a possible match to either. All the distances greater than 20 Å are assigned into the last bin.

3 QSAR MODELING APPROACHES

3.1 3D-QSAR

Two original 3D-QSAR methods, CoMFA (40) and GRID (110), were developed almost simultaneously in the mid- to late-1980s (9). Since its introduction, the CoMFA approach has rapidly become one of the most popular methods of QSAR. Over the years, this approach has been applied to a wide variety of receptor and enzyme ligands [many reviews appeared in a recent monograph (10)]. Undoubtedly, the further development of this and related methods is of great importance and interest to many scientists working in the area of rational drug design.

CoMFA methodology is based on the assumption that because, in most cases, the drug-receptor interactions are noncovalent, the changes in the biological activities or binding affinities of sample compounds correlate with changes in the steric and electrostatic fields of these molecules. In a standard CoMFA procedure, all molecules under investigation are first structurally aligned, and the steric and electrostatic fields around them are sampled with probe atoms, usually sp^3 carbon with a +1 charge, on a rectangular grid that encompasses aligned molecules. The results of the field evaluation in every grid point for every molecule in the data set are placed in the CoMFA QSAR table, which therefore contains thousands of columns (Fig. 2.5). The analysis of this table by the means of standard multiple regression is practically impossible; however, the application of special multivariate statistical analysis routines, such as PLS analysis and LOO cross-validation ensures the statistical significance of the final CoMFA equation. The outcome from this procedure is a cross-validated correlation coefficient R^2 (q^2), which is calculated according to the formula

$$q^2 = 1 - \frac{\Sigma (y_i - \hat{y}_i)^2}{\Sigma (y_i - \bar{y})^2}, \qquad (2.1)$$

where y_i, $\hat{y}_i$, and $\bar{y}$ are the actual, estimated, and averaged (over the entire data set) activities, respectively. The summations in Equation 2.1 are performed over all compounds, which are used to build a model for the training set. The statistical meaning of the q^2 is different from that of the conventional r^2: a q^2 value greater than 0.3 is often considered significant (111).

Despite obviously successful and growing application of CoMFA in molecular design, several problems intrinsic to this methodology have persisted. Studies revealed that CoMFA results can be extremely sensitive to a number of factors, such as alignment rules, overall orientation, lattice placement, step size, and probe atom type (40, 75, 112–114). The problem of three-dimensional alignment has been the most notorious among others. Even with the development of automated or semiautomated alignment protocols such as the Active Analog Approach (108, 115) or DISCO (116) and the opportunity to use, in some cases, the structural information about the target receptor (112, 117) to align molecules, in general there is no standard recipe as to how to align all molecules under consideration in a unique and unambiguous fashion. A QSAR analysis of 60 acetylcholinesterase inhibitors (117) is particularly illustrative with respect to this point. In that study, the combination of structure-based alignment and CoMFA was employed to obtain a QSAR model for 60 chemically diverse inhibitors of acetylcholinesterase (AChE). The great structural diversity of the AChE inhibitors, ranging from choline to decamethonium, made it practically impossible to structurally align all the inhibitors in any unbiased way and generate a unique three-dimensional pharmacophore. X-ray crystallographic analysis of AChE from *Torpedo californica* (EC 3.1.1.7) (118), followed by X-ray determination of the complexes of the enzyme with three structurally diverse inhibitors, tacrine, edrophonium, and decamethonium (119), provided crucial information with respect to the orientation of these inhibitors in the active site of the enzyme. The crystallographic data indicated that each of the three inhibitors had a unique binding orientation in the active site of the enzyme (Fig. 2.6). Their natural structural alignment would probably never have been predicted by any of the existing automated algorithms for ligand align-