Critical Assessment of Side Chain Prediction (CASCP): an in-house evaluation on single-point mutants of lysozyme.

ID - 102

Marabotti Anna[1], Facchiano Angelo[1]

[1]Laboratory of Bioinformatics, Institute of Food Science, CNR, Avellino

Motivation

The simulation of protein structure by computational methods has been exceptionally improved during the last 15 years, as proved by the results from CASP (Critical Assessment of techniques for protein Structure Prediction) competition. However, the prediction of the correct conformation of side chains is still a major problem.

Several programs are available and are widely used to place the side chains on the backbone of a protein structure and to optimize their conformations; however, serious questions can be made on the correctness of their predictions. The scope of the present work is to assess the ability of three widely used and freely available programs for side chain conformation prediction (NCN, SCAP and SCWRL) to simulate the introduction of a single point mutation in a structure and the effects of this mutation on the structure itself.

Methods

We chose as a benchmark the phage T4 lysozyme, for which hundreds of single point mutant structures are present in the PDB database. Starting from the structures of the wild type and of the pseudo-wild-type lysozyme (in which two cysteine residues are mutated to an alanine and a threonine residue), we introduced different single point mutations (for pseudo-wild-type, in addition to the two mutant residues already present), and we optimized the side chain conformations of the mutants with each of the three programs. The resulting structures were compared with the corresponding structures available in PDB using the traditional parameters adopted for this kind of evaluation, i.e. by analyzing the overall and average RMSD, with or without the inclusion of Cbeta atom in calculations, and the differences in chi1 and chi1+2 angles on the side chains.

Results

The results for the predictions on this benchmark appear to be worse than those calculated by the Authors on their benchmarks and published in the papers related to each method. The program NCN appears to be the most accurate for this benchmark both in terms of angle accuracy and RMSD, but in general it predicts no more than 70% of chi1 and chi1+2 dihedrals correctly (the prediction was considered correct for a dihedral angle when the deviation between the predicted and the experimental value was less than 20). The overall RMSD is 1.90 ï¿½when Cbeta atom is not included, 1.68 ï¿½when it is included in calculations. The programs SCAP and SCWRL perform generally worse than NCN both in terms of dihedral angle predictions and RMSD. SCWRL performs slightly better than SCAP in terms of dihedral angle predictions, the opposite is true for RMSD.

We also decomposed our results in terms of side chain exposure to solvent, polarity, size of mutations and influence on the nearest residues to the mutations. We found that generally side chains conformations are better predicted in core residues (with less than 10% solvent accessible area of the side-chain) than in the exposed ones, and the accuracy of prediction is similar for residues near the mutation or far from the site of mutation. The accuracy of prediction is higher when the size of mutant side chain is conserved, and a large to small mutation is better simulated than the opposite. Instead, the polarity of mutation affects the prediction of side chain conformation to a lesser extent.

This study confirms the need for better predictions of side chains conformations; however it also confirms that programs which are parameterized by taking into account not only a large rotamer library but also potential energy functions perform better, although they need a higher amount of computational time.

Email: anna.marabotti@isa.cnr.it