Pacific Symposium on Biocomputing 4:438-443 (1999)

EMPIRICAL VS. "RATIONAL" METHODS OF DISCOVERING NEW DRUGS

PETER GUND
Pharmacopeia Inc., CN 5350
Princeton, NJ 08543, USA
pgund@pharmacop.com

Empirical and theoretical approaches to drug discovery have often been perceived as mutually exclusive. Our experience has rather demonstrated that they can be complementary. The structure-based approach to design of compound libraries is clearly helpful; however, testing large libraries continues to reveal unanticipated actives in many of our programs. A rationale for these observations is offered.

Introduction

Styles in drug discovery have vacillated between empirical and theoretical. From the empirical exploration of herbal extracts by ancient shamans, to the theory-guided concoctions of alchemist-physicians, to the systematic testing of synthetic compounds pioneered by Ehrlich, to the "rational" design of agents based on pharmacophoric, receptor-binding, or mechanistic reasoning, and recently to the empirical high-throughput testing of combinatorial libraries – the pendulum has swung now to favor those diligently trying all possible materials, and then to favor those trying to find the "best" compound quickly and cleverly.

The thesis of this meeting session implies a belief that structure-based, "rational" approaches will eventually win out, if only we can get the physics right. But is this true?

It seems likely that we may eventually learn how to compute the binding energy of ligands to enzyme receptors with tolerable accuracy. We may even hope to accurately model the dynamics of enzyme and inhibitor or substrate as they perform their chemical dance.

But even if we do, we will then only know how to design exquisitely strong enzyme inhibitors and receptor antagonists – and perhaps, eventually, receptor agonists. Such designs are necessary but insufficient for creating a drug that will act in a clinical setting to alleviate human disease. The general ability to "rationally" tailor the other required physical, chemical, and biological properties – specificity, bioavailability, facilitated transport to the site of action, lack of toxicity, long duration of action, chemical stability, etc. – is not in hand.

What is a pharmaceutical researcher, then, to do? Clearly, one should obtain as much information – chemical, biological, physical, and structural – about the system as possible. But one would also be wise to assume there is much that is not known about the system, that might be revealed by testing the effect of large numbers of compounds.

2. Generation of Large Combinatorial Libraries

Pharmacopeia has developed ECLiPS™ technology[1] that allows relatively large combinatorial libraries – typically having 50,000 to 200,000 members – to be conveniently prepared and tested, and the actives quickly identified. Indeed, total compounds prepared in the company's five years of existence totals over 4 million.

The advantages of using such large libraries are severalfold:

- In our experience, there is a significantly increased chance of discovering active compounds, even for difficult targets. This is not surprising; structure-activity relationship (SAR) correlations are incomplete, discontinuous and imperfect, and serendipitous discovery of unanticipated activity remains a major source of new leads. Even when a quantitative SAR has been obtained, jumps in activity may still be found, and enzyme-ligand crystal structures routinely reveal binding patterns which were not anticipated from the binding pattern of analogs or even from computational models.

- Quite often one discovers multiple, chemically unrelated actives. This is advantageous in quickly exploring how much diversity is tolerated by the system, and offers alternative leads for development. This is important: promising leads often run into bioavailability, toxicity or stability roadblocks, or may already be claimed in patents, and an alternative chemical class of leads may offer the only way out of a dead-end project.
- Structure-activity trends are quickly and completely revealed. A large library may be designed to ensure maximum diversity at each site of variation; it can allow preparation of all permutations of all R-groups in all positions for exhaustive structure-activity analysis; and it may reveal much data on inactives which helps define the scope of any correlation.

3. Data Systems for Handling Large Combinatorial Libraries

The synthesis of a combinatorial library requires careful attention to working out the chemistry, choosing a rational set of substituents at each variable position, and generating the library – as indicated in Figure 1.

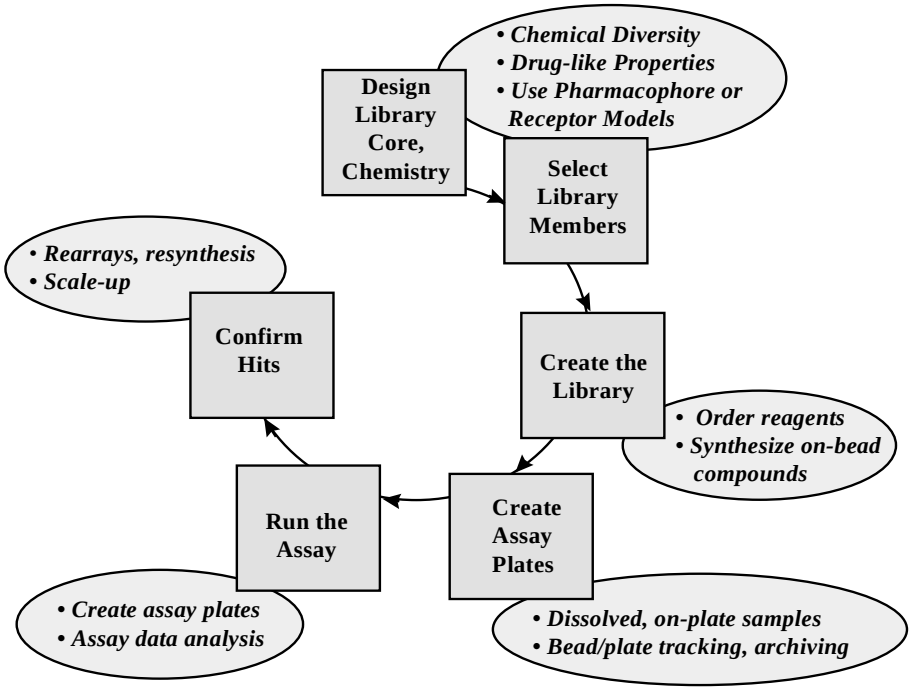

Figure 1. Simplified workflow for creation of a Pharmacopeia ECLiPS™ solid-phase library

Pharmacopeia has developed a set of programs and databases, called Pharmacopeia Information Environment (PIE)[2], to support the design and creation of our combinatorial libraries; the preparation and delivery of assay plates; the follow-up and identification of active leads; and the creation of follow-up libraries (Figure 2). PIE contains chemical and assay data on our 4 million compounds and libraries.

A key component of PIE is our internally developed program for specifying and registering libraries, termed LibDraw. This client-server application allows the chemist to use the commercial structure drawing package, ChemDraw, to draw the library and the component fragments; then registers the members into the PIE database. Released this spring, Pharmacopeia chemists now use this program for defining all combinatorial libraries.

Pharmacopeia uses computational tools to design, specify and register the combinatorial libraries; confirm and identify actives; assist in optimizing leads; and track and analyze the assay results. Several Pharmacopeia projects benefit from available structural information about the target receptor, which information is taken into account during library design.

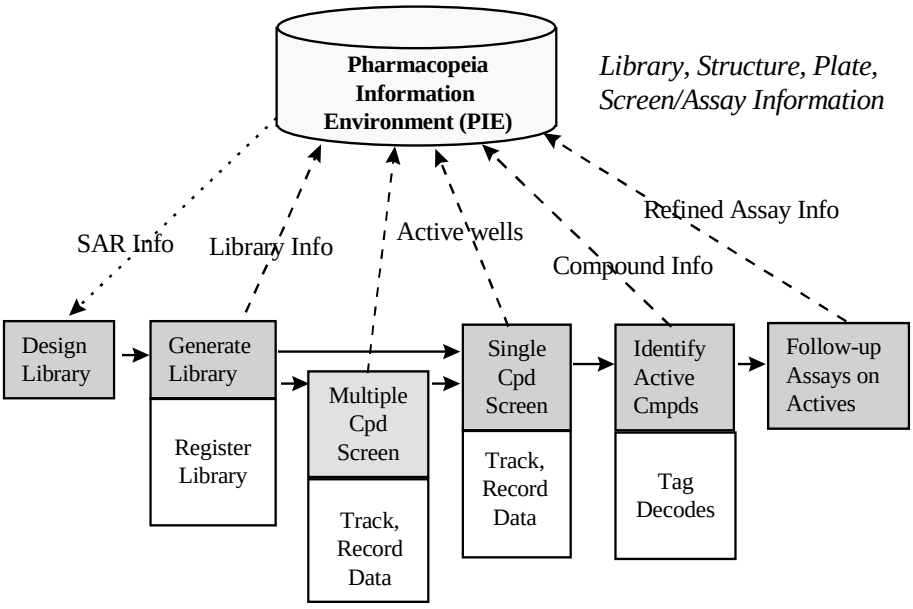

Figure 2. The Pharmacopeia Information Environment (PIE)

Conclusion

We do not see the empirical and theoretical approaches to drug design to be mutually exclusive. Rather, we believe all available information should be used to guide sample preparation into likely chemical domains. Nevertheless, the preparation and testing of very large numbers of combinatorial compounds has been quite effective in generating new and unanticipated information – and has led to the discovery of many active lead structures.

We continue to refine our methods for design of combinatorial libraries and, with our newly acquired partner, Molecular Simulations, expect to increase our use of computational approaches to library design.

References

1. Baldwin, J.J. & Henderson, I. (1997) *Synthesis Of Encoded Small Molecule Combinatorial Libraries Via ECLiPS*™. In J.P. Devlin (Ed.), High Throughput Screening: The Discovery of Bioactive Substances. (pp. 167-190). Marcel Dekker,Inc., New York.

2. *"Technique For Representing Combinatorial Chemistry Libraries Resulting From Selective Combination of Synthons"* - Pharmacopeia, U.S. Patent Pending.

PIE is implemented using Oracle and Daylight tools, hosted on Sun servers.