Abstract: MATH/CHEM/COMP 2008, Dubrovnik, June 16-21, 2008

THE PROTEIN SECONDARY STRUCTURE FLEXIBILITY

Krešimir ŠIKIĆ[1], Branko JEREN[1], Sanja TOMIĆ[2]

[1]*Faculty of Electrical Engineering and Computing, Unska 3, 10000 Zagreb, Croatia*
[2]*Institute Ruđer Bošković, Bijenička c. 54, 10000 Zagreb, Croatia*

kresimir.sikic@fer.hr
branko.jeren@fer.hr
sanja.tomic@irb.hr

Modeling protein flexibility is a long standing challenge in computational biology with a special impact to protein docking. Relating problems are protein structures alignment and identification of flexible and rigid protein regions, as well as a general definition of a region degree of flexibility. Determination of protein flexibility from a single protein conformation, so called *apriori* approach is still computationally demanding and quite unreliable. As a part of our research on finding an accurate procedure for predicting flexible and rigid protein domains and in order to learn as much as possible about protein flexibility we developed a fast and simple *aposteriori* method (method based on comparison of different conformations of the same protein) for characterization of protein regions regarding their flexibility. We applied it to the test set consisting of 220 proteins in which we identified about 6500 on the secondary structure based regions and classified them regarding their flexibility. We investigated possible correlations between the protein flexibility and different 3D structural parameters and found that a region flexibility correlate with population of certain amino acid residues, the protein solvent accessible surface area and B-factors. The set of protein regions characterized by flexibility and the results obtained should be valuable in testing the new *apriori* procedures, but also the procedure itself can be used in discrimination the regions in proteins by flexibility.

Keywords – protein flexibility, proteins alignment, Kabsch algorithm, sliding windows algorithm, Needleman-Wunsch algorithm, Accessible Surface Area, amino acid residue frequency.