CHAPTER II

THEORIES AND METHODS

For the structural investigation of this work, the methodological study can be classified into two parts of the theoretical treatments which are described by the quantum chemical and statistical simulation methods (see Figure 2.1). The quantum chemical theories are used in order to construct an analytical intermolecular interaction energy which is usually called in term "Potential Function". Based on the potential functions, then a Monte Carlo simulation which required the handling of the statistical thermodynamic treatment is carried out in order to evaluate numerical value. The structural study of the interested system will not be completed unless these values are interpreted with the physical-chemical considerations. In this chapter, it is, therefore, involving about how these theories can describe and correspond to the real system. The goal of this chapter is to provide an overview of each method. Thus this text gathers in one place several items which have been described in more details elsewhere [14-16].
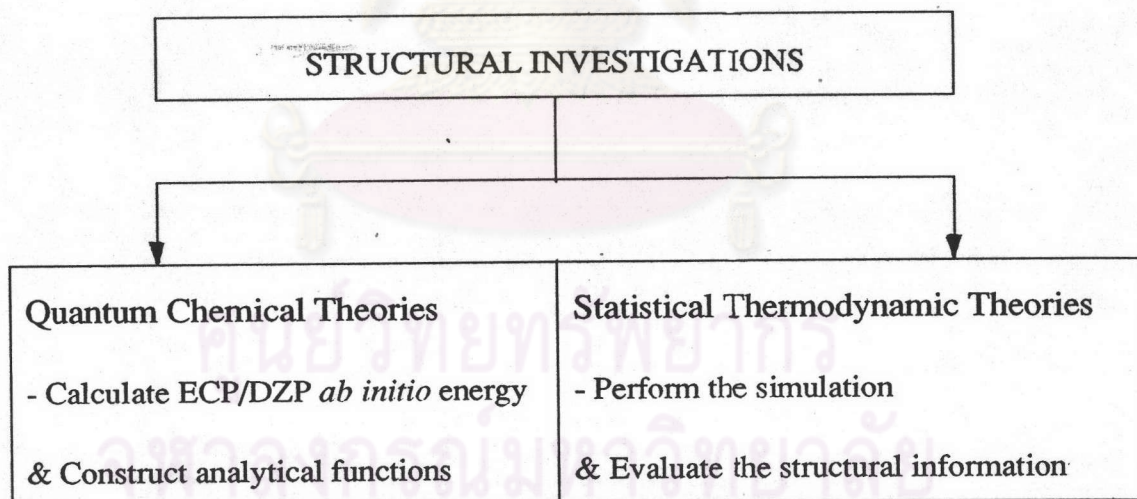

Figure 2.1 The two mainly theoretical treatments for the study of MC method.

Intermolecular Potential Functions

The Monte Carlo (MC) and Molecular Dynamic (MD) simulation techniques have been well established and are used by numerous scientists. The most important factor, the results of such calculations depend undoubtedly on the potential function in use. At the same time, it is necessary to evaluate energy and force values as rapidly as possible in the course of the simulation, due to economy of CPU time.

Given a system of N particles, the total interaction energy can be written as:

$$\Delta E = \sum_{i>j} V(i,j) + \sum_{i>j>k} V(i,j,k) + ... + V(1,2,3,...,n) \qquad (2.1)$$

Where the first term is the two-body term, the second the three-body, and so on till the last term which contains all the bodies of the system. The assumption that ΔE is equal to $\Sigma V(i,j)$ is known as the *pair-wise additive approximation*; the remaining terms of the series are often refered to as *nonadditive corrections*. The basic idea in Eq.*(2.1)* is that, in a many particle system, there are entities recognizable as molecules. It is the quantum mechanical approach that defines an atom or a molecule. In principle, the interaction potential between different molecules can be determined by using quantum chemistry assumptions.

The general procedures for obtaining *ab initio* intermolecular potential function consist of four steps are illustrated in Figure 2.2.

```
┌─────────────────────────────────────┐
│   1. Selection of Conformations      │
└─────────────────────────────────────┘
                  ↓
┌─────────────────────────────────────┐
│   2. Ab initio Computations          │
└─────────────────────────────────────┘
                  ↓
┌─────────────────────────────────────┐
│   3. Selection of functional form    │
└─────────────────────────────────────┘
                  ↓
┌─────────────────────────────────────┐
│   4. Fitting & Testing the function  │
└─────────────────────────────────────┘
```
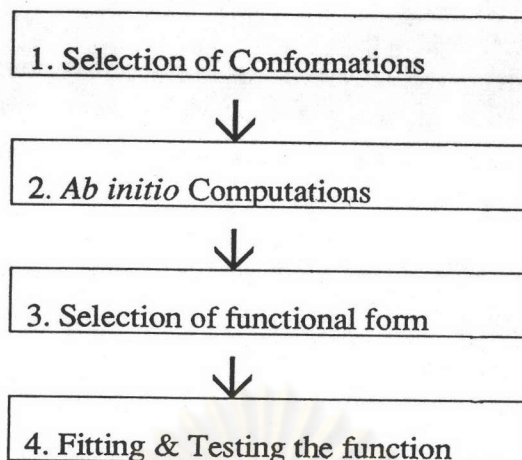
Figure 2.2: The procedure for constructing the function

(Steps 2, 3 are 4 executed more than once.)

To obtain the most accurate results of molecular orbital calculation, several factors must be considered in the construction of potential surface. It should be as efficient as possible, both in terms of time required completing a given task, and in its optimal usage of available computational resources. For instance, the size of basis set used have to be concerned in order to consider an acceptable compromise between computational effort and accuracy. However, the models have to be chosen such as to preserve as close as possible the chemical and physical integrity of the whole.

2.1 Selection of Conformations

The first step involves the selection of the conformations and/or configurations to perform *ab initio* computations. In this task, the graphic drawing represented the geometric conformation can make the complex molecule (ion-molecule complex) easy for looking and advise 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 available in *ab initio* performing.
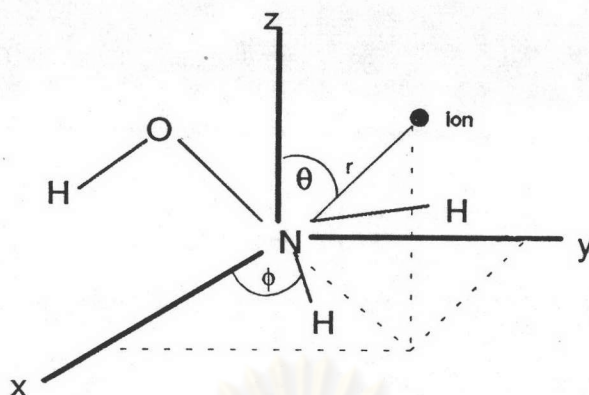
Figure 2.3: Definition of geometric variables (θ,ϕ) for the configuration of ion/NH_2OH

Table 2.1   Coordinates of hydroxylamine molecule (in Å)

Atom	Coordinates		
	x	y	z
H_N	-0.81805	0.60422	0.00000
H_N	0.81805	0.60422	0.00000
N	0.00000	0.00000	0.00000
O	0.00000	-0.55990	1.34079
H_O	0.00000	-1.50331	1.15290

About this work, the positions of the ion illustrated in Figure 2.3 were generated by varying distance $(1 \leq r(\text{Å}) \leq 12)$ and angle $(0 \leq \theta \leq 180, -90 \leq \phi \leq 90)$ whilst the hydroxylamine molecule was fixed within the Cartesian coordinate. By regarding about the molecular structure of hydroxylamine, this molecule can be separated into two kinds of hydrogen atoms. The first one which is adjacent to oxygen atom has been recognized as hydroxyl hydrogen (H_O). Another which bonds with nitrogen atom has been known as an amine group (H_N). The atomic coordinates of hydroxylamine molecule (whose geometry will be presented in the next chapter) were collected in Table 2.1. For the representation of the interaction energy surface, the complex

geometries were selected with special regard to chemically representative conformations and molecular symmetry.

With this selection of conformation, the configurations of Li^{+}-NH_2OH can be produced as many as suitable, and the CPU time reduced to some extent due to the symmetry of hydroxylamine molecule. In addition, the minimum interaction energy can be easily observed by this way.

2.2 *Ab initio* Computations

For the second step, each *ab initio* energy point has been calculated according to its configuration (from the first step), and thus the stabilization energy or the interaction energy has been obtained. An *ab initio* computed has been concerned with the determination of molecular wavefunctions and properties in the framework of the Born-Oppenheimer approximation whereby the molecular wavefunction Φ is written as the product of a nuclear wavefunction X and an electronic wavefunction Ψ which depends on the nuclear coordinates R as parameters, *i.e.* $\Phi(r;R) = \Psi(r;R) \, X(R)$. At all levels of theory the electronic wavefunction $\Psi(r;R)$ is built from molecular orbitals ϕ which are functions of the coordinates of one electron only and are expressed as linear combinations of atomic orbitals χ . The many-electron wavefunction is obtained by solving the time-independent electronic Schrödinger equation,

$$H \, \Psi(r;R) \; = \; E_{elec}(R) \, \Psi(r;R) \qquad (2.2)$$

for a given R where H is the electronic Hamiltonian and E_{elec} is the electronic energy. The square of the wavefunction, Ψ^2, is interpreted as a measure of the probability distribution of the particles within the molecule. The theoretical models presented here are all based on *molecular orbital (MO) theory*. This approximate treatment of electron distribution and motion assign individual electrons to one-electron functions or simplify as the basis function which are the atomic orbitals for the atoms making up the molecule is often described as the linear combination of atomic orbital (LCAO) approximation, and is frequently used in qualitative descriptions of electronic structure.

Given the basis set, the unknown coefficients are determined so that the total electronic energy calculated from the many-electron wavefunction is minimized and , according to the variational theorem, is as close as possible to the energy corresponding to exact solution of the Schrödinger equation.

2.2.1 Theories and Energy Expressions

The electronic wavefunction is written as Slater determinant, an antisymmetrized product of spin-orbital, an ansatz which guarantees that the Pauli exclusion principle is satisfied :

$$\Psi(r_1, r_2, ..., r_n) = (n!)^{-1/2} \sum_P (-1)^P \hat{P}[\phi_1(1)\phi_2(2)...\phi_n(n)] \qquad (2.3)$$

where P is the permutation operator of the electrons among orbitals, and p is the number of transpositions required to restore a given permutation. Note that it is important only whether p is an even or odd number. For a closed shell system in which each spatial orbital ϕ is associated with two electrons with spin α and spin β , the electronic energy is given by

$$E_{elec} = \sum_i 2(i|h|i) + \sum_{ij} \left[2(ii|jj) - (ij|ji) \right] \qquad (2.4)$$

The total energy of the system is equal to the electronic energy plus the electrostatic nuclear repulsion

$$E_{total} = E_{elec} + E_{nuc} \qquad (2.5)$$

$$E_{nuc} = \sum_{A,B}^{Nuc} \frac{Z_A Z_B}{R_{AB}} \qquad (2.6)$$

In Eq. *(2.4)*, $(i/h/i)$ is the matrix element over the orbital ϕ_i of the bare nucleus operator

$$h(1) = -\frac{1}{2}\nabla^2(1) - \sum_{A=1}^{Nuc} \frac{Z_A}{r_{1A}} \qquad (2.7)$$

$$(i/h/i) = \int \phi_i(1)\, h(1)\, \phi_i(1)\, d\tau_1 \qquad (2.8)$$

and (ij/ij) is the electrostatic self-interaction of the $\phi_i(1)\,\phi_j(1)$ charge distribution:

$$(ij/ij) = \iint \phi_i(1)\, \phi_j(1)\, r_{12}^{-1}\, \phi_i(2)\, \phi_j(2)\, d\tau_1\, d\tau_2 \qquad (2.9)$$

The molecular orbitals ϕ are expressed as linear combinations of the atomic orbitals χ according to :

$$\phi = \chi C \qquad (2.10)$$

and are subject to the orthonomality condition :

$$C^\dagger S C = 1 \qquad (2.11)$$

where

$$S_{\mu\nu} = \int \chi_\mu(1)\,\chi_\nu(1)\,d\tau_1 \qquad (2.12)$$

From the application of the variational principle [17], the C coefficients corresponding to the lowest energy are solutions of the Roothaan-Hall [18] equations:

$$F\,C = S\,C\,\varepsilon \qquad (2.13)$$

with

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\rho,\sigma} \mathbf{P}_{\rho\sigma}\left[\ \mu\nu|\rho\sigma - \tfrac{1}{2}\ \mu\rho|\nu\sigma\ \right] \qquad (2.14)$$

in which

$$(\mu\nu|\rho\sigma) = \iint \chi_\mu(1)\,\chi_\nu(1)\,r_{12}^{-1}\chi_\rho(2)\,\chi_\sigma(2)\,d\tau_1\,d\tau_2 \qquad (2.15)$$

$$P_{\mu\nu} = \sum_i^{occ} 2C_{\mu i}C_{\nu i} \qquad (2.16)$$

The matrix of orbital coefficients C is obtained by solving Eq. *(2.13)* which in turn depends on C through F and D. This is the basis of the self-consistent-field (SCF) method. Starting from an inital guess for the molecular orbitals (the bare nucleus hamiltonian may be used, although better guesses can be generated), the density matrix is constructed and combined with the electron repulsion integrals to from the Fock matrix F.

2.2.2 Basis Sets for Computations

The individual molecular orbital can be expressed as linear combinations of a finite set of N prescribed one-electron known as *basis functions* in Eq. *(2.10)*. The molecular orbital expansion coefficients, C, provide the orbital description with some flexibility, but clearly do not allow for complete freedom unless the χ define a complete set. However, the problem of finding the orbitals is reduced from finding complete descriptions of the three-dimensional function ϕ to finding only a finite set of linear coefficients for each orbital.

Two types of atomic basis functions have received widespead use. The first one is *Slater-type atomic orbitals* (STOs) which have exponential radial parts

$$\chi^{STO} = Nr^{(n-1)}\exp(-\zeta r)Y_{l,m}(\theta,\phi) \tag{2.17}$$

where ζ is a constant determining the size of the orbitals. STOs provide reasonable representations of atomic orbitals with standard ζ-values recommended by Slater [19]. They are, however, not well suited to numerical work, and their use in pratical molecular orbital calculations has been limited.

The second type of basis consists of *Gaussian-type atomic functions* [20]. Gaussian-type orbitals are defined as:

$$\chi^{GTO} = Nr^{(n-1)}\exp(-\alpha r^2)Y_{l,m}(\theta,\phi) \tag{2.18}$$

where α is a constant determining the size, that is, radial extent, of the function. Gaussian-type functions are less satisfactory than STOs as representations of atomic orbitals, particularly because they do not have a *cusp* at the origin. Nevertheless, they have the important advantage that all integrals in the computations can be evaluated explicitly without recourse to numerical integration.

By *LCAO-MO* method the molecular orbitals will be constructed from basis functions. A limiting Hartree-Fock treatment would involve an infinite set of basis

functions χ. Moreover, the computational expense of Hartree-Fock molecular orbital calculations is formally proportional to the large total number of basis functions. Therefore, the ultimate choice of basis set size depends on a compromise between accuracy and efficiency. The size of the basis sets became an important consideration for theoretical chemists.

2.2.3 Effective Core Potential (ECP)

The tremendous cost of *ab intio* 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 "forzen 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. Based on these two assumptions, they are given the name as "*effective core potentials (ECP) or pseudopotentials*", reduce the computational problems, especially CPU times, to dealing with valence electron only.

The ECP approach was developed by Kahn *et al.* [21] to replace the inner shell electrons of atoms by the effective potential. This is in line with the accepted idea that in a molecule the electronic structure is dictated mostly by the valence electrons, and that the core orbitals resemble the atomic core orbitals. Therefore, ECP and all-electron results are now acceptable in accurate, even with the advantage to include relativistic corrections for the core orbitals of this study. Moreover, potential and associated valence basis sets have been developed for most of the periodic table, and are tabulated in the program just like the regular basis sets. In particular the potentials and basis set of Stevens *et al* [22]. can be invoked in HONDO by mean of a key work for ease of use, although any other potentials and associated basis set can be used.

2.2.4 Basis Set Superposition Error (BSSE)

When insufficient basis sets are used, an artificial basis set improvement will take place in the complex, leading to an error which has been known as *basis set superposition error*, and the interaction energies are always overestimated [23]. The intermolecular interaction ΔE is obtained as the difference of the separately calculated complex and constituent energies

$$\Delta E = E_{AB}(\chi_{AB}) - [E_A(\chi_A) + E_B(\chi_B)] \tag{2.19}$$

$$\text{where} \quad \chi_{AB} = \chi_A \cup \chi_B$$

Here E_A, E_B, and E_{AB} are energies of constituent monomers A, B and the complex AB, respectively. χ_A and χ_B represent basis sets of monomers A and B. The intermolecular interaction acquired from Eq. *(2.19)* will contain, in addition to the "*real interaction*", the "*non-physical*" energy contribution called as "*basis set superposition error*", which arises from the lack in completeness of the basis set.

One method which has been widely used to estimate BSSE is the "*counterpoise (CP) correction*" proposed by Boys and Bernardi [24]. The CP correction is evaluated from the energy improvement for a free fragment (monomer) of a molecule (complex) when a "*ghost*" basis for the other fragment is added to the calculation. In other word, the monomer energies are evaluated in a dimer basis set, and exactly the same basis is used for both the monomers and dimer as in Eq. *(2.20)*

$$\Delta e = (E_A(\chi_A) - E_A(\chi_{AB})) + (E_B(\chi_B) - E_B(\chi_{AB})) \tag{2.20}$$

and the counterpoise corrected interaction energy is determined from

$$\Delta E^{cp} = \Delta E + \Delta e \tag{2.21}$$

where ΔE is interaction energy which calculate from truncated basis set

2.2.5 Obtaining the Interaction Energy (ΔE_{SCF})

In SCF calculations of the interaction energy ΔE_{SCF} between two systems (A, B) based on the molecular orbital method, the value of ΔE_{SCF} is determined as the difference between the supersystem energy E_{AB} and the sum of the subsystem energies (E_A, E_B).

$$\Delta E_{SCF}(kcal/mol) = (E_{AB} - E_A - E_B) \times 627.5 \qquad (2.22)$$

where

E_{AB} = the supersystem energy (in Hartree)

E_A, E_B = the isolated system or subsystem energy A and B, respectively (in Hartree)

In this work the interaction energy is always obtained by this expression, which seems to be quite simple. However, serious difficulties are encountered in the numerical treatment. The most important difficulty lies in the fact that the supersystem energy and the sum of the energies of the isolated subsystems are very large number compared to the interaction energy. Therefore, the determination of the energy of the systems must be extremely accurate (10^{-5} Hartree). A further 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 polarization functions would be necessary for this.

The interaction energy for each configuration will be used as an input file in the step of fitting potential function

2.2.6 Mulliken Population Analysis

The *electron density function* or *electron probability distribution function*, $\rho(r)$, is a three-dimensional function defined such that $\rho(r)$ dr is the probability of finding an electron in a small volume element, dr

$$\int \rho(r)\, dr \;=\; n \tag{2.23}$$

where n is the total number of electrons. For a single-determinant wavefunction in which the orbitals are expanded in terms of a set of N basis functions, χ_μ, $\rho(r)$ is given by Eq. *(2.24)*,

$$\rho(r) = \sum_\mu^N \sum_\nu^N P_{\mu\nu} \chi_\mu \chi_\nu \tag{2.24}$$

According to the Eq. *(2.12)*, suggestions about how to calculate the electron density, starting from the density matrix ($P_{\mu\nu}$), were made by Mulliken [25].

$$n = \sum_\mu^N \sum_\nu^N P_{\mu\nu} S_{\mu\nu} \tag{2.25}$$

It is desirable to allocate the electrons in some fractional manner among the various parts of a molecule (atoms, bonds, etc.). It will be useful, for example, to define a total electronic charge on a particular atom in a molecule in order that quantitative meaning may be given to such concepts as electron withdrawing or donating ability. In addition, it will be used to represent the interaction potential of the Coulombic force.

2.3 Selection of Functional Form

For the computer simulation of liquids [26], realistic pair potentials between the molecules involved are needed. In the third step, a suitable mathematical function is selected. The analytic potential function can represent the interaction energy between inter-atomic distance. It was assumed in the term called "*pair wise additivity*", so that the total intermolecular interaction energy was written as a sum of atom-atom pair potentials, each of which was considered to be function of the distance r_{ij} of the two atoms concerned, that is

$$\Delta E_{intermolec} = \sum_i \sum_j V_{ij}(r_{ij}) \qquad (2.26)$$

where i : atom ith in first molecule

j : atom jth in second molecule

Of course the true intermolecular potential function is unknown, but the functional form of the pair potentials V_{ij} was chosen with a physical model in mind: there should be a term to describe the *Coulombic interaction* between the atom pairs, and the *Mulliken population analysis* of the *ab initio* calculations was used to get a first estimate of the fractional atomic charges. Then, additional terms of the form B/r^m, *m=2,...6*, (see Eq. *(2.27)*) especially effective at medium distances, should be added to describe phenomena like the distance dependence of these charges, and the mutual disturbance of the charge distributions which results in induced dipole moments and their interaction with the above points charges as well as between them. Finally, there should be a term to represent the electronic repulsion at close distances and take care of the *Pauli exclusion principle*, which effect increases nucleonic repulsion. This term is typically represented by either a polynomial A/r^n, *n = 8 ... 15* or an exponential, $\exp\{-r_{ij}\}$ functional. Therefore a trial potential function ΔE_{ij} was set up as

$$\Delta E_{ij}(r_{ij}) = \frac{A_{ij}}{r_{ij}^n} - \frac{B_{ij}}{r_{ij}^m} + \frac{C_{ij}q_i q_j}{r_{ij}} \qquad (2.27)$$

with the coefficients A_{ij}, B_{ij}, and C_{ij} to be determined so that a best fit of 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 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, and an arbitrary choice of the C_{ij} coefficients may not reproduce the correct limit of the electrostatic energy at large distances.

2.4 Fitting and Testing the Function

The fitting procedure was carried out by minimizing the sum of squared differences between the *SCF* calculated energies and those given by the trial potential function, with respect to the potential parameters. The minimization itself was based on a *multidimensional non-linear Marquard-Levenberg algorithm* [27], which changes softly from an initial steepest descent to a quadratic approximation of the potential parameters near the minimum. 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.

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 fitted energies versus their quantum mechanically calculated values. Moreover, the predictive capabilities of the potential function were then tested according to the procedure suggested by Beveridge et al [28]. Namely, the *SCF* energies (outside the original set) were calculated and compared with the values predicted by the function. These points were then included in the fitting procedure and the whole process was repeated until satisfactory results were obtained.

The Monte Carlo (MC) techniques are generally used to evaluate numerically, by means of random sampling (hence the term "Monte Carlo"), multidimensional integrals which are too difficult to solve with analytical methods. These methods can be applied also to study properties of condensed matters, where the thermodynamical averages obtained over probability distributions are interesting.

The aim of Monte Carlo simulations is to study the microscopic properties of the solution, such as the structural and energetic properties, based on the knowledge of potential functions; for example, how a solute influences the solvent structure or how a solute is solvated by solvent molecules in the solution. This is particularly important for the study of the systems which are difficult to handle in laboratory, because of their instability or the high degree of dilution which are very difficult to observe from spectroscopic measurements.

2.5 Basic Principles of Monte Carlo Methods

Most Monte Carlo calculations in statistical mechanics have been performed using the canonical or (T, V, N) ensemble of Gibbs in which the number of N molecules, the volume V and the temperature T are fixed. Classical statistics is assumed only two-body forces in the potential energy $E(\upsilon)$ calculations. The total energy of the system can be written thus as a sum of pairwise interaction energies between the individual particles of its system,

$$E(\upsilon) = \frac{1}{2} \sum_{i=1, i \neq j}^{N} \sum_{j=1}^{N} V_{ij}(\upsilon) \qquad (2.28)$$

where υ is a configuration coordinate of the particles in a system. The canonical ensemble used for evaluating the average values $<F>$ of the property F of the system can be expressed as,

$$\langle F \rangle \quad = \quad \frac{\int...\int F \, exp\,(-E(\upsilon\ )/kT)\, d\upsilon}{\int...\int \, exp\,(-E(\upsilon\ )/kT)\, d\upsilon} \qquad\qquad (2.29)$$

The above was modified by Metropolis et al in 1953 [4]. In this method, the N configurations are randomly generated but accepted with respect to their probability, $P(\upsilon)$, and they are weighted with $P(\upsilon)$. After taking a quite large number of space points M, Eq.(2.29) can be approximated by the sum,

$$\langle F \rangle \cong \overline{F} = \frac{\displaystyle\sum_{i=1}^{M} F(\upsilon) P^{-1}(\upsilon)\exp(-E(\upsilon)/kT)}{\displaystyle\sum_{i=1}^{M} P^{-1}(\upsilon)\exp(-E(\upsilon)/kT)} \qquad\qquad (2.30)$$

The probability $P(\upsilon)$ in this Metropolis Monte Carlo method is given by a Boltzmann factor:

$$P(\upsilon) = \exp(-E(\upsilon)/kT) \qquad\qquad (2.31)$$

Then Eq.(2.30) can be reduced to a simple form of

$$\overline{F} = \frac{1}{M}\sum_{i=1}^{M} F_i \qquad\qquad (2.32)$$

where F_i is the value of the property F of the system after the ith configuration change.

A key idea in this method is the use of periodic boundary conditions to enhance the ability of small systems to simulate the behavior of large systems. This technique is to consider a certain basic region, usually a cube, containing a certain number N of molecules; and then to imagine the whole of space filled by periodic images of this basic unit. In this way, one can consider configurations of an infinite system (which must of course be periodic) while only considering a limited number N of molecules. The great advantage is that *surface effects*, which would otherwise be very large for small N, are avoided. Often one uses the "*nearest image distance convention*" according to which a given molecule i is supposed to interact only with that periodic image of another molecule j which lies closest to i. In fact, if the range of the molecular interactions is less than half the edge of the cube this includes all interactions; it is convenient to force this by truncating the potential at some distance.

2.6 Conditions of the MC Simulations

2.6.1 Defining The Cubic Size

The cubic box has been often used for holding all particles investigated. (the reason will be explained in the topic "*periodic boundary condition*"). The length of the cubic side $,L$ $(\mathring{A})$, which consists of the N_p particles for m species and the atomic/molecular weight M_p is calculated by

$$L^3 = \frac{\sum_{p=1}^{m} N_p M_p}{0.602 D} \tag{2.33}$$

where D is the density of the solution in g $.cm^{-3}$ for the simulation at temperature T and pressure P.

2.6.2 The Amount of Particles for Simulation

The number of particle in the system investigated depend on the computer efficiency. The result will be more accurate whenever an increasing number of particle is performed. In addition, the CPU time has to be considered between accuracy and time cost. For several studies, they have been shown that the system would consist of about 200 particles for the most suitable condition.

2.6.3 Defining The Starting Configuration

For the beginning of MC simulation, the configuration has offen been generated randomly. However, it would be concerned that the starting configuration should be throughout the cube, for this reason, particles will distribute around the container. Therefore, the system has been equilibrated so fast. The position of each particle can be evaluated by regarding Eq. *(2.34)*

$$l_{min} = \sqrt[3]{\frac{L^3}{N}} \qquad\qquad (2.34)$$

where l_{min} is the shortest distance between particles existed

2.6.4 Periodic Boundary Condition

The problem of surface effects can be overcome by implementing periodic boundary conditions. The cubic box is replicated throughout space to form an infinite "lattice". In the simulation, as a molecule moves in the original box, its periodic image in each of the neighboring boxes moves in exactly the same way. Thus, as a molecule leaves the central box, one of its images will enter through the opposite face. There are no walls at the boundary of the central box, and no surface molecules. This box simply forms a convenient axis system for measuring the coordinates of the N molecules. A two-dimensional version of such a periodic system is shown in Figure 2.4.

The duplicate boxes are labeled A, B, C, etc. . As particle 1 moves through a boundary, its images, $1_A, 1_B$ etc. (where the subscript specifies in which box the image lies) move across their corresponding boundaries. The number density in the central box (and hence in the entire system) is conserved. It is not necessary to store the

coordinates of all the images in a simulation (an infinite number), just that of the molecules in the central box. When a molecule leaves the box by crossing a boundary, attention may be switched to the image just entering. It is sometimes useful to picture the basic simulation box (in the two dimensional examples) as being rolled up to form the surface of a three-dimensional doughnut, when there is no need to consider an infinite number of replicas of the system nor any image particles. This correctly represents the topology of the system, if not the geometry A similar analogy exists for a three-dimensional periodic system.
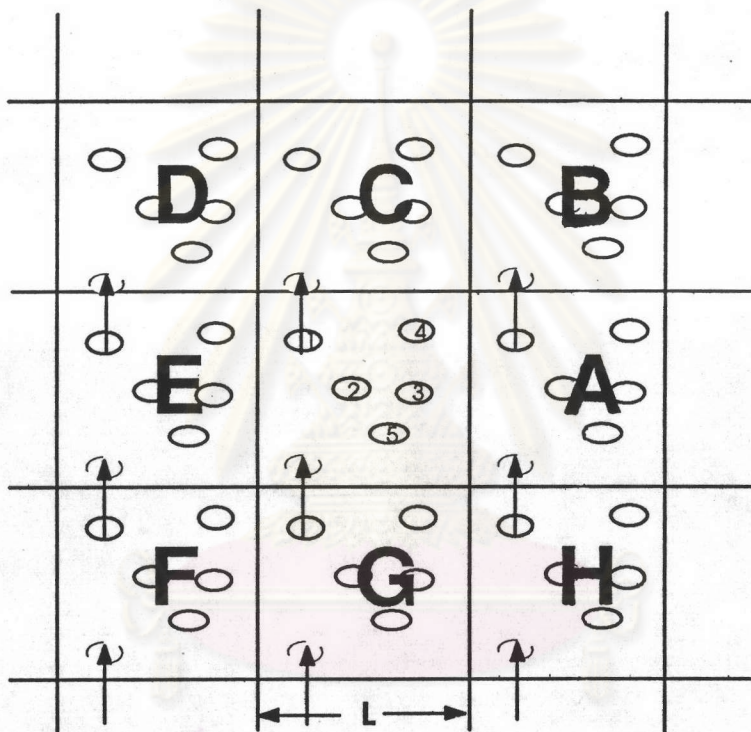

Figure 2.4 : A two-dimensional periodic system.

2.6.5 Spherical Cut-off

In the minimal image convention, then, the calculation of the potential energy due to pairwise-additive interactions involves ½ *N(N-1)* terms. This may still be a very substantial calculation for a system of 1,000 particles. A further approximation significantly improves this situation. The largest contribution to the potential comes from neigbors close to the molecule of interest, and for short-range interactions, a spherical cutoff can be applied. This means setting the pair potential $V(r_{ij})$ to zero for

$r_{ij} \geq r_c$, where r_c is the cutoff distance. The cutoff distance must be no greater than L/2 for consistency with the minimal image convention, and it is applicable only to rapidly decreasing potential terms, e.g. $1/r^6$ or exponential terms.

2.6.6 Long-range Interactions

A long range interaction is defined as one in which the spatial interaction falls off no faster than r^{-d} where d is the dimensionality of the system. In another meaning, long-range interaction is one in which all the particles lying outside the cutoff sphere of a given particle. The charge-charge, charge-dipole, dipole-dipole and charge-quardrupole interactions are the examples of such a strong interaction. Long-range interactions are a serious problem for the computer simulator, since their range is greater than half the box length. So far, *The Ewald summations* [29] and *the reaction field method* [30] have been two widely used methods which can be employed to handle the problem of long-range interaction. *The Ewald summations* procedure which is perfectly consistent with periodic boundary conditions is a technique for efficiently summing the interaction between a particle and all its periodic images. *The reaction field* method introduced without the assumption of the periodicity is to treat all molecules beyond the cutoff sphere of radius from a center charge as forming a continuum with a given dielectric constant; therefore, any charge lying inside the cutoff sphere will polarize the continuum and create a reaction field at the center.

2.7 Steps of Calculations

Assuming N particles in a given configuration and in a given volume at T Kelvin, the Metropolis algorithm consists essentially in iterating the following steps which were illustrated in Figure 2.5

1) Place the initial N particles in any configurations (υ) according to random numbers generated by computer.

2) Calculate the interaction energy ($E(\upsilon$)) based on the potential function obtained from section **2.4** .

3) Move each of the particles (υ --> υ ') in succession according to the following prescription:

$$X \rightarrow X + a\xi_1$$

$$Y \rightarrow Y + a\xi_2$$

$$Z \rightarrow Z + a\xi_3$$

where a is the maximum allowed displacement, which for the sake of this argument is arbitrary, and ξ_1, ξ_2 and ξ_3 are random numbers between (-1) and 1. The maximal displacement within the limit 0.5 for this work will be chosen with some concerning; if it is too large, most of the moves will be forbidden, if too small, the configuration will not change enough. In either case it would take longer to reach equilibrium. If the particles are moved to a new position, one at a time, and a sphere after such a move happens to overlap with another sphere. The particle i will be returned to its former position, and repeat this step.

4) Calculate the new configuration energy ($E'(\upsilon)$) and the change in energy of the system ΔE, which is caused by the move.

5) If $\Delta E < 0$, the move would bring the system to a state of lower energy, the move is allowed and put the particle in its new configuration.

6) If $\Delta E > 0$, the move is allowed with probability $P(\upsilon) = exp(-\Delta E/kT)$, where k is the Boltzmann constant and T is the temperature in Kelvin. A random number ξ_4 is taken between 0 and 1.

7) If $P > \xi_4$, move the particle to its new position, the next procedure is performed by returning to step 3)

8) If $P < \xi_4$, return it to its old position. Then, having attempted to move a particle again by repeating step 3)
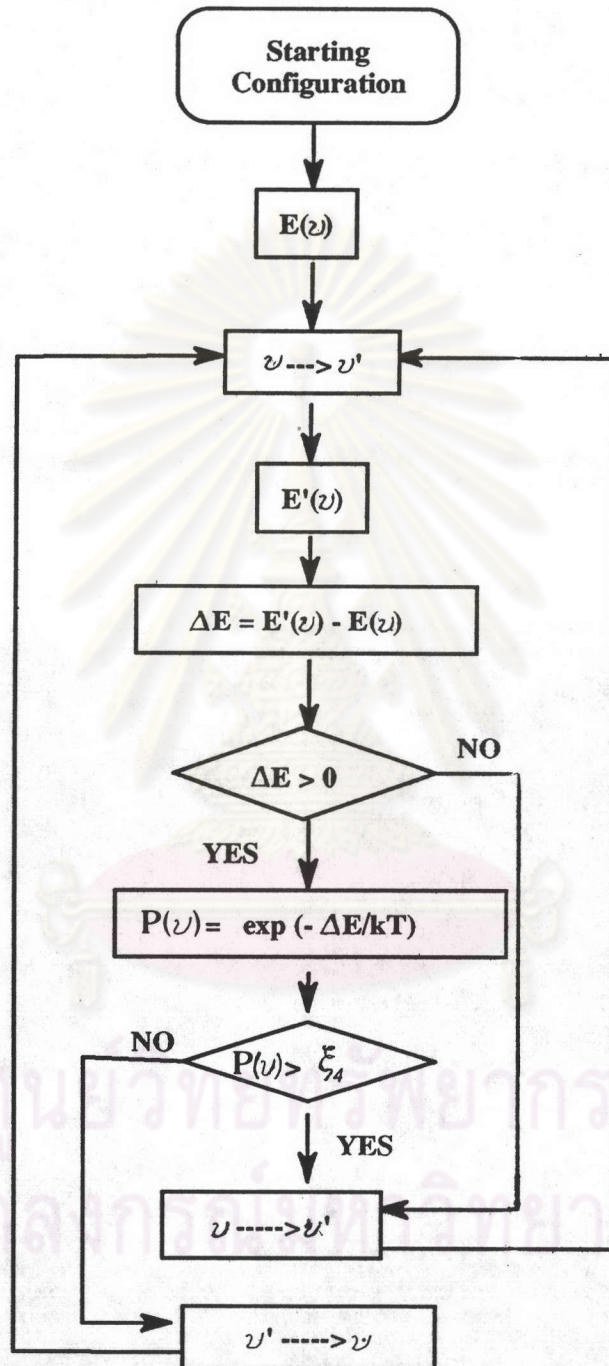
Figure 2.5 Procedure Diagram for MC Simulation

2.8 Radial Distribution Functions (RDFs) and the Integration Numbers

The most common way to analyze the structure of solution is to employ the radial distribution functions ($g(r)$) and the corresponding-running integration number $n(r)$ around various atoms. The radial distribution gives information about the configurationally averaged deviation of the local environment of particle from the values characteristic for the bulk. This function for the N particle system in configuration υ is obtained with r_{ij} as:

$$g(r) = \frac{N(r)}{\rho 4\pi r^2 dr} \tag{2.35}$$

where $N(r)$ is the average number of particles in the spherical shell of width dr at a radial distance r from the central particle. ρ is the number density of the system. Based on the radial distribution functions, solvation shells can be derived from the peaks pronounced over the standard level, and the first solvation number is obtained by the intagration of the function up to the first minimum.

The average number of particles p within a sphere of a given radius can be determined by:

$$n(r) = \rho \int_{o}^{r_m} g(r) 4\pi r^2 \, dr$$

where r_m is often chosen as the radial value of the first or second minimum in $g(r)$.