Novel Ligands for the Chemokine Receptor-3 (CCR3): A Receptor-Modeling Study Based on 5D-QSAR

Angelo Vedani,*,† Max Dobler,† Horst Dollinger,§ Kai-Malte Hasselbach,# Franz Birke,¶ and Markus A. Lill†

Biographics Laboratory 3R, Friedensgasse 35, CH-4056 Basel, Switzerland, and Chemical Research, Lead Discovery, and Pulmonary Research, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach/Riss, Germany

Received April 13, 2004

We recently reported the development of a receptor-modeling concept based on 5D-QSAR (quantitative structure−activity relationships) and which explicitly allows for the simulation of induced fit. In this account, we report its utilization toward the design of novel compounds able to inhibit the chemokine receptor-3 (CCR3). The study was based on a total of 141 compounds, representing four different substance classes. Using the Quasar software, we built two receptor surrogates that yielded a cross-validated r^2 value of 0.950/0.861 and a predictive r^2 of 0.879/0.798, respectively. The model was then employed to predict the activity of 58 hypothetical compounds featuring two variation patterns: lipophilic substitutions and amphiphilic H-bond acceptors. Eleven of the proposed ligands show a calculated binding affinity lower than any compound within the training set; the most potent candidate molecule is expected to bind at an IC_{50} of 0.3 nM.

Introduction

Chemokines—low molecular weight (8−12 kDa) 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.[1−7] Chemokines are classified according to the relative position of the first cysteine found in the primary amino acid sequence. In CXC (α-chemokines), the first pair of cysteines is separated by a single amino acid; CC (*β*-chemokines) feature adjacent cysteines, and in the CX3C (*δ*-chemokines), the first pair is separated by three amino acids. C (*γ*-chemokines) contain only a single cysteine in the homologous position. Chemokines exert their functions through the selective binding to one or more G-protein-coupled receptor (GPCR) 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.[1]

In asthma, airway mast cells eosinophils and $CD4^+$ Th2-type lymphocytes appear to be the major effector cells in the chronic inflammation that underlies the clinical manifestation of the disease.[8] Little details are known about the cellular and molecular mechanisms involved in the recruitment of these cells from the circulation, but two fundamental processes are generally accepted to be involved: leukocyte adhesion involving selectin-mediated tethering and transendothelial migration of the leukocyte into the surrounding tissue up a chemotactic gradient.[1]

Chemokine receptors share common features present in many members of the GPCR superfamily. All chemokine receptors have two conserved Cys residues: one in the N-terminal domain and the other within extracellular loop 3, presumably forming a disulfide bridge that is an integral part of the ligand-recognition site. Except for CXCR3, there are also two conserved acidic residues within the helical bundle. The first, an Asp in helix 2, is intimately associated with receptor activation as is the case with all GPCRs.

Unlike the GPCRs binding monoamine ligands (and featuring a second Asp residue in helix 3 close to the extracellular surface), chemokine receptors have a Glu residue in a similar position but on helix 7. In common with other superfamily members, there is a conserved DRY (Asp-Arg-Tyr) triad at the C-terminus of helix 3, speculatively involved in G-protein binding. Unique to the chemokine receptor subfamily, all have an acidic N-terminus with gross negative charges ranging from −1 to −6, and this might distinguish the initial recognition event from that in other GPCRs.[6]

As no experimental structure is presently available for the CCR3 receptor, s*tructure-based design* is inapplicable, and receptor-modeling approaches must be employed. Receptor modeling is a computational tech-

* To whom correspondence should be addressed. Phone: +41-61-261-4256. Fax: +41-61-261-4258. E-mail: admin@biograf.ch. Internet: http://www.biograf.ch/.
† Biographics Laboratory 3R.
§ Chemical Research, Boehringer Ingelheim Pharma GmbH & Co. KG.
# Lead Discovery, Boehringer Ingelheim Pharma GmbH & Co. KG.
¶ Pulmonary Research, Boehringer Ingelheim Pharma GmbH & Co. KG.

10.1021/jm040827u CCC: $30.25
Published on Web 02/11/2005

nology that allows building atomistic or virtual surrogates for structurally uncharacterized bioregulators. When validated using quantitative structure−activity relationships (QSARs), such models can be used to predict the binding affinity of new compounds. Of particular interest in biomedical research are QSARs based on three-dimensional models (3D-QSAR)[9−12] because they allow for the simulation of directional forces: hydrogen bonds and metal−ligand interactions, entities known to play a key role for both molecular recognition and selective binding.[13−15] While at the true biological receptor only one ligand molecule may bind at the time, a QSAR study is typically based on a series of ligand molecules binding "simultaneously" to the receptor surrogate. In 3D-QSAR, —where each ligand molecule is represented by a single, three-dimensional entity, —the identification of the bioactive conformation, orientation, and, possibly, the protonation state is a crucial step in the procedure. If the underlying pharmacophore hypothesis is based on incorrect assumptions, the resulting surrogate is hardly of any use for predictive purposes. This alignment problem has long been recognized,[9−11] and two principle ways to circumvent it have been explored: in situ generation of the active conformer[16] and multiple ligand representation (4D-QSAR).[17−22] The 4D approaches represent the ligand molecules (of both training and test sets) as an ensemble of conformations, orientations, and protonation states. The most likely bioactive representation may then be genetically evolved from this reservoir, for example, using a Boltzmann-weighted selection criterion.[19−22] An adequate treatment of conformationally flexible H-bond donor or acceptor moieties at the true biological receptor, able to engage in differently directed hydrogen bonds with dissimilar ligand molecules, may be simulated through the definition of "H-bond flip-flop" particles (properties) when using quasi-atomistic receptor models.[19−22]

Even with 4D-QSAR, a major unknown persists: manifestation and magnitude of the induced fit, the ligand-induced adaptation of the binding site to the topology of the small molecule. In the absence of the true biological receptor, these quantities cannot unambiguously be determined; the most recently developed 5D-QSAR tools offer a possible solution to the problem; here, several induced-fit scenarios (cf. Methods) are evaluated simultaneously. 5D-QSAR as implemented in Quasar not only allows for the simultaneous evaluation of different induced-fit scenarios but also allows for the induced-fit crossover and linear combination of the individual hypotheses.[25,26]

More recently, a study on the structure−activity relationship of *N*-(ureidoalkyl)-benzylpiperidines as potent CCR3 inhibitors has been published.[27] Starting from a structurally related series of *N*-(alkyl)-benzylpiperidines CCR1 receptor antagonists, the introduction of the ureidoalkyl moiety improved the binding potency from the micromolar to the low-nanomolar range. Series-1 of our study includes 50 of these ligands, which were synthesized and tested at Brystol-Myers Squibb in Wilmington, DE. Our main data (series-2) comprises 91 ureidoalkyl-piperazines, aminoalkyl-piperazines, and amidoalkyl-piperazines displaying a similar scaffold but synthesized[28] and tested[29] at Boehringer Ingelheim in Biberach, Germany. For series-1, IC_{50} values ranging from 1.0 nM to 19.3 μM are available,[27] while for series-2, K_i values in the range from 50 nM to 63 μM have been experimentally determined.[29]

The ultimate goal of our study was to understand the interaction of CCR3 inhibitors at the molecular level by establishing and validating a family of 3D receptor models for the structurally uncharacterized receptor and to use these models to predict novel, more potent ligands. For this task, we made use of the Quasar technology[24−26] and a total of 141 compounds comprising four substance classes. In our simulations, those compounds were represented by a total of 412 conformers (4D-QSAR) while simultaneously exploring six different induced-fit scenarios (5D-QSAR). The Quasar technology includes an option to determine the contribution of any functional group toward the free energy of ligand binding, ΔG. We made use of this possibility in order to propose new compounds with different chemical and structural features than those used as training and test ligands, respectively.

Methods

A quasi-atomistic binding-site surrogate refers to a high level of model abstraction. The essential information about the hypothetical receptor site is provided by means of a three-dimensional surface that surrounds the ligand molecules (defining the training set) at van der Waals distance and is populated with properties mapped onto it. The topology of this surface mimics the three-dimensional shape of the binding site; the mapped properties represent other information of interest, such as hydrophobicity, partial charge, electrostatic potential, and hydrogen-bonding propensity. A variety of algorithms to generate and validate binding-site models have been described.[18,24−26,28−33] While most approaches are based on a 3D-QSAR concept, more recent algorithms allow for a multiple representation of the ligand conformation (4D-QSAR).[24−26,33,34]

The Quasar concept developed at the Biographics Laboratory 3R allows for a multiple representation of the ligand topology (conformations, orientations, and protonation states, referred to as the fourth dimension in QSAR)[19] as well as a multiple representation of induced-fit hypotheses (referred to as the fifth dimension in QSAR).[24−26] Both ensembles are available throughout the entire simulation, and genetic algorithms are used for selecting the most predictive combination.[25] This approach reduces the bias associated with the choice of the bioactive conformation, the ligand alignment, and the induced-fit model. Quasar explicitly allows for H-bond flip-flop and accounts for solvation phenomena.[24−26] The technical details of model construction in Quasar are published elsewhere[25,35] and shall therefore only be summarized here.

1. Construction of the Receptor Surface. Induced fit may be simulated by adapting a van der Waals surface (generated about all ligands defining the training set) to the topology of each ligand molecule of training, test, and prediction sets. This is achieved by adapting this surface onto the van der Waals surface of each individual ligand molecule. The procedure, mimicking a local induced fit, can be performed using different protocols: isotropically (linearly), anisotropi-

cally (field scaled),[36] or through energy minimization.[25] The corresponding rms shift (the simulated induced fit) is used to estimate the associated energy (cf. eq 1). Typical rms shifts are in the range of 0.4−2.5 Å and are associated with induced-fit energies ranging from 0.2 to 6.0 kcal/mol.

2. Generation of an Initial Family of Parent Structures. Points/domains on the receptor surface are then randomly populated with atomistic properties.[21,25] While the distributed properties are identical for all ligand molecules, their exact location on the envelope varies slightly (rms fluctuations range from 0.5 to 1.5 Å with maximal individual shifts as large as 3.5 Å) depending on the very ligand molecule.

3. Evolution of a Model Family. When we use a genetic algorithm (for a detailed description, see, for example, refs 31 and 37), the initial family of receptor models is evolved simulating crossover events. At each crossover step, there is a small probability (typically 0.01−0.02) of a transcription error, which is expressed by a random mutation. Thereafter, those two individuals of the population with the highest lack-of-fit value[25] are discarded. This process is repeated until a target cross-validated r^2 (typically 0.75−0.95) or the experimental accuracy of the binding data (typically 0.24−0.40 kcal/mol, corresponding to an uncertainty of a factor 0.5−2.0 in the binding affinity) is reached.[25,35]

4. Estimation of Relative Free Energies of Ligand Binding. In our concept,[24−26] we have combined the approach of Blaney et al.[38] with a method of Still et al.[39] for estimating ligand solvation energies and a term to correct for the loss of entropy upon receptor binding following Searle and Williams,[40]

$$E_{\text{binding,ligand}} = E_{\text{ligand-receptor}} - T\Delta S_{\text{binding}} - E_{\text{solvation}} - \Delta E_{\text{internal strain}} - E_{\text{induced fit}} \quad (1)$$

When using a multiple ligand representation, the interactions of all conformations, orientations, and protonation-states are calculated toward all members of the receptor−model family. The contribution of an individual entity to the total energy is determined using a normalized Boltzmann distribution,[19−21,24−26]

$$E_{\text{bdg.tot}} = \sum E_{\text{bdg.ind}} \cdot \exp(-w_i \cdot E_{\text{bdg.ind}}/E_{\text{bdg.ind.lowest}}) \quad (2)$$

where $w_i = (\sum E_{\text{bdg.ind}}/E_{\text{bdg.ind.lowest}})^{-1}$ is the normalizing factor.

Free energies of ligand binding, ΔG_{pred}, are then predicted by means of a linear regression between ΔG_{exp} and E_{binding} (cf. eq 1) using the ligand molecules of the training set:

$$\Delta G_{\text{pred}} = (|a| \cdot E_{\text{binding}}) + b \quad (3)$$

Slope and intercept of eq 3 are inherent to a given receptor model and are subsequently applied to predict the relative binding energy of ligand molecules different from those in the training set. As in Quasar, the receptor surrogate is represented by a family of models (typically 100−1000); this approach allows for a subtler scaling of both the ligand−receptor interactions and the induced-fit scenario.

5. Analysis of the Model Family. A mandatory criterion to validate a family of receptor models is their ability to predict relative free energies of ligand binding for an external set of test ligand molecules, not used during model construction (e.g., its rms deviation or the predictive r^2 value). A more serious challenge to a model family is the so-called scramble test (cf. ref 37). Here, the experimental binding data (i.e., ΔG_{exp}) of the training set are randomly scrambled, and the simulation is repeated under otherwise identical conditions. If, under these circumstances, the ligands of the test set are still predicted correctly (i.e., a predictive $r^2 > 0.5$), the model is worthless, as it is not sensitive toward the biological data (ΔG_{exp}).

The Quasar technology,[24−26] with its underlying 5D-QSAR concept, allows for the simultaneous evaluation of an ensemble of induced-fit hypotheses thus reducing the bias with the choice of the adaptation mechanism. Presently, up to six protocols may be selected: (1) a linear mode, scalable from 0.0 to 1.0 and typically applied at the 0.75 level, with 0.0 referring to "no induced fit" and 1.0 to a maximal adaptation to the topology of the individual molecule; (2−4) adaptations based on the steric-, electrostatic-, and H-bond-field, respectively; (5) energy minimization along the steric-field vectors;[44] (6) adaptation based on the molecular lipophilicity potential.[45,46]

The induced fit in protocols 2−4 and 6 is executed proportional to the effective field acting on the outer surface (accommodating all ligand molecules of the training set) followed by a constrained minimization, to restore equal separations between neighbored grid points defining the adapted surface. While the linear mode behaves isotropically and depends solely on molecular shape, all other options account for molecular properties (presently the steric-field, the electric charge distribution, H-bond donors or acceptors, and the lipophilicity potential) yielding an anisotropic induced-fit model. Figure 1 shows the superimposed induced-fit models obtained for the CCR3 receptor surrogate (series-2) in this study.

During parent generation, all selected induced-fit models are evaluated, and the entity with the lowest lack-of-fit value is selected. This implies that for each of the i models of the surrogate family, n ligands represented by m conformers (typically 4−16) and k induced-fit models (2−6) are evaluated, constituting a genuine 5D-QSAR approach.[47]

The most recent version of Quasar (4.5)[35] promotes the formation of domains with equal (or at least similar) properties—salt bridges, hydrogen bond donors and acceptors, hydrophobic or solvent-accessible regions—as it is observed with true biological receptors. Degrees of "property freedom" (i.e., which property may be deposited on which region in receptor space during the simulated evolution) are now assigned based on very stringent criteria, for example, hydrogen-bonding propensity based on the directionality of hydrogen bonds and the probability to engage in salt bridges on the presence of charged functional groups on the ligands. The gross charge of the receptor surrogate is restricted to ±1.0 for a neutral ligand set and to the opposite ligand charge ±1.5 for all other cases. For the CCR3 receptor with all ligands bearing a formal charge of +1.0

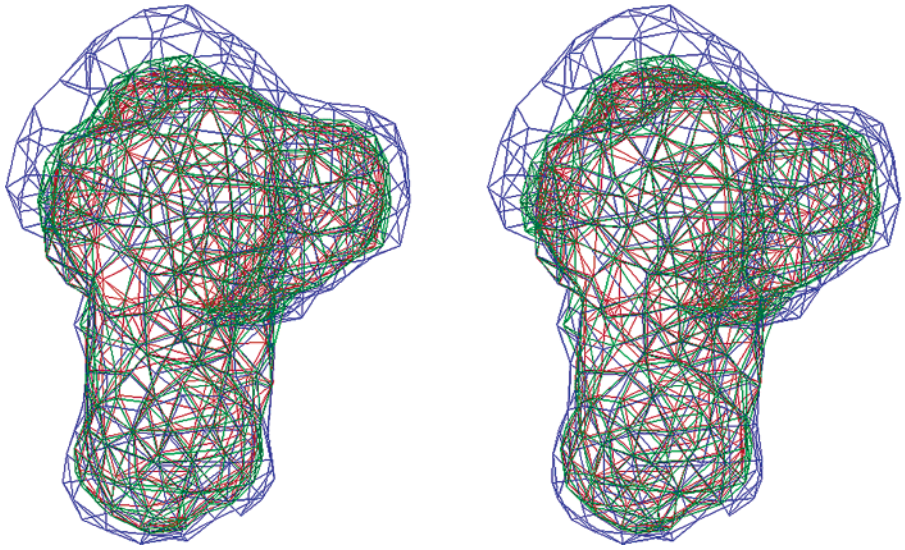


Figure 1. Stereoview of the envelope selection (induced-fit hypotheses) for the most potent antagonist molecule (**1101**) of series-2. For clarity, only 3 of 6 simulated induced-fit scenarios are shown. Color coding is as follows: red, steric-field mode; blue, H-bond-field mode; green, lipophilicity-potential mode.

(the protonated piperazyl or piperidyl N atom), the total charge of the surrogate must lie in the range from −2.5 to +0.5. Finally, the scoring function based on the directional Yeti force field[35] uses a 6/5 potential function for evaluating H-bond interactions instead of the previous 12/10 form. This reduces the sensitivity of the total energy on the H···acceptor distance. The directionality terms remain unaltered. These modifications make it more difficult (i.e., more time-consuming) to identify a robust family of receptor models, the benefit being the clearly safer prediction of affinities of new compounds.

Results and Discussion

Generation of the Ligand Data Set, Conformer, and Training-Set Selection. The three-dimensional structures of all ligand molecules (50 compounds for series-1 and 91 compounds for series-2) were generated using MacroModel 6.5[48] and optimized in aqueous solution based on the AMBER* force field.[49] An extensive conformational search was then performed for at least one compound of any molecular scaffold present in the data set (series-1, 6 compounds; series-2, 35 compounds) and two conceivable protonation states for each ligand, thereby allowing for 20 000 minimized structures each, again using MacroModel and simulating an aqueous environment. Therefore, all conformations within 10 kcal/mol (up to 4611 per molecule) from the lowest-energy conformer were retained to analyze the conformationally accessible space and to identify those entities allowing for a common pharmacophore hypothesis. On the basis of the analyses of these structures, up to eight conformers were then selected for each ligand, thereby defining a 4D data set. Because of the large number of low-energy conformers, we decided to only select entities featuring a similar backbone conformation but allowed for a variation in the side chain substitution. For the 50 ligands of series-1, this protocol resulted in 201 conformers; their internal strain ranged from 0 to 3.1 kcal/mol (average of 0.33 kcal/mol); for the 91 molecules comprising series-2, 210 conformers were selected, ranging from 0.0 to 5.1 kcal/mol in internal strain (average of 0.75 kcal/mol). In Quasar, the internal strain of a ligand is a component of the energy equation (cf. eq 1), which hampers the chance of a "high-energy" conformer to contribute to the Boltzmann-weighted ensemble (cf. eq 2). For each conformer, MNDO/ESP charges were then calculated using MOPAC 6.0.[50] The solvation energy was determined using the approach of Still et al.,[39] and its entropic contribution was determined following the approach of Searle and Williams.[40] The ligand topologies for series-1 are given in ref 27; those for series-2 are shown in Table 1. The training set was manually selected from the whole data set to obtain a maximal diversity based on the 2D substitution pattern. First, all ligands were ranked according to their experimental binding affinity. Then, starting with the most active compound (rank no. 1), any subsequent ligand (rank no. n) was marked "test ligand" if its 2D substitution pattern was present in any combination within the ligands ranked from 1 to (n − 1); otherwise, it was added to the training set.

Boundary Conditions for the 5D-QSAR Simulations. For both series, we made use of all six induced-fit scenarios available in Quasar: a linear induced fit scaled to 75%, four field-based modes (steric, electrostatic, H-bond, lipophilicity), and a protocol based on energy minimization.[25,35] An induced fit simulated using the steric-field as the determinant typically yields the tightest model —[series-1, 1.5 Å (min = 1.4, max = 1.8 Å); series-2, rms induced fit of 1.6 Å (min = 1.4, max = 2.0 Å)]. When we select the H-bond-field, the effect is only moderate —[series-1, 0.3 Å (min = 0.2, max = 0.5 Å); series-2, rms induced fit of 0.3 Å (min = 0.1, max = 0.6 Å)].

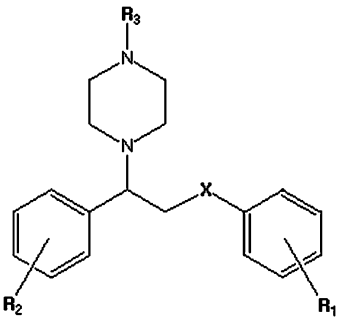
For series-1, the evolution was based on 250 receptor surrogates and simulated for 20 generations (5000 crossovers). The larger series-2 including three substance classes was simulated for 72 generations (18 000 crossovers). For all runs, the default mutation rate of 0.02 was used during transcription of the quasi-atomistic properties. The calculations were performed with Quasar 4.5, the most recent version of the software[35] on Macintosh G4 computer systems. The activity of the

Table 1. Ligand Substitution Pattern in Series-2 (cf. Chart 1)

ligand	R1	X	R2	R3	ligand	R1	X	R2	R3
				Training Set (66)					
1101	3,5-di(CH_3)	−N−(CH_2)$_2$−	4-OCH_3	cyclohexyl	562	4-F	−NH−CO−NH−	4-OCH_3	cyclohexyl
763	3,5-diCl	−N−(CH_2)$_2$−	4-OCH_3	cyclohexyl	597	4-CH_3	−NH−CO−NH−	4-CF_3	cyclohexyl
781	3,5-diCl	−N−(CH_2)$_2$−	H	cyclohexyl	795	4-N(CH_3)$_2$	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl
792	2-OCH_3	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	599	4-F	−NH−CO−NH−	4-CF_3	cyclohexyl
799	3,5-diCl	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	581		−NH−CO−NH−	4-TBu	cyclohexyl
1076	4-CH_3	−NH−CO−NH−	isobutyl[a]	3,4-diCl-benzene	689	4-F	−N−(CH_2)$_2$−	4-CF_3	phenyl
790	3,5-diCl	−N−(CH_2)$_2$−	4-TBu	cyclohexyl	595	4-CF_3	−NH−CO−NH−	4-CF_3	cyclohexyl
785	4-CF_3	−N−(CH_2)$_2$−	4-TBu	cyclohexyl	576	4-CF_3	−NH−CO−NH−		cyclohexyl
1066	3,5-diCl	−NH−CO−NH−	isobutyl[a]	cyclohexyl	558	4-CF_3	−NH−CO−NH−	4-OCH_3	cyclohexyl
1077	3,5-di(CH_3)	−NH−CO−NH−	isobutyl[a]	3,4-diCl-benzene	639	2-OCH_3	−N−(CH_2)$_2$−	2-OCH_3	phenyl
1079	4-CH_3	−NH−CO−NH−	4-phenyl	cyclohexyl	659	4-CF_3	−N−(CH_2)$_2$−	4-N(CH_3)$_2$	phenyl
1080	3,5-di(CH_3)	−NH−CO−NH−	4-phenyl	cyclohexyl	593	4-OCH_3	−NH−CO−NH−	4-CF_3	cyclohexyl
784	4-OCH_3	−N−(CH_2)$_2$−	4-TBu	cyclohexyl	586	4-N(CH_3)$_2$	−NH−CO−NH−	4-TBu	cyclohexyl
789	4-TBu	−N−(CH_2)$_2$−	4-TBu	cyclohexyl	669	4-N(CH_3)$_2$	−N−(CH_2)$_2$−		phenyl
782	H	−N−(CH_2)$_2$−	4-TBu	cyclohexyl	583	4-OCH_3	−NH−CO−NH−	4-TBu	cyclohexyl
685	4-OCH_3	−N−(CH_2)$_2$−	4-CF_3	phenyl	584	4-TBu	−NH−CO−NH−	4-TBu	cyclohexyl
1065	4-Cl	−NH−CO−NH−	isobutyl[a]	cyclohexyl	537	2-OCH_3	−NH−CO−NH−	4-CF_3	methyl
793	4-OCH_3	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	443	4-CH_3	−NH−CO−NH−	2-OCH_3	phenyl
591	H	−NH−CO−NH−	4-CF_3	cyclohexyl	592	2-OCH_3	−NH−CO−NH−	4-CF_3	cyclohexyl
675	2-OCH_3	−N−(CH_2)$_2$−	4-TBu	phenyl	737		−N−(CH_2)$_2$−	4-CF_3	methyl
1085	3,5-di(CH_3)	−N−(CH_2)$_2$−	4-phenyl	cyclohexyl	578	4-CH_3	−NH−CO−NH−		cyclohexyl
796	4-CH_3	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	601	2-OCH_3	−NH−CO−CH2−	2-OCH_3	phenyl
652	4-CH_3	−N−(CH_2)$_2$−	4-OCH_3	phenyl	580	4-F	−NH−CO−NH−		cyclohexyl
753	4-TBu	−N−(CH_2)$_2$−	4-OCH_3	cyclohexyl	517	4-F	−NH−CO−NH−	4-N(CH_3)$_2$	methyl
676	4-OCH_3	−N−(CH_2)$_2$−	4-TBu	phenyl	637	2-OCH_3	−NH−CO−CH2−	4-TBu	phenyl
794	4-CF_3	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	547	4-OCH_3	−NH−CO−NH−	2-OCH_3	cyclohexyl
797	4-F	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	707	4-F	−N−(CH_2)$_2$−	4-OCH_3	methyl
552	3,5-diCl	−NH−CO−NH−	2-OCH_3	cyclohexyl	484	4-OCH_3	−NH−CO−NH−	4-CF_3	phenyl
557	4-TBu	−NH−CO−NH−	4-OCH_3	cyclohexyl	710		−N−(CH_2)$_2$−		methyl
760	4-CH_3	−N−(CH_2)$_2$−	4-OCH_3	cyclohexyl	632	4-CH_3	−NH−CO−CH2−		phenyl
758	4-CF_3	−N−(CH_2)$_2$−	4-OCH_3	cyclohexyl	621	4-CF_3	−NH−CO−CH2−	4-N(CH_3)$_2$	phenyl
585	4-CF_3	−NH−CO−NH−	4-TBu	cyclohexyl	609		−NH−CO−CH2−	4-OCH_3	phenyl
570	3,5-diCl	−NH−CO−NH−	4-N(CH_3)$_2$	cyclohexyl	635	3,5-diCl	−NH−CO−CH2−		cyclohexyl
				Test Set (25)					
783	2-OCH_3	−N−(CH_2)$_2$−	4-TBu	cyclohexyl	560	4-CH_3	−NH−CO−NH−	4-OCH_3	cyclohexyl
786	4-N(CH_3)$_2$	−N−(CH_2)$_2$−	4-TBu	cyclohexyl	516	3,5-diCl	−NH−CO−NH−	4-N(CH_3)$_2$	methyl
754	3,5-diCl	−N−(CH_2)$_2$−	2-OCH_3	cyclohexyl	543	3,5-diCl	−NH−CO−NH−	4-CF_3	methyl
661	4-CH_3	−N−(CH_2)$_2$−	4-N(CH_3)$_2$	phenyl	553	4-F	−NH−CO−NH−	2-OCH_3	cyclohexyl
791		−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	472	4-F	−NH−CO−NH−		phenyl
687	4-N(CH_3)$_2$	−N−(CH_2)$_2$−	4-CF_3	phenyl	470	4-CH_3	−NH−CO−NH−		phenyl
561	3,5-diCl	−NH−CO−NH−	4-OCH_3	cyclohexyl	713	4-CF_3	−N−(CH_2)$_2$−		methyl
798	4-TBu	−N−(CH_2)$_2$−	4-CF_3	cyclohexyl	574	4-OCH_3	−NH−CO−NH−		cyclohexyl
598	3,5-diCl	−NH−CO−NH−	4-CF_3	cyclohexyl	612	4-CF_3	−NH−CO−CH2−	4-TBu	cyclohexyl
755		−N−(CH_2)$_2$−	4-OCH_3	cyclohexyl	542	4-CH_3	−NH−CO−NH−	4-CF_3	methyl
590	4-F	−NH−CO−NH−	4-TBu	cyclohexyl	636		−NH−CO−CH2−	4-TBu	phenyl
579	3,5-diCl	−NH−CO−NH−		cyclohexyl	715	4-CH_3	−N−(CH_2)$_2$−		methyl
776	4-CF_3	−N−(CH_2)$_2$−		cyclohexyl					

[a] The isobutyl group is directly bound to the backbone; that is, no phenyl linker is present.

Chart 1



ligands within series-1 covers a range of 5.4 kcal/mol, −a factor 10 992 in the IC_{50} value; those of series-2 cover a range of 4.2 kcal/mol, −a factor 1256 in the K_i value. The narrower range of activity within series-2 is due to insufficient solubility of low-activity compounds.

Results Obtained for Series-1. The simulation reached the target cross-validated r^2 of 0.950 already at 4658 crossovers (18.6 generations) and resulted in a predictive r^2 of 0.879. Both values reflect quantities averaged over the 250 models that, among themselves, differ in 30% [27−34%] of the mapped 129 properties on the 314 available positions of the surface. The cross-validation was based on five groups comprising eight ligands each ("leave-8-out"). A stereorepresentation of the receptor surrogate is depicted in Figure 2; experimental and calculated IC_{50} values are compared in Table 2 and shown in Figure 3 (left panel). The molecules and their experimental IC_{50} values are given in ref 27 (Tables 5−11).

The Y-shaped receptor model can be best described as a predominantly hydrophobic pocket with two distinct hydrophilic regions: the stem (bottom, cf. Figure 2) may be characterized as a H-bond acceptor-rich area and accommodates the hydrophilic-substituted central aromatic ring of the inhibitors. Its transition into the right lobe is amphiphilic with both H-bond donors and flip-flops; the right lobe itself is hydrophobic and binds

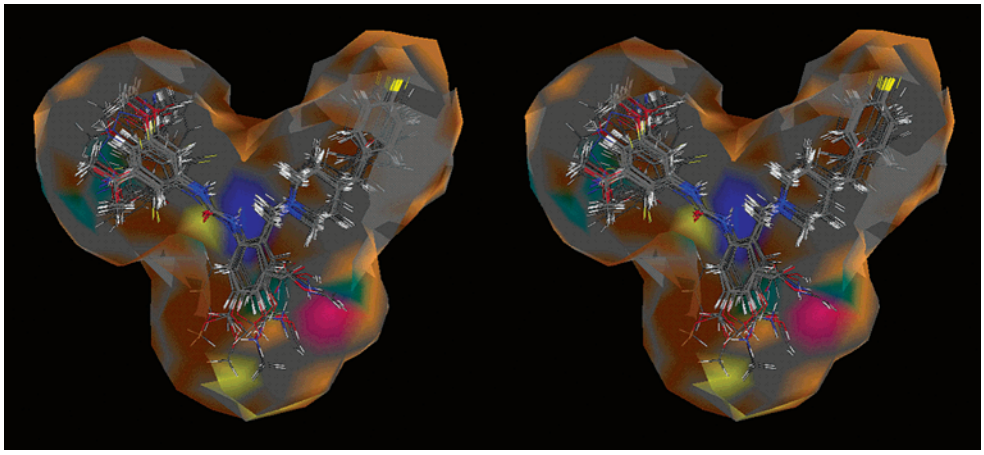


Figure 2. Stereorepresentation of the surrogate for the CCR3 receptor (series-1). The mapped properties are colored as follows: red, positively charged salt bridge (e.g., Arg-Lys); blue, negatively charged salt bridge (e.g., Asp-Glu); green, H-bond donor; yellow, H-bond acceptor; light brown, positively charged hydrophobic; dark brown, negatively charged hydrophobic; gray, neutral hydrophobic.

Table 2. Experimental and Calculated IC_{50} Values for Series-1[a]

ligand[27]	conformers	IC_{50} (exptl)[27]	IC_{50} (calcd)	factor off in IC_{50}	ligand[27]	conformers	IC_{50} (exptl)[27]	IC_{50} (calcd)	factor off in IC_{50}
				Training Set (40)					
113	4	1.0×10^{-9}	$(2.18 \pm 0.44) \times 10^{-9}$	1.2	96	4	6.7×10^{-8}	$(4.03 \pm 0.93) \times 10^{-8}$	0.7
115	8	1.0×10^{-9}	$(9.09 \pm 1.97) \times 10^{-10}$	0.1	93	2	8.5×10^{-8}	$(1.28 \pm 0.32) \times 10^{-7}$	0.5
121	4	1.0×10^{-9}	$(1.38 \pm 0.36) \times 10^{-9}$	0.4	89	4	1.14×10^{-7}	$(1.85 \pm 0.91) \times 10^{-7}$	0.6
111	8	2.0×10^{-9}	$(1.86 \pm 0.75) \times 10^{-9}$	0.1	77	4	1.25×10^{-7}	$(1.62 \pm 0.36) \times 10^{-7}$	0.3
114	8	2.0×10^{-9}	$(2.01 \pm 0.40) \times 10^{-9}$	0.0	95	4	1.41×10^{-7}	$(1.03 \pm 0.29) \times 10^{-7}$	0.4
120	2	2.0×10^{-9}	$(4.34 \pm 1.02) \times 10^{-9}$	1.2	91	4	2.28×10^{-7}	$(5.37 \pm 1.75) \times 10^{-8}$	3.2
102	8	3.0×10^{-9}	$(6.32 \pm 1.55) \times 10^{-9}$	1.1	90	2	2.6×10^{-7}	$(1.45 \pm 0.38) \times 10^{-7}$	0.8
110	4	3.0×10^{-9}	$(2.61 \pm 0.70) \times 10^{-9}$	0.1	76	2	3.1×10^{-7}	$(3.53 \pm 0.77) \times 10^{-7}$	0.1
105	4	7.0×10^{-9}	$(7.53 \pm 1.81) \times 10^{-9}$	0.1	118	2	2.77×10^{-7}	$(3.27 \pm 0.83) \times 10^{-7}$	0.2
101	4	8.0×10^{-9}	$(1.37 \pm 0.37) \times 10^{-8}$	0.7	43	2	4.5×10^{-7}	$(5.91 \pm 1.36) \times 10^{-7}$	0.3
103	8	8.0×10^{-9}	$(9.62 \pm 2.47) \times 10^{-9}$	0.2	44	2	4.5×10^{-7}	$(9.38 \pm 1.82) \times 10^{-7}$	1.1
104	2	9.0×10^{-9}	$(1.49 \pm 0.35) \times 10^{-8}$	0.7	85	2	5.29×10^{-7}	$(3.56 \pm 1.14) \times 10^{-7}$	0.5
106	8	1.1×10^{-8}	$(6.71 \pm 1.67) \times 10^{-9}$	0.6	40	4	6.0×10^{-7}	$(5.48 \pm 1.05) \times 10^{-7}$	0.1
117	2	1.1×10^{-8}	$(5.48 \pm 1.48) \times 10^{-9}$	1.0	41	2	7.5×10^{-7}	$(7.27 \pm 1.52) \times 10^{-7}$	0.0
109	2	1.7×10^{-8}	$(2.07 \pm 1.02) \times 10^{-8}$	0.2	42	2	7.5×10^{-7}	$(6.04 \pm 0.93) \times 10^{-7}$	0.2
119	2	2.1×10^{-8}	$(1.76 \pm 0.54) \times 10^{-8}$	0.2	86	8	7.77×10^{-7}	$(7.19 \pm 2.42) \times 10^{-7}$	0.1
98	8	4.2×10^{-8}	$(1.85 \pm 0.44) \times 10^{-8}$	1.3	38	1	9.0×10^{-7}	$(1.21 \pm 0.18) \times 10^{-6}$	0.3
100	8	4.7×10^{-8}	$(3.96 \pm 1.33) \times 10^{-8}$	0.2	36	3	2.6×10^{-6}	$(3.34 \pm 1.47) \times 10^{-6}$	0.3
108	2	5.3×10^{-8}	$(5.99 \pm 2.22) \times 10^{-8}$	0.1	61	4	4.65×10^{-6}	$(5.03 \pm 1.88) \times 10^{-6}$	0.1
97	8	6.4×10^{-8}	$(4.39 \pm 1.22) \times 10^{-8}$	0.5	62	2	1.1×10^{-5}	$(5.93 \pm 1.78) \times 10^{-6}$	0.9
				Test Set (10)					
112	2	2.0×10^{-9}	$(4.49 \pm 1.11) \times 10^{-9}$	1.2	88	8	3.6×10^{-7}	$(1.77 \pm 1.04) \times 10^{-7}$	1.0
107	8	8.0×10^{-9}	$(3.84 \pm 1.25) \times 10^{-9}$	1.1	84	4	4.23×10^{-7}	$(1.73 \pm 1.04) \times 10^{-7}$	1.5
94	4	4.0×10^{-8}	$(4.74 \pm 1.44) \times 10^{-8}$	0.2	75	1	3.61×10^{-7}	$(1.00 \pm 0.21) \times 10^{-6}$	1.8
99	4	5.5×10^{-8}	$(3.72 \pm 1.24) \times 10^{-8}$	0.5	39	2	1.75×10^{-6}	$(1.12 \pm 0.26) \times 10^{-6}$	0.6
92	2	2.37×10^{-7}	$(1.38 \pm 0.32) \times 10^{-7}$	0.7	37	2	3.4×10^{-6}	$(1.37 \pm 0.38) \times 10^{-6}$	1.5

[a] The standard deviation of the calculated values is derived from the variation over the 250 models comprising the surrogate family. The deviation factor is calculated as $(IC_{50,exptl}/IC_{50,calcd}) - 1.0$ for $IC_{50,exptl}/IC_{50,calcd} > 1.0$ and $(IC_{50,calcd}/IC_{50,exptl}) - 1.0$ otherwise. All values are given in M.

the benzyl or 4-*F*-benzyl portion of the ligand molecules. The left lobe is predominantly hydrophobic with a H-bond-donating domain interacting with the H-bond-accepting substituents on the aromatic ring adjacent to the urea functionality.

The rms deviation for the 40 ligand molecules of the training set is 0.27 kcal/mol (a factor 0.6 off in the IC_{50} value); the maximal individual deviation is 0.84 kcal/mol (3.2 off the IC_{50}). With respect to the induced-fit hypothesis, the simulated evolution converged at the steric-field model.[51] Ten compounds (not used for model construction) were selected for testing the predictive power of the receptor surrogate yielding a predictive r^2 of 0.879. On the average, the predicted IC_{50} value of the test ligands deviates by 0.41 kcal/mol from the experiment (1.0 off the IC_{50}); the maximal observed deviation is 0.59 kcal/mol (1.8 off the IC_{50}). A series of five scramble tests (with an average predictive r^2 of -0.415) demonstrated the sensitivity of the surrogate toward the biological data (Figure 3, right panel).

To identify potential sites and functionalities allowing a further increase of the binding affinity, the individual functional groups of both training and test sets were analyzed for their contribution toward the free energy of ligand binding, ΔG. Table 3 lists all details of enthalpic (electrostatic, van der Waals, H-bond, and polarization terms), entropic, solvation, and induced-fit contributions toward the calculated binding affinity. Figure 4 shows a comparison for the most potent ligand of the series (**113**; $IC_{50} = 1.0$ nM) with an only moderate

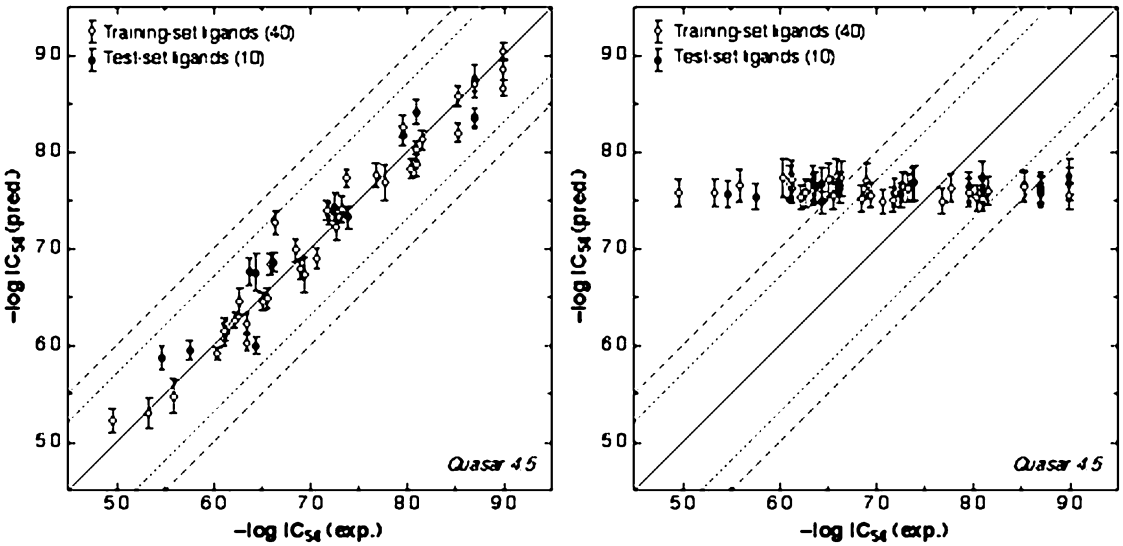


Figure 3. Comparison of experimental and predicted binding affinities for the CCR3 receptor (series-1): correct simulation (left); scramble test (right).

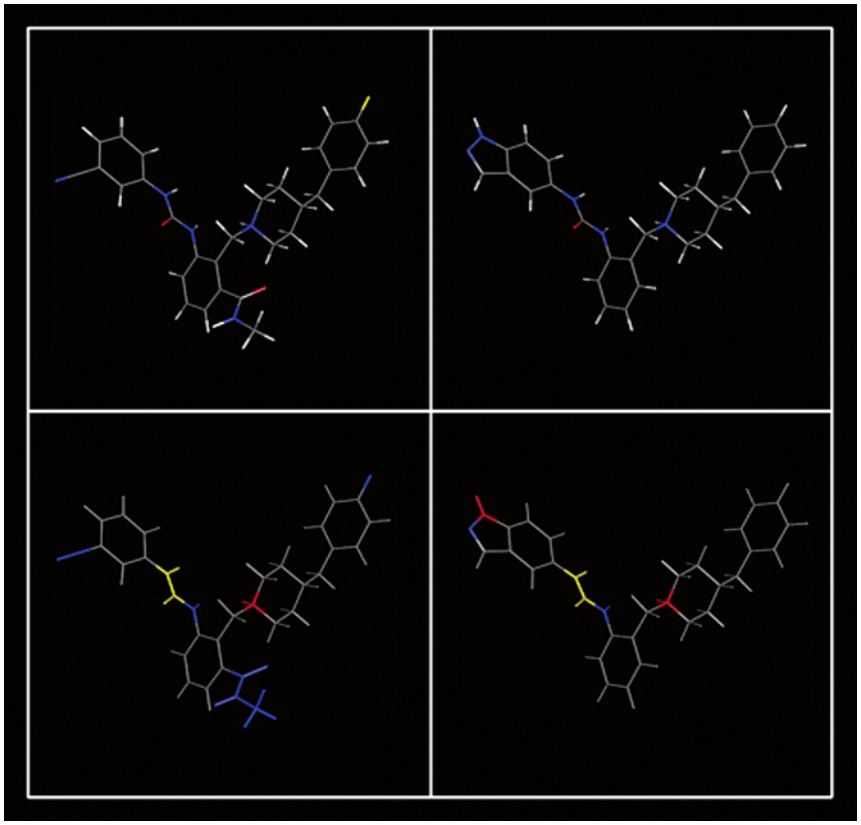


Figure 4. Functional-group analysis of compounds **113** (left) and **43** (right); top, colored by atom type; bottom, contribution of the individual functional groups to ΔG_{pred}. Color coding is as follows: blue, excellent; green, good; white, neutral; yellow, weak; red, poor. Residual fragments (not specifically decomposed interaction scheme) are shown in gray.

binder (**43**; $IC_{50} = 0.45\ \mu M$). The difference would seem to be explainable by the "destabilizing" presence of the H-bond donor functionality in the indole ring of **43**: its interaction with the receptor surrogate yields −1.8 kcal/mol; its contribution to ligand desolvation (4.4 kcal/mol) is not compensated, instead leading to a net unfavorable contribution of $\Delta\Delta G = 3.3$ kcal/mol for this functional group (2.6 kcal/mol thereof is associated with ligand desolvation, 0.7 kcal/mol with the entropic change during ligand binding; the contributions of both induced fit and internal strain are zero). On the other hand, the methyl amide moiety of **113** contributes favorably toward ligand binding: the interaction with the receptor is calculated to −11.6; desolvation contributes with −9.0, yielding a net contribution of $\Delta\Delta G = -20.6$ kcal/mol with the main component being the favorable desolvation of the $-NH-CH_3$ group. The net calculated ΔG is −11.6 kcal/mol for **113** and −8.1 kcal/mol for **43**, respectively.

Results Obtained for Series-2. After 18 000 crossovers (72 generations), the simulation reached the cross-validated r^2 of 0.861 and a predictive r^2 of 0.798. These

Table 3. Functional Group Analysis for Series-1[a]

ligand	total	electrostatic	van der Waals	H-bond	polarization	solvation	$T\Delta S$	internal	ind.fit	ΔG_{calcd}
113	−69.817	−29.578	−20.683	−10.676	−8.881	−55.189	−2.8	0.000	0.218	−11.610
115	−71.191	−29.605	−20.235	−12.106	−9.245	−55.779	−2.8	0.281	0.212	−12.120
121	−71.760	−29.410	−21.949	−10.722	−9.679	−56.585	−2.8	0.285	0.214	−11.876
111	−73.376	−29.126	−18.798	−16.369	−9.083	−58.351	−2.8	0.306	0.217	−11.702
114	−68.901	−29.459	−19.179	−11.125	−9.137	−53.263	−3.5	0.268	0.213	−11.657
120	−70.165	−29.944	−22.509	−8.502	−9.211	−55.937	−2.8	0.000	0.218	−11.210
102	−77.791	−31.682	−23.027	−13.501	−9.581	−63.493	−2.8	0.285	0.222	−10.991
110	−72.174	−30.025	−20.667	−12.453	−9.029	−57.639	−2.8	0.000	0.228	−11.506
105	−69.655	−29.722	−19.748	−11.372	−8.814	−54.910	−3.5	0.136	0.219	−10.889
101	−76.441	−31.947	−23.947	−11.290	−9.258	−62.868	−2.8	0.000	0.233	−10.540
103	−75.255	−34.199	−18.599	−11.866	−10.591	−61.225	−2.8	0.273	0.211	−10.746
104	−66.821	−28.786	−17.886	−11.535	−8.614	−52.472	−3.5	0.135	0.222	−10.492
106	−68.803	−29.674	−18.267	−11.728	−9.135	−53.017	−4.2	0.413	0.217	−10.956
117	−66.495	−29.956	−19.575	−8.407	−8.557	−52.400	−2.8	0.000	0.220	−11.074
109	−72.027	−29.249	−18.986	−15.183	−8.609	−57.331	−3.5	0.665	0.230	−10.300
119	−68.646	−30.776	−20.652	−9.591	−7.627	−54.526	−3.5	0.000	0.226	−10.394
98	−72.886	−31.594	−20.961	−10.756	−9.575	−58.535	−3.5	0.271	0.215	−10.365
100	−70.442	−30.869	−18.127	−12.307	−9.138	−55.095	−4.9	0.302	0.223	−9.923
108	−65.356	−28.901	−16.553	−11.727	−8.176	−51.241	−4.2	0.003	0.229	−9.682
97	−70.582	−32.184	−19.844	−8.957	−9.596	−56.010	−4.2	0.296	0.214	−9.862
96	−71.550	−32.111	−21.789	−8.472	−9.178	−57.919	−3.5	0.000	0.219	−9.913
93	−68.893	−31.064	−20.025	−8.115	−9.689	−56.626	−2.8	0.000	0.226	−9.241
89	−67.797	−30.950	−17.750	−11.092	−8.005	−54.350	−4.2	0.001	0.222	−9.024
77	−68.642	−30.230	−19.498	−10.319	−8.596	−56.250	−2.8	0.269	0.220	−9.103
95	−68.217	−31.526	−18.002	−8.503	−10.187	−54.820	−3.5	0.312	0.221	−9.364
91	−66.981	−31.012	−16.144	−10.454	−9.371	−53.951	−2.8	0.260	0.224	−9.746
90	−67.257	−32.388	−15.614	−9.374	−9.881	−55.064	−2.8	0.000	0.225	−9.167
76	−67.376	−31.260	−19.721	−8.189	−8.206	−55.713	−2.8	0.000	0.215	−8.649
118	−68.081	−32.483	−18.686	−8.983	−7.929	−55.658	−3.5	0.000	0.230	−8.693
43	−65.491	−31.680	−16.063	−9.266	−8.483	−54.115	−2.8	0.000	0.227	−8.349
44	−65.584	−32.021	−15.810	−9.323	−8.429	−54.468	−2.8	0.007	0.228	−8.080
85	−65.750	−31.719	−19.003	−6.617	−8.411	−54.084	−2.8	0.000	0.222	−8.644
40	−62.902	−30.975	−15.420	−8.580	−7.926	−50.483	−3.5	0.303	0.223	−8.393
41	−62.098	−30.809	−15.344	−8.210	−7.736	−50.148	−3.5	0.002	0.220	−8.228
42	−63.741	−30.691	−17.488	−8.037	−7.526	−52.375	−2.8	0.000	0.229	−8.337
86	−65.957	−33.503	−13.920	−9.113	−9.422	−52.968	−4.2	0.337	0.217	−8.235
38	−60.816	−30.079	−15.217	−8.228	−7.292	−49.850	−2.8	0.000	0.232	−7.934
36	−65.347	−27.468	−16.436	−13.402	−8.040	−52.186	−5.6	0.009	0.211	−7.341
61	−66.693	−32.255	−16.328	−9.188	−8.922	−57.164	−2.1	0.095	0.232	−7.102
62	−62.383	−32.050	−14.951	−7.308	−8.074	−53.060	−2.1	0.011	0.205	−7.007
112	−66.885	−28.534	−18.775	−10.903	−8.673	−52.674	−2.8	0.000	0.221	−11.190
107	−71.234	−29.809	−19.204	−13.039	−9.182	−55.760	−3.5	0.481	0.212	−11.280
94	−70.623	−31.090	−19.120	−10.337	−10.077	−57.281	−2.8	0.503	0.221	−9.818
99	−71.455	−31.088	−19.786	−11.791	−8.790	−57.062	−4.2	0.008	0.225	−9.959
92	−65.585	−31.292	−17.110	−8.156	−9.026	−53.362	−2.8	0.000	0.229	−9.195
88	−67.042	−31.065	−16.012	−11.534	−8.431	−52.567	−4.9	0.304	0.221	−9.050
84	−65.343	−31.436	−17.764	−7.298	−8.845	−52.285	−3.5	0.271	0.221	−9.066
75	−64.244	−29.705	−18.437	−8.089	−8.011	−53.174	−2.8	0.000	0.228	−8.042
39	−63.794	−30.690	−16.779	−7.832	−8.494	−52.786	−2.8	0.001	0.228	−7.979
37	−61.087	−30.423	−15.111	−8.154	−7.400	−48.095	−4.2	0.726	0.205	−7.861

[a] Top, training set; bottom, test set. All values are given in kcal/mol.

quantities reflect values averaged over the 250 models that, among themselves, differ in 20% [17−29%] of the mapped 150 properties on 294 available positions of the surface. The cross-validation was based on six groups comprising 11 ligands each ("leave-11-out"). A stereo-representation of the quasi-atomistic receptor surrogate is depicted in Figure 5; experimental and calculated K_i values are compared in Figure 6 (left panel) and Table 4.

The pear-shaped (or truncated Y-shaped) receptor model can be best described as a predominantly hydrophobic pocket with two distinct hydrophilic regions: one opposite to the protonated piperazyl N-atom mimicking the salt-bridge counterpart (Asp or Glu), and the other accommodating the amphiphilic substituents of the central aromatic ring (Figure 5). The subtleties of interacting with the predominantly aromatic portions of the ligand molecules are brought about by the electrostatic distribution within the hydrophobic portion of the surrogate.

The rms deviation for the 66 ligand molecules of the training set is 0.32 kcal/mol (a factor 0.7 off in the K_i value); the maximal individual deviation is 0.78 kcal/mol (2.9 off K_i). With respect to the induced-fit hypothesis, the simulated evolution converged at the model based on the electrostatic field[51] which reflects the ligand data set sparse in H-bond-donating and -accepting functionalities. Twenty-five compounds (not used for model construction) were selected for testing the predictive power of the receptor surrogate yielding a predictive r^2 of 0.798. On the average, the predicted K_i value of the test ligands deviates by 0.33 kcal/mol from the experiment (0.8 off K_i); the maximal observed deviation is 1.12 kcal/ mol (5.8 off K_i). A series of five scramble

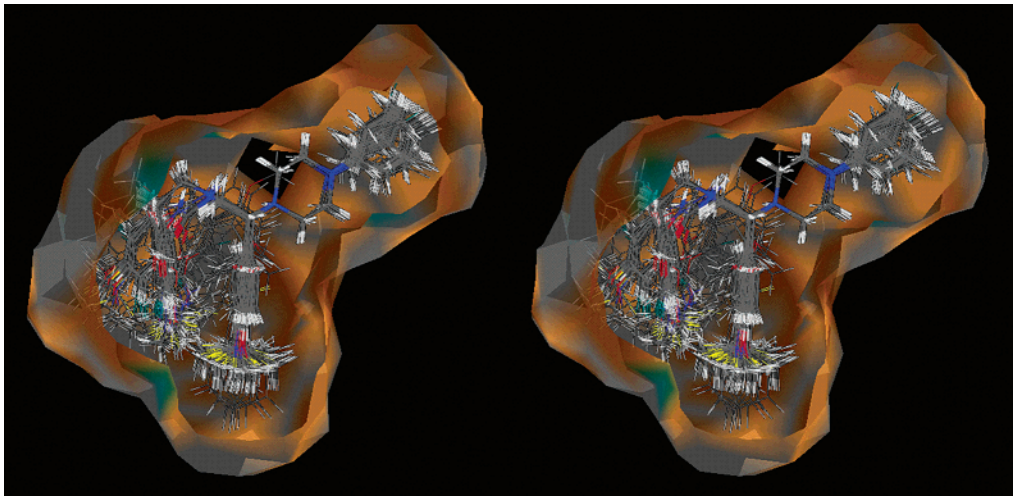


Figure 5. Stereorepresentation of the surrogate for the CCR3 receptor (series-1). Color coding cf. caption to Figure 2.

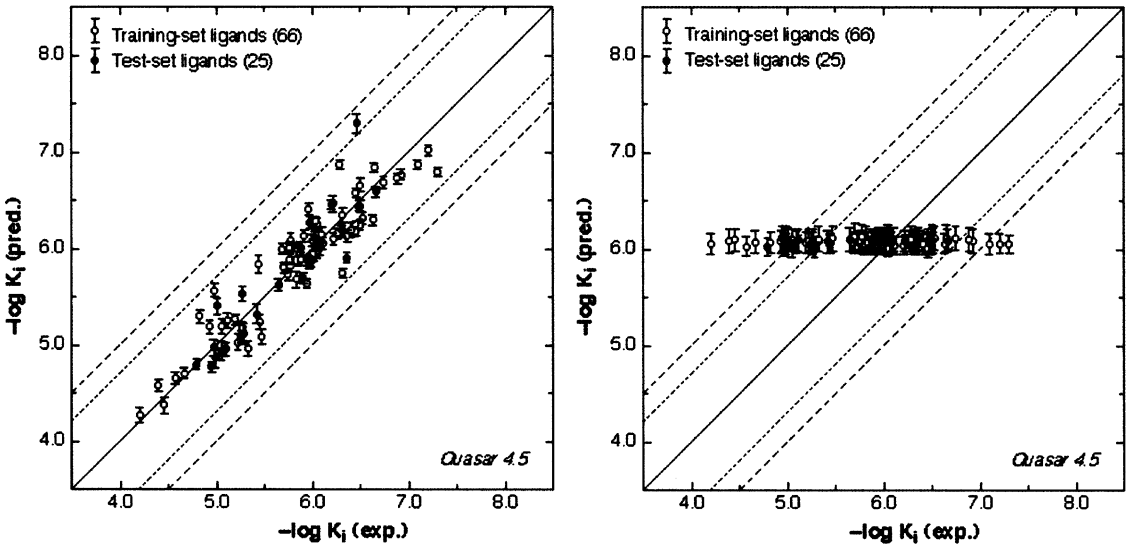


Figure 6. Comparison of experimental and predicted binding affinities for the CCR3 receptor (series-2): correct simulation (left); scramble test (right).

tests (with an average predictive r^2 of −0.133; cf. Figure 6, right panel) demonstrated the sensitivity of the surrogate toward the biological data.

Combination of the Receptor Surrogates. From a pharmacological point of view, both series of compounds used for this study compete for the same binding site at the CCR3 chemokine receptor. Consequently, one needs to probe the spatial matching of the two surrogates, as they were evolved independently from different ligand sets. The reason we did not envision this firsthand is of "technical" nature: the binding affinities of series-1 (measured at Brystol-Myers Squibb) are provided as IC_{50} values while those of series-2 (obtained from Boehringer Ingelheim) are K_i values. The Cheng−Prusoff equation[52] allows the conversion of K_i and IC_{50} for our QSAR study by using $A(K) = 0.15$ nM[27] and $K_p = 0.60$ nM.[52] As the experiments were conducted by different laboratories, the results should be interpreted with caution.

As can be seen from Figures 2 and 5, the 3D space occupied by the ligand molecules of series-2 can be understood as a subset of the volume defined by the ligands of series-1, but indicating a possibly significant amount of induced fit, particularly in the *left-lobe region*.

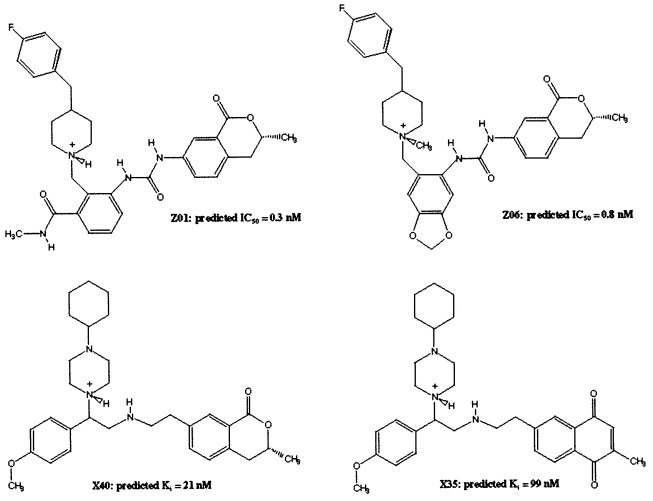


Figure 7. The four most potent proposed CCR3 antagonists: Z01 and Z06 (series-1, top panel, left and right, respectively); X40 and X35 (series-2, bottom panel, left and right, respectively). The given binding affinities represent values predicted by the receptor models and have not yet been experimentally verified. The associated standard deviations are moderate and range from a factor of 0.35 (X40) to a factor of 1.0 (Z01) in the K_i and IC_{50} values, respectively.

Table 4. Experimental and Calculated K_i Values for Series-2[a]

ligand	conformers	K_i (exptl)	K_i (calcd)	factor off in K_i	ligand	conformers	K_i (exptl)	K_i (calcd)	factor off in K_i
				Training Set (66)					
1101	4	5.00×10^{-8}	$(1.61 \pm 0.15) \times 10^{-7}$	2.2	562	2	1.03×10^{-6}	$(1.36 \pm 0.17) \times 10^{-6}$	0.3
763	4	6.20×10^{-8}	$(9.64 \pm 1.21) \times 10^{-8}$	0.6	597	1	1.07×10^{-6}	$(1.45 \pm 0.17) \times 10^{-6}$	0.4
781	2	8.00×10^{-8}	$(1.36 \pm 0.16) \times 10^{-7}$	0.7	795	4	1.08×10^{-6}	$(3.87 \pm 0.59) \times 10^{-7}$	1.8
792	4	1.18×10^{-7}	$(1.71 \pm 0.23) \times 10^{-7}$	0.5	599	1	1.11×10^{-6}	$(1.37 \pm 0.15) \times 10^{-6}$	0.2
799	2	1.29×10^{-7}	$(1.89 \pm 0.23) \times 10^{-7}$	0.5	581	1	1.13×10^{-6}	$(2.30 \pm 0.27) \times 10^{-6}$	1.0
1076	2	1.79×10^{-7}	$(2.04 \pm 0.29) \times 10^{-7}$	0.1	689	2	1.20×10^{-6}	$(7.42 \pm 1.07) \times 10^{-7}$	0.6
790	2	2.24×10^{-7}	$(1.46 \pm 0.15) \times 10^{-7}$	0.5	595	1	1.31×10^{-6}	$(1.29 \pm 0.22) \times 10^{-6}$	0.0
785	2	2.31×10^{-7}	$(5.07 \pm 0.69) \times 10^{-7}$	1.2	576	1	1.46×10^{-6}	$(2.07 \pm 0.40) \times 10^{-6}$	0.4
1066	1	2.96×10^{-7}	$(4.88 \pm 0.67) \times 10^{-7}$	0.6	558	2	1.53×10^{-6}	$(1.04 \pm 0.17) \times 10^{-6}$	0.5
1077	2	3.13×10^{-7}	$(2.18 \pm 0.36) \times 10^{-7}$	0.4	639	8	1.58×10^{-6}	$(1.15 \pm 0.26) \times 10^{-6}$	0.4
1079	1	3.16×10^{-7}	$(5.68 \pm 0.77) \times 10^{-7}$	0.8	659	4	1.67×10^{-6}	$(8.22 \pm 1.33) \times 10^{-7}$	1.0
1080	1	3.26×10^{-7}	$(3.61 \pm 0.48) \times 10^{-7}$	0.1	593	2	1.71×10^{-6}	$(1.28 \pm 0.16) \times 10^{-6}$	0.3
784	4	3.46×10^{-7}	$(5.66 \pm 0.48) \times 10^{-7}$	0.6	586	2	1.72×10^{-6}	$(9.73 \pm 1.24) \times 10^{-7}$	0.8
789	2	3.49×10^{-7}	$(2.65 \pm 0.31) \times 10^{-7}$	0.3	669	4	1.77×10^{-6}	$(1.86 \pm 0.30) \times 10^{-6}$	0.0
782	2	3.50×10^{-7}	$(6.77 \pm 0.73) \times 10^{-7}$	0.9	583	2	1.98×10^{-6}	$(1.58 \pm 0.19) \times 10^{-6}$	0.3
685	4	3.93×10^{-7}	$(6.48 \pm 1.19) \times 10^{-7}$	0.6	584	1	2.05×10^{-6}	$(1.00 \pm 0.12) \times 10^{-6}$	1.0
1065	1	4.29×10^{-7}	$(7.42 \pm 1.01) \times 10^{-7}$	0.7	537	2	3.35×10^{-6}	$(8.18 \pm 1.56) \times 10^{-6}$	1.4
793	4	4.77×10^{-7}	$(4.50 \pm 0.84) \times 10^{-7}$	0.1	443	2	3.54×10^{-6}	$(5.78 \pm 1.09) \times 10^{-6}$	0.6
591	1	4.79×10^{-7}	$(1.81 \pm 0.20) \times 10^{-6}$	2.8	592	2	3.65×10^{-6}	$(1.46 \pm 0.29) \times 10^{-6}$	1.5
675	4	4.88×10^{-7}	$(6.66 \pm 0.92) \times 10^{-7}$	0.4	737	2	4.73×10^{-6}	$(1.10 \pm 0.19) \times 10^{-5}$	1.3
1085	4	5.17×10^{-7}	$(1.35 \pm 0.14) \times 10^{-7}$	2.8	578	1	5.39×10^{-6}	$(6.69 \pm 0.87) \times 10^{-6}$	0.2
796	2	5.45×10^{-7}	$(6.98 \pm 0.81) \times 10^{-7}$	0.3	601	2	6.00×10^{-6}	$(9.36 \pm 1.90) \times 10^{-6}$	0.6
652	4	5.97×10^{-7}	$(8.04 \pm 1.16) \times 10^{-7}$	0.3	580	1	6.36×10^{-6}	$(5.47 \pm 0.66) \times 10^{-6}$	0.2
753	4	6.29×10^{-7}	$(3.56 \pm 0.37) \times 10^{-7}$	0.8	517	2	7.75×10^{-6}	$(5.55 \pm 1.07) \times 10^{-6}$	0.4
676	4	7.68×10^{-7}	$(8.88 \pm 1.30) \times 10^{-7}$	0.2	637	2	8.85×10^{-6}	$(6.43 \pm 1.16) \times 10^{-6}$	0.4
794	2	7.82×10^{-7}	$(7.13 \pm 1.20) \times 10^{-7}$	0.1	547	4	1.05×10^{-5}	$(2.73 \pm 0.44) \times 10^{-6}$	2.9
797	2	8.95×10^{-7}	$(5.63 \pm 0.81) \times 10^{-7}$	0.6	707	4	1.17×10^{-5}	$(6.45 \pm 1.06) \times 10^{-6}$	0.8
552	2	8.97×10^{-7}	$(1.05 \pm 0.21) \times 10^{-6}$	0.2	484	2	1.48×10^{-5}	$(5.09 \pm 0.85) \times 10^{-6}$	1.9
557	2	9.10×10^{-7}	$(9.12 \pm 1.17) \times 10^{-7}$	0.0	710	2	2.18×10^{-5}	$(1.97 \pm 0.25) \times 10^{-5}$	0.1
760	4	9.14×10^{-7}	$(5.26 \pm 0.68) \times 10^{-7}$	0.7	632	1	2.69×10^{-5}	$(2.24 \pm 0.33) \times 10^{-5}$	0.2
758	4	9.29×10^{-7}	$(7.82 \pm 1.50) \times 10^{-7}$	0.2	621	2	3.49×10^{-5}	$(4.24 \pm 0.83) \times 10^{-5}$	0.2
585	1	9.72×10^{-7}	$(1.13 \pm 0.19) \times 10^{-6}$	0.2	609	2	4.02×10^{-5}	$(2.65 \pm 0.37) \times 10^{-5}$	0.5
570	2	9.87×10^{-7}	$(8.21 \pm 1.39) \times 10^{-7}$	0.2	635	1	6.25×10^{-5}	$(5.37 \pm 0.97) \times 10^{-5}$	0.2
				Test Set (24)					
783	4	2.16×10^{-7}	$(2.57 \pm 0.34) \times 10^{-7}$	0.2	560	2	2.25×10^{-6}	$(2.40 \pm 0.32) \times 10^{-6}$	0.1
786	4	3.12×10^{-7}	$(3.68 \pm 0.51) \times 10^{-7}$	0.2	516	2	3.81×10^{-6}	$(4.82 \pm 1.24) \times 10^{-6}$	0.3
754	4	3.44×10^{-7}	$(5.06 \pm 1.20) \times 10^{-8}$	5.8	543	1	5.18×10^{-6}	$(7.68 \pm 2.23) \times 10^{-6}$	0.5
661	4	4.35×10^{-7}	$(1.25 \pm 0.16) \times 10^{-6}$	1.9	553	2	5.38×10^{-6}	$(2.98 \pm 0.53) \times 10^{-6}$	0.8
791	2	4.84×10^{-7}	$(6.52 \pm 0.86) \times 10^{-7}$	0.3	472	1	5.55×10^{-6}	$(8.18 \pm 1.24) \times 10^{-6}$	0.5
687	4	6.05×10^{-7}	$(3.42 \pm 0.70) \times 10^{-7}$	0.8	470	1	8.02×10^{-6}	$(1.10 \pm 0.17) \times 10^{-5}$	0.4
561	2	8.30×10^{-7}	$(9.91 \pm 1.95) \times 10^{-7}$	0.2	713	2	8.97×10^{-6}	$(1.15 \pm 0.26) \times 10^{-5}$	0.3
798	2	8.32×10^{-7}	$(8.90 \pm 1.31) \times 10^{-7}$	0.1	574	2	9.83×10^{-6}	$(3.96 \pm 0.81) \times 10^{-6}$	1.5
598	1	9.32×10^{-7}	$(7.80 \pm 1.26) \times 10^{-7}$	0.2	612	1	1.03×10^{-5}	$(1.37 \pm 0.35) \times 10^{-5}$	0.3
755	4	1.04×10^{-6}	$(5.30 \pm 0.83) \times 10^{-7}$	1.0	542	1	1.07×10^{-5}	$(1.07 \pm 0.20) \times 10^{-5}$	0.0
590	1	1.12×10^{-6}	$(1.32 \pm 0.15) \times 10^{-6}$	0.2	636	1	1.12×10^{-5}	$(1.69 \pm 0.22) \times 10^{-5}$	0.5
579	1	1.25×10^{-6}	$(2.02 \pm 0.26) \times 10^{-6}$	0.6	715	2	1.61×10^{-5}	$(1.59 \pm 0.22) \times 10^{-5}$	0.0
776	2	1.31×10^{-6}	$(9.97 \pm 1.63) \times 10^{-7}$	0.3					

[a] The standard deviation of the calculated values is derived from the variation over the 250 models comprising the surrogate family. The deviation factor is calculated as $(K_{i,exptl}/K_{i,calcd}) - 1.0$ for $K_{i,exptl}/K_{i,calcd} > 1.0$ and $(K_{i,calcd}/K_{i\ exptl}) - 1.0$ otherwise. All values are given in M.

As the two models were evolved independently, we selected the lowest-energy conformers for the pharmacophore construction, respectively (cf. above). The ligands of series-2 can match the volume occupied by those of series-1 far better, if the torsion angle linking the X group with the aromatic ring bearing R_1 (cf. Chart 1) is altered from −60° (gauche) to 180° (trans). The internal strain associated with this change raises by 0.8 to 1.1 kcal/mol, a modest increase which is appropriately considered in the Quasar energy eq 1. The selection of training and test sets was not altered with respect to the individual simulations nor were the boundary conditions of the 5D-QSAR study.

After 8000 crossovers (32 generations), the simulation reached the cross-validated r^2 of 0.907 and a predictive r^2 of 0.899. These quantities reflect values averaged over the 250 models that, among themselves, differ in 27% [21−29%] of the mapped 138 properties on 342 available positions of the surface. The surrogate is depicted in Figure 8; experimental and calculated IC_{50} values are compared in Figure 9. The rms deviation for the 106 ligand molecules of the training set (40 from series-1 and 66 from series-2) is 0.40 kcal/mol (a factor 1.0 off in the IC_{50} value); the maximal individual deviation is 1.2 kcal/mol (7.1 off the IC_{50}). Thirty-five compounds (10 + 25) defined the test set and yielded a predictive r^2 of 0.899. On the average, the predicted IC_{50} value for the test ligands deviates 0.34 kcal/mol from the experiment (0.8 off the IC_{50}); the maximal deviation is 0.88 kcal/mol (3.5 in the IC_{50}).

Novel Compounds. We evaluated a total of 58 novel molecules and tested them in silico against the two receptor surrogates. Thereby, two substitution protocols were applied: (1) lipophilic substituents ($-CH_3$, $-CF_3$, $-F$, $-Cl$, $-C{\equiv}N$) increasing hydrophobic interactions and simultaneously reducing the desolvation energy and

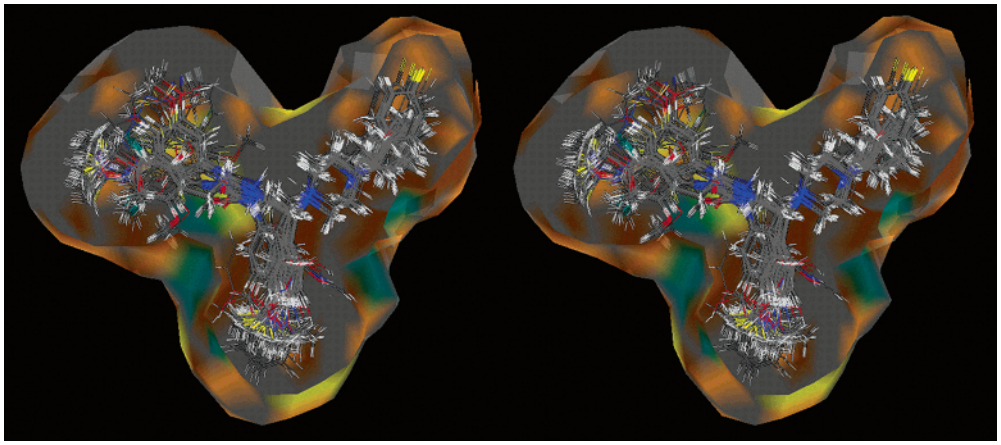


Figure 8. Stereo representation of the surrogate for the CCR3 receptor (combined series). The mapped properties are colored as follows: red, positively charged salt bridge (e.g., Arg-Lys); blue, negatively charged salt bridge (e.g., Asp-Glu); green, H-bond donor; yellow, H-bond acceptor; light brown, positively charged hydrophobic; dark brown, negatively charged hydrophobic; gray, neutral hydrophobic.

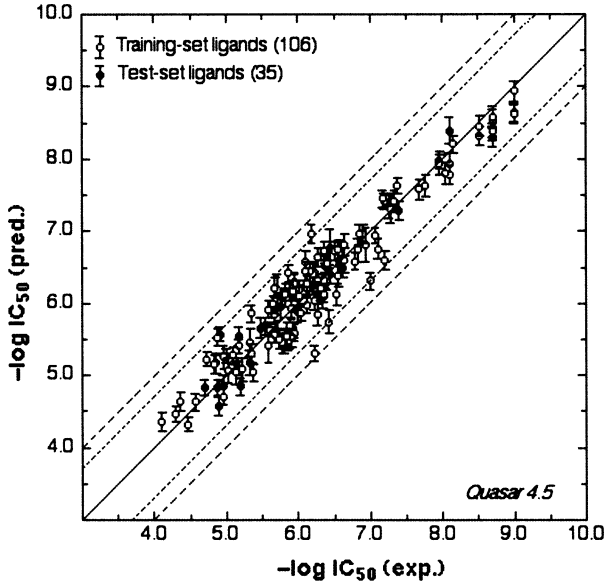


Figure 9. Comparison of experimental and predicted binding affinities for the CCR3 receptor (combined series).

(2) H-bond-accepting moieties (acetate, pyridine, furane, sulfonamide, lactone, ether, isocoumarine, naphthochinone) aimed at strengthening hydrogen-bond interactions but with an only moderate cost in ligand desolvation by using acceptor groups slightly amphiphilic in character.

For series-1, 10 novel ligand molecules were tested using the receptor surrogate. Their calculated IC_{50} value ranges from 0.3 to 16 nM. The two most potent compounds in silico, Z01 (Figure 7, top, left panel) with a predicted IC_{50}value of 0.3 nM and Z06 (Figure 7, top, right panel; IC_{50} = 0.8 nM) bearing an isocoumarine group, would seem to represent realistic assessments of the situation at the true biological receptor as they ideally combine a H-bond-accepting moiety with a delocalized, polarizable ring system. Z01 and Z06 differ solely by their central-ring substitution pattern: a 4-acetamido substituent versus a bridged 3,4-dioxymethylene (cf. Figure 7). Other modifications included a naphthochinone group in place of the isocoumarine functionality (here, the 4-acetamido substitution at the central ring yielded an IC_{50} of 1.6 nM; a 3,4-dioxymethylene moiety yielded an IC_{50} of 3.0 nM), a pyridine ring (5.3 and 16 nM), 3,5-dichloro (3.1 and 8.4 nM), and 3,5-dimethyl (1.4 and 4.2 nM).

For series-2, the 48 tested molecules yielded calculated K_i values from 21 nM to 2.5 μM. The most potent novel ligand X40 (K_i = 21 nM; Figure 7, bottom, left panel) bears the very isocoumarine ring as the compounds leading to the highest predicted affinities within series-1 (Z01 and Z06); additional modifications at the central aromatic ring (4-CF_3, 4-CH_3, 4-$N(CH_3)_2$) yield calculated K_i values ranging from 45 to 82 nM. Compound X35 featuring a naphthochinone group still shows a computed activity within a factor or 2 from the most potent compound of the training set (Figure 7, bottom, right panel; K_i = 99 nM).

Conclusions

In absence of the 3D receptor structure, multidimensional QSAR techniques provide an elegant approach for the estimation of free energies of ligand binding. The receptor modeling software Quasar developed at the Biographics Laboratory 3R is based on 5D-QSAR and explicitly allows for the simulation of induced fit. To determine the ligand−receptor interactions, the scoring function makes use of a directional force field. Free energies of ligand binding are then derived based on the ligand−receptor interaction, ligand desolvation, entropy, as well as terms for internal strain of the ligand molecule and induced fit.

The Quasar concept has been used for the design of novel compounds able in silico to inhibit the CCR3 receptor in the low-nanomolar range. For model validation, we made use of two different series of data including a total of 141 known CCR3 antagonist molecules. The simulations yielded a cross-validated r^2 of 0.950 and 0.861 (106 training ligands) and predictive r^2 of 0.877 and 0.798 (for 35 test ligands), respectively. On the basis of the two surrogate families (comprising 250 models each), we evaluated a total of 58 novel molecules for which various substitution patterns were applied. Nine of the proposed ligands show a calculated IC_{50} < 10 nM (series-1) and K_i < 100 nM (series-2), respectively. The most potent candidate is expected to bind with an IC_{50} of 0.3 nM (series-1) and a K_i of 21 nM (series-2), respectively. With the possibility to calculate

the contribution of each functional toward the free energy of ligand binding, Quasar offers a straightforward way to identify novel, potent ligand molecules and to predict their activity close to experimental uncertainty. By merging the two ligand sets, we could validate the combined model at a high level of confidence (cross-validated r^2 of 0.907, predictive r^2 of 0.899).

Information on Quasar may be obtained at http://www.biograf.ch/software.html and from http://www.biograf.ch/PDFS/Quasar.pdf (program documentation). The Biographics Laboratory 3R is a nonprofit organization aimed at replacing animal models in biomedical research by computational technologies (http://www.biograf.ch).

Acknowledgment. Financial support by Boehringer Ingelheim, Germany, and the Foundation Biographics Laboratory 3R, Switzerland, are gratefully acknowledged.

References

(1) Bertrand, C. P.; Donath, P. D. CCR3 blockade as a new therapy for asthma. *Expert Opin. Invest. Drugs* **2000**, *9*, 43−52.
(2) Baggiolini, M. Chemokine and leucocyte traffic. *Nature* **1998**, *392*, 565−568.
(3) Locati, M.; Murphy, P. M. Chemokines and chemokine receptors: Biology and clinical relevance in inflammation and AIDS. *Annu. Rev. Med.* **1999**, *50*, 425−440.
(4) 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.
(5) Zlotnik, A.; Morales, J.; Hedrick, J. A. Recent advances in chemokine and chemokine receptors. *Crit. Rev. Immunol.* **1999**, *19*, 1−47.
(6) Saunders, J.; Tarby, C. M. Opportunities for novel therapeutic agents acting at chemokine receptors. *Drug Discovery Today* **1999**, *4*, 80−92.
(7) Vaddi, K.; Keller, M.; Newton, R. C. *The Chemokine Facts Book*; Academic Press: San Diego, CA, 1997.
(8) Busse, W. W. Inflammation in asthma: The cornerstone of the disease and target of therapy. *J. Allergy Clin. Immunol.* **1998**, *102*, 18−24.
(9) Kubinyi, H. QSAR and 3-D QSAR in drug design. 1. Methodology. *Drug Discovery Today* **1997**, *2,* 457−467.
(10) Kubinyi, H. QSAR and 3-D QSAR in drug design. 2. Applications and problems. *Drug Discovery Today* **1997**, *2*, 538−546.
(11) Kubinyi, H.; Folkers, G.; Martin, Y. C. QSAR: Current state, scope, and limitations. *Perspect. Drug Discovery Des.* **1998**, *12*, 3−23.
(12) Snyder, J. P.; Rao, S. N.; Koehler, K. F.; Vedani, A. Pseudoreceptors. In *3D QSAR in Drug Design*; Kubinyi, H., Ed.; ESCOM Science Publishers: Leiden, The Netherlands, 1993; pp 336−354.
(13) Murray-Rust, P.; Glusker, J. P. Directional hydrogen bonding to sp^2 and sp^3 hybridized O atoms and its relevance to ligand-macromolecule interactions. *J. Am. Chem. Soc.* **1984**, *106*, 1018−1025.
(14) 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.
(15) Baker, E. N.; Hubbard, R. E. Hydrogen bonding in globular proteins. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.
(16) Sprague, P. W.; Hoffmann, R. Catalyst pharmacophore models and their utility as queries for searching 3D databases. In *Computer-Assisted Lead Finding and Optimization*; van de Waterbeemd, H., Testa, B., Folkers, G., Eds.; VCH: Weinheim, Germany, 1997; pp 223−240.
(17) Hopfinger, A. J.; Wang, S.; Tokarski, J. S.; Jin, B. Q.; Albuquerque, M.; Madhav, P. J.; Duraiswami, C. Construction of 3D-QSAR models using 4D-QSAR analysis formalism. *J. Am. Chem. Soc.* **1997**, *119*, 10509−10524.
(18) So, S. S.; Karplus, M. Three-dimensional quantitative structure−activity relationships from molecular similarity matrices and genetic neural networks. *J. Med. Chem.* **1997**, *40*, 4347−4359.
(19) Vedani, A.; McMasters, D. R.; Dobler, M. Multi-conformational ligand representation in 4D-QSAR: Reducing the bias associated with ligand alignment. *Quant. Struct.-Act. Relat*. **2000**, *19*, 149−161.
(20) Vedani, A.; Dobler, M. Multi-dimensional QSAR in drug research: Predicting binding affinities, toxicity, and pharmacokinetic parameters. In *Progress in Drug Research*; Jucker, E., Ed.; Birkhäuser: Basel, Boston, and Berlin, 2000; pp 105−135.
(21) 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.
(22) Streich, D.; Neuburger-Zehnder, M.; Vedani, A. Induced fit−The key for understanding LSD activity. A 4D-QSAR study on the 5-HT_{2A} receptor system. *Quant. Struct.-Act. Relat*. **2000**, *19*, 565−573.
(23) Vedani, A.; Dobler, M.; Zbinden, P. Quasi-atomistic receptor surface models: A bridge between 3-D QSAR and receptor modeling. *J. Am. Chem. Soc.* **1998**, *120*, 4471−4477.
(24) Dobler, M.; Lill, M. A.; Vedani, A. From crystal structures and their analysis to the in silico prediction of toxic phenomena. *Helv. Chim. Acta* **2003**, *86*, 1554−1568.
(25) Vedani, A.; Dobler, M. 5D-QSAR: The key for simulating induced fit? *J. Med. Chem.* **2002**, *45*, 2139−2149.
(26) Vedani, A.; Dobler, M. Multi-dimensional QSAR: Moving from three- to five-dimensional concepts. *Quant. Struct.-Act. Relat*. **2002**, *21*, 382−390.
(27) DeLucca, G. V.; Kim, U. I.; Johnson, C.; Vargo, B. J.; Welch, P. K.; Covington, M.; Davies, P.; Solomon, K.; Newton, R. C.; Trainor, G. L.; Devicco, C. P.; So, 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.
(28) Boehringer Ingelheim KG. German patent DE 19520499, March 21, 1996.
(29) The CCR3 receptor binding test is based on a K562 cell line (leukemia myelogenic blast cells) transfected with the human chemokine receptor CCR3 (hCCR3-C1). The CCR3 receptor-binding assay with the radioligand ^{125}I-eotaxin-1 was performed in a scintillation proximity assay design. Included were controls for specific binding (no displacer added) and nonspecific binding by adding unlabeled eotaxin or a test compound. Bound radioactivity was determined by scintillation counter. Determination of affinity of test compounds (dissociation constant K_i) was calculated by iterative fitting of experimental data using the law-of-mass action based program EasySys (Schittkowski. *Num. Math*. **1994**, *68*, 129−142).
(30) Hahn, M. Receptor surface models. 1. Definition and construction. *J. Med. Chem.* **1995**, *38*, 2080−2090.
(31) Walters, D. E.; Hinds, R. M. Genetically evolved receptor models: A computational approach to construction of receptor models. *J. Med. Chem.* **1994**, *37*, 2527−2536.
(32) Hahn, M.; Rogers, D. Receptor surface models. 2. Application to quantitative structure−activity studies. *J. Med. Chem.* **1995**, *38*, 2091−2102.
(33) Kurogi, Y.; Guner, O. F. Pharmacophore modeling and three-dimensional database searching for drug design using "Catalyst". *Curr. Med. Chem*. **2001**, *8*, 1035−1055.
(34) Pastor, M.; Cruciani, G.; McLay, I.; Picket, S.; Clementi, S. Grid-independent descriptors (GRIND): A novel class of alignment-independent three-dimensional descriptors. *J. Med. Chem*. **2000**, *43*, 3233−3243.
(35) http://www.biograf.ch/PDFS/Quasar.pdf
(36) The actual field exerted by all atoms of a given ligand molecule onto its surface is determined using a nondirectional force field as, at this point, no atomic properties have been deposited on the receptor surface.
(37) Rogers, D.; Hopfinger, A. J. Genetic function approximation to generate a family of QSAR equations using genetic algorithms. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 854−866.
(38) Blaney, J. M.; Weiner, P. K.; Dearing, A.; Kollman, P. A.; Jorgensen, E. C.; Oatley, S. J.; Burridge, J. M.; Blake, J. F. Molecular mechanics simulation of protein−ligand interactions: Binding of thyroid analogues to prealbumin. *J. Am. Chem. Soc.* **1982**, *104*, 6424−6434.
(39) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(40) Searle, M. S.; Williams, D. H. The cost of conformational order: Entropy changes in molecular associations. *J. Am. Chem. Soc.* **1992**, *114*, 10690−10697.
(41) Vedani, A.; Zbinden, P.; Snyder, J. P.; Greenidge, P. A. Pseudoreceptor modeling: The construction of three-dimensional receptor surrogates. *J. Am. Chem. Soc.* **1995**, *117*, 4987−4994.
(42) Vedani, A.; Huhta, D. W. A new force field for modeling metalloproteins. *J. Am. Chem. Soc.* **1990**, *112*, 4759−4767.
(43) As a virtual particle (VP) in a quasi-atomistic approach has no bonding partners (i.e., unlike functional groups of real molecules it bears no lone-pairs), we apply a reduced function to determine the nonelectrostatic contribution to the H-bond energy involving a VP: For the constellation Don-H···VP, we correct for nonlin-

earity of the Don-H···VP angle (compulsory assuming a perfect directionality at the VP). For the arrangement Acc···VP, we correct for the deviation of the virtual hydrogen bond from the closest lone pair at the acceptor fragment (angle LP−Acc···VP) and assume a perfect linearity of the hydrogen bond. Derivation and calibration of the directional function for H-bond interactions is described in ref 42.

(44) The latter approach searches for an energy minimum while moving along the field lines, —thus, corresponding to a trajectory minimization.

(45) Furet, P.; Sele, A.; Cohen, N. C. 3D molecular lipophilicity potential profiles: A new tool in molecular modeling. *J. Mol. Graphics* **1988**, *6*, 182−189.

(46) Ghose, A. K.; Crippen, G. M. Atomic physicochemical parameters for three-dimensional structure-directed structure−activity relationships I: Partition coefficients as a measure of hydrophobicity. *J. Comput. Chem.* **1986**, *7*, 565−577.

(47) While simulating the evolution, the crossover procedure may be sped up by reducing the number of evaluated induced-fit models, depending on their actual frequency; that is, a model with low frequency is selected fewer times than a model with high frequency. When the speed factor is set to 0.0, all models are fully evaluated at each crossover step; a speed factor of 1.0 would represent a minimal model evaluation. The default is set at 0.9 implying that an induced-fit model "extinct" at a given time is still tested at a 10% frequency.

(48) 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.

(49) 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.

(50) Stewart, J. J. P. MOPAC—A semi-empirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−105.

(51) As in the Quasar concept, the receptor surrogate is represented by a family of models;[25] each of these models can independently select the most predictive induced-fit model. When generating the initial population, each model is analyzed for all selected induced-fit models. When a Boltzmann criterion is used, the model with the lowest lack-of-fit is accepted.

(52) Cheng, H. C. The power issue: Determination of K_B or K_i from IC_{50}. A closer look at the Cheng−Prusoff equation, the Schild plot and related power equations. *J. Pharmacol. Toxicol. Methods* **2001**, *46*, 61−71.

(53) Dairaghi, D. J.; Oldham, E. R.; Bacon, K. B.; Schall, T. J. Chemokine receptor CCR3 function is highly dependent on local pH and ionic strength. *J. Biol. Chem.* **1997**, *272*, 28206−28209.

JM040827U