



# Macromolecular Crystallography in India in the Global Context\*

M. Vijayan

**Abstract** | The most spectacular applications of crystallography are currently concerned with biological macromolecules like proteins and their assemblies. Macromolecular crystallography originated in England in the thirties of the last century, but definitive results began to appear only around 1960. Since then macromolecular crystallography has grown to become central to modern biology. India has a long tradition in crystallography starting with the work of K. Banerjee in the thirties. In addition to their contributions to crystallography, G.N. Ramachandran and his colleagues gave a head start to India in computational biology, molecular modeling and what we now call bioinformatics. However, attempts to initiate macromolecular crystallography in India started only in the seventies. The work took off the ground after the Department of Science and Technology handsomely supported the group at Indian Institute of Science, Bangalore in 1983. The Bangalore group was also recognized as a national nucleus for the development of the area in the country. Since then macromolecular crystallography, practiced in more than 30 institutions in the country, has grown to become an important component of scientific research in India. The articles in this issue provide a flavor of activities in the area in the country. The area is still in an expanding phase and is poised to scale greater heights.

## 1 Introduction

Following the discovery of X-ray diffraction by crystals by von Laue and his colleagues in 1912, Lawrence Bragg determined for the first time a crystal structure, that of sodium chloride, in 1913. That marked the beginning of X-ray crystallography or X-ray crystal structure analysis. Since then X-ray crystallography grew to become the method of choice to determine the atomic and molecular structure of matter. The largest structure to be determined using this method, at the turn of the century, is that of ribosome, an organelle on which protein synthesis occurs. Sodium chloride is made up of just two atoms while ribosome contains a couple of hundred thousand non-hydrogen atoms. The progression from sodium chloride to ribosome demonstrates the progress in X-ray crystallography in about a century and the shift in the emphasis on problems in the area.

In the early years, most of the structures determined were those of minerals or inorganic materials. Organic chemical crystallography gathered momentum in the twenties and the thirties of the last century. The pinnacle of the glory of the structure analysis of organic compounds was the structure determination of vitamin B12 by Dorothy Hodgkin in the fifties. It is often said that she relieved organic chemists of the drudgery of structure determination. With the subsequent advent of direct methods and computer methods based on them, the structure determination of organic compounds has almost become routine. Efforts in inorganic and organic chemical crystallography continue with explorations in new areas such as supramolecular association, crystal engineering, charge density etc. X-ray crystallography is a major tool in the general area of materials science. Undoubtedly the most spectacular applications

*Molecular Biophysics Unit,  
Indian Institute of Science,  
Bangalore 560012, India.  
mv@mbu.iisc.ernet.in*

\* This contribution has considerable overlap with ref. 26 and ref. 27, which are based on lectures delivered by the author.

of crystallography now is in biology, particularly in relation to biological macromolecules such as proteins and their assemblies. Biological macromolecular crystallography is central to structural biology and to modern biology itself.

## 2 Macromolecular Crystallography— The Global Context

The origin of macromolecular crystallography can be traced to the recording of the X-ray diffraction pattern from the crystals of the digestive enzyme pepsin by J.D. Bernal and his graduate student Dorothy Crowfoot (who subsequently became famous as Dorothy Hodgkin) in 1934 at the Cambridge University.<sup>1</sup> Dorothy returned to Oxford and recorded the diffraction pattern from the crystals of insulin in 1935,<sup>2</sup> which marked the beginning of her independent research career. Chymotrypsin and hemoglobin were crystallized, and preliminary X-ray studies were carried out on the crystals soon afterwards by Perutz, Fankuchen and Bernal.<sup>3</sup> Preliminary studies of a few other macromolecular crystals were also carried out during the thirties and the forties. However, the pioneers were well ahead of their times as even the exact chemical nature of proteins was not unambiguously established at that time.

In the meantime, modeling studies of biopolymers with or without the aid of fibre diffraction data were progressing apace. X-ray diffraction studies carried out by Astbury on fibrous proteins have been extensive. Linus Pauling and his colleagues discovered the  $\alpha$ -helical and the  $\beta$ -sheet structures of polypeptide chains in the early fifties using simple modeling.<sup>4,5</sup> The most significant result in the area has been the proposal of the double helical model of DNA by Watson and Crick in 1953.<sup>6</sup> Soon afterwards G.N. Ramachandran and Gopinath Kartha proposed the triple helical coiled coil structure of collagen.<sup>7</sup> The fifties, particularly the first half of the decade, witnessed the triumph of molecular modeling.

A major breakthrough in the X-ray structure analysis of globular proteins occurred when Perutz and his colleagues demonstrated that attachment of one or a few heavy atoms to a protein molecule in the crystal could produce measurable changes in the intensities of diffraction maxima, and hence isomorphous replacement could be used for structure solution.<sup>8</sup> The first definitive results in the area emerged around 1960 when the groups of Kendrew<sup>9</sup> and Perutz<sup>10</sup> determined the crystal structures of myoglobin and hemoglobin respectively. The third protein and the first enzyme to be X-ray analysed was lysozyme, by D.C. Phillips and his colleagues.<sup>11</sup> This analysis was soon followed

by the structure determination of ribonuclease A by Gopinath Kartha and his colleagues at Buffalo.<sup>12</sup> The crystal structures of a number of proteolytic enzymes became available during the second half of the sixties.<sup>13</sup> The structure of insulin was reported from Dorothy Hodgkin's laboratory towards the end of the decade.<sup>14,15</sup>

Efforts to understand and analyse protein structures also started in the sixties. The most important component of this effort was the development of the Ramachandran map by G.N. Ramachandran, C. Ramakrishnan and V. Sasisekharan.<sup>16</sup> The map, which till date remains the simplest descriptor and tool for validation of protein structures, defines the possible conformations of protein structures using simple steric considerations.

Based on the experience gained in the sixties, work in macromolecular crystallography gathered momentum in the seventies and thereafter. Thousands of protein structures have since been determined and the impact of macromolecular crystallography on modern biology has been phenomenal.<sup>17</sup> Major advances in technology involving, for example, synchrotron facilities and position sensitive detectors and recombinant technology, substantially aided the rapid growth of the area. Perhaps, there is no area in modern biology which has remained untouched by macromolecular crystallography.

## 3 Macromolecular Crystallography In India—Early Years

India has a long tradition in crystallography, starting with the work of K. Banerjee, a student of C.V. Raman, at the Indian Association for the Cultivation of Science, Calcutta (now Kolkata) in the thirties. G.N. Ramachandran and his colleagues at Madras (now Chennai), in addition to their contributions to crystallography, gave a head start to computational and theoretical structural biology and what we now call bioinformatics.<sup>18</sup> A few Indians have also been involved in macromolecular crystallographic studies abroad. However, I was the first protein crystallographer to return to India, to the Indian Institute of Science, Bangalore in 1971, after participating in the structure solution of insulin at Oxford. Subsequently, K.K. Kannan, who was involved in the structure determination of carbonic anhydrase at Upsala, joined the Bhabha Atomic Research Centre (Mumbai) in 1978.

My intention on my return was to initiate macromolecular crystallography in India. However, it was impossible to do so primarily on account of the paucity of financial support. I established myself as an independent researcher working on

small molecules. The work involved, among other things, exploration of what is now described as supramolecular association of amino acids and peptides among themselves as well as with carboxylic acids. Macromolecular crystallography in India received a major impetus in 1983 when the Department of Science and Technology (DST) handsomely supported our macromolecular crystallography efforts at Bangalore as part of their Thrust Area Programme. The Bangalore Centre also came to be recognized as a national nucleus for the development of the area in the country.

The early macromolecular efforts at Bangalore were primarily concerned with lectins, protein hydration and plant viruses. The structure analysis of peanut lectin<sup>19,20</sup> and jacalin<sup>21</sup> and that of sesbania mosaic virus<sup>22</sup> and physalis mottle virus<sup>23</sup> were among the major macromolecular crystallographic contributions from India. The work on lectins has had considerable impact on the development of the area in the country. The early Mumbai work was mainly concerned with carbonic anhydrase.<sup>24</sup>

#### 4 The Current Scenario

From the early nineties, macromolecular crystallography began to spread to different laboratories in the country.<sup>25-27</sup> Currently, work in the area is being pursued in more than 30 institutions by groups led mainly by scientists trained at Bangalore and their descendants. There are a few group leaders who have originated from other schools as well. All of them together form a reasonably coherent community which is able to orchestrate together for the common good. The systems pursued by different macromolecular crystallography groups in the country cover a wide range and it is impossible to refer to all of them here. I shall therefore confine myself primarily to a few general remarks pertaining to the papers included in this issue.

A major macromolecular crystallography group that emerged in India in the early nineties is that of T.P. Singh at the All Institute of Medical Sciences in New Delhi. Singh and his colleagues have been extraordinarily prolific during the past two decades. They have worked on a variety of important problems and have made significant contributions in the area of structure-based inhibitor/drug design with particular reference to inflammation. A major strand in their researches has all along been concerned with proteins in animal, including human secretions. The proteins studied by them include lactoferrins, lactoperoxidases, mammary gland/breast cancer regression proteins, matrix melanosomal proteins, and peptidoglycan recognition proteins. The article by

Singh and his colleagues in this issue is concerned with peptidoglycan recognition proteins.

Dinakar Salunke's work at the National Institute of Immunology (NII), New Delhi, gathered momentum at the turn of the century. To start with, the emphasis was on the mimicry of carbohydrates by peptides and other molecules, in their interactions with lectins. This comparatively novel approach led to structural investigations on immunological issues such as antibody maturation. More recently, Salunke's group, at NII and their new abode at the Regional Centre for Biotechnology, National Capital Region on Delhi, have turned their attention to allergenic proteins. This new effort forms part of the subject matter of the article of Salunke and co-authors.

By the turn of the century, macromolecular crystallography in India had come of age and was ready to take up fresh challenges. In particular, there was a new initiative aimed at proteins from microbial pathogens which cause infectious diseases, a problem of considerable relevance to developing countries like India. The initiative was particularly successful in relation to TB proteins. The first structural biology effort in this area was perhaps a modeling study of *M. tuberculosis* RecA (*Mt* RecA) in 1996 at this Institute.<sup>28</sup> This was soon followed by an outstanding bioinformatics effort by one of the contributors to this issue (Shekhar Mande) and his colleague resulting in the identification of a novel class of inositol-1-phosphate synthase enzyme in *M. tuberculosis*.<sup>29</sup> The first crystal structure of a TB protein to be determined in India was that of MtRecA analysed in this laboratory in 2000.<sup>30</sup> Since then many groups have joined the effort. Currently, there are ten institutions in India in which structural biology of mycobacterial proteins is being pursued. Among the total number of TB protein structures determined in the world, more than 10% are from India. Three major participants in the Indian effort on mycobacterial proteins have contributed to this issue.

The contribution of Shekhar Mande in identifying a novel class of inositol-1-phosphate synthase has already been referred to. He was then at the Institute of Microbial Technology, Chandigarh. Subsequently, he spent many years at the Centre for DNA Fingerprinting and Diagnostics, Hyderabad before moving again recently to the National Centre for Cell Science, Pune. All along, he has been actively engaged with the structural biology of TB proteins. The proteins studied by him include Hsp65, chorismate mutase, thioredoxin reductase, YefM antitoxin and cAMP receptor proteins. He has also used graph theoretical analysis of protein-protein interactions and Boolean modeling to

address the onset of latency of *M. tuberculosis*. The article of Mande and his colleagues in this issue is concerned with redox proteins in the pathogen.

The research interests of R. Sankaranarayanan, who has been at the Centre for Cellular and Molecular biology, Hyderabad since 2002, covers a wide spectrum. His major interest has been in editing/proof reading proteins which are important in ensuring translational accuracy during protein synthesis. In collaborative efforts, he has also worked on polyketide and fatty acylsynthases from *M. tuberculosis*, *B. subtilis* lipase, virulence factors of *Xanthomonas oryzae* pv. *oryzae*, RNA degradation machinery in *P. syringae* and calcium binding proteins. The paper by Sankaranarayanan and his colleagues in this issue is concerned with polyketide and acyl synthases from *M. tuberculosis*. These enzymes are particularly important in view of the complex nature of the lipid composition of the pathogen.

Ravishankar Ramachandran has been leading a vibrant macromolecular crystallography group at the Central Drug Research Institute at Lucknow for nearly a decade. Their work has encompassed several pathogens, but the emphasis has been mainly on TB proteins. The TB proteins thoroughly studied by them include the adenylation domain of the NAD<sup>+</sup>-dependent DNA ligase and lysine ε-amino-transferase. They are also seriously into structure based inhibitor design aimed at drug development. The review of Taran Khanam and Ravishankar Ramachandran in the issue highlights the importance of DNA repair system in *M. tuberculosis* in the development of anti-bacterials.

Work of only those contributed to this issue has been referred to above. In addition, lectins continue to be a major area of endeavour in different laboratories. Structural and related studies on proteases and their inhibitors are being pursued in more than one laboratory. Mention must also be made on efforts on a transcription regulator λ CII, the membrane protein OmpC, penicillin V acylase and hydrolase, phospholipases, hemoglobin from different sources, and jack bean urease. Although not as extensively as TB proteins, proteins from *Salmonella typhimurium* are being studied at least in three laboratories. There is also considerable activity on proteins from *Plasmodium falciparum* as well as *Leishmania donovani* and *Entamoeba histolytica*. The work on plant viruses, referred to earlier, continues. Impressive studies have been made in the further work on HIV protease. Notable contributions have emerged from India on the structural biology of proteins from rotavirus as

well. Structural studies on proteins from microbial pathogens thus form an important component of the macromolecular crystallography activities in the country.

## 5 Concluding Remarks

From humble beginnings in the eighties of the last century, macromolecular crystallographic studies in India have grown to become an important component of the Indian scientific effort. Thanks to the support of not only DST, but also of the Department of Biotechnology (DBT) and the Council of Scientific and Industrial Research (CSIR), excellent facilities are now available in most of the concerned laboratories. One major problem in this context is the absence of a state of the art synchrotron facility in the country. This problem has been partly alleviated by part ownership of beamlines elsewhere. Such an arrangement is already functioning well at the European Synchrotron Radiation Facility at Grenoble, France. A similar arrangement would soon become operational at the Elettra Synchrotron Facility at Trieste, Italy. The Indus 2 synchrotron at Indore has also just become operational. Although it is not quite as powerful as most of the machines currently being used elsewhere, efforts are underway to use it as efficiently as possible. Indus 2 is of considerable importance as a technology demonstrator and as a platform for training as well. Happily, there is also a consensus in the scientific community of India that we should soon start building a new, state of the art machine.

Macromolecular crystallography in India, although fairly strong already, is still in the expanding mode. The work on mycobacterial proteins is an example of a loosely coupled, yet concerted effort. More such efforts are likely to emerge. Structure-based inhibitor design aimed at drug development is already being pursued in a couple of laboratories in the country. Work in this direction is expected to gather momentum in the near future. Other new lines of investigation also appear to be on the anvil. Having had occasion to closely observe the efforts of young macromolecular crystallographers in the country, I am confident that they are poised to scale greater heights in the years to come.

Received 12 December 2013.

## References†

1. Bernal, J. D., Crowfoot, D. X-ray photographs of crystal-line pepsin. *Nature* **133**, 794–795 (1934).
2. Crowfoot, D. X-ray single crystal photographs of insulin. *Nature* **135**, 591–592 (1935).

<sup>†</sup> Several hundred publications have arisen from macromolecular crystallography work in the country. Only a few critical ones pertaining to early efforts are referred to here. A full list of publications till 2010 are available at <http://mbu.iisc.ernet.in/~mvlab/history.html> and <http://iris.physics.iisc.ernet.in/ica/mmc.pdf>

3. Bernal J. D., Fankuchen I., Perutz M. F. An X-ray study of chymotrypsin and haemoglobin. *Nature* **141**, 523–524 (1938).
4. Pauling, L., Corey, R. B., Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of polypeptide chain. *Proc. Natl. Acad. Sci. USA* **37**, 205–211 (1951).
5. Pauling, L., Corey, R. B. The pleated sheet. A new layer configuration of polypeptide chains. *Proc. Natl. Acad. Sci. USA* **37**, 251–256 (1951).
6. Watson, J. D., Crick, F. H. C. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
7. Ramachandran, G. N., Kartha, G. Structure of collagen. *Nature* **176**, 593–595 (1955).
8. Green, D. W., Ingram, V. M. and Perutz, M. F. The structure of haemoglobin IV. Sign determination by the isomorphous replacement method. *Proc. Roy. Soc. Series A*, **225**, 287–307 (1954).
9. Kendrew, J. C., Dickerson, R. E., Strandburg, B. E., Hart, R. G., Davies, D. R. Phillips, D. C. and Shore, V. C. Structure of myoglobin. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422–427 (1960).
10. Perutz, M. F., Rossmann, M. G., Cullis Ann F., Muirhead, H., George, W., North, A. C. T. Structure of Hemoglobin. A three dimensional Fourier synthesis at 5.5 Å resolution. *Nature* **185**, 416–422 (1960).
11. Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C., Sarma, V. R. Structure of hen egg white lysozyme: A three dimensional Fourier synthesis at 2 Å resolution. *Nature* **206**, 757–761 (1965).
12. Kartha, G., Bello, J. and Harker, D. Tertiary structure of ribonuclease. *Nature* **213**, 862–865 (1967).
13. Phillips, D., Blow, D., Hartley, B., Lowe, G. A discussion on the structures and functions of proteolytic enzymes. *Phil. Trans. Royal Soc. London Series B* **257**, 63–266 (1970).
14. Adams, M. J., Blundell, T. L., Dodson, E. J., Dodson, G. G., Vijayan, M., Baker, E. N., Harding, M. M., Hodgkin, D. C., Rimmer, B., Sheat, S. Structure of rhombohedral 2 Zinc insulin crystals. *Nature* **224**, 491–495 (1969).
15. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A. and Vijayan, M. Atomic positions in rhombohedral 2 zinc insulin crystals. *Nature* **231**, 506–511 (1971).
16. Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
17. Rose, P. W., Chunxiao, B., Bluhm, W. F., Christie, C. H., Dimitropoulos, D., Dutta, S., Green, R. K., Goodsell, D. S., Prlic, A., Quesada, M., Quinn, G. B., Ramos, A. G., Westbrook, J. D., Young, J., Zardeck, C., Berman, H. M. and Bourne, P. E. The RCSB Protein Data Bank: new resources for research and education. *Nucl. Acids Res.* **41**, D475–D482 (2013).
18. Vijayan, M., Johnson, L. N. Gopalasamudram Narayana Ramachandran. *Biogr. Mem. Fell. R. Soc.* **51**, 367–377 (2005).
19. Banerjee, R., Shekhar, S. C., Ganesh, V., Das, K., Dhanraj, V., Mahanta, S. K., Suguna, K., Surolia, A. and Vijayan, M. Crystal structure of peanut lectin, a protein with an unusual quaternary structure. *Proc. Natl. Acad. Sci. USA* **91**, 227–231 (1994).
20. Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A., Vijayan, M. Conformation, protein-carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin lactose complex. *J. Mol. Biol.* **259**, 281–296 (1996).
21. Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A., Vijayan, M. A novel mode of carbohydrate recognition in jacalin, a *Moraceae* plant lectin with a β-prism fold. *Nat. Struc. Biol.* **3**, 596–603 (1996).
22. Bhuvaneswari, M., Subramanya, H. S., Gopinath, K., Savithri, H. S., Nayudu, M. V. and Murthy, M. R. N. Structure of sesbania mosaic virus at 3.0 Å resolution. *Structure* **3**, 1021–1030 (1995).
23. Sri Krishna, S., Hiremath, C. N., Munshi, S. K., Prahadesswaran, D., Sastri, M., Savithri, H. S., Murthy, M. R. N. Three dimensional structure of physalis mottle virus: Implications for viral assembly. *J. Mol. Biol.* **289**, 919–934 (1999).
24. Chakravarthy, S. and Kannan, K. K. Drug-Protein interactions: Refined structures of three sulfonamide drug complexes of human carbonic anhydrase I enzyme. *J. Mol. Biol.* **243**, 298–309 (1994).
25. Vijayan, M. Macromolecular crystallography in India. A historical overview. *J. Indian Inst. Sci.* **87**, 261–277 (2007).
26. Vijayan, M. The legacy of G. N. Ramachandran and the development of structural biology in India, In: Bansal, M., Srinivasan, N. (eds). Biomolecular forms and functions, IISc Press, World Scientific, Singapore, pp 1–16 (2013).
27. Vijayan, M. Form and function of proteins. Historical background and the Indian effort in macromolecular crystallography. *Nat. Acad. Sci. Lett.* DOI 10.1007/s40009-013-0220-5 (2013).
28. Ajay Kumar, R., Vaze, M. B., Chandra, N. R., Vijayan, M. and Muniyappa, K. Functional characterization of the precursor and spliced forms of RecA protein of *Mycobacterium tuberculosis*. *Biochemistry* **35**, 1793–1802 (1996).
29. Bachhawat, N. and Mande, S. C. Identification of the *INO1* gene of *Mycobacterium tuberculosis* H37Rv reveals a novel class of inositol-1-phosphate synthase enzyme. *J. Mol. Biol.* **291**, 531–536 (1999).
30. Datta, S., Prabu, M. M., Vaze, M. B., Ganesh, N., Chandra, N.R., Muniyappa, K. and Vijayan, M. Crystal structure of *Mycobacterium tuberculosis* RecA and its complex with ADP-AIF<sub>4</sub>: implications for decreased ATPase activity and molecular aggregation. *Nucleic Acids Res.* **28**, 4964–4973 (2000).



**M. Vijayan** was born in 1941. He took his Ph.D degree from the Indian Institute of Science, Bangalore in 1967. He was a post doctoral fellow with Dorothy Hodgkin at Oxford during 1968–71 and a visiting fellow with her during 1976–77, working on the structure of insulin. Except for these stays at Oxford, he has all along been at the Indian Institute of Science. He has been largely responsible for the development of macromolecular crystallography in India. His personal research has encompassed plant lectins, protein hydration, mycobacterial proteins and chemical evolution and origin of life. He is a fellow of all the three science academies of the country and The World Academy of Sciences (TWAS). He has received many awards starting with the Bhatnagar Prize and worked in several national and international bodies. Among other things, he has been the President of the Indian Biophysical Society, founder President of the Indian Crystallographic Association, the Chairman of the IUCr Commission on Biological Macromolecules, member of IUPAB Council, President of the Asian Crystallographic Association and President of the Indian National Science Academy. Currently he is the INSA Albert Einstein Professor.