pattern and candidate structures. Although rigorous, this approach was tedious and required optimization. Lesk (203) devised a method that used the geometric attributes of the query to screen potential candidates. Similarly, Jakes and Willett (204) proposed that screens based on interatomic distances and atom types could considerably augment search efficiency. Furthermore, Jakes et al. (205) showed that methods widely used in two-dimensional structure retrieval could be applied to three-dimensional searches, to remove the vast majority of compounds before more rigorous comparisons. This was validated in test searches against a subset of the CSD. This concept was furthered by Sheridan et al. (200), who included screens based on aromaticity, hybridization, connectivity, charge, position of lone pairs, and centers of mass of rings. To contain this wealth of information, an inverted bit map [the presence or absence of a feature is encoded as a 1 or 0 (bit) at a particular location in a "keyword"] was employed for highly efficient screening, hundreds of thousands of compounds in minutes.

Similar database searching methods have been incorporated into a number of current database searching systems. Programs such as CAVEAT (206), ALADDIN (Abbott) (207), 3DSEARCH (Lederle) (208), MACCS-3D (209), CHEM-X (210), UNITY (211), and others contain considerable functionality useful for such an approach. CAVEAT (206) is designed to assist a chemist in identifying cyclic structures that could serve as the foundation for novel compounds. In particular, it allows an investigator to rapidly search structural databases for compounds containing substituent bonds that satisfy a specific geometric relationship. ALADDIN (207), 3DSEARCH (208), MACCS-3D (209), and CHEM-X (210) are similar, in that geometric relationships between various user-defined atomic components can be used as a query to retrieve matching structures. Features have been included to allow the user to delineate molecular characteristics (atom type, bond angles, torsional constraints, etc.) to ensure the retrieval of relevant compounds. Additional constraints have been incorporated into 3DSEARCH (208) and ALADDIN (207), including the consideration of retrieved ligand-receptor volume complementarity. Furthermore, CHEM-X (210) performs a rule-based conformational search on each structure in the database to account for conformational flexibility. For a comprehensive review of three-dimensional chemical database searching, see Martin et al. (212,213).

Pharmaceutical companies have developed three-dimensional databases for their compound files to help prioritize candidates for screening (210, 214). An essential component in such a system is a method for assessing similarity (212,215). Because most compound databases were entered as two-dimensional structures, this has required conversion to a three-dimensional format. Programs have proved (197–199, 216) useful in generating plausible three-dimensional structures from the connectivity data, as reviewed by Sadowski and Gasteiger (217). Because of the inherent flexibility in most compounds, the use of a single conformation to represent the three-dimensional potential for interaction of a molecule is a clear limitation. Development of three-dimensional databases with a compact, coded representation of the conformational states available to each compound is a logical next step. Efficient use of such a database requires methods for evaluating three-dimensional similarities. In addition to identification of compounds that can present an appropriate three-dimensional pattern, compounds must also fit within the receptor cavity. Based on a shape-matching algorithm, Sheridan et al. (200) screened candidate compounds to select those whose volumes would fit within the combined volumes of known active compounds. Previously, this group used (218) the same algorithm to help identify potential ligands for papain and carbonic anhydrase, by screening compounds from the CSD. Screening of the active site of HIV protease identified (219) haloperidol (Fig. 3.20) as an inhibitor of the enzyme and provided a novel chemical lead for further investigation. Burt and Richards (220) introduced flexible fitting of molecules to a target structure, with assessment of molecular similarity as a means of dealing with the conformational problem.

The use of preliminary screens can eliminate the vast majority of compounds before more rigorous, and computationally demanding, pattern-matching comparisons (212,213).