"Multidimensional analysis of modulated structures of macromolecules using novel phonon corrections on the example of Hyp-1/ANS complex"

Summary

The phenomenon of structure modulation is relatively well understood in small-molecule crystallography, but its occurrence in macromolecular protein crystals has been surprising. Physical manifestations of this phenomenon include the observation of additional reflections between major Bragg peaks in diffractograms. As a result of modulation, the translational symmetry of the crystal is disrupted in three-dimensional space, and the periodicity of the structure is restored only in higher dimensions. This requires specialized methods of multidimensional analysis to correctly index diffraction images and describe the structure. Modulation of the structure can be caused both by periodic changes in the positions of atoms at locations determined by the spatial symmetry of the cell, and by periodic changes in the occupancy of a given crystallographic position.

Existing methods of solving and refining structures routinely used in protein crystallography are inadequate for comprehensive analysis of modulated structures. Restriction to description in three-dimensional space means adopting the measurability of modulation, where translational order is restored after a certain total number of elementary cells. Then the structure has to be analyzed in an enlarged supercell, which leads to a significant increase in the number of parameters for complex protein structures and makes it possible to obtain only an approximate model.

The lack of appropriate tools for the analysis of macromolecular modulated structures leads to serious problems in the proper indexing and processing of diffraction data, and then constructing a complete model of the structure with satisfactory divergence indices. So far, it has been possible to carry out a complete structural analysis only for a few modulated protein crystals. These include complexes of the Hyp-1 protein from St. John's wort (*Hypericum perforatum*) with the fluorescent ligand ANS (8-anilinonaphthalene-1-sulfonate). Depending on crystallization conditions, the Hyp-1/ANS protein complexes can form crystals with seven- (7Hyp/ANS) or nine-fold structure modulation (9Hyp/ANS) along the c direction of the *C*2 space group.

As part of my dissertation, I compared two modulated crystal structures of the Hyp-1/ANS protein complex. The first one, designated 7Hyp/ANS, was obtained, solved and described by Dr. Joanna Sliwiak and Prof. Mariusz Jaskólski's group in 2015 and had seven-fold modulation along the **c**-axis and contained 28 independent protein molecules in an extended supercell. The presence of modulation and tNCS (Translational Non-Crystallographic Symmetry) elements was coupled with tetartohedral twinning, which was an additional difficulty when solving and refining the structure. In the end, the structure was solved and refined in the supercell approach, and then the results were deposited in the PDB (Protein Data Bank) database with the code 4N3E. Shortly thereafter, during the change of crystallization and co-crystallization conditions in the presence of the plant hormone melatonin, another modulated structure of the Hyp-1/ANS complex was obtained. The discussed model of the 9Hyp/ANS complex with nine-fold modulation along the **c**-axis finally consisted of 36 Hyp-1 protein molecules arranged according to a motif containing 4 molecules (2 dimers) of Hyp-1 repeated 9 times along **c** (4 × 9 = 36). The structure of 9Hyp/ANS was solved by Dr. Joanna Sliwiak analogous to the previous case of 7Hyp/ANS, and then refined by me in supercell terms.

The dissertation presented here consists of three main parts and tasks related to the analysis of the crystal structure of the Hyp-1 protein. The first part involves the refinement of the crystal structure model of 9Hyp/ANS in supercell terms using conventional software and the related analysis of the structural elements of the crystal, especially issues related to the packing of atoms within the

elementary cell, perturbations of the periodicity of the structure, comparison with the structure of 7Hyp/ANS and ligand distribution.

In the second part, I made a detailed analysis of the diffraction data of both modulated Hyp-1/ANS crystals, refinement and methodical description and refinement of the 7Hyp/ANS structure in multidimensional space. In the process of refinement of the 7Hyp/ANS structure, proprietary specialized software developed in the Matlab environment was used. Dropping the simplified assumption of modulation dimensionality and introducing additional corrections to account for the disorder in the structure made it possible to obtain new models and improve their divergence rates. The developed package was then extended with further modules that allowed visualization of the data and introduction of corrections related to thermal vibrations of the crystal lattice (phonons).

The third part of the work concerns the application of molecular dynamics methods to the study of the thermal motions and conformations of the side chains of the Hyp-1 protein based on a purely physical approach, in isolation from the constraints imposed during the refinement of the crystal structure. The information obtained made it possible to determine the energetically preferred conformations within the protein and their comparison with the variants selected during the crystal structure refinement process.