

Large-scale virtual screening for discovering leads in the postgenomic era

by B. Waszkowycz
T. D. J. Perkins
R. A. Sykes
J. Li

Virtual screening, or in silico screening, is a new approach attracting increasing levels of interest in the pharmaceutical industry as a productive and cost-effective technology in the search for novel lead compounds. Although the principles involved—the computational analysis of chemical databases to identify compounds appropriate for a given biological receptor—have been pursued for several years in molecular modeling groups, the availability of inexpensive high-performance computing platforms has transformed the process so that increasingly complex and more accurate analyses can be performed on very large data sets. The virtual screening technology of Protherics Molecular Design Ltd. is based on its integrated software environment for receptor-based drug design, called Prometheus. In particular, molecular docking is used to predict the binding modes and binding affinities of every compound in the data set to a given biological receptor. This method represents a very detailed and relevant basis for prioritizing compounds for biological screening. This paper discusses the broader scope of virtual screening and, as an example, describes our recent work in docking one million compounds into the estrogen hormone receptor in order to highlight the technical feasibility of performing very large-scale virtual screening as a route to identifying novel drug leads.

Drug discovery has traditionally made progress by a combination of random screening and rational design.¹ In practice, the latter approach is often frustrated by a paucity of experimental data that define the structure and properties of the biological target. Today, this situation is starting to change, with initiatives such as the Human Genome Project promising access to a vast amount of data detailing the molecular basis of disease.² In parallel we have wit-

nessed the impact of automated laboratory systems in the fields of chemical synthesis, biological assay, drug metabolism, and even protein crystallography, such that high-throughput strategies have come to dominate many areas of drug discovery research.³ Despite these advances (and indeed partly because of them), drug discovery remains a difficult task, but at least we are now gathering together all the pieces of the puzzle, in terms of information, technologies, and expertise.

This paper describes a computational strategy that is beginning to influence traditional routes to drug discovery. *Virtual screening*⁴ is the use of high-performance computing to analyze large databases of chemical compounds in order to identify possible drug candidates, and is a technology that complements current advances in high-throughput chemical synthesis and biological assay. This paper surveys methods and applications of virtual screening and describes its potential for supporting the drug discovery process. In particular, we examine the value of molecular docking as a computational approach to achieve fast and accurate filtering of large data sets and describe as an example a virtual screen against the estrogen receptor (the DockCrunch project).⁵

©Copyright 2001 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

High-throughput screening in drug discovery

For those engaged in drug design, such as medicinal and computational chemists, the research phase can be broken down into two main tasks: identification of new compounds showing some activity against a target biological receptor, and the progressive optimization of these leads to yield a compound with improved potency and physicochemical properties *in vitro*, and, eventually, improved efficacy, pharmacokinetics, and toxicological profiles *in vivo*.

Identification of leads is driven either by random screening or a directed design approach, and traditionally both strategies have been of equal importance, depending on the problem in hand. The directed approach needs a rational starting point for medicinal chemists and molecular modelers to exploit. Examples include the design of analogs of a drug known to be active against a target receptor and mimics of the natural substrate of an enzyme. Increasingly, the three-dimensional structure of many biological targets is being revealed by X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, opening the way to the design of novel molecules that directly exploit the structural characteristics of the receptor binding site. In recent years, this approach of *structure-based design* has had a major impact on the rational design and optimization of new lead compounds in those cases where the receptor structure is well-characterized.⁶⁻⁸

An alternative approach (and one that is essential when no useful starting information is available) is random screening of compound collections. The compounds may be sourced from the in-house database of a pharmaceutical company, or a collection of natural products, or the catalog of a chemical supplier. In the past, this approach represented a moderately fruitful if somewhat long-winded strategy—if you are searching for compounds at random, you may have to assay a very large number of compounds before you find an interesting hit. However, over recent years this strategy has played an increasingly important role because of improvements in high-throughput screening (HTS) technology, and for many pharmaceutical companies, HTS is now an essential component to identify leads.³ Automated compound handling and assaying facilities now perform tasks such as retrieval of compounds from storage, dilution and plating out of samples, and the running of assays of ever-increasing complexity. Such automated systems can handle tens of thousands of compounds per day. At this scale, random screen-

ing has the potential to identify useful numbers of hits on a timely basis.

However, HTS is no guarantee of success, and the problems associated with over-reliance on random HTS are becoming apparent.⁹ Establishing a robust assay for a new target takes time and money. Hit rates against some receptors are reported to be very low, necessitating screening of very large numbers of compounds (tens to hundreds of thousands). Collections of synthesized compounds or natural products often contain far less chemical diversity than is desired, are not bottomless resources, and are very time-consuming to replenish. Techniques such as combinatorial chemistry offer the potential for synthesizing very large libraries of compounds, but in practice this approach is time-consuming for drug-like compounds and may still produce libraries of relatively restricted diversity.

To make matters worse, recent advances in genomics will vastly increase the number of potential therapeutic targets available for research. Although this increase will undoubtedly yield benefits in the long term for the discovery of novel therapeutics, in the short term it poses the problem of how to validate this array of biological targets and how to identify useful lead compounds in an efficient and timely manner.¹⁰

Given these caveats, it is worth evaluating the role of technologies that may complement high-throughput assay and high-throughput synthesis. The term *virtual screening* has been used to describe a process of computationally analyzing large compound collections in order to prioritize compounds for synthesis or assay.⁴ A broad range of computational techniques can be applied to the problem. In our work we have focused on explicit receptor-ligand molecular docking as a means of yielding the most detailed model of the way in which a given ligand will bind to a receptor, and hence the most informative basis on which to assess which ligands are useful candidates for synthesis or assay.

Although the underlying methods of virtual screening have been in use in various guises for several years, it is worth noting the recent impact on molecular modeling made by the increased availability of high-performance computing platforms. Affordable multiprocessor workstations and PC (personal computer) clusters have enabled the modeler to employ computationally demanding algorithms on a routine basis. This change is particularly relevant in

the case of virtual screening, where, as in the work described in this paper, computationally intensive methods such as molecular docking must be applied to very large databases of chemical structures. Indeed, we are now witnessing the expansion and routine application of high-throughput modeling techniques in a way that parallels the rapid progress made over the last decade in high-throughput chemical synthesis and pharmacological screening.

Virtual screening— concepts and feasibility

Virtual screening is a strategy for bringing a more focused approach to HTS by using computational analysis to select a subset of compounds considered to be appropriate for a given receptor. Clearly, this strategy implies that some information is available regarding either the nature of the receptor binding site or the type of ligand that is expected to bind productively, or both. It should be stressed that virtual screening encompasses a variety of computational screens, from the simplistic to the sophisticated, and, hence, can usefully exploit different types of information describing the receptor. Likewise it can be used to produce either a very focused subset of compounds (for example, if only close structural analogs of a lead compound are of interest) or a very open-ended subset (for example, a restriction on size, as described by molecular weight, may be the only constraint applied). Such issues relate to the objectives of a particular research project. For example, an HTS program may be already up and running, and the problem to be addressed is which set of 100 000 compounds from a database of one million should be assayed first. At the other end of the scale, a newly discovered receptor may have a low-throughput binding assay, and a far more focused set of compounds needs to be selected.

In theory, the applicability of virtual screening is limited only by what properties of a compound can be calculated computationally and the perceived relevance of those properties to the problem in hand. On a practical level, further considerations include:

- The timescale for calculation of the properties, which may be considerable for a database of, say, one million compounds
- The accuracy or meaningfulness of the properties, particularly so when computationally cheap methods are applied
- Methods for analysis of the data—not a trivial problem, since high-throughput modeling, as with HTS, generates very large volumes of data

- The software and hardware required to yield a timely answer

The main computational filters that are available to virtual screening (in terms of increasing sophistication and computational cost) can be summarized as:

- Selection on the basis of two-dimensional (2-D) property profiles
- Selection by means of a target-specific pharmacophore
- Selection on the basis of more detailed three-dimensional (3-D) modeling: e.g., receptor–ligand docking

These methods are now described.

Selection on the basis of 2-D property profiles.

Regardless of whether any useful information is available to describe the receptor or known ligands, it is often preferable to limit HTS to “drug-like” compounds. “Drug likeness” is used to indicate a broad range of properties or structural features that are generally important in various stages of drug optimization, such as stability, solubility, and lipophilicity, which can all influence drug absorption and excretion. Selection of drug-like compounds is particularly important when compounds are sourced from suppliers’ catalogs (where many compounds are more reagent-like than drug-like) or from a combinatorial library. The current trend in the pharmaceutical industry is for pharmacokinetic profiles and toxicological profiles to be evaluated at earlier stages in the drug discovery process. Therefore, it is useful to bring some of these concepts into the virtual screening stage, while accepting that the calculable properties currently available are at best only vague indicators of metabolic fate or toxicity.

Lipinski’s “rule-of-five” is a well-known rule-of-thumb that encodes a simple profile for orally bioavailable compounds, basing the classification on a limit on molecular weight, lipophilicity (in terms of the partition coefficient, $\log P$), and hydrophilicity (in terms of counts of hydrogen bond donors and acceptors).¹¹ These properties can be calculated quickly and can be easily applied to filtering a large database. Likewise, filters can be applied on specific chemical substructures, e.g., those associated with problems in chemical stability or toxicity. Examples of this type of filtering process are the REOS (rapid elimination of swill) procedure described by scientists from Vertex Pharmaceuticals as a way of screening out undesirable properties at the start of the pro-

cedure to reduce the number of compounds to be submitted to more intensive computational evaluation.⁴ Other approaches to define drug likeness exploit techniques such as neural networks and genetic algorithms to analyze property profiles in order to construct a scoring or classification scheme that is capable of differentiating known drugs from non-drugs.¹²

In a further stage of this approach, there are examples where it is possible to define 2-D property profiles that are believed to bias the selection toward a particular receptor. For example, analysis of known estrogen receptor agonists may bias the selection toward a specific molecular weight range and a lipophilicity somewhat higher than most drug-like compounds. Such property profiles help to reduce the size of the screening set (sometimes drastically), although smaller size is achieved at the risk of reduced novelty. This approach is useful if the limitations are borne in mind and is particularly appropriate for dividing and prioritizing a very large database into subsets that fulfill different design criteria.

Selection by means of a target-specific pharmacophore. A pharmacophore is a simplified 3-D description of the key structural features of a set of known ligands or of the target receptor. The structural features are usually described in terms of discrete hydrogen bond donors or acceptors, lipophilic centers, ring centroids, and so on, separated in terms of distances (more usually, distance ranges). Usually a small number of such sites is defined, e.g., two to five. Typically these sites are derived from a set of ligands and, hence, represent those features, common to the ligands, that are deemed to be relevant to activity. For example, classical thrombin inhibitors are characterized by a cationic center (binding to the S1 specificity pocket) 9–11 Å (angstroms) away from a lipophilic center (binding to either the S2 or S4 lipophilic pockets). A pharmacophore is readily used to search a database of chemical structures. These structures need to contain 3-D models, and preferably a conformationally flexible search is necessary so that compounds are not rejected on the trivial basis that an inappropriate conformer is stored in the database.

The value of the pharmacophore search is that a reasonably focused query on 3-D structural grounds can be applied relatively quickly to a large database. The limitations are that the query may rapidly become over-defined. It is often the case that a typical three-

or four-point pharmacophore will be too restrictive and yield few interesting hits, whereas a slightly more open-ended query may yield too many hits. Either way, the problem with managing the hit list is that there is usually no effective way of ranking or scoring the hits, other than methods based on evaluation of similarity to the initial ligands. Therefore, there is a tendency for pharmacophore searches to yield solutions similar to those already known, rather than a well-focused set of novel solutions.

Selection by means of receptor–ligand docking. The next stage up from a pharmacophore search in terms of computational expense is explicit docking of the compound database to the biological target of interest. This stage involves objective docking of each compound (either as a rigid or conformationally flexible model) into a model of the receptor that is treated either as rigid or with limited side-chain flexibility. Generally the extents of the expected binding site are defined to limit the search. This virtual screening strategy requires a 3-D database of ligands, a 3-D structure of the target receptor (either derived experimentally or from a model built by homology to related protein structures), and a docking code comprising an efficient searching algorithm and an accurate scoring function. These features will be described in more detail in a subsequent section.

The attraction of explicit receptor–ligand docking is that it represents the most detailed and relevant computational model for identifying a receptor-focused subset. In addition, it is also one of the least-biased approaches. Application of pharmacophore queries or focused 2-D property profiles may significantly inhibit the diversity of the compound subset because they are biased by the properties of known ligands. In contrast, the molecular docking program can process an entire chemical database with minimal prefiltering (e.g., to eliminate unstable or toxic moieties) so that the final selection is based on the quality of the docked models rather than a subjective opinion of what properties are expected in a ligand. This route is a very promising one to finding structurally novel ligands, which may make receptor interactions similar to known ligands or may achieve different interactions within other binding sites.

The output of a docking-based screen is a set of 3-D models of the predicted binding mode of each compound against the receptor, together with a ranking that is a measure of the quality of fit, if not a prediction of the binding affinity itself. The 3-D models

represent the most detailed basis available for determining which molecules are capable of fitting within the very strict structural constraints of the receptor binding site—such a degree of discrimination is not possible using a pharmacophore model, because available searching software cannot handle queries with the number of points needed to adequately represent the complexities of the binding site.

**To perform virtual screening
on a large scale presents a
number of technical
challenges.**

Besides serving as structural filters, the 3-D models are an important basis for scoring the hits, such that a quantitative ranking is achievable using various scoring methods.¹³ These methods are typically empirically derived functions for estimation of free energy of binding, usually calibrated against a data set of heterogeneous receptor–ligand complexes, as determined by X-ray crystallography. For a particular receptor, ligand binding modes may be further classified with respect to the specific receptor contacts that are achieved. Thus, the 3-D models represent a very valuable source of data for understanding the nature of ligand binding in a given receptor and, hence, as a source of inspiration for the design of analogs or indeed novel ligands.

The potential problems associated with molecular docking as a virtual screening strategy are the quality of the docking results and the technical feasibility of processing large databases. For example, there is an issue about the accuracy of the predicted binding modes and energies, particularly when using fast docking protocols to process large data sets: are the methods capable of identifying true hits and rejecting inactive compounds? Although any high-throughput technology must sacrifice some degree of accuracy, there is now enough experience with various docking packages to give confidence that they are *accurate enough* to achieve a meaningful focusing of the database. It should be remembered that the goal of docking is not to pick out conclusively the handful of expected hits, but rather to pick a subset of compounds (perhaps 1 to 10 percent of the database) that will contain significantly more hits than a randomly chosen set.¹³

More of a concern is the computational expense to dock and analyze a large database. For virtual screening to be a useful sister technology to HTS, it should be capable of processing data sets of at least the order of a million compounds. (This amount represents roughly the number of commercially available drug-like molecules or the historic collection of a pharmaceutical company and is a large underestimate of the numbers potentially available via combinatorial synthetic routes.) Molecular docking is the most computationally expensive of all the virtual screening methods discussed in this paper. However, advances in the efficiency of algorithms and hardware have now reduced the computational cost typically to the order of one minute of CPU time per ligand per processor. Some docking programs are appreciably faster, mainly those that treat ligands as conformationally rigid, but they run the risk of missing hits because of an inadequate representation of conformational space. Even accepting the computational cost of flexible ligand docking, analysis of very large databases on reasonably priced supercomputers (or, increasingly, on clusters of PCs) is now a technically feasible task. Over the last year, a number of workers have published virtual screening studies using a variety of docking programs (Table 1^{14–22}), demonstrating the level of interest in this approach and its broad applicability.

To perform virtual screening on a large scale presents a number of technical challenges. Not only are there the more obvious questions of how quickly and how accurately a large database can be evaluated, but there is also the need to automate the procedure, including preprocessing of the chemical database and postprocessing of the results. For virtual screening to become a technology of day-to-day applicability and ease-of-use, it is essential that modelers are provided with the tools to support and facilitate each stage. To demonstrate this point, we recently completed the DockCrunch project⁵ (in collaboration with Silicon Graphics Inc. [SGI]) as a worked example of virtual screening in an industrial setting, and the results of this exercise will form the basis for a more detailed examination of the virtual screening process in the following section.

DockCrunch—a feasibility study on large-scale virtual screening

The DockCrunch project involved the virtual screening of a large database of commercially available compounds against the estrogen receptor. The main objective of the study was to demonstrate the tech-

Table 1 Summary of recently published virtual screening studies based on receptor–ligand docking

Target	Software Package	Ligand Data Set	Reference
Thymidylate synthase	DOCK	153 000 ACD compounds	14
FK506-binding protein	SANDOCK	ACD and Cambridge Crystallographic Database	15
Retinoic acid receptor	ICM	153 000 ACD compounds	16
HIV-1 RNA transactivation response element	ICM	153 000 ACD compounds	17
Farnesyl transferase	EUDOC	67 928 ACD compounds	18
DNA gyrase	LUDI, CATALYST	350 000 ACD + in-house compounds	19
Kinesin	DOCK	110 000 ACD compounds	20
Hypoxanthine-guanine-xanthine phosphoribosyl transferase	DOCK	599 compound virtual library	21
Thrombin; factor Xa; estrogen receptor	PRO_LEADS	10 000 ChemBridge compounds	22
Estrogen receptor (agonist and antagonist forms)	PRO_LEADS	1.1 million ACD-SC compounds	5

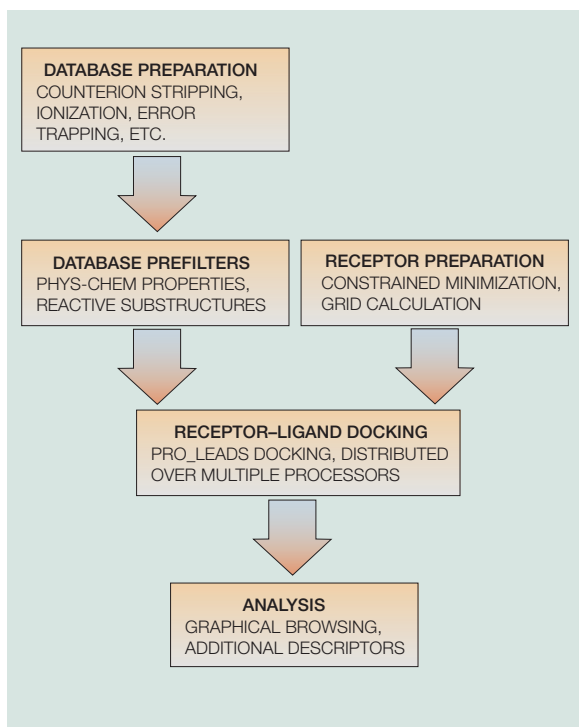
nical feasibility of performing a virtual screening exercise on such a large scale, using a computationally intensive approach (receptor–ligand docking) to analyze a database of over one million compounds. The project served as a test case to address the many technical issues that arise when implementing a high-throughput virtual screening program, such as the development of tools for automated preprocessing and filtering of chemical databases, parallelization of the docking simulation on a multiprocessor machine, and visualization tools for analysis of the large body of results. The various modeling tasks are executed within a proprietary software environment (Prometheus). In this way the majority of the tasks can be performed within an integrated platform rather than by accessing a number of disparate software packages, which is an important practical consideration for a molecular modeling team. (Commercial software is used for specialized tasks such as structural refinement of the receptor and 2-D-to-3-D conversion of the compound database.) An overview of the process is depicted in Figure 1.

Completing the project in an appropriate timescale was an important objective in order to demonstrate that such a computationally intensive strategy is a

relevant adjunct to a modern HTS program. However, equally important was to validate the accuracy of the method by demonstrating that the docking and scoring strategy was able to retrieve known ligands seeded in the random data set, and that novel high-scoring compounds identified by the docking were in fact potent ligands.

Data set definition and preparation. We chose the estrogen receptor as an example of a well-understood target of therapeutic relevance for which crystallographic structural data were publicly available. The estrogen receptor is an example of the family of *nuclear hormone receptors*.²³ The binding of an *agonist* (i.e., a ligand capable of activating the receptor) induces a particular fold in the vicinity of the binding site that promotes association with other recognition proteins; the receptor–ligand complex is then transported into the cell nucleus, binds to a specific sequence of DNA (deoxyribonucleic acid) and thereby modulates gene expression. Other ligands that bind to the receptor elicit no such response and are classed as *antagonists*. From X-ray crystallography it is observed that the 3-D fold of the receptor is somewhat different upon binding an antagonist. The critical helix that moves to bind snugly around a small agonist

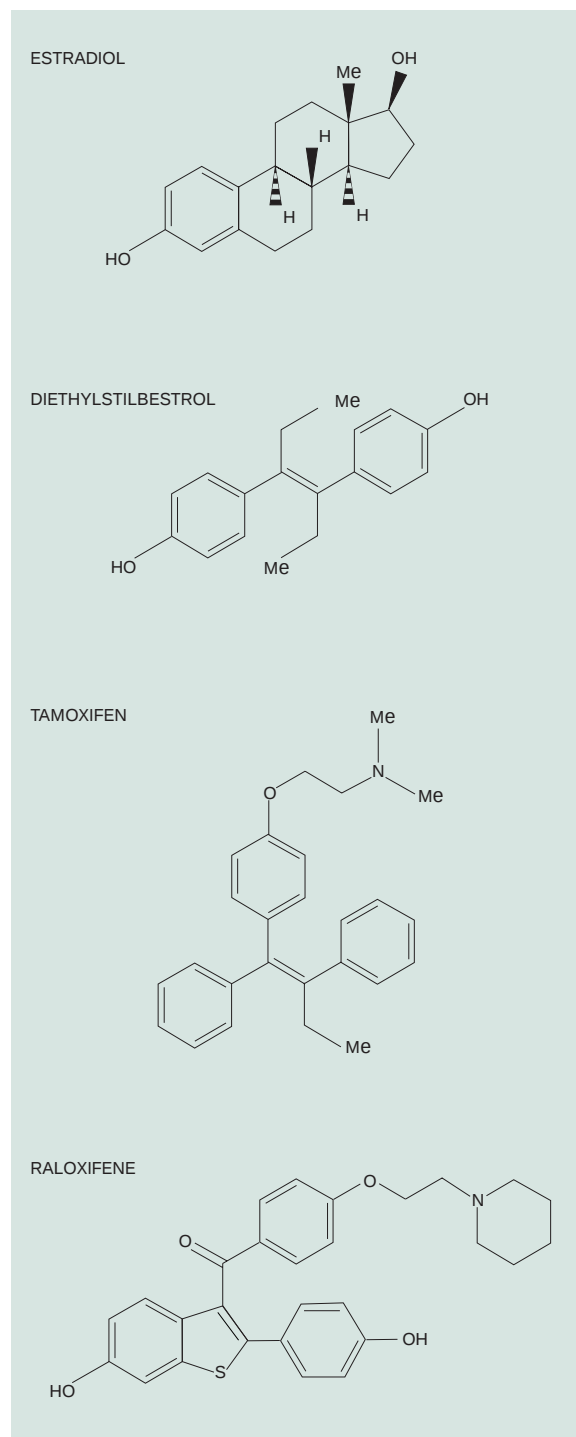
Figure 1 Summary of virtual screening protocol as implemented in Prometheus



ligand is forced away by the characteristic large side-arm structure that distinguishes antagonist compounds (such as tamoxifen or raloxifene) from agonists (e.g., estradiol) (Figures 2 and 3).²⁴

Given that the agonist and antagonist forms of the receptor exhibit some key structural differences, the DockCrunch project involved evaluation of the compound database against both forms of the receptor in order to explore whether the virtual screening strategy was capable of distinguishing potential antagonists from agonists. (Note that the terms agonist and antagonist as used here reflect the gross structural differences between the ligands. In a pharmacological sense, agonism and antagonism are more accurately functionally defined with reference to effects observed within a given animal or tissue model. *Antagonist* is used here to describe ligands that may actually exhibit partial agonist activity and are more correctly referred to as *selective receptor modulators* in the pharmacological literature.²⁵) The crystallographic structures used were those of the human α -estrogen receptor complexed with estradiol (Brookhaven code 1ERE) and with raloxifene

Figure 2 Examples of structures of estrogenic compounds within the reference set: estradiol and diethylstilbestrol (agonists), raloxifene and tamoxifen (antagonists)



(code 1ERR).²⁴ The crystallographic structures required minimal refinement before docking: i.e., the addition of hydrogen atoms and a constrained geometry refinement of the residues forming the main binding pocket, i.e., residues within 15 Å of the ligand (using InsightII software of Molecular Simulations Inc. [MSI]).²⁶

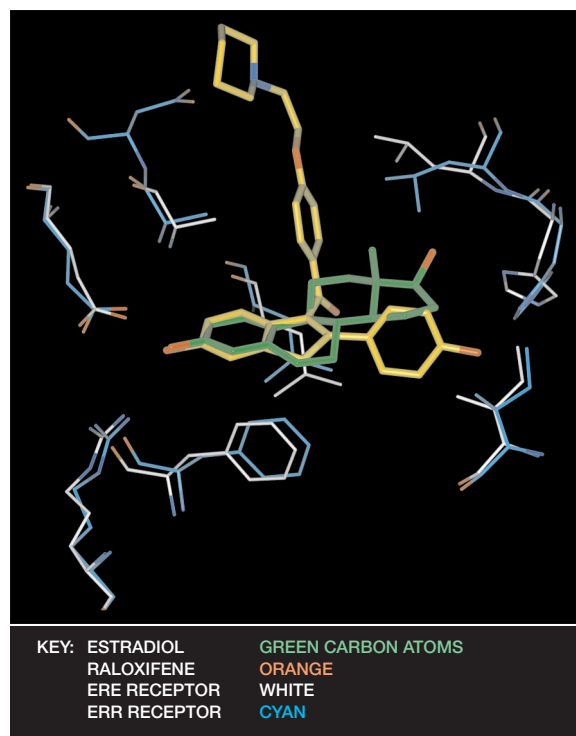
For the compound data set we chose the Available Chemicals Directory-Screening Compounds (ACD-SC) from Molecular Design Ltd. (MDL).²⁷ This represents over 1.5 million commercially available compounds collected from the major chemical suppliers and, as such, is a suitable starting point for a virtual screening project. Prior to docking, the data set was processed as follows⁵ (unless otherwise stated, all the database processing tasks were performed with our in-house software Prometheus):

- Database preprocessing:
 - Correction or standardization of the 2-D structural representations
 - Removal of counter-ions
 - Protonation of acids or bases as predicted at physiological pH
 - Calculation of physicochemical properties (e.g., molecular weight, log *P*, hydrogen bond counts, flexibility, counts of chiral centers, etc.)
- 2-D property filters:
 - Removal of inorganic compounds
 - Removal of compounds displaying nondrug-like physicochemical properties, e.g., molecular weight <200 or >600, calculated log *P* > 7, number of hydrogen bond donors >8 or acceptors >8
- Conversion into 3-D using the MSI Converter:
 - One conformer per molecule, plus additional enantiomers for undefined chiral centers

The end result of this task was to yield a data set of 3-D structures for 1.1 million drug-like compounds. The definition of a drug-like set was generous so as not to limit novelty by imposing too restrictive a physicochemical profile. Hence, the physicochemical profiles were used to reject compounds lying at the extremes of property ranges. In the end, only around 15 percent of the database was excluded on the basis of 2-D properties.⁵

A small number of known ligands was added to the database in order to test whether the docking process could reliably retrieve known hits from the large number of random compounds. The set of 20 agonists included analogs of estradiol, diethylstilbestrol,

Figure 3 Comparison of the 3-D receptor-bound conformations of estradiol and raloxifene as derived from the coaligned X-ray structures 1ERE and 1ERR



and various environmental estrogenic compounds, and the 20 antagonists comprised analogs of raloxifene, tamoxifen, and other literature compounds.²⁵

Docking and scoring. Docking was performed using our in-house package PRO_LEADS, which has been developed to objectively dock flexible ligands into a rigid receptor model.²⁸ The main features of a docking package are an efficient searching algorithm to explore a very large number of potential docking conformations and orientations per ligand, and a scoring function that can accurately identify the correct binding mode and yield a reliable estimate of the binding affinity. PRO_LEADS uses a tabu searching strategy that in effect performs a random search in the conformational or orientational space of the ligand but maintains a tabu list of visited solutions to drive the search toward novel regions of space. Typically 10^5 to 10^6 moves are required within the tabu search to find the global minimum on what is an extremely complicated energy landscape. This procedure is usually repeated several times from a

Figure 4 Frequency distribution of DockedEnergy for docking the one million random compound set (ACD-SC) and reference set (agonists) against the agonist conformation of the estrogen receptor

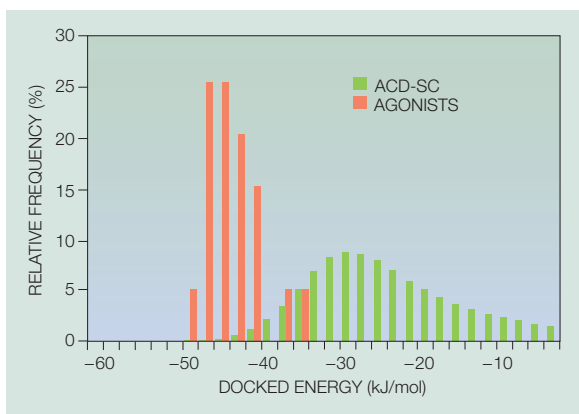
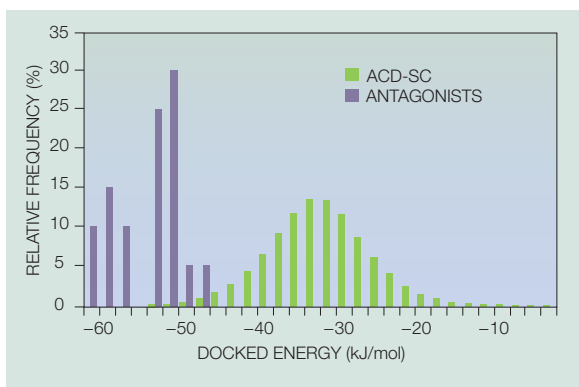


Figure 5 Frequency distribution of DockedEnergy for docking the one million random compound set (ACD-SC) and reference set (antagonists) against the antagonist conformation of the estrogen receptor



random starting position of the ligand, so that in the current example, five attempts were made per ligand, each resulting in a saved solution. The best scoring solution was then selected from the five attempts.

The energy function, ChemScore, is an empirical scoring function that gives a direct prediction of free energy of binding.²⁹ The scoring is based on an evaluation of the strength of the ligand–receptor interaction in terms of the geometry of hydrogen bond contacts, lipophilic contacts, and metal binding con-

tacts, together with an entropic penalty to describe the freezing of ligand conformational flexibility upon binding. This scoring function is rapidly evaluated using precalculated receptor-specific grids to serve as look-up tables of component atomic binding energies or to point to lists of neighboring receptor atoms. The scoring function has been calibrated against a large data set of heterogeneous crystallographic ligand–receptor complexes. The accuracy of prediction is ~ 8.7 kJ/mol (kiloJoules per mole) (i.e., the cross-validated standard error for an 82-complex set), equating to around one to two orders of magnitude in binding affinity.²⁹ For use as a fitness function for docking, the original ChemScore function is modified to introduce longer-range hydrogen bonding terms and to include terms to describe ligand–receptor clash and ligand strain energy.²⁸ In a validation study, the lowest energy binding mode predicted by PRO_LEADS was within 2 Å root-mean-square (RMS) deviation of the crystallographically determined binding mode for 79 percent of ligands in a 70-complex test set.²²

In this example, grid boxes were built within Prometheus to cover the complete binding site as occupied by estradiol in 1ERE or raloxifene in 1ERR. As the binding site is quite enclosed in the estrogen receptor, a grid box extending 4 Å around the bound ligand in the crystal structure encompasses all the accessible binding pockets.

Results of DockCrunch

In this section, we discuss results obtained in the DockCrunch project.

Energy profiles and prediction of selectivity. The predicted binding modes for the reference ligands were in agreement with the available crystal data. For example, estradiol and raloxifene were predicted to within 1 Å heavy atom RMS deviation of the crystallographic solutions. For the other reference ligands the true binding modes are unknown, but it is assumed that they will bind in a broadly similar manner. In each case, PRO_LEADS yielded a plausible solution.

Figures 4 and 5 summarize the distribution of Docked-Energy for the million compound set docking to the two forms of the estrogen receptor. (DockedEnergy is a prediction of the free energy of binding, and, therefore, a more negative energy represents a more favorable binding interaction.) For comparison, the graphs also show the energy distribution for the sets

of known agonists and antagonists. The energies of the one million random compounds binding to the agonist receptor are very broadly distributed (Figure 4), with a median of -23 kJ/mol and with only the top 1 percent of the random compounds scoring below -40 kJ/mol. The broad tail at high (positive) energies reflects a number of compounds docking very poorly, i.e., large compounds that are not well accommodated within the enclosed binding pocket and therefore incur a large clash penalty. (Note that this clash function is softer than the typical van der Waals energy used in molecular mechanics.) In contrast, the set of 20 known agonists is markedly clustered around energies of -33 to -46 kJ/mol; the poorer scores are associated with nonsteroidal ligands, whereas estradiol analogs are more tightly clustered at -41 to -46 kJ/mol. Around 10 percent of the random compounds score better than -33 kJ/mol (i.e., the maximum score for a known agonist), whereas only 1 percent score better than -40 kJ/mol. Therefore, there is a good separation between known ligands and random compounds on the basis of predicted binding affinity.

A similar result is seen for compounds docking to the antagonist receptor (Figure 5). The random compounds are again broadly distributed, now with a somewhat greater median value of -30 kJ/mol. The known antagonists, scoring from -44 to -62 kJ/mol, are very well separated from the random compounds, with only around 1 percent of the random set scoring at a similar level. It is perhaps not surprising that better discrimination is seen for the antagonist receptor, since the known ligands are relatively large compounds, and it is unlikely that the random data set will contain many hits possessing all the structural features required for good binding. In contrast, the agonists are smaller and comparatively simple compounds, and it is more likely that the random data set will contain small lipophilic compounds capable of mimicking a steroid skeleton.

It was noted above that the random compounds appeared to show more negative binding energy to the antagonist receptor compared with the agonist receptor. This difference is highlighted as a scatter plot in Figure 6. This plot demonstrates the degree of selectivity shown by the random data set between the two forms of the receptor. It can be seen that a large number of points lie along the main diagonal—they represent compounds binding with approximately similar energies to the two forms of the receptor. In addition, there is a significant cloud of points lying below the diagonal, i.e., compounds dem-

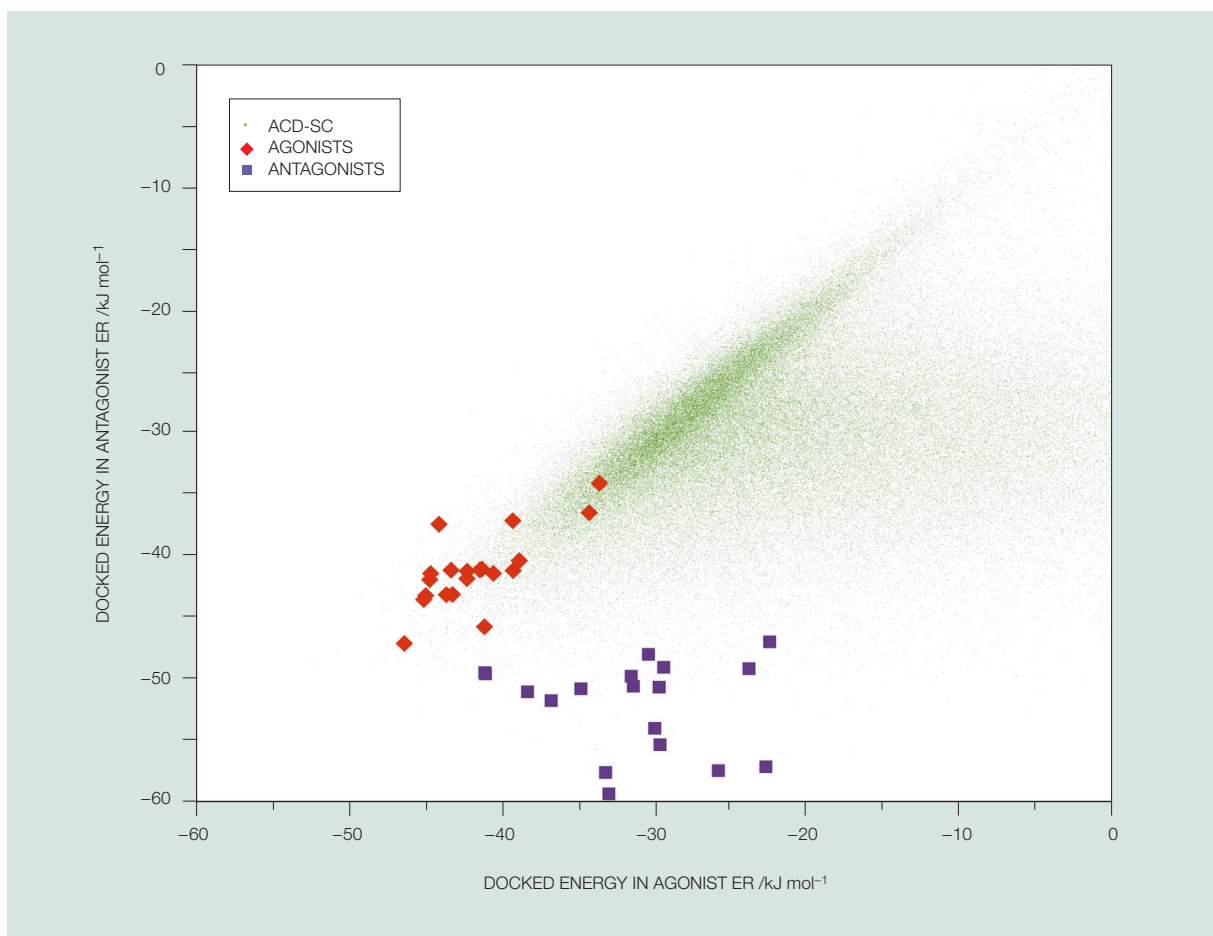
onstrating greater binding energies to the antagonist receptor compared with the agonist receptor. These compounds are typically large ligands that bind well to the additional binding pocket in the antagonist receptor but bind poorly (because of clashing) to the agonist receptor. These trends are highlighted by the reference ligands. The known antagonists lie well below the diagonal, whereas the known agonists lie on the diagonal.

It may have been expected that the agonists would show a preference in energy for the agonist receptor. However, the nature and number of the contacts that these small ligands make to the two receptors are very similar, and the scoring function is not sensitive enough to differentiate between them. This does not imply that these compounds would actually act as antagonists, since the antagonist form of the receptor is presumably only stabilized by virtue of binding to an appropriate ligand.

Thus, the DockedEnergy is not only able to separate known ligands from random compounds for a particular receptor but is also able to give an indication of receptor selectivity. This feature is important in drug design, where toxicity of a drug may be related to promiscuous binding across a number of structurally related receptors.

Supplementary descriptors: Measures of surface complementarity. One important benefit of performing virtual screening by means of receptor–ligand docking is that a large body of data is generated for assessing the quality of the models. Although the above analyses have focused on the total DockedEnergy, there are also component energy terms (e.g., hydrogen bonding and lipophilic terms), as well as a wide range of structural descriptors that may be calculated from the 3-D models of the receptor–ligand complexes. These additional descriptors have been found to improve the discrimination between known ligands and random ligands (reflecting the fact that most scoring functions tend toward overprediction of random compounds—an unavoidable consequence of calibrating the functions against training sets of tight-binding ligands). In the present example, these supplementary descriptors included a measure of shape complementarity (StericPenalty³⁰), and also the degree of mismatch between polar and lipophilic surface areas on ligand and receptor, which are features not adequately described in the ChemScore energy function.

Figure 6 Scatter plot of DockedEnergy for the agonist estrogen receptor (ER) versus the antagonist ER, for the set of random compounds (ACD-SC) and reference agonists and antagonists

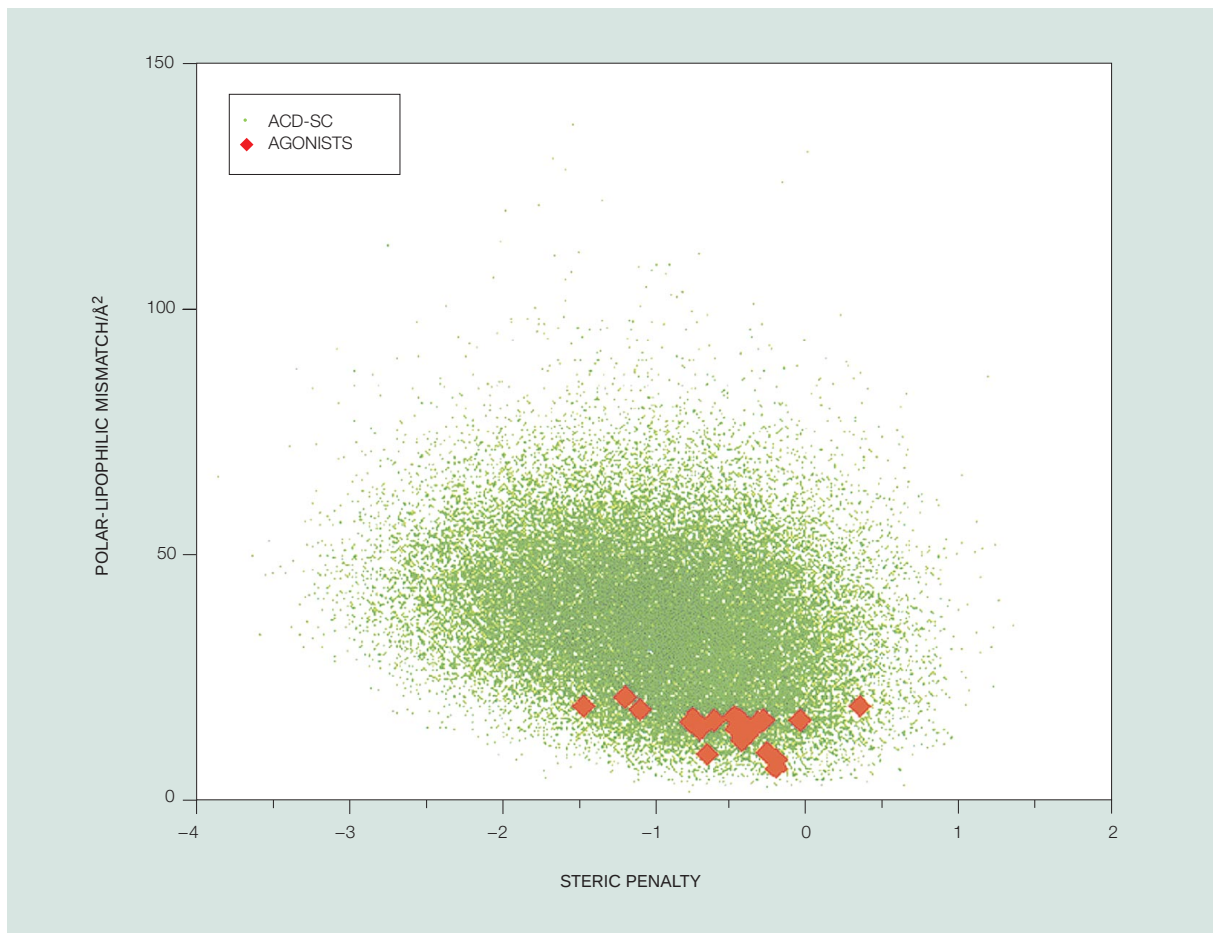


For most of these structural descriptors, it is necessary to define preferred ranges by analysis of the docking modes of known ligands because we have found that different combinations of descriptors may be effective for different types of receptor. For example, given the nature of the estrogen receptor (i.e., an enclosed lipophilic pocket), it may be expected that useful descriptors would include measures of clashing with the receptor and the presence of inappropriate polar functionality in the ligand. Although this implies that the analysis is to some extent tailored to the specific example, this condition should not be seen as a limitation of the method. The analysis protocol is very similar across all receptor types, but some of the finer details of the analysis are varied to exploit specific properties or knowl-

edge of the structure-activity relationship (SAR) of a particular receptor.

The StericPenalty and PolarLipophilicMismatch terms are plotted in Figure 7 for the known agonists and the higher-scoring random ligands (i.e., DockedEnergy < -30 kJ/mol and HbondEnergy < -3 kJ/mol). The plot shows that the known agonists cluster more compactly than the random set in this particular property space, and hence, these supplementary properties are useful in differentiating between good ligands and random compounds. Ideally, these supplementary descriptors would be incorporated into the ChemScore predicted binding energy. The problem is in assigning an energy value to these descriptors that is compatible with ChemScore, par-

Figure 7 Scatter plot of StericPenalty vs PolarLipoMismatchArea for the sets of high-scoring random compounds (DockedEnergy < -30 kJ/mol, HbondEnergy < -3 kJ/mol) and known agonists, docked against the agonist receptor



ticularly as the random data set typically displays values that can be well outside the limits seen for the ChemScore training set.

Subset selection: Enrichment of subsets with known ligands. The objective of performing virtual screening is to select subsets of compounds for assay that are more likely to contain active hits than a sample chosen at random. For the validation of the Dock-Crunch exercise, we have included sets of 20 known agonists and 20 known antagonists to act as our reference hits. The *enrichment rate* is the increase in the proportion of hits found in any given sample of compounds, compared with the proportion expected from a random sample. (For example, the hit rate for a randomly chosen sample is 20 known hits out

of 1.1×10^6 random compounds, or 0.002 percent.) In terms of the practical value of selecting compounds for high-throughput screening, achieving a good enrichment rate of active compounds in the chosen subset is more important than predicting the binding affinity of individual compounds to high accuracy.

Table 2 presents examples of enrichment rates that can be obtained by making a focused selection on the basis of DockedEnergy or other properties. Note that in practice the enrichment rates will vary according to the nature of the receptor and the ligands and to the selection criteria applied, so the data are presented as a representative example.

Table 2 Examples of enrichment rates for selection of high-scoring ligands to the estrogen agonist receptor (enrichment rate is the relative increase in the proportion of reference ligands recovered in a given sample compared to a random sample)

Subset Selection Criteria	Percent of Total Database Selected	Number of Reference Ligands Recovered	Enrichment Rate
None	100.0	20	1
DockedEnergy < -30 kJ/mol	19.0	20	5.2
DockedEnergy < -40 kJ/mol	1.0	15	76
DockedEnergy < -30 kJ/mol + additional complementarity descriptors*	0.9	20	108

*The additional complementarity descriptors were: HydrogenBondEnergy < -3 kJ/mol, StericPenalty -1.5 to +0.5 and PolarLipoMismatchArea < 25 Å²

Table 3 An example of the filtering protocol applied to the selection of potential estrogen receptor agonists from docking 1.1 million ACD-SC compounds

Selection Criteria	Subset Size
None	1 152 379
DockedEnergy	73 961
Energy components	12 265
Receptor-ligand complementarity	2 571
2-D property profiles	1 520
3-D similarity profiles	293

DockedEnergy was defined as < -30 kJ/mol. *Energy components* included limits on hydrogen bonding (-12 to -5 kJ/mol), lipophilic (-40 to -20 kJ/mol) and clash energy (< -10 kJ/mol) terms. *Receptor-ligand complementarity* included the StericPenalty (-2 to +1) and PolarLipophilic Mismatch Area (<25 Å²) properties described in the text. *2-D property profiles* comprised a more focused limit on molecular weight (350 to 450), log *P* (1 to 6), and counts on rings (≤3), chiral centers (≤3) and polar atoms (≤5 acceptors, ≤3 donors), while excluding high structural similarity on the basis of 2-D substructures (Tanimoto coefficient <0.9 based on 2-D MACCS keys). *3-D similarity profiles* were measures of similarity in binding mode versus models of known ligands docked in the receptor; these were calculated on the basis of volume overlap with a docked reference ligand, and also by a bit-string comparison of the receptor atoms forming favorable hydrophobic or hydrogen-bonding contacts with the ligand.

When the selection is made on the basis of DockedEnergy alone, it is seen that by applying more stringent energy criteria (e.g., -30, -35, and -40 kJ/mol), smaller subsets of the database are defined (approximately 20 percent, 5 percent, and 1 percent of the total). In each case the majority of known ligands are recovered, although some of the reference ligands are lost when using the most stringent cutoff. However, it is the case that using DockedEnergy as the single exclusion criterion yields an effective increase in the enrichment rate, so that when the top-scoring 1 percent of the data set is chosen, the proportion of known ligands in the sample is 76-fold that of a random sample. However, Table 2 also demonstrates the value of using other properties to yield a more focused selection. Thus by supplementing a more generous energy cutoff with other energy terms and descriptors of receptor-ligand complementar-

ity, a small subset is again derived (~1 percent of the total) without the loss of the poorer-scoring reference ligands. A similar result was obtained for the antagonist receptor.

Identification of novel estrogenic compounds. As a further validation of the DockCrunch project, a number of compounds were selected from the high-scoring random compounds and assayed for binding affinity against the human alpha estrogen receptor. Following the general selection protocols outlined above, a number of preferred compounds were selected for each receptor using criteria such as total and component docked energies, receptor complementarity, 2-D physicochemical properties, similarity of binding mode to known ligands, and dissimilarity of 2-D structure to known ligands. An example of the selection process for the agonist receptor is given in Table 3.

Where a high-throughput assay is available, it would generally be more productive to search for novel hits by assaying a large number of compounds—say, in the thousands—rather than use all the above selection criteria to produce too small and narrow a subset. However, in the present example, compounds were to be assayed in an external low-throughput assay, and therefore, a small focused subset was selected. As shown in Table 3, the application of energy components and receptor complementarity descriptors reduced the data set to about 2500 compounds as potential agonists. At this stage, further refinement was achieved by applying a more restrictive set of 2-D physicochemical properties in order to select small lipophilic but nonsteroidal compounds. 3-D similarity descriptors were used to select compounds binding in a mode approximately similar to known ligands. This left around 300 compounds for the agonist receptor. At this stage, the

list was inspected to select a diverse set of compounds, rejecting those with potentially problematic chemistries (e.g., potentially insoluble compounds), resulting in a list of about 120 compounds.

A similar process was applied to the ligands binding to the antagonist receptor. In this case, although there are many high-scoring random ligands, only a small proportion of these ligands have an appropriate range of complementarity properties and 2-D properties, and from a short list of 118 compounds a manual selection of 32 compounds was made. Since not all compounds were readily available from the supplier, the final set comprised 37 compounds (21 from the agonist list, 16 from the antagonist list), which were assayed in a standard competitive radioligand binding assay.³¹

Of these 37 compounds, 21 exhibited an inhibition constant (K_i) of 300 nM (nanomolar) or less, with the best compounds at 8 nM (Table 4). A set of control compounds, possessing similar drug-like features but poor docking energies, were shown to be inactive (<25 percent inhibition at the maximum assayed concentration of 10 μ M, or micromolar). Given the structural novelty of the hits (compounds known to possess estrogenic activity or to be structurally similar to known ligands were excluded from assay), this represents a very positive result, demonstrating that virtual screening can readily identify potent ligands from a variety of structural classes.

Analysis tools. The management and analysis of the data generated within virtual screening is a problem in its own right and one that parallels the problems in data management faced by HTS scientists. In the present example we saved five docked solutions for each of the 1.1 million ligands against both forms of the estrogen receptor, resulting in 11 million 3-D models of ligand-receptor complexes, each supplemented by docked energies, component energies, and other 2-D and 3-D descriptors.

Navigation of this large body of data is achieved by using an in-house molecular graphics package developed to integrate interactive tools for data analysis, structure browsing, and subset selection. The tools include, for example:

- Real-time display and navigation of very large data sets, including molecular structures, property histograms, and scatter plots
- Interactive subset selection using desired property ranges

Table 4 Binding affinities (K_i) against the human alpha-estrogen receptor for 37 compounds selected from the virtual screen

Binding Affinity Criterion	Number of Compounds
<10 nM	2
<100 nM	14
<300 nM	21
>25 percent inhibition*	27
None	37

*Percent inhibition at maximum assayed concentration (10 μ M) for compounds for which a full K_i was not determined.

- Display of receptor-bound ligand structures, with the ability to couple the viewing orientation of multiple receptors (e.g., to compare different ligand sets in the same receptor or the same ligands bound to different receptors)
- Calculation of additional properties, e.g., specific receptor-based contact distances or measures of 3-D similarity in binding mode to a reference ligand (usually submitted as a background calculation)

The impetus behind the visualization software is the need for an interactive platform to allow the modeler to explore the property distributions, construct various subsets based on different combinations of selection criteria, and visualize the docked models relating to the chosen subsets. Once the modeler is confident about how best to analyze a particular data set, the selection criteria can be automatically applied to other related data sets using the scripting routines within Prometheus.

One advantage of constructing an in-house software package is the degree of integration between different modeling tasks that can be achieved. This integration contrasts with the environment usually encountered in industry, where a modeling group will be running different codes from different sources to perform the various tasks in library construction, property profiling, docking, analysis, etc. For example, within Prometheus the docking code produces as output a single file of the best docked conformations of each ligand (in the MDL ISIS[®], or Integrated Scientific Information System, *sdf* format), which also contains all the docked energies and other properties (such as receptor complementarity) calculated during docking, as well as the 2-D properties calculated in the original preprocessing of the data set. This *sdf* file comprises the basic input required for the graphics visualization package, thus avoiding

problems of compiling and reformatting data from different sources.

Timescales. A range of docking protocols were evaluated for the estrogen receptor to determine the number of iterations (energy evaluations) required to accurately dock known ligands. The preferred protocol docks a single ligand in approximately 30 seconds on an SGI R10000** processor. Therefore, the 1.1 million data set took six days on a 64-processor SGI Origin 2000**.

Docking protocols that we have optimized for other receptors tend to be somewhat longer, e.g., we may use a fivefold longer protocol (in terms of tabu searching steps) for the enzyme thrombin. This greater length reflects the nature of the receptor site—enzymes such as thrombin tend to display a more open (solvent exposed) binding site, with many accessible subsites, not all of which are essential for binding any particular ligand. In addition, some subsites can consist of relatively deep or restricted pockets and, therefore, the searching algorithm has a more difficult task in locating the correct solution. In contrast, the binding site of the estrogen receptor consists of a single well-enclosed pocket that does not accommodate many orientations or conformations of a ligand.

Docking methods such as PRO_LEADS that treat the ligand as conformationally flexible are inevitably slower than methods that keep the ligand rigid. However, exploring ligand flexibility allows a more objective and accurate prediction of the bound geometry within the receptor and, therefore, the increased cost in docking is justified by the quality of the predicted conformation and binding energy. The elapsed time for the DockCrunch experiment (six days for one receptor) is a very acceptable timescale for this type of task within a drug discovery program, and compares favorably with the start-up times required in other areas of a screening program such as assay automation and validation, compound acquisition, or synthetic route finding and library synthesis (i.e., in a combinatorial chemistry program).

It should be mentioned that other hardware platforms can achieve comparable or improved throughput for virtual screening. For example, we have recently installed a PC cluster running the Linux** operating system, based on 100 Intel Pentium** III 750 MHz processors. This installation is an attractive solution for cost-effective computing in the present example of receptor–ligand docking, since

this application is straightforward to parallelize across a multiprocessor machine. Thus a daemon, written in Perl and backed by a keyed database, distributes batches of ligands to as many individual processors on the multiprocessor or cluster as the user requires. Additional batches of ligands are automatically launched as processors become free. This approach enables the modeler to readily manage machine load, submit across heterogeneous architectures, and track the progress of the job.

Conclusions

Virtual screening is increasingly gaining acceptance in the pharmaceutical industry as a cost-effective and timely strategy for analyzing very large chemical data sets. Although the number of therapeutic targets that have been fully characterized by crystallography is currently limited, this situation is set to change significantly in the immediate future as structural genomics initiatives begin to yield fruit. Accordingly, the work involved to validate all these potential targets, to demonstrate their therapeutic relevance, and to find effective ligands will become heavily dependent on the new high-throughput screening technologies.

We have used explicit molecular docking to yield detailed models of over one million drug-like compounds bound to two conformations of the human estrogen receptor. This procedure is computationally intensive for analyzing a large database but provides the most detailed basis for determining which compounds are likely to be potent ligands. Our results demonstrate that the docking program can accurately reproduce the known binding modes of reference ligands, and is capable of discriminating between reference ligands and random compounds on the basis of predicted binding affinity and other structure-derived descriptors.

We have already demonstrated that existing (and moderately inexpensive) hardware can perform a very detailed computational analysis (molecular docking) on data sets as large as a million compounds. In the near future we can foresee that PC clusters of a few hundred processors will be able to process millions of compounds a week and to explore selectivity across large families of structurally related receptors. The successful integration of such analyses with high-throughput library synthesis and evaluation will have a significant effect on the optimization of drug discovery strategies in coming years.

Acknowledgments

For assistance with the modeling, programming, and administrative tasks involved in the DockCrunch project, we acknowledge the contributions of Carol Baxter, Chris Murray, Paul Greaney, Andrew Crew, and Stephen Young at Protherics; and also the assistance of Stavros Taraviras, Pam Bremmer, and Ulrich Meier (SGI Europe), Paul Collins (MDL, UK), and Cait Murray (MSI).

**Trademark or registered trademark of MDL Information Systems, Inc., Linus Torvalds, Intel Corporation, Technologies, Inc., or Silicon Graphics Inc.

Cited references and note

1. J. Drews, "Drug Discovery: A Historical Perspective," *Science* **287**, 1960–1964 (2000).
2. L. J. Beeley, D. M. Duckworth, and C. Southan, "The Impact of Genomics on Drug Discovery," *Progress in Medicinal Chemistry* **37**, 1–43 (2000).
3. B. Cox, J. C. Denyer, A. Binnie, et al., "Application of High-Throughput Screening Techniques to Drug Discovery," *Progress in Medicinal Chemistry* **37**, 83–133 (2000).
4. W. P. Walters, M. T. Stahl, and M. A. Murcko, "Virtual Screening—An Overview," *Drug Discovery Today* **3**, 160–178 (1998).
5. DockCrunch website: <http://www.protherics.com/crunch>.
6. J. Greer, J. W. Erickson, J. J. Baldwin, and M. D. Varney, "Application of the Three-Dimensional Structure of Protein Target Molecules in Structure-Based Drug Design," *Journal of Medicinal Chemistry* **37**, 1035–1054 (1994).
7. R. S. Bohacek, C. McMartin, and W. C. Guida, "The Art and Practice of Structure-Based Drug Design: A Molecular Modeling Perspective," *Medicinal Research Reviews* **16**, 3–50 (1996).
8. B. K. Shoichet and D. E. Bussiere, "The Role of Macromolecular Crystallography and Structure for Drug Discovery: Advances and Caveats," *Current Opinion in Drug Discovery and Development* **3**, 408–422 (2000).
9. R. Lahana, "How Many Leads from HTS?" *Drug Discovery Today* **4**, 447–448 (1999); C. S. Ramesha, "How Many Leads from HTS?—Comment," *Drug Discovery Today* **5**, 43–44 (2000).
10. T. S. Peat, J. M. Newman, and D. E. Bussiere, "Structural Genomics in the Post-Genomics Era: The Shape of Things to Come," *Current Opinion in Drug Discovery and Development* **3**, 399–407 (2000).
11. C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings," *Advances in Drug Delivery Research* **23**, 3–25 (1997).
12. D. E. Clark and S. D. Pickett, "Computational Methods for the Prediction of Drug-Likeness," *Drug Discovery Today* **5**, 49–58 (2000).
13. P. S. Charifson, J. J. Corkery, M. A. Murcko, and W. P. Walters, "Consensus Scoring: A Method for Obtaining Improved Hit Rates from Docking Databases of Three-Dimensional Structures into Proteins," *Journal of Medicinal Chemistry* **42**, 5100–5109 (1999).
14. D. Tondi, U. Slomczynska, M. P. Costi, et al., "Structure-Based Discovery and In-Parallel Optimization of Novel Competitive Inhibitors of Thymidylate Synthase," *Chemistry and Biology* **6**, 319–331 (1999).
15. P. Burkhard, U. Hommel, M. Sanner, and M. D. Walkinshaw, "The Discovery of Steroids and Other Novel FKBP Inhibitors Using a Molecular Docking Program," *Journal of Molecular Biology* **287**, 853–858 (1999).
16. M. Schapira, B. M. Raaka, H. H. Samuels, and R. Abagyan, "Rational Discovery of Novel Nuclear Hormone Receptor Antagonists," *Proceedings of the National Academy of Sciences (USA)* **97**, 1008–1013 (2000).
17. A. V. Filikov, V. Mohan, T. A. Vickers, et al., "Identification of Ligands for RNA Targets via Structure-Based Virtual Screening," *Journal of Computer-Aided Molecular Design* **14**, 593–610 (2000).
18. E. Perola, K. Xu, T. M. Kollmeyer, et al., "Successful Virtual Screening of a Chemical Database for Farnesyltransferase Inhibitor Leads," *Journal of Medicinal Chemistry* **43**, 401–408 (2000).
19. H.-J. Boehm, M. Boehringer, D. Bur, et al., "Novel Inhibitors of DNA Gyrase: 3D Structure Based Biased Needle Screening, Hit Validation by Biophysical Methods, and 3D Guided Optimization. A Promising Alternative to Random Screening," *Journal of Medicinal Chemistry* **43**, 2664–2674 (2000).
20. S. C. Hopkins, R. D. Vale, and I. D. Kuntz, "Inhibitors of Kinesin Activity from Structure-Based Computer Screening," *Biochemistry* **39**, 2805–2814 (2000).
21. A. M. Aronov, N. R. Munagala, P. R. Ortiz de Montellano, I. D. Kuntz, and C. C. Wang, "Rational Design of Selective Submicromolar Inhibitors of *Tritrichomonas foetus* Hypoxanthine-Guanine-Xanthine Phosphoribosyltransferase," *Biochemistry* **39**, 4684–4691 (2000).
22. C. A. Baxter, C. W. Murray, B. Waszkowycz, et al., "New Approach to Molecular Docking and Its Application to Virtual Screening of Chemical Databases," *Journal of Chemical Information and Computer Science* **40**, 254–262 (2000).
23. J. I. MacGregor and C. Jordan, "Basic Guide to the Mechanisms of Antiestrogen Action," *Pharmacological Reviews* **50**, 152–196 (1998).
24. A. M. Brzozowski, A. C. Pike, Z. Dauter, et al., "Molecular Basis of Agonism and Antagonism in the Oestrogen Receptor," *Nature* **389**, 753–758 (1997).
25. T. A. Grese and J. A. Dodge, "Selective Estrogen Receptor Modulators," *Current Pharmaceutical Design* **4**, 71–92 (1998).
26. *Insight II v 95.0*, Molecular Simulations Inc., San Diego, CA.
27. *ISIS/Base 2.2*, MDL Information Systems, Inc. (MDL), San Leandro, CA.
28. C. A. Baxter, C. W. Murray, D. E. Clark, D. R. Westhead, and M. D. Eldridge, "Flexible Docking Using a Tabu Search and an Empirical Estimate of Binding Affinity," *Proteins: Structure, Function, and Genetics* **33**, No. 3, 367–382 (1998).
29. M. D. Eldridge, C. W. Murray, T. R. Auton, et al., "Empirical Scoring Functions: I. The Development of a Fast Empirical Scoring Function to Estimate the Binding Affinity of Ligands in Receptor Complexes," *Journal of Computer-Aided Molecular Design* **11**, 425–445 (1997).
30. StericPenalty is derived from measuring the contact distance between each point on the van der Waals surface of the ligand and the nearest accessible receptor atom. On subtracting the van der Waals radius of the receptor atom, a value d is obtained that is zero for two atoms separated by the sum of their van der Waals radii, and positive for atoms separated by a greater distance. The median of the set of d values for a given ligand was found to be a useful measure of the re-

ceptor–ligand steric complementarity. As this median value was found to be correlated with ligand size, the value is normalized by ligand surface area with respect to the set of receptor–ligand complexes used for the calibration of the ChemScore scoring function. The normalized value, StericPenalty, has a value of zero for ligands as tightly bound as the average of the reference set, has a negative value for ligands more tightly bound (e.g., clashing), and a positive value for ligands less tightly bound.

31. J. D. Oburn, N. J. Koszewski, and A. C. Notides, “Hormone- and DNA-Binding Mechanisms of the Recombinant Human Estrogen Receptor,” *Biochemistry* **32**, 6229–6236 (1993).

Accepted for publication January 22, 2001.

Bohdan Waszkowycz *Protherics Molecular Design Ltd., Beechfield House, Lyme Green Business Park, Macclesfield, Cheshire SK11 0JL, United Kingdom (electronic mail: Bohdan.Waszkowycz@protherics.com).* Dr. Waszkowycz received a B.Sc. degree in pharmacy and a Ph.D. degree in computational chemistry at the University of Manchester, then joined Proteus Molecular Design Ltd. (now Protherics) in 1990. He served as a molecular modeler on a number of structure-based drug design projects, most recently on the design of tryptase inhibitors, before leading the computational team in the development of the in-house software suite, Prometheus. The recent focus of the group has been the validation of software for high-throughput virtual screening, and Dr. Waszkowycz is currently responsible for establishing collaborative projects on virtual screening with a number of pharmaceutical and biotechnology companies.

Tim D. J. Perkins *Protherics Molecular Design Ltd., Beechfield House, Lyme Green Business Park, Macclesfield, Cheshire SK11 0JL, United Kingdom.* Dr. Perkins graduated from Cambridge with a B.A. degree in natural sciences and received a Ph.D. degree in medicinal chemistry from the University of London. He joined the Drug Design Group in the Pharmacology Department at Cambridge with Dr. Philip Dean, where he researched computational methods for conformational analysis and molecular superposition. This group formed the basis of the TeknoMed drug design collaboration with Rhône-Poulenc Rorer. On joining Protherics in 1998, Dr. Perkins has worked on software development within Prometheus, particularly in implementing tools for facilitating high-throughput virtual screening, including novel methods for analysis of receptor–ligand complementarity.

Richard A. Sykes *Protherics Molecular Design Ltd., Beechfield House, Lyme Green Business Park, Macclesfield, Cheshire SK11 0JL, United Kingdom.* Mr. Sykes received a B.Sc. degree in logic with mathematics from the University of Sussex and worked as a programmer for a number of years before returning to academia to research the theory and application of functional programming languages at the University of London. Since joining Protherics in 1991, he has been the senior programmer involved in the development of Prometheus. He has a particular interest in the development of the scripting language Global and the design and implementation of graphical user interfaces to support the requirements of the group’s structure-based design and virtual screening efforts.

Jin Li *Protherics Molecular Design Ltd., Beechfield House, Lyme Green Business Park, Macclesfield, Cheshire SK11 0JL, United Kingdom.* Dr. Li is Head of Computational Chemistry at Protherics. He received a B.Sc. degree from Sichuan University,

China, and a Ph.D. degree in macromolecular sciences from Aston University, followed by postdoctoral research in theoretical biochemistry at the University of Manchester with Dr. Barry Robson. He joined Protherics in 1990 to undertake research into computer simulation of protein folding and protein structure prediction. Since 1994 he has led the computational chemistry team in developing methods and software for molecular design, particularly in the areas of *de novo* design, molecular docking, and combinatorial library design. More recently Dr. Li was responsible for initiating and directing the DockCrunch project, involving a million compound virtual screen versus the estrogen receptor.