X-ray Crystallography

Accuracy of Refined Biomolecular X-ray Structures in Doubt

X-ray crystallography is the most widely applied experimental method for biomolecular structure determination and is routinely used to ascertain how a pharmaceutical drug interacts with its protein target. However, a recent study published in *Nature Communications* suggests that current software programs used to predict a protein's structure from X-ray crystallography data dramatically underestimate the level of protein dynamics and do not give a realistic picture of the protein's structural heterogeneity.

For X-ray crystallography, proteins are purified and induced to form a crystal, each of which contains millions of copies of the same protein. In the second step, the crystal is placed in a beam of X-rays, which scatter when striking the crystal. By measuring the angles and intensities of these diffracted beams and applying a set of well-defined mathematical operations, one can determine the electron density distribution of the protein in question and, through it, its 3-dimensional structure with atomistic resolution. The whole process is carried out using highly advanced, specialized software and has, unlike protein purification and crystallization, in many cases reached the point of being almost routine undertaking.

Since each crystal contains a large, dynamic collection of proteins, crystallographic observables are only averages over time and space. The prominent way of obtaining and analyzing microscopic dynamics and heterogeneity in biomolecular X-ray crystallography is through isotropic and anisotropic B-factors. These are also referred to as temperature or Debye-Waller factors and they account for the thermal motion of an atom. B-factors are commonly used to distinguish mobile parts from rigid ones in X-ray structures and are, in the absence of other effects such as lattice imperfections, rigid-body motions or refinement artifacts, thought to be a quantitatively accurate measure of positional fluctuations of atoms.

**Status Quo Reconsidered**
Antonija Kuzmanic and Bojan Zagrovic, computational biologists from the Max F.
Perutz Laboratories of the University of Vienna and the Medical University of Vienna, and Navraj S. Pannu, a theoretical crystallographer from Leiden University, now provide a strong critique of this widely held assumption [1]. They investigate how averaging over many structurally different conformers in a protein crystal affects isotropic and anisotropic B-factors obtained in crystallographic refinement using state-of-the-art simulation and refinement approaches. In particular, they analyze a system in which, unlike in a typical crystallographic experiment, one has full control over microscopic variability and can directly manipulate it. More specifically, they use molecular dynamics simulations to manipulate in a controlled fashion the level of microscopic heterogeneity of one-of-the-largest protein crystal systems simulated to date: 27 unit cells with a total of 216 copies of villin headpiece with explicit solvent. Villin headpiece is a 35-residue 3-helix bundle protein and is a widely studied model systems in protein biophysics due to its small size and fast-folding properties [2]. Furthermore, they use rigorous methods to predict time- and ensemble-averaged X-ray structure factors from the simulated crystal. A comparison between the model derived from these structure factors using molecular replacement model refinement with the actual ensemble present in the simulated crystal allows the computational biologists to determine how accurately X-ray refinement capture the true structural diversity present at the microscopic level.

Remarkably, even at the high resolution of 1 ångström (Å) and an extremely low R \text{free} value of 5.9% (a measure of refinement quality calculated from a subset of reflections that were not included in refinement), the isotropic and anisotropic B-factors of atoms in several well-resolved residues drastically underestimate their true value known from simulation even up to six-fold with an average level of underestimation of approximately three-fold over the entire set. Most importantly, the actual B-factors are correctly captured by refinement only for highly immobile
atoms with B-factor values below approximately 10 Å². The simulation setup used allows the researchers to fully exclude the effect of other potential contributions to B-factors such as lattice defects or rigid body motions, suggesting that the above discrepancy is an intrinsic consequence of time- and ensemble-averaging implicit in crystallographic structure determination in combination with inadequate treatment of correlated motion in standard X-ray structure refinement.

**Implications for the Future**
The principal implication of these results is that in some cases the level of dynamics and structural heterogeneity in biomolecular crystals may be considerably larger than previously thought. This potentially carries great significance for the way we understand biomolecular structure and function in different context ranging from enzymatic catalysis to allosteric communication to non-covalent biomolecular interactions. Furthermore, the results relate to all situations in which experimental B-factors have been compared against other estimates and measures of protein dynamics.

The results of the study show that the models used to describe the richness of atomic displacements in crystallographic refinement may be improved, as suggested in other studies [3,4], and that there is potential for improvement with the help of molecular dynamics simulations. Regarding the recent advances at the hardware and software level in this field [5,6], rich structural ensembles of a given crystal could be readily created and used in the refinement. In the long run, a more realistic picture of protein structure and dynamics may be obtained, which could significantly improve our ability to develop potent medicines to modify protein function.

**References**

**Authors**
Bojan Zagrovic, Department of Structural and Computational Biology, Max F. Perutz Laboratories, Vienna
Contact

Max F. Perutz Laboratories
Dr. Bohrgasse 9 /6
1030 Wien
Österreich