search of pharmacophore space is not done, so any number of valid solutions can be missed.

It is important to recognize that all of the aforementioned methods are designed to operate on relatively small numbers of high-affinity ligands, yet there is sometimes a need to make sense of data arising from hundreds or thousands of heterogeneous compounds of varying activity. In the mid- to late 1990s, recursive partitioning[36] was increasingly recognized for its ability to identify patterns in complex chemical data,[37–41] and pharmacophore-based analysis of large data sets was a natural application for this technique. SCAMPI (Statistical Classification of Activities of Molecules for Pharmacophore Identification)[42] leverages the power of recursive partitioning to construct decision trees that subdivide compounds according to their activities and the pharmacophores they contain. Each node in a tree adds a feature to a pharmacophore from a parent node, and compounds are split into left and right child branches according to whether they contain the proposed pharmacophore. Experimental activities are incorporated into the process to ensure that a given splitting condition yields a statistically significant separation of compounds into less active and more active groups. Addition of features to a particular branch terminates when further separation by activity cannot be achieved. A SCAMPI decision tree produces a number of pharmacophores, usually containing two or three features, which rationalize the observed activities and suggest possible binding modes.

More recently developed pharmacophore perception methods include GALAHAD[12] and PHASE.[7] GALAHAD improves and extends methodologies introduced in GASP by incorporating a multiobjective Pareto scoring function to balance pharmacophore consensus, shape consensus, and conformational energy, while relaxing the requirement that every ligand match all features in the pharmacophore. PHASE provides exhaustive exploration of common pharmacophore space by way of a novel distance-based partitioning algorithm and takes an eclectic approach to scoring with user-adjustable terms to optimize alignment and orientation of features, shape overlap, pharmacophore selectivity, reference ligand conformational energy, and reference ligand activity.

Finally, it should be noted that the past three decades have seen the emergence of numerous other important pharmacophore-based systems, including ALADDIN,[43] DANTE,[21] CAVEAT,[44] APEX-3D,[31] and CHEM-X,[18] to name just a few. Many of the principles on which these systems are based are used routinely in other software packages and will be covered in subsequent sections of this chapter.

PHARMACOPHORE MODEL DEVELOPMENT

Pharmacophore models are created using a variety of methods and workflows, including manual construction, automated perception from ligand structure alone, and receptor-based inference from a crystallographic structure. The particular method or workflow used depends on any number of factors, including the amount and quality of available experimental data, computational resources, and the ultimate goals and expectations placed on the pharmacophore model itself. The next several sections are concerned with the methodological details and applicability of these different approaches to pharmacophore model development.

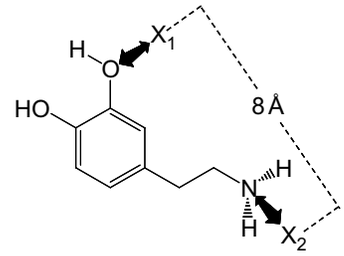

Figure 9.2. Dopamine D_2 agonist pharmacophore model from Seeman et al.[45] Two receptor sites, X_1 and X_2, separated by about 8Å, form hydrogen bonds to an acceptor and a basic center, respectively, on the ligand.

MANUAL CONSTRUCTION

The simplest and probably the most widely employed method to create a pharmacophore model is to construct it by hand, using the structure of a known active or based on general characteristics of known actives. Figure 9.2 illustrates the classic Seeman model[45] of dopamine D_2 agonists, which consists of an aromatic ring bearing a hydrogen bond acceptor and a basic nitrogen that form hydrogen bonds with complementary receptor sites X_1 and X_2, respectively, which are separated by about 8Å. In practical applications, such as 3D searching, one would normally designate some variability in the distances between pharmacophore features, or positional tolerances that constrain how far a matching feature may deviate from the corresponding feature in the model after an alignment is performed. Depending on the software used, it may also be possible to define tolerances on the orientation of features, such as a range in the allowed angles between the hydrogen bond acceptor axis and the plane of the ring.

A manually constructed pharmacophore can be quite advantageous, particularly if it's derived from the x-ray structure of a ligand or from a ligand with a rigid backbone. In either case, the locations of pharmacophore features are essentially pinned down, so one of the biggest uncertainties in pharmacophore model development, conformational flexibility, is eliminated. There is still the question of the particular features to incorporate in the model, which is not always easy to infer without additional information, such as the structure of a ligand/receptor complex, activities from a well-designed SAR series, or data from mutagenesis experiments.

AUTOMATED PERCEPTION FROM LIGAND STRUCTURE

A review of the available methods for automated perception of common pharmacophores has already been provided,