Using PyMOL as a platform for computational drug design

Article *in* Wiley interdisciplinary reviews: Computational Molecular Science · January 2017

DOI: 10.1002/wcms.1298

CITATIONS
10

READS
2,082

3 authors:

Shuguang Yuan
Shenzhen Institutes of Advanced Technology Chinese Academy of Sciences
57 PUBLICATIONS **872** CITATIONS
SEE PROFILE

Zhenquan Hu
Hefei Institute of Physical Sciences, Chinese Academy of Sciences
34 PUBLICATIONS **140** CITATIONS
SEE PROFILE

H. C. Stephen Chan
26 PUBLICATIONS **223** CITATIONS
SEE PROFILE

Software Focus

Using PyMOL as a platform for computational drug design

Shuguang Yuan [1]* H.C. Stephen Chan[2] and Zhenquan Hu[3]

PyMOL, a cross-platform molecular graphics tool, has been widely used for three-dimensional (3D) visualization of proteins, nucleic acids, small molecules, electron densities, surfaces, and trajectories. It is also capable of editing molecules, ray tracing, and making movies. This Python-based software, alongside many Python plugin tools, has been developed to enhance its utilities and facilitate the drug design in PyMOL. To gain an insightful view of useful drug design tools and their functions in PyMOL, we present an extensive discussion on various molecular modeling modules in PyMOL, covering those for visualization and analysis enhancement, protein–ligand modeling, molecular simulations, and drug screening. This review provides an excellent introduction to present 3D structures visualization and computational drug design in PyMOL.

How to cite this article:
WIREs Comput Mol Sci 2017, e1298. doi: 10.1002/wcms.1298

INTRODUCTION

PyMOL is an open-source molecular visualization system created by Warren Lyford DeLano and commercialized initially by DeLano Scientific LLC. In 2010, Schrödinger Inc. reached an agreement to acquire PyMOL. From then on, Schrödinger has taken over the development and maintenance, as well as support and sale of PyMOL, including all current subscription. PyMOL uses OpenGL Extension Wrangler Library (GLEW) and Free OpenGL Utility Toolkit (Freeglut). PyMOL uses cross-platform widget toolkit (Tk) for the GUI widgets. PyMOL can produce high-quality movies and images of macromolecules in different representations including ribbons, cartoons, dots, surfaces, spheres, sticks, and lines (Figure 1). At present, PyMOL is one of the most widely used macromolecular visualization tools.

Since PyMOL is written in Python, one of the most popular programming languages, it can be extended to Python plugins easily. Apart from discussing the visualization and the enhanced analysis functions in PyMOL, our topics also extend to the protein–ligand modeling, molecular simulations (MS), and virtual screening (VS) unities in PyMOL. The computational drug discovery function of PyMOL has been successfully applied to find new drug candidates for various targets. These include the discovery of a potent small molecule inhibitor for gankyrin,[1] lead optimization for Cytochrome P450 enzymes,[2] VS of new drug candidates for the tumor suppressor protein P53,[3] and so on.

In this Software Focus article, we cover the usage of PyMOL and its plugins as a platform for computational drug design. We have summarized PyMOL plugins and their functions in Table 1.

*Correspondence to: shuguang.yuan@gmail.com

[1]Laboratory of Physical Chemistry of Polymers and Membranes, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[2]Faculty of Life Sciences, University of Bradford, Bradford, UK

[3]High Magnetic Field Laboratory, Chinese Academy of Science, Hefei, China

Conflict of interest: The authors have declared no conflicts of interest for this article.

VISUALIZATION AND ANALYSIS ENHANCEMENT FOR DRUG DESIGN

Rational drug design is the inventive process of finding new compounds, based on the knowledge of a biological target.[4,5] Nowadays, this process usually relies heavily on the computer visualization and modeling, and is therefore known as the computer-aided drug design (CADD).[6] CADD involves small

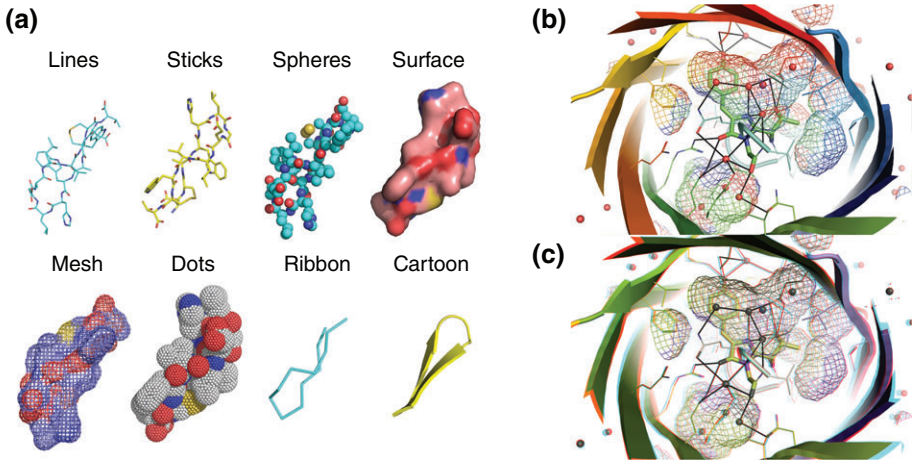

FIGURE 1 | Representations of macromolecules and the anaglyph stereo mode in PyMOL. (a) PyMOL can display macromolecules in different styles including lines, sticks, spheres, surface, mesh, dots, ribbon, and carton. (b) The binding pocket of the green fluorescent protein (pdb: 1EMA) displayed in nonstereo cartoon and mesh. (c) The binding pocket of the green fluorescent protein (pdb: 1EMA) displayed in anaglyph stereo mode. The stereo effect can be visualized by a pair of red–blue anaglyph stereo glasses.

molecular preparation [such as generating the three-dimensional (3D) structure from a two-dimensional (2D) scheme, ligand conformation search, and energy minimization], homology modeling, protein preparation (such as pKa calculation, Asn, Gln, and His side-chain flipping, H-bond network optimization, and energy minimization), lead design (such as docking, design *in situ*, pharmacophore model, and VS), and MS [molecular dynamics (MD), simulation, water network prediction, and free energy calculation] (Figure 2). In the following paragraphs, we will discuss how PyMOL can be applied to each stage of CADD.

Macromolecular Visualization

Visualization of molecules is the preliminary task of CADD. PyMOL has been widely used for 3D visualization of macromolecules and it becomes one of the most popular tools for preparing high-resolution images of macromolecules for publications.[7] However, many users may ignore or are not aware of many other useful modules and functions, which perfectly facilitate the CADD. For instance, PyMOL supports 12 different stereo visualization modes, which enables the users to conceive much more 3D information for the atom positions in space. This useful function, inspiring new ideas for the compound modification (Figure 1), are seldom used. Instead, many users only stick to the 2D mode, rotating the models with their mouse. The stereo view is of great importance for *in situ* design. For instance, to obtain detailed information of the binding pocket for the design of novel molecules, the protein surface can be shown as a mesh (Figure 1(b) and (c)). It conveys many useful messages, including the size of binding pocket and the electrostatic properties which can be generated by the APBS plug-in.[8] This information is particularly useful for designing new compounds. However, the mesh is so condensed that it is very difficult to visualize each atom in nonstereo mode (Figure 1(b)). In contrast, viewing the stereo mode (Figure 1(c)) with 3D glasses could overcome this weakness of 2D mode and facilitate the *in situ* design.

Movie Making

Under the movie menu, PyMOL has a default functionality to make a basic movie, based on either a single static structure or a series of frames from simulations. The eMovie plugin[9] in PyMol further provides a more user-friendly way to create movies. Modular actions such as zooms, rotations, fading, and morphs can be inserted to any frame in the movie, whereas the actions comprising the movie can be reviewed in list-format by viewing the eMovie storyboard.[9] The storyboard also allows for deletion or reinsertion. Finally, movies can be generated and exported easily in eMovie.

Macromolecule Editing

PyMOL is capable of editing a macromolecule including changing the dihedral angles, constructing a peptide, mutating a residue, and so on. One can also easily build a molecule from scratch with the Builder toolbox in PyMOL (Figure 2). Unfortunately, many users do not even realize the existence of such helpful functions for drug design in PyMOL. Instead,

TABLE 1 | Summary of PyMOL Plugins and Their Functions

PLUGINS	Description	Website	License
Visualization and analysis enhancement			
BONDPACK	A collection of PyMOL plugins to visualize atomic bonds	https://github.com/rasbt/BondPack	Open source
DYNOPLOT	Generating Ramachandran (ϕ/ψ) plots for proteins	http://www.pymolwiki.org/index.php/DynoPlot	Open source
SCULPTING	Real-time energy minimizer	http://pymolwiki.org/index.php/Molecular_Sculpting	Open source
CMPYMOL	Analyzing protein interactions	https://github.com/emptyewer/CMPyMOL	Open source
MOLE	A Voronoi diagram-based explorer of molecular channels	http://webchem.ncbr.muni.cz/Platform/App/Mole	Free to academy
CAVER	Calculation and visualization of tunnels	http://www.caver.cz/index.php?sid=199	Free to academy
PYANM	Protein elastic network models analysis	http://pymolwiki.org/index.php/PyANM	Open source
BNITOOLS	Enhanced visualization toolkits	https://sourceforge.net/projects/bni-tools/	Open source
DSSP and STRIDE	Secondary structure prediction	http://www.biotec.tu-dresden.de/~hongboz/dssp_pymol/dssp_pymol.html	Free to academy
APBS	Electrostatic map calculation	http://www.poissonboltzmann.org/examples/comp_tut	Free to academy
MIPTOOL	Lipophilic surface visualization	http://mlptools.altervista.org/	Open source
AZAHAR	Analyze and model glycans and glycoconjugated molecules	https://pymolwiki.org/index.php/Azahar	Open source
PYTMS	Common posttranslational modifications	https://pymolwiki.org/index.php/Pytms	Open source
PLIP	Protein–ligand interaction	https://projects.biotec.tu-dresden.de/plip-web/plip	Free to academy
EMOVIE	Movie making toolkits	https://pymolwiki.org/index.php/EMovie	Open source
PROMOL	Predict the function of proteins	http://www.promol.org	Open source
AVERAGE3D	Calculate the average structure of multiple frames	http://muralab.org/~cmura/PyMOL/	Open source
Protein–ligand modeling			
OPTIMIZE	Ligand conformation generation and energy minimization	http://www.pymolwiki.org/index.php/Optimize	Open source
T-CONCOORD-GUI	Ligand conformation generation	http://wwwuser.gwdg.de/~dseelig/cncplugin.html	Open source
LIGAIGN	Flexible ligand alignment	http://compbio.cs.toronto.edu/ligalign/	Open source
KABSCH	Structural alignment	http://www.pymolwiki.org/index.php/Kabsch	Open source
PHARMDOCK	Pharmacophore modeling, protein–ligand docking and virtual screening	http://people.pharmacy.purdue.edu/~mlill/software/pharmdock/	Open source
MLP GOLD	Protein–ligand docking	http://mlptools.altervista.org/	Additional license needed for Gold docking
LIQUID	Pharmacophore modeling	http://gecco.org.chemie.uni-frankfurt.de/liquid/index.html	Open source
ALIGN-IT	Pharmacophore modeling and ligand alignment	http://silicos-it.be.s3-website-eu-west-1.amazonaws.com/software/align-it/1.0.4/align-it.html	Open source
DRUGON	Pharmacophore modeling and structure optimization	http://bioinfoteam.com/software.html	Free to academy
POCKETPICKER	Binding site prediction	http://gecco.org.chemie.uni-frankfurt.de/pocketpicker/index.html	Open source

(*continued overleaf*)

TABLE 1 | Continued

PLUGINS	Description	Website	License
LIGALIGN	Automated ligand-based active site alignment	http://compbio.cs.toronto.edu/ligalign/	Open source
AUTODOCK/ VINA	Protein–ligand docking	http://wwwuser.gwdg.de/~dseelig/adplugin.html	Open source
FLEXAID	Protein–ligand docking	http://bcb.med.usherbrooke.ca/FlexAID.php	Free to academy
PYMOD	Homology modeling	http://schubert.bio.uniroma1.it/pymod/index.html	Free to academy, additional license needed for Modeller
PYROSETTA	Homology modeling, protein–ligand docking	http://www.pyrosetta.org/	Free to academy, additional license needed for PyRosetta
Virtual screening			
LISICA	2D and 3D ligand based virtual screening	http://insilab.org/lisica-plugin/	Free to academy
PHARMDOCK	Pharmacophore modeling, protein–ligand docking and virtual screening	http://people.pharmacy.purdue.edu/~mlill/software/pharmdock/	Open source
DRUGON	Pharmacophore modeling and structure optimization	http://bioinfoteam.com/software.html	Free to academy
Molecular simulation			
DYNAMICS	MD simulations with Gromacs	https://github.com/tomaszmakarewicz/Dynamics	Open source
AMBER	MD simulations with Amber	http://people.pharmacy.purdue.edu/~mlill/software/pymol_plugins/install.shtml	Additional license needed for Amber
LOAD_TRAJ	Load MD trajectory file	http://www.pymolwiki.org/index.php/Load_Traj	Open source
GTKDYNAMO	QM or QM/MM calculation	https://sites.google.com/site/gtkdynamo/home	Open source, additional license needed for Amber
WATSITE	Water hydration energy and water network calculation	http://people.pharmacy.purdue.edu/~mlill/software/watsite/	Open source

MD, molecular dynamics; MM, molecular mechanical; QM, quantum mechanical.

they seek to other software for these purposes. Similarly, the PyTMs plugin enables users to easily introduce a set of common posttranslational modifications, including acetylation, carbamylation, phosphorylation, and so on, into protein models.[10]

In addition, the real-time energy minimization by Sculpting (Table 1) in PyMOL is another helpful but ignored function. It returns local atomic geometries (bonds, angles, chirality, and planarity) to the configuration the molecules possess. Atomic crashes and contacts can also be shown in real-time when the users adjust the atomic positions. However, Sculpting module has a weakness that force fields are not implemented. Fortunately, the Optimize[11] plug-in, which supports MMF94/GAFF/FF/Ghemical force fields, could be a complement of the Sculpting module. PyMOL becomes particularly useful if one combines the stereo mode, Builder toolbox, Sculpting, and the Optimize modules for the new compound design *in situ*.

Protein–Ligand Interaction Analysis

Protein–ligand interactions are essential to almost all biological processes occurring in living organisms.[12] Ligand-mediated signal transmission via molecular complementarity is essential to all life processes; these chemical interactions comprise biological recognition in atomic level.[13] Thus, the understanding of the mechanisms responsible for the protein–ligand recognition and binding are of great importance for the discovery and design of new drugs.[14] Besides the preset ligand site analysis module in PyMOL, the BNI-tools plugin adds additional functionalities and presets to the PyMOL GUI. These include creating H-bond interactions between protein and ligand in

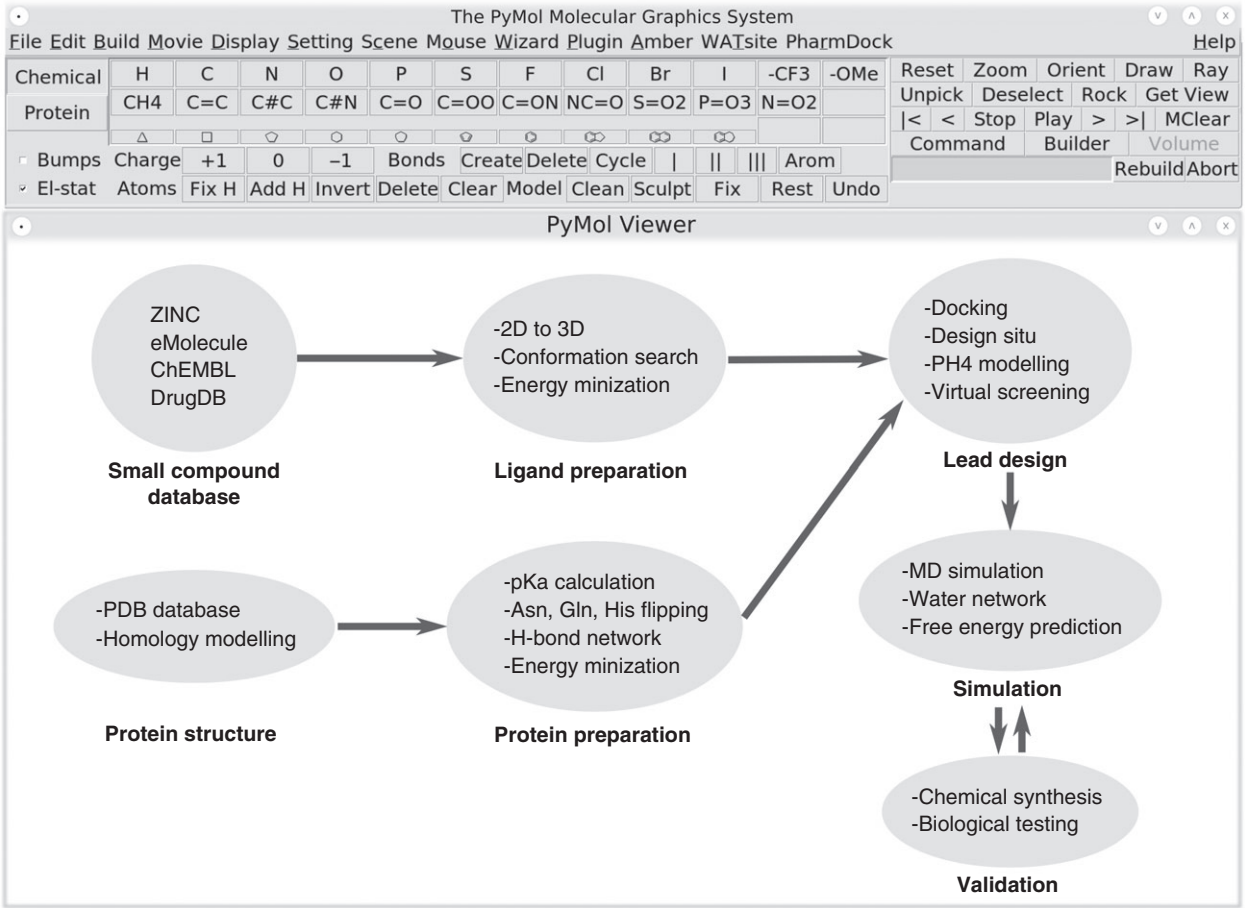

FIGURE 2 | The interface of PyMOL and the workflow of modern computational drug design. Upper panel: The PyMOL interface of main menu and that of the macromolecular builder tools. Lower panel: the workflow of modern computational drug design.

an easy way, creating a box around a macromolecule facilitating the rendering, coloring residues according to their hydrophilicities, creating a pseudo atom in the mass center of the selected atoms. More importantly, the PLIP plugin[15] generates automatically the protein–ligand interaction in a more sophisticated way: besides the H-bond interactions, other interactions including halogen bonds, hydrophobic interactions, π–π stacking, π-cation interactions, and water bridges can also be shown in PLIP(Figure 3(a)).

Protein Structure Analysis

There are many other plug-ins enhancing the structural analysis in PyMOL. As an example, the DSSP[17] and Stride[18] Plugin for PyMOL provides a graphic interface for coloring proteins according to their secondary structures assigned by DSSP or Stride. Moreover, tunnels and channels facilitate the transport of small molecules and solvent in various proteins. The Caver[19] and Mole[20] plug-ins in PyMOL can predict these structural features of a macromolecule. The characteristics of individual transport pathways, including their geometries and dynamic properties, can guide the design of novel drug candiates.[19] Additionally, the PyANM plugin allows its users to build and visualize anisotropic network models (ANM), a member of the elastic network models family which have been successful in reproducing fluctuations for proteins of native conformations[21] (Figure 3(b)). Finally, the Average3D command line plugin permits users to calculate the average structure of a given structure file, consisting of multiple 3D structural models of the same protein (e.g., NMR structures or MD simulation frames). The root-mean-square deviation of atomic positions could also be generated between the average structure and each individual frame.

PROTEIN–LIGAND MODELING

The major goal of protein–ligand modeling is to investigate the essential interactions between a ligand

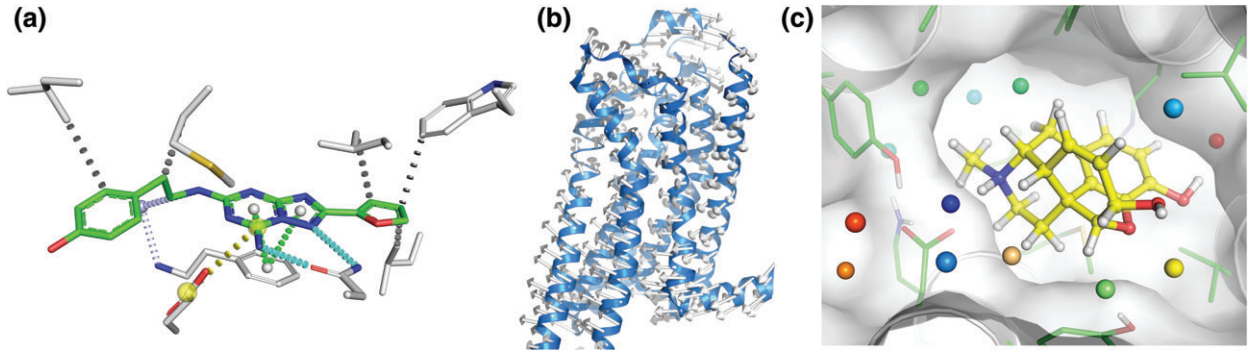

FIGURE 3 | Exemplified plugins in PyMOL. (a) Protein–ligand interaction of Adenosine 2A receptor ($A_{2A}R$, pdb: 4EIY) generated by the PLIP plugin. Gray stick: $A_{2A}R$ sidechain; green stick: ligand molecule; black dash: hydrophobic contact; cyan dash: H-bond interaction; yellow dash: ion lock interaction; blue dash: water bridge; and green dash: π–π stacking. (b) The anisotropic network models (ANM) movements of $A_{2A}R$ (pdb: 4EIY) generated by the PyANM plugin. Gray arrow: ANM movement directions. (c) The predicted hydration sites by the WATsite tool for the μ-opioid receptor (μOR).[16] The hydration sites are shown as small spheres and colored according to their desolvation energy ΔG values. Yellow stick: a morphine molecule in the binding pocket. Green stick: μOR sidechain. Red→Cyan→Green→Blue represents ΔG values range from favorable to unfavorable values. When designing new drug candidates, additional function groups can be introduced in these regions with unfavorable ΔG values, to increase the ligand binding affinity.

and a receptor.[22] It is an important step in modern drug discovery (Figure 2). Protein–ligand modeling includes protein homology modeling, protein preparation, ligand alignment, ligand preparation, and protein–ligand docking. All these functions are accessible through PyMOL plugins.

Homology Modeling

Homology modeling is to construct an atomic-resolution model of the targeted protein from its sequence and a template structure of the homologous protein.[23] The quality of a homology model relies on the quality of the template and the sequence identity between the targeted protein and the template.[24] Modeller[25] is a widely used tool for homology modeling. Moreover, Modeller can perform additional tasks including *de novo* modeling of protein loops, optimization of a flexibly region, multiple alignment of protein sequences and/or structures, clustering, searching of sequence databases, comparison of protein structures, and so on.[25] The PyMod[26] plugin in PyMOL combines Modeller and several other bioinformatics tools including PSI-BLAST,[27] Clustal Omega,[28] MUSCLE,[29] and PSIPRED.[30] Users can take advantage of PyMod to carry out the homology modeling process in three steps: template searching, target–template sequence alignment, and model building. Moreover, sequence similarity searches, multiple sequence–structure alignments, and evolutionary conservation analysis can also be performed in PyMod.[26]

In contrast, the PyRosetta[31] package is a Python-based command line collections of the Rosetta modeling suite.[32] It enables users to create custom molecular modeling algorithms with Rosetta sampling and scoring functions with Python scripting. The PyRosetta Toolkit is the PyMOL plugin of PyRosetta. PyRosetta Toolkit[33] creates and runs protocols in PyRosetta for molecular modeling in the backend, whereas it analyses the results in PyMOL interface in real-time.[34] Identical to PyRosetta, the PyRosetta Toolkit plugin permits users to perform *de novo*/comparative homology modeling, protein–ligand docking, protein–protein docking, energy minimization, and rational enzyme redesign.[33,34]

Protein Preparation

Crystal structures and homology models contain many defects, including the wrong conformation of Asn, Gln, and His, atom crashes, and the missing pKa information.[35] To obtain reliable and accurate computational results, protein preparation is a compulsory step before any other molecular modeling task. Thanks to the Amber plugin,[36] one can prepare the protein structure in PyMOL. This includes optimizing residue Asn, Gln and His conformation/protonation state, refining H-bond network and energy minimization. The residue protonation states can be prepared with PropKa[37] plugin in PyMOL.

Ligand Alignment

Given a number of molecules known to bind with a particular receptor, it is of great interest to determine their structural similarity. The LigAlign plugin[38] in PyMOL automates flexible ligand alignment and

analysis. When performing rigid alignments, LigAlign can produce results which are consistent with manually annotated structural motifs. In flexible alignments, LigAlign is capable of producing biochemically reasonable superposition of ligand and subsequently identifying conserved structural motifs that are rarely detected by rigid alignment.[38]

Ligand Preparation

In structure-based protein–ligand modeling, all ligands must be well prepared with correct 3D geometries, proper bond orders, accessible tautomer, and proper ionization states prior to any other modeling tasks.[39,40] The Optimize plugin provides a PyMOL graphic interface to some of the molecular mechanics features available in OpenBabel.[11] It allows users to generate 3D coordinates and minimize the energy of any molecule in PyMOL. Optimize is also capable of performing conformation searches for a specified molecule.

Protein–Ligand Docking

The goal of protein–ligand docking is to predict the position and orientation of a ligand in the binding pocket of a specific receptor. Protein–ligand docking has been widely used in both VS and lead optimization steps in modern drug discovery. The Autodock/Vina[41] plugin enables users to perform various tasks including defining binding sites and exporting to Autodock and VINA input files, performing receptor and ligand preparation automatically, executing docking with Autodock or VINA, visualizing grid maps generated by autogrid in PyMOL, handling multiple ligands, setting up VSs, and setting up docking with flexible sidechains.[41] In contrast, NRGsuite[42] is another protein–ligand docking PyMOL plugin which permits the detection of surface cavities in proteins, the structure refinements, the calculation of volume, and using featured binding-sites for docking with FlexAID.[43] Furthermore, the PyRosetta Toolkit[33] is also capable of protein–ligand docking by running PyRosetta in the backend.

VIRTUAL SCREENING

VS is a computational technique used in drug discovery to search libraries of small molecules in order to find new compounds which can bind to a specific protein.[44] VS can efficiently identify potential drug candidates from compound library with millions of candidates within a few weeks, and is much faster than the experimental high throughput screening. Currently, three VS tools, including Lisica,[45] Pharmdock,[46] and DrugOn,[47] have been developed as PyMOL plugins. LiSiCA is for 2D and 3D ligand-based VS,[45] whereas both Pharmdock and Drugon are pharmacophore models-based VS tools.[46,47]

MOLECULAR SIMULATION

MS study the physical movements of atoms and molecules. It mimics the physiological environment of macromolecules by constructing water or lipid atomic models. MS provides reliable information for drug discovery from various aspects, including the ligand-induced protein conformation changes,[48] the ligand binding energy estimation,[36] thermodynamic profile of individual water,[49] and enzyme catalytic mechanisms.[50] All these functions are now available in PyMOL plugins.

MD Simulation

The Dynamics[48] plugin allows researchers to perform MD simulations directly from the PyMOL software by GUI-based interface of Gromacs.[51] Dynamics supports both explicit solvent model and implicit solvent model simulations in the current new version.[48] Moreover, normal mode analysis and principle component analysis can also be performed in Dynamics. The contact map and cross-correlations can be visualized in PyMOL through Dynamics plugin. In contrast, the Amber plugin[36] in PyMOL uses Amber[52] as MD simulation engine in backend. Protein–ligand binding energy estimation can also be performed with Amber plugin by SIE method.[53] Finally, the internal command line 'load_traj' in PyMOL is capable of reading various types of trajectory files directly, including .xtc, .gro, .pdb, .dcd, .trj, and .dtr. This feature facilitates MD trajectory analysis and MD trajectory-based movie making in PyMOL.

Hydration Site Prediction

Water molecules mediating protein–ligand interactions or releasing from the binding pocket can contribute both enthalpically and entropically to the ligand binding energy.[49] Thus, the calculation of thermodynamic profile of individual water molecules and their potential contribution to ligand binding are of great importance for novel drug design. The WATsite plugin (Figure 3(c)) in PyMOL estimates the free energy profile of each hydration site by computing the enthalpy and entropy of the water molecules that occupy a hydration site, through the MD

simulation in Gromacs. Moreover, WATsite is also able to estimate the desolvation free energy for any specified ligand.

QM/MM Calculation

Hybrid quantum mechanical (QM)/molecular mechanical (MM) potentials are very powerful tools for MS. They are especially useful for studying processes in condensed phase systems, such as chemical reactions that involve electron transfer. GTKDynamo[50] links the PyMOL visualization platform and the pDynamo QM/MM simulation engine. GTKDynamo uses Amber[52] as a simulation engine in backend. The MD simulation system can be built with GTKDynamo plugin directly which evokes the AmberTools.[52] GTKDynamo supports various simulations including QM calculations of the reaction pathways and the organic reactions, QM/MM calculations of enzyme reactions, and the free energy calculations by umbrella sampling.[50]

CONCLUSION

PyMOL is a cross-platform molecular graphic tool and has been widely used for 3D visualization of macromolecules. The utilities of PyMOL have been extensively enhanced by various plugins, including macromolecular analysis, homology modeling, protein–ligand docking, pharmacophore modeling, VS, and MD simulations. These features facilitate PyMOL as a platform for modern computational drug design.

REFERENCES

1. Thakur PK, Hassan I. Discovering a potent small molecule inhibitor for gankyrin using de novo drug design approach. *Int J Comput Biol Drug Des* 2011, 4:373–386.
2. Danielson ML, Desai PV, Mohutsky MA, Wrighton SA, Lill MA. Potentially increasing the metabolic stability of drug candidates via computational site of metabolism prediction by CYP2C9: the utility of incorporating protein flexibility via an ensemble of structures. *Eur J Med Chem* 2011, 46:3953–3963.
3. Pereira D, Lima RT, Palmeira A, Seca H, Soares J, Gomes S, Raimundo L, Maciel C, Pinto M, Sousa E, et al. Design and synthesis of new inhibitors of p53–MDM2 interaction with a chalcone scaffold. *Arab J Chem*. In press.
4. Lee CJ, Liang X, Wu Q, Najeeb J, Zhao J, Gopalaswamy R, Titecat M, Sebbane F, Lemaitre N, Toone EJ, et al. Drug design from the cryptic inhibitor envelope. *Nat Commun* 2016, 7:10638.
5. Kingsley LJ, Wilson GL, Essex ME, Lill MA. Combining structure- and ligand-based approaches to improve site of metabolism prediction in CYP2C9 substrates. *Pharm Res* 2015, 32:986–1001.
6. Baig MH, Ahmad K, Roy S, Ashraf JM, Adil M, Siddiqui MH, Khan S, Kamal MA, Provaznik I, Choi I. Computer aided drug design: success and limitations. *Curr Pharm Des* 2016, 22:572–581.
7. Rigsby RE, Parker AB. Using the PyMOL application to reinforce visual understanding of protein structure. *Biochem Mol Biol Educ* 2016, 44:433–437.
8. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 2001, 98:10037–10041.
9. Hodis E, Schreiber G, Rother K, Sussman JL. eMovie: a storyboard-based tool for making molecular movies. *Trends Biochem Sci* 2007, 32:199–204.
10. Warnecke A, Sandalova T, Achour A, Harris RA. PyTMs: a useful PyMOL plugin for modeling common post-translational modifications. *BMC Bioinformatics* 2014, 15:370.
11. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: an open chemical toolbox. *J Cheminform* 2011, 3:33.
12. Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL, Liu SQ. Insights into protein-ligand interactions: mechanisms, models, and methods. *Int J Mol Sci* 2016, 17.
13. Dunn MF. Protein–ligand interactions: general description. In: Dunn MF, ed. *Encyclopedia of Life Sciences*. Chichester, UK: John Wiley & Sons; 2001.
14. Held M, Metzner P, Prinz JH, Noe F. Mechanisms of protein-ligand association and its modulation by protein mutations. *Biophys J* 2011, 100:701–710.
15. Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. PLIP: fully automated protein-ligand interaction profiler. *Nucleic Acids Res* 2015, 43: W443–W447.
16. Yuan S, Palczewski K, Peng Q, Kolinski M, Vogel H, Filipek S. The mechanism of ligand-induced activation or inhibition of mu- and kappa-opioid receptors. *Angew Chem Int Ed Engl* 2015, 54:7560–7563.
17. Joosten RP, te Beek TA, Krieger E, Hekkelman ML, Hooft RW, Schneider R, Sander C, Vriend G. A series

of PDB related databases for everyday needs. *Nucleic Acids Res* 2011, 39:D411–D419.

18. Frishman D, Argos P. Knowledge-based protein secondary structure assignment. *Proteins* 1995, 23:566–579.
19. Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* 2012, 8:e1002708.
20. Sehnal D, Svobodova Varekova R, Berka K, Pravda L, Navratilova V, Banas P, Ionescu CM, Otyepka M, Koca J. MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *J Cheminform* 2013, 5:39.
21. Sanejouand YH. Elastic network models: theoretical and empirical foundations. *Methods Mol Biol* 2013, 924:601–616.
22. Yuan S, Filipek S, Vogel H. A gating mechanism of the serotonin 5-HT3 receptor. *Structure* 2016, 24:816–825.
23. Yuan S, Peng Q, Palczewski K, Vogel H, Filipek S. Mechanistic studies on the stereoselectivity of the serotonin 5-HT1A receptor. *Angew Chem Int Ed Engl* 2016, 55:8661–8665.
24. Raval A, Piana S, Eastwood MP, Dror RO, Shaw DE. Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins* 2012, 80:2071–2079.
25. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci* 2007, 50:1–31.
26. Bramucci E, Paiardini A, Bossa F, Pascarella S. PyMod: sequence similarity searches, multiple sequence-structure alignments, and homology modeling within PyMOL. *BMC Bioinformatics* 2012, 13(suppl 4):S2.
27. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, 25:3389–3402.
28. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011, 7:539.
29. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004, 32:1792–1797.
30. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 1999, 292:195–202.
31. Chaudhury S, Lyskov S, Gray JJ. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* 2010, 26:689–691.
32. Simons KT, Bonneau R, Ruczinski I, Baker D. Ab initio protein structure prediction of CASP III targets using ROSETTA. *Proteins* 1999, 3(suppl 3):171–176.
33. Adolf-Bryfogle J, Dunbrack RL Jr. The PyRosetta toolkit: a graphical user interface for the Rosetta software suite. *PLoS One* 2013, 8:e66856.
34. Baugh EH, Lyskov S, Weitzner BD, Gray JJ. Real-time PyMOL visualization for Rosetta and PyRosetta. *PLoS One* 2011, 6:e21931.
35. Xiang Z. Advances in homology protein structure modeling. *Curr Protein Pept Sci* 2006, 7:217–227.
36. Lill MA, Danielson ML. Computer-aided drug design platform using PyMOL. *J Comput Aided Mol Des* 2011, 25:13–19.
37. Olsson MH, Sondergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *J Chem Theory Comput* 2011, 7:525–537.
38. Heifets A, Lilien RH. LigAlign: flexible ligand-based active site alignment and analysis. *J Mol Graph Model* 2010, 29:93–101.
39. Sastry GM, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 2013, 27:221–234.
40. Chan HC, Filipek S, Yuan S. The principles of ligand specificity on beta-2-adrenergic receptor. *Sci Rep* 2016, 6:34736.
41. Seeliger D, de Groot BL. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des* 2010, 24:417–422.
42. Gaudreault F, Morency LP, Najmanovich RJ. NRGsuite: a PyMOL plugin to perform docking simulations in real time using FlexAID. *Bioinformatics* 2015, 31:3856–3858.
43. Gaudreault F, Najmanovich RJ. FlexAID: revisiting docking on non-native-complex structures. *J Chem Inf Model* 2015, 55:1323–1336.
44. Rollinger JM, Stuppner H, Langer T. Virtual screening for the discovery of bioactive natural products. *Prog Drug Res* 2008, 65: 211, 213–249.
45. Lesnik S, Stular T, Brus B, Knez D, Gobec S, Janezic D, Konc J. LiSiCA: a software for ligand-based virtual screening and its application for the discovery of butyrylcholinesterase inhibitors. *J Chem Inf Model* 2015, 55:1521–1528.
46. Hu B, Lill MA. PharmDock: a pharmacophore-based docking program. *J Cheminform* 2014, 6:14.
47. Vlachakis D, Fakourelis P, Megalooikonomou V, Makris C, Kossida S. DrugOn: a fully integrated

pharmacophore modeling and structure optimization toolkit. *PeerJ* 2015, 3:e725.

48. Makarewicz T, Kazmierkiewicz R. Improvements in GROMACS plugin for PyMOL including implicit solvent simulations and displaying results of PCA analysis. *J Mol Model* 2016, 22:109.
49. Hu B, Lill MA. WATsite: hydration site prediction program with PyMOL interface. *J Comput Chem* 2014, 35:1255–1260.
50. Bachega JF, Timmers LF, Assirati L, Bachega LR, Field MJ, Wymore T. GTKDynamo: a PyMOL plug-in for QC/MM hybrid potential simulations. *J Comput Chem* 2013, 34:2190–2196.
51. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 2013, 29:845–854.
52. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *Wiley Interdiscip Rev: Comput Mol Sci* 2013, 3:198–210.
53. Lill MA, Thompson JJ. Solvent interaction energy calculations on molecular dynamics trajectories: increasing the efficiency using systematic frame selection. *J Chem Inf Model* 2011, 51:2680–2689.