Published online 11 May 2009 *Nucleic Acids Research, 2009, Vol. 37, Web Server issue* ***W559–W564***
doi:10.1093/nar/gkp359

SLITHER: a web server for generating contiguous conformations of substrate molecules entering into deep active sites of proteins or migrating through channels in membrane transporters

Po-Hsien Lee[1], Kuei-Ling Kuo[2], Pei-Ying Chu[1], Eric M. Liu[2] and Jung-Hsin Lin[1,2,3,*]

[1]Division of Mechanics, Research Center for Applied Sciences, Academia Sinica, [2]School of Pharmacy, National Taiwan University and [3]Institute of Biomedical Science, Academia Sinica, Taipei, Taiwan

Received March 8, 2009; Revised April 19, 2009; Accepted April 23, 2009

ABSTRACT

Many proteins use a long channel to guide the substrate or ligand molecules into the well-defined active sites for catalytic reactions or for switching molecular states. In addition, substrates of membrane transporters can migrate to another side of cellular compartment by means of certain selective mechanisms. SLITHER (http://bioinfo.mc.ntu.edu.tw/slither/or http://slither.rcas.sinica.edu.tw/) is a web server that can generate contiguous conformations of a molecule along a curved tunnel inside a protein, and the binding free energy profile along the predicted channel pathway. SLITHER adopts an iterative docking scheme, which combines with a puddle-skimming procedure, i.e. repeatedly elevating the potential energies of the identified global minima, thereby determines the contiguous binding modes of substrates inside the protein. In contrast to some programs that are widely used to determine the geometric dimensions in the ion channels, SLITHER can be applied to predict whether a substrate molecule can crawl through an inner channel or a half-channel of proteins across surmountable energy barriers. Besides, SLITHER also provides the list of the pore-facing residues, which can be directly compared with many genetic diseases. Finally, the adjacent binding poses determined by SLITHER can also be used for fragment-based drug design.

INTRODUCTION

In order to achieve specific recognition or to enhance association efficiency, many proteins use long- or half-channels to guide the substrates or ligands into deep catalytic sites or binding pockets for catalytic reactions or for switching molecular states. For example, the acetylcholinesterase uses a 20 Å aromatic gorge (1) to guide the substrate acetylcholine into the catalytic triad of this enzyme. The histone deacetylase (HDAC) uses a 14 Å tunnel (2,3) to assist the catalysis of the removal of acetyl groups for the ε-amino group of lysine residues of nucleosomal histones. G-protein coupled receptors used a half-channel to guide the ligand into a deep binding pocket and thereby change the molecular states (4). Membrane transporters generally use a curved tunnel to migrate substrates into the other side of the plasma membrane in a selective manner (5).

Recognizing the significance of channels within the protein structure, various algorithms, e.g. HOLE (6,7) and MOLE (8), have been developed to characterize the geometrical properties of these channels or to determine the pore lining residues (8). Although geometric dimensions may be useful characteristics for channels of ions, most molecular ligands or substrates are distinct from such simple objects. For a given protein structure, one often wonders whether this structure depicts a conducting or non-conducting state for its substrate or ligand, which cannot be directly judged by the pore radius profile. For most molecular substrates or ligands with shapes other than the simple sphere, it is of great interest to see whether and how the substrates or ligands can crawl through the channel by crossing the surmountable barriers. SLITHER is therefore developed to provide the free energy profile along the access channel, the list of pore-lining residues, and the binding poses of the ligand with these residues.

METHODS

An iterative docking scheme is employed to generate contiguous conformations and their corresponding binding

*To whom correspondence should be addressed. Tel: +886 2 2782 3212, ext. 886; Fax: +886 2 2782 3060; Email: jlin@ntu.edu.tw; jhlin@gate.sinica.edu.tw

free energies along a curved tunnel inside a protein. The kernel of this scheme is the docking calculation, and both AutoDock (9,10) and MEDock (11) have been adopted in this SLITHER web server. In the preparatory stage the server will generate the energetic grid maps, which store the potential energies probed by various atom types of the ligand molecule on the three-dimensional (3D) meshes in a specific region of the receptor molecule. The potential energy terms include the van der Waals interactions, hydrogen bond interactions, electrostatic interactions, and desolvation free energy. The free energy calculation implemented in AutoDock 3.05 (9) and AutoDock 4 (10) is based on semi-empirical formula derived from multivariate linear regression, using above potential energy terms and a conformational entropy term. The details of the free energy formula can be found in the original articles of AutoDock 3.05 (9) and AutoDock 4 (10).

Because the grid map files will limit the search ranges in the docking calculations, it is recommended to assign the grid center and the grid box dimensions with caution. For transmembrane proteins, it is suggested that the grid box cover the entire membrane-embedding region of the protein. Transmembrane regions of a protein can be predicted by, e.g. TMHMM (12). Either the Lamarckian genetic algorithm (LGA) in AutoDock (9) or the evolutionary Gaussian algorithm (EGA) in MEDock (11) is supposed to find out the global minimum of the free energy in the grid space, which is depicted by the location, orientation and conformation of the ligand with the lowest free energy determined by the empirical scoring function. In the iterative docking scheme adopted in our SLITHER server, the energies of the grid maps will be elevated for 5 kcal/mol at the grid points that are within a $(2.25\,\text{Å})^3$—cube centered at the ligand atoms, and then a subsequent docking calculation is performed to find out the new global minimum. The users can specify the overlapping parameter, which is the fraction of ligand atoms whose neighboring grid points will be raised. The allowed overlapping parameter is between 0 and 0.8. The overlapping parameter of zero indicates that the ligand atoms between adjacent iterations will not have any spatial overlap. On the other hand, the overlapping parameter of 0.8 indicates that 80% of the ligand atoms can be overlapped with those in the subsequent iteration. It should be stressed that 'can be' overlapped does not mean 'will be' overlapped. The users can also specify the number of iterations to be conducted. Finally, clustering of ligand conformations will be performed. Two ligand conformations will be assigned to the same cluster if the nearest distance of the atoms in these two conformations is <4 Å. After all the ligand conformations have been clustered, the largest cluster is outputted. The ligand conformations of this cluster are also used as the probes to detect the pore facing residues in the receptor. Pore facing residues are those residues whose heavy atoms are within a specified distance of any ligand atom. Finally, the binding free energy profile will be constructed according to the position along the channel pathway.

The core program of SLITHER is written in Perl 5.8.8, while the web interface is written in PHP (5.1.6). Rasmol (http://www.bernstein-plus-sons.com/software/RasMol_2.7.2.1/) was used to generate the animation of ligands slithering in the protein. In addition, user can manipulate the resulting 3D structures interactively with the JmolApplet embedded in the result displaying pages. JmolApplet is the Java applet of Jmol (http://www.jmol.org/), which is an open-source Java viewer for chemical structures in 3D. Gnuplot (http://www.gnuplot.info/) is used to generate the Slithering Energetics plots.

INPUT, OUTPUT AND OPTIONS

The input window of the SLITHER web server is shown in Figure 1. The channel or half-channel of protein in question needs to be aligned to the *y*-axis, by, e.g. Chimera (13), and the transformed coordinates in the PDB format be outputted. AutoDockTools (ADT) (14) can then be used to visually specify the coordinates of the grid center and the dimensions of the grid box. The dimensions of the grid box should be at least slightly larger than those of the channel or the half-channel. The default input file format is the so-called PDBQ format, which is an extension of the PDB format. In the PDBQ file, all the hydrogen atoms have been added and the partial charges have been assigned to all the atoms. The PDBQ file of the ligand molecule can be generated by many chemical software or web servers, e.g. Dundee's PRODRG server (15) (http://davapc1.bioch.dundee.ac.uk/programs/prodrg/). In the ligand PDBQ file, the fixed portion of the atoms in the ligand molecules are grouped into 'root', from which rotatable 'branches' sprout. 'Torsions' are special cases of 'branches', where the two atoms at either end of the rotatable bond have only two nearest neighbors. Different torsional angles on the 'branches' and 'torsions' will be evaluated according to a simplified torsional energy function implemented within AutoDock (9). The PDBQ file for proteins can be generated by the PDB2PQR (16) (http://agave.wustl.edu/pdb2pqr/) and a simple awk or perl script. The PDB2PQR server will provide predictions of the protonation states of the ionizable residues in a protein at a given pH value. Our SLITHER web server will perform automatic conversion of the PDB or the PQR format into the PDBQ format, if the users do not have their preferred procedures for the required conversion. Our web server also provides the 'flexible receptor' mode and the 'relaxed receptor' mode to accommodate the receptor flexibility in the iterative docking scheme. In the 'flexible receptor' mode, users can select the residues on which they would like to observe the effect of flexibility, which is a new feature of AutoDock 4. In the 'relaxed receptor' mode, users can upload a set of receptor conformations, either generated from molecular dynamics simulations or other conformation sampling techniques, and then the iterative docking scheme in the SLITHER server will be performed on each of these conformations. This mode can be considered the extended version of the relaxed complex scheme (17,18), in which the protein flexibility has been accommodated by molecular dynamics simulation for high

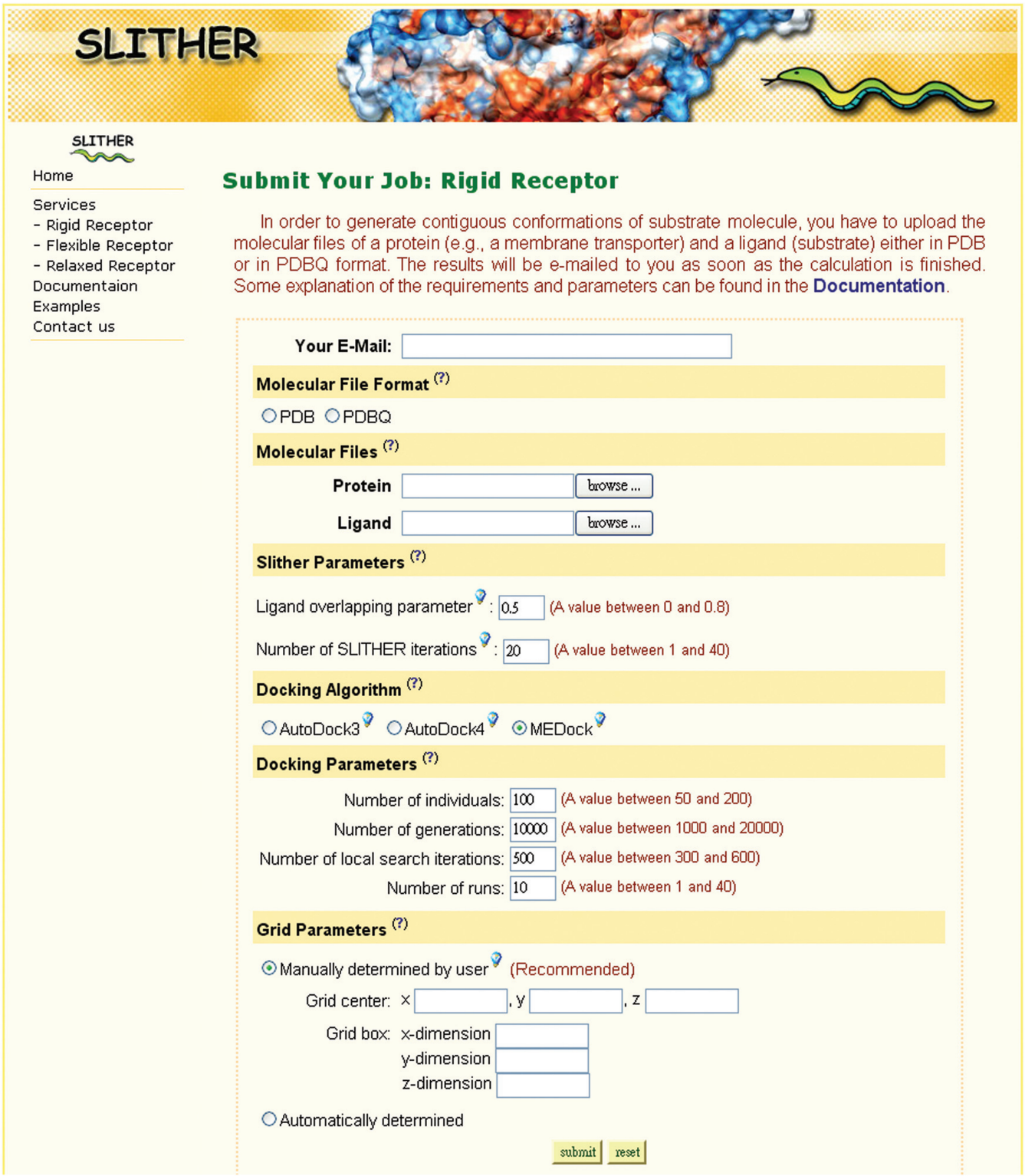

Figure 1. Screenshot of the SLITHER input page.

resolution drug design. Pore-facing residues will be listed according to their amino acid sequence. In the relaxed receptor mode, the 'consensus' pore-facing residues, i.e. the pore-facing residues that are found in all uploaded conformations, will be listed.

RESULTS AND DISCUSSION

To demonstrate that our SLITHER server can also be used to explain some human genetic diseases, we chose the human glucose transporter Glut1 as an example. The homology model of this transporter (19) was retrieved

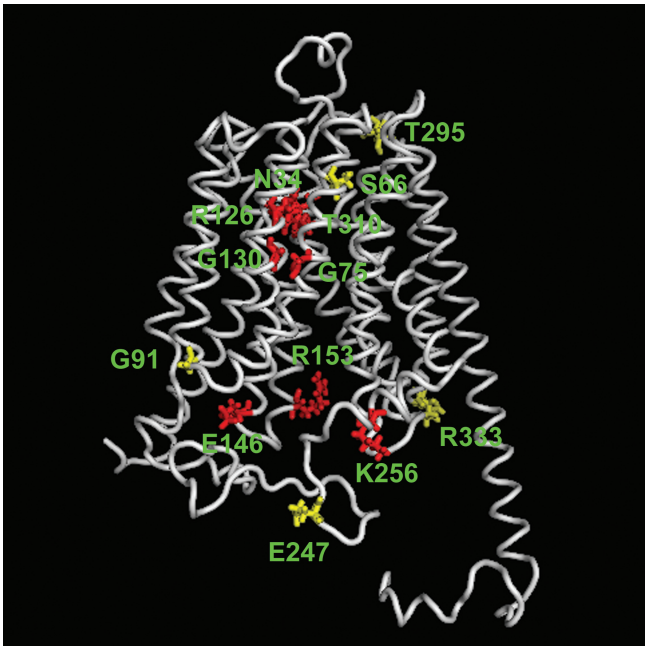

Figure 2. Red: the pore-facing residues corresponding to the missense mutation of GLUT1 related to the Glut1 deficiency syndrome. Yellow: residues related to the missense mutation but not along the access channel.

from the Protein Databank (PDB ID: 1SUK) (20). Some known missense mutations of the well-studied Glut1 deficiency syndrome (21) can be associated with the pore-facing residues in the human glucose transporter Glut1. In Figure 2, the residues colored in red are the pore-facing residues corresponding to the missense mutation of GLUT1 related to the Glut1 deficiency syndrome, while the residues colored in yellow are those related to the missense mutation but not along the access channel. It is shown that the known residues responsible for the Glut1 deficiency syndrome have significant overlaps with the pore-facing residues. It should be noted that the correlation of these missense mutations with such pore-facing residues also depends on the homology modeling methodology and the sequence similarity of the structural template to Glut1. Figure 3a is the slab view of the molecular surface representations of the human glucose transporter Glut1. The geometry of the access channel is depicted by HOLE. Figure 3b is similar to Figure 3a, but the substrate conformations along the access channel are generated by SLITHER. Because the structural template of this human Glut1 homology model is the X-ray crystallographic structure of glycerol-3-phosphase transporter (GlpT) at the inward-facing conformation (22), i.e. its access channel should be closed from the periplasmic side, it can be seen from the HOLE plot that the pore radius is only 1 Å at the 31.2 Å position. Our SLITHER free energy profile also indicates a major barrier peaked at 32.3 Å. Although the geometry of the channel seems to be a good indicator for the possible steric barrier for the substrate entrance pathway, the free energy profile generated by the SLITHER server showed that other energetic contribution may also play important roles. As depicted in Figure 3c and d, the locations of the peaks of the energetic profiles are not the always correlated with the locations of the valleys of the pore radius profile. This is understandable because in the energetic analysis, not only the steric effects (e.g. van der Waals interactions), is taken into account. Electrostatic interactions, hydrophobic interactions, cation–π and π –π interactions, among many others, can also contribute to the binding free energies. Admittedly, the accuracy of current SLITHER energetic analysis is limited by the scoring functions used in the iterative docking procedure. On the other hand, it should also be noted that neither the pore radius profile nor the slither energetic profile from a single static structure should be considered the actual free energy profile when a substrate is crawling through the entrance pathway. It has been gradually established that at least for some transport protein, e.g. AcrB, the peristaltic motions are required for allowing the substrates to pass the channel (23). The slithering energetic should be considered a conceptual device to describe the energetic profile for a given structure. It will be ideal if an ensemble of structures or trajectories of molecular dynamics can be combined with such the slithering energetic analysis. The changes of energetic profiles upon the conformational changes will provide a physical explanation why some conformational changes are beneficial for the entrance or exit of substrates.

With the aid of programs for analyzing protein–ligand interactions, e.g. LigPlot (24), MOE (25), etc. it is possible to further dissect the types of molecular interactions, e.g. cation–π, π–π, ionic, van der Waals or hydrophobic interactions, with the pore-facing residues. Such information will be useful to explain the mutagenesis data to design new mutagenesis experiments and also to elucidate the roles of some single nucleotide polymorphisms (SNPs). On the other hand, the non-overlapping adjacent low energy binding modes of two identical molecules or two different molecules can be used to design composite molecules that are made by linking these two molecules with appropriate linkers, which is one of the simplest forms of the fragment-based drug design. The composite compound will generally exhibit stronger binding affinity than the original individual compound if the linkers are not too inappropriate.

SLITHER can also be regarded as an extension or an enhanced version of docking programs. Docking programs promise to predict the ligand binding pose with the lowest free energy, usually done with a global optimization algorithm. However, suitable parameters are required for a given algorithm to accurately predict the global minima, and many repeated runs are often necessary for confirming its identification. Since SLITHER is enforced to explore the space other than that occupied by the previously found 'global minimum', if a lower free energy is found in the subsequent iteration, it indicates that the docking parameters need to be improved. With a proper set of docking parameters, the free energies of different iterations should follow an ascending order.

Finally, we would like to stress that the strength of our SLITHER web site is for ligand or substrate molecules with conformational variability, not for ligands like ions

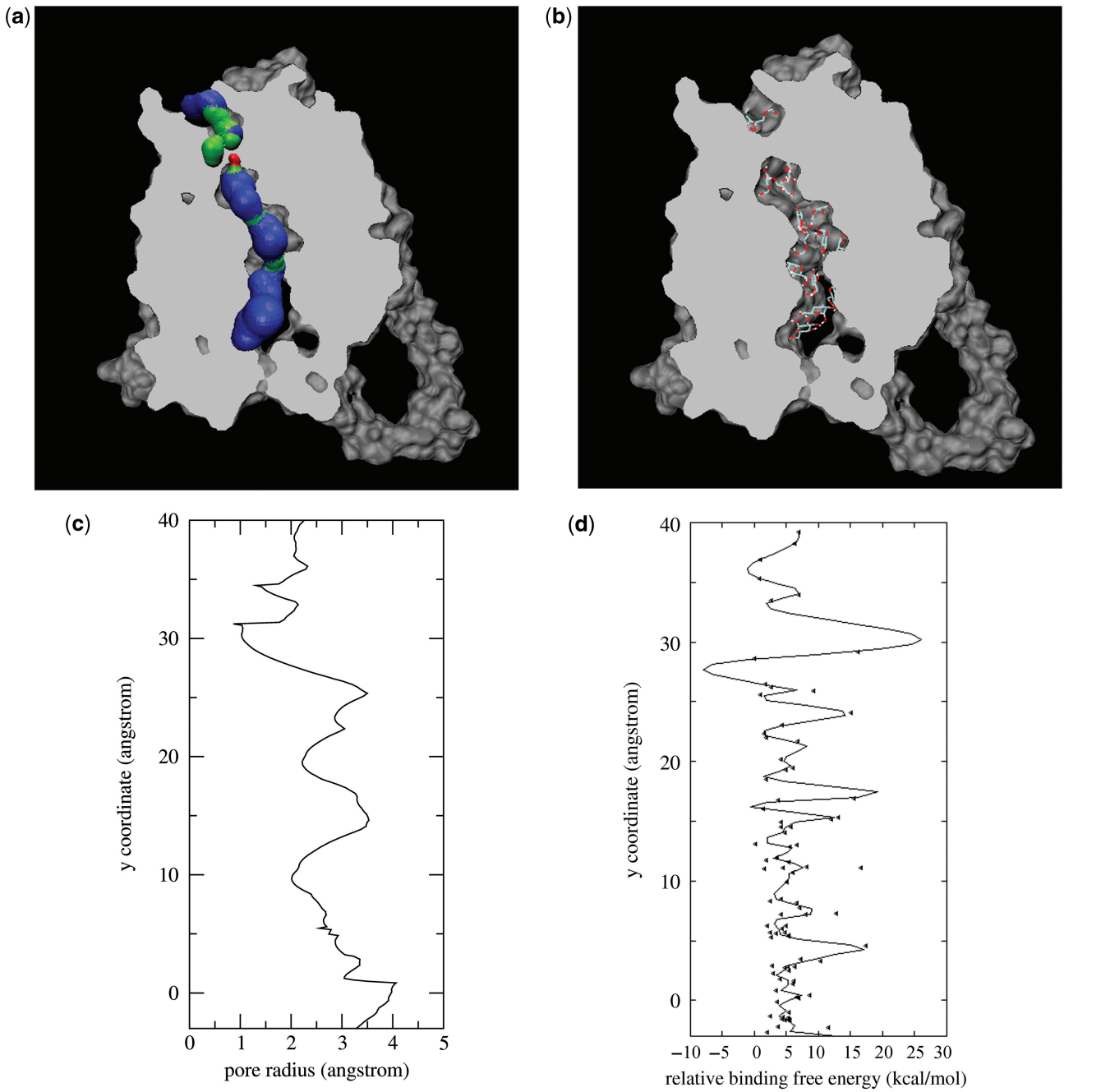

Figure 3. (**a**) A slab view of the molecular surface representations of the human glucose transporter. The geometry of the access channel is depicted by HOLE. (**b**) Similar to (a), but the substrate conformations along the access channel are generated by SLITHER. (a) and (b) were created with DINO (http://www.dino3d.org/). (**c**) Pore radius profile along the access channel, generated by HOLE. (**d**) Free energy profile along the access channel, generated by SLITHER.

with simple spherical shape. In principle, one could apply rigorous statistical mechanics-based potential of mean force calculations using molecular dynamics simulations with explicit solvent (or also with explicit lipid bilayer) for the free energy profiles of a substrate molecule entering the channel within a protein, but they are typically too time-consuming and computationally demanding, and the convergence of such calculations is always a major concern. Although the scoring function used in the SLITHER web server is of semi-empirical nature, it provides a rapid, alternative route for the qualitative description of the conformational influences on the free energy profiles.

CONCLUSION

Our SLITHER web server can be used to predict the free energy profile along the access channel within a protein, the pore-lining residues, and the binding poses of the ligand with these residues. Such information can be used

to explain the mutagenesis data, to design new mutagenesis experiment, and to elucidate the molecular origins of genetic diseases. This web server can also be used to perform fragment-base drug design.

ACKNOWLEDGEMENTS

Support from Research Center for Applied Science, Academia Sinica is greatly acknowledged.

FUNDING

NSC 95-2112-M-002 -007, 96-2627-B-002-00, 97-2323-B-002-015 and 97-2323-B-077-001. Funding for open access charge: Research Center for Applied Sciences, Academia Sinica.

Conflict of interest statement. None declared.

REFERENCES

1. Sussman,J.L., Harel,M., Frolow,F., Oefner,C., Goldman,A., Toker,L. and Silman,I. (1991) Atomic-structure of acetylcholinesterase from torpedo-californica—a prototypic acetylcholine-binding protein. *Science*, **253**, 872–879.
2. Vannini,A., Volpari,C., Filocamo,G., Casavola,E.C., Brunetti,M., Renzoni,D., Chakravarty,P., Paolini,C., De Francesco,R., Gallinari,P. *et al*. (2004) Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor. *Proc. Natl Acad. Sci. USA*, **101**, 15064–15069.
3. Finnin,M.S., Donigian,J.R., Cohen,A., Richon,V.M., Rifkind,R.A., Marks,P.A., Breslow,R. and Pavletich,N.P. (1999) Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature*, **401**, 188–193.
4. Marinissen,M.J. and Gutkind,J.S. (2001) G-protein-coupled receptors and signaling networks: emerging paradigms. *Trends Pharmacol. Sci*., **22**, 368–376.
5. Hediger,M.A., Romero,M.F., Peng,J.B., Rolfs,A., Takanaga,H. and Bruford,E.A. (2004) The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins—introduction. *Pflugers Arch.-Eur. J. Physiol*., **447**, 465–468.
6. Smart,O.S., Goodfellow,J.M. and Wallace,B.A. (1993) The pore dimensions of gramicidin—A. *Biophys. J*., **65**, 2455–2460.
7. Smart,O.S., Neduvelil,J.G., Wang,X., Wallace,B.A. and Sansom,M.S.P. (1996) HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph*., **14**, 354–360.
8. Petrek,M., Kosinova,P., Koca,J. and Otyepka,M. (2007) MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure*, **15**, 1357–1363.
9. Morris,G.M., Goodsell,D.S., Halliday,R.S., Huey,R., Hart,W.E., Belew,R.K. and Olson,A.J. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem*., **19**, 1639–1662.
10. Huey,R., Morris,G.M., Olson,A.J. and Goodsell,D.S. (2007) A semiempirical free energy force field with charge-based desolvation. *J. Computat. Chem*., **28**, 1145–1152.
11. Chang,D.T.H., Oyang,Y.J. and Lin,J.H. (2005) MEDock: a web server for efficient prediction of ligand binding sites based on a novel optimization algorithm. *Nucleic Acids Res*, **33**, W233–W238.
12. Krogh,A., Larsson,B., von Heijne,G. and Sonnhammer,E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol*., **305**, 567–580.
13. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem*., **25**, 1605–1612.
14. Sanner,M.F. (1999) Python: a programming language for software integration and development. *J. Mol. Graph*., **17**, 57–61.
15. Schuttelkopf,A.W. and van Aalten,D.M.F. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. Sect. D—Biol. Crystallogr*., **60**, 1355–1363.
16. Dolinsky,T.J., Czodrowski,P., Li,H., Nielsen,J.E., Jensen,J.H., Klebe,G. and Baker,N.A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res*., **35**, W522–W525.
17. Lin,J.H., Perryman,A.L., Schames,J.R. and McCammon,J.A. (2002) Computational drug design accommodating receptor flexibility: the relaxed complex scheme. *J. Am. Chem. Soc*., **124**, 5632–5633.
18. Lin,J.H., Perryman,A.L., Schames,J.R. and McCammon,J.A. (2003) The relaxed complex method: accommodating receptor flexibility for drug design with an improved scoring scheme. *Biopolymers*, **68**, 47–62.
19. Salas-Burgos,A., Iserovich,P., Zuniga,F., Vera,J.C. and Fischbarg,J. (2004) Predicting the three-dimensional structure of the human facilitative glucose transporter Glut1 by a novel evolutionary homology strategy: Insights on the molecular mechanism of substrate migration, and binding sites for glucose and inhibitory molecules. *Biophys. J*., **87**, 2990–2999.
20. Berman,H.M., Westbrook,J., Feng,Z., Gilliland,G., Bhat,T.N., Weissig,H., Shindyalov,I.N. and Bourne,P.E. (2000) The protein data bank. *Nucleic Acids Res*., **28**, 235–242.
21. Klepper,J. and Leiendecker,B. (2007) GLUT1 deficiency syndrome—2007 update. *Dev. Med. Child Neurol*., **49**, 707–716.
22. Huang,Y.F., Lemieux,M.J., Song,J.M., Auer,M. and Wang,D.N. (2003) Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science*, **301**, 616–620.
23. Seeger,M.A., Schiefner,A., Eicher,T., Verrey,F., Diederichs,K. and Pos,K.M. (2006) Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science*, **313**, 1295–1298.
24. Wallace,A.C., Laskowski,R.A. and Thornton,J.M. (1995) Ligplot—a program to generate schematic diagrams of protein ligand interactions. *Protein Eng*., **8**, 127–134.
25. Clark,A.M. and Labute,P. (2007) 2D depiction of protein—ligand complexes. *J. Chem. Inform. Model*., **47**, 1933–1944.