Integration of virtual and physical screening

Dan C. Fara[1], Tudor I. Oprea[1], Eric R. Prossnitz[3], Cristian G. Bologa[1],
Bruce S. Edwards[2], Larry A. Sklar[2,*]

[1]Department of Biochemistry and Molecular Biology, Division of Biocomputing, University of New Mexico, Health Sciences Center, Albuquerque, NM 87131, USA
[2]Cancer Research and Treatment Center and Department of Pathology, University of New Mexico, Health Sciences Center, Albuquerque, NM 87131, USA
[3]Department of Cell Biology and Physiology, University of New Mexico, Health Sciences Center, Albuquerque, NM 87131, USA

High-throughput screening (HTS) represents the dominant technique for the identification of new lead compounds in current drug discovery. It consists of physical screening (PS) of large libraries of chemicals against one or more specific biological targets. Virtual screening (VS) is a strategy for *in silico* evaluation of chemical libraries for a given target, and can be integrated to focus the PS process. The present work addresses the integration of both PS and VS, respectively.

Section Editors:
Tudor Oprea – University of New Mexico School of Medicine, Albuquerque, USA
Alex Tropsha – University of North Carolina, Chapel Hill, USA

Virtual versus physical (high-throughput) screening

The random physical screening (PS) of compound collections, that is the in-house database of a pharmaceutical company, or the catalogs of various chemicals' vendors, represented a long-range strategy and was regarded as a substitute for serendipity [1].

Automated compound handling and assaying facilities currently enable tasks such as compound retrieval from storage, dilution and plating of samples; therefore, PS is reaching its potential to identify hits on a timely basis. While random screening has enabled some successes, it has not fulfilled initial expectations. A gross increase in the number of assayed compounds does not guarantee better productivity *per se* [2].

Also, PS results are not free from errors and different assay formats for the same target can give different results [3,4].

The *in silico* screening of (virtual) libraries of compounds is conceptually and economically attractive, as it makes possible the evaluation of an almost unlimited number of chemical structures, only a subset of which will be selected and subsequently assayed in a PS experiment. Typical virtual screening (VS) methods involve, for example, filtering of libraries for compounds containing toxic, reactive, or otherwise undesirable groups, or, by contrast, the search for molecules with preferred lead- or drug-like properties and desired activity [5]. In fact, VS is emerging as a key strategy to help filter out those compounds with poor potency, and biological or pharmacological properties [6–8]. Recent reviews suggest that while VS is often presented as an alternative to PS, both strategies are highly complementary [5,9].

Integrating virtual and physical screening

The current trend in the pharmaceutical industry is to integrate computational and experimental technologies early in the drug discovery process [10–13]. For instance, it is considered that a better and earlier utilization of information,

---

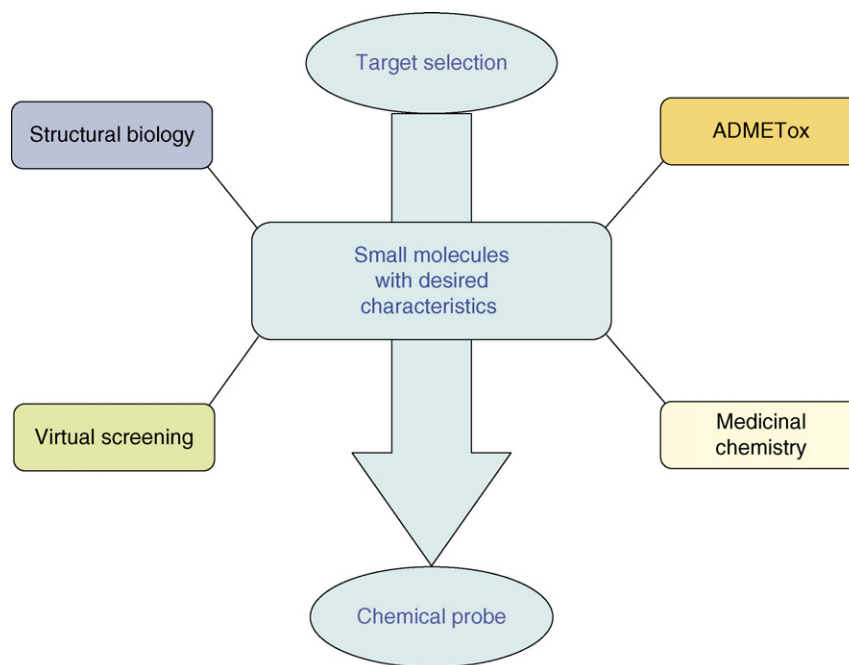*Corresponding author: L.A. Sklar (lsklar@salud.unm.edu)

Figure 1. Integration of virtual chemistry and screening. A general scheme that shows how a continuous information exchange between different areas could lead to chemical libraries with more desirable chemical and biological properties. Abbreviation: ADMETox: absorption, distribution, metabolism, excretion and toxicity evaluation.

that is (i) genomic, (ii) chemical, (iii) biological, (iv) structural, and (v) molecular property data, would lead to chemical libraries with more desirable chemical and biological properties [14,15]. This can be done by integrating information from different areas that include: (i) analysis of the gene or protein family for target selection; (ii) absorption, distribution, metabolism, excretion and toxicity evaluation (ADME/Tox); (iii) structural biology; (iv) VS; and (v) medicinal chemistry with parallel synthesis (Fig. 1).

Literature reports from 2005 to 2006 show an increased dependence on the integration VS and PS for identifying leads in the drug discovery process. Twenty-three representative studies are summarized in Table 1, which lists (i) the target studied, (ii) the strategy applied for both VS and PS methods, (iii) the size of the library evaluated, and (iv) the results achieved in each project. The last column indicates the original references.

Table 1 reveals the VS workflow used in most of the studies: (i) preliminary filtering of a virtual library (tens to hundreds of thousands of compounds) using various criteria (e.g. Lipinski's rule of five); (ii) homology and pharmacophore modeling; combined with (iii) docking and scoring of the protein–ligand complexes. For PS methods, assays that are particular to each specific target were applied, most of them being based on fluorescence detection. The outcome of these studies indicates that preliminary VS not only drastically reduces the number of chemicals that are physically screened, but also increases the hit rates.

Integration strategy

Target selection

VS is a knowledge-driven approach that depends on the amount and quality of information available about the system under investigation. Knowing the structure of the biological target macromolecule offers many advantages in comparison with the situation where only information about the geometry of a bioactive (reference) ligand is available. When structural information about the target is available from either fact or inference, we consider this target-based, or structure-based, virtual screening (TBVS). All other cases represent ligand-based virtual screening (LBVS), where the reference compound(s) are known substrates, inhibitors, agonists or antagonists, among others. An important issue in target selection is the druggability of the molecule under consideration, and whether the target is amenable to therapeutic intervention via small molecules. One approach is to assess whether members from the same gene family show similar binding affinity towards drug-like ligands with related physicochemical properties [16]; the other approach is sometimes referred as 'target fishing' (see J. L. Jenkins, this issue).

With the advent of structural genomics and homology modeling initiatives [17], the number of potential targets is expected to grow. For instance, over the last year (2005–2006), 26 new targets were screened at the NIH Molecular Libraries Initiative (formerly the Molecular Libraries Screening Center Network) (http://mli.nih.gov/); more are expected to be screened within next 2 years.

Table 1. Overview of recent publications that integrate virtual and physical screening (VS and PS)

Nos.	Target	Strategy		Data set (molecules)		Outcome	Refs
		VS	PS	VS	PS		
1	FPR[a]	Homology modeling + docking + pharmacophore selection	Hypercyt[TM] – high-throughput flow cytometry	~480,000	4324	30 potential lead cpds (K_i^b = 1 ÷ 32 μM); 4 partial agonists; 26 antagonists; VS improved the PS hit rate by fourfold	[18]
2	GPR30[c]	2D + shape + field similarity	Hypercyt[TM] – high-throughput flow cytometry	9283	100	1 low nanomolar GPR30 selective agonist	[19]
3	Caspase-3	Self-organizing Kohonen maps	Fluorescence + microtitration	18,073	1233	6 quinoline derivatives as inhibitors; IC_{50}^d = 4 ÷ 30 nM	[31]
4	β-Amyloid Peptide	Docking + scoring	ThT[e] fluorescence	~278,000	125	1 potent inhibitor of $A\beta_{1-42}$ fibrillation (concentration-dependent manner); IC_{50} = 35.6 μM	[32]
5	[f]SARS-CoV M[pro]	Docking + scoring + analogue searching	Fluorescence resonance energy transfer (FRET[g])	58,855	50	21 inhibitors; IC_{50} = 0.3 ÷ 50 μM	[33]
6	P-12-LOX[h]	Homology modeling + docking + scoring	Direct measurement of the 12(S)-HETE[i] formation	80	10	4 synthetic curcuminoids as inhibitors; IC_{50} = 1.7 ÷ 66 μM	[34]
7	FAAH[j], MGL[k]-like in rat cerebellar membranes	Homology modeling + docking + scoring	Radiolabeled ligand (FAAH) + HPLC analysis (MGL)	2376	62	9 FAAH inhibitors; IC_{50} = 0.52 ÷ 44 μM; 1 MGL-like inhibitor; IC_{50} = 31 μM	[35]
8	MDM2[l]-p53[m] interaction	Druglike filtering + pharmacophore searching + docking + scoring	Fluorescence – polarization-based (FP-based)	150,000	67	10 inhibitors (K_i < 10 μM); the most potent has K_i = 120 nM	[36]
9	11β-HSD1[n], 11β-HSD2[o]	Pharmacophore modeling + data mining + docking	Radiolabeled ligand in (i) lysates of HEK-293 cells[p], and (ii) intact transfected HEK-293 cells	1,776,579	30	7 cpds inhibited more than 70% of the activity of 11β-HSD1 (IC_{50} < 10 μM); 5 out of those 7 cpds showed significant inhibition of 11β-HSD2 (IC_{50} < 10 μM)	[37]
10	CypA[q]	Focused combinatorial library design + docking + scoring	Surface plasmon resonance (SPR[r])	85,000	40	4 potent inhibitors of CypA; IC_{50} = 2.5 ÷ 6.2 μM	[38]
11	CypA	Pharmacophore selection + docking + scoring	Chymotrypsin-coupled colorimetric assay	296,387 (3129)	31	5 cpds inhibited more than 80% of the isomerase activity of CypA; IC_{50} = 0.3 ÷ 5 μM	[39]
12	Cysteine protease from *P. falciparum* (falcipain-2, falcipain-3) and *L. donovani*	Filtering + docking + scoring + pharmacophore modeling	Fluorescence + flow cytometry	355,000	100	18 cpds active against falcipain-2 (IC_{50} = 1.4 ÷ 54.3 μM and falcipain-3 (IC_{50} = 11.4 ÷ >50 μM); 4 cpds active against *Leishmania donovani* cysteine protease (IC_{50} = 23.5 ÷ 43.0 μM)	[40]
13	12-hLO[s], 15-hLO[t]	Homology modeling + docking + scoring	Measurement of the rate of formation of the conjugated diene products at 234 nm	50,000	20	3 low molar inhibitors of 15-hLO (IC_{50} = 6.8 ÷ 18.8 μM) and 2 out of 3 are low molar inhibitors of 12-hLO (IC_{50} = 12.3 ÷ 30.7 μM)	[41]
14	ALK[u]	Pharmacophore selection + docking + scoring + homology modeling	Amplified luminescent proximity homogeneous assay	60,000 724	2677 24	5-Aryl-pyridone-3-carboxamide derivatives as novel ALK inhibitors; IC_{50} = 0.4 ÷ 19.5 μM	[42]
15	ACE2[v]	Pharmacophore modeling + docking + scoring	Fluorescence	63,307	17	6 cpds showed an inhibitory effect on ACE2; IC_{50} = 62 ÷ 179 μM	[43]
16	GSK-3β[w]	Docking + scoring + pharmacophore selection	Kinase-Glo[TM] luminescent kinase assay + ATP[x] titration	16,299 (only 5904 were docked)	16,299	90 validated hits (6 clusters) by PS; VS identified 25–33% of actives in clusters 1–4 and failed to pick up any hit from clusters 5 to 6	[44]
17	DPP-IV[y]	Filtering + pharmacophore selection + docking + scoring	Fluorescence	800,000	4000	51 active compounds (%inhibition ranges from 30 to 81.9 at 30 μM); the hit rate for the actives is 1.28% (VS) and 0.012% (PS)	[45]

Table 1 (*Continued*)

Nos.	Target	Strategy		Data set (molecules)		Outcome	Refs
		VS	PS	VS	PS		
18	COX-2[z]	Pharmacophore modeling + support vector machine classification + docking	ELISA[aa] + monoclonal antibody + microtitration	~2,7 millions	22 (13 cherry picked)	3 cpds exhibited an inhibitory effect in the activity assay (the most potent had $IC_{50} = 0.2 \pm 0.3$ μM)	[46]
19	AChE[bb]	Pharmacophore modeling + docking + scoring	Extract screening (LC-MS[cc]) + Modified Ellman's method	47	5	2 cpds were confirmed as weak actives: (i) 8-deoxylactucin ($IC_{50} = 308.1$ μM), and (ii) lactucopicrin ($IC_{50} = 150.3$ μM)	[47]
20	HRV[dd] coat protein	Pharmacophore modeling + docking + scoring + PCA[ee]-based clustering	Multiple-cycle CPE[ff] inhibition + microscopic + spectrophotometric	~60,000	6	6 promising inhibitors; [gg]$EC_{50} = 7.3 \div 247.1$ μmol/L (microscopic), $EC_{50} = 4.3 \div 245.5$ μmol/L (spectrophotometric)	[48]
21	Stat3[hh]	Filtering + docking + scoring	Stat3-dependent luciferase reporter + gel EMSA[ii]	~429,000	100	1 natural product, a deoxytetrangomycin, an angucycline antibiotic, showed remarkable inhibition of Stat3 dimerization, DNA binding, and nucleus translocation as well as the Stat3-regulated genes such as Bcl-X$_L$[jj] and cyclin D1	[49]
22	EphB2[kk]	Filtering + docking + scoring + pharmacophore modeling	Frontal affinity chromatography with mass spectrometry detection (FAC-MS[ll])	50,452	468	12 potential inhibitors; $IC_{50} = 5.2 \div 250$ μM	[50]
23	Alpha1A[mm]	Filtering + pharmacophore modeling + docking + scoring	Radioligand displacement assay	23,000	80	37 hits confirmed ($K_i < 10$ μM) of which 3 cpds showed affinity between 1 and 10 nM	[51]

[a] Formyl peptide receptor.
[b] Inhibition constant.
[c] G protein-coupled receptor 30.
[d] The half maximal inhibitory concentration, represents the concentration of an inhibitor that is required for 50% inhibition.
[e] Thioflavin T.
[f] Severe acute respiratory syndrome coronavirus main protease.
[g] Fluorescence resonance energy transfer.
[h] Platelet 12-lipoxygenase.
[i] 12(S)-hydroxyeicosatetraenoic acid.
[j] Fatty acid amide hydrolase.
[k] Monoglyceride lipase.
[l] Murine double minute 2.
[m] Tumor protein 53.
[n] 11β-hydroxysteroid dehydrogenase type 1.
[o] 11β-hydroxysteroid dehydrogenase type 2.
[p] Human embryonic kidney epithelial cell line.
[q] Cyclophilin A.
[r] Surface plasmon resonance.
[s] Human platelet-type 12-lipoxygenase.
[t] Human reticulocyte 15-lipoxygenase-l.
[u] Anaplastic lymphoma kinase.
[v] Angiotensin converting enzyme 2.
[w] Glycogen synthase kinase-3β.
[x] Adenosine triphosphate.
[y] Dipeptidyl peptidase IV.
[z] Cyclooxygenase 2.
[aa] Enzyme-linked immunosorbent assay.
[bb] Acetylcholinesterase.
[cc] Liquid chromatography–mass spectrometry.
[dd] Human rhinovirus.
[ee] Principal components analysis.
[ff] Cytopathic effect.
[gg] The concentration of a compound where 50% of its maximal effect is observed.
[hh] Signal transducers and activators of transcription 3.
[ii] Electrophoretic mobility shift assay.
[jj] An anti-apoptotic protein of the Bcl-2 (*B-cell lymphoma 2*) gene family.
[kk] Erythropoietin-producing hepatocellular B2.
[ll] Frontal affinity chromatography with mass spectrometry detection.
[mm] Adrenergic receptor subtype A.

Table 2. Screening of various class and target focused libraries: brief comparison

Library	Library type	Compounds	Confirmed hits <30 μM	Best hit (μM)	% Hits
PCL[a]	Specialty set (marketed drugs)	880	1	24	0.11
MLSCN 10K[b]	Diverse (85% DC[c], 10% TL[d], 5% SS[e])	9993	17	4	0.17
GPCR[f]	Class focused	4959	12	3	0.24
FPR[g]	Target focused (pharmacophore)	4324	30	1	0.70

[a] A library from the Prestwick Chemical Library.
[b] A library from the NIH small molecule repository.
[c] Diverse collection.
[e] Specialty sets.
[d] Targeted libraries.
[f] Library focused for G protein-coupled receptors.
[g] Library focused for formyl peptide receptor.

Our Center (New Mexico Molecular Libraries Screening Center: http://screening.health.unm.edu/) has undertaken projects on several classes of receptors, which include two GPCRs listed in Table 1 and an integrin, LFA-1. For the formyl peptide receptor, we compared several strategies including screening of class-focused (GPCR) and target-focused (FPR) libraries (Table 2). In the latter case, we used a homology model with pharmacophore docking to select a compound library [18]. We recorded active molecules at the rate of 1/880 from the Prestwick Chemical Library (http://www.prestwick-chemical.com/chem_lib.htm), 17/9993 in a diverse set from the NIH Small Molecule Repository (http://mlsmr.glpg.com/MLSMR_HomePage/), 12/4959 in a GPCR class focused library from ChemDiv (http://www.chemdiv.com/en/products/screening/), and 30/4324 in a target focused library [18].

In the case of GPR30, where no Target-specific ligand was known, we performed LBVS on the latter two libraries, using 17β-estradiol as the reference ligand. We ranked the top 100 structures for PS, finding one selective nM affinity agonist for GPR30 [19] and two ERα/ERβ that do not bind to GPR30 (C.G. Bologa, unpublished).
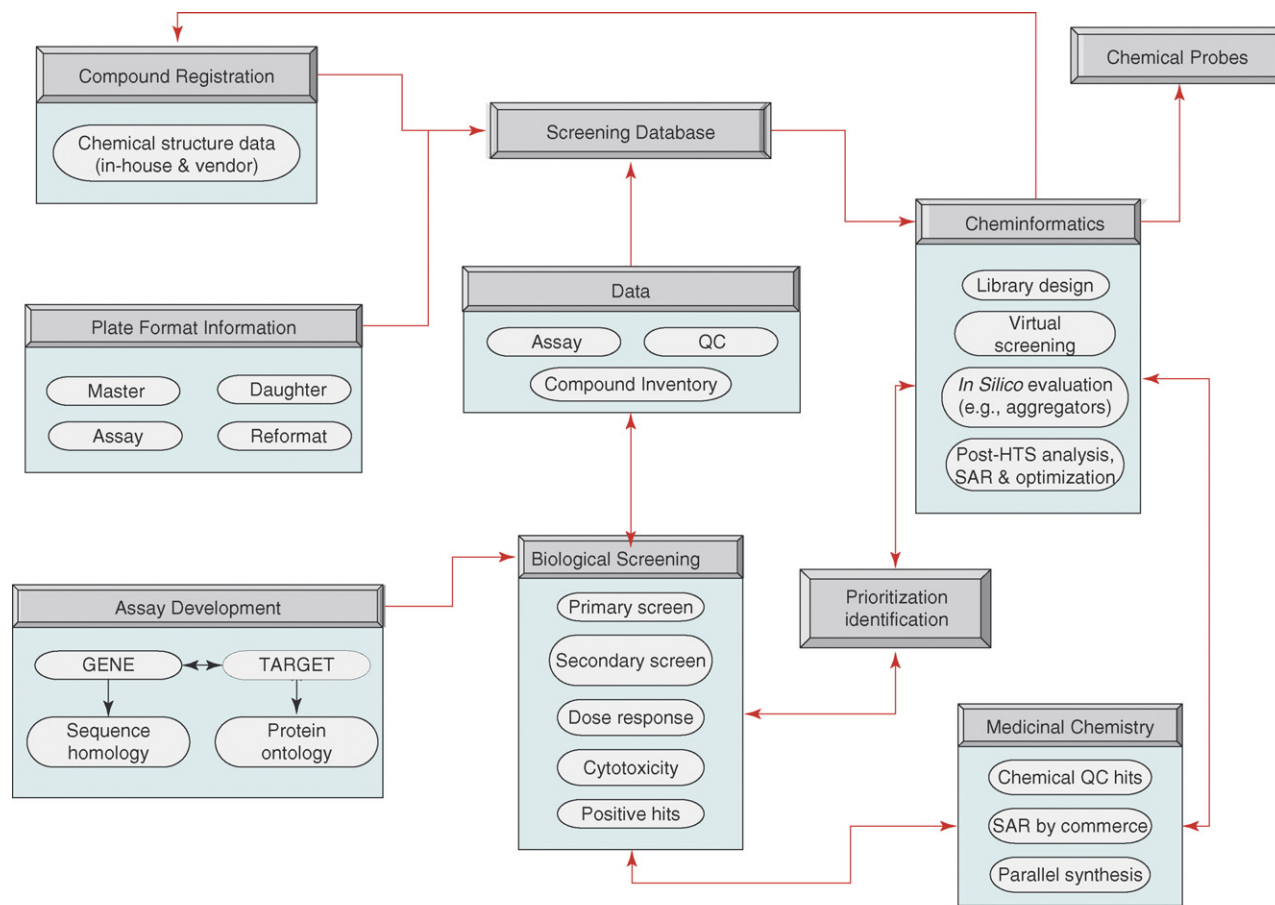
Workflow
The generic workflow shown in Fig. 2 is currently implemented at the New Mexico Molecular Libraries Screening Center, and follows the general template presented in Fig. 1. The core of the system is represented by the 'screening database' where complete information related to the (i) chemical structure data of the compounds, that is existing in-house or purchased from different vendors, and (ii) bioactivity data, that is plate format information and biological screening outcome, is recorded. The workflow steps are well defined. The biological team provides the target and the corresponding assay to the screening team, which is responsible for the PS. Once compounds are registered into the database by the cheminformatics team, plates are created and recorded in the database. The medicinal chemistry team provides chemistry follow-up and property optimization.

There is a continuous information exchange between the screening, cheminformatics, and medicinal chemistry teams. Based on the outcome from the screening team, the cheminformatics team applies VS and post-high-throughput screening (HTS) analyses to further identify and prioritize compounds (or chemotypes) for further evaluation. The medicinal chemistry team verifies the chemical qualitative chemistry (QC) of the proposed hits and gives feedback to both the screening and cheminformatics teams. The final result is represented by the so-called 'chemical probe', which can be an inhibitor, activator or modulator of the studied target.

VS methods have been comprehensively reviewed [20] (see I. Muegge, this issue). Therefore, only a brief description is given here. Our VS workflow is shown in Fig. 3. There are two main types of VS: (i) target or structure-based virtual screening (TBVS, or SBVS); and (ii) LBVS.
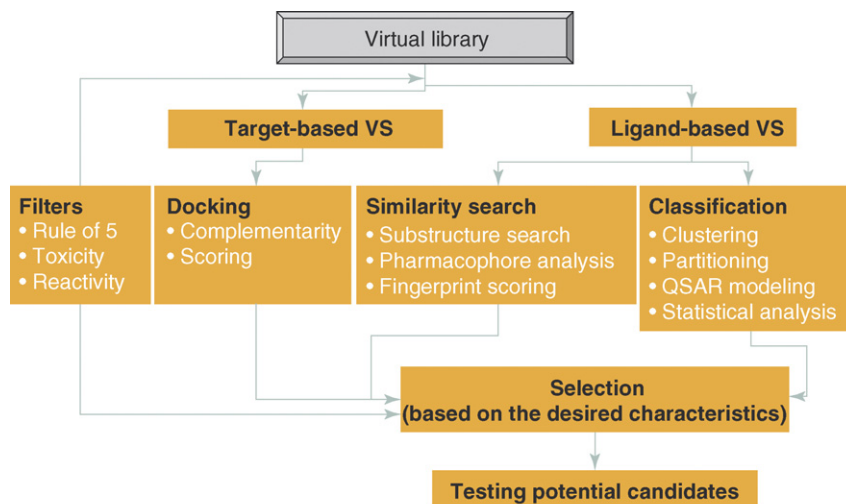
TBVS is applied when the target (protein, enzyme) structure is (i) known, based on the crystallographic or other structural methods, or (ii) built, using homology modeling. The procedure consists of docking a large number of molecules into the active site of the target, then scoring the binding affinity. The limiting step in this strategy is accurate scoring. There are four categories of scoring functions:

(i) Knowledge-based methods [21] based on Boltzmann-weighted Potentials of Mean Force (PMF) derived from statistical analyses of ligand–receptor interatomic contacts from Protein Data Bank (PDB: http://www.rcsb.org/) complexes.

(ii) 'Master equation' approaches [22] estimate the energetic contributions of various interaction types in a semi-quantitative manner.

(iii) Regression-based methods [23,24] train on bioactives with known binding mode using sets of ligand–receptor complexes from the PDB.

(iv) ZAP is a Poisson-Boltzmann equation solver [25] that scores ligands while incorporating solvent effects.

Figure 2. Workflow. The strategy implemented at the New Mexico Molecular Libraries Screening Center for the integration of virtual and physical screening. Abbreviations: QC: qualitative chemistry; SAR, structure–activity relationships.

Figure 3. Virtual library. An example of virtual screening (VS) workflow as implemented at the University of New Mexico, Department of Biochemistry and Molecular Biology, Division of Biocomputing (http://biocomp.health.unm.edu/).

Because none of the above schemes has been shown to be general, it is preferable to use consensus-scoring methods [26].

LBVS methods start from known bioactives, which can be inhibitors or activators. Similarity search [27] and classification [28] methods can be used to select novel scaffolds (see K. V. Balakin, this issue). Filtering can be applied (i) forward or (ii) backward to the hits obtained in the VS campaign. In forward filtering [29], the selection criteria are used to reduce the size of the initial library, that is from several millions to several hundreds or hundreds of test compounds further to be docked. The backward procedure uses filtering criteria to prioritize candidates for PS during post-HTS analysis [30].

Testing strategies for the selected compounds can include does–response characterization and cytotoxicity of the selected compounds. In our Center, many of the assays are being set up as high throughput multiplexes so that selectivity and specificity information against families of targets is available for individual compounds. The physical properties of the components to be tested contribute to the screening strategy because the presence of molecules which are insoluble, are aggregators, or are fluorescent, can interfere with the identification of their activities in the assays.

When screening 100,000 to more than 1,000,000 compound libraries, it is paramount that the assay be optimized for complexity/simplicity (minimal additions, no wash steps, simple detection schemes), volume (typically no more than 10 μL, permitting use of 384 or 1536 well plates) and reproducibility (because the library is only screened once or twice). However, if VS technologies can reduce the library size to 10's or 100's of compounds, many of these constraints are lifted allowing the use of more physiologically relevant or complex assays, potentially including whole animal-based assays, which are recognized as a viable alternative in today's search for novel pharmaceuticals. Another issue related to assay implementation is the selection of *in vitro* vs. *in vivo* assays. Many of the initial hits in enzyme/protein-based *in vitro* assays may fail at the cellular or animal level due to stability or metabolism issues. However, with the development of VS and post-HTS analysis methods, this can be minimized. Finally, assay selection has a strong influence on the outcome of the physical screen. For example, the use of a simple binding assay screen can yield both agonists and antagonists of target activity. If only agonists are desired, a functional assay based on target activity might be preferentially selected.

Post-HTS analysis

The results of both VS and PS can yield a large number (often in the order of 100–1000) of interesting hits that warrant further attention. The issue of what molecules to select for dose–response confirmation is often left with the medicinal chemist or the biologist, and requires cheminformatics support. In our center, we apply the following post-HTS

Box 1. Evaluation criteria to select molecules for chemistry follow-up

(i) **Chemotype evaluation**: This criterion gives higher priority (a) to chemotypes that occur more in active compounds, compared to the overall number of tested chemotypes; (b) to chemotypes that are absent or less present in patents, disclosures and medicinal chemistry (for intellectual property reasons); and (c) to chemotypes that are generally regarded as safe, or are less frequent in toxicity databases.

Chemotype evaluation is used to rank families of HTS hits, and is further applied to (ii) **Individual molecule evaluation**: Our scheme gives higher priority (a) to molecules that are in the desired physico-chemical property range (using methods to compute, e.g. solubility and permeability); (b) to molecules that have high(er) activity, compared to those that are less active. To the above, (c) we add the chemotype score computed earlier, for a final composite score that ranks all confirmed or presumed actives.

prioritization scheme [30] for chemotype as well as individual molecule evaluation (See Box 1).

Thus, post-HTS analysis is essentially a practical step designed to assist the decision-makers to evaluate compounds for further experiments. It should be applied only to confirmed hits, both at the structure and purity, as well as at the dose–response level. The final score captures information related not only to actives, but also inactives from the same chemical family, while the intellectual property and toxicity evaluation schemes are aimed at encoding information related to individual chemotypes. At the individual molecule level, the use of estimated physico-chemical properties can assist the final prioritization score.

Summary and outlook

The integration of VS and PS technologies is attractive for both scientific and economic reasons. Scientifically, one can reduce the search space to rapidly find a solution; economically, one is no longer required to screen millions of compounds physically, before identifying hits. Negative aspects of this integration effort are of theoretical and practical nature. Theoretically, the streamlining of PS may result in testing the wrong library subset. This relates to both the limitation of theoretical methods, and to the inappropriate use of VS technologies. Practically speaking, it is possible that the entire effort leads to naught; in this case, doubt is typically cast over *in silico* approaches, although experimental techniques are not without flaw either. Continuous information exchange and effective team communication, as illustrated in Fig. 2, will avoid such negative situations.

Integration of these technologies is further supported by the increasing amount of valuable information being deposited in target and bioactivity databases (see T.I. Oprea, this issue). Conceivably, one can apply existing information to improve the success rate by using, for example, machine learning techniques to develop target-specific libraries; or to include target-specific or ligand-specific information in

the post-HTS analysis process, to give higher priority to high-quality probes; or perhaps to download the entire matrix of target/bioactivity data and use it to profile compounds.

Acknowledgements

References

1 Kubinyi, H. (1999) Chance favors the prepared mind – from serendipity to rational drug design. *J. Recept. Signal Tr. R.* 19, 15–39

2 Fox, S. *et al.* (2002) High throughput screening 2002: moving toward increased success rates. *J. Biomol. Screen.* 7, 313–316

3 Caron, P.R. *et al.* (2001) Chemogenomic approaches to drug discovery. *Curr. Opin. Chem. Biol.* 5, 464–470

4 Sills, M.A. *et al.* (2002) Comparison of assay technologies for a tyrosine kinase assay generates different results in high throughput screening. *J. Biomol. Screen.* 7, 191–214

5 Bajorath, J. (2002) Integration of virtual and high-throughput screening. *Nature Rev. Drug. Discov.* 1, 882–894

6 Toledo-Sherman, L.M. and Chen, D. (2002) High-throughput virtual screening for drug discovery in parallel. *Curr. Opin. Drug Discov. Develop.* 5, 414–421

7 Walters, W.P. *et al.* (1998) Virtual screening – an overview. *Drug Discov. Today* 3, 160–178

8 Oprea, T.I. (2002) Virtual screening in lead discovery: a viewpoint. *Molecules* 7, 51–62

9 Mestres, J. (2002) Virtual screening: a real screening complement to high-throughput screening. *Biochem. Soc. Trans.* 30, 797–799

10 Beresford, A.P. *et al.* (2002) The emerging importance of predictive ADME simulation in drug discovery. *Drug Discov. Today* 7, 109–116

11 Johnson, D.E. and Wolfgang, G.H.I. (2000) Predicting human safety: screening and computational approaches. *Drug Discov. Today* 5, 445–454

12 Bajorath, J. (2001) Rational drug discovery revisited: interfacing experimental programs with bio- and chemo-informatics. *Drug Discov. Today* 6, 989–995

13 Terstappen, G.C. and Reggiani, A. (2001) *In silico* research in drug discovery. *Trends Pharmacol. Sci.* 22, 23–26

14 Drews, J. (2000) Drug discovery: a historical perspective. *Science* 287, 1960–1964

15 Leach, A.R. and Hann, M.M. (2000) The *in silico* world of virtual libraries. *Drug Discov. Today* 5, 326–336

16 Hopkins, A.L. and Groom, C.R. (2002) The druggable genome. *Nat. Rev. Drug Discov.* 1, 727–730

17 Stevens, R.C. *et al.* (2001) Global efforts in structural genomics. *Science* 294, 89–92

18 Edwards, B.S. *et al.* (2005) Integration of virtual screening with high-throughput flow cytometry to identify novel small molecule formylpeptide receptor antagonists. *Mol. Pharmacol.* 68, 1301–1310

19 Bologa, C.G. *et al.* (2006) Virtual and biomolecular screening converge on a selective agonist for GPR30. *Nature Chem. Biol.* 2, 207–212

20 Klebe, G. (2006) Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today* 11, 580–594

21 Sippl, M.J. (1995) Knowledge-based potentials for proteins. *Curr. Opin. Struct. Biol.* 5, 229–235

22 Ajay, and Murcko, M.A. (1995) Computational methods to predict binding free energy in ligand-receptor complexes. *J. Med. Chem.* 38, 4953–4967

23 Bohm, H.J. (1994) The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three-dimensional structure. *J. Comput. Aid. Mol. Des.* 8, 243–256

24 Head, R.D. *et al.* (1996) VALIDATE: A new method for the receptor-based prediction of binding affinities of novel ligands. *J. Am. Chem. Soc.* 118, 3959–3969

25 Grant, J.A. *et al.* (2001) A smooth permittivity function for Poisson-Boltzmann solvation methods. *J. Comp. Chem.* 22, 608–640

26 Feher, M. (2006) Consensus scoring for protein–ligand interactions. *Drug Discov. Today* 11, 421–428

27 Hert, J. *et al.* (2006) New methods for similarity-based virtual screening. *Chemogenomics* 133–156

28 Schneider, G. and Nettekoven, M. (2003) Ligand-based combinatorial design of selective purinergic receptor (A2A) antagonists using self-organizing maps. *J. Comb. Chem.* 5, 233–237

29 Oprea, T.I. (2000) Property distribution of drug-related chemical databases. *J. Comput. Aid. Mol. Des.* 14, 251–264

30 Oprea, T.I. *et al.* (2005) Post-HTS compound prioritization analysis: An empirical evaluation scheme. *J. Biomol. Screen.* 10, 419–425

31 Kravchencko, D.V. *et al.* (2006) Design and synthesis of new nonpeptide caspase-3 inhibitors. *Pharma. Chem. J.* 40, 127–131

32 Liu, D. *et al.* (2006) Inhibitor discovery targeting the intermediate structure of β-amyloid peptide on the conformational transition pathway: implications in the aggregation mechanism of β-amyloid peptide. *Biochemistry* 45, 10963–10972

33 Lu, I.-L. *et al.* (2006) Structure-based drug design and structural biology study of novel nonpeptide inhibitors of severe acute respiratory syndrome coronavirus main protease. *J. Med. Chem.* 49, 5154–5161

34 Jankun, J. *et al.* (2006) Synthetic curcuminoids modulate the arachidonic acid metabolism of human platelet 12-lipoxygenase and reduce sprout formation of human endothelial cells. *Mol. Cancer Ther.* 5, 1371–1382

35 Saario, S.M. *et al.* (2006) Fatty acid amide hydrolase inhibitors from virtual screening of the endocannabinoid system. *J. Med. Chem.* 49, 4650–4656

36 Lu, Y. *et al.* (2006) Discovery of a nanomolar inhibitor of the human murine double minute 2(MDM2)-p53 interaction through an integrated, virtual database screening strategy. *J. Med. Chem.* 49, 3759–3762

37 Schuster, D. *et al.* (2006) The discovery of new 11β-hydroxysteroid dehydrogenase type 1 inhibitors by common feature pharmacophore modeling and virtual screening. *J. Med. Chem.* 49, 3454–3466

38 Li, J. *et al.* (2006) Discovering novel chemical inhibitors of human cyclophilin A: virtual screening, synthesis, and bioassay. *Bioorg. & Med. Chem.* 14, 2209–2224

39 Guichou, J.-F. *et al.* (2006) Structure-based design, synthesis, and biological evaluation of novel inhibitors of human cyclophilin A. *J. Med. Chem.* 49, 900–910

40 Desai, P.V. *et al.* (2006) Identification of novel parasitic cysteine protease inhibitors by use of virtual screening. 2. The Available Chemical Directory. *J. Med. Chem.* 49, 1576–1584

41 Kenyon, V. *et al.* (2006) Novel human lipoxygenase inhibitors discovered using virtual screening with homology models. *J. Med. Chem.* 49, 1356–1363

42 Li, R. *et al.* (2006) Design and synthesis of 5-aryl-pyridone-carboxamides as inhibitors of anaplastic lymphoma kinase. *J. Med. Chem.* 49, 1006–1015

43 Rella, M. *et al.* (2006) Structure-based pharmacophore design and virtual screening for novel angiotensin converting enzyme 2 inhibitors. *J. Chem. Inf. Model.* 46, 708–716

44 Polgár, T. *et al.* (2005) Comparative virtual and experimental high-throughput screening for glycogen synthase kinase-3β inhibitors. *J. Med. Chem.* 48, 7946–7959

45 Ward, R.A. *et al.* (2005) Structure-based virtual screening for low molecular weight chemical starting points for dipeptidyl peptidase IV inhibitors. *J. Med. Chem.* 48, 6991–6996

46 Franke, L. *et al.* (2005) Extraction and visualization of potential pharmacophore points using support vector machines: application to ligand-based virtual screening for COX-2 inhibitors. *J. Med. Chem.* 48, 6997–7004

47 Rollinger, J.M. *et al.* (2005) Application of the *in combo* screening approach for the discovery of non-alkaloid acetylcholinesterase inhibitors from *Cichorium intybus*. *Curr. Drug. Discov. Technol.* 2, 185–193

48 Steindl, T.M. *et al.* (2005) Pharmacophore modeling, docking, and principal component analysis based clustering: combined computer-assisted approaches to identify new inhibitors of the human rhinovirus coat protein. *J. Med. Chem.* 48, 6250–6260

49 Song, H. *et al.* (2005) A low-molecular-weight compound discovered through virtual database screening inhibits Stat3 function in breast cancer cells. *Proc. Natl. Acad. Sci. U S A* 102, 4700–4705

50 Toledo-Sherman, L. *et al.* (2005) Frontal affinity chromatography with MS detection of EphB2 tyrosine kinase receptor. 2. Identification of small molecule inhibitors via coupling with virtual screening. *J. Med. Chem.* 48, 3221–3230

51 Evers, A. and Klabunde, T. (2005) Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the Alpha1A adrenergic receptor. *J. Med. Chem.* 48, 1088–1097