Design and development of protein crystallography beamline at Indus-2 synchrotron and structural studies of a bacterial prolidase and human PSP94

By

Ashwani Kumar

(PHYS01201004008)

High Pressure & Synchrotron Radiation Physics Division

Bhabha Atomic Research Centre,

Mumbai, India 400085

A thesis submitted to the Board of Studies in Physical Sciences

In partial fulfillment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE

July, 2016

Homi Bhabha National Institute

Recommendations of the Viva Voce Board

As members of the Viva Voce Board, we certify that we have read the dissertation prepared by Ashwani Kumar entitled "Design and development of protein crystallography beamline at Indus-2 synchrotron and structural studies of a bacterial prolidase and human PSP94" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman: Prof. Surnider M. Sharma Date:

Convener: Dr. Mukesh Kumar Date:

Member: Dr (Smt.) R. Mukhopadhyay Date:

Member: Dr. A. K. Varma Date:

Member: Dr. Ravindra D. Makde Date:

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.
I hereby certify that I have read this dissertation prepared under my direction and recommend that it may be accepted as fulfilling the dissertation requirement.

Dr. Mukesh Kumar Date: Place:
(Guide)

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

(Signature)

Ashwani Kumar

(Name of the student)

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

(Signature)

Ashwani Kumar

(Name of the student)

PHYS01201004008

(Registration Number)

July 20, 2016

List of publications

- **Papers published in refereed journals:**

1. "Protein crystallography beamline (PX-BL21) at Indus-2 synchrotron", **A. Kumar**, B. Ghosh, H. K. Poswal, K. K. Pandey, Jagannath, M. V. Hosur, A. Dwivedi, R. D. Makde and S. M. Sharma, Journal of Synchrotron Radiation. 2015 Oct 08, Journal of Synchrotron Radiation, Volume 23, Part 2, March 2016, 629–34
2. "Crystallization and preliminary X-ray diffraction analysis of human seminal plasma protein PSP94", M. Kumar, D. D. Jagtap, S. D. Mahale, V. Prashar, **A. Kumar**, A. Das, S. C. Bihani, J. L. Ferrer, M. V. Hosur and M. Ramanadhama, Acta Crystallographica Section F: Structural Biology and Crystallization Communications, Apr 1, 2009; 65(Pt 4): 389–391
3. "Crystal Structure of Prostate Secretory Protein PSP94 Shows an Edge-to-Edge Association of two Monomers to Form a Homodimer", **A. Kumar**, D. D. Jagtap, S. D. Mahale, M. Kumar, Journal of Molecular Biology, Volume 397, Issue 4, 9 April 2010, Pages 947–956
4. "Crystallization and preliminary X-ray diffraction analysis of Xaa-Pro dipeptidase from *Xanthomonas campestris*", **A. Kumar**, V. N. Are, B. Ghosh, U. Agrawal, S. N. Jamdar, R. D. Makde and S. M. Sharma, Acta Crystallographica Section F: Structural Biology and Crystallization Communications, 2014 Sep 1;70 (Pt 9): 1268–71

- **Technical Report:**

1. "Conceptual Design Report for the High-Throughput Macromolecular Crystallography Beamline at the Indus-2", By **A. kumar** and Jagannath, BARC (External) Report 2007

(Signature)
Ashwani Kumar

To My Family
whose dedication, care and blessings I cherish...

Acknowledgements

This thesis is culmination of years of work the with help and guidance from several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study. It is a pleasure to convey my gratitude to them all in my humble acknowledgment.

My deepest gratitude is to my advisor, Dr. Mukesh Kumar for his supervision, advice, and guidance from the very early stage of this research as well as giving me extraordinary experiences throughout the work. Above all and the most needed, he provided me unflinching encouragement and support in various ways.

I would like to thank Dr. S.M. Sharma for being generous on me during this work.

I would like to express my special gratitude to Dr. Ravindra D. Makde and Dr. Biplab Ghosh, for their extraordinary support throughout the beamline installation, commissioning and operation. I would also like to thank them for guidance and support in carrying out research on prolidases.

My sincere thanks to my examiners, Dr. A. K. Varma and Dr (Smt.) R. Mukhopadhyay for their useful suggestions after each progress seminar.

I am extraordinarily fortunate in having some colleagues whom I will never forget. They were helpful in a variety of ways and I thank each of them. Special thanks to Dr. H. K. Poswal, Dr. K. K. Pandey, Dr. Jagannath, Dr. Ajay Mishra, Shri A. Dewedi, Shri Rahul Singh, Dr. Sahayog N. Jamdar and Shri. Venkata Narayana Are for encouraging me all the time through stimulating discussions and criticisms at various occasions. My sincere thanks for all the staff members of the HP&SRPD office (BARC). I would also like to thank Dr. Vishal Prashar, Dr. Amit Das, and Dr. S. C. Bihani for many useful discussions; and Shri. S. R. Jadhav for the technical help.

The front-end of PX beam-line was installed by Indus Synchrotron Radiation Utilization Division (ISUD) of RRCAT, Indore. Some of the beam-line components like primary slit, water cooled compact slit, mirror chambers and Hexapods are designed and fabricated by CDM (group leader Sh. A. K Sinha). I would like to thank ISUD and CDM scientific staff for there participation in beamline development.

Where would I be without my family? My parents deserve special mention for their inseparable support and prayers. Their blessing has always been my strength. Their hard working, never say die attitude with humble nature encourages me to be a good hard

working human being.

Words fail me to express my appreciation to my wife Queeni, whose dedication, love and persistent confidence in me, has taken the load off my shoulder. Finally a big thanks for my Son Agastya for being so cute and entertaining.

(Ashwani Kumar)

Table of Contents

Synopsis

This thesis presents the design, development and commissioning of protein crystallography beamline (PX-BL21) at 1.5 T bending magnet (BM) port of Indian synchrotron (Indus-2). The thesis also includes the structural studies on a prolidase from *Xanthomonas Campestris* (Xaa-Pro di-peptidase, PepQxc) and human PSP94. The structural studies on PepQxc enzyme were carried out using recently installed PX-BL21 beamline at the Indus-2 synchrotron.

PX-BL21 at Indus-2 Synchrotron

The Indus-2 synchrotron at Raja Ramanna Center for Advanced Technology (RRCAT), Indore, India, is an electron storage ring [1] designed for $2 - 2.5\ GeV$ beam energy and with stored beam current upto 200 mA. The synchrotron has been on regular operation at 2.5 GeV beam energy with $\sim 150\ mA$ stored beam current for the last couple of years. The protein crystallography beamline (PX-BL21) on the Indus-2 is a newly constructed bending magnet beamline with working energy range of $5 - 20\ keV$. The main beamline components are a collimating mirror (CM), Double crystal monochromator (DCM) and a Toroidal mirror (TM). The end-station of the beamline is equipped with a single axis goniometer, Rayonix MX225 CCD detector, fluorescence detector, cryogenic sample cooler and automated sample changer. The beamline has capability to perform a single crystal monochromatic diffraction experiments, anomalous diffraction experiments (SAD, MIR, MAD), Laue diffraction and temperature dependent diffraction (100 K-325 K) studies. The Chapter 1 of the thesis gives a brief background of the work that is being reported in the thesis. It gives an overview of synchrotrons and different PX beamlines around the world in context of the development work that has been envisaged for the PX beamline at Indus-2. It also reviews the current research on bacterial prolidases and human PSP94 with a particular focus on the significance of their structural work that is being reported in this thesis.

The design and development of PX-BL21 was carried out taking into consideration the user requirements and synchrotron capabilities. Also, as it is being planned to shift the beamline to a Wiggler source in the future, the mandate was to design the beamline components suitable for future upgradation to Wiggler source from the current 1.5T BM

source. Being the first protein crystallography beamline developed here in India, the options have been provided to users to carry out a broad range of protein crystallography experiments. For this, various beamline configurations were explored and CM-DCM-TM combination with TM in 2 : 1 configuration has been chosen for the PX-BL21. The availabile beamline length is 36 m from the tangent point of BM source and for 2 : 1 configuration, a toroidal mirror is placed at a distance of 24 m from the source that will focus the beam at sample placed at 12 m distance from the toroidal mirror. The schematic layout of the protein crystallography beamline is shown in Fig. 1.

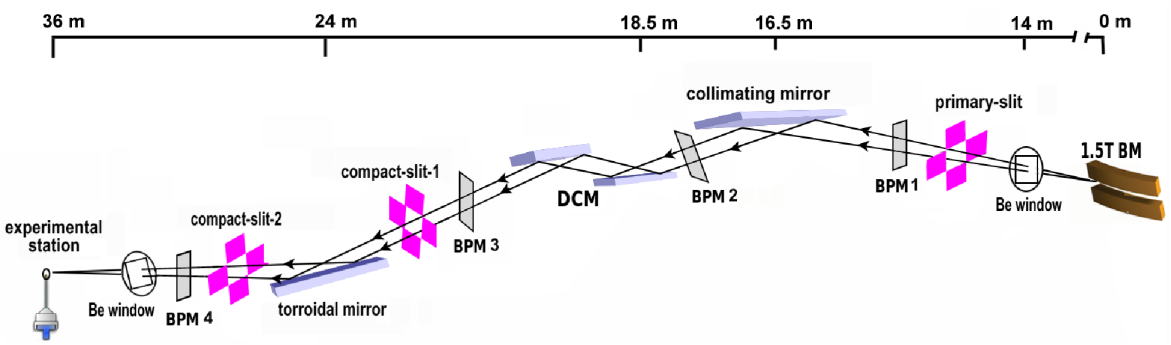

Figure 1: Schematic of the optical layout of the beamline. The collimating mirror, at a distance of 16.5 m from the tangent point, is vertically collimating the beam. The double crystal monochromator (DCM) is in non dispersive mode and the toroidal mirror at 24.0 m from the tangent point is focusing horizontally as well as vertically the X-ray beam.

Chapter 2 of the thesis describes the ray tracing calculations that have been used to finalize positions and optical specifications of different optical components of the PX-BL21. The calculated spot size at sample for the BM source is $\sim 0.5 \times 0.4\ mm^2$, while it is $\sim 1.2 \times 0.5\ mm^2$ for the Wiggler source. The required tangential radius of curvature for the CM mirror is 10.5 km, and for toroidal mirror the major and minor radii of curvature are 8.5 km and 48.5 mm, respectively, to achieve the calculated spot size at the sample position.

Chapter 2 also describes the heat load calculations for the CM and the DCM, using ANSYS code for both the BM and the Wiggler source. Most of the heat load from the synchrotron radiation will be shared by the first and the second optical components i.e., the CM and the first crystal of DCM of the PX-BL21. The calculated heat load for 2 $mrad \times 0.15\ mrad$ beam from central part of emission cone from the Wiggler source is about 1600 W. Approximately 200 W power will be absorbed by the CM and 1 kW by the first crystal of the DCM. Therefore, it requires careful consideration of cooling

schemes and design in order to run the beamline smoothly and to avoid damage of the downstream optical components due to high heat load. Various cooling schemes have been considered for the CM, and the side water cooling with optimal contact-cooling (OCC) scheme was found suitable and is discussed in Chapter 2.

The heat load on first crystal of DCM will be significant, therefore multiple cooling scheme have been discussed in the chapter. It was found that water cooling will be insufficient to take care of the rate of heat dissipation (1 kW) at first crystal of DCM for a Wiggler source. Therefore, the thermal-structural calculations have been performed using liquid nitrogen cooling for the first crystal of DCM. For the bending magnet source water could be used as a coolant instead of liquid nitrogen for the same design as the calculations show acceptable temperature rise with slope error of less than 6 μrad for a beam energy of 5 keV. Therefore, the design of the cooling system for the DCM compatible with both the coolants, viz., liquid nitrogen as well as water, have been proposed in chapter, where a water cooling has been provided for the BM source with the compatibility of LN2 cooling for a Wiggler source.

Chapter 3 of thesis describes the design and fabrication of beamline components. Many beamline components were designed and fabricated in-house. These beamline components are Beam position monitors (4 No.s), one primary X-ray slit, two no.s of water cooled compact slits, X-ray pipes, Mirror chambers (2 no.s) and Hexapods (2 no.s). The beam position monitors (BPMs) have a phosphor coated water cooled copper (Cu) plate for X-ray beam viewing, and an electrically isolated, 500 μm tungsten wire placed at bottom of Cu plate. As discussed in the chapter, these BPMs have six-way cross as shown in Fig. 2a. The X-ray beam viewed from the BPM camera is shown in Fig. 2b, while the beam profile measured from the tungsten wire is shown in Fig. 2c.

Chapter 3 of the thesis also presents the installation and commissioning of beamline. All beamline components were installed at optical line of the BL-21 port at Indus-2. Installation was followed by leak testing of individual beamline components to check the ultra-high vacuum compatibility. Subsequently, shielding hutch was installed to carry out hot beamline alignment. Hot alignment of beamline components in X-rays was performed using BPMs. Both mirrors were aligned with an accuracy of few microns at 3 $mrad$ glancing angle in X-rays. Software for calibration of DCM energy and absorption edge scanning using fluorescence detector has been developed and deployed. The measured spot size at sample position is 0.5 $mm(V) \times$ 0.5 $mm(H)$, which is close to the calculated spot size of 0.4 $mm(V) \times$ 0.4 $mm(H)$. The flux, measured by a calibrated photo-diode, at sample position with 0.4 $mm \times$ 0.4 mm slit was 6.4 $\times$ 10^9 ph/sec for 80 mA storage

ring current at 2.5 GeV. Initial data were collected on lysozyme crystals with good statistics, suggesting the viability of beamline. Anomalous diffraction data from other protein crystals were collected later and analyzed satisfactorily. There is also an ancillary biochemistry laboratory associated with this beamline to provide necessary infrastructure to the beamline users for crystal preparation, manipulation, soaking and optimization, if required.

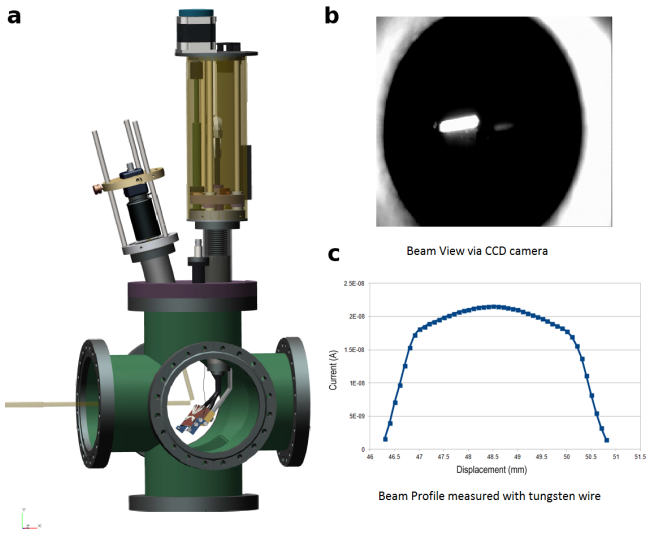

Figure 2: (a) Beam Position monitor with six way cross, (b) X-ray beam viewed from BPM's camera, (c) X-ray beam profile measured from the tungsten wire

The last two chapters (Chapter 4 and 5) ot the thesis describe the determination of crystal structures of two proteins, namely Xaa-Pro dipeptidase from *Xanthomonas campestris* (PepQxc) and human PSP94, using synchrotron beamline. The structural studies on PepQxc were carried out using the PX-BL21 beamline and its associated biochemical facility at Indore, whereas the diffraction data from human PSP94 were collected at PXIII beamline of Swiss Light Source. Details of these works are summarized in the following sections.

Structural studies of PepQxc

The prolidase (EC 3.4.13.9) or Xaa-Pro dipeptidase or PepQ cleaves the trans Xaa-Pro peptide bond in a dipeptide with a prolyl residue at the carboxy terminus. The prolidase is found in mammals, bacteria and archaea [2]. Functionally, prolidase is well characterized in human and less is known about the prolidase in archaea and bacteria. The prolidase

has been shown to be involved in degradation of collagen [3] and its deficiency in human causes a recessive disorder, which is characterized by skin ulceration, mental retardation and recurrent infections [4]. The recombinant prolidase is used commercially in the food and dairy industries for improving flavor and texture of the food [2]. The prolidase and other peptidases of M24B family has also been shown to degrade toxic organophosphorus (OP) compounds, which include pesticides and nerve agents, by cleaving the P-F and P-O bonds [5]. Thus, the prolidase has the huge potential applications in detoxifying OP nerve agents, and as a biosensors for the OP compound detection.

The prolidase in present study is a 399 residues long protein from a *Xanthomonas campestris* bacterium (protein annotated as XPD43 or PepQxc; GenBank accession, NP-637763). The PepQxc orthologs are quite unique among M24B family of peptidases, where the strictly conserved glycine and tyrosine residues (*Gly385* and *Tyr387*; numbering as per *E. coli* Aminopeptidase-P) are changed to methionine and valine, respectively. The *Tyr387* of *E. coli* Aminopeptidase-P (APPro) form the conserved hydrogen bond network in *Asp260-Arg404-Tyr387* motif, which has been proposed to shuttle a proton from bulk solvent to the leaving peptide [6].

Chapter 4 describes the cloning, expression, purification, crystallization and structure determination of the native and R372A mutant of PepQxc protein. A total of six (four native and two with R372A mutation), high resolution PepQxc crystal structures were determined. Also, detailed enzyme assays were performed to study the substrate specificity of PepQxc in presence of various Xaa-Pro substrate and metal ions. The present study gives insight into the role of conserved residues in Xaa-Pro dipeptidase class of proteins. The structural analysis of PepQxc shows that the protein is dimeric in nature. The monomer of PepQxc consist of two domains, as shown in Fig. 3. The active site is characterized by di-nuclear metal center containing two manganese ions.

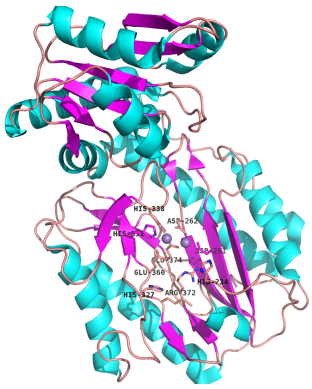

Figure 3: Cartoon representation of the PepQxc monomer with helixes in cyan, β-sheets in magenta and two manganese ions in light brown color.

The crystal structure of native and R372A mutated PepQxc protein showed that the

conserved *Tyr* residue, which is a part of hydrogen bond networking motif in *E. coli* APPro is replaced by *Met* residue in PepQxc. Interestingly, *Met* being a non-polar hydrophobic residue cannot participate in proton shuttling with the *Arg372* and the *Asp251* in the PepQxc enzyme. Also, the R372A mutated structure of the PepQxc showed that the overall structure of active site is intact and the *Arg372* residue is not critically required for metal center stability.

The enzymatic activity assay of PepQxc shows a concentration dependent inhibitory effects by some dipeptide substrates and phosphate molecules. The loss in enzymatic activity in presence of phosphate ions has been explained in the chapter on the basis of structure of the PepQxc1-Mn and PepQxc1-Zn complexes.

Structural studies of human PSP94

The prostate-secretory protein of 94 residues (PSP94), also known as β-microseminoprotein (β-MSP), is one of the most abundant proteins in human semen [7]. It is a small non-glycosylated protein that is rich in cysteines. It is secreted at a high level in the seminal fluid by the epithelial cells of the prostate gland; however, the protein is not prostate-specific as its presence has been reported in several other tissues [8]. Although the exact biological function of PSP94 remains elusive, various systemic and confined roles for this protein have been reported, including as a binder of immunoglobulins [9], as a motility inhibitor of sperm [10] and as a regulator of calcium levels during hypercalcaemia of malignancy [11]. Two proteins that bind to PSP94 in seminal fluid [12] and in blood (PSP94-binding protein; [13]) have also been identified, which is likely to drive further investigations in the search for more biological functions of this protein. PSP94 serum measurements have potential clinical utility in prostate-cancer management [14].

Although no crystal structures of any members of this family of proteins have been reported, two laboratories ([15] and [16]) earlier reported the NMR structure of PSP94. The NMR structure shows that the protein contains two domains bridged by a disulfide linkage. Interestingly, the secondary structures of the individual domains reported by the two laboratories were similar but the orientation of the two domains was very different, giving rise to two distinct overall shapes for this small protein.

Chapter 5 describes the crystal structure of PSP94, which was determined by single isomorphous replacement with anomalous scattering (SIRAS) method using crystals soaked in uranyl nitrate. The overall structure of PSP94 is similar to the NMR structure reported by *Ghasriani et al.*, except the difference in relative orientation of the two do-

mains by $\sim 20^{\circ}$ [16]. The crystal structure of PSP94 shows that the edges from two PSP94 monomers associate to form a dimer. Further, the present results showed for the first time a pH-induced transition of PSP94 from dimeric to monomeric form. This assumes biological significance as PSP94 would function as a dimer in the human semen (pH $7.5 - 8.0$). The structure shows that the amino and the carboxyl ends of the polypeptide chain are held in close proximity facing each other. A strong hydrogen bond between these ends, which are located respectively on the first and the last -strands, leads to formation of an almost straight edge in PSP94 structure as shown in Fig. 4a.

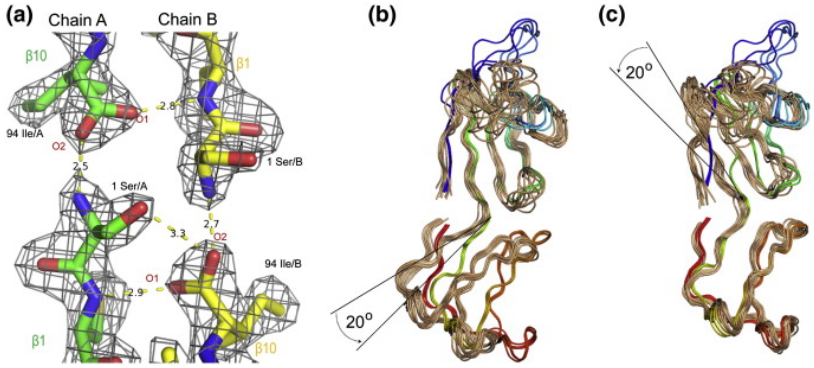

Figure 4: PSP94-dimer to monomer. (a) The stick representation of the terminal residues from two different polypeptide chains (with green and yellow carbons) at the dimeric interface is shown along with their electron density maps ($2F_O - F_C$) contoured at 1.5σ. Oxygen and nitrogen atoms are colored red and blue, respectively. The hydrogen-bond distances are marked in angstroms. Superposition of either (b) the C-terminal domain or (c) the N-terminal domain of the NMR models (from PDB file 2iz3 and shown in brown) onto the corresponding domain of the crystal structure (four polypeptide chains in rainbow color) showing a shift in the relative orientation of two domains.

Further analysis shows that the relative orientation of the two domains of PSP94 in the present dimeric structure is slightly different from that in the NMR structure (PDB ID 2IZ3). When all the C_α atoms of one of the domains in the NMR structure are superposed on the corresponding atoms in the present structure, the other domain in the NMR structure appears shifted as compared to the present structure (Fig. 4b and 4c) and the shift is of a pure rotation of $\sim 20^{\circ}$ around an axis located between the two domains.

The interaction of PSP94 with several CRISPs and IgG molecules has also been proposed in the chapter. The present work thus provides valuable insight into the structure of PSP94 as well as its dimeric form and its possible mode of interaction with different binding proteins, which, in turn, may provide useful clues to the biological functions of this important family of proteins.

In conclusion, this thesis presents the design, development and commissioning of a protein crystallography beamline (PX-BL21) at RRCAT Indore. The beamline is primarily aimed to cater the experimental needs of about 100 independent research groups working in the area of structural biology in India. This beamline is a national facility and the construction, operation and the maintenance of the facility is funded by the Department of Atomic Energy (DAE), Government of India. The last two chapters of the thesis describe the crystal structure of two proteins and discuss their biological significance.

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview of Synchrotron and Beamlines

Synchrotron radiation sources started as a parasitic sources to accelerators used for high energy and nuclear physics studies [17], during 60s and 70s. But, the researchers quickly realized the importance of synchrotron source due to a number of characteristic features of synchrotron radiations, such as;

- The availability of wide spectrum of photon energy from the synchrotron source.
- Highly collimated beam in vertical plane with flux order of magnitude greater than the conventional X-ray source.
- Linearly and elliptically polarized light from the source.
- Pulsed nature of synchrotron radiation, which can be useful for time resolved studies

During 80s, the second generation synchrotron radiation sources became available. A few examples are, 2 GeV SRS at Daresbury, 2.5 GeV Photon factory in Japan, 2.5 GeV at NSLS and SPEAR2 at SSRL. These dedicated facilities have bending magnet (BM) as a source for the synchrotron radiation. Soon, it was found that further enhancement in X-ray beams can be achieved by inserting magnetic array devices into the straight sections between the bending magnets (BMs) [18]. These magnetic array devices are Wigglers and undulators that generate the alternating (Sinusoidal type) magnetic field along the path length of relativistically moving bunch of electrons in the storage ring. The alternating magnetic field is produced by placing small high power magnets or electromagnets of

alternating polarity in an array. In last two decades, new third generation synchrotron radiation sources increasingly become available to users, notables are 6 GeV ESRF in France, 7 GeV APS at Argonne Photon Source in US and 8 GeV SPring8 in Japan. The third generation sources have Wigglers and undulators as X-ray sources. Recently some new synchrotron source, mostly in 2.5–3 GeV energy range has become operational like, 2.4 GeV SLS (Switzerland), 2.8 GeV Soleil (France), 2.9-GeV CLS (Canada), 3 GeV DIAMOND (UK), 3 GeV Australian Light Source, 6 GeV PETRA-III (Germany) and 2.5 GeV Indus-2 (India).

An accelerating charge particle emits the radiation [19]. For energy > 1 GeV, the relativistic electron ($\beta = v/c \approx 1$) emits the radiation in a narrow almost pencil shape cone of an angle $\approx \gamma^{-1}$, in the direction of motion [20]. Where, γ is lorentz factor, defined as $\gamma = 1/\sqrt{1-\beta}$, and v is velocity of electron, and c is velocity of the light. Therefore, higher the electron energy in storage ring, higher will be the collimation of the emitted radiation beam. In BM, the electron moves in a circular orbit under the influence of the magnetic field of the BM, and the radiation is emitted tangentially to the electron orbit. The radiation from BM source has wide energy spectrum and high angular collimation, the former is due to the short pulse duration of the radiation and later is because of the angle dependent relativistic Doppler shift [21].

Spectral brightness or brilliance term is often used to characterize the synchrotron beam. Brilliance is defined as number photons per second per mm^2 per $mrad^2$ in 0.1% bandwidth [22]. For a 2.5 GeV Indus-2 synchrotron at 100 mA with 1.5 T BM source, the brilliance is of the order of 10^{14} $ph/sec/mm^2/mrad^2/0.1\%bandwidth$. The source brilliance can be further improved by inclusion of the insertion devices (Wiggler and undulator) in the storage ring. Wigglers and undulators have periodic magnetic structure, these devices are installed at the straight section of a storage ring. A Wiggler simply can be considered as mulitiple of the BM source that can be seen from the spectral characteristics of Wiggler radiation, which is similar to BM radiation with high flux. The brilliance of a 1.5 T Wiggler with Indus-2 parameters will be order of 10^{15} $ph/sec/mm^2/mrad^2/0.1\%bandwidth$. On the other hand, the undulator has the coherent constructive superposition of emitted radiation from successive poles that results into a quasi-monochromatic and highly collimated beam of the high brilliance, for example the undulator with Indus-2 parameters will have the brilliance of order 10^{17} $ph/sec/mm^2/mrad^2/0.1\%bandwidth$.

The synchrotron radiation emitted from these sources (BM, Wiggler and undulator) are utilized by the user for their experiments using a beamline. The beamline efficiently

transports the X-rays from the synchrotron ring and focus it on the sample position through a series of optical elements, designed and installed at calculated positions, for providing the best possible match to the sample requirements. A typical synchrotron beamline is aimed to exploit two main features of synchrotron source: high brilliance of beam that is good for weakly diffracting crystals and the energy tunability that is essential for determining phases. First protein crystallography beamline became operational in May 1974, at SSRL [23] that used to be a first generation synchrotron source. At present, almost all operational synchrotrons have dedicated protein crystallography beamlines. Some of these operational beamlines are optimized for particular type of experiments in protein crystallography for example, micro-focussed beamlines have focussed X-ray to sub-micron ($< 10\mu m$) spot size to diffract small size crystals. The 24ID-E beamline at APS, I24 at DIAMOND, ID13 at ESRF and BL32XU at SPring8 are a few examples of micro-focussed beamlines. Then there are fixed energy beamlines to collect the diffraction data at particular (usually between $1 - 0.8$ $\mathring{A}$) wavelength, examples are; ID23-2 at ESRF is fixed energy (14.20 keV, 0.873 $\mathring{A}$) beamline, beamline 14.3 at Helmholtz-Zentrum Berlin is a fixed energy beamline operating at 13.87 keV (0.89 $\mathring{A}$). But, most protein crystallography beamlines have the double crystal monochromator and bendable X-ray mirrors, to tune the X-ray energy and change the shape of the spot at sample position. This type of beamline can have focussed as well as parallel (low divergence) X-ray beam at sample position. Some of the examples are; XRD2 beamline at Elettra, BM-14 at ESRF, 08B1-1 at CLSI and BL24XU at SPring8 etc. A detail information on protein crystallography beamlines can be accessed online from http://biosync.sbkb.org/.

1.1.1 Protein Crystallography using Synchrotron Radiation

A living organism has thousands of different proteins, and each one folds into a specific 3D structure to perform a specific cellular function [24]. Determining the three-dimensional (3D) structure of the proteins is crucial for understanding their associated cellular function at an atomic detail, which eventually may help us to understand the biological processes and control several diseases. There are a number of techniques, which can be used to determine the 3D-structure of proteins for example; X-ray diffraction of protein crystals, cryo-electron microscopy [25] and NMR [26]. Out of these listed techniques, the X-ray diffraction of protein in crystalline form is highly utilized [23] because of the usage of the synchrotron radiation sources. The protein crystals are weakly diffracting in nature due to presence of low Z elements like carbon, nitrogen and oxygen. Also, the protein

crystals have high content of disordered solvent, which contribute towards noise in X-ray diffraction data. In addition, micron size protein crystals with less number of repeating units (unit cell), often impaired by imperfections, leads to poor quality diffraction data. Therefore, the high brilliance X-ray source becomes the paramount requirement for the X-ray diffraction data collection on protein crystals. In a modern 3^{rd} generation synchrotron with micro-focus beamline on an undulator source, data can be collected quickly and efficiently on poorly diffracting microns size crystals with good statistics, which would be impossible with conventional lab source.

Moreover, the experimental determination of 3D structure of a single crystal of macromolecule depends on the acquisition and processing of diffraction data. Each diffraction spot captured in an image frame, has associated with it both an amplitude and a phase. However, the phases are not directly measurable in a typical diffraction experiment and therefore must be estimated by indirect means. Methods used for derivation of phases are : Molecular Replacement (MR) [27], Multiple Isomorphous Replacement (MIR) [28], Single and Multiwavelength Anomalous Diffraction (SAD and MAD), [29] and [30]. The MIR, SAD and MAD methods use anomalous scattering of heavy atoms embedded in protein crystals for the estimation of phases. Dispersive and anomalous scattering effects measured at several X-ray wavelengths can provide a direct solution to the crystallographic phase problem.

Advent of tunable and reliable brilliant synchrotron radiation sources make the anomalous scattering method more powerful as measurements can be made at a precisely chosen wavelength. Anomalous scattering was first recognized as a source of phase information by Bijvoet [31], and has been employed since the early days of macromolecular crystallography. Anomalous scattering has been used to locate positions of anomalous scatterer [32], usually a heavy atom, to supplement phase information from isomorphous replacement [33, 34] and to identify the enantiomorph of the heavy-atom partial structure in multiple isomorphous replacement (MIR) phasing [35]. Anomalous scattering at a single wavelength was the sole source of phase information in the structure determination of crambin [36] protein, an important precursor to development of MAD [29].

1.2 Protein Crystallography Beamline at Indus-2 Synchrotron

The number of protein structures deposited in Protein Data Bank (PDB) [37] has increased exponentially over the last decade. Availability of protein crystallography beamlines on synchrotron rings around the world has been responsible in large measure for this phenomenon. The high brilliance of X-ray beams delivered on the sample crystals has enabled scientists to collect accurate diffraction data from crystals too small to be useful earlier. The energy-tunability of the X-ray beams has enabled scientists to exploit the phenomenon of anomalous diffraction to speedily solve the crystallographic "phase problem". The quality of diffraction data collected on a synchrotron source is generally also so much higher than that obtainable on a home source, because of several reasons, that one prefers to use a synchrotron source for serious data collection.

The Indus-2 synchrotron is an electron storage ring [1] designed for $2-2.5\ GeV$ beam energy and stored current upto 200 mA [38] at Raja Ramanna Center for Advanced Technology (RRCAT), Indore, India. The synchrotron has been on regular operation at 2.5 GeV beam energy with $\sim 150\ mA$ stored beam current; it is operated in decay mode with typical lifetime of $\sim 22\ h$ at 100 mA stored beam current [39]. Recently, Indus-2 has been upgraded to a third generation source with the commissioning of two undulators [40]. The protein crystallography beamline (PX-BL21) is a newly constructed bending magnet beamline on Indus-2 synchrotron with working energy range of $5-20\ keV$.

The design and development of PX-Bl21 was carried out keeping the user requirements and synchrotron capabilities as our priorities. Since, it will be the first protein crystallography beamline in India, the option for broad range of experiments in protein crystallography have been provided to users. The beamline has capability to perform a single crystal monochromatic diffraction experiment, a anomalous diffraction experiment (SAD, MIR, MAD), Laue diffraction and temperature dependent diffraction (100 K-325 K) experiments. Key design elements in PX-BL21 are collimating mirror (CM) Double crystal monochromator (DCM) Toroidal mirror (TM), with TM in $2:1$ configuration, which provides the stable and aberration free X-rays at the sample. Both mirrors are bendable to optimize the X-ray beam spot at end-station. The appropriate cooling has been applied, based on the heat load calculations, to the optical elements of the beamline. The Chapter 2 of the thesis describes the design, development and commissioning of PX-BL21. Ray tracing and heat load calculations has been presented in the context of

beamline design and development. The heat load calculations using Wiggler as a source have shown that liquid nitrogen cooling will be required for the DCM. While, the water cooling will be adequate for X-ray mirrors of the beamline. Also, the Chapter 3 describes the hot alignment of beamline components in X-rays by using the beam position monitors (BPMs). Many beamline components like, BPMs (4 No.s), one primary X-ray slit, two no.s of water cooled compact slits, X-ray pipes, Mirrors chambers (2 no.s) and Hexapods (2 no.s), has been designed and fabricated indigenously. Furthermore, the beamline optimization and validation has been reported in the chapter.

In future, the PX beamline will be shifted to a 2.5 T superconducting multipole Wiggler (MPW) port and current beamline design includes the necessary requirements for the same. There is also an ancillary biochemistry laboratory associated with this beamline to facilitate the beamline users at various stages of crystal preparation and optimization. This beamline is a national facility and the Department of Atomic Energy (DAE), Government of India, has funded for the construction and operation of this beamline. The PX-BL21 is primarily aimed to cater the need of about 100 independent research groups in India working in the area of structural biology. The first diffraction data sets on single crystals of Lysozyme protein was recorded during mid-2012 and since then the beamline has been used by several users.

Beamline development is a large scale project, involving multiple discipline ranging from engineering, instrumentation to basics sciences. It involves people from diverse background like, mechanical, instrumentation, software, alignment, vacuum, health and radiation shielding, biologist and physicist etc. Therefore, this project would not have been possible without the contribution from many people from different divisions within the BARC and the RRCAT. However, I will briefly state my contribution in the beamline project. I was involved in the beamline project from the beginning. I was involved in each aspect of beamline design, development, commissioning. I have done the ray tracing and heat load calculations, prepared the specification of optical components and participated in procurement of each beamline components. I have also participated in the indigenous development of rest of the beamline components. Finally, I was involved in the installation, X-ray alignment and commissioning of the protein crystallography beamline.

Determination of the structure of two proteins, namely Xaa-Pro dipeptidase from *Xanthomonas campestris* (PepQxc) and human PSP94, using synchrotron beamline will also be presented in this thesis. The structural studies on PepQxc was carried out using the PX-BL21 and its associated biochemical facility at Indore. Whereas, the protein crystallography beamline at SLS and the biochemical facility at Bhabha Atomic Research

Centre (BARC) was used to study human PSP94. Details of these works are summarized in the following sections.

1.3 Xaa-Pro dipeptidase from *Xanthomonas campestris*

The prolidase (EC 3.4.13.9) or Xaa-Pro dipeptidase cleaves the trans Xaa-Pro peptide bond in a dipeptide with a prolyl residue at carboxy terminus. The prolidase are found in mammals, bacteria and archaea [2]. Functionally, prolidase are well characterized in human and less is known in archaea and bacteria. The prolidase are shown to involved in degradation of collagen [3] and its deficiency in human is a recessive disorder, which is characterized by skin ulceration, mental retardation and recurrent infections [4]. The recombinant prolidase has been used commercially in the food and dairy industries for improving flavor and texture of the food [2]. The prolidase and other peptidases of M24B family has also been shown to degrade toxic organophosphorus (OP) compounds that include pesticides and nerve agents, by cleaving the P-F and P-O bonds [5]. Thus, it has huge potential applications in detoxifying OP nerve agents, and as biosensors for OP compound detection.

The present study of prolidase is from the *Xanthomonas campestris* bacteria. The selected protein from *Xanthomonas campestris* is 399 (42.8 kDa; XPD43 or PepQxc; GenBank accession, NP-637763) residue long. The XPD43 (PepQxc) orthologs are quite unique among M24B family of peptidases, where the strictly conserved glycine and tyrosine (*Gly385*, *Tyr387*; numbering as per Escherichia coli Aminpeptidase-P) is changed to methionine and valine, respectively (Fig. 4.1). The *Tyr387* of APPro has been proposed to be involved in the conserved hydrogen bond network in *Asp260-Arg404-Tyr387* motif, which is able to shuttle a proton from bulk solvent to the leaving peptide [6].

To study the role of these conserved residues in prolidase family, the recombinant native and R372A mutated PepQxc protein has been cloned, purified and crystallized. A total of six (four native and two with R372A mutation), high resolution PepQxc crystal structures were determined. Also, the detailed enzymatic assays were performed to study the substrate specificity of PepQxc towards the Xaa-pro substrates in presence of different metal ions. The present study will give some insight into role of these conserved residues in Xaa-Pro dipeptidase class of proteins. Also, the enzyme assay for PepQxc shows the concentration dependent inhibitory effects by some the dipeptides and phosphate molecules. The cause of inhibitory effect on PepQxc activity by phosphate and some dipeptide, will be discussed based on the crystal structure of the enzyme.

1.4 Human PSP94

The prostate-secretory protein of 94 residues (PSP94), also known as β-microseminoprotein (β-MSP), is one of the most abundant proteins in human semen ([7] and [41]). It is a small non-glycosylated protein that is rich in cysteines. It is secreted at a high level in the seminal fluid by the epithelial cells of the prostate gland; however, the protein is not prostate-specific as its presence has been reported in several other tissues ([42] and [8]). Although the exact biological function of PSP94 remains elusive, various systemic and confined roles for this protein have been reported, including as a modulator of follicle-stimulating hormone (FSH) level [43], as a binder of immunoglobulins [9], as a motility inhibitor of sperm [10], as a growth regulator and inducer of apoptosis in prostate cancer cells in vitro and in vivo [44] and as a regulator of calcium levels during hypercalcaemia of malignancy [11]. Two proteins that bind to PSP94 in seminal fluid [12] and in blood (PSP94-binding protein; [13]) have also been identified, which is likely to drive further investigations in the search for more biological functions of this protein. PSP94 serum measurements have potential clinical utility in prostate-cancer management [14].

In addition to humans, PSP94 homologues have been identified in many other mammals ([45], [46] and [47]), ostrich [48] and recently in the serum of a reptile, the Japanese viper [49]. Interestingly, one of the PSP94 homologues (SSP-2) from this reptile binds to the snake-venom-derived Ca^{2+}-channel blocker (triflin), perhaps imparting self-protection from its own venom to the snake [49]. There is a large degree of sequence diversity among different PSP94 homologues, but the positions of the ten cysteine residues are well conserved throughout, suggesting that the overall tertiary structures of this family of proteins should remain conserved. No crystal structures of any members of this family of proteins have been reported to date. However, two laboratories ([15] and [16]) have reported the NMR structure of PSP94, showing that this small protein contains two domains that are bridged by a disulfide linkage. Interestingly, the secondary structures of the individual domains reported by the two laboratories were similar but the orientation of the two domains was very different, giving rise to two distinct overall shapes for this small protein.

The crystal structure of PSP94 has been determined by single isomorphous replacement with anomalous scattering (SIRAS) method using crystals soaked in uranyl nitrate. The overall structure of PSP94 is similar to the NMR structure reported by *Ghasriani et al.*, except the difference in relative orientation of the two domains by $\sim 20^{\circ}$ [16]. The difference in orientation seems to have functional implications and will be discussed in the thesis. The PSP94 structure also gives some insight into the possible mode of interaction

of PSP94 with CRISP and IgG molecules.

1.5 Objective of the thesis

The objective of thesis is to the design, develop and commission the protein crystallography beamline (PX-BL21) at RRCAT Indore. Moreover, our aim is to keep the future upgradation of X-ray source a possibility, therefore, thermal design of individual optical components will be optimized for the higher heat load from the Wiggler source. The optical design of PX-BL21 will be optimized by the ray tracing and heat load calculations. Scope of present thesis will also includes the design, development and fabrication of various beamline components; Beam position monitors (4 No.s), one primary X-ray slit, two no.s of water cooled compact slits, X-ray pipes, Mirrors chambers (2 no.s) and Hexapods (2 no.s). The installation and hot alignment of beamline components in X-rays using BPMs, design and deployment of shielding hutch, development of software for calibration of DCM energy and absorption edge scanning using fluorescence detector will be the part of the thesis.

Also, the structural studies on the human PSP94 protein and the Xaa-Pro di-peptidase (PepQxc) from *Xanthomonas Campestris* using auxiliary biochemical facility of PX beamline at RRCAT Indore will be part of the thesis. The present thesis aim to study the prolidase enzyme from *Xanthomonas campestris* bacteria. In this study, the crystal structures of PepQxc1, PepQxc1-Mn, PepQxc1-Zn, PepQxc2, PepQxcR372A and Zn-PepQxcR372A have been determined using X-ray crystallography techniques. The thesis will explore the role of conserved residue in active site of PepQxc enzyme. The structural basis for concentration dependent inhibitory effects by some dipeptides and phosphate molecule will be discussed in present work. The structural studies of PSP94 will include the crystal structure solution using SIRAS method. It will be shown how the edges from two PSP94 monomers associate to form a dimer. Based on the structure, the interaction of PSP94 with several CRISPs and IgG molecules will be studied.

1.6 Outline of the thesis

This thesis includes the design, development and commissioning of a protein crystallography beamline (PX-BL21) at RRCAT Indore. Also, the thesis includes the structural studies on the human PSP94 protein and the PepQxc from *Xanthomonas Campestris*. To

present all the work in a clear and concise manner, the thesis is divided into six chapters.

- **Chapter 1** presents a brief review on synchrotron, beamlines and protein crystallography. Also, the review of the current research in PepQxc enzyme and human PSP94 with particular focus on the problems to be addressed in this thesis has been reported.
- **Chapter 2** will present the ray tracing and heat load calculations for beamline components. The conceptual design of PX-BL21 will be outlined in this chapter.
- **Chapter 3** outlines the development of beamline components. Furthermore, installation and commissioning of PX-BL21 will be presented.
- **Chapter 4** present the structural and biochemical studies on PepQxc enzyme.
- **Chapter 5** describe the structural studies on human PSP94.
- **Chapter 6** concludes the thesis with a brief summary of our contributions along with the further scope of work.

Chapter 2

Conceptual design of protein crystallography beamline

First crystal structure of protein hemoglobin [50] was solved by Max Perutz in 1950s using the X-ray crystallography techniques. At that time, solving the protein structure of even 100 atoms was a difficult task, whereas the hemoglobin contain 2291-non H atom in asymmetric unit. It required 20 years of unparalleled patience and hard work from Max Perutz to solve the hemoglobin structure. Comparing the present status to 1960 and 1970, it now requires just a few days to solve a protein structure instead of several years. Although, the whole process still follows the same stages of crystallization, data collection, phasing, model building and refinement and presentation of results, today these steps can be performed much more easily and quickly. Primary contribution to this advancement is due to enormous progress in computational capabilities and availability of synchrotron sources for the high intensity X-rays. At present, most of synchrotron beamlines utilize the radiation from the insertion devices, such as Wiggler or undulator reducing the data collection time substantially compared to that obtained with a bending magnet source.

This chapter will describe the optical design concept, the beamline layout and the expected performance of the beamline. The PX-BL21 has the collimating mirror for vertical collimation of beam, the double crystal monochromator for energy selection, and the toroidal focusing mirror for focusing beam both in vertical and horizontal directions. The End-station or experimental station of the beamline is last component, where the data collection takes place. The End-station of the beamline is equipped with a CCD detector. The CCD detectors are characterized by high spatial resolution, low intrinsic

noise and short read out time. The detector have large front windows for capturing wide angle.

Insertion devices are planned in the second phase of commissioning of Indus-2 synchrotron. Thus, the PX-BL21 has been designed in such a way that it can be shifted to the insertion device with the least number of modifications. In this chapter, the ray tracing results of PX-BL21 for the BM and the Wiggler source have been presented. Since, the Wiggler specifications are not finalized yet, the Elettra Wiggler [51] specifications has been taken as a source for the heat load calculations with the Indus-2 ring parameter.

Number of codes have been used in designing of the beamline. A brief overview of those codes are presented here. A XOP and SHADOWVUI has been used for ray tracing calculations of the bemaline optical components [52]. While for heat load calculations, the ANSYS software package from ANSYS, Inc [53] has been used. XOP (X-ray Oriented Programs) is a widget based, collection of programs that provides the common front-end interface to various useful synchrotron related codes, such as codes for calculation of synchrotron radiation source spectra (bending magnet, Wigglers and undulators), codes for calculation of characteristics of optical devices (mirror, filters, crystals, multilayers, etc.) and codes for multipurpose data visualizations and analyses [52]. The XOP includes the several extension, including SHADOWVUI for ray tracing calculations. The SHADOWVUI is widget based interface for the SHADOW ray-tracing code [52]. The interface has menu options to prepare input data, and to run the SHADOW code, subsequently, the output is displayed in plots and graphs for easy visualization.

The SHADOW is a open source, multifaceted software package that is used for modeling optical systems [54]. The program is optimized for the designing of optical systems in the synchrotron radiation beamlines. The code perform the calculations in two steps; first step is source generation and second step is tracing of the rays that constitute the source. These steps are executed by parts of code called SOURCE and TRACE along with some pre- and post-processor utilities [54]. The program SOURCE generates the source for ray tracing calculations, and the source in SHADOW is a collection of rays originating in a region of space that propagate in straight line without changes until intersected by a obstacle or optical surface. The SHADOW uses the stochastic process to generate the source i.e., the rays are all generated with equal amplitude but with a spatial frequency determined by the model describing the source. The SHADOW can handle maximum 25000 rays. Each generated ray has 18 variables like, three coordinates, three momentum etc. The program TRACE does the ray tracing of the source containing rays through the optical elements. Each optical element is defined individually and the ray tracing

proceeds in series with one element at a time used for ray tracing calculations.

The SPECTRA code [55] has been used for the calculation of synchrotron source properties, especially flux and power calculations for the BM and the Wiggler source. The SPECTRA can be used to calculate optical properties of synchrotron radiation emitted from bending magnets, Wigglers (conventional and elliptical) and undulators (conventional, helical and elliptical). The SPECTRA have graphical user interfaces (GUIs) that is used to input the electron beam and the source parameters as well as to show the calculation result graphically. The ANSYS code is used for thermal and structural analysis of X-ray mirrors and double crystal monochromator, under the high heat load from the Wiggler. The ANSYS is a finite element analysis software package which is used for analyzing various engineering applications like, structural, thermal, electromagnetic and fluid dynamics [53].

2.1 Photon Source

Indus-2 is a 2.5 GeV synchrotron radiation source with critical wavelength of about 2 $\mathring{A}$ [56]. Indus-2 lattice is designed in such a way so as to give low beam emittance and high brightness. The lattice is a double bend acromat with zero dispersion function along the long straight section. It has eight super periods, each having two dipole bending magnets, six sextupoles, four focusing and five defocussing quadrupoles as shown in Fig. 2.1.

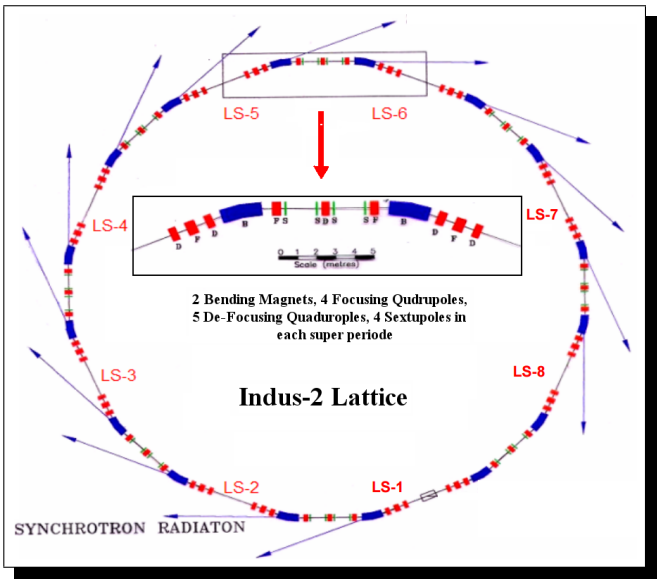

Figure 2.1: Indus-2 lattice. (inset) Magnified image of Super period showing bending magnets, quadrupoles and sextupoles

Of the eight long straight sections, three are used for injection and RF cavities and the remaining five will be used for insertion devices [57]. Basic design parameters of Indus-2 are mentioned in Table. 2.1.

Table 2.1: Design parameters of Indus-2

Energy	2.5 GeV
Current	300 mA
Bending Field	1.502 T
Critical Wavelength (λ_0)	1.986 A^0 (BM) 0.596 A^0 (WS) 1.659 A^0 (MW)
Circumference	172.4743 m
Beam emittance δ_x δ_y	5.81×10^{-8} $mrad$ 5.81×10^{-9} $mrad$
Electron beam size and divergence σ_x, σ_y Centre of bending magnet σ'_x, σ'_y	$0.234, 0.237$ mm $0.359, 0.062$ $mrad$
Centre of insertion section σ_x, σ_y σ'_x, σ'_y	$0.903, 0.108$ mm $0.064, 0.054$ $mrad$
Bunch length	2.23 cm
Beam life time ($1\%coupling$)	18 $hours$
Energy spread	9.0×10^{-4}
Damping times (t_x, t_y, t_e)	$4.74, 4.62, 2.28$ ms
Revolution frequency	1.738 MHz
RF frequency	505.812 MHz
Harmonic number	291
Power loss	186.6 kW (BM) 6.7 kW (WS) 8.1 kW (MW)

BM : Bending Magnet MW: Multipole Wiggler ($1.8T$) WS :Wavelength shifter ($5T$)

The data from Table. 2.1 were used to calculate some useful characteristics of the source. The calculations were performed using the *bm* 1.1 utility of the XOP 2.1 package

[52] and the results are given in Table. 2.2.

Table 2.2: Calculated characteristics of a bending magnet source at Indus-2

Magnetic radius $\rho[m]$	5.552
Magnetic Field $B_0[T]$	1.502
$E/(mc^2)$ γ	4892.378
Critical Wavelength λ_0 $[A^0]$	1.985
Critical Energy $E_c[eV]$	6242.929
Total Power (2π) $[kW]$	186.979
Total Power $[Watts/mrad]$	29.758

Table 2.3: Elettra Wiggler specifications with INDUS-2 ring parameter

Beam energy	2.5 GeV
Magnetic Field $B_0[T]$	1.50200
Period length	14 cm
Overall length	4.5 m
Number of poles	57
Critical Energy	6.375 keV
Deflection parameter K	20
Total Power	11.7 kW

The radiated power from the horizontal 2 $mrad$ cone (with full vertical spectrum) will be approximately 60 W, at a stored current of 300 mA. The photon flux from a 1.5 T bending magnet (BM) source, using the XOP program [52], are shown in Fig. 2.2. The vertically integrated photon flux generated from the BM source is greater than $1 \times 10^{12} photons/(sec.mrad.0.1\%BW)$ at 20 keV.

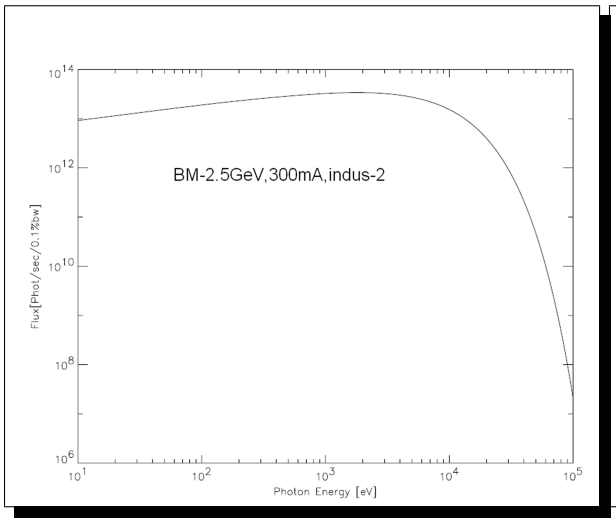

Figure 2.2: Photon flux from 2.5 GeV Indus-2 for $1.5T$ BM

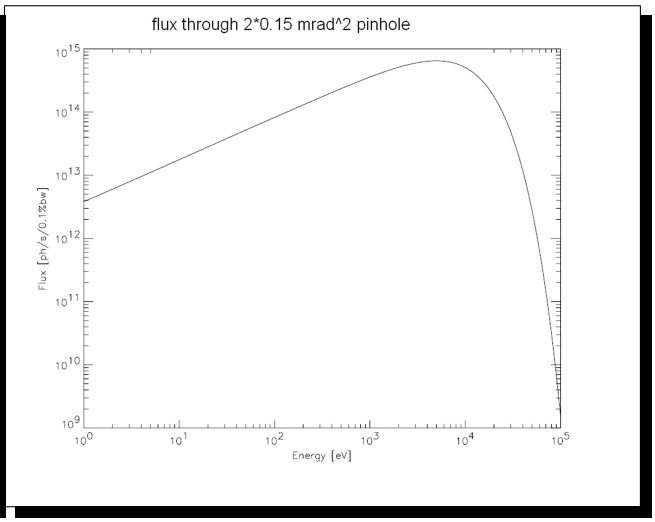

Figure 2.3: Photon flux from 2.5 GeV Indus-2 for Wiggler

The Wiggler at Elettra consists of 57 magnetic poles with a period of 14 cm and a magnetic field strength of 1.5 Tesla, which gives horizontal radiation fan of $\pm$ 4.5 $mrad$. Approximately, a 11 kW power will be radiated by the Wiggler source at 300 mA storage

ring current. The radiated power from the central cone of dimension 2 $mrad$ (H) $\times$ 0.15 $mrad$ (V) will be 1600 W at a stored current of 300 mA, this emitted power will be used for the heat load calculations and thermal designing of the beamline components. The Fig. 2.3 shows the photon flux, through a pin hole of dimension 2 $mrad$ (H) $\times$ 0.15 $mrad$ (V), generated from the Wiggler source. The photon flux generated from the Wiggler is greater than 2×10^{13} $photons/$ $(sec.mrad.0.1\%BW)$ at $20keV$ as shown by the Fig. 2.3.

2.2 Beamline requirement

Requirements of a beamline, depend on the kind of experiments to be performed. Basically, the PX-BL21 will offer four broad categories of experiments.

- X-ray studies of small protein crystals.
- Variable wavelength optimized anomalous dispersion experiments.
- Laue experiments for time resolved studies.
- Routine data collection at a chosen wavelength.

The diffracting power of the biological samples is known to be very weak, mainly because of large unit cell, less ordered lattice packing and high water content. Depending on the experiment planned, the requirements may vary significantly. For example, studies on small crystals require high flux density and for Indus-2, this requires large source demagnification. Large demagnification causes large divergence of beam at a sample and hence reduces the unit cell size resolvable for a given detector. On the other hand, for a crystal with very large cell parameters like virus etc., high collimating X-ray beam is preferred. For MAD (multi-wavelength anomalous dispersion) experiments, difference in the intensities of the Friedel-pairs is an important parameter. The spectral resolution should be comparable to the line width of a absorption edge, for maximizing the difference in the intensities of Fiedel-pairs. Therefore, the required energy resolution should be order of $(\delta E/E)$: 2.5×10^{-4} and the beamline energy must be rapidly tunable. Time resolved studies are usually done through Laue experiments; a range of spectral bandwidth is taken, which is wide enough to record a large number of reflections simultaneously. The Laue diffraction data are collected at selected time intervals that can be used to study the time evolution of a process under study. In case of routine data collection, high flux is required

at an appropriate wavelength. Accordingly, the beamline performance requirements of the proposed protein crystallography beamline are listed in Table. 2.4.

Table 2.4: The PX-BL21 requirement

Energy Range	$5 - 20\ keV$
Energy resolution ($\delta E/E$)	$< 2.5 \times\ 10^{-4}$ for Si (111)
Spot Size (mm^2)	1.0 (H) $\times$ 0.5 (V)
Photon flux at sample	$> 10^{11}\ photons/sec$
Intensity instability	$< 0.5\%$
Beam divergence at sample	$< 4\ mrad$

2.3 Beamline Optics

Primary objective of beamline optics is to use synchrotron radiation efficiently and effectively. For this, the white beam of photons diverging from the source must be collected by the beamline elements, and should be focused to a size comparable to the sample size. The selection of optical components of the beamline (mirrors, monochromator and slits etc.) has been finalized based on following factors:

- Energy resolution and size of the photon beam at the sample.
- Energy spectrum of photon beam.
- Reflectivity of optical elements.

Various beamline configurations are possible, each having their own advantages and disadvantages. Therefore, the optimum configuration based on our experimental requirements need to be selected. Some of the possible beamline configurations, which has been implemented at various protein crystallography beamlines, are listed here [58]:

2.3.1 DCM and Toroidal Focusing Mirror

This type of beamline configuration consists of one plane double crystal monochromator (DCM) and one toroidal focusing mirror (TM) located at where one-to-one focusing is achieved. The X-Ray Diffraction (XRD-1) beamline at Elettra operates in this configuration as well. However, the energy resolution is limited by the divergence of a incident

photon beam falling on a TM mirror. Thus, the design is not suitable for MAD experiments, where the energy resolution requirements are of the order of $2 - 4 \times 10^{-4}$. Therefore, this beamline configuration has not been selected for the PX-BL21.

2.3.2 Gradient DCM and Toroidal Focusing Mirror

This type of beamline consists of one gradient DCM and one toroidal focusing mirror located at where one-to-one focusing is achieved. This design can improve the energy resolution limitation that is inherent in the above mentioned design configuration 2.3.1. But, the fabrication of a good gradient crystal monochromator is still difficult. In future, if manufacturing improvement in fabrication of a gradient crystal takes place, then this configuration will be the prefered one.

2.3.3 DCM with Toroidal Collimating and Focusing Mirrors

This type of beamline consists of one toroidal collimating mirror, one plane DCM and one toroidal focusing mirror. The design improves the energy resolution by collimating the X-ray beam in meridional plane. Therefore, the MAD experiments can be performed in this type of beamline configuration. The major disadvantage of beamline configuration is in difficulty of aligning optical axes of two toroidal mirrors accurately.

2.3.4 Collimating Mirror, Sagittal Focusing DCM and Focusing Mirror

The configuration consists of one vertical collimating mirror, one sagittal focusing DCM, and one cylindrical vertical focusing mirror. This design can deliver higher photon flux at the sample positon, almost double to that of the configuration 2.3.3 and 2.3.5. The very thin and dynamic nature of the sagittally focusing crystal is a major drawback, because this design would require significant operational care to ensure peak effectiveness all the time, especially during MAD scans requiring frequent energy changes. Therefore, this beamline configuration has not been selected for the PX-BL21. Examples of this type of configuration adopted at various beamline are; MX beamline in ALBA, biocrystallography beamline at SOLEIL and beamline 14.2 at SRS.

2.3.5 Curved Single Crystal Monochromator and Focusing Mirror

This type of configuration has one curved single crystal (Si or Ge) serving as both; a monochromator and a horizontal focusing mirror, along with vertical focusing mirror. This design provides the high photon flux, because of large horizontal capturing of white beam. However, it is also difficult to use this type of beamline for the applications that require the energy scan during experiments and data collection. An appropriately sized good crystal is required to capture wide beam cross-section by the monochromator. Also, the X-ray beam falling on sample has large divergence because of wide beam cross section. The beamline 9.6 at SRS is one of the example of this type of configuration.

2.3.6 Collimating Mirror, DCM and Toroidal Focusing Mirror

This type of configuration has one vertical collimating mirror, one plane DCM and one toroidal focusing mirror. The design can effectively provide the energy tunability, without disturbing the beam size and beam position stability at the sample position. It is currently one of the most widely used design for the experiments requiring multi-energy scan, such as XAS and MAD measurements. Therefore, this configuration has been adopted for the PX beamline. The beamline 5.0.2 at ALS, MX beamline CANDLE and BL-5A at Photon factory are some of the examples of this type of configuration.

2.4 Description of optics

Primary optical components of the beamline are; X-ray mirrors and crystals. The mirrors are used for focusing and collimating the X-ray beam. Interaction of X-rays with matter can be described in terms of complex refractive index,

$$\eta = \eta' - i\beta = 1 - \delta - i\beta \tag{2.1}$$

Where β and δ are the optical constants that account for dispersive and absorptive properties of sample material, respectively [59]. They can be expressed in terms of X-ray wavelength and materials parameter,

$$\delta = \frac{r_o}{2}\lambda^2 N_A \frac{\rho Z}{A} \quad and \quad \beta = \frac{\mu}{2\pi}\lambda, \tag{2.2}$$

Where r_0 is the classical electron radius, μ is the absorption coefficient, N_A is the Avogadro number, Z is the atomic number, A is the atomic weight, ρ is the mass density and λ is the wavelength of X-rays. Typical values of δ ranges from $10^{-2} - 10^{-5}$ and that of β from $10^{-2} - 10^{-6}$. From Snells law, the condition for total external reflection [60],

$$n = \frac{\sin(90 - \theta_c)}{\sin 90} = \cos\theta_c, \tag{2.3}$$

Where n is real part of refractive index, θ_c is critical angle for reflectivity.
This implies $1 - \delta = \cos\theta_c$. Since $\delta << 1$, expanding $\cos\theta_c$ in power series up to second order will give [60],

$$\theta_c = \sqrt{2\delta} \;\; or\ in\ practical\ units \;\; \theta_c(mrad) = 2.33\sqrt{\frac{Z\rho}{A}}\ \lambda(A^0) \tag{2.4}$$

It can be seen from Eq. 2.4 that the high density materials have large critical angle, therefore, mirrors are usually coated with heavier elements like Platinum and Rhodium. The heavy element coated mirror will efficiently reflect all the X-ray energies up to critical energy corresponding to critical angle as per Eq. 2.4, and all the X-rays with energy above the critical energy will be absorbed at mirror surface. This mechanism can be used to cut down shorter wavelength or higher energy radiation from the beam, thus mirror can be used as low pass filter. This is particularly useful in reducing the heat load on first crystal of DCM by cutting unwanted high energy X-rays. Also, the length of mirror required to reflect the X-rays will be lower in case of high Z coated material due to higher critical angle.

The other important component of beamline is the crystal that is used for monochromatization of white synchrotron beam. Besides, the monochromator can be used for focusing of the beam. The principle of a monochromator is based on the well known Braggs law [61],

$$2d\sin\theta = m\lambda \tag{2.5}$$

X-ray diffraction in crystalline material provides an efficient mechanism for wavelength selection. Type of crystal used as a monochromator depends upon range and resolution of energy required for the experiment. Silicon and germanium crystals are most commonly used in beamline to monochromatize the X-rays because of their high thermal conductivity with low expansion coefficient. Also, the good quality polished crystal with low surface roughness and small slope error are available due to the presence of mature manufacturing

and fabrication process, in both Si and Ge.

The energy resolution of a crystal depends upon; its rocking width, source divergence in dispersing plane and angular size of a entrance slit of monochromator subtended at the source. The resolution may be increased by selecting a diffracting plane of smaller lattice d-spacing, at the cost of lesser reflectivity. The X-ray flux and the energy resolution, both are the important factor in an experiment. The $Si(111)$ has higher reflectivity compared to $Si(220)$ with reasonable resolution, therefore, the $Si(111)$ crystal is most widely used in second generation synchrotron sources. In third generation sources, where flux is large at higher energy ($> 15keV$), $Si(220)$ will be more useful. The energy resolution of crystal is given by Eq. 2.6 [62],

$$\left(\frac{\delta E}{E}\right)_{total} = \frac{\delta\lambda}{\lambda} = \delta\theta.\cot\theta_B = \sqrt{\left(\frac{\delta E}{E}\right)^2_{crystal} + [\delta\theta^2_{sourcesize} + \delta\theta^2_{slit}]\cot^2\theta_B} \qquad (2.6)$$

Where $\delta\theta_{sourcesize} = v/r_1$ and $(\delta\theta)_{slit} = s/r_2$, v is the vertical size of the photon source, r_1 is the distance between the crystal and photon source, s is the entrance slit opening (or vertical acceptance distance) of the crystal, and r_2 is the distance between the slit and photon source. The resolution for Si (111), Si (220) and Ge (111) are about 1.3×10^{-4}, 5.6×10^{-5}, *and* 3.2×10^{-4}, respectively.

The layout of the protein crystallography beamline is shown in Fig. 2.4. It consists the following major optical components; Collimating mirror, Double Crystal Monochromator and Toroidal focusing mirror.

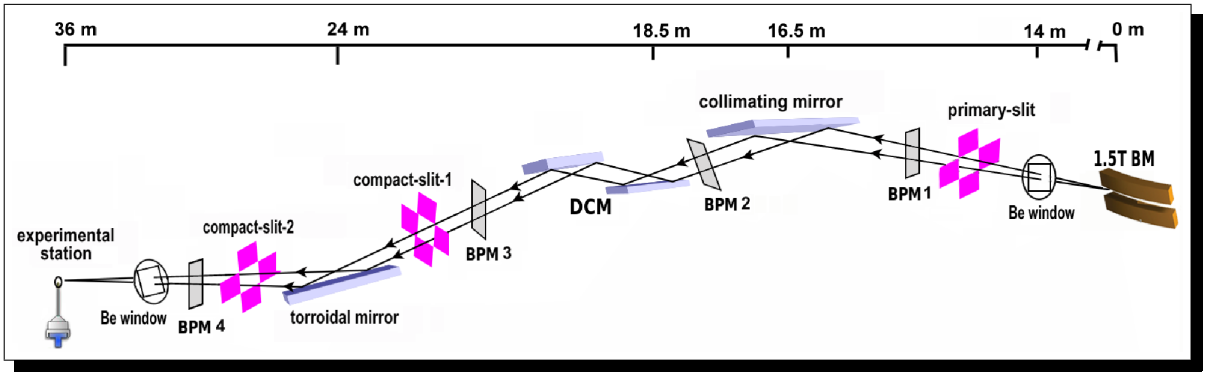

Figure 2.4: Schematic of the optical layout of the beamline. The pre-mirror, at a distance of $16.5\ m$ from the tangent point, is vertically collimating. The double crystal monochromator (DCM) is in non dispersive mode and the toroidal mirror at $24.0\ m$ from the tangent point is focusing horizontally as well as vertically.

2.4.1 Collimating mirror

The collimating mirror (CM) is a plane bendable mirror, which collimates the white beam in meridional plane that improves the energy resolution. The CM is placed at 16.5 m from the source and have dynamic radius of curvature in meridional plane. The radius of curvature in meridional plane can be varied from 9 km to 14.0 km by using a bender. This is the first optical element of the beamline, which acts as low pass filter. It cuts off all short wavelengths or high-energy beam, depending upon the type of coating on a mirror surface and an angle of incidence of the X-ray beam on mirror.

Choice of a mirror coating depends upon required energy range and type of experiments. In our case, the required energy range is $5-20\ keV$, where most of the absorption edges of heavy metals fall. The reflectivity as a function of energy for Platinum (Pt) and Rhodium (Rh) coated surface as shown in Fig. 2.5. In case of Rh coated mirror, the reflectivity response curve (Fig. 2.5) is flat compared to Pt coated mirror in the required energy range. Eq. 2.4 shows that smaller the incidence angle, larger the cut-off energy. For example, the cut off energy for Rh coated X-ray mirror, at 3 $mrad$ angle of incidence, is 22 keV. The cut-off energy can be shifted to 25 keV by selecting Pt coating, but it is the uniform reflectivity of Rh coating compared to the Pt that is more important for an experiment. Moreover, the Pt coating on mirror surface induces intensity variations near its absorption edge, as shown in Fig. 2.5. This makes experiments using Pt as a heavy atom in protein sample, much more difficult. The cut-off energy range can also be increased by using the combination of coatings, choosing base layer of Pt and top layer of Rh of appropriate thickness.

2.4.2 Monochromator

A double crystal monochromator (DCM) configuration has been selected for the PX-BL21. A typical DCM select and transmit X-rays of desired photon energy from an incident white synchrotron radiation beam. There are two types of crystal arrangements for the DCM. One is called parallel or non-dispersive $(+,-)$ [63], where reflecting surface of crystal face each other, and the other is anti-parallel or dispersive $(+,+)$ arrangement [64] in which reflecting surface of crystals are placed side by side. The $(+,+)$ mode is referred to as the high resolution mode. Although, $(+,+)$ is high resolution mode, it is difficult to move whole end station during energy scan. In $(+,-)$ mode, a outgoing beam is parallel to a incoming beam, which is fixed spatially and directionally, during the energy scan. Therefore, the $(+,-)$ mode is selected for the DCM in the PX-BL21.

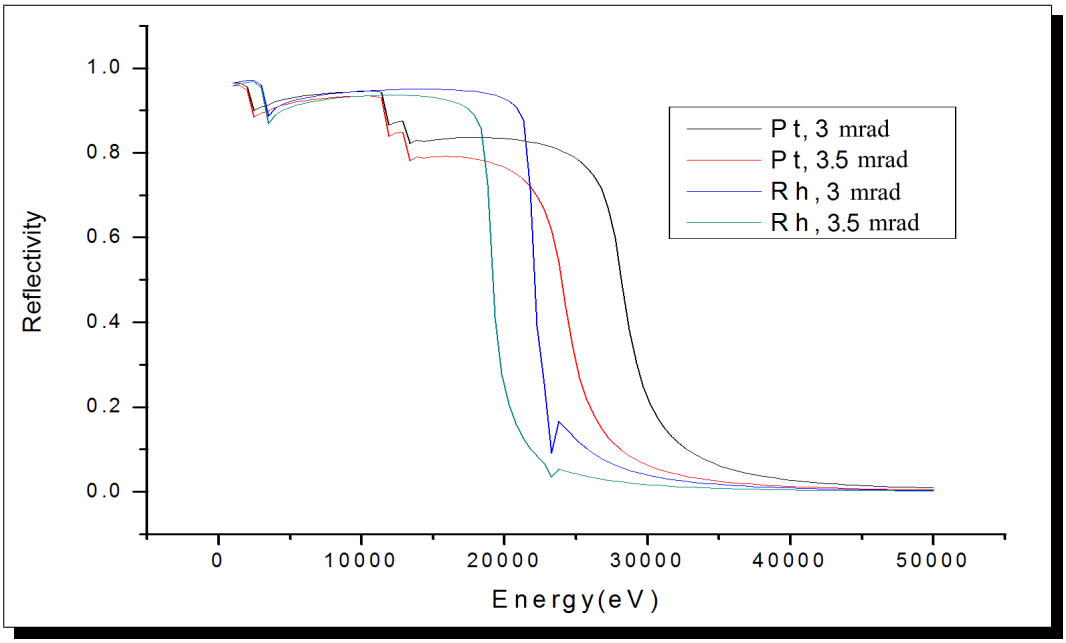

Figure 2.5: Reflectivity of Pt and Rh surfaces as a function incidence energy for different glancing angle

A monochromator with $(+,-)$ configuration can have two options; either a channel cut monochromator or two independent crystal monochromator. Both have their advantages and disadvantages. A channel cut monochromator is single block of crystal with a channel cut through it, therefore, the parallelism between the reflecting surfaces is automatically ensured. Hence, the procedure for wavelength modification becomes fast and straightforward. In addition, a channel cut monochromator is a single block of material, and hence cooled from the same source, eliminating the mismatches between separated diffracting surfaces that can occur at high power synchrotron sources. The major disadvantage of the channel-cut option is the non-fix exit. The offset between the incident white beam and the exit monochromatic beam is almost twice the gap d between the reflecting surfaces (Offset $= 2d\cos\theta$). This offset, then progressively changes throughout the whole energy range due to change in θ value. Besides, cutting and polishing of reflecting surfaces is difficult in a channel-cut block.

In case of two independent crystals, cutting and polishing is easy and the required crystal surface with the desired accuracy can be obtained. But the monochromator with two independent crystal must have an accurate mechanism to adjust the gap for a selected wavelength, while maintaining a fixed exit beam position. The gap and the parallelism between the two crystals have to be checked and tuned carefully after each energy modification. Since, the precision motor drives are now easily available, crystal motion can

be controlled with desired precision and accuracy. Therefore, 'two independent crystal' monochromator for the PX-BL21 has been chosen, as shown in Fig. 2.6.

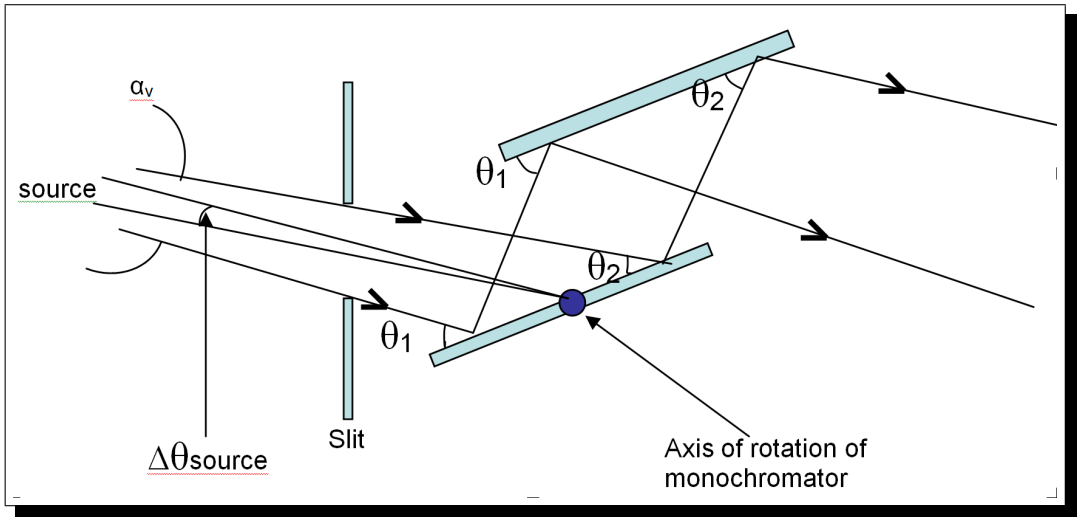

Figure 2.6: The contribution of source divergence α_v and size $\delta\theta_{source}$ to the range of energies reflected by a $(+,-)$ double crystal monochromator.

Thus, the DCM with $(+,-)$ configuration and two independent crystals, with two pairs of flat parallel crystals (Si (111) & Si (220)) will be used for PX-BL21. First crystal will disperse incoming white beam and the second crystal will redirect dispersed beam in the parallel direction of incoming beam with fixed offset. The second crystal will also be slightly detuned to reject higher harmonics. Since the DCM is not used for focusing, theoretically it can be placed anywhere between a collimator and a toroidal mirror.

2.4.3 Toroidal mirror

A toroidal mirror (TM) has been used for focusing the X-ray beam at the sample. It effectively focuses in both horizontal as well as vertical direction. Parameters of mirror are given in Table. 2.5. Although, certain aberrations are associated with toroidal focusing and can be reduced for certain configurations of demagnification. The toroidal mirror focusing has many advantages:

- A single mirror can effectively focus the photon beam in both horizontal and vertical direction,
- A mirror based focusing is more simple and robust compared to crystal based focusing,

- Since only single mirror is used for focusing, it is much cheaper compared to Kirkpatrick and Baez (KB) mirror based focusing. A KB-mirror system has two reflecting surfaces due to which, flux is less as the reflectivity is not 100%. Therefore, for BM & Wiggler beamline toroidal mirror is the ideal choice; while for the Undulator beamlines (where flux is much higher) KB-mirror based focusing will be more useful,

- A toroidal mirror is cheaper to manufacture compared to any other asymmetric mirrors which can focus photon beams in both directions,

Focusing formula for a toroidal mirror in tangential and sagittal direction are given by Eq. 2.7 [65].

$$\frac{1}{r_t} + \frac{1}{r'_t} = \frac{2}{R\cos\alpha} \quad and \quad \frac{1}{r_s} + \frac{1}{r'_s} = \frac{2\cos\alpha}{\rho} \tag{2.7}$$

Where r is object distance and r' is image distance from mirror. R, ρ and α represent the major radius, the minor radius and the angle of incidence measured from the normal, respectively. In $1:1$ toroidal mirror configuration, for point source to point image, the chromatic aberrations (sagittal and tangential) vanishes, which results into the high image quality.

Alastair A. et.al [65] has shown that the same beneficial effect occurs under the following conditions,
$r_s = arbitrary, \quad r_t = \infty, \quad r'_s = r_s/2, \quad r'_t = r'_s$
Here the source is collimated in meridonal plane (source vertical line lies at infinity) and have finite divergence in saggital plane (source horizontal line lies at finite distane from the mirror). It has also been shown that the coma terms of the optical-path-function expansion under these magic conditions are as per equation Eq. 2.8,

$$\begin{aligned} F_{120} &= \frac{\sin\alpha}{2r'^2}\left(\frac{r'}{r_s}+1\right)\left(\frac{r'}{r_s}-\frac{1}{2}\right) \qquad (Astigmatic\ coma), \\ F_{300} &= \frac{\cos^2\alpha\sin\alpha}{2r'^2} \qquad (Primary\ coma), \end{aligned} \tag{2.8}$$

Where r' is the image distance and r_s is the object distance, α is the incidence angle to the normal. The principal aberration (the astigmatic coma) vanishes at $r' = r_s/2$. The primary coma, although not exactly zero, can be sufficiently small due to grazing incidence of angle, as can be seen from Eq. 2.8. Therefore, the toroidal mirror in $2:1$ configuration for protein crystallography beamline will be the best option. The total available length for the beamline is 36 m from source to sample position, hence the torroidal mirror has

been positioned in such a way that the image distance will be $r' = r_s/2$, $r' = 12\ m$ and source distance will be $r_s = 24\ m$.

2.4.4 Ray Tracing Calculations

The toroidal mirror with 2 : 1 configuration has been chosen for PX-BL21 beamline. A toroidal mirror is placed at a distance of 24 m from the source, to focus the beam at sample placed at 12 m distance from the toroidal mirror utilizing the total beamline length of 36 m from the tangent point of the BM source. The SHADOW [54], a ray tracing software, has been used for the simulation of X-ray beam interaction with the optical elements of the beamline. 15000 rays of 10 keV energy with 20 eV band width, were used for ray tracing simulation. Ray tracing was performed for collimator-DCM-toroidal (CM-DCM-TM) configuration. Sources generated by SHADOW for the BM and Wiggler source are shown in Fig. 2.7. and Fig. 2.8, respectively. The calculated BM source size is $\sim 0.5 \times 0.5\ mm^2$, while it is $\sim 2.1 \times 0.3\ mm^2$ for the Wiggler source. The calculated spot size, flux at sample position, energy resolution, radius of collimating and focusing mirror are given in Table. 2.5.

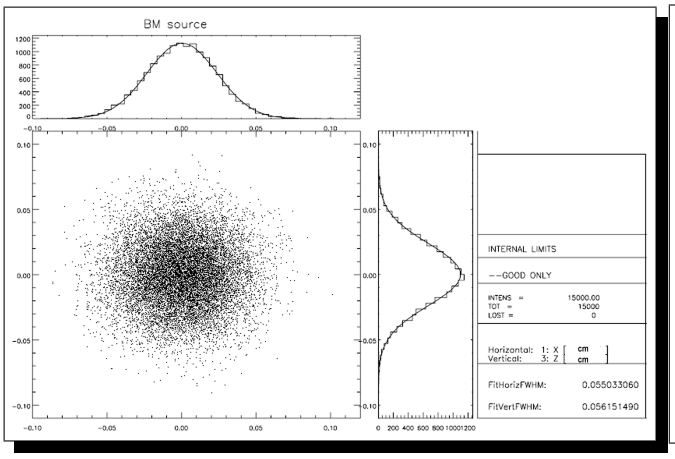

Figure 2.7: 1.5 T BM Source Size

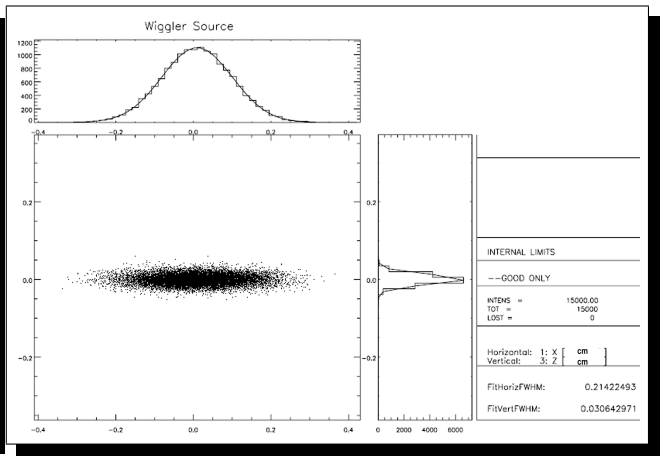

Figure 2.8: 1.5 T, 57 pole Wiggler Source Size

Ray tracing has been done for perfect optics i.e. optics without any slope error and surface aberrations. Finite dimensions of optical elements were taken into consideration in these calculations. Reflectivity of Rh coating (mirror) and silicon crystal were calculated using shadow preprocessor PREREFL and BRAGG [54], respectively. The calculated spot size at the sample position, for the bending magnet and Wiggler source, are shown in Fig. 2.9 and Fig. 2.10, respectively. The calculated spot size at sample for the BM source is $\sim 0.5 \times 0.4\ mm^2$, while it is $\sim 1.2 \times 0.5\ mm^2$ for the Wiggler source. The

required tangential radius of curvature for the CM mirror is 10.5 km, and for toroidal mirror the major and minor radii of curvature are 8.5 km and 48.5 mm, respectively, to achieve the calculated spot size at the sample position.

Table 2.5: Shadow ray tracing results

				Collimator-DCM-Toroidal (2:1)
Energy range				5-20 keV for Rh coating
Flux BM (10keV) at sample (ph/sec) Flux Wig (10keV) at sample (ph/sec)				8.03×10^{11} 9.5×10^{12}
Energy Resolution (Δ E/ E) for BM		5keV		1.2×10^{-4}
		10keV		1.8×10^{-4}
		18keV		3.2×10^{-4}
Energy Resolution (Δ E/ E) for Wig		5keV		1.16×10^{-4}
		10keV		1.4×10^{-4}
		18keV		2.15×10^{-4}
Spot size (FWHM) **mm(h) x mm(v)**		BM		0.5×0.5
		Wig		1.2×0.5
Beam divergence (FWHM) **mrad(h) x mrad(v)**		BM		3.3×0.3
		Wig		3.3×0.4
Collimator	Radius of curvature			10.5km
	Angle of Incidence (mrad)			3
Toroidal	Radius of curvature	Major		8.1km
		Minor		48.5mm
	Angle of Incidence (mrad)			3

2.5 Heat load on optical elements

Most of the heat load from the synchrotron radiation will be shared by the first and the second optical component i.e., the collimating mirror and the first crystal of DCM of the PX-BL21. The collimating mirror will remove high energy X-rays ($>$ 22 keV) from beam, while the remaining power will mostly be absorbed by the first crystal of DCM. Consequently, both components require the efficient cooling mechanism, which may be either water cooling or liquid gallium cooling or liquid nitrogen cooling. The calculation of convection coefficient (h) for heat transfer during cooling, is done by following equations

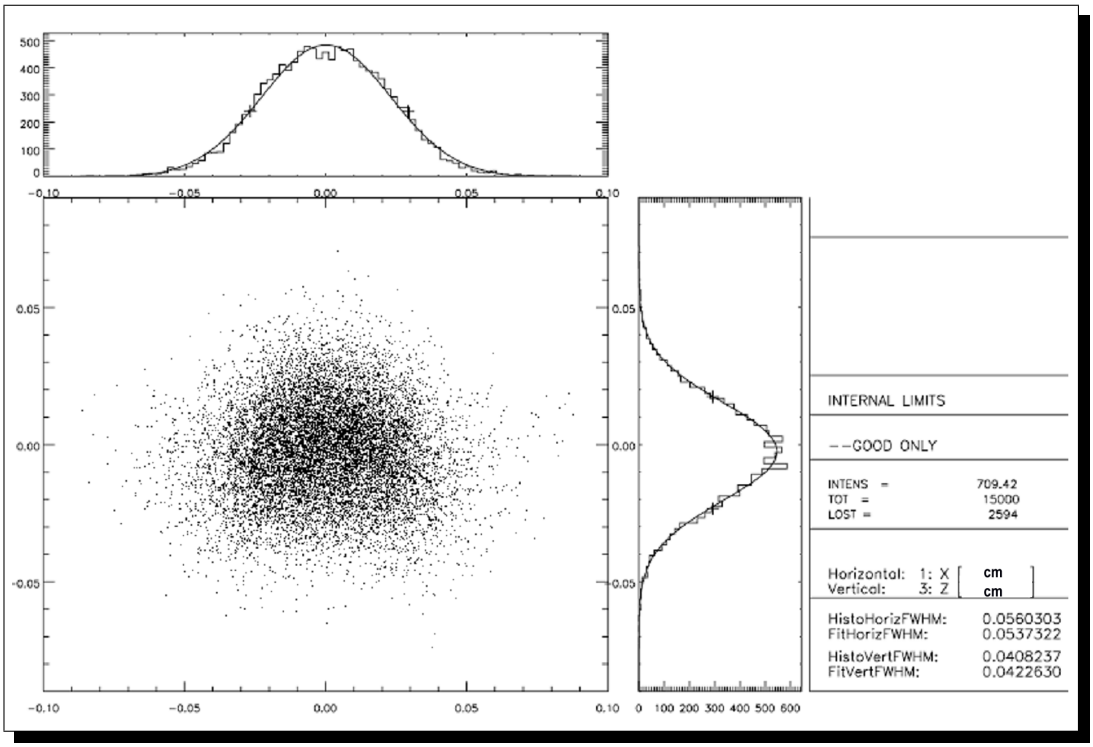

Figure 2.9: Spot size at sample position at $36\ m$ from bending magnet source

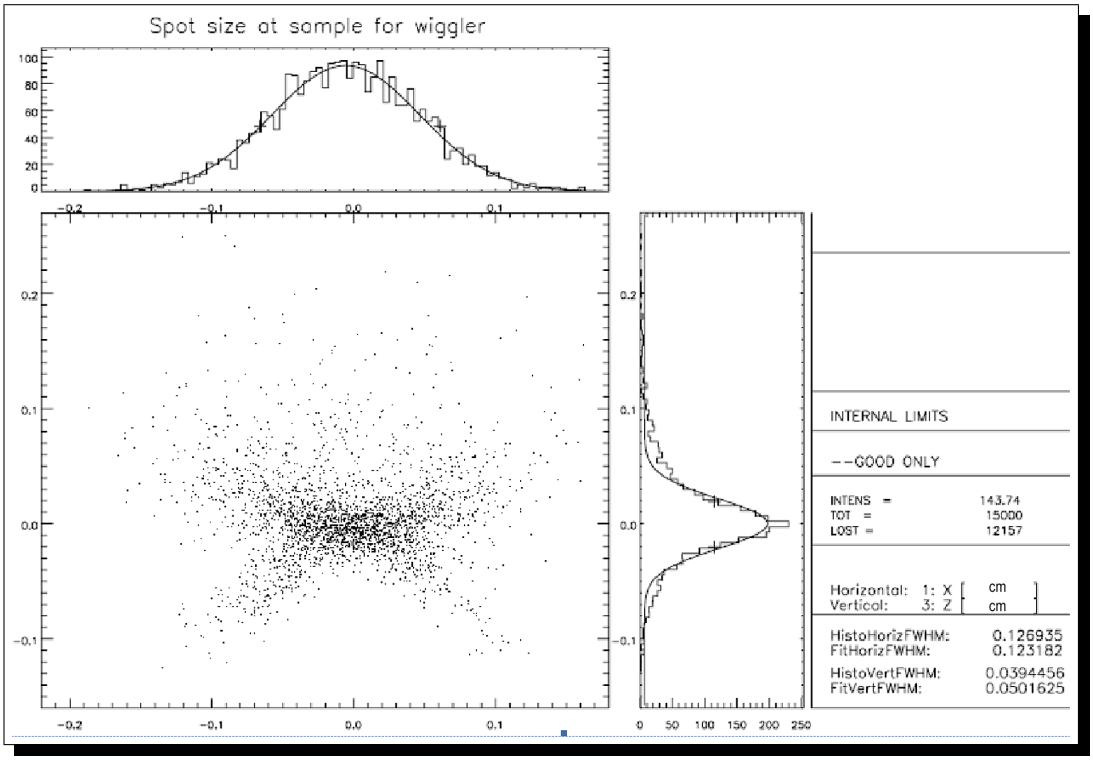

Figure 2.10: Spot size at sample position at $36\ m$ from Wiggler source

[66]:

$$h(T) = \frac{Nuk(T)}{D} \tag{2.9}$$

Where Nu is Nusselt number, D is hydraulic diameter and k is thermal conductivity of the coolant. For liquid gallium Nu is given by Eq. 2.10 and for water by the Eq. 2.11 [66]:

$$Nu = 7.0 + 0.025Re^{0.8}Pr^{0.8} \tag{2.10}$$

$$Nu = \frac{(Re - 1000)Prf/2}{1 + 12.7(Pr^{2/3} - 1)\sqrt{f/2}} \tag{2.11}$$

Here Re is Reynold number $Re(= vD/\mu)$, with v the coolant velocity; μ and ρ are the coolant dynamic viscosity and density, respectively; $Pr(= C_p\mu/k)$ is the Prandlt number, with C_p the coolant heat capacity; f is the friction coefficient given by Eq. 2.12.

$$f = (1.58\ln(Re) - 3.28)^{-2} \tag{2.12}$$

Hydraulic diameter (D) represents characteristic size of the flow channel. It is defined as four times the cross-section area divided by the perimeter of the flow channel [67].

$$D = \frac{4A_c}{p} \quad : A_c\ is\ area\ of\ cross\ section, p\ perimeter \tag{2.13}$$

$$D = Diameter, \qquad ForCircular \tag{2.14}$$

$$D = \frac{2ab}{a+b} \qquad For\ rectangle \tag{2.15}$$

Where a and b are the channel width and height, respectively. In case of cryo cooling, the APS cooling design with average convection coefficient of 10000 $Watt/m^2K$ in side wall of crystal has been employed [68]. Estimation of heat load on optical components has been carried out by XOP [52] and SPECTRA [55] codes.

2.5.1 Heat load due to BM Source.

Table. 2.6 shows the calculated heat load on the CM and the first crystal of the DCM, for the bending magnet source. Only 27 W of power comes through $2mrad \times 0.15mrad$ radiation cone, out of which 4.6 W, 3.7 W and 18.6 W is absorbed by the Be filter, mirror and crystal, respectively. Thermal and structural analysis has been carried out using ANSYS [53] code. Temperature distribution and slope error of crystal are shown in

Fig. 2.11 and Fig. 2.12 respectively.

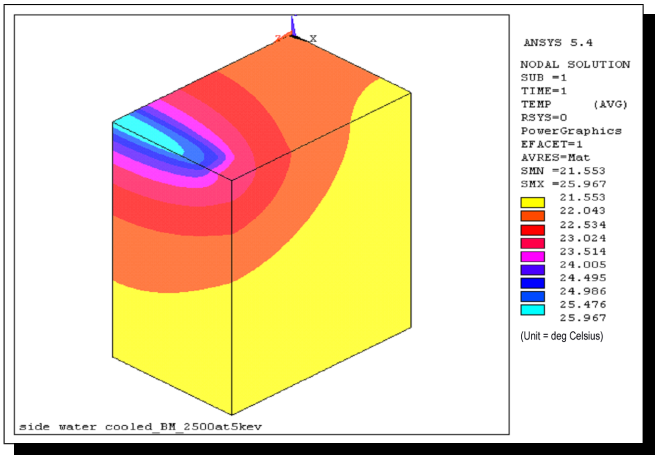

Figure 2.11: Temperature distribution side water cooled crystal for BM source at $5keV$

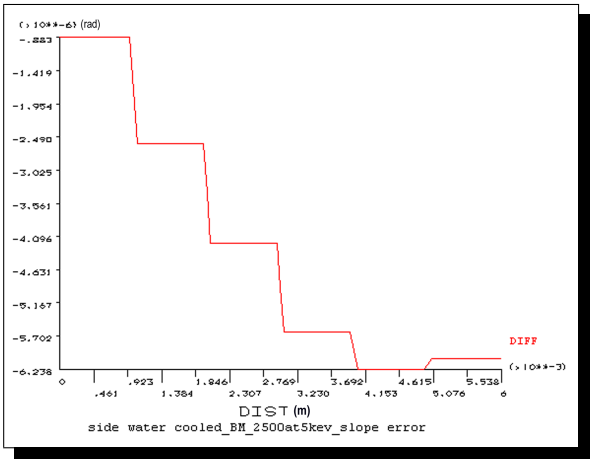

Figure 2.12: Slope error of side water cooled crystal for BM source at $5keV$

Table 2.6: Thermal loading on each of optical element of the beamline with the storage ring current of 300 mA for bending magnet. Copper mask and primary slit restricts the radiation collected to 2 $mrad$ $\times$ 0.15 $mrad$ and the total thermal power is about 27 W.

	Be-window	Collimating mirror	DCM	Toroidal Mirror
Location (m)	-	16.5	18.5	24
Power absorbed (W)	4.6	3.7	18.6	-

2.5.2 Heat load due to Wiggler Source

As mentioned earlier, the Elettra Wiggler with Indus-2 parameters has been considered for the estimation of heat absorption by the optical elements of the beamline. This heat load is calculated for 2 $mrad$ $\times$ 0.15 $mrad$ beam from central part of emission cone and 1600 W of power comes through the defined radiation cone. In order to keep heat load on first crystal of the DCM less than 1 kW, 0.1 mm thick Carbon-filters (C-filters) has been used upstream of the CM for thermal-structural calculations of the optical components. With 0.1 mm C-filters, approximately 200 W power will be absorbed by the mirror and 1 kW by the first crystal of the DCM as shown in Table. 2.7. This requires careful consideration of cooling schemes and design in order to run beamline smoothly and to avoid damage of downstream optical components due to high heat load.

Our design will be based on heat load calculation of Wiggler source. Major optical components like mirror and crystal will be designed to take care heat load from Wiggler source; therefore, design will automatically be suitable for the BM source. It is clear that large amount of thermal load will be absorbed by the optical elements due to which following problems may arise;

- Overall bending of the crystal and the mirror,
- Development of a thermal bump on the surface of optical element due to non-uniform power profile of beam causing non-uniform power absorption on surface,
- Permanent damage to crystal because of high thermal stress caused by the heat load.

Bending of the 1^{st} crystal (due to heat load) will change its lattice d spacing with respect to the 2^{nd} crystal of the DCM. This will reduce transmission efficiency of DCM due to mismatch of Braggs angle for the selected energy.

Table 2.7: The thermal loading on each of the optical elements of the beamline with the storage ring current of 300 mA for a Wiggler. Copper mask and primary slit restricts the radiation collected to 2 $mrad$ $\times$ 0.15 $mrad$, and the total thermal power is about 1600 W.

	Carbon filter (0.1 mm)	Be-window (0.2 mm)	Collimating Mirror	DCM	Toroidal Mirror
Location (m)	-	-	16.5	18.5	24
Power absorbed (W)	400	50	220	930	-

2.5.2.1 Mirror

Thermal-structural analysis of beamline optical components has been done using ANSYS [53] software. The CM works at a grazing incidence angle, and 200 W of thermal power absorbed by mirror is distributed on large surface area, therefore, a side-cooling with water is considered for analysis. For this analysis, one half of the mirror with the side water cooling through a copper plate attached to side of the mirror is used. Also, a gallium-indium eutectic (Ga-In) was applied between copper and Si-substrate to increase

the thermal transmission efficiency. Silicon is chosen as the substrate for the mirror. It has high thermal conductivity and therefore, most of the heat will be transferred to the cooling copper plate. For heat load analysis, the surface absorption (i.e heat flux is applied on a surface) of power coming from the X-ray beam has been considered, on a mirror with dimensions $1000(L) \times 80(W) \times 60(T)\ mm^3$. Heat load was applied on the $750 \times 30\ mm^2$ footprint area of the mirror. Now, this heat will conduct through the silicon substrate, then to Ga-In layer, then to copper block, followed by convection through the surface of copper by water. Since, the thermal resistance of Si-GaIn-Cu layer is small, it can be considered as a single conductive layer.

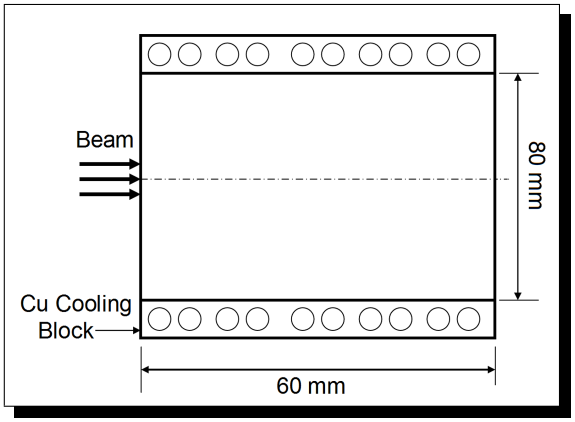

Figure 2.13: Cross-section of side cooled mirror

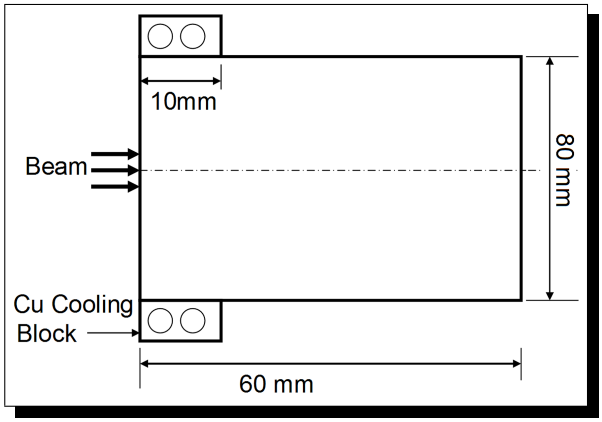

Figure 2.14: Cross-section of upper side cooled mirror

A schematic diagram of side water-cooled mirror is shown Fig. 2.13. The thermal analysis results of CM with side water cooling are shown in Figs. 2.15. The maximum temperature difference on the mirror surface is less than 2.6 °C, as shown in Fig. 2.15. The maximum slope deviation compared to the original surface (along the centerline on the mirror surface) is about 10 μrad, as shown in Fig. 2.16.

The side cooling with water is not an adequate choice for the collimating mirror (CM) due to high slope error (10 μrad). Some modifications are necessary. Instead of cooling all of the side, if only upper side surface close to reflecting surface [69] is cooled, as shown in Fig. 2.14, there will be reverse bending of bottom surface. The optimal contact-cooling (OCC) scheme consists of two optimally sized cooling blocks positioned on the sides of the mirror and flushed with the reflecting surface. The incident beam upon striking the surface, deforms the mirror into a convex shape. But within a few minutes, a certain thermal profile is established across the mirror cross section that counteracts the-convex deformation, making the mirror almost flat across its length. This process is entirely due to

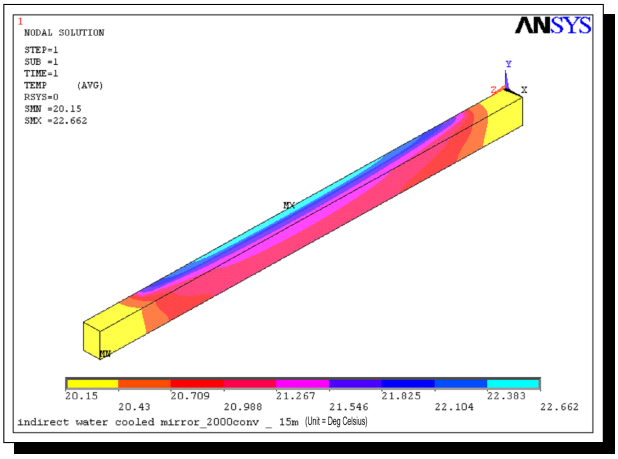

Figure 2.15: Temperature distribution of indirect side water cooled mirror for Wiggler source

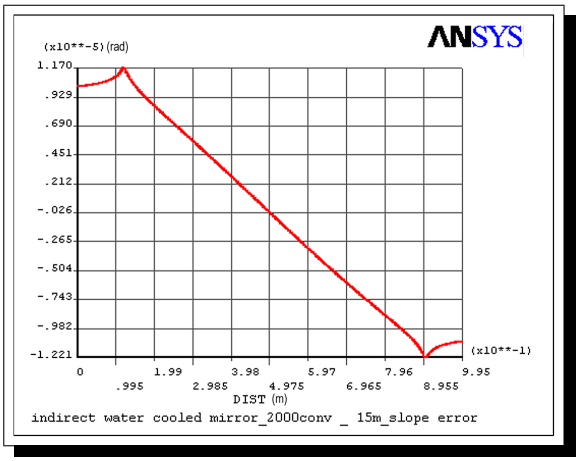

Figure 2.16: Slope error of indirect side water cooled mirror for Wiggler source

thermal moments. This will significantly counter the up bending of the reflecting surface. In this cooling scheme, the 10 mm thick upper side surface out of 60 mm, is cooled with convection coefficient 2000 W/m^2K. Maximum slope error in this case is approximately 4 μrad as shown in Fig. 2.18. Also with an improved convection coefficient 5000 W/m^2K, the slope error reduces to less than 3 μrad as shown in Fig. 2.20. Stress and slope error results are shown in Table. 2.8. Computed stress distributions for convection coefficient 2000 W/m^2K and 5000 W/m^2K are shown in Fig. 2.21 and Fig. 2.22, respectively. Maximum stress values are less than 1 MPa for the given cooling scheme, which is well within the upper safety limits of 7 MPa, the safety factor as 10 is considered.

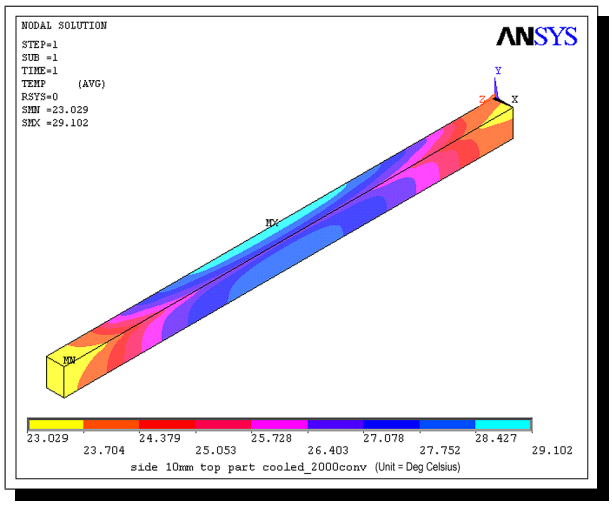

Figure 2.17: Temperature distribution of indirect water cooled mirror for Wiggler source with convection coefficient 2000 W/m^2K

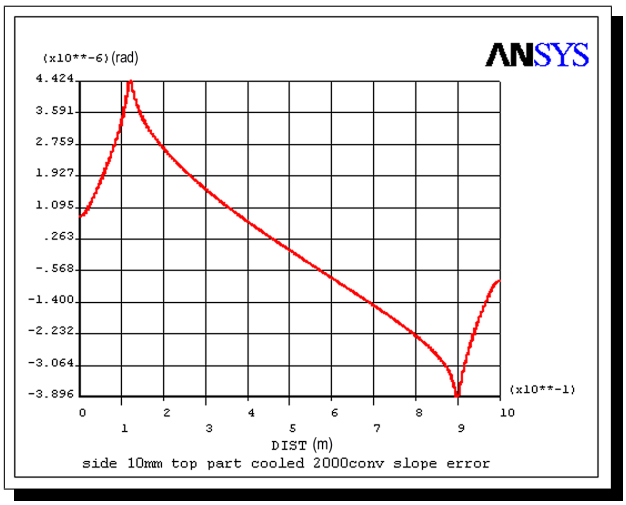

Figure 2.18: Slope error of indirect water cooled mirror for Wiggler source with convection coefficient 2000 W/m^2K

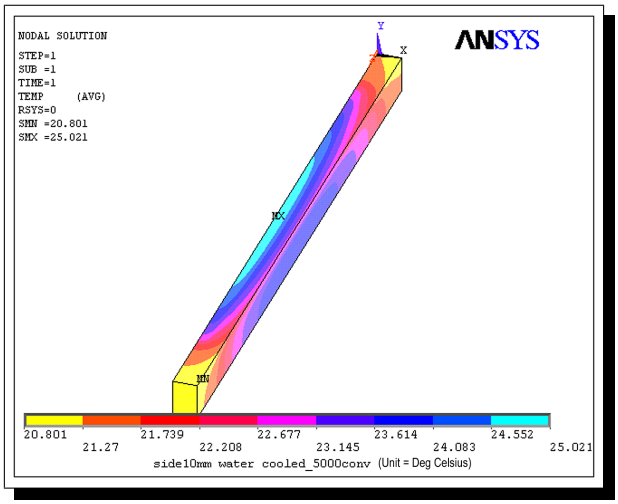

Figure 2.19: Temperature distribution of indirect water cooled mirror for Wiggler source with convection coefficient $5000\ W/m^2K$

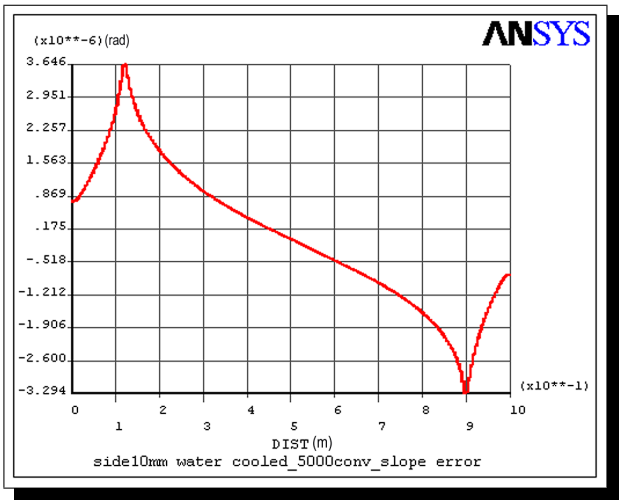

Figure 2.20: Slope error of indirect water cooled mirror for Wiggler source with convection coefficient $5000\ W/m^2K$

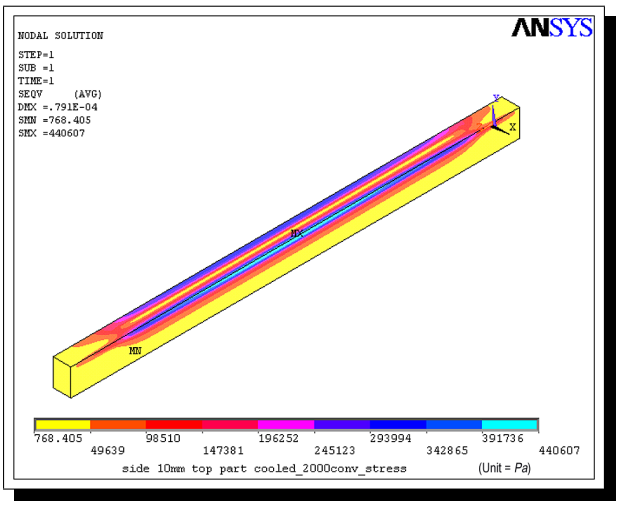

Figure 2.21: Stress distribution of indirect water cooled mirror for Wiggler source with convection coefficient $2000\ W/m^2K$

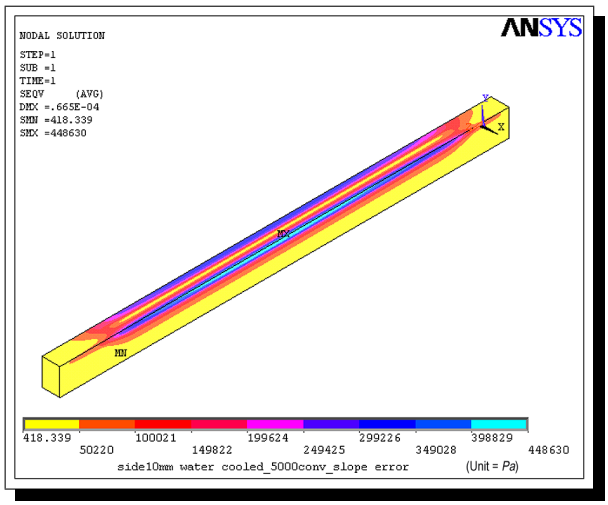

Figure 2.22: Stress distribution of indirect water cooled mirror for Wiggler source with convection coefficient $5000\ W/m^2K$

Another efficient cooling scheme is by cooling the mirror through copper plates inserted into two grooves along the sides of the mirror [70], shown in Fig. 2.23. The contact between the mirror and the copper plate is through a gallium-indium eutectic bath in the groove. This configuration provides efficient thermal contact between the mirror substrate and the cooling block without applying stress to the mirror from the cooling block. This cooling scheme is also suggested for the future upgradation of beamline to Wiggler source.

In white beam mode operation, the DCM will be removed from the beam path and

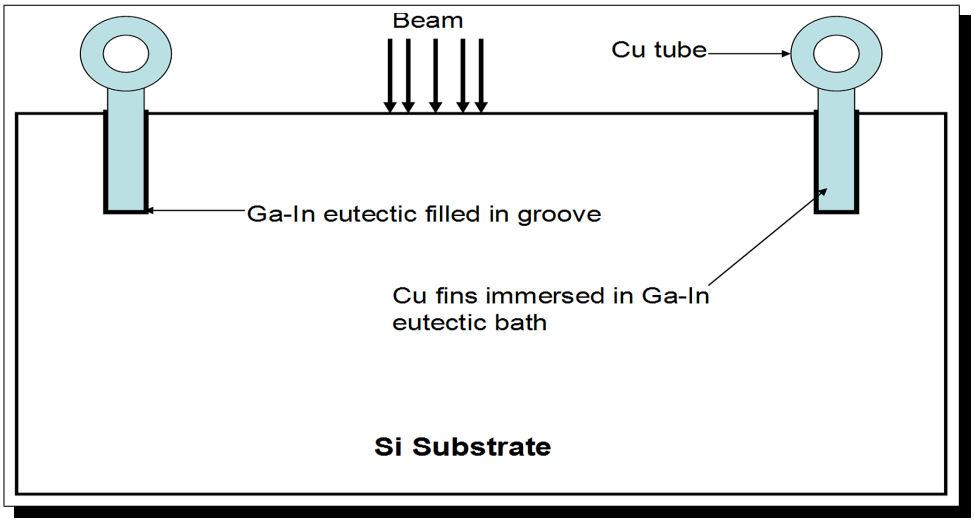

Figure 2.23: Cross-section of side water cooled mirror with groove

white beam will be allowed to fall directly on toroidal mirror. Approximately 57 W of power will be absorbed by toroidal mirror. Therefore, cooling will be required for the toroidal mirror. Same cooling geometry will be used as for the collimating mirror. Since, the adapted cooling geometry for CM can handle 200 W of heat load, the same cooling geometry has been utilized for the toroidal mirror. The finite element results of two water cooling scheme for the collimating mirror are summarized in Table .2.8.

Table 2.8: Finite element analysis results of two water cooling scheme for the collimating mirror.

	Convection coefficient ($Watt/m^2K$)	Filter (mm)	Min Temp ($^\circ C$)	Max Temp ($^\circ C$)	Max. Slope error (μrad)	Max. Stress (MPa)
Indirect side water cooling	2000	0.2Be+0.5C	20	22.6	10	0.284
Indirect side 10mm water cooling	2000	0.2Be+0.5C	23	29	4	0.44
	5000	0.2Be+0.5C	20	25	3	0.44

2.5.2.2 Double Crystal Monochromator

The first crystal of double crystal monochromator bear approximately 1kW of heat load. The DCM crystal of dimensions $60(L) \times 40(W) \times 40(T)\ mm^3$, has been used for the thermal analysis. Power absorption has been considered over whole volume, because the amount of heat falling on the first crystal is large, and it is absorbed at larger angles

compared to that of a mirror (where the angle is 3 $mrad$). Therefore, the penetration depth for heat absorption is significant, justifying the treatment as a volume absorption. Only one quarter of the crystal is used for the analysis due to symmetric distribution of power in X-ray beam. Finite element analysis shows that the amount of heat load will lead to significant amount of thermal distortion.

Analysis has been carried out for the Bragg angle (θ_C), corresponding to 5 keV X-ray energy for silicon (111) with a 0.2 mm Be and 0.1 mm carbon filter upstream to the DCM. The cooling geometry of rectangular water-cooling channels [66], below the reflecting surface, along the length of the crystal are used for heat load calculations (Fig. 2.24). A simple water-cooling can only limit the slope deviation of the crystal, caused by thermal load, to 170 μrad, which is much larger than the Darwin width ($\sim$ 20 μrad) of Si (111) in the energy range from 6 keV to 19 keV. These results for various cooling schemes are summarized in Table. 2.9.

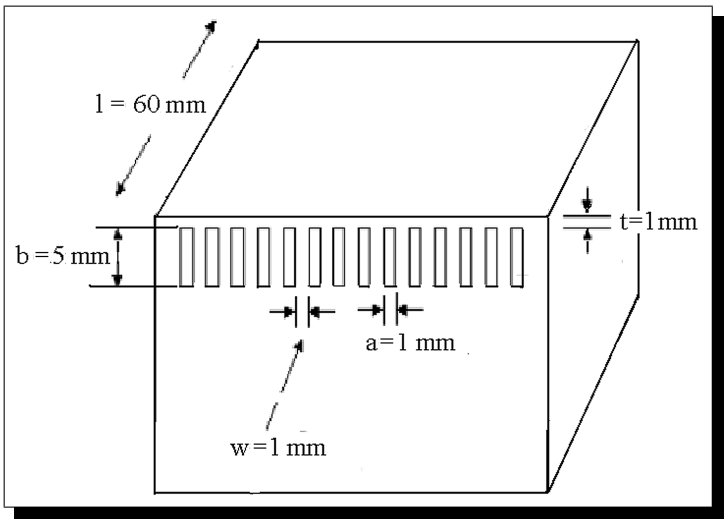

Figure 2.24: Schematic of water channel cooled crystal

Same crystal cooling geometry was also checked with liquid gallium as a coolant, for $t =$ 1 mm and $t =$ 0.5 mm thick top reflecting surface, refer to Fig. 2.24. Operating temperature of gallium is taken 50 $^{\circ}C$. The slope error at 5 keV, in Si (111), for $t =$ 1 mm thick top reflecting surface are 36μrad, similarly for $t =$ 0.5 mm thick top surface, the slope error is 10 μrad. Although, the liquid gallium cooling is giving acceptable results, the operation and handling of liquid gallium is still difficult. Therefore, direct liquid gallium cooling of first crystal of DCM is not considered.

Above analysis shows that simple direct water-cooling of crystal is not adequate. Therefore, carbon filters of various thickness need to be used upstream to reduce the

Table 2.9: Various cooling configurations for first crystal of DCM are considered for Wiggler source at 5 keV, one forth of crystal is used for simulation.

Cooling schemes	**Source type**	**Energy** (keV)	**Filter** (mm)	**Min Temp**	**Max Temp**	**Slope error** (μrad)
Direct water cooling	Wig	5	0.2Be+0.1C	20 $°C$	140 $°C$	172
Liquid Gallium cooling (1 mm)	Wig	5	0.2Be+0.1C	50 $°C$	85 $°C$	36
Liquid Gallium cooling (0.5 mm)	Wig	5	0.2Be+0.1C	50 $°C$	72 $°C$	10
Indirect Liquid Nitrogen cooling	Wig	5	0.2Be+0.1C	97 K	134 K	22
Indirect Liquid Nitrogen cooling	Wig	5	0.2Be+0.5C	91 K	110 K	9.6
Indirect Liquid Nitrogen cooling	Wig	10	0.2Be+0.5C	90 K	107 K	8.2
Indirect Liquid Nitrogen cooling	Wig	18	0.2Be+0.5C	90 K	102 K	2.5
Indirect side water cooling	BM	5	0.2Be	20 $°C$	25 $°C$	6

heat load on optical components. Table. 2.10, shows the power absorption and the heat load on downstream optics with different carbon foil thickness. The slope error and the deformation of first crystal is calculated for various carbon foil thickness and accordingly different cooling schemes are considered, viz., direct water cooling, micro channel cooling, liquid nitrogen cooling. Micro channel cooling [71] and cryo cooling [72] have shown acceptable results. Problem with micro channel cooling lies essentially with the manufacturing difficulties that almost always leave the crystal highly strained. Liquid nitrogen cooling has many advantages over water cooling due the excellent thermal properties of Silicon at a low temperature. Comparison of the Silicon properties at low and room temperature are shown in Table. 2.11.

Silicon has high thermal conductivity (k (W/cmK)) and low negative thermal expansion coefficient (α ($10^{-6}K^{-1}$)) below 125 K, as shown in Table. 2.11. Therefore, the temperature gradient in the crystal causes smaller distortions on surface of silicon crystal, which

Table 2.10: Power absorbed by mirror and DCM for different carbon foil thickness. Power absorbed is out of 1600 W in unit of $Watt$ at each components Using XOP.

C filter thickness ⇒	100 μm	200 μm	300 μm	400 μm	500 μm
Mirror (Rh coated)	213	205	199	194	190
Power on DCM	936	845	780	730	687

leads to better performance of monochromator compare to that at room a temperature.

Table 2.11: Properties of Silicon at Room temperature and at Cryogenic temperatures.

Material	Temp. (K)	k (W/cmK)	α $(10^{-6}K^{-1})$
Silicon	300	1.4	2.6
Silicon	100	13.4	-0.5

It is found that for 500 μm carbon filter and indirect liquid nitrogen cooling give acceptable results. All the results are summarized in Table. 2.9. For indirect liquid nitrogen cooling with 0.5 mm carbon filter, maximum slope error is less than 10 μrad and maximum temperature goes up to 108 K, at 5 keV energy as shown in and Fig. 2.26 and Fig. 2.25, respectively.

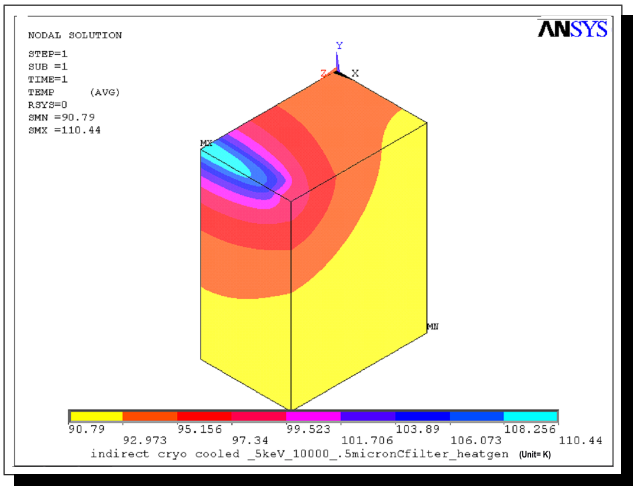

Figure 2.25: Temperature distribution of indirect liquid nitrogen cooled crystal with 0.5 mm C filter for Wiggler source at 5 keV

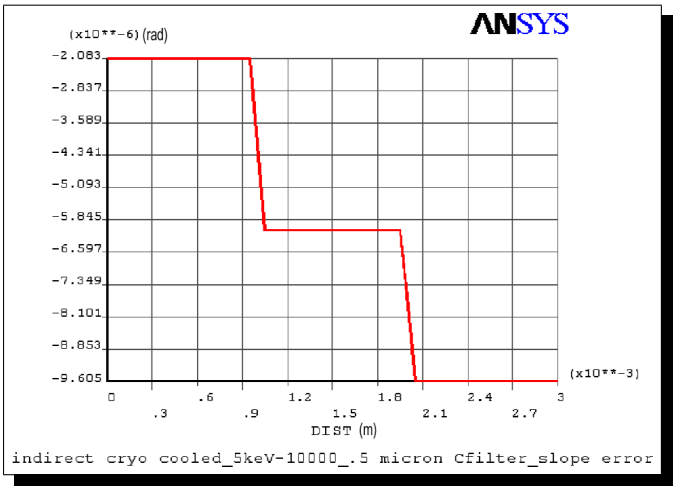

Figure 2.26: Slope error of indirect liquid nitrogen cooled crystal with 0.5 mm C filter for Wiggler source at 5 keV

The slope error at 10 keV is 8.2 μrad as shown in Fig. 2.27, which is less than the Darwin width of Si (111) i.e., 27 μrad. Similarly, at 18 keV the slope error is 2.5 μrad (Fig. 2.28), which is less than the Darwin width of Si (111) i.e., 15 μrad at same energy. The maximum stress generated due to thermal load at 5 keV energy, for indirect cryo cooling, is 1.8 MPa, as shown in Fig. 2.29, which is well below the safety limit of 7 MPa for the silicon crystal. Therefore, based of thermal-structural calculations, the indirect liquid nitrogen cooling has been suggested for the first crystal of DCM of PX-BL21. If in future, the beamline is shifted to a Wiggler source, a 500 μm carbon filter should be used upstream. About, 700 W of heat will be absorbed by the these upstream filters, therefore, there design requires a careful consideration. Since, Wiggler specification is not finalized yet, final decision on thickness and design of carbon filters will be finalized later.

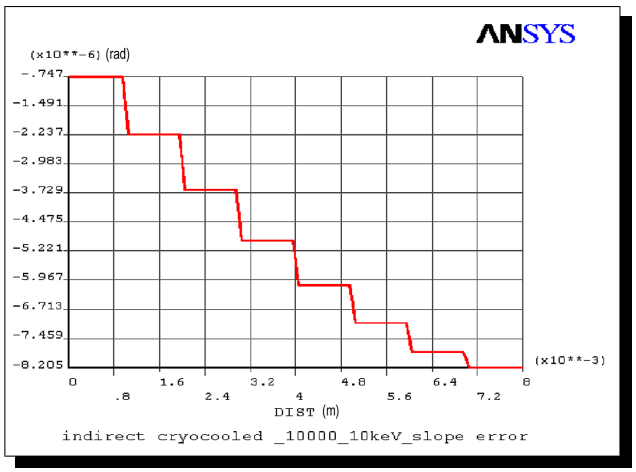

Figure 2.27: Slope error of indirect liquid nitrogen cooled crystal with 0.5 mm C filter for Wiggler source at 10 keV

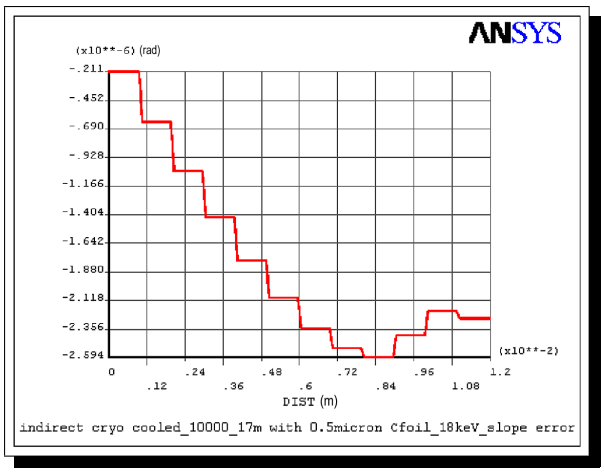

Figure 2.28: Slope error of indirect liquid nitrogen cooled crystal with 0.5 mm C filter for Wiggler source at 18 keV

Cooling system design of DCM should be compatible with both coolant, liquid nitrogen and water. For bending magnet source, same design with water coolant instead of LN2 could be used. The temperature distribution on DCM crystal due to BM flux at 5 keV energy is shown in Fig. 2.11, maximum temperature rise on surface of crystal is 5 $^{\circ}C$. The maximum slope error in this case is 6 μrad for 5 keV as shown in Fig. 2.11. Therefore, water cooling of DCM crystal is adequate for BM source.

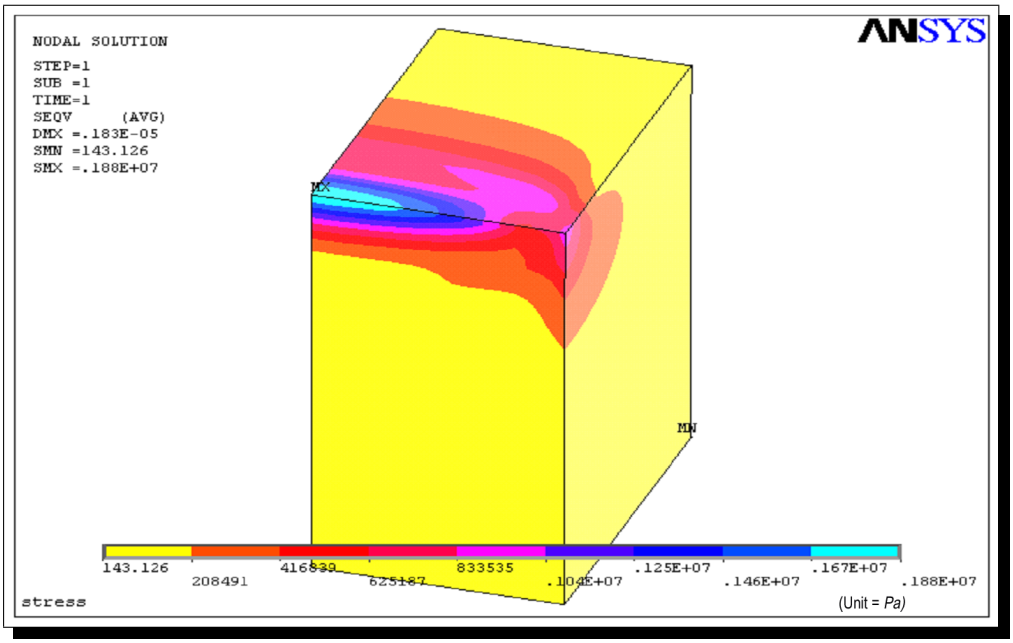

Figure 2.29: Stress distribution of indirect liquid nitrogen cooled crystal for Wiggler source at $5\ keV$

2.6 End-station

End station will be the state of art hutch for the protein crystallography experiments. Nowadays, macromolecular crystallography increasingly becoming automated. Sample mounting, data acquisition, data processing all steps in crystal structure determination are mostly automated. Our aim is to include all latest possible developments, to provide users with state of the art facility. End-station will have following major components,

2.6.1 Sample goniometer

The end-station will have a single-axis goniometer with 3-axis translational stages to permit efficient measurement of complete diffraction data sets. Goniometer will also include the slits and one fast shutter.

2.6.2 Fluorescence detector

A fluorescence detector is an important component of the end station that will aid in calibrating the monochromator and in the selection of wavelengths for MAD data collection.

2.6.3 Sample viewing optics

High quality visualization system will be required to permit accurate alignment of the sample to the X-ray beam and clear discrimination between crystal, surrounding solvent and mounting loop.

2.6.4 Detector

Broadly two type of technology is used for area detectors, viz, a direct conversion and a indirect conversion type detectors. The CCD detector is indirect conversion type, where X-ray photons first interact with phosphor and emit photons in visible region, which in-turn is detected by CCD chip. The CCD based detectors are well known and widely used in protein crystallography beamlines. These type of detectors have fast read out time of order ~ 1 *sec*. A hybrid pixel detector is a direct conversion type detector, which works in a single counting mode [73]. Key advantages of the hybrid pixel detector are direct detection of X-rays in single-photon-counting mode, zero readout noise and zero dark current. Also, the read out time is much less (~ 10 ms) than that of the CCD detector. The PX-BL21 will be installed at the BM source, where flux is less compared to undulator based beamline, therefore, data collection time per frame will be in range of a few seconds. In that case, the hybrid pixel detector would not be used at its full potential. Also, the cost of CCD based detector in much less than the hybrid pixel detector. Keeping these constraints in picture, a large (< 220 mm) and fast (< 4 s) readout area CCD detector will be suitable for, collecting high resolution data from a large unit cell, reducing time dependent decay of the sample and allowing for effective use of beam time. This heavy detector will be mounted on pseudo 2θ arm; it will be able to move 45° up and -5° down. This will permit higher resolution data to be recorded at long wavelengths, if required. Translational motion of detector i.e sample to detector distance will be 50 mm to 600 mm.

2.6.5 Low temperature device

A good quality diffraction data from protein crystal, with better signal to noise ratio, can be obtained by increasing the exposure time, which is also increases the chances of radiation damage. The radiation damage of protein crystals can be minimized by collecting data at cryo conditions. Therefore, end station will be equipped with a sample cooling device with cooling range from 90 K- 300 K. This device should not couple vibration into the goniometer, and should operate at base temperature without the need of crystal

shielding.

2.7 Summary

In this chapter, a conceptual design of protein crystallography beamline has been outlined. Beamline has been designed to accommodate insertion device (Wiggler) as a future source. Various beamline configuration were explored, and it has been found that CM-DCM-TM (sec. 2.3.6) with TM in 2 : 1 configuration will be the suitable choice for Protein crystallography beamline. Ray tracing calculation were performed to optimize the spot size at sample position. Also, based on ray tracing calculations, positions and optical specifications of X-ray mirrors and DCM has been optimized. Since, beamline would be shifted to Wiggler source, the high heat load generated by Wiggler has been taken into thermal-structural calculations of the optical elements. It was proposed that side groove cooling with water as a coolant, will be sufficient for both mirrors. In case of the DCM, the $\sim 1kW$ heat load will be absorbed by the first crystal for a Wiggler source. Therefore, a cryo cooling has been proposed for DCM with carbon filter upstream. Also, a large 2D area detector (CCD) has been proposed for high resolution and fast data collection. A robotic sample changer that will reduce the sample changing time and a fluorescence detecter that will complement the CCD detector by measuring the anomalous signal from protein crystals has been proposed for the PX-BL21.

Chapter 3

Beamline development, installation and commissioning

The development of the protein crystallography beamline (PX-BL21) at Indus-2 synchrotron at Raja Ramanna Centre for Advanced Technology (RRCAT), Indore, was taken up by High Pressure & Synchrotron Radiation Physics Division (HP&SRPD), Bhabha Atomic Research Centre (BARC) under X-XI plan project. The PX-BL21 on Indus-2 synchrotron is aimed at catering to more than 50 independent research groups in India, working in the area of structural biology.

The PX beamline installed at 1.5 T bending magnet source and is located at port no. 21 (5° port) of Indus-2 synchrotron. The beam-line "Front-end" (FE) is 14 m long and connects the storage ring to a beamline. The FE consists of a 2.2 $mrad$ (H) beam defining copper-mask and a water-cooled shutter. The FE employs a fast-closing shutter and a delay line for ultra high vacuum (UHV) protection of the storage ring. A 250 μm thick Beryllium window at the end of FE separates the storage ring vacuum from the beam-line. It also acts as a high pass filter, removing all the photons having energy below the 5 keV, thereby reducing the heat load on downstream optical components.

Most of the components of the beamline, such as water cooled primary slit, water cooled compact slits, mirror chambers along with hexapods, beam pipes, support tables and beam position monitors (BPM) were developed indigenously. The main optical components (two no.s of 1.2 m long mirrors and a DCM) were installed and connected with slits, pipes and BPMs. Ultra high vacuum (UHV) of order of $10^{-8} torr$ were achieved along 24 m length of the beamline. The alignment of each optical component was carried

out in steps with the help of in-house developed BPMs, which measure the beam position and the beam profile of the X-ray beam. The hot alignment of the beamline was completed with respect to the emitted synchrotron radiation from the bending magnet port 21. The optical components were adjusted and aligned to bring the X-ray beam at experimental station. The experimental station of beamline is the state art equipment, which was installed at the end of the beamline, at 36 m from source. The experimental station includes a large CCD detector for fast (1 sec) data collection and a cryo-system for protection of sample from radiation damage along with a robotic sample changer for automatic loading and unloading of protein crystals under cryo conditions.

3.1 Beamline development

Many beamline components were designed and fabricated at in-house facility of BARC. The Centre of Design and Manufacturing (CDM), a division in BARC participated in design and fabrication of Primary slits, water cooled compact slits, X-Ray Mirror chambers and Hexapods. Beam position monitors were designed and then fabricated locally at M/s Excel Instruments Ltd. The software development has been carried out for automated DCM energy calibration. Also for fluorescence scanning, a server-client program has been developed in Microsoft Visual Basic 6.0 (VB-6.0) to facilitate anomalous diffraction experiments.

3.1.1 Beam Position Monitors

The fluorescent based X-ray beam position monitors (BPM) measure a position and a profile of an incident X-rays ($5 - 25$ keV). Four of these BPM are designed, fabricated and installed at the PX-BL21. Detailed assembly drawings were prepared based on the requirements of the beamline. Further design optimization was carried out during fabrication stage by incorporation of a tungsten wire monitor. These BPMs have a phosphor coated water cooled copper (Cu) plate for X-ray beam viewing, and a electrically isolated, 500 μm tungsten wire placed at bottom of Cu plate. The stepper motors of BPMs have accuracy better than 5 μm. The UHV compatible fluorescent screen assembly is water cooled. Therefore, a BPM can be used with both; white beam and monochromatic beam. The specification of the BPM is summarized in Table. 3.1

Reference drawing for Beam Position Monitor is shown in Fig. 3.1. The whole assembly fits into a six-way cross, as shown in Fig. 3.2. The center of screen moves upward

Table 3.1: Beam Position Monitor Technical Specifications

Mechanism mounting flange	150 CF
Screen coating	Europium phosphor $P22R$
Copper plate dimensions	$80(h) \times 40(v)$ mm^2
Actuator type	Stepper motor
Operating Pressure Range	1×10^{-9} mbar to Atm
Leak Rate	$< 1 \times 10^{-10}$ $mbar\ l/s$
Material	Stainless Steel $304L$
Wire Material	500 μm tungsten

(out of beam path) by 70 mm. The screen plate is coated with a fluorescent material to convert X-rays to visible light. Remotely the Cu plate can be moved in and out of the beam by a control PC. A cctv camera coupled with manual zoom lens, looks through a viewport directly onto the grid plate of the screen. The tungsten wire scans through X-ray beam and a picometer connected to the wire system measures the induced photocurrent. The plot of photocurrent vs motors steps, provides the beam dimension and profile. This unique ability of BPM to locate beam using camera, and quantitative profile and position measurement using integrated wire monitor, has allowed the fast and accurate alignment of beamline components.

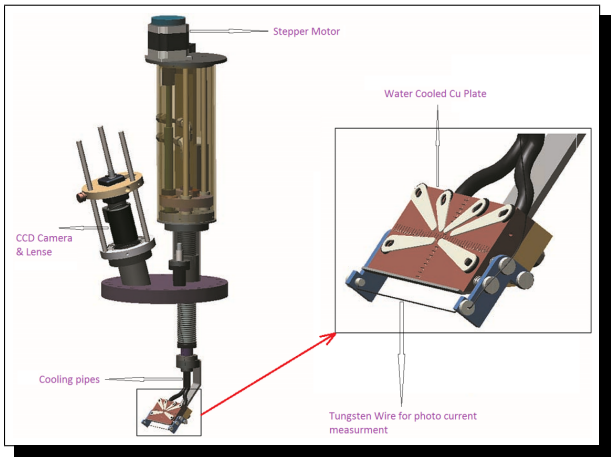

Figure 3.1: Beam Position monitor

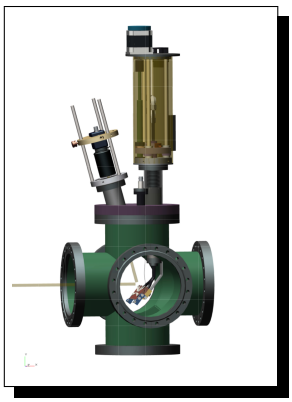

Figure 3.2: Beam Position monitor with six way cross assembly

3.1.2 Primary slit and water cooled compact slits

X-ray slits or beam defining slits are the essential components of a beamline. These are placed at various location of the beamline for X-ray beam conditioning and designed to work under ultra high vacuum (UHV) in high radiation zones. Slits are water cooled to operate in white beam mode. The criticality of the slit assembly lies in its positional accuracy, provision of heat sink, remote operation ability, UHV compatibility and compactness of the instrument. Slits are required to control the aperture of the synchrotron beam to suit the acceptance of successive instruments, like a DCM and X-ray mirrors etc. This is achieved by the synchronized motion of usually two (Fig. 3.4) or four (Fig. 3.3) independent metal blades that blocks the X-ray beam.

Slit blades are fabricated from tungsten carbide and attached to copper plates at backside, which in turn are brazed with copper pipes for cooling. Three slits are installed at the PX-BL21. The first slit is called primary slit, it is the first component of beamline. The primary slit has four independent slit blades for each direction viz. top, bottom, left and right, as shown in Fig. 3.3. All four blades are separately water cooled and have 5 μm accuracy in motor movement. The maximum and minimum slit opening are 40 mm (H) $\times$ 20 mm (V) and -0.5 mm $\times$ -0.5 mm, respectively. The length of the primary slit assembly in 0.5 m long with the option of viewing ports on both sides. Each slit mechanism has linear encoder for positional feedback and limit switches for mechanical safety of the instrument.

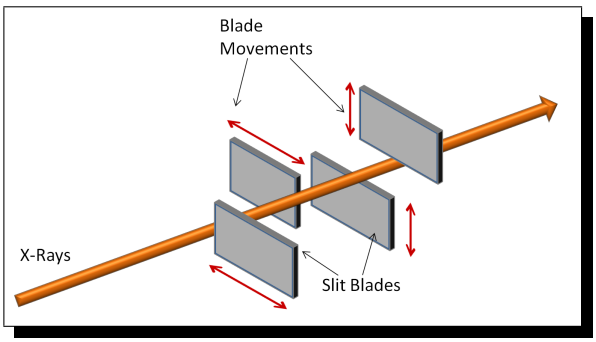

Figure 3.3: Schematic of primary slit

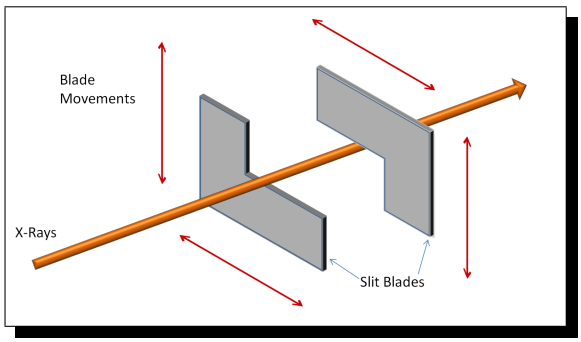

Figure 3.4: Schematic of compact slit

The second and third water cooled slit installed at the PX-BL21, are identical and have smaller dimension compared to primary slit, thus referred as Water cooled compact Slits. These precision slits are fully automated and remotely operated, similar to primary slit. Instead of four independent slit blades as compared to primary slit, the compact slit have two "L-Shaped" blades as shown in Fig. 3.4. The slit motor mechanism aligns and

moves these two "L-Shaped" blades with slider mechanism in a planar spatial space by two degree of freedom. The slider mechanisms has high flexural rigidity with positional accuracy of 2 μm. The compact slit aperture can be varied from $-0.5\ mm\ \times\ -0.5\ mm$ to 40 mm (H) $\times$ 20 mm (V) with 10 μm accuracy.

3.1.3 Mirror Chambers and Hexapod

3.1.3.1 Mirror Chambers

Two ultra high vacuum (UHV) X-ray mirror chambers were fabricated by the CDM. These mirror chambers are made of $SS304$, which is a UHV compatible stainless steel alloy. The collimating mirror (CM) chamber is 1650 mm (L) $\times$ $571mm$ (W) $\times 832$ mm (H) in dimension, as shown in Fig. 3.5. The toroidal mirror (TM) chamber has smaller dimension of 1650 mm (L) $\times$ $571mm$ (W) $\times 577$ mm (H), as shown in Fig. 3.6. The toroidal mirror has upward facing configuration, therefore, bender is below the mirror that results in compact overall dimension. While, in case of CM chamber, bender system is above the mirror due to downward facing configuration, this increase the mirror's chamber size compared to TM mirror chamber. Both chambers have two parts, a bottom flat rectangular base and a top hollow rectangular cover. The bottom flat base has openings for ion pump and three supporting rods, these three rods will support the X-ray mirror block. The top hollow rectangular cover has, beam-in and beam-out CF flange, side CF flanges for vacuum gauges, beam viewing ports and feed through for bender assembly operations. The top part of the chamber sits on the bottom part, both are separated by the metal gasket for vacuum isolation. These chambers were chemically cleaned and electro-polished to remove any surface impurity. Subsequently, the UHV testing of both chambers were carried out for leak detection in absence of the X-ray mirrors.

3.1.3.2 Hexapod

A hexapod also known as parallel manipulator [74] consist of six actuators sandwiched between two plates. Bottom plate is fixed and motion of six actuators along with a top plate, provide the six degree of freedom (6 DOF) with respect to bottom plate. The hexapod have many advantages over the serial mirror manipulator.

- It has high rigidity due to object weight shared by six legs,
- This high stiffness that provides the high accuracy,

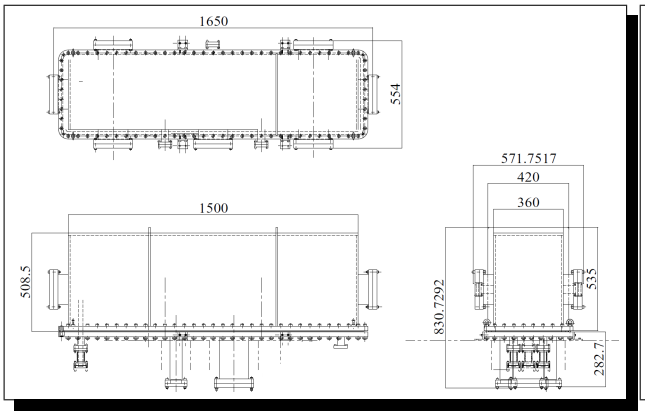

Figure 3.5: Collimating Mirror Chamber

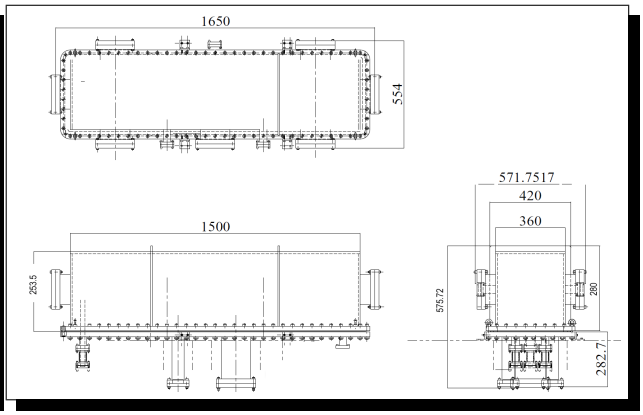

Figure 3.6: Toroidal Mirror Chamber

- Center of rotation is variable and can be defined as per the user requirement,

Two hexapod were designed and fabricated at CDM for the manipulation of collimating and toroidal mirror. The hexapod has been tested thoroughly for its motion and installed successfully (Fig. 3.7) at the PX beamline Indus-2. Each hexapod has $\pm$ 45 mm movement in each $x,\ y,\ z$ direction with $< 10\ \mu m$ accuracy. Also, it has $\pm\ 2^{\circ}$ rotational range with $\sim 1\ arcsec$ accuracy. Top and bottom plates are the granite blocks, which isolate and absorbs the surrounding vibrations. Top plate has numbers of hole to accomodate the mirror supporting rods, cooling pipes and a sputter ion pump (SIP) for the chamber. Design is such that mirror directly rest upon the top granite block via three supporting rods, without any contact with chamber. This effectively isolate the mirror from any vibrations coming through beamline components.

3.1.4 Beam line automation and software development

MAD differs from other methods and uses anomalous scattering at several wavelengths for complete phase determination without approximations or simplifying assumptions. Anomalous scattering violates Friedel's law, *i.e.* the reflections (h, k, l) and $(-h, -k, -l)$ have different intensities; the intensity difference is however very small (about few % only). Every MAD experiment requires three dataset to be taken at different energies. The energies should be chosen precisely from the absorption energy scan of a heavy element in the protein.

Anomalous data collection therefore requires 1) a precise energy from DCM and 2) determination of absorption edge of heavy element in protein crystal. In order to get precise energy from DCM, energy calibration is essential. The different wavelengths can

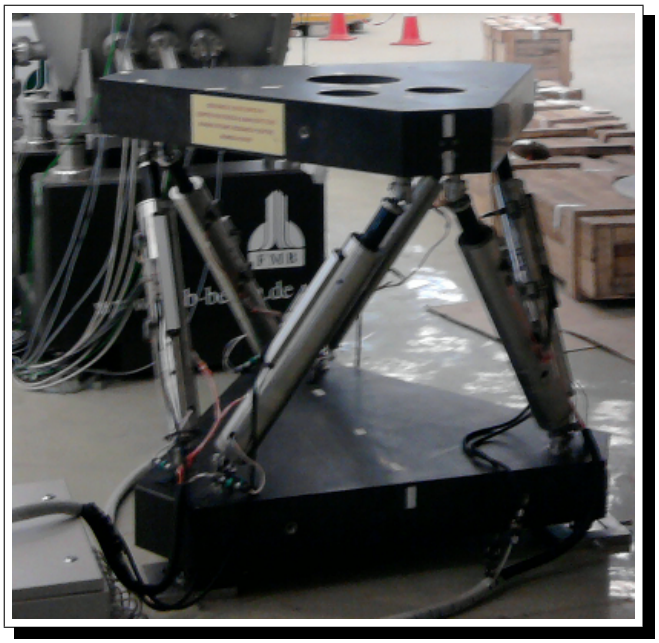

Figure 3.7: Hexapod installed at PX beamline

be selected from the synchrotron spectrum using the installed Double Crystal Monochromator (DCM) at PX Beamline, Indus-2. However, the DCM installed at BL-21 works in isolation, *i.* e. energy from the DCM can only be chosen manually through the software. The precise measurement of absorption edge of heavy element is typically done through the fluorescence measurement by counting the fluorescent photons over the wavelength range in the neighborhood of the absorption edge of the heavy atom. For this, X-ray energy is varied in small steps, typically $\sim 1\ eV$. This energy scanning step needs to be automated as one has to perform about 100 measurements by changing the X-ray energy using DCM and also manual data taking may introduce error in the measurements.

Two softwares have been developed: 1) DCM energy calibration software and 2) Absorption edge scanning software using fluorescence. These software have been tested extensively and are in use at the PX beamline of Indus-2. The layout for connection between the different hardware and software programs is shown in Fig. 3.8.

3.1.4.1 Development of automated DCM energy calibration software

Measurement of absorption edge of a known element is required to calibrate the DCM energy. The initial calibration of the DCM changes due to the change in the incident beam angle at the first crystal of DCM, thus DCM energy calibration is important before taking actual data. This process has been made automated by developing a software

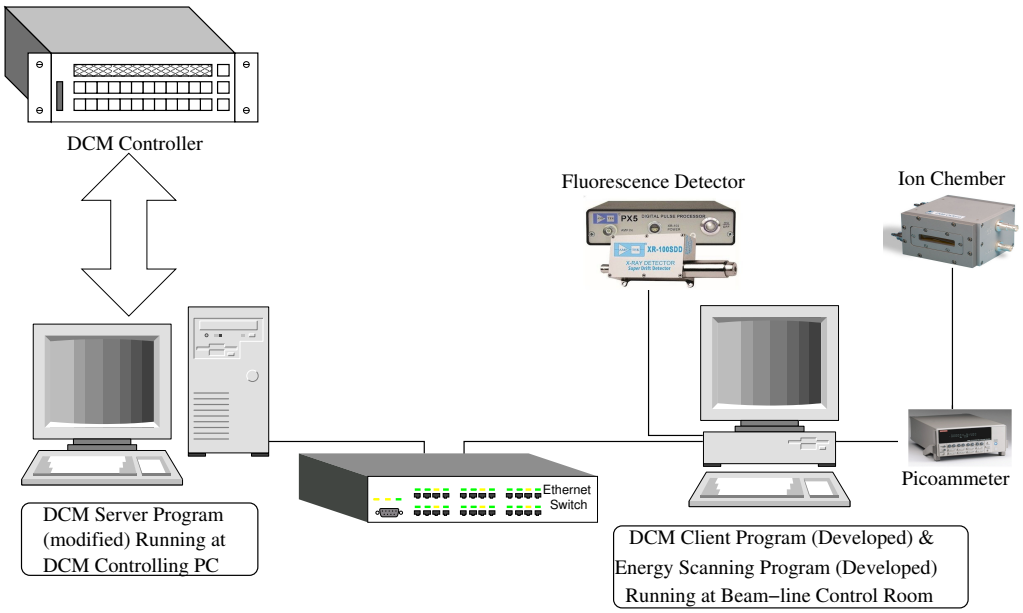

Figure 3.8: Showing various hardware components and their connectivity for the energy scanning experiment. The DCM software (written in VEE Pro 8.5 by FMB, Berlin) has been modified to work as server program. This server program drives the DCM components according to the chosen energy. The client computer kept at the beamline control room is directly connected to the ion-chamber through picoammeter and fluorescence detector via controller.

shown in Fig. 3.9.

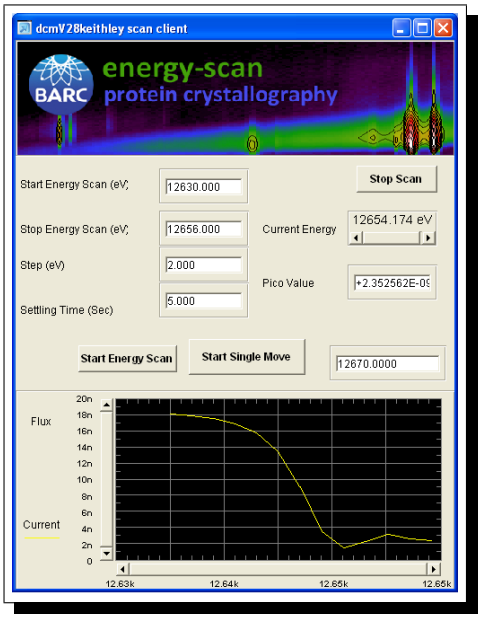

Figure 3.9: Graphical User Interface (GUI) of the absorption edge scanning software. Foils of known elements (Selenium in the present case) are exposed to X-rays and intensity of X-rays transmitted through the foil is measured using an ionization chamber and a picoammeter. Picoammeter reading is displayed with the X-ray energy.

The existing DCM software, written (FMB, Berlin) in VEE Pro 8.5 language does not have any provision to communicate through external program or remote computer.

Therefore it has been modified by us, so as it works as server program and accept command from remote computer. A client program (Fig. 3.9) has been developed in VEE Pro 8.5 language. The client can communicate with the modified DCM server program through the $TO - FROM$ socket method of VEE PRO-8.5. The $TO - FROM$ socket method is a server-client based protocol over Ethernet. To use the existing DCM program as server, the $TO - FROM$ socket has been added in the DCM software. Henceforth, this will be called as "DCM server" program. While modifying the DCM server program particular emphasis was given in maintaining the existing structure of the program so that its functionality, safety features and error handling capability does not get compromised. DCM server program can perform three different actions based on the request string it receives from DCM client program:

- Moves the DCM to given energy value as requested by client
- Sends status report to the client on completion of job
- Reads the current DCM energy and pass it on to the client

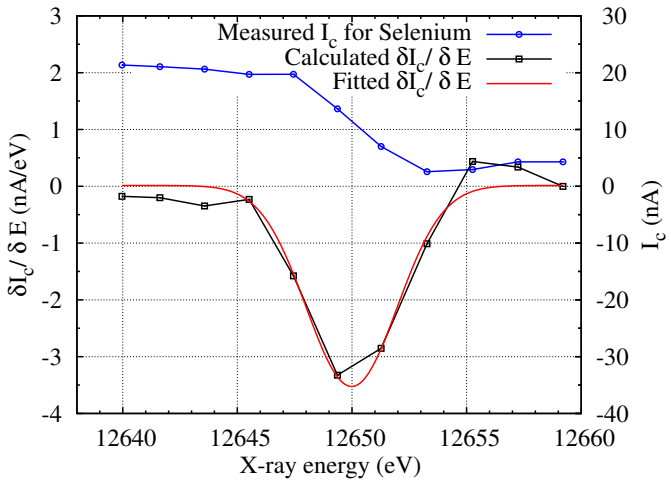

Figure 3.10: The ionization current (I_c) measured by the ion-chamber (for a given voltage, $180\ V$) is shown. To locate the inflection point precisely, I_c has been differentiated with the X-ray energy (E).The differentiated data is then fitted with a Gaussian and the peak of the Gaussian ($12650\ eV$) precisely gives the inflection point or the absorption edge energy.

The client program, on the other hand, takes four inputs through the GUI shown in Fig. 3.9. They are the start and end energies, energy step size for movement and the settling time. The client first sends request to the server to move DCM to the starting

energy. DCM server then sends back the job completion flag to the client. The ionization current reading (I_c), measured by the picoammeter, is stored against the current energy. The starting energy is then incremented with the step size and the client again sends request to DCM server. This is continued until the end energy is reached. After each energy movement, the server waits for some time (settling time) to allow for DCM energy stabilization. Following this procedure, the ionization current (I_c) as a function of X-ray energy (E) can be measured. The client program also has provision of real-time plotting of the ionization current (I_c) with X-ray energy.The absorption edge of a given foil can be found out from the inflection point as explained in the Fig. 3.10.

The Selenium K-edge energy is 12658 eV, however the measured value from DCM is coming at 12650 eV. So, the DCM needs to be corrected for this offset in energy by driving the DCM to 12650 eV for measurements at Selenium edge.

3.1.4.2 Absorption edge scanning software using fluorescence

Anomalous scattering changes the atomic scattering factor for the atoms in question. The atomic scattering factor - typically called $f(S)$ - is a real quantity that falls off in a Gaussian type way with resolution ($i.e$ $|S|$) and is usually wavelength independent. However in the case of anomalous scattering it is changed to: $f(\lambda, S) = f_0(S) + f'(\lambda) + if''(\lambda)$. Where $f_0(S)$ is the usual energy independent value (i.e. away from the absorption edge). Note that the f' and f'' components do not fall off with resolution unlike the atomic scattering factor. The symbol S refers to the diffraction vector whose length is characteristic of resolution (actually $1/d$). A fluorescence scan is required to verify the presence and location of the absorption edge for the crystal under consideration. f'' can be measured using fluorescence scan and f' can be calculated from f'' using the Kramers-Kronig equation.

For fluorescence or MAD scan, the DCM is moved in very fine steps (typically 1 eV) and the total counts in the selected range of channels (of fluorescence detector) at the fluorescence peak region ($i.e$ area under the fluorescence peak) are measured for a given time (typically few seconds) and for each energy steps. This process has been automated by developing a server-client based software written in Microsoft Visual Basic 6.0 (VB-6.0) and VEE Pro 8.5. The GUI of the existing fluorescence detector software has been extended to add all the required features as shown in Fig. 3.11. This program can be used for both, the DCM energy calibration and fluorescence scan, by choosing an appropriate option.

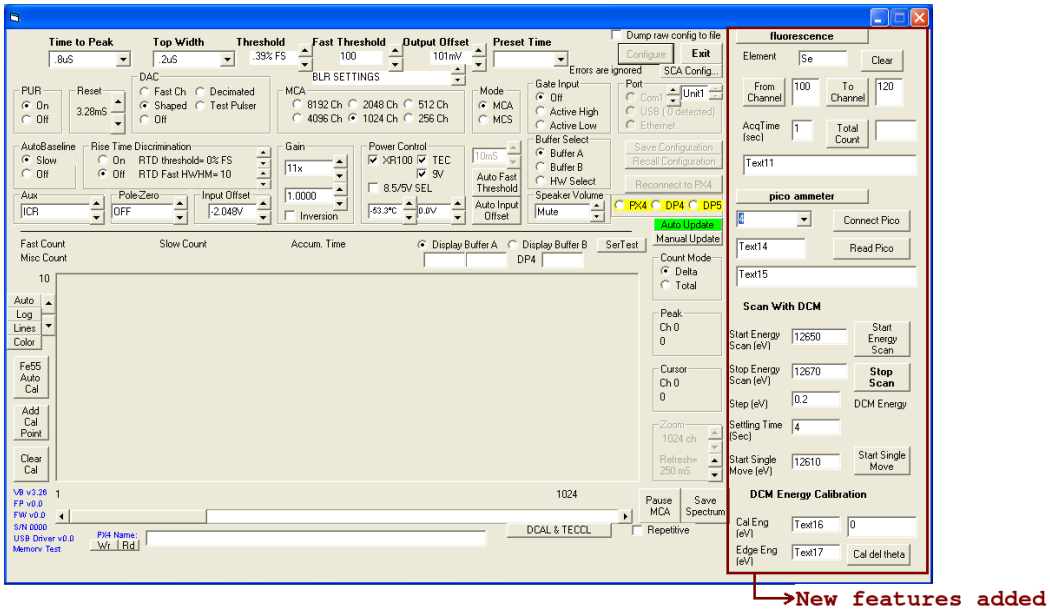

Figure 3.11: GUI of the fluorescence and absorption edge scanning software developed in VB.

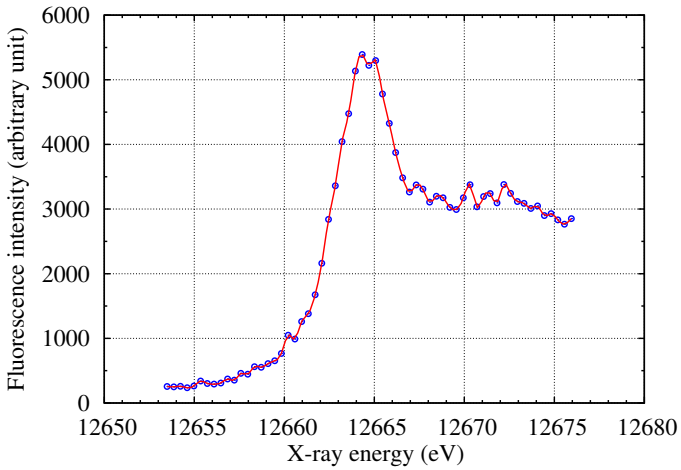

Figure 3.12: Fluorescence scan of Se-element in a selenomethionine protein sample.

This software has three parts: 1) energy scanning program (written in VB-6.0), shown in Fig. 3.11, 2) client program for DCM (written in VEE Pro 8.5) and 3) DCM server program. The DCM client program can communicate with both the DCM server program (via $TO-FROM$ socket) and the scanning program (via callable server ActiveX). Hence the DCM client program works as a bridge program. The working of DCM client-server program is similar to that as explained earlier. The DCM client passes the value of current energy to the scanning program. The scanning software program communicates with the fluorescence detector (Amptek make) through the USB interface and picoammeter

(Keithley make) via RS-232 serial port. Fluorescence detector counts, over a selected range, are recorded at the current energy for a chosen acquisition time. DCM energy is then updated with the given step size and the scanning program communicates this to the DCM server via DCM client. The fluorescence data recorded during absorption edge scanning of a selenomethionine protein sample is shown in Fig. 3.12. The anomalous scattering factors are calculated by using the program CHOOCH [75] and shown in Fig. 3.13.

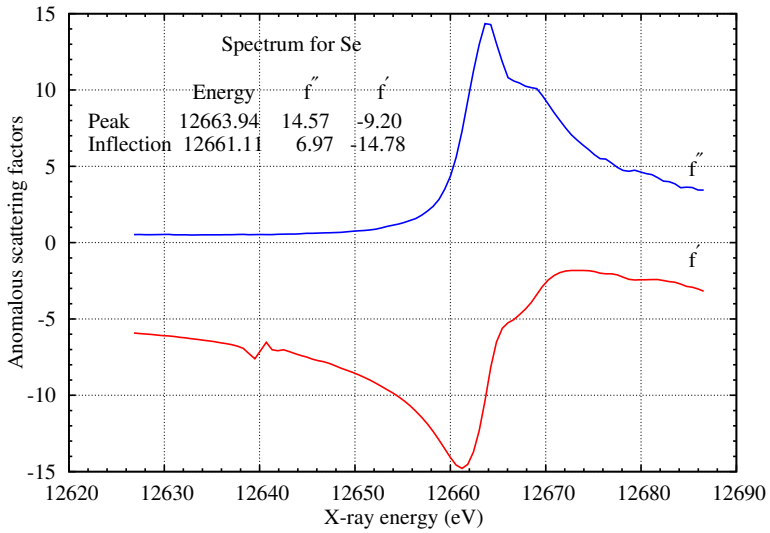

Figure 3.13: Variation of the normal scattering (f') and the anomalous scattering (f'') of selenium as a function of energy of incident X-ray photons, going through the K absorption edge. The quantity plotted as (f') is the loss of normal scattering power. f' and f'' have been calculated using the program CHOOCH [75] with the fluorescence data (shown in Fig. 3.12) recorded at BL-21 using the developed software being reported in this article. The three energies for MAD data collection should be chosen by looking at this plot: 1) Peak wavelength is maximum of f'' (top of the fluorescence scan), 2) High energy remote wavelength is $150-250\ eV$ higher than the peak wavelength and 3) Inflection wavelength is half-way down the f'' absorption edge ($i.e.$ the inflection point of the curve). As can be seen from this plot that the inflection point is only about $2\ eV$ down the peak, the DCM should have the required energy resolution, $\Delta E/E \sim 10^{-4}$.

The DCM scan with picoammeter can be used for the calibration of DCM energy. The calibration is done by measuring the ionization current as explained earlier (see Fig. 3.9). The angular offset in the DCM can be calculated using the relation

$$\Delta\theta = \theta' - \theta = sin^{-1}(\lambda'/2d) - sin^{-1}(\lambda/2d).$$

Where, λ is the actual absorption edge wavelength and λ' is the measured value corre-

sponding to the energy E' ($\lambda'(\text{Å}) = 12398.4/E'$ (eV)). For our DCM crystals, Si(111), the value of $2d$ is 6.271142Å. This angular offset should be applied for each DCM energy value during fluorescence scanning. The client program applies this angular offset automatically to each energy step during the fluorescence scanning.

3.1.5 Beam line Shielding

There are several sources of radiation at synchrotron beamlines, which can be hazardous to user and beamline workers. Therefore, the proper shielding at beamline is necessary requirement for optimum utilization of synchrotron facility. The possible sources of radiations are

- Synchrotron radiation
- Primary bremsstrahlung radiation
- Secondary bremsstrahlung radiation

Primary bremsstrahlung radiation is generated by the interaction of electrons with the residual gas, inside the storage ring. The front end of beamline interfaces with the storage ring via a thick shutter. Both, the synchrotron and the primary bremsstrahlung radiations penetrate the experimental hutch when the shutter is open. Subsequently, primary bremsstrahlung radiations fall on optical elements inside the hutch, which generates the secondary bremsstrahlung radiation. The dose generated at a synchrotron beamline is a cumulative effect of all of these three radiations viz. synchrotron, primary bremsstrahlung and secondary bremsstrahlung.

D Lahiri *et.al* [76] has done the detail shielding calculations for Indus-2 synchrotron beamlines. Also shielding calculation has been done specifically for the PX-BL21 [77]. Based on these calculations, the beamline shielding proposed as per the Table. 3.2.

Table 3.2: Proposed shielding for PX beamline

Wall	**Optics Hutch**	**Experimental Hutch**
Front Wall	3 mm Pb	3 mm Pb
Side Wall	2 mm Pb	2 mm Pb
Roof	1 mm Al	-
Direct gas bremssrahlung stop	10 cm Pb(50 cm $\times$ 100 cm)	15 cm Pb(100 cm $\times$ 100 cm)

3.2 Beamline Installation

Beamline installation has been carried out in parts. First, the DCM was installed at 18.5 m from tangent point of bending magnet source and tested for vacuum. Subsequently, both mirrors were installed in beamline. The collimating mirror (CM) is placed at 16.5 m and toroidal mirror at 24 m from the source. During installation each beamline component were checked for leaks before connecting to other components. This had simplified the leak testing process for whole beamline. Shielding hutch installation was deferred to last, to allow easy access from both side during beamline component installation.

3.2.1 Monochromator installation

Double crystal monochromator (DCM) has been supplied by M/s FMB-Oxford GmbH as per our specifications. The DCM is equipped with a high-resolution rotation axis (Type 420 UHV, Huber GmbH) mechanism and a high-resolution rotary encoder (RON 905 UHV, Heidenhain GmbH) for sensing Bragg angle. The main rotation axis have Bragg angle range from -3° *to* 50° with resolution and repeatability of 0.1 $acrsec$ each. The DCM has two sets of crystals ($Si(111)$ *and* $Si(220)$) which can be exchanged in vacuum, in a few seconds. High resolution translational and rotational motions of these crystals are achieved by precision stepper motors (Phytron GmbH), Pico motors (NFO-8321-V) and piezo crystals (PA 100, Piezosysteme Jena) with controllers drivers and encoders (Renishaw plc). These translational stages have resolution of 0.25 $\mu m/half\ step$ with repeatability of 1 μm. The first crystal (white beam facing) and white beam mask are water cooled. Installation has been accomplished by FMB-Oxford's engineers and scientific staff from BARC and RRCAT. The DCM installation has been performed with following course of action:

- Footprint marking at installation site for DCM's mineral cast or granite block,
- Placement of mineral cast and DCM chamber on marked position,
- Chamber's height adjustment and leveling using theodolite at X-ray beam level,
- Electrical wire connection and testing of each motor, encoder and limit switches for motion,
- Placement of first crystal; its height, roll, pitch and yaw adjustment by theodolite and auto collimator,

- Placement of second crystal; its roll, pitch and yaw adjustment by auto collimator,
- UHV and water cooling testing before coupling the system with rest of the beamline components,

3.2.2 Mirror, chamber and hexapod installation

Installation of two X-ray mirror (CM and TM) along with there chambers and support have been accomplished at PX-BL21. The X-ray mirrors with bender and cooling system as per our specifications were supplied by M/s SESO France. Like DCM installation, similar course of actions have been followed. At first step, empty mirror chambers were installed along with hexapod to check the UHV compatibility and movement mechanism of system. In parallel, the X-ray mirror and bender mechanism has been assembled and kept inside the laminar flow to minimize the contamination of reflecting surface from the environment. The CM and TM are shown with benders in Fig. 3.14 and Fig. 3.15, respectively. Then assembled mirrors and bender were placed inside the mirror chamber. Cooling is made through plates of copper plunging in two grooves along the sides of the mirror (groove cooling). Contact between the mirror and the copper plate is done with Gallium bath for groove cooling. These Copper plates are brazed on a tube, inside which water flows. Each of the cooling tube terminates with CF flange for connection with external cooling circuit.

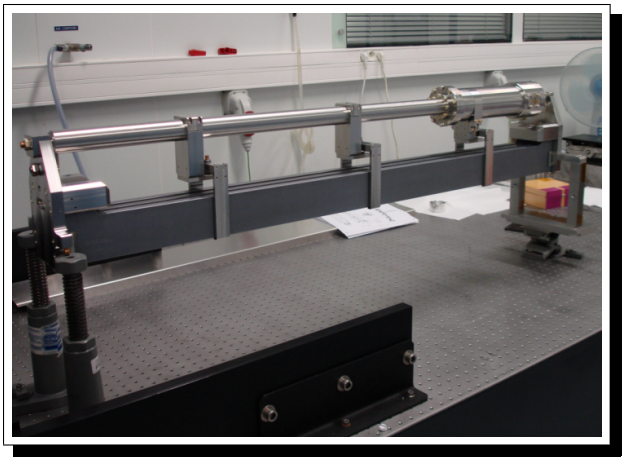

Figure 3.14: Collimating Mirror with bender

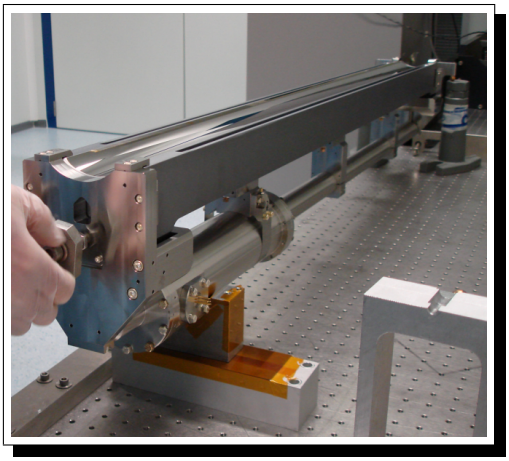

Figure 3.15: Toroidal Mirror with bender

Both X-ray mirrors have $Rh(\sim 500\ \mathring{A})$ coating on surface with chromium binding. X-ray mirrors are 1.2 m long and UHV compatible installed along with cooling and bender

system for optimum operation of beamline. Detail specifications of CM and TM, are presented in Table. 3.3. Both mirrors are equipped with "U-bender" [78] with gravity compensation to optimize the spot shape and size at sample position.

Table 3.3: X-ray mirrors specification

	Collimating Mirror (CM)	**Toroidal Mirror (TM)**
Length	1.2 m	1.2 m
Width	80 mm	80 mm
Useful width	60 mm	60 mm
Thickness	60 mm	60 mm
Shape	Cylindrical (with bender)	Toroidal (with meridional bender)
Meridional Radii	$9000 < R(m) < 15000$	$7000 < R(m) < 9000$
Sagittal Radius	∞	$\rho = 48.5(\pm 0.05)mm$
Position	Downward	Upward
Material	Silicon (Si)	Silicon (Si)
Coating	$Rh(\sim 500\ \mathring{A})$	$Rh(\sim 500\ \mathring{A})$
Roughness	$< 3\ \mathring{A}\ rms$	$< 3\ \mathring{A}\ rms$
Slope errors - Tangential	$< 2.5\ \mu rad\ rms$	$< 2.5\ \mu rad\ rms$
Slope errors- Sagittal	$< 25\ \mu rad\ rms$	$< 25\ \mu rad\ rms$
Cooling	Groove Cooling	Groove Cooling

3.2.3 Installation of rest of the beamline components

Fig. 3.16 shows the mechanical lay out of beamline. Beamline starts at $\sim 14\ m$ from the source, connected to the fron-end via 250 μm beryllium (Be) window. Primary slit is first component of beamline connected to Be-window through 100 ID SS304L bellow. The primary slit is followed by the the BPM and the CM. Another BPM is sandwiched between the DCM and the CM. The CM is at 16.5 m, while DCM is placed at 18.5 m from source. A pair of a BPM and a water cooled compact slit is placed before and after the TM mirror. The TM mirror is located at 24 m from source. Therefore, each optical component is preceded and followed by a BPM.

Five numbers of UHV gate-valves were installed at PX beamline for vacuum isolation of the beamline components. One gate-valve, on each mirror chamber (CM and TM), at beam-in and beam-out port, and one after the Double crystal monochromator (DCM) has

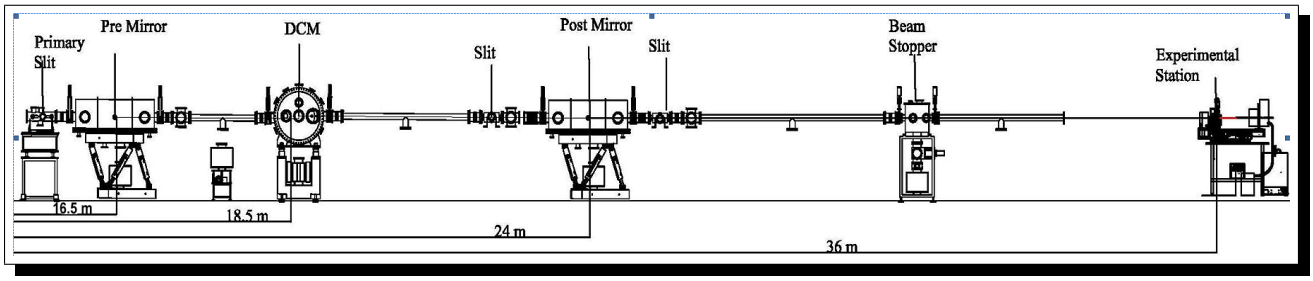

Figure 3.16: Mechanical layout of PX beamline

been installed. Compressed air supply system were laid using polyurethane (PU-8) pipes and push-fit connectors, which is utilized for operation of pneumatics, for opening and closing of gate-valves. Similarly, a water supply system were laid down by utilizing PU-8, PU-10 pipes and connectors. Water supply network were built for cooling down the Be-window, CM, TM, DCM, X-ray slits and BPMs. CAT-6 and Cat-5, networking cables were laid down along the length of beamline to network the controllers and computers, connected to beamline components. The UHV compatibility is an essential requirement for beamline operation, therefore, each and every component were tested for vacuum leak individually, and in connection with each other. Liquid acetone is used for gross leak detection, while helium gas has been used for fine leak detection.

The experimental station is installed at 36 m from the BM source, which is at focal point of the toroidal mirror. The experimental station (mardtb) has been supplied by the M/s Marresearch GmbH. The mardtb is a state-of-art system that can handle delicate and temperature sensitive protein crystals. The experimental station shown in Fig. 3.17, comprises of a cooled MX225 CCD detector (Rayonix USA) with detection area of 225 $mm \times 225$ mm and pixel size of 72 μm (2×2 $binning$), for fast (~ 1 sec) data acquisition. To avoid radiation damage in the protein samples, a $100K$ LN_2 cryo-streamer (marcsc, marresearch GmbH) has been coupled with the experimental station. Automatic sample loading/unloading is also possible by the attached robotic sample changer (marcsc, marresearch GmbH). A fluorescence detector (Amptek, USA) has been integrated with the station for recording fluorescence signal from heavy atoms in crystals, which can enable anomalous diffraction experiments. The available sample to detector length is 50 mm to 650 mm. By default single axis goniometer of mardtb take the spine standard pin.

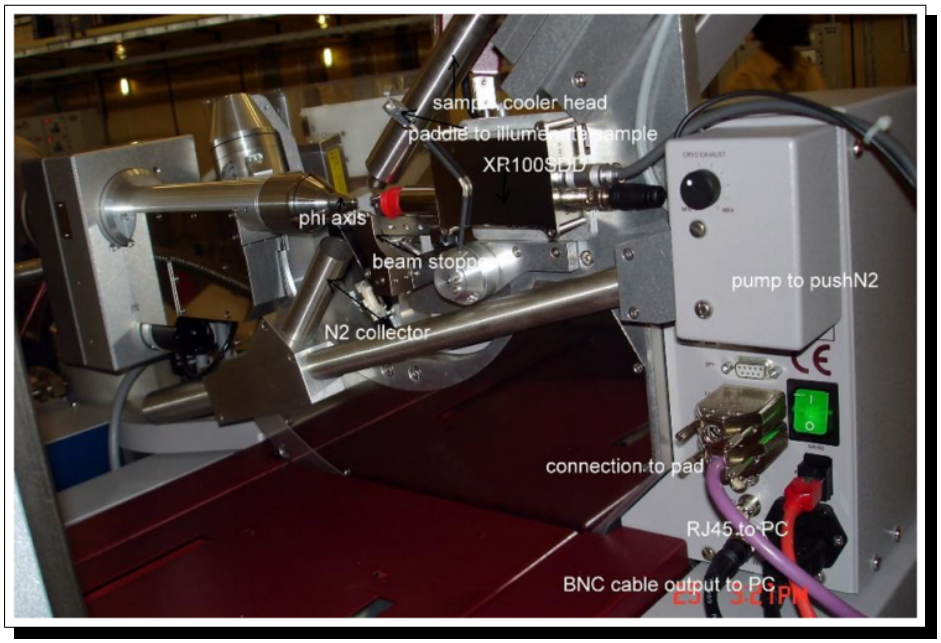

Figure 3.17: Experimental Station

3.2.4 Shielding hutch installation

Based on the proposed specifications (Table. 3.2), a shielding hutch has been designed to encapsulate the PX beamline. The shielding hutch has been divided in to 4 sections as shown in Fig. 3.18. A optics hutch-1 encloses the primary slit, CM and DCM. A optics hutch-2 contain TM and BPM's. The user sitting area is marked as a computer room in schematic. The hutch is a freestanding, self-supported structure fixed to the floor. The assembled sections (made out of CRCA steel sheets) create a structural system capable of supporting the hutch panels and attach utilities.

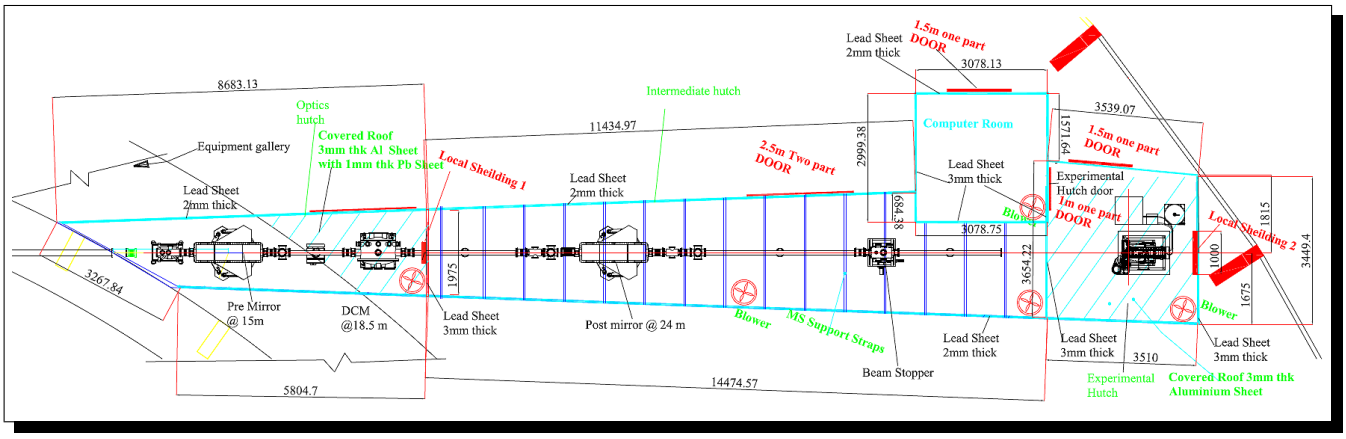

Figure 3.18: Schematic of shielding hutch

These panels has been manufactured as a lead sheet sandwich between two CRCA steel sheets. Side walls have 2 mm Pb sheets while walls perpendicular to X-ray beam contains 3 mm Pb sheets. The hutch has internal height of 2.5 m, and the floor area (including wall thickness) is $\sim$ 64 m^2. The hutch includes standard features such as single and double-sliding doors (manually operated) with lead glass windows, and penetrations for the several types of chicanes, fluids, electrical and exhaust ventilation. All doors are manually operated with electromagnetic latch and industrial limit switches. Two local shielding has been installed to further protect against bremsstrahlung scattering. As shown in the drawing Fig. 3.18, first local shielding placed inside optics hutch-1 of 1000 $mm \times$ 1000 $mm \times$ 100 mm thickness with center height at 1420 mm. The local shielding-2 with size of 1000 $mm \times$ 1000 $mm \times$ 100 mm and center height at 1375 mm, is placed after experimental station.

3.3 Beamline Commissioning

3.3.1 Beamline alignment and optimization

The beamline alignment and a spot optimization are crucial for maximizing the flux at a sample position. Four wire-cum-fluorescence BPM are used for the alignment of optical components. The indigenously developed BPM have two components: 1) a tungsten wire to measure photocurrent ($\sim$ nA range) and 2) a phosphor coated water cooled copper plate for beam viewing, which can also act as a beam stopper. The intensity profile of the X-ray beam is calculated from the measured photocurrent with the stepper motor position (in the vertical direction), as shown in Fig. 3.19. The X-ray beam is visualized (Fig. 3.20), through the fluorescence from the phosphor coated plate, by using a CCTV camera. Thus, the combination of wire current and fluorescence, provide both qualitative and quantitative information of the X-ray beam. Both CM and TM, were aligned using the respective downstream BPM. Both the mirrors are dynamically bent for matching source and focal point. The bending curvature and incident angle of TM are used to achieve optimal focus size at sample. Also, TM is placed in $2:1$ configuration, which minimizes the chromatic aberrations (sagittal and tangential). The combination of Be-window, which act as low pass filter, and the Rh coated mirrors that act as high pass filter, gives the beamline working range of $5 - 20$ keV. This energy range covers the most of the heavy element absorption edges. The measured spot size at sample position is 0.5 $mm(V) \times$ 0.5 $mm(H)$, which is close to the calculated spot size of 0.4 $mm(V) \times$ 0.4 $mm(H)$. The flux, mea-

sured by a calibrated photo-diode, at sample position with 0.4 $mm \times$ 0.4 mm slit was 6.4 $\times$ 10^9 ph/sec for 80 mA storage ring current at 2.5 GeV.

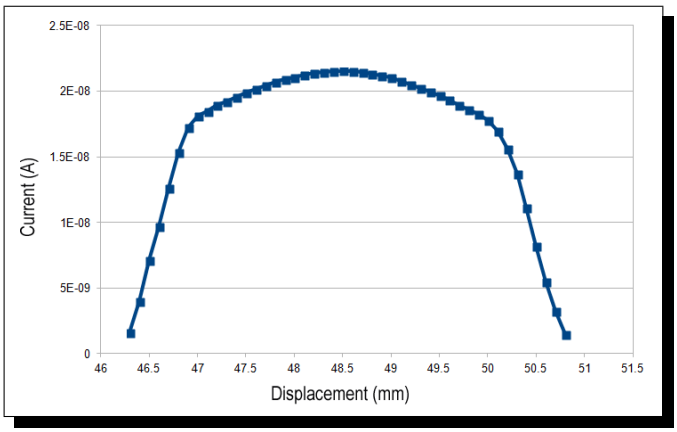

Figure 3.19: X-ray beam profile measured from BPM2 after CM

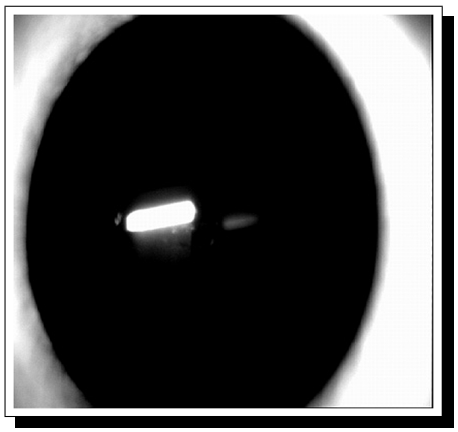

Figure 3.20: The view of white X-ray beam captured by BPM

The alignment of PX-BL21 has been carried out in steps. First, the X-ray beam position and its direction were determined at Be window, to align front-end mask and Be-window. Then, primary slit, BPM-1, collimating mirror, BPM-2 and DCM were connected with FE in optics hutch-1. The X-ray beam was scanned through BPM-1 wire to insure that only central Gaussian part of beam was coming through primary slit. The 2 mm vertical central peak of X-ray beam was selected from the primary slit. The BPM-1 and the BPM-2 used for alignment of collimating mirror in X-ray beam. The alignment of X-ray mirrors were commenced through following steps:

- X-ray beam profile is measured using BPM-1 and central portion of beam is selected by primary slit,
- Mirror is removed completely from the beam path and beam is allowed to fall on BPM-2,
- X-ray beam is roughly half cut by mirror,
- Mirror is rocked around the glancing angle to determine the extreme angle, at which X-ray beam is either completely cut, or just fully reflected,
- Average of the determined extreme angles makes the mirror parallel to X-ray beam,
- Parallel mirror is moved vertically to cut X-ray beam at half,

- Required glancing angle is set in mirror using hexapod, which reflect the X-ray beam to downstream beamline component,

The aligned beam coming through collimating mirror was allowed to fall on first crystal of DCM. The beam intensity was optimized by movement of second crystal using piezo motor. Once beam came outside the DCM in optic hutch-1, the rest of beamline component connected with optics-hutch-1 components. Similar procedure was used for toroidal mirror alignment with horizontal and vertical focusing at sample position. Before any data collection, the DCM is calibrated by scanning the absorption edge of standard metal foil.

3.3.2 First data collection

Initial X-ray diffraction data were recorded and indexed from lysozyme crystals (Fig. 3.21) and some organic crystals. Lysozyme data collected, at 12.0 KeV X-ray energy, at 40 mA beam current for 60 sec per frame. The crystals were diffracted up to 1.8 $\mathring{A}$ with $I/\sigma =$ 24.8 and $R_{merge} =$ 0.052. The Diffraction data collection statistics are shown in Table. 3.4. Subsequently, fine tuning and spot optimization reduced the data collection time to 10 sec per frame with improved statistics.

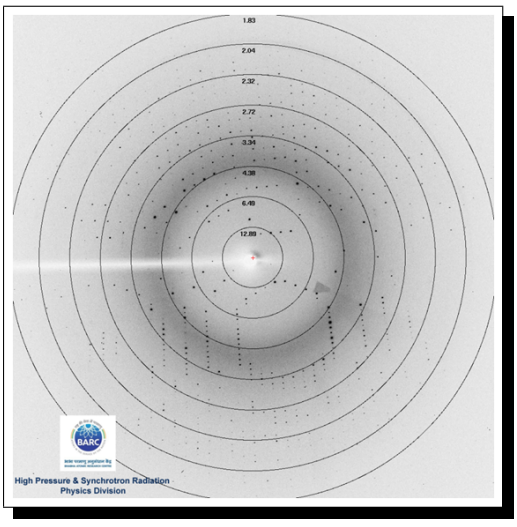

Figure 3.21: Lysozyme X-ray diffraction image

Table 3.4: Data collection statistics of Lysozyme

Crystal	**Hen egg lysozyme**
Data collection statistics	
Space group	$P4_32_12$
a, b, c (Å)	81.69, 81.69, 39.48
α, β, γ (°)	90, 90, 90
Resolution (Å)	23.31-1.83 (1.931.83)
Reflections measured	86874 (12277)
Unique reflections	12111 (1708)
Completeness (%)	99.1 (98.4)
Average I/σ	24.8 (11.9)
R_{merge}	0.052 (0.146)
Multiplicity	7.2 (7.2)

3.3.3 Anomalous data collection on PepQdr protein

Anomalous diffraction data has been collected on PepQdr crystals and the structure was determined using single-wavelength anomalous diffraction (SAD) method using Selenium K-edge. The PepQdr (monomeric mass $\sim$ 37 kDa) is a Xaa-Pro dipeptidase (PepQdr) from *Denociccous radiodurans* (GenBank accession, NP_{294970}) and consists of 349 residues. The six sulphur atoms of the native -methionine residues in the protein were replaced by selenium-methionine using recombinant DNA technology at the beamline associated biochemical laboratory. The selenomethionyl labelled protein crystals (size appx 200 μm) were cryo-protected and flash frozen into liquid nitrogen before exposing to X-rays at the PX beamline. The diffraction data were collected at 12666 eV energyand the ring current of 80-90 mA st 2.5 GeV and slit size 0.3 $mm \times 0.3$ mm. The other diffraction experiment parameters were as follows: Temperature: 100 K; Exposure: approx 10 sec (dose mode of MARdtb); $\delta\phi$: 1°; distance: 200 mm; total ϕ collected: 90°. The data were processed using XDS [79] and scaled using AIMLESS [80] in CCP4 suite [81], data collection statistics is presented in Table. 3.5. The hkl2map/ShelX [82] [83] was used for SAD phasing using this dataset. The $< d''/\sigma F >$ plot from ShelxC is shown in Fig. 3.22. The value of $< d''/\sigma F >$ greater than 1.3 suggests significant anomalous signal in the dataset. The selenium atoms (six no.) and their occupancies were determined using ShelXD [84] (Fig. 3.23) with CCall > 45%.

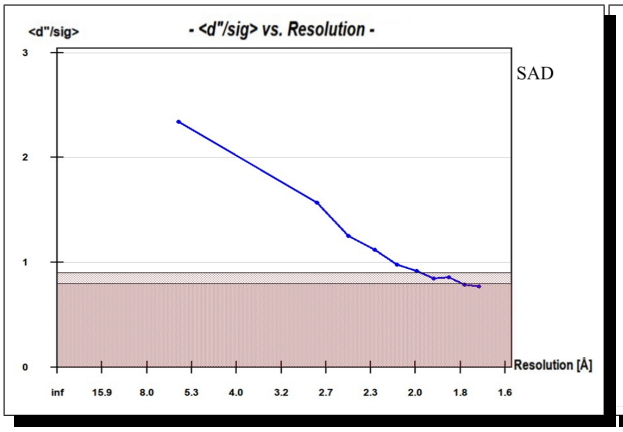

Figure 3.22: Anomalous signal $< d''/\sigma F >$ of the PepQdr

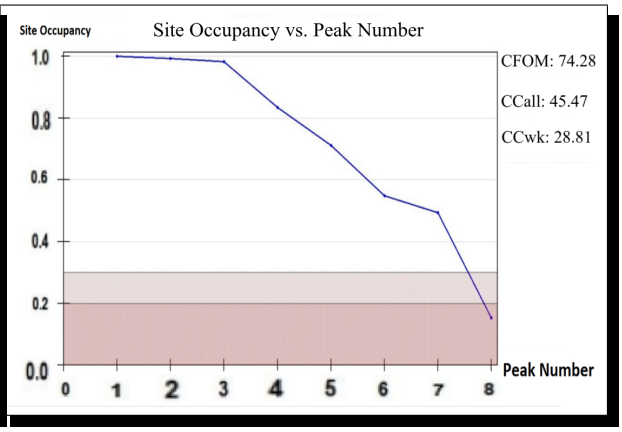

Figure 3.23: Selenium site and their occupancy in PepQdr

The initial map was significantly improved by density modification using ShelXE [85] suite. The map clearly distinguished between the correct and the wrong enantiomer. The residues were auto-build into density modified map using ARP/aWRP [86] Classic software. The structure was refined using PHENIX [87] refinement with intermittent

manual building using COOT [88] suites. The structure was refined to R_{work}/R_{free} value of 19.6/23.2%. with good stereochemistry as shown in Table. 3.6. The experiment was subsequently reproduced on the another crystal of the same protein.

Table 3.5: Data collection statistics of PepQdr (Se-SAD)

Crystal	**PepQdr (Se-SAD)**
Data collection statistics	
Space group	$P4_32_12$
a, b, c (Å)	60.64, 60.64, 202.8
α, β, γ (°)	90, 90, 90
Resolution (Å)	45.14-1.80 (1.841.83)
Reflections measured	238913
Unique reflections	35744
Completeness (%)	98.6 (85.5)
AverageI/σ	16.4 (1.9)
R_{merge}	0.087 (0.692)
Multiplicity	6.7 (4.4)
Mid-Slope of Anom Normal Probability: 1.301	

Table 3.6: Data Refinement statistics of PepQdr (Se-SAD)

Crystal	**PepQdr (Se-SAD)**
Refinement statistics	
Resolution (Å)	33.7-1.8
Total non-H atoms	2941
Protein	2593
Water	348
R_{cryst}	0.1959
R_{free}	0.2319
Residues in favored regions of the Ramachandran plot (%)	97.39
Residues in allowed regions of the Ramachandran plot (%)	2.61
Rmsd ideal bond length ($\AA$)	0.010
Rmsd ideal bond angles (°)	1.320
Average B factor	23.32

3.4 Summary

In summary, the protein crystallography beamline has been installed and commissioned at 5° port of 1.5T BM source at Indus-2, RRCAT Indore. Many beamline components were designed and fabricated locally. These beamline components are Beam position monitors (4 No.s), one primary X-ray slit, two no.s of water cooled compact slits, X-ray pipes, Mirrors chambers (2 no.s) and Hexapods (2 no.s). The double crystal monochromator and X-ray mirrors were procured based on the finalized design of PX beamline. Each component of PX-BL21 has been tested thoroughly for the UHV compatibility. Shielding hutch specifications were finalized and assembly drawings has been prepared based on

radiation dose calculations. Finally, the experimental station (mardtb) and the shielding hutch has been installed at beamline.

Hot alignment of each optical components in X-rays is carried out by using BPMs. Both mirrors are aligned in the X-ray with micron accuracy and operated at glancing angle of $3 - 3.5$ *mrad*. Software for, calibration of DCM energy and absorption edge scanning using fluorescence detector, has been developed and deployed at the beamline. Networking, communication and functioning of different hardware component along with the software has been developed and tested successfully. The DCM is now routinely calibrated by using standard metal foils. First data were collected on lysozyme crystals with good statistics, which display the viability of beamline. Anomalous diffraction of a protein sample has been recorded and analyzed.

Chapter 4

Xaa-Pro dipeptidase from *Xanthomonas campestris*

Peptidases are enzymes that hydrolyze a peptide bond of a protein/peptide substrate. But only a few peptidases hydrolyze a peptide bond between the caboxy terminal of an aminoacid and the amino terminal of a proline residue due to the cyclic structure of prolines. Xaa-Pro dipeptidase (XPD, EC 3.4.13.9) specifically cleaves a trans Xaa-Pro peptide bond (Xaa is any amino acid) in a dipeptide substrate with a prolyl residue at carboxy terminus (MEROPS database, [89]). The enzyme is also referred as a proline dipeptidase, or a prolidase, or a peptidase-Q (PepQ). XPD belongs to M24B family of peptidase, which also includes Aminopeptidase-P (APPro, EC 3.4.11.9), capable of hydrolyzing a trans Xaa-Pro peptide bond at the N-terminus of polypeptide. The prolidase is ubiquitous in nature and has been isolated from mammals, bacteria and archaea [2]. In archaea and bacteria, although the function of prolidase is not well elucidated, it has been suggested to be involved in the recycling of proline [90]. On the other hand, in humans, the role of prolidase is well established in different patho physiological conditions. For instance, the enzyme is shown to be involved in the final stages of the degradation of dietary as well as endogenous proteins particularly collagen. Moreover, in humans, mutations in XPD gene causes a rare autosomal recessive disorder, prolidase deficiency (PD), which exhibits phenotypes like skin ulceration, mental retardation and recurrent infections ([4] and [91]). The enzyme is also important commercially in the food and dairy industries for improving flavor and texture of the food [2]. For example, during cheese ripening, action of this enzyme on proline dipeptides leads to reduction in the bitterness of the product due to release of proline residues. In addition to the dipeptidase activity, XPD and other peptidases of M24B family display fortuitous activity against toxic

organophosphorus (OP) compounds that include pesticides and nerve agents, by cleaving the P-F and P-O bonds [5].

The genome analysis of *Xanthomonas campestris* showed the presence of three different genes, encoding peptidases of M24B family; two are annotated as XPD while third one as APPro. The size of the polypeptide chain encoded by two XPD genes are 441 (48 kDa; XPD48; GenBank accession, NP-638603) and 399 (42.8 kDa; XPD43; GenBank accession, NP-637763) residues long, and they share $\sim 24\%$ sequence identity. It was further analyzed that the individual orhologs of all the three genes are also present in other *Xanthomonas spp.* as well as in most of the members of gamma proteobacteria, which includes many pathogenic bacteria. A detailed structural and functional studies of a XPD48 ortholog has been reported from a gamma-proteobacteria, *Alteromonas sp.* strain *JD*6.5 (PDB entry, 3L24; [92]) and *E. Coli* (PDB entry, 4QR8), while functionally, its orthologs have been characterized from *Escherichia coli* and *Xanthomonas maltophilia* ([93] and [94]). On the other hand, structural and functional characterization of XPD43 (PepQxc) orthologs have only been reported from the two archaea, *Pyrococcus furiosus* and *Pyrococcus horikoshii* (PDB entry, 1PV9, 2HOW; [95] and [96]). Although, the structures of three XPD43 orthologs from eubacteria like *Mycobacterium ulcerans*, *Bacillus anthracis* and *Thermotoga maritima*, (PDB entries, 4EGE, 3Q6D, 2ZSG) have been reported by structural genomics consortia, none of them have been functionally characterized so far. Moreover, a detailed information on both structural and functional aspects of XPD43 like proteins from entire proteobacteria phylum is lacking.

The bioinformatics analysis showed that the PepQxc from *Xanthomonas spp*, are quite unique among M24B family of peptidases, where the strictly conserved glycine and tyrosine residues (*Gly385*, *Tyr387*; numbering as per *Escherichia coli* Aminpeptidase-P) are changed to methionine and valine respectively (Fig. 4.1). The *Tyr387* in APPro of *Escherichia coli* has been proposed to be involved in the conserved hydrogen bond network in *Asp260–Arg404–Tyr387* motif, which shuttles a proton from bulk solvent to the leaving peptide [6] during peptide hydrolysis. Moreover, the *Tyr387*, the *Arg404* and the *His350* might be important residues for the proline specificity ([97] and [98]). The site-directed mutagenesis of *Tyr387* and *Arg404* residues resulted in several fold reduction in the aminopeptidases activity [6]. It was, therefore, felt necessary to characterize PepQxc structurally as well as functionally from *Xanthomonas campestris*. The structural analysis of PepQxc would provide new insights into the role of the conserved residues towards catalysis. Further structure-function analysis may shed light on the significance of two isoforms of prolidase in *Xanthomans spp.* and other members of gamma-proteobacteria.

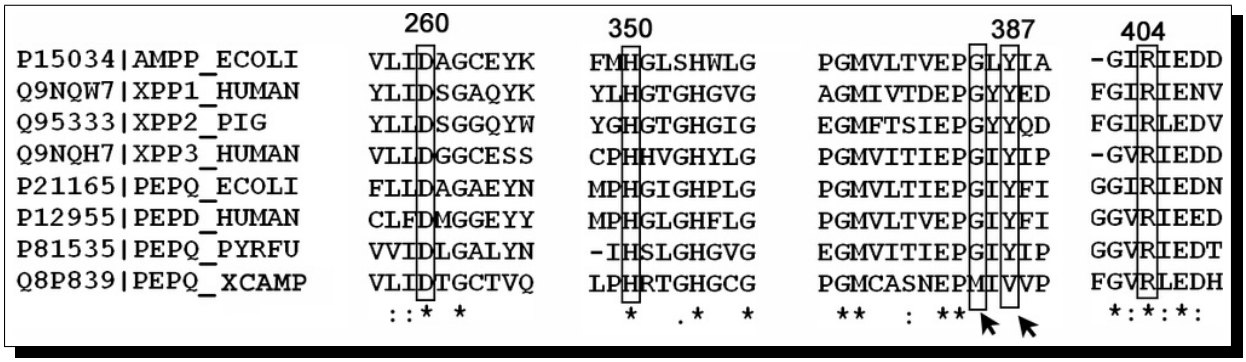

Figure 4.1: Multiple sequence alignment of the peptidases from M24B (MEROPS) family (only relevant blocks of sequences are shown). Uniprot accession numbers are mentioned. The conserved residues important in hydrogen-bonding network and proline specificity are marked in boxes. The substitutions in XPD43 (Uniprot accession Q8P839) are marked by arrows. The numbering on the top row are as per the *Escherichia coli* APPro (Uniprot accession, P15034).

The recombinant PepQxc protein has been cloned, purified and crystallized to study the role of conserved residues in prolidase family. Furthermore, the mutant R372A PepQxc has also been prepared to explore the role of conserved *Arg372* residue in the active site. The native PepQxc protein crystallized in two different crystallization conditions, and henceforth referred as PepQxc1 and PepQxc2 respectively, corresponding to each crystallization condition. Two more variants of PepQxc1 were crystallized by adding Mn and Zn in protein buffer, and will be referred as PepQxc1-Mn and PepQxc1-Zn, respectively. In order to study the substrate binding at active site, the crystals of R372A mutated protein has been soaked with di-peptide substrate (His-Pro) to obtain the complex, and will be referred as PepQxcR372A. Also, the crystal structure of $1mM\ ZnCl_2$ soaked PepQxcR372A has been solved using $Zn-SAD$ method, and henceforth will be referred as a PepQxcR372A-Zn. In this study, the crystal structures of PepQxc1, PepQxc1-Mn, PepQxc1-Zn, PepQxc2, PepQxcR372A and PepQxcR372A-Zn have been solved and refined. The crystallization condition for all four wild type and two mutatant PepQxc proteins have been summarized in Table. 4.1.

Table 4.1: Crystallization condition of four wild type and two mutated PepQxc proteins

Protein (native/mutated)	Protein Buffer	Crystallization Condition
PepQxc1 (native)	20 mM TrisHCl pH 8.0, 200 mM NaCl	40 mM KH_2PO_4, 15% glycerol, 12% (w/v) polyethylene glycol (PEG) 8000
PepQxc1-Mn (native)	20 mM TrisHCl pH 8.0, 200 mM NaCl + 1mM Mn	40 mM KH_2PO_4, 20% glycerol, 14% (w/v) polyethylene glycol (PEG) 8000
PepQxc1-Zn (native)	20 mM TrisHCl pH 8.0, 200 mM NaCl + 1mM Zn	40 mM KH_2PO_4, 20% glycerol, 14% (w/v) polyethylene glycol (PEG) 8000
PepQxc2 (native)	20 mM TrisHCl pH 8.0, 200 mM NaCl	0.1 M Citric acid pH 5.0, 20% $PEG6000$
PepQxcR372A (mutated)	20 mM TrisHCl pH 8.0, 200 mM NaCl + 1mM His-Pro	0.1M Bis-Tris pH 5.5, 0.2M $LiS0_4$ and 25% $PEG3350$
PepQxcR372A-Zn (mutated)	20 mM TrisHCl pH 8.0, 200 mM NaCl + 1mM Zn	0.1M Bis-Tris pH 5.5, 0.2M $LiS0_4$ and 25% $PEG3350$

4.1 Materials and methods

4.1.1 Cloning, Expression and Purification of PepQxc1, PepQxc2, PepQxcR372A and PepQxcR372A-Zn

The coding DNA sequence (CDS) of a gene encoding PepQxc protein (GenBank accession, NP-637763) was from genomic DNA of *Xanthomonas campestris*, ATCC 33913 (NCBI taxonomy ID, 190485). The oligos, opx15 (cgtggatccagcacgcagatcggcgggatg) and opx16 (cgtaagcttatgcaaacggttgatcgatcgccac) comprising *BamHI* and *HindIII* restriction enzyme sites, respectively, were used to amplify the CDS in a PCR using thermostable Pfu DNA polymerase (Table. 4.2). The CDS was cloned into pST50Tr, a T7-promoter-based expression plasmid [99] so as to form in-frame translational fusion protein with streptavidin-His6-TEV tag (STRHISTEV). The cloning and expression hosts used in the present studies were XL-1 Blue and BL21(DE3)plysS strains of *Escherichia coli*, respectively. The construct was verified by restriction digestion analysis, and DNA sequencing using T7 and T7-term oligos. The expressing clone was grown in 3l of 2xTY broth at 310 K till $OD600$ reaches to $\sim$ 0.4, afterward, the culture was shifted to 291 K for further growth at lower temperature till $OD600$ reaches to 0.8. At this point, the culture was

induced by addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 18 h. The culture was harvested at 7000 rpm (Sorvall-GSA rotor) at ambient temperature and the cell pellet was suspended in $P300$ buffer (50 mM sodium phosphate pH 7.0, 300 mM NaCl). The desired protein was purified from the soluble fraction of the cell lysate using Ni-metal ion affinity (IMAC) and gel filtration chromatography techniques. Typically, the soluble lysate containing desired protein (STRHISTEV-XPD43) was mixed with 3 ml of pre-equilibrated ($P300$ containing 20 mM imidazole) Ni-sepaharose resin (GE Healthcare) and the bound protein was eluted with $P300$ buffer containing 500 mM imidazole by batch method at 277 K. The eluted protein STRHISTEV-XPD43 was subsequently subjected to digestion by tobacco etch virus protease (TEV) (1:50 molar ratio) at 293 K for 24 h, while dialyzing against $T200$ (20 mM TrisHCl pH 8.0, 200 mM NaCl). The desired protein (PepQxc) was separated from STRHISTEV peptide and undigested STRHISTEV-XPD43 protein using His-trap (5 ml) column on AKTA system (GE Healthcare) at room temperature. The protein was concentrated to $\sim$ 15 mg/ml using 10 kDa cut-off ultrafiltration unit (Sartorius); glycerol was added up to 20% (v/v), and stored at 203 K freezer till further use.

The R372A mutation in the peQxc gene was introduced using site directed mutagenesis. The forward (GCGTTCGGTGTGgcgCTGGAAGACCATTTC) and reverse (GAAATGGTCTTCCAGcgcCACACCGAACGC) primer has been used for the site directed mutagenesis. Further, the mutation in plasmid was confirmed by DNA sequencing. The expression and purification of PepQxcR372A has been carried out as per the method described for native PepQxc previously.

4.1.2 Protein crystallization

For crystallization, the required amount of stored (in 203 K) wild type PepQxc protein was subjected to size exclusion column chromatography. The chromatography was carried out using superdex 200 10/300 GL (GE Healthcare) with $T200$ buffer for removal of impurities, heterogeneity and buffer equilibration. The eluted protein was concentrated to $\sim$ 12 mg/ml and used for setting up crystallization at 294 K employing under-oil microbatch method [100]. Typically, 2 μl of protein solution was mixed with 2 μl of crystallization solution in 96-well U-bottom plates and overlaid with 50 μl of Al's oil (Hampton Research). Initial crystallization screening was performed using Index (Hampton Research) and JCSG-plus (Molecular Dimensions) crystal screens. The crystal hits were observed in several conditions of JCSG-plus and Index, but all of them were restricted to $pH \leq 5$.

Table 4.2: Macromolecule production information

Source organism	*Xanthomonas campestris*
DNA source	genomic DNA
Forward primer	cgtggatccagcacgcagatcggcgggatg
Reverse primer	cgtaagcttatgcaaacggttgatcgatcgccac
Cloning vector	None
Expression vector	pST50Tr
Expression host	*Escherichia coli*, BL21(DE3) plysS
Complete amino acid sequence of the construct produced	MASWSHPQFEKGSSHHHHHHSSGSGGGGGENLYFQGSSTQIGGMSLD QARTQLAPWTQRAAPIGADEYQQRIERARVLMRAQGVDALLIGAGTS LRYFSGVPWGASERLVALLLTTEGDPVLICPAFEEGSLDAVLQLPVR KRLWEEHEDPYALVVQAMDEQHAHALALDPGIAFAVHTGLRAHLGTAI RDAGAIIDGCRMCKSPAELALMQQACDMTLLVQRLAAGIAHEGIGTD QLVRFIDEAHRALGADNGSTFCIVQFGHATAFPHGIPGVQHLRAGEL VLIDTGCTVQGYHSDITRTWIYGTPSDAQQRIWELELAAQAAAFAAVR PGVACEAVDQAARAVLQAAGLGPDYRLPGLPHRTGHGCGLAIHEAPY LVRGNRQPLQPGMCASNEPMIVVPGAFGVRLEDHFYVTDTGAQWF TPPSVAIDQPFA

However, the best diffraction quality crystals were grown from the crystallization conditions containing 40 mM KH_2PO_4, 15% glycerol, 12% (w/v) polyethylene glycol (PEG) 8000 for PepQxc1 and 0.1 M Citric acid pH 5.0, 20% $PEG6000$ for PepQxc2. These crystals were of rod shaped and typically grew to a size of about 0.1 $\times$ 0.1 $\times$ 0.5 mm^3 in 45 days (Fig. 4.3, inset).

Similar to native PepQxc crystallization protocol, the required amount of mutated PepQxcR372A protein was purified by size exclusion column chromatography in $T200$ buffer. The eluted protein was concentrated to $\sim$ 15 mg/ml and crystallization trials were set up at 294 K with under-oil microbatch method. Initial crystallization screening was performed with JCSG plus screen, using 1 μl protein solution with 1 μl crystallization solution in 96-well U-bottom plates and overlaid with 50 μl of Al's oil (Hampton Research). The U-bottom plates were kept at incubation temperature of 21 $^{\circ}C$ for crystal growth. Best crystals of PepQxcR372A appeared in new condition: 0.1M Bis-Tris pH 5.5, 0.2M $LiS0_4$ and 25% $PEG3350$, different than the native PepQxc crystallization condition. The crystals were rectangular solid with typical size of about 0.2 $\times$ 0.2 $\times$ 0.1 mm^3. The crystallization condition for four wild type and two mutated PepQxc proteins are summarized in Table. 4.1.

4.1.3 Data collection and structure determination

The PepQxc1 crystals were directly frozen in liquid nitrogen (LN2) without any additional cryoprotectant. The PepQxc2 crystals were first treated with cryoprotectant solution (mother liquer + 20% glycerol) before freezing into LN2. Crystals of the pepqxcR372A were soaked in $1mM$ His-Pro substrate with varying time intervals (10s, 1min, 2 min, and 30 min). The soaked crystals were then coated with paratone oil (Hampton Research) for cryo-protection before freezing into LN2. Similarly, the PepQxcR372A-Zn crystals were prepared by soaking the PepQxcR372A crystals in $1mM\ ZnCl_2$ solution for different time intervals, from 10 *sec* to 30 *mins*. Subsequently, crystals were dipped into paratone oil for cryo-protection before freezing into LN2. Crystals of the PepQxc1-Mn and the PepQxc1-Zn, were obtained by soaking the PepQxc1 crystal in $1mM\ MnCl_2$ and $1mM\ ZnCl_2$ solution, respectively, then the soaked crystlas were cyro-protected with paraton oil before freezing with LN2. All crystals were flash frozen in liquid nitrogen and stored in LN2 dewar till the diffraction studies.

X-ray diffraction data were collected using crystals of all six variants of the protein (PepQxc1, PepQxc1-Mn, PepQxc1-Zn, PepQxc2, PepQxcR372A and PepQxcR372A-Zn) on recently commissioned bending magnet protein crystallography beamline (PX-BL21) of Indus-2 synchrotron (India). Diffraction data were collected on cryo-cooled crystals (100 K) with an oscillation of 1° at wavelength of 0.97947 $\mathring{A}$, and for PepQxcR372A-Zn crystals at wavelength of 1.2822 $\mathring{A}$, using MAR225 CCD (Rayonix) detector. The data were collected up to 1.83 $\mathring{A}$ resolution with completeness better than 98%. The collected data were indexed and integrated using XDS [101] and subsequently, scaled using AIMLESS [81]. The data statistics are summarized in Table. 4.6.

All six crystals were belonged to $P2_12_12_1$ space group. The unit-cell parameters of PepQxc1, PepQxcR372A and PepQxcR372A-Zn, are in same range with c-axis 111.5$\mathring{A}$, while the a-axis length varies from 84.3 *to* 81.09 $\mathring{A}$ and the b-axis changes from 105.5 *to* 101.9 $\mathring{A}$, as shown in Table. 4.6. Since, the crystallization condition of PepQxc2 is different from the PepQxc1, as expected, the unit cell parameters ($a = 58.5$, $b = 105.5$ and $c = 111.4$ $\mathring{A}$) are different. The crystal structure of PepQxc1, PepQxc1-Mn, PepQxc1-Zn, PepQxc2 and PepQxcR372A, were solved by molecular replacement method (MR). The Matthews coefficients [102] were calculated to be 2.9 $\mathring{A}^3\ Da^{-1}$ with a 58% solvent content. The two monomers were present in the asymmetric unit of the cell. The standard protein BLAST search using the protein sequence at NCBI (http://blast.ncbi.nlm.nih.gov) against the Protein Data Bank (PDB) database fetched a homolog from an archaea, *Ther-*

mococcus sibiricus (PDB entry, 4FKC) as the best structural match for PepQxc ($\sim$ 29% identity). The model based on the monomer of PDB entry 4FKC was used for molecular replacement solution using PHASER suite [103] on the CCP4 platform [81]. The convincing molecular replacement solution with final log-likelihood gain (LLG) score of 191 with two monomers in the asymmetric unit was obtained. Initial atomic model were build by the AutoBuild Wizard from Phenix [87]. Almost 85% to 90% of residues of protein chain were built by AutoBuild Wizard with a R_{work}/R_{free} of 26/30%. Further model building were carried out manually in Coot [88] by utilizing 2mFobs-DFmodel maps from Phenix.refine [87]. Final refinement for four native and two mutated crystals were converge to R_{work}/R_{free} of $< 20/22\%$, as shown in Table. 4.6.

Zn-SAD structure of the mutant R372A PepQxc (PepQxcR372A-Zn): The crystal structure of PepQxcR372A-Zn has been solved using single wavelength anomalous diffraction (Zn-SAD) method. The X-ray diffraction data on crystals soaked in $1mM$ $ZnCl_2$ were collected at at peak energy (max f'') of 9669 eV for the Zinc metal by tuning the double crystal monochromator of the beamline. The heavy metal soaked crystal diffracted up to 1.85 $\AA$ resolution with overall I/σ of 20. The collected data showed good anomalous signal (Mid-Slope of Anom Normal Probability = 1.256). Furthermore, the HySS routine in autosol [87] has located 12 Zn atoms in a unit cell, out of which, four Zn metal atoms with full occupancy were present at the active site of the enzyme and three metal atom with low occupancy were found at the dimeric interface of protein. Therefore, the seven numbers of Zn metal ions have been located in final refined model of protein with R_{work} = 18.46% and R_{free} = 22.42% (Table. 4.6).

4.1.4 Enzyme Kinetics

Prolidase assay: The 250 μl reaction mixture containing 100 mM Tris buffer (pH 8.0), 2 mM $MnCl_2$, 2 mM substrate and approximately 2 μg enzyme were prepared. The reaction mixture was incubated for 10 min at 50 $^{\circ}C$. After the incubation, 250 μl glacial acetic acid was added to stop the reaction. Then 250 μl ninhydrin reagent (3% [wt/ vol]) was added to the mixture and boiled for 10 min. The absorbance (OD) was measured with UV-Visible spectrophotometer (Jsco V-630 spectrophotometer). The OD was determined at 515 nm with an extinction coefficient $4,570$ $M^{-1}cm^{-1}$ for substrate-ninhydrin complex. One unit of prolidase activity is defined as the amount of enzyme that liberates one micromole of proline per minute.

4.2 Results

The typical yield of recombinant PepQxc was found to be $\sim 90\ mg$ from $3l$ of bacterial culture. The purified protein showed a single symmetrical peak at molecular weight corresponding to $\sim 70\ kDa$ on size-exclusion chromatography (Fig. 4.2), while the expected monomeric mass of the molecule deduced from the sequence is $42.8\ kDa$. The result thus suggests that the PepQxc1 exists as a dimer in the given condition. The protein peak is eluted somewhat later than the expected elution time possibly due to more compact and globular nature of dimer, or due to weak interaction of the protein with Superdex 200 column. As mentioned earlier, PepQxc1 and peQxc2, both crytallized in orthorhombic space group $P2_12_12_1$ with different unit cell dimensions. The crystal structures were determined with molecular replacement method. The structure shows that there are two molecules in the crystallographic asymmetric unit. The atomic coordinates and structure amplitude file of PepQxc1 and PepQxc2 has been deposited in Protein Data Bank (PDB) with accession code 4R60 and 5CIK respectively. The structure of PepQxcR372A-Zn has been determined using SAD method as described earlier. The atomic coordinates of PepQxcR372A-Zn has been deposited in PDB with accession code 5CDE.

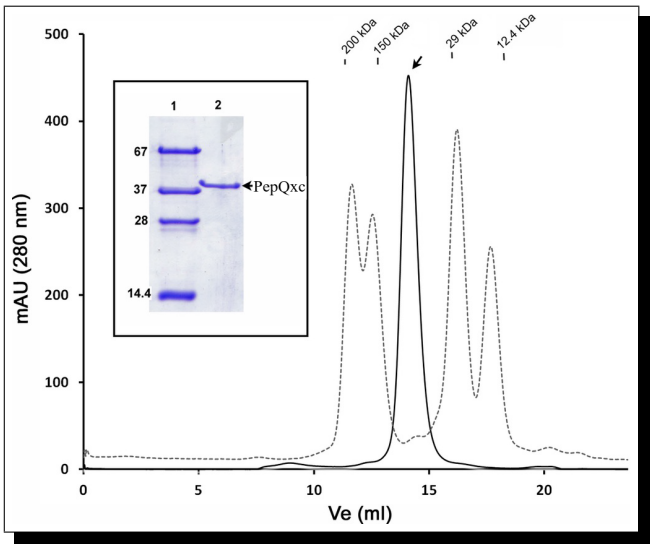

Figure 4.2: The size-exclusion chromatography profile of purified PepQxc (marked by an arrow) on pre-calibrated superdex 200 using AKTA system (GE Healthcare) with buffer 20 mM, Tris-HCl pH 8.0, 200 mM NaCl. A 15% SDS-PAGE (inset) showing molecular mass markers in kDa (lane 1) and purified PepQxc (lane 2).

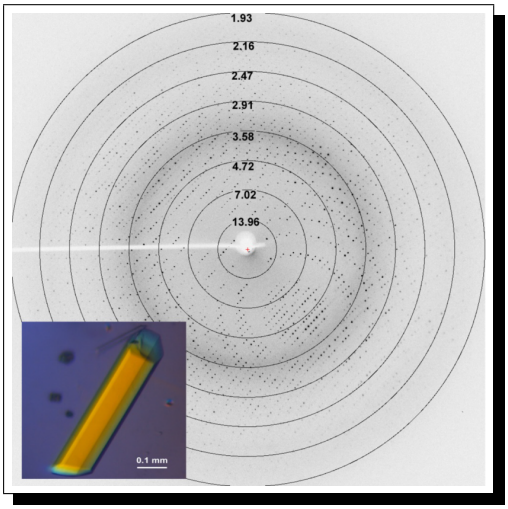

Figure 4.3: Diffraction pattern recorded at PX-BL21, Indus-2 on a PepQxc crystal (inset) at exposure of 30 s and 1° oscillation. The crystal was grown from 40 mM KH2PO4, 15% glycerol, 12% PEG-8000 in 45 days at 294 K temperature (inset).

4.2.1 Overall structure of PepQxc

The structural analysis of PepQxc1 shows that the protein is dimeric in nature. The polypeptide fold is shown in Fig. 4.4a, each monomer is composed of two domains; a N-terminal from residues $27-166$ and a C-terminal pita-bread domain with residues from $2-25$ and $167-399$. The C-terminal domain is shared with the other prolidases (EC 3.4.13.9), aminopeptidases (EC 3.4.11.9) and methionine aminopeptidases (EC 3.4.11.18). The N-terminal domain is connected with C-terminal domain via two random coil (residues $26-28$ and $165-167$) linkers. The N-terminal has six mixed central β-strands flanked by six α-helices (four on one side and two on other). The C-terminal domain is composed of six α-helices and twelve β-strands. The active site is an oval depression in β-sheet part of the C-domain as shown in Fig. 4.4. The β-strands surrounding the active site is flanked by four long parallel running α-helices (two on each side). A loop of residues from $233-235$ make a partial flap over the active site and will be referred as the "PHG" loop. The active site is characterized by di-nuclear metal center containing two Mn ($Mn1$ and $Mn2$) ions. These metal ions are $3.3\mathring{A}$ apart and interacts with surrounding residues as shown in Fig. 4.5a. The side chains of two aspartate residues (251 and 262), two glutamate residues (360 and 374) and a histidine (331) participate in co-ordination bond with Mn atoms at the active site. The $Mn1$ is co-ordinated by three protein oxygen atoms (residues; $262, 360$ and 374) and one nitrogen atom from residue 331, while, $Mn2$ is co-ordinated by four protein oxygen atoms (residues; $251, 262$ and 374). The carboxylate group of Asp-262 and Glu-374 bind as bridging ligands to both metal ions. The Asp-251 act as bi-dentate ligand to $Mn2$ ion. Overall structure of all six (PepQxc1, PepQxc1-Mn, PepQxc1-Zn, PepQxc2, PepQxcR372A and PepQxcR372A-Zn) types are identical (Fig. 4.4b), PepQxc1-Mn and PepQxc1-Zn structures are identical to the PepQxc1 structure with pairwise rmsd of $0.06\mathring{A}$. While, the structures of PepQxc2 and PepQxcR372A are similar to the PepQxc1 with pair wise RMSD of $0.46\mathring{A}$ and $0.17\mathring{A}$, respectively.

PepQxc1-Mn and PepQxc1-Zn were crystallized in $P2_12_12_1$ space group and exist as dimer in an asymmetric unit. Interestingly, a phosphate molecule was found in the active site of both the enzymes, and it interacts with the two metal ions at the active site Fig. 4.6. The prolidase activity is carried out by an activated nucleophilic water molecule [97], which is located at mid-plane of di-nuclear metal ions at the active site. This nucleophilic water molecule in the active site has been displaced by the phosphate molecule in both the PepQxc1-Mn and the PepQxc1-Zn. The extra electron desity is found at the dimeric interface of the both PepQxc1-Mn and PepQxc1-Zn at the N-terminal side of dimer, as

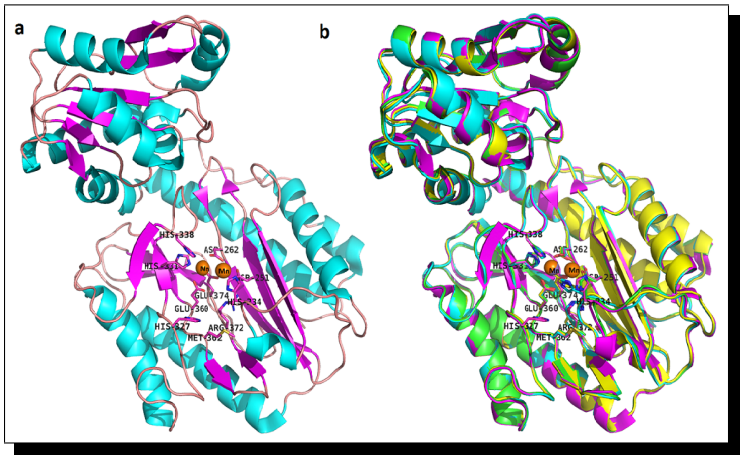

Figure 4.4: (a) Cartoon represetation of the PepQxc monomer with helixes in cyan, β-sheets in magenta and two Mn atoms in orange color. (b) Cartoon representation of the superposed monomer of PepQxc1(colored in magenta), PepQxc2(colored in cyan), PepQxcR372A (green colored) and PepQxcR372A-Zn (yellow colored). Mn atoms at active site are from the PepQxc1, shown for reference

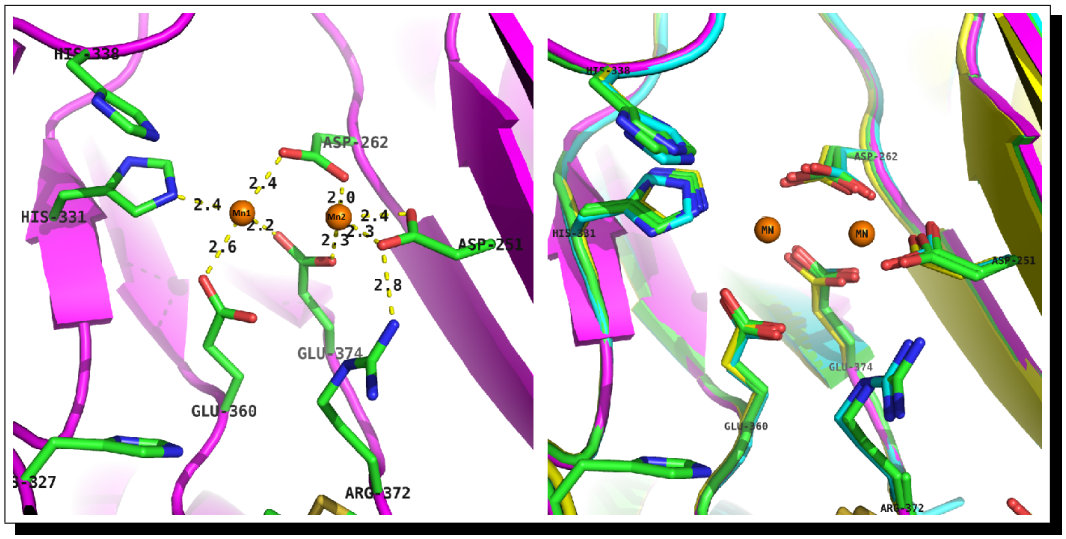

Figure 4.5: A detail view of acitive site containing di-nuclear metal center. (a) PepQxc1 active site: the atoms are color coded: **Mn** in ornage, **O** in red, **C** in geen and **N** in red. Atom to atom distance is represented in $\AA$ units. (b) Superposed active sites of PepQxc1(colored in magenta), PepQxc2 (colored in cyan), PepQxcR372A (green colored) and PepQxcR372A-Zn (yellow colored). Mn atoms at active site are from PepQxc1, shown for reference

shown in Fig. 4.7. The crystallization condition of both enzymes have glycerol, phosphate and PEG. Attempts were made to refine the extra electron density with these compounds. But, none of these compounds and their by-products were able to account for the extra electron density.

PepQxc dimer; The biological active unit of PepQxc is a dimer and crystallizes in $P2_12_12_1$ space group with two molecule in asymmetric unit, as shown in Fig. 4.8b. Both molecules in the asymmetric unit are structurally similar, and pair wise root mean square

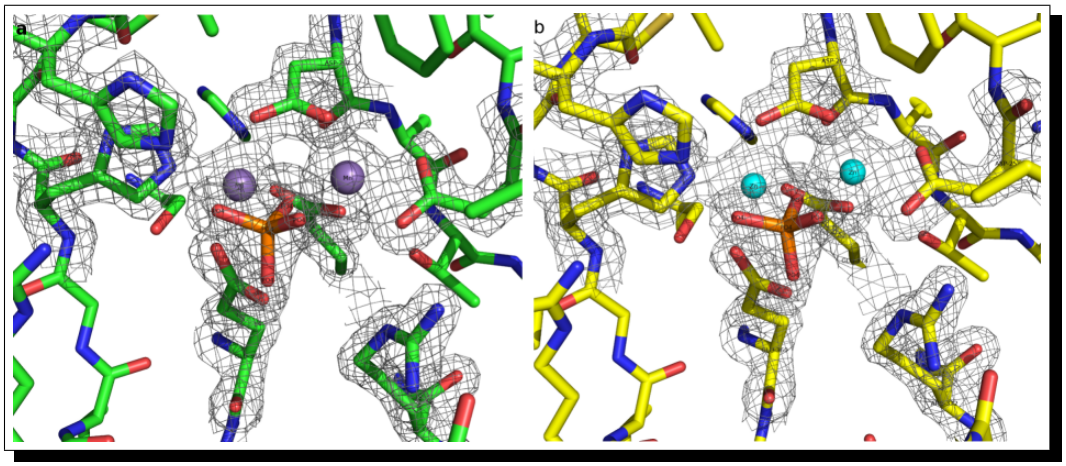

Figure 4.6: A detail view of acitive site containing dinuclear metal center and phosphate ion (Colored in Orange) with 2Fo-Fc Fourier map (Colored in gray) contoured at 2.0 sigma. (a) PepQxc1-Mn active site: the atoms are color coded: **Mn** in purple, **O** in red, **C** in geen and **N** in blue. (b) PepQxc1-Zn active site: the atoms are color coded: **Zn** in cyan, **O** in red, **C** in yellow and **N** in blue.

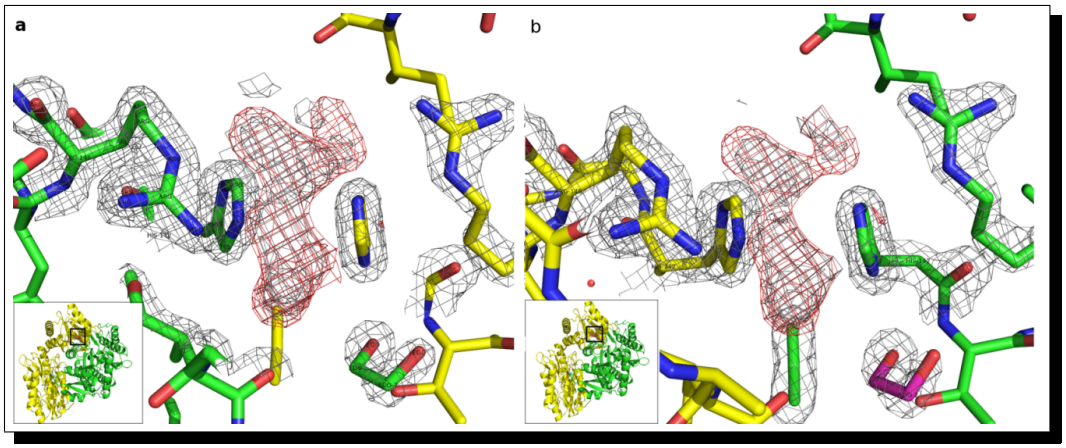

Figure 4.7: A detail view of the dimeric interface containing extra unknown electron density shown with 2Fo-Fc Fourier map (Colored in gray) contoured at 2.0 sigma and Fo-Fc Fourier map (Colored in red) contoured at 5.0 sigma. (a) PepQxc1-Mn dimeric interface: Chain A is represented in green and Chain B is shown in yellow. (b) PepQxc1-Zn dimeric interface: Chain A is represented in yellow and Chain B is shown in green.

deviation (RMSD) between molecules in asymmetric unit is 0.33$\mathring{A}$. The homodimer has approximate dimension of 73 $\times$ 70 $\times$ 60 $\mathring{A}^3$ with relatively significant buried area of 1629 $\mathring{A}^2$ between subunits of the dimer. The dimer has non-crystallographic dyad axis perpendicular to the z-axis, which orients the N-terminal domain of the monomer on the active site of C-terminal domain of the other monomer, as can be seen in Fig. 4.8b.

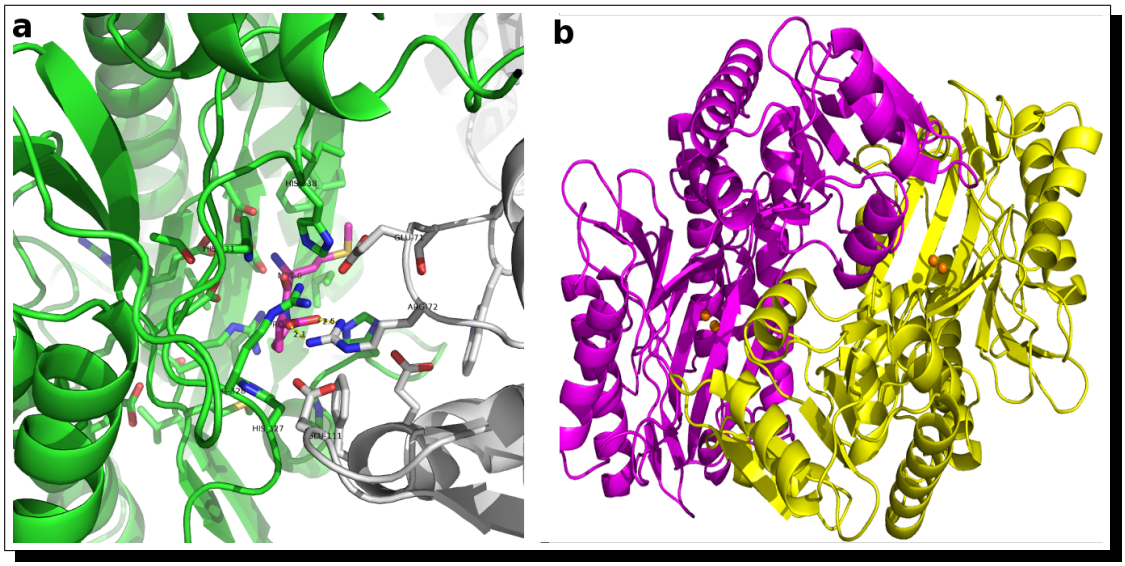

Figure 4.8: (a)Met-pro (colored magenta) docked at active site of peQxc chainA (colored green) via alignment with ValProLeu (PDB entry 2BHA) substrate from *E. Coli* APPro. Carbonyl group of Met-Pro make salt bridge with *Arg72* from chainB (Colored grey) (b)PepQxc dimer viewed along the local 2-fold axis coming out of the plane of paper. One subunit is colored yellow and other represented in magenta color. Metal ions are shown in orange color.

4.2.2 Citrate bound structure of PepQxc2

The PepQxc2 exist as a dimer in asymmetric unit and both molecules of dimer are structurally similar. There are no metal ions present at active site, instead, a citrate molecule is bound at active site as shown in Fig. 4.10. The citrate molecule interacts with numbers of residue at active site. The O5 atom of citrate makes a hydrogen bond with N^E atom of $His234$. The O4 and O2 oxygen atoms of citrate makes hydrogen bonds with N atom of $Arg372$ and N^D atom of $His331$ atoms respectively. The citrate molecule is present at same location in both chains of a homodimer. Interestingly, *His234* exists in different conformations in in two chains as shown in Fig. 4.9a. In chain A, the *His234* interacts with citrate, while in chain B, it interacts with *Trp67* of chain A from the dimer.

4.2.3 PepQxc structure with R372A mutation

The R372A mutant (PepQxcR372A protein) is structurally similar to PepQxc as can be seen from superposed structure model of PepQxcR372A with that of other three molecules (PepQxc1, PepQxc2 and PepQxcR372A-Zn) in Fig. 4.4b. Also, the active site of PepQxcR372A is identical to native PepQxc as superposed structures are shown

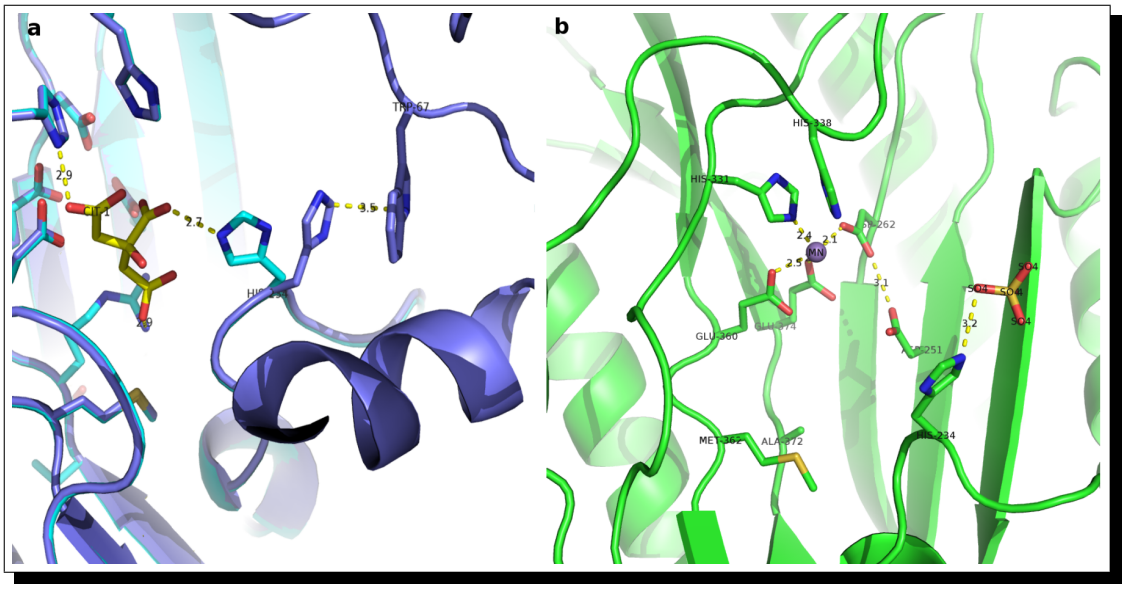

Figure 4.9: (a) PepQxc2: Superposed active sites of chain A (colored in Cyan) and chain B (colored in purple) of PepQxc1. Residue *His234* exist in different conformation in chain A and chain B. It interacts with citrate in chain A and in chain B, it interacts with residue *Trp67* of chain A from the dimer. (b) PepQxcR372A: Active sites of PepQxcR372A (colored in green) shows only one Mn atom in active site, **O** in red, **Mn** in purple and **N** in blue. The sulphate molecule is interacting with *His234* residue.

in Fig. 4.5c. The PepQxcR372A was found to be biologically inactive and incapable of hydrolysing the Xaa-Pro substrate. Therefore, an attempt was made to obtain the crystal structure of Xaa-Pro and PepQxcR372A complex by soaking experiments. The crystals were soaked in $1mM$ His-Pro dipeptide substrate for various time period but no extra density has been located at active site which can be ascribed to His-Pro compound. A single metal ion ($Mn1$) was found to be bound at the active site of the mutated PepQxcR372A, instead of two as in the native PepQxc1 protein. Even for the PepQxc1, the second metal ion(Mn2) binds at active site with less affinity as compared to the first metal ion (Mn1). The absence of second metal ion at the active site could be due to R372A mutation, or some unknown reasons.

To further investigate the cause of absence of the second metal ion from the active site, the PepQxcR372A crystals were soaked with $1mM$ $ZnCl_2$ solution and the structure was solved by Zn-SAD method. Total seven Zn atoms were located in asymmetric unit, out of which, two each were present at active sites of homodimer. The anomalous difference Fourier map confirms the presence of Zn atom at the active site with full occupancy as shown in Fig. 4.11. Three Zn atoms are present at the dimeric interface, and two of it can be seen in Fig. 4.11. Active site Zn metal ions are placed at identical position as in the

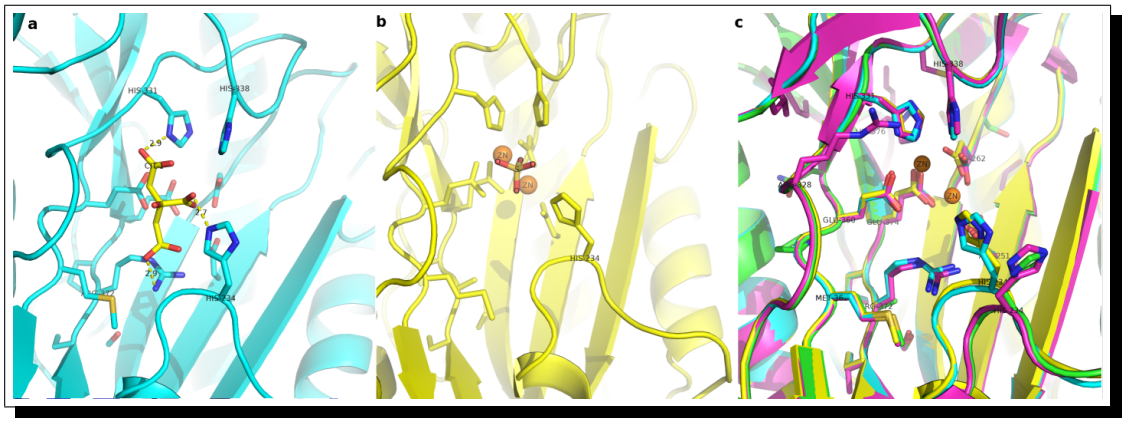

Figure 4.10: A detail view of active site with substrate binding. (a) PepQxc2 active site: the atoms are color coded: **Citrate** in yellow, **O** in red, **C** in cyan and **N** in blue. Atom to atom distance is in $\mathring{A}$ units. The O5 (Citrate) atom makes a hydrogen bond with 234 $His\ N^E$ atom. The O4 and O2 atoms of Citrate molecule makes hydrogen bonds with 372 $Arg\ N$ and 331 $His\ N^D$ atoms, respectively. (b) The active site of PepQxcR372A-Zn (colored in yellow), Sulphate bound at same location in active site, instead of Citrate as in the PepQxc2 . Two Zn ions are present at active site with full occupancy. (c) Superposed active sites of the PepQxc1 (colored in magenta), the PepQxc2 (colored in cyan), the PepQxcR372A (green colored) and the PepQxcR372A-Zn (yellow colored), showing loop (residues:233-235) positions. Zn atoms at active site from the PepQxcR372A-Zn are shown for reference.

native PepQxc1. The Zn atoms located at the dimeric interface make the hydrogen bonds with three nearby residues. For example, the *His234* of chain A and the *Glu92* of chain B make hydrogen bonds with one the Zn atom at dimeric interface, similarly another Zn atom make hydrogen bonds with the *His338* residue of chain A and the *Glu71* residue of chain B, as shown in Fig. 4.11. The active site of PepQxcR372A-Zn also contains the sulphate molecule at the same location where a citrate molecule binds in the PepQxc2, as can be seen in Fig. 4.10b. The sulphate molecule interact with the *His234* and *His338* residues at active site.

4.2.4 Hydrolysis of Xaa-pro dipeptides by PepQxc

The substrate specificity of the PepQxc in presence of the metal ion Zn^{2+} and the Mn^{2+} was determined using Xaa-Pro dipeptide substrates. The substrate specificity for the Mn^{2+} and the Zn^{2+} metal ion containing PepQxc is listed in Table. 4.3. These measurements have been carried out at pH 8.0 and 50°C. The PepQxc has shown the activity for all the tested dipeptides in the presence of various metal ions. In general, the PepQxc has shown higher activity in presence of Mn^{2+} ion with some exceptions. The comparative

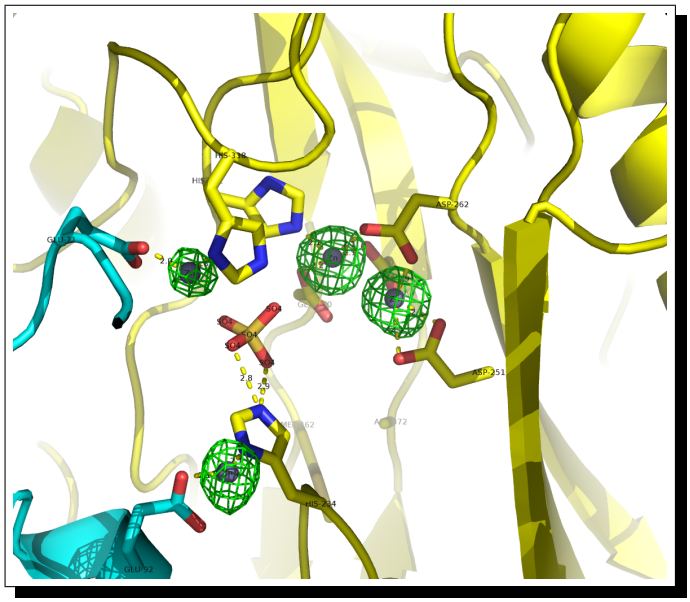

Figure 4.11: PepQxcR372A-Zn: Active sites of PepQxcR372A-Zn shows two Zn atoms in active site, The anomalous difference Fourier map (Colored in green) contoured at 5.0. **O** in red, **Zn** in grey and **N** in blue. The sulphate molecule interacting with the *His234* residue, also two Zn atoms can be seen to bind at the dimeric interface i.e., between chain A (yellow) and chain B (cyan).

chart of the PepQxc substrate specificity for different metal ions (Mn^{2+}, Zn^{2+} and Co^{2+}) were shown in Fig. 4.12. The substrate specificity of the enzyme differs in presence of metal ions. For example, the Leu-Pro exhibited better activity only in presence of Zn^{2+}, and lesser in Mn^{2+} or Co^{2+}. The His-Pro substrate showed the lesser activity in presence of Co^{2+}. Activity towards Ala-Pro, Gly-Pro, Val-Pro and Ile-Pro could be better observed only in the presence of Mn^{2+} ion. The Phe-Pro is the best substrate irrespective of metal ion used, and has highest K_{cat}/K_m value of $1.11 \times 10^5 (/M/s)$ for the Mn^{2+} ion.

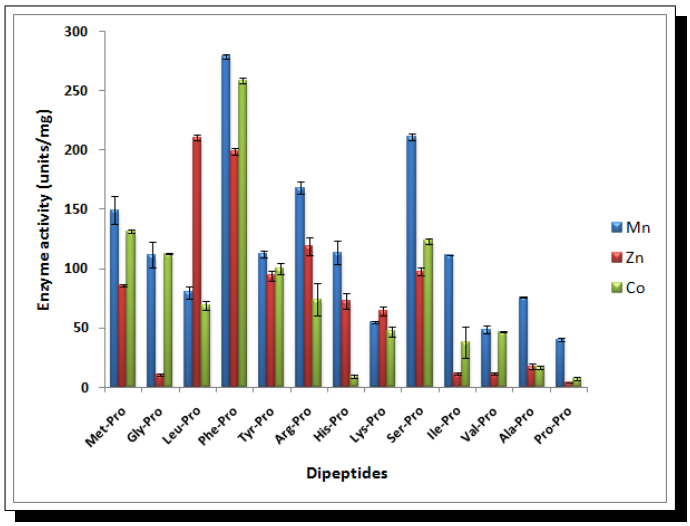

Figure 4.12: PepQxc substrate specificity in the presence of different metal ions (Co^{2+}, Zn^{2+} and Mn^{2+}). Phe-Pro is the best substrate irrespective of metal ion used.

The optimum concentration of metal ion is different for different metal ions as shown in Fig. 4.13. The enzyme shows the maximum activity, at $1mM$ metal ion concentration for Mn^{2+} ions. The enzyme is activated by several metal ions viz., Fe^{2+}, Fe^{3+}, Mg^{2+},

Ca^{2+} and Ni^{2+} with maximum activity at $0.005mM$ metal ion concentration, as seen in the Fig. 4.13b. The presence of Mn^{2+} ion, improves the thermal stability of the enzyme, as shown in Fig. 4.14b. The phosphate concentration has the inhibitory effect on the enzymatic activity of PepQxc (Fig. 4.14a).

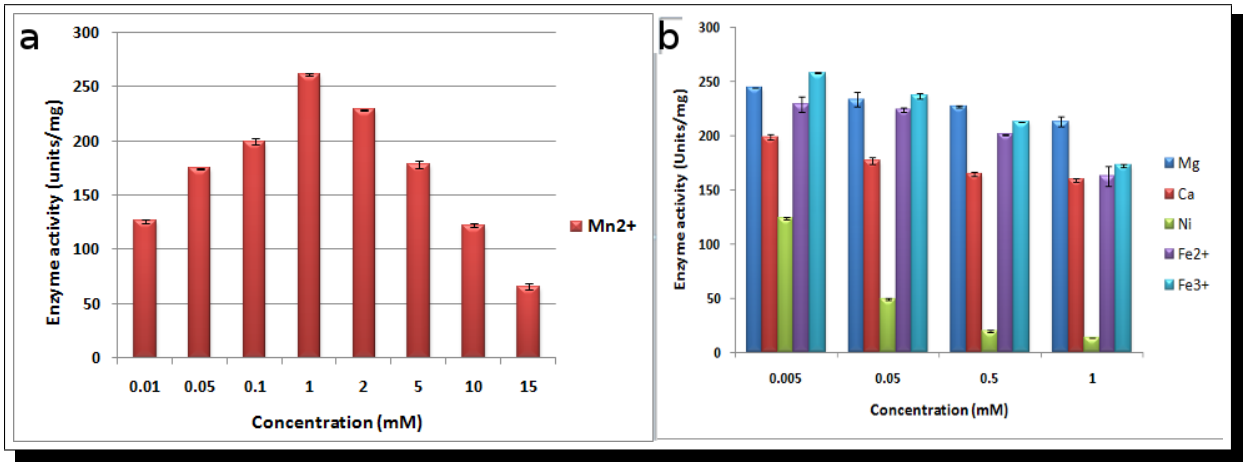

Figure 4.13: PepQxc enzymatic activity against metal ion concentration. The optimum concentration of metal ion is different for different metal ions. (a) Enzymatic activity variation with Mn^{2+} ion concentration. (b) Enzymatic activity variation with Fe^{2+}, Fe^{3+}, Mg^{2+}, Ca^{2+} and Ni^{2+} ions concentration.

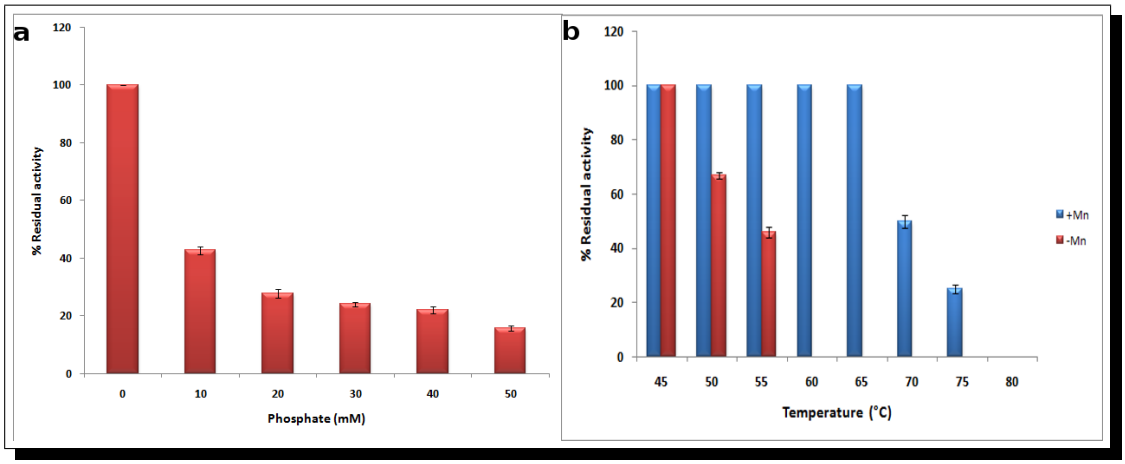

Figure 4.14: (a) PepQxc activity inhibition against phosphate ion concentration. (b) Temperature stability of PepQxc's activity in presence and absence of Mn^{2+}.

The enzyme is strictly dipeptidase. Interestingly, the PepQxc shows the concentration dependent substrate inhibition for a few dipeptides, as shown in the Table. 4.4. At $3mM$ concentration of the Pro-Pro and the His-Pro substrate, the enzyme shows the complete inhibition of the activity. But, the Met-Pro, the Phe-Pro, the Gly-Pro and the Ala-Pro,

does not shows any inhibitory effect on the enzymatic activity, these di-peptides has either non-polar or small or sulphur containing first residue. It is also found that the presence of inhibitory concentration of one substrate, does not allow the enzyme to cleave other substrates, which are not inhibitory even at higher concentration, suggesting that the inhibition is non-competitive. Also, the inhibitory effect is reversed by dilution, which suggest the reversible binding at the inhibitory site.

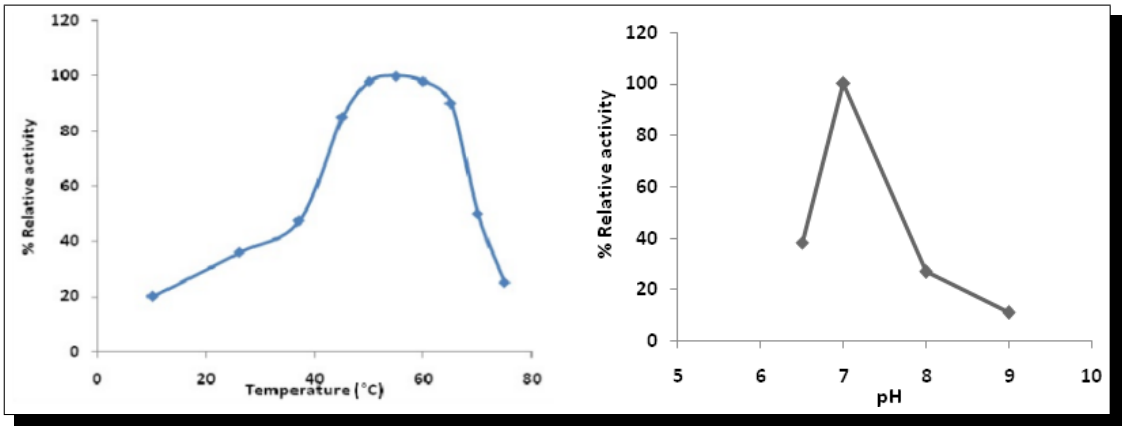

Figure 4.15: PepQxc relative enzymatic activity against temperature and pH variations.

4.3 Discussion

The crystal structure of four native and two mutant (R372A) prolidase from *Xanthomonas campestris* has been determined to high resolution (Table.4.6). The four native crystal structures has been determined from two different crystallization conditions, crystallization conditions of native and mutant proteins are listed in Table. 4.1. Also, the detailed enzyme kinetics has been performed on the PepQxc protein (Table.4.3). Substrate specificity of the PepQxc enzyme for various dipeptides has been measured in presence of Mn^{2+} and Zn^{2+} metal ions. The enzymatic activities in presence of various metal ions have been determined. The temperature and pH dependence of the PepQxc enzyme were measured to find out optimal functional working range of the enzyme.

4.3.1 Role of *Arg372* in PepQxc

Although, the R372A mutation in PepQxc leads to inactive enzyme, the crystal structure of PepQxcR372A and the PepQxcR372A-Zn shows that the R372A mutation keep the overall structure of the active site identical to native PepQxc enzyme, as shown in Fig.4.5b.

Table 4.3: Kinetic constants for the hydrolysis of Xaa-Pro dipeptides at pH 8.0 and 50^0C (metal ion: Mn^{2+} and Zn^{2+})

	Mn^{2+}			Zn^{2+}		
Substrates	K_m (mM)	K_{cat}/K_m (/M/s)	K_{cat}(/sec)	K_m (mM)	K_{cat}/K_m (/M/s)	K_{cat}(/sec)
Phe-Pro	3.45	1.11×10^{05}	384.09	2.471	1.21×10^{05}	298.18
Met-Pro	2.99	5.34×10^{04}	160.05	1.448	4.00×10^{04}	57.97
Leu-Pro	1.61	5.51×10^{04}	88.86	3.404	8.13×10^{04}	276.91
Gly-Pro	3.32	3.24×10^{04}	107.83	*N.A.*†	N.A	N.A
Tyr-Pro	2.32	5.16×10^{04}	119.69	5.199	4.59×10^{04}	238.68
Arg-Pro	6.05	5.77×10^{04}	348.85	6.235	3.54×10^{04}	220.76
His-Pro	1.76	2.95×10^{04}	51.85	1.897	3.65×10^{04}	69.23
Lys-Pro	3.02	3.32×10^{04}	100.43	1.9	3.13×10^{04}	59.52
Ser-Pro	*N.D.**	*N.D.*	*N.D.*	2.058	5.43×10^{04}	111.75
Ile-Pro	2.67	4.57×10^{04}	121.97	N.A	N.A	N.A
Ala-Pro	5.65	2.96×10^{04}	167.17	3.31	2.17×10^{04}	71.85
Val-Pro	1.94	2.72×10^{04}	52.95	N.A	N.A	N.A
Pro-pro	5.65	1.41×10^{04}	79.69	N.A	N.A	N.A

* N.D.: Activity was present but not fitting with given curve fitting equation
† N.A.: No Activity

This implies that the *Arg372* residue is not required for the structural integrity of the active site. The *Arg404* (equivalent *Arg372* in PepQxc) positive charge could provide the electrostatic repulsion between guanidinium group and substrate, but R404K mutation [6] rendered the APPro from the *Escherichia coli* inactive. Therefore, the loss of activity could not only be due to positive charge contribution from guanidinium group. In the crystal structure of PepQxcR372A, only one Mn^{2+} ion present at the active site (Fig. 4.9) and absence of second metal ion could be due to the R372A mutation. As can be seen in the PepQxc structure, the *Arg372* make a hydrogen bond with the *Asp251*, which coordinates in bidentate fashion with $Mn2$ atom at the active site. Thus, the loss of *Arg372* could lead to destabilization of *Asp251* residue and loss of second metal ion at the active site. But, the Zn-SAD structure of $1mM\ ZnCl_2$ soaked PepQxcR372A crystals shows that the Zn^{2+} ion exist at both the metal center site with full occupancy

Table 4.4: Pepqxc Xaa-Pro dipeptides concentration for substrate inhibition

Substrates	Inhibitory effect observed at concentration (mM)	Minimum concentration for complete inhibition (mM)
Leu-pro	2.5	6
Tyr-Pro	5	7
Arg-Pro	2.5	6
His-Pro	2	3
Lys-Pro	2	4
Ser-Pro	3	5
Ile-Pro	2.5	5.5
Pro-Pro	2	3

as shown in Fig. 4.11. Therefore, the *Arg372* residue may not be critically required for metal ions binding at the active site of the enzyme. The role of *Arg372* residue in proton shuttling is not clear from PepQxc structure. Since, *Met352* can not participate in proton transfer and other way could be direct proton shuttle from bulk to *Arg372* to *Asp251* and then to leaving group of substrate. It could be that there are multiple pathways from which proton can be shuttled to substrate [97]. Summarizing the role of *Arg372* residue in PepQxc: The residue does not affect the overall structure of active site, not critically required for metal center stability, the positive charge contribution from guanidinium group is not necessary for catalytic activity and, it is not entirely clear that it participate in proton shuttling from bulk to substrate. Therefore, the other possible role of *Arg372* (or equivalent conserved residues in related enzyme) residue in prolidase could be in stabilization and proper orientation of transition state complex in active site, to facilitate nucleophilic attack by activated water molecule on peptide bond of Xaa-Pro substrate.

The *Met362* residue in PepQxc is structurally positioned at the place of *Tyr* residue from M24B family of peptidases. Analysis of the PepQxc structure shows that the active site of enzyme is structurally similar to other prolidases from M24B family (Fig. 4.16). The clustering of conserved residues (*Asp251, Asp262, Glu360, Glu374, His331, His338* and *Arg372*) around the dinuclear metal center is identical to other prolidases from M24B family. However, the strictly conserved glycine and tyrosine (*Gly385*, *Tyr387*; numbering as per *Escherichia coli* APPro) of M24B family of peptidases is changed to methionine

and valine, respectively, (Fig. 4.1), in the PepQxc enzyme. These conserved residues are proposed to participate in proton shuttling from bulk solvent to leaving peptide through the conserved hydrogen bond network *Asp260–Arg404–Tyr387* motif [6], in APPro of *Escherichia coli*. But the hydroxyl group of *Tyr*, which participate in hydrogen bonding with guanidinium group of *Arg* residue, is replaced by sulphur of *Met* residue in PepQxc, as shown in Fig. 4.16. The *Met* is a non-polar hydrophobic residue and cannot participate in proton shuttling with *Arg372* and *Asp251*.

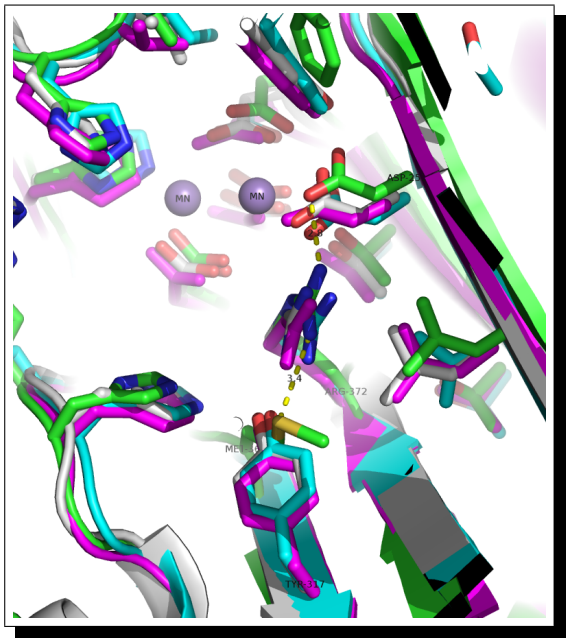

Figure 4.16: Superposition of proposed hydrogen bond network *Asp260–Arg404–Tyr387* motif of active site from PepQxc (pdb id:5CIK, Green colour), Human Prolidase (pdb id:2IW2, Grey colour), *P. furiosus* prolidase (pdb id:1PV9, Cyan colour) and Proline specific aminopeptidase (pdb id: 1JAW, Magenta colour). Atom to atom distance is represented in $\mathring{A}$ units. The **O** atom of conserved residue Tyr, which makes a hydrogen bond with **N** atom of Arg residue, is replaced by **S** atom of Met residue in PepQxc.

4.3.2 Role of the PHG loop in substrate binding

The residues from 233–235 make a partial flap over the active site. The PHG loop containing these residues is highly flexible and refined with relatively high B-factor for all six structures. Interestingly, this particular PHG loop found in open conformation for the PepQxc1 and the PepQxcR372A, and in closed conformation for the PepQxcR372A-Zn and the PepQxc2, as can be seen in Fig. 4.10c. The main chain of PHG loop moves toward the active site by 2.8$\mathring{A}$ (measured at C_α of *His234* residue) while, simultaneously C_α of the *Pro233* residue moves by 2.4$\mathring{A}$ toward the active site for closed conformation. The closed conformation of PHG loop in the PepQxc2 is accompanied with a *citrate* molecule at active site, the *Citrate* molecule makes a hydrogen bond with the *His234* residue, as shown in Fig. 4.10a. While, in case of the PepQxcR372A-Zn, a *citrate* is replaced by a *sulphate* molecule at the active site (Fig. 4.10b), which coordinates with the *His234*

residue. For the open conformation of PHG loop in the PepQxc1 and the PepQxcR372A, no substrate or ligand present at the active site. The PHG loop conformations for four PepQxc protein are listed in Table. 4.5. Therefore, it can be concluded that as long as active site is empty, the PHG loop remains in open conformation to allow substrate entry at active site and, as soon as substrate arrives at the active site, the PHG loop flip over to active site and changes to closed conformation, which lock the substrate in active site for the enzyme activity.

Table 4.5: PHG loop conformations in PepQxc

Protein (native/mutated)	Substrate at active site	PHG loop conformations
PepQxc1 (native)	No substrate	Open
PepQxc2 (native)	Citrate	Close
PepQxcR372A (mutated)	No substrate	Open
PepQxcR372A-Zn (mutated)	Sulphate	Close

His243 residue in APPro (Equivalent *His234* in PepQxc) is proposed [98] to involved in binding stabilization of substrate through a hydrogen bonding between imidazole group of the *His243* and the carbonyl group of the substrate. Subsequently, it was shown that the activity loss due to H243A mutation has no effect on binding of the Val-Pro-Leu substrate in the active site [104]. Therefore, the hydrogen bonding can not be essential for substrate stabilization [104]. Also, the *His243* was suggested to be one of the residue, which can shuttle a proton from the dinuclear center to solvent [97]. The structure of PepQxc2 shows the *His234* residue in two different conformation (for PHG loop in closed conformation) for chainA and chainB as shown in Fig. 4.9a. In one conformation the *His234* interacts with the bound *citrate* molecule in the active site, while in other conformation, it interacts with the *Tyr67* residue from other chain of dimer, as shown in Fig. 4.9a. Therefore, it further supports the previous conclusion [104], that the *His234* (and equivalent conserved residues in related enzymes) residue is not essential for substrate stabilization via hydrogen bonding. But *His243* might be able to shuttle proton from metal center to solvent, owing to its ability to change the conformation without disturbing the binding substrate at the active site.

4.3.3 Structural basis for substrate selectivity in PepQxc

PepQxc enzyme is strictly a dipeptidase, structurally the active site of enzyme cannot accommodate more than two residue long substrate. As can be seen in Fig. 4.8a, the carboxyl group of *proline* residue of the Met-Pro substrate interacts with the *Arg72* residue of the neighbouring chain. The *Arg72* residue is essential for enzymatic activity, as R72A mutation (Our unpublished result) in PepQxc leads to loss of the enzymatic activity in PepQxc. Furthermore, the residues (*Arg328, His338, His331,and His327* from chainA, *Glu71, and Glu111* from chainB) from both the chains of dimer are packed in the active site in such a way that it hinders the binding of longer than two residue substrate.

The PepQxc enzyme shows the substrate specificity against the various Xaa-pro dipeptides in presence of Mn^{2+} and Zn^{2+} metal ions as listed in Table. 4.3. The enzyme activity observed in presence of many metal ion viz., Co^{2+}, Fe^{2+}, Fe^{3+}, Mg^{2+}, Ca^{2+} and Ni^{2+}, is shown in Fig. 4.13b. The enzyme has maximum activity at $pH7.0$ (Fig. 4.15b) and at temperature range from $50°C$ to $60°C$ as shown in Fig. 4.15a. The enzyme is more efficient in hydrolyzing the Xaa-pro substrate with Mn^{2+} ion than Zn^{2+} ion (Table. 4.3) at active site except in Leu-Pro case, where the activity is more in presence of Zn^{2+} metal ion. Similarly, the activity towards Ala-Pro, Gly-Pro, Val-Pro and Ile-Pro could be better observed only in the presence of Mn^{2+} ion. The PepQxc shows higher enzymatic activity for Xaa-Pro substrates with first bulky residue like, *Phe, Arg, His* and *Tyr* in presence of both Mn^{2+} and Zn^{2+} metal ions (Table. 4.3). While, the Ser-Pro substrate show higher activity against all three metal (Mn^{2+}, Co^{2+} and Zn^{2+}) ions as shown in Fig. 4.12. Following conclusions can be drawn from the enzyme kinetic data:

- The Xaa-pro substrate with branched C-beta first residue, like *Ile*, *Val* or *Pro* at binding site of enzyme, has the less activity because of stereochemical hindrance created by branching near the metal ion.

- The active site of the PepQxc containing Mn^{2+} and Zn^{2+} metal ions is identical in composition and placement of residues. The activity difference for the same Xaa-pro residue could only be due to subtle change in coordination chemistry of metal ions. Therefore, in current studies no structural explanation for activity difference in Mn^{2+} and Zn^{2+} co-ordinated active site can be ascribed.

4.3.4 Phosphate and substrate inhibition in PepQxc

The enzymatic activity of PepQxc was inhibited by phosphate concentration, as shown in Fig. 4.14a. At $10mM$ phosphate concentration, the activity is reduced by 60% and at $50mM$ concentration, only 20% residual activity remained. The crystal structure of PepQxc1-Mn and PepQxc1-Zn shows (Fig.4.6) that a phosphate molecule is binding at the known position of activated nucleophilic water molecule [97]. Therefore, the replacement of activated water by phosphate molecule will inhibit the enzyme activity and as the concentration of phosphate is increased it will competitively displace the activated water from the active site.

The inhibitory effects were also observed in the PepQxc enzyme for some of the dipeptide substrates with change in concentration, as listed in the Table. 4.4. The Lys-Pro, Pro-Pro and His-Pro substrate completely inhibit the enzyme activity even at $3mM$ to $4mM$ substrate concentration. The inhibition effect is non-competitive and reversible in nature. The non-competitive inhibition by PepQxc enzyme suggest that there could be additional substrate binding sites on protein. An extra non-proteinous electron density has been found in PepQxc1-Mn and PepQxc1-Zn at the dimeric interface on N-terminal side as shown in Fig. 4.7. Similarly, two Zn^{2+} metal ions are found to be at dimeric interface of PepQxcR372A-Zn protein as shown in Fig. 4.11. Therefore, these additional sites at dimeric interface might act as a binding pocket for a substrate at higher concentration. The binding at dimeric interface could reduce the relative movement of chains in dimer, which inturn could inhibit the substrate movement in vicinity of the active site.

4.4 Summary

A total of six (four native and two with R372A mutation), high resolution PepQxc crystal structures were determined. Also, the detailed enzyme assay were performed to investigate the substrate specificity of the PepQxc in presence of various Xaa-pro substrates and metal ions. In PepQxc structure, the *Tyr* residue which participate in hydrogen bond networking motif is structurally replaced by *Met* residue as shown in Fig. 4.16. Therefore, being a non-polar hydrophobic, the *Met* residue cannot participate in proton shuttling with *Arg372* and *Asp251* in PepQxc. Also, the R372A mutated structure of PepQxc showed that the overall structure of active site is intact and the *Arg372* residue is not critically required for metal center stability. The crystal structure of PepQxc1, PepQxcR372A, PepQxcR372A-Zn and PepQxc2 has different conformations of PHG loop, the PHG loop

is open for PepQxc1 and PepQxcR372A, the loop is closed for PepQxcR372A-Zn and PepQxc2 as shown in Fig. 4.10c. The closed conformation of PHG loop has a ligand bound at the active site, while in open loop conformation active site was free from ligand. Therefore, it is proposed that loop facilitate the binding and release of substrate and product during the catalysis, respectively. The *His234* residue has two different conformations in PepQxc2 structure, without any structural change in binding of a Citrate substrate at the active site. These conformations of the *His234*, suggest that residue is not essential for the substrate stabilization via hydrogen bonding. The PepQxc enzyme is strictly dipeptidase and based on crystal structure it was shown that longer than two residue substrate are difficult to bind at active site. The enzymatic assay performed on the PepQxc shows the activity in presence of various Xaa-Pro substrate. Also, the enzyme can be activated in presence of various metal ions (Fig. 4.13b). The Xaa-Pro substrate with branched C-beta first residue, like *Ile*, *Val* or *Pro* at binding site of enzyme, has the less activity because of the stereochemical hindrance created by branching near the metal ion at the active site. The substrate concentration dependent inhibitory effects were observed for the PepQxc enzyme as shown in Table. 4.4. These inhibitory effects were non-competitive and reversible in nature. The phosphate also has the concentration dependent inhibitory effect on the enzymatic activity of PepQxc, as shown in Fig. 4.14a. The reduction in enzyme activity in presence of phosphate has been explained by structure of PepQxc1-Mn and PepQxc1-Zn protein (Fig. 4.6). A phosphate molecule replaces the activated nucleophilic water molecule, which results into loss of enzymatic activity, as the activated water molecule is responsible for the Xaa-Pro peptide bond cleavage.

Table 4.6: Data collection statistics of PepQxc1, PepQxc2, PepQxcR372A, PepQxcR372A-Zn

Crystal	**PepQxc1**	**PepQxc1-Mn**	**PepQxc1-Zn**	**PepQxc2**	**PepQxcR372A**	**PepQxcR372A-Zn**
Data collection statistics[a]						
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a, b, c (Å)	84.32, 105.5, 111.4	81.96, 104.28, 112.32	82.03, 104.08, 112.41	58.18, 103.7, 136.1	83.01, 105.0, 111.9	81.09, 101.95, 111.51
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	47.68-1.83 (1.86-1.83)	47.29-1.85 (1.89-1.85)	47.22-1.95 (1.99-1.95)	48.44-2.20 (2.27-2.20)	47.53-1.80 (1.83-1.80)	46.36-1.85 (1.89-1.85)
Reflections measured	617282	560427	432574	519091	582206	539390
Unique reflections	88190	82603	70802	42175	90983	77891
Completeness (%)	99.9 (98.8)	99.8 (98.4)	100 (99.9)	98.9 (91.2)	99.8 (97.5)	98.0 (79.8)
AverageI/σ	18.2 (2.1)	14.0 (2.4)	13.6 (2.4)	43.6 (20.8)	25.6 (1.8)	20.1 (2.1)
$CC^b_{1/2}$	0.999 (0.716)	0.998 (0.704)	0.997 (0.722)	0.998 (0.995)	0.999 (0.608)	0.999 (0.800)
R^c_{merge}	0.068 (0.553)	0.102 (0.871)	0.107 (0.798)	0.051 (0.099)	0.055 (0.876)	0.073 (0.686)
R^d_{meas}	0.079 (0.720)	0.110 (0.942)	0.117 (0.872)	0.053 (0.106)	0.060 (0.985)	0.080 (0.759)
Multiplicity	7.0 (4.3)	6.8 (6.8)	6.1 (6.1)	12.3 (8.7)	6.4 (4.7)	6.9 (5.5)
Overall B factor from Wilson plot ($\mathring{A}^2$)	16.0	16.9	17.0	9.8	20.4	21.7
Refinement statistics						
Resolution (Å)	46.46-1.83	47.29-1.85	46.36-1.95	44.25-2.20	38.60-1.80	34.92-1.85
No. of Reflections	88025	82460	70658	40022	90848	77681
Completeness (%)	99.90	99.67	99.84	98.78	99.66	97.74
Total non-H atoms	6714	6548	6467	6610	6530	6537
Protein	6041	5969	5969	6048	5980	5982

Continued on next page

Table 4.6 – *Continued from previous page*

Crystal	**PepQxc1**	**PepQxc1-Mn**	**PepQxc1-Zn**	**PepQxc2**	**PepQxcR372A**	**PepQxcR372A-Zn**
Water	673	561	480	523	527	535
R^e_{cryst}	0.181	0.169	0.172	0.147	0.205	0.184
R_{free}	0.206	0.200	0.207	0.181	0.226	0.224
Ramachandran plot: Favored (%)	97.23	97.83	98.09	96.40	96.96	97.85
Ramachandran plot: Allowed (%)	2.77	2.17	1.91	3.47	2.78	2.02
Rmsd ideal bond length ($\mathring{A}$)	0.007	0.007	0.008	0.017	0.007	0.007
Rmsd ideal bond angles (°)	1.080	1.116	1.125	1.627	1.119	1.097
Average B factor	24.57	21.13	23.09	13.46	31.34	26.12

a Data for the highest-resolution shell are given in parentheses.

b $CC_{1/2}$ is defined as correlation coefficient between two random half data sets [105].

c $R_{merge} = \frac{\sum_{hkl}\sum_j |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$, where $I_{hkl,j}$ is the intensity of an individual reflection and $< I_{hkl} >$ is the mean intensity of multiple observations of symmetry-related reflections.

d $R_{meas} = \frac{\sum_{hkl}\sqrt{\frac{n}{n-1}}\sum_{j=1}^{n} |I_{hkl,j} - \langle I_{hkl}\rangle|}{\sum_{hkl}\sum_j I_{hkl,j}}$,

e $R_{cryst} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$; 4.95% randomly omitted reflections used for R_{free}.

Chapter 5

Human PSP94

The human MSMB gene encoding β-microseminoprotein has attracted much attention recently since the two genome-wide association studies have independently identified it to be linked with prostate cancer susceptibility ([14] and [106]), which were lately confirmed by other studies ([107] and [108]). The MSMB gene is expressed abundantly by prostate epithelial cells and the encoded protein is, therefore, also called the prostate secretory protein of 94 residues (PSP94). During the development of prostate cancer from early to late stages, expression of PSP94 progressively decreases [109]. The loss of PSP94 expression possibly contributes to the development of prostate cancer as PSP94 has been reported to suppress tumor growth and metastasis [11].

Emanating from the prostatic secretion, PSP94 forms an abundant constituent ($\sim 1.0\, mg/ml$) of the human semen. Although exact biological function of PSP94 remains unknown, several potential roles of PSP94 have been reported during the last three decades. *Kamada et al.* have shown that PSP94 binds to human immunoglobulin G (IgG) and referred to it as immunoglobulin binding factor [110]. It was suggested that a high amount of PSP94 present in the human semen might be involved in regulation of immune response in the female reproductive tract against the allogeneic sperm [110], [111] and [112]. PSP94 has also been shown to be a motility inhibitor of sperm [10] and an inhibitor of sperm acrosome reaction [41]. Recently it has been reported that PSP94 has a fungicidal activity in low-calcium environments [113]. Apart from reproductive tissues, PSP94 has also been detected in several nonreproductive tissues, such as the respiratory and gastric tract tissues, and in both males and females [8]. PSP94 homologues have also been identified in several other mammals [45], [46] and [47] and nonmammalian species [48], [49] and

[114]. There is a large degree of sequence diversity among these homologues, but a striking feature of PSP94 family of proteins is the presence of 10 highly conserved cysteine residues.

Efforts to understand the biological function of PSP94 led to the identification of cysteine-rich secretory proteins (CRISPs) as potential natural binding partners of PSP94. It has been shown that PSP94 binds to human CRISP-3 in seminal plasma [12] as well as PSP94 binding protein (PSPBP or CRISP-9) in blood [13]. A recent report has revealed that PSP94 from human as well as porcine species form high-affinity complexes even with evolutionary diverse CRISPs present in several snake venoms [115]. These snake venom CRISPs have been shown to be ion channel blockers [116], while the mammalian CRISPs are thought to be involved in sperm maturation, gamete fusion, and host defense [117]. The fact that PSP94 from human and porcine species vary in their amino acid sequences (51% identity) and both of them bind to CRISPs of diverse species like snakes suggests that the interaction between these two families of proteins may be quite general. It is likely that some of the functions of PSP94 might actually be orchestrated through its binding to CRISPs, but how PSP94 binds to a variety of CRISPs, including those from snake venoms, is not known.

PSP94 per se is a non-glycosylated protein synthesized as a precursor polypeptide of 114 residues with a cleavable signal peptide (residues $1 - 20$). The mature protein (94 residues) has a molecular mass of 10.7 kDa; however, it migrates as a 16- to 18-kDa band on SDS-PAGE. *Mori et al.* had speculated earlier that PSP94 is secreted as a homodimer that does not bind IgG, but these dimers get activated to monomeric form in female reproductive tract to interact with IgG [111]. However, these aspects have not been investigated further and the biologically relevant oligomeric state of PSP94 remains unknown. Several recent reports have, however, shown that PSP94 elutes as a single peak of ~ 21 kDa from gel-filtration columns corresponding to a possible homodimer [12], [115] and [118]. However, no dimers were observed in the NMR structures of PSP94.

The solution structure of PSP94 by NMR reported earlier by two laboratories [15] and [16] shows that PSP94 has two distinct domains bridged by a disulfide bond. Although the secondary structure of the individual domains reported by the two laboratories was similar, the relative orientation of the domains was very different ($\sim 90^\circ$ apart), giving rise to two distinct overall shapes (globular versus elongated) for this small protein. The discrepancy seems to be due to the different interpretations of the 10 specific nuclear Overhauser effects, which were probably weak. As there are no crystal structures known for any member of this family of proteins, the relative orientation of two domains of PSP94

could not be confirmed independently. We, therefore, crystallized PSP94 recently [119], and subsequently, *Kumar et al.* also crystallized this protein under a different crystallization condition and in a different space group [118]. Though the unit cell dimensions along the a- and b-axis reported by *Kumar et al.* are related to ours, it is different along the c-axis [118]. Here, the crystal structure of PSP94 has been determined by single isomorphous replacement with anomalous scattering (SIRAS) method using crystals soaked in uranyl nitrate. The overall structure of PSP94 reported here is similar to the NMR structure reported by *Ghasriani et al.*, but the relative orientation of the two domains is shifted by $\sim 20^{\circ}$ [16]. The reason for such a shift seems to have functional implications and is discussed. The present structure further gives insight into the possible mode of interaction of PSP94 with CRISP and IgG molecules.

5.1 Materials and Methods

5.1.1 Protein purification, crystallization, and data collection

PSP94 was purified from human seminal plasma using a well-established protocol [112]. Crystals were grown by hanging drop vapor diffusion method [119]. Briefly, drops were prepared by mixing 2 μl of protein solution (10 mg/ml in water) with 2 μl reservoir solution [0.1 M sodium acetate, pH 4.5, 0.2 M lithium sulfate, and $44-47\%$ (v/v) PEG 400] and 0.4 μl additive solution (γ-butyrolactone) and equilibrated against 1 ml reservoir solution at 293 K in a sealed well. The crystallization buffer, which contained $\sim 45\%$ (v/v) PEG 400, was directly used for flash freezing the crystals in liquid nitrogen without the addition of any further cryoprotectant solution. Heavy-atom derivatives were prepared by soaking PSP94 crystals overnight in crystallization buffer containing 1 mM uranyl nitrate $[UO_2(NO_3)_26H_2O]$. The soaked crystals were backsoaked in crystallization buffer lacking uranium before freezing. X-ray diffraction data on native crystals as well as uranium derivatives were collected at X06DA (PXIII) beamline of the Swiss Light Source, Paul Scherrer Institut, Villigen, Switzerland. Diffraction data were processed using MOSFLM [120] and SCALA [81] or XDS [79].

5.1.2 Structure determination

The crystal structure of PSP94 was determined using SIRAS using the uranium derivative. Table. 5.1 contains data collection and refinement statistics. The program SOLVE

as implemented in the PHENIX suite [87] was used to locate six U atoms, refine their positions, and calculate experimental phases. Phase extension, density modification, and automatic tracing of the polypeptide chain were carried out in RESOLVE/PHENIX. Almost 80% of the residues in four polypeptide chains were built by RESOLVE. Further model buildings were carried out manually in Coot [88] and O [121] using the density-modified map produced by RESOLVE. Few more residues were added during iterative rounds of refinement in CNS [122] followed by rebuilding in O. Finally, the refinement converged to R_{work} and R_{free} of 21.3% and 26.6%, respectively. Illustrations were made in PyMOL (DeLano Scientific).

5.2 Results

PSP94 was purified from human seminal plasma and crystallized in tetragonal space group $P4_12_12$ with unit cell dimensions $a = b = 107.9$ $\mathring{A}$ and $c = 82.1$ $\mathring{A}$ [119]. There are four molecules per asymmetric unit. PSP94 structure was determined from diffraction data extending to 2.3 $\mathring{A}$ resolution. Experimental phases were obtained from SIRAS using crystals soaked in uranyl nitrate. Experimental electron density maps, improved by density modification, allowed unambiguous placement of most of the residues. The final, refined atomic model contains 369 out of 376 amino acid residues in four polypeptide chains of PSP94. The missing region is a flexible loop (amino acids $10 - 16$) in chain A. The atomic coordinates have been deposited in the Protein Data Bank (PDB) with the accession code **3IX0**.

5.2.1 Overall structure of PSP94 monomer

PSP94 monomer has a long extended structure, rich in β-sheets (Fig. 5.1). There are two distinct domains: an N-terminal domain from residues $1 - 52$ and a C-terminal domain from residues $53 - 94$. The two domains are held together by a disulfide bond between *Cys37* and *Cys73*. The N-terminal domain has four antiparallel β-strands (β1: residues $1 - 6$, β4: residues $30 - 34$, β5: residues $38 - 42$, and β6: residues $46 - 51$) arranged in the form of Greek key motif and two small antiparallel β-strands (β2: residues $18 - 21$ and β3: residues $23 - 27$) forming a flap on top of the Greek key motif. There are three disulfide bonds in the N-terminal domain. The disulfides *Cys2-Cys50* between strands β1 and β6 and *Cys40-Cys49* between strands β5 and β6 make the Greek key structure rigid while the third disulfide *Cys18-Cys42* between β2 and β5 orients the flap onto the

Greek key motif. The C-terminal domain has two double-stranded antiparallel β-sheets. The strand $\beta 7$ (residues $55-58$), which is an extension of the N-terminal strand $\beta 6$, is associated with the C-terminal β-strand $\beta 10$ (residues $90-94$). Interestingly, $\beta 10$ seems to get extended further by the first strand $\beta 1$ of the N-terminal domain. The other two β-strands in the C-terminal domain, $\beta 8$ (residues $64-70$) and $\beta 9$ (residues $74-79$), are longer and separated from the first pair. The arrangement of these four β-strands in the C-terminal domain is unusual and gives rise to a unique fold. The lone disulfide *Cys64-Cys87* in the C-terminal domain brings rigidity to the loop structure. One of the remarkable features of PSP94 monomer is the close proximity of the amino and carboxyl ends of the polypeptide chain, which are facing each other with a strong hydrogen bond between the main-chain nitrogen atom of the first residue and one of the carboxyl oxygen atoms of the last residue (N-O distances, $2.45-2.66$ $\AA$). The significance of this hydrogen bond is discussed in the next section.

All the four molecules in the asymmetric unit are structurally very similar. The superposition of their C_α atoms shows that the polypeptide chain from residues $1-6$ and $32-94$ in the four molecules follow the same course whereas the polypeptide chain from residues $7-31$ in one of the molecule (chain D) follows a slightly different course. The peptide segment $8-17$ contains mostly hydrophilic residues and is flexible with poor electron density. The pairwise root-mean-square deviation (RMSD) between the four molecules in the asymmetric unit is $0.6-0.8$ $\AA$ when C_α atoms of the peptide segments $1-6$ and $18-94$ were used in the superposition by least squares.

5.2.2 Structural similarity search

The structural similarity search for PSP94 shows that the C-terminal domain has no structural similarity with any known proteins, but residues 15-52 in the N-terminal domain have a fold similar to the fibronectin type I (FnI) module (Fig. 5.1b). The RMSD of these aligned residues between PSP94 and different FnI structures are $\sim 2.1-3.0$ $\AA$. The structure-based sequence alignment further shows that the four conserved cysteine residues in the FnI structures, which make two disulfide bonds and hold the β-strands together, are present at the corresponding positions even in the PSP94 structure (Fig. 5.1c). PSP94, however, has two extra cysteine residues (*Cys37* and *Cys50*) in this region of the sequence, which make additional disulfide bonds. *Cys37* makes disulfide bond with *Cys73* located in the C-terminal domain whereas *Cys50* makes disulfide bond with *Cys2* located on the first strand. Further, although the first strand, $\beta 1$, of the Greek key motif

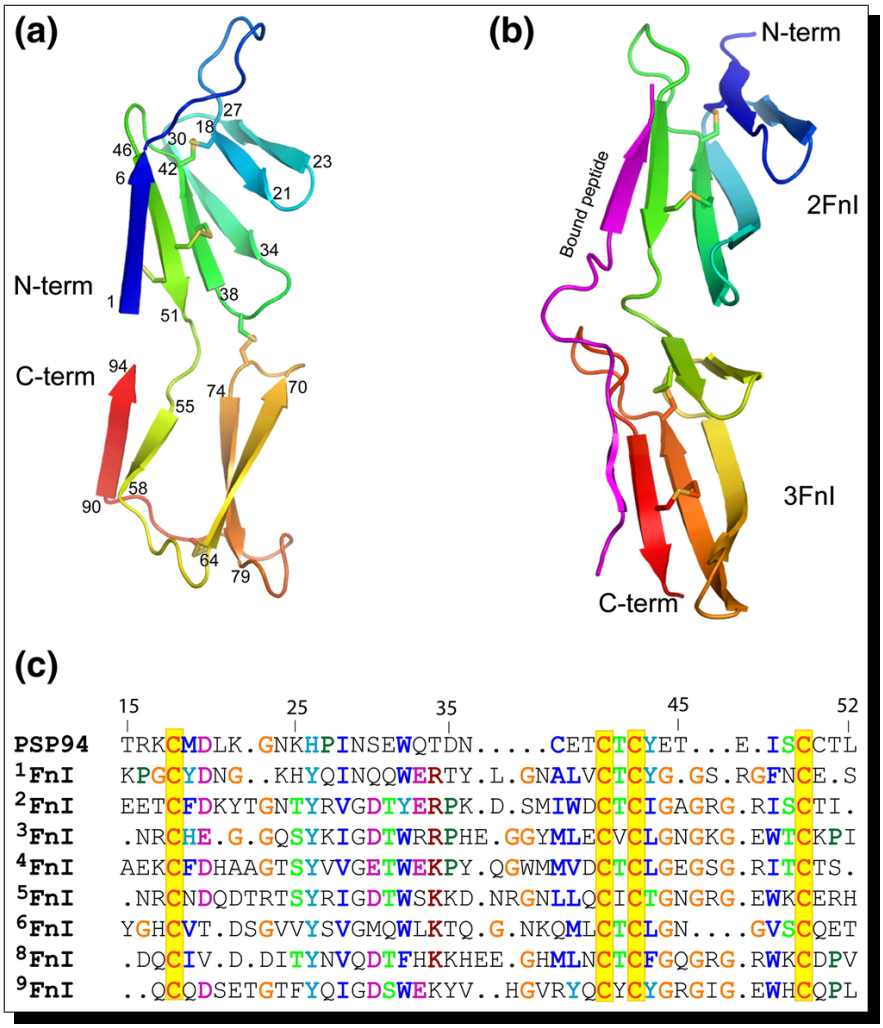

Figure 5.1: PSP94 monomer and its comparison with FnI modules. (a) Cartoon representation of PSP94 with polypeptide chain in rainbow color from blue (N-terminus) to red (C-terminus). Five disulfides are shown in stick representation. (b) Cartoon representation of fibronectin module pair (^{2}FnI ^{3}FnI from PDB ID 2rkz) with polypeptide chain in rainbow color and a bound peptide, making an antiparallel β -strand, in magenta color. (c) Structure-based sequence alignment of PSP94 with different fibronectin FnI modules. The PDB files used for structural superposition of different fibronectin modules are as follows: 1o9a for ^{1}FnI, 3cal for ^{2}FnI and ^{3}FnI, 2rky for ^{4}FnI and ^{5}FnI, 1e8b for ^{6}FnI, and 3ejh for ^{8}FnI and ^{9}FnI. The completely conserved cysteine residues making disulfides are highlighted in yellow, and the other partially conserved residues are shown in different colors. The residue numbering indicated at the top is with respect to PSP94 sequence.

of PSP94 is absent in the native FnI structures, comparison with the FnI bound to a cognate peptide either from a fibronectin binding protein of a pathogenic bacterium [123] and [124] or from the α1 chain of type I collagen [125] is noteworthy. The bound peptide,

which forms an additional antiparallel β-strand with the FnI module pair, compares well with the first strand, β1, of PSP94 (Fig. 5.1b).

5.2.3 PSP94 dimer

The arrangement of four molecules of PSP94 in the crystallographic asymmetric unit suggests that they form two dimers. These two dimers are related to each other by 1/2 unit cell translation along the z-axis. Within a dimer, there is a non-crystallographic dyad axis perpendicular to the z-axis that orients the N-terminal domain of one monomer adjacent to the C-terminal domain of the other monomer (Fig. 5.2a). In this arrangement, the β-sheet of the N-terminal Greek key motif in each monomer gets extended across the dimer interface by the C-terminal strands β10 and β7 of the adjacent monomer. This leads to the formation of two six-stranded β-sheets in the PSP94 dimer (Fig. 5.2b). Because the dimer is formed by edge-to-edge interaction of the two monomers, the surface area buried between them (960 $\mathring{A}^2$) is relatively less. The dimeric association is stabilized by a number of interactions that include eight interchain backbone hydrogen bonds, a hydrogen bond between O^{γ} atom of *Ser1* from one of the monomer and the terminal carboxyl oxygen atom (O2) of *Ile94* from the other monomer, and 66 other interchain contacts shorter than 4.0 $\mathring{A}$. The dimeric interaction further brings the two N- and two C-terminal ends of the monomers in close proximity to each other (Fig. 5.2 and Fig. 5.3a).

5.2.4 pH-driven monomer/dimer transition

To test whether PSP94 dimer is a manifestation of crystal packing or whether it exists even in solution, a gel-filtration chromatography has been performed. PSP94 eluted as a single peak corresponding to a molecular mass of $\sim$ 21 kDa (dimer) when elution was carried out with 50 mM Tris at pH 8.0. This was consistent with several independent reports published earlier [12], [115] and [118]. However, when elution was carried out at the pH of crystallization with 50 mM sodium acetate, pH 4.5, PSP94 eluted as a single peak corresponding to a molecular mass of $\sim$ 10 kDa (monomer) (Fig. 5.2c). The transition of PSP94 from dimer to monomer seems to occur by subtle changes in the solution environment. The fact that a dimer has been observed in the crystal seems to suggest that either the precipitant [45% polyethylene glycol (PEG) 400] in the crystallization buffer or the crystalline environment itself might have coaxed the monomers to associate in the form of native dimers while packing the molecules in the crystal.

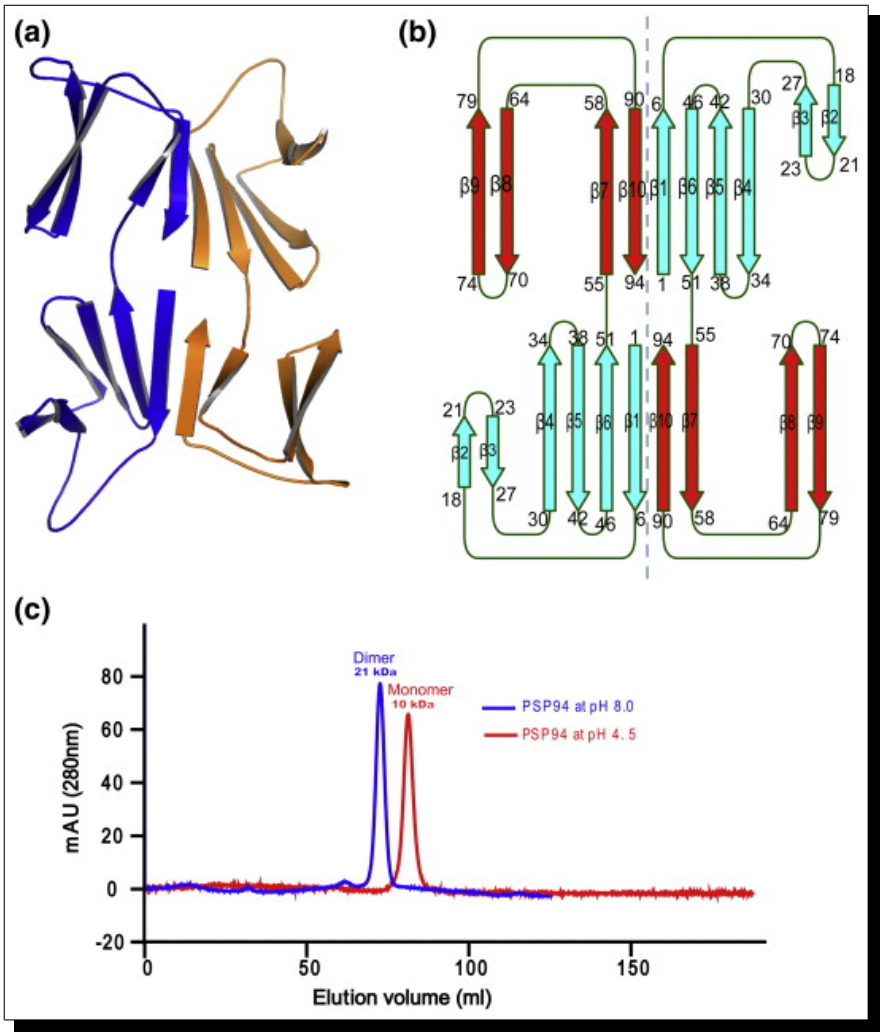

Figure 5.2: PSP94 dimer. (a) Cartoon representation showing the edge-to-edge interaction of two PSP94 monomers with their polypeptide chains in blue and orange colors. (b) Topology diagram of dimeric interaction showing extension of β-sheet across the dimer interface (broken vertical line). The β-strands in the N-terminal domain are shown in cyan, and those in the C-terminal domain are shown in red. (c) Elution profile of monomeric and dimeric PSP94. Purified PSP94 ($400-500\ \mu g$) was subjected to gel-filtration chromatography using a Superdex G-75 column (Hi-Load $16/60$, Amersham) at a flow rate of $0.5\ ml/min$. The elution buffers used were as follows: $50\ mM$ Tris, $pH\ 8.0$ (in blue), and $50\ mM$ acetate, $pH\ 4.5$ (in red).

5.3 Discussion

5.3.1 Mechanism of monomer/dimer transition

The edge-to-edge interaction of β-sheets from two different polypeptide chains is an important mode of protein-protein interaction and has been observed in several protein

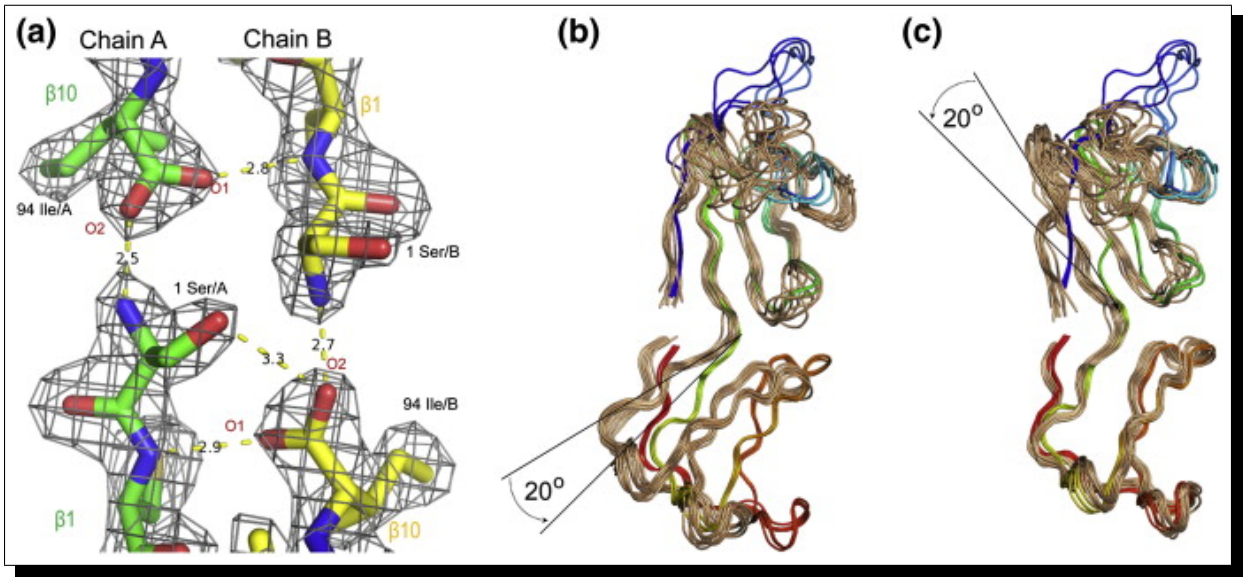

Figure 5.3: PSP94-dimer to monomer. (a) The stick representation of the terminal residues from two different polypeptide chains (with green and yellow carbons) at the dimeric interface is shown along with their electron density maps ($2F_O - F_C$) contoured at 1.5σ. Oxygen and nitrogen atoms are colored red and blue, respectively. The hydrogen-bond distances are marked in angstroms. Superposition of either (b) the C-terminal domain or (c) the N-terminal domain of the NMR models (from PDB file 2iz3 and shown in brown) onto the corresponding domain of the crystal structure (four polypeptide chains in rainbow color) showing a shift in the relative orientation of two domains.

structures either between the two identical polypeptide chains resulting in the formation of homodimers or between two different polypeptide chains as in the heterodimer complexes ([126] and [127]). The crystal structure reported here shows a rather unique edge-to-edge interaction of two PSP94 monomers. This "edge-on" interaction is facilitated by the formation of an almost straight edge in PSP94 monomer containing two terminal strands, $\beta 1$ and $\beta 10$. These two strands, which terminate in the amino and the carboxyl ends of the polypeptide chain facing each other, are held in place mainly due to a strong hydrogen bond between the main-chain nitrogen atom of the first residue, *Ser1*, and one of the oxygen atoms (O2) of the terminal carboxyl group of the last residue, *Ile94* (Fig. 5.3a). The environment around the terminal carboxyl group suggests that it is ionized in the present structure. The ionized terminal carboxyl group seems to be essential for holding the two strands together and providing a straight edge for the dimeric interactions. However, at the acidic pH, close to the pKa (~ 4.0) of free carboxyl group, one of the oxygen atoms (O2) of the terminal carboxyl group can get protonated and as the protonated hydroxyl (-OH) oxygen of the carboxyl group cannot accept a hydrogen bond

from a potential donor group (NH_3^+ of *Ser1*) nearby, the hydrogen-bonding interaction holding the amino and the carboxyl ends of the polypeptide chain will be broken. The terminal residues (*Ser1* and *Ile94*) would then move away from each other to avoid close contacts resulting in distortion of the straight edge, which would lead to disruption of the β-sheet interactions at the dimer interface and subsequent dissociation of the dimer. Interestingly, in the solution structure of PSP94 by NMR [16], a similar movement of the terminal residues is seen where the main-chain nitrogen atom of *Ser1* and the carboxyl oxygen atom of *Ile94* in different models are far apart ($6.0 - 8.9$ $\mathring{A}$). Thus, the NMR structure, wherein the samples were prepared in water at pH 6.0 without any buffer, would represent the monomeric form of PSP94. In the present structure, the crystalline environment seems to stabilize the dimeric form even at lower pH. Further analysis shows that the relative orientation of the two domains of PSP94 in the present dimeric structure is slightly different from that in the NMR structure (PDB ID 2IZ3). When all the C_α atoms of one of the domains in the NMR structure are superposed on the corresponding atoms in the present structure, the other domain in the NMR structure appears shifted as compared to the present structure (Fig. 5.3b and 5.3c) and the shift is of a pure rotation of $\sim 20^\circ$ around an axis located between the two domains. The domain shift seems to be needed to avoid close contacts between the amino and carboxyl ends of the polypeptide chain without altering the β-sheet structure in the individual domains. Due to the shift in the orientation of two domains, the strands $\beta 1$ and $\beta 10$ appear twisted in the NMR structure and no longer form a straight edge that is needed to make the dimeric interaction.

5.3.2 PSP94-CRISP interaction

PSP94 and CRISP families of proteins are present in several organisms, and the fact that there is a high-affinity binding between them ([12], [13] and [115]) suggests that these interactions must be of considerable physiological relevance and some of the functions of PSP94 might actually be mediated through its interaction with CRISPs. PSP94 has a partial but striking structural similarity with FnI modules with conserved disulfides as described in the previous section. Different proteins are thought to bind fibronectins by adding a β-strand in antiparallel fashion to the existing β-sheet structure of fibronectin modules [124]. Whether different CRISPs would also bind to the PSP94 family of proteins in a similar fashion remains a distinct possibility. Interaction of PSP94 with human CRISP-3 has recently been studied by NMR [128]. It was found that the residues in PSP94

that get affected upon complex formation are located on the two terminal β-strands (β1 and β10) as well as two strands (β6 and β7) that are associated with these terminal strands. Based on this observation, the authors proposed a model wherein the first β-strand (β1) of PSP94 binds to an accessible β-strand of CRISP-3, while the last β-strand (β10) of PSP94 binds to a different small β-strand in CRISP-3, making two parallel -sheet interactions at the PSP94–CRISP-3 interface. However, our analysis shows that such an interaction make a knot between the polypeptide chains of PSP94 and CRISP-3. In the crystal structure reported here, it is found that the terminal strands (β1 and β10) of PSP94 are involved in the formation of a dimer with antiparallel association of β-strands at the protein–protein interface. Since these terminal strands get affected upon PSP94–CRISP-3 complex formation, it is possible that CRISP-3 would replace one of the PSP94 molecules from the dimer and form a PSP94–CRISP-3 heterodimer complex. The 1 : 1 stoichiometry in the PSP94–CRISP complex has also been suggested recently in the binding study of porcine PSP94 with different snake venom CRISPs [115]. We, therefore, attempted to model PSP94–CRISP interaction in such a way that the terminal β-strands, β1 and β10, of PSP94 interact with a single accessible β-strand of CRISP-3 in an antiparallel fashion (Fig. 5.4a), similar to the protein-protein interaction observed in the formation of the PSP94 dimer. In this model, a part of the sperm-coating protein domain in CRISP wraps around one side of PSP94 to increase the intermolecular interaction, which would result in the formation of a higher affinity PSP94–CRISP complex than the formation of the PSP94–PSP94 homodimer, consistent with the experimental observation reported earlier [12]. Thus, the straight edge of PSP94 containing two terminal β-strands seems to be 'sticky' as it is involved in binding to different CRISPs, and when the binding partners are not present, two of them associate to form a PSP94 dimer. Further, the interaction between the secondary structural elements of PSP94 and CRISPs, as proposed here, would be tolerant to sequence variations among members in these two families of proteins and thus would provide an ingenious solution in maintaining high affinity between proteins from these two families.

5.3.3 PSP94-IgG interaction

PSP94 binds to different human IgG [110] and it has been speculated that PSP94 may be involved in regulation of immune response in the female reproductive tract [111]. The binding of PSP94 to human IgG remains unaffected even when all the five disulfides in PSP94 are reduced, suggesting the involvement of sequential epitopes of PSP94 in

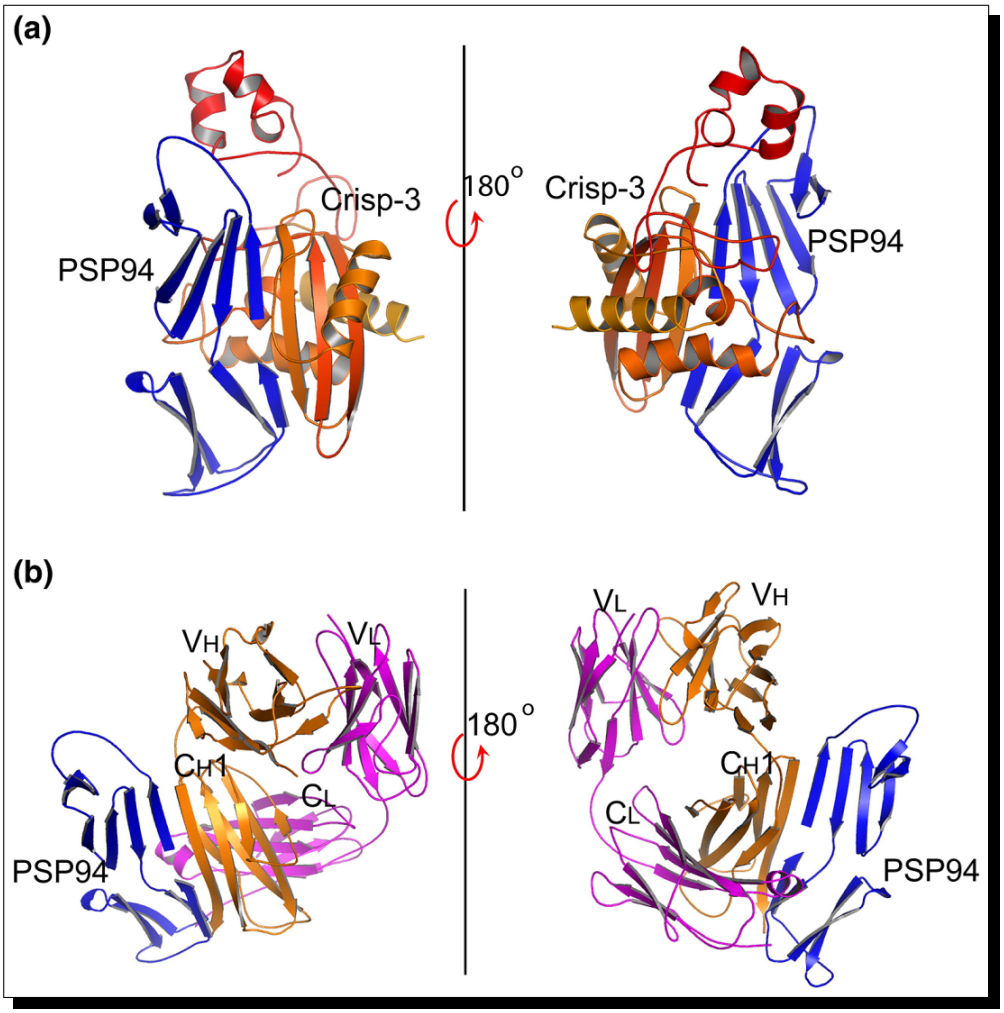

Figure 5.4: PSP94 binding partners. (a) A model for interaction of PSP94 (blue) with CRISP-3 (orange). CRISP-3 model was built by homology modeling based on triflin structure. Right panel shows another view rotated by 180° around a vertical axis as shown. (b) A model for interaction of PSP94 (blue) with Fab domain of *IgG*. The Fab model was taken from PDB file 1igc. The variable and constant domains of the heavy (in orange) and the light (in pink) chains of Fab are labeled. Right panel shows another view rotated by 180° around a vertical axis as shown.

IgG binding [112]. Further, our recent results suggest that PSP94 binds to Fab as well as Fab_2 domains and not to the Fc domain of human IgG [129]. Interestingly, similar binding characteristics have been reported in binding of protein G, a cell surface protein from *Streptococcus*, with IgG [130]. Crystal structure of protein G with Fab fragment shows that the outer β-strand in protein G forms an antiparallel interaction with the last β-strand in the constant heavy-chain domain (CH_1) of IgG, leading to an extension of β-sheet across the binding interface. It is likely that PSP94 would also bind to IgG in a similar fashion with its 'sticky' edge, containing two terminal β-strands $\beta 1$ and $\beta 10$,

interacting with the last β-strand in the CH_1 domain of Fab in an antiparallel fashion (Fig. 5.4b). Such interaction between the secondary structural elements of PSP94 and the CH_1 domain would again be tolerant to variations among different IgG molecules.

5.4 Summary

PSP94 has been attracting interest for the last three decades due to its abundant presence in human semen. Several biological roles have been proposed since then [48], but the exact function of PSP94 still remains elusive. The overall structure of PSP94 reported here is similar to the NMR structure reported by Ghasriani et al., but the relative orientation of the two domains is shifted by $\sim 20^\circ$ [16]. The crystal structure reported here shows how the edges from two PSP94 monomers associate to form a dimer. Further, the present results shows for the first time a pH-induced transition of PSP94 from dimeric to monomeric form. This assumes biological significance as PSP94 would function as a dimer in the human semen (pH $7.5 - 8.0$), whereas the dimers may dissociate into 'active' monomers when it comes to vagina (pH $3.8 - 5.0$). The concept of inactive dimers and the active monomers, in the context of PSP94 binding to IgG [111], was proposed earlier, but the role of vaginal pH in 'activating' PSP94 had not been investigated there. Further, based on several observations, it is being proposed that the interaction of PSP94 with several CRISPs and IgG molecules may also occur through the N- and C-terminal β-strands of PSP94. The present report thus provides valuable insight into the structure of PSP94 as well as its dimeric form and its possible mode of interaction with different binding proteins, which, in turn, may provide useful clues to the biological functions of this important family of proteins.

Table 5.1: PSP94 Data collection and refinement statistics

Crystal	**Native**	**U derivative**
Data collection statistics[a]		
Space group	$P4_12_12$	$P4_12_12$
Cell dimensions		
a, b, c (Å)	107.88, 107.88, 92.13	107.90, 107.90, 91.76
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	$50.0 - 2.3$ $(2.42 - 2.30)$	$39.31 - 2.50$ $(2.64 - 2.50)$
Reflections measured	$198,419$ $(28,313)$	$137,140$ $(19,260)$
Unique reflections	$24,778$ (3528)	$19,314$ (2767)
Completeness (%)	100 (100)	99.9 (99.8)
Average I/σ	16.1 (4.1)	13.3 (3.6)
R^b_{merge}	0.104 (0.460)	0.114 (0.538)
Multiplicity	8.0 (8.0)	7.1 (7.0)
Refinement statistics		
Resolution (Å)	$39.43 - 2.30$	
No. of reflections	$23,955$	
Completeness (%)	96.87	
Total non-H atoms	3176	
Protein	2938	
Water	238	
R_{cryst}/R_{free} [c]	0.213/0.266	
Residues in favored regions of the Ramachandran plot (%)	96.7	
Residues in allowed regions of the Ramachandran plot (%)	3.0	
RMSD ideal bond length (Å)	0.007	
RMSD ideal bond angles (°)	1.0	

a Data for the highest-resolution shell are given in parentheses.
b $R_{merge} = \sum |I_i - < I > | / \sum I_i$, where I_i is the intensity of an individual reflection and $< I >$ is the mean intensity of multiple observations of symmetry-related reflections.
c $R_{cryst} = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$; 6.2% randomly omitted reflections used for R_{free}.

Chapter 6

Conclusions

The title of thesis "Design and development of protein crystallography beamline at Indus-2 synchrotron and structural studies of a bacterial prolidase and human PSP94" summarizes the work of the thesis. The protein crystallography beamline (PX-BL21) at Indian synchrotron (Indus-2) has been made operational and is available to users now. The beamline is a national facility and the Department of Atomic Energy (DAE), Government of India, funded the construction and operation of this beamline. The PX-BL21 is primarily aimed to cater the needs of about 100 independent research groups in India working in the area of structural biology. The first diffraction data sets on single crystals of Lysozyme protein was recorded during mid-2012 and since then the beamline has been used by several users. Apart from beamline developments, structural studies were also carried out on two proteins. The structural studies on human PSP94 protein has been carried out at SSPD BARC, under the guidance of Dr. Mukesh Kumar. While, the structural studies Xaa-Pro di-peptidase (PepQxc) from *Xanthomonas Campestris* has been completed at auxiliary biochemical facility of PX beamline at RRCAT Indore, under the guidance of Dr. Ravindra D. Makde. The main results of the thesis are as follows:

- *The PX-BL21 has been designed keeping the option of future upgradation to Insertion Device source from the current* $1.5T$ *BM source. Ray tracing and heat load calculations were performed to optimize the optical and thermal design of beamline.*

- *Various beamline configurations were explored and CM-DCM-TM combination with TM in* $2:1$ *configuration was chosen for the PX-BL21. The positions for X-ray mirrors and DCM and their optical specifications were optimized using ray tracing*

calculations. Since beamline would be shifted to Wiggler source, the heat load calculation were performed on CM and DCM. The heat load calculations using Wiggler as a source have shown that liquid nitrogen cooling will be required for the DCM. While, the water cooling will be adequate for X-ray mirrors of the beamline.

- *Many beamline components were designed and fabricated indigenously. These beamline components are Beam position monitors (4 No.s), one primary X-ray slit, two no.s of water cooled compact slits, X-ray pipes, Mirrors chambers (2 no.s) and Hexapods (2 no.s).*

- *All beamline components were installed at optical line of BL21 port at Indus-2. Installation was followed by leak testing of individual beamline components to check the UHV compatibility. Subsequently, shielding hutch was installed to carry out hot beamline alignment.*

- *Hot alignment of beamline components in X-rays were performed using BPMs. Both mirrors were aligned with micron accuracy at* 3 *mrad glancing angle in X-rays. Software for calibration of DCM energy and absorption edge scanning using fluorescence detector has been developed and deployed. First data were collected on lysozyme crystals with good statistics showing the viability of beamline.*

- *The X-ray diffraction data on several PepQXc crystals were collected on the beamline. The crystal structure of native and R372A mutated PepQxc showed that a conserved residue Tyr which participate in the hydrogen bond networking motif in other prolidases is replaced by Met residue in PepQXc. Interestingly, Met being a non-polar hydrophobic residue cannot participate in proton shuttling with the Arg372 and the Asp251 in the PepQxc enzyme. Also, the R372A mutated structure of the PepQxc showed that the overall structure of active site remains intact and the Arg372 residue is not critically required for metal-center stability.*

- *The crystal structure of PepQxc1, PepQxcR372A, PepQxcR372A-Zn and PepQxc2 has different loop (residues from* 233–235*) conformations, loop is open for PepQxc1 and PepQxcR372A, loop is closed for PepQxcR372A-Zn and PepQxc2. The closed loop conformation has a ligand bound at the active site, while, in open loop conformation active site was free from ligand. Therefore, it was proposed that loop facilitate the binding and release of substrate and product during catalysis, respectively.*

- *The His234 residue in PepQxc2 structure was found in two different conformations without any structural change in binding of Citrate substrate at the active site. Based on these conformations of His234, it was concluded that residue was not essential for substrate stabilization via hydrogen bonding.*

- *The phosphate has concentration dependent inhibitory effect on the enzymatic activity of the PepQxc protein. The loss of enzymatic activity in presence of phosphate has been explained by the structure of PepQxc1-Mn and PepQxc1-Zn. A phosphate molecule replaces the activated nucleophilic water molecule, which results in loss of enzymatic activity by the PepQxc, as the activated water molecule is responsibles for the Xaa-Pro peptide bond cleavage.*

- *The crystal structure of PSP94 showed how the edges from two PSP94 monomers associate to form a dimer. Further, the present results showed for the first time a pH-induced transition of PSP94 from dimeric to monomeric form. This assumes biological significance as PSP94 would function as a dimer in the human semen (pH* $7.5 - 8.0$*)*

- *It is being proposed that the interaction of PSP94 with several CRISPs and IgG molecules may also occur through the N- and C-terminal* β*-strands of PSP94. The present report thus provides valuable insight into the structure of PSP94 as well as its dimeric form and its possible mode of interaction with different binding proteins, which in turn may provide useful clues to the biological functions of this important family of proteins.*

The ray tracing calculations show an improvement of two orders of magnitude in photon flux when the source is upgraded from 1.5 T BM to the superconducting MPW. Therefore, the beamline has been designed to accommodate such future upgradation with minimal changes in beamline components. The structural studies on a prolidase have been carried out on the newly commissioned beamline at Indus-2. Although crystal structure of several variants of the protein has been determined, the origin of proline specificity of the Xaa-Pro prolidase is not very clear from these structures and further studies need to be carried out. The crystal structure of human PSP94 reported here is the only crystal structure reported for this ubiquitous family of proteins which are present across many species. The crystal structure of PSP94 shows that the two PSP94 protomers interact edge-on to form a PSP94 dimer. It was found that these dimers dissociate at lower pH and this finding has significant biological implications. The biological significance arises as PSP94

would function as a dimer in the human semen (pH $7.5 - 8.0$), whereas the dimers may dissociate into 'active' monomers when it comes to vagina (pH $3.8 - 5.0$). Based on crystal structure of PSP94, a model has been proposed for its interaction with CRISP and IgG molecules, however, determination of crystal structures of these complexes will be an important milestone in understanding the biological functions of this important family of proteins.

Bibliography

[1] A. D. Ghodke, R. Husain, G. Singh, et al. "Progress in Commissioning of Indus-2". *Beam Dynamics Newsletter*, (41), 77 – 95 (2006).

[2] C. Theriot, S. Tove, and A. Grunden. "Chapter 3 Biotechnological Applications of Recombinant Microbial Prolidases". *Advances in Applied Microbiology*, **68**, 99–132 (2009).

[3] P. Ledoux, C. R. Scriver, and P. Hechtman. "Expression and molecular analysis of mutations in prolidase deficiency." *American journal of human genetics*, **59**(5), 1035 (1996).

[4] F. Endo, I. Matsuda, A. Ogata, et al. "Human erythrocyte prolidase and prolidase deficiency". *Pediatric Research*, **16**(3), 227–231 (1982).

[5] T.-C. Cheng, S. Harvey, and A. Stroup. "Purification and properties of a highly active organophosphorus acid anhydrolase from Alteromonas undina". *Applied and Environmental Microbiology*, **59**(9), 3138–3140 (1993).

[6] S.-C. Jao, L.-F. Huang, S.-M. Hwang, et al. "Tyrosine 387 and arginine 404 are critical in the hydrolytic mechanism of Escherichia coli aminopeptidase P". *Biochemistry*, **45**(6), 1547–1553 (2006).

[7] H. Lilja and P.-A. Abrahamsson. "Three predominant proteins secreted by the human prostate gland". *Prostate*, **12**(1), 29–38 (1988).

[8] M. Baijal-Gupta, M. Clarke, M. Finkelman, et al. "Prostatic secretory protein (PSP94) expression in human female reproductive tissues, breast and in endometrial cancer cell lines". *Journal of Endocrinology*, **165**(2), 425–433 (2000).

[9] Z. Liang, M. Kamada, S. Mitsudo, et al. "Immunoglobulin binding factor of seminal plasma: A secretory product of human prostate". *Archives of Andrology*, **28**(3), 159–164 (1992).

[10] C.-F. Chao, S.-T. Chiou, H. Jeng, et al. "The porcine sperm motility inhibitor is identical to β-microseminoprotein and is a competitive inhibitor of Na+,K+-ATPase". *Biochemical and Biophysical Research Communications*, **218**(2), 623–628 (1996).

[11] N. Shukeir, A. Arakelian, S. Kadhim, et al. "Prostate secretory protein PSP-94 decreases tumor growth and hypercalcemia of malignancy in a syngenic in Vivo model of prostate cancer". *Cancer Research*, **63**(9), 2072–2078 (2003).

[12] L. Udby, . Lundwall, A. Johnsen, et al. "β-Microseminoprotein binds CRISP-3 in human seminal plasma". *Biochemical and Biophysical Research Communications*, **333**(2), 555–561 (2005).

[13] J. Reeves, J. Xuan, K. Arfanis, et al. "Identification, purification and characterization of a novel human blood protein with binding affinity for prostate secretory protein of 94 amino acids". *Biochemical Journal*, **385**(1), 105–114 (2005).

[14] R. Eeles, Z. Kote-Jarai, G. Giles, et al. "Multiple newly identified loci associated with prostate cancer susceptibility". *Nature Genetics*, **40**(3), 316–321 (2008).

[15] I. Wang, Y.-C. Lou, K.-P. Wu, et al. "Novel solution structure of porcine β-microseminoprotein". *Journal of Molecular Biology*, **346**(4), 1071–1082 (2005).

[16] H. Ghasriani, K. Teilum, Y. Johnsson, et al. "Solution Structures of Human and Porcine β-Microseminoprotein". *Journal of Molecular Biology*, **362**(3), 502–515 (2006).

[17] N. Marks. "Synchrotron radiation sources". *Radiation Physics and Chemistry*, **45**(3), 315–331 (1995).

[18] H. Winick and S. Doniach. "An overview of synchrotron radiation research". pp. 1–10 (1980).

[19] D. J. Griffiths and R. College. *Introduction to electrodynamics*, vol. 3 (prentice Hall Upper Saddle River, NJ, 1999).

[20] S. Mobilio and A. Balerna. "Introduction to the main properties of Synchrotron Radiation". In "CONFERENCE PROCEEDINGS-ITALIAN PHYSICAL SOCIETY", vol. 82, pp. 1–24 (Editrice Compositori; 1999, 2003).

[21] G. Margaritondo. "A primer in synchrotron radiation: everything you wanted to know about SEX (Synchrotron Emission of X-rays) but were afraid to ask". *Journal of synchrotron radiation*, **2**(3), 148–154 (1995).

[22] D. Mills, J. Helliwell, Å. Kvick, et al. "Report of the Working Group on Synchrotron Radiation Nomenclature–brightness, spectral brightness or brilliance?" *Journal of synchrotron radiation*, **12**(3), 385–385 (2005).

[23] Z. Dauter, M. Jaskolski, and A. Wlodawer. "Impact of synchrotron radiation on macromolecular crystallography: a personal view". *Journal of synchrotron radiation*, **17**(4), 433–444 (2010).

[24] B. Schwanhäusser, D. Busse, N. Li, et al. "Global quantification of mammalian gene expression control". *Nature*, **473**(7347), 337–342 (2011).

[25] W. Kühlbrandt. "The resolution revolution". *Science*, **343**(6178), 1443–1444 (2014).

[26] K. Wüthrich. "The way to NMR structures of proteins". *Nature Structural & Molecular Biology*, **8**(11), 923–925 (2001).

[27] M. G. Rossmann and D. M. Blow. "The detection of sub-units within the crystallographic asymmetric unit". *Acta Crystallographica*, **15**(1), 24–31 (1962).

[28] D. Blow and M. Rossmann. "The single isomorphous replacement method". *Acta Crystallographica*, **14**(11), 1195–1202 (1961).

[29] W. Hendrickson. "Determination of macromolecular structures from anomalous diffraction of synchrotron radiation". *Science*, **254**(5028), 51–58 (1991).

[30] B. Matthews. *Heavy-atom location and phase determination with single-wavelength diffraction data* (Springer, 2001).

[31] J. M. Bijvoet. "Structure of optically active compounds in the solid state". *Nature (London)*, **173**, 888 – 891 (1954).

[32] M. G. Rossmann. "The position of anomalous scatterers in protein crystals". *Acta Cryst.*, **14**, 383 – 388 (1961).

[33] A. C. T. North. "The combination of isomorphous replacement and anomalous scattering data in phase determination of non-centrosymmetric reflexions". *Acta Cryst.*, **18**, 212 – 216 (1965).

[34] B. W. Matthews. "The determination of the position of the anomalously scattering heavy atom groups in protein crystals". *Acta Cryst.*, **20**, 230 – 239 (1966).

[35] B. W. Matthews. "The extension of the isomorphous replacement method to include anomalous scattering measurements". *Acta Cryst.*, **20**, 82 – 86 (1966).

[36] W. A. Hendrickson and M. M. Teeter. "Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur". *Nature(London)*, **290**, 107–113 (1981).

[37] H. Berman, K. Henrick, and H. Nakamura. "Announcing the worldwide protein data bank". *Nature Structural & Molecular Biology*, **10**(12), 980–980 (2003).

[38] A. D. Ghodke and P. R. Hannurkar. "Indus-2 reached beam current of 200 mA at 2.5 GeV". *RRCAT Newsletter*, **27**(2), 3 (2014).

[39] P. Kumar, A. D. Ghodke, and G. Singh. "Beam lifetime measurement and analysis in Indus-2 electron storage ring". *Pramana*, **80**(5), 855 – 871 (2013).

[40] P. R. Hannurkar. "Indus-2: upgradation with two insertion devices: A step towards third generation regime". *RRCAT Newsletter*, **28**(1), 3–4 (2015).

[41] N. Franchi, C. Avendao, R. Molina, et al. "β-microseminoprotein in human spermatozoa and its potential role in male fertility". *Reproduction*, **136**(2), 157–166 (2008).

[42] H. Weiber, C. Andersson, A. Murne, et al. "β-Microseminoprotein is not a prostate-specific protein: Its identification in mucous glands and secretions". *American Journal of Pathology*, **137**(3), 593–603 (1990).

[43] A. Sheth, N. Arabatti, M. Carlquist, et al. "Characterization of a polypeptide from human seminal plasma with inhibin (inhibition of FSH secretion)-like activity". *FEBS Letters*, **165**(1), 11–15 (1984).

[44] S. Garde, V. Basrur, L. Li, et al. "Prostate secretory protein (PSP94) suppresses the growth of androgen- independent prostate cancer cell line (PC3) and xenografts by inducing apoptosis". *Prostate*, **38**(2), 118–125 (1999).

[45] P. Fernlund, L. Granberg, and P. Roepstorff. "Amino acid sequence of β-microseminoprotein from porcine seminal plasma". *Archives of Biochemistry and Biophysics*, **309**(1), 70–76 (1994).

[46] M. Mäkinen, C. Valtonen-Andrè, and A. Lundwall. "New World, but not Old World, monkeys carry several genes encoding β- microseminoprotein". *European Journal of Biochemistry*, **264**(2), 407–414 (1999).

[47] J. Xuan, J. Kwong, F. Chan, et al. "cDNA, genomic cloning, and gene expression analysis of mouse PSP94 (prostate secretory protein of 94 amino acids)". *DNA and Cell Biology*, **18**(1), 11–26 (1999).

[48] C. Lazure, M. Villemure, D. Gauthier, et al. "Characterization of ostrich (Struthio camelus) β-microseminoprotein (MSP): Identification of homologous sequences in EST databases and analysis of their evolution during speciation". *Protein Science*, **10**(11), 2207–2218 (2001).

[49] N. Aoki, A. Sakiyama, M. Deshimaru, et al. "Identification of novel serum proteins in a Japanese viper: Homologs of mammalian PSP94". *Biochemical and Biophysical Research Communications*, **359**(2), 330–334 (2007).

[50] M. F. Perutz, M. G. Rossmann, A. F. Cullis, et al. "Structure of hæmoglobin: a three-dimensional Fourier synthesis at 5.5-Å. resolution, obtained by X-ray analysis". *Nature*, **185**, 416–422 (1960).

[51] C. Poloni, R. Bracco, B. Diviacco, et al. "Status of development of the insertion devices for ELETTRA". In "Particle Accelerator Conference, 1991. Accelerator Science and Technology., Conference Record of the 1991 IEEE", pp. 2712–2714 vol.5 (1991).

[52] M. Sanchez del Rio and R. J. Dejus. "Status of XOP: an x-ray optics software toolkit" (2004).

[53] ANSYS-5.7. "Engineering analysis systems (Houston, PA: Swanson Analysis Systems Inc) 15342" (1998).

[54] C. Welnak, G. Chen, and F. Cerrina. "SHADOW: A synchrotron radiation and X-ray optics simulation tool". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **347**(13), 344 – 347 (1994).

[55] T. Tanaka and H. Kitamura. "*SPECTRA*: a synchrotron radiation calculation code". *Journal of Synchrotron Radiation*, **8**(6), 1221–1228 (2001).

[56] A. S. Rajarao. "Indus-2 project report". Tech. rep., Raja Ramanna Centre for Advanced Technology (1998).

[57] V. Sahni. "Commissioning of Indus-2 Storage Ring". In "Proceedings of the 4th Asian Particle Accelerator Conference (APAC2007), RRCAT, Indore", vol. 29, pp. 61–65 (2007).

[58] Y. C. J. P. C. Tseng. "Conceptual Design Report for the High-Throughput Biological Crystallography Beamlines". Tech. rep., Synchrotron Radiation Research Center, Taiwan (2002).

[59] F. Ciocci. *Insertion Devices for Synchrotron Radiation and Free Electron Laser*, chap. 9, pp. 287–352 (World Scientific Publishing Company, 2001).

[60] D. Attwood. *Soft x-rays and extreme ultraviolet radiation: principles and applications* (Cambridge university press, 1999).

[61] W. L. Bragg. "The diffraction of short electromagnetic waves by a crystal". In "Proceedings of the Cambridge Philosophical Society", vol. 17, p. 4 (1913).

[62] B. E. Warren. *X-ray Diffraction* (Courier Corporation, 1969).

[63] A. Compton and S. Allison. *X-Rays in Theory and Experiment by Arthur H. Compton and Samuel K. Allison*. X-rays and electrons by Arthur H. Compton was 1st edition (Van Nostrand, 1957).

[64] J. W. DuMond. "Theory of the use of more than two successive x-ray crystal reflections to obtain increased resolving power". *Physical Review*, **52**(8), 872 (1937).

[65] A. A. MacDowell, R. S. Celestre, M. Howells, et al. "Suite of three protein crystallography beamlines with single superconducting bend magnet as the source". *Journal of Synchrotron Radiation*, **11**(6), 447–455 (2004).

[66] L. Assoufid, W.-K. Lee, and D. M. Mills. "Analysis of Room Temperature Silicon Crystals for the Advanced Photon Source Bending Magnet and Insertion Device Beams". Tech. rep., ANL/APS/TB-19 (1994).

[67] Q. Shen. "A5:High Heat Load Design". Tech. rep., CHESS (2000).

[68] A. Chumakov, R. Rüffer, O. Leupold, et al. "Performance of a cryogenic silicon monochromator under extreme heat load". *Journal of Synchrotron Radiation*, **11**(2), 132–141 (2004).

[69] Y. Li, A. M. Khounsary, J. Maser, et al. "Cooled mirror for a double-undulator beamline" (2004).

[70] Y. R. Jaski, M. Meron, and P. J. Viccaro. "Performance analysis of a side-cooled vertical-focusing mirror for the undulator beamline at APS" (1998).

[71] J. Arthur, W. H. Tompkins, C. Troxel, et al. "Microchannel water cooling of silicon xray monochromator crystals". *Review of Scientific Instruments*, **63**(1), 433–436 (1992).

[72] D. Bilderback. "The potential of cryogenic silicon and germanium x-ray monochromators for use with large synchrotron heat loads". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **246**(1), 434–436 (1986).

[73] C. Broennimann, E. Eikenberry, B. Henrich, et al. "The PILATUS 1M detector". *Journal of synchrotron radiation*, **13**(2), 120–130 (2006).

[74] D. Stewart. "A platform with six degrees of freedom". *Proceedings of the institution of mechanical engineers*, **180**(1), 371–386 (1965).

[75] G. Evans and R. F. Pettifer. "CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra". *Journal of Applied Crystallography*, **34**(1), 82–86 (2001).

[76] D. Lahiri and S. M. Sharma. "Shielding report for Indus-2 beamlines". Tech. rep., Bhabha Atomic Research Centre (2007).

[77] J. G. A. K. . G. H. Saleem Khan, M K Nayak. "Radiation Shielding Report on Protein Crystallography Beam Line($BL-21$), Indus-2 SRS". Tech. rep., Raja Ramanna Center for advanced Technology, Indore (2013).

[78] J.-J. Ferme and G. M. Dahan. "Bendable mirrors" (1998).

[79] W. Kabsch. "Xds". *Acta Crystallographica Section D: Biological Crystallography*, **66**(2), 125–132 (2010).

[80] P. R. Evans and G. N. Murshudov. "How good are my data and what is the resolution?" *Acta Crystallographica Section D: Biological Crystallography*, **69**(7), 1204–1214 (2013).

[81] M. D. Winn, C. C. Ballard, K. D. Cowtan, et al. "Overview of the CCP4 suite and current developments". *Acta Crystallographica Section D: Biological Crystallography*, **67**(4), 235–242 (2011).

[82] G. M. Sheldrick. "A short history of SHELX". *Acta Crystallographica Section A: Foundations of Crystallography*, **64**(1), 112–122 (2007).

[83] T. Pape and T. R. Schneider. "HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs". *Journal of applied crystallography*, **37**(5), 843–844 (2004).

[84] T. R. Schneider and G. M. Sheldrick. "Substructure solution with SHELXD". *Acta Crystallographica Section D: Biological Crystallography*, **58**(10), 1772–1779 (2002).

[85] G. M. Sheldrick. "Experimental phasing with SHELXC/D/E: combining chain tracing with density modification". *Acta Crystallographica Section D: Biological Crystallography*, **66**(4), 479–485 (2010).

[86] G. Langer, S. X. Cohen, V. S. Lamzin, et al. "Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7". *Nature protocols*, **3**(7), 1171–1179 (2008).

[87] P. D. Adams, P. V. Afonine, G. Bunkóczi, et al. "PHENIX: a comprehensive Python-based system for macromolecular structure solution". *Acta Crystallographica Section D: Biological Crystallography*, **66**(2), 213–221 (2010).

[88] P. Emsley and K. Cowtan. "Coot: model-building tools for molecular graphics". *Acta Crystallographica Section D: Biological Crystallography*, **60**(12), 2126–2132 (2004).

[89] N. Rawlings, M. Waller, A. Barrett, et al. "MEROPS: The database of proteolytic enzymes, their substrates and inhibitors". *Nucleic Acids Research*, **42**(D1), D503–D509 (2014).

[90] M. Ghosh, A. Grunden, D. Dunn, et al. "Characterization of native and recombinant forms of an unusual cobalt- dependent proline dipeptidase (prolidase) from the hyperthermophilic archaeon Pyrococcus furiosus". *Journal of Bacteriology*, **180**(18), 4781–4789 (1998).

[91] A. Forlino, A. Lupi, P. Vaghi, et al. "Mutation analysis of five new patients affected by prolidase deficiency: The lack of enzyme activity causes necrosis-like cell death in cultured fibroblasts". *Human Genetics*, **111**(4-5), 314–322 (2002).

[92] N. Vyas, A. Nickitenko, V. Rastogi, et al. "Structural insights into the dual activities of the nerve agent degrading organophosphate anhydrolase/prolidase". *Biochemistry*, **49**(3), 547–559 (2010).

[93] M.-S. Park, C. Hill, Y. Li, et al. "Catalytic properties of the PepQ prolidase from Escherichia coli". *Archives of Biochemistry and Biophysics*, **429**(2), 224–230 (2004).

[94] K. Suga, T. Kabashima, K. Ito, et al. "Prolidase from Xanthomonas maltophilia: Purification and characterization of the enzyme". *Bioscience, Biotechnology and Biochemistry*, **59**(11), 2087–2090 (1995).

[95] M. Maher, M. Ghosh, A. Grunden, et al. "Structure of the Prolidase from Pyrococcus furiosus". *Biochemistry*, **43**(10), 2771–2783 (2004).

[96] J. Jeyakanthan, K. Takada, M. Sawano, et al. "Crystal structural and functional analysis of the putative dipeptidase from Pyrococcus horikoshii OT3". *J. Biophys.*, **2009** (2009).

[97] M. Wilce, C. Bond, N. Dixon, et al. "Structure and mechanism of a proline-specific aminopeptidase from Escherichia coli". *Proceedings of the National Academy of Sciences of the United States of America*, **95**(7), 3472–3477 (1998).

[98] S. Graham, P. Lilley, M. Lee, et al. "Kinetic and crystallographic analysis of mutant Escherichia coli aminopeptidase P: Insights into substrate recognition and the mechanism of catalysis". *Biochemistry*, **45**(3), 964–975 (2006).

[99] S. Tan, R. Kern, and W. Selleck. "The pST44 polycistronic expression system for producing protein complexes in Escherichia coli". *Protein Expression and Purification*, **40**(2), 385–395 (2005).

[100] N. Chayen, P. Shaw Stewart, and D. Blow. "Microbatch crystallization under oil - a new technique allowing many small-volume crystallization trials". *Journal of Crystal Growth*, **122**(1-4), 176–180 (1992).

[101] W. Kabsch. "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry land cell constants". *Journal of Applied Crystallography*, **26**(pt 6), 795–800 (1993).

[102] B. Matthews. "Solvent content of protein crystals". *Journal of Molecular Biology*, **33**(2), 491–497 (1968).

[103] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, et al. "Phaser crystallographic software". *Journal of applied crystallography*, **40**(4), 658–674 (2007).

[104] S. C. Graham and J. M. Guss. "Complexes of mutants of Escherichia coli aminopeptidase P and the tripeptide substrate ValProLeu". *Archives of biochemistry and biophysics*, **469**(2), 200–208 (2008).

[105] P. A. Karplus and K. Diederichs. "Linking crystallographic model and data quality". *Science*, **336**(6084), 1030–1033 (2012).

[106] G. Thomas, K. Jacobs, M. Yeager, et al. "Multiple loci identified in a genome-wide association study of prostate cancer". *Nature Genetics*, **40**(3), 310–315 (2008).

[107] H. Lou, M. Yeager, H. Li, et al. "Fine mapping and functional analysis of a common variant in MSMB on chromosome 10q11.2 associated with prostate cancer susceptibility". *Proceedings of the National Academy of Sciences of the United States of America*, **106**(19), 7933–7938 (2009).

[108] B.-L. Chang, S. Cramer, F. Wiklund, et al. "Fine mapping association study and functional analysis implicate a SNP in MSMB at 10q11 as a causal variant for prostate cancer risk". *Human Molecular Genetics*, **18**(7), 1368–1375 (2009).

[109] L. Beke, M. Nuytten, A. Van Eynde, et al. "The gene encoding the prostatic tumor suppressor PSP94 is a target for repression by the Polycomb group protein EZH2". *Oncogene*, **26**(31), 4590–4595 (2007).

[110] M. Kamada, H. Mori, N. Maeda, et al. "β-Microseminoprotein/prostatic secretory protein is a member of immunoglobulin binding factor family". *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, **1388**(1), 101–110 (1998).

[111] H. Mori, M. Kamada, M. Maegawa, et al. "Enzymatic activation of immunoglobulin binding factor in female reproductive tract". *Biochemical and Biophysical Research Communications*, **246**(2), 409–413 (1998).

[112] D. Jagtap, A. Narahari, M. Swamy, et al. "Disulphide bond reduction and S-carboxamidomethylation of PSP94 affects its conformation but not the ability to bind immunoglobulin". *Biochimica et Biophysica Acta - Proteins and Proteomics*, **1774**(6), 723–731 (2007).

[113] S. Sutcliffe, A. M. De Marzo, K. S. Sfanos, et al. "MSMB variation and prostate cancer risk: clues towards a possible fungal etiology". *The Prostate*, **74**(6), 569–578 (2014).

[114] N. Aoki, H. Matsuo, M. Deshimaru, et al. "Accelerated evolution of small serum proteins (SSPs)-The PSP94 family proteins in a Japanese viper". *Gene*, **426**(1-2), 7–14 (2008).

[115] K. Hansson, M. Kjellberg, and P. Fernlund. "Cysteine-rich secretory proteins in snake venoms form high affinity complexes with human and porcine β-microseminoproteins". *Toxicon*, **54**(2), 128–137 (2009).

[116] Y. Yamazaki and T. Morita. "Structure and function of snake venom cysteine-rich secretory proteins". *Toxicon*, **44**(3), 227–231 (2004).

[117] G. Gibbs, K. Roelants, and M. O'Bryan. "The CAP superfamily: Cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins - Roles in reproduction, cancer, and immune defense". *Endocrine Reviews*, **29**(7), 865–897 (2008).

[118] V. Kumar, Y. Roske, N. Singh, et al. "Purification and preliminary X-ray crystallographic studies of β-microseminoprotein from human seminal plasma". *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, **65**(5), 518–521 (2009).

[119] M. Kumar, D. Jagtap, S. Mahale, et al. "Crystallization and preliminary X-ray diffraction analysis of human seminal plasma protein PSP94". *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, **65**(4), 389–391 (2009).

[120] A. G. Leslie. "The integration of macromolecular diffraction data". *Acta Crystallographica Section D: Biological Crystallography*, **62**(1), 48–57 (2005).

[121] T. Jones, J. Zou, S. Cowan, et al. "Improved methods for building protein models in electron density maps and the location of errors in these models." *Acta Crystallographica Section A: Foundations of Crystallography*, **47**, Pt 2/– (1991).

[122] A. Brunger, P. Adams, G. Clore, et al. "Crystallography and NMR system: A new software suite for macromolecular structure determination". *Acta Crystallographica Section D: Biological Crystallography*, **54**(5), 905–921 (1998).

[123] R. Bingham, E. Rudio-Piera, N. Meenan, et al. "Crystal structures of fibronectin-binding sites from Staphylococcus aureus FnBPA in complex with fibronectin domains". *Proceedings of the National Academy of Sciences of the United States of America*, **105**(34), 12254–12258 (2008).

[124] U. Schwarz-Linek, J. Werner, A. Pickford, et al. "Pathogenic bacteria attach to human fibronectin through a tandem -zipper". *Nature*, **423**(6936), 177–181 (2003).

[125] M. Erat, D. Slatter, E. Lowe, et al. "Identification and structural analysis of type i collagen sites in complex with fibronectin fragments". *Proceedings of the National Academy of Sciences of the United States of America*, **106**(11), 4195–4200 (2009).

[126] Y. Dou, P.-F. Baisne, G. Pollastri, et al. "ICBS: A database of interactions between protein chains mediated by -sheet formation". *Bioinformatics*, **20**(16), 2767–2777 (2004).

[127] H. Remaut and G. Waksman. "Protein-protein interaction through -strand addition". *Trends in Biochemical Sciences*, **31**(8), 436–444 (2006).

[128] H. Ghasriani, P. Fernlund, L. Udby, et al. "A model of the complex between human β-microseminoprotein and CRISP-3 based on NMR data". *Biochemical and Biophysical Research Communications*, **378**(2), 235–239 (2009).

[129] D. D. Jagtap, D. N. Modi, M. Kumar, et al. "High-affinity binding of seminal plasma PSP94 to human immunoglobulin is through the Fab domain". *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, **1804**(9), 1889–1894 (2010).

[130] J. Derrick and D. Wigley. "Crystal structure of a streptococcal protein G domain bound to an Fab fragment". *Nature*, **359**(6397), 752–754 (1992).