Small angle X-ray scattering data assisting protein structure prediction with coarse-grained UNRES simulations.

Cezary Czaplewski[1], Agnieszka S. Karczyńska[1], Magdalena A. Mozolewska[1,2], Paweł Krupa[1,3], Artur Giełdon[1], Adam Liwo[1]

[1]Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, Gdańsk 80-308, Poland
[2]Institute of Computer Science, Polish Academy of Sciences, Jana Kazimierza 5, Warsawa 01-248, Poland
[3]Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, Warsawa 02-668, Poland

A new approach to small angle X-ray scattering (SAXS) assisted protein-structure prediction is presented, which is based on running multiplexed replica exchange molecular dynamics simulations with the coarse-grained UNRES force field. Distance distribution derived from SAXS data is included as a maximum-likelihood penalty term with the coarse-grained UNRES energy function. To improve the quality of the resulting models and to speed up the search, we also use restraints derived from knowledge-based models of simulated proteins.

The coarse-grained UNRES model developed in our laboratory is a highly reduced model of polypeptide chains, with two interaction sites per residue: united side chains and united peptide groups. The UNRES force field is physics-based and originates from the potential of mean force of polypeptide chains in water, in which the degrees of freedom not present in the model have been averaged out. The UNRES force field is based on Kubo's cluster-cumulant expansion and includes the multibody terms responsible for the formation of regular secondary structures. The complete pseudo-energy function used in simulations consists of the UNRES energy function, the template-restraint penalty term, and the SAXS-restraint term. We use the distance distribution derived from the SAXS data to restraint the simulations instead of the scattering intensity usually employed in SAXS penalty terms. The distance distribution is more closely related to a conformation than its Fourier transform, the scattering intensity, and, therefore, seems to be more sensitive to conformation changes. Calculating distance distribution we approximated the contribution from a given pair of residues as a Gaussian distribution with the standard deviation calculated from the Stokes radii of these residues. The Gaussian approximation to the distance distribution due to a pair of coarse-grained sites partially accounts for the solvation-shell water.