Editorial

Philip E. Bourne and Keith Watenpaugh

The International Union of Crystallography (IUCr) Macromolecular Crystallography Computing School was held at Western Washington University in Bellingham, Washington between August 17 and 22, 1996. The School was the seventh in a series of IUCr Crystallographic Symposia. There were 106 attendees from 16 countries; of these 38 were either speakers or tutors. Of the remaining 68 attendees, 29 were graduate students, 12 were postdoctoral fellows, 9 were faculty, and 18 were from industry or elsewhere. The format of the School was formal lectures in the morning, tutorials in the afternoon, and software demonstrations and more lectures in the evening.

Draft proceedings were distributed to attendees of the School and final papers appear in the published proceedings. For the first time at a Computing School full Proceedings are available electronically via the World Wide *Web (http://www.sdsc.edu/Xtal/IUCr/CC/School96/IUCr.html*) or via anonymous ftp from *ftp.sdsc.edu* in the directory */pub/sdsc/societies/IUCr/School96.*

The full program which left both the organizers and attendees exhausted, reflects the current state of excitement in the field of macromolecular structure determination using the technique of X-ray crystallography. Our understanding of biological processes has gained immeasurably from work already done, and as shown in Figure 1 of the paper presented by Stewart, we are climbing rapidly along an exponential growth curve that promises many more major discoveries before the next IUCr Congress in 1999. It is estimated that at the 1999 Congress the approximately 5,000 structures solved at the time of the 1996 School will have grown to over 20,000. The new and improved technologies and techniques described in these Proceedings are contributing to that growth and at the same time, as pointed out in the paper given by Sussman, creating challenges for the Protein Data Bank (PDB).

As the School progressed, we were struck by the similarities to events which took place in small molecule crystallography beginning some 20 to 25 years ago. Growth then was fueled by the advent of new algorithms, affordable computer hardware, and good software. So it is today for macromolecular crystallography, but with the added bonus of the Internet which is changing how we conduct our research. Flack presented this view as part of his on-going contribution to how crystallographers use the Internet.

As the number of available structures increase, thoughts turn to relative accuracy and common and different features between structures. The relative ease with which the majority of structures are solved today, effects how we work and how we think about each individual structure and structures en masse. Thornton and Wodak in there papers focused on comparative analysis of proteins with regard to accuracy, and Thornton, Bryant, and Bourne considered the problem of classification and comparison of proteins. With so many structures, databases are becoming vital weapons in the crystallographers armory. Gilliland , Sussman , and John Westbrook described their work with the crystallization database, 3DB, and the Nucleic Acid Database, respectively. Good databases require a complete and computer usable representation and hence Fitzgerald and John Westbrook provided an overview of the macromolecular Crystallographic Information File (mmCIF) and associated tools, respectively.

After presentations discussing structures *en masse* we returned to the more traditional mode of presentation which parallels the determination of a single macromolecular structure: data collection - phasing- model building and visualization - refinement.

Howard provided attendees with an overview of the steps in the collection of macromolecular data. This was followed by a presentation from Minor discussing recent advances and the even more important issue of accuracy. Related to accuracy (and sanity) is the development of graphical user interfaces (GUIs) which reduce mistakes when collecting data. Sweet described the GUI they have developed for the X12-C line at Brookhaven National Laboratory. Finally Mary Westbrook described the software system being developed for data acquisition at the Structural Biology Center (SBC), located at Argonne National Laboratory's Advanced Photon Source, which was just coming on-line at the time of her presentation.

The discussion on phasing techniques for macromolecules began with the Plenary lecture delivered by Hauptman who explained how least squares can lead to a formulation of the minimal principle and effectively replaces the phase problem by one of constrained global optimization. The Shake-and-Bake software addresses this reformulation and was the subject of a presentation by Weeks who highlighted the successful application of Shake-and-Bake in *ab initio* phasing of proteins of 600 non-hydrogen atoms or less where data to greater than 1.2 Å are available. Hauptman went on to describe how the use of single wavelength anomalous scattering (SAS) diffraction data leads to yet another formulation in global optimization and hence solution to the phase problem. Giacovazzo took a more optimistic view of the power of direct methods as applied to phasing protein structures, and went as far as to suggest that direct methods is competitive with single isomorphous replacement techniques. If low resolution phases are available, Main described how they could be used in a rotation function yielding better results that the standard rotation function that operates in vector space. Also on the topic of phase extension Collins described work using an objective function and Bricogne described maximum likelihood heavy atom parameter refinement for the MIR and MAD methods.

On the subject of model building and visualization, a paper delivered by Edwards on behalf of Oldfield described the procedures available in QUANTA 96 for the semi-automated fitting of electron density maps, a time consuming step in the structure solution process. Finzel described the use of existing structural information for map fitting in the LORE program and Nicholls described the representation and the beginnings of the comparative analysis of protein surfaces. Features embodied in the program GRASP.

On the subject of refinement Sheldrick described the use of SHELXL in the refinement of proteins using high resolution data. He touched upon issues of constraints versus restraints, restrained anisotropic refinement, the role of Rfree, disorder, the solvent model, radius of convergence and esd's. Ten Eyck described the role of full matrix least squares refinement for macromolecules while Bricogne described the use of maximum likelihood as opposed to least squares in refinement using a combination of the programs BUSTER and TNT. Chapman returned us to the use of real space refinement, first pioneered by Diamond in 1971, but now using stereochemical restraints in the refinement of large virus structures and fitting know structures to low resolution electron microscope reconstructions. Finally Tronrud introduced the notion of joint refinement using Thermolysin inhibitors. Restraints are constructed based on regions of the multiple structures which are invariant versus those regions which are variant - an increasingly common occurrence.

The latest developments in software used by many of us were described by their authors. McRee described Xtalview, Steigemann described the PROTEIN system, concentrating on molecular replacement and density modification in real space, Dodson described the CCP-4 package, Turk described MAIN, and Sheldrick described the application of SHELX to macromolecules.

Finally, a session was devoted to either new approaches or ancillary techniques. In the former category Johnson described work in crystallographic topology, and in the latter Moss described the application of object oriented programming to crystallography, Wild described an AVS interface to the CCP-4 programs, and Ramanadham described the parallelization of crystallographic codes.

Knowing that crystallographers excel at both work and play, Nicholls and Bourne used the receptive post-banquet audience to reminisce about advances in the field. Two apologies are due, one to Monty Python for modyfying their "Three Yorkshireman" sketch, and one to Sheldrick - we never would have guessed he was from Yorkshire.

These Computing Schools remain the most important function of the IUCr Computing Commission in educating young crystallographers in the foundations of computational crystallography and in also bringing us not-so-young crystallographers up to speed on the latest developments. This school was not meant to train crystallographers in the use of normally available crystallographic software, but to introduce them to the latest developments in the field and what new directions are being pursued. It is our hope that seeds of new ideas were planted in the minds of the attendees and that the students of this school will be the tutors and lecturers of the subsequent schools presenting exciting new developments. Also, it is our hope that these Proceedings will enable those not in attendence to develop future contributions to computational crystallography and to share the excitement generated at the School. Finally, we thank the authors and attendees for making this an outstanding School.