W550–W556 *Nucleic Acids Research, 2016, Vol. 44, Web Server issue* *Published online 5 May 2016*
doi: 10.1093/nar/gkw393

ACFIS: a web server for fragment-based drug discovery

Ge-Fei Hao[1,†], Wen Jiang[1,†], Yuan-Nong Ye[2], Feng-Xu Wu[1], Xiao-Lei Zhu[1], Feng-Biao Guo[2,*] and Guang-Fu Yang[1,3,*]

[1]Key Laboratory of Pesticide & Chemical Biology, Ministry of Education, College of Chemistry, Central China Normal University, Wuhan 430079, P.R.China, [2]Center of Bioinformatics and Key Laboratory for NeuroInformation of the Ministry of Education, University of Electronic Science and Technology of China, Chengdu 610054, P.R. China and [3]Collaborative Innovation Center of Chemical Science and Engineering, Tianjing 300072, P.R.China

Received January 31, 2016; Revised April 20, 2016; Accepted April 28, 2016

ABSTRACT

In order to foster innovation and improve the effectiveness of drug discovery, there is a considerable interest in exploring unknown 'chemical space' to identify new bioactive compounds with novel and diverse scaffolds. Hence, fragment-based drug discovery (FBDD) was developed rapidly due to its advanced expansive search for 'chemical space', which can lead to a higher hit rate and ligand efficiency (LE). However, computational screening of fragments is always hampered by the promiscuous binding model. In this study, we developed a new web server Auto Core Fragment *in silico* Screening (ACFIS). It includes three computational modules, PARA_GEN, CORE_GEN and CAND_GEN. ACFIS can generate core fragment structure from the active molecule using fragment deconstruction analysis and perform *in silico* screening by growing fragments to the junction of core fragment structure. An integrated energy calculation rapidly identifies which fragments fit the binding site of a protein. We constructed a simple interface to enable users to view top-ranking molecules in 2D and the binding mode in 3D for further experimental exploration. This makes the ACFIS a highly valuable tool for drug discovery. The ACFIS web server is free and open to all users at http://chemyang.ccnu.edu.cn/ccb/server/ACFIS/.

INTRODUCTION

It is widely recognized that drug discovery is a time-consuming, expensive and complex process. The search for new drugs is mainly plagued by the rising cost and low success rate in Research and Development (R&D) (1). The goal of drug lead discovery is to identify candidate molecule with improved biological potency and physiochemical properties. However, the number of compounds needed to be synthesized to deliver one new lead is continuously rising in recent years (2). Both quality and quantity of hit classes available to medicinal chemists are primary drives for discovering best-in-class leads, which makes hit identification a crucial step to reduce attrition and therefore improve the overall R&D productivity (3). Thus, how to improve the efficiency and reduce the time required for successful hit identification will contribute greatly to drug discovery.

Over the past decade, high-throughput screening (HTS) of corporate compound decks has become the major paradigm for hit discovery in big pharma. Up to a few million compounds can be screened against the target of interest. But challenges facing traditional HTS technologies include low hit rates and hit molecule with low synthetic feasibility and nondrug-like properties are often identified. Compared with HTS, fragment-based drug design (FBDD) has some significant advantages (4). First, FBDD identifies ligand using ligand efficiency (LE) rather than potency alone to judge the relative order of molecules with different sizes, which offers more efficient and fruitful optimization campaigns. Second, FBDD only requires screening a small number of fragments to reach a more expansive chemical space, which leads to a much broader range of activity. Hence, the hit rate of FBDD is typically much higher than that observed with HTS (5). However, the specialized and expensive instruments, such as NMR spectroscopy, X-ray diffraction, surface plasmon resonance or mass spectrometry are always required to detect fragment binding due to the relative weak affinity (6,7). In addition, a large amount of purified proteins (>10 mg) and high concentrations of fragment molecules are always essential to achieve this task, which are very difficult to achieve in most cases (8,9).

*To whom correspondence should be addressed. Tel: +86 27 67867800; Fax: +86 27 67867141; Email: gfyang@mail.ccnu.edu.cn
Correspondence may also be addressed to Feng-Biao Guo. Email: fbguo@uestc.edu.cn
† These authors contributed equally to the work as first authors.

Hence, various computational methods have been developed to perform FBDD (10–12). Molecule docking has been used as a potentially attractive way to prioritize fragments from the much larger commercially available data set (11). Nevertheless, the problem of fitting small fragments targeting sub-pockets within the active site of the target protein is different from fitting a larger 'druglike' molecule. The binding mode of a fragment is relatively more difficult to be predicted by docking due to the smaller size. Meanwhile, scoring functions optimized according to drug-like molecules may be less accurate for predicting fragment affinities, which makes fragment docking problematic and triggers critical discussions (11). Furthermore, commercial softwares like LUDI (13), GLIDE (14), etc. are proved to be powerful enough to place fragments into the correct pocket of the active site. In addition, there is more and more free softwares dedicated to *de novo* drug design compared to other cheminformatic tools. For example, LigBuilder was developed for structure-based *de novo* drug design and optimization (15), and S4MPLE was developed as a conformational sampling tool for *in silico* FBDD (16). However, most of these tools are not web server facility, which makes them not easily practicable for non computational scientists. Only a minority of the web servers are dedicated to *in silico* drug discovery such as (17): MTiOpenScreen (http://bioserv.rpbs.univ-paris-diderot.fr/services/MTiOpenScreen/) (18), e-LEA3D (http://bioinfo.ipmc.cnrs.fr/lea.html) (19), iScreen (http://iScreen.cmu.edu.tw/) (20), VSDocker (http://www.bio.nnov.ru/projects/vsdocker2/) (21), however, they perform virtual screening of molecular libraries rather than fragments using docking algorithm. This prompts us to develop a web server to perform FBDD.

Herein, we present Auto Core Fragment *in silico* Screening (ACFIS), the first web server to perform computer-aided fragment-based drug discovery. It can generate core fragment structure based on the primary active molecule and perform automated fragment linking and virtual screening. The computation is based on the pharmacophore-linked fragment virtual screening (PFVS) method (22). Several libraries containing fragments extracted from known bioactive compounds are provided for screening. The core fragment generation tool uses fragment deconstruction analysis to evaluate the contribution of each fragment of the primary active molecule to the binding affinity. Another tool assists the user in the linking of new fragments to the selected core fragment and in the evaluating of the contribution to the binding potency. Thus, ACFIS enables researcher to perform computer-aided fragment-based drug discovery on traditional or more challenging protein targets.

MATERIALS AND METHODS

Fragment library

Fragment library design is crucial for the effectiveness of fragment-based virtual screening. The virtual library can be designed as a set of compounds consisting of building blocks for synthetic feasibility, a set of structures obtained by decomposition of larger molecules, or a set of commercially available molecules. The Rule of Three (Ro3) is acknowledged as the best definition of a fragment, i.e. a low molecular weight (MW) $\leq$ 300 Da, a reduced lipophily index clog$P \leq 3$, a number of hydrogen bond donors and acceptors $\leq$3, low flexibility represented by a number of rotatable bonds $\leq$3 and a polar surface area (PSA) $\leq$ 60 Å^2 (23,24).

The decomposition approach was used to design combinatorial fragment libraries by the deconstruction of known bioactive compounds (FDA approved drug and commercial pesticide molecules). Decomposition and Identification of Molecules (DAIM) program was used to perform decomposition of small molecules (25). The fragment size is controlled by the number of bond breakages. The breaking rules are based on retrosynthetic analysis. The fragment library derived from decomposition was filtered by Ro3. In addition to certain physicochemical property filters, the aspect of availability is also important and should be checked individually. However, the fragments from decomposition should already feature certain synthetic accessibility. Depending on the application, two fragment libraries (drug and pesticide fragment database) are supplied and will be continuously improved to expand its diversity.

Fragment deconstruction analysis

The FBDD approach can create novel lead compounds and allow the exploration of a larger chemical space. It either grows functional fragments from the starting fragment or pieces several fragments together into novel scaffold, which are generally defined as growing or linking algorithms. Growing algorithms start from a core fragment placed in the binding site and other fragments which are linked to improve activity. Linking algorithms, on the other hand, involve docking of functional fragments to each sub-pockets followed by linking adjacent fragments together. The ACFIS web server creates new molecules based on growing algorithm. First, the core fragment structure is determined. Then, automatic fragment linking is performed to improve the binding affinity. Hence, how to select a core fragment structure is a major challenge.

Studies of drugs entering the market show that known drugs, clinical candidates and bioactive compounds represent very attractive and valuable starting point for new lead compound discovery (26). Therefore, selecting core fragment structures from the known bioactive compounds seems to be an effective strategy. The fragment deconstruction analysis is performed with a three-step computational protocol shown in Supplementary Figure S1: A three-step minimization procedure was performed on the binding conformation of protein–ligand complex. (i) First, movement was allowed only for the ligand molecule with a harmonic constraints (100 kcal/mol•Å^2) applied to the complex. Second, the mainchain atoms of the protein were fixed and other atoms were allowed to move. Finally, all atoms were minimized with no restraint to a convergence of 0.01 kcal/(mol·Å). (ii) Ligand structure binding in the pocket is deconstructed into fragments according to the retrosynthetic analysis by using DAIM software. Single bond is broken and hydrogen is used to link with heavy atom to make the total charge value of each 'piece' integer. In addition, everything will be reparameterized including the partial charges according to new generated protein-fragment

structure. (iii) The binding free energy (ΔG) is calculated as previously described by the combination of the MM_PBSA method (27) for the enthalpy and an empirical method for the entropy (see details in the Supplementary Material) (28) for each protein–fragment structure. The computational procedure used to evaluate the binding free energy includes binding energy, solvation entropy and conformational entropy (Equation 1), which is the same as that described in our previous publication (29).

$$\Delta G_{\text{bind}} = \Delta E_{bind} - T\Delta S_{\text{solv}} - T\Delta S_{\text{conf}} = \Delta E_{bind} - T\Delta S_{\text{solv}} + w(\Delta N_{rot}) \quad (1)$$

The ranking of fragments is sorted according to LE defined as ΔG divided by the heavy atom count (HAC), LE $= -\Delta G_{\text{cal}}/\text{HAC}$. The core fragment is derived from the reassembly of the adjacent fragments based on the binding contribution. Hence, the obtained core fragment structure has a highly conserved binding conformation and efficient contribution to the entire binding affinity. Last, junction on the core fragment is determined according to the surrounding space (volume ≥ 50 Å^3).

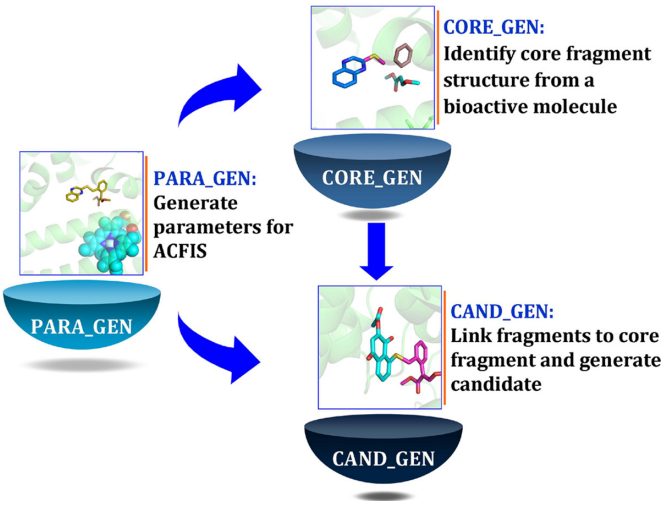

Figure 1. The interplay of the three ACFIS modules. PARA_GEN runs in a serial mode, whereas CORE_GEN and CAND_GEN run in parallel mode.

Auto core fragment *in silico* screening

New fragment is linked to the junction of core fragment structure placed in the binding site using a modified version of AutoGrow (30). The orientation of growing fragment was optimized with minimum steric clashes (overlap volume ≤ 4 Å^3) to the surrounding residues. Because the PFVS method reasonably relies on the resemblance assumption that the structural perturbations associated with fragment changing are relatively small, which would only lead to local changes of the protein structure and would not significantly change the backbone structure of protein. This procedure preserves the binding conformation of the core fragment. The energy minimization of each new ligand is achieved in four steps using the Sander module of Amber program (31). First, the fragment is minimized with the core fragment and the protein fixed. Then, the ligand is minimized with the protein fixed. Subsequently, the backbone atoms of the protein are fixed and other atoms are relaxed. The final minimization is performed with both the ligand and protein relaxed. In each step, the energy minimization is executed by using the steepest descent method for the first 2000 cycles and the conjugated gradient method for the subsequent 3000 cycles with a convergence criterion of 0.1 kcal mol^{-1} Å^{-1}. Finally, the MD simulation is performed using the generalized Born continuum solvent model to further relax the conformation of the growing fragment. The last snapshot of the MD simulation was minimized to a convergence criterion of 0.1 kcal mol^{-1}Å^{-1} and used for ΔG calculation. Finally, hit candidates were selected according to the ΔG value.

WEB SERVER

Web server configuration

ACFIS contains three computational modules. PARA_GEN is a tool to generate parameters for ACFIS, CORE_GEN is a tool to derive core fragment structure from a bioactive molecule, and CAND_GEN is a tool to link fragments to the core fragment structure and generate candidates. The inputs of all modules are easy to prepare and some items are optional to meet the purpose of individual projects. Output structure files and related data can be downloaded. Figure 1 shows how PARA_GEN, CORE_GEN and CAND_GEN work together to make online FBDD possible. In order to make this server more user friendly, these modules are connected in the 'primary mode'. After the submission of a complex pdb file, the server will generate several cores and select the top-ranked core fragment to perform fragment virtual screening automatically. To avoid the bugs of the uploaded pdb file, an initial file checking module is also developed in 'primary mode'. The three modules can also be independently used in the 'Advanced Mode'. ACFIS server runs on a dedicated Linux machine at the Supercomputer Cluster. The web application uses PHP (version 5.0), HTML and Java script to serve web pages. Related messages and results of each task are stored in a database implemented using MySQL (version 14.12). The web server runs on the apache HTTP server version 2.0.51, and the JSmol interactive molecular viewer applet (http://www.jmol.org/) is used for structure visualization. Chrome and Firefox are recommended explorers for our server. Screen with resolution higher than 1440 × 900 are needed for the web pages.

For PARA_GEN, 'your job' will start to run after you submit your job with a runtime of seconds to minutes based on the size of your molecule and the charge method. CORE_GEN will also manage your job immediately and finish in a few hours. CAND_GEN is a lengthy step depending on the size of the protein, Some tasks may take a few days. A job management system based on PHP and MYSQL has been developed to manage all the submitted jobs in this server.

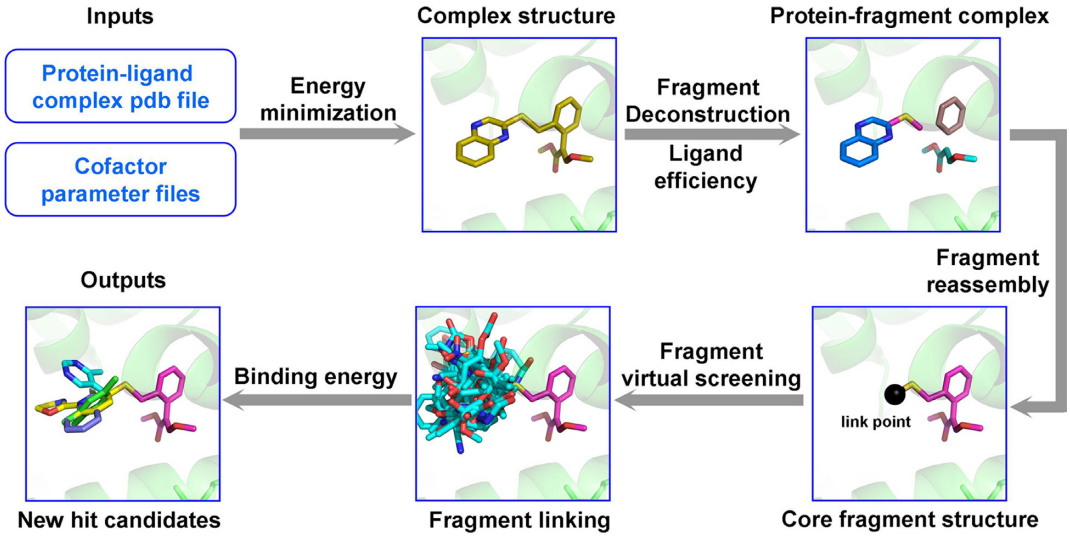

Figure 2. Workflow of ACFIS. The user input is shown in blue. The arrows denote the computational process.

PARA_GEN: a parameter generation tool

Input. In the front-page, the user is invited to upload a structure file of a protein cofactor in mol2 or pdb format. If there is no cofactor, this module can be skipped. A guideline at the bottom of the page explains how to use PARA_GEN to set exact parameters for a selected ligand. Hydrogen atoms should be correctly assigned, which can be done with most visualization software like AutoDock Tools, CHIMERA and PYMOL. Total formal charge should be correctly assigned, which is an essential parameter for the calculation of atomic charge. If the total formal charge is uncertain, it can be assigned automatically. Then a charge method from AM1-BCC, Muliken and Gasteiger options can be selected. And finally, an email address may be submitted to receive the final result files.

Server workflow. Once a valid structure file of cofactor is uploaded, PARA_GEN program starts to run (Figure 2). First, your uploaded pdb file will be converted into mol2 file by using openbabel2.3.1 (32). Then, the total formal charge will be calculated, if it is not already specified. Parameter files for AMBER force field are then prepared by AMBERTOOLS.

Output. Result of PARA_GEN contains all the parameter files of AMBER force field required by MD simulation and a log file recording the calculation process, which will be available after your job successfully finished. If the uploaded structure file is invalid, the job will be terminated and an 'error' will be reported on the web page. Moreover, it is also linked with CORE_GEN and CAND_GEN module in the result page to start the following computation directly.

CORE_GEN: a core fragment generation tool

Input. A protein–ligand complex in pdb format is required as an input of CORE_GEN. This can be obtained from RCSB protein data bank (http://www.rcsb.org/pdb/home/home.do) or acquired from docking calculation. Additionally, the name of ligand in the complex structure file should be referred. Parameters for the cofactor will be required, if it is in the complex structure. At last, a password should be assigned to make your job confidential and an email notice is optional.

Server workflow. After the initial data validation, CORE_GEN started to run with a three-step computational protocol (Figure 2): (i) The minimization procedure was performed for the complex structure. (ii) Ligand structure was deconstructed into fragments by using DAIM program. (iii) The ΔG was calculated for each protein-fragment complexes. The ranking of fragments is sorted according to LE. Comparison based on LE rather than potency alone could be useful in deciding the potential of fragments.

Output. The users are guided to the result page after they submit the job successfully. Messages about all submitted jobs are printed in this page. Once the job is completed, results of CORE_GEN are presented to the user via a web page by a click on the ID of the job (the user is invited to insert a password if it was assigned before). The summary table shows the information related to each generated core fragment structure (results shown in Supplementary Figure S2 in the Supplementary Material). Due to the importance of physicochemical properties for fragment selection, a link to mol-inspiration prediction server (http://www.molinspiration.com/) is created. The 3D structure of each protein-core fragment complex is shown with JSmol by a click on the structure, which can also be downloaded by clicking on the link. The junction of the core fragment structure is marked with a hydrogen atom shown as a ball in the 3D view mode. In addition, CORE_GEN is also linked with CAND_GEN module in the result page to start the following computation directly.

CAND_GEN: a hit generation tool

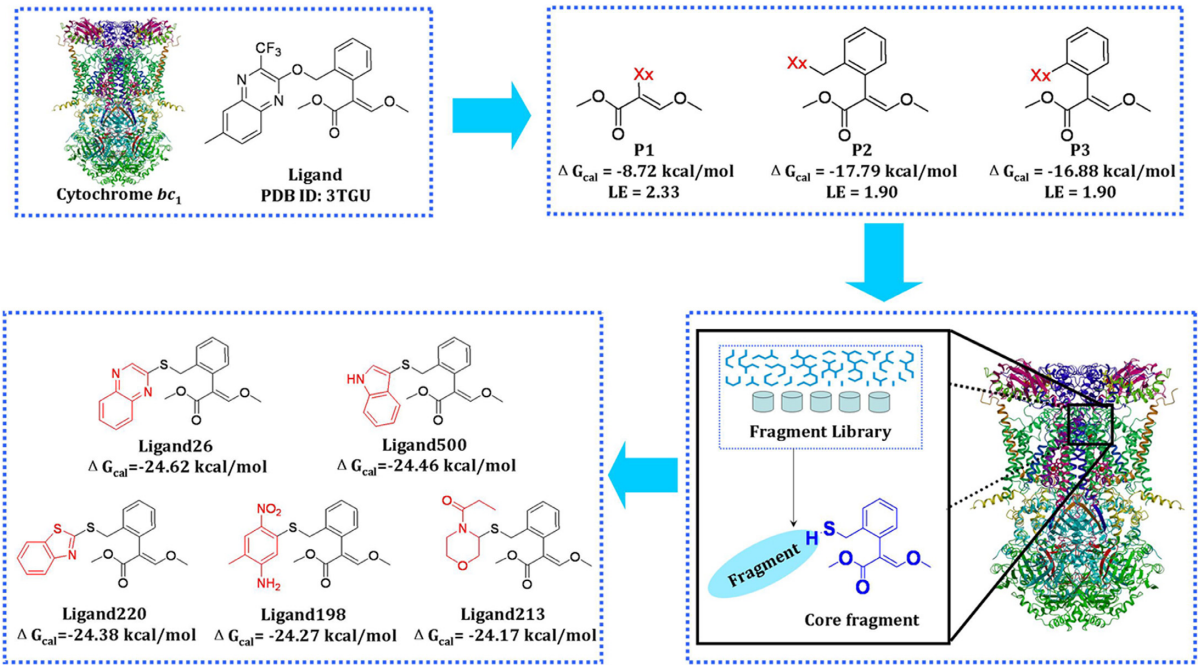

Figure 3. A case study of cytochrome bc_1 complex inhibitors design with ACFIS. Input is the inhibitor WF3 presented in the crystallized structure of cytochrome bc_1 complex (PDB code: 3TGU). The composite computation includes parameter generation, fragment deconstruction, core fragment identification, fragment linking, structure optimization, molecular dynamic simulation and binding free energy calculation, which can be processed automatically by this web server. Outputs are the top five hit molecules.

Input. The fragments in the library will be linked to the junction of the core fragment structure by CAND_GEN module. Parameters for cofactor will be required, if it is in the complex structure. The time of MD simulation for each ligand is selected between a minimum of 5 ps and maximum of 20 ps. The number of reported hits should be selected from the list. In addition, the binding free energy calculation method MM_PBSA or MM_GBSA is chosen to sort the order of the hit candidates. At last, a password should be assigned to make the job confidential and submission of an email address is arbitrary.

Server workflow. The workflow of CAND_GEN is in three steps (Figure 2): (i) A minimization procedure is performed. (ii) A short time MD simulation is performed on each minimized structure. (iii) The last snapshot from MD simulation is minimized to perform ΔG calculation.

Output. The running status of all jobs will be shown in the job management system (results shown in Supplementary Figure S2 in the Supplementary Material). Once the job is completed, a click on the ID of the job will lead the user to a web page that contains a summary table showing the binding free energy of hit candidates. Physicochemical properties reported by mol-inspiration prediction server will also guide the process of hit candidates selection. The 3D structure of the final protein–ligand complex can also be viewed through JSmol. They can also be downloaded for further analysis.

Performance of ACFIS

We assessed the performance of ACFIS on several classes of important therapeutic protein targets and 78 crystal structures of target–ligand complexes were taken from the PDBbind database (http://www.pdbbind-cn.org/) (33). This data set has been widely used previously for benchmark. In addition several classes of important pesticide protein targets have also been used for accuracy validation, which include 26 protein–ligand complexes. The total validation data set is 104 including 49 positive and 55 negative samples. The composition of the validation data set are listed in Supplementary Table S1. The ability of ACFIS to discriminate high binding affinity compounds (positive sample) from low binding affinity compounds (negative sample) in the validation data set has been assessed. Interestingly, ACFIS identified the right binding for 41 out of the 49 positive samples assessed. For 42 out of the 55 negative samples assessed, they are not among the top generated hits. In addition, ACFIS also showed very good performance on binding pose prediction (see Supplementary Table S1): the correct binding poses (RMSD < 1.5Å compared with the crystal structure) were obtained for 77 out of the 84 crystal available ligands (91.7%). The detailed results of validation are summarized in Supplementary Tables S2–S4 and Supplementary Figures S3–S5. As shown in Supplementary Table S4, the sensitivity, specificity and precision of ACFIS is 83.7%, 76.4% and 75.9%. It also shows in Supplementary Figure S5 that the AUCs (Area under operation character curve) is 82.2%. Hence, it is concluded, our server could accurately predict fragment binding. We retain a high sensitivity (83.7%) but a lower specificity (76.4%) because the number of non-binding fragment is large that even a low error rate

of negative samples may cause many false positive predictions.

Taking cytochrome bc_1 complex inhibitor as an example, the parameters of cofactor heme were generated by PARA_GEN, the crystal complex structure of WF3-bound bc_1 (PDB code: 3TGU) was downloaded to perform CORE_GEN calculation. As shown in Figure 3, the top three core fragments were generated (P1, P2 and P3) with ΔG_{cal} of −8.72, −17.79 and −16.88 kcal/mol. Obviously, the calculations systematically overestimated the absolute values of the binding affinities. However, the calculated ΔG_{bind} values can still qualitatively reflect the relative order of the binding affinities of these core fragments. The error induced by the molecular size can be eliminated by using LE rather than potency. As shown in Figure 3, the LE of P1 (LE = 2.33) is relatively higher than the other two fragments P2 (LE = 1.90) and P3 (LE = 1.90). However, comprehensively considered with ΔG_{bind}, LE and new fragment growing spaces, P2 was selected as a core fragment to perform CAND_GEN calculation. In addition, a sulfur atom was set as the linker between the core fragment and the new fragment for synthetic feasibility (Figure 3). Finally, the top five new hit candidates selected according to the binding free energy sorting were shown in Figure 3 with the original ligand scaffold (Ligand26) included. It needs to be emphasized that the synthesis of the candidate (Ligand220), the inhibitory kinetics with cytochrome bc_1 complex (K_i = 31.10 nM), and the co-crystal structure study have been completed (34), which further validated the predicted results of ACFIS.

CONCLUSIONS

The discovery of new hit molecules remains a crucial step for drug and agrochemical discovery. We presented ACFIS, a computer-aided FBDD web server, to provide solutions for fragment virtual screening and to assist scientists in the new hits identification. It is differentiated from other web services in which ligand screening was presented in docking based manner. In the future, we will perform a continual update of the fragment library of ACFIS to offer broader and more diverse selection of databases to start a fragment-based drug discovery task.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Prof. Amedeo Caflisch for the permission to use DAIM program on our server and the developers of open source softwares (Jsmol, OpenBabel, *et. al.*). This research was supported by the Special Fund for Agro-scientific Research in the Public Interest (No. 201203022)and the National Natural science Foundation of China (No. 21332004 and 31470068).

FUNDING

Special Fund for Agro-scientific Research in the Public Interest [No. 201203022]; National Natural science Foundation of China [21332004 and 31470068]. Funding for open access charge: Special Fund for Agro-scientific Research in the Public Interest [No. 201203022]; National Natural science Foundation of China [No. 21332004 and 31470068].
Conflict of interest statement. None declared.

REFERENCES

1. Dickson,M. and Gagnon,J.P. (2004) Key factors in the rising cost of new drug discovery and development. *Nat. Rev. Drug Discov.*, **3**, 417–429.
2. Lamberth,C., Jeanmart,S., Luksch,T. and Plant,A. (2013) Current challenges and trends in the discovery of agrochemicals. *Science*, **341**, 742–746.
3. Frearson,J.A. and Collie,I.T. (2009) HTS and hit finding in academia - from chemical genomics to drug discovery. *Drug Discov. Today*, **14**, 1150–1158.
4. Murray,C.W. and Rees,D.C. (2009) The rise of fragment-based drug discovery. *Nat. Chem.*, **1**, 187–192.
5. Hann,M.M., Leach,A.R. and Harper,G. (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.*, **41**, 856–864.
6. Erlanson,D.A., McDowell,R.S. and O'Brien,T. (2004) Fragment-based drug discovery. *J. Med. Chem.*, **47**, 3463–3482.
7. Rees,D.C., Congreve,M., Murray,C.W. and Carr,R. (2004) Fragment-based lead discovery. *Nat. Rev. Drug Discov.*, **3**, 660–672.
8. Fruh,V., Zhou,Y., Chen,D., Loch,C., Ab,E., Grinkova,Y.N., Verheij,H., Sligar,S.G., Bushweller,J.H. and Siegal,G. (2010) Application of fragment-based drug discovery to membrane proteins: identification of ligands of the integral membrane enzyme DsbB. *Chem. Biol.*, **17**, 881–891.
9. Congreve,M., Rich,R.L., Myszka,D.G., Figaroa,F., Siegal,G. and Marshall,F.H. (2011) Fragment screening of stabilized G-protein-coupled receptors using biophysical methods. *Methods Enzymol.*, **493**, 115–136.
10. Villar,H.O. and Hansen,M.R. (2007) Computational techniques in fragment based drug discovery. *Curr. Top. Med. Chem.*, **7**, 1509–1513.
11. Chen,Y. and Shoichet,B.K. (2009) Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.*, **5**, 358–364.
12. Loving,K., Alberts,I. and Sherman,W. (2010) Computational approaches for fragment-based and de novo design. *Curr. Top. Med. Chem.*, **10**, 14–32.
13. Bohm,H.J. (1992) The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *J. Comput. Aided. Mol. Des.*, **6**, 61–78.
14. Sandor,M., Kiss,R. and Keseru,G.M. (2010) Virtual fragment docking by Glide: a validation study on 190 protein-fragment complexes. *J. Chem. Inf. Model.*, **50**, 1165–1172.
15. Yuan,Y.X., Pei,J.F. and Lai,L.H. (2011) LigBuilder 2: a practical de novo drug design approach. *J. Chem. Inf. Model.*, **51**, 1083–1091.
16. Hoffer,L., Renaud,J.-P. and Horvath,D. (2013) In silico fragment-based drug discovery: setup and validation of a fragment-to-lead computational protocol using S4MPLE. *J. Chem. Inf. Model.*, **53**, 836–851.
17. Villoutreix,B.O., Lagorce,D., Labbe,C.M., Sperandio,O. and Miteva,M.A. (2013) One hundred thousand mouse clicks down the road: selected online resources supporting drug discovery collected over a decade. *Drug Discov. Today*, **18**, 1081–1089.
18. Labbe,C.M., Rey,J., Lagorce,D., Vavrusa,M., Becot,J., Sperandio,O., Villoutreix,B.O., Tuffery,P. and Miteva,M.A. (2015) MTiOpenScreen: a web server for structure-based virtual screening. *Nucleic Acids Res.*, **43**, W448–W454.
19. Douguet,D. (2010) e-LEA3D: a computational-aided drug design web server. *Nucleic Acids Res.*, **38**, W615–W621.
20. Tsai,T.-Y., Chang,K.-W. and Chen,C.Y.-C. (2011) iScreen: world's first cloud-computing web server for virtual screening and de novo drug design based on TCM database@Taiwan. *J. Comput. Aided. Mol. Des.*, **25**, 525–531.
21. Prakhov,N.D., Chernorudskiy,A.L. and Gainullin,M.R. (2010) VSDocker: a tool for parallel high-throughput virtual screening using AutoDock on Windows-based computer clusters. *Bioinformatics*, **26**, 1374–1375.

22. Hao,G.F., Wang,F., Li,H., Zhu,X.L., Yang,W.C., Huang,L.S., Wu,J.W., Berry,E.A. and Yang,G.F. (2012) Computational discovery of picomolar Q(o) site inhibitors of cytochrome bc(1) complex. *J. Am. Chem. Soc.*, **134**, 11168–11176.
23. Congreve,M., Carr,R., Murray,C. and Jhoti,H. (2003) A 'rule of Three' for fragment-based lead discovery? *Drug Discov. Today*, **8**, 876–877.
24. Jhoti,H., Williams,G., Rees,D.C. and Murray,C.W. (2013) The 'rule of three' for fragment-based drug discovery: where are we now? *Nat. Rev. Drug Discov.*, **12**, 644–644.
25. Kolb,P. and Caflisch,A. (2006) Automatic and efficient decomposition of two-dimensional structures of small molecules for fragment-based high-throughput docking. *J. Med. Chem.*, **49**, 7384–7392.
26. Teague,S.J. (2011) Learning lessons from drugs that have recently entered the market. *Drug Discov. Today*, **16**, 398–411.
27. Hou,T.J., Wang,J.M., Li,Y.Y. and Wang,W. (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.*, **51**, 69–82.
28. Pan,Y., Gao,D. and Zhan,C.G. (2008) Modeling the catalysis of anti-cocaine catalytic antibody: competing reaction pathways and free energy barriers. *J. Am. Chem. Soc.*, **130**, 5140–5149.
29. Hao,G.F., Zhu,X.L., Ji,F.Q., Zhang,L., Yang,G.F. and Zhan,C.G. (2009) Understanding the Mechanism of Drug Resistance Due to a Codon Deletion in Protoporphyrinogen Oxidase through Computational Modeling. *J. Phys. Chem. B*, **113**, 4865–4875.
30. Durrant,J.D., Amaro,R.E. and McCammon,J.A. (2009) AutoGrow: a novel algorithm for protein inhibitor design. *Chem. Biol. Drug Des.*, **73**, 168–178.
31. Case,D.A., Berryman,J.T., Betz,R.M., Cerutti,D.S., Cheatham,T.E.I., Darden,T.A., Duke,R.E., Giese,T.J., Gohlke,H., Goetz,A.W. *et al.* (2014) *AMBER14*. University of California, San Francisco.
32. O'Boyle,N.M., Banck,M., James,C.A., Morley,C., Vandermeersch,T. and Hutchison,G.R. (2011) Open Babel: an open chemical toolbox. *J. Cheminform.*, **3**, 33.
33. Wang,R.X., Fang,X.L., Lu,Y.P. and Wang,S.M. (2004) The PDBbind database: collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.*, **47**, 2977–2980.
34. Hao,G.F., Yang,S.G., Huang,W., Wang,L., Shen,Y.Q., Tu,W.L., Li,H., Huang,L.S., Wu,J.W., Berry,E.A. *et al.* (2015) Rational design of highly potent and slow-binding cytochrome bc1 Inhibitor as fungicide by computational substitution optimization. *Sci. Rep.*, **5**, 13471.