ICA News letter, 2003-2004

Macromolecular Crystallography at the Saha Institute of Nuclear Physics

Jiban K. Dattagupta

Crystallography and Molecular Biology Division
Saha Institute of Nuclear Physics
1/AF Bidhannagar
Kolkata 700 064

The other members of the laboratory are: Debi Choudhury, Jhimli Dasgupta, Udayaditya Sen, Piyali GuhaThakurta, Sampa Biswas, Chandana Chakrabarti, Susmita Khamrui

Ongoing Research Work

Hemoglobin variants isolated from the blood samples of β-thalassemic patients

Hemoglobin A_2 ($\alpha_2\delta_2$) is a naturally occurring hemoglobin, expressed at low concentration (~2-3%) in normal individuals but with exceptionally high level in patients having β-thalassemia, for whom it is used as a diagnostic. Thalassemias, the most common inherited genetic disorder of blood, are characterized by the absence or reduced production of globin chains, resulting in a microcytic anemia of varying degree. In β-thalassemia major, though there is a complete failure of β chain production, the α chain production continues at a near normal rate, and it stimulates the production of δ and γ chains. The δ and γ chains thus expressed, combine with α chains to raise the level of hemoglobin A_2 (HbA_2) and hemoglobin F (HbF). HbF has a very high affinity for oxygen and is a poor oxygen deliverer[1]. Hence, the predominant functional hemoglobin present in the patients of β-thalassemia major is HbA_2.

The δ chain of HbA_2 differs from the β chain of adult hemoglobin (HbA) at ten sites[2] in the amino acid sequence. The oxygen binding property of HbA_2 was a matter of dispute[3, 4] but it has been shown that recombinant HbA_2 has a higher oxygen affinity than HbA and its response to the allosteric regulators like 2, 3-diphosphoglycerate is lower than the corresponding properties for the adult Hemoglobins[5]. Apart from this, HbA_2 could potentially be used for genetic manipulation in the treatment of patients with sickle cell anemia as it has a considerable inhibitory property towards the polymerization of sickle cell hemoglobin (HbS).

β-thalassemia, when combined with splice site mutations, causes severe clinical consequences[6]. HbE is a very common abnormal hemoglobin, caused by splice site mutation, with a single β chain substitution (GAG → AAG) where Glutamic acid is replaced by Lysine at 26th position (E26K)[6]. Heterozygous HbE (Hb AE) and homozygous HbE (Hb EE) are both benign disorder with mild microcytic anemia, but the offspring who coinherit hemoglobin E and β thalassemia have a more severe microcytic anemia. Hence to explore the roles of HbA_2 and HbE in β thalassemias at the atomic level, we undertook a project that included the purification, crystallization and three dimensional structural studies of these important hemoglobin variants.

Purification of HbA_2 and HbE were done by cation exchange column chromatography in presence of KCN from the blood samples of individuals suffering from β-thalassemia minor and Eβ-thalassemia respectively. Both HbA_2 and HbE crystallized in space group $P2_12_12_1$ and the asymmetric unit in each case contains one Hb tetramer in R_2 state. X-ray diffraction data of HbA_2 and HbE were collected upto 1.85Å and 1.73Å respectively.

Molecular replacement calculations for HbA_2 and HbE were done using AMoRe, CCP4[7] with data between 10-3.5Å, using the coordinates of liganded human carbonmonoxy adult Hb (in R2-state) as search model (PDB code 1BBB[8];). A clear-cut solution was obtained for HbA_2 with a correlation coefficient of 64.5% and R factor of 37.1%. Subsequent rigid body refinement using CNS[9] treating each subunit as a separate rigid body, further reduced the R factor to 33.1% (R_{free}=35.2%). Electron density map, calculated at this stage, showed clear side chain densities for the ten δ chain residues that differ from β chain. For HbE, solution was obtained with correlation coefficient of 73.1% and R-factor of 32% which, upon rigid body refinement, using CNS, further reduced the R-factor to 31.1% (R_{free}=34.2%). A composite OMIT map was calculated at this stage which showed electron density for the side chain of $K26^{\beta}$ residue unequivocally. Both the structures are now refined. In case of HbA_2, the final R factor is 17.7 % (R_{free}=19.9 %) whereas for HbE, it is 19.1% (R_{free}=21.1%). Currently we are analysing the structures.

References

1. James A. Frier and M. F. Perutz. (1977). Structure of Human Foetal Deoxyhaemoglobin, *J. Mol. Biol*. **112,** 97-112.
2. I. M. Russu, A. K. Lin, S. Ferro-Dosch, C. Ho. (1984). A proton nuclear magnetic resonance investigation of human hemoglobin A_2, *Biochim et Biophys Acta* **785(3),** 123-131.
3. Huisman, T.H.J., Dozy, A.M., Nechtman, C. & Thompson, R.B. (1962). Oxygen Equilibrium of Hemoglobin A_2 and Its Variant Hemoglobin A_2 (or B_2') *Nature* **195** 1109-1110.
4. DeBruin, S.H. & Jansen, L.H.M.. (1973). Comparison of the oxygen and proton binding behavior of human hemoglobin A and A_2, *Biochem. Biophys. Acta*. **295** 490-494.
5. K. Inagaki, J. Inagaki, A. Dumoulin, J. C. Padovan, B. T. Chait, A. Popowicz, L. R. Manning, J. M. Manning. (2000). Expression and Properties of Recombinant HbA_2 ($\alpha_2\delta2$) and Hybrids Containing δ-β sequences, *J. Protein Chem.* **19(8)**, 649-662.
6. H. Shirohzu, H. Yamaza, Y. Fukumaki. (2000). Repression of aberrant splicing in beta-globin pre-mRNA with HbE mutation by antisense oligoribonucleotide or splicing factor SF2/ASF, *Int J. Hematol*. **72(1)** 28-33.
7. Collaborative Computational Project, Number 4. (1994). The CCP4 suit: Programs for protein crystallography, *Acta Cryst.* **D50** 760-763.

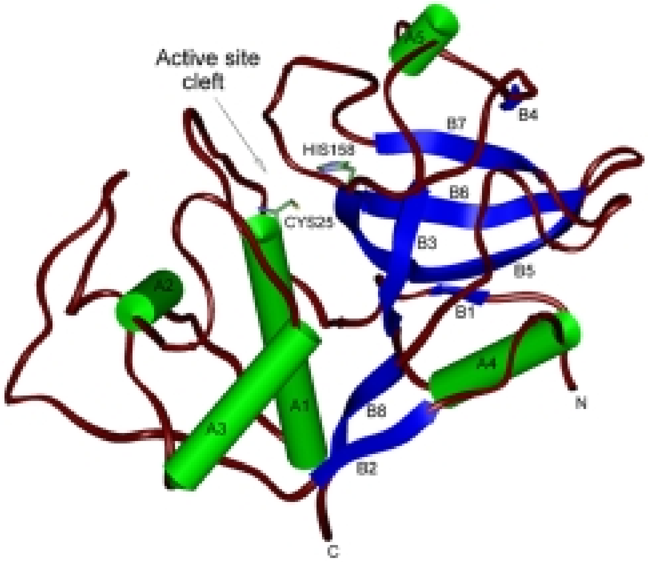
8. M. M. Silvia, P. H. Rogers, A. Amone. (1992) A third quaternary structure of human hemoglobin A at 1.7 Å resolution, *J. Biol. Chem*. **267,** 17248-17256.

9. A. T. Brünger, P. D. Adams, G. M. Clore, W. L. Delano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren. (1999). Crystallography & NMR system: A new software suit for macromolecular structure determination, *Acta Cryst*. **D54,** 905-921.

The ervatamins : cysteine proteases from *Ervatamia coronaria*

Papain-like cysteine protease is the class of enzymes whose proteolytic activity is characterized by the presence of a catalytic diad of cysteine and histidine. The members of the papain-like cysteine protease family share similar amino acid sequences and a charecteristic overall fold [1]. A number of papain-like cysteine proteases, ervatamins, have been isolated [2,3] from the latex of the medicinal plant *Ervatamia coronaria.* Ervatamins exhibit some novel properties that are distinctly different from papain and the other members of the family and these enzymes also differ among themselves in many respects. Relatively higher stability of ervatamins to pH, chemical denaturants, temperature and organic solvents and differential specificity towards various synthetic substrates make them excellent candidates to provide better insights into the structure function relationship of the cysteine proteases in general.

The crystal structures of two stable ervatamins, ervatamin B and ervatamin C, were solved at 1.63[4] and 1.9Å resolution respectively. The two proteases share the same overall fold observed in the papain family of cysteine proteases. The active site is in a cleft at the interface between the two domains. The left or the L-domain is mainly made up of α-helices and the right or the *R*-domain has mainly an anti-parallel β-sheet structure. The N-terminus crosses over from the right to the left domain, and the C-terminus from the left to the right, strapping the two domains tightly together.



For both the proteins, only 21 N-terminal amino acid residues were known to us, and the full primary sequence could be determined from the well developed electron density map. Some important substitutions of conserved residues at both the left and the right domains and also at the region of the interdomain cleft in comparison to other members of the papain family were observed. The increase in the number of hydrogen bonds due to these substitutions seems to be responsible for their higher thermal stability. Moreover, in ervatamin C, a unique fourth disulphide bond, present in the right domain in addition to the other three conserved disulphide bridges, coupled with a low gap volume, high aliphatic index and low instability index make it more stable than ervatamin B and others.

Another important feature observed in the ervatamins is their altered substrate specificity caused by natural substitutions of amino acid residues at the active site compared to other members of the family. In ervatamin B[4] these substitutions reduce the volume of the S2 subsite, thereby changing its substrate specificity. This is in accordance with the biochemical observation that though it can accommodate smaller side chains like alanine, upto leucine, it is too shallow for bulkier ones. In ervatamin C, unique natural substitutions at the active site cleft alters its substrate specificity towards polypeptide substrates by eliminating sufficient hydrophobic interactions at the S3-P3 site. Molecular modelling studies with a protein inhibitor docked at the active site region reveals that these substitutions do not affect the binding of a natural protein substrate to the enzyme significantly, as there exist a number of other non specific interactions at the enzyme-substrate interface.

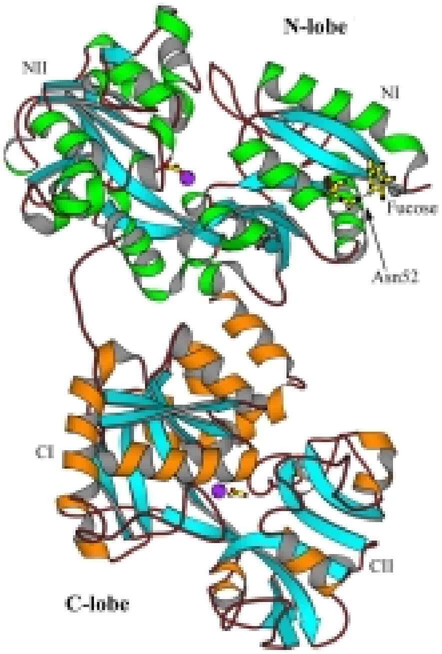
References

1. Kamphuis, I. G., Drenth, J. & Baker, E.N. Tniol Proteases:Comparitive Studies Based on the High-resolution Structures of Papain and Actinidin, and on Amino Acid Sequence Information for Cathepsin B and H, and Stem Bomelain. (1985). *J. Mol. Biol.* **182**, 317-329.
2. Sundd, M., Kundu, S., G.P. Pal & Medicherla, J.V. (1998). Purification and Characterization of a Highly Stable Cysteine Protease from the latex of *Ervatamia coronaria*. *Biosci.Biotechnol. Biochem*. **62**, 1947-1955.
3. Kundu, S., Sundd, M. & Jagannadham, M.V. (2000). Purification and Characterization of a Stable Cysteine Protease Ervatamin B, with two disulphide bridges from the latex of *Ervatamia coronaria*. *J. Agric. Food Chem*. **48**, 171-179.
4. Biswas, S., Chakrabarti, C. Kundu, S., Jagannadham, M.V. & Dattagupta, J.K. (2003). Proposed amino acid sequence and the 1.63Å X-ray crystal structure of a plant cysteine protease, Ervatamin B: some insights into the structural basis of its stability and substrate specificity. *PROTEINS: Structure, Function, and Genetics* **51**, 489-497.

Hen Serum Transferrins

Transferrins, the major proteins involved in iron regulation and transport in vertebrates and some invertebrates, are mainly subdivided into two branches, soluble glycoproteins and membrane melanotransferrin. The soluble glycoproteins include serum transferrin found in blood, ovotransferrin found in avian egg white, and lactoferrin found in numerous extracellular fluids and in the specific granules of polymorphonuclear lymphocytes. These proteins are monomeric glycoproteins with a molecular weight of ~80kDa, organized as two terminal lobes (N- and C-). The two lobes, endowed with high degree of similarity, arose from gene duplication and are further divided into two similar sized domains, NI & NII in the N-lobe and CI & CII in the C-lobe. Each lobe contains an iron binding cleft in which Fe^{3+} is coordinated to two tyrosine residues, one histidine residue, one aspartate residue and a synergistic bicarbonate or carbonate adjacent to an arginine. The two domains of each transferrin are believed to open to allow the entry or release of Fe^{3+}. Serum transferrins transport iron from the blood stream to the cytosol in a pH dependent manner by receptor-mediated endocytosis. Both the avian transferrins, ovo and serum, have the same amino acid sequence, and differ only in their attached carbohydrate moiety[1]. On the contrary, the sequence identity between hen serum transferrin and other mammalian serum transferrins e.g. rabbit and porcine, is only 51%. Although all cell binding studies[2,3] have utilized ovotransferrin with the assumption that it is identical to its serum counterpart an *in vivo* iron-transport function characteristic of serum transferrin has not been proved for hen ovotransferrin (hOT) [4].

We have isolated, purified and determined the three dimensional structure of hen serum transferrin in its diferric form (hST) by X-ray crystallography at 2.8Å resolution. The structure has been compared with the three dimensional structure hOT[5], and also with structures of some other transferrins viz. rabbit serum transferrin (rST) and human lactoferrin (hLF). The overall conformation of the hST molecule is essentially the same as that of other transferrins. However, the relative orientation of the two lobes, which is related to the species-specific receptor recognition property of transferrins, has been found to be different in hST from that of hOT, rST and hLF. A number of additional hydrogen bonds between the two domains in the N- and C-lobes have been identified in the structure of hST compared to hOT, which indicate more compactness of the lobes of hST than hOT. The nature of interdomain interactions of hST are observed to be more closer to rST rather than hOT. A putative carbohydrate-binding site has been identified in the N-lobe of hST at Asn52 and a fucose molecule could be modeled at the site.



We are also in the process of solving the three dimensional strucuture of iron-free (apo-) form of hen serum transferrin at 3.45Å. This will be the first structure to date of any intact molecule of apo- serum transferrin.

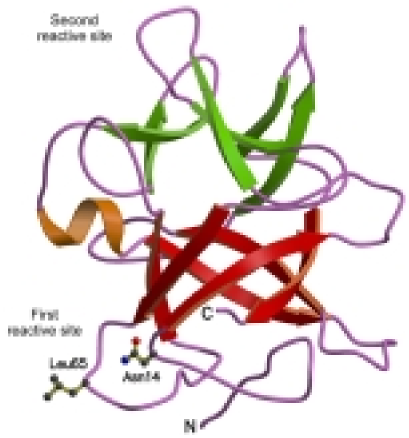
References:

1. Thibodeau, S.N., Lee, D.C. & Palmiter, R.D. (1978). Precursor of egg white ovomucoid. Amino acid sequence of an NH2-terminal extension. *J. Biol. Chem.* **253**, 3771-3774.
2. Mason, A.B., Brown, S. A. & Church, W. R. (1987). Monoclonal antibodies to either domain of ovotransferrin block binding to transferrin receptors on chick reticulocytes. *J. Biol. Chem.* **262**, 9011-9015.
3. Mason, A.B., Woodworth, R. C., Oliver, R. W. A., Green, B. N., Lin, L-N., Brandts, J. F., Savage, K. J., Tam, B. M. & MacGillivray, R. T. A. (1996). Association of the two lobes of ovotransferrin is a prerequisite for receptor recognition. Studies with recombinant ovotransferrins. *Biochem. J.* **319**, 361-368.
4. Mizutani, K., Muralidhara, B. K., Yamashita, H., Tabeta, S., Mikami, B. & Hirose, M. (2001). Anion mediated Fe^{3+} release mechanism in ovotransferrin C-lobe: A structurally identified SO_4(2-) binding site and its implications for kinetic pathway. *J. Biol. Chem*, **276**, 35940-35946.
5. Kurokawa, H., Mikami, B. & Hirose, M. (1995). Crystal structure of diferric hen ovotransferrin at 2.4Å resolution. *J. Mol. Biol*. **254**, 196-207.

Winged Bean Chymotrypsin Inhibitor protein

a) In-silico mutations and molecular dynamics studies

The Winged bean Chymotrypsin Inhibitor protein, WCI, belongs to the Kunitz (STI) family of serine protease inhibitors and inhibits α-chymotrypsin in a 1:2 molar ratio[1]. The single polypeptide chain of WCI (183 amino acid residues, M_r=20.2 kDa) has a characteristic β-trefoil fold (Figure 1) and two reactive sites, Gln63-Phe68 and Asn38-Leu43 (designated as 'first' and 'second' site respectively) are situated on surface exposed loops[2]. The scaffolding residue Asn14, conserved in the Kunitz (STI) family, intrudes into the first reactive site loop and forms hydrogen bonding interactions with the loop residues. To understand the contribution of this



scaffolding residue on the stability of the reactive site loop, Asn14 in WCI was mutated to Lys (N14K) and Asp (N14D) for X-ray structural studies. The crystal structure analysis of these mutants (N14K at 2.05Å and N14D at 1.9 Å resolution) showed no major alteration in the reactive site loop conformation which was reflected in a minor difference in dissociation constants (rWCI 1.14x10^{-9} M, N14K 2.4x10^{-9} M, N14D 2.58x10^{-9} M)[3]. So we felt the necessity of replacing this residue with few other amino acids having different sizes and charges, which is likely to throw more light on the electrostatic nature of the loop. Accordingly we decided to mutate it *in-silico* to Gly, Ala, Ser, Thr, Leu and Val with an idea to run molecular dynamics simulations on the mutants using DISCOVER-III module of Insight-II.

The simulated structures of the mutants were then analyzed in terms of hydrogen bonding interactions, dihedral angles and water-orientation in the loop region. In each case, we collected twenty snapshots from the trajectories of the last 100ps production run. Then we docked each snapshot at the active site groove of Chymotrypsin, using the program MULTIDOCK[4], to quantify the deviations in the canonical conformation in terms of the interaction energy and occurrence of standard hydrogen bonding interactions at P3, P1 and P2´ positions as seen in the serine protease-protease inhibitor complexes.

Results of the MD simulations revealed the conformational variability and range of motions possible for the reactive site loop of different mutants. The N-terminus side of the scissile bond, which is close to a β-barrel, is conformationally less variable, while the C-terminus side, which is relatively far from any such secondary structural element, is more variable and needs stability through hydrogen bonding interactions. Moreover, results of the MD simulations coupled with docking studies showed that the hydrophobic residues like Leu and Val at the 14th position are disruptive for the integrity of the reactive site loop whereas residue like Thr, that can stabilize the C-terminus side of the scissile bond, can be predicted at this position. However, the size and charge of an Asn residue made it most suitable for the best maintenance of the integrity of the reactive site loop explaining its conserved nature in the family.

References:

1. Kortt, A. A. (1980) Isolation and properties of a chymotrypsin inhibitor from winged bean seed (Psophocarpus tetragonolobus (L) Dc.). *Biochim. Biophys. Acta*, **624**, 237-248.
2. Dattagupta JK, Podder A, Chakrabarti C, Sen U, Mukhopadhyay D, Dutta SK, Singh M. (1999). Refined crystal structure (2.3 A) of a double-headed winged bean alpha-chymotrypsin inhibitor and location of its second reactive site. *PROTEINS: Structure, Function, and Genetics*, **35(3)**:321-31.
3. S. Ravichandran, J. Dasgupta, C. Chakrabarti, S.Ghosh, M.Singh and J.K.Dattagupta (2001) The role of Asn 14 in the stability and conformation of the reactive-site loop of winged bean chymotrypsin inhibitor : Crystal structures of two point mutants Asn14→ Lys and Asn14 →Asp . Protein Engineering **14**, 349-357.
4. Jackson, R. M., Gabb, H. A. and Sternberg, M. J. E. (1998). Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. *J. Mol. Biol.* **276(1)**, 265-285.

b) Cloning expression and purification of some mutants

To explore the effect of Asn14 in the loop stability of WCI, we felt the necessity of making some more mutations at this site. With this objective, three mutant inserts N14G, N14T and first 14-deletion were prepared through PCR based site directed mutagenesis and were cloned in pGEM-T vector (T/A cloning vector, having M13 primer site) using JM109 competent cells. Desired clones were selected through blue-white screening using IPTG and X-gal in presence of Ampicillin and the mutations were confirmed by DNA sequencing. For bacterial overexpression, the gene of N14T mutant was subcloned in pET28a+, a vector containing a N-terminal His tag/Thrombin/T7-tag just before multiple cloning sites. The mutant protein N14T was then expressed in BL21 bacterial cell using IPTG up to a final concentration of 0.1mM at OD_{600} of 0.5.

After initial purification of the protein, using Ni-NTA column, the N-terminal tag was cleaved by Thrombin. This mutant protein N14T was further purified from the contaminants/Thrombin/His-tag using G-75 gel filtration chromatography column which showed a single band in SDS-PAGE.

The mutant protein N14T was checked for its inhibitory activity against Chymotrypsin. Supporting the results of MD simulation[1], N14T formed 1:2 complex with Chymotrypsin. The expression, purification and assay of other mutants are underway.

Reference

1. Dasgupta, J., Sen, U. and Dattagupta J.K. (2003). *In-silico* mutations and molecular dynamics studies on a Winged bean Chymotrypsin Inhibitor protein. *Protein Engineering*, **16(7)**, 489-496.

Recent publications related to the research work in this report:

1. Dattagupta, Jiban K., Podder, Aloka, Chakrabarti, Chandana, Sen, Udayaditya, Mukhopadhyay, Debashis, Dutta, Samir K. and Singh, Manoranjan (1999). Refined Crystal Structure (2.3 Å) of a Double-Headed Winged Bean α-Chymotrypsin Inhibitor and Location of Its Second Reactive Site. *PROTEINS : Structure, Function, and Genetics*, **35(3)**, 321-331.
2. Chandana Chakrabarti, Sampa Biswas, Suman Kundu, Monica Sundd, Jagannadham V. Medicherla and Jiban Kanti Dattagupta (1999). Crystallization and preliminary X-ray analysis of ervatamin B and C, two thiol proteases from *Ervatamia coronaria. Acta Cryst.* **D55**, 1074-1075.
3. Debashis Mukhopadhyay and Jiban K. Dattagupta (1999). Synchrotron radiation in Molecular Biology. *Science & Culture*, May-June, 128-132.
4. J. Dasgupta, S. Ravichandran, D. Mukhopadhyay, U.Sen, A. Podder, C. Chakrabarti and J.K. Dattagupta (1999). A Serine Protease Inhibitor Protein : From X-ray Structure to Protein Engineering. In Volume published in honour of Prof. G.N. Ramachandran during International Biophysics Congress at New Delhi in September, 1999. Book name: "*Perspectives in Structural Biology*" Editors:M.Vijayan, N.Yathindra, A.S.Kolaskar, Page 75-82.

5. S. Ravichandran, U. Sen, C. Chakrabarti and Dattagupta, J.K. (1999). Cryocrystallography of a Kunitz-type serine protease inhibitor : The 90 K structure of winged bean chymotrypsin inhibitor (WCI) at 2.13 Å resolution. *Acta Cryst.* **D55(11)**, 1814-1821.

6 Mukhopadhyay D. (2000) The molecular evolutionary history of a winged bean alpha-chymotrypsin inhibitor and modeling of its mutations through structural analyses. *J Mol Evol.* **50(3**), 214-23.

7. S. Ravichandran, J. Dasgupta, C. Chakrabarti, S.Ghosh, M.Singh and J.K.Dattagupta (2001) The role of Asn 14 in the stability and conformation of the reactive-site loop of winged bean chymotrypsin inhibitor : Crystal structures of two point mutants Asn14→ Lys and Asn14 →Asp . *Protein Engineering* **14**, 349-357.

8. Debi Choudhury, Piyali Guha Thakurta, Rakhi Dasgupta, U.Sen, S. Biswas, C. Chakrabarti and J.K. Dattagupta (2002) Purification and Preliminary X-ray Studies on Hen Serrotransferrin in Apo- and Holo- form. *Biochemical and Biophysical Research Communications,* **295**, 125-128.

9. Sampa Biswas, Chandana Chakrabarti, Suman Kundu, Medicherla V. Jagannadham and Jiban K. Dattagupta (2003) Proposed Amino Acid Sequence and the 1.63 Å X-ray Crystal Structure of a Plant Cysteine Protease, Ervatamin B : Some Insights into the Structural Basis of its Stability and Substrate Specificity. . *PROTEINS : Structure, Function, and, Genetics* **51**, 489-497.

10. Jhimli Dasgupta, Udayaditya Sen, Debi Choudhury, Poppy Datta, Abhijit Chakrabarti, Sudipa Basu Chakrabarty, Amit Chakrabarty and J.K. Dattagupta (2003) Crystallization and preliminary X-ray structural studies of Hemoglobin A2 and Hemoglobin E, isolated from the blood samples of Beta-thalassemic patients. *Biochemical and Biophysical Research Communications*, **303**, 619-623.

11. Jhimli Dasgupta, Udayaditya Sen and J.K.Dattagupta (2003) *In-silico* mutations and molecular dynamics studies on a Winged bean Chymotrypsin Inhibitor protein. *Protein Engineering,* **16(7)**, 489-496.

12. Piyali Guha Thakurta, Debi Choudhury, Rakhi Dasgupta and J.K.Dattagupta (2003) Crystal structure of an avian serum transferrin and possible explanations for the functional differences between same gene products. *Acta Cryst*. D (in press)

13. Udayaditya Sen, Sampa Biswas, Chandana Chakrabarti and J. K. Dattagupta (2003) Structure-function of a few plant proteins using X-ray crystallography. *Indian Journal of Physics* Special Issue (in press).