optimization time robotics has been used. In this situation complete assay systems can be run where a robotic arm is programmed to sample devices, leave them for a set running time and then read the devices. Using a CRS 7 joint robot arm it has been possible to set up an automated system capable of assaying eight devices/min for 24 h per day (see Figure 4). This approach reduces resource requirements and increases throughput 6-fold.

Summary
The use of HTS systems at various stages in the product-development cycle can dramatically speed up the time to market for a new product. Many industries are under commercial pressure to innovate and bring products to market quickly. Where as new drugs can have lifetimes of over 15 years many consumer products typically have shorter product lifetimes. Therefore, a short product-development cycle is essential for consumer-based products.

References

Received 9 April 2002

Virtual screening: a real screening complement to high-throughput screening
J. Mestres
Computational Medicinal Chemistry, Organon Laboratories Ltd., Newhouse, Lanarkshire ML1 5SH, U.K.

Abstract
Virtual screening is being routinely used as an integral part of today's hit-identification strategies for, on one hand, prioritizing large corporate screening collections and, on the other hand, to extend the scope of screening to external databases. A brief description of the essential elements required for virtual screening and an application example to the identification of agonist hits for the oestrogen receptor subtype ERα are presented.

Introduction
In the last few years, the pharmaceutical industry has relied heavily on high-throughput screening (HTS) technology for the identification of hits. Large investments and logistic efforts have been put into setting up robotic equipment and in acquiring, assembling and maintaining big corporate compound collections. Within this period, HTS has indeed been able to prove itself as a technology capable of supplying novel valuable hits to our drug-discovery pipelines. However, alongside its now established value, it has been widely recognized that the assay-to-lead attrition rate is higher than initially anticipated and that, even in those cases where leads are generated from hit-optimization programmes, the quantity, quality and diversity of those leads have been below original expectations [1]. In addition, a number of important quality issues have emerged, namely, quality of the assays (leading often to poor quality of data, large numbers of false positives and tedious active confirmation process), quality of the collection (purity and stability of compounds, as well as composition and origin of the collection) and quality of the logistics (purchase, tracking, handling and storage of compounds and data management), which have increased significantly the cost of maintenance of HTS.

In spite of the relatively low hit rates and the quality issues highlighted above, HTS is still considered nowadays to be the main source of hits in pharmaceutical companies. However, at the same time, it has become apparent that other techniques and/or strategies for hit identification are needed to keep up with current demands in drug discovery. Within this scenario, computational methods for the virtual screening of compound libraries offer a time-efficient and cost-effective alternative or complement to HTS [2].

Virtual screening
Virtual screening can be defined as the process of reducing a library containing an unmanageable number of compounds (available or virtual) to a limited number of potentially promising compounds for the target (or target family) of interest by means of computational methods that exploit
existing knowledge. The application of virtual screening techniques previous to or in parallel with HTS has three clear objectives: first, to extend the scope of the screening to external databases; second, by doing so, to identify more and more diverse hits; and third, as a consequence, help to reduce the assay-to-lead attrition rate observed from HTS.

Despite being done in silico instead of in vitro, the basic requirements for a virtual screening show a good parallelism with those needed to perform an experimental (referred to as 'real' in this article) screening. This is illustrated in Scheme 1. In essence, there are four components that need to be in place for any screening to be performed. (i) A prerequisite for performing any type of real screening is the development of a biological assay for the target. The related prerequisite in virtual screening is the availability of a given methodology based on the target and/or ligands associated with it [3]. From a practical point of view, virtual screening offers the advantage that methodologies tend to be general and readily applicable in principle to every target. (ii) In a real screening, a compound collection needs to be physically available. As highlighted above, putting a compound screening collection together is expensive, requires demanding logistic efforts, and has associated with it a number of quality issues. In contrast, in virtual screening, electronic databases are easy to set up, maintain and ensure for quality [4]. In addition, virtual screening is not restricted to the compounds physically available in corporate screening collections, but can be easily extended to compounds available from external suppliers and in principle to any non-available but potentially synthesizable compound. (iii) In a real screening, compounds are ranked on the basis of their response to an assay, which is generally expressed as a one-point measurement at a given concentration of the compound. The data produced by this means are not of high quality, but are often useful to reduce dramatically the number of interesting compounds for which more careful testing might be required. In virtual screening, ranking of compounds is done according to a given scoring function [5]. Numerous scoring functions are currently available but, despite the efforts, there is still significant room for improvement. (iv) Finally, the ultimate confirmation that a compound is binding in the receptor cavity is obtaining structural information on the protein–ligand complex. Experimentally, this information can be provided by X-ray crystallography. Unfortunately, approx. 45% of today's drugable targets are G-protein-coupled receptors, targets that are still a challenge to crystallize. Alternatively, in the virtual world, structural information is gained by applying protein and molecular modelling techniques [6], providing that structural templates for a close member of the same protein family and/or active ligands, respectively, are available. In the end, any screening approach ultimately aims to identify hits for the target of interest and, in this respect, they all (real and virtual) should be synergistic.

**Scheme 1**

Comparison of the basic requirements to perform real and virtual screenings

- **Real**
  - Assay
  - Collection
  - Testing
  - Crystallography
  - Hits

- **Virtual**
  - Methodology
  - Database
  - Scoring
  - Modelling

**Hit identification**

As an application example, we embarked on the search for alternative hits as agonists for the oestrogen receptor subtype ERα. The methodology employed in this case was a ligand-based flexible superposition technique (as implemented in an in-house program called MIMIC) [7], the database screened was the Available Chemicals Directory (ACD), and the scoring function used was a combination of the three-dimensional steric and electrostatic similarity with respect to a reference compound. The procedure followed is illustrated in Figure 1. Basically, the adopted conformation of diethylstilbestrol when forming a complex with ERα (PDB code 2ERD) was taken as the reference structure. Then, the steric and electrostatic fields derived from that conformation were used to, on one hand, obtain the best superimposition with MIMIC and, on the other hand, produce a ranking reflecting the goodness of the fit for each compound in the ACD database with respect to diethylstilbestrol.

The result of the virtual screening was a list of top-scoring compounds that were subsequently
subjected to visual inspection. A final selection of them was purchased from external suppliers. As an example, we report here the identification of bisphenol-Z (ACD code MFCD00019341) as a novel submicromolar ERα agonist ($EC_{50} = 0.2 \mu M$). Interestingly, bisphenol-A (having a dimethyl substituent in place of the cyclohexyl) has been reported to be inactive ($pK_a = -0.164$) [8]. Also, bisphenol-Z shows some degree of similarity to cyclofenyl, a well-known ERα agonist ($EC_{50} = 0.1 \mu M$). Finally, it is worth noting that the hit identified contains a central scaffold that is essentially different from the reference structure used. As shown recently in a comparative study between receptor-based and ligand-based methods, this highlights the fact that three-dimensional ligand-based flexible superposition methods are indeed able to identify alternative diverse hits while retaining some degree of similarity with respect to the reference active compound [9]. This finding is specially relevant for applying virtual screening to the many drugable targets for which little or no structural information is available and thus to which receptor-based methods are not applicable.

**Conclusions**

Any corporate screening library is intrinsically incomplete. While the estimated number of synthetizable compounds is in the order of $10^{48}$, corporate screening collections contain in the order of $10^8$ compounds. It seems therefore naïve to think that compounds possessing the optimal features arranged in the optimal way around an optimal core structure to bind to our protein target of interest will habitually be present in our screening collection. As shown recently, the probability of a compound having the right key functional groups arranged optimally decreases dramatically as the complexity of the receptor increases [10]. The realization of these facts explains current trends in the pharmaceutical industry of making use of virtual screening to explore external libraries of compounds (available or virtual) other than corporate screening collections. Virtual screening has become part of today’s lead discovery arsenal for the identification of more, and more diverse, hits and has been established as a real screening complement to HTS.

**References**


Received 9 April 2002