Raptor: Combining Dual-Shell Representation, Induced-Fit Simulation, and Hydrophobicity Scoring in Receptor Modeling: Application toward the Simulation of Structurally Diverse Ligand Sets

Markus A. Lill,* Angelo Vedani, and Max Dobler

Biographics Laboratory 3R, Friedensgasse 35, 4056 Basel, Switzerland

Received April 28, 2004

We present a novel receptor-modeling approach (software Raptor) based on multidimensional quantitative structure−activity relationships (QSARs). To accurately predict relative free energies of ligand binding, it is of utmost importance to simulate induced fit. In Raptor, we explicitly and anisotropically allow for this phenomenon by a dual-shell representation of the receptor surrogate. In our concept, induced fit is not limited to steric aspects but includes the variation of the physicochemical fields along with it. The underlying scoring function for evaluating ligand−receptor interactions includes directional terms for hydrogen bonding and hydrophobicity and thereby treats solvation effects implicitly. This makes the approach independent from a partial-charge model and, as a consequence, allows one to smoothly model ligand molecules binding to the receptor with different net charges. We have applied the new concept toward the estimation of ligand-binding energies associated with the chemokine receptor-3 (50 ligands: $r^2 = 0.965$; $p^2 = 0.932$), the bradykinin B_2 receptor (52 ligands: $r^2 = 0.949$; $p^2 = 0.859$), and the estrogen receptor (116 ligands: $r^2 = 0.908$; $p^2 = 0.907$), respectively.

Introduction

In the last two decades, a large body of computer-aided drug design (CADD) concepts have been devised and matured into powerful tools for drug-design purposes. On the basis of the three-dimensional structure of macromolecular drug targets, structure-based design has become a widespread tool to identify and optimize drug candidates. To assess the pharmacological suitability of a given compound in silico, a reliable prediction of relative free energies of binding, e.g. by using free energy perturbation calculations,[1] is mandatory. Unfortunately, the associated computational procedures are computationally very demanding and often limited to the comparison of affinities of structurally similar compounds. Although the Protein Data Bank[2] nowadays comprises 25000 protein structures (3000−4000 thereof may be considered unique), for the majority of systems of biomedical interest, the three-dimensional structure of the associated bioregulator is not available to atomistic resolution, implying that structure-based design cannot be applied here.

Quantitative structure−activity relationships (QSARs) aim to identify a correlation between the binding affinity and structural features for a series of ligand molecules binding to a common target. Of particular interest for the biomedical research are QSARs based on three-dimensional models (3D-QSAR, e.g. CoMFA,[3] CoMSIA,[4] QSiAR,[5] GRID/GOLPE,[6] or Quasar[7,8]). They generate a rational model of the binding site, and allow for the quantification of the interactions between ligand molecules and protein at an atomistic level, by simulating electrostatic forces, hydrogen bonds, or van der Waals interactions.

In contrast to the true biological receptor, where the binding site is defined by a 3D arrangement of amino acids, most 3D-QSAR models typically represent this binding site by mapping physicochemical properties onto a surface or a grid surrounding the ligand molecules, superimposed in 3D space (pharmacophore hypothesis). As such a model interacts with all ligands simultaneously, it represents but an averaged surrogate; a fundamental shortcoming as receptor−ligand adaptation (the specific alteration of protein conformations induced by the individual ligand) cannot be simulated. This adaptation leads to different physicochemical fields experienced by the individual ligands, a fact that cannot be simulated with an averaged model. The Quasar methodology[7,8] is one of the few QSAR approaches which account for ligand-triggered induced fit by specifically allowing for a topological adaptation of the receptor surrogate to the individual ligand molecules. At the true biological receptor, however, this adaptation is not controlled only by steric factors; instead it may be governed by the simultaneous movement of amino acid domains within the binding pocket. In addition, amino acid residues bearing a conformationally flexible H-bond donor or acceptor moiety (Ser, Thr, Tyr, Cys, His, Asn, and Gln) may engage in differently oriented H-bonds (flip-flop) with dissimilar ligand molecules, an effect that cannot be simulated with an averaged receptor model. As we shall discuss in the following, induced fit leads to different physicochemical fields about each compound once bound to the protein. As a consequence, we present a dual-shell concept of an in-depth formulation of the binding site, based on experimental observations and explored by extensive MD simulations.

To quantify ligand−receptor interactions, we introduce an empirical scoring function which is based on

* Author to whom correspondence should be addressed. Tel: ++41 61 2614259. Fax: ++41 61 2614258. E-mail: markus@biograf.ch.

10.1021/jm049687e CCC: $27.50
Published on Web 10/28/2004

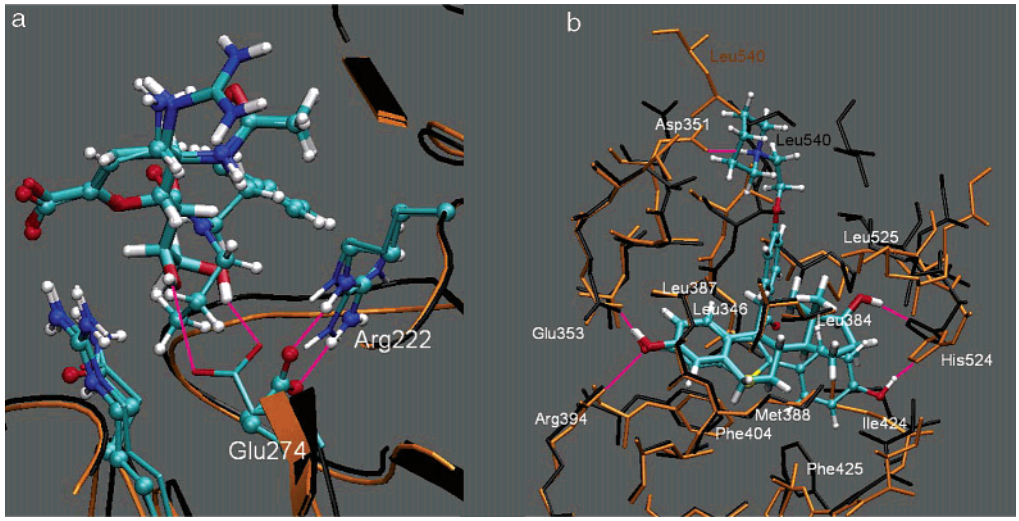

Figure 1. Small-molecule protein structures with different ligands solved by means of X-ray crystallography: (a) superposition of zanamivir (licorized) and 2-dihydropyran-phenethyl-propyl-carboxamide (ball-and-stick) bound to neuroaminidase; (b) superposition of 17β-estradiol (protein, black; ligand, licorized) and raloxifene (protein, orange; ligand, ball-and-stick) bound to human estrogen receptor.

hydrophobicity and hydrogen bonding, including terms for induced fit and entropic cost during ligand binding, thus avoiding problems associated with opposing partial charge models and salvation effects triggered by differently charged ligand molecules, e.g. neutral and formally charged species binding to the same bioregulator.

Analysis of Protein−Ligand Structures. To demonstrate that induced fit may result in individual physicochemical fields for each ligand molecule, we have analyzed protein structures for which different ligand complexes have been determined by means of X-ray crystallography. Figure 1 shows a superposition of zanamivir[9] and 2-dihydropyran-phenethyl-propyl-carboxamide[10] (2-PPC) bound to the enzyme neuraminidase[11] on the one hand and 17β-estradiol[12] and Raloxifene[13] bound to the estrogen receptor[14] on the other. While zanamivir engages with one (and, possibly, two) of its hydroxyl groups in hydrogen bonds with Glu 274, 2-PPC—lacking appropriately positioned hydrogen-bond donors—forces Glu 274 into a different conformation, forming a salt bridge to Arg 222. This conformational change results in a more lipophilic pocket to accommodate the propyl group of 2-PPC. In the complex of 17β-estradiol with the estrogen receptor, the ligand forms hydrogen bonds via both hyroxyl groups with Glu 353/Arg 394 and His 524, respectively. The residual aromatic/aliphatic portion is accommodated by the extended hydrophobic pocket comprising Leu 346, Leu 384, Leu 387, Met 388, Phe 404, Ile 424, Phe 425 and Leu 525. When complexed with raloxifene, the estrogen receptor opens a small channel near the center of the binding pocket[14] by translocating Leu 540 approximately 10 Å. Within this channel the alkylaminoethoxy side chain of raloxifene is accommodated. Additionally, Asp 351 is rotated toward the protonated piperidyl N atom ($pK_a = 11.2$) of raloxifene, forming a salt bridge. Both examples support the fact that the binding site experiences conformational changes induced by the individual ligands. As a result, both hydrophobic field and hydrogen-bond propensity spawned by the binding site are altered.

As of today, no generally available receptor-modeling tool would seem to allow for induced fit extending beyond the topology of the protein binding site; instead the generally available tools typically make use of a fixed-feature mapping, thereby representing a mean model.

Structures obtained by X-ray crystallography are averaged over time and space and do not disclose inherent fluctuations of the physicochemical properties, including hydrophobicity and hydrogen-bonding propensity, although, in reality, a ligand experiences time-dependent physicochemical fields. To assess the magnitude of such fluctuating fields, we have modified a concept previously developed by D. A. Pearlman[15] which uses a floating independent reference frame (FIRF).[16] Therein, a grid is placed around each ligand molecule. During the molecular-dynamics (MD) simulation, each grid point moves coupled to the ligand atom closest[17] to it. This allows for the monitoring of both the mean value and the fluctuation of the fields the ligand experiences during the simulation. MD simulations (100 ps equilibration and 400 ps data collection each) were run for 27 different small-molecule estrogen receptor complexes.[18] For analyzing the various fields using FIRFs, snapshots were taken every 2.5 ps. Figure 2a and Figure 2b show the hydrophobic fields of the amino acid residues of the receptor (averaged over the MD simulation) projected onto the topologically changing FIRF grid around the compounds diethylstilbestrol (DES) and bisphenol B (BPB). DES experiences hydrophilic fields around the two phenolic moieties of the molecule, which are hydrogen bonded to Glu 353 and Arg 394 on the one side and to a His 524 on the other (Figure 2c). While one phenolic oxygen atom of BPB is also hydrogen bonded to Glu 353 and Arg 394, the other points into a different octant of 3D space and hydrogen bonds to the side chain hydroxyl group of Thr 40, which results in an additional hydrophilic-field component (Figure 2d). In contrast hereto, the ethene group of DES located at this position would not seem to experience any hydrophilic field at all.

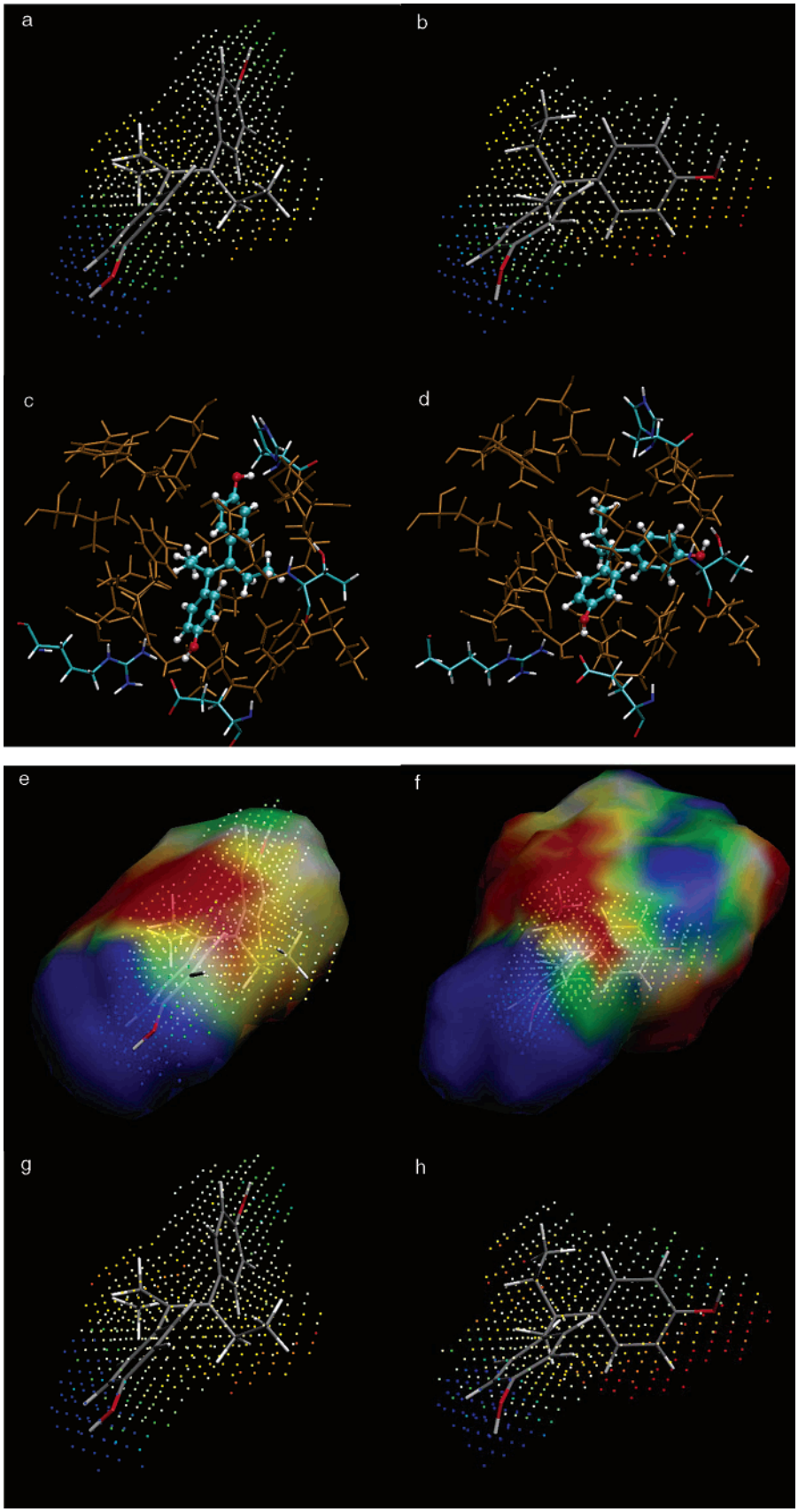

Figure 2. Hydrophobic fields averaged over the MD simulation projected onto the FIRF grid coupled to the actual ligand topology of (a) diethylstilbestrol (DES) and (b) bisphenol B (BPB). (c) Final structure of the MD simulations of ER (brown) with DES and (d) with BPB (both ball-and-stick), with H-bonding groups of the binding pocket colored by atom. (e) Inner shell with bound DES depicting the hydrophilic field (blue = strong, green = moderate) both hydroxyl groups experience and the hydrophobic field (red = strong, yellow = moderate) around the residual portion of the ligand molecule. (f) Outer shell with bound BPB modeling the additional hydrophilic field the second hydroxyl group of BPB experiences. (g) The field projected from the dual-shell representation to the grid points for DES and (h) for BPB, showing a qualitative agreement with the results from the MD simulations (a and b).

Methods

Dual-Shell Representation. To allow for ligand-dependent induced fit in receptor-modeling concepts based on QSAR, we have devised the Raptor technology using a dual-shell representation of the binding site. Within this representation, the inlying layer maps the fields which a substance would feel if it were to fit snugly into the binding pocket. Another compound, which features additional groups that reach deeper into the protein, may experience different fields as a consequence of induced fit. In contrast to other approaches, in Raptor the fields generated by the altered binding site can be modeled by a second, outer layer. As the inner shell is optimized using the most potent ligands of the training set, other compounds may have portions of matter located in the interstice between the two shells. Fields associated with such parts are interpolated between the two layers.[19]

To support the claim that a dual-shell representation of the binding pocket is sufficient to describe the individual adaptation of each compound, we have conducted a search to identify a distribution of properties on the two shells which is in qualitative agreement with the projected properties collected during the MD simulations (see above). As shown in Figure 2e−g, such a distribution can indeed be identified for the systems simulated in this study. This suggests that our approach may be sufficient for describing the physicochemical fields experienced by individual compounds during a MD simulation of the equilibrium binding state on an atomistic scale.

Scoring Function. The major problem associated with most existing scoring functions used to estimate the free energy of ligand binding is the weighting of solvation effects, particularly if the data set includes ligand species with different net charge. Hydrogen bonding, including the formation or disruption of water structures, is a major driving force during ligand binding. Substances that are predominantly hydrophobic in nature cannot engage in hydrogen bonds with the surrounding water and therefore tend to aggregate or to bind to hydrophobic portions of the binding pocket in macromolecular targets. Changes in entropy are part of this phenomenon, particularly the term arising from the loss or gain of ordered water structure. During ligand binding, the solvent structure of the ligand as well as solvent occupying the binding pocket of the protein in the uncomplexed state is displaced into disordered bulk solvent, resulting in an increase of entropy.[20] This is even true for predominantly hydrophobic compounds, where the surrounding water molecules cannot engage in hydrogen bonds with the ligand. Clearly, the binding process is governed by the tendency of both ligand and binding pocket to associate with a similar environment (hydrophilic or hydrophobic). The corresponding tendency of a ligand molecule is typically measured by 1-octanol/water partition coefficient (log $P_{o/w}$ (log P)). Kellogg and co-workers,[21] for example, developed an empirical force field (HINT) using log P as the only quantity and applied it successfully for structure-based design[22] and receptor modeling.[23] While interactions between hydrophobic entities are an (experimentally observed) consequence of the tendency of water to saturate its hydrogen-bond propensity, interactions between hydrophilic partners are in addition governed by an immediate physical force, the hydrogen bond itself. The latter interaction is clearly directional in nature.[24] To join these two concepts, our scoring function comprises hydrophobicity (ΔG_{HO}) and hydrogen bonding (ΔG_{HB}), as well as terms for the cost of the topological adaptation (ΔG_{IF}) and the changes in entropy ($\Delta G_{T\Delta S}$)[25] upon ligand binding:

$$\Delta G = \Delta G_{const} + f_{HO}\Delta G_{HO} + f_{HB}\Delta G_{HB} + f_{IF}\Delta G_{IF} + f_{T\Delta S}\Delta G_{T\Delta S} \quad (1)$$

ΔG_{const} is a contribution to the binding energy rationalizable as an overall loss of translational and rotational entropy of the ligand or overall gain of entropy due to desolvation of the binding pocket. f_{HO}, f_{HB}, f_{IF}, and $f_{T\Delta S}$ are scaling factors which are inherent to a given receptor model; they are optimized during the simulation (see below) for each specific drug target and typically constrained to specific intervals (e.g. $f_{HO} = 0.75-1.25$).

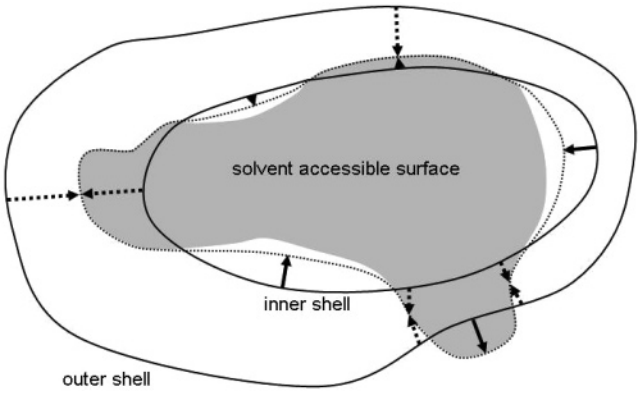

Figure 3. Sketch of a ligand molecule (represented as its SAS; gray surface) in the dual-shell representation of the receptor surrogate (solid lines). During the steric adaptation process, the fields generated by the protein binding site onto the ligand's SAS are computed by linear interpolation between inner and outer shell, if the ligand's SAS lies between those two shells (dashed arrows). For surface points located inside the inner layer, the latter may adapt only in part to the ligand topology (solid arrows; dotted line = topologically adapted receptor surface).

Induced Fit. In Raptor, the topology of a ligand molecule is represented by its solvent accessible surface ($r_{probe} = 1.4$ Å). The fields (hydrogen bonding and hydrophobicity) spawned by the ligand atoms are projected onto this surface; more precisely, onto discrete points located on equidistant positions thereon.[26]

The adaptation of both field and topology of the receptor surrogate to each ligand is achieved by combining a steric adjustment to the topology of the very ligand (Figure 3) and a component due to the attraction or repulsion between ligand and receptor surrogate. The latter is obtained by correlating their physicochemical properties (hydrophobicity and hydrogen-bond propensity) in 3D space.

During the steric adaptation process, the fields generated by the protein binding site onto the ligand's solvent-accessible surface (SAS) are computed by linear interpolation between inner and outer shell, if the ligand's SAS has portions located in there (Figure 3: dashed arrows; cf. above). For surface points located inside the inner layer (very small ligands), the latter may adapt only in part to the ligand topology, where the magnitude of the induced fit is proportional to the distance and orientation of the normal vectors of these surfaces (solid arrows in Figure 3).

In addition, the adaptation between ligand and surrogate is dependent on the strength of their interaction and is modeled by two parameters δ and ω, where the fields at the surface point α for every individual ligand are given by eqs 2 and 3:

$$HO_\alpha = HO^0_\alpha \frac{1}{1 + \exp(\delta_{ho,\alpha}(-ho_\alpha - \omega_{ho,\alpha}))} \quad (2)$$

$$HB_\alpha = (HB^0_{Acc,\alpha} - HB^0_{Don,\alpha}) \frac{1}{1 + \exp(\delta_{hb,\alpha}(hb_\alpha - \omega_{hb,\alpha}))} + HB^0_{Don,\alpha}. \quad (3)$$

HO^0_α, $HB^0_{Acc,\alpha}$, and $HB^0_{Don,\alpha}$ are hydrophobicity and hydrogen-bond-acceptor and -donor propensities, respectively, of the two layers linearly interpolated to each ligand surface point α.[27] The second term explicitly allows for hydrogen-bond flip-flop, which implies that, for example, a ligand donor can induce a change of the hydrogen-bond propensity of the binding site from "donating" to "accepting".

The interaction energy for each surface point is then calculated by multiplying its hydrophobicity ho_α with the hydrophobicity of the field of the receptor model HO_α for hydrophobic surface points or, otherwise, by multiplying its hydrogen-bonding propensity with that of the receptor model. The energy associated with the topological adaptation to the ligand's surface is assumed to be proportional to the magnitude of the simulated effect.

Optimization Algorithm. To obtain the best agreement between calculated and experimental binding affinities for the ligands of the training set, the underlying algorithm in Raptor optimizes the nature and distribution of the properties mapped onto the two shells representing the receptor surrogate. Those properties include the hydrophobicity, the hydrogen-bonding propensity, and the variables describing the adaptation process $\delta_{ho,\alpha}$, $\omega_{ho,\alpha}$, $\delta_{hb,\alpha}$, $\omega_{hb,\alpha}$, as well as a scaling factor for the energy associated with induced fit $c_{IF,\alpha}$. As quality criterion in the optimization process, the following expression is minimized:

$$Q = \sum_{\text{ligands}} (\Delta G_{\text{calc}} - \Delta G_{\text{exp}})^4 \qquad (4)$$

The fourth power was deliberately chosen as it weights small deviations (which typically come close to experimental uncertainty) less.

As at the true biological receptor the properties are not "randomly" scattered across the binding site, the algorithm in Raptor forces the formation of domains with similar properties: instead of altering the variables on each point independently, the quantities HO^0_α, $HB^0_{Acc,\alpha}$, $HB^0_{Don,\alpha}$, $\delta_{ho,\alpha}$, $\omega_{ho,\alpha}$, $\delta_{hb,\alpha}$, $\omega_{hb,\alpha}$, $c_{IF,\alpha}$ are optimized for each domain. Additional variables Δ are defined to describe the anisotropy of these properties over the domain. The hydrophobicity of a point within a domain, for example, is calculated by the expression

$$HO^0_\alpha = (1 - \Delta d)HO^0_{\text{domain}} \qquad (5)$$

where d is the distance from this point to the center of the domain.

Since the total number of variables is large (5−6 parameters for each of the 50−100 domains, i.e., HO^0_α, $\delta_{ho,\alpha}$, $\omega_{ho,\alpha}$, $c_{IF,\alpha}$, Δ_α for hydrophobic domains and $HB^0_{Acc,\alpha}$, $HB^0_{Don,\alpha}$, $\delta_{hb,\alpha}$, $\omega_{hb,\alpha}$, $c_{IF,\alpha}$, Δ_α for hydrogen-bonding domains), Raptor uses a multistep optimization protocol including domain assignment, tabu search,[28] and local search. To avoid overfitting, the number of optimizing variables is increased continuously; i.e. at the beginning (i.e. domain assignment and tabu search) only HO_α and HB_α are variable (for each domain) with all others added subsequently (i.e. $\delta_{ho,\alpha}$, $\omega_{ho,\alpha}$, Δ_α, or $\delta_{hb,\alpha}$, $\omega_{hb,\alpha}$, Δ_α during the first half of local search optimization, and $c_{IF,\alpha}$ during the final half of local optimization). In detail, the hydrophobic and hydrogen-bonding fields are projected onto the two layers weighted by ΔG of the corresponding ligand. Domain centers are assigned to the position with unambiguous manifestation of the projected hydrophobicity/hydrogen-bonding propensity. The assignment of each layer point to a domain is decided by comparing its field with that experienced at the center of the domain. At the beginning, only values for HO^0_α, $HB^0_{Acc,\alpha}$, and $HB^0_{Don,\alpha}$ are assigned to the domain centers and then optimized using tabu search.[28,29] Then, the individual domains are optimized using a local-search protocol. At this stage, all quantities are variable and treated continuously. If a better correlation is obtained, a given point may migrate to a neighboring domain.

Since the mapping of properties onto the shells is not unambiguously determinable—different models with similar predictive power can be identified—Raptor, therefore, generates a family of receptor models. Such surrogate families may be interpreted to represent the various configuration states of the true biological receptor. The obtained binding affinities are typically averaged over the individual models. Moreover, properties which are common to most models may be interpreted as chemically relevant features for ligand binding.

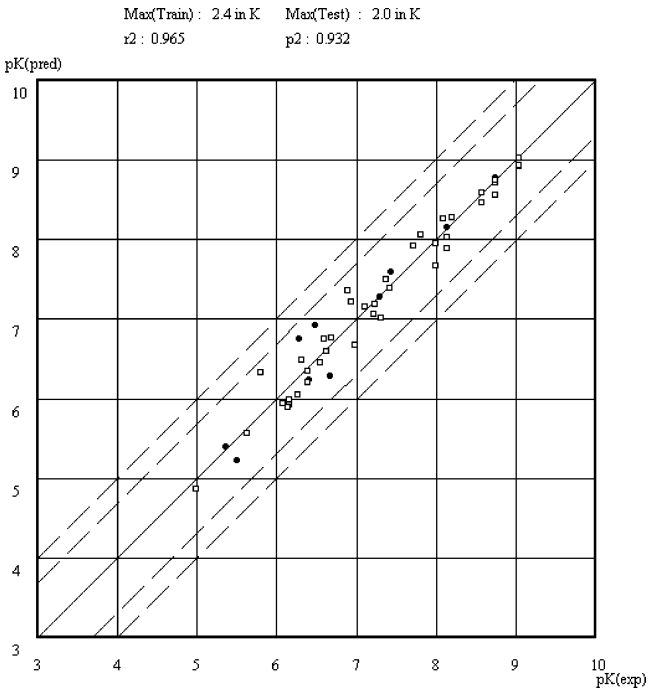

Figure 4. Predicted versus experimental binding affinities for ligand molecules binding to the CCR-3 receptor (training set, open squares; test set, filled circles).

Results

The Raptor concept was applied toward the prediction of binding affinities associated with the bradykinin B_2 system (34 training and 18 test compounds), the CCR-3 system (40/10), and the estrogen receptor (93/23), respectively. The latter includes eight substance classes as well as positive, neutral, and negatively charged compounds binding all to this receptor.

Chemokine Receptor (CCR-3). Chemokines—low molecular weight (8000−12000) chemotactic cytokines—are structurally related proteins that participate in the activation, proliferation, and differentiation of leukocytes and play a key role in the control of basal leukocyte trafficking and recruitment of leukocytes during inflammation. During inflammatory processes chemokines act via chemoattraction and activation of leukocytes. In response to certain stimuli or insult to the immune system, chemokines are secreted by proinflammatory cells, leukocytes, or endothelial cells to recruit new leukocytes from the circulation across the lumen and into the tissue.[30] Chemokines exert their functions through the selective binding to one or more G-protein-coupled receptors (GPCRs) differently expressed on leukocytes. More recent results suggested an important role for chemokines in a variety of pathophysiological processes including acute and chronic inflammation, infectious diseases, and modulation of angiogenesis and fibrosis.

On the basis of 50 selected compounds (40 training and 10 test substances) synthesized and tested at Brystol-Myers Squibb[31] we performed a Raptor study. The three-dimensional structure of all ligand molecules was generated using MacroModel 6.5[32] and optimized in aqueous solution on the basis of the AMBER 4.0 force field.[33] An extensive conformational search was then performed for at least one compound of any molecular scaffold present in the data set (6 compounds in total),

Table 1. Experimental and Calculated IC_{50} for CCR-3

name	IC_{50} expt	IC_{50} calc	factor off in IC_{50}
	Test Set		
112	2.0×10^{-9}	$1.88 \times 10^{-9} \pm 4.63 \times 10^{-10}$	0.1
107	8.0×10^{-9}	$8.26 \times 10^{-9} \pm 3.38 \times 10^{-9}$	0.0
94	4.0×10^{-8}	$3.01 \times 10^{-8} \pm 1.40 \times 10^{-8}$	0.3
99	5.5×10^{-8}	$6.48 \times 10^{-8} \pm 3.61 \times 10^{-8}$	0.2
92	2.4×10^{-7}	$6.12 \times 10^{-7} \pm 2.07 \times 10^{-7}$	1.6
88	3.6×10^{-7}	$1.38 \times 10^{-7} \pm 5.59 \times 10^{-8}$	1.6
84	4.2×10^{-7}	$6.85 \times 10^{-7} \pm 3.04 \times 10^{-7}$	0.6
75	5.6×10^{-7}	$1.96 \times 10^{-7} \pm 4.28 \times 10^{-8}$	1.9
39	1.8×10^{-6}	$5.15 \times 10^{-7} \pm 9.50 \times 10^{-8}$	2.4
37	3.4×10^{-6}	$7.42 \times 10^{-6} \pm 4.01 \times 10^{-6}$	1.2
	Training Set		
113	1.0×10^{-9}	$1.37 \times 10^{-9} \pm 3.78 \times 10^{-10}$	0.4
115	1.0×10^{-9}	$1.35 \times 10^{-9} \pm 4.88 \times 10^{-10}$	0.4
121	1.0×10^{-9}	$1.07 \times 10^{-9} \pm 3.25 \times 10^{-10}$	0.1
111	2.0×10^{-9}	$2.30 \times 10^{-9} \pm 8.95 \times 10^{-10}$	0.2
114	2.0×10^{-9}	$2.05 \times 10^{-9} \pm 4.61 \times 10^{-10}$	0.0
120	2.0×10^{-9}	$3.09 \times 10^{-9} \pm 8.04 \times 10^{-10}$	0.5
102	3.0×10^{-9}	$4.04 \times 10^{-9} \pm 1.63 \times 10^{-9}$	0.4
110	3.0×10^{-9}	$3.11 \times 10^{-9} \pm 1.43 \times 10^{-9}$	0.0
105	7.0×10^{-9}	$5.87 \times 10^{-9} \pm 1.57 \times 10^{-9}$	0.2
101	8.0×10^{-9}	$1.15 \times 10^{-8} \pm 5.81 \times 10^{-9}$	0.4
103	8.0×10^{-9}	$1.52 \times 10^{-8} \pm 6.12 \times 10^{-9}$	0.9
104	9.0×10^{-9}	$6.10 \times 10^{-9} \pm 1.79 \times 10^{-9}$	0.5
106	1.1×10^{-8}	$1.29 \times 10^{-8} \pm 3.89 \times 10^{-9}$	0.2
117	1.1×10^{-8}	$2.41 \times 10^{-8} \pm 5.96 \times 10^{-9}$	1.2
109	1.7×10^{-8}	$1.02 \times 10^{-8} \pm 4.12 \times 10^{-9}$	0.7
119	2.1×10^{-8}	$1.56 \times 10^{-8} \pm 1.05 \times 10^{-8}$	0.4
98	4.2×10^{-8}	$4.98 \times 10^{-8} \pm 2.66 \times 10^{-8}$	0.2
100	4.7×10^{-8}	$3.92 \times 10^{-8} \pm 2.21 \times 10^{-8}$	0.2
108	5.3×10^{-8}	$1.11 \times 10^{-7} \pm 3.42 \times 10^{-8}$	1.1
97	6.4×10^{-8}	$7.95 \times 10^{-8} \pm 3.86 \times 10^{-8}$	0.2
96	6.7×10^{-8}	$1.01 \times 10^{-7} \pm 4.40 \times 10^{-8}$	0.5
93	8.5×10^{-8}	$8.25 \times 10^{-8} \pm 3.19 \times 10^{-8}$	0.0
89	1.1×10^{-7}	$2.54 \times 10^{-7} \pm 9.67 \times 10^{-8}$	1.2
77	1.3×10^{-7}	$7.00 \times 10^{-8} \pm 2.41 \times 10^{-8}$	0.8
95	1.4×10^{-7}	$4.80 \times 10^{-8} \pm 1.15 \times 10^{-8}$	2.0
91	2.3×10^{-7}	$1.97 \times 10^{-7} \pm 6.88 \times 10^{-8}$	0.2
90	2.6×10^{-7}	$3.12 \times 10^{-7} \pm 1.44 \times 10^{-7}$	0.2
76	3.1×10^{-7}	$4.24 \times 10^{-7} \pm 1.85 \times 10^{-7}$	0.4
118	2.8×10^{-7}	$2.16 \times 10^{-7} \pm 9.46 \times 10^{-8}$	0.3
43	4.5×10^{-7}	$7.06 \times 10^{-7} \pm 2.03 \times 10^{-7}$	0.6
44	4.5×10^{-7}	$5.47 \times 10^{-7} \pm 3.07 \times 10^{-7}$	0.2
85	5.3×10^{-7}	$3.80 \times 10^{-7} \pm 1.82 \times 10^{-7}$	0.4
40	6.0×10^{-7}	$1.01 \times 10^{-6} \pm 2.99 \times 10^{-7}$	0.7
41	7.5×10^{-7}	$1.19 \times 10^{-6} \pm 4.32 \times 10^{-7}$	0.6
42	7.5×10^{-7}	$1.30 \times 10^{-6} \pm 3.05 \times 10^{-7}$	0.7
86	7.8×10^{-7}	$1.51 \times 10^{-6} \pm 5.72 \times 10^{-7}$	0.9
38	9.0×10^{-7}	$1.28 \times 10^{-6} \pm 3.03 \times 10^{-7}$	0.4
36	2.6×10^{-6}	$3.11 \times 10^{-6} \pm 1.16 \times 10^{-6}$	0.2
39	4.7×10^{-6}	$4.99 \times 10^{-6} \pm 2.50 \times 10^{-6}$	0.1
40	1.1×10^{-5}	$1.60 \times 10^{-5} \pm 5.97 \times 10^{-6}$	0.5

thereby allowing for 20000 minimized structures each, again using MacroModel and simulating an aqueous environment. Thereof, all conformations within 10 kcal/mol from the lowest-energy conformer were retained. To allow for a comparison with the Quasar study (cf. below), we used those conformers selected by Quasar.[35]

In Raptor, the binding affinities were averaged over 20 individual models and yielded an r^2 of 0.965 for the training and a predictive r^2 of 0.932 for the test set, where the maximal deviation of any ligand from the experimental value is a factor of 2.4 in the IC_{50} for the training and 2.0 for the test set (Figure 4 and Table 1), respectively. In the Quasar study (5D-QSAR),[34] a cross-validated r^2 of 0.950 and a predictive r^2 of 0.879 were obtained. The maximal deviation from the experiment was a factor of 0.6 and 3.2 in the IC_{50} value for the training and test sets, respectively. For this particular system, Raptor yields more accurate predictions for the test ligands when compared with Quasar. Model generation required 8 h of CPU time on a 1.8 GHz Pentium 4 computer; evaluation of the test compounds required 5−10 s per molecule.

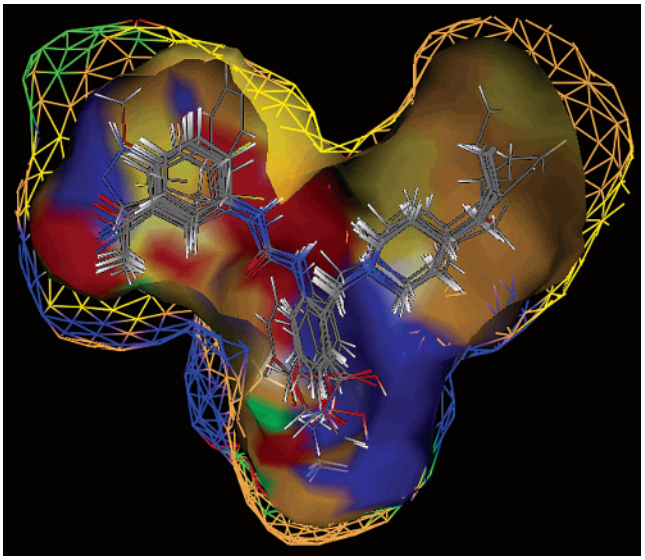

Figure 5. Dual-shell representation of one of the individual models for the CCR-3 receptor (hydrophobic fields, beige; hydrogen-bond-donating propensity, blue; hydrogen-bond-accepting propensity, red; hydrogen-bond flip-flop, green).

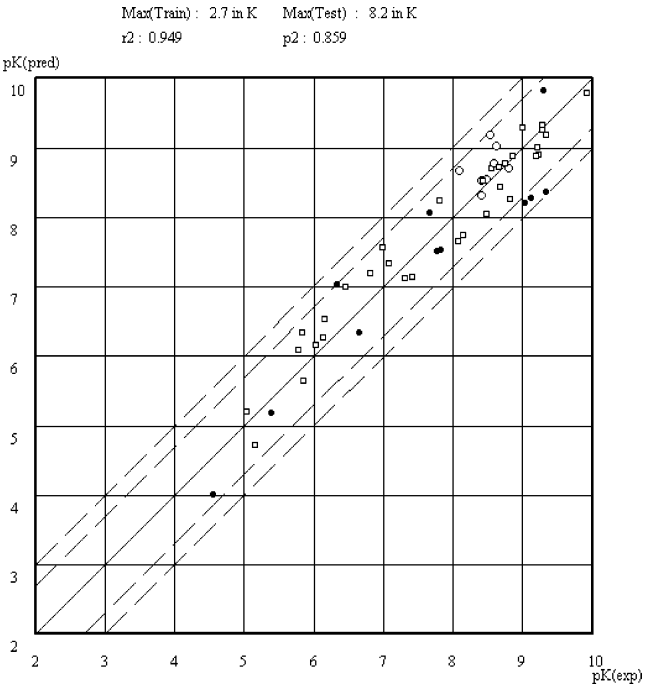

Figure 6. Computationally predicted versus experimentally measured binding affinities for compounds binding to the bradykinin B_2 receptor (training set, open squares; test set, filled circles; novel compounds (see text), open circles).

The dual-shell representation of one of the individual models is shown in Figure 5. The receptor model depicts a Y-shaped form. One lobe (top right, Figure 5) is mostly hydrophobic in nature, the stem shows large parts of hydrogen-bond-donating and -accepting areas, and the second lobe shows mixed hydrophobic/hydrophilic characteristics.

Table 2. Experimental and Calculated IC_{50} for the Bradykinin B_2 Receptor

name	IC_{50} expt	IC_{50} calc			factor off in IC_{50}
		Prediction Set			
50a	1.8E−09	2.24E−09	±	2.61E−09	0.3
17b	2.7E−09	1.10E−09	±	7.31E−10	1.4
50b	2.9E−09	1.90E−09	±	3.15E−09	0.5
23b	3.3E−09	7.51E−10	±	1.03E−09	3.3
17a	3.7E−09	3.33E−09	±	2.36E−09	0.1
17c	4.4E−09	5.63E−09	±	3.84E−09	0.3
23a	4.3E−09	3.38E−09	±	9.27E−09	0.3
23c	9.1E−09	2.46E−09	±	3.28E−09	2.6
		Test Set			
14a	5.1E−10	4.66E−09	±	2.50E−09	8.2
54b	5.5E−10	1.65E−10	±	7.26E−11	2.3
52e	8.3E−10	5.72E−09	±	6.13E−09	5.9
57a	1.0E−09	6.73E−09	±	3.60E−09	5.8
45a	1.6E−08	3.27E−08	±	2.89E−08	1.1
63c	1.8E−08	3.34E−08	±	5.43E−08	0.9
53	2.3E−08	9.47E−09	±	4.84E−09	1.4
6b	2.3E−07	5.13E−07	±	2.43E−07	1.2
37e	5.0E−07	1.03E−07	±	8.00E−08	3.8
5b	4.5E−06	7.10E−06	±	5.84E−06	0.6
5a	3.1E−05	1.07E−04	±	1.95E−04	2.5
		Training Set			
90c	1.3E−10	1.78E−10	±	9.74E−11	0.4
88a	5.0E−10	7.18E−10	±	1.44E−10	0.4
52a	5.6E−10	5.14E−10	±	2.96E−10	0.1
87b	5.6E−10	6.02E−10	±	2.46E−10	0.1
50a	6.4E−10	1.38E−09	±	2.49E−10	1.2
18c	6.6E−10	1.07E−09	±	2.84E−10	0.6
81a	6.9E−10	1.47E−09	±	2.91E−10	1.1
75a	1.1E−09	5.64E−10	±	2.25E−10	0.9
190997	1.5E−09	1.46E−09	±	5.00E−10	0.0
52f	1.6E−09	6.00E−09	±	2.93E−09	2.7
74a	1.9E−09	1.82E−09	±	7.41E−10	0.0
29a	2.3E−09	4.00E−09	±	1.10E−09	0.8
75b	2.4E−09	2.06E−09	±	1.16E−09	0.2
16	3.1E−09	2.16E−09	±	4.41E−10	0.4
63b	3.5E−09	9.91E−09	±	2.45E−09	1.9
74b	4.1E−09	3.25E−09	±	7.92E−10	0.3
48d	7.8E−09	2.02E−08	±	4.36E−09	1.6
52a	9.1E−09	2.43E−08	±	4.47E−09	1.7
23a	1.7E−08	6.29E−09	±	1.44E−09	1.7
29a	4.2E−08	8.14E−08	±	3.37E−08	1.0
31a	5.4E−08	8.19E−08	±	3.98E−08	0.5
52 g	9.0E−08	5.14E−08	±	2.51E−08	0.8
56	1.1E−07	3.04E−08	±	6.22E−09	2.6
58	1.7E−07	7.20E−08	±	9.03E−09	1.4
51	3.8E−07	1.10E−07	±	2.86E−08	2.4
45b	7.7E−07	3.17E−07	±	5.11E−08	1.4
45c	8.1E−07	6.02E−07	±	1.25E−07	0.3
52h	1.0E−06	7.56E−07	±	5.06E−07	0.3
30a	1.5E−06	2.53E−06	±	8.49E−07	0.7
45d	1.6E−06	5.00E−07	±	1.36E−07	2.2
32a	1.8E−06	9.02E−07	±	2.70E−07	1.0
7	7.5E−06	2.11E−05	±	6.08E−06	1.8
5d	1.0E−05	7.01E−06	±	2.81E−06	0.4

Bradykinin B_2 Receptor. Bradykinin (BK) is an endogenous nonapeptide (Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg) produced by proteolytic cleavage of the high molecular weight kininogen by plasma kallikreins. Because of its highly potent proinflammatory activity, BK has been implicated in a variety of pathophysiological responses, including pain, inflammation, asthma, rhinitis, and hypotension.[35] Two types of BK receptors, referred to as B_1 and B_2, have been identified by molecular cloning and pharmacological means.[36] B_2 receptors are expressed constitutively in many tissues and are thought to mediate most of the biological actions of BK. Because of the pathophysiological role of bradykinin, antagonists to BK are of biomedical interest.

On the basis of the IC_{50} values of 43 ligands synthesized and biologically assessed during the development of highly affine antagonists at Fujisawa Pharmaceuticals,[37] we performed a QSAR study using Raptor. The three-dimensional structures of all ligand molecules were generated using MacroModel 6.5 and optimized in aqueous solution on the basis of the AMBER 4.0 force field. On the basis of a pharmacophore hypothesis generated at the University of Düsseldorf,[38] we performed a receptor-mediated alignment protocol by means of the PrGen software.[44] Thirty-two molecules thereof were assigned to the training set and the remaining 11 used as test compounds. Most recently, an additional set of compounds has been published,[39] which feature additional functionalities, i.e. imidazolyl, pyrazolyl, triazolyl, diethylamine, and phenyl-substituted quinoline derivatives. To challenge the approach, we included only one of the additional compounds in the training set featuring the phenyl substituent, while eight were assigned to the test set.

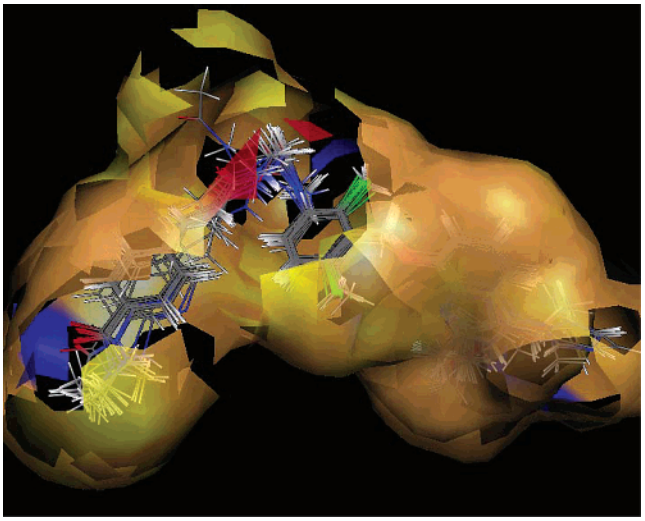

Figure 7. Common features determined by the Raptor models for the bradykinin B_2 receptor (hydrophobic fields, beige; hydrogen-bond-donating propensity, blue; hydrogen-bond-accepting propensity, red).

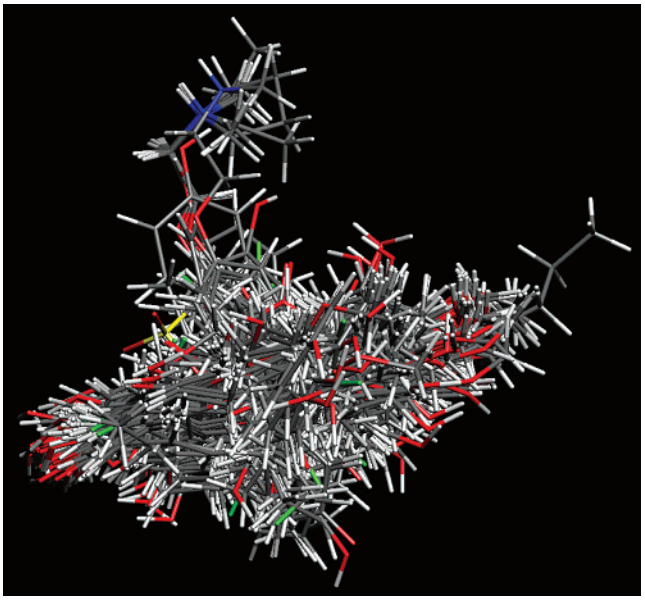

Figure 8. Pharmacophore hypothesis for the 116 superimposed ligand molecules binding to the estrogen receptor.

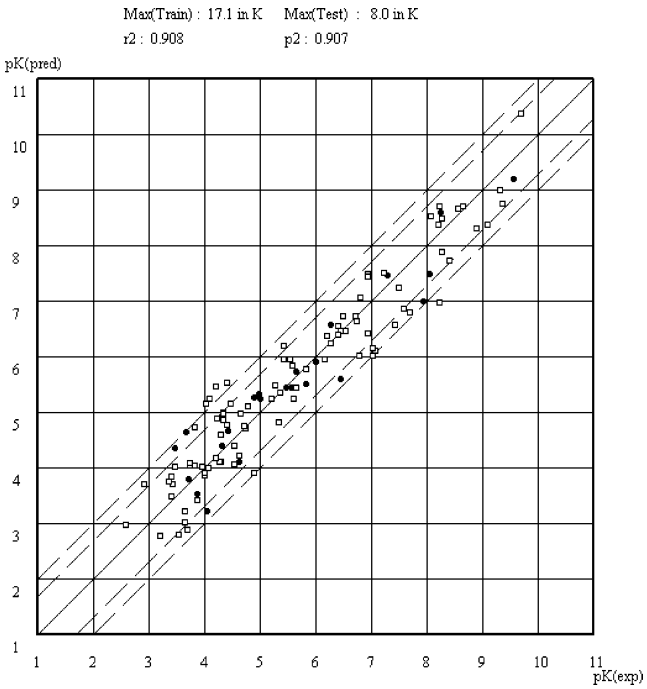

Figure 9. Computationally predicted versus experimentally measured binding affinities for ligand molecules binding to the estrogen receptor (training set, open squares; test set, filled circles).

A family of 10 receptor models was created, resulting in an averaged r^2 of 0.949 for the training set (32 + 1 compounds) and a predictive r^2 of 0.859 for the test set (11 + 8). The maximal deviation of any ligand from the experimental value is a factor of 1.8 in the IC_{50} for the training and 8.2 for the test set, respectively (Figure 6 and Table 2). The nine new ligands[39] are predicted close to experimental error, with a maximal deviation of 3.3 in the IC_{50}. The physicochemical properties, in common at the 70% level for the 10 models, are shown in Figure 7. The receptor model is "V-shaped", with the benzo-condensated heterocyclic substituents (e.g. quinolines; right part in Figure 4) being mainly hydrophobic in nature. Around the central amide bond a dominant hydrogen-bond-acceptor region is observed, and a hydrogen-bond-donor domain is observed around the terminal amide oxygen atom. Model generation required 7 h of CPU time on a 1.8 GHz Pentium 4 computer; evaluation of the test compounds required 5−10 s per molecule.

Estrogen Receptor. The estrogen receptor (ER) is a ligand-activated transcription factor. Upon binding of estrogen, it undergoes a conformational change which promotes its homodimerization and high-affinity binding to specific sites on DNA−estrogen response elements. Once bound to DNA, the estrogen receptor complex regulates the expression of estrogen-responsive genes. The effects of this are very broad and have an influence on secondary sexual characterization for both male and female, bone maintenance, or the cardiovascular system. The ER has a broad solvent-accessible binding site which allows for structurally diverse chemicals to bind to it.

Endocrine disruptors, exogenous substances that cause adverse health effects due to changes in the endocrine system, mediate these effects by binding to specific receptors including the estrogen receptor. Studies[40] suggest that exposure to endocrine disruptors may contribute to the development of hormone-dependent cancers (e.g. breast, prostate, testicular) and to the disorder of the male reproductive tract, as well as compromise reproductive fitness by decreasing sperm counts and semen volume.

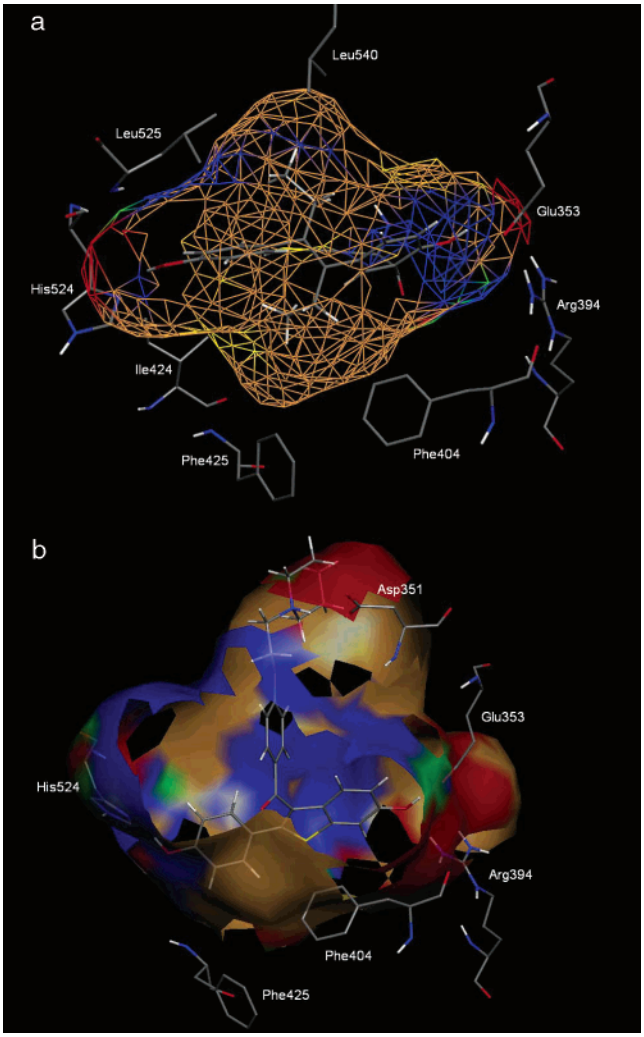

Figure 10. Common features determined by the Raptor models for the estrogen receptor (hydrophobic fields, beige; hydrogen-bond-donating propensity, blue; hydrogen-bond-accepting propensity, red) with bound (a) DES and (b) raloxifene. For comparison the binding mode of DES and raloxifene in the true biological receptor is shown together with some protein residues (3ERD and 1ERR, respectively).

Due to its broad chemical spectrum of binding compounds, including neutral substances as well as positively and negatively charged ions, predicting the binding affinities of these ligands by computational approaches is a big challenge for every method. Since several three-dimensional structures of small-molecule ER complexes are deposited with the PDB,[2] we used a receptor-mediated alignment protocol to create a pharmacophore hypothesis for a total of 116 compounds (93 training and 23 test ligands).[41] Again, the ligands were generated using MacroModel, minimized in an implicit water solution using the AMBER force field, and an extended conformational search was performed. All conformers within 50 kJ/mol of the lowest-energy structure were retained and aligned using Rfit.[42] For the neutral ligands, the receptor-mediated alignment

Table 3. Experimental and Calculated IC_{50} for the Estrogen Receptor

name	log RBA expt	IC_{50} expt	IC_{50} calc			factor off in IC_{50}
Test Set						
meso-hexestrol	2.48	2.9E−10	6.96E−10	±	6.46E−10	1.4
dimethylstilbestrol	1.16	6.2E−09	2.87E−09	±	3.92E−09	1.1
monomethyI ether hexestrol	0.97	9.5E−09	3.61E−08	±	2.17E−08	2.8
estrone	0.86	1.2E−08	1.11E−07	±	1.39E−07	8.0
tamoxifen	0.21	5.5E−08	3.89E−08	±	8.87E−08	0.4
monohydroxy methoxychloroolefin	−0.63	3.8E−07	2.89E−06	±	7.42E−06	6.6
6,4′-dihydroxyflavone	−0.82	5.9E−07	2.91E−07	±	4.41E−07	1.0
bisphenol B	−1.07	1.1E−06	1.36E−06	±	4.00E−06	0.3
diethylstilbestrol dimethyl ether	−1.25	1.6E−06	3.43E−06	±	5.70E−06	1.2
2,5-dichloro-4′-biphenylol	−1.44	2.5E−06	2.13E−06	±	5.33E−05	0.2
nonylphenol	−1.53	3.0E−06	4.06E−06	±	7.12E−06	0.3
kaempferol	−1.61	3.6E−06	4.00E−06	±	1.69E−05	0.1
heptylparaben	−2.09	1.1E−05	6.29E−06	±	4.90E−06	0.7
bisphenol A	−2.11	1.2E−05	5.26E−06	±	2.20E−06	1.2
3-deoxyestrone	−2.2	1.4E−05	6.20E−06	±	4.93E−06	1.3
4,4′-dihydoxybenzophenone	−2.46	2.6E−05	8.68E−05	±	1.11E−04	2.4
4′-hydroxyflavanone	−2.65	4.0E−05	2.43E−05	±	1.97E−05	0.6
2-chloro-4-biphenylol	−2.77	5.3E−05	4.46E−05	±	5.61E−05	0.2
p-phenylphenol	−3.04	9.8E−05	6.72E−04	±	3.46E−04	5.9
ethylparaben	−3.22	1.5E−04	3.37E−04	±	1.79E−04	1.3
4-*sec*-butylphenol	−3.37	2.1E−04	1.78E−04	±	1.48E−04	0.2
6-hydroxyflavone	−3.41	2.3E−04	2.56E−05	±	5.70E−05	8.0
4-*tert*-butylphenol	−3.61	3.7E−04	5.05E−05	±	2.20E−05	6.2
Training Set						
diethylstilbestrol (DES)	2.6	2.2E−10	4.67E−11	±	1.07E−11	3.8
ethynyl estradiol	2.28	4.7E−10	2.01E−09	±	8.95E−10	3.3
4-OH-tamoxifen	2.24	5.1E−10	1.15E−09	±	3.45E−10	1.3
17β-estradiol	2	8.9E−10	4.79E−09	±	1.50E−09	4.4
4-OH-estradiol	1.82	1.3E−09	5.57E−09	±	4.96E−09	3.2
dienestrol	1.57	2.4E−09	2.15E−09	±	3.11E−09	0.1
2-OH-estradiol	1.47	3.0E−09	2.42E−09	±	4.87E−09	0.2
diethylstilbestrol monomethyl ether	1.31	4.4E−09	2.11E−08	±	9.55E−09	3.8
3,3′-dihydroxyl hexestrol	1.19	5.7E−09	3.63E−09	±	5.80E−09	0.6
droloxifene	1.18	5.9E−09	1.44E−08	±	8.86E−09	1.5
moxestrol	1.14	6.4E−09	2.21E−09	±	1.08E−08	1.9
17-deoxyestradiol	1.14	6.4E−09	1.17E−07	±	6.29E−08	17.1
2,6-dimethylhexestrol	1.11	6.9E−09	4.60E−09	±	6.24E−09	0.5
estriol	0.99	9.1E−09	3.29E−09	±	4.09E−09	1.8
p-(α,β-diethyl-*p*-methyl-phenethyl)-meso-phenol	0.6	2.2E−08	1.79E−07	±	5.82E−08	7.0
17α-estradiol	0.49	2.9E−08	1.57E−07	±	1.11E−07	4.5
dihydroxymethoxychloroolefin	0.42	3.4E−08	6.54E−08	±	2.59E−07	0.9
mestranol	0.35	4.0E−08	3.04E−07	±	2.09E−07	6.7
toremifene	0.14	6.5E−08	3.46E−08	±	4.92E−08	0.9
α-α-dimethyl-*b*-ethyl allenolic acid	−0.02	9.3E−08	8.54E−07	±	5.37E−07	8.2
coumestrol	−0.05	9.9E−08	7.85E−07	±	7.87E−07	6.9
4-ethyl-7-OH-3-(*p*-methoxyphenyl)-dihydro-l-benzopyran-2-one	−0.05	9.9E−08	1.07E−06	±	4.91E−07	9.7
clomiphene	−0.14	1.2E−07	4.20E−07	±	2.04E−07	2.4
nafoxidine	−0.14	1.2E−07	3.57E−08	±	1.33E−08	2.4
6α-OH-estradiol	−0.15	1.3E−07	4.04E−08	±	1.67E−08	2.1
3-hydroxy-estra-1,3,5(10)-trien-16-one	−0.29	1.7E−07	9.73E−08	±	6.67E−08	0.8
3-deoxyestradiol	−0.3	1.8E−07	1.08E−06	±	5.51E−07	5.1
7,3′,4′-trihydroxyisoflavone	−0.35	2.0E−07	2.57E−07	±	5.45E−07	0.3
3,6,4′-trihydroxyflavone	−0.35	2.0E−07	2.57E−07	±	5.08E−07	0.3
genistein	−0.36	2.0E−07	2.13E−07	±	4.11E−07	0.0
4,4′-dihydroxystibene	−0.55	3.2E−07	3.94E−07	±	8.66E−07	0.2
HPTE	−0.6	3.6E−07	2.09E−07	±	1.16E−07	0.7
2,3,4,5-tetrachloro-4′-biphenylol	−0.64	3.9E−07	3.75E−07	±	6.40E−07	0.0
norethynodrel	−0.67	4.2E−07	4.49E−07	±	4.36E−07	0.1
2,2′,4,4′-tetrahydroxybenzil	−0.68	4.3E−07	3.12E−07	±	4.66E−07	0.4
equol	−0.82	5.9E−07	6.38E−07	±	1.06E−06	0.1
monohydroxy methoxychlor	−0.89	6.9E−07	4.83E−07	±	7.02E−07	0.4
3β-androstanediol	−0.92	7.4E−07	1.22E−06	±	8.57E−07	0.7
4,2′,4′-trihydroxychalcone	−1.26	1.6E−06	1.92E−06	±	2.17E−06	0.2
4,4′-(1,2-ethanediyl)bisphenol	−1.44	2.5E−06	4.14E−06	±	1.65E−05	0.7
16β-hydroxy-16-methyl-3-methylether estradiol	−1.48	2.7E−06	6.55E−06	±	4.96E−06	1.4
aurin	−1.5	2.8E−06	1.61E−06	±	1.35E−06	0.8
apigenin	−1.55	3.2E−06	1.27E−06	±	5.60E−06	1.5
daidzein	−1.65	4.0E−06	7.09E−07	±	4.45E−07	4.6
3-methylestriol	−1.65	4.0E−06	1.27E−06	±	8.95E−07	2.1
4-dodecylphenol	−1.73	4.8E−06	5.08E−06	±	4.98E−06	0.1
ethylhexylparaben	−1.74	4.9E−06	1.73E−05	±	1.55E−05	2.5
4-*tert*-octylphenol	−1.82	5.9E−06	3.59E−06	±	3.55E−06	0.6

Table 3. (Continued)

name	log RBA expt	IC_{50} expt	IC_{50} calc			factor off in IC_{50}
Training Set (Continued)						
phenolphthalein	−1.87	6.6E−06	6.55E−06	±	3.96E−06	0.0
4-chloro-4′-biphenylol	−2.18	1.4E−05	1.42E−04	±	5.81E−05	9.5
octylphenol	−2.31	1.8E−05	8.60E−06	±	6.11E−06	1.1
fisetin	−2.35	2.0E−05	2.20E−05	±	3.99E−05	0.1
biochanin A	−2.37	2.1E−05	2.03E−05	±	4.37E−05	0.0
4′-hydroxychalcone	−2.43	2.4E−05	1.17E−05	±	8.01E−06	1.1
2,2′-methylenebis(4-chlorophenol)	−2.45	2.5E−05	6.66E−05	±	7.70E−05	1.6
benzylparaben	−2.54	3.1E−05	9.78E−05	±	8.48E−05	2.2
4-hydroxychalcone	−2.55	3.2E−05	4.54E−05	±	5.07E−05	0.4
2,4-dihydroxybenzophenone	−2.61	3.6E−05	7.99E−06	±	3.82E−06	3.6
3α-androstanediol	−2.67	4.2E−05	3.31E−06	±	1.83E−06	11.6
4-phenethylphenol	−2.69	4.4E−05	1.93E−05	±	9.65E−06	1.3
doisynoestrol	−2.74	4.9E−05	1.52E−05	±	3.23E−05	2.2
5,4′-dihydroxy-7-methoxyiso-flavone	−2.74	4.9E−05	1.12E−05	±	9.55E−06	3.4
myricetin	−2.75	5.0E−05	1.20E−05	±	1.35E−05	3.2
triphenylethylene	−2.78	5.4E−05	8.79E−05	±	7.07E−05	0.6
3′-hydroxyflavanone	−2.78	5.4E−05	2.86E−05	±	4.79E−05	0.9
chalcone	−2.82	5.9E−05	8.60E−05	±	9.85E−05	0.5
o,*p*′-DDT	−2.85	6.3E−05	1.46E−05	±	1.06E−05	3.3
4-heptyloxyphenol	−2.88	6.8E−05	7.54E−05	±	4.05E−05	0.1
dihydrotestosterone (DHT)	−2.89	6.9E−05	3.82E−06	±	2.39E−06	17.2
formononetin	−2.98	8.5E−05	6.42E−06	±	2.28E−06	12.3
bis(4-hydroxyphenyl)methane	−3.02	9.4E−05	1.15E−04	±	8.18E−05	0.2
6-hydroxyflavanone	−3.05	1.0E−04	8.08E−06	±	2.65E−06	11.4
4,4′-sulfonyldiphenol	−3.07	1.1E−04	1.51E−04	±	9.80E−05	0.4
butylparaben	−3.07	1.1E−04	1.37E−04	±	8.65E−05	0.3
diphenolic acid	−3.13	1.2E−04	1.09E−04	±	2.33E−04	0.1
propylparaben	−3.22	1.5E−04	4.24E−04	±	2.07E−04	1.9
3,3′,5,5′-tetrachloro-4,4′-biphenyldiol	−3.25	1.6E−04	2.14E−05	±	1.51E−05	6.4
phenolred	−3.25	1.6E−04	1.04E−04	±	1.29E−04	0.5
4-*tert*-amylphenol	−3.26	1.6E−04	2.10E−05	±	1.07E−05	6.8
balcalein	−3.35	2.0E−04	1.04E−04	±	1.45E−04	0.9
morin	−3.35	2.0E−04	9.32E−05	±	1.58E−04	1.2
4-chloro-3-methylphenol	−3.38	2.1E−04	1.46E−03	±	5.71E−04	5.8
3-phenylphenol	−3.44	2.5E−04	1.05E−03	±	8.81E−04	3.3
methylparaben	−3.44	2.5E−04	6.72E−04	±	5.20E−04	1.7
2-*sec*-butylphenol	−3.54	3.1E−04	1.83E−03	±	1.13E−03	4.9
2,4′-dichlorobiphenyl	−3.61	3.7E−04	1.09E−04	±	3.39E−05	2.3
2-chloro-4-methylphenol	−3.66	4.1E−04	2.20E−04	±	1.94E−04	0.9
phenolphthalin	−3.67	4.2E−04	1.59E−04	±	1.24E−04	1.6
4-chloro-2-methylphenol	−3.67	4.2E−04	3.71E−04	±	4.95E−04	0.1
7-hydroxyflavanone	−3.73	4.8E−04	1.97E−04	±	1.67E−04	1.5
3-ethylphenol	−3.87	6.6E−04	1.85E−03	±	1.18E−03	1.8
4-ethylphenol	−4.17	1.3E−03	2.25E−04	±	9.64E−05	4.9
4-methylphenol	−4.5	2.8E−03	1.20E−03	±	3.79E−04	1.4

protocol was executed using the 3ERD structure,[43] for the charged species the 1ERR,[13] by means of the PrGen software[44] (Figure 8).

A family of 10 receptor models was created, resulting in an averaged r^2 of 0.908 and in a predictive r^2 of 0.907. The maximal deviations from the experimental binding affinities are a factor of 17.1 and 8.0 off in IC_{50}, whereby only five compounds of 93 training and none of the 23 test ligands are off more than a factor of 10. Binding of neutral as well as positively and negatively charged ligands was all simulated equally well. This suggests that Raptor not only seems to reproduce the experimental binding affinities of the training set but also gives reliable prediction for the ligands of the test set. The experimental and calculated IC_{50} values are compared in Figure 9 and Table 3. The molecules shown along with their experimentally given IC_{50} values are given in ref 45 (Tables 5−11). Model generation required 10 h of CPU time on a 1.8 GHz Pentium 4 computer; evaluation of the test compounds required 5−10 s per molecule.

A representation of the common features of the receptor surrogates is depicted in Figure 10 (Figure 10a, inner layer; Figure 10b, outer layer). A comparison with the optimized experimental structure with bound DES[43] (Figure 10a) shows that the major features of the true biological receptor are mapped in the Raptor model, i.e. the central hydrophobic barrel as well as the hydrophilic domains about the phenolic hydroxyl groups of DES.

A comparison with the X-ray crystal structure with bound raloxifene[13] (Figure 10b) shows Asp351 interacting with the protonated nitrogen of the antagonist molecule. This feature could only be mapped on the second, outer shell of the Raptor model and demonstrates the necessity for anisotropic modeling in this very case.

Conclusions

To quantitatively predict relative free energies of ligand binding in receptor-modeling studies, it is of utmost importance to simulate induced fit. In a novel approach (software Raptor), we explicitly and anisotro-

pically allow for it by a dual-shell representation of the receptor surrogate. In contrast to other approaches in the field, induced-fit simulation is not restricted to topological aspects but includes the variation of the physicochemical fields along with it. The underlying scoring function includes directional terms for hydrogen bonding and treats solvation effects implicitly. Electrostatic terms are deliberately omitted; instead, the function empirically assesses the very behavior by means of terms quantifying hydrophobicity and hydrogen bonding. This makes it independent from a partial-charge model and, as a consequence, allows one to model ligand molecules binding to the receptor with different net charges.

The requirements on data size and affinity range are comparable to the requirements in grid-based 3D-QSAR approaches, as the shell representation in Raptor lowers the available degrees of freedom resulting in a reduction of the 3D optimization problem to an effective 2D surface approach. In addition, combining the receptor presenting points on the shells in domains lessens the number of degrees of freedom further. We suggest using a minimal size of 30 compounds with a range in affinity larger than a factor of 1000.

We have applied the new concept toward the estimation of ligand-binding energies associated with the chemokine receptor-3 (50 ligands), the bradykinin B_2 receptor (52), and the estrogen receptor (116). The obtained results for the ligand molecules of the test sets (predictive r^2 ranging from 0.859 to 0.932) suggest that this approach can be used for drug-design studies using very diverse data sets. Updated information may be continuously obtained from http://www.biograf.ch; the Biographics Laboratory 3R is a nonprofit organization aimed at the replacement of animal models in biomedical research.

References

(1) Reddy, M. R., Erion, M. D., Eds. *Free energy calculations in rational drug design*; Kluwer Academic Publishers: Dordrecht, 2001.

(2) www.rcsb.org and Berman et al.: Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28,* 235−242.

(3) Cramer, R. D.; Patterson, D. E., III; Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(4) Klebe, G.; Abraham, U.; Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological potency. *J. Med. Chem.* **1994**, *37*, 4130−4146.

(5) Kubinyi, H.; Hamprecht, F. A.; Mietzner, T. Three-dimensional quantitative similarity-activity relationships (3D QsiAR) from SEAL similarity matrices. *J. Med. Chem.* **1998**, *41,* 2553−2564.

(6) Baroni, M.; Costantino, G.; Cruciani, G.; Riganelli, D.; Valigi, R.; Clementi, S. Generating Optimal Linear PLS Estimations (GOLPE): An advanced chemometric tool for handling 3D-QSAR problems. *Quant. Struct.-Act. Relat.* **1993**, *12,* 9−20.

(7) Vedani, A.; Briem, H.; Dobler, M.; Dollinger, K.; McMasters, D. R. Multiple conformation and protonation-state representation in 4D-QSAR: The neurokinin-1 receptor system. *J. Med. Chem.* **2000**, *43*, 4416−4427.

(8) Vedani, A.; Dobler M. 5D-QSAR: The key for simulating induced fit? *J. Med. Chem.* **2002**, *45*, 2139−2149.

(9) PDB code: 1A4G.

(10) PDB code: 1A4Q.

(11) Taylor, N. R.; Cleasby, A.; Singh, O.; Skarzynski, T.; Wonacott, A. J.; Smith, P. W.; Sollis, S. L.; Howes, P. D.; Cherry, P. C.; Bethell, R.; Colman, P.; Varghese, J. Dihydropyrancarboxamides related to zanamivir: A new series of inhibitors of influenza virus sialidases. 2. Crystallographic and molecular modeling study of complexes of 4-Amino-4*H*-pyran-6-carboxamides and sialidase from influenza virus types A and B. *J. Med. Chem.* **1998**, *41*, 798−807.

(12) PDB code: 1ERE.

(13) PDB code: 1ERR.

(14) Brzozowski, A. M.; Pike, A. C. W.; Dauter, Z.; Hubbard, R. E.; Bonn, T.; Engstrom, O.; Ohman, L.; Greene, G. L.; Gustaffson, J.-A.; Carlquist, M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **1997**, *389,* 753−758.

(15) Pearlman, D. A. Free energy grids: a practical qualitative application of free energy perturbation to ligand design using the OWFEG method. *J. Med. Chem.* **1999**, *42,* 4313−4324.

(16) In structure-based design, calculating free energies of binding based on a static X-ray structure, e.g. using docking in combination with scoring, typically does not yield accurate predictions. Therefore, free-energy perturbation, linear-interaction energy methods, etc. sample many possible configurations throughout a Monte Carlo or molecular-dynamics simulation, resulting in a more reliable estimation of binding affinities. On the basis of this principle, we make use of MD simulations in combination with the modified FIRF approach to obtain a more accurate observation of the strength of the physicochemical fields the ligands experience by the receptor.

(17) To monitor the physicochemical fields a substance is exposed to during a simulation, a grid is constructed around the starting configuration of the ligand. Since the conformation of a nonrigid molecule is changing throughout the simulation, the surrounding grid is deformed as well. Within the floating independent reference frames (FIRF), each grid point is linked to its nearest ligand's heavy atom and a coordinate system is defined for each grid point based on that heavy atom, its bond partner, and the atacedent bonded partner. Each grid-point vector undergoes translation and rotation along with the atom it is coupled to. In this way, the reference frame for two adjacent grid points can differ, and an appropriate grid can be produced even when parts of the ligand are flexible.

(18) Each compound was docked into the structure of the estrogen receptor (PDB code: 1ERE). A solvation cap of 30 Å around the geometric center of the binding pocket was created, and the whole system was minimized using AMBER 7.0 (Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7*; University of California: San Francisco, 2002). During the MD simulation (100 ps equilibration, slowly heating from 100 to 300 K, followed by 400 ps of data collection) all atoms more than 25 Å away from the center were constrained.

(19) A grid with 1 Å spacing is placed around the superposition of the ligands. A value of occupancy of a ligand's volume is assigned to each grid point using the Wyvill function [Wyvill, B.; Wyvill, G. Field functions for implicit surfaces. *Visual Comput.* **1989**, *5*, 75−82] $f(R) = 1 - (22/9)R^2 + (17/9)R^4 - (4/9)R^6$ where $R = r/r_{max}$, r is the distance from grid point to the closest atom, and r_{max} is the solvent-accessible radius plus an additional 0.5 Å for smearing purposes. Next, a triangulated surface is created on an isocontour level using the marching cubes algorithm [Lorensen, W. E.; Cline, H. E. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. *Comput. Graphics (Proc.SIGGRAPH '87)* **1987**, *21,* 163−169]. For the inner shell, only the experimentally most affine ligand is taken into account, while for the outer layer, all ligands from the training set are included. This way, a smoothed solvent-accessible surface results for for the inner envelope.

(20) Dunitz, J. D. The entropic cost of bound water in crystals and biomolecules. *Science (Washington, D.C.)* **1994**, *264,* 670.

(21) Kellogg, G. E.; Burnett, J. C.; Abraham, D. J. Very empirical treatment of solvation and entropy: a force field derived from $LogP_{o/w}$. *J. Comput.-Aided Mol. Des.* **2001**, *15,* 381−393.

(22) (a) Gunnarsson, T.; Desai, U. R. Designing small, nonsugar activators of antithrombin using hydropathic interaction analyses. *J. Med. Chem.* **2002**, *45,* 1233−1243. (b) Marabotti, A.; Balestreri, L.; Cozzini, P.; Mozzarelli, A.; Kellogg, G. E.; Abraham, D. J. HINT Predictive analysis of binding between retinol binding protein and hydrophobic ligands. *Bioorg. Med. Chem. Lett.* **2000**, *10,* 2129−2132.

(23) (a) Galatin, P. S.; Abraham, D. J. QSAR: Hydropathic analysis of inhibitors of the p53-mdm2 interaction. *Proteins: Struct., Funct., Genet.* **2001**, *45,* 169−175. (b) Pajeva, I. K.; Wiese, M. Interpretation of CoMFA results—A probe set study using hydrophobic fields. *Quant. Struct-Act. Relat.* **1999**, *18,* 369−379.

(24) (a) Baker, E. N.; Hubbard, R. E. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **1984**, *44,* 97−179. (b) Murray-Rust, P.; Glusker, J. P. Directional hydrogen binding in sp^2- and sp^3-hybridized O atoms and its relevance to ligand−macromol-

ecule interactions. *J. Am. Chem. Soc.* **1984**, *106*, 1018−1025. (c) Taylor, R.; Kennard, O. Hydrogen bonding geometry in organic crystals. *Acc. Chem. Res.* **1984**, *17,* 320−326. (d) Tintelnot, M.; Andrews, P. Geometries of functional group interactions in enzyme-ligand complexes: Guides for receptor modeling. *J. Comput.-Aided Mol. Des.* **1989**, *3,* 67−84. (e) Vedani, A.; Dunitz, J. D. Lone-pair directionality in H-bond potential functions for molecular mechanics calculations. *J. Am. Chem. Soc.* **1985**, *107*, 7653−7658.

(25) Searle, M. S.; Williams, D. H. The cost of conformational order: Entropy changes in molecular associations. *J. Am. Chem. Soc.* **1992**, *114*, 10690−10697.

(26) The hydrophobicity and hydrogen-bond propensity on a surface point α are given by $ho_\alpha = \sum_{i=0}^{N} ho_i\, f_{Fermi}(r_{i\alpha};1.5,2.5)$ and $hb_\alpha = \sum_{i=0}^{N} hb_i\, f_{Fermi}(r_{i\alpha};0.5,2.3) f_{Fermi}(\theta;10°,50°)$, where $f_{Fermi}(x;a,b) = 1/1 + \exp[(1/a)(x - b)]$ is the Fermi function; N is the number of atoms; $r_{i\alpha}$ is the distance between atom i and surface point α; and θ is the angle between atom i, its lone-pair, and the surface point if i is an acceptor or between the bond partner of atom i, the atom i, and the surface point if i is a donating hydrogen. Each atom can either be hydrogen-bond donor/acceptor or neither, i.e. $hb_i \in \{0;1\}$. The hydrophobicity of an atom is defined by its partial propensity toward water or octanol, defined by an atomistic log P value. An atom type classification system based on its 2D connectivity from Wildman and Crippen [Wildman, S. A.; Crippen, G. M. Prediction of physicochemical parameters by atomic contributions. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 868−873] is used for the assignment of a partial log P to each atom type, ranging from negative values (log $P < -1.0$) for net charged groups to positive values for hydrophobic atom types.

(27) For the first expression it is twice the hydrophobicity the ligand surface point feels, if no interaction-driven induced fit is possible, i.e. $\delta_{ho,\alpha} = 0$.

(28) Glover F. Future paths for integer programming and links to artificial intelligence. *Comput. Oper. Res.* **1986**, *5,* 533−549.

(29) At this stage, the values can have up to eight discrete values. To accelerate the optimization the fields of the ligands are precomputed for the domains. Tabu search (TS) is a heuristic global optimization technique, which is an iterative procedure for exploring the search space by a sequence of moves, making use of an adaptive form of memory. In each tabu step, the properties of 25−50% of the domains are changed. Within one step, the properties can change their value only to their discrete neighboring level. To avoid being trapped in a local optimum, each step is taken as starting point for the next move. To improve the efficiency of the exploration process, the information from previous steps is taken into account in addition to the actual move. This is done by introducing a tabu list, which prevents a domain from assuming the same value as at the starting point of the previous step. By rejecting these moves (=tabu) the algorithm is forced to find new solutions and avoids cycling in the neighborhood of suboptimal solutions. An aspiration criterion is defined which allows one to overcome any tabu, if this move leads to an unprecedented best solution. Besides the short-time memory of a tabu list, a long-time memory is used storing a few best local minima encountered so far. The tabu search is stopped after a given number of moves or, alternatively, when no better solution is found for some time. Then, the long-time memory is used as restart point for local search which thoroughly examines promising areas. The tabu search is repeated several times from different starting points in the search space, avoiding any configuration stored in the long-time memory.

(30) (a) Bertrand, C. P.; Donath, P. D. CCR3 blockade as a new therapy for asthma. *Exp. Opin. Invest. Drugs* **2000**, *9*, 43−52. (b) Baggiolini, M. Chemokine and leucocyte traffic. *Nature* **1998**, *392*, 565−568. (c) Locati, M.; Murphy, P. M. Chemokines and chemokine receptors: biology and clinical relevance in inflammation and AIDS. *Annu. Rev. Med.* **1999**, *50*, 425−440. (d) Wells, T. N. C.; Proudfoot, A. E. I.; Power, C. A. Chemokine receptors and their role in leukocyte activation. *Immunol. Lett.* **1999**, *65*, 35−40. (e) Zlotnik, A.; Morales, J.; Hedrick, J. A. Recent advances in chemokine and chemokine receptors. *Crit. Rev. Immunol.* **1999**, *19*, 1−47. (f) Saunders: J.; Tarby, C. M. Opportunities for novel therapeutic agents acting at chemokine receptors. *DDT* **1999**, *4*, 80−92. (g) Vaddi, K.; Keller, M.; Newton, R. C. *The chemokine facts book*; Academic Press: New York, 1997.

(31) De Lucca, G. V.; Kim, U. T.; Johnson, C.; Vargo, B. J.; Welch, P. K.; Covington, M.; Davies, P.; Solomon, K. A.; Newton, R. C.; Trainor, G. L.; Decicco, C. P.; Ko, S. S. Discovery and structure−activity relationship of *N*-(ureidoalkyl)-benzyl-piperidines as potent small molecule CC chemokine receptor-3 (CCR3) antagonists. *J. Med. Chem.* **2002**, *45,* 3794−3804.

(32) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. MacroModel−An integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990**, *11*, 440−467.

(33) Weiner, S. J.; Kollmann, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. A new force field for molecular-mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **1984**, *106*, 765−784.

(34) Vedani, A.; Dobler, M.; Dollinger, H.; Hasselbach, K.-M.; Birke, F.; Lill, M. A. Novel ligands for the chemokine receptor-3 (CCR-3): A receptor-modeling study based on 5D-QSAR. *J. Med. Chem.*, submitted.

(35) (a) Regoli, D.; Barabe, J. Pharmacology of bradykinin and related kinins. *Pharmacol. Rev.* **1980**, *32,* 1−46. (b) Marceau, F.; Lussier, A.; Regoli, D.; Giroud, J. P. Pharmacology of kinins: Their relevance to tissue injury and inflammation. *Gen. Pharmacol.* **1983**, *14*, 209−229. (c) Proud, D.; Kaplan, A. P. Kinin formation: Mechanisms and role in inflammatory disorders. *Annu. Rev. Immunol.* **1988**, *6*, 49−83. (d) Dray, A.; Perkins, M. Bradykinin and inflammatory pain. *Trends Neurosci*. **1993**, *16*, 99−104. (e) Greaves, M. W. Inflammation and mediators. *Br. J. Dermatol.* **1988**, *119*, 419−426. (f) Bhoola, K. D.; Figueroa, C. D.; Worthy, K. Bioregulation of kinins: Kallikreins, kininogens and kininases. *Pharmacol. Rev.* **1992**, *44*, 1−80. (g) Farmer, S. G.; Burch, R. M. The pharmacology of bradykinin receptors. In *Bradykinin Antagonists: Basic and Clinical Research*; Burch, R. M.; Ed.; Marcel Dekker: New York, 1991; pp 1−31. (h) Burch, R. M.; Farmer, S. G.; Steranka, L. R. Bradykinin receptor antagonists. *Med. Res. Rev.* **1990**, *10*, 237−269.

(36) (a) McEachern, A. E.; Shelton, E. R.; Bhakta, S.; Obernolte, R.; Bach, C.; Zuppan, P.; Fujisaki, J.; Aldrich, R. W.; Jarnagin, K. Expression cloning of a rat B2 bradykinin receptor. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7724−7728. (b) Hess, J. F.; Borkowski, J. A.; Young, G. S.; Strader, C. D.; Ransom, R. W. Cloning and pharmacological characterization of a human bradykinin (BK-2) receptor. *Biochem. Biophys. Res. Commun.* **1992**, *184*, 260−268. (c) Menke, J. G.; Borkowski, J. A.; Bierilo, K. K.; MacNeil, T.; Derrick, A. W.; Schneck, K. A.; Ransom, R. W.; Strader, C. D.; Linemeyer, D. L.; Hess, J. F. Expression cloning of a human B1 bradykinin receptor. *J. Biol. Chem.* **1994**, *269*, 21583−21586.

(37) (a) Abe, Y.; Kayakiri, H.; Satoh, S.; Inoue, T.; Sawada, Y.; Inamura, N.; Asano, M.; Aramori, I.; Hatori, C.; Sawai, H.; Oku, T.; Tanaka, H. A novel class of orally active non-peptide bradykinin B2 receptor antagonists. 4. Discovery of novel frameworks mimicking the active conformation. *J. Med. Chem.* **1998**, *41,* 4587−4598. (b) Abe, Y.; Kayakiri, H.; Satoh, S.; Inoue, T.; Sawada, Y.; Inamura, N.; Asano, M.; Aramori, I.; Hatori, C.; Sawai, H.; Oku, T.; Tanaka, H. A novel class of orally active non-peptide bradykinin B2 receptor antagonists. 3. Discovering bioisosteres of the imidazo[1,2-*a*]pyridine moiety. *J. Med. Chem.* **1998**, *41,* 4062−4079. (c) Abe, Y.; Kayakiri, H.; Satoh, S.; Inoue, T.; Sawada, Y.; Inamura, N.; Asano, M.; Hatori, C.; Sawai, H.; Oku, T.; Tanaka, H. A novel class of orally active non-peptide bradykinin B2 receptor antagonists. 2. Overcoming the species difference between guinea pig and man. *J. Med. Chem.* **1998**, *41*, 4053−4061. (d) Abe, Y.; Kayakiri, H.; Satoh, S.; Inoue, T.; Sawada, Y.; Imai, K.; Inamura, N.; Asano, M.; Hatori, C.; Katayama, A.; Oku, T.; Tanaka, H. A novel class of orally active non-peptide bradykinin B2 receptor antagonists. 1. Construction of the basic framework. *J. Med. Chem.* **1998**, *41,* 564−578. (e) Asano, M.; Hatori, C.; Sawai, H.; Johki, S.; Inamura, N.; Kayakiri, H.; Satoh, S.; Abe, Y.; Inoue, T.; Sawada, Y.; Mizutani, T.; Oku, T.; Nakahara, K. Pharmacological characterization of a nonpeptide bradykinin B2 receptor antagonist, FR165649, and agonist, FR190997. *Br. J. Pharmacol.* **1998**, *124,* 441−446. (f) Asano, M.; Hatori, C.; Inamura, N.; Sawai, H.; Hirosumi, J.; Fujiwara, T.; Nakahara, K. Effects of a nonpeptide bradykinin B2 receptor antagonist, FR167344, on different in vivo animal models of inflammation. *Br. J. Pharmacol*. **1997**, *122,* 1436−1440.

(38) Gurrath, M. Personal communication.

(39) Sawada, Y.; Kayakiri, H.; Abe, Y.; Imai, K.; Mizutani, T.; Inamura, N.; Asano, M.; Aramori, I.; Hatori, C.; Katayama, A.; Oku, T.; Tanaka, H. A new series of highly potent non-peptide bradykinin B2 receptor antagonists incorporating the 4-heteroarylquinoline framework. Improvement of aqueous solubility and new insights into species difference. *J. Med. Chem.* **2004**, *47,* 1617−1630.

(40) (a) Iguchi, T. Environmental endocrine disruptors. *Jpn. J. Clin. Med.* **2000**, *56,* 2953−2962. (b) Zacharewski, T. Identification and assessment of endocrine disruptors: Limitations of in vivo and in vitro assays. *Environ. Health Perspect.* **1998**, *106,* 577−582.

(41) Emery, F. Construction and validation of an estrogen-receptor model using 5D-QSAR. M.Sc. Thesis, 2002.

(42) Schmid, S. Unpublished.

(43) Shiau, A. K.; Barstad, D.; Loria, P. M.; Cheng, L.; Kushner, P. J.; Agard, D. A.; Greene, G. L. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* **1998**, *95*, 927−937.

(44) Zbinden, P.; Dobler, M.; Folkers, G.; Vedani, A. PrGen: Pseudoreceptor modeling using receptor-mediated ligand alignment and pharmacophore equilibration. *Quant. Struct.-Act. Relat*. **1998**, *17*, 122−130.

(45) Blair, R. M.; Fang, H.; Branham, W. S.; Hass, B. S.; Dial, S. L.; Moland C. L.; Tong, W.; Shi, L.; Perkins, R.; Sheehan, D. M. The estrogen receptor relative binding affinities of 188 natural and xenochemicals: structural diversity of ligands. *Toxicol. Sci.* **2000**, *54*, 138−53.

JM049687E