Closely related to selectivity is the ability of a pharmacophore to distinguish actives from inactives when the two classes of compounds are structurally similar and/or come from the same structure/activity series. A tacit assumption here is that the inactive compounds fail to bind because they lack one or more key features not because of steric clashes with the receptor, poor solubility, or other factors that cannot be explained by the pharmacophore model itself. SCAMPI[42] builds pharmacophore models, feature by feature, incorporating a Student's *t*-test on the separation of actives and inactives directly into the feature selection process. APEX-3D[31] identifies pharmacophores common to the most active compounds using clique detection and then employs Bayesian statistics to eliminate pharmacophores whose frequencies in the active and inactive populations are not sufficiently different. CATALYST/HYPOGEN[6] constructs a series of HIPHOP pharmacophore models from the most active compounds in a training set, eliminates those that are found in more than half of the inactives, and then refines each model by adding, subtracting, and moving features so the degree to which each compound fits the pharmacophore correlates with experimental activity. Thus all of these methods use activity to drive development of pharmacophore models that can distinguish actives and inactives, while discarding those that do not. PHASE[7] uses a more passive approach, scoring each common pharmacophore model according to how well it matches a set of known inactives but ultimately deferring to the user the decision about which pharmacophores are most relevant.

Reasonable superposition of the ligand features that map to a common pharmacophore model is normally a given, but this does not guarantee that the ligand structures themselves or their overall superpositions will be satisfactory. A number of pharmacophore development approaches[7,11,12,29,60] incorporate a procedure for eliminating or penalizing unrealistic high-energy structures, although methods that accept external conformers[7,9,10,31] will accomplish essentially the same objective if an appropriate energy filter is applied to the conformers before they are supplied. To achieve consensus in the alignment of chemical features that do not contribute to the pharmacophore, the scoring process may favor conformers that match the pharmacophore and yield superior overlap of molecular volumes throughout each ligand structure.[7,11] Volumes may be distinguished by atom type to help ensure that chemically similar fragments in different ligands are superimposed.[7]

RECEPTOR-BASED PHARMACOPHORE MODELS

Previous sections focused on the identification of common feature pharmacophores within flexible ligand structures, which is often a necessary exercise in the absence of crystallographic data. But when explicit knowledge of the receptor binding site is available, it can be a tremendous advantage in pharmacophore model development. Though it is certainly possible to visually inspect a ligand/receptor complex to identify key interactions, and manually construct a pharmacophore model that encodes those interactions, automated procedures to achieve this task are in high demand. A number of important receptor-based pharmacophore techniques have been reported involving ligand docking,[61,62] fragment docking,[63,64] and molecular dynamics simulations,[65,66] but this section is concerned primarily with methodologies that provide an alternative to what are essentially products of structure-based and de novo design.

A fundamental step in developing a receptor-based pharmacophore model is an analysis of the binding site to identify potential interaction points. In structure-based focusing,[67] a sphere with user-adjustable location and radius is used to mark key residues in the binding site, and a LUDI[68] interaction map is generated to describe favorable interactions in which a ligand is expected to engage. The interaction map is translated to an interaction model, which consists of a set of complementary points in the binding pocket, representing possible locations of pharmacophore features on the ligand. A user-defined density controls the number of points created, but it is usually quite large, so hierarchical clustering is performed to select a smaller number of representative points, typically on the order of a dozen. Normally, this is still too many interaction points for a single ligand, so a series of pharmacophore models containing subsets of the representative features is constructed, with restrictions on minimum and maximum separations between points. Excluded volumes (see next section) are added to each model to represent the receptor surface, and a 3D database of known actives is searched to determine which pharmacophore models are most frequently matched.

LIGANDSCOUT[69] takes a somewhat more direct approach, deriving a pharmacophore model from a single ligand/receptor complex. After perceiving hybridization, unsaturated bonds, and aromatic rings, the resulting ligand structure is analyzed for the presence of features encoding hydrogen bonding, hydrophobic character, and charge transfer. Features are mapped in general accordance with CATALYST rules, with customization to allow certain atoms to be associated with more than one type of pharmacophoric feature. Whether a feature is incorporated into the pharmacophore model depends on its location relative to a complementary site on the receptor. For example, a hydrogen bond donor is included if the associated heavy atom X is 2.5–3.8Å from an acceptor atom Y in the receptor and the X−H−Y angle is within 34° of colinearity. Incorporation of hydrophobic and ionic features, which are nondirectional, depends only on a user-defined distance range from a compatible interaction site on the receptor. If the receptor site is hydrophobic, a steric constraint is added to the pharmacophore model by creating a series of excluded volume spheres in the vicinity. A LIGANDSCOUT model normally needs some manual refinement (removal of features,