

W0308

Teaching Elves to Collect Data: An Analysis of the Last Million Diffraction Images from ALS 8.3.1.
James Holton, Physical Biosciences, Lawrence Berkeley National Laboratory, Berkeley, CA 94720.

Most X-ray data sets collected at synchrotron sources do not produce usable results. An analysis of data collected in 2003 at the ALS beamline 8.3.1 shows that 2346 datasets were collected and 41 structures were deposited in the PDB. Although it is understandable that not every dataset leads to a published structure, it is troubling that ~98% of them do not. This large gap between collected data and useful results is not unique to 8.3.1. The 28 operating American PX beamlines collect ~100,000 datasets/year. This suggests that a great deal of improvement in scientific productivity can be attained if the reasons for failed projects are better understood.

I will describe the efforts at the ALS beamline 8.3.1 to address this gap. We are developing rapid and reliable means of measuring the success rate of each stage in the structure determination process. These statistics are helping us identify the nature and relative importance of stumbling blocks and design new strategies and algorithms that avoid or address the most important problems. The fruits of these efforts are the glue for integrating BLU-ICE, Elves, robotics and other "off the shelf" components to provide a maximally smooth and productive user experience.

Preliminary results suggest that rapid radiolysis of heavy atom sites is one of the most significant reasons for the failure of de-novo structures. Another prevalent problem is insufficient signal-to-noise ratio to determine heavy atom phases. Spot overlap resulting from a long unit cell plays the most significant role in the failure of data reduction. Data collection strategies for avoiding and/or minimizing these high impact problems will be discussed.