wwLigCSRre: a 3D ligand-based server for hit identification and optimization

O. Sperandio[1,2], M. Petitjean[3] and P. Tuffery[1,4,*]

[1]MTi, INSERM UMR-S973, [2]CDithem, Université Paris Diderot - Paris 7, F75013, Paris, [3]CEA/DSV/iBiTec-S/ SB2SM (CNRS URA 2096), Gif-sur-Yvette and [4]RPBS, Université Paris Diderot-Paris 7, F75013, Paris, France

ABSTRACT

The wwLigCSRre web server performs ligand-based screening using a 3D molecular similarity engine. Its aim is to provide an online versatile facility to assist the exploration of the chemical similarity of families of compounds, or to propose some scaffold hopping from a query compound. The service allows the user to screen several chemically diversified focused banks, such as Kinase-, CNS-, GPCR-, Ion-channel-, Antibacterial-, Anticancer- and Analgesic-focused libraries. The server also provides the possibility to screen the DrugBank and DSSTOX/ Carcinogenic compounds databases. User banks can also been downloaded. The 3D similarity search combines both geometrical (3D) and physi-cochemical information. Starting from one 3D ligand molecule as query, the screening of such databases can lead to unraveled compound scaffold as hits or help to optimize previously identified hit molecules in a SAR (Structure activity relationship) project. wwLigCSRre can be accessed at http://bioserv. rpbs.univ-paris-diderot.fr/wwLigCSRre.html.

INTRODUCTION

Ligand-based screening methods are now being conventionally used in the early stages of a variety of drug discovery projects to mine chemical databases with the aim of identifying new hit compounds or optimizing leads. They are based on the assumption that similar compounds are likely to exhibit similar biological activities on a given target. Depending on the data available about active compounds, different strategies have been proposed. QSAR techniques attempt to build on top of a significant collection of compounds of known activity, a statistical model that will be used to predict the activity of new compounds (1). Similarity search techniques explicitly address the question of the similarity between compounds. Similarity can be addressed at different levels. Pharmacophore techniques (2) attempt to derive key features that are shared by active compounds for further use in the mining of collections. 2D similarity techniques (3,4) use the topological description of compounds to search for shared chemical patterns. 3D similarity search, such as ROCS (5), MedSuMoLig (6), shaEP (7) or Superimposé (8) explicitlysearch for common shapes in the compound conformations with in some cases a consideration for chemistry. While 2D similarity tools, such as LigandInfo (3) or ChemMine (4) represent efficient tools to screen large chemical databases, several studies suggest that the legitimate wish for scaffold hopping is more expected to be granted using 3D molecular similarity screening tools (8,9).

An important aspect of *in silico* approaches to drug design is the amount of potential compounds to deal with. A growing number of chemical suppliers provide free access to their numeric catalog (10) (e.g. Aurora Fine Chemical http://www.aurorafinechemicals.com/, Asinex http://www.asinex.com/, Chembridge http://www.chembridge.com, MayBridge http://www.maybridge.com/). Recent initiatives have even provided to the community a numeric collection of the main commercial catalogs (11). Above all, the complexity and size of the chemical space, estimated to more than 10^{60} molecules (12), makes it impossible to tackle with exhaustive search tools. More and more studies (13–15) suggest that an efficient way to identify new chemical entities in drug discovery projects is to design smaller and smarter chemical libraries that should be chemically diversified and representative of protein families.

In terms of tools available online, various web servers such as sMOL (16), eMolecules www.emolecules.com, QueryChem (17) already exist to mine chemical databases. Yet, most of them use only 1D (SMILE) or 2D (SDF) query-based searches including the chemical compound providers themselves directly from their website e.g. Aurora Fine Chemicals, MayBridge. FTrees (18–20) can screen databases with more complex approach than just 2D similarity. FTrees is able to condense molecular descriptions into a graph object and to search for actives in large databases using graph similarity. Finally, several web servers use a 3D description of chemical compounds.

*To whom correspondence should be addressed. Tel: +331 57 27 83 74; Fax: +331 57 27 83 72; Email: pierre.tuffery@univ-paris-diderot.fr

ValLigURL (21) can be used by crystallographers to validate ligand structure or to locate all the instances of one ligand within the structural databases (PDB). Pharmagist (22) can perform pharmacophore construction starting with several known drug-like binders of a given protein target. The well-established PubChem server (23) has very recently offered the possibility to screen very rapidly the whole PubChem compound collections starting with a 3D compound query. It does not however allow to upload user compound collections and does not provide either a superimposition of the hit molecules onto the query compound. Finally, a dedicated 3D molecular similarity screening service is Superimposé (8) which proposes two different superimposition algorithms that can be used to screen three different databanks, Superdrug, LigandDepot and the NCI collection. Superdrug (24) represents a collection of active ingredients of essential marketed drugs, LigandDepot (25) contains ligands that are present within the PDB and NCI represents the 04 September 2007 version of the National Cancer Institute database. However, Superimposé execution times remain consequent, on the order of 60 h in some cases for large banks.

wwLigCSRre online facility is built on top of a powerful 3D molecular similarity search engine (LigCSRre) to screen focused chemical libraries. LigCSRre relies on the CSR algorithm (26) that searches for the maximal common substructure between two sets of unordered coordinates. In LigCSRre, the search is in addition driven by the information about the atomic nature, bond and connectivity information, to define which pairings are possible based on some physicochemical properties. In order to keep execution times reasonably low (not exceeding the order of 1 h), wwLigCSRre strategy is to screen focused libraries of small size but preserving chemical diversity. wwLigCSRre presently allows the screening of 12 thematically focused libraries, including family-based focused chemical libraries (GPCR, CNS, Ion Channel, Kinase, Anticancer, Analgesic, and Antibacterial), a chemically diverse subset of the diversity set of the Chembridge database, three sets from the Drugbank, (Small molecule, Approved, and Withdrawn) and finally the CPDB Summary Tables (Carcinogenic Potency Database) subset of the DSSTOX database. This design makes wwLigCSRre particularly well suited for fast scaffold hopping, lead identification and optimization.

CONCEPTS AND METHODS

Maximal substructure search

The similarity search engine, LigCSRre is an evolution of the CSR algorithm originally developed by Petitjean (26) that searches for the maximal 3D motif—or maximal substructure (MSS)—common to two sets of coordinates. The CSR algorithm searches for the largest set of atom pairings between two clouds of atom coordinates—no *a priori* pairings or *a priori* rules such as the knowledge of the neighbors are required—using an iterative and stochastic procedure. Each iteration starts from a random initial superimposition, and iteratively maximizes the number of pairings. Pairings are based on distance sort of the $N1^*N2$ inter-atomic distances between the N1 atoms of the molecule 1 and the N2 atoms of the molecule 2. The array of the $N1^*N2$ distances is sorted by increasing values. The first atom pair, corresponding to the smallest distance, is always included in the common motif. Next pairs are included until a member of a pair already included in the common motif occurs. The latter pair is not included in the motif and the pairing terminates. Then the complete sets of coordinates are best superimposed from the current pairings and the whole process distance sorting/best fit is iterated until no new pair is accepted. This whole process is performed for a series of random starting points and CSR returns the largest motif identified.

In LigCSRre, we have implemented several additional particular features. First, LigCSRre extends the set of pairings at search convergence, in order to overcome a limitation of the CSR algorithm that makes possible that CSR stops to enlarge its MSS for an atom already paired, but hiding subsequent pairs. Second, it considers the fact that atomic properties of the pairs must be compatible. For this, LigCSRre embeds a regular expression formalism that allows, for each atom, to define which pairings are possible based on some physicochemical properties (see next paragraph). This results in smaller search space and increases search efficiency—LigCSRre usually requires much less iterations than CSR.

To take into account the fact that authorized or forbidden pairings must be defined in a way flexible enough to be adapted to a particular chemical context, we use the mol2 atom types, as assigned by open-babel (27). We use a three level mechanism of regular expressions to define atomic types compatible for pairing. The first one is the default level: default regular expressions authorize atoms to be paired only with atom having the same exact atomic type. The second level is the generic level: equivalence classes are defined. It is for instance possible to assert that a carbon atom could be paired with any carbon, but not oxygen or sulfur. A third level is the specific level: regular expressions specific of a particular atom can be defined. For instance, it would be possible to impose, for a specific carbon of known importance for chemical activity, that it could match only an aromatic carbon or a Nitrogen. The precedence order gives the higher priority to the specific level, then to the generic level, then to the default level. In wwLigCSRre, we use a minimal set of rules for the generic level. More in detail, the equivalence classes correspond (i) to carbons but carbo-cations; (ii) sp2 Oxygen (0.2 and 0.co2 mol2 types); (iii) sulfoxide and sulfone Sulfur (S.o and S.o2); (iv) sp2 and sp3 Sulfur; and (v) Nitrogen. The exact detail of these equivalences is available on the wwLigCSRre help page. For other atomic types, the default rules only accept pairings for atoms of identical atomic types. This set of rules has shown efficient to search large collections of compounds and retrieving compounds of known similar activity (Quintus,F. *et al.*, submitted for publication).

Bank screening

The LigCSRre algorithm is iteratively applied to each compound of the bank. For each, it stores the number of bonds (nB) identified as shared, where a bond is denoted as shared if both two atoms at bond extremities are paired on exit of LigCSRre. It also stores the size of the pairing set (nP), and the RMS deviation (RMSd) associated with these pairs. Once all the compounds have been screened, they are sorted using a cascading procedure: according to nB, then to RMSd. We also calculate a z-score to roughly assess the significance of the match. It is based on the number of bonds shared. $Z = (nB_{obs} - nB_{exp})/\sqrt{(nB_{exp})}$, where nB_{obs} is the observed number of bonds paired, nB_{exp} is the expected number of bonds paired by chance, that is calculated as $nB_{comp}{}^{*}\mu$, where nB_{comp} is the actual number of bonds of the query compound, and μ the average number of shared bonds over unrelated compounds. Figure 1 illustrates the z-scores distributions for unrelated compounds and 47 compounds known to share activities.

3D conformational variability

The performances of ligand screening based on 3D conformation depend on the number of conformations associated with each compound. However, one must consider the balance between conformational diversity, computational cost and search efficiency. Based on a previous study (Quintus,F. *et al.*, submitted for publication), we use a maximum of 50 conformations per compound here.

Bank preparation

In order to limit the size of the banks to a tractable number in the context on an online service, we generated subsets of the collections when required. When chemical diversity was required, like for the family-based focused libraries, a combination of the tools Cactvs (28) (http://www.xemistry.com) and Subset (29) was used. The fingerprints calculated within Cactvs were submitted to a diversity criterion such as having no pair of ligands with a Tanimoto coefficient above 0.85/0.8/0.75 depending on the databank using the tool Subset. The slight Tanimoto coefficient variation was adjusted so to select a number of compounds on the same magnitude per bank. From the Aurora Fine Chemicals electronic catalog we have integrated the following focused chemical libraries: Analgesic, Antibacterial, Anticancer, Ion channel and Kinase. From the Chembridge databases we have extracted the CNS, GPCR and diversity subset to create chemically diversified subsets. No chemical diversity criterion has been applied to construct the three subsets of the DrugBank, but they were filtered for ADME properties using the tool FAFDrugs2 (30) so to remove nondrug-like compounds like injectable or non orally bioavailable drugs. Finally, the CPDB subset from the DSSTOX database was taken as it is. All the subsets were generated in 3D with Frog (31) and a maximum of 50 best scored conformations per molecule was generated using the tool MS-Dock (32). Table 1 gives an overview of the different databases available on the wwLigCSRre server.

INPUT/OUTPUT

The web interface proposed by wwLigCSRre is very simple. Basically, the user will input a query, select a bank to mine or upload it and run the search for similar compounds. Despite for computational cost reasons, the size of the collections available online in wwLigCSRre is limited, depending on the nature of the query, the total
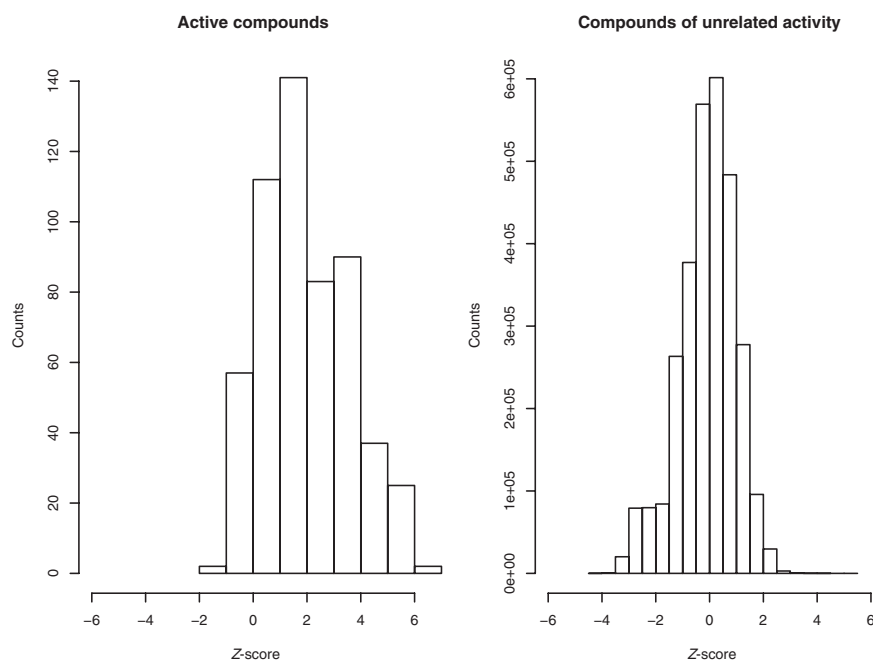
Figure 1. *Z*-score distributions for active and unrelated compounds. For active compounds, the data plotted corresponds to 47 compounds actives on cyclin-dependent kinase 2 (CDK2), neuraminidase, RNAse, coagulation factor Xa and thymidine kinase. Unrelated compounds correspond to the ChemBridge diversity set (http://www.chembridge.com/), ADME/tox filtered (38 000 compounds).

Table 1. Banks presently implemented in wwLigCSRre

Bank	Number of compounds	Tanimoto	ADME-Filter
AFC			
Analgesic	1587	0.8	No
Antibacterial	2069	None	No
Anticancer	2048	0.8	No
Ion channel	775	0.8	No
Kinase	2283	0.8	No
CB			
GPCR	940	0.85	No
CNS	2363	0.75	No
Diversity	2880	0.8	Yes
DB			
Small molecule	942	None	Yes
Approved	409	None	Yes
Withdrawn	24	None	Yes
DSSTOX			
CPDB	1547	None	None

[a]Provider: Aurora Fine Chemicals (AFC), ChemBridge (CB), DrugBank (DB), DSSTOX. For each bank, we report the number of compounds selected, the Tanimoto Diversity criterion threshold (Tanimoto). Some banks were filtered for ADME-Tox.

number of conformations of the bank selected, and server load, execution times may vary from several minutes up to hours. So it is possible (but not mandatory) to specify an email to get an alert on request termination.

The query consists in a single compound 3D conformation that must be in the mol2 format. Since the search is sensitive to atom types we prefer not to perform any automated format interconversion that could result in unexpected atom typings. The query will be confronted to one collection of conformations that can be either one of the proposed banks or a user upload, up to a maximum of either 500 compounds or 10 000 conformations, whichever is limiting. The proposed compound collections comprise two small test sets that are made available for demonstration and 12 focused collections (Table 1), a number likely to evolve in the future. Finally, the possibility to switch the use of the generic rules off and to revert to strict atom type pairing is left, although such option is strongly discouraged since this impacts the pruning of the MSS algorithm and results in increased CPU cost. It is also possible to specify the maximal number nMax of compounds to get returned in their 3D conformations.

At program termination, a list of the compounds ranked by decreasing order is returned. Since the ranking is performed on the basis on the number of bonds and atoms paired, detailed results are provided per compound, along with the Z-score value associated with the number of bonds shared with the query. The server will also supply the 3D visualization of the superimpositions of the 10 best compounds based on the JMol applet (http://www.jmol.org). Finally, it is possible to download a mol2 file of the nMax best compounds superimposed onto the query, for further investigation. For each compound, it contains only the conformation being the most similar to the query.

CASE STUDIES

We illustrate the use wwLigCSRre in three different contexts.
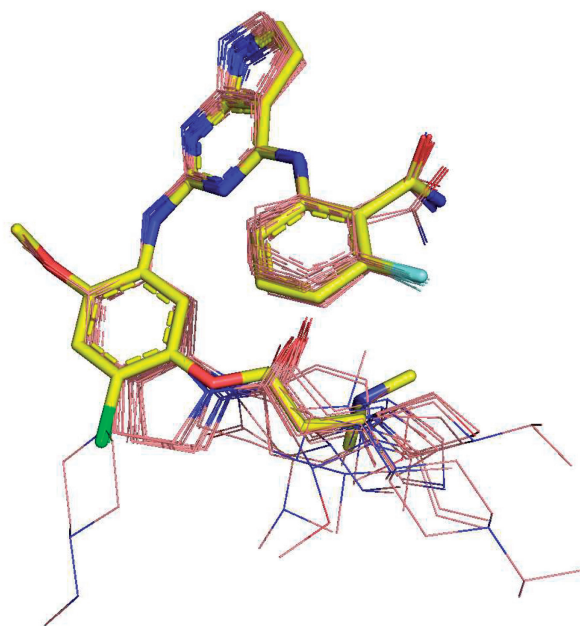
Figure 2. Superimposition of SAR results on IGF-1R.

Structure activity relationship on IGF-1R

Structure activity relationship (SAR) studies rely on the principle that similar compounds, and more specifically molecules sharing a similar scaffold, are likely to share binding properties to the same protein target. A recent study (33) has identified several IGF-1R (insulin-like growth factor-1 receptor) tyrosine kinase inhibitors. About 30 new compounds have been identified sharing a 4,6-*bis*-anilino-1H-pyrrolo[2,3-d]pyrimidine scaffold with various peripheral groups and displaying different binding activities and specificities toward the related JNK1 enzyme. When the shared scaffold can display various conformations in the presence of different peripheral groups it becomes a difficult task to evaluate the correct disposition of them, making the impact on activity of these groups harder to figure out. Figure 2 shows the optimized superimposition by wwLigCSRre of all the SAR structures on one of them used as the query ligand. The present figure helps to anticipate without the structure of the bound protein target, which peripheral groups are rather to be available for protein binding and those that will generate steric stringency within the compound. Although for most of the SAR cases 2D R-group analysis can be sufficient to assess R-group variation, the use of 3D methods such as wwLigCSRre can provide a useful superimposition to visualize the space occupancies of the peripheral groups as it is done in 3D-QSAR projects.

Screening of the Aurora Fine Chemicals Kinase focused library

As mentioned above, the use of 3D ligand-based tools on family-based focused chemical libraries can represent an efficient way to tackle the difficult problem of rapidly identifying new molecule binders on a given protein target. In the case represented in Figure 3, we illustrate the

(a)

(b)

(c)
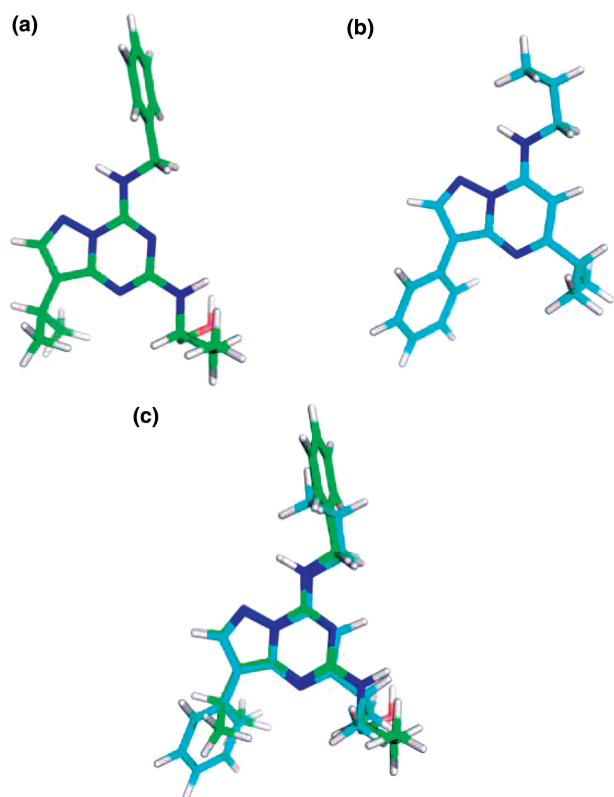
Figure 3. A wwLigCSR run carried out using a Purine bioisostere ligand as the query (**a**) detected a molecule hit (**b**) within the Aurora Fine Chemicals Kinase focused library and proposed the corresponding superimposition (**c**) with a *z*-score of 2.3.

pertinence of using wwLigCSRre on the Aurora Fine Chemicals Kinase focused library. A recent study (34) has discovered a purine bioisostere Roscovitine analog that we ran through wwLigCSRre on the Aurora Fine Chemicals Kinase focused library. Among several interesting hits, the first ranked hit was correctly superimposed onto the query ligand with an interesting level of shared features such as the pyrazolo[1,5-a]-1,3,5-triazine scaffold and some hydrophobic groups superimposed logically onto the hydrophobic groups of the Roscovitine analog.

Example of scaffold hopping on CDK2 inhibitors

Finally, to illustrate the scaffold hopping capacities of wwLigCSRre, we ran a quick run using the co-crystallized ligand of PDB structure 1E9H (a CDK2 inhibitor) as the query on 47 diversified ligands including 9 CDK2, 9 coagulation factor Xa, 10 neuraminidase, 8 ribonuclease and 10 thymidine kinase inhibitors. The results shown on Figure 4 illustrate the three first CDK2 inhibitor hits along with their *z*-scores. The three CDK2 hits, while having 3 different scaffolds, are superimposed onto the indirubin-based ligand of 1E9H. Interestingly, this superimposition corresponds exactly to the crystallographic alignment of the ligands in the crystals. This shows the interest of using 3D methods over regular 2D similarity search when expecting scaffold hopping.

DISCUSSION AND FUTURE WORK

The goal of wwLigCSRre is to provide an online versatile facility to explore the 3D the chemical diversity of compounds. Basically, it can be used for the 3D superimposition of small compounds. Applied to chemically diversified focused banks, it can be used to provide some

(a)

(c)

(b)

(d)

(e)
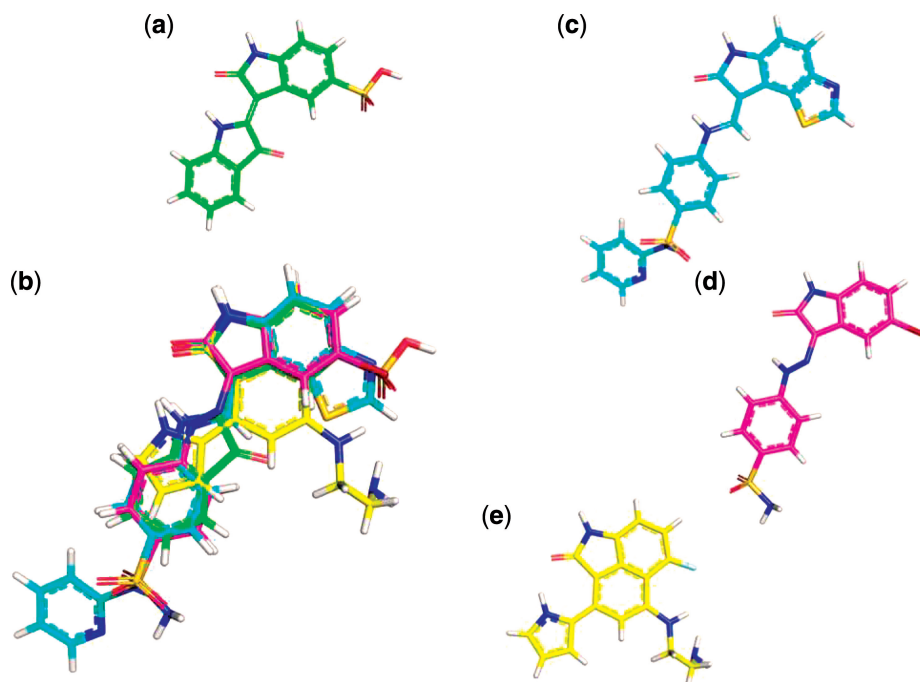
Figure 4. wwLigCSR run on CDK2. 1E9H ligand was used as the query (**a**) and three known CDK2 inhibitors were correctly detected and superimposed onto the query molecule (**b**), with respective *z*-scores of 3.353 (**c**), 2.708 (**d**) and 1.740 (**e**).

scaffold hopping, as introduction to further investigations. Applied to toxic or withdrawn compound libraries, it could also help to identify compounds likely to exhibit undesirable properties that could not easily be detected using classical ADME/Tox facilities. Finally, applied to patented compounds, it could help to identify compounds too similar to these. One limitation of wwLigCSRre comes from the limited number of compounds of the banks proposed. Better rules to prune pairing exploration could result in increased search efficiency and allow for larger banks. Also, LigCSRre presently consider molecular similarity alone. The consideration of both molecular similarities and discrepancies could result in increased performance in bank mining.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Esposito,E.X., Hopfinger,A.J. and Madura,J.D. (2004) Methods for applying the quantitative structure-activity relationship paradigm. *Methods Mol. Biol.*, **275**, 131–214.
2. Khedkar,S.A., Malde,A.K., Coutinho,E.C. and Srivastava,S. (2007) Pharmacophore modeling in drug discovery and development: an overview. *Med. Chem.*, **3**, 187–197.
3. von Grotthuss,M., Koczyk,G., Pas,J., Wyrwicz,L.S. and Rychlewski,L. (2004) Ligand.info small-molecule meta-database. *Comb. Chem. High Throughput Screen.*, **7**, 757–761.
4. Girke,T., Cheng,L.-C. and Raikhel,N. (2005) Chemmine. a compound mining database for chemical genomics. *Plant Physiol.*, **138**, 573–577.
5. Hawkins,P.C.D.A. (2006) A comparison of structure-based and shape-based tools for virtual sreening. *Abstracts of Papers*, 231st ACS National Meeting, Atlanta, GA, United States, March 26–30.
6. Sperandio,O., Andrieu,O., Miteva,M.A., Vo,M.-Q., Souaille,M., Delfaud,F. and Villoutreix,B.O. (2007) Med-sumolig: a new ligand-based screening tool for efficient scaffold hopping. *J. Chem. Inf. Model.*, **47**, 1097–1110.
7. Vainio,M.J., Puranen,J.S. and Johnson,M.S. (2009) Shaep: molecular overlay based on shape and electrostatic potential. *J. Chem. Inf. Model.*, (last accessed date February 5).
8. Bauer,R.A., Bourne,P.E., Formella,A., Frömmel,C., Gille,C., Goede,A., Guerler,A., Hoppe,A., Knapp,E.-W., Pöschel,T. *et al.* (2008) Superimpose: a 3d structural superposition server. *Nucleic Acids Res.*, **36**, W47–W54.
9. Thimm,M., Goede,A., Hougardy,S. and Preissner,R. (2004) Comparison of 2d similarity and 3d superposition. application to searching a conformational drug database. *J. Chem. Inf. Comput. Sci.*, **44**, 1816–1822.
10. Jónsdóttir,S., Jørgensen,F.S. and Brunak,S. (2005) Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates. *Bioinformatics*, **21**, 2145–2160.
11. Irwin,J.J. and Shoichet,B.K. (2005) Zinc–a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.*, **45**, 177–182.
12. Dobson,C.M. (2004) Chemical space and biology. *Nature*, **432**, 824–828.
13. Orry,A.J.W., Abagyan,R.A. and Cavasotto,C.N. (2006) Structure-based development of target-specific compound libraries. *Drug. Discov. Today*, **11**, 261–266.
14. Schnur,D.M. (2008) Recent trends in library design: 'rational design' revisited. *Curr. Opin. Drug. Discov. Dev.*, **11**, 375–380.
15. Gregori-Puigjané,E. and Mestres,J. (2008) Coverage and bias in chemical library design. *Curr. Opin. Chem. Biol.*, **12**, 359–365.
16. Ingsriswang,S. and Pacharawongsakda,E. (2007) smol explorer: an open source, web-enabled database and exploration tool for small molecules datasets. *Bioinformatics*, **23**, 2498–2500.
17. Klekota,J., Roth,F.P. and Schreiber,S.L. (2006) Query chem: a google-powered web search combining text and chemical structures. *Bioinformatics*, **22**, 1670–1673.
18. Rarey,M. and Dixon,J.S. (1998) Feature trees: a new molecular similarity measure based on tree matching. *J. Comput. Aided Mol. Des.*, **12**, 471–490.
19. Boehm,M., Wu,T.-Y., Claussen,H. and Lemmen,C. (2008) Similarity searching and scaffold hopping in synthetically accessible combinatorial chemistry spaces. *J. Med. Chem.*, **51**, 2468–2480.
20. Gerlach,C., Broughton,H. and Zaliani,A. (2008) Ftree query construction for virtual screening: a statistical analysis. *J. Comput. Aided Mol. Des.*, **22**, 111–118.
21. Kleywegt,G.J. and Harris,M.R. (2007) Valligurl: a server for ligand-structure comparison and validation. *Acta Crystallogr. D Biol. Crystallogr.*, **63**, 935–938.
22. Schneidman-Duhovny,D., Dror,O., Inbar,Y., Nussinov,R. and Wolfson,H.J. (2008) Pharmagist: a webserver for ligand-based pharmacophore detection. *Nucleic Acids Res.*, **36**, W223–W228.
23. Sayers,E.W., Barrett,T., Benson,D.A., Bryant,S.H., Canese,K., Chetvernin,V., Church,D.M., DiCuccio,M., Edgar,R., Federhen,S. *et al.* (2009) Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **37**, D5–D15.
24. Goede,A., Dunkel,M., Mester,N., Frommel,C. and Preissner,R. (2005) Superdrug: a conformational drug database. *Bioinformatics*, **21**, 1751–1753.
25. Feng,Z., Chen,L., Maddula,H., Akcan,O., Oughtred,R., Berman,H.M. and Westbrook,J. (2004) Ligand depot: a data warehouse for ligands bound to macromolecules. *Bioinformatics*, **20**, 2153–2155.
26. Petitjean,M. (1998) Interactive maximal common 3d substructure searching with the combined sdm/rms algorithm. *Comp. Chem.*, **22**, 463–465.
27. Guha,R., Howard,M.T., Hutchison,G.R., Murray-Rust,P., Rzepa,H., Steinbeck,C., Wegner,J. and Willighagen,E.L. (2006) The blue obelisk-interoperability in chemical informatics. *J. Chem. Inf. Model.*, **46**, 991–998.
28. Ihlenfeldt,W.-D., Voigt,J.H., Bienfait,B., Oellien,F. and Nicklaus,M.C. (2002) Enhanced cactvs browser of the open nci database. *J. Chem. Inf. Comput. Sci.*, **42**, 46–57.
29. Voigt,J.H., Bienfait,B., Wang,S. and Nicklaus,M.C. (2001) Comparison of the nci open database with seven large chemical structural databases. *J. Chem. Inf. Comput. Sci.*, **41**, 702–712.
30. Lagorce,D., Sperandio,O., Galons,H., Miteva,M.A. and Villoutreix,B.O. (2008) Faf-drugs2: free adme/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics*, **9**, 396.
31. Bohme Leite,T., Gomes,D., Miteva,M.A., Chomilier,J., Villoutreix,B.O. and Tufféry,P. (2007) Frog: a free online drug 3d conformation generator. *Nucleic Acids Res.*, **35**, W568–W572.
32. Sauton,N., Lagorce,D., Villoutreix,B.O. and Miteva,M.A (2008) Ms-dock: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. *BMC Bioinformatics*, **9**, 184.
33. Chamberlain,S.D., Redman,A.M., Patnaik,S., Brickhouse,K., Chew,Y.-C., Deanda,F., Gerding,R., Lei,H., Moorthy,G., Patrick,M. *et al.* (2009) Optimization of a series of 4,6-bis-anilino-1h-pyrrolo[2,3-d]pyrimidine inhibitors of igf-1r: elimination of an acid-mediated decomposition pathway. *Bioorg. Med. Chem. Lett.*, **19**, 373–377.
34. Popowycz,F., Fournet,G., Schneider,C., Bettayeb,K., Ferandin,Y., Lamigeon,C., Tirado,O.M., Mateo-Lozano,S., Notario,V., Colas,P. *et al.* (2009) Pyrazolo[1,5-a]-1,3,5-triazine as a purine bioisostere: Access to potent cyclin-dependent kinase inhibitor (r)-roscovitine analogue. *J. Med. Chem.*, **52**, 655–663.