Receptor Surface Models. 1. Definition and Construction

Mathew Hahn*

Molecular Simulations Incorporated, 16 New England Executive Park, Burlington, Massachusetts 01803-5297

Received December 6, 1994®

A receptor site model is a hypothetical model that characterizes the putative active site of a receptor. This paper describes a type of receptor site model called a *receptor surface model*, which is based on the construction of surfaces to represent spatial and electrostatic properties of the receptor active site. A receptor surface model is visually intuitive and is modifiable as the hypothesis is refined. It allows computations comparable to those that can be performed with traditional atomistic models. Structures can be energy minimized within the receptor surface model to arrive at conformations that are consistent with the model, and interaction energies can be estimated. Such calculations facilitate the evaluation of new candidate structures and provide a means to assess the predictive ability of a model.

1. Introduction

In drug discovery, it is common to have measured activity data for a set of compounds acting upon a particular protein but not to have knowledge of the three-dimensional structure of the protein active site. In the absence of such three-dimensional information, one can attempt to build a hypothetical model of the receptor site that can provide insight about receptor site characteristics. This hypothetical model must be deduced primarily from the set of compounds known to bind to the receptor. Such a model is known as a receptor site model. To be useful, the model should be consistent with known data. Ideally, the model should be predictive when evaluating new compounds and provide the medicinal chemist direction in the design of novel compounds.

Receptor site models can be distinguished from *pharmacophore models*. Pharmacophore models postulate that there is an essential three-dimensional arrangement of functional groups that a molecule must possess to be recognized by the receptor. These models are often generated by finding the chemically-important functional groups that are common to the molecules that bind. Receptor site models, in contrast, attempt to postulate and represent the essential features of a receptor site itself, rather than the common features of the molecules that bind to it.

In the absence of direct knowledge of the receptor site, the creation of receptor site models relies on the assumption of an underlying complementarity between the shape and properties of the receptor and the compounds that bind. A molecule and a receptor "see" each other through characteristics presented on the accessible surface of the other, such as the functional groups exposed and the associated molecular fields of the molecule and receptor. Representations of the receptor binding surface can contain detailed information relevant to the binding of a wide variety of molecules with differing features and topologies; a single pharmacophore model has difficulty representing this variety of features and topologies. Further, receptor models can easily and directly represent information such as excluded areas and the shape of hydrophobic regions that are difficult or impossible to represent using pharmacophore models.

A number of methods for constructing receptor site models have been described. The hypothetical active-site lattice (HASL)[1,2] approach represents the molecules inside an active site as a collection of grid points. (Strictly speaking, HASL models are not receptor site models since they characterize molecules and not the active site.) The RECEPS program by Itai and co-workers[3,4] represents the shape around one or more template molecules as a set of grid points tagged with chemical properties. Crippen and co-workers[5] use voronoi polyhedra to build active site models composed of distinct binding regions. Vedani and co-workers[6] have described the generation of full atomistic models of the active site and refer to these models as pseudoreceptors or minireceptors. Comparative molecular field analysis (CoMFA) models[7,8] are effectively receptor site models that represent the three-dimensional field properties around a set of superimposed molecules as a set of grid-based probe interaction energies. Recently, Walters and Hinds[9] described the use of a genetic algorithm to optimally place atoms around a set of superimposed molecules to arrive at a predictive receptor site model.

A critical component of a receptor site model is a representation of the shape of the active site surface. Shape can be defined either implicitly or explicitly. Field-based approaches represent shape implicitly; most other techniques represent shape explicitly. Atomistic van der Waals surfaces are the most common explicit representation. Solvent-accessible surfaces can be used to represent the shape of both small and large molecules.[10,11] Molecular surfaces can be constructed from electron density data.[12] Splined surfaces have been used to define both rigid and malleable surfaces.[13] Surface shape has also been described in terms of spherical harmonics.[14] Explicit surface representations of receptor site models possess several advantages as they are more tangible and correspond more closely to the physical and conceptual notion of a receptor.

A new type of receptor site model, called a *receptor surface model*, is described herein. A receptor surface model is a nonatomistic model that uses explicit surfaces to characterize the shape of the active site. Receptor surface models are simple to understand, are easy to

* The author can be reached via Internet as matth@msi.com.
® Abstract published in *Advance ACS Abstracts*, May 15, 1995.

0022-2623/95/1838-2080$09.00/0

visualize and display, convey important information in an intuitive manner, allow computations comparable to those that can be performed with traditional atomistic models, and can provide predictive capability for evaluating new compounds.

This paper defines receptor surface models, explains how they are generated, describes how energetics calculations can be performed with the model and a molecule, and proposes how they can be applied to qualitative and quantitative studies of a receptor site. The application of these models will be addressed in the following paper,[15] which describes the application of receptor surface models in quantitative structure–activity studies.

2. Definition of a Receptor Surface Model

A receptor surface model represents essential information about the hypothetical receptor site as a three-dimensional surface with associated properties mapped onto the surface model. The location and shape of the surface represent information about the steric nature of the receptor site; the associated properties represent other information of interest, such as hydrophobicity, partial charge, electrostatic potential, and hydrogen-bonding propensity.

The surface is represented internally as a set of points organized in a triangle mesh. The density of the mesh is set when the model is constructed and determines the amount of detail in the representation. The associated property data is stored with the points which comprise the mesh.

While this mesh uses discrete elements to represent a continuous surface, in practice both the visualization and computational uses of the surface model are unaffected by the discrete nature of the internal storage of information. For nearly all purposes, the underlying mesh is invisible to the user.

Finally, the surface model can be either *open* or *closed*. A closed model completely encloses some region of space; an open model has "holes" in the surface. These openings may represent solvent-accessible regions or regions about which nothing is known. In fact, the receptor surface model many not even be continuous; instead, it could be composed of a number of smaller surface patches which represent information about known regions, while leaving unknown regions open and undefined.

3. Generating Receptor Surface Models

Generation of a receptor surface model starts with a series of molecules with associated binding activities. Some number of the most active molecules are aligned. A steric surface is generated to enclose the aggregate of aligned molecules. Scalar properties corresponding to putative properties in the receptor are associated with each surface point. Finally, regions of the receptor surface model are removed to reflect corresponding openings in the receptor site. This section will discuss each of these steps in turn.

A receptor surface model is generated from a set of one or more aligned structures, usually some subset of the most active. If possible, the conformations of the structures should reflect any knowledge of their active conformations in the actual receptor site. Using the set of aligned structures, a receptor surface model is generated over all or some subregion of the structures.

Selecting the appropriate conformations and obtaining an alignment is a complex matter. While there are a number of good techniques for aligning molecules,[16–22] arriving at an alignment model is often not trivial. Errors in the alignment model can lead to models that are incorrect or are poorly predictive.

Once the alignment model is generated for the chosen subset of compounds, a surface is generated to represent their aggregate molecular shape. The surface encloses a volume common to all the aligned molecules. The approach is conceptually similar to the active analog approach,[23] where the union volume is constructed over a set of the most active structures. The shape mapped out by the active structures is assumed to be complementary to the shape of the receptor site itself.

To generate the surface, a volumetric field, characterizing molecular shape, is constructed for each aligned structure. These fields are known as *shape fields*. The shape fields from each individual structure are combined to produce a final volumetric shape field from which an explicit surface is generated. (The shape fields described here differ from the steric fields generated by probe-based approaches like CoMFA[7] or GRID,[24] in which each point in the field corresponds to the steric energy of a probe atom at that point interacting with the structure.)

The technique for generating shape fields and for constructing explicit objects from shape fields is based on work in the computer graphics world of "soft objects".[25] Here, large aggregate objects are constructed from the interacting fields of a number of smaller objects. In this approach, a set of field sources are placed in space, and then a field value is computed at different points in space (typically on a three-dimensional grid). Each field source gives rise to a local field defined by some distance-dependent function. The overall shape field is based on a contribution from each field source. In this work, each field source corresponds to an atom.

Once a combined shape field has been created, an *isosurface* of the field can be computed to create an explicit object with well-defined shape. Given a scalar field $f(x,y,z)$, an isosurface is defined as the set of points satisfying $f(x,y,z) = c$, where c is a constant known as the *isovalue*. A number of techniques can be used to generate isosurfaces from volumetric field data. We use the widely-known marching cubes algorithm.[26] The application of isosurfaces to receptor surface modeling is similar to generating isosurfaces from a field of electron-density data.[12] Isofurface generation is a powerful tool for visualizing and understanding three-dimensional scalar fields and has been used to characterize molecular shape and visualize molecular properties.[27]

There are many field functions which can be used to create a shape field around one or more objects. We describe two-field functions which generate different types of shape fields. Each field, when isosurfaced, yields surfaces with different shape characteristics. The first function, termed the *van der Waals field function*, gives rise to surfaces that characterize the van der Waals shape of molecules well. Atom positions of the aligned molecules are clearly defined. The second

function, termed the Wyvill field function, is taken from the work of Wyvill on soft objects.[25] This function yields surfaces that are a more abstract representation of shape. The surfaces are much smoother and hide individual atom details.

The van der Waals field function is

$$V(r) = r - \mathrm{VDWr} \quad (1)$$

where "r" is the distance from the point to the atom and VDWr is the van der Waals radius of the atom. This function has the property that at the van der Waals surface the value $V(r)$ is zero. Inside the van der Waals volume, the value is negative; outside the volume, the value is positive. This function is computed for every grid point around an atom. If a grid point is encountered that already has a value computed for a different atom, the smaller of the two valves is kept. This function allows a surface to be created inside or outside the van der Waals surface of a set of atoms, at a specified distance from van der Waals surface. Since this function is unbounded, a cutoff distance is employed to calculate field values only within some fixed radius of an atom.

The Wyvill function is also a function of distance r but is bounded and decays completely in finite distance R

$$V(r) = -{}^{4}/_{9}r^6/R^6 + {}^{17}/_{9}r^4/R^4 - {}^{22}/_{9}r^2/R^2 + 1 \quad \text{for } r > 0 \text{ and } r < R \quad (2)$$

This function has the properties that $V(0) = 1$, $V(R) = 0$, and $V(R/2) = {}^{1}/_{2}$. This causes a smooth blending of surfaces. A field value is the sum of the field values contributed by each atom. If a point is outside of R, is not computed. R specifies the distance at which the field value decays to zero. In this work, R varies with atom type; we typically use a value of twice the van der Waals radius of an atom.

The two functions give different representations of molecular shape. Figure 1 shows two surfaces constructed over the same molecule using each function.

The isovalue at which the surface is created is known as the *surface fit* and is used to change the overall size of the surface. With the van der Waals field function, the surface fit value generates a surface that is directly related to the distance from a molecule's van der Waals surface. The ability to construct a surface that is some distance away from the van der Waals surface of a set of atoms corresponds qualitatively to the notion that a ligand binds to a receptor with some amount of tolerance. With the Wyvill function, the surface fit value is directly related to volume enclosed. The surface fit value allows the tightness or looseness of the receptor surface model to be adjusted during model refinement. For example, early in the development of a receptor surface model, a loose fit might be used to represent the gross overall shape of the receptor. As the model is refined, the fit may be tightened to obtain a model that has a more realistic representation of the receptor site.

The marching cubes isosurface algorithm produces a set of triangulated surface points. As part of the isosurface calculation, normals are determined for each point. The normals specify the direction of the interior and exterior of the surface. The normals are stored with the receptor surface and are used in lighting calculations for graphics display and for quickly testing whether atoms are inside or outside the surface during molecule evaluation.

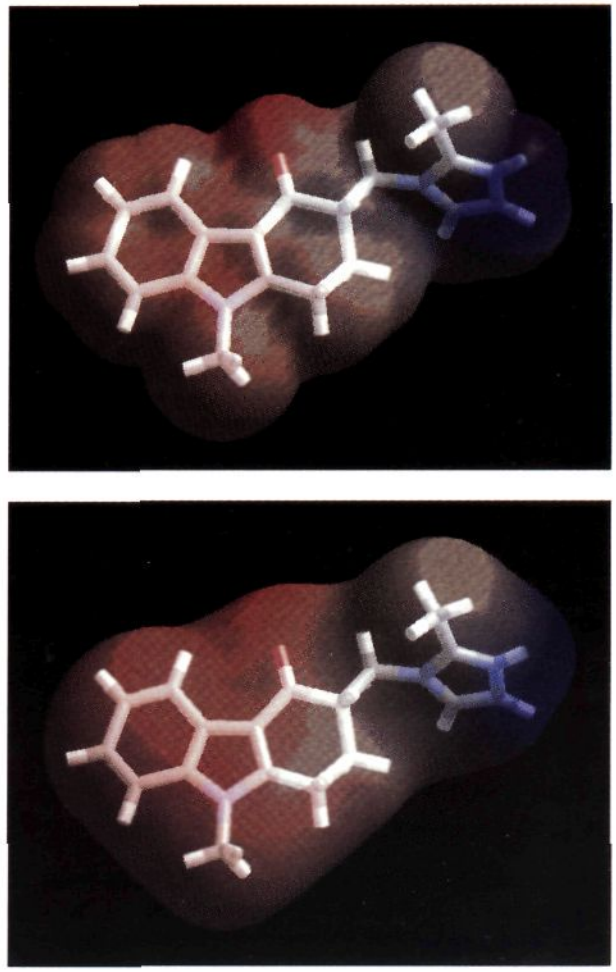

Figure 1. Two different surface models of ondansetron. The top model shows the surface model generated with the van der Waals field function; the bottom model shows the surface model generated with the Wyvill field function.

The generated surface points have a consistent average point density over all regions of the model, though neighboring points are not necessarily evenly spaced. The point density is determined by the initial grid density of the field volume. The default grid spacing of 0.5 Å yields an average surface density of 6 points/Å^2. This gives an average distance between neighboring points (points in the same triangle) of about 0.47 Å.

A receptor surface contains information besides molecular shape. After a surface is created, information corresponding to putative chemical properties of the receptor are associated with each surface point. These properties include partial charge, electrostatic potential, hydrogen-bonding propensity, and hydrophobicity. A scalar value for each of these properties is calculated and stored with every surface point in the model. This information serves two purposes. First, it is used during display to visually convey active site characteristics in an intuitive fashion. Second, it is used when calculating interaction energies between a molecule and a surface model. This subsection describes how these properties are calculated and assigned to the surface points to yield a plausible electrostatic description of the receptor.

As described, a receptor surface model is generated by combining the volumetric shape fields of individual molecules. Likewise, volumetric property fields are calculated for each individual molecule and are combined to produce a final set of property fields describing the regions surrounding the set of aligned molecules. When combining volumes, the final property value at a single volume vertex is the average of the corresponding vertex values from the individual volumes. After the shape field is used to generate a surface, a set of property values for each surface point is calculated by interpolating from eight volume vertex points of the property fields. This interpolated value is then stored with the surface point.

The first property stored with each surface point is the partial atomic charge which would be desirable at a particular position in the receptor site. The assumption is made that the charge on the receptor surface is complementary to the partial atomic charge of any atom in contact with the surface. If the surface model is constructed over a single molecule, each surface point is given a charge which is equal to but opposite in sign to the charge of the closest atom in the molecule. If the site model is constructed over a set of molecules, each surface point is given a charge which is equal and opposite to the average partial atomic charge of the set. The average for a point is found by summing the partial atomic charges of the closest atom in each molecule and dividing by the number of molecules. This assumes that each molecule contributes equally to the description of the model.

The second property stored with each surface point is the electrostatic potential expected to be seen by a point at that location in the receptor. Again, an assumption is made that the electrostatic fields of molecules that bind to a receptor, and the receptor site itself, are complementary. When a surface is constructed from a single molecule, each surface point is given an electrostatic potential value which is equivalent to but opposite in sign to the distance-dependent electrostatic potential at that point, calculated by summing Q_aQ_p/r^2 for all atoms (where Q_a and Q_p represent atom charge and point charge, respectively). As before, if a surface is constructed over a set of molecules with activity data, a point's electrostatic potential can be calculated as the average of the corresponding point in each individual volume.

A hydrogen bond property is also stored with each surface point and corresponds to the tendency of the point to be involved in a hydrogen bond. For hydrogen bonding, a value of 1.0 is assigned to volume vertex points where a hydrogen bond acceptor would be desirable in the receptor. A value of -1.0 is assigned to volume vertex points where a hydrogen bond donor hydrogen might be desirable. These vertex points are found by projecting a cone away from each hydrogen bond donor or acceptor atom. (Since it is assumed that the alignment procedure has oriented functional groups, including hydrogen bonding hydrogens, there is no explicit manipulation of hydrogen positions to take into account hydrogen rotomeric states.) Hydrogen bond acceptors are defined as any oxygen or nitrogen atom with a free lone pair of electrons. Hydrogen bond donors are any hydrogens attached to oxygen or nitrogen. The hydrogen bond values from different volumes are averaged, and a surface point is given a value by interpolating from the final volume. The resulting surface values are in the range $[-1.0,1.0]$.

A final property, corresponding to hydrophobicity, is stored with every surface point. Each point is classified as being either hydrophobic or nonhydrophobic using the other calculated properties. Hydrophobicity is a binary property; a value of one is assigned to points in hydrophobic regions and zero to all other points. A hydrophobic point is a point with a low partial charge (absolute value less than 0.15), a low electrostatic potential (absolute value less than 0.01), and a low hydrogen bond-donating or -accepting propensity (absolute value less than 0.1).

The isosurface procedure produces a surface that entirely encloses the molecules over which it is generated. The surface has no holes and is known as a *closed* model. Often, a receptor site is not best represented by an enclosed volume. For example, some portion of a ligand may be exposed to solvent. Such a receptor site is better represented by a surface with an open cavity. Alternatively, a user may want to generate a receptor surface model that only characterizes distinct subregions of a receptor site. The regions may be small, distinct surface patches with some relative orientation in space. Both of these latter models are known as *open* models.

An open receptor surface model is generated from a closed model by cutting away regions of the surface. Several types of solid modeling operations used in mechanical CAD are supported. A surface can be intersected with various primitive object types to cut away regions that are either inside or outside the primitive object. There are currently three primitive objects types: spheres, cylinders, and planes. Figure 2 shows two examples of open receptor surface models.

Since a surface is comprised of a list of triangles, the intersection routine does not simply discard triangles that have a vertex that is within the primitive object. Discarding such a triangle will leave the surface with jagged edges. Figure 3 (left) shows such a surface model around propane. The central atom defines a sphere which is used to cut the surface, and in the absence of smoothing, the model is jagged.

While the existence of jagged edges does not affect the integrity of the model, it is unsightly. To correct this problem, the intersection routine smooths the intersection region by generating new triangles. Smoothing is performed when intersecting with any primitive object type. Figure 3 (right) shows the same receptor surface model around propane after the smoothing operation is performed.

4. Using Receptor Surface Models

A receptor surface model is designed to allow the information about the shape and chemical properties in a hypothetical active site to be used and displayed in an intuitive fashion. The model can be used to visually convey active site characteristics. The model can be used to calculate energetics information about the binding of a molecule into the model and can be used to minimize a molecule into a conformation consistent with the conformations of the molecules used to generate the receptor surface model. The model can be refined and changed from the results of new experiments

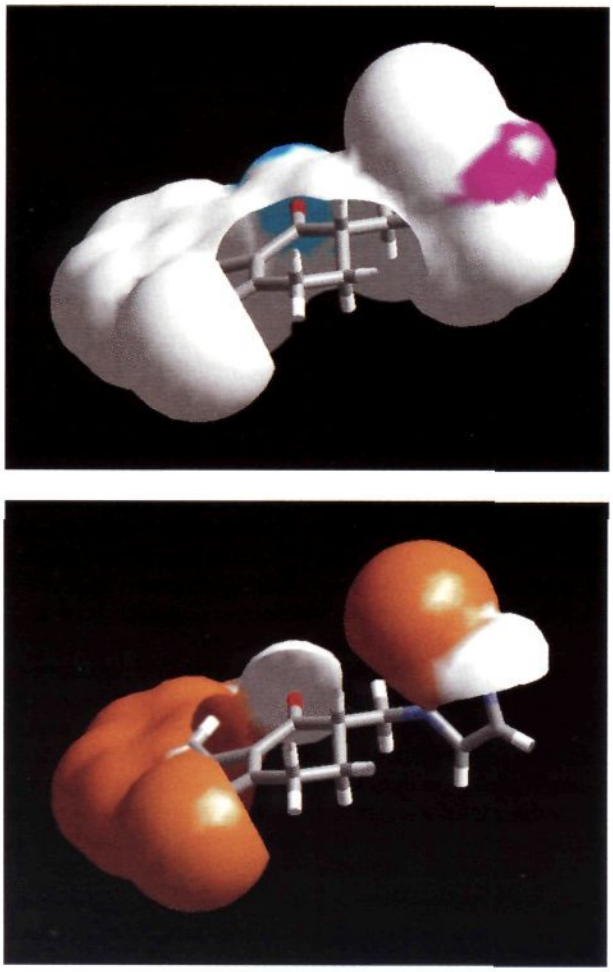

Figure 2. Two different open receptor surface models constructed from ondansetron. The upper model shows a surface model containing a single open region. The low surface model is comprised of two disconnected surface patches. The upper model has been color coded on the basis of hydrogen-bonding propensity (cyan represents hydrogen bond donor; magenta represents hydrogen bond acceptor). The lower model has been color coded according to hydrophobicity (brown indicates hydrophobic regions).

or to reflect new information from other sources. Finally, the model can be applied to a number of tasks, including quantitative structure–activity relationship modeling,[15] alignment of molecules, and de novo molecule design. The following subsections discuss each of these points in turn.

Receptor site information is conveyed visually by mapping properties onto the surface. Regions of the surface are color coded to indicate particular chemical properties. The intensity of the color on the surface corresponds to the magnitude of the property. For example, assume that a receptor surface model is constructed from six aligned molecules and each of the molecules position a hydrogen acceptor in the same location. Three of the molecules position a second hydrogen bond acceptor in a different location. If hydrogen bonding propensity is mapped onto the surface, the region adjacent to the six acceptors will show a full intensity color, indicating a strong likelihood of a hydrogen bond donor existing at that location. The region adjacent to the three hydrogen bond acceptors will show the same color at half the intensity. Since the receptor surface model is hypothetical, it must be remembered that the property characteristics mapped may not always reflect properties of the actual receptor. Figures 1 and 3 show receptor surface models with different properties mapped. Color mapping only displays a single property at one time. Another graphical display technique, texture mapping, could also be used and would allow multiple properties to be displayed simultaneously.[28]

Receptor surface models can be displayed semitransparent, as shown in Figure 1. This allows one to see inside the surface and facilitates docking or modifying a structure within the context of the model.

The receptor surface model supports computations that are analogous to those which can be performed with an atomistic model of a receptor site. A structure can be docked into the model. Energetics calculations can be performed to minimize the structure with respect to the model. Energetic information like the strain energy of the structure of the "bound" state and the interaction

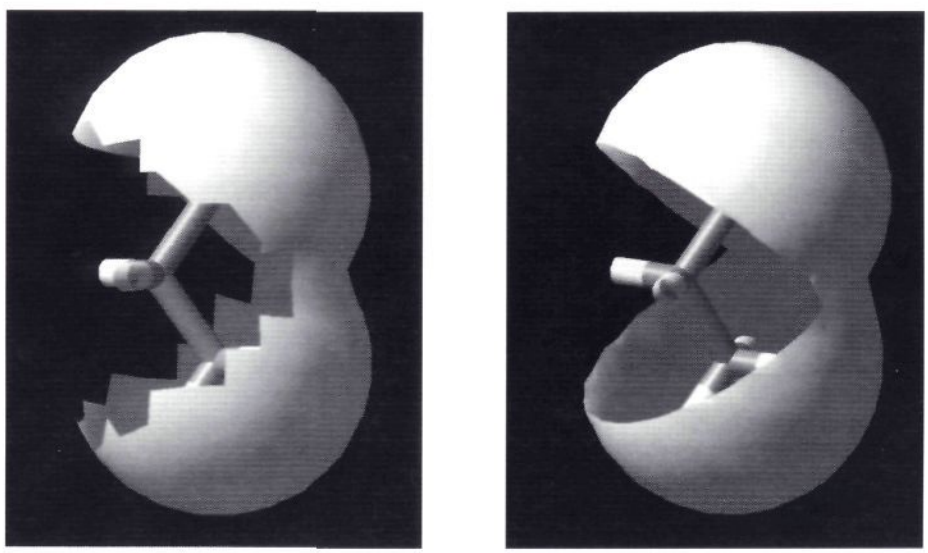

Figure 3. A receptor surface model after cutting with a sphere primitive object. Left, the jagged edges are caused by the removal of the triangles that have a vertex within the sphere primitive. Right, a receptor surface model after cutting with the sphere primitive. The smoothing operation repairs the triangles that were cut.

energy between the structure and the model is available for evaluation. This information can be used in a qualitative fashion to rank potential test compounds or used quantitatively as descriptors for a QSAR analysis.[15]

A unique feature of the receptor surface model is that a molecule can be energy minimized in the context of the model, where the molecule "feels" the surface of the model. The energetics calculations rely on a fast, approximate force field, termed *Clean*. The force field quickly calculates reasonable geometries and energies of drug size molecules, either in the presence or absence of a receptor surface model. The force field is described fully in the Appendix. A description of how Clean is used to minimize a molecule against a receptor surface model follows.

As discussed earlier, each site model surface point has a number of surface property values associated with it, including point position, point charge, and the direction of surface interior (normal). This is used to calculate nonbonded interaction energy between the atoms in a structure and the points on the surface. The total interaction energy contribution of a single atom is calculated by summing the individual interaction energies of all valid atom/point pairs. An atom/point pair is valid if the distance between the pair is less than a 6 Å cutoff distance and if the atom is inside the surface with respect to the point. Testing whether an atom is inside or outside the surface is computed with the point normal.

The nonbonded energy between an atom and the surface is composed of a *van der Waals* term, an *electrostatic* term, and a *desolvation energy correction* term. The van der Waals term is a scaled Lennard-Jones expression:

$$E_{(\mathrm{vdw})} = K((\mathrm{RA}/r)^{12} - 2(\mathrm{RA}/r)^{6})D \quad (3)$$

$$\mathrm{RA} = \mathrm{VDWr}C_{\mathrm{h}} \quad (4)$$

RA is the hybridization corrected van der Waals radius for the atom, r is the distance between the atom and the surface point (the radius of the surface point is implicitly zero), K is the well depth constant and is set to 0.1 for all van der Waals atom/point interactions, and D is an empirically derived point-density scaling factor, which scales the van der Waals energy and forces so that ideal atom/surface interactions yield a van der Waals value of 0.0125 kcal/Å^2 of surface contact. D is set to 0.01 for the default grid resolution of 0.5 Å and surface point density of 6 points/Å^2.

The electrostatic term is a monopole−monopole Coulombic function and is calculated using eq 5

$$E_{(\mathrm{ele})} = (322.1Q_{\mathrm{A}}Q_{\mathrm{P}}/r)DS(r) \quad (5)$$

r is the distance between the atom and the surface, Q_{A} is the partial atomic charge of the atom (Gasteiger[29] charges are used by default), and Q_{P} is the charge of the surface point. The point charge can be obtained from either the partial charge associated with point or the electrostatic potential value associated with point. D is the same point density correction factor used in the van der Waals calculation, and $S(r)$ is an atom-based switching function and is described in the Appendix.

The desolvation energy term is a penalty function. It introduces a penalty when polar atoms are placed in hydrophobic regions of the receptor surface model. If the fraction of hydrophobic points to total points in proximity to a polar atom is greater than 90%, then a desolvation correction energy is added. This energy is proportional to the exposed surface area of the polar atom. A value of 0.3 kcal/mol Å^2 is used. The use of the desolvation energy term is based on the simplifying assumption that any electrostatic (including H-bond) interactions between a molecule and surrounding water molecules will be replaced by similar electrostatic interactions when the molecule is bound. The desolvation energy term is added to the total energy after the minimization process is complete (the term is not treated as a nonbonded interaction).

The calculation of molecule/surface interaction energy involves summing many atom−point pair interactions. For each atom, there may be several hundred surface points within the cutoff range. Calculating an interaction energy by summing all valid pair interactions can be slow. To speed up the atom−surface calculation, a virtual grid of precomputed energies and forces is built up and stored during minimization. An atom's position is mapped to a volume element (voxel). The voxel index is a 3-tuple of x,y,z values, clamped by the grid resolution. The eight vertices of the voxel are determined. for each voxel vertex, a hash lookup is performed to see if interaction energy information for the vertex has already been computed. If it has not, then the van der Waals energy and forces and the electrostatic energy and forces are calculated for that voxel vertex. This information is then stored in the hash table.

The electrostatic energy and force for a voxel vertex are calculated only once. A unit positive charge is assigned to the voxel vertex for the calculation. The energy and force for the positive charge interacting with the surface is calculated and stored. When the interaction energy and force is to be calculated for an atom with some partial charge, the energy and forces for the vertex are retrieved and then scaled by the partial charge. For van der Waals terms, a separate van der Waals entry is stored for every atomic van der Waals radius encountered, since the van der Waals force and energy is radius dependent and cannot be scaled.

Once all eight voxel vertex energies and derivatives have been calculated, the energy and force on the atom inside the voxel is calculated by trilinear interpolation of the eight voxel points. This result is an approximation, since neither the van der Waals nor electrostatic energy is a linear function. At the standard grid resolution used (0.5 Å), reasonable energy estimates are produced, and the difference between interpolated and full-interaction results is generally small.

A molecule is minimized using interpolated energies and forces until a termination condition is reached. This is currently 300 cycles of steepest descent or a minimum energy change of 0.01 kcal. After the minimization is terminated, five additional cycles of full interaction minimization is performed to obtain the final, full interaction energy. The minimization procedure is rapid. Typically, a structure can be evaluated with respect to a surface in several seconds.

During minimization, all surface points are fixed. The process models a flexible ligand inside a rigid

receptor site. The structure being minimized, therefore, may be perturbed significantly by the procedure, since the geometry of the structure will adopt a conformation consistent with the shape of the surface. For example, if a surface is created over a chair cyclohexane and a boat conformation structure is minimized against the surface, the boat conformation can be flipped to chair in the process. Sometimes a structure will assume a geometry lower in energy than the starting structure. Often, however, a structure will be forced to adopt a geometry higher in energy than the initial geometry because of the shape of the surface. The van der Waals term can induce bond and angle distortions. To detect conformation strain introduced by the minimization, a second minimization is performed on the structure in the absence of the surface. This second minimization will bring the structure to a nearby minimum energy conformation.

The minimizations produce three energy values. The first value is the nonbonded interaction energy between the structure and the surface, termed E_{interact}. The second value is the internal strain energy of the structure with respect to the surface. This is the energy of the "bound" conformation and is the sum of all bond, angle, torsion, inversion, and intramolecular nonbonded energies. This value is termed E_{inside}. The third value is the internal energy of the structure, after it has been allowed to relax without feeling the surface. This value is termed E_{relax} and will always be less than or equal to E_{inside}.

The E_{interact}, E_{inside}, and E_{relax} values can be quickly inspected to facilitate an evaluation of goodness of fit. Evaluation is typically based upon two criteria: E_{interact} and the difference between E_{inside} and E_{relax}. The more negative E_{interact} is, the better the complementarity between the molecule and the model.

The difference between E_{inside} and E_{relax} is a measure of strain energy between the bound conformation and a nearby relaxed conformation. The smaller value, the less strain introduced by the minimization within the model. This strain estimate indicates nothing about the difference between the bound conformation and the global energy minimum. If a conformational search has previously been performed on the structure, then E_{relax} can be replaced with the global energy minimum (or lowest minimum found) to give a better estimate of strain energy.

The energetic results can also be visualized by mapping energy of interaction onto the surface. This allows the user to see where favorable and unfavorable interactions are present. Van der Waals energies can be mapped to see where steric groups "bump" into the receptor surface model. Electrostatic energies can be mapped to see good and bad charge interactions. Figure 4 shows a model mapping steric energy.

Because a structure can be minimized quickly, with the results displayed in color on the surface, a user can quickly test a hypothesis by editing the molecule to see if changes can be made that strengthen the interaction energy without introducing significant strain in the structure. In addition, because the user can always map the initial receptor properties (charge, H-bonding, hydrophobicity), the user can be guided in terms of what editing changes to make in various regions of the model.

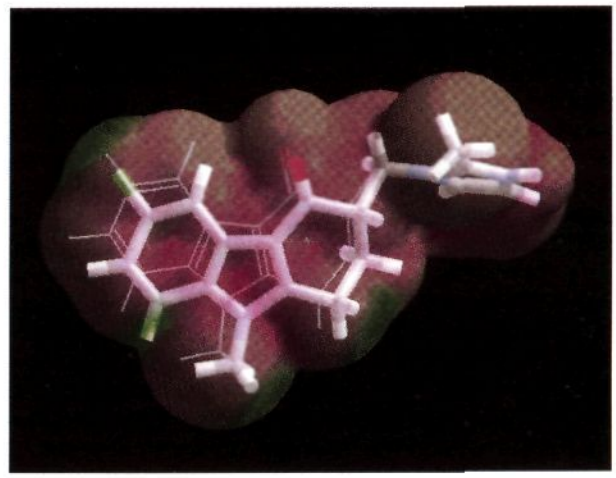

Figure 4. A receptor surface model with VDW steric energy mapped onto it. A chlorine atom has been added to the phenyl ring of ondansetron (shown in wireframe representation). Minimization shifts the new molecule (shown in cylinder representation) and introduces strain energy, which is displayed as green "hot spots" on the receptor surface model. Purple regions indicate areas of favorable VDW contact.

Once a reasonable receptor surface model has been defined, a series of structures can be evaluated against the model. The results of the minimization procedure can be used as descriptors either to refine the model or to predict activity. These descriptors can also be combined with other three-dimensional or two-dimensional descriptors in a QSAR analysis.

A receptor surface model is a hypothesis and as such much be validated and refined over time. The receptor surface model supports modification operations that allow such refinement.

A receptor surface model mimics a rigid receptor site. The surface does not deform during minimization against a molecule. However, it is possible to interactively modify the model prior to minimization. There are several ways that a receptor surface model can be modified. One approach is to change the surface charge characteristics. Particular electrostatic interactions can be emphasized or deemphasized by scaling specific point charges.

A second approach is to change the overall size of the model but not change the general shape of the model. This is accomplished by changing the fit parameter and regenerating a larger or smaller surface at a new isovalue.

A third approach is to modify the actual shape of a model. The use of soft objects to define surfaces allows the steric field itself to be changed interactively. Dragging a set of one or more field sources (atoms) inside the field can either add or remove local field density. A new surface can be quickly regenerated (at some isovalue) every time the field is modified. This has the effect of making the surface malleable. New bumps, protrusions, or depressions can, therefore, be easily added to any region of a model.

Finally, open receptor surface models can be created by using solid modeling operations. Regions of the surface can be removed by intersecting the surface with spheres, planes, or cylinders of various size. Leaving a region open allows other parts of the model to be preferentially refined. As molecules are evaluated

against the model, atoms that lie in any open regions do not contribute to the interaction energy.

5. Applications of Receptor Surface Models

A receptor surface model provides a framework for representing, in an abstract nonatomistic fashion, a receptor active site. Since the model is a hypothesis, it can be tested to see if it is predictive. If a model is unacceptable it can be rejected or refined.

Once a receptor surface model has been constructed, it can be used in much the same way that an atomistic model of a receptor site can be used. A model can be displayed and manipulated graphically in three dimensions. Binding properties of different regions in the model can be visualized. The surface can be made transparent so that it is possible to see inside the model. Molecules can be brought into the site model and can be evaluated with respect to the model.

Molecules interact with the surface model in a realistic fashion. Structures are evaluated against the model by energy minimizing the structure against the model. This process is analogous to minimizing a structure in an actual receptor, holding the receptor atoms fixed. The assumption that the receptor site remains fixed in geometry is a limitation, but is not without experimental support. Studies of HIV-1 protease bound to a set of inhibitors indicates that the geometry of the receptor remains relatively constant even when there is significant structural diversity in the inhibitors.[30]

The ability to minimize a structure against a model allows one to flexibility fit a structure into the model. The minimization can also be used as a shape- (and not atom-) based alignment technique. Applying the fitting process to a set of molecules will force all the molecules to adopt a shape that is consistent with the model.

The results of minimization yield several energetic descriptors corresponding to molecule/site interaction energy and strain energies of the bound and unbound structure. These values can be used as three-dimensional descriptors in QSAR studies. Hopfinger advocates using binding energetics as QSAR descriptors when the receptor is known.[31,32] Even when the receptor is unknown, using binding energetics from a hypothetical receptor surface model can be a useful predictive tool.

After the minimization of a molecule, information about location-specific van der Waals and electrostatic interactions is maintained. The receptor surface model is then visually marked to indicate where the molecule is making favorable and unfavorable contacts with the site. This information can guide users as they manually edit, design, or optimize a structure inside the site.

Finally, a receptor surface model can be used for automatic de novo design experiments. A model characterizes the boundary of a receptor site, and it supports energetic calculations. Molecules can be grown into the site using atom or fragment-based de novo algorithms, guided by the energetics results. The ability to construct de novo into a receptor site model is useful, even if the underlying model is not strictly correct. If the model is constructed from one or more structures known to bind, the structures generated would be expected to have a similar shape and spatial distribution of properties.

6. Discussion

Receptor surface models differ from pharmacophore models in that the former try to capture essential information about the receptor, while the latter capture information about the commonality of compounds that bind. Pharmacophore models generally represent some minimal set of features present in the actives and postulate that those features, in some configuration, are required for binding. Since these models do not usually represent the receptor boundary, molecules that fit the model can still be inactive because of additional regions of the molecule that are sterically unfavorable. Pharmacophore models, therefore, tend to be geometrically underconstrained (while topologically overconstrained); this steric underconstraint leads to false positives, that is, compounds that are deemed active by the model but which are inactive when tested.

Receptor surface models, on the other hand, tend to be geometrically overconstrained (and topologically neutral) since, in the absence of steric variation in a region, they assume the tightest steric surface which fits all training compounds. This may be significantly more restrictive than the actual boundaries of the receptor. This means they are prone to false negatives: new actives (not used in creating the model) may map out new regions of the active site and thus may evaluate poorly against the model. This is illustrated by the *o*piate analgetics. Generation of a receptor surface model from molecules such as morphine, meperidine, and levorphanol (all having an *N*-methyl group) would indicate that a meperidine analog where the *N*-methyl is extended by a phenyl butyl side chain would be inactive. In fact, this analog has 100−1000 times the activity of morphine. In such cases, as new information is obtained, the receptor surface model can be modified to extend the surface into new regions; pharmacophore models, since they do not directly represent steric boundaries, are less suitable for such modification.

This paper has described the creation of receptor surfaces from one or more ligand molecules, where the surface is assumed to have characteristics that are complementary to the ligands. When the receptor structure is known, it is also possible to construct a receptor surface model from the actual receptor. The same field functions and isosurface routines can be used to generate a triangulated surface from the atoms in the receptor. The normals of the surface must be flipped so that the direction of inside and outside the receptor site is consistent. Rather than mapping molecule complementary properties to the surface, we map the actual properties of the atoms found in the receptor. Further, just as it is possible to construct a surface from several superimposed molecules, it is possible to superimpose two or more protein homologs and to construct an aggregate site model. A surface constructed from a receptor should be comparable to a model constructed from a set of structures that bind to the receptor.

We described two functions for generating shape fields from which surfaces are generated. Many different functions can be envisioned. Gaussian or exponential decay functions will generate surfaces with different shape characteristics. The approach described here can be used with any technique for generating a steric or shape field from which an isosurface can be generated.

This paper has not addressed how best to optimize both the shape of the surface and the charge distribution on the surface to get predictive models. One can envision both manual and automatic procedures to accomplish this. Initial assignment of charges to the surface is crude and is based on the assumption of atom charge or electrostatic potential complementarity. The evaluation of a structure within a model is fairly simple and could be extended, for example, by adding solvation and entropic effects or by replacing the Clean force field with a fully parameterized force field like CHARMM[33] (at the expense of evaluation speed).

7. Conclusions

A novel form of receptor site model, called a *receptor surface model*, has been described. A receptor surface model is generated from a series of aligned molecules with associated binding activities. A steric surface is generated to enclose the aggregated aligned molecules, and scalar properties corresponding to putative properties in the receptor are associated with each surface point. Regions of the receptor surface model can be removed to reflect corresponding openings in the receptor site or areas of the receptor site about which nothing is known.

The receptor surface model has characteristics that make it a desirable representation for receptor site hypotheses. The models are intuitive and visually appealing. The receptor surface model supports energetics calculations for the interactions of molecules with the model. The model uses a unique force field, termed *Clean*, which is optimized for speed and accuracy when used with the receptor surface model representation. The model provides interactive and qualitative feedback for evaluating and testing new structures. The model generates quantitative information that is available for QSAR analysis. The models are easily modified as the active site hypothesis is refined.

This combination of qualitative and quantitative properties makes the receptor surface model a strong candidate for many applications in receptor site modeling. The following paper presents one such application, the use of receptor surface models in QSAR studies.[15]

Acknowledgment. The author would like to thank Dr. David Rogers for his insightful contributions and help in the preparation of this manuscript. The author would also like to thank Molecular Simulations Incorporated for supporting the research that led to this publication. The methods described in this paper are implemented and available as part of MSI's Drug Discovery Workbench.[34]

Appendix

Description of the Clean Force Field. The energetics calculations can minimize a structure in the presence or absence of a receptor surface model. The calculations employ a fast, approximate force field, termed the Clean force field. The force field quickly calculates reasonable geometries and energies of drug size molecules. Clean shares commonality with both the Drieding force field[35] and the Chem-X force field[3]16 in that it does not rely on extended force field atom types. Only element type, hybridization, and bond type are used in calculating the energy of a system. This allows the energy of a molecule to be recalculated quickly after an atom type or bond type modification, without applying any force field atom typing rules.

Table 1. Default Bond Radius and van der Waals Radius Parameters

element	bond radius	VDW radius
H	0.33	1.20
C	0.77	1.95
N	0.70	1.83
O	0.66	1.70
F	0.61	1.73
Cl	0.997	1.97
Br	1.17	2.10
S	1.04	2.00
P	0.89	2.08

The overall form of the energy expression is

$$E_{(\text{total})} = E_{(\text{bnd})} + E_{(\text{ang})} + E_{(\text{tor})} + E_{(\text{inv})} + E_{(\text{vdw})} + E_{(\text{ele})} \quad (6)$$

where

$E_{(\text{bnd})}$ = the bond stretch term

$E_{(\text{ang})}$ = the angle bending term

$E_{(\text{tor})}$ = the torsion twist term

$E_{(\text{inv})}$ = the out of plane bending term

$E_{(\text{vdw})}$ = the van der Waals nonbonded term

$E_{(\text{ele})}$ = the electrostatic nonbonded term

The Bond stretch term is a simple harmonic:

$$E_{(\text{bnd})} = K(R - R_0)^2 \quad (7)$$

The force constant, K, is determined by the bond type. Single bonds have $K = 100$ kcal/mol, double bonds $K = 200$ kcal/mol, aromatic $K = 400$ kcal/mol, and triple bonds $K = 500$ kcal/mol. R is the distance between the two atoms. R_0 is the equilibrium bond length. The value for R_0 assumes the additivity of bond radii and is corrected based upon both hybridization and bond type information:

$$R_0 = (R_a C_h + R_b C_h) C_b \quad (8)$$

where R_a and R_b are the bond radii of the two bonding atoms, C_h is a correction term for each bond radius based on the hybridization of each atom, and C_b is a correction term for the total bond length based upon the bond type of the bond. The expression reproduces the alternate short/long bond length of allene structures, which Drieding does not. The hybridization correction term C_h is 1.0 for sp^3-hybridized atoms, 0.95 for sp^2-hybridized atoms, and 0.90 for sp hybridization. The total bond length correction C_b is 1.0 for single bonds, 0.91 for double bonds, 0.93 for aromatic bonds and 0.87 for triple bonds.

The default bond radius parameters for common atoms are shown in Table 1.

This bond stretch term in Clean is unique in that the equilibrium bond length is found by scaling bonding radii by hybridization type and then scaling the overall bond length by the bond type. This method produces bond lengths that agree well with bond lengths produced

Table 2. Comparison of Bond Lengths Predicted by Different Force Fields

bond type	Clean	MMFF	CHARMm	DREIDING
Csp^3-Csp^3	1.54	1.53	1.53	1.55
Csp^3-Csp^2	1.49	1.50	1.51	1.48
Csp^2-Csp^2	1.46	1.44	1.47	1.40
$Csp^2=Csp^2$	1.33	1.34	1.34	1.40
Csp-Csp(s)	1.39	1.37	1.45	1.19
Csp#Csp(t)	1.21	1.20	1.19	1.19
Csp^3-Nsp^3	1.46	1.44	1.46	1.43
Csp^3-Nsp^2	1.42	1.44	1.45	1.40
Csp^2-Nsp^2	1.39	1.37	1.35	1.34
$Csp^2=Nsp^2$	1.27	1.29	1.35	1.34
Csp#Nsp(t)	1.15	1.16	1.16	1.15
Csp^3-Osp^3	1.42	1.42	1.41	1.43
Csp^2-Osp^3	1.39	1.37	1.36	1.36
$Csp^2=Osp^2$	1.23	1.22	1.21	1.25

Table 3. Default Angle Parameters

hybridization	$u°$ (deg)	K
sp^3	109.5	60
sp^2	120	70
sp	180	80
sp^3-3mem	80	60
sp^3-4mem	100	60

Table 4. Periodicity (N), Barrier to Rotation (K), and Equilibrium Angle ($u°$) for Torsions

bond hybridization	N	K	$u°$ (deg)
sp^3-sp^3	3	1.0	180
sp^3-sp^2	-2	0.25	0
sp^2-sp^2	-2	15.0	180
sp^2-sp2(single)	-2	2.0	0

by more rigorously parametrized methods. A comparison of Clean against CHARMM,[33] DRIEDING,[35] and MMFF[37] was performed on a set of small, unstrained molecules, with various bonding and hybridization topologies. The results of this comparison indicate that the simple scheme used in Clean is quite effective. Table 2 shows the bond lengths generated by Clean, MMFF, CHARMM, and DRIEDING.

The angle bending term is also harmonic and is parameterized only on the hybridization of the central atom of the angle.

$$E_{(ang)} = K(\cos u - \cos u^0)^2 \quad (9)$$

There are two special hybridization types for 3- and 4-membered ring atoms. Table 3 lists the angle parameters.

The torison term is a standard periodic function. The parameterization is simplified in that it uses only the central two atoms of the torsion

$$E_{(tor)} = K(1 - \cos N(u - u^0)) \quad (10)$$

where K is the barrier of rotation, N is the periodicity of the torsion, and u^0 is the equilibrium torsional angle. The periodicity and barrier is based upon the hybridization and bond type of the two central atoms. These values are shown in Table 4.

The out-of-plane bending term uses the same form as the torsion term but is applied to the improper torsion angle. This is the same approach as used in AMBER.[38] The out-of-plane term is only calculated for planar (sp^2) atoms with three attached atoms.

The nonbonded van der Waals term is a standard 12-6 Lennard-Jones expression

$$E_{(vdw)} = K(((R1 + R2)/r)^{12} - 2((R1 + R2)/r)^6) \quad (11)$$

$$R1 = \mathrm{VDWr}C_h \quad (12)$$

where K is the well depth constant and is set to 0.1 for all van der Waals interactions and $R1$ and $R2$ are the corrected van der Waals radii of each atom. The radii of each atom are corrected using the same hybridization values that the bond stretch term uses: for sp^2, C_h = 0.95; for sp, C_h = 0.90. A united atom model is used for all carbon atoms to reduce the total atom number of nonbonded interactions. A value of 0.08 Å is added to the hybridization-corrected radius of carbon for each hydrogen attached to a carbon. Hydrogens on non-carbon atoms are explicitly considered. The van der Waals parameters for the most common elements are shown in Table 1.

The electrostatic term is a standard monopole–monopole Coulombic function. The electrostatic energy is calculated using

$$E_{(ele)} = (332.1Q_iQ_j/r_{ij})S(r) \quad (13)$$

Q_i and Q_j are the atom charges and r is the distance between the atoms. Gasteiger[29] charges are calculated for all atoms by default. Nonbonded van der Waals and electrostatic energies are not calculated for bonded atoms (1,2 interactions) or in an angle (1,3 interactions). The van der Waals energy is calculated for 1,4 interactions, but the van der Waals radii are scaled by 0.75. A nonbond cutoff distance of 8 Å is used for intramolecular atom pairs. This is a fairly tight cutoff distance, and so the energy for forces are switched (from 7 to 8 Å). The switching function $S(r)$ is an atom-based switching function that is employed in CHARMM[33]

$$S(r) = (r^2_{off} - r^2)^2(r^2_{off} + 2r^2 - 3r^2_{on})/(r^2_{off} - r^2_{on})^3 \quad \text{for } r_{on} < r < r_{off} \quad (14)$$

$$S(r) = 1 \quad \text{for } r < r_{on} \quad (15)$$

This cubic function of r^2 yields a continuous potential energy and force. Currently, r_{on} = 7 Å and r_{off} = 8 Å. All pairs of atoms with a distance greater than this cutoff are neglected in the nonbonded calculation.

The minimizer is steepest descent. There are two stop conditions: if the energy converges (ΔE < 0.001 kcal) or if a maximum number of iterations have been performed (default 300 steps).

References

(1) Doweyko, A. M. The Hypothetical active site lattice. An approach to modelling active sites from data on inhibitor molecules. *J. Med. Chem.* **1988**, *31*, 1396–1406.

(2) Wiese, M. The Hypothetical Active-Site Lattice. In *3D QSAR in Drug Design: Theory Methods and Applications*; Kubinyi, H., Ed.; Escom: Leiden, 1993; pp 80–116.

(3) Kato, Y.; Itai, A.; Iitaka, Y. A novel method for superimposing molecules and receptor mapping. *Tetrahedron* **1987**, *43*, 5229–5236.

(4) Kato, Y.; Inoue, A.; Yamada, M.; Tomioka, N.; Itai, A. Automatic superposition of drug molecules based on their common receptor site. *J. Comput. Assisted Mol. Design* **1992**, *6*, 475–486.

(5) Srivastava, S.; Richardson, W. W.; Bradley, M. P.; Crippen, G. M. Three-Dimensional Receptor Modeling Using Distance Geometry and Voronoi Polyhedra. In *3D QSAR in Drug Design: Theory Methods and Applications*; Kubinyi, H., Ed.; Escom: Leiden, 1993; pp 80–116.

(6) Snyder, J. P.; Rao, S. N.; Koehler, K. F.; Vedani, A. Minireceptors and Pseudoreceptors. In *3D QSAR in Drug Design: Theory Methods and Applications*; Kubinyi, H., Ed.; Escom: Leiden, 1993; pp 336–354.
(7) Cramer, R. D.; 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.
(8) Cramer, R. D.; DePriest, S. A.; Patterson, D. E.; Hecht. The Developing practice of Comparative Molecular Field Analysis. In *3D QSAR in Drug Design: Theory Methods and Applications*; Kubinyi, H., Ed.; Escom: Leiden, 1993; pp 443–485.
(9) Walters, D. E.; Hinds, R. M. Genetically Evolved Receptor Models: A Computational Approach to Construction of Receptor Models. *J. Med. Chem.* **1994**, *37*, 2527–2535.
(10) Connolly, M. L. Analytical molecular surface calculation. *J. App. Crystallogr.* **1983**, *16*, 548–558.
(11) Connolly, M. L. Sovlent-accessible surface of proteins and nucleic acid. *Science* **1983**, *221*, 709–713.
(12) Purvis, G. D. On the use of isovalued surfaces to determine molecule shape and reaction pathways. *J. Comput.-Aided Mol. Design* **1991**, *5*, 55–80.
(13) Klein, T. E.; Huang, C. C.; Pattersen, E. F.; Couch, G. S.; Ferrin, T. E.; Langridge, R. A real-time malleable surface. *J. Mol. Graphics* **1990**, *8*, 16–24.
(14) Leicester, S. E.; Finney, J. L.; Bywater, R. P. Description of molecular surface shape using Fourier descriptors. *J. Mol. Graphics* **1988**, *6*, 104–108.
(15) Hahn, M. A.; Rogers, D. Receptor surface models. 2. Application to Quantitative Structure Activity Relationship Studies. *J. Med. Chem.* **1995**, *38*, 2091–2102.
(16) Kearsely, S. K.; Smith, G. M. An alternative method for the alignment of molecular structures: Maximizing electrostatic and steric overlap. *Tetrahedron Comput. Methodol.* **1990**, *3*, 615–633.
(17) Dammkoehler, R. A.; Karasak, S. F.; Berkely Shands, E. F.; Marshall, G. R. Constrained search of conformation hyperspace. *J. Comput.-Aided Mol. Design* **1989**, *3*, 3–21.
(18) Perkins, T. D.; Deam, P. M. An exploration of a novel strategy for superimposing several flexible molecules. *J. Comput.-Aided Mol. Design* **1993**, *7*, 155–172.
(19) Blaney, J. M.; Dixon, J. S. A good ligand is hard to find: Automatic docking methods. *Perspect. Drug Disc. Design* **1993**, *1*, 301–319.
(20) Martin, Y. C.; Bures, M. G.; Danahar, E. A.; DeLazzar, J.; Lico, I.; Pavlik, P. A. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. *J. Comput.-Aided Mol. Design* **1993**, *7*, 83.
(21) Hoffmann, R.; Langer, T. Use of the CATALYST program as a new alignment tool for 3D QSAR. In *Proceedings of the 10th European Symposium on Structure-Activity Relationships: QSAR and Molecular Modeling*; Prous Science Publishers: Barcelona, Spain, 1994.
(22) Barnum, D.; Greene, J.; Smellie, A. Identification of Common Functional Configurations. *J. Chem. Inf. Comput. Sci.* In press.
(23) Marshall, G. R. Binding Site Modeling of Unknown Receptors. In *3D QSAR in Drug Design: Theory Methods and Applications*; Kubinyi, H., Ed.; Escom: Leiden, 1993; pp 80–116.
(24) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857.
(25) Wyvill, G.; McPheeters, C.; Wyvill, B. Data Structures for Soft Objects. *The Visual Computer* **1986**, *2*, 227–234.
(26) Lorensen, W. E.; Cline, H. E. Marching Cubes: A High Resolution 3-D Surface Construction Algorithm. *Comput. Graphics (Proc. SIGGRAPH, 21)* **1987**, *4*, 163–169.
(27) Heiden, W.; Schlenkrich, M.; Brickmann, J. Triangulation algorithms for the representation of molecular surface properties. *J. Comput.-Aided Mol. Design* **1990**, *4*, 225–269.
(28) Teschner, M.; Henn, C.; Vollhardt, H.; Reiling, S.; Brickmann, J. Texture Mapping: A new tool for molecular graphics. *J. Mol. Graphics* **1994**, *12*, 98–105.
(29) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219.
(30) Appelt, K. Crystal Structures of HIV-1 protease-inhibitor complexes. *Perspect. Drug Disc. Design* **1993**, *1*, 23–48.
(31) Hopfinger, A. J.; Nakata, Y.; Max, N. Quantitative structure-activity relationship of anthracycline antitumor activity and cardiac toxicity based upon intercalation calculations. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981; p 431.
(32) Hopfinger, A. J.; Kawakami, Y. QSAR analysis of a set of benzothiopyranoindazole anti-cancer analogs based on their DNA intercalation properties as determined by molecular dynamics simulation. *Anti-Cancer Drug Design* **1992**, *7*, 203–217.
(33) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
(34) Available from Molecular Simulations Incorporated, 16 New England Executive Park, Burlington, MA 94107.
(35) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: A generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.
(36) Davies, K. E.; Murall, N. W. How accurate does a force field need to be? *Comput. Chem.* **1989**, *13*, 149–156.
(37) Halgren, T. A. Merck Molecular Force Field: I-IV. *J. Comput. Chem.* Submitted.
(38) Weiner, J.; Kollman, P. A.; Nguyen, T.; Case, A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **1986**, *7*, 230–252.

JM940814O