

## W0188

**PDB Depositions and Statistics From Synchrotron Sources.** Jiansheng Jiang. Biology Dept., Brookhaven National Laboratory, 50 Bell Ave., Upton, NY 11973 USA.

An analysis of Protein Data Bank (PDB) depositions from synchrotron radiation facilities (<http://asdp.bnl.gov/asda/Libraries/>) shows that worldwide, every year since 1999, more than 50% of the deposited x-ray structures have used the synchrotron facilities, reaching 75% by 2003. Structures from synchrotron source are general larger than those from home source (738 residues verse 453 residues in average). The impact of structural genomics and high-throughput methods are observed. Statistics regarding to the quality of experimental data and x-ray structures are presented. The average resolution is 2.2Å with a standard deviation of 0.5Å over 17,144 structures. Only 6% structures have the higher resolution than 1.5Å and 4% structures have the resolutions lower than 3.0Å. The reported R-values have a mean of 0.20 ( $\sigma=0.03$ ), and the reported Free R-values have a mean of 0.25 ( $\sigma=0.04$ ). The distribution of mean Free R-values is found best to fit to a simple exponential function. An empirical formula as a function of resolution is proposed for estimating the expected Free R-value and structure validation.

$$R_{free} = R_t \exp(-b/d_{min}),$$

where  $R_t=0.40$ ,  $b=0.96$ , and  $d_{min}$  is the resolution (Å). The structures with the Free R-value that exceeds  $R_{free} \pm 3\sigma$  should be validated.