

Macromolecular Crystallography in India, IUCr, 2017

Macromolecular Crystallography, a key tool of Structural Biology, is well entrenched in India, and a large number of research groups within both government and private sectors are actively engaged in it. It is estimated that there are > 50 independent investigators in India whose core themes of research utilize macromolecular crystallography (MX) to enlighten their interests in modern biology. Highly diverse sets of biological problems, from equally diverse organisms, are being actively pursued by the Indian MX community. Their achievements have been summarised on different occasions, and several of our esteemed colleagues have written wonderful summaries of Indian presence in this field. Many of these documents are available on this and other sites (IUCr 2017 site: [India/Crystallography/Perspectives/Publications](#), and 1-4).

Below we have distilled some of the key thrusts areas from four independent research laboratories in India. These are exemplary examples of a balance between fundamental and translational research using macromolecular crystallography. We apologize to all other MX members of India whose work we have been unable to highlight here due to paucity of space and time. The selected areas of macromolecular crystallography research highlighted here simply reflect the overall achievements, breadth, depth and focus of the Indian MX community presently.

Chiral proofreading during protein translation– Dr. Rajan Sankaranarayanan, CCMB

It may seem incredible yet evidence suggests that the ubiquitous enzyme family of aminoacyl-tRNA synthetases can every now and then attach a chirally wrong amino acid to tRNA during the evolutionarily conserved process of protein translation. Hence, nature, never behind us, has devised proofreading mechanisms that correct these errors. The structural biology of enzymes and domains that are key players in this job of cleaning up mistakes made by housekeeping enzymes during protein translation has been of deep interest to Dr. Sankaranarayanan's group. He has been at the Centre for Cellular and Molecular Biology (CCMB) Hyderabad for over a decade, and during this time has made numerous exceptionally insightful observations of the structural underpinnings of proofreading in biology. For example, early on, his group made a fundamental discovery by

identifying a D-aminoacyl-tRNA deacylase (DTD) scaffold that was attached to the protein translational apparatus (5-7). This allowed an understanding of the selectivity of L-amino acids over D-amino acids during early evolution of the genetic code. His group addressed the prevalent dogma of the field that proofreading was due to rejection of cognate substrates from binding, and revealed that 'functional positioning' rather than 'steric exclusion' was the vital discriminatory force, thus providing an alternative mechanistic scenario to the text book 'Double-sieve model' for proofreading (5-7). His group has also uncovered the long sought after mechanism of D-amino acid rejection during protein synthesis that prevents the opposite chiral molecules from infiltrating the translational machinery. The work revealed an unprecedented mechanism by which DTD performs this '*Chiral Proofreading*' task, a term they have introduced to the biological text (5-8). The study provides deeper insights into how a single DTD deals with the incredible challenge of being an absolute D-configuration specific enzyme, thus excluding the vast excess of diverse L-aminoacyl-tRNA pool present in the cellular context. More recently, his group addressed the birth of first protein enzymes at the RNA-protein/peptide interface. They identified that the DTD-fold, present as the proofreading domain of a tRNA synthetase in archaea, does not use side chains for the two hallmark activities of enzymes, i.e. catalysis and substrate specificity. This presented the first ever evidence of an enzyme that differentially remodels its RNA-protein interface for performing its enzymatic function (9). Very recently, his group has shown that DTDs '*Chiral Proofreading*' site is designed only for rejection of L-amino acids attached to tRNA and not for selection of D-aminoacyl-tRNA. Remarkably, it has been shown using both *in vitro* and *in vivo* studies that 'achiral' glycyl-tRNA is misedited as effectively as the D-aminoacyl-tRNA substrates resulting in glycine '*misediting paradox*'. The work demonstrates the need for tighter regulation of DTD levels in all organisms without which the protection of glycyl-tRNA(Gly) offered by Elongation Factor(Tu) will be relieved leading to cellular toxicity (10).

DNA modification enzymes - Saikrishnan Kayarat, IISER Pune

Dr. Saikrishnan Kayarat's in IISER, Pune, is focused on how modular and multifunctional enzymes, often referred to as macromolecular machines, orchestrate their varied activities to carry out a specific cellular function. For the past six years, his group has established the nucleoside triphosphate (NTP)-dependent restriction-modification (RM) systems as model systems to study the mechanism of helicase-like NTPase-driven macromolecular machines involved in nucleic acid transactions. NTP-dependent RM enzymes are one of the most prominent bacterial defense systems. These enzymes are complex multifunctional protein machines that recognize and bind to a specific DNA sequence, perform sequence-specific methylation (modification), and nucleolytically cleave the DNA (restriction) subsequent to translocation of the DNA by hydrolysis of NTP. Modification of host DNA protects it from restriction, while unmodified foreign DNA (bacteriophage DNA) is cleaved. The DNA target recognition domain, methyltransferase, DNA translocation motor (NTPase) and nuclease constitute the RM system and are present in many other protein machines which are involved in varied activities across the three domains of life.

Structure determination of a whole NTP-dependent RM enzyme had been an unaccomplished challenge primarily due to its enormous size and dynamic nature of the enzyme complex. Recently, his group has determined the crystal structure of NTP-dependent RM enzymes LlaBIII and LlaGI bound to DNA substrate mimics (11,12). LlaBIII-DNA complex revealed an ATP-dependent RM enzyme bound to a dsDNA containing the target DNA sequence. Upon binding, the DNA at the target sequence partially unwinds facilitating the reading of the 6 bp long sequence by the enzyme. The adenine base of the target sequence that undergoes methylation was flipped out into the active site of the methyltransferase. The conformational change in the DNA steered it towards the helicase-like ATPase motor of the enzyme. This snapshot provided unique insights into the molecular basis of how target recognition is coupled to the methyltransferase and translocase activity of the enzyme through the conformational change in the substrate DNA (11).

A surprising finding was that the nuclease and the ATPase domain were located upstream of the direction of translocation and cleavage, indicating that the mechanism of

DNA cleavage was different from the earlier proposal that the ATPase loops out DNA resulting in convergence of two enzymes bound to the DNA (the convergence results in dimerization of the two nucleases that nick each of the strands to cause dsDNA break). However, the structure clearly suggested that the convergence of the enzymes would leave the nucleases distal from one another. In collaboration with Mark Szczelkun at University of Bristol, Sai's group carried out single-molecule magnetic tweezers experiments, and sequencing the products of single cleavage events revealed that the enzymes made multiple nicks on the DNA that resulted in dsDNA break formation (11, 13). Physiologically, this mode of cleavage is more potent as it is more difficult to repair such breaks. The consensus code obtained by this study can serve as a platform in engineering new specificities in Type IIS and related types of restriction enzymes and methyltransferases. This structure also provided the first snapshot for the mode of binding of the ATPase to dsDNA. The SF2 motors power a variety of cellular processes, including chromatin remodeling, transcription modulation and DNA repair. His group is now working towards deciphering the mechanism of translocation of these motors on a dsDNA track using a combination of X-ray crystallography and electron cryo-microscopy, complemented by single-molecule biophysical studies.

Biology of G Protein-Coupled Receptors – Arun Shukla, IIT, Kanpur

Cells in our body continuously encounter numerous stimuli to which they respond by eliciting appropriate signalling and cellular responses. Integral membrane proteins embedded in the plasma membrane bilayer are major nodes for receiving a diverse array of “signals” and transmitting the “message” across the membrane. The extensively studied class of integral membrane proteins is G Protein-Coupled Receptors (GPCRs), also known as seven transmembrane receptors (7TMRs) constitute the largest family of cell surface receptors in the human genome. GPCRs are involved in almost every physiological and pathophysiological process in our body and about ~40-50% of the marketed drugs target these receptors (14).

Agonist stimulation of GPCRs leads to coupling of heterotrimeric G proteins, generation of secondary messengers (such as cyclic AMP, inositol phosphates and Ca^{2+}) and onset of multiple signalling pathways. Activated GPCRs are phosphorylated by G Protein-Coupled Receptor Kinases (GRKs) followed by recruitment of multifunctional scaffolding proteins, β -arrestins (β arrs) which sterically hinders further G protein coupling and result in GPCR desensitization. Overall architecture, signalling and regulatory paradigms of GPCRs are remarkably well conserved and have three distinct hallmarks, i) seven transmembrane (7TM) helix architecture, ii) signalling through heterotrimeric G Proteins and iii) down regulation and signaling through β arrs.

The major focus of Dr. Arun Shukla's group in Department of Biological Sciences and Bioengineering, IIT, Kanpur is to obtain high-resolution crystal structures of selected GPCRs, understand their dynamic process of activation and to decipher the biophysical details of their interaction with effectors. Dr. Shukla's group has recently analyzed high-resolution crystal structures of GPCRs bound to a variety of ligands ranging from antagonists, agonists, biased ligands to allosteric ligands (15) and identified the diversity in ligand binding pockets of different receptors. This diversity in receptor-ligand interaction results in significantly different conformational states of the receptors which in turn is responsible for determining the differential functional outcomes and signaling. His group has also discovered recently, contrary to generally believed notion in the field of GPCR biology, that even partially engaged GPCR-effector complexes can be functionally competent (16). Based on the structural model of a GPCR- β arr complex, the 3rd intracellular loop of β 2 adrenergic receptor (β 2AR) was truncated and it resulted in generation of a β 2AR- β arr complex engaged solely through the

carboxyl-terminus of the receptor. Surprisingly, this partially engaged complex robustly mediated receptor endocytosis and ERK MAP kinase signaling. This changes the long standing dogma in the field that complete engagement of GPCRs with β arrs is essential for triggering functional responses. Furthermore, Dr. Shukla's group has also observed that GPCR ligands capable of eliciting selective signaling through β arr pathway, refer to as biased ligands, trigger the formation of partially engaged complexes in cellular context. Going forward, his group is generating synthetic protein binders such as antibody fragments against GPCRs and their effectors (e.g. β arrs) as a novel strategy to allosterically modulate their activation and signaling in cellular context. These synthetic binders may also serve as crystallization chaperones by stabilizing these receptors and promoting their crystallogenesis.

Aminoacyl tRNA synthetases as druggable targets – Amit Sharma, ICGEB, New Delhi

Several eukaryotic parasites cause diseases like malaria, leishmaniasis, toxoplasmosis, cryptosporidiosis and coccidiosis. These together affect ~600 million people worldwide with estimated deaths of ~500,000 annually. These diseases are caused by parasites such as Plasmodium, Leishmania, Toxoplasma, Cryptosporidium and Eimeria. Within the genomes of these parasites a high degree of conservation is evident in their housekeeping genes, and especially in their tRNA synthetases. This thus underlies the Dr Sharma's thrust that few or several of these parasitic diseases can be controlled via the inhibition of orthologous protein targets that are in themselves druggable. Although seemingly a facile idea, in practise gathering structural and functional evidence for this has been tedious. In 2009, Dr Sharma's group published a report highlighting why the entire tRNA synthetase family from malaria parasites could serve as very valuable new drug targets (17). This seminal paper thence not only set up the field of drugging malaria parasite tRNA synthetases, but also extrapolating these studies to related and unrelated eukaryotic parasites. A series of breakthroughs, both in Dr. Sharma's group and in many other labs worldwide have since validated the insights of 2009 paper (17). Indeed today, the tRNA synthetase enzyme family is one of the most sought after in terms of proving new lead molecules that can be considered for further development as anti-malarials, and indeed in some cases as anti-infectives. Since 2009, Dr. Sharma's group has focused on numerous tRNA synthetases from *P. falciparum* and has provided the structural basis for either their exploitation for drugging or for their interactions with drug-like molecules (18-28). In 2016, a breakthrough was achieved in collaboration with several international groups and Amit's laboratory in targeting the phenylalanyl tRNA synthetase with a novel compound that possess very potent anti-malarial activity (28).

An early example of this is the work on cladosporin - a fungal sourced metabolite found in *Aspergillus flavus* and *Cladosporium cladosporioides* that inhibits the parasite lysyl-tRNA synthetase with high potency and selectivity. Crystal structures of lysyl-tRNA synthetase (KRS) in complex with cladosporin and L-lysine from *Plasmodium falciparum* and *Loa Loa* have been determined, which revealed that cladosporin is housed in the canonical ATP binding pocket (20, 21, 27). Two residues, V328 and S344, at the bottom of the ATP-binding pocket interact with the methyl-tetrahydropyran group of cladosporin and seem critical for its specificity for KRS over other class II tRNA synthetase families. In Plasmodium,

Loa Loa and *Schistosoma mansoni* these positions are occupied by a valine and serine residue respectively, but in humans with larger residues and hence the selectivity. Based on the identity of the Plasmodium amino acid residues, smaller side chains are favoured due to decreased steric hindrance (27). The cooperative binding of cognate L-Lys amino acid and cladosporin together allow for highly specific inhibition through the universal pocket for ATP.

A second example is based on the compound febrifugine from the Chinese herb *Dichroa febrifuga*. Many groups worldwide have shown that the active component febrifugine and its derivatives such as halofuginone are potent anti-malarials. Amit's more recent cell-, enzyme-, and structure-based data on febrifugine scaffold inhibitors strengthens the case for their inclusion in anti-malarial and anti-parasitic drug development efforts (22, 23). His group has determined structures for ~12 parasitic prolyl-tRNA synthetases in complexes with halofuginone or its derivatives. Halofuginone occupies the 3' end tRNA A76 and L-proline sites as it is a dual site inhibitor. These studies suggest that a common scaffold like febrifugine may in its derivatized forms provide drug-like molecules that are capable of binding to several eukaryotic prolyl-tRNA synthetases, cause their inhibition, stop protein synthesis and eventually lead to cell (parasite) death.

Amit's group has additionally provided data on moon-lighting functions of two *Plasmodium falciparum* tRNA synthetases - in addition to their canonical housekeeping protein translation functions. The chemi-luminescence of parasite culture supernatants and confocal microscopy data revealed that tyrosyl-tRNA synthetase (PfTyrRS) exited from the parasite cytoplasm into the infected red blood cell (iRBC) cytoplasm, from where it could be released into the extracellular medium on iRBC lysis. The interactions and internalization of PfTyrRS into host macrophage and alterations in the pro-inflammatory cytokines like TNF- α , IL-1 and IL-6 levels were also established (18). Similarly, the hemin binding ability of *Plasmodium falciparum* arginyl-tRNA synthetase (PfRRS) was established recently, where this interaction inhibits the arginyl-tRNA synthetase which thus dampens protein synthesis by the parasite (24).

References

1. Vijayan, M. The legacy of G. N. Ramachandran and the development of structural biology in India. **Current Science** 2016 110, 535-542.
2. Vijayan, M. Macromolecular Crystallography in India in the historical context. **J Proteins Proteomics** 2013. 4, 83-84.
3. Vijayan, M. Macromolecular Crystallography in India in the Global Context. In: **J. Indian Inst. Sci.** 2014, 94 (1, SI). pp. 103-108.
4. Arora A, et al. Structural biology of Mycobacterium tuberculosis proteins: the Indian efforts. **Tuberculosis** 2011, 456-468.
5. Dwivedi S, et al. A D-amino acid editing module coupled to the translational apparatus in archaea. **Nature Struct. Mol. Biol.** 2005, 12, 556-557.
6. Hussain T, et al. Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea. **EMBO J.** 2006, 25, 4152-4162.
7. Hussain T, et al. Mechanistic insights into cognate substrate discrimination during proofreading in translation. **Proc. Natl. Acad. Sci. (USA)** 2010, 107, 22117-22121.
8. Ahmad S, et al. Mechanism of chiral proofreading during translation of the genetic code, **eLife** 2013, 2:e01519.
9. Ahmad S, et al. Specificity and catalysis hardwired at the RNA-protein interface in a translational proofreading enzyme. **Nature Commun.** 2015, 6:7552.
10. Routh SB, et al. Elongation factor Tu prevents misediting of Gly-tRNA(Gly) caused by the design behind the chiral proofreading site of D-aminoacyl-tRNA deacylase. **PLOS Biol.** 2016, 14(5) e1002465, 1-22.
11. Chand, MK, et al. Translocation-coupled DNA cleavage by the Type ISP restriction-modification enzymes. **Nature Chemical Biology** 2015, 11, 870-877.
12. Kulkarni, M, et al. Structural insights into DNA sequence recognition by Type ISP restriction-modification enzymes. **Nucleic Acids Res.** 2016, 44, 4396-4408.
13. van Aelst, K, et al. Mapping DNA cleavage by the Type ISP restriction-modification enzymes following long-range communication between DNA sites in different orientations. **Nucleic Acids Res.** 2015, 43, 10430-10443.
14. Kumari P, et al. Emerging Approaches to GPCR Ligand Screening for Drug Discovery. **Trends Mol Med.** 2015 Nov;21(11):687-701.
15. Ghosh E, et al. SnapShot: GPCR-Ligand Interactions. **Cell** 2014 Dec 18;159(7):1712, 1712.e1.
16. Kumari P, et al. Functional competence of a partially engaged GPCR- β -arrestin complex. **Nat Commun.** 2016 Nov 9;7:13416
17. Bhatt TK, et al. A genomic glimpse of aminoacyl-tRNA synthetases in malaria parasite Plasmodium falciparum. **BMC Genomics** 2009 Dec 31;10:644.
18. Bhatt TK, et al. Malaria parasite tyrosyl-tRNA synthetase secretion triggers pro-inflammatory responses. **Nat Commun.** 2011 Nov 8;2:530.
19. Khan S, et al. An appended domain results in an unusual architecture for malaria parasite tryptophanyl-tRNA synthetase. **PLoS One** 2013 Jun 12;8(6):e66224.
20. Khan S, et al. Structural analysis of malaria-parasite lysyl-tRNA synthetase provides a platform for drug development. **Acta Crystallogr D Biol Crystallogr.** 2013 May;69(Pt 5):785-95.

21. Khan S, et al. Structural basis of malaria parasite lysyl-tRNA synthetase inhibition by cladosporin. *J Struct Funct Genomics* 2014 Jun;15(2):63-71.
22. Jain V, et al. Structural and functional analysis of the anti-malarial drug target prolyl-tRNA synthetase. *J Struct Funct Genomics* 2014 Dec;15(4):181-90.
23. Jain V, et al. Structure of Prolyl-tRNA Synthetase-Halofuginone Complex Provides Basis for Development of Drugs against Malaria and Toxoplasmosis. *Structure* 2015 May 5;23(5):819-29.
24. Jain V, et al. Dimerization of Arginyl-tRNA Synthetase by Free Heme Drives Its Inactivation in Plasmodium falciparum. *Structure* 2016 Sep 6;24(9):1476-87.
25. Hussain T, et al. Inhibition of protein synthesis and malaria parasite development by drug targeting of methionyl-tRNA synthetases. *Antimicrob Agents Chemother.* 2015 Apr;59(4):1856-67.
26. Sharma A, Sharma A. Plasmodium falciparum mitochondria import tRNAs along with an active phenylalanyl-tRNA synthetase. *Biochem J.* 2015 Feb 1;465(3):459-69.
27. Sharma A, et al. Protein translation enzyme lysyl-tRNA synthetase presents a new target for drug development against causative agents of Loiasis and Schistosomiasis. *PLoS Negl Trop Dis.* 2016 Nov 2;10(11):e0005084.
28. Kato N, et al. Sharma A, Winzeler EA, Wirth DF, Scherer CA, Schreiber SL. Diversity-oriented synthesis yields novel multistage antimalarial inhibitors. *Nature* 2016 Sep 7;538(7625):344-349.

Authors:

Manickam Yogavel and Amit Sharma

Affiliation:

Molecular Medicine Group, ICGEB, New Delhi

Emails:

myogavel@gmail.com and amit.icgeb@gmail.com

For IUCr 2017