CHAPTER II

Intermolecular Potential Functions

2.1 Ab Initio Calculation Method

2.1.1 General

The most commonly used molecular quantum-mechanical methods can be classified as either **ab initio or semiempirical** type. **Semiempirical** methods use a simpler Hamiltonian and use parameters whose values are adjusted to fit experimental data or the the results of ab initio calculations. In contrast, an **ab initio** calculation uses the full Hartree-Fock Hamiltonian and does not use experimental data other than the values of the fundamental physical constants. A Hartree-Fock SCF calculation seeks the antisymmetrized product Φ of one-electron functions that minimizes $\int \Phi^* H \Phi \, d\tau$,where H is the Hamiltonian. (Ab initio is Latin for "from the beginning" and indicates a calculation based on first principles.)

The ab initio methods have risen in popularity in recent years because of the availability of high-speed digital computers and the development of theoretical and computational methods. The main factor determining the accuracy of an ab initio calculation is the size and type of the basis set used. Several types of basis sets will be presented in this chapter including the ECP (Effective Core Potential) basis sets used in this work.

Ab Initio Molecular Orbital Theory

Any physically observable molecular property can be calculated quantum mechanically if the corresponding Schroedinger equation can be solved.

$$H\Psi = E\Psi \qquad (2.1)$$

According to this equation the total energy is obtained by

$$E = \langle \Psi | H | \Psi \rangle \qquad (2.2)$$

where Ψ is the normalized total wave function of the system.

The principle of the ab initio approach is based on the three following step:

i) Define an approximate Hamiltonian operator for the system.

ii)Define one-electron functions $\psi_i(q_i)$ as the trial wave functions,which are composed from basis functions ϕ_j according to $\psi_i = \Sigma_j C_{ij}\phi_j$

iii) Minimize the total energy , (eq. 2.2) with respect to variations of the coefficients C_{ij}

$$E = \frac{\int \Psi(q_1, q_2, q_3, \ldots) H \Psi(q_1, q_2, q_3, \ldots)}{\int \Psi(q_1, q_2, q_3, \ldots) \Psi(q_1, q_2, q_3, \ldots)} \qquad (2.3)$$

More details of the above consecutive steps will be briefly summarized as the following.

The Molecular Hamiltonian Operator

If we are interested in molecular systems, the total Hamiltonian operator of the system will be given as sum of all possible Coulombic interactions plus kinetic energy of the electrons and nuclei.

$$H^{core}(i) = -1/2 \sum_i \nabla_i^2 - \sum_A Z_A / r_{iA}$$

H^{el} is a one-electronic Hamilton operator, and the total electronic Hamiltonian is composed of the 1 electron operators for all electrons.

The Hartree-Fock Wave Functions

a.) Independent Electron Model

The typical approximation to construct the many-electron wave function is the Independent Particle Approximation, based on a factorization of Ψ into one-electron spin orbitals(MOs), being themselves products of a spatial orbital and a spin function. The most convenient way to represent a trial wave function for a 2n-electron-closed shell system is to use a single-determinantal wave function, also called a Slater determinant (14,15) given in eq.(2.6), in order to observe the antisymmetry principle.

$$\Psi = \frac{1}{\sqrt{2n!}} \left| \psi_1(1)\, \bar{\psi}_1(2)\, \psi_2(3)\, \bar{\psi}_2(4) \ldots \psi_n(2n-1)\, \bar{\psi}_n(2n) \right| \qquad (2.6)$$

the spin orbitals ψ_i and ψ_i correspond to $\psi(\alpha)$ and $\psi(\beta)$, where α and β denote the spin functions.

This approximation automatically leads to a split-up of the Hamiltonian into a sum of one electron operators (Fock Operator)

$$H^{core}(i) = -1/2 \sum_i \nabla_i^2 - \sum_A Z_A / r_{iA}$$

H^{el} is a one-electronic Hamilton operator, and the total electronic Hamiltonian is composed of the 1 electron operators for all electrons.

The Hartree-Fock Wave Functions

a.) Independent Electron Model

The typical approximation to construct the many-electron wave function is the Independent Particle Approximation, based on a factorization of Ψ into one-electron spin orbitals(MOs), being themselves products of a spatial orbital and a spin function. The most convenient way to represent a trial wave function for a 2n-electron-closed shell system is to use a single-determinantal wave function, also called a Slater determinant (14,15) given in eq.(2.6), in order to observe the antisymmetry principle.

$$\Psi = \frac{1}{\sqrt{2n!}} \left| \psi_1(1)\, \bar{\psi}_1(2)\, \psi_2(3)\, \bar{\psi}_2(4) \ldots \psi_n(2n-1)\, \bar{\psi}_n(2n) \right| \qquad (2.6)$$

the spin orbitals ψ_i and ψ_i correspond to $\psi(\alpha)$ and $\psi(\beta)$, where α and β denote the spin functions.

This approximation automatically leads to a split-up of the Hamiltonian into a sum of one electron operators (Fock Operator)

The matrix of the elements of the core Hamiltonian, H^{core}, contains the elements for the core-electron Hamiltonians, H^c, for electrons moving in the field of nuclei :

$$H_{\mu}^{core} = < \phi_{\mu}(i) \mid H^{c} \mid \phi (i) > \qquad (2.10)$$

The second term in eq.(2.9) is the two-electron part. The elements of the density matrix $P_{\lambda\sigma}$ and the two-electron integrals, $(\mu\nu/\lambda\sigma)$ and $(\mu\sigma/\nu\lambda)$, are given by

$$P_{\lambda\sigma} = 2 \sum_{i}^{occ} C^{*}_{\lambda i} C_{\sigma i} \quad \textbf{(for closed shell system)} \qquad (2.11)$$

$$(\mu\nu|\lambda\sigma) = \int\int \phi^{*}_{\mu}(1)\, \phi_{\nu}(1) \frac{1}{r_{12}} \phi^{*}_{\lambda}(2)\, \phi_{\sigma}(2)\, d\tau_1\, d\tau_2 \qquad (2.12)$$

and

$$(\mu\sigma|\nu\lambda) = \int\int \phi^{*}_{\mu}(1)\, \phi_{\sigma}(1) \frac{1}{r_{12}} \phi^{*}_{\nu}(2)\, \phi_{\lambda}(2)\, d\tau_1\, d\tau_2 \qquad (2.13)$$

From the first approximation of the Fock operator matrix F, a second approximation of C can be obtained by solving of the Roothaan equation

$$FC = SCE \qquad (2.14)$$

The matrix of the elements of the core Hamiltonian, H^{core}, contains the elements for the core-electron Hamiltonians, H^c, for electrons moving in the field of nuclei :

$$H_{\mu}^{core} = < \phi_{\mu}(i) | H^{c} | \phi (i) > \qquad (2.10)$$

The second term in eq.(2.9) is the two-electron part. The elements of the density matrix $P_{\lambda\sigma}$ and the two-electron integrals, $(\mu\nu/\lambda\sigma)$ and $(\mu\sigma/\nu\lambda)$, are given by

$$P_{\lambda\sigma} = 2 \sum_{i}^{occ} C^{*}_{\lambda i} C_{\sigma i} \quad \textbf{(for closed shell system)} \qquad (2.11)$$

$$(\mu\nu|\lambda\sigma) = \int\int \phi^{*}_{\mu}(1)\, \phi_{\nu}(1) \frac{1}{r_{12}} \phi^{*}_{\lambda}(2)\, \phi_{\sigma}(2)\, d\tau_1\, d\tau_2 \qquad (2.12)$$

and

$$(\mu\sigma|\nu\lambda) = \int\int \phi^{*}_{\mu}(1)\, \phi_{\sigma}(1) \frac{1}{r_{12}} \phi^{*}_{\nu}(2)\, \phi_{\lambda}(2)\, d\tau_1\, d\tau_2 \qquad (2.13)$$

From the first approximation of the Fock operator matrix F, a second approximation of C can be obtained by solving of the Roothaan equation

$$FC = SCE \qquad (2.14)$$

2.1.2 Supermolecule approach

The **supermolecule method** is the usual quantum chemical approach, used within the framework of the discrete molecular complex models. The supermolecule procedure consists of the solution of the Schrodinger equation in the "clamped nuclei" (Born-Oppenheimer) approximation for the whole molecular complex, e.g. for a solution containing one solute and a certain number of solvent molecules. In general, within this approach, the total interaction energy, ΔE, is defined as the difference between the total energy of the supermolecule and the sum of energies of the isolated constituents (one solute and N solvent) molecules :

$$\Delta E = E^{(sup)} - \sum_{I} E_{I}^{0} \qquad (2.15)$$

In this work the interaction energy will be determined by this expression, which seems to be quite simple. However, serious difficulties are encountered in numerical treatment. The most important difficulty lies in the fact that the supersystem energy and the sum of the energy of each isolated constituent molecules are very large numbers compared to the interaction energy. Therefore, the determination of the energy of the system must be accurate (10^{-5} Hartree). Another factor influencing the ΔE value is the effect of the basis set on the value of the SCF interaction energy. In general, the basis set employed should correctly express the multipole moments and the polarizability of the systems. Inclusion of polarization functions would be essential for this.

Basis Set Superposition Error (BSSE)

In calculation of the interaction energy for complex systems, the basis set employed should be sufficiently large and correctly express the multipole moments

2.1.2 Supermolecule approach

The **supermolecule method** is the usual quantum chemical approach, used within the framework of the discrete molecular complex models. The supermolecule procedure consists of the solution of the Schrodinger equation in the "clamped nuclei" (Born-Oppenheimer) approximation for the whole molecular complex, e.g. for a solution containing one solute and a certain number of solvent molecules. In general, within this approach, the total interaction energy, ΔE, is defined as the difference between the total energy of the supermolecule and the sum of energies of the isolated constituents (one solute and N solvent) molecules :

$$\Delta E = E^{(sup)} - \sum_{I} E_{I}^{0} \qquad (2.15)$$

In this work the interaction energy will be determined by this expression, which seems to be quite simple. However, serious difficulties are encountered in numerical treatment. The most important difficulty lies in the fact that the supersystem energy and the sum of the energy of each isolated constituent molecules are very large numbers compared to the interaction energy. Therefore, the determination of the energy of the system must be accurate (10^{-5} Hartree). Another factor influencing the ΔE value is the effect of the basis set on the value of the SCF interaction energy. In general, the basis set employed should correctly express the multipole moments and the polarizability of the systems. Inclusion of polarization functions would be essential for this.

Basis Set Superposition Error (BSSE)

In calculation of the interaction energy for complex systems, the basis set employed should be sufficiently large and correctly express the multipole moments

2.1.3 Basis Functions

In all molecular quantum mechanical methods one has to choose in the begining of the calculation the basis set. The use of an adequate basis set is an essential requirement for success of the calculation.

There are several types of atomic orbital function (basis set) :

Slater Type Orbital (STO) :

$$\chi^{STO} = Nr^{(n-1)} \exp(-\beta r)\, Y_{l,m}(\theta, \phi) \qquad (2.18)$$

STOs were originally introduced by Slater (17) as nodeless approximations of hydrogen-like atomic orbitals according to empirical rules. They were mostly used for the calculations of small molecules. Their advantage is that only few functions are needed for a good description, but the integration of higher functions is largely time consuming and increasingly inaccurate due to numerical integration errors.

Gaussian Type Orbital (GTO)(18) are defined as:

$$\chi^{GTO} = N \exp(-\beta r^2) Y_{l,m}(\theta, \phi) \qquad (2.19)$$

GTOs are more suitable and mostly used nowadays. The integration of such functions is easier and can be performed simply. However, due to the different shape of the GTO, for the region near the nucleus, a combination of GTOs with different exponents is required to obtain equivalent results. Contraction is applied to avoid a too large size of the combined basis set. For example, the symbol STO-3G for a basis set means that each STO is approximated by a linear combination of three GTOs, the coefficients being chosen to minimize the difference between the STO and its GTO

approximation. There are many GTO expansions in current use which have been optimized for molecular calculations

Gaussian Lobe Orbitals (GLO)(19,20) are the most simple form of basis functions ;

$$\chi^{GLO} = N \exp(-\beta r^2) \qquad (2.20)$$

The angular part has been omitted, and GLOs located at different points in space are used to reproduce the conventional orbital symmetry. Used in sufficient large number, they also give sufficient accuracy.

An orbital review of basis sets has been given in ref.(21)

In general,one distinguishes minimal and extended basis sets or classifies them according to the STO representations per valence electron type, eg.

Single zeta 1 STO per s,p,d,...function

Double zeta 2 STO per s,p,d,...function

Extended basis sets usually contain also polarization functions ,e.g. p-functions for hydrogen, d and f functions for higher atoms. These functions give a more flexible shape to the molecular orbitals and help them to describe polarization effects related to chemical binding.

2.1.4 ECP (Effective Core Potential)

The tremendous cost of ab initio calculations has motivated many attempts to find computational shortcuts. One such approach is based on the observation that core orbitals are relatively inert to changes in chemical bonding (the so called "frozen core approximation"). Another observation is that the effect of core electrons on the valence electron can be treated through the use of a potential energy term expressed as the sum of local functions multiplied by projection operators. Based on these two assumptions effective core potentials (ECP's) or pseudopotentials, as they are sometimes called,reduce the computational problems to dealing with valence electron

only. Most early results obtained from ECP's compared favorably with results obtained from all- electron calculations, although there was a tendency to find shorter bond lengths and somewhat deeper potential energy curves. However, recently developed ECP's (22,23) have solved these problem so that ECP and all-electron results are now in nearly exact agreement, even with the advantage to include relativistic corrections for the core orbitals of heavier atoms.

The generation of ECP's is performed as following :

1) The "core" orbitals to be replaced and the remaining "valence" orbitals are defined.

2) Numerical valence orbitals (ϕ_l') are obtained from self-consistent Hartree-Fock calculations for l = 0,1,...,L, where L, in general, is one greater than the highest angular momentum quantum number of any core potential.

3) Smooth,nodeless pseudo-orbitals (ϕ_l')are derived from the Hartree-Fock orbitals (ϕ_l') in a manner so that ϕ_l' behaves as closely as possible to ϕ_l in the outer,valence region of the atom.

4) Numerical effective core potentials U_l are derived for each l by demanding that ϕ_l is a solution in the field of U_l with the same orbital energy ϵ_l as the orbital ϕ_l.

5) The numerical potentials are fitted in analytic form with Gaussian functions. The total potential is represented as

$$U(r) = U_L(r) + \sum_{l=0}^{L-1} [\, U_l(r) - U_L(r) \,] \hat{p}_l \qquad (2.21)$$

where

U_L, U_l .. Gaussian functions

p_l .. projection operators

r .. distance between atom

6) The numerical pseudo-orbitals are also fitted to Gaussian functions to obtain basis sets for molecular calculations.

2.2 Potential Functions

The reliability of results obtained from a computer simulation of Monte Carlo or Molecular Dynamics type depends mainly on the quality of the potential functions used. A potential function describes the interaction energy among particles in that system. In general, the total interaction energy of N particles system can be written as ;

$$\Delta E = \sum_{i} \upsilon_1(r_i) + \sum_{i}\sum_{j>i} \upsilon_2(r_i, r_j) + \sum_{i}\sum_{j>i}\sum_{k>j>i} \upsilon_3(r_i, r_j, r_k) + \ldots \qquad (2.22)$$

where the first term in eq. 2.22, $\upsilon_1(r_i)$,represents the effect of an external field(e.g. the container walls) on the system.The second term, υ_2, the pair potential, is the most important one. In practice ,one assumes ΔE to be equal to this term only - known as the **pair-wise additive approximation** - ,whereas the remaining terms are often referred to as **nonadditive corrections**.The third term,υ_3 called three-body interaction sometimes becomes very significant in case of condensed systems. In most of published simulations, however, they were not included, due to the large computer time needed. Four-body (and higher) terms are expected to be small in comparison to υ_2 and υ_3 .

2.2.1 Pair Potentails

The majority of simulations approximate ΔE simply by the pair potential term. So far, there have been several types of pair potentials used in computer simulations. The Lennard-Jones 12-6 pair potential is one of the commonly used simple pair potentials which has the form

$$v^{LJ}(r) = 4\epsilon((\sigma/r)^{12} - (\sigma/r)^{6}) \tag{2.23}$$

This potential has an attractive tail of the form $-1/r^6$, the term $1/r^{12}$ represents repulsive interaction, a negative well of depth ϵ, and a steeply rising repulsive wall at distances less than $r \sim \sigma$.

The other forms of pair potential, which are very simple and convenient to use in computer simulation and in liquid-state theory, are :

a) The hard-sphere potential

$$v^{HS}(r) = \begin{cases} \infty & (r < \sigma) \\ 0 & (\sigma \leq r) \end{cases} \tag{2.24}$$

b) The square-well potential

$$v^{SW}(r) = \begin{cases} \infty & (r < \sigma_1) \\ -\epsilon & (\sigma_1 \leq r < \sigma_2) \\ 0 & (\sigma_2 \leq r) \end{cases} \tag{2.25}$$

c) The soft-sphere potential

$$v^{SS}(r) = \epsilon(\sigma/r)^{\kappa} = ar^{-\kappa} \tag{2.26}$$

where κ is a parameter, often chosen to be an integer. Soft-sphere potentials

contain no attractive part. The illustration of these three potentials are shown in figure 2.1 .

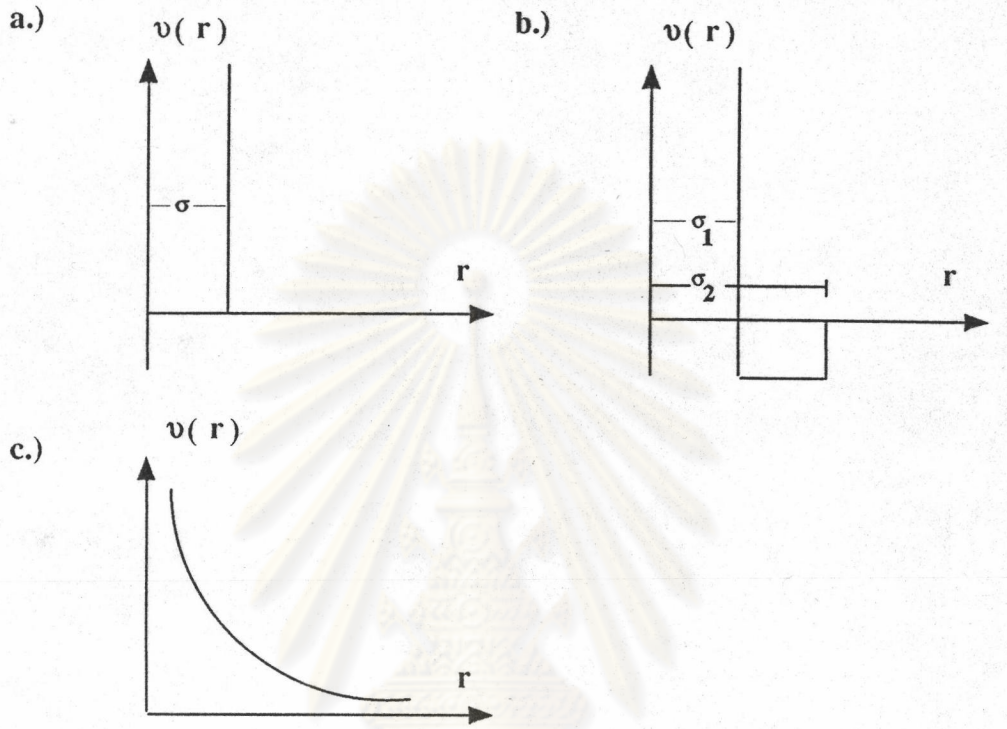

Figure 2.1 : a). The hard-sphere potential; b). The square well potential; c). The soft-sphere potential with replusion parameter $\kappa = 1$.

2.2.2 N-body Corrections

N-body correction are "classically" evaluated form SCF calculations on all possible configurations of a supermolecule consisting of N particles. The resulting energy surface is fitted to an analytical correction potential. The large number of possible configurations and the increasing effort for SCF computations of such large system makes these corrections enormously time-consuming and expensive. For strongly interacing systems at least 3-body terms become essential, however, and the simple pair approximation alone is no more sufficient for obtaining correct structural data in liquid state.

2.2.3 Analytical Form of Potential Functions

In the energy fitting step, a suitable mathematical function is selected. Among the many possible analytical forms of potential functions, one needs to construct the most reliable function in order to use it in Monte Carlo simulation. Given two molecules M and N, the pair interaction potential V(M,N) might contain explicitly both angular and radial dependency. In practical way, however, one selects to use only the radial functions depending on the interatomic distances r(i,j) with i and j being atoms belonging to molecules M and N , respectively. For accuracy and flexibility, one would like to use a fairly long series of terms. On the other hand, the longer the series, the larger the number of associated fitting parameters, and the larger the number of machine cycles required to compute interaction energies in the simulation. In general, the form of potential function can be expressed as

$$V(M,N) = \sum_{ij}\left[\frac{A_{ij}}{r_{ij}^{x}} + \frac{B_{ij}}{r_{ij}^{y}} + C_{ij}\exp(D_{ij}r_{ij}) + \frac{F\,q(i)\,q(j)}{r_{ij}}\right] \qquad (2.27)$$

where A,B,..,F are fitting parameters for atom i and charge q(i) in molecule M and for atom j and charge q(j) in molecule N. The parameter F (often equal to 1) represents an average correction to the charges q(i) and q(j) obtained from computation of molecular wave functions of molecules M and N at infinite separation. It is obvious that the value of the net charge is a function of the internuclear separation and relative orientation of M and N. The exponential term formally describes the short-range inteaction (repulsive interaction).

Once the analytical form of the potential has been chosen, a fitting procedure then will be carried out. The form of the selected analytical function and the algorithm used in the fitting procedure are both important.