

Electrostatics in Chemistry

4. Electrostatic Models for Weak Molecular Complexation

Shridhar R Gadre and K Babu

Introduction

We have seen in the earlier parts of this series that many important facets of structure, bonding and energetics in supramolecular chemistry can be understood on the basis of simple electrostatic concepts. The basic rules of the ‘electrostatic game’ were presented in Part 1¹. Features of the atomic and molecular electrostatic potentials (MESP) were described with examples in the next two parts. In particular, the role of topography as a tool to extract essential features of MESP was highlighted in Part 3. We now discuss how electrostatic effects can be modeled qualitatively as well as semi-quantitatively to rationalize and predict many subtle interactions in chemistry.

Non-covalent interactions play a crucial role in many chemical and biological processes ranging from chromatography to enzyme catalysis. Given a pair of molecules likely to form a weak complex, one would like to be able to predict the strength of the interaction as well as the preferred geometry of the complex. On the other hand elaborate quantum chemical *ab initio* methods, which reproduce non-covalent interactions sufficiently accurately, are too time-consuming to be used routinely on large systems of chemical interest. The word ‘*ab initio*’ in Latin means ‘from the beginning’, i.e. these methods rigorously solve the Schrödinger equation from scratch, without using any empirical parameters. The only experimental data used are the basic physical constants such as Planck’s constant, velocity of light, electronic charge, etc. However, several approximations are applied to make it practically feasible. The most popular *ab initio* methods, as of today, are those based on Gaussian functions, wherein the atomic and molecular orbitals are represented as

1. Basic Principles, *Resonance*, Vol.4, No.2, 11–19, 1999.

2. Electrostatic Potentials of Atoms, Ions and Molecules, *Resonance*, Vol.4, No.5, 40–51, 1999.

3. Molecular Electrostatic Potential: Visualization and Topography, Vol.4, No.7, 14–23, 1999.



**Box 1. Quantum
Chemistry Nobel Prize
for 1998.**

Popularising gaussian-based ab initio methods through a set of elaborately developed Fortran codes called 'GAUSSIAN' yielded John Pople his Nobel Prize for Chemistry in the year 1999 shared by Walter Kohn for his contribution to the density functional theory (DFT). The GAUSSIAN package now includes a DFT option apart from other quantum-chemical methods for effecting reliable molecular calculations. See [2] for an introduction to ab initio methods and DFT.

combinations of standard Gaussian functions. This makes the method computationally feasible due to inherent properties of Gaussian functions (see *Box 1*).

Since electrostatics is often the dominant factor in weak complexation, simpler models, which reproduce the magnitude of electrostatic interactions, may prove to be quite powerful. Developing and testing such models represents an active area of research in itself. In this part, some of the more widely used models are described. Some specific examples are considered to highlight the power of these models.

Qualitative Model of Legon and Millen (L-M)

In 1980's, Legon and Millen proposed the following set of qualitative empirical rules that help explain the interactions in complexes of type B...HX, observed experimentally. Here, B denotes an electron donor molecule viz. one possessing at least one pair of non-bonding or π electrons, and HX a molecule with electropositive hydrogen.

- i) The electron pair of B involved in hydrogen bonding will be collinear with the H-X bond.
- ii) If π -electron pair of B is involved in hydrogen bonding, then the H-X bond would lie perpendicular to the nodal plane.
- iii) Non-bonding electron pairs are preferred over π -electrons for hydrogen bond formation.

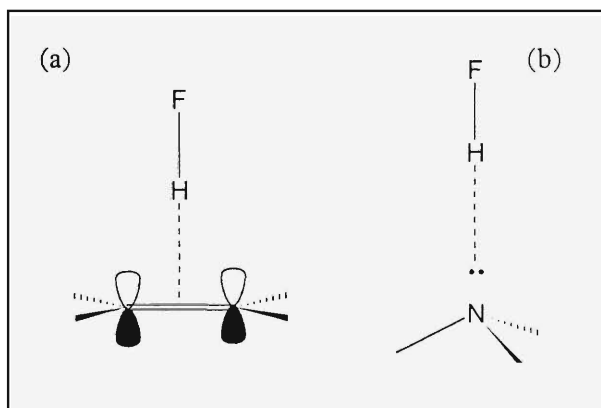


Figure 1. Qualitative prediction for structures of (a) ethylene...HF and (b) NH₃...HF complexes based on L-M rules.

With this set of rules one could successfully predict the directionality of hydrogen bonds in complexes. *Figure 1* shows the structures of ethylene...hydrogen fluoride and ammonia ... hydrogen fluoride complexes as predicted by the rules 1 and 2 above. The drawback of L-M model is that it does not say anything about the energetics or relative strengths of hydrogen bonds.

Quantitative Models

Buckingham and Fowler (B-F) were the first to devise a quantitative method for predicting structures of weakly interacting binary complexes. In this model, the individual geometries and wave functions of the two interacting species are obtained at some standard *ab initio* level by solving the Schrödinger equation. This wave function is then employed to obtain distributed multipole moments (DMM). In order to achieve this, the $1/|\mathbf{r}' - \mathbf{r}|$ term involved in the definition of MESP (see (7) in Part 3 of this series) is expanded in a Taylor's series and the first few significant terms are retained. These terms individually resemble the potentials due to a single point charge, a dipole, a quadrupole, etc., respectively and are in general termed as multipoles. The Coulombic interaction energy between the two molecules for a given relative orientation is then obtained as

$$U = \sum_{a \in A} \sum_{b \in B} \frac{1}{4\pi\epsilon_0} \left\{ \frac{q^a q^b}{R} - \frac{R_x (\mu_x^b q^a - \mu_x^a q^b)}{R^3} + \frac{(3R_x R_y - R^2 \delta_{xy})(q^a \theta_{xy}^b + q^b \theta_{xy}^a - 3\mu_x^a \mu_y^b)}{3R^5} + \dots \right\}$$

where R defines the vector joining atoms 'a' of molecule A and 'b' of molecule B. q , μ and θ are the zeroth (charge), first (dipole) and second (quadrupole) order moments located at the atomic positions of the molecules. Subscripts x , y etc. stand for the Cartesian components. The interaction energy U is then minimized by rolling one molecule over the other, while internal geometry of each species is kept fixed. In order that the molecules do not penetrate each other, appropriate molecular sur-

Box 2. van der Waals (vdW) Radii

The Dutch physicist J H van der Waals suggested, in order to explain certain properties of gases, that molecules have a well-defined size. Later, appropriate radii have been assigned to atoms in molecules (by Pauling and others) using the contact distances from crystal structure data. These 'van der Waals' radii extracted from the contact distances in crystals typically lie between 1.0 and 2.2 Å and are approximately 0.8 Å longer than the respective covalent radii. See *Table 1* for a tabulation of vdW radii.



Atom	vdW radius (Å)
H	1.10
N	1.50
O	1.40
F	1.35
P	1.90
S	1.85
Cl	1.80

Table 1. van der Waals (vdW) radii of some common atoms.

faces are constructed employing atomic van der Waals radii (see *Box 2*). This model was tested on several binary complexes and has been found to be quite successful.

Recently, Stone and co-workers have developed a similar modeling program called 'ORIENT', which also makes use of the DMM's for getting the electrostatic interaction energy. This program also incorporates the dispersion and repulsion energies of the form

$$U_{\text{exp-6}} = \sum_{i,j} k \exp[-\alpha_{ij}(R_{ij} - \rho_{ij})] - \frac{c_{ij}^6}{R_{ij}^6},$$

where, α , ρ and c are empirical parameters.

The 'Molecular Mechanics for Clusters' model developed by Dykstra also makes use of the DMM's for obtaining the Coulombic interaction energy. In addition to this, Dykstra used Lennard-Jones potential term to represent non-Coulombic interactions, which is expressed as

$$V_{A-B} = \sum_i^A \sum_j^B \left(\frac{d_i d_j}{R_{ij}^{12}} - \frac{c_i c_j}{R_{ij}^6} \right),$$

with R_{ij} being the distance between a pair of atoms i of A and j of B. Variables c and d can be parametrized for specific types of systems. The form of 6-12 potential in V_{A-B} above is popularly employed in many modeling programs.

Alhambra, Luque and Orozco developed a specific model for hydration of molecules. The molecular solvation potential (MSP) defined by them is primarily Coulombic in nature, augmented with a 6-12 potential which also includes empirical parameters to represent the atomic hardness. As in the previous models, the VMSP here is optimized by rolling the water molecules over the solute molecules.

All these models do have their own limitations. Models like B-F invariably make use of DMM's and cannot be easily applied to clusters with large numbers of molecules, as that would be computationally expensive. Solvation model of Alhambra and



others is the most successful of the models mentioned above. This, in the present form, is applicable only to hydration of molecules. For a detailed description of these models, see [1].

The EPIC Model

An electrostatics-based model, EPIC (electrostatic potential for intermolecular complexation) was developed a few years ago by the authors' group. This is designed to make use of the *ab initio* wave functions of molecules in the complex. The *ab initio* electrostatic interaction energy between two molecules is obtained on integrating the product of charge distribution of one species with the complete MESP distribution due to the other. Such a calculation would be time-consuming since it involves a three dimensional quadrature. Therefore, in EPIC we use a set of point charges, derived from the fit of MESP in the outer region of the molecule, to represent the charge distribution. With this, a good approximation to the *ab initio* (Hartree–Fock) interaction energy can be obtained as a summation of discrete point charges on one molecule multiplied by the MESP at those points due to the other molecule, viz.

$$E_{\text{int}} = 1/2 \left\{ \sum_i V_{A,i} q_{B,i} + \sum_j V_{B,j} q_{A,j} \right\}$$

The second molecule is then moved around the first, without changing the internal geometry of either of them, to minimize the value of the E_{int} (more negative the value of E_{int} , greater will be the attractive electrostatic interaction between the two molecules). This process can be referred to as 'docking'. To prevent the molecules from running into each other, a strategy similar to the one in the B–F model above is employed. The docking process in the search for the minimum energy geometry employs standard optimization algorithms such as cyclic relaxation, simulated annealing, genetic algorithm, etc. This model has been successfully tested on several weak binary complexes. The geometries predicted by EPIC are quite close to those obtained with *ab initio* calculations.



Recently, this model has been extended to deal with complexes of more than two molecules. In such cases, a few more assumptions need to be made. The MESP at any point in the complex is taken to be the additive sum of MESP's due to all the individual molecules in the complex at their respective geometries. Though this looks like a gross approximation, it does result in fair prediction. This is due to the fact that complex formation changes the charge distribution of the individual molecules only in the regions where the actual binding takes place and such sites are not accessed by any new molecule binding with the complex. This adds to the simplicity of the model as the wave functions of individual species can still be used for MESP evaluation of the complex. Thus the total interaction energy will be summation of all the E_{int} terms, one for each pair, and is given by

$$E_{\text{int}} = 1/2 \left\{ \sum_A^N \sum_{\substack{B \neq A \\ i \in B}}^N V_{A,i} q_{B,i} \right\}$$

Hydration of Small Molecules via EPIC Model

We have chosen three types of systems, polar (formaldehyde), totally non-polar (cyclopropane) and a system with both polar and non-polar ends (methanol).

Formaldehyde is well known to science students in the form of formalin, a 40% solution of formaldehyde in water used for preserving biological specimens and is the smallest analogue in the family of aldehydes. The first step for studying intermolecular interactions using MESP would be to obtain the topographical features of MESP for the systems involved. The MESP-derived point charges of all the molecules considered here are given in *Table 2*. Water has two MESP minima at positions corresponding to the conventional 'rabbit ears' (lone pairs) of oxygen. Formaldehyde also shows two such CPs (refer to Part 3 of this series for MESP features of these molecules).

Several guess geometries for the complex $\text{HCHO} \dots \text{H}_2\text{O}$ were

generated keeping in mind that the hydrogen of one molecule should be put closer to the most negative valued MESP CPs of the other. This is to make use of the lock-and-key mechanism operative on the MESP of the binding molecules. MESP minima in formaldehyde are stronger than those in water. Hence, hydrogen bond between H of water and O of HCHO will be stronger than that between H of HCHO and O of water. When a structure with both these bonds was generated and docked with EPIC the latter type of hydrogen bond was found to be longer than the former. Subsequent Hartree–Fock optimization of the complex led to the structure in *Figure 2c* which is very similar to the one predicted by EPIC (see *Figure 2a*). A few other possible structures were also generated and docked with EPIC followed by *ab initio* optimization. But the interaction energy is most negative for the above mentioned structure. The same procedure is followed for the HCHO...2H₂O complex and the structure predicted by EPIC is again not different from that obtained by complete *ab initio* optimization (see *Figures 2e* and *f*).

Figure 2 depicts the isosurface (surface of constant function value) of HCHO...*n*H₂O (*n*=1 and 2) with a view to bringing out the features of electrostatic complementarity. The MESP isosurfaces of value -126 (purple) and 788 kJ mol⁻¹ (blue) of the individual molecule at the EPIC-optimized orientation are shown in *Figure 2a*. The corresponding MESP isosurfaces for the EPIC and *ab initio* optimized geometries respectively, are shown in *Figures 2b* and *2c*. The lock and key function of the molecule is shown in *Figure 2a*. *Figures 2b* and *2c* bring out the loss of negative valued isosurface of value -126 kJ mol⁻¹ in the bonding region. The EPIC MESP distribution seems to mimic the corresponding *ab initio* one quite well. Similar features are exhibited by HCHO...2H₂O system in *Figures 2d* and *2f*, respectively. Employing similar methodology, further hydration with HCHO...6H₂O has been investigated.

Molecule	Atom	Charge
Water	O	-0.840
	H	0.420
Formaldehyde	O	-0.492
	C	0.466
	H	0.013
Cyclopropane	C	-0.380
	H	0.190
Methanol	H ^a	0.448
	O	-0.745
	C	0.346
	H ^b	0.028
	H ^c	-0.039

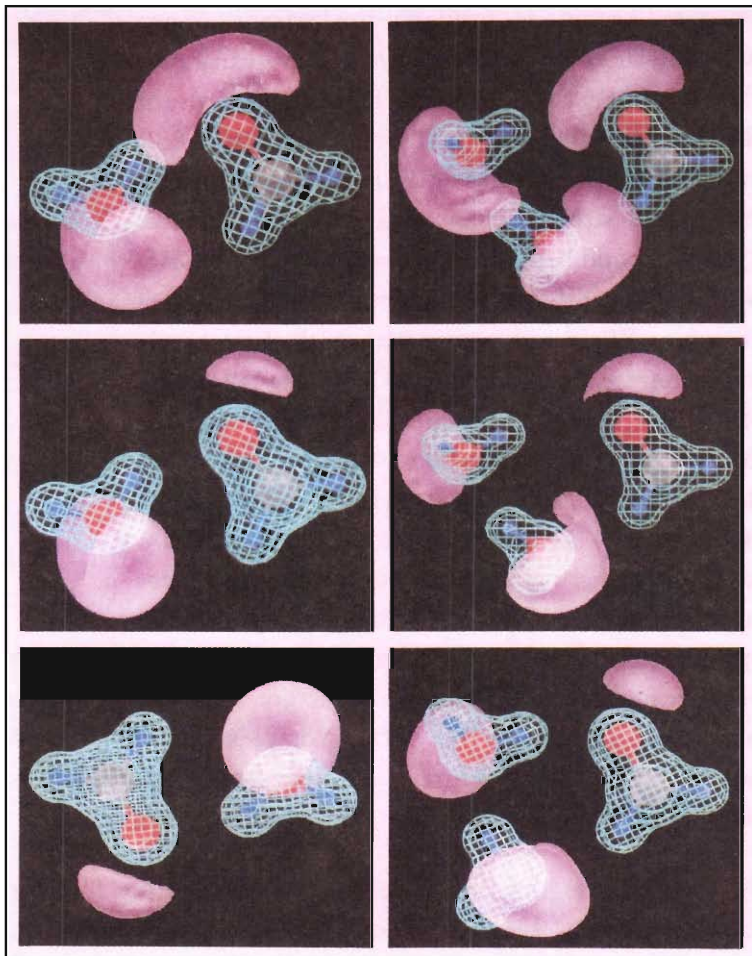
Table 2. MESP-derived point charges (in a.u.) for symmetry-distinctive atoms of formaldehyde and water. *a*, *b* and *c* denote hydrogens attached to O, anti to OH and gauche to OH in methanol, respectively.



Figure 2a and d. MESP isosurfaces of values -126 kJ mol^{-1} (purple) and 788 kJ mol^{-1} (blue) of formaldehyde and water placed in the EPIC-optimized orientation of $\text{HCHO}\dots\text{H}_2\text{O}$ and $\text{HCHO}\dots 2\text{H}_2\text{O}$ complexes, respectively. This figure brings out the lock-and-key role played by the MESP.

Figure 2b and e. MESP isosurfaces of values -126 kJ mol^{-1} and 788 kJ mol^{-1} , obtained on addition of ESPs of individual molecules, at EPIC-optimized geometries for $\text{HCHO}\dots\text{H}_2\text{O}$ and $\text{HCHO}\dots 2\text{H}_2\text{O}$ complexes, respectively.

Figure 2c and f. MESP isosurfaces of values -126 kJ mol^{-1} and 788 kJ mol^{-1} for *ab initio* optimized structures of $\text{HCHO}\dots\text{H}_2\text{O}$ and $\text{HCHO}\dots 2\text{H}_2\text{O}$ complexes, respectively.



Cyclopropane is also known for its use as a surgical anaesthetic. It may be interesting to explore the interactions of cyclopropane with the major component of any biological system and the universal solvent, water. Cyclopropane along with other hydrocarbons is conventionally known to be a non-polar system, which does not interact with polar molecules. However, mapping the MESP topography of cyclopropane, or any other hydrocarbon for that matter, reveals that such molecules do have electron localization leading to sites that can weakly bind to electrophiles. Cyclopropane has two types of MESP minima on the C_2 axis of symmetry, one near the C-C bond centre and the other between two hydrogens of CH_2 . Docking a water molecule with cyclopropane using EPIC, followed by *ab initio* optimiza-

tion leads to a geometry with H of water binding to the most negative MESP CP of cyclopropane (see *Figure 3a*) and not H of cyclopropane with CP of water as would be expected without any insight into the MESP topography. Similar procedure for $C_3H_6 \dots 2H_2O$ complex yields a cyclic hydrogen bonded structure. However, in this case the EPIC-predicted geometry differs from the *ab initio* one in the orientation of the hydrogen atoms. A general agreement between the EPIC electrostatic distribution and its *ab initio* counterpart is noticed, although sometimes the detailed features differ to some extent. *Figures 3a* and *b* bring out the MESP isosurface of values -39 and -53 kJ mol^{-1} , respectively, for cyclopropane with one and two water molecules, respectively.

Methanol is a unique organic solvent, being highly polar, the least volatile of common organic solvents, completely miscible with water and economically viable. It is used as a solvent in several organic reactions. Studying the hydration of methanol would definitely add to our understanding of reaction mechanisms. As in the previous cases, a sample study of interaction of CH_3OH with one and two water molecules was carried out. The interaction energies for the geometries obtained by the above mentioned method are tabulated in *Table 3*. The isosurfaces of

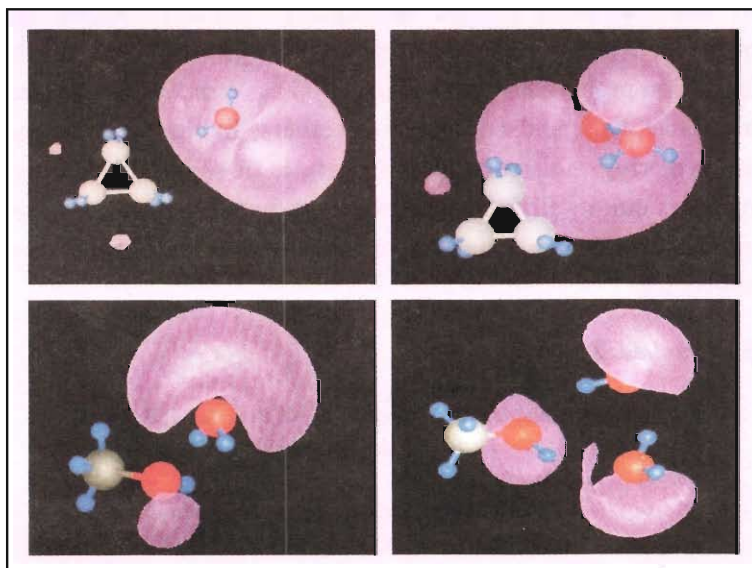


Figure 3a and b. The *ab initio* optimized geometry of (a) $C_3H_6 \dots H_2O$ with isosurface of MESP value -39 kJ mol^{-1} (b) $C_3H_6 \dots 2H_2O$ with the isosurface of MESP value -53 kJ mol^{-1} .

Figure 3c and d. The *ab initio* optimized geometry of (c) $CH_3OH \dots H_2O$ and (d) $CH_3OH \dots 2H_2O$ with the isosurface of MESP value -105 kJ mol^{-1} .

Table 3. Interaction energies (in kJ mol^{-1}) given by EPIC and those obtained with *ab initio* optimization for complexes of formaldehyde, cyclopropane and methanol with one and two water molecules.

Complex	E_{int} (EPIC)	E_{int} (<i>ab initio</i>)
HCHO...H ₂ O	-24.41	-19.34
HCHO...2H ₂ O	-44.47	-53.05
C ₃ H ₆ ...H ₂ O	-10.72	- 5.11
C ₃ H ₆ ...2H ₂ O	-37.93	-28.26
CH ₃ OH...H ₂ O	-24.95	-24.03
CH ₃ OH...2H ₂ O	-65.61	-60.08

MESP value -105 kJ mol^{-1} , for these species, are shown in *Figure 3c* and *d*, respectively. These patterns will govern further hydration of the species.

The cover page picture of this issue depicts the isosurface of values of a water cluster, $(\text{H}_2\text{O})_6$ obtained using a treatment similar to the one discussed above.

Conclusions

In this part we have provided an introduction to some simple models useful for investigation of weak-intermolecular complexation. Out of these, an electrostatics-based model EPIC, which is quite reliable for reproducing *ab initio* quality results is discussed in detail. The application of the methodology for understating the hydration of a few small molecules is described. The simplicity and reliability suggest applications to more complex supramolecular systems. The cooperative-electrostatics in the hydration process brought out by EPIC has already been applied in exploring the hydration of crown ethers. In view of its success in predicting the structure and energies of supramolecular assemblies quite well, it may be expected that electrostatics-based models would play a significant role in studying aggregates of molecules.

Suggested Reading

- [1] **S R Gadre; P K Bhadane; S S Pundlik and S S Pingale** in **J S Murray and K D Sen Ed., *Molecular Electrostatic potentials: Concepts and Applications*, Elsevier, Amsterdam, 1996.**
- [2] Please refer to the following articles for an introduction to *ab initio* methods including DFT. (a) **S K Ghosh, *Resonance*, Vol. 4 No. 5, 1999.** (b) **K D Sen, *Resonance*, Vol. 4 No. 4, 1999.**

Address for Correspondence

Shridhar R Gadre and K Babu
Department of Chemistry
University of Pune
Pune 411 007, India.