Review

Strategies for Indirect Computer-Aided Drug Design

Gilda H. Loew,1,2 Hugo O. Villar,1 and Ibon Alkorta1

INTRODUCTION

Computers have become powerful tools in all areas of scientific research. Pharmaceutical and medicinal chemistry are no exception to the rule. They have profited from the use of the methods of theoretical chemistry to understand the structure and mechanism of action of biological systems, as well as to design new compounds that can be used to further this understanding or be investigated as potentially useful therapeutic agents.

The appropriate strategy to use in the design of novel drugs depends on the available knowledge about the structure of the macromolecular target. A “direct” strategy can be used if the three-dimensional structure of the binding sites is known, allowing explicit characterization of ligand–receptor interactions, for example, for the design of drugs that are enzyme inhibitors, since there are many enzymes with known structures (1–3). Such knowledge can be achieved either from appropriate experimental techniques, such as X-ray crystallography or NMR, or from homology modeling that uses theoretical tools to deduce the three-dimensional (3D) structure of a protein given structural data for a highly homologous one (4). If the 3D structure of the macromolecule is not known, then the clues for the design of new ligands for it are more “indirect” and are based on the analysis of the molecular properties of compounds known to have some interaction with it, resulting in diverse pharmacological activities.

In this review, we focus on strategies appropriate for the design of ligands when the 3D structure of the biological target is not known as is the case for CNS active drugs. Two qualitatively different “indirect” approaches are described here. One, called 2D-QSAR, is briefly reviewed. It is based on delineating regression relationships between a specified biological end point and properties of the compounds eliciting it. The other, based on pharmacophore development, constitutes the main part of this review. Several levels of pharmacophore development are described, which differ in the extent to which they encompass fundamental molecular properties that are determinants of receptor recognition and activation. The strengths and limitations of each procedure are discussed and illustrated by examples. Two methods for obtaining model receptor structures are then briefly described. Both rely on the prior success of the indirect methods in obtaining ligand properties that modulate receptor recognition and activation. These emerging capabilities have the potential to bridge the gap between indirect and direct methods of drug design, since, if successful, the design process can continue in a direct mode using explicit characterization of drug–receptor interactions. Strategies for hypothesis validation and use of hypothesis for drug design and discovery are also briefly reviewed. The final sections of this review describe specific computational tools such as molecular mechanics and quantum mechanical methods used to characterize and identify relevant molecular properties and indicate some areas for future development of computational chemistry methods that could increase its effectiveness in the design of novel drugs.

KEY WORDS: drug design; pharmacophore development; QSAR; molecular mechanics; quantum mechanics.

1 Molecular Research Institute, 845 Page Mill Road, Palo Alto, California 94304.
2 To whom correspondence should be addressed.
The overall strategy involved in indirect design of ligands is summarized in Fig. 1. In any such approach, initial pharmacological data for hypothesis development must be obtained for a set of ligands for the system for which novel compounds are to be designed. The initial data should be homogeneous, i.e., obtained using uniform protocols and, ideally, from a single source. Otherwise, the data could have differences that may be misleading. If the data set contains only binding data for antagonists, then the hypothesis can include only molecular requirements for recognition. If the data set contains binding data only for active compounds, then the hypothesis will encompass molecular determinants for both recognition and activation but will not be able to distinguish between them. Only if the data set contains agonists and antagonists, identified as such by an activation end point, and that have different affinities for the receptor, will it be possible for the hypothesis to encompass separate determinants of recognition and activation. In parallel with the experimental effort, the techniques of computational chemistry should be used to calculate molecular properties of the same compounds. From an analysis of the relationship between the molecular properties and the pharmacological profile for each compound, a hypothesis is developed, which, in turn, permits the selection or design of novel ligands for the macromolecular target.

The validity of the initial hypothesis of the mechanism by which the compounds elicit their effect can be verified by acquisition or synthesis of the compounds selected or designed and their subsequent pharmacological evaluation. If the compounds tested have the predicted pharmacological profile, then they could be novel probes of mechanism or clinically useful drugs. Alternatively, if the compound does not have the profile expected, then the results can still be used to refine working hypothesis.

In the following section of this Review, different approaches to hypothesis development are described, and their strengths and limitations are discussed and illustrated by examples. In the next section, strategies for hypothesis validation and use of initial hypothesis for drug design and discovery are briefly reviewed.

In the fourth section, two strategies for obtaining model receptor structures are briefly described. Both rely on the prior success of the indirect methods in obtaining ligand properties that modulate receptor recognition and activation. These emerging capabilities have the potential to bridge the gap between indirect and direct methods of drug design, since, if successful, the design process can continue in a direct mode using explicit characterization of drug–receptor interactions.

The next part of this Review describes several specific computational tools, specifically, molecular mechanical, and quantum mechanical methods used to characterize and identify relevant molecular properties. Finally, we indicate some areas for improvement of computational chemistry methods that could increase their effectiveness when applied to design of novel compound.

This Review has three aims: (i) to serve as an introductory guide to how a working hypothesis can be developed if the 3D structure of the biological target is not known, (ii) to describe the tools required for such a purpose, and (iii) to indicate how to proceed to the design of drugs once the initial hypothesis has been formulated. The literature referenced is by no means exhaustive and it is provided with the main purpose of providing introductory material to the field.

HYPOTHESIS DEVELOPMENT

2-D QSAR

A typical 2D-QSAR procedure assumes that, under certain conditions, the relationship between a biological end point and the molecular properties that determine it can be described in terms of a linear-free energy equation, for any congeneric set of drugs (5–7). A typical equation has the form

\[
\log A = c_h f_{\text{Hyd}}(X_h) + c_e f_{\text{elec}}(X_e) + c_s f_{\text{st}}(X_s) + \text{constant}
\]

where \(A\) is related to either the receptor binding affinity or a specific biological activity, and each of the terms is a congeneric property that can affect either receptor recognition or activation. Typically, hydrophobic \(f_{\text{Hyd}}\), electronic \(f_{\text{elec}}\), and steric \(f_{\text{st}}\) properties of the ligands are used. Each term is a function of the corresponding parameter, \(X\), which may have a linear or quadratic representations.

Although a large number of parameters have been used in connection with this approach, the most widely used set is (i) the octanol water partition coefficient for a hydrophobic term, (ii) Taft \(E_a\) quantities for steric effects, (iii) Hammett constants to describe electronic effects, and (iv) the molar refractivity to account for dispersion forces. Since these are all empirical parameters, this 2D-QSAR or "Hansch" ap-