Comparative Molecular Field Analysis

Comparative molecular field analysis (CoMFA) is another promising approach developed in recent years for QSAR study. CoMFA is a molecular modeling technique for the determination of molecular steric and electrostatic force fields (Tripos, 1992). It has been successfully used in deriving molecular descriptors for prediction of the bioactivity of steroids (Cramer et al., 1988), molecular flux through a polymer membrane (Liu and Matheson, 1994), and metabolism and cytochrome p450 enzyme activities (Long and Walker, 2003).

To compare the steric effects of different molecules in the CoMFA study, molecules must be displaced in space according to a set of specified rules (Cramer et al., 1988; Tripos, 1992; Liu and Matheson, 1994). In the study of predicting the solubility in isopropyl alcohol, the optimal molecular models for aromatic compounds were aligned in reference to the benzene molecule with assigned carbon atom positions according to the following rules:

1. For the monosubstituted compounds, the aromatic carbon atom connected to the functional group was selected as the prime atom and was assigned as position 1.
2. For multisubstituted compounds, the ring carbon atom connected to the functional group with the greatest effects on solubility, based on fragmental coefficients determined in a previous study (Hu, 1990), was selected as the prime atom and assigned as position 1.
3. All other six-membered rings were aligned with the reference benzene ring in the way as described earlier, with the added provision that alkyl groups and bulky groups were aligned in the direction of carbon atoms 4 and 5 on the reference molecule.

Following the fitting of each molecule to the reference benzene using the least squares method, electrostatic and steric fields of a molecule were calculated in a region ranging from −14.0 to +14.0 Å along the *x*- and *y*-axes and from −10 to +10 Å along the *z*-axis. A probe atom of sp^3 carbon with a charge of +1 was placed in the defined region, and the field energy was calculated with the probe atom moving in a spacing of 2 Å along the axes. The resultant steric and electrostatic field energies, together with the melting point (mp) and an intramolecular hydrogen-bond indicator (IHB), were then used for QSSR study, using a partial least squares (PLS) method in the SYBYL package. It should be noted that, owing to the nature of molecular field analysis, there would be hundreds or thousands of descriptors for the molecular steric and electrostatic force fields (Table 3.5), and it is inconvenient and unnecessary to write down the QSSR equation in a CoMFA study since these descriptors are generated and directly used to predict targeted properties using the same software package. Statistical results for the final models correlating the mole fraction solubility in isopropyl alcohol with the desired parameters for 60 aromatic and heteroaromatic crystalline compounds are SD = 0.243, $r^2 = 0.942$, and $F = 146$. The contribution of each parameter in the two CoMFA models is listed in Table 3.5. The experimental and calculated mole fraction solubilities by Model 2 using CoMFA, mp, and IHB as predictors are listed in Table 3.6.

TABLE 3.5
Relative Contribution of Predictors in CoMFA for Solubility

	Model 1		Model 2	
Predictors	**Normal Coefficient**	**Contributions (%)**	**Normal Coefficient**	**Contributions (%)**
CoMFA (864 vars) (Steric)	1.469	40.1	0.942	31.0
CoMFA (864 vars) (Electrostatic)	1.575	43.0	1.092	35.9
100/MP	0.622	17.0	0.656	21.6
1000*IHB			0.351	11.5