Combining 4D Pharmacophore Generation and Multidimensional QSAR: Modeling Ligand Binding to the Bradykinin B_2 Receptor

Markus A. Lill* and Angelo Vedani

Biographics Laboratory 3R, Friedensgasse 35, 4056 Basel, Switzerland, and Institute of Molecular Pharmacy, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland

Received May 12, 2006

We recently reported the development of two receptor-modeling concepts (software Quasar and Raptor) based on multidimensional quantitative structure−activity relationships (QSAR) and allowing for the explicit simulation of induced fit. As the identification of the bioactive configuration of ligand molecules in such studies is all but unambiguous, each compound may be represented by an ensemble of different conformations, orientations, stereoisomers, and protonation states, leading to a 4D data set. In this account, we present a novel technology (software Symposar) allowed to automatically generate a 4D pharmacophore as input for multidimensional QSAR. Symposar aligns ligands utilizing fuzzylike 2D-subfeature mapping and, subsequently, a Monte Carlo search on a 3D similarity grid. The two-step concept (4D pharmacophore generation and quantification of ligand binding by multidimensional QSAR) was applied to 186 compounds binding to the bradykinin B2 receptor. The prediction of their binding affinity by means of the Quasar and Raptor technologies allowed for consensus scoring and generated topologically and quantitatively consistent receptor models. These converged at a cross-validated r^2 of 0.752 and 0.815 and yielded a predictive r^2 of 0.784 and 0.853 for a test set (for Quasar and Raptor, respectively).

INTRODUCTION

Nowadays, the structures of 36 000 macromolecules are deposited with the Protein Data Bank—still, for many proteins of biomedical interest, the three-dimensional structure is not available at atomic resolution. A wealth of quantitative structure−activity relationships (QSAR) demonstrated that these techniques are efficient in predicting quantities such as the binding affinity, the toxic potential, and pharmacokinetic parameters of existing or hypothetical molecules. The idea behind QSAR is that structural features can be correlated with biological activity. Of particular interest for biomedical research are QSAR based on three-dimensional models (3D-QSAR).[1−4] They generate a rational model of the target protein and allow for the quantification of the mutual interactions, including electrostatic forces, hydrogen bonds, and hydrophobic contacts, the forces known to play a key role for both substrate recognition and specificity. In contrast to the true biological receptor, where the binding site is characterized by a 3D arrangement of amino acids, 3D-QSAR models typically represent this entity by mapping physicochemical properties onto a surface or a grid surrounding the ligand molecules, superimposed in 3D space according to a pharmacophore hypothesis.

A critical issue of classical 3D-QSAR techniques is the identification of the bioactive conformation and the alignment of the ligand molecules. Particularly, in the absence of the 3D structure information of the target protein, the identification of the bioactive configuration of the ligand is all but unambiguous. 4D-QSAR concepts[5,6] incorporate molecular and spatial variety by representing each molecule in different conformations, orientations, stereoisomers, or protonation states (4D data set). The underlying algorithm then selects the combination of conformations most consistent with the structure−activity relationship. The identified conformation of each molecule and the resulting alignment are then assumed to best represent the bioactive conformation.

Various methods for creating 3D molecular alignments have been developed.[7−17] These mainly differ in their treatment of conformational flexibility (e.g., rigid versus flexible alignment) and the definitions of molecular similarity (e.g., functional-group similarity and the similarity of various physicochemical fields). In this account, we present a novel technique (software *Symposar*) aiming to create a consistent data set, that is, with a high similarity of physicochemical properties in 3D space among the ligands but allowing for alternative configurations of the latter (4D data set).The ligand molecules are superimposed onto one or several template molecules, first, on the basis of fuzzylike 2D substructure similarities and, subsequently, in 3D space with respect to their similarity of physicochemical fields. This two-step process combines the speed of a 2D similarity search with the accuracy and authenticity of protein−ligand interactions in 3D space. The molecules are thereby treated as flexible and are fully relaxed at the end of the alignment process.

METHODS

A. 4D Pharmacophore Generation. Scheme 1 displays the general procedure of the two-step alignment process combining 2D- and 3D-similarity searches. The molecule to be aligned (hereafter referred to as the *query* molecule), possibly represented by different stereoisomers and protonation states, is first aligned in 2D space on one or more

* Corresponding author phone: +41-61-2614259; fax: +41-61-2614258; e-mail: markus@biograf.ch and markus.lill@unibas.ch.

10.1021/ci6001944 CCC: $33.50
Published on Web 08/18/2006

Scheme 1. Two-Step Alignment Procedure as Implemented in Symposar

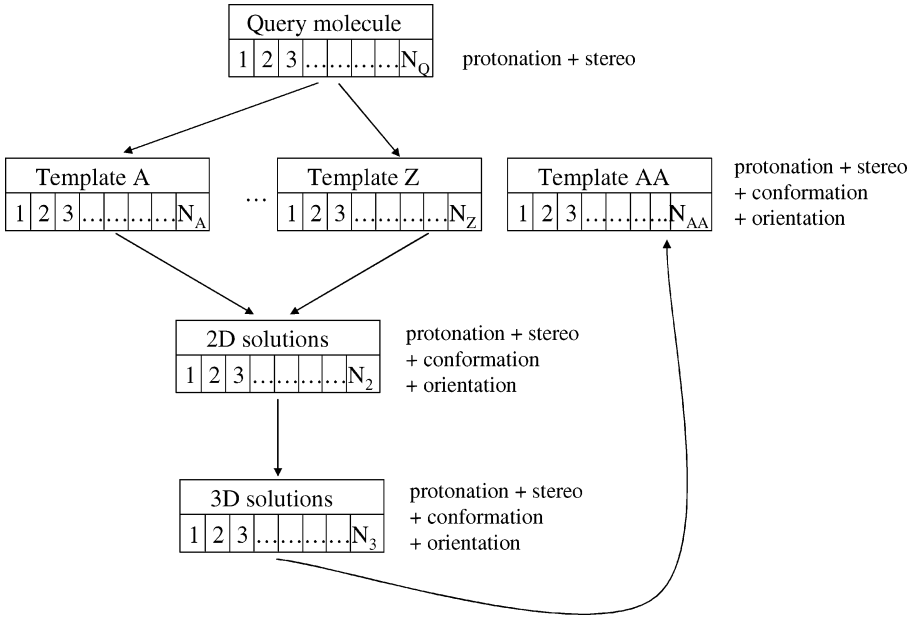


Table 1. Criteria for Matching Pharmacophore Features in the First-Level Substructure-Finding Process (Match, 1; No Match, 0)

	Ring	RingAcc	RingDon	RingD/A	Any	HBAcc	HBDon	HBD/A
Ring	1	1	1	1	0	0	0	0
RingAcc	1	1	0	1	0	1	0	1
RingDon	1	0	1	1	0	0	1	1
RingD/A	1	1	1	1	0	1	1	1
Any	0	0	0	0	1	0	0	0
HBAcc	0	1	0	1	0	1	0	1
HBDon	0	0	1	1	0	0	1	1
HBD/A	0	1	1	1	0	1	1	1

template molecules. Each template may be represented by different protonation states, stereoisomers, conformations, and orientations, which accounts for uncertainties in identifying the binding mode of the template compounds. After aligning the query molecule in 3D space onto the structure of the template on the basis of common-substructure criteria, it is minimized. On the basis of the refined solutions, a Monte Carlo search with subsequent minimization is performed to identify optimal 3D matches between the pharmacophore features of query and template molecules. Optionally, the best 3D solution (maximum similarity to the best-matching template) may be used as template for aligning the query molecules to follow. If this option is selected, the molecules are aligned in order of their experimental binding affinity to the target protein.

A.a. 2D Subfeature Mapping. A query molecule is prealigned to a template molecule in 2D space with respect to common substructure elements. Common substructures are sets of atoms that query molecule Q and template molecule T share. These substructures can be expressed as cliques ("maximal completely connected subgraphs") of non-hydrogen atoms, that is, a list of atoms with the following properties:

1. Corresponding atoms in Q and T must display the same associated physicochemical pharmacophore features. In our approach, we use features representing atoms that are members of rings (short notation: Ring); ring atoms which are hydrogen-bond acceptors (RingAcc), donors (RingDon), or both (RingD/A); atoms that are part of acyclic structures and are hydrogen-bond acceptors (HBAcc), hydrogen-bond donors (HBDon), or both, for example, hydroxyl groups (HBD/A); and such with no hydrogen-bond functionality (Any). Table 1 shows the criteria that two pharmacophore elements match.

2. The topological distances between the atoms in Q must be the same as the distances between the corresponding atoms in T. The topological distance between atoms is the distance in bonds along the shortest path connecting the atoms.

We used a modified clique-finding algorithm[18,19] (Chart 1 and Figure 1) to identify such common features of query and template molecules. The algorithm starts with all possible pairings of features F_Q^0 of the query molecule (e.g., molecule **1** in Figure 1) and features F_T^0 of the template molecule (e.g., molecule **2** in Figure 1) (Chart 1, line 1). If both features (2a) and their possible stereochemistry (2b) match (Table 1), these features are stored as the first pair in the clique (5) and are removed (4) from the list of available features in the clique-finding procedure (3). For example, in Figure 1, feature **A** for molecule **1** and feature **A′** in molecule **2** may form the first pair in the clique. The next pairs of features that can form a clique with the previous pair (9a) are identified among the features bonded to the first pair in the clique (9b). They are added to the clique (13), if the properties of the features (10a), their possible stereochemistry (10b), and the distances to the previously assigned pairs in the clique (10c) match. If there are different matching possibilities at this step, the number of matches of the next layer (i.e., atoms that are connected to the actual pair under consideration) are taken into account (11). In our example, (**B1**, **B1′**) and (**B2**, **B2′**) or (**B1**, **B2′**) and (**B2**, **B1′**) would

Chart 1. Algorithm to Identify Common Substructures between Two Small Molecules (First Level)

for (all pairs (F_Q^0, F_T^0))	1
if ($P(F_Q^0, F_T^0) = 1 \wedge$	2a
$S(F_Q^0, F_T^0) = 1$)	2b
$M_Q(\{F_Q^0\}) = 0$, $M_T(\{F_T^0\}) = 0$	3
$M_Q(F_Q^0) = 1$, $M_T(F_T^0) = 1$	4
$R_Q(1) = F_Q^0$, $R_T(1) = F_T^0$	5
number_pairs = 1	6
$F_Q^0 \rightarrow F_Q^P$, $F_T^0 \rightarrow F_T^P$	7
for (all pairs (F_Q^P, F_T^P))	8
for (all pairs (F_Q, F_T) with	9a
F_Q bonded to $F_Q^P \wedge F_T$ bonded to $F_T^P \wedge$	9b
$M_Q(F_Q) = 0 \wedge M_T(F_T) = 0$)	9c
if ($P(F_Q, F_T) = 1 \wedge$	10a
$S(F_Q, F_T) = 1 \wedge$	10b
topological distance of F_Q to	10c
all features $R_Q(.)$ matches	
distances of F_T to $R_T(.)$)	
if (more than one matching combinations)	11a
check for best matching neighbors	11b
$M_Q(F_Q) = 1$, $M_T(F_T) = 1$	12
R_Q(number_pairs) = F_Q, R_T(number_pairs) = F_T	13
number_pairs +1	14
$F_Q \rightarrow F_Q^P$, $F_T \rightarrow F_T^P$	15
$\rightarrow 8$	16
check for quality and similarity to previous solutions $\rightarrow$ store in data set	17

represent possible pairs. The algorithm then chooses the pairs (**B1**, **B1′**) and (**B2**, **B2′**), because the next layer would yield three additional pairs (**C1**, **C1′**), (**C2**, **C2′**), and (**C3**, **C3′**) instead of two pairs, for example, (**C2**, **C3′**) and (**C3**, **C2′**) for (**B1**, **B2′**) and (**B2**, **B1′**).

After finding the matching pairs of features in this layer of connectivity, the next layer of features may be added to the clique (15 and 16), (**C1**, **C1′**), (**C2**, **C2′**), and (**C3**, **C3′**) in our example. Finally, the algorithm will identify the red portions in Figure 1 as the largest connected substructure common in molecules **1** and **2**. It will also store alternative common substructures in order of decreasing quality of the match that are dissimilar to the solutions with higher quality.

Common substructures might be interrupted by elements which do not match in both molecules (e.g., Figure 2). Releasing the condition that the features F_Q and F_T have to be bonded to F_Q^P and F_T^P (9b) would allow the algorithm to identify the additional common features (**d**, **d′**), (**e**, **e′**), (**f**, **f′**), (**g**, **g′**), (**h**, **h′**), and (**i**, **i′**) and the atoms of the terminal aromatic ring (**k**, **k′**). In addition, matching features can differ in their topological distances to the previously assigned common substructure but align reasonably in 3D space [Figure 1, (**b**, **c′**) and (**c**, **b′**)]. The protein can adapt its conformation to different molecules; for example, side chains may flip their hydrogen-bonding properties (Figure 3), which leads to a relative superposition of the hydrogen-bond donor and acceptor of the two molecules. Protein cavities will be occupied with groups of two molecules, which differ in their physicochemical properties. To be able to deal with these situations, we developed a fuzzylike feature-mapping algorithm (second-level substructure-finding) which allows for gaps of ± 1 in the topological connectivity of common features (Chart 2).

The search starts from the common substructures obtained from the first-level search (1′, 2′). For all possible remaining pairs (F_Q, F_T) (3′), the algorithm identifies the topological distance to the closest feature $F_Q^{first,near}$ and it's bonded atoms, previously assigned (first-level query) (4′). If this distance does not differ by more than one unit from the corresponding distance of F_T to $F_T^{first,near}$ and its neighbors (5′), the fuzzylike similarity *Simil* is calculated for the pair (F_Q, F_T) including its bonded atoms (6′), not yet assigned, making use of the similarity matrix in Table 2. The list of all possible pairs is sorted with respect to their group similarity Simil (8′). The group of features with the highest value of Simil (9′) is added to the common substructure (10′−13′), and the group similarity is updated using the features, not yet assigned, only (14′). Finally, the algorithm identifies the common substructures colored in red, green, and cyan in Figures 1−3 as common substructures.

A.b. Alignment in 3D Space. The alignment based on 2D similarities does not recognize different hybridization states or different conformational preferences of different chemical groups in the same hybridization state: Ether and thioether groups, for example, differ significantly in their conformation because of longer S−C bond lengths and a smaller angle subtended at the S atom; similarly, the conformation of ortho-substituted biphenyls depends on the size of the substituents. As the strength of protein−ligand interactions is determined by the configurations of the protein and ligand in 3D space, we added a similarity search in 3D space with a relaxation of the ligand conformation at its end (Chart 3).

The conformations of the query molecule as obtained by the 2D similarity search (1*) are minimized with harmonic restraints to the atoms of the identified common substructure (2*). Typically, the force constant of the restraint for the atoms of the first-level substructure is larger than those of the second-level substructure. The atoms not identified as common substructure elements are let free. Starting from the relaxed conformation, a Monte Carlo search is performed aiming to optimize the similarity in 3D space (3*). A new conformation is created by local translation of the center of mass of the molecule (5a*) and the rotation around it (5b*), as well as by local rotation and global flip (cis−trans flip for sp^2-sp^2 torsions, trans−gauche+−gauche− for sp^3-sp^3 torsions) of the accessible torsion bonds (5c*); during the first half of the Monte Carlo search, torsions sharing two central atoms in the common substructure are fixed. All other torsions are accessed during the Monte Carlo process step by step, starting from the torsion closest to the common substructure (4*). In Figure 4, for example, the red portion displays a common substructure. In the first $^1/_6$ (in the second half of the Monte Carlo search, all torsions are accessed $\rightarrow$ $^1/_2 \times {}^1/_3$: three levels of torsion not part of the substructure) of the Monte Carlo search, the values of torsion angle **1**, in the second $^1/_6$, those of levels **1** and **2**, and in the third, $^1/_6$, those of levels **1**, **2**, and **3** are modified.

To score the quality of a new conformation, the similarity of the atoms, which are part of the common 2D substructure or of torsions that are accessed at the current status of the Monte Carlo search, to the template molecules is calculated in 3D space (6*). Similar to FlexS,[7] we used common intermolecular interactions with a hypothetical protein as the similarity measure. The intermolecular interactions are of

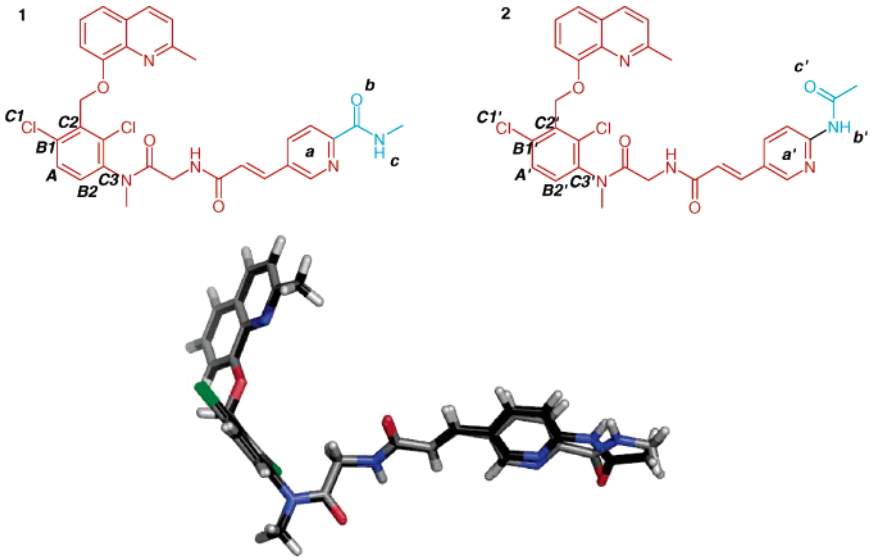


Figure 1. Two molecules binding to the bradykinin B_2 receptor [top panel, 2D representation with common pharmacophore features in red (first-level search) and cyan (second-level search); lower panel, 3D alignment as obtained with Symposar].

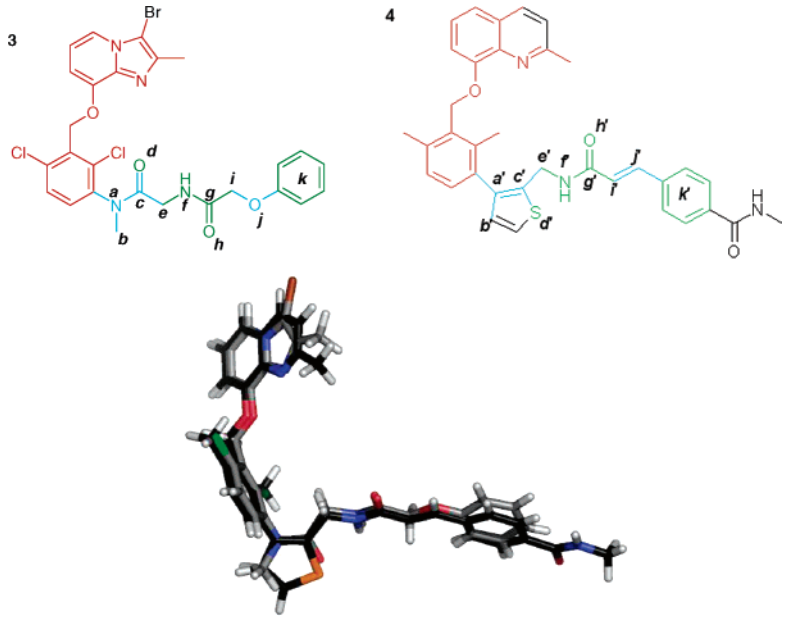


Figure 2. Two molecules binding to the bradykinin B_2 receptor [top panel, 2D representation with common pharmacophore features in red (first-level search) and cyan (second-level search); lower panel, 3D alignment as obtained with Symposar].

steric fit and hydrogen-bonding type. To calculate the similarity in steric fit of the query and template molecules, a Gaussian function $A \exp(-Ar^2)$, centered at the atomic center of non-hydrogen atoms and hydrogen atoms involved in hydrogen bonds, is used to distribute the steric properties of the template molecule on a rectangular grid. The parameter A, determining the height and width of the Gaussian, is inverse proportional to the square of the van der Waals radius of the atom and a scaling factor f_{scale}. Initially set to 1 f_{scale}, it decreased to 0.5 during the last $^1/_4$ of the Monte Carlo search, to tighten the alignment of similar functional groups. Additional Gaussians are placed on the end of the lone-pair vectors and hydrogen atoms of donating functional groups to include hydrogen-bond directionality in the similarity measure. The similarity of the query molecule is then calculated by summing up the physicochemical properties (steric fit and hydrogen-bond properties) of the template molecules on the grid points adjacent to the atomic positions of the query molecule (including linear interpolation of the grid points to the actual position of an atom of the query molecule). As a final score S of the quality of the alignment of the query to the template molecule, the internal energy of the query molecule is added to the similarity measure (6*). The conformation is accepted as new solution, if it is dissimilar to previous ligand configurations (of the same initial 2D alignment) or the score is better than those of similar solutions (7*). The same check is performed with respect to all initial 2D alignments after the Monte Carlo search has been terminated (8*). To obtain relaxed conformations, the best solutions are finally minimized using intramolecular properties as well as interactions with the physicochemical properties on the grid, which represents the hypothetical binding site of the target protein (9*). To avoid excessive computing time or jeopardizing the QSAR algorithm, the number of conformations comprising the 4D data set as input for multidimensional QSAR is preferably limited

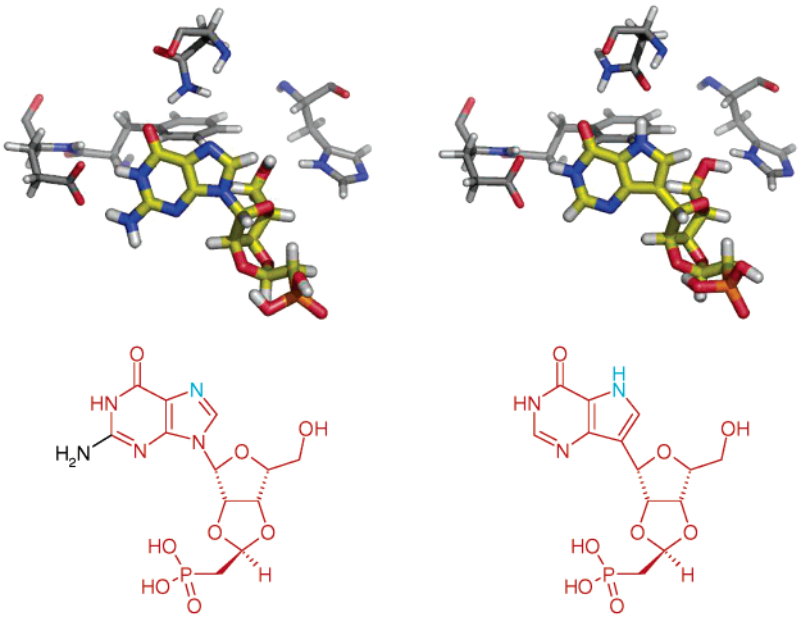


Figure 3. Two molecules binding to purine-nucleoside phoshorylase [top panel, minimized X-ray structures demonstrating the hydrogen-bond flip of Asn 235;[20] lower panel, 2D representation with common pharmacophore features in red (first-level search) and cyan (second-level search)].

Chart 2. Algorithm to Identify Fuzzylike Common Substructures between Two Small Molecules (Second Level)

for (best substructures first level $\{(F_Q, F_T)\}$)	1'
$F_Q \rightarrow F_Q^{first}$, $F_T \rightarrow F_T^{first}$	2'
for (all pairs $(F_Q, F_T) \notin \{(F_Q^{first}, F_T^{first})\}$)	3'
smallest topological distance to any $F_Q^{first} \rightarrow F_Q^{first, near}$	4'
if (topological distance of F_Q to $\{F_Q^{first, near}$ + neighbors$\}$ matches	5'
distance of F_T to $\{F_T^{first, near}$ + neighbors$\} \pm 1$)	
calculate similarity of substructure groups $\{(F_Q, F_T)\}^G$:	6'
$\text{Simil} = C(F_Q, F_T) + \sum_{\text{neighbors}} C(F_Q^{\text{neighbor}}, F_T^{\text{neighbor}})$	
else	
Simil = 0	7'
sort $\{(F_Q, F_T)\}^G$ with respect to Simil	8'
while ($\exists$ groups $\{(F_Q, F_T)\}^G \notin \{(F_Q^{first}, F_T^{first})\}$ with Simil > 0,	9a'
starting with highest Simil)	9b'
for(all features of group $\{(F_Q, F_T)\}^G$)	10'
$M_Q(F_Q) = 1$, $M_T(F_T) = 1$	11'
R_Q(number_pairs) = F_Q, R_T(number_pairs) = F_T	12'
number_pairs +1	13'
update Simil (2'-7')	14'
check for quality and similarity to previous solutions → store in data set	15'

to 4−16. In addition, a user-defined range determines how much the scoring value *S* of a ligand's conformation is allowed to deviate from the highest scored one to be included in the 4D data set. The scale of the 3D similarity measure and consequently of the scoring value *S* has been tuned to provide values in the order of typical energies associated with ligand−protein interactions.

B. Multidimensional QSAR. Quasar[6,21−23]—a receptor-modeling concept developed at our laboratory—is based on 6D-QSAR and explicitly allows for the simulation of induced fit. It generates a family of quasi-atomistic receptor-surface models that are optimized by means of a genetic algorithm. The hypothetical receptor site is characterized by a three-dimensional surface with atomistic properties mapped onto it: hydrophobic character, electrostatic potential, and hydrogen-bonding propensity.

The fourth dimension (→ 4D-QSAR) refers to the possibility of representing each ligand molecule as an ensemble of conformations, orientations, and protonation states, thereby reducing the bias in identifying the bioactive conformer. The underlying algorithm selects the combination of conformations most consistent with the structure−activity relationship: Within the ensemble of conformers, the contribution of an individual entity to the total energy is determined by a normalized Boltzmann weight

$$\Delta G = \frac{\sum_{i=1}^{N_{\text{Confomers}}} \Delta G_i \exp(-\Delta G_i / k_B T)}{\sum_{i=1}^{N_{\text{Confomers}}} \exp(-\Delta G_i / k_B T)} \quad (1)$$

where the energy of binding is determined as follows:

$$\Delta G = \Delta E_{\text{forcefield}} + \Delta E_{\text{polarization}} - \Delta G_{\text{desolvation}} - T\Delta S - \Delta E_{\text{internal}} - \Delta G_{\text{inducedfit}} \quad (2)$$

The selected conformation of each molecule and the resulting alignment are then assumed to represent the bioactive conformation.

The fifth dimension (→ 5D-QSAR) allows the simultaneous evaluation of different induced-fit protocols, while the sixth dimension (→ 6D-QSAR) allows the simultaneous consideration of different solvation models.

Raptor[24−26] explicitly and anisotropically allows for induced fit by a dual-shell representation of the receptor surrogate, mapped with physicochemical properties (hydrophobic character and hydrogen-bonding propensity) onto it. In Raptor, 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

Table 2. Similarity Score to Identify Common Pharmacophore Features in the Second-Level Substructure-Finding Process

	Ring	RingAcc	RingDon	RingD/A	Any	HBAcc	HBDon	HBD/A
Ring	0.5	0.5	0.5	0.5	0.3	0.3	0.3	0.3
RingAcc	0.5	3.0	1.0	3.0	0.3	3.0	0.5	3.0
RingDon	0.5	1.0	3.0	3.0	0.3	0.5	3.0	3.0
RingD/A	0.5	3.0	3.0	5.0	0.3	3.0	3.0	5.0
Any	0.3	0.3	0.3	0.3	0.5	0.3	0.3	0.3
HBAcc	0.3	3.0	0.5	3.0	0.3	3.0	0.5	3.0
HBDon	0.3	0.5	3.0	3.0	0.3	0.5	3.0	3.0
HBD/A	0.3	3.0	3.0	5.0	0.3	3.0	3.0	5.0

Chart 3. Algorithm to Identify Most Similar Solutions of Query and Template Ligand in 3D Space

```
for (best 2D solution {(F_Q, F_T)})                                   1*
    Minimize with releasing restraints to substructure template       2*
    (restraints on first-level substructures larger than on second-level ones)
    for (number of Monte-Carlo steps)                                 3*
        change zone of interest                                       4*
        local translation                                             5a*
        local rotation                                                5b*
        local torsional rotation & torsion flips                      5c*
        calculate score S = similarity + internal energy in zone      6*
        check for quality S and similarity to previous solutions      7*
        → store in data set
    check for quality and similarity to previous solutions            8*
    → store in data set
Minimize in similarity grid                                           9*
```

ligand−receptor interactions includes directional terms for hydrogen bonding and hydrophobicity and thereby treats solvation effects implicitly.

$$\Delta G = \Delta G_{\text{H-bond}} + \Delta G_{\text{hydrophobic}} - T\Delta S - \Delta G_{\text{inducedfit}} \quad (3)$$

This makes the approach independent from a partial-charge model and, as a consequence, allows smooth modeling of ligand molecules binding to the receptor with different net charges.

RESULTS AND DISCUSSION

Evaluation of Symposar. To evaluate the Symposar concept, we tested its accuracy in predicting binding modes of known ligands. For this purpose, we attempted to reproduce the binding geometry of various ligands with different sizes and flexibility, as obtained by X-ray crystallography. We employed the validation set used for evaluating FlexS[27] and compared our results with those obtained with FlexS.[7] We analyzed a set of 13 proteins (the data for thrombin in ref 21 differs from that in ref 7) for which several protein−ligand complexes have been determined crystallographically. Lemmen et al.[7] superimposed these complexes by minimizing the positional differences between the backbone C_α atoms. The reference alignment of two ligands is then obtained by extracting the ligands from the complexes in their conformation and relative orientation in space.

As in the validation studies on FlexS,[7] we superimposed those pairs of ligands where the volume portion of the test ligand intersecting with the reference ligand had to be at least 60%, and the number of non-hydrogen atoms in the test ligand located outside the intersection volume is limited to 10. Each reference ligand was taken in its crystallographically determined binding conformation. The computed alignment of the test ligand was compared to the ligand configuration in its reference alignment. As a measure for the goodness of fit, we used the root-mean-square deviation (RMSD) in atomic coordinates excluding all hydrogen atoms. The ligands were divided into three classes. The first class (I) contained molecules with fairly rigid skeletons, including steroid-type and carbohydrate-type ligands, binding to an immunoglobulin, to streptavidin, trypsin, glycogen phosphorylase, and concanavalin. The second class (II) comprised ligands of considerable size, containing a large variety of organic building blocks besides peptidic portions, including compounds binding to dihydrofolate reductase, human rhinovirus, and fructose bisphosphatase. The third class (III) contained large and very flexible peptidelike molecules binding to endothiapepsin, HIV protease, elastase, thermolysin, and carboxypeptidase. Lemmen et al.[7] associated a RMSD < 1.5 Å with a sufficiently precise reproduction of the experimental situation. A deviation between 1.5 and 2.0 Å indicates that the orientation and conformation of a molecule is overall correctly reproduced. With a RMSD > 2.0 Å, major differences occur, at least in some portions of the ligand.

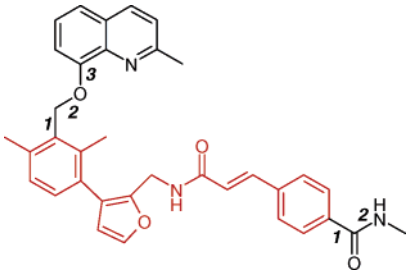


Figure 4. Step-by-step releasing of torsion degrees of freedom during the Monte Carlo search. The query molecule substructure, which is common to the template, is colored in red.

Similar to FlexS, Symposar yields multiple solutions. To use the ensemble of these solutions as 4D input to multidimensional QSAR, their size should practically be limited to 4−16 entities. Thus, it is central that the computed ligand configurations, which show the best agreement with the experimental results, are ranked among the best solutions of the alignment process. In the statistical analysis of the alignment fit, Lemmen et al.[7] limited themselves to the 10 best solutions. In our validation study, we defined a maximum of 10 solutions in Symposar. The statistical analysis of the entire set of 272 alignment pairs (Table 3 and Figure 5a) shows that, in 56% (45% for FlexS) of all cases, solutions are identified within a RMSD < 1.5 Å and, in 67% (58%), within a RMSD < 2.0 Å. Analyzing the results of the three different subsets (Table 3 and Figure 5b) shows that Symposar's slightly superior behavior is mainly

Table 3. Individual Statistics on Symposar and FlexS Performance in Reproducing the Binding Geometry of Various Protein-Bound Ligands

protein	number of pairs	probability for placement with RMSD < 0.5 Å		RMSD < 1.0 Å		RMSD < 1.5 Å		RMSD < 2.0 Å	
		Symposar	FlexS	Symposar	FlexS	Symposar	FlexS	Symposar	FlexS
immunoglobulin	16	0.00	0.06	0.38	0.06	0.63	0.31	0.94	0.69
streptavidin	20	0.00	0.00	0.55	0.65	1.00	1.00	1.00	1.00
trypsin	27	0.07	0.15	0.59	0.44	0.78	0.78	0.93	0.96
glycogen phosphorylase	6	0.17	0.00	1.00	1.00	1.00	1.00	1.00	1.00
concanavalin	2	0.50	0.00	1.00	1.00	1.00	1.00	1.00	1.00
class I	71	0.06	0.07	0.58	0.48	0.83	0.76	0.96	0.92
dihydrofolate reductase	2	0.00	0.00	0.00	0.00	0.50	0.50	1.00	1.00
human rhinovirus	56	0.13	0.04	0.34	0.36	0.45	0.38	0.55	0.43
fructose bisphosphatase	2	0.00	0.00	0.00	0.50	1.00	0.50	1.00	1.00
class II	60	0.12	0.03	0.32	0.35	0.47	0.38	0.58	0.47
endothiapepsin	7	0.00	0.00	0.00	0.00	0.43	0.43	1.00	0.57
HIV protease	41	0.00	0.00	0.07	0.00	0.27	0.05	0.39	0.24
elastase	14	0.00	0.00	0.36	0.14	0.50	0.43	0.57	0.43
thermolysin	65	0.03	0.00	0.26	0.22	0.45	0.32	0.52	0.48
carboxypeptidase	14	0.00	0.00	0.71	0.50	1.00	0.93	1.00	1.00
class III	141	0.01	0.00	0.25	0.16	0.45	0.32	0.56	0.46
all	272	0.05	0.03	0.35	0.29	0.56	0.45	0.67	0.58

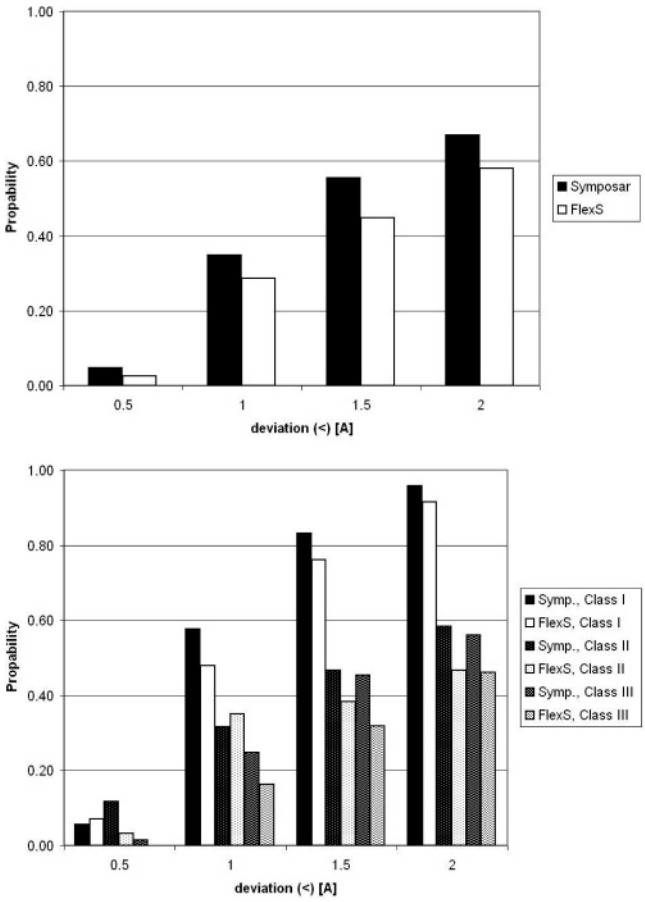


Figure 5. Histogram analysis of Symposar and FlexS results to reproduce the binding geometry of protein-bound ligands: (a) overall agreement with experiment and (b) each class separately.

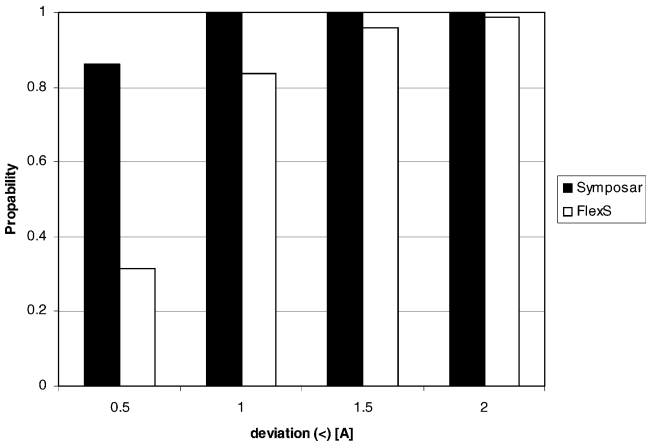


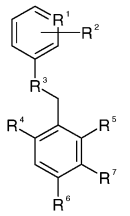
Figure 6. Histogram analysis of Symposar and FlexS results of self-fit validation studies.

pronounced for the larger ligands in classes II (47% vs 38% for RMSD < 1.5 Å and 58% vs 47% for RMSD < 2.0 Å) and III (45% vs 32% for RMSD < 1.5 Å and 56% vs 46% for RMSD < 2.0 Å), while both approaches provide reasonable alignments for class I ligands (83% vs 76% for RMSD < 1.5 Å and 96% vs 92% for RMSD < 2.0 Å). The analysis further indicates some limitations of our approach. Similar to FlexS, it typically fails to predict inverse or alternative binding modes of structurally related compounds. This is also the main reason for the low probability for placements agreeing with the experimental data for HIV protease, elastase, thermolysin, and human rhinovirus. In all other systems, Symposar produces alignments with a probability > 90% to be in agreement with the experimental data.

The statistical analyses so far (Table 3 and Figure 5) do not include the data of self-fit results, that is, the alignment of a molecule to a structurally identical template. Figure 6 shows the results of all 73 self-fits. While FlexS reaches probabilities of 30% for a RMSD < 0.5 Å, 84% for a RMSD < 1.0 Å, and 96% for a RMSD < 1.5 Å, Symposar produces in 86% of all cases alignments within a RMSD < 0.5 Å and a RMSD < 1.0 Å for all 73 self-fits. This superior behavior of Symposar is mainly due to the primary 2D alignment process and may become important for series of compounds with structurally similar backbones or even congeneric series of molecules respecting the same binding mode. The RMSD > 0.5 Å for 14% of all cases is caused by the difference in force-field parametrizations between Symposar (using the Yeti force field[28]) and that used throughout the refinement process of X-ray crystallography.

Combination with Multidimensional QSAR. We applied our 4D pharmacophore-generation technology Symposar to

Chart 4. Ligands Binding to the Bradykinin B_2 Receptor. **D50a**, **C90c**, **E17b**, and **G22b** Were Used as Templates for Alignment with Symposar



R^1: >N–, =N–, >C–, =CH–

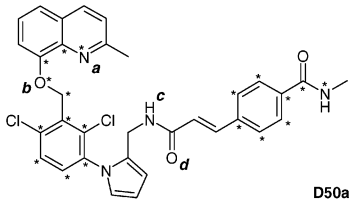
R^2: attached ring with additional functional groups (typically fused 5- or 6-membered ring)

R^3: mainly –O–, but also –NH–

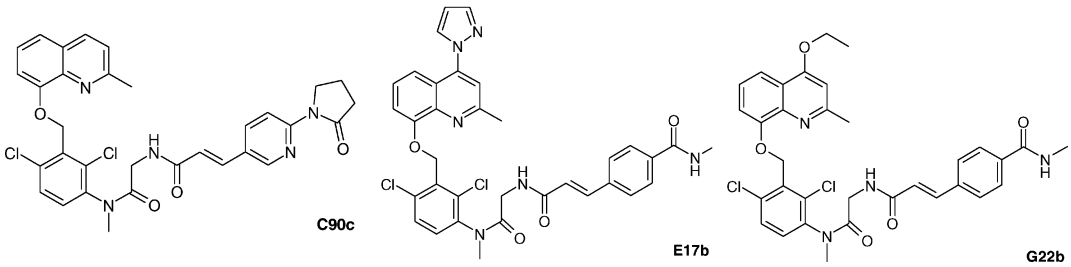
R^4/R^5: –H, –Cl, –CH_3, –OCH_3

R^6: –H or –Cl

R^7: miscellaneous



D50a



C90c E17b G22b

a large series of molecules binding to the bradykinin B_2 receptor. The endogenous agonist, bradykinin, is a 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 high proinflammatory activity, bradykinin receptors have been implicated in a variety of pathophysiological responses, including pain, inflammation, asthma, rhinitis, and hypotension.[29] Two types of receptors, referred to as B_1 and B_2 subtypes, have been identified by molecular cloning and pharmacological means.[30] B_2 receptors are expressed constitutively in many tissues and are thought to mediate most of the biological actions of bradykinin receptors. Because of the pathophysiological role of bradykinin, antagonists of bradykinin receptors are of biomedical interest.

Our simulation was based on 186 ligands (Chart 4 and Supporting Information) synthesized and biologically assessed during the development of highly affine antagonists at Fujisawa Pharmaceuticals.[31] The three-dimensional structures of all ligand molecules were generated using MacroModel 6.5[32] and optimized in aqueous solution on the basis of the AMBER*[33] force field. Next, CM-1 atomic partial charges were generated using the AMSOL package.[34] To explore the configurational space, we performed a conformational search for **D50a** using MacroModel 6.5 and selected a conformation that provides access to the functional groups **a**, **b**, **c**, and **d** by hydrogen-bonding groups of the protein suspected to line the binding pocket. Analyzing the SAR qualitatively, groups **a**−**d** would seem to be relevant for interacting with the protein. The energy of this conformation is 5.0 kcal/mol above that of the energetically lowest ligand configuration. All energetically lower conformations represented back-folded structures with intramolecular interactions among the various hydrophobic portions (cf. Chart 4). This would obscure the access to one or more of the hydrogen-bond donating or accepting groups **a**, **b**, **c**, and **d**. We then aligned three affine ligands (**C90c**, **E17b**, and **G22b**) with additional functional groups onto this template using Symposar. These four molecules served as the secondary template for superimposing all of the other 182 compounds (Figure 7), resulting in a total of 269 representations (→ 4D data set) of the 186 ligands.

On the basis of this alignment, we performed multidimensional QSAR simulations using Quasar[6,21−23] and Raptor[24−26] using a consensus-scoring protocol. A total of 147 of the 186 molecules were composed into the training set; the remaining 39 ligands defined the test set. In the Raptor simulation, a family of 10 receptor models was created, resulting in a cross-validated r^2 of 0.815 for the training set and a predictive r^2 of 0.853 for the test set (Figure 8a). The maximal deviation of any ligand from the experimental value is a factor 11.7 in the IC_{50} for the training set and 14.9 for the test set.

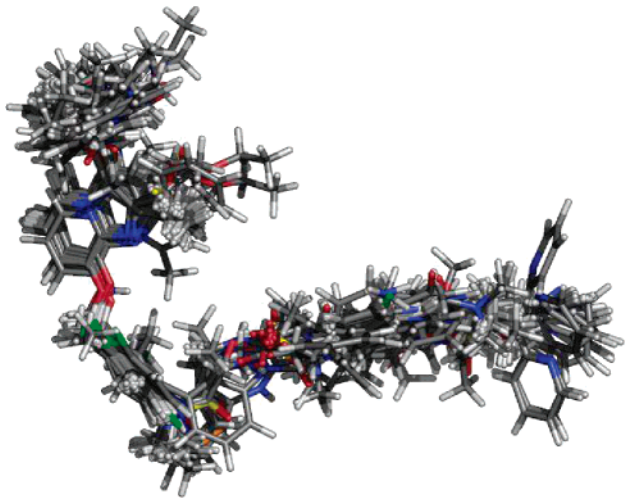


Figure 7. Superposition of the molecules binding to the bradykinin B_2 receptor as obtained with Symposar.

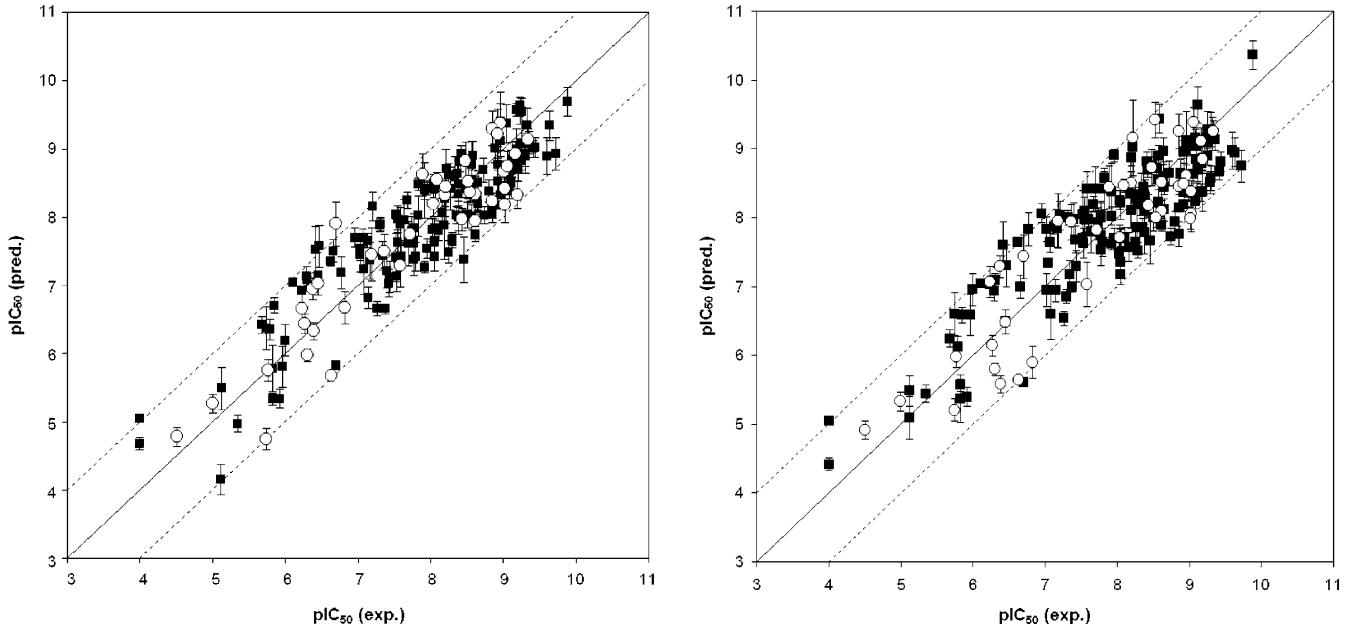


Figure 8. Comparison of experimental and predicted pIC_{50} values for the bradykinin B_2 receptor as generated by the (a) Quasar and (b) Raptor technologies, respectively (filled squares, training set; open circle, test set).

The physicochemical properties—in common at the 70% level for the 10 models—are shown in Figure 9a with the most affine ligand **C90c**. The portions of the receptor model enclosing the quinoline group and the chloro-substituted benzene are predominantly hydrophobic in nature (color: beige-brown). Additional hydrophobic portions are observed around the 6-(2-oxo-pyrrolidin)pyridin group. Around the central amide bond, a hydrogen-bond-acceptor region is observed (red), and hydrogen-bond-donor domains (blue) around the oxygen atoms of the *N*-methylated amide and the benzyloxy group and the nitrogen atoms of the pyridin and quinoline are observed.

In the Quasar simulation, a family of 250 receptor models was created, resulting in a cross-validated r^2 of 0.752 for the training set and a predictive r^2 of 0.784 for the test set (Figure 8b). The maximal deviation of any ligand from the experimental value is a factor 14.1 in the IC_{50} for the training set and 9.6 for the test set. The mean receptor model (Figure 9b) is predominantly hydrophobic in nature (colors: brown and gray). Similar to Raptor, the model displays a hydrogen-bond-acceptor region (red) at the central amide bond, hydrogen-bond donor domains (blue) around the oxygen atoms of the methylated amide and the benzyl-oxy group, and the nitrogen atoms of the pyridin and quinoline. Thus, Symposar yielded a 4D data set that produced topologically and quantitatively consistent results within the Raptor and Quasar technologies, although both approaches are based on different concepts including alternative scoring functions.

Varying the parameters in the QSAR simulations, for example, the number of models and length of optimization, resulted in variations less than 0.04 in the cross-validated r^2 and less than 0.03 in the predicted r^2 for the test set when compared to the original simulation. The most influential step in the procedure is the alignment. We shortened the three-step alignment protocol (selection of conformations for the first template **D50a**, building second template, and superimposing all other 182 molecules) to a two-step process (selection of the template and superimposing all other 185 molecules) using either **D50a**, **C90c**, **E17b**, or **G22b** as the template. The QSAR analyses based on these alignments yielded similar cross-validated r^2 values for the training set but smaller predictive r^2 values for the test set by about 0.07—

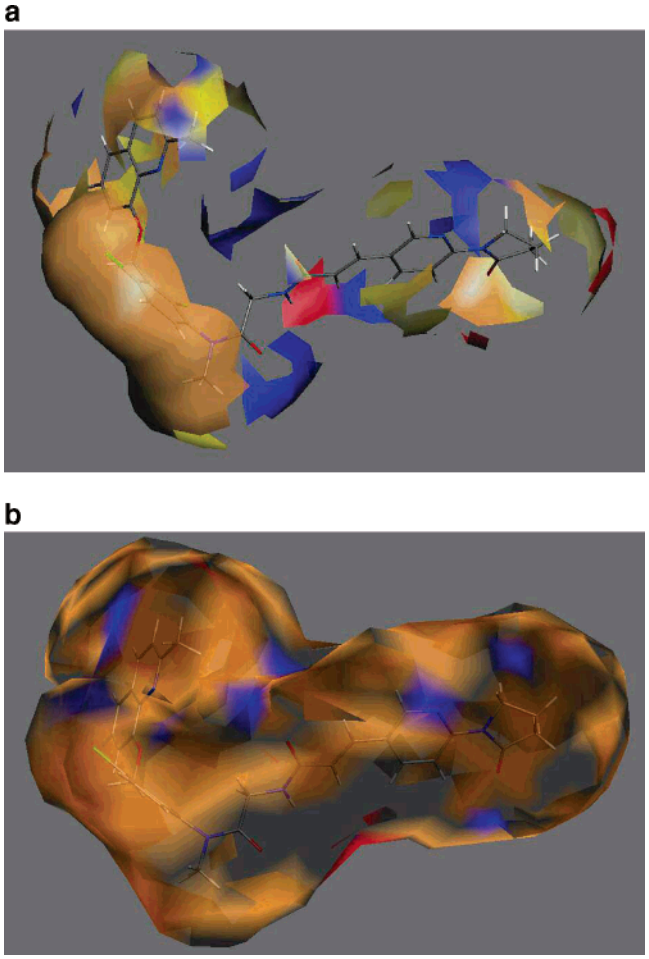


Figure 9. Receptor models for the bradykinin B_2 receptor obtained using (a) Raptor (beige-brown, hydrophobic properties; red, hydrogen-bond acceptor; and blue, hydrogen-bond donor) and (b) Quasar (gray, neutral and hydrophobic; brown, charged and hydrophobic; red, hydrogen-bond acceptor; and blue, hydrogen-bond donor).

0.12. The most probable reason might be that the functional groups, provided by the additional templates **C90c**, **E17b**, and **G22b** in the original approach, are less coherently aligned in the "one-molecule-as-template" attempt.

To establish the supremacy of the Symposar concept as an alignment tool prior to the QSAR analysis on the test, we compared the results based on the Symposar alignment with those obtained using a rigid superposition (using the Kabsch[35] algorithm). The rigid alignment was again based on the very template **D50a** in the previously described conformation that provides access to the functional groups **a**, **b**, **c**, and **d** by hydrogen-bonding groups of the protein suspected to line the binding pocket (cf. Chart 4). We then aligned three affine ligands (**C90c**, **E17b**, and **G22b**) onto this template, performing a rigid alignment by minimizing the geometric distances between defined atoms of **D50a** (see asterisk symbols in Chart 4) and corresponding atoms in **C90c**, **E17b**, and **G22b**. These four molecules served as the secondary template for rigidly superimposing the remaining 182 compounds. For the rigid alignment, we used the conformations obtained by minimizing the ligands with MacroModel 6.5 in aqueous solution on the basis of the AMBER* force field. On the basis of the alignment, we then performed QSAR simulations using Quasar and Raptor with identical settings as used previously with the Symposar alignment. The Raptor models resulted in a cross-validated r^2 of 0.592 for the training set and a predictive r^2 of 0.096 for the test set, while the Quasar even failed to yield a positive cross-validated r^2 for the training set. These statistically significant inferior results based on the rigid alignment suggest the robustness of the Symposar concept as shown independently for two different QSAR analysis tools.

CONCLUSION

We describe a novel technology that automatically superimposes ligand molecules binding to a common target protein combining 2D and 3D similarity criteria (software Symposar). It has been designed to generate a consistent 4D data set as input for multidimensional QSAR and aims to identify conformations and orientations most similar among the various ligands but allows for diversity in the ligand's configurations, respecting the inherent ambiguity throughout the identification of the bioactive form. Thus, it may be used as input for 3D QSAR but is predestinated for multidimensional QSAR, which allows for a representation of each compound by a whole ensemble of different conformations, orientations, stereoisomers, and protonation states (4D data sets). This rationale was demonstrated by applying Symposar in conjunction with the multidimensional QSAR techniques Quasar and Raptor (allowing for consensus scoring) to 186 compounds binding to the bradykinin B2 receptor.

In validation studies on the accuracy in predicting binding modes of protein-bound ligands, Symposar was demonstrated to provide decent agreement with the experimental results (RMSD < 2.0 Å) for 67% of all 272 cases. For eight of the 13 ligand sets, which do not display alternative or inverse binding modes for structurally similar compounds, the success rate even exceeds 90%. On the other hand, our studies also indicate a limitation of Symposar for identifying the correct binding mode: If two structurally similar compounds bind to the same target protein in different binding orientations or conformations, Symposar will infrequently be able to identify the bioactive form of the ligand. We therefore currently extend the algorithm to allow for the identification of inverse binding modes as additional input for multidimensional QSAR. Symposar also demonstrated excellent behavior for self-fit tests, proving its usefulness to align congeneric series of molecules, under the premise that similar compounds display similar binding modes.

ACKNOWLEDGMENT

Financial support from the Foundation Biographics Laboratory 3R Basel, Switzerland, is gratefully acknowledged.

Note Added after ASAP Publication. This article was released ASAP on August 18, 2006 with minor errors in molecule **4** of Figure 2. The correct version was posted on September 15, 2006.

Supporting Information Available: Experimental and predicted pIC_{50} values in tabular format for the 186 compounds binding to the bradykinin B_2 receptor used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Cramer, R. D., III.; Patterson, D. E.; 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.
(2) 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.
(3) 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.
(4) 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.
(5) Esposito, E. X.; Hopfinger, A. J.; Madura, J. D. In *Handbook of Chemoinformatics − From Data to Knowledge*; Gasteiger, J., Engel, T., Eds.; Wiley-VCH: Weinheim, Germany, 2003; Volume 4, pp 1576−1603.
(6) 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.
(7) Lemmen, C.; Lengauer, T.; Klebe, G. FLEXS: A method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41*, 4502−4520.
(8) Lemmen, C.; Lengauer, T. Computational methods for the structural alignment of molecules. *J. Comput.-Aided. Mol. Des.* **2000**, *14*, 215−232 and references therein.
(9) Fradera, X.; Knegtel, R. M. A.; Mestres, J. Similarity-driven flexible ligand docking. *Proteins* **2000**, *40*, 623−636.
(10) Mills, J. E. J.; de Esch, I. J. P.; Perkins, T. D. J.; Dean, P. M. SLATE: A method for the superposition of flexible ligands. *J. Comput.-Aided. Mol. Des.* **2001**, *15*, 81−96.
(11) Labute, P.; Williams, C.; Feher, M.; Sourial, E.; Schmidt, J. M. Flexible alignment of small molecules *J. Med. Chem.* **2001**, *44*, 1483−1490.
(12) Pitman, M. C.; Huber, W. K.; Horn, H.; Kramer, A.; Rice, J. E.; Swope, W. C. FLASHFLOOD: A 3D field-based similarity search and alignment method for flexible molecules. *J. Comput.-Aided. Mol. Des.* **2001**, *15*, 587−612.
(13) Jewell, N. E.; Turner, D. B.; Willett, P.; Sexton, G. J. Automatic generation of alignments for 3D QSAR analyses. *J. Mol. Graphics Modell.* **2001**, *20*, 111−121.
(14) Kramer, A.; Horn, H. W.; Rice, J. E. Fast 3D molecular superposition and similarity search in databases of flexible molecules. *J. Comput.-Aided. Mol. Des.* **2003**, *17*, 13−38.
(15) Korhonen, S. P.; Tuppurainen, K.; Laatikainen, R.; Perakyla, M. FLUFF−BALL, a template-based grid-independent superposition and QSAR technique: Validation using a benchmark steroid data set. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1780−1793.
(16) Girones, X.; Carbo-Dorca, R. TGSA-flex: Extending the capabilities of the topo-geometrical superposition algorithm to handle flexible molecules. *J. Comput. Chem.* **2004**, *25*, 153−159.

(17) Tervo, A. J.; Ronkko, T.; Nyronen, T. H.; Poso, A. BRUTUS: Optimization of a grid-based similarity function for rigid-body molecular superposition. 1. Alignment and virtual screening applications. *J. Med. Chem.* **2005**, *48*, 4076−4086.
(18) Miller, M. D.; Kearsley, S. K.; Underwood, D. J.; Sheridan, R. P. FLOG: A system to select quasi-flexible ligands complementary to a receptor of known three-dimensional structure. *J. Comput.-Aided. Mol. Des.* **1994**, *8*, 153−174.
(19) Sheridan, R. P.; Miller, M. D. A method for visualizing recurrent topological substructures in sets of active molecules. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 915−924.
(20) Toms, A. V.; Wang, W.; Li, Y.; Ganem, B.; Ealick, S. E. Novel multisubstrate inhibitors of mammalian purine nucleoside phosphorylase. *Acta Crystallogr., Sect. D* **2005**, *61*, 1449−1458.
(21) Vedani, A.; Dobler, M. 5D-QSAR: The key for simulating induced fit? *J. Med. Chem.* **2002**, *45*, 2139−2149.
(22) Vedani, A.; Dobler, M.; Lill, M. A. Combining protein modeling and 6D-QSAR − Simulating the binding of structurally diverse ligands to the estrogen receptor. *J. Med. Chem.* **2005**, *48*, 3700−3703.
(23) User and Reference Manual: Quasar 5.0. http://www.biograf.ch/downloads/quasar.pdf (accessed Aug 2006).
(24) Lill, M. A.; Vedani, A.; Dobler, M. *Raptor*—Combining dual-shell representation, induced-fit simulation, and hydrophobicity scoring in receptor modeling: Application towards the simulation of structurally diverse ligand sets. *J. Med. Chem.* **2004**, *47*, 6174−6186.
(25) Lill, M. A.; Dobler, M.; Vedani, A. Prediction of small-molecule binding to Cytochrome P450 3A4: Flexible docking combined with multidimensional QSAR. *ChemMedChem.* **2006**, *1*, 73−81.
(26) User's Guide: Raptor 2.0. http://www.biograf.ch/downloads/raptor.pdf (accessed Aug 2006).
(27) FlexS-77 data set collected by C. Lemmen, G. Klebe, and M. Böhm, first published in Lemmen, C.; Lengauer, T.; Klebe, G. FlexS: A method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41*, 4502−4520.
(28) Vedani, A.; Huhta, D. W. A new force field for modeling metalloproteins. *J. Am. Chem. Soc.* **1990**, *112*, 4759−4767.
(29) (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.
(30) (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.
(31) (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) Sawada, Y.; Kayakiri, H.; Abe, Y.; Mizutani, T.; Inamura, N.; Asano, M.; Hatori, C.; Aramori, I.; Oku, T.; Tanaka, H. Discovery of the first non-peptide full agonists for the human bradykinin B(2) receptor incorporating 4-(2-picolyloxy)-quinoline and 1-(2-picolyl)benzimidazole frameworks. *J. Med. Chem.* **2004**, *47*, 2853−2863. (f) Sawada, Y.; Kayakiri, H.; Abe, Y.; Mizutani, T.; Inamura, N.; Asano, M.; Aramori, I.; Hatori, C.; Oku, T.; Tanaka, H. A new class of nonpeptide bradykinin B(2) receptor ligand, incorporating a 4-aminoquinoline framework. Identification of a key pharmacophore to determine species difference and agonist/antagonist profile. *J. Med. Chem.* **2004**, *47*, 2667−2677. (g) 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.
(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) Cramer, C. J.; Truhlar, D. G. AM1-SM2 and PM3-SM3 parametrized SCF solvation models for free energies in aqueous solution. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 629−666.
(35) Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr., Sect. A* **1976**, *32*, 922−923.

CI6001944