 6.9.

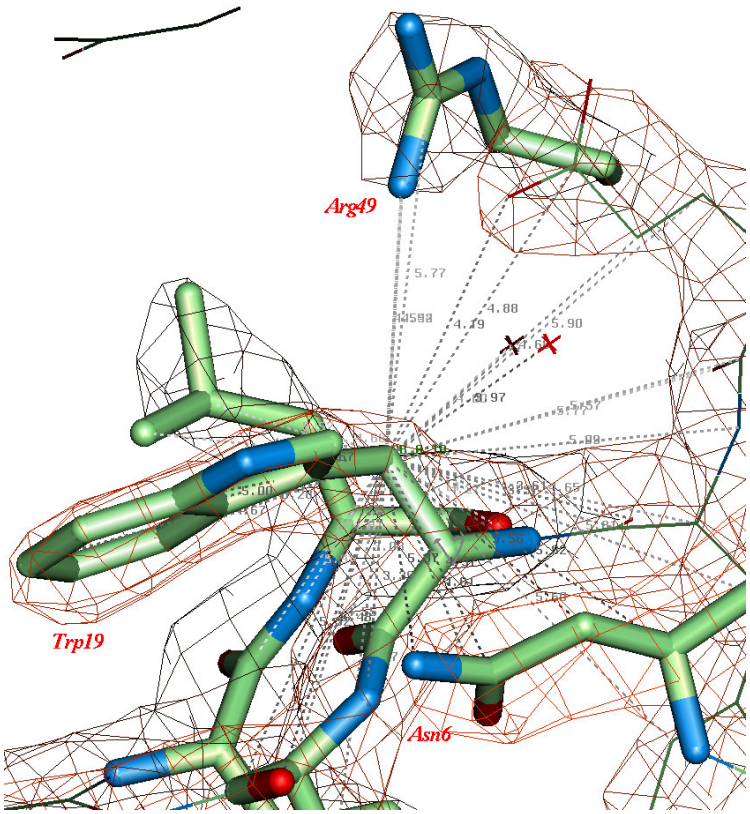


Fig. 6.9 Refined model of AT*B*XYL-C (along with (2Fo-Fc) map contoured at 2σ) showing the microenvironment of Trp-19. Electropositivity of the environment is due to Asn-6, nitrogen atoms from main chain and Arg-49.

6. 6 AT*B*XYL-C complexed with xylotriose in tetragonal crystals

AT*B*XYL-C catalyses the random cleavage of the β-1, 4-glycosidic bond of the xylan backbone with the retention of anomeric configuration. This reaction is believed to occur through a double-displacement reaction. During the process a covalent glycosyl-enzyme intermediate is formed and subsequently hydrolysed *via* a oxocarbenium-ion-like transition state. Such a mechanism requires a minimum of two catalytic carboxylates. One, an acid/base group that can function initially as a Bronsted acid, protonating the glycosidic bond to assist leaving-group departure and a second group that functions as a nucleophile forming a covalent glycosyl-enzyme intermediate of inverted configuration. The catalytic intermediate is then hydrolysed by a water molecule, activated by

deprotonation by the acid/base group, which now functions as a Bronsted base. In *Bacillus circulans* xylanase, the acid/base residue has been identified and the nucleophile has been trapped using xylo-oligosaccharide analogs (Wakarchuk *et al*, 1994). The catalytic mechanism is believed to feature 'pKa cycling' that regulates the pKa of the acid/base along the reaction trajectory as has been demonstrated on the *Bacillus circulans* enzyme using ^{13}C nuclear magnetic resonance (NMR) spectroscopy (Melntosh *et al.,* 1996).

Xylotriose and xylobiose were introduced into the mother liquor directly as solids as their use in solution form resulted in crystal cracking. Addition as powder reduced variation in the native crystallization condition in the hanging drop, which otherwise induces cracking. Soaking with powdered xylotriose not only avoids crystal cracking but also improves the diffraction quality of the crystal from 3.5Å to 2.4Å. Despite several crystallization trials, native tetragonal crystals failed to diffract beyond 3.5Å. Similar experiment involving addition of substrates to the orthorhombic crystal form did not result in a complex. Tetragonal crystals were stable for more than 4 days after xylotriose addition and data was collected on the fourth day from a single crystal at 120 K, to a resolution of 2.40 Å. The processed data had a R_{merge} of 5.2% and completion of 99.6%. After a few cycles of refinement of the structure the difference Fourier map showed electron density for xylotriose at the active site of AT*B*XYL-C. However, the vacant electron density was in two fragments, complelling us to interpret them as products of ligand hydrolysis (xylotriose⟹xylobiose + xylose). The disposition of the electron density showed that the xylobiose and xylose are hydrogen bonded with residues in the active site. Observed electron density for the xylobiosyl-enzyme intermediate and xylose is shown in Fig. 6.10. The two catalytic glutamate residues are located on opposite sides of the active-site cleft. The implications of the observation of xylobiosyl-enzyme in the crystal structure are discussed below (Fig. 6.11). The xylobiosyl-enzyme structure permits description of the protein-ligand interactions in the -2 and -1 subsites and xylose near the active site residues (Fig. 6.10 & 6.11). The bound sugar at subsite -2 is in the chair conformation while the occupant at subsite -1 is in boat conformation. They are bound *via* aromatic stacking 'below' the aromatic ring plane of Trp19, a residue

conserved throughout family 11 and family 12 xylanases (Balakrishnan *et al.,* 2006). The O3 hydroxyl group (-2 subsite) interacts both with the main chain carbonyl group of Pro133 and with the NH1 group of Arg129 (Table 6.5). The O2 hydroxyl group of subsite -2 also makes a hydrogen bond with the hydroxyl group of Tyr85. The O2 and O4 atoms of sugar molecule at subsite -1 form a hydrogen bond with the OE1 atom of nucleophile Glu94. By analogy with other systems it seems likely that the OEl…O2 interaction plays an important role in transition-state stabilization (Namchuck, 1995). The sugar hydroxyl group at C2 interacts with the NH1 atom of Arg129 and is 3.4Å from the carboxyl oxygen (OEl) of the nucleophile (all these distance are in reference to A subunit of the tetragonal unit cell) (Table 6.5). The catalytic acid/base residue Glu184 sits near to the pyranoside endocyclic O5-Cl bond (Varrot *et al.,* 1999) where it interacts with a solvent water molecule poised for nucleophilic attack at the anomeric centre. This water molecule is at 3.4 Å from the Cl of sugar and the Glu94 (OE2)-Cl-water angle is 168º, consistent with an 'in-line' nucleophilic attack (Sabini *et al.,* 1999). The 'catalytic water' of AT*B*XYL-C interacts with Asn45 and the OE2 of the acid/base is 4.6Å from C1 –1 subsite making angle 138º and not with Tyr96 and OEl of the acid/base as reported in the case of *B. circulans* enzyme. The carboxylate oxygens from Glu184 (OE2) interacts with Asn143 and Tyr96, respectively (Table 6.5 & Fig. 6.11). Close van der Waals' contacts are observed also with Leu47. These might play a role both in conferring the specificity for xylose-configured substrates (Sabini *et al.,* 2001). The structure of xylobiosyl-enzyme intermediate is very similar to the native structure. The overall r.m.s deviation between the molecules in native orthorombic unit cell and in the tetragonal complex structure is 0.24 Å for all Cα atoms. The only significant difference in the active site is that Tyr186, which showed two distinct conformations in the native structure in orthorhombic unit cell, adopts just one of these conformations in the complex tetragonal unit cell. One of the two conformations in the native structure occupies a position equivalent to the +l subsite in the complex. The structure of xylobiosyl-enzyme intermediate is very similar to the crystal structure of xylanase from *B.agaredhearens.* The co-ordinates and structural factors of orthorombic and tetragonal forms were deposited in PDB with I.D names 2F6B and 2NQY, respectively.

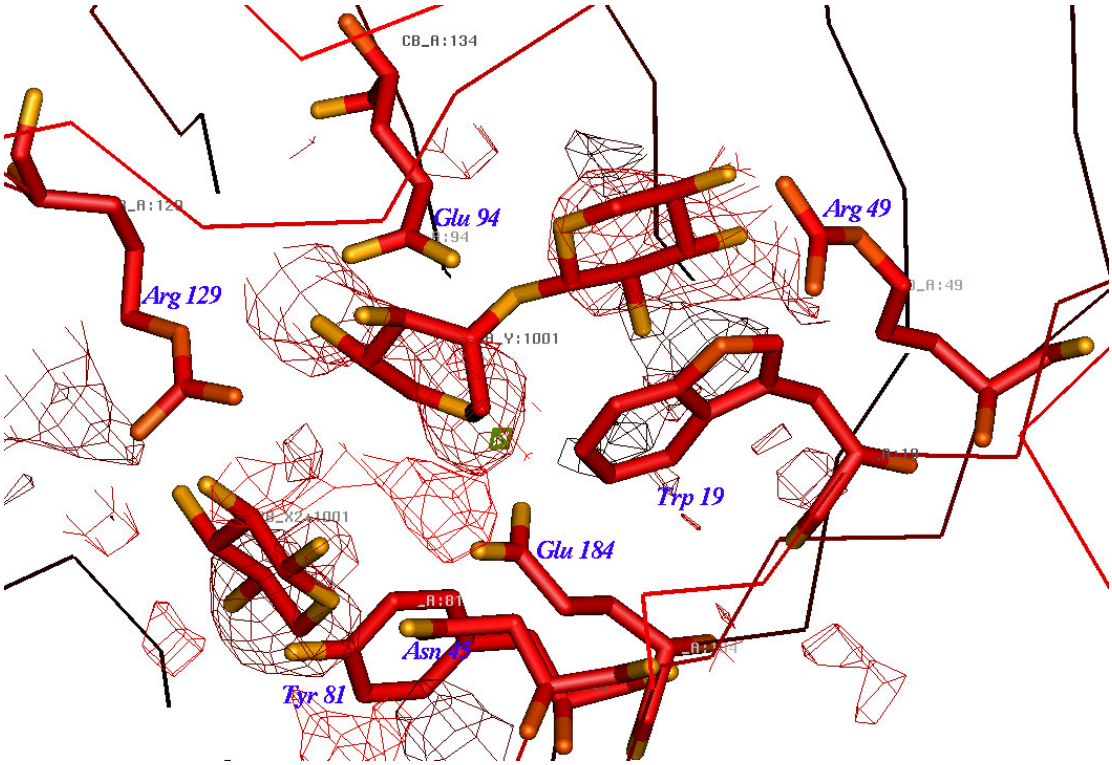


Fig. 6.10 AT*B*XYL-C complexed with cleavage products of xylotriose (xylobiose and xylose). (Fo-Fc) map is contoured at 2σ.

Table 6.5 Hydrogen bonds involved in substrate binding

	Residues involved in hydrogen bonding	Ligand	Distance (Å)
1	O2 and NH2 of Arg 129	-1 subsite	3.291
2	O2 and NH1 of Arg 129	-1 subsite	3.178
3	O3 and NH2 of Arg 129	-1 subsite	3.217
4	O2 and OE1 of Glu 94	-1 subsite	2.676
5	O4 and OE1 of Glu 94	-1 subsite	3.36
6	O3 and NH2 of Arg 49	-2 subsite	2.904
7	O3 and NH1 of Arg 49	-2 subsite	2.801
8	O3 and CO of main chain Pro 133	-2 subsite	2.715
9	O2 and OH of Tyr 85	-2 subsite	3.328
10	O2 and NH1 Arg 129	Xylose	3.291
11	O3 and H2O (162)	Xylose	3.347
12	O5 and NH Asn 45	Xylose	3.202
13	O2 and NH1 Arg 129	Xylose	3.291
14	OH of Tyr 85 and OE2 of Glu 94	-	2.513
15	NH of Asn 143 and OE2 Glu 94	-	2.668
16	OE1 of Glu 184 and OH of Tyr 96	-	2.868
17	OE1 of Glu 184 and OH2 (163)	-	2.577
18	ND2 Asn 45 and H2O (163)	-	2.643

***-1 and -2 subsites indicate xylose rings 1 and 2, respectively in a molecule of xylobiose**

***14-18 are intramolecular H-bonds in the active site**

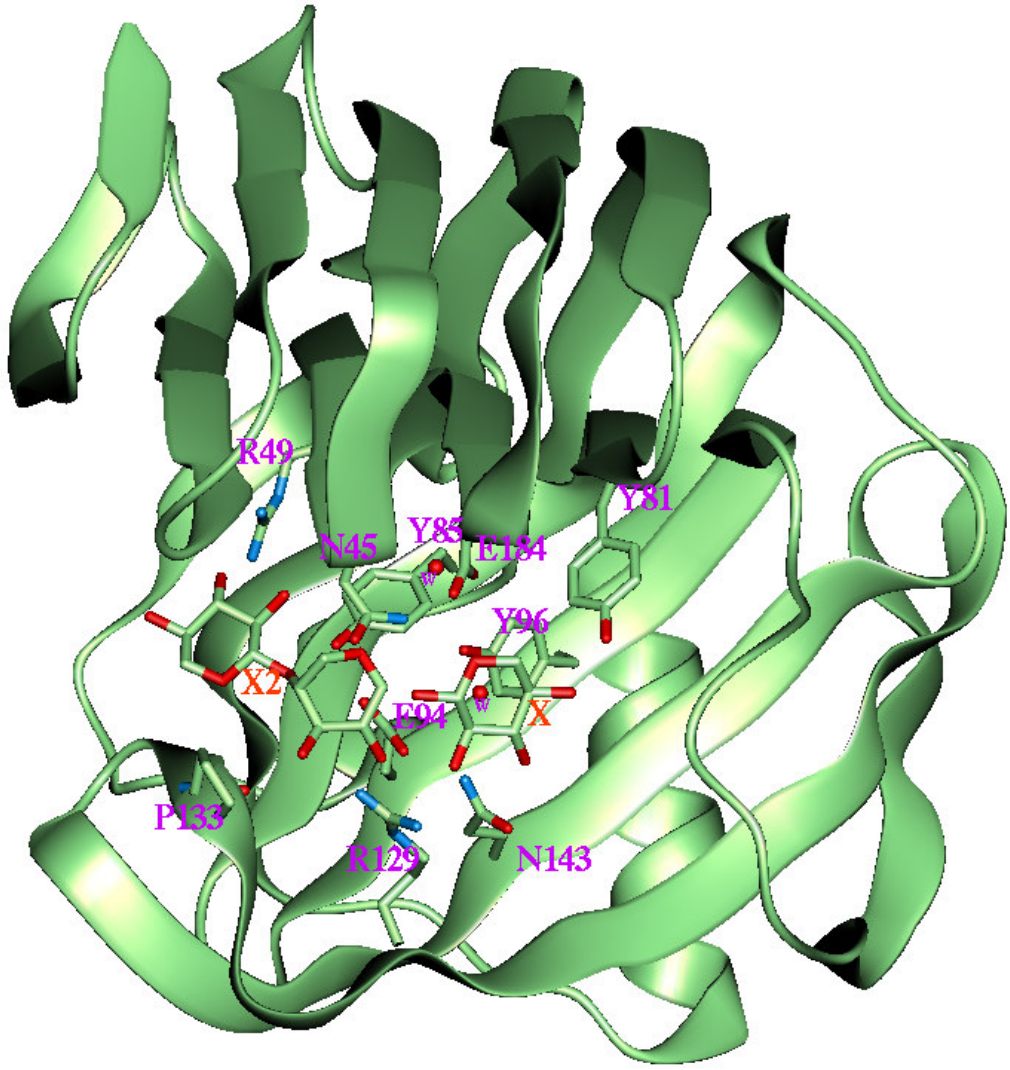


Fig. 6.11 Residues involved in hydrogen bonding at active site of AT*B*XYL-C complexed with xylobiose and xylose. Amino acid residues, Pro 133 (P133), Arg 129 (R129), Asn 143 (N143), Tyr 81 (Y81), Glu 184 (E184), Tyr 85 (Y85), Asn 45 (N45), Arg 49 (R49), Glu 94 (E94), Tyr 96 (Y96), H_20 (W) are labeled in magenta and xylobiose and xylose (X2 and X) in yellow

6.7 Structural comparison of family 11 xylanases

A number of crystal structures of family 11 xylanases are now available for mesophiles, thermophiles and acidophiles, except alkalophile. We have made a detailed comparison of the structures of these enzymes with AT*B*XYL-C. Twelve structures used in the comparison are summarized in Table 6.6 and the sequence alignment is shown in Fig. 6.12. Based on sequence homology family 11 xylanases can be divided into four groups (Table 6.6). The first group is formed by acidophilic xylanases ANX, AKX, and TRX-I. The second group contains alkalophilic AT*B*XYL-C and highly thermophilic DTX (sequence identity 58%). The third group is formed from thermophilic NFX and mesophilic BCX (sequence identity 59%). The fourth group contains mesophilic THX and TRX-II together with thermophilic PVX, TLX, and CTX. AT*B*XYL-C, DTX, BCX, and NFX are from bacterial sources, whereas the others are from fungal enzymes. Rmsd (root-mean-square deviation) values of superimposed structures of family 11 xylanases correlate well with the sequence similarities. The sequence identity range for different family 11 xylanases was 32–91% and rmsd range was 0.2–1.4 Å. As ligand binding can induce some structural changes, only native structures were chosen for comparison. The superimposition of three-dimensional structures confirmed the classification of xylanases based on sequence similarities. For example, the lowest rmsd value of thermophilic NFX is with mesophilic BCX (0.78 Å), both belonging to group 3. In group 1, AT*B*XYL-C has the lowest rmsd with thermophilic DTX (0.871 & 0.876 Å between molecule A & B of AT*B*XYL-C and molecule A & B of DTX, respectively). As the crystal structures of mesophilic and thermophilic xylanases are very similar, it is likely that an array of minor modifications forms the structural basis for enhanced stability for alkalophilic and thermophilic xylanases. Therefore, several factors, which are thought to be responsible for alkalo- and thermo-stability were compared between alkalophilic, thermophilic and mesophilic family 11 xylanases. The AT*B*XYL-C was not included in the same group as other mesophilic xylanases, because its functional properties seem to be different; for instance, *Bacillus sp. (NCL 86-6-10)* grows optimally at high pH (pH 10) on wheat bran medium and AT*B*XYL-C produced has an optimum activity at pH 8. On the other hand, acidophilic TRX I, AKX and ANX would be considered as a separate group of acidic

xylanases, but in our comparisons they were included in the mesophiles as a large number of mesophilic xylanases are slightly acidic in their activity profiles.

Table 6.6 Summary of the crystal structures used in comparison

Organism	Code	PDB Code	Temp. Pref.	PH Pref.	Res. (Å)	Data col. T (K)	Reference
Bacillus sp (NCL 86-10-6)	AT*B*XYL-C	2F6B & 2NQY	Thermophile	Alkalophile	2.8	100	Balakrishnan *et al*
T. reesei	TRX-I	1XYN		Acidophile		295	Torronen *et al*
A. kawachii	AKX	1BK1		Acidophile	2.0 2.4	295	Fushinobu *et al.*
A. niger	ANX	1UKR		Acidophile	2.0	295	Krengel *et al*
N. flexuosa	NFX	1M4W	Thermophile		2.1	120	Hakulinn *et al*
C. thermophilum	CTX	1H1A	Thermophile		1.75	120	Hakulinn *et al*
D. thermophilum	DTX	1F5J	Thermophile		1.8	110	Macarthy *et al*
T. lanuginosus	TLX	1YNA	Thermophile		1.55	295	Gruber *et al.*
P. varioti	PVX	1PVX	Thermophile		1.6	295	Kumar *et al*
T. reesei	TRX II	1ENX			1.5	295	Torronen *et al*
B. circulans	BCX	1XNB			1.5	295	Campbell *et al*
T. harzianum	THX	1XND			1.8	295	Campbell *et al*

6.8 Factors responsible for higher stability of family 11 xylanases

6.8.1 Amino acid composition and sequence: Frequencies of all 20 amino acids were computed for akalophilic, thermophilic and mesophilic family 11 xylanases (Table 6.9). An increased occurrence of arginines was found in the alkalophilic and thermophilic xylanases. Large-scale sequence comparisons have shown that thermophilic proteins contain more arginines on the protein surface than mesophilic proteins (Menendez-Arias *et al*., 1989; Vogt *et al*., 1997; Kumar *et al.*, 2000). The effect of a large-scale increase in the number of arginines was tested experimentally in *T. reesei* xylanase II (Turunen *et al*., 2002). These results showed that the introduction of five arginines into the Ser/Thr surface considerably increased the thermotolerance of the molecule. Another trend observed in thermophilic xylanases was the decrease in frequency of Ser and corresponding increase in the frequency of Thr (Table 6.9). Ser to Thr mutation was one of the stabilizing mutations found by Argos *et al*., (1979) in their early studies. Similarly, another study reported a decrease in the frequency of Ser rather than an increase of Thr for thermophilic proteins (Kumar *et al.*, 2000). It was also found that in thermophilic proteins Arg and Tyr were more frequent, while Cys and Ser were less frequent. One possible explanation for this observation in xylanases is that the increase in Thr : Ser ratio in β-strands improves the β-forming propensities. Over half of the residues in the family 11 xylanases are located in β-strands. In thermophilic xylanases, the frequency of asparagines is slightly lower (Table 6.9). Asn has a low β-forming propensity, and thus might be avoided in the β-strands of thermophilic xylanases. The highly alkalophilic xylanase AT*B*XYL-C and highly thermostable DTX showed lesser frequency of Gly compared to both mesophilic and other thermophilic xylanases. Absence of Gly probably increases the rigidity of the loop regions. However, there is no general trend toward decreased frequency of Gly among thermophilic family 11 xylanases. Pro does not seem to play any general role in the thermostabilization of these enzymes. Thermophilic xylanases have substantially less Val (Table 6.9). Although Val has a good β-forming propensity, still its frequency is lower in the β-strands of the thermophilic xylanases (Table 6.9), indicating the increase in the β-forming propensity of residues is not of primary importance in xylanases if some other property is more critical for

thermostability. In addition, thermophilic xylanases contain more charged amino acid residues than mesophilic xylanases (Table .9). The higher frequency of charged residues can naturally increase the number of polar interactions.

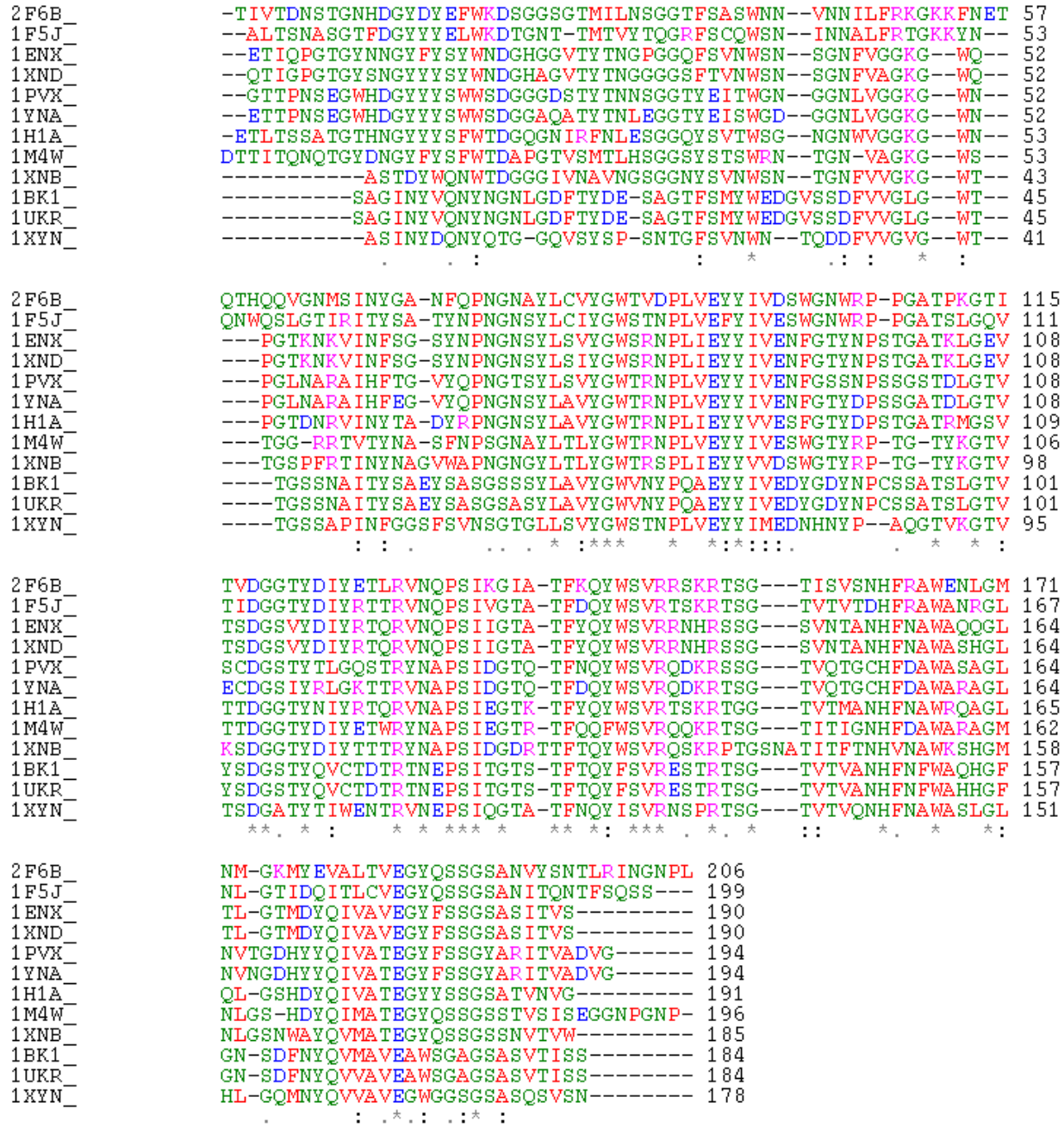


Fig. 6.12 Sequence alignment of family 11 xylanases: The identical amino acids (conserved) are marked with a star '*' and conservatively substituted amino acids are marked with a colon ':'. The alignment was prepared wtih ClustalW v.1.82.

6.8.2 Secondary structures: Facchiano *et al*., 1998 observed that 69% of the α-helices of thermophilic proteins are more stable than their mesophilic counterparts. The stabilizing factor was the intrinsic helical propensity of the constitutent amino acids. Lack of β-branched residues (Val, Thr, Ile) correlated significantly with thermostability. In the case of xylanases, there is only one α-helix in the structure. The α-helix of thermophilic xylanases showed a higher frequency of Asp and Arg. In thermophilic xylanases, an average of 123 residues (range 121–128) occurs in the β-strands and the corresponding number in the case of mesophilic xylanases was 114 residues (range 106–118). This result indicates that major part of the sequence in β-strand rigidify the protein and thereby enhance its thermostablity. AT*B*XYL-C has as many as 131 residues in its β-strands, which indicates that the overall stability of the β-strands may be important for the alkalitolerance of family 11 xylanases. The alkalophilic AT*B*XYL-C and thermophiles have an additional β-strand B1 at the N-terminus, which could have a stabilizing effect. However, mesophilic TRX II and THX also have this additional β-strand. The highly thermostable DTX has a clearly longer β-strand B3 and C-terminal β-strand A4, which most likely provides stability to the structure. The C-terminal β-strand A4 gives additional hydrogen bonding with β-strand A5 and the extension of β-strand B3 interacting with a β-strand B4. Even AT*B*XYL-C has a longer C-terminal β-strand A4 followed by a short additional β-strand. Superimposition of three-dimensional structures of all xylanases revealed a striking feature that alkalophilic AT*B*XYL-C and thermostable DTX had a long insertion between β-strands B3 and A5 in addition to the longer stretches at the terminals. According to McCarthy *et al*., (2000), the loop between B3 and A5 combined with extended C-terminus of DTX gives additional hydrogen bonding and hydrophobic packing. They suggest that these factors may account for the enhanced thermal stability. In fact the loop of DTX includes some regular secondary structures: the extension of β-strand B3 and 3_{10}-type helix. Structurally similar AT*B*XYL-C has a short α-helix, but there is no extension of β-strand B3. Furthermore, the structures of cord, thumb and loop regions vary among family 11 xylanases. It has been shown that these areas are flexible both in crystals and in the molecular dynamics simulations (Mailu *et al*., 1998). Loops are typically the regions with the largest temperature factors, indicating that they might unfold first during thermal denaturation (Daggett *et al.*, 1993). However,

the overall temperature factors of mesophilic and thermophilic xylanases were not comparable because the diffraction data for different xylanases have been collected either at 120 K or at room temperature. Temperature factors are dependent, in addition to the resolution and the programs used in the refinements, on the temperature at which data is collected. Several xylanases have short insertions or deletions in the loops (Fig.6.12). Mesophilic BCX has a short insertion between β-strands B7 and A6 and mesophilic ANX and AKX have short insertions between β-strands A3 and B3, but there is no clear trend that shortened loops would be associated with thermostability. The other highly thermostable xylanase, DTX, has a deletion in this loop.

Table 6.7 Sequence and residues present in the β stands of AT*B*XYL-C

No	Start	End	Residues	Sequence
1	Ile2	Val3	**2**	IV
2	Ser7	His11	5	STGNH
3	Tyr14	Lys20	7	YDYEFWK
4	Ser25	Leu30	6	SGTMIL
5	Phe36	Asn41	6	FSASWS
6	Asn45	Lys53	9	NILFRKGKK
7	Met66	Ser75	10	MSISYGANFQ
8	Gly77	Val89	12	GNAYLCVYGWTV
9	Val93	Trp101	9	VEYYIVDSW
10	Gly113	Val117	5	GTITV
11	Gly120	Gln132	13	GTYDIYETDRTNQ
12	Ala139	Arg148	10	ATFSQYWSTR
13	Ser154	Ile157	4	SGTI
14	Lys175	Ile191	17	DMYEVALTVEGYQSSGS
15	Asn193	Ile201	9	NVYSNTLRI

6.8.3 Disulfide bridges: Thermophilic PVX and TLX have one disulfide bridge connecting the C-terminus of the β-strand B9 with the N-terminus of the α-helix. According to the experimental data, introduction of disulfide bridges *via* site-directed mutagenesis has increased the thermostability xylanases from *T. reesei* and *B. circulans*. Disulfide bridges at the N-terminus or in the α-helix region improves the thermostability

by 10–15 ºC (Turunen, *et al*., 2001; Daggett, 1993; Wakarchuk, 1994; Sung, 2000). However, the disulfide bridge alone cannot be the crucial factor for enhanced thermal stability of xylanases due to the fact that highly thermostable DTX, NFX and AT*B*XYL-C do not contain disulfide bonds. In DTX, two of the cysteines are close enough to form a disulfide bridge between β-strands B5 and B4 in the catalytic area, but no bridge was seen in the electron density map (MaCarthy *et al*., 2000). In addition, the mesophilic AKX and ANX have a disulfide bond between the cord and the β-strand B8, indicating that stabilization of the α-helix region as well as other weak areas like N-terminus by various strategies is more important for the thermostability and alkalophilicity than the introduction of disulfide bridge alone.

6.8.4 Salt bridges and hydrogen bonds: There is increasing evidence that salt bridges and hydrogen bonds can affect protein stabilization (Vogt *et al*., 1997; Kumar *et al.,* 2000). In family 11 xylanases, the number of salt bridges varies between 2 and 12. There is one completely conserved salt bridge between the C-terminal Asp99 of β-strand B6 and Arg152 of the loop between β-strands B7 and A6. Thermophilic xylanases tend to have more salt bridges than mesophilic homolgues. Alkalophilic AT*B*XYL-C has the largest number of salt bridges, while acidophilic xylanases have the lowest number. Apparently, there could be a correlation between alkali-tolerance and salt bridges. The catalytic site of alkaline AT*B*XYL-C contains more charged residues, with Asp15, Glu17, and Glu184 corresponding to Tyr16, Ser18, and Tyr186 in *Streptomyces sp*. S38 Xyl1 (Wouters *et al*., 2001), and it has been suggested that these residues influence the pH optima of XylJ (Tamanoi *et al*., 1998). Interestingly, some of these charged residues are present in the amino acid sequences of *Bacillus pumilus* A-30 XYNA, and to a lesser extent, in *Dictyoglomus thermophilum* Rt46B.1 XynB. Comparison of amino acid sequence and structure of AT*B*XYL-C with those of other family 11 xylanases indicate that the configuration of the thumb of AT*B*XYL-C contain an eight-residue-longer loop between β-strands B3 and A5, whereas this loop is three residues shorter in other members of family 11 xylanases like *Dictyoglomus thermophilum* Rt46B.1 XynB. In this longer loop, Lys53 protrudes into the catalytic cavity, whereas Glu56 found in AT*B*XYL-C points towards the thumb and is connected by a salt-bridge to the lysine residue in position 136.

Inspection of the catalytic site of AT*B*XYL-C showed that Glu17, Arg49, Lys53 and Glu178 are connected by four salt-bridges and are located in the substrate binding area. AT*B*XYL-C contains a total of 10 salt-bridges, whereas other members of the family have lesser number of salt bridges. Higher number of salt bridges AT*B*XYL-C may be responsible for its alkali-tolerance.

Table 6.8 Comparison of ionic interactions in alkalophilic and thermophilic members of family 11 xylanases

Salt-bridge	AT*B*XYL-C	XynB
1	Glu17-Lys53	(Glu16/Lys51)
2	Glu17-Arg49	Glu16-Arg47
3	Arg49-Glu178	(Arg47/Glu174)
4	Glu56-Lys136	(-/Val132)
5	Asp99-Arg152	Glu95-Arg148
6	Arg105-Asp123	Arg101-Asp119
7	Lys112-Glu126	(Leu108/Arg122)
8	Asp123-Arg149	(Asp119/Asp138)
9	Glu126-Lys142	Arg122-Asp138
10	Lys53-Glu178	Lys51-Asp173

***Residues in the bracket in case of XynB are putative salt bridges.**

6.8.5 Packing and thermal stability: It has been proposed that thermophilic proteins have a tighter internal packing with smaller and less numerous cavities than mesophilic proteins (Russell *et al.*, 1997; DeDecker *et al*., 1996). Because thermophilic xylanases have more atoms, the void volumes were divided by the total number of atoms to normalize them. Both the protein density and void volume values for thermophilic and mesophilic xylanases were similar. Only alkalophilic AT*B*XYL-C and highly thermophilic DTX have slightly higher protein density and lower void volumes indicating better packing. By comparing structures in PDB Karshikoff and Ladenstein (1998) observed that proteins

from thermophilic and mesophilic organisms essentially do not differ in packing. They suggest that neither the reduction in packing density nor the reduction of the packing defects can be considered as a common mechanism for increasing thermal stability. On the other hand Chen *et al*., (2000) observed in a mutagenesis study that the stabilizing mutations in *Staphylococcal* nuclease resulted in improved packing, with the volume of the mutant's hydrophobic cores decreasing as protein stability increased. Apparently, a few protein families or some members in them may use better packing to improve the thermostability. Our study indicates that in AT*B*XYL-C adaptation to alkaline pH and high thermostability in DTX benefit from better packing.

6.8.6 Hydrophobicity and surface characteristics: Because protein cores are typically hydrophobic, increased packing efficiency is often correlated with increased hydrophobicity. Tighter packing can be achieved through the formation of hydrophobic clusters and enhanced van der Waals interactions. Increased hydrophobicity is usually associated with decreased accessible surface areas and a higher percentage of buried atoms (Chan *et al*., 1995). As alkalophilic and thermophilic xylanases contain more amino acid residues than mesophilic xylanases, they also have larger accessible surface areas (Table 6.9). So far, all family 11 xylanases are reported to be monomers, therefore the solvent accessible areas are not buried by oligomerization. When accessible surface area is counted per atom, it appears that AT*B*XYL-C and DTX may benefit from increased hydrophobicity as well as better packing. In addition, these two xylanases on an average have longer side-chains (atoms per residue) than the other family 11 xylanases studied. One type of hydrophobic interaction is the closely packed aromatic ring–ring interaction, which has been estimated to contribute nonbonded potential energies between 1 and 2 kcal/mol (Burley *et al*., 1988). Additional aromatic–aromatic interactions are also believed to contribute to the increased stability (Kannan *et al*., 2000). Introduction of a single tyrosine into the N-terminal region has been reported to considerably improve the thermostability and thermophilicity of *Streptomyces* xylanase (Georis *et al*., 2000). However, AT*B*XYL-C did not show any general trend towards increased proportion of aromatic residues. It is thought that the increased fractional polar surface, which results in added hydrogen bonding to water, contributes to the greater stability (Vogt *et al*., 1997).

Table 6.9 Amino acid composition of Family 11 xylanases

	ATBXYL-C	*CTX*	*DTX*	*TLX*	*PVX*	*NFX*	*thermo*	*BCX*	*TRX II*	*THX*	*TRX I*	*AKX*	*ANX*	*meso*
% Ala	3.9	5.2	4.5	6.7	4.6	4.2	5.1	4.8	3.7	4.7	5.1	7.6	8.2	5.7
% Val	6.8	6.8	5.0	6.7	6.2	4.2	5.8	7.5	7.4	6.8	10.1	8.2	8.7	8.1
% Leu	4.3	3.1	5.0	4.1	3.6	2.6	3.7	2.1	2.6	2.6	3.4	2.2	2.2	2.5
% Ile	5.8	3.1	5.5	3.6	3.1	4.2	3.9	3.2	4.7	5.3	3.4	2.7	2.7	3.7
% Pro	3.4	2.6	2.5	3.1	3.1	2.6	2.8	3.2	3.7	3.2	3.4	1.6	1.6	2.8
% Met	2.4	1.0	0.5	0.0	0.0	1.6	0.6	1.1	0.5	0.5	1.1	1.1	0.5	0.8
% Phe	3.4	2.6	3.5	2.6	2.6	3.7	3.0	2.1	4.2	3.7	3.4	4.9	4.9	3.9
% Trp	3.4	3.7	4.0	4.1	4.1	4.2	4.0	5.9	3.2	3.2	3.4	2.7	2.7	3.5
% Gly	11.6	14.1	9.5	14.9	16.0	13.6	13.6	13.4	14.2	14.2	12.4	10.3	10.3	12.5
% Ser	8.2	8.9	10.1	6.7	11.3	10.5	9.5	9.6	11.6	12.6	12.9	15.8	15.2	13.0
% Thr	7.7	12.6	14.1	9.3	10.8	15.2	12.4	13.4	8.4	8.9	10.1	10.9	10.9	10.4
% Cys	0.5	0.0	1.5	1.0	1.0	0.0	0.7	0.0	0.0	0.0	0.0	1.1	1.1	0.4
% Tyr	6.3	9.4	7.0	8.8	8.8	8.4	8.5	8.0	8.9	9.5	5.6	9.2	9.2	8.4
% Asn	10.6	8.4	8.5	6.2	7.2	5.8	7.2	9.6	10.5	10.0	10.1	6.5	6.5	8.9
% Gln	4.3	4.2	5.5	4.1	3.6	4.2	4.3	2.7	5.3	3.2	6.2	3.3	2.7	3.9
% Asp	3.9	3.1	3.5	6.2	5.2	3.7	4.3	3.7	2.1	2.1	2.8	4.9	4.9	3.4
% Glu	3.4	2.6	2.0	4.1	2.6	3.1	2.9	1.1	2.1	2.1	2.8	4.3	4.3	2.8
% Lys	4.3	1.6	2.0	1.5	1.0	1.6	1.5	2.7	2.1	2.1	0.6	0.0	0.0	1.2
% Arg	3.9	5.2	5.0	4.1	3.1	5.2	4.5	3.7	3.2	3.2	1.7	1.6	1.6	2.5
% His	1.9	1.6	0.5	2.1	2.1	1.6	1.6	1.1	1.6	2.1	1.7	1.1	1.6	1.5
Total	206	191	199	194	194	191	194	187	190	190	178	184	184	186
%Non-polar	33.3	27.7	30.6	30.9	27.3	27.7	28.8	31	30	30	33.1	31	31.5	31.1
% Polar	49.3	58.7	56.3	51	58.8	58.2	56.6	56.7	58.9	58.4	57.3	57.1	56	57.4
% Charged	17.4	13.8	13.1	18.1	13.9	14.1	14.6	12.3 1	11.1	1.6	9.6 1	1.9	12.5	11.5

REFERENCES

- Abad-Zapatero, C., Fox J.L. and Hackert, M.L. (1977) The quaternary structure of a unique phycobiliprotein: phycoerythrin from *Porphyridium cruentum*. *Biochem. Biophys. Res. Cannon*. **78**, 266-272.

- Adir, N. and Lerner N. (2003) The crystal structure of a novel unmethylated form of C-phycocyanin, a possible connector between cores and rods in pycobilisomes. *J. Biol. Chem.* **278,** 25926-25932.

- Adir, N., Vainer, R. and Lerner, N. (2002) Refined structure of C-phycocyanin from the cyanobacterium *Synechococcus vulcanus* at 1.6 Å insights into the role of solvent molecules in thermal stability and co-factor structure. *Biochem. Biophys. Acta.* **1556,** 168-174.

- Amemiya, Y., Zhou, J.M., Fan, Y.X., Kihara, H. and Kimura, K. (1997) Unfolding of dimeric creatine kinase in urea and guanidine hydrochloride as measured using small angle X-ray scattering with αsynchrotron radiation. *FEBS Lett.* **415,** 183-185.

- Anamika Patel., Sandhya Mishra., Richa Pawar and Ghosh, P.K.(2004) Purification and charactarization of C-Phycocyanin from cyanobacterial species of marine and freshwater habitat. *Protein Expr. Purif.* **40,** 248–255.

- Anna, T., Hadar, F., Rotem, G., Aviva, L., Adra, M. and Vivian, S. (1997) Crystallization and preliminary X-ray analysis of the thermostable alkaline xylanase from *Bacillus stearothermophilus*. *Crystallogr. D*. **53,** 608–611.

- Apt, K.E. and Collier, J.L. and Grossman, A.R. (1995) Evolution of the phycobiliproteins. *J. Mol Biol.* **248,** 79-96.

- Arakawa, T. and Timasheff, S.N. (1985) Mechanism of poly(ethylene glycol) interaction with proteins. *Biochemistry.* **24,** 6756-6762.

- Arase, A., Yomo, T., Urabe, I., Hata, Y., Katsube, Y. and Okada, H. (1993) Stabilization of xylanase by random mutagenesis. *FEBS Lett.* **316,** 123-127.

- Argos, P., Rossman, M.G., Grau, U.M., Zuber, H., Frank, G. and Tratschin, J.D. (1979) Thermal stability and protein structure. *Biochemistry* **18,** 5698–5703.

- Arribas, A., Martin Luengo, C., Rodriguez, J., Pabon, P., Nieto, F., Diego, M., Castano, L., Santos, I., Cascon, M. and Sala, A.(1994) The "arrhythmic pattern": a new method with prognostic significance in myocardial infarct. *Rev Esp Cardiol.* **47,** 284-293.

- Astier, J.P., Veesler, S. and Boistelle, R. (1988) Protein crystals orientation in a magnetic field. *Acta Crystallogr. D.* **54**, 703-706.

- Ay, J., Gotz, F., Borriss, R. and Heinemann, U. (1998) Structure and function of the *Bacillus* hybrid enzyme GluXyn-1: nativelike jellyroll fold preserved after insertion of autonomous globular domain. *Proc. Natl. Acad. Sci. USA* ***95*****,** 6613–6618.

- Balakrishnan, H., Duttachoudhary, M., Srinivasan, M.C., Rele, M.V. (1992) Cellulase free xylanase production from an alkalophilic *Bacillus species*. *World J. Microbiol Biotechnol.* **8,** 627–631.

- Balakrishnan, H., Kamal Kumar, B., Dutta-Choudhury, M. and Rele, M.V. (2002) Characterization of alkaline thermoactive cellulase-free xylanases from alkalophilic *Bacillus (NCL 87-6-10). J. Biochem Mol Biol Biophys.* **6,** 325-334.

- Balakrishnan, H., Satyanarayana, L., Gaikwad, SM. and C.G. Suresh (2006). Structural and active site modification studies implicate Glu, Trp and Arg in the activity of xylanase from alkalophilic *Bacillus sp. (NCL 87-6-10). Enzyme and Microbial Tecnology.* 39, 67-73.

- Beauchamp, J.C. and Isaacs, N.W. (1999) Methods for X-ray diffraction analysis of macromolecular structures. *Curr Opin Chem Biol.* **3,** 525-529.

- Bedford, M.R. and Classen, H.L. (1992) Reduction of intestinal viscosity through manipulation of dietary rye and pentosanase concentration is effected through changes in the carbohydrate composition of the intestinal aqueous phase and results in improved growth rate and food conversion efficiency of broiler chicks. *J. Nutr.* **122,** 560-569.

- Bennett, A. and Bogorad, L. (1973) Complementary chromatic adaptation in a filamentous blue green alga. *J. Cell Biol.* **58,** 419–435.

- Berns, D.S. and MacColl, R. (1989) Phycocyanin in physical chemical studies. *Chem. Rev*. **89,** 807-825.

- Bhat, S., Hutson, RA., Owen, E. and Bhat, M.K. (1997) Determination of Immunological Homology Between Cellulosome Subunits and Cloned Endoglucanases and Xylanase of *Clostridium thermocellum*. *Anaerobe.* **3,** 347-352.

- Bhat, V.B. (2001) Biomodulating properties of C-phycocyanin from *Spirulina platensis*, PhD thesis, JNCASR. Bangalore. India.

- Bhat, V.B. and Madyastha, K.M. (2000) C-Phycocyanin: a potent peroxyl radical scavenger in vivo and in vitro. *Biochem. Biophys. Res. Commun*. **275,** 20-25.

- Bhat, V.B., Gaikwad, N.W. and Madyastha, K.M. (1998) Hepatoprotective effect of C-phycocyanin: protection for carbon tetrachloride and R- (+)-pulegone-mediated hepatotoxicty in rats. *Biochem. Biophys. Res. Commun.* **249,** 428-431.

- Biely, P. (1985) Microbial xylanolytic systems. *Trends Biotechnol.* **3,** 288-290.

- Biely, P., Kratky, Z. and Vrsanska, M. (1981) Substrate-binding site of endo-1, 4-beta-xylanase of the yeast *Cryptococcus albidus*. *Eur. J. Biochem*. **119,** 559–564.

- Biely, P., Kremnicky, L., Alfoldi, J. and Tenkanen, M. (1994) Stereochemistry of the hydrolysis of glycosidic linkage by endo-β-1,4-xylanases of Trichoderma *reesei. FEBS Lett.* **356**, 137-140.

- Boussiba, S. and Richmond, AE. (1979) Isolation and characterization of phycocyanins from the blue-areen alga *Spirulina platensis*. *Arch. Microbiol.* **120,** 155-159.

- Brand, L. and Witholt, B. (1967) Fluorescence measurements**.** *Methods Enzymol*. ***11*****,** 776-856.

- Bray, M.R. and Clarke, A.J. (1990) Essential carboxy groups in xylanase A. *Biochem. J.* **270,** 91-96.

- Bray, M.R. and Clarke, A.J. (1994) Identification of a glutamate residue at the active site of xylanase A from *Schizophyllum commune*. *Eur J. Biochem.* **219,** 821-827.

- Bray, M.R., Carriere, A.D. and Clarke, A.J. (1994) Quantitation of tryptophan and tyrosine residues in proteins by fourth-derivative spectroscopy. *Anal Biochem.* **221,** 278-284.

- Brejc, K., Ficner, R., Huber, R. and Steinbacher, S. (1995) Isolation, crystallization, crystal strucrure analysis and refinement of allophycocyanin from the cyanobacterium *SpiruIina platensis* at 2.3Å resolution. *J. Mol. Bio.* **249**, 424-440.

- Bricogne, G. (1997) Bayesian statistical viewpoint on structure determination: basic concepts and examples. *Methods in Enzymol.* **276,** 361-423.

- Brunger, A.T. (1992) Free R-value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature.* **355,** 472-474.

- Brunger, A.T., Adams, P.D., and Rice, L.M. (1997) New applications of simulated annealing in x-ray crystallography and solution NMR. *Structure.* **5,** 325-336.

- Brunger, A.T., Krukowski, A. and Erickson, J. (1990) Slow-Cooling Protocols for Crystallographic Refinement by Simulated Annealing. *Acta crystallog. A.* **46,** 585-593.

- Brunger, A.T., Kuriyan, J. and Karplus, M. (1987) Crystallographic R factor Refinement by Molecular Dynamics. *Science.* **235,** 458-460

- Bryant, D.A. (1991) The Photosynthetic Apparatus, Molecular Biology and Operation, Cell Culture and Somatic Cell Genetics *of Plants*, Vol. 7B, edited by L. Bogorad and I. K. Vasil, New York: Academic Press pp. 255-298.

- Bryant, D.A., Glazer, A.N. and Eiserling, F. (1976) Characterization and structural properties of the major biliproteins of *Anabaena sp. Arch. Microbiol.* **110,** 61-75.

- Bryant, D.A., Guglicmi, C., Tandeau de Marsac N., Casters, A.M. and Cohen-Bazire, G. (1979) The structure of cyanobecterial phycobilisomes: A model. *Arch. Microbiol.* **123,** 113-121.

- Burke, M.J., Pratt, D.C. and Moskowitz, A. (1972) Low-temperature absorption and circular dichroism studies of phytochrome. *Biochemistry.* **11,** 4025-4031.

- Burla, M.C., Camalli, M., Carrozzini, B., Cascarano, G.L., Giacovazzo, C., Polidori, G. and Spagna, R. (2000) SIR2000, a Program for the automatic ab initio crystal structure solution of proteins. *Acta Crystallogr. A*. **56,** 451-457.

- Burley, SK. and Petsko, G.A. (1985) Aromatic–aromatic interaction: a mechanism of protein structure stabilization. *Science* **229,** 23–28.

- Campbell, R.L, Rose, D.R., Wakarchuk, W.W., Sung, W. and Yaguchi, M. (1993) A comparison of the structure of 20 kDa xylanase from *Trichoderma harzianum* and *Bacillus circulans*. In: Proceedings of the Second TRICEL Symposium on *T. reesei* cellulase of other hydrolases. pp. 72–83.

- Campbell,R.L., Rose,D.R.,Wakarchuk,W.W.,To,R.J., Sung,Z. and Yagushi, M. (1993) High resolution structures of xylanases from *Bacillus circulans* and *Trichoderma harzianum* identify a new folding pattern and implications for the atomic basis of the catalysis. Foundation for biotechnical and industrial fermentation research In: *Trichoderma reesei* Cellulases and Other Hydrolases (Souminen,P .and Reikainen,T., Eds.),pp. 63–72.Espoo,Finland.

- Canals, A., Vega, M.C., Gomis-Ruth, F.X., Diaz, M., Santamaria, R.R. and Coll, M. (2003) Structure of xylanase Xys1delta from Streptomyces halstedii. *Acta Crystallogr. D.* **59,** 1447–1453.

- Carter, C.W Jr. and Carter, C.W. (1979) Protein crystallization using incomplete factorial experiments. *J. Biol. Chem.* **254,** 12219-12225.

- Chan, M.K., Mukund, S., Kletzin, A., Adams, M.W.W. and Rees, D.C. (1995) Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* **267,** 1463– 1469.

- Chang, W.R., Jiang, T., Wan, Z.L., Thang. J.P., Yang, Z.X. and. Liang, D.C. (1996) Crystal structure of R-phycoeryhrin from *Polysiphonia urceolata* at 2.8Å resolution. *J. MoI. Biol.* **262,** 72l-731.

- Chen, F., Zang, Y.M. and Guo, S.Y. (1996) Growth and phycocyanin formation in photoheterotrophic culture, *Biotechnol. Lett.* **18,** 603–608.

- Clement, G. (1975) Production and characteristic constituents of the algae *Spirulina platensis* and *S. maxima. Ann. Nutr. Aliment.* **29,** 477-488.

- Collaborative Computational Projcct, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D.* **50,** 760-763.

- Collins, T., Gerday. C. and Feller, G. (2005) Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol Rev.* **29,** 3-23.

- Contreras-Martel, C., Martinez-Oyanedel, J., Bunster, M., lrgrand, P., Piras, C., Vernede, X. and Fontecilla-Camps, J.C. (2001) Crystallization and 2.2Å resolution structure of R-phycoerythrin from *Gracilaria chilensis*: a case of perfect hemihedral twinning. *Acta Crystallogr. D,* **57,** 52-60.

- Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-active enzyme server (CAZY) at URL: http://afmb.cnrs-mrs.fr/~cazy/CAZY/.

- Crowther, R.A. (1972) The fast rotation function. In: Molecular replacement method. Rossmann MG. Gordon and Breach eds. New York. pp: 173-178.

- Crowther, R.A. and Blow, D.M. (1967) A method of positioning a known molecule in an unknown crystal structure. *Acta Crystallogr.* **23,** 544-548.

- Daggett, V. and Levitt, M. (1993) Protein unfolding pathways explored through molecular dynamics simulations. *J. Mol. Biol.* **232,** 600–619.

- Dale, R.E. and Teale, F.W.J. (1970) Number and distribution of chromophore types in native phycobiliproteins. *Photochem. Photobiol.* **12,** 99-117.

- Dauter, Z. and Dauter, M. (2001) Entering a new phase. Using solvent ions in structure determination. *Stucture.* **9,** 21-26.

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

- Davies, G.J. and Withers, S.G. (1998) Snapshots along an enzymatic reaction coordinate: analysis of a retaining 6-glycoside hydrolase. *Biochemistry* **37,** 1707-1713.

- Debeire, P., Priem, B., Strecker, G. and Vignon, M. (1990) Purification and properties of an endo-1,4-xylanase excreted by hydrolytic thermophilic anaerobe *Clostridium thermolacticum. Eur. J. Biochem.* **187,** 573-580.

- Debreczeny, M.P., Sauer, K., Zhou, J. and Bryant, D.A. (1993) Monomeric C-Phycocyanin at Room Temperature and 71 K. Resolution of the Absorption and Fluorescence Spectra of the Individual Chromophores and the Energy-Transfer Rate Constants. *J. Phys Chem.* **97,** 9852-9862.

- Debreczeny, M.P., Sauer, K., Zhou, J. and Bryant, D.A.(1995b) Comparison of Calculated and Experimentally Resolved Rate Constants for Excitation Energy transfer in C-phycocyanin. 2.Trimers. *J. Phys. Chem.* **99,** 8420-8431.

- Debreczeny, M.P., Sauer, K., Zhou, J. and Bryant, DA. (1995a) Comparison of Calculated and Experimentally Resolved Rate Constants for Excitation Energy transfer in C-phycocyanin, l. Monomers. *J. Phys. Chem*. **99,** 8412-8419.

- Dekker, H., and Richards, G.N. (1976) Hemicellulases: Their occurrence, purification, properties and mode of action. *Adv. Carbohydr. Chem. Biochem.* **32,** 277-352.

- Derewenda, U., Swenson, L., Green, R., Wei, Y., Morosoli, R., Shareck, F., Kluepfel, D. and Derewenda, Z.S. (1994) Crystal structure, at 2.6-Å resolution, of the *Streptomyces lividans* xylanase A, a member of the F family of beta-1,4-D-glycanases. *J. Biol. Chem.* **269,** 20811–20814.

- Deshpande, V., Hinge, J., and Rao, M. (1990) Chemical modification of xylanases: evidence for essential tryptophan and cysteine residues at the active site**.** *Biochim. Biophys. Acta*. **1041**, 172-177.

- Deshpande, V., Lachke, A., Mishra, C., Keskar, S. and Rao, M. (1986) Mode of action and properties of xylanase and L- xylosidase from *Neurospora crassa*. *Biotechnol.Bioeng.* **26,** 1832-1837.

- Dey, D., Hinge, J., Shendye, A. and Rao, M. (1992) Purification and properties of extracellular endoxylanases from an alkalophilic thermophilic *Bacillus sp*. *Can. J. Microbiol.* **38,** 436-442.

- Dobler, M., Dever, S.D., Laves, K., Binder, A. and Zuber, H. (1972) Crystallization and preliminary crystal data of C-phycocyanin. *J. Mol Biol.* **71,** 785-787.

- Dominguez, R., Souchon, H., Spinelli, S., Dauter,Z., Wilson, K.S., Chauvaux, S., Beguin, P. and Alzari, P.M. (1995) A common protein fold and similar active site in two distinct families of betaglycanases. *Nat. Struct. Biol.* **2,** 569–576.

- Donald Berns, S. and Robert MacColl. (1989) Phycocyanin in physical chemical studies Chem. **89,** 807-825.

- Drenth, J. and Haas, C. (1998) Nucleation in protein crystallization. *Crystallogr. D*. **54,** 867-872.

- Ducret, A., Sidler, W., Frank, G. and Zuber, H. (1994) The complete amino acid sequence of R-phycocyanin alpha and beta subunits from the red alga *Porphyridium cruentum*. Structural and phylogenetic relationships of the phycocyanins within the phycobiliprotein families. *Eur J Biochem.* **221,** 563-580.

- Ducruix, A. and Giege, R., (1992). Crystallization of nucleic acids and proteins: A practical approach. Eds. IRL Press. Oxford, United Kingdom.

- Duerring, M., Huber, R. and Bode, W. (1988) The structure of gamma-N-methylasparagine in C-phycocyanin from *Mastigocladus laminosus* and *Agmenellum quadruplicatum. FEBS Let.* **236,** 167-170.

- Duerring, M., Huber, R., Bode, W., Ruembeli, R. and Zuber, H. (1990) Refined three-dimensional structure of phycoerythrocyanin from the cyanobacterium *Mastigocladus laminosus* at 2.7 Å. *J. Mol. Biol.* **211**, 633-644.

- Duerring, M., Schmidt, G.B. and Huber, R. (1991) Isolation, crystallization, crystal structure analysis and refinement of constitutive C-phycocyanin from the chromatically adapting cyanobacterium *Fremyella diplosiphon* at 1.66 Å resolution. *J. Mol Biol.* **217,** 577-592.

- Dupont, C., Kluepfel, D. and Morosoli, R. (1996) In Lysozymes Model enzymes in Biochemistry and Biology (ed. P. Jollès) pp. 411-423.

- Eftink, M.R. and Hagaman, K.A. (1985) Fluorescence quenching of the buried tryptophan residue of cod parvalbumin. *Biophys Chem.* **22,** 173-180.

- Eyzaguirre, J. (1986) In *Chemical Modification of Enzymes*, Active Site Studies, (Ed. Eyzaguirre, J.) John Wiley and Sons, N.

- Facchiano, A.M., Colonna, G. and Ragone, R. (1998) Helix stabilizing factors and stabilization of thermophilic proteins: an X-ray based study. *Protein Eng.* **11,** 753–760.

- Farrell, R.L. and Skerker, P.S. (1992) *In Xylan and Xylanases Prog. Biotechnol. vol 7* (Visser, J., Beldman, G., Kusters-van Someren, M .A., and Voragen, A. G.J., eds) pp 1-15, Elsevier, Amsterdam.

- Feher, G. and Kam, Z. (1985) Nucleation and growth of protein crystals: general principles and assays. *Methods Enzymol.* **114,** 77-112.

- Ficner, R. and Huber, R. (1993) Refined crystal structure of Phycoerythrin from *Porphyridium cruentum* at 0.23-nm resolution and localization of the gamma subunit. *Eur. J. Biochem.* **218,**103-106.

- Ficner, R., Lobeck, K., Schmidt, G. and Huber, R. (1992) Isolation, crystallization, crystal structure analysis and refinement of B-phycoerythrin from the red alga *Porphyridium sordidum* at 2.2 Å resolution. *J. Mol. Biol.* **228,** 935-950.

- Fisher, R.G., Woods, N.E., Fuchs, H.E. and Sweet, R.M. (1980) Three-dimensional structures of C-phycocyanin and B-phycoerythrin at 5-Å resolution. *J. Biol. chem.* **255,** 5082- 5089.

- Forster, T. (1965) Delocalized excitation and excitation transfer, in sinanoglu O. (Ed.), Modern Quantum chemistry. Vol. III, pp. 93-137. Academic press, New York.

- Frank, G., Sidler, W., Widmer, H. and Zuber, H. (1978) The complete amino acid sequence of both subunits of C-phycocyanin from the cyanobacterium *Mastigocladus laminosus*. *Hoppe Seyler Z. Physiol. Chem.* **359,** 1491-1507.

- Fuglistaller, P., Suter, F. and Zuber, H. (1983) The complete amino acid sequence of both subunits of phycoerythrocyanin from the thermophilic cyanobacterium *Mastigocladus laminosus. Hoppe Seyler Z. Physiol. Chem*.**364,** 691-712.

- Fujimoto, Z., Kuno, A., Kaneko, S., Yoshida, S., Kobayashi, H., Kusakabe, I. and Mizuno, H. (2000) Crystal structure of *Streptomyces olivaceoviridis* E-86 beta-xylanase containing xylan-binding domain. *J. Mol. Biol.* **300,** 575–585.

- Furnham, N., Dore, A.S., Chirgadze, D.Y., de Bakker P.I., Depristo, M.A. and Blundell, T.L.(2006). Knowledge-based real-space explorations for low-resolution structure determination. *Structure.* **14,** 1313-1320.

- Fushinobu, S., Ito, K., Konno, M., Wakagi, T. and Matsuzawa, H. (1998) Crystallographic and mutational analyses of an extremely acidophilic and acid-stable xylanase: biased distribution of acidic residues and importance of Asp37 for catalysis at low pH. *Protein Eng.* **11,** 1121–1128.

- Gantt, E. (1989) Phycobilisomes, Encyclopedia of Plant Physiology, Pirson A. Gottingen. Zimmermann MH. and Harvard. eds, Springer Verlag, Germany.

- Gantt, E. (1990) Biology of Red Algae, edited by K.M.Coles & R.G.Sheath, Cambridge University Press. pp. 203- 219.

- Gantt, E. and Conti, S.F. (1965) The ultrastructure of *Porphyridium cruentum, J. Cell Biol.* **26,** 365-38l.

- Gantt, E. and Conti, S.F. (1969) Ultrastructure of blue-green algae, *J. Bacteriol.* **97,**1486-1493.

- Gantt, E. and Lipschultz, C.A. (1973) Energy transfer in phycobilisomes from phycoerythrin to allophycocyanin. *Biochim. Biophys. Acta.* **292,** 858-861.

- Gantt, E. and Lipshultz, C.A.(1977) Probing phycobilisome structure by immuno-electron microscopy, *J. Phycol.* **13,** 185-192.

- Gantt, E., Edwards, M.R. and Provasoli, L. (1971) Chloroplast structure of the Cryptophyceae. Evidence for phycobiliproteins within the intrathylakoidal spaces, *J. Cell Biol.* **48,** 280-290.

- Gantt, E., Lipschultz, C.A. and Zilinskas, B. (1976) Further evidence for a phycobilisome model from selective dissociation, fluorescence emission, immunoprecipitation, and electron microscopy, *Biochem. Biophys. Acta.* **430,** 375-388.

- Garman, E.F. and Schneider, T.R. (1997) Macromolecular Cryocrystallography, *J. Appl. Cryst.* **30,** 211-237.

- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, RD. and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31,** 3784-3788.

- Gawande, P.V. and Kamat, M.Y. (1990) Chemical modification of xylanases: evidence for essential tryptophan and cysteine residues at the active site. *Biochim. Biophys. Acta.* **1041,** 172-177.

- Gawande, P.V. and Kamat, M.Y. (1999) Purification of *Aspergillus sp* xylanase by precipitation with an anionic polymer Eudragit S100**.** *Process Biochem*. **34**, 577-580.

- George, S.P. and Rao, M. (2001) Conformation and polarity of the active site of xylanase I from *Thermomonospora sp.* as deduced by fluorescent chemoaffinity

labeling. Site and significance of a histidine residue. *Eur. J. Biochem.* **268,** 2881-2888.

- George, S.P., Ahmad, A., and Rao, M.B. (2000) A novel thermostable xylanase from *Thermomonospora sp.* influence of additives on thermostability. *Biores. Technol., 78,* 221-224.

- Georis, J., de Lemos Esteves, F., Lamotte-Brasseur, J., Bougnet, V., Devreese, B., Giannotta, F., Granier, B. and Frere, J.M. (2000) An additional aromatic interaction improves the thermostability and thermophilicity of a mesophilic family 11 xylanase: structural basis and molecular study. *Protein Sci.* **9,** 466-475.

- Ghatge, M.S. and Deshpande, V.V. (1993) Evidence for Specific Interaction of Guanidine Hydrochloride with Carboxy Groups of Enzymes/Proteins. *Biochem. Biophys. Res. Commun*., **193**, 979-984

- Giddings, T.H., Wasman, C. and Staehlin, L.A. (1983) Structure of the thylakoids and envelope membranes of the cyanelles of *Cyanophora paradoxa*. *Plant Physiol*.**71,** 409-419.

- Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller, R.C. and Warren, R.A.J. (1991) Domains in L-1,4-glycanases: Se-quence conservation, function and enzyme families, *Microbiol. Rev,* **55**, 303-315.

- Gilliland, G.L., Tung, M., Blakeslee, D.M. and Ladner, J. (1994) The Biological Macromolecule Crystallization Database, Version 3.0: New Features, Data, and the NASA Archive for Protein Crystal Growth Data. *Acta Crystallogr. D.* **50,** 408-413.

- Glazer, A.N. (1982) Phycobilisomes: structure and dynamics. *Annu. Rev. Microbiol.* **36,** 173- 198.

- Glazer, A.N. (1983) Comparative biochemistry of photosynthetic light-harvesting systems. *Annu. Rev. Biochem*. **52,** 125-151.

- Glazer, A.N. (1985) Light harvesting by phycobilisomes. *Annu. Rev. Biophys. Biophys. Chem.* **14,** 47-77.

- Glazer, A.N. (1988) Phycobiliproteins. *Methods Enzymol.* **167,** 303-3 13.

- Glazer, A.N. (1989) Lighr guides. Directional energy transfer in a photosynthetic antenna. *J. Biol. Chem. 264,* l-4.

- Glazer, A.N. and Hixson, C.S. (1977) Subunit structure and chromophore composition of rhodophytan phycoerythrins. *Porphyridium cruentum* B-phycoerythrin and b-phycoerythrin. *J. Biol. Chem.* **252.** 32-42.

- Glazer, A.N. and Stryer, L. (1984) Phycofluor probes. *Trends Biochent. Sci.* **9,** 423-427.

- Glazer, A.N., Chan, C., Willims, R.C., Yeh, S.W. and Clark, J.H. (1985) Kinetics of energy flow in the phycobilisome core. *Science.* **227,** 105l -423.

- Glazer, A.N., Yeh, S.W., Webb, S.P. and Clark, J.H. (1985) Disk-to-disk transfer as the rate-limiting step for energy flow in phycobilisomes. *Science.* **227,** 419-423.

- Gosalbes, M.J., Perez-Gonzalez, J.A., Gonzalez, R. and Navarro, A. (1991) Two β-glycanase genes are clustered in Bacillus polymyxa: molecular cloning, expression, and sequence analysis of genes encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase. *J. Bacteriol.,* **173**, 7705–7710.

- Grabski, A.C., Jeffries, T.W. (1991) Production, Purification, and Characterization of beta-(1-4)-Endoxylanase of Streptomyces roseiscleroticus. *Appl Environ Microbiol.* **57,** 987-992.

- Green, B.R. (2001) Was "molecular opportunism" a factor in the evolution of different photosynthetic light-harvesting pigment systems? *Proc. Natl. Acad. Sci. (U. S.A.).* **98,** 2119- 2121.

- Green, D.W., Ingram, V.M. and Perutz, M.F. (1954) The structure of haemoglobin IV. Sign determination by the isomorphous replacement method. *Proc. Roy. Soc. Lond., A.* **225**, 287- 307.

- Gross, N. (1996) Extreme enzymes *Business Week,* April 1, 1996.

- Grossman, A.R., Bhaya, D., Apt, K.E. and Kehoe, D.M. (1995) Light-harvesting complexes in oxygenic photosynthesis: diversity, control, and evolution. *Annu. Rev. Genet.* **29,** 231-288.

- Gruber, K., Klintschar, G., Hayn, M., Schlacher, A., Steiner, W. and Kratky, C. (1998) Thermophilic xylanase from *Thermomyces lanuginosus*: high-resolution X-ray structure and modeling studies. *Biochemistry* **37,** 13475–13485.

- Gruninger, H. and Fiechter, A. (1986) A novel, highly thermostable D-xylanase. *Enzyme Microb. Technol.,* **8,** 309-314.

- Guerin-Dumartrait, E. and Moyse, A. (1975) Biological characteristics of the *spirulines*. *Ann. Nutr. Aliment.* **29,** 489-496.

- Gust, D., Moore, T.A., and Moore, A.L. (2001) Mimicking photosynthetic solar energy transduction. *Acc. Chem. Res.* **34.** 40-48.

- Hackert, M.L., Abacl-Zapatero, C., Stevens, S.E. and Fox, L.I. (1977) Crystallization of C-phycocyanin from the marine blue-green alga *Agmenellum quadruplicatum*. *J. Mol. Biol.* **111,** 365-369.

- Haltrich, D., Nidetzky, B., Kulbe, K.D., Steiner, W. and Zupancic, S. (1996) Production of fungal xylanases. *Biores. Technol*., **58,** 137-161.

- Hardy, L.W. and Poteete, A.R. (1991) Reexamination of the role of Asp20 in catalysis by bacteriophage T4 lysozyme. *Biochemistry* **30**, 9457-9463.

- Harris, G.W., Jenkins, J.A., Connerton, I., Cummings, N., Lo Leggio, L., Scott, M., Hazlewood, G.P., Laurie, J.I., Gilbert, H.J. and Pickersgill, R.W. (1994)

Structure of the catalytic core of the family F xylanase from Pseudomonasfluorescens and identification of the xylopentaose-binding sites. *Structure* **2,** 1107–1116.

- Harris, G.W., Pickersgill, R.W., Connerton, I., Debeire, P., Touzel, J.P., Breton, C. and Perez, S. (1997). Structural basis of the properties of an industrially relevant thermophilic xylanase. *Proteins*, **29**, 77–86.

- Hattori, A.and Fujita, Y. (1959) Crystalline phycobilin chromopeptides obtained from a bluegreen alga, *Tolypotluix tenuis. J. Biochem. (Tokyo).* **46,** 633-644.

- Havukainen, R., Torronen, A., Laitinen, T. and Rouvinen, J. (1996) Covalent bindingof three epoxyalkyl xylosides to the active site of endo-1, 4-xylanase II from *Trichoderma reesei*. *Biochemistry,* **35,** 9617–9624.

- Hazlewood, G.P. and Gilbert, H J. (1993) In: *Hemicelluloses and Hemicellulases* pp. 103-126 Coughlan, M. P. and Hazlewood, G. P., Eds., Portland Press, London.

- Hegde Subray, S., Kumar Ameeta, R., Ganesh Krishna, N. and Khan Islam, M. (1998) Catalytic thiol and carboxylate: role of cysteine and glutamic acid in the xylosidic activity of endoxylanase from *Chainia* sp. (NCL 82-5-1). *Arch Biochem Biophys* **355,** 154–159.

- Heightman, T.D.and Vasella, A.T. (1999) Recent insights into inhibition, structure and mechanism of configuration-retaining glycosidases. *Angew. Chem.* Int Ed, **38**, 750-770.

- Hendrickson, W.A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science.* **254,** 5l-58.

- Hendrickson, W.A., Horton, J.R. and LeMaster, D.M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD):

a vehicle for direct determinarion of three-dimensional strucrure. *EMBO J.* **9,** 1665-1672.

- Hendrickson, W.A., Smith, J.L. and Sheriff, S. (1985) Direct phase determination based on anomalous scattering. *Methods Enzymol.* **115,** 41-55.

- Henrissat, B., Claeyssens, M., Tomme, P., Lemesle, L. and Mornon, J.P. (1989) Cellulase families revealed by hydrophobic cluster analysis. Gene 81, 83–95.

- Heocha, C.O. (1965) Biliproteins of Algae. *Ann. Rev. Plant. Physiol***.76**, 415-434.

- Higgins, D., Thompson, J., Gibson, T., Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res* **22,** 4673–4680.

- Horikoshi, K. and Atsukawa, Y. (1973) Xylanase produced by alkalophilic *Bacillus* No. C-59-2, *Agric. Biol. Chem,* **37**, 2097-2103.

- Hu, X. and Schulten, K. (1997) How nature harvests sunlight. Physics Today. **50,** 28-34.

- Huber, R. in Nobel Lectures in Chemistry 1981-1990. edited by B. G. Molmstrom. (World Scientific, London 1991).

- Huber, R. Nobel lecture. (1989) A structural basis of light energy and electron transfcr biology. *EMBO J.* **8,** 2125-2111.

- Huber, R. Nobel lecture. (1989) A structural basis of light energy and electron transfer biology. *Biosci. Rep*. **9,** 635-613.

- Jancarik, J. and Kim, S. (1991) Sprase matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* **24,** 409-411.

- Jiang, T., Zhang, J. and Liang, D. (1999) Structure and function of chromophores in R-Phycoerythrin at 1 .9 Å resolution. *Proteins.* **34,** 224-231.

- Kamal Kumar, B., Balakrishnan, H. and Rele, M.V. (2001) Compatibility of alkaline xylanases from an alkaliphilic *Bacillus* NCL (87-6-10) with commercial detergents and proteases. *J. Ind Microbiol Biotechnol.* **31,** 83-87.

- Kannan, N. and Vishveshwara, S. (2000) Aromatic clusters: a determinant of thermal stability of thermophilic proteins, *Protein Eng.*, **13**, 753–761.

- Karshikoff, A. and Ladenstein, R. (1998) Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. *Protein Eng.*, **11**, 867–872.

- Katusbe, Y., Hata, Y., Yamaguchi, H., Moriyama, H., Shinmyo, A. and Okuda, H. (1990) Estimation of the xylanase active site from crystalline structure. *Protein Eng*, **3**, 289–290.

- Kelly, C.T., O' Matory, M.R., and Fogarty, W.M. (1989) Extracellular xylanolytic enzymes of *Paecilomyces varioti* . *Biotechnol. Letts.* **13**, 885-890.

- Keskar, S.S., Srinivasan, M.C. and Deshpande, V.V. (1989) Chemical modification of a xylanase from a thermotolerant *Streptomyces.* Evidence for essential tryptophan and cysteine residues at the active site, *Biochem. J.* **261**, 49-55.

- Keskar, SS., Rao, MB., Deshpande, VV. (1992). Characterization and sequencing of an active-site cysteine-containing peptide from the xylanase of a thermotolerant *Streptomyces. Biochem J.* **281,** 601-605.

- Khasin, A., Alchanti, I. and Shoham, Y. (1993) Purification and characterization of a thermostable xylanase from *Bacillus stearothermophilus* T-6. *Appl. Environ. Microbiol*,. **59,** 1725-1730.

- Kleywegt, G.J. (1996) Use of Non-crystallographic symmetry in protein structure refinement. *Acta Crystallogr. D.* **52,** 842-857.

- Kleywegt, G.J. and Brunger, A.T. (1996) Checking your imagination: Applications of the free R value. *Structure.* **4,** 897-904.

- Kleywegt, G.J. and Jones, T.A. (1998) Databases in protein crystallography. *Acta crystallogr. D.* **54,** 1119 - 1131.

- Knox, R.S. (1975) In Bioenergetics of photosynthesis (Govindjee. ed.), pp. l83-2211. Academic Press. New York.

- Ko, E.P., Akatsuka, H., Moriyama, H., Shinmyo, A., Hata, Y., Katsube, Y., Urabe, I. and Okada, H. (1992) Site-directed mutagenesis at aspartate and glutamate residues of xylanase from *Bacillus pumilus. Biochem J.* 288, 117-121.

- Krengel, U.B and Dijkstra, W. (1996) Three dimensional structure of *endo*-1,4-β-xylanase I from *Aspergillus niger*: molecular basis for its low pH optimum, *J. Mol Biol*,**263,**70–78.

- Krishnamurthy, S. and Vithayathil, P.J. (1989) Purification and characterization of endo-1,4-β-xylanase from *Paecilomyces varioti* banier, *J. Ferment. Bioeng.*, **67**, 77-82.

- Kubuckova, M., Kuracsonyi, S., Bilisics, L. and Toman R. (1978) Some properties of an *endo*-1,4-β-xylanase from the ligniperdous fungus *Trametes hirsute*. *Folia Microbiol,* **23**, 202–209.

- Kumar, P.R., Eswaramoorthy, S., Vithayathil, P.J. and Viswamitra, M.A. (2000) The tertiary structure at 1.59 Å resolution and the proposed amino acid sequence of a family-11 xylanase from the thermophilic fungus *Paecilomyces varioti* bainier. *J. Mol. Biol.* **295,** 581–593.

- Kumar, S., Tsai, C.J. and Nussinov, R. (2000) Factors enhancing protein thermostability. *Protein Eng.* **13,** 179–191.

- Kylin, H. (1910) Uber Phycoerytrhin und phycocyanin bei *Ceramium rubrum* (Huds.) *Ag.Hoppe Seyler's Z. Physiol. Chem.* **69**, 169-239.

- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* **227**, 680-685.

- Lemberg, R. (1928) Die Ctronropeptide der Rotalgen. *Ann. Chem.* **461**, 46-89.

- Lin, L.L. and Thomson, J.A. (1991) An analysis of the extracellular xylanases and cellulases of *Butyrivibrio fibrisolvens* H17c. *FEMS Microbiol Lett.* **68,** 197-203.

- Lin, L.L. and Thomson, J.A. (1992) Characterization and sequencing of an active-site cysteine-containing peptide from the xylanase of a thermotolerant *Streptomyces*. *Biochem J.* **281,** 601-605.

- Liu, J.Y., Jiang T, Zhang, J.P. and Liang, DC. (1999) Crystal structure of allophycocyanin from red algae *Porphyra yezoensis* at 2.2-Å resolution. *J. Bio. Chem.***274,** 16945-16952.

- Lo Leggio, L., Kalogiannis, S., Bhat, M.K. and Pickersgill, R.W. (1999) High resolution structure and sequence of T. aurantiacus xylanase I: implications for the evolution of thermostability in family 10 xylanases and enzymes with (beta)alpha-barrel architecture. *Proteins* **36,** 295–306.

- Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951) Protein measurement with the Folin phenol reagent. *J. Biol Chem.* **193,** 265-275.

- Luthi, E., Bhana, J.M. and Bergquist, P.L. (1990) Xylanase from the extremely thermophilic bacterium *Caldocellum saccharolyticum*: over expression of the gene in Escherichia coli and characterization of the gene product. *Appl. Environ. Micribiol,* **56**, 2677-2683.

- Luthi, E., Reif, K., Jasmat, NB. and Bergquist, PL. (1992) In vitro mutagenesis of a xylanase from the extreme thermophile *Caldocellum saccharolyticum*. *Appl Microbiol Biotechnol.* **36,** 503-506.

- Luzzatt, P.V. (1952) Traitment Statistique des Erreurs dans la Determination des Structures Cristallines, *Acta crystallogr.* **5,** 802-810.

- Maat, J., Roza, M., Verbakel, J., Stam, H., Santos da Silva, M.J., Bosse, M., Egmond, M.R., Hagemans, M.L.D., Gorcom, R.F.M., Hessing, J.G.M., Hondel, C. and Rotterdam, C. (1992) In: *Xylan and Xylanases* Prog. Biotechnol. vol 7 (Visser, J., Beldman, G., Kusters-van Someren, MA., and Voragen, AGJ., eds. pp 349-360, Elsevier, Amsterdam.

- MacColl, R. (1998) Cyanobacterial phycobilisomes. *J. Struct. Bio*. **124**, 311-334.

- MacColl, R. and Guard-Friar, D. (1987) Phycobiliproteins, CRC Press, Boca Raton, FL.

- MacColl, R., Williams, O., Eisele, L.E. and Berns D.S. (1994) Spectroscopic changes for C-phycocyanin and phycoerythrin 545 produced by ferric ion. *Biochem Biophys. Acta*. **118**, 398- 404.

- MacLeod, AM., Lindhorst, T., Withers, S.G. and Warren, R.A. (1994) The acid/base catalyst in the exoglucanase/xylanase from *Cellulomonas fimi* is glutamic acid 127: evidence from detailed kinetic studies of mutants. *Biochemistry.* **33,** 6371-6376.

- Main, P. (1979) A theoretical comparison of B.G and 2Fo-Fc syntheses. *Acta Crystallogr. A***. 35**, 779- 785.

- Maio, S., Ziser, L., Aebersold, R. and Withers, S.G. (1994) Identification of glutamic acid 78 as the active site nucleophile in *Bacillus subtilis* xylanase using electrospray tandem mass spectrometry. *Biochemistry.* **33,** 7027– 7032.

- Malkin, A.J.(1995a). Mechanisms of growth for protein and virus crystals. *Nat. Struct. Biol*. **2**, 956.

- Malkin, A.J.(1995b). Investigation of virus crystal growth by in situ atomic force microscopy. *Phys. Rec. Lett.* **75**, 2778.

- Marui, J., Kato, M., Kobayashi, T. and Tsukagoshi, N. (2003) Upregulation of promoter activity of the *Aspergillus oryzae* xylanase gene by site-directed mutagenesis. *Biotechnol Lett.* **25,** 371-374.

- Matthews, B.W. (1968) Solvent content of protein crystals, *J. Mol. Biol.* **33**, 49l-497.

- Maxson, P., Sauer, K., Zhou, J.H., Bryant, D.A. and Glazer, A.N., (1989) Spectroscopic studies of cyanobacterial phycobilisomes lacking core polypeptides. *Biochim. Biophys. Acta.* **977,** 40- 51.

- McAllister, T.A., Inglis, G.D., Popp, A.P., Selinger, L.B., Kawchuk, L.M. and Gaudet, D.A. (2000) Production of cellulases and xylanases by low-temperature basidiomycetes. *Can J. Microbiol.* **46,** 860-865.

- McCarter, J.D. and Withers, S.G. (1994) Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol*., **4**, 885–892.

- McCarthy, A.A., Morris, D.D., Bergquist, P.L. and Baker, E.N. (2000) Structure of XynB, a highly thermostable beta-1,4- xylanase from *Dictyoglomus thermophilum* Rt46B.1, at 1.8 Å resolution. *Acta Crystallogr. D*. **56,** 1367–1375.

- McCarthy, A.J., Peace, E. and Broda, P. (1985) Studies on the extracellular xylanase activity of some thermophilic actinomycetes. *Appl. Microbiol. Biotechnol.* **21**, 238-244.

- McCleary, B.V. (1986) Enzymatic modification of plant polysaccharides. *Int. J. Macromol.*, **8**, 349-354.

- McIntosh, L.P., Hand, G., Johnson, P.E., Joshi, M.D., Korner, M. and Plesniak, L.A. (1996) The p*K*a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a 13C NMR study of *Bacillus circulans* xylanase. *Biochemistry*, 35, 9958–66.

- McPherson, A. (1976) The growth and preliminary investigation of protein and nucleic acid crystals for X-ray diffraction techniques. *Methods Biochem. Anal*. **23**, 2249.

- McPherson, A. (1999) Crystallization of Biological Macromolecules. Eds. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

- Means, G.E., Feeney, R.E. (1971) Affinity labeling of pancreatic ribonuclease. *J. Biol Chem.* **246,** 5532-5533.

- Mechaly, A., Teplitsky, A., Belakhov, V., Baasov, T., Shoham, G. and Shoham, Y. (2000) Overproduction and characterization of seleno-methionine xylanase T-6. *J. Biotechnol*. **78**, 83–86.

- Menendez-Arias, L. and Argos, P. (1989) Engineering protein thermal stability. Sequence statitistics point to residue substitutions in alpha-helices. *J. Mol. Biol.,* **206**, 397–406.

- Miao, S., Ziser, L., Aebersold, R. and Withers, S.G. (1994) Identification of glutamic acid 78 as the active site nucleophile in *Bacillus subtilis* xylanase using electrospray tandem mass spectometry. *Biochemistry,* **33,** 7027-7032.

- Michael Burke, J., Douglas Pratt, C. and Albert Moscowitz. (1972) Low-temperature absorption and circular dichroism studies of phytochrome. *Biochemistry.* **11,** 4025-4031.

- Miki, K., Ezoe, T., Masui, A., Yoshisaka, T., Mimuro, M., Fujiwara Arasaki, T. and Kasai, N. (1990) Crystallization and preliminary X-ray diffraction studies of C-phycocyanin Porphyra tenera. *J. Biochem.* **108**, 646-649.

- Mimuro, M., Fuglistaller, P., Rumbeli, R. and Zuber, H. (1986) Functional assignment of chromophores and energy transfer in C-phycocyanin isolated from the thermophilic cyanobaterium *Mastigocladus luminosus*. *Biochim. Biophys. Acta*. **848**, 155-166.

- Mishra, C., Keskar, S. and Rao, M. (1984) Production and properties of extracellular endoxylanases from *Neurospora crassa. Appl. Environ. Microbiol.* **48,** 224-228.

- Morcos, N.C., and Henry, W.L. (1992) Medical treatment of tumors with phycocyanin , united States Patent# 05I 63898.

- Moreau, A., Shareck, F., Kluepfel, D. and Morosoli, R. (1994) Increase in catalytic activity and thermostability of the xylanase A of *Streptomyces lividans* 1326 by site-specific mutagenesis. *Enzyme Microb. Technol.,* **16,** 420-424.

- Morisset, W., Wehrmeyer, W., Schirmer, T. and Bode W. (1984) Crystallization and preliminary ray diffraction data of the cryptomonad biliprotein phycocyanin-645 from a *Chroomonas Arch. Microbiol.* **140**. 202-205.

- Moriyama, H., Hata, Y., Yamaguchi, H., Sato, M., Shinmyo, A., Tanaka, N., Okada, H. and Katsube, Y. (1987) Crystallization and preliminary X-ray studies of *Bacillus pumilus* IPO xylanase. *J. Mol. Biol.*, **193,** 237-238.

- Morosoli, R.J., Bertrand, J.L., Mondou, F., Shareck, F. and Klupfel, D. (1986) Purification and properties of a xylanases from *Streptomyces lividans*. *Biochem. J*. **239**, 587-592.

- Muilu, J., Torronen, A., Perakyla, M. and Rouvinen, J. (1998) Functional conformational changes of endo-154-xylanase II from *Trichoderma reesei*: a molecular dynamics study. *Proteins* **31,** 434–444.

- Murshudov, G.N, Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallogr. D*, **53**, 240–55.

- Nakamura, S., Wakabayashi, K., Nakai, R., Aono, R. and Horikoshi K. (1993) Purification and some properties of an alkaline xylanase from alkaliphilic *Bacillus sp.* strain 41M-1. *Appl Environ Microbiol.* **59,** 2311-2316.

- Nakanishi, K., Kubo, I. and Miura, I. (1976) The structure of xylomollin, a secoiridoid hemiacetal acetal. *J. Am Chem Soc.* **98,** 6704-6705.

- Namchuk, M.N., and Withers, S.G. (1995) Mechanism of *Agrobacterium* 6-glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate Interactions, *Biochemistry*, **34,** 16194-l 6202.

- Natesh, R., Bhanumoorthy, P., Vithayathil, P.J., Sekar, K., Ramakumar, S. and Viswamitra, M.A. (1999) Crystal structure at 1.8 Å resolution and proposed amino acid sequence of a thermostable xylanase from *Thermoascus aurantiacus. J. Mol. Biol.* **288,** 999–1012.

- Navaza, J. (1994) AMoRe: An automated packagc for molecular replacement. *Acta Crystallogr. A.* **50**, 157-163.

- Navaza, J. and Saludjian, P. (1997) AMoRe: An automated molecular replacement program package. *Methods Enzymol.* **276**, 581-594.

- Nguyen, T.Q., Wu, J., Doan V.V., Schwartz, B.J. and Tolbert, S.H. (2000) Control of energy transfer in oriented conjugated polymer-mesoporous silica composites. *Science.* **288,** 652-656.

- Nield, J., Bibby, T.S., Chen, M., Larkum, A.W. and Barber, J. (2003) Structure of a photosystem II supercomplex isolated from *Prochloron didemni* retaining its chlorophyll a/b light-harvesting system. *Proc Natl Acad Sci U S A.* **100,** 9050-9054.

- Nield, J., Rizkallah, P.J., Barber, J. and Chayen, N.E. (2003) The 1.45 A three-dimensional structure of C-phycocyanin from the thermophilic cyanobacterium *Synechococcus elongatus. J. Struct. Biol.* **14,** 149-155.

- Nikolova, P.V., Creagh, A.L., Sheldon, J.B., Du¡, C.A. and Haynes, D. (1997) Thermostability and irreversible activity loss of exoglucanase/xylanase Cex from *Cellulomonas fimi. Biochemistry*, **36**, 1381-1388.

- Notenboom, V. and Withers, S.G. (1998) Insights into transition state stabilisation of the α-l ,4 glycosidase Cex by covalent intermediate accumulation in active site mutants. *Nat. Struct. Biol.* **5**, 812-818.

- Oakley, A.J., Heinrich, T., Thompson, C.A. and Wilce, M.C. (2003) Characterization of a family 11 xylanase from *Bacillus subtillis* B230 used for paper bleaching. *Acta Crystallogr. D.* **59,** 627–636.

- Ohkushi, A., Kudo, T., Mase, T. and Horikoshi, K. (1985) *Agric. Biol. Chem.* **49**, 3037-3038.

- Ohmiya, K., Sakka, K., Karita, S. and Kimura, T. (1997) Structure of cellulases and their applications. *Biotechnol. Genet. Eng. Rev.* **14,** 365-414.

- Okazaki, W., Akiba, T., Horikoshi, K. and Akahoshi, E. (1984) Production and properties of two types of xylanases from alkalophilic thermophilic *Bacillus sp. Appl. Microbiol. Biotechnol.* **19**, 335-340.

- Otwinowski, Z. (1993) *Proceedings of the CCP4 study weekend*. 56-62.

- Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol***. 276**, 307 -326.

- Padyana, A.K., Bhat, V.B., Madyastha, K.M., Rajashankar, K.R. and Ramakumar, S. (2001) Crystal Structure of a Light-Harvcsting Protein C-Phycocyanin from *Spirulina platensis*. *Biochem. Biophys. Res. Commun.* **282**, 893-898.

- Paice, M.G., Bernier, R Jr. and Jurasek, L. (1988) Viscosityenhancing bleaching of hardwood kraft pulp with xylanase from a cloned gene. *Biotechnol. Bioeng.*, **32**, 235-239.

- Panbangred, W.A., Shinmyo, S., Kinoshita, S. and Okada, H. (1983) Molecular cloning of a xylanase fom *Bacillus pumilus* in Escherichia coli. *Mol. Gen. Genet.*, **192**, 335-341.

- Pannu, N.S. and Read, R.J. (1996) Improved structure refinement through maximum likelihood. *Acta Crystallogr. A*. **52,** 659-668.

- Pastore, A. and Lesk, A.M. (1990) Comparison of the.structures of globins and phycocyanins: evidence for evolutionary relationship. *Proteins.* **8,** 133-155.

- Paul, J. and Varma, A.K. (1990) Influence of sugars on endoglucanase and beta-xylanase activities of a *Bacillu*s strain. *Biotechnol. Lett.*, **12**, 61-64.

- Pederson, L. S. (1989) On the use of Pulpenzyme for bleach boosting, Novo.

- Pell, G., Szabo, L., Charnock, S.J., Xie, H., Gloster, T.M., Davies, G.J. and Gilbert, H.J. (2004) Structural and biochemical analysis of Cellvibrio japonicus xylanase 10C: how variation in substrate-binding cleft influences the catalytic profile of family GH-10 xylanases. *J. Biol. Chem.* **279,** 11777–11788.

- Pell, G., Taylor, E.J., Gloster, T.M., Turkenburg, J.P., Fontes, C.M., Ferreira, L.M., Nagy, T., Clark, S.J., Davies, G.J. and Gilbert, H.J. (2004) The mechanisms by which family 10 glycoside hydrolases binddecorated substrates. *J. Biol. Chem.* **279,** 9597–9605.

- Pereira, N Jr., Lemos, J.L. and Fontes, M.C. (2001) Xylanase production by Aspergillus awamori in solid-state fermentation and influence of different nitrogen sources. *Appl Biochem Biotechnol.* **93,** 681-689.

- Peter Biely., Lubomír Kremnický., Juraj Alföldi. and Maija Tenkanen. (1985) Microbial xylanolytic systems. *Trends Biotech.* **3,** 286-290.

- Pflugrath, J.W. (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr. D.* **55,** 1718-1725.

- Phillips, S.E. and Schoenborn, B.P. (1981) Neutron diffraction reveals oxygen-histidine hydrogen bond in oxyrnyoglobin. *Nature.* **292,** 81.

- Pickersgill, R.W., Debeire, P., Debeire-Gosselin, M. and Jenkins, J.A. (1993) Crystallization and preliminary analysis of a thermophilic *Bacillus* xylanase. *J. Mol Biol*, **230**,664–666.

- Pizarro, S.A. and Sauer, K. (2001) Spectroscopic study of the light-harvesting prorein C-phycocyanin associated with colorless linker peptides. *Photochem. Photobiol*. **73**, 556-563.

- Polson, A., Potgieter, G.M. (1964) Diffusion constants of proteins determined by a multi-unit analytical method. *Nature.* **24,** 204-379.

- Ponder, J.W. and Richards, F.M. (1987) Internal packing and protein structural classes. *Cold Spring Harb Symp Quant Biol.* **52,** 421-428.

- Ponder, J.W. and Richards, F.M. (1987) Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J Mol Biol.* **193,** 775-791.

- Poon, D.K., Webster, P., Withers, S.G. and McIntosh, L.P. (2003) Characterizing the pH-dependent stability and catalytic mechanism of the family 11 xylanase from the alkalophilic *Bacillus agaradhaerens. Carbohydr. Res.* **338,** 415-421.

- Ramachandran. G. N. and Sasishekaran. V. (1968). Conformation of polypeptides and proteins. *Adv. Protein. Chem*., **23**; 283-438.

- Rao, M., Khadilkar, S., Bandivadekar, K. and Deshpande, V. (1996) Structural environment of an essential cysteine residue of xylanase from *Chainia sp*. (NCL 82.5.1). *Biochem. J*., **316**, 771-775.

- Ratanakhanokchai, K., Kyu, K.L. and Tanticharoen, M. (1999) Purification and properties of a xylan-binding endoxylanase from alkaliphilic *Bacillus sp.* strain K-1. *Appl. Environ. Microbiol.*, **65,** 694–697.

- Read, R.J. (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A*. **42,**140- 149.

- Read, R.J. (1997) Model phases: probabilities and bias. *Methods Enzymol*. **278**, 110-128.

- Reuter, W., Wiegand, G., Huber, R. and Than, M.E. (1999) Structural analysis at 2.2Å of orthorhombic crystals presents the asymmetry of the allophycocyanin-linker complex, AP.LC 7.8, from phycobilisomes of *Mastigocladus laminosus*. *Proc. Natl. Acad. Sci. (U.S.A).* **96**,1363-1368.

- Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M. and Stanier, R.Y. (1979) Generic assignments, strain histories and properties of the pure culture of cyanobacteria*, J. Gen. Microbiol.* **111,** 1–61.

- Ritter, S., Hiller, R.G., Wrench, P.M., Welte, W. and Diederichs, K. (1999) Crystal structure of a phycourobilin-containing phycoerythrin at 1.90Å resolution. *J.Struct. Biol.* **126**, 86-97.

- Roberge, M., Shareck, F., Morosoli, R., Kluepfel, D. and Dupont, C. (1997) Characterization of two important histidine residues in the active site of xylanase A from *Streptomyces lividans*, a family 10 glycanase. *Biochemistry.* **36,** 7769-7775.

- Romay, C., Armesto, J., Remirez, D., Gonzalez, R., Ledon, N. and Garcia, I. (1998) Antioxidant and anti-inflammatory properties of C-phycocyanin from blue-green algae. *Inflammation Research.* **47**, 36-41

- Rossmann, M.G. (1972) In The Molecular replacement method. A collection of papers on the use of NON-CRYSTALLOGRAPHIC Symmetry. Ed. Rossmann MG. Gordon & Breach, Science Publishers, Inc. New York.

- Rossmann, M.G. and Blow, D.M. (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr.* **15**, 24-31.

- Rudiger. W. (1992) Events in the phytochrome molecule after irradiation. *Photochem. Photobiol*. **56**, 803-809.

- Rumbeli, R., Suter, F., Wirth, M., Sidler, W. and Zuber, H. (1987a). Gamma-N-Methylasparagine in phycobiliproteins of the thermophilic cyanobacterium *Mastigocladus laminosus* and *Calothrix. FEBS Lett,* **221**.1-2.

- Rumbeli, R., Suter, F., Wirth, M., Sidler, W. and Zuber, H. (1987b) Isolation and localization of N-4-gamma-methylasparagine in phycobiliproteins of the thermophilic cyanobacterium *Mastigocladus laminosus*. *Biol. Chem.* Hoppe-Sevler. **368**, 1401-1406.

- Russell, R.J.M., Ferguson, J.M.C., Hough, D.W., Danson, M.J.and Taylor, G.L. (1997) The crystal structure of citrate synthase from hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry* **36,** 9983–9994.

- Sabini, E., Sulzenbacher, G., Dauter, M., Dauter, Z., Jorgensen, P.L., Schulein, M. (1999) Catalysis and specificity in enzymatic glycoside hydrolysis: a 2,5 B conformation for the glycosyl–enzyme intermediate revealed by the structure of the *Bacillus agaradhaerens* family 11 xylanase. *Chem. Biol*. **6,** 483–492.

- Sabini, E., Wilson, K.S., Danielsen, S., Schulein, M. and Davies, G.J. (2001). Oligosaccharide binding to family 11 xylanases: both covalent intermediate and mutant product complexes display (2, 5) B conformations at the active centre. *Acta Crystallogr. D.* **57,** 1344–1347.

- Satyanarayana, L., Suresh, C.G., Anamika Patel, Sandhya Mishra, and Pushpito Kumar Ghosh.(2005) X-ray crystallographic studies on C-phycocyanins from cyanobacteria from different habitats: marine and freshwater. *Acta Crystallgr. F.* **61,** 844–847.

- Sauer, K. (1975) in Bioenergetics of Photosynthesis (Govindjee, ed.), pp. 115-181. Academic Press. New York.

- Sauer, K. and Scheer. (1988) Excitation transfer in C-Phycocyanin. Forster rate and exciton calculations based on new crystal structure data for C-phycocyanins

from *Agmenellum quadruplicatum* and *Mastigocladus laminosus*. *Biochim. Biophys. Acta.* **936**, 157-170.

- Sauer, K., Scheer, H. and Sauer, P. (1987) Forster Transfer Calculations Based on Crystal Structure Data from *Agmenellum quadruplicatum* C-Phycocyanin. *Photochem. Photobiol*. **46**, 427-440.

- Saul, D. J., Williams, L.C., Reeves, R.A., Gibbs, M.D. and Bergquist, P.L. (1995) Sequence and expression of a xylanase gene from the hyperthermophile *Thermotoga sp.* strain FjSS3-B.1 and characterization of the recombinant enzyme and its activity on kraft pulp *Appl. Environ. Microbiol.* **61**, 4110-4113.

- Schirmer, T. and Vincent, M.G. (1987) Polarized absorption and fluorescence spectra of single crystals of C-phycocyanin *Biochim. Biophys. Acta*. **893**, 379-385.

- Schirmer, T., Bode, W. and Huber, R. (1987) Refined three-dimensional structures of two cyanobacterial C-phycocyanins at 2.1 and 2.5Å resolution. A common principle of phycobilin-protein interaction. *J.Mol. Biol*. **196**, 677- 695.

- Schirmer, T., Bode, W., Huber, R., Sidler, W. and Zuber, H. (1985) X-ray crystallographic structure of the light-harvesting biliprotein C-phycocyanin from the thermophilic cyanobacterium *Mastigocladus laminosus* and its resemblance to globin structures. *J. Mol. Biol.* **184,** 257- 277.

- Schirmer, T., Huber, R., Schneider, M., Bode, W., Miller, M. and Hackert, M.L. (1986) Crystal structure analysis and refinement at 2.5 Å of hexameric C-phycocyanin from the cyanobacterium *Agmenellum quadruplicatum*. The molecular model and its implications for light harvesting. *J. MoI. Biol.* **188,** 651-676.

- Schmidt, A., Schlacher, A., Steiner, W., Schwab, H.and Kratky, C. (1998) Structure of the xylanase from *Penicillium simplicissimum*. *Protein Sci.* **7,** 2081–2088.

- Shao, W., DeBlois, S. and Wiegel, J. (1995) A high-molecular- weight, cell-associated xylanase isolated from exponentially growing Thermoanaerobacterium sp. strain JW/SLYS485. *Appl. Environ. Microbiol.* **61**, 937-940.

- Shibuya, H., Kaneko, S. and Hayashi, K. (2000) Enhancement of the thermostability and hydrolytic activity of xylanase by random gene shuffling *Biochem J*. **349,** 651-656.

- Sidler W., Fuglistller, P., Gysi, J., Isker, E. and Zuber, H. (1981) In photosynthesis III. Structure and Molecular organization of the Photosynthetic Apparatus (Akoyunoglou G.ed.), pp: 583-594, Balaban lnternational Science Services, Philadelphia.

- Sidler, W., Gysi, J., Isker, E. and Zuber, H. (1981) The complete amino acid sequence of both subunits of allophycocyanin: a light harvesting protein-pigment complex from the cyanobacterium *Mastigocladus laminosus*. Hoppe Seyler Z. *Physio. Chem.* **362**, 611-628.

- Siiman, O., Wilkinson, J., Burshreyn, A., Roth, P. and Ledis. (1999) Fluorescent neoglycoproreins: antibody-aminodextran-phycobiliprotein conjugates. *Bioconjug Chem.***10**, 1090-1106.

- Simpson, H.D., Haufer, U.R. and Daniel, RM. (1991) An extremely thermostable xylanase from the thermophilic eubacterium *Thermotoga. Biochem. J***. 277**, 413-417.

- Sinha, U. and Brewer, J.M. (1985). A spectrophotometric method for quantitation of carboxyl group modification of proteins using Woodward's Reagent K. *Anal Biochem.* **151**, 327-333.

- Sinnot, M.L. (1990) Catalytic mechanisms of enzymic glucosyl transfer. *Chem. Rev*., **90**, 1171-1202.

- Spande T.F, and Witkop B. (1967) Determination of the tryptophan content of proteins with *N* -bromosuccinimide. *Methods Enzymol*, **11**,498–506.

- Spurway, T.D., Morland, C., Cooper, A., Sumner, I., Hazlewood, G.P., O'Donnell, A.G., Pickersgill, R.W. and Gilbert, H.J. (1997). Calcium protects a mesophilic xylanase from proteinase inactivation and thermal unfolding. *J. Biol Chem.* **272,** 17523-17530.

- Stec, B., Troxler, R.F. and Teeter, M.M. (1999) Crystal structure of C-phycocyanin from cyanidium caldarium provides a new perspective on phycobilisome assembly. *Biophys J.* **76**, 2912- 2921.

- Stutzenberger, F.J. and Bodine, A.B. (1992) Xylanase production by *Thermomonospora curvata. J. Appl. Bacteriol.* **72**, 504-511.

- Sung, W. L., Yaguchi, M. and Ishikawa, K. (1998) US Patent 5759840.

- Sung, W.L. and Tolan, J.S. (2000) Thermostable xylanases. WO Patent 00/29587.

- Sung, W.L., Yaguchi, M. and Ishikawa, K. (1998) US Patent 5759840.

- Sung, W.L., Yaguchi, M., and Ishikawa, K. (1998) Modification of xylanase to improve thermophilicity, alkophilicity and thermostability for pulp bleaching. Patent US-5759840,

- Sunna, A. and Antranikian, G. (1997) Xylanolytic enzymes from fungi and bacteria. *Crit Rev Biotechnol.* **17,** 39-67.

- Sunna, A., Prowe, S.G., Stoffregen, T., Antranikian, G. (1997) Characterization of the xylanases from the new isolated thermophilic xylan-degrading *Bacillus thermoleovorans* strain K-3d and *Bacillus flavothermus* strain LB3A. *FEMS Microbiol Lett.* **148,** 209-216.

- Suter, F., Fuglistaller, P., Lundell, D.J., Glazer, AN. and Zuber, H. (1987) Amino acid sequences of alpha-allophycocyanin B from *Syncchococcus* 6301 and *Mastigocladus laminosus*. *FEBS Lett*. **217**, 279-282.

- Svedberg, T., and Katsurai, T. (1929) The molecular weights of phycocyanin and of phycoerythrin from *Porphyra tenera* and of phycocyanin from *Aphanizomenon flos aquae*. *J. Am. Chem. Soc.* **51**, 3573-3583.

- Sweet, R.M. and Eisenberg, D. (1983) Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J. Mol. Biol*. **171**, 479-488.

- Sweet, R.M., Fuchs, H.E., Fisher, R.G., and Glazer, A.N. (1977) Preliminary crystallographic investigations of two phycobiliproreins. *J. Biol. Chem.* **252**, 8258-8260.

- Szabo, A.G., Stepanik, T.M., Wayner, D.M. and Young, N.M. (1983) Conformational heterogeneity of the copper binding site in azurin. A time-resolved fluorescence study. *Biophys. J.,* **41**, 233-244.

- Takenishi, S., Tsujisaka, Y. and Fukumoto, J. (1973). Studies on hemicellulases. IV. Purification and properties of the xylosidase produced by *Aspergillus niger* van Tieghem. *J. Biochem (Tokyo).* **73,** 335-343.

- Tamanoi, H., Kasahara, S., Kuroda, T., Kubo, T., Nakai, R., Namba, K., Wakabaashi, K., and Nkamura, S. 1998. Role of acidic aminoacids in the alkaline pH optimum of xylanase J from alkaline *Bacillus sp. strain* 41M-1. *Nucleic Acids symp*. Ser., **39,**205-206.

- Tan, L.U., Rose, D.R., Birnbaum, G.I. and Saddler, J.N. (1987) Crystallization and preliminary X-ray diffraction study of a xylanase from *Trichoderma harzianum. J. Mol Biol.* **194,** 755-756.

- Teplitsky, A., Shulami, S., Moryles, S., Shoham, Y. and Shoham, G. (2000) Crystallization and preliminary X-ray analysis of an intracellular xylanase from *Bacillus stearothermophilus* T-6. *Acta Crystallogr. D.* **56,**181–184.

- Thibault, F. (1992) Optimizing protein crystallization by aggregate size distribution analysis using dynamic light scattering. *J. Cryst. Growth.* **122**, 50.

- Thomson, J.A. (1993) Molecular biology of xylan degradation. *FEMS Microbiol. Rev.,* **104,** 65-82.

- Tollin, P. and Rossmann, M.G. (1966). A description of various rotation function programs. *Acta Crystallogr*. **21**, 872-876.

- Tomitani, A., Okada, K., Miyashita, H., Matthijs, H.C., ohno, T., and Tanaka, A. (1999) Chlorophyll b and phycobilins in the common ancestor of cyanobacteria and chloroplasts. *Nature.* **400**, 159-162.

- Tong, L. and Rossmann, M.G. (1990) The locked rotation function. *Acta cryst. A.* **46**, 783-792.

- Torronen A, and Rovinen J. (1997) Structural and functional role of low molecular weight *endo*-1,4-β xylanases. *J. Biotechnol*, **57**,137–149.

- Torronen A. and Rovinen J. (1995) Structural comparison of two major *endo*-1,4-β-xylanases from *Trichoderma reeseei*. *Biochemistry*, **34**,847–856.

- Torronen, A., Harkki, A. and Rouvinen, J. (1994) Threedimensional structure of endo-1,4-beta-xylanase II from *Trichoderma reesei*: two conformational states in the active site. *EMBO J.* **13,** 2493–2501.

- Torronen, A., Mach, R.L., Messner, R., Gonzalez, R., Kalkkinen, N., Harkki, A. and Kubicek C.P. (1992) The two major xylanases from *Trichoderma reesei*: characterization of both enzymes and genes. *Biotechnology* (NY). **10,** 1461-1465.

- Turkenberg, J.P. and Dodson, E.J. (1996) Modern developments in molecular replacement. *Curr. Opin. Struct. Biol.,* **6,** 604-610.

- Turunen, O., Etuaho, K., Fenel, F., Vehmaanpera, J., Wu, X., Rouvinen, J. and Leisola, M. (2001) Combination of weakly stabilizing mutations with a disulfide-bridge in the α-helix region of *Trichoderma reesei* endo-1,4-b-xylanase II increases the thermal stability through synergism. *J. Biotechnol.,* **88**, 37–46.

- Turunen, O., Vuorio, M., Fenel, F. and Leisola, M. (2002) Engineering of multiple arginines into the Ser/Thr surface of *Trichoderma reesei* endo-1,4-beta-xylanase II increases the thermotolerance and shifts the pH optimum towards alkaline pH. *Protein Eng.*, **15,** 141–145.

- Uber, D.C, Jaklevic, J.M, Theil, E.H., Lishanskaya, A. and McNeely, M.R. (1991) Application of robotics and image processing to automated colony picking and arraying. *Biotechniques.* 11, 642-647.

- Van Grondelle, R. (1985) Excitation energy transfer, trapping and annihilation in photosynthetic systems. *Biochem. Biophys. Acta.* **811,** 147-195.

- Van Grondelle, R., Dekker, J.P., Gillbro, T., and Sundstrom, V. (1994) Energy transfer and trapping in Photosynthesis. *Biochem. Biophys. Acta*. **1187**, 1-65.

- Vellieux, F.M. (1998) A comparison of two algorithms for electron-density map improvement by introduction of atomicity: skeletonization, and map sorting followed by refinement. *Acta Crystallogr. D.* **54,** 81-85.

- Viikari, L., Linko, M., and Suihko, M.L. (1984) Hydrolysis of xylan and fermentation of xylose to ethanol. *Biotechnol Adv.* **2,** 233-252.

- Vogt, G., Woell, S. and Argos, P. (1997) Protein thermal stability, hydrogen bonds, and ion pairs. *J. Mol. Biol*., **269,** 631–643.

- Vonshak, V. (1997) *Spirulina platensis* (Arthrospira) Physiology, cell-biology and biotechnology, Taylor and Francis Ltd.

- Wakarchuk, W.W, Campbell, R.L, Sung, W.L, Davoodi, J and Yaguchi, M. (1994) Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci*. **3**, 467–475.

- Wakarchuk, W.W., Sung, W.L., Campbell, R.L., Cunningham, A., Watson, D.C. and Yaguchi, M.(1994) Thermostabilization of the *Bacillus circulans* xylanase by the introduction of disulfide bonds. *Protein Eng.* **7,** 1379–1386.

- Wang, X.Q., Li, L.N., Chang, W.R., Zhang, J.P., Gui, L.L., Guo, B.J., and Liang, D.C. (2001) Structure of C-phycocyanin from *Spirulina platensis* at 2.2 Å resolutions: a novel monoclinic crystal form for phycobiliproteins in phycobilisomes. *Acta Crystallogr. D.* **57**, 784-792.

- Warren, R.A.J.(1996) Microbial hydrolysis of polysaccharides. *Annu. Rev. Microbiol*. **50,** 183-212.

- Wassenberg, D., Schurig, H., Liebl, W. and Jaenicke, R. (1997) Xylanase XynA from the hyperthermophilic bacterium *Thermotoga maritima*: Structure and stability of the recombinant enzyme and its isolated cellulose-binding domain. *Protein Sci*. **6**, 1718-1726.

- Weeks, C.M. and Miller, R. (1999) Optimizing Shake-and-Bake for proteins. *Acta Crystallogr. D.* **55**, 492-500.

- White, A., Withers, S.G., Gilkes, N.R. and Rose, D.R. (1994) Crystal structure of the catalytic domain of the beta-1,4- glycanase cex from *Cellulomonas fimi. Biochemistry* **33,** 12546–12552.

- Wilk, K.E., Hanop, S.J., Jankova, L., Edrer, D., Keenan, G., Sharples, F., Hiller, R.G. and Curmi, P.M. (1999) Evolution of a light-harvesting protein by addition of new subunits and rearrangement of conserved elements: crystal structure of a cryptophyte phycoerythrin at 1.63 Å resolution. *Proc. Natl. Acad. Sci. (U.S.A).* **96**, 890l-8906.

- William, G.T., Mark, WM., John, P.M. and Allnutt. F.C.T. (2001). Cryptomonad algal phycobiliproteins as fluorochromes for extracellular and intracellular antigen detection by flow cytometry. *Cytometry.* **44,** 16-23.

- Winterhalter, C. and Liebl, W. (1995) Two extremely thermostable xylanases of the hyperthermophilic bacterium *Thermotoga maritima* MSB8. *Appl. Environ. Microbiol.*, **61**, 1810-1815.

- Withers, S.G and Aebersold RR.(1995) Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci*, 4, 361–372.

- Wong, K KY. and Saddler, JN. (1992) In: *Xylan and Xylanases* Prog. Biotechnol. vol 7 (Visser, J., Beldman, G., Kusters-van Someren, M .A., and Voragen, A. G.J., eds) pp 171-186. Elsevier, Amsterdam.

- Wong, K.K.Y., Tan, L.U.L. and Saddler, J.N.(1988) Multiplicity of β-1,4-xylanase in microorganism: Functions and applications. *Microbiol. Rev.* **52**, 305-317.

- Wouters, J., Georis, J., Engher, D., Vandenhaute, J., Dusart, J., Frere, J.M., Depiereux, E. and Charlier, P. (2001) Crystallographic analysis of family 11 endo-beta-1,4-xylanase Xyl1 from *Streptomyces sp*. S38. *Acta Crystallogr. D.* **57,** 1813–1819.

- Yu, M.H. and Glazer, A.N. (1982) Cyanobacterial phycobilisomes. Role of the linker polypeptides in the assembly of phycocyanin. *J. Biol. Chem.* **257,** 3429-3433.

- Zou, J.Y. and Mowbray. SL. (1994) An eavaluation of the use of databases in protein strucrure refinement. *Acta Crystallogr. D.* **50**, 237-249.

Publications

L. Satyanarayana, C.G.Suresh, Anamika Patel, Sandhya Mishra, and Pushpito Kumar Ghoshb (2005).
X-ray crystallographic studies on C-phycocyanins from cyanobacteria from different habitats: marine and freshwater. *Acta Crystallogr*. (2005). F61, 844–847.

H. Balakrishnan [1], **L. Satyanarayana** [1], S.M. Gaikwad, C.G. Suresh (2006).
Structural and active site modification studies implicate Glu, Trp and Arg in the activity of xylanase from alkalophilic Bacillus sp. (NCL 87-6-10). Enzyme and Microbial Tecnology. (Vol-39, Issue 1.June 2006, Pages 67-73). **(PDB ID: *2F6B*)**
([1] both authors contributed equally)

L. Satyanarayana, S.M. Gaikwad, H. Balakrishnan, C.G. Suresh (2006)
Fluorescence and structural studies indicates presence of two tryptophan populations (hydrophobically buried and partially exposed to solvent) and an electropostive active site in xylanase from *Bacillus sp*. (Manuscript under preparation)

L. Satyanarayana S.M. Gaikwad, H. Balakrishnan, C.G. Suresh (2006)
Crystal structure of alkaline thermophlic xylanase from Bacillus sp. (NCL 87-6-10) with complex xylotriose: Xylotriose cleaved to xylobiose and xylose. **(PDB ID: *2NQY*)**
(Manuscript under preparation)

L. Satyanarayana, Anamika Patel, Sandhya Mishra, P.K Ghosh and C.G.Suresh, (2006).
Crystal structure of C-phycocyanin from *Phormidium, Lyngbya spp*. (Marine) and *Spirulina sp.* (Fresh water) shows two different ways of energy transfer between two hexamers. (Manuscript under preparation).

Poster and Oral Presentations

Purification and crystallization of C-phycocyanin from *Phormidium, Lyngbya spp.* (Marine) and *Spirulina sp.* (fresh water)**.**
L.Satyanarayana, A.Patel , S. Mishra, P.K. Ghosh, C.G.Suresh.
National Seminar on Crystallography (NSC-33). Pune. India Jan 7-10 2004

Structural and active site modification studies implicate Glu, Trp and Arg in the activity of xylanase from alkalophilic *Bacillus sp*. (NCL 87-6-10).
L. Satyanarayana S.M. Gaikwad, H. Balakrishnan, C.G. Suresh.
National Seminar on Crystallography (NSC-34) Guwahati, India. Jan 10-12, 2005

Structural studies of C-phycocyanin from Marine and Fresh water forms of Cyanobacteria: Energy transfer is more efficient in fresh water than marine. **L.Satyanarayana**, A. Patel, S. Mishra, P.K. Ghosh, C.G.Suresh. Oral presentation at Raman Memorial Conference (RMC) at Physics Dept. Pune University, Pune-411 007, India. 25 Feb 2006.