BIOINFORMATICS APPLICATIONS NOTE Vol. 29 no. 6 2013, pages 805–806
doi:10.1093/bioinformatics/btt032

Structural bioinformatics

Advance Access publication February 3, 2013

BION web server: predicting non-specifically bound surface ions

Marharyta Petukh[1,†], Taylor Kimmet[1,2,†] and Emil Alexov[1,*]

[1]Computational Biophysics and Bioinformatics, Department of Physics and Astronomy, Clemson University, Clemson, SC 29634, USA and [2]Department of Computer Sciences, Clemson University, Clemson, SC 29634, USA

Associate Editor: Anna Tramontano

ABSTRACT

Motivation: Ions are essential component of the cell and frequently are found bound to various macromolecules, in particular to proteins. A binding of an ion to a protein greatly affects protein's biophysical characteristics and needs to be taken into account in any modeling approach. However, ion's bounded positions cannot be easily revealed experimentally, especially if they are loosely bound to macromolecular surface.

Results: Here, we report a web server, the BION web server, which addresses the demand for tools of predicting surface bound ions, for which specific interactions are not crucial; thus, they are difficult to predict. The BION is easy to use web server that requires only coordinate file to be inputted, and the user is provided with various, but easy to navigate, options. The coordinate file with predicted bound ions is displayed on the output and is available for download.

Availability: http://compbio.clemson.edu/bion_server/

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Contact: ealexov@clemson.edu

Received on October 19, 2012; revised on January 14, 2013; accepted on January 17, 2013

1 INTRODUCTION

Ions are important component for all living cells, and they are involved in different physiological processes, such as maintaining the osmotic balance (Williams, 1970) and the general ionic environment (Bhattacharya, 2005) of the cells, participating in signal transduction (McAinsh *et al.*, 1997) and many others. Because of that predicting ion's positions is important for understanding various biological reactions. The need for predictions is dictated by the difficulties different experimental techniques experience in determining surface bound ions, as the residential time and the occupancy may not be large; thus, the signal could be weak.

Although there are many approaches for predicting specifically bound ions as methods that are based on coordination numbers of ions (Katz *et al.*, 1996), geometries (Babor *et al.*, 2008; Levy *et al.*, 2009), binding site preferences (Dudev *et al.*, 2003), typical valence of metal ions (Muller *et al.*, 2003) and many others, to the best of our knowledge, there is no web server for non-specifically bound surface ions predictions. Here, we report such a web server, which is based on our new methodology using Delphi (Li *et al.*, 2012) and calculates the electrostatic potential map in conjunction with an in-house clustering algorithm to predict non-specific ion binding sites (Petukh *et al.*, 2012). This method is implemented into a web server, the BION web server, which allows predicting non-specific bound ions on the surface of proteins. It is easy to use and intuitive for broad range of users. As result of calculations the user receives a file in Protein Data Base (PDB) format of protein with predicted ions position ranked according to the confidence of the predictions.

2 METHODS

2.1 BION web server algorithm

To successfully run the algorithm, the user should (i) upload protein structure file in PDB format (or the ID number of corresponding protein in PDB), (ii) choose the type of ion for which the position will be predicted and (c) select the number of ion positions to be predicted. The next step involves in-house algorithm, which fixes structural defects in the protein with *profix* software from *Jackal* package developed in Honig's laboratory (http://wiki.c2b2.columbia.edu/honiglab_public/index.php/Software website: JACKAL) (Xiang, 2002) and protonates the proteins, i.e. generates missing hydrogen atoms, with MCCE software (Georgescu *et al.*, 2002) [AMBER force field parameters (Wang *et al.*, 2004) are used]. The electrostatic potential distribution is calculated with the DelPhi software (Li *et al.*, 2012) with the following parameters: scale = 1 $Å^{-1}$, perfil = 70%, internal dielectric constant of the protein 4, external dielectric constant of the water phase 80 and ionic strength 0.5 mol/l. The details of the applied methodology are described in the Supplementary Material and are similar, but not identical, to those described in our original work (Petukh *et al.*, 2012).

2.2 Benchmarking parameters

Rank: The predictions are based on the magnitude of the electrostatic potential at selected surface-bound grid points. The representative grid points are sorted in descending order (by absolute value) of the potential (positive for negatively charged ions and negative for positive one), and the position of a given point within this list is termed *Rank*. Thus, a representative grid point being on third position within the ordered list of N representative points is considered to be *Rank* = 3.

Receiver operating characteristic (*ROC*) *curves*: The ability to predict the experimental ion's position and accuracy of the BION algorithm was analyzed by generating *ROC curves*. The *x*-axis represents the *Rank* of the closest representative grid point to the experimentally determined ion position; whereas the *y*-axis represents the number of successful predictions (true predictions) in percentage of all predictions. A prediction is considered to be true if the distance between the predicted representative grid point and the actual experimental ion position is <10 Å. This criterion is chosen because of the parameters of the clustering algorithm, resulting in that the shortest distance between representative grid points is ≥10 Å.

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author 2013. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com

Downloaded from http://bioinformatics.oxfordjournals.org/ at :: on August 30, 2016

Time of calculations: Total time that is required to perform all calculations to predict the ion positions.

2.3 Web server architecture

The design of the BION server can be described by two basic parts, the first of which is the client or user interface, and the second of which is the process that occurs on the server or back end. The client interface is implemented using HTML (http://www.w3.org/TR/html5/) and JavaScript (http://www.ecma-international.org/publications/standards/Ecma-262.htm). It is here that a user will fill out various parameters for a job to be submitted to the server, as well as receive the results of the specified job. The back end request handling is implemented using Python (http://python.org/). Later in the text we describe each part of the processes in detail.

Client interface. This is where the user specifies all necessary information to complete a server job. The most important piece of information the user must supply is a PDB file. A user may choose either to upload a PDB file from their local file system, or they may enter the name of a PDB file located in the RCSB PDB, which will then be downloaded by the server. The next job parameter is the type of ion that will be predicted, and it is selected using a simple drop down menu. The final parameter, which is not necessary for the user to change, is the number of binding sites that will be predicted and will be appended to the PDB file.

Back end. The server end or back end is where the information specified by the user is processed. Both the request handling method and the calculation method are implemented in Python.

Results. If all user input was valid, then the results of their specified job are returned to them. The basic output from the server includes a summary of the job and a fixed version of the user's submitted PDB file with calculated ion binding sites appended. If the user requested the results to be e-mailed, then this information will be sent to them in an e-mail to which the fixed PDB file is attached. From the actual client interface, the user will have access to all of this returned information, as well as a Jmol (http://www.jmol.org/) visualization of the protein, which also depicts potential binding sites for the ion (Supplementary Fig. S1).

3 RESULTS AND DISCUSSION

The BION web server is based on the BION program that allows predicting the position of non-specific bound ions on the surface of protein. The key properties of the server, the accuracy and the time, are analyzed. We benchmarked the results against the database of proteins with experimentally determined surface ions (for detailed description of dataset pruning for each type of ions see Petukh, *et al*., 2012). The following types of ions were tested: CA, CL, FE, K, NA, MG, MN and ZN (the list of PDB files is available at the front page of the server).

3.1 Test of BION accuracy

For each type of ion, we constructed *ROC curve* by varying the *Rank*, which essentially means by varying the number of predictions (Supplementary Fig. S2). In ~50% of the cases for all tested types of ions, there is a prediction within 10 Å away from the experimental ion's position, which is found within the first 10 ranked points. In addition, enrichment curves are shown in Supplementary Figure S3.

3.2 Total time of calculations

One of the most important parameter of any server performance is how quick the user receives the results. According to the Supplementary Figure S4, the average time is ~10 min.

3.3 Conclusions

The BION web server is a simple to use tool that uses new algorithm for prediction of non-specific bound surface ions based on electrostatic properties of the protein.

ACKNOWLEDGEMENT

The authors thank Chuan Li and Lin Li for the help in the development.

Funding: National Institutes of Health (R01 GM093937).

Conflict of Interest: none declared.

REFERENCES

Babor,M. *et al*. (2008) Prediction of transition metal-binding sites from apo protein structures. *Proteins*, **70**, 208–217.

Bhattacharya,P.K. (2005) *Metal Ions in Biochemistry*. Alpha Science International, Ltd, Oxford, UK.

Dudev,T. *et al*. (2003) First-second shell interactions in metal binding sites in proteins: a PDB survey and DFT/CDM calculations. *J. Am. Chem. Soc*., **125**, 3168–3180.

Georgescu,R.E. *et al*. (2002) Combining conformational flexibility and continuum electrostatics for calculating pK(a)s in proteins. *Biophys. J*., **83**, 1731–1748.

Katz,A.K. *et al*. (1996) Calcium ion coordination: a comparison with that of beryllium, magnesium, and zinc. *J. Am. Chem. Soc*., **118**, 5752–5763.

Levy,R. *et al*. (2009) Prediction of 3D metal binding sites from translated gene sequences based on remote-homology templates. *Proteins*, **76**, 365–374.

Li,L. *et al*. (2012) DelPhi: a comprehensive suite for DelPhi software and associated resources. *BMC Biophys*., **5**, 9.

McAinsh,M.R. *et al*. (1997) Calcium ions as second messengers in guard cell signal transduction. *Physiol. Plant*., **100**, 14.

Muller,P. *et al*. (2003) Is the bond-valence method able to identify metal atoms in protein structures? *Acta Crystallogr. D*, **59**, 32–37.

Petukh,M. *et al*. (2012) Predicting nonspecific ion binding using DelPhi. *Biophys. J*., **102**, 2885–2893.

Wang,J. *et al*. (2004) Development and testing of a general amber force field. *J. Comput. Chem*., **25**, 1157–1174.

Williams,R.J.P. (1970) Tilden lecture. The biochemistry of sodium, potassium, magnesium, and calcium. *Q. Rev. Chem. Soc*., **24**, 331–365.

Xiang,J.Z. and Honig,B. (2002) *JACKAL: A Protein Structure Modeling Package*. Columbia University and Howard Hughes Medical Institute, New York.

Downloaded from http://bioinformatics.oxfordjournals.org/ at :: on August 30, 2016