pharmacophore generation varies depending on the training set. The use of these programs for pharmacophore generations was recently reviewed in detail (103). Here we focus on common features of these programs. All programs use algorithms that identify common pharmacophore features in the training set molecules; they use scoring functions to rank the identified pharmacophores. The following features are identified in each molecule: hydrogen-bond donors, hydrogen-bond acceptors, negative and positive charge centers, and surface accessible hydrophobic regions that can be aliphatic, aromatic, or nonspecific. Most of the programs consider ligand flexibility when generating pharmacophores because compounds might not bind to the protein in the minimum energy conformation.

2.3.4.4 Receptor-Based pharmacophore Generation. If the 3D structure of a receptor is known, a pharmacophore model can be derived based on the receptor active site. Biochemical data can be used for identifying key residues that are important for substrate and/or inhibitor binding. This information can be used for building pharmacophores targeting the region defined by key residues or for choosing among pharmacophores generated by an automated program. This can greatly improve the chance of finding small molecules that inhibit the protein because the search is focused on a region of the binding site that is crucial for binding substrates and inhibitors. Many ligands bind to proteins through nonbonded interactions such as hydrogen bonds and hydrophobic interactions. Programs such as LUDI (104–106) or POCKET (107) can use the structure of the protein to generate interaction sites or grids to characterize favorable positions that ligand atoms should occupy. Four types of interaction sites are characterized: hydrogen-bond donors, hydrogen-bond acceptors, and hydrophobic groups that can be lipophilic-aliphatic or lipophilic-aromatic. LUDI-generated interaction maps for Cerius^2 Strudure-Based Focusing (108) do not differentiate between aliphatic and aromatic interaction sites. This is based on the observation by Burley and Petsko (109) that, besides aromatic side chains, aliphatic and aromatic side chains also pack closely to form the hydrophobic core of proteins. Because proteins are not rigid, Carlson et al. (110) proposed using molecular dynamics simulation for generating a set of diverse protein conformations to include protein flexibility in the pharmacophore development. In this case distance ranges between pharmacophores are obtained by examining several conformations of the protein. This technique is similar to the one used for the generation of flexible pharmacophores (Fig. 6.16), based on active compounds, when several conformations of the compound and/or many compounds are considered for pharmacophore mapping.

2.3.5 Pharmacophore-Based Virtual Screening. Pharmacophore-based virtual screening is the process of matching atoms and/or functional groups and the geometric relations between them to the pharmacophore in the query. Examples of programs that perform pharmacophore-based searches are 3Dsearch (111), Aladdin (53), UNITY (112), MACCS-3D (113), Catalyst (114), and ROCS (102). There are also web-based applications (115,116) that can perform pharmacophore searches. Usually pharmacophore-based searches are done in two steps. First, the software checks whether the compound has the atom types and/or functional groups required by the pharmacophore; then it checks whether the spatial arrangement of these elements matches the query. The fastest approach used in the matching step is considering rigid compounds. Because molecules that are not rigid might have a conformation that matches the pharmacophore, flexibility of the ligands should be considered. Flexible 3D searches identify a higher number of hits than rigid searches do (117). However, flexible searches are more time consuming than rigid ones. There are two main approaches for including conformational flexibility into the search: one is to generate a user-defined number of representative conformations for each molecule when the database is created; the other is to generate conformations during the search. By use of the first approach, any rigid search program can be used for doing a flexible search; however, generating the database takes more time and disk space. The second approach gives more flexibility to the user, given that a larger number of conformations can be generated for each