

CHAPTER IV

DETAILS OF THE CALCULATIONS

4.1 Development of Intermolecular Potential Function

It is generally accepted that a deficiency of solute-solute and solute-solvent potential functions is the main obstacle to the rapid progress of simulation techniques in solution chemistry. The reliability of the simulation depends strongly on the quality of the potential functions describing the interaction energy among particles in the system. One obvious way to develop such functions is to use quantum chemical calculations.

Given a system of N particles, the total configurational interaction-energy $E(v)$ can be obtained from:

$$E(v) = \sum_{i>j} V(i,j) + \sum_{i>j>k} V(i,j,k) + \dots + V(1,2,3,\dots,n) \quad [4.1],$$

where the first term in the right hand side denotes the two-body interactions, the second the three-body, and so on until the last term which contains all the bodies of the system. The assumption that $E(v)$ is equal to $\sum V(i,j)$ is known as the *pairwise additive approximation* and the remaining terms of the series are often referred to as *nonadditive corrections*. The basic idea is that the total configurational energy of the many particle system is approximately written as a sum of the individual interaction energies between the particles.

The general procedure for obtaining the *ab initio* intermolecular potential function consists of five consecutive steps:

- a) selection of representative geometries of the pairs,
- b) selection of a reliable basis set for the SCF calculations,
- c) performance of the SCF calculations,
- d) fitting of the computed interaction energies to a functional form,
and
- e) testing the quality of the function.

In this section, the derivation of the intermolecular potential function for 1,4,7,10-tetraazacyclododecane (cyclen) and ammonia is reported, based on the *ab initio* computed energies, with the STO-3G basis set.

4.1.1 Selection of the geometries of the pairs

In this task, a graphic drawing representing the geometric conformation of cyclen-ammonia was monitored (Figure 4.2). Such picture not only makes the solute-solvent molecules easy to visualize but also helps suggest which geometry should be considered and chosen. A point to be kept in mind, during the selection, is that both attractive and repulsive interaction-energy configurations must be included.

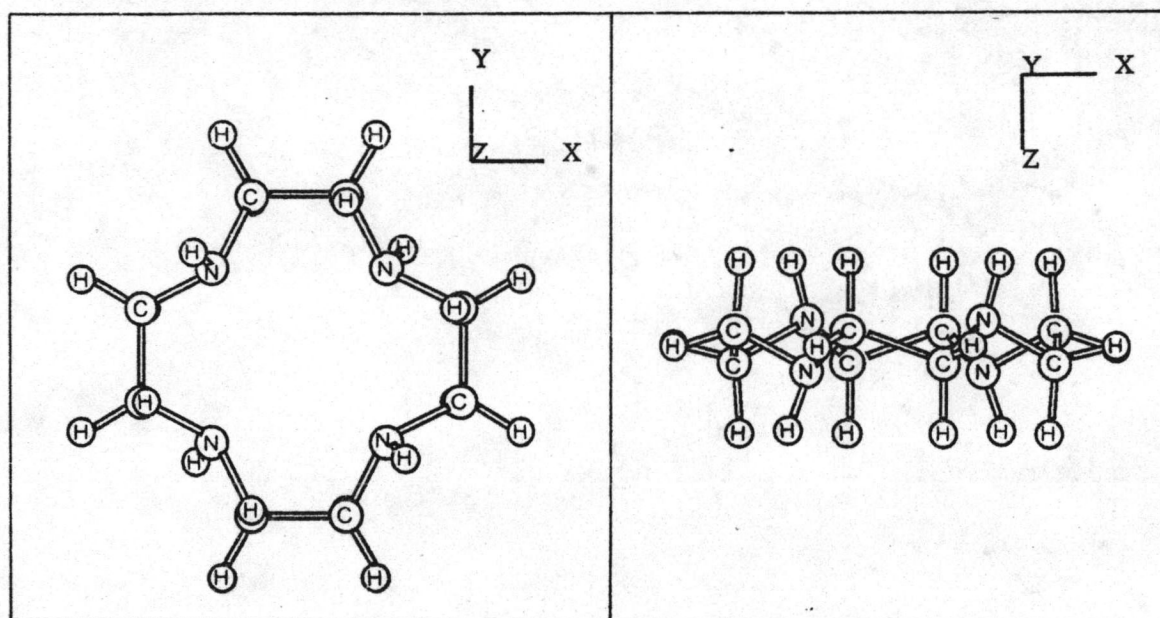


Figure 4.1 The *alternate form* of 1,4,7,10-tetraazacyclododecane.

The cyclen and ammonia molecules were treated as rigid throughout the calculations. The geometry of the ligand molecule was taken from previous optimizations (44,45). For the ammonia molecule, the experimental values for the gas phase, with an N-H distance of 1.0124 Å and an H-N-H angle of 106.67°, are used (46). The cyclen molecule is in the *alternate form* (44,45) and has C_{2v} -symmetry (Figure 4.1). Its optimized coordinates are given in Table 4.1.

Considering the data from a Mulliken population analysis (40) in the SCF calculations, the hydrogen atoms of cyclen were distinguished into three groups: H_N for hydrogen atoms binding to the N atoms, and H_C and $H_{C'}$ for the hydrogens binding to the C atoms in and perpendicular to the molecular plane, respectively.

The pair geometries were selected with special regard to chemically representative conformations and molecular symmetry. The ammonia molecule was placed at numerous positions within the space around the cyclen, which was and centered at the origin of spherical coordinates, where $0^\circ \leq \phi \leq 90^\circ$, $0^\circ \leq \theta \leq 180^\circ$ and $2.0 \leq r \leq 10.0 \text{ \AA}$ (Figure 4.2).

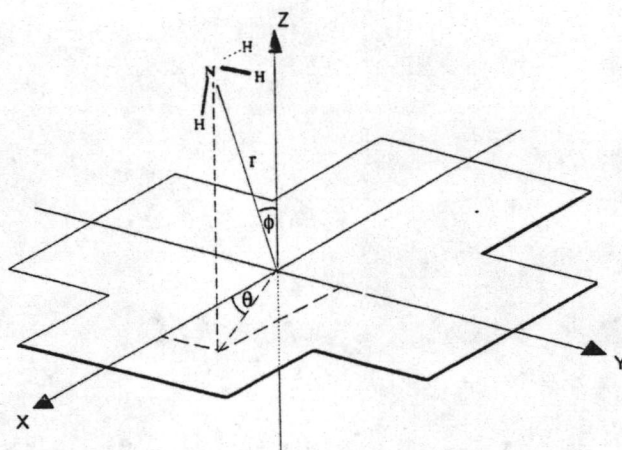


Figure 4.2 The optimized coordination of solvent molecule interacting with 1,4,7,10-tetraazacyclododecane.

With this consideration 600 configurations of the ammonia molecule have been selected and used. The CPU time has been reduced to some extent due to the symmetry of the cyclen molecule. In addition, the minimum interaction energy can be easily observed.

Table 4.1 Optimized Cartesian coordinates of 1,4,7,10-tetraazacyclododecane using GLO basis set (44,45) (in Angstroms).

atoms	X	Y	Z
N	-1.34168	-1.34168	-0.39390
N	1.34167	-1.34168	0.39390
N	1.34168	1.34167	-0.39390
N	-1.34166	1.34168	0.39390
C	-0.72621	-2.50763	0.25518
C	0.72623	-2.50763	-0.25518
C	2.50762	-0.72624	-0.25518
C	2.50763	0.72622	0.25518
C	0.72624	2.50762	0.25518
C	-0.72621	2.50762	-0.25519
C	-2.50762	0.72624	-0.25518
C	-2.50763	-0.72621	0.25518
H _C	-1.21458	-3.44214	-0.03509
H _C	1.21461	-3.44214	0.03510
H _C	3.44213	-1.21462	0.03509
H _C	3.44214	1.21459	-0.03510
H _C	1.21463	3.44212	-0.03510
H _C	-1.21459	3.44214	0.03509
H _C	-3.44212	1.21462	0.03509
H _C	-3.44214	-1.21458	-0.03509
H _C	-0.72960	-2.41776	1.34513
H _C	0.72961	-2.41776	-1.34513
H _C	2.41776	-0.72962	-1.34514
H _C	2.41777	0.72961	1.34513
H _C	0.72963	2.41776	1.34513
H _C	-0.72960	2.41776	-1.34514
H _C	-2.41775	0.72962	-1.34514
H _C	-2.41776	-0.72960	1.34514

Table 4.1 (continued)

atoms	X	Y	Z
H _N	1.66177	-1.60058	1.32010
H _N	1.60058	1.66208	-1.32011
H _N	-1.66208	1.60058	1.32010
H _N	-1.60059	-1.66209	-1.32010

4.1.2 Selection of a reliable basis set for the SCF calculations

To obtain the most accurate results for the molecular orbital calculations, several factors must be considered. The calculated procedure should be as efficient as possible, both in terms of time required to complete a given task and optimal usage of available computational resources. For instance, in choosing the size of the basis set, we have to find an acceptable compromise between computational effort and accuracy. However, the models have to be chosen so as to preserve as closely as possible the chemical and physical integrity of the whole system.

The SCF calculations have been performed for the cyclen-ammonia dimer in the configuration given in Figure 4.2, where $\theta = \phi = 0^\circ$. The STO-3G, 3-21G, 6-21G, 6-31G, DZ, DZV and DZP basis sets have been used. The N-origin distance (Figure 4.2) has been optimized in order to search for the optimal geometry for each basis set.

The STO-3G set, for which each Slater type function contains 3 primitives, is (6s3p/3s)/[2s1p/1s] as in ref. (47). That means that the 6s and 3p primitives for C, N and O atoms and s- and p-shells are contracted to 2s and 1p, respectively. For the hydrogen atom, 3 primitives on s-shells are contracted to 1s.

Split-valence basis sets, 3-21G, 6-21G and 6-31G, have been applied. The corresponding contractions are (6s3p/3s)/[3s2p/2s], (9s3p/3s)/[3s2p/2s] and (10s4p/4s)/[3s2p/2s], respectively (48-52).

The last set which is taken into consideration are the double zeta quality sets of the form (10s5p/4s)/[3s2p/2s], (9s5p/4s)/[4s2p/2s] and

(9s5p1d/4s1p)/[4s2p1d/2s1p]. The corresponding notations are DZV, DZ and DZP, respectively (53). The exponents of the STO functions, together with their coefficients, are summarized in Appendix (II).

4.1.3 Performance of the SCF calculations

Using selected configurations, as illustrated in section (4.1.1) and the selected STO-3G basis set (for a detailed discussion, see chapter V), the interaction energy ΔE_{SCF} between cyclen (*L*) and ammonia (*A*) is determined as the difference between the supersystem energy, E_{LA} , and the sum of the subsystem energies (E_L , E_A),

$$\Delta E_{SCF}(kcal/mol) = \{E_{LA} - (E_L + E_A)\} \times 627.5 \quad [4.2],$$

where E_{LA} , E_L and E_A are in atomic units.

The interaction energy, in this study, is always obtained by this expression, which seems to be quite simple. However, serious difficulties are encountered in the numerical treatment. The most important one lies in the fact that the supersystem energy and the sum of the energies of the isolated subsystems are very large compared to the interaction energy. Therefore, the determination of the total energy of the systems must be extremely accurate, to within 10^{-5} Hartree or better.

The Gaussian 92 program (54) on the IBM RISC 6000/53H Workstation was used for the SCF calculations with varying basis sets and the HONDO program (55) on the DEC 3100 Workstation was employed using the selected STO-3G basis set for the 600 cyclen-ammonia configurations. Both computers are in the Computational Chemistry Unit Cell and the Austrian-Thai Centre for Computer-Assisted Chemical Education and Research at Chulalongkorn University.

4.1.4 Fitting of the pair interaction energies to a functional form

In this step, a suitable mathematical function is selected. The total intermolecular interaction energies between the two molecules were written as a sum of atom-atom pair potentials, each of which was considered to be a function of the distance r_{ij} of the two atoms concerned. Of course the true intermolecular potential function is unknown. The functional form of the pair potentials is normally chosen, with a physical meaning in mind, to consist of a term to describe the *Coulombic interaction* between the atom pairs. The *Mulliken population analysis* from the *ab initio* calculations was used to get a first estimate of the fractional atomic charges, q_i . Then, additional terms of the form A/r^m , $m = 2 \dots 6$, especially effective at medium distances, should be added to describe phenomena like the distance dependence of these charges. Finally, there should be a term to represent the electronic repulsion at close distances and to take care of the Pauli exclusion principle, which effect increases nucleonic repulsion. This term is typically represented by a polynomial B/r^n , $n = 8 \dots 15$, function. Therefore, a trial potential function was set up as

$$\Delta E(r_{ij}) = \sum_i \sum_j \frac{A_{ij}}{r_{ij}^m} + \frac{B_{ij}}{r_{ij}^n} + \frac{q_i q_j}{r_{ij}} \quad [4.3],$$

with the coefficients A_{ij} and B_{ij} to be determined so that a best fit to the SCF calculated energies is achieved. However, it should be emphasized that these coefficients should be treated with some care and not only as mathematical objects, especially if the obtained potential function is to be used in a subsequent simulation study. For instance, negative values for the A_{ij} coefficients may result in the corresponding atomic potential function turning negative at very small distances.

The fitting procedure was performed starting from an initial guess for the parameters and iterating until the standard deviation was minimized, whereby constancy of the fitting parameters was reached. To improve the quality of the fit, especially in the most important configurations (which correspond to the lowest lying values of the SCF energy points), additional weight factors were introduced, and very repulsive configurations were excluded from the process.

4.1.5 Testing the quality of the function

The quality of the fit for the obtained potential function was judged by its statistical characteristics (standard deviation, residuals etc.), the values and positions of energy minima for the chemically important configurations, and a graphical representation of the fitted energies vs. their quantum mechanically calculated values. Moreover, the predictive capabilities of the potential function were then tested according to the procedure suggested by Beveridge (56). An additional set of certain, for example 50, randomly distributed SCF points was calculated. The ability of the function to reproduce the energies of these 50 points was tested by comparing the standard deviation of these points with that of the previously calculated points. Afterwards, these 50 points were included into the fit and the parameters of the function were reoptimized. Then, the standard deviation was, again, determined. This procedure was repeated until constancy of the fitting parameters was obtained within a range of $\pm 5\%$ (57).

4.2 Specification of the model system

Taking into account the whole procedure of the fit, a variety of functional forms were considered. The first trial potential consisted of 300 SCF data points. An additional set of 50 points was used to test the quality of the fit and included into the function afterwards. The final form of the intermolecular potential function was obtained on the basis of 600 points on the energy surface.

The prevailing functional form describing the SCF energy surface for cyclen-ammonia is

$$\Delta E(L, A) = \sum_{i=1}^4 \sum_{j=1}^{32} \frac{A_{ij}^{ab}}{r_{ij}^6} + \frac{B_{ij}^{ab}}{r_{ij}^{12}} + \frac{q_i q_j}{r_{ij}} \quad [4.4],$$

where A_{ij} and B_{ij} are the fitting constants, r_{ij} is the distance between the i th atom of ammonia and the j th atom of cyclen, in Angstroms, and q_i and q_j are the corresponding atomic net charges, in atomic units, obtained from the Mulliken population analysis of the isolated molecules. Superscripts a and b indicate, as

mentioned before, that the coefficients not only varied with atomic number, but also among atoms of equal atomic number but different environmental conditions. Weighting factors were introduced to give special emphasis to ΔE_{SCF} near the energy minima (-2.51 kcal/mol). Values above 10 kcal/mol were excluded.

4.3 Monte Carlo studies of 1,4,7,10-tetraazacyclododecane in aqueous ammonia

4.3.1 *Characteristics of the simulations*

The simulations have been performed, starting from a randomly generated configuration, for a very dilute solution of cyclen in 18.45 mol% aqueous ammonia. That is, the periodic cube contains one cyclen molecule that is fixed at the center of cube, and 37 ammonia and 164 water molecules. An experimental density of 0.9307 g/cm³ of the 18.45 mol% aqueous ammonia solution at 298 K and 1 atm has been used (48). The resulting side length of the periodic box is 18.5581 Å. Half of this length was chosen as the spherical cut-off. All molecules were assumed to be rigid. The Metropolis sampling algorithm (1) was applied. The MCY (27) and the HISH potentials (28) were used to describe water-water (*WW*) and ammonia-ammonia (*AA*) interactions, respectively.

The MCY potential (ΔE in kcal/mol, r in Angstroms) is

$$\begin{aligned} \Delta E(WW) = & \frac{Q^2}{4\pi\epsilon_0} \left[\left(\frac{1}{r_{13}} + \frac{1}{r_{23}} + \frac{1}{r_{14}} + \frac{1}{r_{24}} \right) + \frac{4}{r_{78}} - 2 \left(\frac{1}{r_{28}} + \frac{1}{r_{18}} + \frac{1}{r_{37}} + \frac{1}{r_{47}} \right) \right] \\ & + a_1 \exp(-b_1 r_{56}) + a_2 [\exp(-b_2 r_{13}) + \exp(-b_2 r_{14}) \\ & + \exp(-b_2 r_{23}) + \exp(-b_2 r_{24})] + a_3 [\exp(-b_3 r_{16}) + \exp(-b_3 r_{26}) \\ & + \exp(-b_3 r_{35}) + \exp(-b_3 r_{45})] - a_4 [\exp(-b_4 r_{16}) + \exp(-b_4 r_{26}) \\ & + \exp(-b_4 r_{35}) + \exp(-b_4 r_{45})] \end{aligned} \quad [4.5],$$

where Q is a positive charge situated on each hydrogen atom and a_1 - a_4 and b_1 - b_4 are fitting parameters. The hydrogen atoms are labelled 1-4 and the oxygen atoms

are labelled 5 and 6. The labels 7 and 8 refer to point charges of $-2Q$ situated along the axis of the dipole-moment vector.

The HISH potential for ammonia-ammonia interactions (ΔE in kcal/mol, r in Angstroms) is

$$\Delta E_{NN}(r) = \frac{q_N q_H}{r} + \frac{A}{r^{12}} - \frac{B}{r^6} \quad [4.6a],$$

$$\Delta E_{NH}(r) = -\frac{q_N q_H}{r} + C\{\exp[-2D(r-E)] - 2\exp[-D(r-E)]\} \quad [4.6b],$$

$$\Delta E_{HH}(r) = \frac{q_H q_H}{r} + F \exp(-Gr) \quad [4.6c],$$

$$\Delta E(AA) = \Delta E_{NN}(r) + \Delta E_{NH}(r) + \Delta E_{HH}(r) \quad [4.6d].$$

The TR ammonia-water interactions energy (AW) (29) is employed (ΔE in kcal/mol, r in Angstroms),

$$\Delta E(AW) = \sum_{i<j} \frac{q_i q_j}{r_{ij}} + \sum_{m<n} [A_{mn} \exp(-B_{mn} r_{mn}) - C_{mn} \exp(-D_{mn} r_{mn})] \quad [4.7],$$

where i and j refer to charges and m and n to atoms of the water and ammonia molecules, respectively.

The HR potential for cyclen-water interactions (LW) [Hannongbua, S. V., and Rode, B. M. (63)] (ΔE in kcal/mol, r in Angstroms) is:

$$\Delta E(LW) = \sum_{i=1}^3 \sum_{j=1}^{32} \left[\frac{|A_{ij}^{ab}|}{r_{ij}^6} + \frac{|B_{ij}^{ab}|}{r_{ij}^{12}} + \frac{q_i q_j}{r_{ij}} (|C_{ij}^{ab}| + \frac{|D_{ij}^{ab}|}{r_{ij}}) \right] \quad [4.8],$$

where detailed variable are equivalent to those of the cyclen-ammonia function. The fitting parameters for all pairs are summarized in the Appendix (II).

4.3.2 Defect in the HR potential for cyclen-water interactions

After some million configurations of the simulations, a cyclen-water interaction energy of -15.8 kcal/mol, which is lower than that of -11.3 kcal/mol for the optimal configuration of the dimer (Figure 4.3a) reported in ref. (63), was surprisingly detected. It was found after investigating in detail that the corresponding STO-3G stabilization energy for such an unwanted configuration (Figure 4.3b) is $+6.0$ kcal/mol. The big difference of about 22 kcal/mol is surely unacceptable. In the optimal configuration (Figure 4.3a), the STO-3G stabilization energy of -5.5 kcal/mol is also uncomparable to that of -11.3 kcal/mol from the HR function. These data together with the shift of the optimal O-origin distance from 2.5 Å for the HR function to 2.7 Å for the STO-3G calculations suggested that the cyclen-water function had to be newly developed.

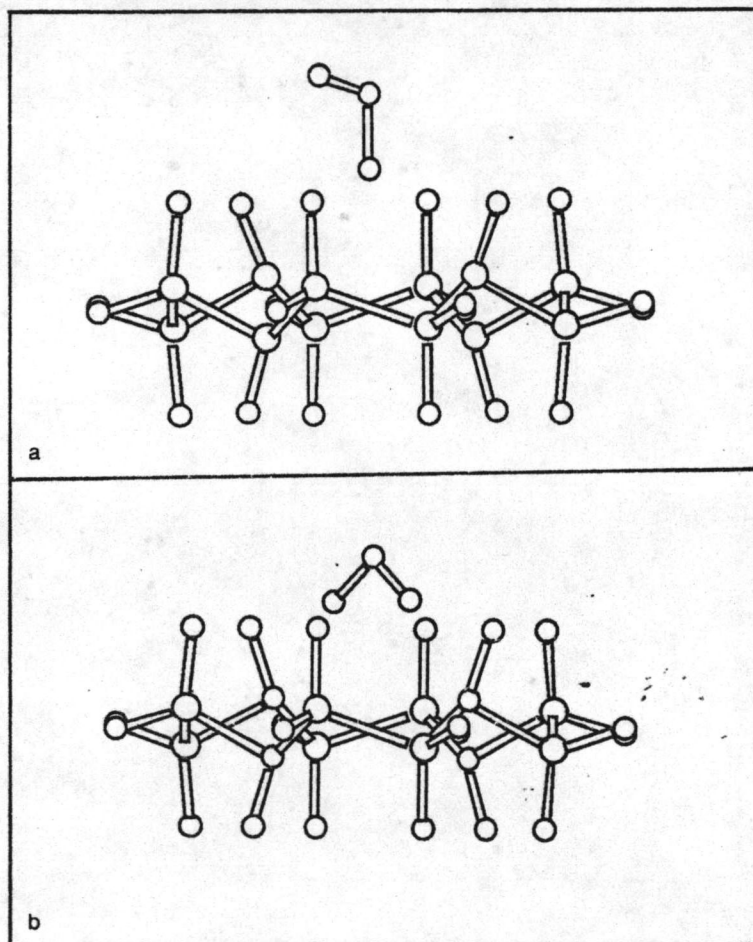


Figure 4.3 Cyclen-water configurations where
a) the global minimum and b) the unwanted minima take place.

4.3.3 *Development of the cyclen-water potential function*

A new cyclen-water potential function has been developed. The cyclen-water interactions have been generated by a geometrical selection of chemically representative conformations. An intermolecular pair potential of this interaction has been derived based on these energies. The experimental geometries of water with known bond lengths [O-H=0.9572 Å and H-O-H=104.5° (64)], have been used and kept constant. The interaction procedure, as in the cyclen-ammonia case, has been applied. The same series of the basis sets was tested and the STO-3G was selected. The fit started with 300 data points and N_{test} was 50. Weighting factors were introduced to give special emphasis to values near global and local energy minima. Values above 10 kcal/mol were excluded. The final form of the cyclen-water function, based on 600 data points, is given in equation [5.1]), with a final standard deviation of 0.75 kcal/mol.

4.3.4 *Monte Carlo simulations with the new cyclen-water function*

The Monte Carlo simulations were carried out in continuation from the last configuration obtained with the old cyclen-water function. It was found, after equilibration, that the shape of the radial distribution functions and the corresponding integration numbers are inconsistent between the two sets of about other million configurations. Water molecules which are situated under the first peak of the RDF are found to move in and out. In order to achieve periodicity in these results, the simulations were performed for 35 million configurations. All quantities were evaluated from 28 million configurations after equilibrium was reached. Of course, it is not possible to collect history file - the file where the Cartesian coordinates of all atoms are summarized - for the whole run (42 Mbyte for one million configurations). Therefore, all the RDFs and integration numbers were averaged from 28 million configurations, while the other results were investigated from a history file consisted of the last 3 million configurations.

4.4 BSSE for cyclen-water interactions

In order to estimate the possible influence of basis set superposition error, additional calculations have been made including the Boys-Bernardi counterpoise correction. The selected direction is the configuration where the global minimum of the stabilization energy takes place ($\theta = \phi = 0^\circ$ in Figure 4.2). The calculations are limited only to the STO-3G and the DZP basis sets since they are expected to provide reasonable characterization of the dimer (a detailed discussion is given in Chapter V).