JOURNAL OF CHEMICAL INFORMATION AND MODELING

Application Note

pubs.acs.org/jcim

ProBiS-CHARMMing: Web Interface for Prediction and Optimization of Ligands in Protein Binding Sites

Janez Konc,[†,‡] Benjamin T. Miller,[§] Tanja Štular,[‡] Samo Lešnik,[†] H. Lee Woodcock,[▽] Bernard R. Brooks,[§] and Dušanka Janežič*[,‡]

[†]Laboratory for Molecular Modeling, National Institute of Chemistry, Hajdrihova 19, SI-1000, Ljubljana, Slovenia

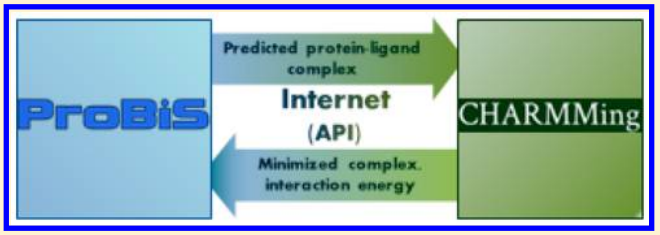
[‡]Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Glagoljaška 8, SI-6000, Koper, Slovenia

[§]Laboratory of Computational Biology, Biochemistry and Biophysics Center, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, United States

[▽]Department of Chemistry, University of South Florida, 4202 East Fowler Ave., Tampa, Florida 33620, United States

Supporting Information

ABSTRACT: Proteins often exist only as apo structures (unligated) in the Protein Data Bank, with their corresponding holo structures (with ligands) unavailable. However, apoproteins may not represent the amino-acid residue arrangement upon ligand binding well, which is especially problematic for molecular docking. We developed the ProBiS-CHARMMing web interface by connecting the ProBiS (http://probis.cmm.ki.si) and CHARMMing (http://www.charmming.org) web servers into one functional unit that enables prediction of protein−ligand complexes and allows for their geometry optimization and interaction energy calculation. The ProBiS web server predicts ligands (small compounds, proteins, nucleic acids, and single-atom ligands) that may bind to a query protein. This is achieved by comparing its surface structure against a nonredundant database of protein structures and finding those that have binding sites similar to that of the query protein. Existing ligands found in the similar binding sites are then transposed to the query according to predictions from ProBiS. The CHARMMing web server enables, among other things, minimization and potential energy calculation for a wide variety of biomolecular systems, and it is used here to optimize the geometry of the predicted protein−ligand complex structures using the CHARMM force field and to calculate their interaction energies with the corresponding query proteins. We show how ProBiS-CHARMMing can be used to predict ligands and their poses for a particular binding site, and minimize the predicted protein−ligand complexes to obtain representations of holoproteins. The ProBiS-CHARMMing web interface is freely available for academic users at http://probis.nih.gov.



INTRODUCTION

Predicting protein−ligand binding is one of the most challenging problems in biochemistry, with profound implications for pharmaceutical chemistry and the discovery of protein function. Many approaches have been developed for protein−ligand binding prediction, the most prominent being molecular docking.[1] In template-free docking, however, every new molecule must be docked *ab initio*, and information from existing similar protein−ligand complexes is not considered. Recently, we proposed a different approach that, contrary to docking, predicts ligands for a protein by their transposition between similar binding sites.[2] In this approach, available at the ProBiS-ligands web server,[3] we assume that similar binding sites are likely to bind similar ligands. An existing ligand bound to a protein A can be predicted to bind to protein B if the binding sites in protein A and protein B are similar. The confidence level of this prediction is assessed by the binding sites' similarity score (Z-score) calculated from RMSDs and Blosum62 matrices[4] between superimposed residues. Transposition of ligands, especially between nonhomologous proteins, is dependent on accurate local three-dimensional superimpositions of binding sites' amino acid residues, usually not detectable by standard sequence or structure alignment approaches.

There are many approaches to protein similarity searching, ranging from sequence[5]- or structure-homology-based approaches[6,7] that compare global properties of proteins to substructure or fingerprint searches[8−11] that focus on smaller protein regions (for example, binding sites). Since binding sites are evolutionarily more conserved than complete proteins, comparison of protein binding sites has precedence over sequence or structural homology-based methods when analyzing protein−ligand interactions. The ProBiS algorithm[12] compares protein binding sites represented as protein graphs, irrespective of sequence or global fold similarity of proteins, in a

Received: August 24, 2015
Published: October 28, 2015

DOI: 10.1021/acs.jcim.5b00534
J. Chem. Inf. Model. 2015, 55, 2308−2314

pairwise fashion using a fast maximum clique algorithm.[13,14] It finds sets of residues that are physicochemically and geometrically related. Querying a binding site or protein structure against a database of template protein structures, ProBiS retrieves proteins with similar binding sites. From the resulting alignments, it calculates degrees of structural evolutionary conservation (shown as different colors) for all surface amino acid residues of the query protein.

The approach for ligand prediction, which was implemented in the ProBiS-ligands web server,[3] relies only on the geometric and evolutionary criteria between proteins for ranking the predicted ligands; i.e., no energy minimization is performed in the transposition process. Consequently, predicted ligands can have steric clashes with the atoms of the query protein, and it is possible for electrostatically unfavorable conformations to be found. It would thus be beneficial to be able to automatically refine the predicted poses in the protein−ligand complexes and calculate interaction energies.

The CHARMMing (CHARMM interface and graphics)[15] web interface facilitates access to the powerful and widely used molecular software package, CHARMM.[16] CHARMMing allows for the execution of tasks such as energy calculation, minimization, dynamics, multiscale modeling, structure mutation, advanced visualization using WebGL, and other techniques commonly used by computational life scientists. Of these, support for both additive and subtractive multiscale techniques is especially strong due to CHARMM's support for these methods (see below). Furthermore, model building for the structure−activity relationship and the quantitative structure−activity relationship has recently been added.[17] Also, a fully featured module that implements an automated procedure for fragment-based docking has been integrated, which is a highly relevant technique in computational drug discovery.[18] All inputs and outputs from CHARMMing are downloadable by the end-user, allowing them to modify the produced CHARMM scripts and run them locally. The Python source code for CHARMMing itself is available in the public domain and is under active development. In addition, CHARMMing has a strong educational component, comprising of several lessons that can be used to learn basic techniques and procedures commonly used in molecular simulations.[19−21] Several important pieces of CHARMMing's functionality such as structure building and energy minimization are exposed via a Representational State Transfer (RESTful) Web-services Application Programming Interface (API) that has recently been implemented and which allows the ProBiS-ligands server to make calls on functionality provided by CHARMMing.

The CHARMMing application itself implements only simple analysis methods on static structures. Any simulation method that requires the calculation of forces or potential energies is passed to a separate simulation package. As the name implies, CHARMMing currently uses CHARMM[16] as its primary simulation package. CHARMM was developed by a network of collaborators who worked in conjunction with Professor Martin Karplus's research group at Harvard University. It supports most commonly used simulation methods including energy minimization, molecular dynamics, Monte Carlo simulation and replica exchange. Recent developments include enhanced sampling methods such as self-guided Langevin Dynamics[22] and new free-energy methods such as the combination[23,24] of Enveloping Distribution Sampling[25] with Hamiltonian replica exchange.[26] Substantial work has also been done to improve CHARMM's performance, which is on par with other major biophysical simulation packages.[27] In addition, CHARMM supports diverse multiscale techniques[28,29] that are also made available through CHARMMing.

In this work, we develop a new ProBiS-CHARMMing web interface that allows the prediction of ligands, as well as their energy minimizations and interaction energy calculations (Figure 1). This web interface addresses the main existing shortcoming of ProBiS-ligands, which is the inability to remove steric clashes between predicted ligands and protein, which can be the cause of unrealistic models, and the lack of energy-based scores to assess the strength of ligand binding. The web interface also facilitates the creation of CHARMM friendly protein−ligand systems including CHARMM input scripts for further modeling studies. We present the usage of the newly developed interface by predicting energy minimized protein−peptide, protein−small molecule, and protein−ion complexes for unligated (apo) query protein structures. The ProBiS-CHARMMing web interface provides an interactive environment in which users can explore the predicted protein−ligand complexes and calculate and compare their energetic properties.

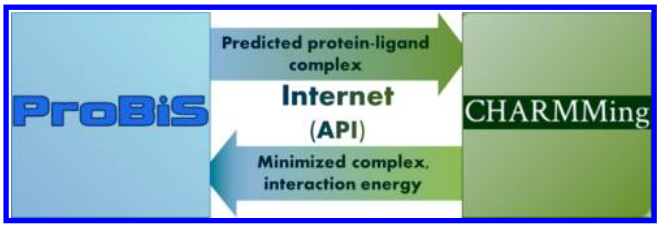


Figure 1. ProBiS-CHARMMing web interface organization, accessible at http://probis.nih.gov.

DESIGN AND IMPLEMENTATION

Ligands Prediction. The input to the ProBiS-CHARMMing web interface is a query protein structure, provided with a PDB/chain identifier or with an uploaded PDB model. Optionally, a particular binding site may be selected on the query protein. Using this query protein structure (or query binding site), the ProBiS algorithm identifies template proteins with similar three-dimensional surface amino acid patterns by comparing the query protein with every nonredundant template (to >95% sequence identity) protein chain in the PDB. The ProBiS-ligands algorithm then transposes the existing ligands bound to protein matches that were found during the aforementioned search. This is done by rotating and translating ligand coordinates according to the superimposition matrix obtained from aligning similar query and template protein binding sites. In contrast to global protein alignment approaches, ProBiS performs local comparisons of the query protein structure or query protein binding site with the template structures in the nonredundant PDB database; consequently, it may discover proteins with similar binding sites but different global folds to that of the query protein. Protein surface amino acids are treated as protein graphs, and a fast maximum clique algorithm[13,14] is used to perform the alignment of the compared protein graphs. Each maximum clique represents a rigid body alignment of two protein structures or binding sites, and the vertex-to-vertex alignment that is obtained from the maximum clique is used to superimpose the matched parts of the two compared protein structures.

DOI: 10.1021/acs.jcim.5b00534
J. Chem. Inf. Model. 2015, 55, 2308−2314

Figure 2. Minimization setup in the ProBiS-CHARMMing web interface. On the right side of the screen is the list of predicted ligands with the selected ligand's row highlighted in orange. On the left side is the JSmol viewer that contains the three-dimensional pose of the predicted ligand before (sticks, CPK colored) and after (balls and sticks, CPK colors) the minimization in the query protein (cartoon model, structural conservation colored from white [unconserved] to green [conserved]). The Menu window at the top-center contains different drop-down menus organized with letters A, S, H, L, C, and M and another window (summoned upon clicking M and then Minimize) where the user can set the minimization parameters by filling the form. On the far top-left, the Minimization window containing the interaction energy of the protein−ligand complex after minimization is displayed.

In addition to the live job calculation described above, which takes time and uses computer resources, the search bar at the top of the web interface also allows users to access precalculated ligand predictions through its connection to the ProBiS-Database,[30] using a PDB ID as the query. This database currently consists of precalculated structural similarity profiles for ∼37 000 protein chains in the nonredundant (>95% sequence identity) PDB.

Minimization and Interaction Energy Calculation of Predicted Ligands. The new ProBiS-CHARMMing user interface allows for minimization and energy calculation of the predicted ligands (see Figure 2). The interface is organized as follows: on the left side of the screen is a Jsmol[31] molecular viewer, on the right side there is a list of predicted ligands, and in the center, there is a Menu window (that can be clicked on or relocated) that lists the currently viewed molecules in the Jsmol viewer and allows their manipulation, including energy minimization. When the user selects a ligand by clicking the "View 3D" button in the "Ligand" column, the selected ligand's predicted three-dimensional (3D) pose is loaded into the Jsmol viewer. At the same time, the ligand's name appears in the Menu window, where the user can select the menu option corresponding to the selected ligand: "A" (Action), "S" (Show), "H" (Hide), "L" (Label), or "M" (Modeling).

Clicking on "M" in the same row as the ligand's name and then selecting the "Minimize" option opens a new window, where it is possible to set up the minimization parameters, such as the gradient tolerance, the number of steps for the Steepest Descent (SD) minimization, and the number of steps for the adopted basis Newton−Raphson method (ABNR) minimization, as well as request the minimization with or without implicit solvent. Upon clicking the "Start Minimization" button, the request is sent to the CHARMMing web server through the RESTful interface, where minimization of the coordinates of the ligand bound to the query protein, and subsequent calculation of their interaction energy is performed. After CHARMMing completes the job, it returns the minimized coordinates of the ligand and the protein, together with their interaction energy to the ProBiS-CHARMMing web interface, where the coordinates are loaded into the Jsmol molecular viewer, and the Mini1 and Mini2 entries are added to the Menu window corresponding to the minimized ligand and protein, respectively. In addition, another Minimization window is created, which displays the interaction energy between the ligand and the protein, as well as containing a link that allows users to download the CHARMM scripts, parameter files, and all associated files to recreate the minimization, or run longer simulations on a local computer.

ProBiS-CHARMMing Integration through RESTful API. The ProBiS web application communicates with CHARMMing via a recently implemented RESTful API. REST is an acronym for REpresentational State Transfer and is a popular method for building expandable web services.[32] Clients make requests to the server via standard HTTP POST or GET requests (the same functionality is exposed via both methods). The desired functionality is specified by the URL to which the request is sent, as well as the parameters of the call. The parameters for a request to a particular URL are given as the string representation of a JavaScript Object Notation (JSON) data structure, which is interpreted by the server. Three major pieces

DOI: 10.1021/acs.jcim.5b00534
J. Chem. Inf. Model. 2015, 55, 2308−2314

of functionality are exposed by API calls: energy calculation, energy minimization, and interaction energy calculation. Input parameters include the starting PDB file for the calculation, whether implicit solvent is to be used, and (for optimization) the number of minimization steps and the tolerance. After the requested calculation has been completed, the server returns another JSON object to the caller that encapsulates a status (success or an error code with a human-readable error string); if the calculation was successful, the returned JSON data structure also contains the results of the calculation. Example Python code demonstrating calls to CHARMMing's API is given in the Supporting Information. The ProBiS web server then makes use of the energy and minimization functions provided by the CHARMMing API, as described in the previous section.

In order to improve the response time of the ProBiS server, an option was added to the CHARMMing API to allow for asynchronous calls. The ProBiS application becomes unresponsive to user input while an API call is being made, so it is imperative for API calls to return as soon as possible; however, on the CHARMMing side, the energy and minimization calls can take some time to run the necessary calculations. When an asynchronous call is made, the server creates the CHARMM input files for the calculation, submits them to a batch queue system using CHARMMing's standard job submission mechanism as described in ref 15, and returns a reference to the job (called a handle) to the caller. Handles are designed so as not to be easily guessable by another API client, even if they are familiar with the algorithm by which the handles are generated. The client can send a new query to the server referencing that handle, and the server will return updated status information. If the job is complete, the server will return the results to the client.

CHARMMing Web Server Energy Calculation. As indicated above, the CHARMMing web server supports numerous different types of calculations, and three of these in particular—energy calculation, geometry optimization via energy minimization, and interaction energy calculation—are exposed via the RESTful API; the latter two calculation types are used by the ProBiS-CHARMMing web interface. All protein energy calculations are done using the CHARMM36 force field for proteins,[33] which can also calculate energies for nucleic acids, while the CHARMM General Force Field (CGenFF)[34] is used for small molecules. CHARMM version c37b2[16] is currently employed on the CHARMMing web server. For each type of calculation, obtaining the result is done similarly to the method described in the reference CHARMMing paper.[15] However, the process is also outlined briefly below.

Initially, the starting structure is parsed into its constituent segments by the CHARMMing parser, which is now available as part of a standalone project called pychm.[35,36] Ligand segments that contain residues not supported by the CHARMM36 force field (so-called bad HETATM residues) are dispatched to ParamChem[37−39] to generate topology and parameter files. For each segment, the CHARMM protein structure files (PSFs) are then built and combined into the final PSF for the entire structure; in addition, the coordinates are written out in the CHARMM card format. The potential energy of the structure is then calculated using the CHARMM ENERgy command. If geometry optimization via energy minimization is requested, a CHARMM input script is constructed that reads in the PSF and coordinate files with all necessary topology and parameter information, and then the CHARMM MINImization command is used. The caller may specify the number of steps for steepest descent (default 100, maximum 1000) and adopted basis Newton Rhapson (default 100, maximum 10 000) minimization that are to be used and the tolerance of the root-mean-squared gradient that is expected (default: 0.1 kcal/mol). If the interaction energy between two segments is requested, either with or without minimization, it is calculated via the CHARMM INTE command. Example CHARMM input scripts for each type of calculation are given in the Supporting Information.

USAGE EXAMPLES

The ProBiS-CHARMMing web interface provides an interactive environment with molecular modeling capabilities. Here, we present its use on a series of examples involving proteins, small molecules, and single-atom ionic ligands. The web interface helps answer the question how a ligand, for which the ligand−protein complex structure exists in the PDB, binds to another protein, for which the complex structure with this ligand is unavailable. The server facilitates creation of protein−ligand complex models, as well as estimation of their interaction energies.

Construction of Protein−Protein Complexes. We present an example of a protein−protein complex that exists in nature, but its complex structure, consisting of two binding partners, is not available in the PDB. It is formed by the N-terminal SH3 domain protein and a decapeptide derived from Son of sevenless (SOS).[40] However, the complex structure is not available for the particular N-terminal SH3 domain from *Drosophila melanogaster*, which exists in the PDB only in its apo form (PDB ID: 2AZS) and is used here as the query protein for prediction of its complex with the SOS peptide. Using the ProBiS part of the ProBiS-CHARMMing web interface, we predicted a possible binding partner (SOS peptide) of the query protein from another crystal structure (PDB ID: 1GBQ) that has <30% sequence identity to the query protein (Figure 3). Based on the superimposition of the binding sites (Z-score of 3.19, indicating a significant similarity), we transposed this SOS peptide to its binding position on the query protein structure. Then, to get an optimized model of the SOS peptide binding to the query protein, we minimized the energy of the predicted SOS peptide-N-terminal SH3 protein domain complex and calculated the interaction energy using the CHARMMing side of the web interface. We minimized the complex using 1000 SD and 10000 ABNR minimization steps without implicit water. The resulting energy of the minimized complex was −224 kcal/mol, while prior to minimization the calculated energy was on the order of $+10^6$ kcal/mol, because of steric hindrances. The minimization and energy calculation process required only a few mouse clicks from the user. After minimization, the interactions between the SOS peptide and the query N-terminal SH3 protein domain resembled the interactions observed in other crystal structures of the complex (e.g., PDB ID 1AZE, shown in Figure S1 in the Supporting Information). The predicted minimized protein−ligand model thus resembled the actual complex, and could be subjected to further molecular dynamics studies to get an even more precise picture of this protein−peptide complex.

Construction of Protein−Small Molecule Complexes. Similar to the previous example, the ProBiS-CHARMMing web interface can be used for the construction of protein−small molecule complexes. Apoprotein binding sites often differ from their holo counterparts, because of induced fit effects upon

DOI: 10.1021/acs.jcim.5b00534
J. Chem. Inf. Model. 2015, 55, 2308−2314

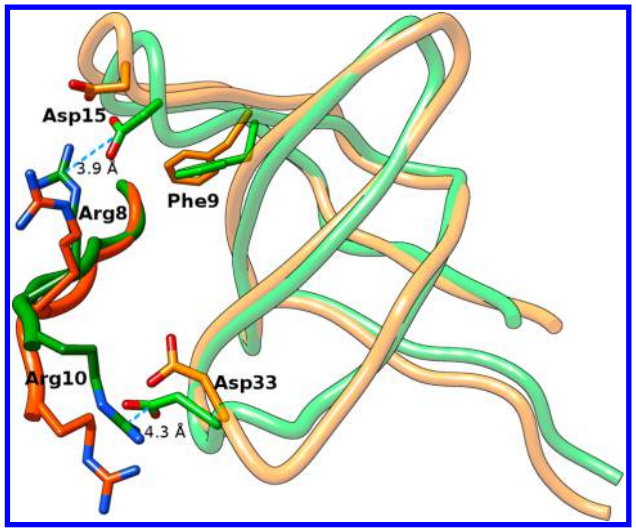


Figure 3. A protein−peptide complex predicted and simulated using ProBiS-CHARMMing. The N-terminal SH3 domain (PDB code: 2AZS) before (pale orange) and after minimization (pale green) is shown as ribbon models on the right side. The SOS peptide before (darker orange) and after minimization (darker green) with the interacting amino acid residues (shown as sticks) and salt bridge interactions (shown as blue dashed lines) is on the lefthand side.

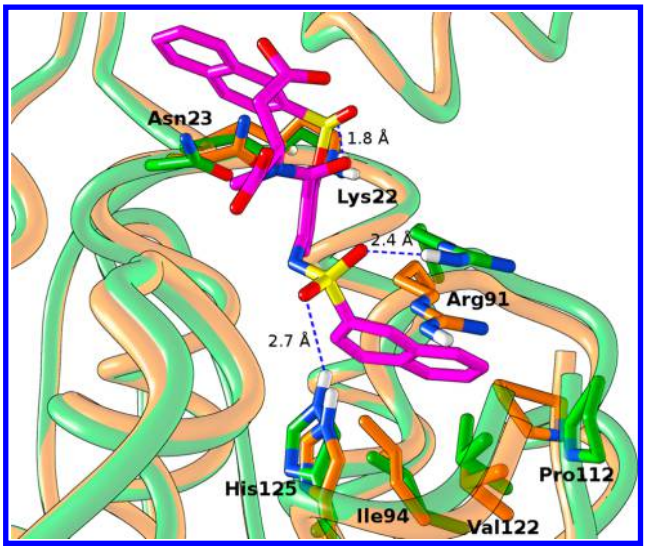


Figure 4. Protein−small molecule model obtained with ProBiS-CHARMMing. Important amino acid residues before minimization (orange) and after minimization (green) are shown as sticks. The TAV ligand is shown as purple sticks. Possible hydrogen bonds are shown as blue dashed lines.

ligand binding, which alters the spatial arrangement of the binding site's residues. It has been shown that molecular docking to an holoprotein's binding site results in higher enrichment, i.e., ability to discriminate between active and inactive compounds, compared to when docking is performed to apo structures.[41,42] The binding site cavity in an apoprotein can often be too small to accommodate drug-sized molecules, which prevents docking altogether.[43,44] When only the apoprotein is available in the PDB, it would thus be of great benefit to docking if we could make the binding site more similar to a holoprotein. One method to accomplish this is to predict the binding pose of a ligand from another similar crystal structure, dock it to the binding site of the apoprotein, and then minimize the geometries of both the ligand and the protein to make it resemble the interactions in the holo form.

The UDP-*N*-acetylglucosamine enolpyruvyl transferase enzyme (MurA) crystal structure from *Escherichia coli*, which is involved in bacterial cell wall synthesis, is an example of such a crystal structure, currently only available in the PDB in its apo form (PDB ID: 1UAE). As an attractive target for antimicrobial design, we used ProBiS-CHARMMing to simulate the conformational change of its binding site upon ligand binding. The ligand (PDB ligand code: TAV) from a similar protein (>90% sequence identity) structure (PDB ID: 1YBG) from the *Enerobacter cloacae* was transposed to the apo-MurA binding site, and, initially, this ligand's pose sterically clashed with the residues of the enzyme. To resolve the clashes, we minimized (1000 steepest descent, 10000 ABNR steps, without implicit water) the TAV ligand's predicted pose with the enzyme and observed the interactions forming between the ligand and the binding site amino acid residues (see Figure 4). The possible interactions in the minimized structure closely resembled those observed in a similar holoprotein (1YBG); the RMSD between the binding amino acid residues of the superimposed template with the minimized simulated structure lowered to 2.0 Å, compared to the RMSD of 2.3 Å between the template 1YBG and the original apo 1UAE structure. The two naphthalene rings of the TAV ligand exhibited possible stacking and hydrophobic interactions with the amino acid residues Arg91, Ile94, Pro112, Val122 and Asn23, which have also been recognized as important in ref 45. Hydrogen bonds formed between oxygen atoms of the ligand and the side chains of Lys22, Arg91, and His125, although, in our predicted complex, Arg91 hydrogen-bonded with the sulfonamidic oxygen of the ligand, while in the 1YBG structure the bond forms with the amidic oxygen. Using ProBiS-CHARMMing, we made the apoprotein binding site more closely resemble the holoprotein binding site, which is presumably more suitable for docking, and will ideally facilitate the discovery of novel MurA inhibitors.

Construction of Protein−Ion Complexes. ProBiS-CHARMMing is also able to simulate protein−ion complexes, and we show this on an example of diphtheria toxin repressor (DtxR) crystal structure (see Figure S2 in the Supporting Information). This protein contains a Zn^{2+} binding site, and its structure is deposited in the PDB database both in its apo (1BI2) and holo (1BI3) form, therefore being a good test of the ability of our method to predict realistic protein−ion complexes. Using the apo form as input, we identified, among the highest scoring protein structures, the holo structure of the same protein. With the superimposition of the two protein structures, we obtained an initial prediction of the Zn^{2+} position within the apoprotein. We then performed 1000 steepest descent and 10 000 ABNR minimization steps, resulting in a minimized Zn^{2+} complex. The distances between the Zn^{2+} and the nearest atoms of the binding site amino acid residues of our predicted minimized complex were calculated and compared to those of the original holo structure. The average distance differences were found to be only 0.26 Å after minimization, while, previously, they were 0.41 Å. This shows that ProBiS-CHARMMing was able to simulate the metal binding site of the apoprotein structure with high precision.

■ CONCLUDING REMARKS

In this work, we developed ProBiS-CHARMMing, a web-based interface between two web servers: ProBiS and CHARMMing.

DOI: 10.1021/acs.jcim.5b00534
J. Chem. Inf. Model. 2015, 55, 2308−2314

The ProBiS side of the interface performs a comparison of the query protein surface against a nonredundant protein database to determine similar binding sites, and transposes co-crystallized ligands found in these sites to the query binding site. Provided with the initial prediction of the protein−ligand model, the CHARMMing side of the interface then refines and assesses this model by minimization and subsequent interaction energy calculation. The computational time is dependent mainly on the number of jobs in a queue. Once a job is started, the ProBiS structural similarity calculation and ligand prediction can require up to 10 min to complete; each CHARMMing minimization initiated by the user typically lasts a few minutes, depending on the chosen settings (implicit solvent, number of minimization steps). We have shown how this interface can be used for the prediction of protein−protein, protein−small molecule, and protein−ion complexes. We emphasize its usefulness in target preparation for molecular docking, as it presents a fast method for approximating holoprotein binding sites, when only the apoprotein structures are available.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00534.

Python code example, which demonstrates calls to CHARMMing's API; CHARMM scripts examples (ZIP); protein−peptide model comparison with the 1AZE crystal structure; protein−ion complex example (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: dusanka.janezic@upr.si.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

The financial support through Grant Nos. P1-0002 and J1-6743 of the Slovenian Research Agency is acknowledged. J.K. acknowledges the Fulbright Scholarship at the National Heart, Lung, and Blood Institute, NIH. H.L.W. would like to acknowledge the DOE (No. DE-SC0011297) and NSF (No. CHE-1464946) for Funding. Support by the Intramural Research Program of the National Heart, Lung, and Blood Institute, NIH, is also acknowledged.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank Frank C. Pickard IV and Vinushka Schalk for their assistance with pychm development.

■ REFERENCES

(1) Yuriev, E.; Ramsland, P. A. Latest Developments in Molecular Docking: 2010−2011 in Review. *J. Mol. Recognit.* **2013**, *26*, 215−239.

(2) Konc, J.; Janežič, D. ProBiS-Ligands: A Web Server for Prediction of Ligands by Examination of Protein Binding Sites. *Nucleic Acids Res.* **2014**, *42*, W215−W220.

(3) *ProBiS-ligands Server*; http://probis.cmm.ki.si/ligands/ (accessed Sept. 24, 2015).

(4) Eddy, S. R. Where Did the BLOSUM62 Alignment Score Matrix Come From? *Nat. Biotechnol.* **2004**, *22*, 1035−1036.

(5) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403−410.

(6) Holm, L.; Kääriäinen, S.; Wilton, C.; Plewczynski, D. Using Dali for Structural Comparison of Proteins. In *Current Protocols in Bioinformatics*, Unit 5.5; John Wiley and Sons: New York, 2006; Chapter 5, pp 5.5.1−5.5.24 (DOI: 10.1002/0471250953.bi0505s14).

(7) Nadzirin, N.; Firdaus-Raih, M. Proteins of Unknown Function in the Protein Data Bank (PDB): An Inventory of True Uncharacterized Proteins and Computational Tools for Their Analysis. *Int. J. Mol. Sci.* **2012**, *13*, 12761−12772.

(8) Lee, H. S.; Im, W. Ligand Binding Site Detection by Local Structure Alignment and Its Performance Complementarity. *J. Chem. Inf. Model.* **2013**, *53*, 2462−2470.

(9) Lee, H. S.; Im, W. Identification of Ligand Templates Using Local Structure Alignment for Structure-Based Drug Design. *J. Chem. Inf. Model.* **2012**, *52*, 2784−2795.

(10) Shatsky, M.; Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. J. The Multiple Common Point Set Problem and Its Application to Molecule Binding Pattern Detection. *J. Comput. Biol.* **2006**, *13*, 407−428.

(11) Shulman-Peleg, A.; Shatsky, M.; Nussinov, R.; Wolfson, H. J. MultiBind and MAPPIS: Webservers for Multiple Alignment of Protein 3D-Binding Sites and Their Interactions. *Nucleic Acids Res.* **2008**, *36*, W260−W264.

(12) Konc, J.; Janežič, D. ProBiS Algorithm for Detection of Structurally Similar Protein Binding Sites by Local Structural Alignment. *Bioinformatics* **2010**, *26*, 1160−1168.

(13) Konc, J.; Janezic, D. An Improved Branch and Bound Algorithm for the Maximum Clique Problem. *MATCH* **2007**, *58*, 569−590.

(14) Depolli, M.; Konc, J.; Rozman, K.; Trobec, R.; Janezic, D. Exact Parallel Maximum Clique Algorithm for General and Protein Graphs. *J. Chem. Inf. Model.* **2013**, *53*, 2217−2228.

(15) Miller, B. T.; Singh, R. P.; Klauda, J. B.; Hodoscek, M.; Brooks, B. R.; Woodcock, H. L. CHARMMing: A New, Flexible Web Portal for CHARMM. *J. Chem. Inf. Model.* **2008**, *48*, 1920−1929.

(16) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(17) Weidlich, I. E.; Pevzner, Y.; Miller, B. T.; Filippov, I. V.; Woodcock, H. L.; Brooks, B. R. Development and Implementation of (Q)SAR Modeling within the CHARMMing Web−User Interface. *J. Comput. Chem.* **2015**, *36*, 62−67.

(18) Pevzner, Y.; Frugier, E.; Schalk, V.; Caflisch, A.; Woodcock, H. L. Fragment-Based Docking: Development of the CHARMMing Web User Interface as a Platform for Computer-Aided Drug Design. *J. Chem. Inf. Model.* **2014**, *54*, 2612−2620.

(19) Miller, B. T.; Singh, R. P.; Schalk, V.; Pevzner, Y.; Sun, J.; Miller, C. S.; Boresch, S.; Ichiye, T.; Brooks, B. R.; Woodcock, H. L., III. Web-Based Computational Chemistry Education with CHARMMing I: Lessons and Tutorial. *PLoS Comput. Biol.* **2014**, *10*, e1003719.

(20) Pickard, F. C., IV; Miller, B. T.; Schalk, V.; Lerner, M. G.; Woodcock, H. L., III; Brooks, B. R. Web-Based Computational Chemistry Education with CHARMMing II: Coarse-Grained Protein Folding. *PLoS Comput. Biol.* **2014**, *10*, e1003738.

(21) Perrin, B. S., Jr; Miller, B. T.; Schalk, V.; Woodcock, H. L.; Brooks, B. R.; Ichiye, T. Web-Based Computational Chemistry Education with CHARMMing III: Reduction Potentials of Electron Transfer Proteins. *PLoS Comput. Biol.* **2014**, *10*, e1003739.

(22) Wu, X.; Brooks, B. R. Self-Guided Langevin Dynamics Simulation Method. *Chem. Phys. Lett.* **2003**, *381*, 512−518.

(23) Lee, J.; Miller, B. T.; Damjanovic, A.; Brooks, B. R. Constant pH Molecular Dynamics in Explicit Solvent with Enveloping Distribution

DOI: 10.1021/acs.jcim.5b00534
J. Chem. Inf. Model. 2015, 55, 2308−2314

Sampling and Hamiltonian Exchange. *J. Chem. Theory Comput.* **2014**, *10*, 2738−2750.
(24) Lee, J.; Miller, B. T.; Damjanovic, A.; Brooks, B. R. Enhancing Constant-pH Simulation in Explicit Solvent with a Two-Dimensional Replica Exchange Method. *J. Chem. Theory Comput.* **2015**, *11*, 2560−2574.
(25) Christ, C. D.; van Gunsteren, W. F. Simple, Efficient, and Reliable Computation of Multiple Free Energy Differences from a Single Simulation: A Reference Hamiltonian Parameter Update Scheme for Enveloping Distribution Sampling (EDS). *J. Chem. Theory Comput.* **2009**, *5*, 276−286.
(26) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional Replica-Exchange Method for Free-Energy Calculations. *J. Chem. Phys.* **2000**, *113*, 6042−6051.
(27) Hynninen, A.-P.; Crowley, M. F. New Faster CHARMM Molecular Dynamics Engine. *J. Comput. Chem.* **2014**, *35*, 406−413.
(28) Woodcock, H. L.; Hodoscek, M.; Gilbert, A. T. B.; Gill, P. M. W.; Schaefer, H. F.; Brooks, B. R. Interfacing Q-Chem and CHARMM to Perform QM/MM Reaction Path Calculations. *J. Comput. Chem.* **2007**, *28*, 1485−1502.
(29) Woodcock, H. L.; Miller, B. T.; Hodoscek, M.; Okur, A.; Larkin, J. D.; Ponder, J. W.; Brooks, B. R. MSCALE: A General Utility for Multiscale Modeling. *J. Chem. Theory Comput.* **2011**, *7*, 1208−1219.
(30) Konc, J.; Česnik, T.; Konc, J. T.; Penca, M.; Janežič, D. ProBiS-Database: Precalculated Binding Site Similarities and Local Pairwise Alignments of PDB Structures. *J. Chem. Inf. Model.* **2012**, *52*, 604−612.
(31) *Jmol: An Open-Source Java Viewer for Chemical Structures in 3D*; http://jmol.sourceforge.net/ (accessed May 28, 2015).
(32) Richardson, L.; Ruby, S.; Hansson, D. H. *RESTful Web Services*, 1st Edition; O'Reilly Media: Farnham, Surrey, U.K., 2007.
(33) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(34) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671−690.
(35) Jarvis, R. M.; Broadhurst, D.; Johnson, H.; O'Boyle, N. M.; Goodacre, R. PYCHEM: A Multivariate Analysis Package for Python. *Bioinformatics* **2006**, *22*, 2565−2566.
(36) Schalk, V. *Pychm*; https://git.lobos.nih.gov/vinushka/pychm (accessed May 24, 2015).
(37) Vanommeslaeghe, K.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144−3154.
(38) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155−3168.
(39) *CGenFF*; https://cgenff.paramchem.org/ (accessed May 24, 2015).
(40) Vidal, M.; Goudreau, N.; Cornille, F.; Cussac, D.; Gincel, E.; Garbay, C. Molecular and Cellular Analysis of Grb2 SH3 Domain Mutants: Interaction with Sos and dynamin. *J. Mol. Biol.* **1999**, *290*, 717−730.
(41) McGovern, S. L.; Shoichet, B. K. Information Decay in Molecular Docking Screens against Holo, Apo, and Modeled Conformations of Enzymes. *J. Med. Chem.* **2003**, *46*, 2895−2907.
(42) Fischer, M.; Coleman, R. G.; Fraser, J. S.; Shoichet, B. K. Incorporation of Protein Flexibility and Conformational Energy Penalties in Docking Screens to Improve Ligand Discovery. *Nat. Chem.* **2014**, *6*, 575−583.
(43) Vankayala, S. L.; Hargis, J. C.; Woodcock, H. L. Unlocking the Binding and Reaction Mechanism of Hydroxyurea Substrates as Biological Nitric Oxide Donors. *J. Chem. Inf. Model.* **2012**, *52*, 1288−1297.
(44) Vankayala, S. L.; Hargis, J. C.; Woodcock, H. L. How Does Catalase Release Nitric Oxide? A Computational Structure−Activity Relationship Study. *J. Chem. Inf. Model.* **2013**, *53*, 2951−2961.
(45) Eschenburg, S.; Priestman, M. A.; Abdul-Latif, F. A.; Delachaume, C.; Fassy, F.; Schönbrunn, E. A Novel Inhibitor That Suspends the Induced Fit Mechanism of UDP−*N*-Acetylglucosamine Enolpyruvyl Transferase (MurA). *J. Biol. Chem.* **2005**, *280*, 14070−14075.

DOI: 10.1021/acs.jcim.5b00534
J. Chem. Inf. Model. 2015, 55, 2308−2314