Pseudoreceptor Modeling: The Construction of Three-Dimensional Receptor Surrogates

Angelo Vedani,*,† Peter Zbinden,† James P. Snyder,‡ and Paulette A. Greenidge§

Contribution from the Biographics Laboratory, Swiss Institute for Alternatives to Animal Testing (SIAT), Missionsstrasse 60, CH-4055 Basel, Switzerland, Istituto di Richerche di Biologia Moleculare (IRBM), P. Angeletti S.p.A., Via Pontina Km. 30.600, I-00040 Pomezia (Roma), Italy, and Department of Pharmacy, ETH Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Received December 6, 1994[⊗]

Abstract: Pseudoreceptor modeling allows the construction of a receptor surrogate for a structurally uncharacterized bioregulator (an enzyme or receptor) based on the structures of known ligand molecules. Although, in general, a pseudoreceptor and its natural counterpart will bear little structural resemblance, they should accommodate a series of ligand molecules in a relatively similar binding sense. A pseudoreceptor validated using a representative series of ligand molecules may subsequently be used to estimate relative free energies for binding for novel ligand molecules. A pseudoreceptor-modeling concept developed at our laboratory allows the generation of a three-dimensional peptidic receptor model (a miniprotein) about any molecular framework of interest. The concept was validated by constructing pseudoreceptors for the enzyme human carbonic anhydrase, the dopaminergic receptor, and the β_2-adrenergic receptor. Predicted differences in free energy of ligand binding toward the pseudoreceptor, $\Delta(\Delta G°_{calc})$, and experimental values determined toward the biological receptor, $\Delta(\Delta G°_{exp})$, agree to within 0.6 and 1.2 kcal/mol.

Introduction

A prerequisite for stimulating a side-effect free response to a drug is believed to be the stereospecific and selective binding of a small molecule to a bioregulatory macromolecule. Soluble enzymes, antibodies, receptor proteins embedded in cell-protective lipid membranes, glycosylated marker proteins projecting similarly from enveloped viruses, ion channels, and various forms of RNA and DNA represent some of the regulatory targets. The contribution of computational methodologies to the identification of small molecules binding to such bioregulators takes many forms. The modeling component is ordinarily represented by approaches that fall into the categories of *receptor fitting* or *receptor mapping*. In the former instance, the three-dimensional structure of the bioregulator or a close homolog is known. At the other extreme, a paucity of information concerning receptor structures has spawned techniques that project the properties of a sample of bioactive ligands into three dimensions about their appropriately superimposed molecular framework. The resulting map provides steric, electrostatic, and lipophilic profiles for subsequent molecular modeling.

More recently, a number of efforts have begun to explore the bridge between structure-based fitting and property-based mapping. Receptor models of various types have been constructed around single or multiple ligands with predefined geometries.[1-6] The corresponding complexes are termed *pseudoreceptors* and permit the receptor-mapping results to be exploited in a receptor-fitting context, for example, the examination of key ligand–receptor interactions, the simulation of molecular dynamics, and the estimation of relative free binding energies.

Momamy and co-workers were among the first to use the term pseudoreceptor and to apply it to a model of an unknown bioregulator.[2] Still and co-workers have combined both computational design and laboratory synthesis in an attempt to generate synthetic enantioselective peptidic pseudoreceptors.[3] Snyder and Rao made use of highly potent cyclopropyl glutamic acids to develop an *ad hoc* NMDA (*N*-methyl-D-aspartate) pseudoreceptor. Subjected to a free-energy perturbation treatment, the polycyclic receptor surrogate was shown to be capable of reproducing relative free binding energies to within 0.3 and 0.8 kcal/mol for a variety of NMDA agonists.[6]

Klebe compiled *composite crystal-field environments* about different functional groups[7] using small-molecule data retrieved from the Cambridge Structural Database (CSD[8]). The spatial distribution of these groups can subsequently be used to map out putative interaction sites, e.g. about amino acid residues oriented toward the binding pocket of a given protein. The information contained in the different composite crystal field environments can be translated into rules which may serve as guidelines for the automated docking for small-molecule fragments into macromolecular binding sites. This approach has been developed for the *de novo* design of protein ligands rather than for the construction of receptor surrogates and, as a

† Swiss Institute for Alternatives to Animal Testing.
‡ Istituto di Richerche di Biologia Moleculare.
§ ETH Zürich.
⊗ Abstract published in *Advance ACS Abstracts*, April 1, 1995.

(1) Snyder, J. P.; Rao, S. N.; Koehler, K. F.; Vedani, A. In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; ESCOM Science Publishers B. V.: Leiden NL, 1993; pp 336–354.
(2) Momamy, F.; Pitha, R.; Klimkowsky, V. J.; Venkatachalam, C. M. In *Expert Systems Applications in Chemistry*; Hohne, B. A., Pierce, T. H., Eds.; ACS Symp. Ser. No. 408; American Chemical Society: Washington, DC, 1989; pp 82–91.
(3) Hong, J.-L.; Namgoong, S. K.; Bernardi, A.; Still, W. C. *J. Am. Chem. Soc.* **1991**, *113*, 5111–5112.
(4) Snyder, J. P.; Rao, S. N. *Chem. Design Automation News* **1989**, *4*, 13–15.
(5) Frühbeis, H.; Klein, R.; Wallmeier, H. *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 403–418.
(6) Snyder, J. P.; Rao, S. N.; Koehler, K. F.; Pellicciari, R. In *Trends in Receptor Research*; Angeli, P., Gulini, U., Quaglia, W., Eds.; Elsevier Science Publishers: Amsterdam, 1992; pp 367–403.
(7) Klebe, G. *J. Mol. Biol.* **1994**, *237*, 212–235.
(8) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr., Sect. B* **1979**, *B35*, 2331–2339.

0002-7863/95/1517-4987$09.00/0 © 1995 American Chemical Society

consequence, lacks a function for quantitatively evaluating ligand–receptor interactions. However, it is conceivable that the information obtained with this approach could serve as input for pseudoreceptor modeling.

Methods

The philosophy underpinning the *pseudoreceptor concept* is to engage the bound species in sufficient, specific non-covalent binding so as to mimic the essential ligand–macromolecule interactions at the true biological receptor. Although, in general, a pseudoreceptor and its natural counterpart will bear little structural resemblance, they should accommodate a series of ligands in a relatively similar binding sense.[1,6] Generation of a refined pharmacophore model is the first step. In the most limited application, a single rigid molecule might serve as the receptor-mapping template. More frequently, an ensemble of structurally similar but conformationally mobile active analogs are composed into a pharmacophore based on predetermined alignment rules.[1-6] A subsequent step involves the construction of a pseudoreceptor, an explicit molecular binding pocket, accommodating a series of ligand molecules in their bioactive conformation. After validation, the pseudoreceptor may be used for receptor-fitting purposes.

As a preliminary to the discussion, it is useful to make a distinction between two qualitatively different constructs that have attempted to unite the fitting/mapping extremes. On the one hand, active-site ligands have been surrounded with separate small molecules as surrogates for receptor side chains. We refer to the unconnected set of contact groups with the ligand as a *minireceptor*. On the other hand, some complexes employ a fully-linked polypeptide or nonpeptide receptor to create a binding pocket. These entities we designate as *pseudoreceptors*. Obviously, hybrid model receptors with intermediate degrees of connectivity are possible.

Directionality of Molecular Interactions. Essential elements to mimic the true biological receptor at the molecular level include ion-pair formation, metal–ligand interactions, hydrogen bonding, and hydrophobic clustering. As the forces controlling these phenomena are more or less directional in nature (cf. below), a strong orientational component had to be incorporated into the mapping concept.[1,9]

Murray-Rust and Glusker as well as Taylor and Kennard were among the first to analyze the directionality of hydrogen bonds in organic molecules. Using data retrieved from the CSD, they showed that H-bond donors are concentrated in directions commonly ascribed to the lone-pair orbital of the O acceptor atom.[10,11] Dunitz and Vedani studied H-bonding patterns at hydroxyl and sulfonamide O atoms as well as N-acceptor atoms in aromatic five- and six-membered rings,[12] confirming the results from refs 10 and 11 but observing an even stronger directionality at aromatic N-acceptor atoms. Alexander and co-workers investigated the stereochemistry of phosphate–Lewis acid interactions with particular reference to different interaction schemes of metal ions and H-bond donors interacting with the phosphate moiety.[13] Baker and Hubbard published a detailed study on the geometry of hydrogen bonds in high-resolution protein structures (retrieved from the Brookhaven Protein Data Bank (BPDB)[14]) where they observed a similar but less pronounced directionality.[15]

To allow for directionality of hydrogen bonds in molecular simulations, Vedani, Dunitz, and Huhta proposed an extended potential function, including as an additional variable the angular deviation of the hydrogen bond from the closest lone-pair direction at the H-bond acceptor atom.[12,16] The term defining the penalty for deviation from ideal directionality was calibrated for a total of 14 different H-bond acceptor types to give the best possible agreement with the corresponding experimental distribution.

Similarly, Vedani and Huhta studied the geometries of metal–ligand complexes, again, using data retrieved from the CSD and proposed a new potential function for modeling metal centers in macromolecules.[16] The new function includes terms for covalent-bond character, metal–ligand separation, symmetry at the metal center, and directionality of metal–ligand interactions. In addition, electrostatic contributions to the total energy are evaluated using a function for dynamical metal–ligand charge transfer.

Directional preferences of hydrophobic interactions were analyzed by Tintelnot and Andrews using data retrieved from the BPDB.[17] The experimental distribution of amino acids around phenyl moieties, for instance, suggested that hydrophobic directionality is clearly less pronounced than the directionality of hydrogen bonds. Klebe and Diederich showed that the structural orientation in host–guest complexes between benzene and cyclophane receptors closely resembles the structural motifs observed in crystalline benzene.[18] Klebe analyzed the spatial distribution of methyl groups about phenyl rings using data retrieved from the CSD and observed a clear agglomeration of methyl groups above and below the center of the benzene ring.[7] Although there would not seem to be a need for an additional term in the force-field energy expression, "hydrophobic directionality" is useful for the initial positioning of hydrophobic amino acid residues in the mapping process (cf. below).

The *Yak* Concept. The pseudoreceptor-modeling software *Yak* (named after the animal native to the high planes of Central Asia), developed at the SIAT Biographics Laboratory, allows the construction of a three-dimensional peptidic pseudoreceptor about any molecular framework of interest, e.g. a pharmacophore.

Prior to the construction of a pseudoreceptor, the molecular framework defining the training set must be prepared. In a pseudoreceptor context, this refers typically to a pharmacophore, an ensemble of ligand molecules superimposed in their bioactive conformation. In general, the following steps are involved in pharmacophore construction: (1) identification of a series of ligand molecules with high affinity (preferably in the nano- to micromolar range) toward the target receptor; (2) retrieval of the ligand structures from a structural database (e.g. the CSD) or, alternatively, assemblage of the ligand structures from molecular fragments thereof, followed by structure refinement; (3) identification of the bioactive conformation by conformational-search procedures; and (4) superposition of the ligand molecules to yield the pharmacophore. Pharmacophore construction is one of the critical steps in the mapping process. Attempts to derive a pseudoreceptor based on an incomplete (not all functional groups necessary to explain the biological activity being present) or an incorrect pharmacophore (some ligands present in other than their bioactive conformation) may not lead to a receptor model consistent with biological data. Pharmacophore construction in general is discussed, for example, in ref 19 or in a pseudoreceptor context in refs 1–6.

As the precision of any force field underlying the mapping process is limited, the ligands defining the training set should preferably span a range of three orders of magnitude in K (the binding constant toward the true biological receptor), corresponding to a difference in free energy of binding, $\Delta(\Delta G°)$, of 4.0 kcal/mol.

In the following, the basic steps of the *Yak* algorithm are discussed.

1. Identification of Functional Groups of the Pharmacophore, Relevant for Receptor Binding. For this purpose, the program generates three types of vectors associated with directional interactions: hydrogen-extension vectors (HEVs), lone-pair vectors (LPVs), and hydrophobicity vectors (HPVs). Hydrogen-extension vectors originate at H-bond donors; their endpoints mark the ideal position for a H-bond acceptor fragment relative to the H-bond donor. Lone-pair vectors originate at H-bond acceptors; their endpoints mark the ideal

(9) Vedani, A.; Zbinden, P.; Snyder, J. P. *J. Recept. Res.* **1993**, *13*, 163–177.
(10) Murray-Rust, P.; Glusker, J. P. *J. Am. Chem. Soc.* **1984**, *106*, 1018–1025.
(11) Taylor, R.; Kennard, O. *Acc. Chem. Res.* **1984**, *17*, 320–326.
(12) Vedani, A.; Dunitz, J. D. *J. Am. Chem. Soc.* **1985**, *107*, 7653–7658.
(13) Alexander, R. S.; Kanyo, Z. F.; Chirlian, L. E.; Christianson, D. W. *J. Am. Chem. Soc.* **1990**, *112*, 933–937.
(14) Bernstein, K.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F., Jr. *J. Mol. Biol.* **1977**, *112*, 535.
(15) Baker, E. N.; Hubbard, R. E. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97–179.
(16) Vedani, A.; Huhta, D. W. *J. Am. Chem. Soc.* **1990**, *112*, 4759–4767.
(17) Tintelnot, M.; Andrews, P. J. *Computer-Aided Mol. Design* **1989**, *3*, 67–84.
(18) Klebe, G.; Diederich, F. *Phil. Trans. R. Soc. London, Ser. A* **1993**, *345*, 37–48.
(19) Ramsden, C. A., Ed. *Quantitative Drug Design, Comprehensive Medicinal Chemistry*; Pergamon Press: Oxford, 1990; Vol. 4.

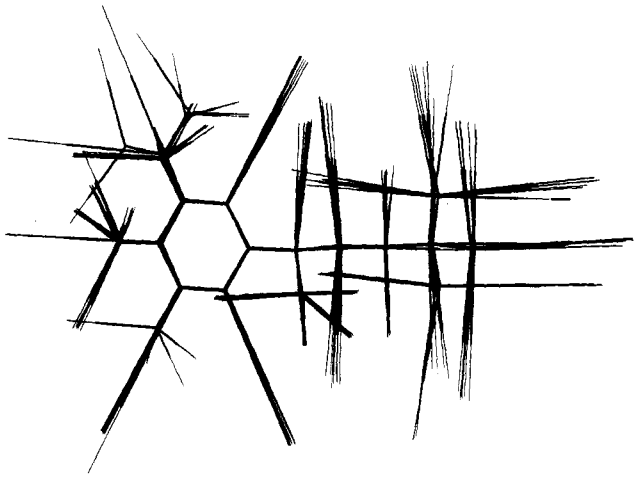

Figure 1. HEVs, LPVs, and HPVs clustering about the pharmacophore used to construct the β_2-adrenergic receptor model. The pharmacophore is rendered in thick-line style and the vectors in thin-line style.

position of a H-bond donor or a metal ion relative to the acceptor atom. Hydrophobicity vectors originate at apolar H atoms; their endpoints mark the approximate position for another hydrophobic moiety relative to the apolar H atom. In three-dimensional space, HEVs, LPVs, and HPVs define the midpoint of the experimental distribution about a given functional group (cf. above and refs 10–13 and 15–17). Next, all vectors of the ligand molecules comprising the pharmacophore are analyzed for clustering in three-dimensional space. The vector density is computed as a function of number, type, separation, and orientation of vectors within the cluster volume, typically 1.0–3.0 Å^3. The idea behind this approach is that HEVs/LPVs/HPVs—in contrast to atoms—are indicative of molecular interactions and, thus, may be used to identify type and approximate position of receptor residues (functional groups), interacting with the ligands at the true biological receptor. Regions of three-dimensional space with a high vector density indicate possibly important interaction sites at the receptor,[20] while spacial zones where low clustering is observed would seem to imply a lower probability for binding to a common set of ligands (cf. Figure 1).

2. Selection of Suitable Receptor Residues (Amino Acids, Metal Ions, Solvent) To Interact with the Functional Groups of the Pharmacophore. This second critical step is supported by a database including frequently observed ligand–receptor residue pairs (compiled from data by Narayana and Argos,[21] augmented by searches in the CSD and BPDB) and by the estimation of the molecular lipophilicity potential.[22,23] For selecting an initial amino acid side-chain conformation, *Yak* makes use of the Ponder–Richards rotamer library.[24] For the simulation of metal-binding sites, the *Yak* database includes Zn^{II}, Mg^{II}, and Ca^{II} templates.

3. Residue Retrieval, Docking, and Optimization. Residue templates are retrieved from the database, automatically docked, and oriented, and the growing pharmacophore–pseudoreceptor complex is optimized by a conformational search protocol combined with energy minimization. For this purpose, *Yak* uses the all-atom type version of the *Yeti* force field[16] and a conjugate-gradient minimizer. This step is repeated until all functional groups of the ligands are saturated or, alternatively, until spatial requirements forbid the further attachment of receptor residues. The receptor may also be extended independently from interactions with the ligands. This allows for introduction of loops, helices, and β sheets.

4. Receptor Solvation. To analyze solvent accessibility of the binding pocket or to identify H-bond networks and solvent channels, the pharmacophore–pseudoreceptor complex may be solvated, i.e. surrounded with both structural water molecules and bulk solvent. The corresponding algorithm is based on ref 25. For the "solvation" of receptors known to occur in a hydrophobic environment (e.g. membrane-bound receptors), we use a *virtual Lennard-Jones liquid* instead, following Kern and co-workers.[26] This liquid consists of spheres of varying size and polarizability, depending on the local molecular lipophilicity potential.[22]

5. Estimation of Differences in Free Energy of Ligand Binding: $\Delta(\Delta G°)$. For each ligand molecule of the pharmacophore, *Yak* calculates an interaction energy, ΔE_{calc}, toward the pseudoreceptor. To derive differences in free energy of ligand binding, $\Delta(\Delta G°)$, from the differences of the individual interaction energies, $\Delta(\Delta E_{calc})$, we make use of an approximation, proposed by Blaney and co-workers.[27] In a pseudoreceptor context, it may be written as:

$$\Delta(\Delta G°_{calc}) \approx \Delta(\Delta E_{calc}) - \Delta(T\Delta S_{binding}) - \Delta(\Delta G_{solv,lig}) \quad (1)$$

This approximation is based on the assumption that the ligands do not bind at or near the receptor surface and, hence, differences in the solvation energy of the ligand–receptor complexes are neglectable. To calculate individual free energies of ligand solvation, $\Delta G_{solv,lig}$, we use the approach of Still and co-workers;[28] $T\Delta S_{binding}$ may be estimated following refs 29 and 30.

A more complex situation arises if the ligands bind near the receptor surface or to a solvent-accessible binding site. In principle, an additional term, $\Delta(\Delta G_{solv,lig-rec})$, might be added to eq 1, correcting for differences in solvation of the individual ligand–pseudoreceptor complexes. However, as (in general) pseudoreceptor and biological receptor share only a few structural elements, this approach would not seem to be promising. Consequently, solvent-accessible receptors represent a limitation to our pseudoreceptor approach if the ligands differ significantly in number and types of functional groups exposed to the solvent.

Quantitatively, a pseudoreceptor hardly discriminates to the same extent between the individual ligands as the true biological receptor. Corrected values may be obtained by means of a linear regression (LR) using $\Delta(\Delta G°_{calc})$ and $\Delta(\Delta G°_{exp})$ data of the training set:

$$\Delta(\Delta G°_{corr}) = \text{slope}^{LR} \cdot \Delta(\Delta G°_{calc}) + \text{intercept}^{LR} \quad (2a)$$

As $\Delta(\Delta G°)$ data are calculated toward a reference compound for which, by definition, $\Delta(\Delta G°)$ is set equal to zero, eq 2a simplifies to:

$$\Delta(\Delta G°_{corr}) = \text{slope}^{LR} \cdot \Delta(\Delta G°_{calc}) \quad (2b)$$

6. Pseudoreceptor Analysis. After completion of the mapping process, it is of great importance to analyze the pseudoreceptor for its biophysical relevance. Among others, criteria include the following: the semiquantitative reproduction of relative free energies of binding for a set of ligand molecules different from the training set, the secondary structure of the pseudoreceptor, the distribution of hydrophilic and hydrophobic residues, and the solvent accessibility of the binding site.

As the *Yak* algorithm generates an *averaged* receptor model about the pharmacophore, the validity of our approach would seem to be limited to receptors where the conformation of the binding pocket does not change specifically upon ligand binding, i.e. to systems where no

(20) Clusters mainly comprised of a single vector type (i.e. functional groups common to all ligand molecules) would seem to indicate interaction sites relevant for recognition, while clusters comprised of different vector types (i.e. portions of the pharmacophore where the individual ligand molecules differ in functionality) may be interpreted as indicating interaction sites responsible for selectivity.

(21) Narayana, S. V. L.; Argos, P. *Int. J. Peptide Protein Res.* **1984**, *24*, 25–39.
(22) Furet, P.; Sele, A.; Cohen, N. C. *J. Mol. Graphics* **1988**, *6*, 182–189.
(23) Ghose, A. K.; Crippen, G. M. *J. Comp. Chem.* **1986**, *7*, 565–577.
(24) Ponder, J. W.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 775–791.
(25) Vedani, A.; Huhta, D. W. *J. Am. Chem. Soc.* **1991**, *113*, 5860–5862.
(26) Kern, P.; Brunne, R. M.; Rognan, D.; Folkers, G. Submitted for publication.
(27) Blaney, J. M.; Weiner, P.; Dearing, A.; Kollman, P.; Jorgensen, E.; Oatley, S.; Burridge, J.; Blake, C. *J. Am. Chem. Soc.* **1982**, *104*, 6424–6434.
(28) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
(29) Searle, M. S.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 10690–10697.
(30) Andrews, P. R.; Craik, D. J.; Martin, J. L. *J. Med. Chem.* **1984**, *27*, 1648–1657.

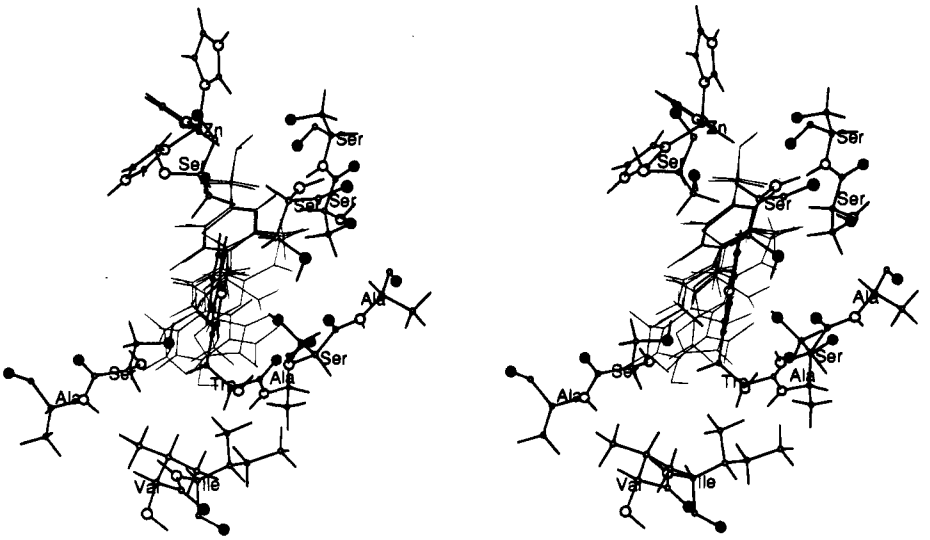

Figure 2. Stereoview of the carbonic anhydrase pseudoreceptor. O atoms are represented as filled circles, N atoms as medium-size open circles, and C atoms as small open circles. Amino acid residues are labeled at the α-C atom position. For clarity, the pharmacophore is rendered in thin-line style with atom labels omitted.

induced-fit occurs. Unfortunately, this information is not available when pseudoreceptor modeling is needed. Induced fit may be simulated starting from any averaged receptor model; the transfer of corresponding results to the biological receptor, however, remains questionable as the surrogate and its natural counterpart, in general, share only a few structural elements (cf. below).

Results and Discussion

To test the *Yak* concept and algorithm, we have performed a series of simulations, aimed at reproducing relative free energies of ligand binding, $\Delta(\Delta G°)$. In this paper, we discuss pseudoreceptors for the enzyme carbonic anhydrase, the dopaminergic receptor, and the β_2-adrenergic receptor, respectively. Each of the receptor models was constructed using a representative series of ligand molecules and subsequently tested using ligand molecules different therefrom.

Carbonic Anhydrase. Carbonic anhydrase, a zinc-containing enzyme, is an extremely efficient catalyst of the reversible hydration of carbon dioxide. The crystal structure of the native enzyme has been determined to a resolution of 2.0 Å by Kannan and co-workers.[31,32] Various mechanistic models for the catalytic reaction have been put forward based on experimental and theoretical data.[32-37,16] The structure of the complex between carbonic anhydrase and the sulfonamide inhibitor 2-acetamido-1,3,4-thiadiazole-5-sulfonamide was determined to a resolution of 3.0 Å.[31,38] Clinical applications of the inhibition of carbonic anhydrase focus on the treatment of glaucoma, epilepsy, and acute mountain sickness.[31,32]

The construction of a receptor surrogate was based on a total of nine sulfonamide inhibitors (cf. Table 1). The structures of the individual ligand molecules were either retrieved from the CSD or generated using molecular fragments thereof. All inhibitor molecules contain either an aromatic or a heterocyclic ring. The possible conformations of their (often hyperconjugated) side chains were assessed through searches in the CSD. Protonation state and stereochemistry of the sulfonamide group were adapted from refs 39 and 40. Experimental free energies of ligand binding, $\Delta G°_{exp}$, were derived from thermodynamic and kinetic data.[41-43]

Solvation energies of the individual ligand molecules were calculated using a semianalytical approach following Still et al.;[28] the atomic partial charge model was determined using the method of Rappe and Goddard.[44] The inhibitor molecules were superimposed using five atoms as reference points[45] with acetazolamide (2-acetamido-1,3,4-thiadiazole-5-sulfonamide) serving as the template. Relative free energies of ligand binding were calculated according to eq 1 and corrected according to eq 2b. Although in this case the true biological receptor (carbonic anhydrase) has been determined to atomic resolution by means of X-ray crystallography,[31,32] the only information used for the mapping process was that (1) carbonic anhydrase is a zinc-requiring enzyme and (2) that a reversible proton-relay network is essential for the catalytic mechanism.

The pseudoreceptor consists of a Zn^{II} template (Zn[imidazole]$_3$) and 12 amino acid residues, representing a truncated protein. The pharmacophore–pseudoreceptor complex is shown in Figure 2; calculated and experimental differences in free energy of ligand binding are compared in Table 1. The correlation coefficient for $\Delta(\Delta G°_{exp})$ vs $\Delta(\Delta G°_{calc})$ is 0.989, indicating an excellent agreement between calculated and experimental differences in free energy of ligand binding. The ranking is correctly reproduced for all ligands. $\Delta(\Delta G°_{calc})$

(31) Kannan, K. K. In *Biophysics and Physiology of Carbon Dioxide*; Gros, H., Bartels, H., Eds.; Springer: Berlin, 1979; pp 184–205.
(32) Tashian, R. E.; Hewett-Emmett, D., Eds. Biology and Chemistry of the Carbonic Anhydrases. *Ann. N.Y. Acad. Sci.* **1984**, *429*.
(33) Pocker, Y.; Sarkanen, S. *Adv. Enzymol.* **1978**, *47*, 149–274.
(34) Silverman, D. N.; Vincent, S. H. *Crit. Rev. Biochem.* **1983**, *14*, 207–255.
(35) Merz, K.; Hoffmann, R.; Dewar, M. J. S. *J. Am. Chem. Soc.* **1989**, *111*, 5636–5649.
(36) Tu, C.; Silverman, D. N.; Forsman, C.; Jonsson, B. H.; Lindskog, S. *Biochemistry* **1989**, *28*, 7913–7918.
(37) Vedani, A.; Huhta, D. W.; Jacober, S. P. *J. Am. Chem. Soc.* **1989**, *111*, 4075–4081.
(38) Kannan, K. K.; Vaara, I.; Nostrand, B.; Lowgren, S.; Borell, A.; Fridborg, K.; Petef, M. In *Drug Action at the Molecular Level*; Roberts, C. G. K., Ed.; McMillan: London, 1977; pp 73–91.
(39) Mukherjee, J.; Rogers, J. I.; Khalifah, R. G.; Everett, G. W., Jr. *J. Am. Chem. Soc.* **1987**, *109*, 7232–7233.
(40) Everett, G. W., Jr., 1989, personal communication.
(41) Taylor, P. W.; King, R. W.; Burgen, A. S. V. *Biochemistry* **1970**, *9*, 2368.
(42) Kakeya, N.; Aoki, A.; Kamada, A.; Yata, N. *Chem. Pharm. Bull.* **1969**, *17*, 1010.
(43) Sprague, J. M. In *Topics in Medicinal Chemistry*; Rabinowitz, J. L., Myerson, R. M., Eds.; Interscience: New York, 1968; pp 1–63.
(44) Rappe, A. K.; Goddard, W. A., III *J. Phys. Chem.* **1991**, *95*, 3358–3363.
(45) The sulfonamide molecules were superimposed using five reference points: the sulfonamide N and S atom, two C (or N) atoms of the aromatic (heterocyclic) ring and the first C (or N) atom of the side chain. For benzenesulfonamide (lacking a ring substituent), only four reference points were used.

Table 1. Comparison of Calculated and Experimental Relative Free Binding Energies (in kcal/mol) for the Carbonic Anhydrase System with the Reference Compound Given in Italics

ligand[c]	$\Delta\Delta G°_{exp}$	$\Delta\Delta G°_{ave}$	$\Delta\Delta G°_{ind.fit}$	$\Delta\Delta G°_{avd} - \Delta\Delta G°_{exp}$	$\Delta\Delta G°_{ind.fit} - \Delta\Delta G°_{exp}$
			Training Set[a]		
ETZA	0.00	0.00	0.00	0.00	0.00
SABS	+0.69	+0.75	+0.97	+0.06	+0.28
MTZ	+1.16	+0.83	+0.53	−0.33	−0.63
AAA	+1.27	+1.00	+0.85	−0.27	−0.42
DBSA	+2.15	+1.85	+2.08	−0.30	−0.07
SBSA	+2.66	+2.94	+2.34	+0.28	−0.32
MBSA	+3.36	+3.24	+3.46	−0.12	+0.10
BSA	+4.00	+4.09	+4.31	+0.09	+0.31
FURO	+4.70	+4.25	+3.36	−0.45	−1.34
			Test Set[b]		
BAAA	+1.33	+0.36	−0.03	−0.97	−1.36
LBSA	+2.52	+2.80	+2.97	+0.28	+0.45
CBSA	+2.99	+3.60	+3.16	+0.61	+0.17
YBSA	+3.23	+2.34	+2.57	−0.89	−0.66
SAM	+4.75	+3.45	+3.66	−1.30	−1.09

[a] Training set (nine molecules): The correlation coefficient for $\Delta\Delta G°_{exp}$ vs $\Delta\Delta G°_{calc}$ is 0.989 for the averaged receptor model and 0.944 for the induced-fit simulation; the RMS deviation of $\Delta\Delta G°_{calc}$ and $\Delta\Delta G°_{exp}$ is 0.25 and 0.55 kcal/mol, respectively. [b] Test set (five molecules): The RMS deviation of $\Delta\Delta G°_{corr}$ and $\Delta\Delta G°_{exp}$ is 0.88 kcal/mol for the averaged receptor model and 0.86 for the induced-fit simulation, respectively. [c] ETZA: 6-ethoxybenzothiazole-2-sulfonamide (ethoxzolamide). SABS: *p*-(salicyl-5-azo)benzenesulfonamide. MTZ: 2-acetimido-3-methyl-1,3,4-thiadiazole-5-sulfonamide (metazolamide). AAA: 2-acetamido-1,3,4-thiadiazole-5-sulfonamide (acetazolamide). DBSA: 3,5-dichlorobenzenesulfonamide. SBSA: benzene-1,4-disulfonamide. MBSA: 4-methylbenzenesulfonamide. BSA: benzenesulfonamide. FURO: 4-chloro-*N*-furfuryl-5-sulfamoylanthranilic acid (furosemide). BAAA: 2-butylamido-1,3,4-thiadiazole-5-sulfonamide. LBSA: 4-chlorobenzenesulfonamide. CBSA: 4-carboxybenzenesulfonamide. YBSA: 4-cyanobenzenesulfonamide. SAM: 4-aminobenzenesulfonamide (sulfanilamide).

values corrected according to eq 2b lead to a root-mean-square (RMS) deviation of 0.25 kcal/mol.

Five ligand molecules different from the training set were then added to the pseudoreceptor and their relative free energies of binding calculated. The conformation of the pseudoreceptor was not altered in order to obtain $\Delta(\Delta G°_{calc})$ values toward the averaged receptor model. The RMS deviation for these predicted differences is 0.88 kcal/mol, corresponding to an uncertainty of a factor of 4.5 in the binding constant K (cf. Table 1). Within this margin of error, the binding strengths of LBSA, CBSA, and YBSA were predicted correctly, while BAAA and SAM were calculated to bind too strongly.

To allow for an induced-fit mechanism, we split the pharmacophore−pseudoreceptor complex into fourteen (1:1) ligand−pseudoreceptor complexes and refined them individually. To mimick the (non-existing) residual protein, we surrounded the truncated protein with a virtual Lennard-Jones liquid, following Kern et al.[26] Differences in free energy of binding, $\Delta(\Delta G°)$, were calculated from the individual ΔE_{calc} following eq 1 and corrected according to eq 2b. For the training set, the linear regression yielded a correlation coefficient of 0.944 (compared with a value of 0.989 for the averaged receptor model). For the test set, the RMS deviation is 0.86 kcal/mol (0.88 kcal/mol). These results (cf. also Table 1) are not significantly different from those obtained with the averaged receptor model and—in agreement with experimental data[31,32,38]—do not suggest the presence of an induced-fit mechanism.

We have also constructed a single-strand polypeptide pseudoreceptor, using an appropriate algorithm for linking key amino acid residues responsible for recognition and selectivity. This miniprotein consists of a Zn^{II} template (Zn[imidazole]$_3$) and 16 amino acid residues. For the training set, the correlation coefficient is 0.907; the RMS deviations are 0.97 kcal/mol for the training set and 1.75 kcal/mol for the test set. Differences in free energy of ligand binding for ligands BAAA, LBSA, and CBSA are predicted about correctly, while ligands YBSA and SAM were both predicted to bind too strongly.

Dopaminergic Receptor. Dopamine is an important neurotransmitter both in the central nervous system and in peripheral tissues. Malfunction of the dopaminergic system has been proposed to play a major role in diseases like schizophrenia and parkinsonism.[46] The molecular substructure responsible for the dopaminomimetic activity of apomorphines and related molecules—the rigid dopamine—has been established by Cannon and others.[47−50] Böttcher and co-workers identified and synthetized a new class of compounds, 3-(1,2,3,6-tetrahydro-1-pyridylalkyl)indoles, not sharing the structural features of the dopamine agonists.[51] To understand the structure−activity relationship in this class of dopamine agonists, they investigated the structural features necessary for this unique dopamine agonist profile and applied molecular-modeling techniques.

In our study, nine of these 3-pyridylalkyl indoles defined the training set; five were used to test the predictive power of the pseudoreceptor. Model building and superposition of the agonists are described in refs 51 and 52. Solvation energies of the individual ligand molecules were calculated using a semi-analytical approach following Still et al.;[28] the atomic partial charge model was determined using the method of Rappe and Goddard.[44] Experimental free energies of ligand binding, $\Delta G°_{exp}$, were derived from binding data in ref 51. Relative free energies of ligand binding were calculated according to eq 1 and corrected according to eq 2b. The pseudoreceptor consists of 19 amino acid residues and represents a truncated protein. It is shown in Figure 3; experimental and predicted free energies of ligand binding are compared in Table 2.

For the training set, the correlation coefficient for $\Delta(\Delta G°_{exp})$ vs $\Delta(\Delta G°_{calc})$ is 0.954, indicating a very good agreement between calculated and experimental differences in free energy of ligand binding. The ranking of $\Delta(\Delta G°)$ is correctly reproduced for all but one ligand. The RMS deviation of calculated and experimental differences in free energies for the training set, $\Delta(\Delta G°)$, is 0.50 kcal/mol.

Five ligand molecules different from the training set were then added to the pseudoreceptor and their relative free energies of binding calculated. The conformation of the pseudoreceptor was not altered in order to obtain $\Delta(\Delta G°_{calc})$ values toward the averaged receptor model. The RMS deviation for these predicted differences is 1.20 kcal/mol, corresponding to an uncertainty of a factor of 7.8 in the binding constant K (cf. Table 2). Within this margin of error, the binding strengths of C16, C26, and C17 were predicted correctly, while C25 and C11 were calculated to bind too weakly.

The induced-fit simulation yielded a correlation coefficient (for the training set) of 0.949, compared with a value of 0.954 for the *averaged* receptor model. The RMS deviation for the

(46) Kaiser, C.; Jain, T. *Med. Res. Rev.* **1985**, *5*, 145−229.
(47) Cannon, J. G. In *Progress in Drug Research*; Jucker, E., Ed.; Birkhäuser: Basel, 1985; Vol. 29, pp 303−413.
(48) Miller, D. D. *Fed. Proc., Fed. Am. Soc. Exp. Biol.* **1978**, *37*, 2392−2395.
(49) Neumeyer, J. L.; Neustadt, B. R.; Oh, K. H.; Weinhardt, K. K.; Boyce, C. B.; Rosenberg, F. J.; Teiger, D. G. *J. Med. Chem.* **1973**, *16*, 1223−1228.
(50) Neumeyer, J. L.; Dafeldecker, W. P.; Costall, B.; Naylor, R. *J. Med. Chem.* **1977**, *20*, 190−196.
(51) Böttcher, H.; Bamickel, G.; Hausberg, H.-H.; Haase, A. F.; Seyfried, C. A.; Eiermann, V. *J. Med. Chem.* **1992**, *35*, 4020−4026.
(52) Böttcher, H.; Bamickel, G.; Rippmann, F.; Greiner, H. E.; Seyfried, C. A. XIIIth International Symposium on Medicinal Chemistry, Paris, 1994.

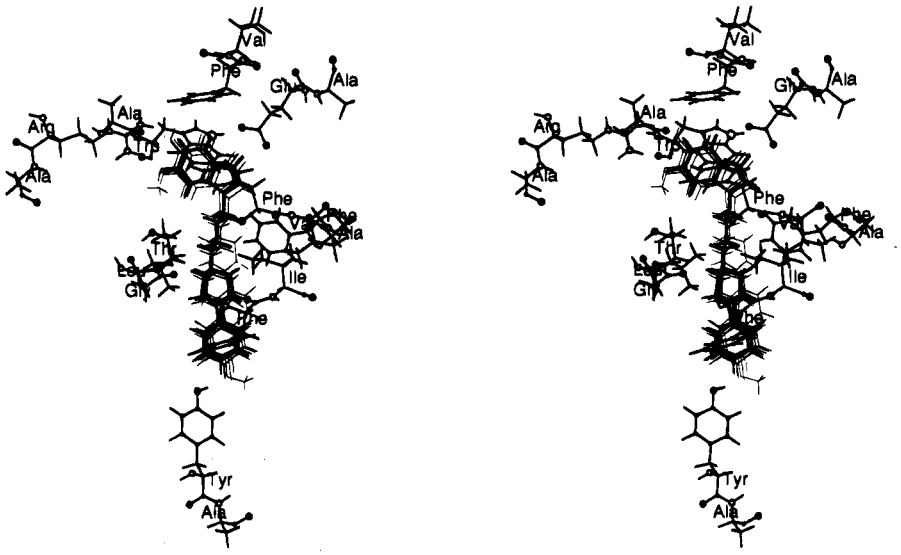

Figure 3. Stereoview of the dopaminergic pseudoreceptor. O atoms are represented as filled circles, S atoms as large open circles, N atoms as medium-size open circles, and C atoms as small open circles. Amino acids residues are labeled at the α-C atom position. For clarity, the pharmacophore is rendered in thin-line style with atom labels omitted.

Table 2. Comparison of Calculated and Experimental Relative Free Binding Energies (in kcal/mol) for the Dopaminergic Receptor System with the Reference Compound Given in Italics

ligand[c]	$\Delta\Delta G°_{exp}$	$\Delta\Delta G°_{ave}$	$\Delta\Delta G_{ind.fit}$	$\Delta\Delta G°_{ave} - \Delta\Delta G°_{exp}$	$\Delta\Delta G°_{ind.fit} - \Delta\Delta G°_{exp}$
			Training Set[a]		
C28	0.00	0.00	0.00	0.00	0.00
C24	+0.35	+1.06	+0.91	+0.71	+0.56
C29	+0.61	+1.35	+0.55	+0.74	−0.06
C28P	+0.91	+1.39	+1.26	+0.48	+0.35
C5	+1.15	+1.40	+1.22	+0.25	+0.07
C15	+1.65	+2.45	+2.21	+0.80	+0.56
C19	+2.24	+1.84	+1.47	−0.40	−0.77
C21	+2.76	+3.05	+2.54	+0.29	−0.22
C20	+4.10	+4.30	+4.18	+0.20	+0.08
			Test Set[b]		
C25	+0.73	+2.37	+1.10	+1.64	+0.37
C16	+0.88	+1.72	+1.48	+0.84	+0.60
C26	+0.94	+1.49	+1.72	+0.55	+0.78
C17	+2.30	+2.66	+2.29	+0.36	−0.01
C11	+2.63	+4.47	+3.34	+1.84	+0.71

[a] Training set (nine molecules): The correlation coefficient for $\Delta\Delta G°_{exp}$ vs $\Delta\Delta G°_{calc}$ is 0.954 for the averaged receptor model and 0.949 for the induced-fit simulation; the RMS deviation of $\Delta\Delta G°_{calc}$ and $\Delta\Delta G°_{exp}$ is 0.50 and 0.40 kcal/mol, respectively. [b] Test set (five molecules): The RMS deviation of $\Delta\Delta G°_{corr}$ and $\Delta\Delta G°_{exp}$ is 1.20 kcal/mol for the averaged receptor model and 0.56 for the induced-fit simulation, respectively. [c] C28: 5-hydroxy-3-[4-(1,2,3,6-tetrahydro-4-phenyl-1-pyridyl)butyl]indole hydrochloride. C24: 5-fluoro-3-[4-(1,2,3,6-tetrahydro-4-phenyl-1-pyridyl)butyl]indole hydrochloride. C29: 6-hydroxy-3-[4-(1,2,3,6-tetrahydro-4-phenyl-1-pyridyl)butyl]indole hydrochloride. C28P: 5-hydroxy-3-[4-(4-phenyl-1-piperazyl)butyl]indole hydrochloride. C5: 3-[4-(1,2,3,6-tetrahydro-4-phenyl-1-pyridyl)butyl]-indole hydrochloride. C15: 3-[4-(1,2,3,6-tetrahydro-4-(4′-fluorophenyl)-1-pyridyl)butyl]indole hydrochloride. C19: 3-[4-(1,2,3,6-tetrahydro-4-(4′-methoxyphenyl)-1-pyridyl)butyl]indole hydrochloride. C21: 3-[4-(4-phenyl-1-piperidyl)butyl]indole hydrochloride. C20: 3-[4-(1,2,3,6-tetrahydro-1-pyridyl-3-ene)butyl]indole hydrochloride. C25: 5-chloro-3-[4-(1,2,3,6-tetrahydro-4-phenyl-1-pyridyl)butyl]indole hydrochloride. C16: 3-[4-(1,2,3,6-tetrahydro-4-(2′-fluorophenyl)-1-pyridyl)butyl]indole hydrochloride. C26: 5-methoxy-3-[4-(1,2,3,6-tetrahydro-4-phenyl-1-pyridyl)butyl]indole hydrochloride. C17: 3-[4-(1,2,3,6-tetrahydro-4-(4′-chlorophenyl)-1-pyridyl)butyl]indole hydrochloride. C11: 3-[2-(1,2,3,6-tetrahydro-4-phenyl-1-pyridyl)ethyl]indole hydrochloride.

test set decreased significantly (0.56 kcal/mol vs 1.20 kcal/mol). This result suggests that induced fit might indeed play a role for the binding of 3-pyridylalkyl indoles to the dopaminergic receptor.

β_2-Adrenergic Receptor. The β_2-adrenergic receptor is a member of the class of G protein-coupled receptors. The rational design of both potent and selective β_2-adrenergic agonists—of particular interest for the clinical treatment of asthma[53]—would be facilitated by the availability of the three-dimensional structure of the receptor-binding pocket. In the past few years, many efforts have been undertaken in order to derive three-dimensional models for integrated membrane proteins (see, for example, ref 54). Recently, much attention has been concentrated upon the building of β_2-adrenergic receptor models in which the X-ray structure of bacteriorhodopsin[55] was used either directly or indirectly, but much controversy exists over the validity of such models.[54,56]

In contrast hereto, we seek to derive the receptor-binding pocket by means of pseudoreceptor modeling. In our study, nine adrenergic agonists defined the training set. The agonists were superimposed upon each other using a RMS overlay-fitting procedure in which ligand AH3 served as the template ligand. The atomic partial charge model was based on MNDO electrostatic potential charges, derived using MOPAC 6.0;[57,58] free energies of ligand solvation were calculated using a semianalytical approach following Still and co-workers.[28] Experimental free energies of ligand binding, $\Delta G°_{exp}$, were taken from ref 59.[60] Relative free energies of ligand binding were calculated according to eq 1 and corrected according to eq 2b. The pseudoreceptor consists of 15 amino acid residues and represents a truncated protein. It is shown in Figure 4; experimental and predicted free energies of ligand binding are compared in Table 3.

The correlation coefficient for $\Delta(\Delta G°_{exp})$ vs $\Delta(\Delta G°_{calc})$ is 0.919, indicating a good agreement between calculated and

(53) Main, B. G. In *Comprehensive Medicinal Chemistry*; Emmett, J. C., Volume Ed.; Pergamon Press: Oxford, 1990; pp 187−228.
(54) Kontoyianni, M.; Lybrand, T. P. *Perspect. Drug Discovery Design* **1993**, *1*, 291−300.
(55) Henderson, R.; Baldwin, J. M.; Ceska, T. A.; Zemlin, F.; Beckmann, E.; Downing, K. M. *J. Mol. Biol.* **1990**, *213*, 899−923.
(56) Hoflack, J.; Trumpp-Kallmeyer, S. *TIPS* **1994**, *15*, 7−9.
(57) Distributed by QCPE, University of Indiana, Bloomington, IN.
(58) Stewart, J. J. P. *J. Comp.-Aided Mol. Design* **1990**, *4*, 1−105.
(59) Donné-Op den Kelder, G. M.; Bultsma, T.; Timmerman, H. *J. Med. Chem.* **1988**, *31*, 1069−1079.
(60) For constructing our model, we used the set of high-affinity binding constants corrected for differently active isomers and ionic forms of the β_2-adrenergic agonists (cf. Table 1 in ref 59).

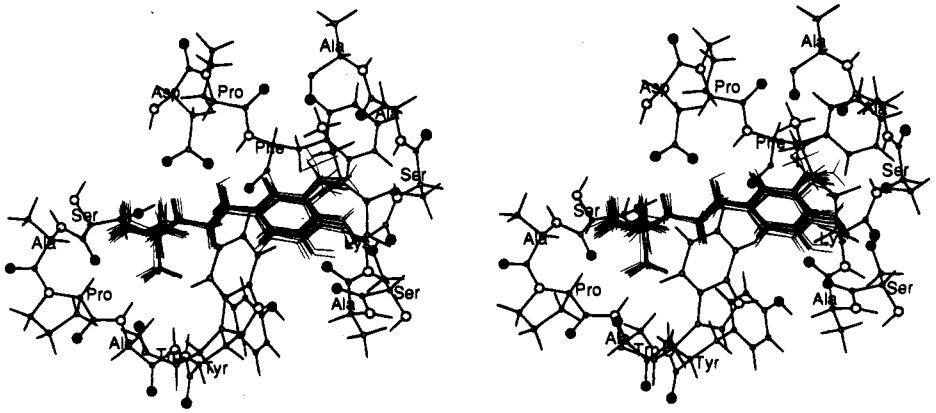

Figure 4. Stereoview of the β_2-adrenergic pseudoreceptor. O atoms are represented as filled circles, N atoms as medium-size open circles, and C atoms as small open circles. Amino acids residues are labeled at the α-C atom position. For clarity, the pharmacophore is rendered in thin-line style with atom labels omitted.

Table 3. Comparison of Calculated and Experimental Relative Free Binding Energies (in kcal/mol) for the β_2-Adrenergic Receptor System with the Reference Compound Given in Italics

ligand[c]	$\Delta\Delta G°_{exp}$	$\Delta\Delta G°_{ave}$	$\Delta\Delta G°_{ind.fit}$	$\Delta\Delta G°_{ave} - \Delta\Delta G°_{exp}$	$\Delta\Delta G°_{ind.fit} - \Delta\Delta G°_{exp}$
			Training Set[a]		
AH3	0.00	0.00	0.00	0.00	0.00
TER	+0.27	+0.43	+0.40	+0.16	+0.13
NAB	+0.29	+0.67	+0.47	+0.38	+0.18
TBF2	+0.67	+0.36	+0.24	−0.31	−0.43
ISO	+0.86	+0.97	+1.05	+0.11	+0.19
SAL	+1.31	+0.54	+0.55	−0.77	−0.76
NIS	+1.39	+1.64	+1.73	+0.25	+0.34
ADR	+2.06	+1.81	+1.61	−0.25	−0.45
NOR	+2.66	+2.47	+2.39	−0.19	−0.27
			Test Set[b]		
CLB	−0.47	+0.53	+0.40	+1.00	+0.87
ORC	+1.26	+0.90	+0.84	−0.36	−0.42
2CL	+1.48	+1.60	+1.48	+0.12	0.00
DU3	+1.55	+0.66	+0.60	−0.89	−0.95
SKF	+2.39	+2.11	+1.55	−0.28	−0.84
ISOP	+3.33	+3.08	+3.05	−0.25	−0.28

[a] Training set (nine molecules): The correlation coefficient for $\Delta\Delta G°_{exp}$ vs $\Delta\Delta G°_{calc}$ is 0.919 for the averaged receptor model and 0.912 for the induced-fit simulation; the RMS deviation of $\Delta\Delta G°_{calc}$ and $\Delta\Delta G°_{exp}$ is 0.34 and 0.37 kcal/mol, respectively. [b] Test set (six molecules): The RMS deviation of $\Delta\Delta G°_{corr}$ and $\Delta\Delta G°_{exp}$ is 0.59 kcal/mol for the averaged receptor model and 0.66 for the induced-fit simulation, respectively. [c] AH3: 1-(3-amido-4-hydroxyphenyl)-2-(*tert*-butylamino)ethanol. TER: 1-(3,4-dihydroxyphenyl)-2-(*tert*-butylamino)ethanol. NAB: 1-(4-amino-3,5-dichlorophenyl)-2-(isopropylamino)ethanol. TBF2: 1-(3,5-dihydroxyphenyl)-2-(*tert*-butylamino)ethanol; ISO: (−)-1-(3,4-dihydroxyphenyl)-2-(isopropylamino)ethanol. SAL: 1-(4-hydroxy-3-(hydroxymethyl)phenyl)-2-(*tert*-butylamino)ethanol. NIS: 1-(3-hydroxyphenyl)-2-(isopropylamino)ethanol. ADR: 1-(3,4-dihydroxyphenyl)-2-(methylamino)ethanol. NOR: 1-(3,4-dihydroxyphenyl)-2-aminoethanol. CLB: 1-(4-amino-3,5-dichlorophenyl)-2-(*tert*-butylamino)ethanol. ORC: 1-(3,5-dihydroxyphenyl)-2-(isopropylamino)ethanol. 2Cl: 1-(2-chlorophenyl)-2-(*tert*-butylamino)ethanol. DU3: 1-(4-hydroxy-3-aminophenyl)-2-(*tert*-butylamino)ethanol. SKF: 1-(4-hydroxy-3-(aminomethyl)phenyl)-2-(*tert*-butylamino)ethanol. ISOP: (+)-1-(3,4-dihydroxyphenyl)-2-(isoproylamino)ethanol.

experimental differences in free energy of ligand binding. The ranking is correctly reproduced for all but two ligands. The RMS deviation of calculated and experimental differences in free energies for the training set, $\Delta(\Delta G°)$, is 0.34 kcal/mol.

Six ligand molecules different from the training set were then added to the pseudoreceptor and their relative free energies of binding calculated. The conformation of the pseudoreceptor was not altered in order to obtain $\Delta(\Delta G°_{calc})$ values toward the averaged receptor model. The RMS deviation for these predicted differences is 0.59 kcal/mol, corresponding to an uncertainty of a factor of 2.7 in the binding constant K (cf. Table 3). Within this margin of error, the binding strengths of ORC, 2CL, SKF, and ISOP were predicted correctly, while CLB and DU3 were calculated to bind too weakly and too strongly, respectively. Remarkably, the relative free energy of binding for ISOP, a stereoisomer of ISO, is predicted to within 0.25 kcal/mol of the experimental value.

The induced-fit simulation yielded a correlation coefficient (for the training set) of 0.912, compared with a value of 0.919 for the *averaged* receptor model. The RMS deviation for the test set decreased slightly (0.59 kcal/mol vs 0.66 kcal/mol). These results do not suggest the presence of an induced fit for the β_2-adrenergic receptor.

Limitations of Pseudoreceptor Modeling. Our pseudoreceptor approach would presently seem to be limited by four factors: the pharmacophore model, the presence of a ligand-specific induced-fit mechanism, solvation phenomena, and the underlying force field.

As we have outlined above, a validated pharmacophore model is the key element for deriving a receptor surrogate, able to reproduce relative free energies of binding in a semiquantitative fashion. Many approaches have been validated for pharmacophore construction (cf., for example, ref 19), therefore, this should not represent a principal limitation of the pseudoreceptor concept. On the other hand, the possibility of a ligand-specific induced-fit mechanism might seriously interfere with the philosophy behind an *averaged* receptor model as, in general, a given pseudoreceptor and its natural counterpart share only a few structural elements. Although our study showed that better results were obtained for the dopaminergic receptor when allowing for induced fit (the results for carbonic anhydrase and the β_2-adrenergic receptor were indifferent in this respect), it may not be safe to conclude that such a mechanism is relevant for binding to the biological receptor. The presence of solvation phenomena could jeopardize the validity of eq 1 or, at least, increase the systematic error in the estimation of relative free binding energies.

As in any other biophysical simulation, the quality of the underlying force field is the ultimate discriminator. The force field used in *Yak* includes directional terms for hydrogen bonds and metal−ligand interactions. Although they would seem to be mandatory for the mapping procedure, we are aware that our force field[16] does not handle all interaction phenomena equally well. For example, it lacks general terms for polarization and (except for metal−ligand interactions) charge transfer.

In summary, the "RMS resolution" of our approach might lie in the range of 0.8 to 1.2 kcal/mol, corresponding to an

uncertainty in the binding constant of a factor of 4.0 to 8.0. However, in the context of *drug design*, an uncertainty of a factor of 10 in the prediction of relative binding constants for novel ligand molecules would seem to be acceptable, particularly since experimental binding data (e.g. from cell lines) are often associated with similarly large errors. Finally, a pseudoreceptor might be used as input for free-energy perturbation simulations, where semiquantitative agreements have been achieved as accurate as 0.3 kcal/mol.[6]

Conclusions

Pseudoreceptor modeling provides a means for bridging structure-based receptor fitting and property-based receptor mapping by allowing the construction of a three-dimensional receptor surrogate, based solely on the structures of known ligand molecules. Although, in general, a pseudoreceptor and its natural counterpart share only a few structural elements, they should accommodate a series of ligand molecules in relatively similar binding sense. Validated pseudoreceptors can, therefore, be used to estimate relative free energies, $\Delta(\Delta G^\circ)$, for novel ligand molecules.

Based on the directionality of molecular interactions we have developed a concept and an algorithm to construct such three-dimensional receptor surrogates. Our approach has been validated by constructing models for the binding site of the enzyme carbonic anhydrase, the dopaminergic, and the β_2-adrenergic receptor. For a small series of test molecules, predicted differences in free energy of ligand binding toward the pseudoreceptor were shown to agree to within 0.88, 1.20, and 0.59 kcal/mol with the experimental data (determined toward the true biological receptor), respectively.

Upon request, the coordinates of the models described in this paper are available for distribution. Details for the distribution of the program should be requested from the corresponding author.

The Swiss Institute for Alternatives to Animal Testing (SIAT) foundation is dedicated to research, education, and consultation in the field of **Alternatives to Animal Testing** according to the **3Rs** (refinement, reduction, replacement; cf. refs 61 and 62).[63] The SIAT is a non-profit organization, controlled by a Board of Trustees and the Swiss Department of the Interior.

Acknowledgment. We express our gratitude to Professors Max Dobler and Hans Dutler (Department of Organic Chemistry, ETH Zürich) and to Professor Gerd Folkers (Department of Pharmacy, ETH Zürich) for both challenging and stimulating discussions. We are much indebted to Dr. Gerhard Barnickel (E. Merck Darmstadt/Germany, Preclinical Pharmaceutical Research) for providing us with the pharmacophore model for the dopaminergic receptor system. Financial support from the Foundation SIAT (Zürich, Switzerland), the *Margaret and Francis Fleitmann Foundation* (Lucerne, Switzerland), and the Swiss National Science Foundation (Projects 31-32395.91 and 31-39229.93) is gratefully acknowledged.

JA9439480

(61) Russel, W. M. S.; Burch, R. L. *The Principles of Humane Experimental Techniques*; UFAW: Potters Bar, 1959/1992.

(62) Reinhardt, C. A., Ed. *Alternatives to Animal Testing*; Chemie: Weinheim, 1993.

(63) It is our aim to develop pseudoreceptor modeling toward the reduction of animal models in toxicological research by constructing *toxoreceptors*, 3-dimensional surrogates of receptors responsible for the initial binding of toxicologically active substances. The "resolution" of our approach achieved for pharmacologically relevant receptor systems of 0.6−1.2 kcal/mol (corresponding to an uncertainty in the binding constant of a factor of 3 to 8) would seem to be sufficient to identify strongly toxic molecules and discard them from the screening process well before animal testing is applied. A first model for the aromatic hydrocarbon (AH) receptor has been constructed based on TCDD and a series of 9 polychlorinated biphenyls (PCB). For a test set of 8 PCB molecules, the RMS deviation of calculated and experimental differences in free energy of binding was 0.8 kcal/mol, corresponding to an uncertainty in the binding constant of a factor of 4.