Phosphate-binding sites identification in unbound protein structures

Parca L, Gherardini PF, Helmer-Citterich M, Ausiello G

Motivation

Most important cellular processes, reactions and interactions involve molecules containing at least one phosphate group. Different bioinformatic metodologies exists to predict metal, carbohydrate or nucleotidebinding sites in protein structures but none is available so far for the identification of phosphate-binding sites. Here we present a new method to search in protein structures the binding sites of phosphate groups both in ion form and as a part of different ligands like nucleotides. Functional annotation procedures, docking methods and drug design methodologies can take advantage of the prediction made by this method.

Methods

The method uses the Query3D (Ausiello, 2005) local structural comparison program to search the query protein structure for the presence of a number of structural motifs identified for their ability to bind the phosphate chemical group (Ausiello, 2009). The predicted phosphate groups are then added to the protein structure and evaluated using information from the solvent excluded surface and cavities of the protein.

Results

The method has been trained on a dataset composed of 39 high quality protein-nucleotide complexes (Zhao, 2001). It correctly identified all but one binding sites with an average of 8.8 false positive predictions per structure, considering a prediction to be correct whenever a phosphate is placed within a 5.0 Å distance from its true position. The method has then been tested on an independent set of 54 proteins crystallized in both apo and holo form (Dessailly, 2008) and demonstrated to be only slightly affected by the conformation change induced by the bound ligands. We obtained at least one correct prediction in the 67% of the holo structures and in 63% of the apo structures with a PPV of respectively 16% and 14%.

Contact e-mail

luca.parca@uniroma2.it

Centre of Molecular Bioinformatics, Department of Biology, University of Tor Vergata, Roma

Supplementary information

References

[1] Ausiello, G., Via, A., and Helmer-Citterich, M. (2005). Query3d: a new method for high-throughput analysis of functional residues in protein structures. BMC Bioinformatics 6, Suppl 4:S5. [2] Ausiello, G., Gherardini, P.F., Gatti, E., Incani, O., and Helmer-Citterich, M. (2009). Structural motifs recurring in different folds recognize the same ligand fragments. BMC Bioinformatics 10:182. [3] Zhao, S., Morris, G.M., Olson, A.J., and Goodsell, D.S. (2001). Recognition templates for predicting adenylate binding sites in proteins. J. Mol. Biol. 314, 1245-1255. [4] Dessailly, B.H., Lensink, M.F., Orengo, C.A., and Wodak, S.J (2008). LigASite—a database of biologically relevant binding sites in proteins with known apo-structures. Nucleic Acids Research 36, 667-673