How many leads from HTS? – Comment

In reply to the Editorial by Roger Lahana (*Drug Discovery Today* 4, 447–448), the purpose of high-throughput screening/ultra-high-throughput screening (HTS/uHTS) is not to identify the molecule, but to identify good leads from which potential clinical candidates could be synthesized using a more rational and/or classical approach. If such exercises fail to generate leads, it means only two things: either the assays were not set up properly or the quality of the library was poor. To obtain 'true' hits from HTS, I would say that leads have been obtained.

Why is a hit not a lead?

Assuming that the activity of the hit is reproducible, the following questions should be asked before it is called a hit.

- Did the hit display activity against many diverse targets?
- Did analogs of the hit display no structure–activity relationship?
- Are chemically reactive functions present in the hit?

A 'yes' answer to any one of these three questions disqualifies the compound from being called a hit; a 'no' answer means it would be a lead. Whether the lead was made by combinatorial chemistry (CC) or classical medicinal chemistry (MC) is immaterial, as is the route (CC or MC) taken to optimize activity. However, it must be highlighted that the chemical tractability (suitability of a compound for chemical modification) of the compound often stops a real hit from becoming a lead. Although chemical tractability is subjective, in my view and experience, this is the single major reason why a hit from HTS is not advanced. How to address this issue is beyond the scope of this discussion.

Differences between HTS and CC/MC

Let us also not confuse HTS with CC/MC. They serve entirely different purposes and complement each other in the drug discovery process. As already mentioned, the purpose of HTS is to identify potential lead candidates as fast as possible, whereas the purpose of CC and MC is to improve the activity of chemical leads as fast as possible. Again, CC and MC serve different functions. Some of the leads are ideally suited to modification by CC, while others require modification by classical MC as they are not amenable to current CC technologies.

If HTS is not producing the expected results, the composition of the chemical library should be examined. After all, the lead that is going to be identified (by HTS) will be as good as (or as bad as) the library that is being used for screening. Therefore, the importance of the quality of the screening library cannot be emphasized enough, while quantity is only secondary. Until recently, the major source of compounds for many major pharmaceutical companies was their internal collection of compounds synthesized over the past 10–50 years. These collections include compounds made not only for human use but also for veterinary and agricultural use. Although a good percentage of these molecules are drug-like, a greater percentage of them are often unstable, reactive, chemically not tractable, or known toxins or carcinogens, making them unsuitable for screening. To eliminate or minimize wasteful efforts, it is essential that these unsuitable compounds are excluded from screening libraries as much as possible.

Yet another important factor that contributes to the quality of the screening library is the chemical diversity: the more diverse the library, the better the chance of identifying hits. Historically, pharmaceutical companies have developed drugs for up to 100 major biochemical targets. Current screening libraries are composed mainly of analogs of compounds made for these targets, this representing a small fraction with regards to chemical diversity. It is my opinion that CC and natural product chemistry could, and should, have a major impact in this respect by adding chemically diverse drug-like molecules to screening libraries. These are not necessarily the compounds that are easy to make but should include the 'right' ones.

Role of HTS and computational screening in drug discovery

The question we need to consider is not whether virtual screening and other computational methods are better than HTS or *vice versa*, but what roles these technologies play in drug discovery. HTS/uHTS is often viewed as a novel science of this decade, but it is not. Biological screening is a proven technique by which most of the currently marketed drugs have been identified and developed. HTS is purely a high-tech, sophisticated version of biological screening, in which cost is reduced and speed is increased by several orders of magnitude. The advancement in HTS means the availability of new tools, while the basic science remains the same. There are no hypotheses or assumptions in HTS: if a compound is active, it is active while if it is not, then it is not. Therefore, to be successful using low-throughput screening will mean success using HTS, provided that the biology is correct and a 'good' library is used. There are several examples where random HTS alone has generated leads,

1359-6446/00/$ – see front matter PII: S1359-6446(99)01444-0

the analogs of which are in the clinic today. However, random HTS (screening all compounds against all targets) might not be the correct approach for the current and future requirements. With a constantly increasing chemical library size as well as screenable targets, a more rationalized HTS is necessary. It is apparent from many of the recent conferences that I have attended that such approaches are beginning to be implemented.

By contrast to HTS, virtual screening and other computational methods are relatively new technologies in drug discovery. All virtual screening methods rely on a subjective model based on a set (or sets) of calculated and/or perceived data derived from primary information such as protein sequences, homology, 3D-structures and biological (including HTS) activities. Therefore, the accuracy of this calculation or perception is the greatest limiting factor in virtual screening. Only when the calculations and presumptions are accurate will the subjective model become objective. For virtual screening to be successful, the model should be built using as many data sets as possible, and using all the information present in the data sets (there is nothing called improper information). Unfortunately, current data sets are very limited and, more importantly, the tools available to extract the information content (data mining) from the data are too primitive.

There is a tremendous effort to develop more valuable and 'right' tools. Without the correct tools, the correct information cannot be extracted (and so remains mostly unknown) and, without the correct information, the correct tools cannot be developed. This is not to say that a dead-end has been reached, but to emphasize the fact that the process of developing the right tools is an evolving process. This involves many reiterating cycles of presumption, testing (mainly biologically) and refinement until it develops into the 'right' tools. It is only then that true objective calculations can begin and the 'right' models can be developed. This process has only just begun. Considering the complexity of the information content in biological and chemical data, and considering how little of the biological and chemical data are known today, it is my view that the quantity of knowledge in knowledge-based computational methods is only at a minimum.

The point that I am trying to emphasize is that, at the present time, virtual screening has not evolved enough or acquired enough sophistication to be successful by itself. It does not, however, mean that virtual screening and other computational methods should be ignored. The limited knowledge acquired by computational methods has to be used in combination with other non-computational technologies (such as HTS, CC and MC). This is how computational methods can make a significant impact on the rationalization of HTS, CC, MC and alike and, in doing so, can increase the knowledge-base of these screening methods. This is the only objective method of carrying out lead discovery. In my opinion, the screening technologies (such as HTS, CC and computational methods) are only the tools. Provided these tools are used intelligently, they can increase speed and reduce costs. It is only then that there will be a significant impact on drug discovery.

Chakkodabylu S. Ramesha
Head, High-Throughput Screening
Inflammatory Diseases Business Unit
Roche Bioscience, Palo Alto
CA 94303, USA
tel: +1 650 855 5847
fax: +1 650 354 7393
e-mail: chakk.ramesha@roche.com

Dimeric malarial drugs for enhanced activity

US chemists have synthesized new dimeric derivatives of an extract from an ancient Chinese herbal remedy that have high activity against malaria. Herbal extracts of the Chinese wormwood, *Artemesia annua*, have been used for centuries to treat fever, including malaria-related fever. The active compound, qinghaosu (known in the West as artemisini) is a natural trioxane containing an unusual seven-membered C–O ring bridged by an O–O peroxide unit, thought to be central to its biological activity. The compound has shown great promise in continuing trials against malaria in China and elsewhere.

Mechanism of action of artemisinin

Gary Posner's team at the Department of Chemistry (Johns Hopkins University, Baltimore, MD, USA) has been studying the mode of action of artemisinin (Fig. 1) against the malarial parasite, *Plasmodium falciparum*, to find a simpler compound with similar or improved activity but without the chemical complexity, hence reducing manufacturing costs. In 1996, Posner's team, working with colleagues at the Weizmann Institute (Rehovot, Israel), found that the essential structure of the artemisinin molecule required for activity is an oxygen triad (trioxane)[1,2].

Malarial parasites obtain their nutrients from the human host's haemoglobin and

 1359-6446/00/$ – see front matter PII: S1359-6446(99)01453-1