ELSEVIER

feature

State-of-the-art in ligand-based virtual screening

Peter Ripphausen, Britta Nisius and Jürgen Bajorath, bajorath@bit.uni-bonn.de

Virtual screening is a much discussed topic in chemoinformatics and medicinal chemistry, and widely applied in pharmaceutical research. Here, we provide an in-depth analysis of currently available ligand-based virtual screening applications. We formulate several scientific quality criteria for prospective ligand-based virtual screens and analyze, in detail, the information provided by currently available peer-reviewed publications. The results presented herein provide a detailed view of the current state-of-the-art in this field and point at several problematic issues but also opportunities for further advances.

High-throughput screening (HTS) and virtual screening (VS) essentially have the same goal (i.e. the identification of novel hits). However, HTS and VS are based on different philosophies. Whereas HTS aims at experimental testing of increasingly large numbers of compounds in the most efficient manner, VS departs from this 'let's screen them all' paradigm and attempts to rationalize compound selection to reduce the number of candidates for experimental evaluation as much as possible [1]. Despite the crucial role of HTS in drug discovery settings, the many technical advances and the enormous growth in HTS capacity, VS has become an established computational discipline in pharmaceutical research [2]. Of course, the development and application of computational methods to search through millions of virtually formatted compounds and prioritize small sets of candidates for testing is also attractive from an academic point of view and, hence, there is considerable interest in VS methodologies outside drug discovery environments. However, it is currently difficult to judge VS performance, given the plethora of available computational methods, the many benchmark investigations that still dominate this field [3] and the lack of generally accepted community standards for method evaluation [3,4]. There continues to be much debate on how best to evaluate computational screening methods and benchmark calculations and establish meaningful statistical measures [4]. Moreover, going beyond computational studies and benchmarking, the key measure of success for VS approaches, regardless of the methods that are applied, is the experimental evaluation of computational predictions. Recently, we have carried out a comprehensive literature survey of prospective VS applications (i.e. studies where candidate compounds were predicted and experimentally tested) [5]. This survey has identified the scientific journals where VS applications are mostly published and revealed the most popular VS target protein families and the potency distribution of VS hits [5]. Furthermore, it was determined that approximately three times as many prospective structure-based VS (mostly ligand-target docking) publications than ligand-based VS reports are currently available [5].

This statistical analysis of prospective VS applications raised two other questions concerning VS performance that we felt were important to address. First: once candidate compounds are predicted, how are the predictions evaluated and ultimately 'validated'? Specifically, how rigorous and scientifically sound are the experimental procedures applied to confirm predicted hits? Second: how significant are the findings? In other words, to what extent is the compound information provided by VS approaches relevant or novel? Because our own research activities in the VS arena are, at present, mostly focused on molecular similarity methods for ligand-based virtual screening (LBVS), we have investigated these questions by an in-depth analysis of currently available reports of prospective LBVS applications. In LBVS similarity-based computational methods are generally used to extrapolate from reference compounds with known activity and identify structurally diverse compounds having similar activity.

The results reported herein provide a detailed view of the state-of-the-art of publicly available practical LBVS applications.

1359-6446/06/$ - see front matter doi:10.1016/j.drudis.2011.02.011

Prospective LBVS applications

We have screened peer-reviewed computational, chemical and life science journals that are operational at present and that had a 2009 impact factor of at least 2.0 for papers describing prospective LBVS applications. Including our most recent update (1 November 2010), a total of 115 relevant publications were identified. These publications are provided as: Appendix BAppendix B in Supplementary information. The majority of these LBVS publications have appeared in only five journals including *Journal of Medicinal Chemistry*, *Bioorganic Medicinal Chemistry*, *Bioorganic Medicinal Chemistry Letters*, *Journal of Chemical Information and Modeling* and *ChemMedChem*.

Analysis scheme

These 115 publications were analyzed in detail, as described below. We first determined whether the screening targets were defined, the computational approaches were truly ligand-based and appropriate assays were used. For publications meeting these requirements, we then closely inspected the computational protocols, analyzed how the hits were characterized in detail and determined whether novel compound information was provided. Finally, we identified those studies that met all, or almost all, of our quality control criteria. Because this analysis ultimately required studying all 115 publications in detail to extract the relevant information, the results reported herein have a high level of confidence.

It is important to note that even a thorough analysis of currently available LBVS studies does not provide a full account of the principal opportunities and limitations of LBVS, for two reasons at least. First: only 'successful' applications are published and, without knowledge of negative results, general success rates cannot be rigorously determined. From this point of view, the analysis of original literature principally overestimates the performance of LBVS applications. Second: many of the VS applications carried out in pharmaceutical companies remain proprietary, probably including the most interesting ones. From this point of view, the analysis underestimates the opportunities of LBVS. Despite these general limitations, a comprehensive analysis of currently available peer-reviewed publications is the most rigorous way to obtain a detailed picture of the current state-of-the-art in applied LBVS.

Basic scientific criteria

Initially, we determined the potency distribution of the most active hit(s) identified by LBVS in each study. We note that ~60% of these studies reported hits with <1 μM potency and ~30% hits with <10 μM potency. Thus, the majority of these LBVS hits were relatively potent. In fact, in only two cases, hits with borderline detectable potency of >100 μM were reported. This LBVS hit distribution has been an encouraging finding that, as we have shown previously, is not representative of the VS field as a whole [5].

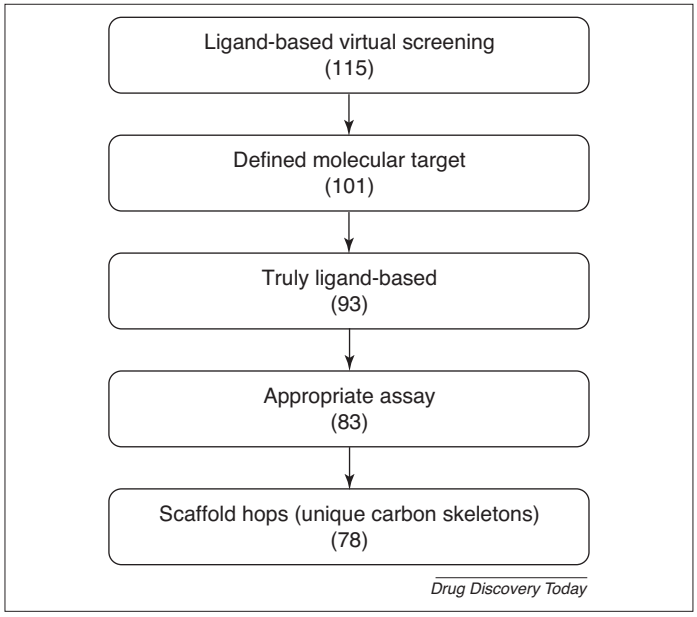

FIGURE 1

Assessment of prospective LBVS applications. A flow chart is shown that represents a sequence of analysis steps applied to evaluate the basic scientific quality of LBVS investigations. Studies passing these filters were considered successful LBVS applications. The number of qualifying publications at each stage is reported in parentheses.

We next evaluated our pool of 115 prospective LBVS publications following a four-step procedure that is summarized in Fig. 1.

Initially, 14 studies were omitted from further consideration that did not address a defined target. Here, compounds were computationally selected on the basis of specifically active reference compounds and screened on bacterial strains or cancer cell lines without proposed molecular targets. Although such 'phenotypic' screening efforts for broad-spectrum antibacterial or anticancer activities are not unreasonable *per se*, we considered them less important for the assessment of LBVS than target-specific screening efforts. We then studied whether the computational approaches were truly ligand-based and identified eight studies where it remained unclear whether, or to what extent, target structure information was also taken into consideration. For the remaining 93 publications, we then analyzed whether appropriate assays were used to test candidate compounds. For example, the experimental evaluation was considered inappropriate if no assay details were provided. In addition, it was determined whether defined and reproducible potency end points (i.e. K_d, K_i, IC_{50}, EC_{50} or ED_{50} values) were reported rather than, for example, % relative inhibition. Furthermore, for enzymes for which standard direct inhibition assays were available (e.g. proteases, cyclooxygenases or lipoxygenases) it was considered inappropriate to 'confirm' the activity of candidate compounds by testing them on cell lines (for which proposed target-specific inhibition remained ambiguous). Of course, in other cases, such as G-protein-coupled receptors, ion channels or transcription factors, assays using membrane preparations or other cell-based assays must be used. On the basis of these criteria, assay results in ten studies were considered questionable, and these studies were also not further considered.

Thus, taken together, 32 of 115 peer-reviewed prospective LBVS applications (27.8%) were identified that lacked well-defined targets or experimentally evaluated LBVS candidate compounds in a non-rigorous or questionable manner.

Structural novelty

The identification of compounds having a desired activity that structurally depart from known active reference compounds and represent new chemotypes is the primary goal of LBVS. By contrast, the identification of analogs of known actives or very similar compounds is not challenging and can be attempted, for example

by simple substructure searching (we have only identified one study where substructure searching was used as an LBVS approach to identify a novel scaffold) [6]). For the 83 studies that passed our target and experimental filters we systematically analyzed the structural novelty of confirmed active compounds step "appropriate assay" in Fig. 1. Therefore, we extracted hierarchical scaffolds [7] from reported hits by removing all substituents from ring systems, but retaining linkers between rings and further transforming these scaffolds into carbon skeletons (CSKs) by converting all heteroatoms to carbon and setting all bond orders to one. Unique CSKs represent topologically distinct molecular scaffolds. For each hit taken from a publication, all compounds having a corresponding activity annotation were collected from ChEMBL [8] and BindingDB [9] – the major repositories of bioactive molecules from medicinal chemistry sources – and also transformed into CSKs. In each case, care was taken to select only correspondingly active database compounds reported before publication of the LBVS study. The CSKs from LBVS hits and known active compounds were then compared. This analysis was carried out with in-house-generated Pipeline Pilot scripts [10]. We determined that 78 of the remaining 83 LBVS studies indeed reported active compounds that corresponded to previously unobserved (topologically unique) CSKs and, hence, represented novel active structures. This was a rather encouraging result. It should also be noted that two of the five studies that were omitted owing to the lack of structural novelty explicitly stated analog searching as a primary goal.

Quality criteria

The 78 publications reporting structurally novel hits for well-defined targets that were experimentally confirmed in a rigorous manner were ultimately considered to be 'successful' LBVS applications. Next, we analyzed these studies using additional quality criteria, as summarized in Fig. 2. First, we studied the LBVS protocols in detail and determined whether the applied methods and search calculations were clearly described and in sufficient detail to reproduce the calculations. For example, we examined whether the reference compounds were provided, the search methods and descriptors, and that the similarity metrics and threshold criteria were defined. Here we detected shortcomings in a total of nine publications. In several instances, these were letter-type reports containing only very brief methodological descriptions. Regardless, the absence of clear or sufficiently detailed descriptions of computational procedures in several successful LBVS applications was considered non-satisfactory from a methodological point of view.

In the context of evaluating computational protocols, we also analyzed an additional aspect that should be of interest to virtual screeners using fingerprints for similarity searching. We investigated which Tanimoto similarity threshold values were applied in fingerprint searching as an indicator of biological activity similarity between reference and database compounds. Eleven relevant publications were identified and Tanimoto similarity ranges were applied in only two of these studies. The remaining nine studies defined Tanimoto coefficient cutoff values above which (or, in three cases, below which) compounds were considered as potential hits. Most of these Tanimoto similarity threshold values were set to 0.75 or 0.85, regardless of the fingerprints that were utilized. Here, it should be noted that we have recently shown that it is not possible to derive Tanimoto similarity threshold values for different fingerprints that are reliable indicators of biological activity similarity and that Tanimoto coefficient ranges where bioactive compounds might be enriched in moderately sized database selection sets significantly vary for fingerprints of different design [11].

Additional quality criteria were also considered for the experimental evaluation of hits. For example, 38 studies reported a multistep experimental validation approach, including, for example, binding/inhibition assays followed by a functional assay to characterize initial hits

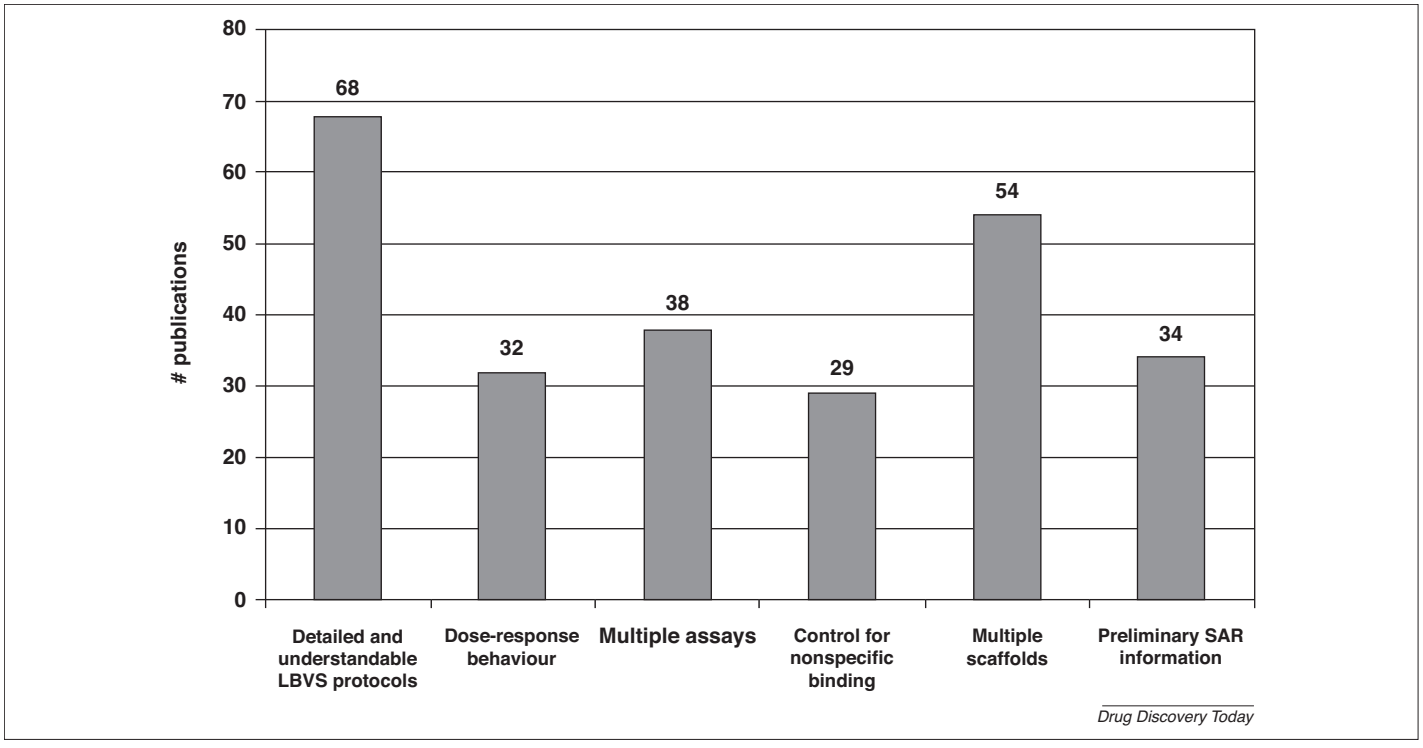

FIGURE 2

Additional quality criteria. The number of studies meeting each of six additional criteria applied to evaluate successful LBVS applications is reported.

further, which was clearly a plus. Furthermore, 32 studies showed detailed dose–response curves for the newly identified active compounds. Additionally, 29 studies reported additional controls for non-specific binding.

Finally, quality criteria assessing the compound information and novelty were applied. In 54 of the 78 reports more than one new active chemotype was identified, and in 34 publications preliminary SAR information was also provided for novel hits by testing analogs of these hits or other structurally similar compounds.

Advanced LBVS applications

The additional criteria discussed above helped to identify LBVS applications of comparably high scientific quality. However, only four studies were identified that met all of the six criteria summarized in Fig. 2. These studies are listed in Table 1 [Refs. 12–15]. Obtaining preliminary SAR information might be considered as the initial step to bridge between LBVS and hit-to-lead projects. If we did not require the presence of preliminary SAR information, five additional studies did meet the remaining five criteria, and these studies are also highlighted in Table 1 [Refs. 16–20]. These advanced LBVS applications also included short letters demonstrating that comprehensive and detailed information can be readily provided in a short publication format. Taken together, on the basis of our analysis, the investigations reported in Table 1 currently represent the overall most-advanced prospective LBVS applications. Of course, other quality criteria might also be considered and/or selected criteria might be weighted in different ways. For example, one might also be interested in 'hit rates' of LBVS calculations (i.e. the ratio of confirmed active over experimentally tested compounds). However, in our analysis, the calculation of such rates was often not very informative because, in many cases, only subsets of computationally selected compounds were available for testing or 'subjective' selection criteria were applied to arrive at final compound testsets. Regardless, the investigations in Table 1 are recommended as further readings and instructive examples of advanced LBVS applications.

Summary and outlook

Herein we have presented a thorough analysis of the state-of-the-art of LBVS that has taken into account currently available peer-reviewed prospective LBVS applications. As discussed, given the grey zone of unpublished failures and proprietary applications, the analysis does not provide a measure of the true potential of LBVS. However, it mirrors the level of scientific rigor and quality in this field and also provides insights into what might be expected from practical LBVS efforts. In addition to LBVS experts, the results presented herein should also be helpful to practitioners in pharmaceutical research who might occasionally apply commercial LBVS tools. There are no reasons why LBVS studies might not be carried out successfully by bench scientists providing the requirements and limitations of LBVS predictions and their experimental evaluation are carefully considered.

On the basis of our analysis, there are good news and bad news. To start with the latter, we first note that approximately a third of all published prospective LBVS studies were found to have, more or less, significant shortcomings in the way computational predictions were assessed experimentally. Furthermore, ~10% of preselected successful LBVS applications lacked sufficiently detailed computational descriptions to understand fully and/or reproduce the calculations. Both these issues will require attention as the field moves forward. Our analysis of computational protocols also pointed at another potentially problematic issue (i.e. methodological relevance), which cannot be fully addressed at present because of the lack of generally accepted calculation standards. In many instances, it was not possible to evaluate on the basis of original publications which, of many, calculations that were carried out were decisive for the identification of novels hits. This problem might be addressed in the future, for example by requiring simple similarity search calculations as controls for more-complex LBVS protocols to determine whether the applied LBVS procedures were indeed essential for hit identification. On the positive side, we have also found that the majority of scientifically sound LBVS studies have indeed identified novel active compounds whose scaffolds were not present in major repositories of bioactive molecules before publication. In addition, many LBVS hits were fairly potent, with ~30% falling into the submicromolar range. This is all good news for the LBVS field as it evolves further. Taken together, the results of our analysis revealed that many practical LBVS applications resulted in the identification of new and interesting active compounds. However, further improvements in the presentation and relevance assessment of computational approaches and the experimental evaluation of candidate compounds will be required.

Finally, we would anticipate that a systematic analysis of prospective ligand–protein docking studies, conceptually similar to the one presented herein and carried out by leading investigators in this area, might also yield interesting results. Such an analysis would certainly represent another large-magnitude task. However, it should be timely and useful to assess and perhaps further advance the structure-based virtual screening field.

TABLE 1

Advanced LBVS publications

Prospective LBVS applications meeting additional quality criteria
12. Carosati, E. *et al*. (2010) *Bioorg. Med. Chem.* 18, 7773–7785
13. Franke, L. *et al.* (2007) *J. Med. Chem.* 50, 2640–2646
14. Noeske, T. *et al*. (2007) *ChemMedChem* 2, 1763–1773
15. Shoda, M. *et al*. (2004) *J. Med. Chem.* 47, 4286–4290
16. Stumpfe, D. *et al*. (2010) *ACS Chem. Biol.* 5, 839–849
17. Peng, Y. *et al*. (2010) *Bioorg. Med. Chem. Lett.* 20, 4825–4830
18. Foloppe, N. *et al*. (2009) *Bioorg. Med. Chem. Lett.* 19, 4183–4190
19. Naylor, E. *et al*. (2009) *Nat. Chem. Biol.* 5, 220–226
20. Fan, Y. *et al*. (2008) *Bioorg. Med. Chem. Lett.* 18, 5789–5791

Reported are prospective LBVS studies that meet all quality criteria according to Fig. 2 (12–15) or all criteria except 'preliminary SAR information' (16–20). Studies 12–15 and 16–20 are ranked by the year of publication.

Conflicts of interest statement

The authors declare no competing financial interests.

Acknowledgements

We thank Eugen Lounkine for helpful discussions. P.R. is supported by a post-graduate program of the *Deutsche Forschungsgemeinschaft* (GRK804) and B.N. by Bayer Healthcare AG, Wuppertal.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.drudis.2011.02.011.

References

1 Bajorath, J. (2002) Integration of virtual and high-throughput screening. *Nature Rev. Drug Discov.* 1, 882–894
2 Schneider, G. (2010) Virtual screening: an endless staircase. *Nature Rev. Drug Discov.* 9, 273–276
3 Geppert, H. *et al.* (2010) Current trends in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and

performance evaluation. *J. Chem. Inf. Model.* 50, 205–216
4 Jain, A. and Nicholls, A. (2008) Recommendations for evaluation of computational methods. *J. Comput.-Aided Mol. Des.* 22, 133–139
5 Ripphausen, P. *et al.* (2010) Quo vadis, virtual screening? A comprehensive survey of prospective applications. *J. Med. Chem.* 53, 8461–8467
6 Breault, G.A. *et al.* (2008) Exploring 8-benzyl-6,7-diones as inhibitors of glutamate racemase (Murl) in Gram-positive bacteria Bioorg. *Med. Chem. Lett.* 18, 6100–6103
7 Bemis, G.W. and Murcko, M.A. (1996) The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* 39, 2887–2893
8 ChEMBL, (2010) *European Bioinformatics Institute (EBI), Cambridge.* http://www.ebi.ac.uk/chembldb/
9 Liu, T. *et al.* (2007) BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res.* 35, 198–201
10 *Pipeline Pilot* (student edition, version 6.1), Accelrys, San Diego, CA
11 Vogt, M. *et al.* (2010) Scaffold hopping using two-dimensional fingerprints: true potential, black magic, or a hopeless endeavor? Guidelines for virtual screening. *J. Med. Chem.* 53, 5707–5715
12 Carosati, E. *et al.* (2010) Ligand-based virtual screening and ADME-Tox guided approach to identify triazolo-quinoxalines as folate cycle inhibitors Bioorg. *Med. Chem.* 18, 7773–7785
13 Franke, L. *et al.* (2007) Identification of natural-product-derived inhibitors of 5-lipoxygenase activity by ligand-based virtual screening J. *Med. Chem.* 50, 2640–2646
14 Noeske, T. *et al.* (2007) *Virtual screening for selective allosteric mGluR1 antagonists and structure-activity relationship investigations for coumarine derivatives ChemMedChem* 2, 1763–1773
15 Shoda, M. *et al.* (2004) Identification of structurally diverse growth hormone secretagogue agonists by virtual screening and structure-activity relationship analysis of 2-formylaminoacetamide derivatives. *J. Med. Chem.* 47, 4286–4290
16 Stumpfe, D. *et al.* (2010) Targeting multifunctional proteins by virtual screening: structurally diverse cytohesin inhibitors with differentiated biological functions. *ACS Chem. Biol.* 5, 839–849
17 Peng, Y. *et al.* (2010) Discovery of novel alpha7 nicotinic receptor antagonists Bioorg. *Med. Chem. Lett.* 20, 4825–4830
18 Foloppe, N. *et al.* (2009) Discovery and functional evaluation of diverse novel human CB1 receptor ligands Bioorg. *Med. Chem. Lett.* 19, 4183–4190
19 Naylor, E. *et al.* (2009) Identification of a chemical probe for NAADP by virtual screening Nat. *Chem. Biol.* 5, 220–226
20 Fan, Y. *et al.* (2008) The identification of neurotensin NTS1 receptor partial agonists through a ligand-based virtual screening approach Bioorg. *Med. Chem. Lett.* 18, 5789–5791

Peter Ripphausen,
Britta Nisius,
Jürgen Bajorath

Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität, D-53113 Bonn, Germany