7816706

SWAMINATHAN, SUNDARAMOORTHI
THEORETICAL STUDIES OF WATER AND DILUTE AQUEOUS SOLUTIONS.

CITY UNIVERSITY OF NEW YORK, PH.D., 1978

THEORETICAL STUDIES OF WATER AND DILUTE AQUEOUS SOLUTIONS

By

S. SWAMINATHAN

A dissertation submitted to the Graduate Faculty in Chemistry in partial fullfillment of requirements for the degree of Doctor of Philosophy, The City University of New York.

1978

This manuscript has been read and accepted for the Graduate Faculty in Chemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

May 17, 1978
Date

David L. Beveridge
Professor David L. Beveridge
Chairman of Examining Committee

May 17, 1978
Date

Leonard H. Schwartz
Professor Leonard H. Schwartz
Executive Officer

Bernard J. Bulkin
Professor Bernard J. Bulkin

A. M.
Professor Avigdor Ronn

Supervisory Committee

The City University of New York

Abstract

THEORETICAL STUDIES OF WATER AND DILUTE AQUEOUS SOLUTIONS

By

S. Swaminathan

Adviser: Professor David L. Beveridge

Monte Carlo computer simulation of liquid water has been performed using a quantum mechanical potential function that gives a radial distribution function in close accord with experiment. The results of the simulation were analysed in terms of quasicomponent distribution functions (QCDF). The predominance of four-coordinate quasicomponent in the diffusionally average structure of the system is confirmed. Analysis of the results in terms of the distribution of binding energies favors the idea of a continuous distribution structures rather than a mixture model of liquid water, and included in the continuum of structures are contributions from 3,5 and 6 coordinate species as well as the four-coordinate component of "Pople-water". A fully ab initio calculation of the free energy of liquid water has been performed. A heuristic method has been described for the determination of intermolecular potential functions from quantum mechanical calculations and applied to the formaldehyde-water and methane-water systems. Monte carlo computer simulation on the dilute aqueous solution of methane at 25 C using quantum mechanical potential functions reveal the local solution environment to be that of a distorted defective continuum pentagonal dodecahedral clathrate structure. Increased four-coordination and stronger binding are induced in the solvent water in comparison with the bulk liquid. The use of QCDFs is seen to be a powerful means for the analysis of structural chemistry of the statistical state of the system.

Acknowledgements

It has been a privelege to have had the opportunity to study with Professor David L. Beveridge over the past five years. His contribution to the completion of this study and my own scientific development during that process is immeasurable. His effusive enthusiasm for research and the lively discussions we have had, made this work possible. His active support and constant encouragement are greatfully acknowledged.

The support and help received from my colleagues, Mr. Elias Guth and Drs. Gary Schneulle, Mihaly Mezei, Shirley Harrison, Rosemary Whitehead, Hans J. Nolte and P.K. Mehrotra are greatly valued.

The advice and support of Profs. B.J. Bulkin, L. Massa and A. Ronn made this study possible.

The support of my brothers, sisters and mother was instrumental in my undertaking this study.
This study is dedicated to them.

Table of Contents

I INTRODUCTION

The objective of this project is to study the structure of liquid water and aqueous solutions using statistical mechanical computer simulation. The advent of third generation digital computer hardware together with recent advances in computational chemistry based on molecular quantum mechanics and statistical mechanics combine to make the structure of molecular liquids accessible to theoretical study at new levels of rigor.

Water has been studied for a long time using statistical mechanics based on ad hoc models[1]. In the ad hoc methods, a statistical mechanical model is developed and thermodynamic properties are expressed in terms of disposable parameters. The disposable parameters are fitted to experimental data and therefore agreement with experiment is not a real test of the model. This has resulted in a controversy over the evaluation of the various models proprosed so far.

Ab initio theories of liquids assume a certain form for the configurational energy and obtain rigorously the distribution functions using analytical or numerical methods based on classical statistical mechanics. There remain significant assumptions in the ab initio approach to molecular liquids since only pairwise potentials are known well, but the assumptions enter at a lower point in the theoretical heirarchy than in ad hoc methods. There is a comprehensive

review of the various ab initio theories of liquids by Barker and Henderson.[1]

The method of Integral Equations suggested independently by Yvon[3] and Kirkwood[4] is one of the early ab initio methods suggested to study the structure of liquids. This method involves finding an analytical or numerical solution to an integral equation obtained using a closure relationship between pair correlation function and the three particle distribution function. The more recent methods of this type are the Percus-Yevick approximation[5] and the hyper-netted chain approximation[6]. The integral equation methods are well-behaved mathematically for weak, spherical potentials only up to moderate densities and thus have a limited range of applicability.

An alternative approach is based on perturbation theory. This method was originally used by Zwanzig[7] to rederive van der Waals' theory. This method involves partitioning the interaction potential into the repulsive or the hard part and the attractive or the soft part. The solution is obtained using the soft part as a perturbation on the result for the hard part. The presently well studied versions of this method are due to Barker and Henderson[8] and Weeks,Chandler and Anderson[9]. This method cannot also be easily used to study liquids like water, since the soft or attractive part of their interaction potentials is relatively large.

The computer simulation methods presently offer a lot of promise in the study of the structure of molecular liquids and

solutions. The most fundamental approach from this point of view is to treat each system as simply an assembly of molecules interacting via a configurational potential, and attempt to solve the problem in classical statistical mechanics or kinetic theory. The statistical mechanics computer simulation is effectively a numerical integration of the configurational integral. The kinetic theory method requires the simultaneous solution to Newton-Euler equations to calculate the individual molecular trajectories after which equilibrium properties are computed by time averaging. These two methods are widely known as the Monte Carlo and Molecular Dynamics computer simulations, respectively. The computer simulation methods are tractable even for systems with a strong potential and another advantage of the simulation methods is the possibility of computer experiments to elucidate the effects of various definable characteristics on the results.

In the method of Molecular Dynamics, direct solutions to equations of motion of a system of N particles of O(100) in a box with periodic boundary conditions are obtained. This method was first used by Alder and Wainwright[10] for simple liquids. Rahmann and Stillinger[11] applied this to simulate liquid water using an empirical potential, and this was the first simulation of a molecular liquid using Molecular Dynamics. Recently, MacDonald and Klein[12] simulated liquid ammonia using a central force model potential. One of the major advantages of the Molecular Dynamics method is that one

can calculate various correlation functions and even non-equilibrium properties. A comprehensive review of this method for simple fluids is due to Alder and Hoover[13] and the review by Kushick and Berne[14] details it for anisotropic fluids.

The Monte-Carlo method[15] is an integration over the configuration space of N molecules to obtain ensemble averages with Boltzman weighting, using an importance sampling method suggested by Metropolis et al.[16] The sampling involves a stochastic walk through the low energy regions of configuration space. This method was first applied to molecular liquids by Barker and Watts.[17] They simulated liquid water using an empirical potential. Later, Watts[18] made a study of liquid water using various potentials. Clementi and co-workers determined a water-water potential representative of self consistent field calculations using a large basis set,[19] and also analytical potentials including electron correlation energy obtained empirically[20] and using configuration interaction.[21] The analytical potential obtained using configuration interaction calculations produced a radial distribution function in good agreement with the experimental result.[22] The Monte Carlo method is described in further detail in the following chapter and a detailed review has been done by Wood.[23]

The availability of a good analytical potential able to reproduce the two particle distribution in liquid state has made the study of the structure of water and aqueous solutions based on this potential an exciting prospect. The research

described herein uses the Monte Carlo simulation incorporating this potential to study the structure of liquid water and dilute aqueous solutions. In view of the diversity involved in the whole study each chapter will independently include the relevant introductory material needed. A brief resume of what is to come in each chapter is given below.

In Chapter II, the statistical mechanical basis of the project is developed, followed by a description of the theory and methodology involved in each calculation. Chapter II concludes with a survey of the analytical tools used to analyze the structure of the system.

A detailed study of liquid water is reported in Chapter III. This chapter will describe two simulations of liquid water in the canonical ensemble using two different analytical potential functions. The results of these two simulation will be used to evaluate the various ad hoc models for liquid water that have been proposed so far. A calculation of Helmoltz free energy for liquid water using Kirkwood's method is also described.

The development of a reliable solute-solvent analytical potential for a simulation of a dilute aqueous solution is discussed in Chapter IV. This chapter reviews the field of potential surfaces for intermolecular interactions and describes a new methodology that has been developed to obtain analytical potentials. The latter part of this chapter will deal with the application of this methodology to develop pairwise potential functions for the formaldehyde-water and

methane-water interactions.

The statistical thermodynamic computer simulations of a dilute aqueous solution of methane form Chapter V. This chapter will deal with the description and results from three different simulations using three different potentials. The development of a methodology to study the influence of solute on solvent structure will be discussed.

II THEORY AND METHODS

a) Classical Statistical Mechanics

Motional degrees of freedom of molecules in liquids mandates theoretical studies of liquids to be problems in statistical mechanics. The theoretical description of a system of N molecules in a volume V at a temperature T follows from the partially classical partition function for the (T,V,N) canonical ensemble,

$$Q(T,V,N) = \frac{q^N}{(8\pi^2)^N \Lambda^N N!} \int \cdots \int \exp\left(-\beta E\left(\underset{\sim}{X}^N\right)\right) d\underset{\sim}{X}^N \qquad (1)$$

Here q is the internal partition function and Λ is the one-dimensional translational partition function for a single particle. The integration ranges over all configurational coordinates $\underset{\sim}{X}^N$ of the N particle system,

$$\underset{\sim}{X}^N = \left\{ \underset{\sim}{X}_1 , \underset{\sim}{X}_2 , \cdots , \underset{\sim}{X}_N \right\} \qquad (2)$$

where $\underset{\sim}{X}_i^N$ represents the configurational coordinates of the i-th particle consisting of a specification of position $\underset{\sim}{R}_i$ and orientation Ω_i. $\underset{\sim}{X}_i = (\underset{\sim}{R}_i , \underset{\sim}{\Omega}_i)$. The quantity $E(\underset{\sim}{X}^N)$ is the configurational energy of the system with β representing $(kT)^{-1}$. The multiple integration in equation 1 is referred to as configurational averaging. Any average property of the system is expressed as a configurational average. i.e. if $F(\underset{\sim}{X}^N)$

defines the property of the system in configuration $\underset{\sim}{X}^N$, then $\overline{F}$, the configurational average for that property is written as

$$\overline{F} = \int \cdots \int F(\underset{\sim}{X}^N)\, P(\underset{\sim}{X}^N)\, d X^N \qquad (3)$$

where $P(\underset{\sim}{X}^N)\, d\underset{\sim}{X}^N$ is the probability of finding the system in configuration $\underset{\sim}{X}^N$,

$$P(\underset{\sim}{X}^N) = \frac{\exp(-\beta E(\underset{\sim}{X}^N))}{\int \cdots \int \exp(-\beta E(\underset{\sim}{X}^N))\, d\underset{\sim}{X}^N} \qquad (4)$$

Monte Carlo computer simulations focus on the determination of average properties of the system such as configurational internal energy U,

$$U = \int \cdots \int E(\underset{\sim}{X}^N)\, P(\underset{\sim}{X}^N)\, d X^N \qquad (5)$$

Expressions for other average properties of the system such as distribution functions can be formulated in a manner analogous to equation 3.

b) Computer Simulation and Monte Carlo Methodology

Computer simulation in the Monte Carlo[15] sense is the numerical integration of equation 5.

$$U = M^{-1} \sum_{i=1}^{M} P(\underset{\sim}{X}_i^N)\, E(\underset{\sim}{X}_i^N) \qquad (6)$$

with the M configurations $\underset{\sim}{X}_i^N$ generated at random. In practice

the numerical process represented by equation 6 is known to be very slowly convergent. An expediant suggested by Metropolis et. al. involves carrying out the calculation by means of a stochastic walk through the configuration space, selecting points $\underset{\sim}{X}_j^N$ which enter the averaging process with a frequency $P(X_j^N)$. The determination of average quantities reduces to a simple summation, i.e.

$$U = M^{-1} \sum_{j=1}^{M} E(\underset{\sim}{X}_j^N) \qquad (7)$$

where the configurations $\underset{\sim}{X}_j^N$ are chosen by the Metropolis method.

The Metropolis method[16] is an importance sampling method, involving a stochastic walk through the configuration space and the configurations are chosen according to their probability. The method as it was described by Metropolis et.al. is the following:

First, an initial configuration $\underset{\sim}{X}_i^N$ is chosen, such as for example the molecules deployed randomly in crystal lattice sites. Then one of the molecules is moved and one of its axes rotated randomly. This creates a new configuration $\underset{\sim}{X}_j^N$. The decision whether to include the old configuration again or to accept the new configuration is taken on the basis of their relative probabilities. If the energy of the new configuration $E(\underset{\sim}{X}_j^N)$ is less than the energy of the old configuration $E(\underset{\sim}{X}_i^N)$, the new configuration is accepted. If not, the ratio of the Boltzman probabilties P_{ji},

$$P_{ji} = \exp\left(-\beta E\left(\underset{\sim}{X}_j^N\right)\right) \Big/ \exp\left(-\beta E\left(\underset{\sim}{X}_i^N\right)\right) \quad (8)$$

is compared with a random number ξ ($0 \leqslant \xi \leqslant 1$) and the new configuration is accepted if P_{ji} is greater than ξ.ie the new configuration is accepted with a probability P given by

$$P = \min\left(1, P_{ji}\right) \quad (9)$$

Metropolis showed that this procedure leads to configurations chosen according to their Boltzman probabilties and is ergodic as well.

In the computer simulation, the N-particle system is given a condensed phase environment by the appropriate choice of T and V and the use of image cells, i.e. periodic boundary conditions.[25] The cell of volume V is surrounded by identical cells with the same configuration. Each move of a particle in the central cell is also done automatically by its images in the neighbouring cells. If a particle leaves the central cell, its image enters from the corresponding neighbouring cell. The shape of the unit cell is chosen to be a regular packing polyhedron. The most commonly used unit cell is the simple cubic cell.

The principal variables in a Monte Carlo computer simulation are T, V, N and the means of determining the configurational energy of the system $E(\underset{\sim}{X}^N)$. This quantity can

be expressed as

$$E(\underset{\sim}{X}^N) = \sum_{i<j} E^{(2)}(\underset{\sim}{X}_i, \underset{\sim}{X}_j) + \sum_{i<j<k} E^{(3)}(\underset{\sim}{X}_i, \underset{\sim}{X}_j, \underset{\sim}{X}_k) + \cdots \qquad (10)$$

where $E^{(n)}$ are n-body contributions to the interaction. The conventional way to proceed in computer simulations is to obtain relatively simple algebraic expressions for the $E^{(n)}$ and program this into the overall calculation. Most simulations on molecular liquids have considered only the pairwise term(n=2) with others either neglected entirely or considered to be mapped on to an effective pairwise term in some average way. The analytical potential functions representative of quantum mechanical calculations are directly identified with $E^{(2)}$ by definition, i.e.

$$E(\underset{\sim}{X}^N) \simeq \sum_{i<j} E^{(2)}(\underset{\sim}{X}_i, \underset{\sim}{X}_j) \qquad (11)$$

The effect of higher order terms has been studied especially for multiple water interactions by Hankins, Moscowitz and Stillinger[26] and Lentz and Scheraga[27]; higher order terms can contribute errors in the interaction energy of the order of 10-15%. All calculations to be decribed in forthcoming chapters assume pairwise additivity and this should be considered as a possible source of error in the results.

The energy of the configuration $E(\underset{\sim}{X})^N$ is calculated as the sum of pairwise interactions, using analytical form for $E_{ij}^{(2)}(\underset{\sim}{X}_i, \underset{\sim}{X}_j)$. There are two conventions followed in this procedure;

1) minimum image convention and 2) spherical cutoff. In the minimum image convention all the interactions of a given particle with the nearest images of all the other particles are evaluated and summed up. In the spherical cutoff method, the energy evaluation is done for all interactions within a given radius R . If R is chosen to be less than or equal to half the edge of the cubic cell, only interactions of a molecule with one of the images of the other molecules have to be evaluated.

The error estimates for the calculation are calculated using the method suggested by Wood. This procedure involves dividing the whole Monte Carlo[23] run after it has attained stabilisation into n different segments. The average property is evaluated for each of these segments $\bar{F}_i$, called "control functions". Then the standard deviation is estimated as

$$\sigma(\bar{F}) = \sqrt{\frac{1}{n(n-1)} \sum_i (\bar{F} - \bar{F}_i)^2} \qquad (12)$$

where $\bar{F}$ is the average for all the segments. The control functions for energy averages are used as an indicator of convergence. i.e. if the control functions for energy are randomly distributed about a mean value for a relatively long stretch, the calculation is assumed to have converged.

c) Distribution Functions

Three types of distribution functions are discussed here: radial distribution functions, closely related to the second

order molecular distribution function and observable from diffraction experiments on the liquid, quasicomponent distribution functions, used in the analysis of structure and difference quasicomponent distribution functions, used in comparing structural analysis on different but related systems.

i) Radial Distribution Function

The radial distribution function g(R) is obtained as

$$g(R) = \frac{N(R)}{4\pi R^2 \Delta R \cdot \rho} \qquad (13)$$

where ρ is the number density and N(R) is the average number of particles in a spherical shell of width ΔR at a radial distance R from another particle; the function g(R) gives the configurationally averaged deviation of the local microscopic environment of the particle from a value characteristic of the bulk density. The average coordination number can be determined from g(R) as the average number of particles within a given sphere of radius R_M

$$N_c = \rho \int_0^{R_M} 4\pi R^2 g(R)\, dR \qquad (14)$$

The parameter R_M can be chosen to be the radial value of the first minimum in g(R) to correspond to common chemical concepts of coordination number.

ii) Quasicomponent Distribution Functions(QCDF)

A microscopic theoretical analysis of the system can be

developed in terms quasicomponent distribution functions (QCDF).[24] QCDF with respect to coordination number and binding energy are the principal means of analysis of the computer simulation results in the forthcoming studies. The QCDF for coordination number $x_c(K)$ is given by

$$x_c(K) = \frac{N_c^{(1)}(K)}{N} = \frac{1}{N} \int \cdots \int P(\underset{\sim}{R}^N) \sum_i \delta \left[C_i(\underset{\sim}{R}^N) - K \right] d\underset{\sim}{R}^N \tag{15}$$

where $N_c^{(1)}(K)$ is the average number of particles in the system with coordination number K and $\delta\left[C_i(\underset{\sim}{R}^N)-K\right]$ is a Dirac delta counting function for the number of particles with a coordination number K in configuration . The quantity $x_c(K)$ can be viewed as a component of the vector

$$x_c = \left\{ x_c(0), x_c(1), x_c(2) \cdots \right\} \tag{16}$$

the elements of which define the entire composition of the system in terms of coordination number. The average coordination number of equation 9 in terms of $x_c(K)$ is

$$N_c = \sum_{K=0}^{\infty} K \, x_c(K) \tag{17}$$

The QCDF with respect to binding energy B for particle i follows from the definition

$$B_i = E\left(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \cdots \underset{\sim}{X}_{i-1}, \underset{\sim}{X}_i, \underset{\sim}{X}_{i+1}, \cdots \underset{\sim}{X}_N\right) - E\left(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \cdots \underset{\sim}{X}_{i-1}, \underset{\sim}{X}_{i+1}, \cdots \underset{\sim}{X}_N\right) \quad (18)$$

The mole fraction of particles with binding energy between and $\nu+d\nu$ is given by

$$x_B(\nu) = \frac{N_B^{(1)}(\nu)}{N} = \frac{1}{N} \int \cdots \int P(\underset{\sim}{X}^N) \sum_i \delta\left[B_i(\underset{\sim}{X}^N) - \nu\right] d\underset{\sim}{X}^N \quad (19)$$

where $\delta\left[B_i(\underset{\sim}{X}^N) - \nu\right]$ is a Dirac delta counting function for particles with binding energy between and +d . The vector

$$x_B = \left\{ x_B(\nu_0), x_B(\nu_1), \cdots \right\} \quad (20)$$

defines the composition of the entire system as a function of configurationally averaged binding energies for individual particles. The relationship between $x_B(\nu)$ and configurational internal energy U is

$$U = \int_{-\infty}^{+\infty} \nu \, x_B(\nu) \, d\nu \quad (21)$$

Note that $x_c(K)$ is a discrete function whereas $x_B(\nu)$ is a continuous function in ν.

Finally the analysis of the system in terms of QCDF can be extended to display the distribution of binding energies as a function of coordination number,

$$x_{B,C}(\nu, K) = \frac{(8\pi^2)^N}{N} \int \cdots \int P(\underset{\sim}{X}^N) \sum_i \delta\left[C_i(\underset{\sim}{X}^N) - K\right] \sum_j \delta\left[B_j(\underset{\sim}{X}^N) - \nu\right] d\underset{\sim}{X}^N \quad (22)$$

and define a compositional vector as

$$x_{B,C} = \left\{ x_{B,C}\,(V, K) \right\} \qquad (23)$$

Further characteristics of the system can be displayed in terms of QCDF formulated on the basis of considerations analogous to those defined above.

iii) Difference Quasicomponent Distribution Functions.

This section deals with the perturbations on solvent structure due to the addition of a solute molecule, developed in terms of aqueous solution characteristics. The structure of water in an aqueous solution can be defined in terms of QCDF for coordination number K and binding energy v, as described in equations 15 and 19 . In a solution problem, one wishes to determine such the quantities such as

$$\frac{d\,N_C^{(1)}(K)}{d\,N_S} \qquad \text{and} \qquad \frac{d\,N_B^{(1)}(V)}{d\,N_S} \qquad (24)$$

the change $dN_C^{(1)}$ in the first order QCDF with respect to change in the number of solute molecules dN_S.

Taking coordination number as an example and developing the derivative in terms of finite differences, we have

$$\frac{d\,N_C^{(1)}(K)}{d\,N_S} \simeq \frac{\Delta\,N_C^{(1)}(K)}{\Delta\,N_S} = \Delta\,N_C^{(1)}(K) \qquad (25)$$

where $\Delta\,N_S$ has been assigned the value of unity since our

system involves the addition of only one solute particle. Now for calculations based on N_w water molecules in the pure liquid and N solvent water molecules in the solution,

$$\Delta N_c^{(1)}(K) = N_c^{(1)}(K) - N_c^{(1)}(K)^w \qquad (26)$$

where $x_c(K)$ is the mole fraction of water molecules as a function of K for solvent water in a solution and $x_c(K)^w$ is the corresponding quantity for pure water. In our treatment of the problem, the number of solvent waters N is equal to $N_w - 1$ one of the waters having been replaced by a solute. Using this relationship for N, expanding equation 26 and substituting back in equation 25, we have

$$\frac{d N_c^{(1)}(K)}{d N_s} \cong N_w \Delta x_c(K) \quad , N_w \gg 1 \qquad (27)$$

where $\Delta x_c(K)$ is the difference quantity.

$$\Delta x_c(K) = x_c(K) - x_c(K)^w \qquad (28)$$

Analogously for binding energy,

$$\frac{d N_B^{(1)}(\nu)}{d N_s} \cong N_w \Delta x_B(\nu) \quad , N_w \gg 1 \qquad (29)$$

where

$$\Delta x_B(\nu) = x_B(\nu) - x_B(\nu)^w \qquad (30)$$

Note that $\Delta x_c(K)$ alone is a quantity independent of the number of particles in the system whereas the difference QCDFs $dN_c^{(1)}/dN_s$ and $dN_6^{(1)}/dN_s$ are effectively independent of the number of solute particles in the system and can be used for absolute comparison among systems.

d) Free Energy

In the dicussion of the Monte Carlo method it was mentioned that it was useful for calculating average properties of the system. The ensemble average expression for the free energy can be expressed as

$$\Delta A = -\beta^{-1} \langle \ln \left(\exp \left(\beta E(\underset{\sim}{X}^N) \right) \right) \rangle \qquad (30)$$

This expression is not convenient for computational procedures due to the ill-conditioned nature of the integrand and concomitant convergence problems. The various possible computational approaches to free energy considered to date have been reviewed recently by Valleau and Torrie,[28] who have discussed the theoretical or computational limitations of each method. An alternative approach follows early work by Kirkwood,[29] wherein the Helmoltz configurational free energy is given by

$$\Delta A = \int_0^1 U(\lambda)\, d\lambda \qquad (31)$$

Here the integrand $U(\lambda)$ can be expressed as an ensemble

average

$$\mathcal{U}(\lambda) = \int \cdots \int E(\underset{\sim}{X}^N)\, P(\underset{\sim}{X}^N, \lambda)\, d\underset{\sim}{X}^N \qquad (32)$$

where $E(\underset{\sim}{X}^N)$ is the configurational and $P(\underset{\sim}{X}^N, \lambda)$ is the probability of observing the system in configuration X, conditional upon the auxiliary parameter λ,

$$P(\underset{\sim}{X}^N, \lambda) = \frac{\exp\left(-\beta E(\underset{\sim}{X}^N, \lambda)\right)}{\int \cdots \int \exp\left(-\beta E(\underset{\sim}{X}^N, \lambda)\right) d\underset{\sim}{X}^N} \qquad (33)$$

with β representing $(kT)^{-1}$.

When the auxiliary parameter serves to couple the system according to the expression

$$E(\underset{\sim}{X}^N, \lambda) = \lambda\, E(\underset{\sim}{X}^N) \qquad (34)$$

the free energy is defined with respect to an ideal gas reference state of liquid density and is equivalent to the general class of "thermodynamic integrations" for free energy, the main disadvantage stems from problems of integration over λ. This can be done using a small order Gaussian quadrature. The simple extra computational effort required may be complicated by convergence problems for small λ and possible discontinuities in $U(\lambda)$. Calculations based on this approch are reported for liquid water in the following chapter.

III LIQUID WATER

Introduction

This chapter describes a theoretical analysis of the equilibrium, diffusionally averaged structure of liquid water at 25° C based on a Monte Carlo statistical thermodynamic calculations based on pairwise additive analytical potential functions representative of ab initio quantum mechanical calculations of the interaction of water dimer. The analysis of the structure is carried out in terms of quasicomponent distribution functions(QCDF). A comparison of results based on potential functions derived from calculations based at the Hartree-Fock level and from calculations including configuration interaction permit a determination of the effect of electron correlation on the structure of liquid water. A fully ab initio calculation of the free energy of liquid water is also reported.

The background for this chapter is developed in section b The calculations and results are presented in section c. The discussion of the results take place in section d. A summary of the chapter is given in the final section e.

b) Background

Early thoughts about the structure of liquid water developed as variations of ideas on the structure of ice. The conception of the structure of water as a quasi-solution in

equilibrium with a truly fluid polymorph can be traced as far back as Roentgen in 1892.[30] Modern quantitative treatements of the problem date from the classic paper of Bernal and Fowler in 1933,[31] featuring an ice-like component consisting of a network of tetrahedrally coordinated water molecules. The theoretical development of the subject from that point can be traced in terms of essentially ad hoc and ab initio approaches to the problem

The ad hoc approach involves a statistical thermodynamic treatment of a viable model for the entire system, with disposable parameters chosen empirically or for optimum agreement between calculated and experimentally observed properties. Representative examples of ad hoc models for liquid water are the interstitial models developed by Samoilov,[32] Mikhailov[33] and Narten, Danford and Levy,[34] the clathrate model due to Pauling[35] and Frank and Quist,[36] the two state model described by Davis and Litovitz,[37] the three state model of Vand and Senior,[38] the fickering cluster model due Frank and Wen[39] and Hagler, Nemethy and Scheraga[40] and the significant structure approach by Jhon, Grosh, Ree and Eyring.[41] Alternatively, the structure of liquid water has been viewed as a continuous distribution of structures involving bent hydrogen bonds as proposed by Pople,[42] and elaborated into the continuum model for the system.[44]

All of the early theoretical work on the structure of water was based on ad hoc methods, and much of our current thinking about the structural chemistry of molecular liquids

developed from these studies. The fundamental idea that the anomolous behavior of liquid water as compared with normal liquids in the behavior of density and compressibility with respect to temperature was due to contributions from a low density four-coordinate ice-like structure with anomolously low energy was well established in the early work. Detailed reviews of ad hoc models are available in the monographs of Eisenberg and Kauzman,[1] Ben Naim,[24] and the review articles by Davis and Jarzynski[43] and by Hagler, Nemethy and Scheraga[40] on mixture models and Kell[44] on continuum models. Interesting additional perspective on the problem is due to Kamb,[45] Stevenson[46] and Frank.[47] Extensive experimental data is available on the system, but the collected results are unable to differentiate unequivocaly between mixture models and continuum models or among the various mixture models. The experimental evidence for mixture models based on vibrational spectra has been given particularly by Walrafen,[48] but questioned by Scherer;[49] the arguments from diverse view points are given in two of the review articles referred to previously.

A considerable literature exists on computer simulation of model fluids, but the application of computer simulation techniques to molecular liquids is at a relatively early stage of development at present. The predominant number of published studies deal with liquid water,[11,17-22] with simulations on other pure liquids and some solutions just now appearing.[12] The original molecular dynamics study of liquid water is due to Rahmann and

Stillinger[11] and incorporates the Ben-Naim-Stillinger(BNS)[50] potential. The first Monte Carlo simulations liquid water were reported by Barker and Watts[17] based on BNS potential.

Concurrently an important series of papers on analytical potential functions representative of ab initio quantum mechanical calculation of pairwise interaction energies by Clementi and coworkers have appeared. Potentials for the water-water interaction were reported at the level of Hartree-Fock(HF) molecular orbital theory[19] and with a number of alternative corrections for electron correlation,[20] the best quality function being representative of moderately large configuration interaction(CI) calculations on the water dimer.[21] Each of these potentials was incorporated in Monte Carlo simulations and calculations on a wide variety of properties of liquid water were reported. The HF potential produced an oxygen-oxygen radial distribution function with serious discrepancies between calculated and experimental values beyond the first hydration shell.[51] The CI potential with only intermolecular correlation corrections included gave results for the oxygen-oxygen radial distribution function reasonably close to the observed values for the entire range of inter-particle separations,[22] and is currently considered the best available pairwise potential function for the water-water interaction.

A satisfactory quantitative agreement between experimental radial distribution functions and corresponding quantities calculated for liquid water in computer simulation

incorporating the CI potential having been demonstrated, it is of interest to analyse the results of the simulation with a view towards obtaining an ab initio theoretical idea of the nature of the structure of liquid water. Moreover the availability of both the HF and CI potential functions permits a comparison of the results obtained with and without electron correlation and a study of the effect of electron correlation on the structure of liquid water.

The calculations described herein are used as the basis for an analysis of the structure of liquid water and the calculation of observable properties such as internal energy and free energy. The ensemble average expression for free energy is inconvenient for computational procedure due to ill-conditioned nature of the integrand and concomitant convergence problems as discussed in the previous chapter. In view of the central importance of liquid water in the chemistry and physics of liquids, the calculation of the free energy was carried out using the Kirkwood method as described in the previous chapter in spite of the extra computational effort.

c) Calculations and Results

Programs based on the theoretical development discussed in the previous chapter were coded and implemented on the IBM 370/168 computer. Calculations were carried out on liquid water by means of a 125 particle simulation at 25 ^{o}C and a volume commensurate with 1 gm/cm^{3} as in the Rahmann-Stillinger[11] calculation. Stepsize in the stochastic walk in the Metropolis

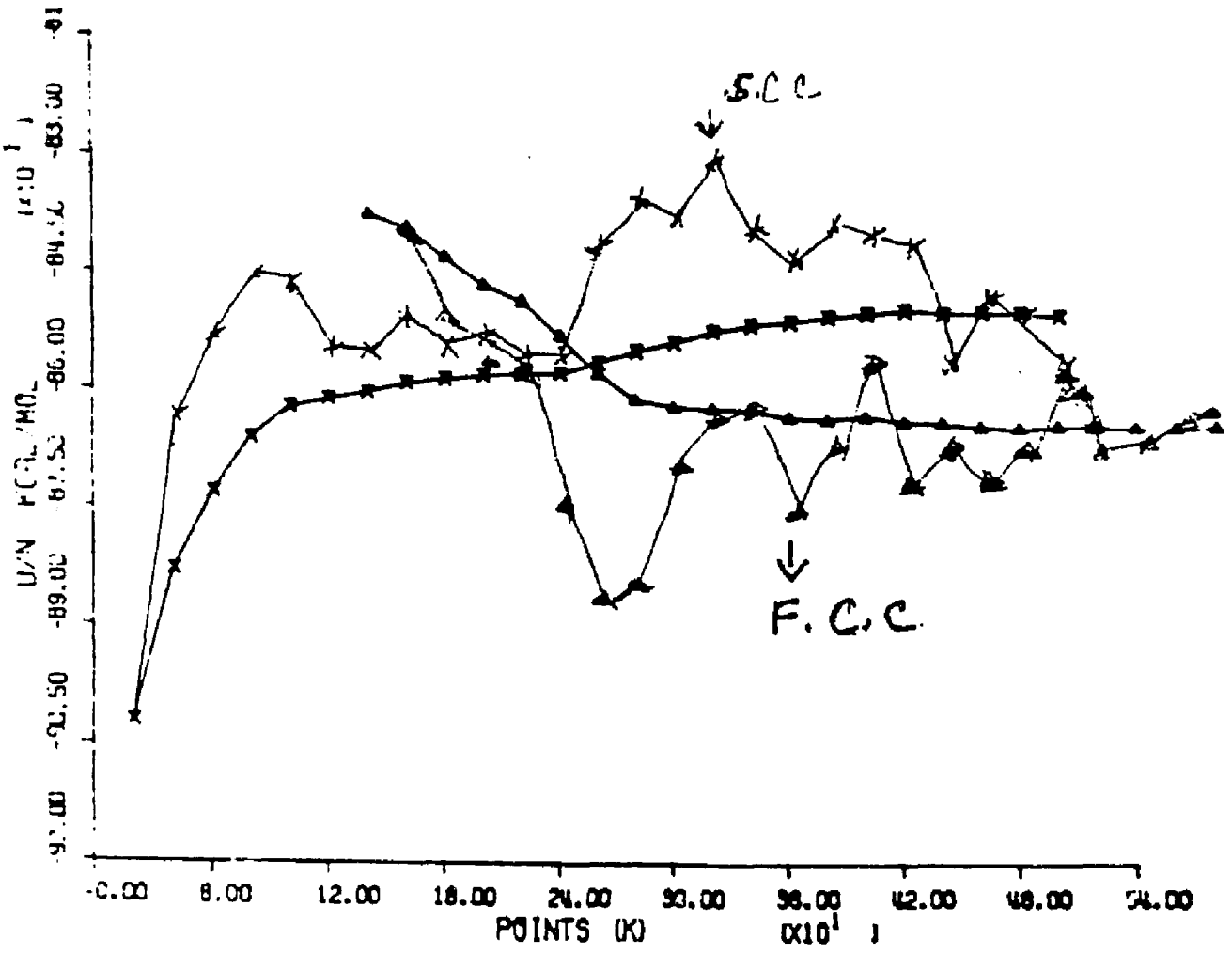


Figure 1. The control function plots for internal energy per molecule vs. number of Monte Carlo points for two simulations. (X) denotes simulation of 125 particles in a simple cubic periodic boundary and (Δ) denotes a simulation of 100 particles in a face centred cubic boundary.

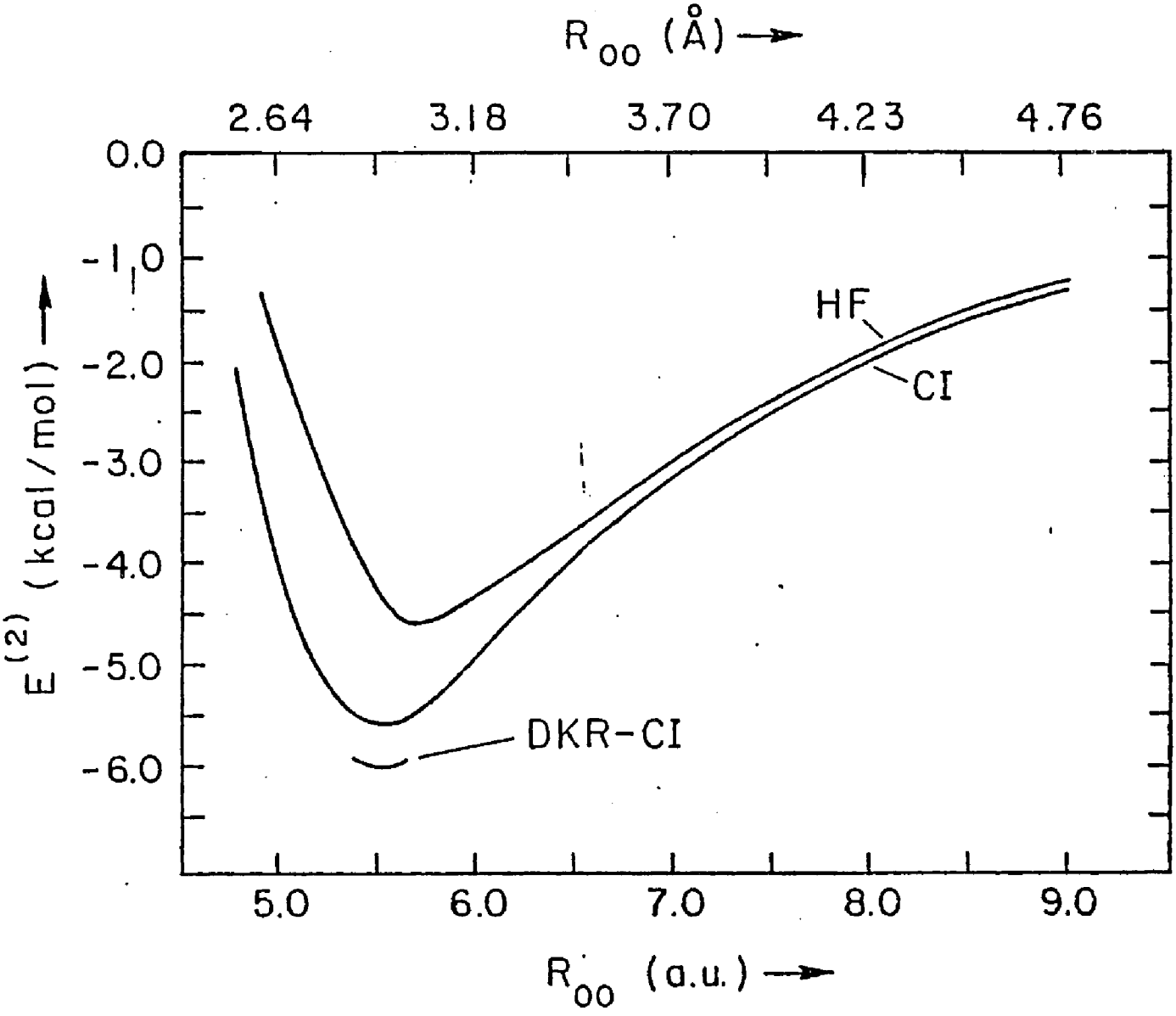


Figure 2. Calculated pairwise interaction energy $E^{(2)}$ as a function of inter-oxygen separation R for the "linear" water dimer; adapted from Fig. 6 of Ref 22.

procedure[16] was selected to produce an approximately 50% acceptance rate. The spherical cutoff for the intermolecular potential was chosen to be half the box size. All calculations were found to be well converged after 500K steps of the Monte Carlo process, 200K having been initially discarded.

The convergence in the Metropolis sampling procedure was tested out by two independent runs of computer simulation of 125 water molecules under simple cubic periodic boundary conditions and 100 water molecules under face centered cubic periodic boundary conditions. They both converged to the same values as shown in Fig. 1. This ensures that the simulations described above have not encountered any practical non-ergodicities.

One of the computer simulations carried out was based on the HF potential and the other based on the CI potential. The difference between these functions is displayed on one representative slice of the potential energy hypersurface in Fig 2. Here the interaction energy for the linear water dimer is plotted as a function of the oxygen-oxygen separation ; this is the optimum geometry for hydrogen bonding. The HF potential has a minimum at 3.01 Å and and a binding energy of -4.55 kcal/mol. The CI potential has a minimum at 2.92 Å and a binding energy of -5.62 kcal/mol. The CI potential gives results parallel to and within 0.43 kcal/mol of the most extensive calculation to date on the water reported by Dierksen, Kramer and Roos,[52] labelled "DKR-CI" in Fig. 2. The

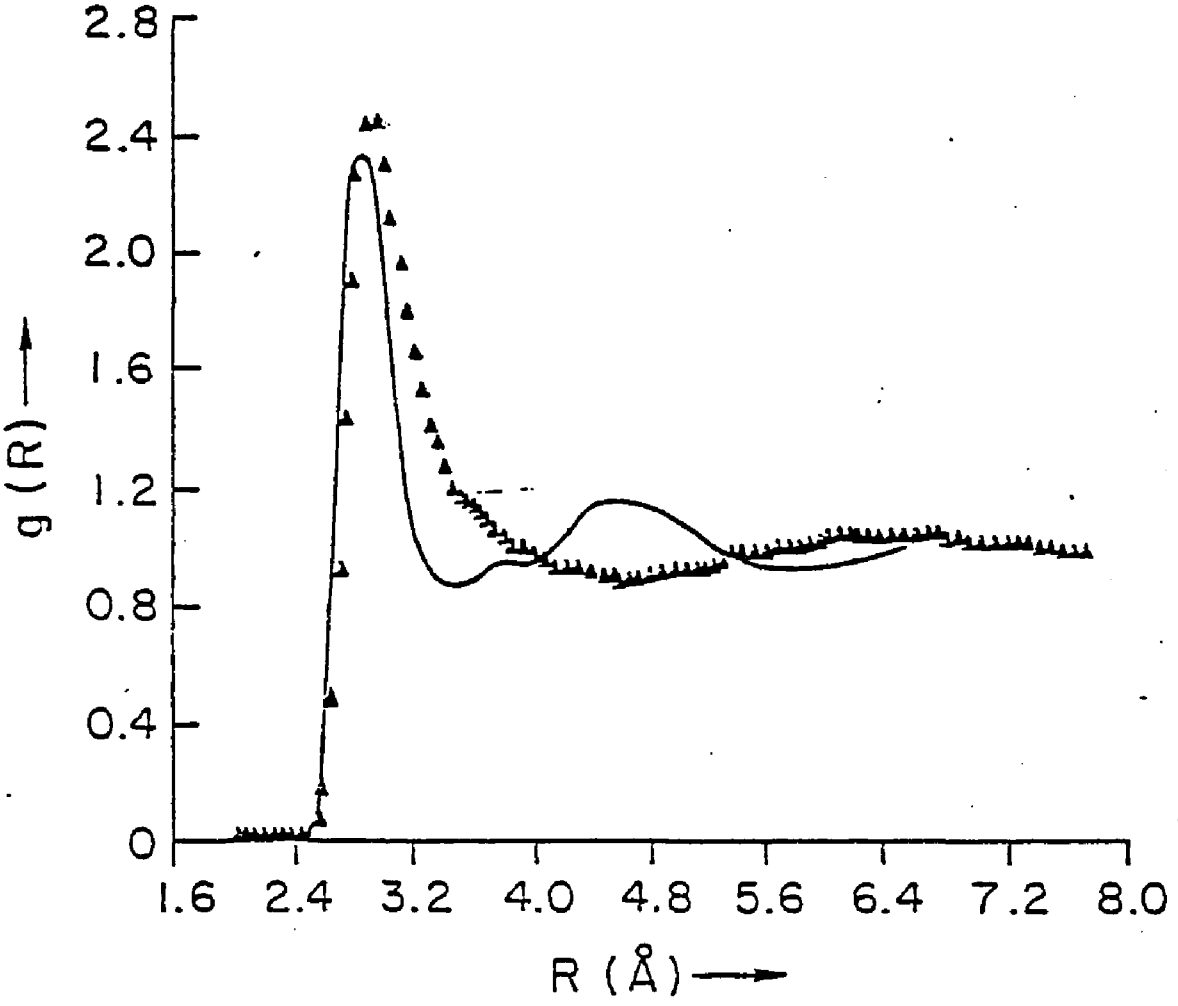


Figure 3. A comparison of the experimentally observed radial distribution function for liquid water (solid line) with the points on the radial distribution function calculated from a Monte Carlo simulation of liquid water based on the HF potential. The experimental curve is redrawn from Fig. 2 of Ref 23.

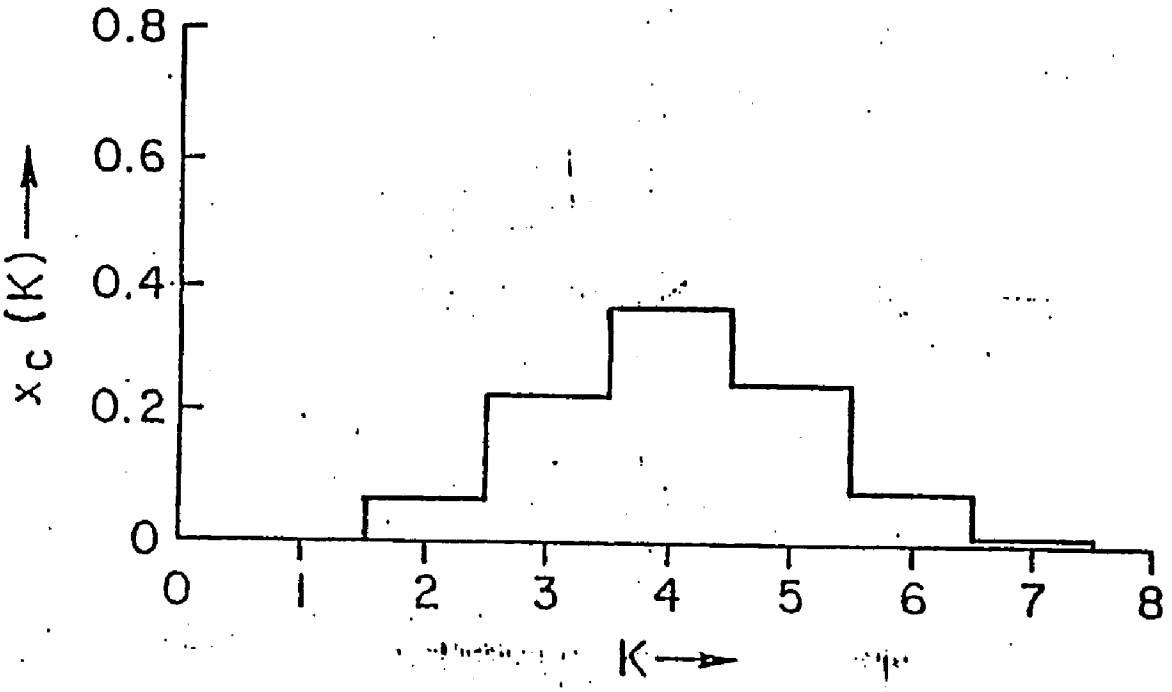


Figure 4. Calculated quasicomponent distribution function x_c (K) vs. coordination number K for liquid water based on the HF potential.

effect of correlation on the pairwise interaction is to stabilise the energy by approximately 1 kcal/mol and reduce the minimum energy intermolecular separation by 0.09 Å.

The calculated radial distribution function for the computer simulation incorporating the HF potential is shown in comparison with the corresponding experimentally observed result in Fig 3. This reproduced Fig 10 of Kistenmacher, Popkie, Clementi and Watts and serves as confirmation of the computer program used. Comparing calculated and observed results, the HF potential shows significant deviations from the experimentally determined function, with the first peak(first hydration shell) biased toward larger distances and the position of the second shell well displaced form the observed values; in fact the HF potential produces a minimum in g(R) where the experimental value has a maximum. The calculated configurational internal energy is -6.9 kcal/mol with a standard deviation of 0.03 kcal/mol in comparison to the experimental value of -9.9 kcal/mol. This dicrepancy is attributed mainly to electron correlation, the assumption of pairwise additivity and basis set truncation errors. The calculated heat capacity is 18 cal/deg.mol in exact(fortuituously) agreement with the experimental value at 25°C.

An analysis of the computer simulation based on the HF potential was carried out in terms of the QCDF for coordination number and binding energy. The structure of "HF water" in terms coordination number is shown in Fig. 4. Here

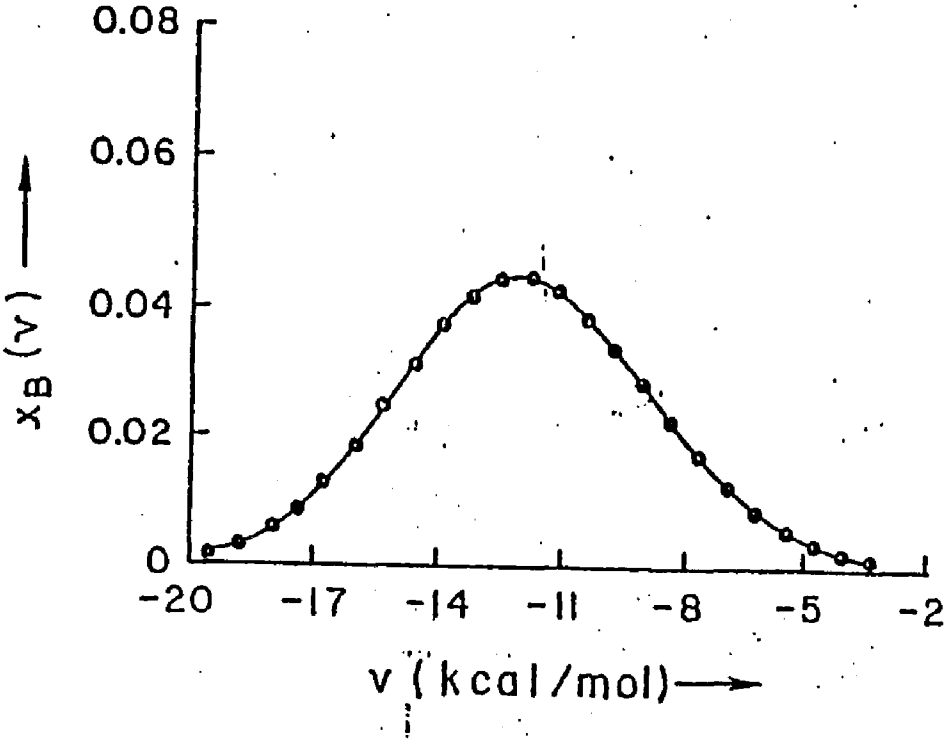


Figure 5. Calculated quasicomponent distribution function $x_B(\nu)$ vs. binding energy ν for liquid water based on the HF potential.

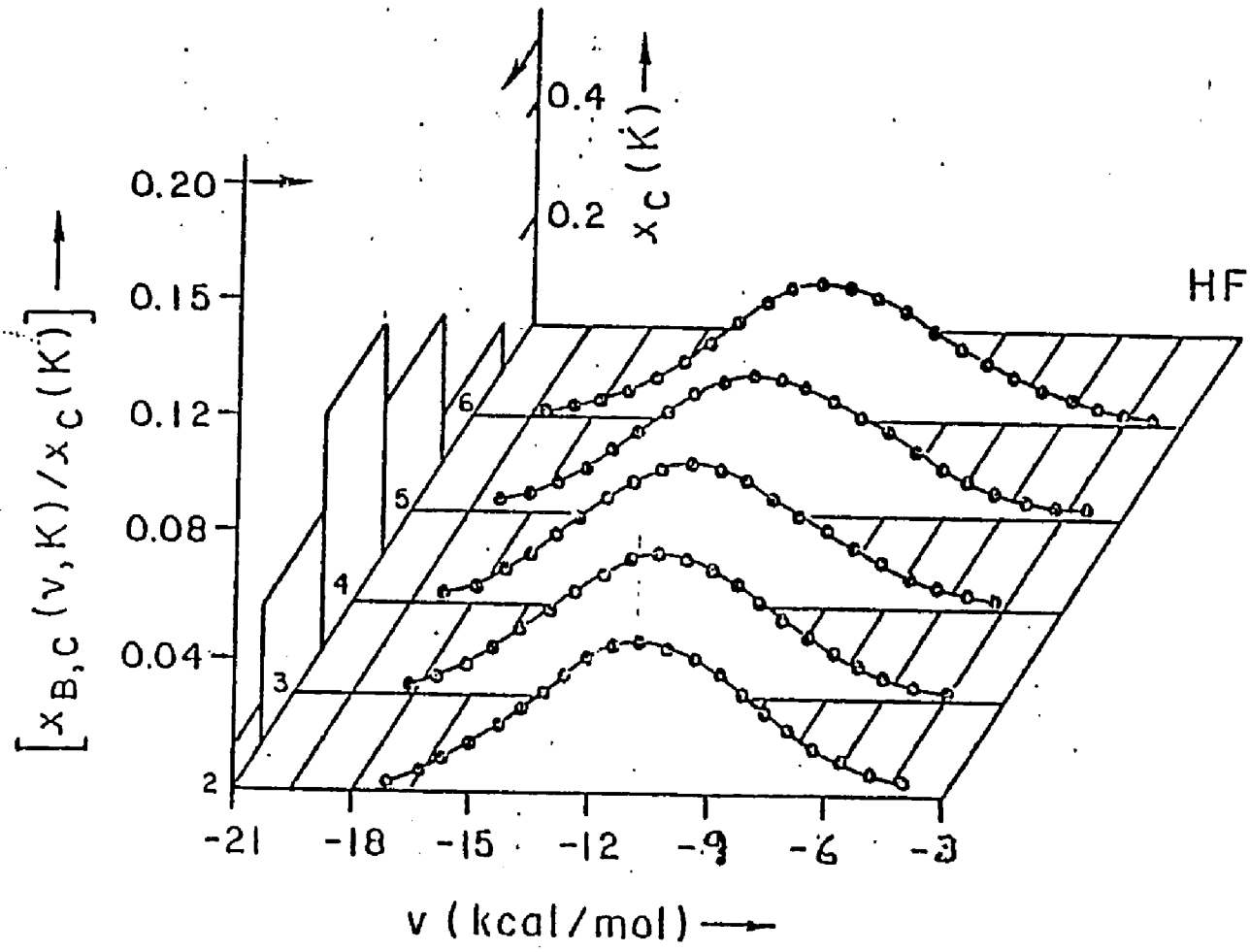


Figure 6. Calculated quasicomponent distribution function $x_{B,C}(\nu,K)$ vs. binding energy ν and coordination number K for liquid water based on the HF potential. Each curve is normalised by $x_C(K)$ and must be weighted by this quantity to assess its relative contribution to the statistical state of the system. To aid in the comprehension of this plot, $x_C(K)$ from Fig. 4 is plotted along the K axis.

the mole fraction $X_c(K)$ based on R = 3.3 Å is plotted vs K in histogram form. The predominant coordination number is K=4, coming in at 37 %. The distribution is markedly symmetric about K=4, with K=3 and K=5 represented to the extent of 23% and K=2 and K=6 at about 7%. Approximately 2% of the molecules have a coordination number of 7 in the equilibrium statistical state of the system. The average coordination number is 4.05 compared with an experimental value of 4.4.

The structure of HF water in terms of binding energy is shown in Fig. 5. Here the mole fraction $x_B(\nu)$ is displayed as a function of ν. We find a continuous distibution of binding energies in the system essentially symmetric about -12 kcal/mol In Fig. 6, the distribution of particles as a function of both coordination number and binding energy, $X_{B,C}(\nu$,K) vs. and K is presented. The curves are individually normalised and must be multiplied by x_c (K) to assess their respective contributions to the statistical state of the system. The $x_c(K)$ histogram is displayed along the K axis of the figure to facilitate comprehension of the plots. Considering first $x_{B,C}(\nu,K)$ vs ν. each of the distributions are individually continuous and essentially symmetric about a single maximum value. The behavior of $x_{B,C}(\nu,K)$ vs K shows a single minimum at K=4, the most predominant coordination number for the system. The binding energies for K=4 are lower in energy than those for K$\rangle$4.

The calculated and observed radial distibution functions from the simulation of liquid water using the CI potential is

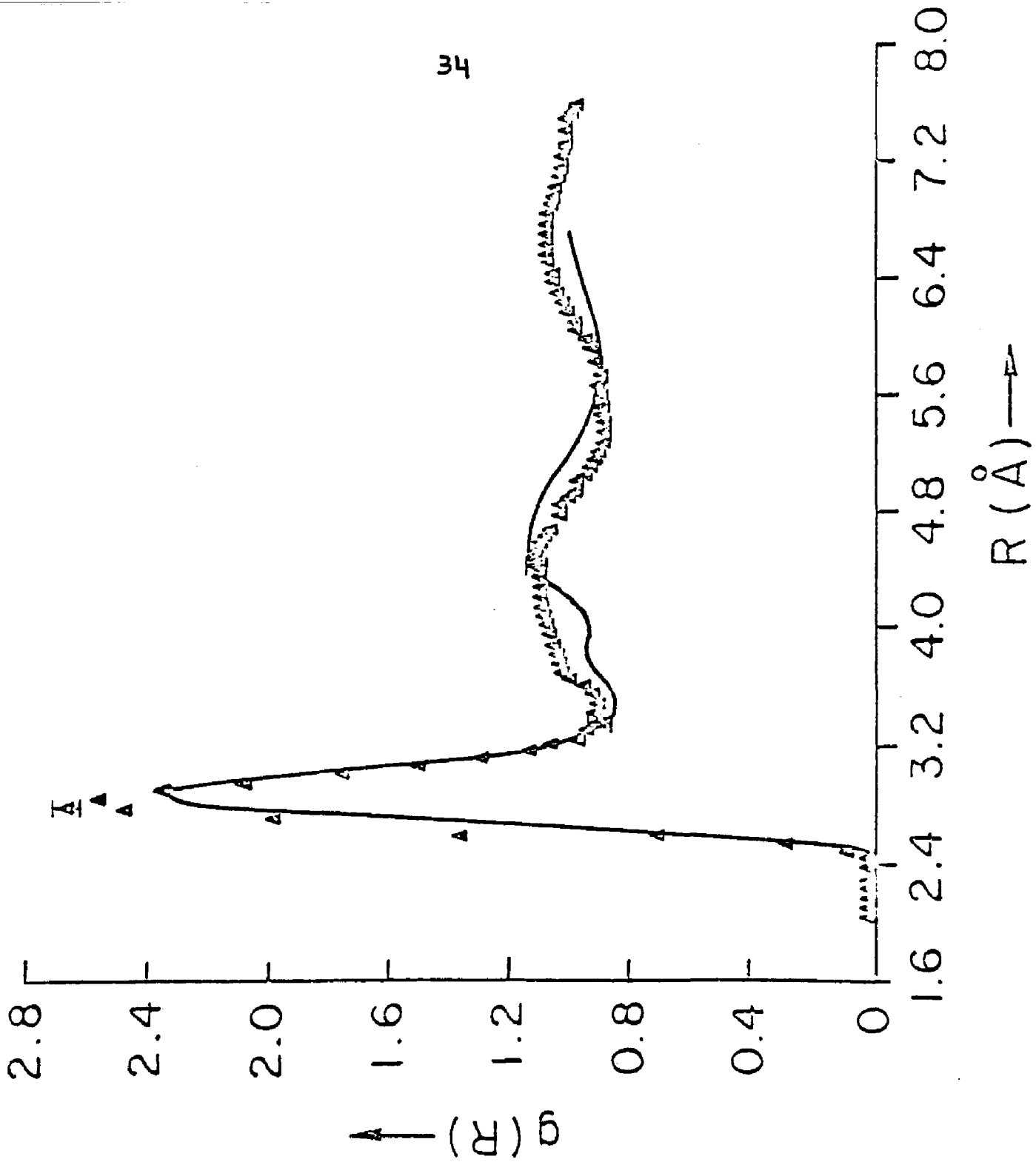


Figure 7. A comparison of the experimentally observed radial distribution function for liquid water (solid line) with the points on the radial distribution function calculated from a Monte Carlo simulation of liquid water based on the CI potential. The experimental curve is redrawn from Fig. 2 of Ref 23.

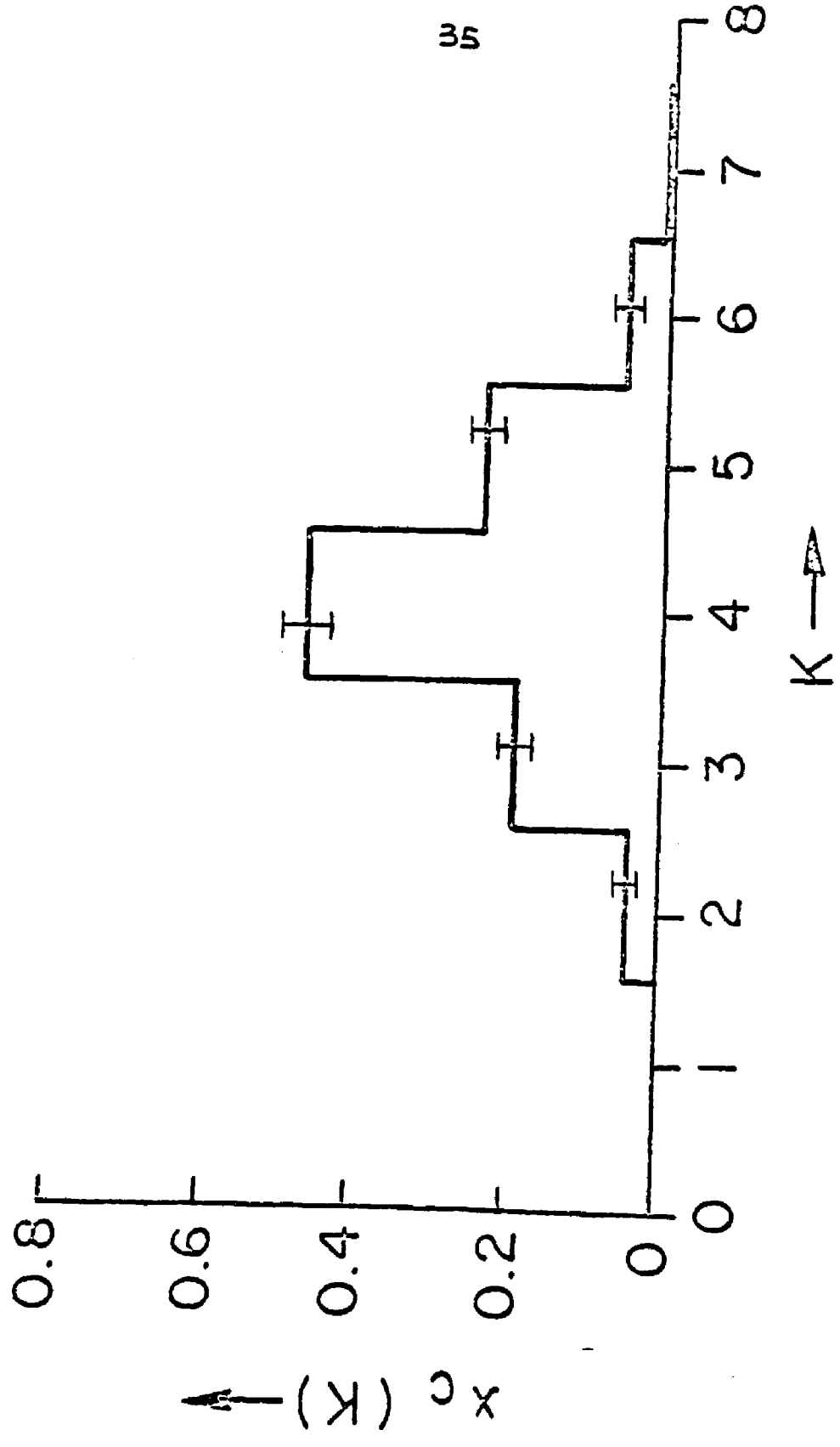


Figure 8. Calculated quasicomponent distribution function x_c (K) vs. coordination number K for liquid water based on the CI potential.

shown in Fig. 7; c.f. Fig 2a of Ref. 23. There is quite good accord between the between calculated and observed values for the position of all the three main peaks. In the region of the first hydration shell agrees well with the experiment although the calculated peak is slightly too high. For the second hydration shell the calculated position of maximum is in g(R) is in close accord with experiment[54], but the shape is biased towards short distances. The small peak or shoulder at ca. 3.5 Å, much discussed by previous investigators, is clearly evident in the calculated g(R). As noted by Lie et. al., overall the results are in satisfactory accord with experiment and as good or better than corresponding results based any of the available empirical potential functions. The calculated internal energy is -8.57 kcal/mol, some 16% above the observed value. The standard deviation on the calculated internal energy is 0.03 kcal/mol. This discrepancy is attributed mainly to the assumption of pairwise additivity in the potential function, and appears to be the order of magnitude expected for this effect. The calculated heat capacity is 17.9 cal/deg.mol.

The analyses of the computer simulation of liquid water based on the CI potential in terms QCDF's are displayed in Figs. 8, 9 and 10. The calculated structure of water in terms coordination number for R = 3.3 Å is shown in Fig. 8. The predominant coordination number is K=4 at 47%. The The distribution of coordination numbers is biased towards K=4, with K=5 at 24% and K=6 at 6% compared with K=3 at 19% and K=2

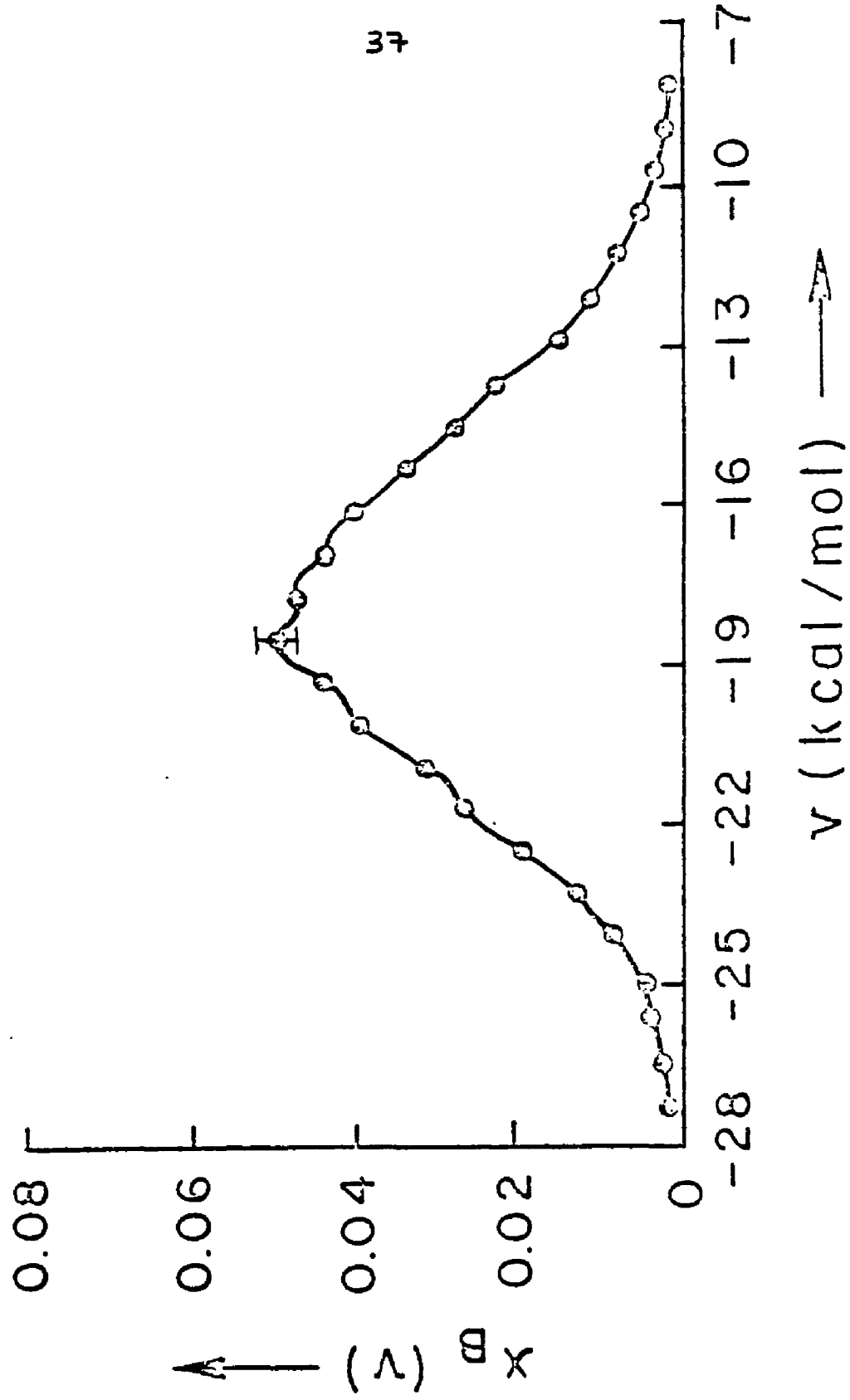


Figure 9. Calculated quasicomponent distribution function $x_\beta(\nu)$ vs. binding energy ν for liquid water based on the CI potential.

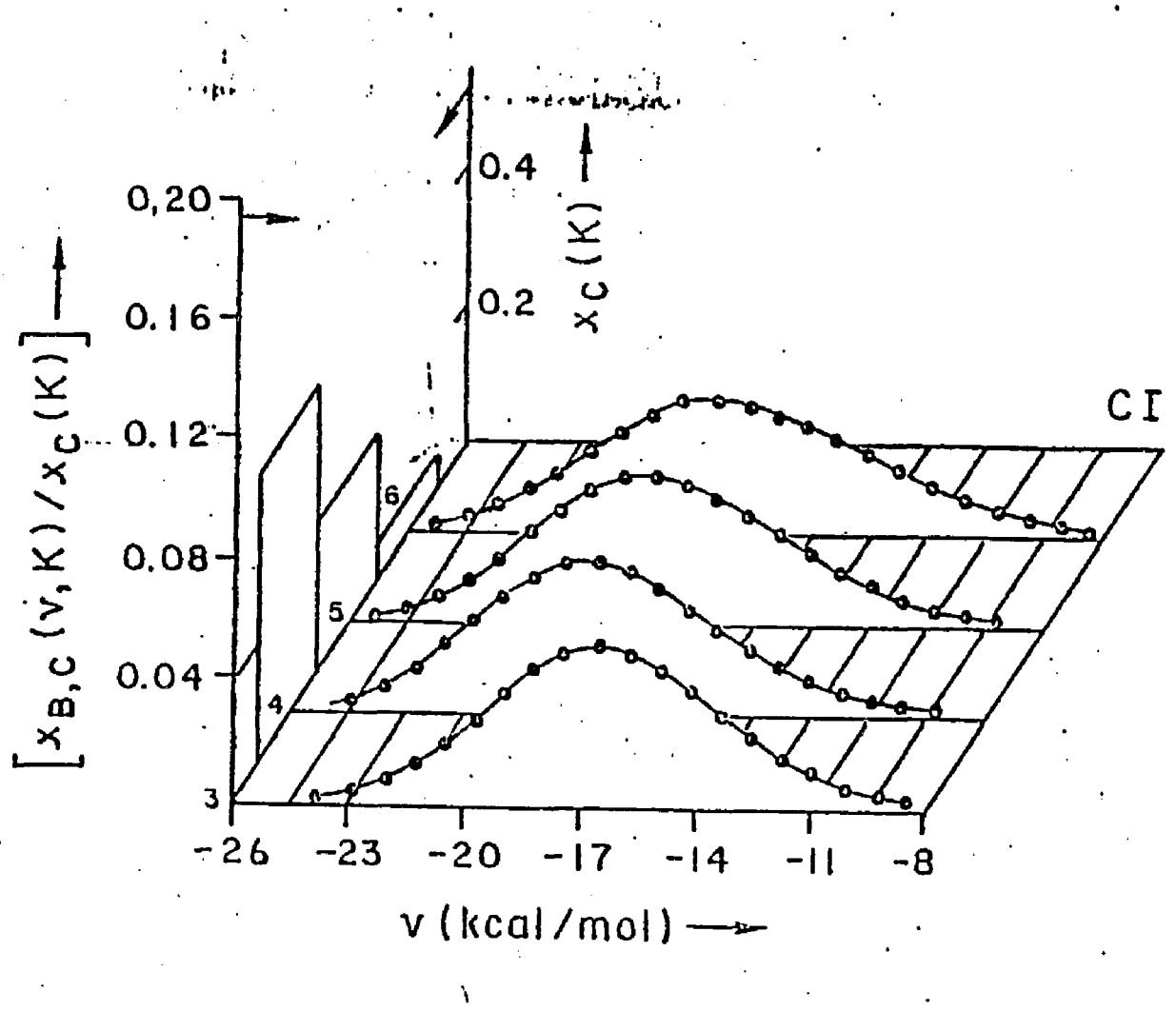


Figure 10. Calculated quasicomponent distribution function $x_{B,C}(\nu,K)$ vs. binding energy ν and coordination number K for liquid water based on the CI potential. Each curve is normalised by $x_C(K)$ and must be weighted by this quantity to assess its relative contribution to the statistical state of the system. To aid in the comprehension of this plot, $x_C(K)$ from Fig. 8 is plotted along the K axis.

at 4%. The calculated average coordination number is 4.1

The calculated structure of water in terms of binding energy is shown in Fig. 9. There is a major maximum in the plot at -17.7 kcal/mol, a relatively symmetric distribution of binding energies about this value and a slight but clearly discernable bias toward lower binding energies. There is evidence of some structure in the curve, but overall the calculated distribution is continuous.

The calculated distribution of water molecules as a function of both the binding energy and coordination number is shown in Fig. 10. Considering $x_{B,C}(\gamma,K)$ vs. γ, each of the curves is individually continuous about a single maximum value. The behavior of $x_{B,C}(\gamma,K)$ vs. K shows a single maximum at K=4. The binding energy curves for K$\leq$4 are displaced toward lower energies relative to those for K$>$4. The anomolously low binding energy of the four coodinate quasicomponent is clearly evident in Fig.10.

The analyses described above must be considered in perspective of the calculated error bounds on each quantity. A detailed error analysis based on the procedure described by Wood was carried out on the g(R) and $x_C(K)$ for the CI function. The standard deviation on g(R) was found to be a maximum of 0.04 in the region of the first peak and typically of the order of 0.01 to 0.02 elsewhere. The standard deviation for in $x_C(K)$ was at its maximum for K=4 at value of 0.007. Similar statistics are implied for the HF results.

The calculation of the free energy was carried out using

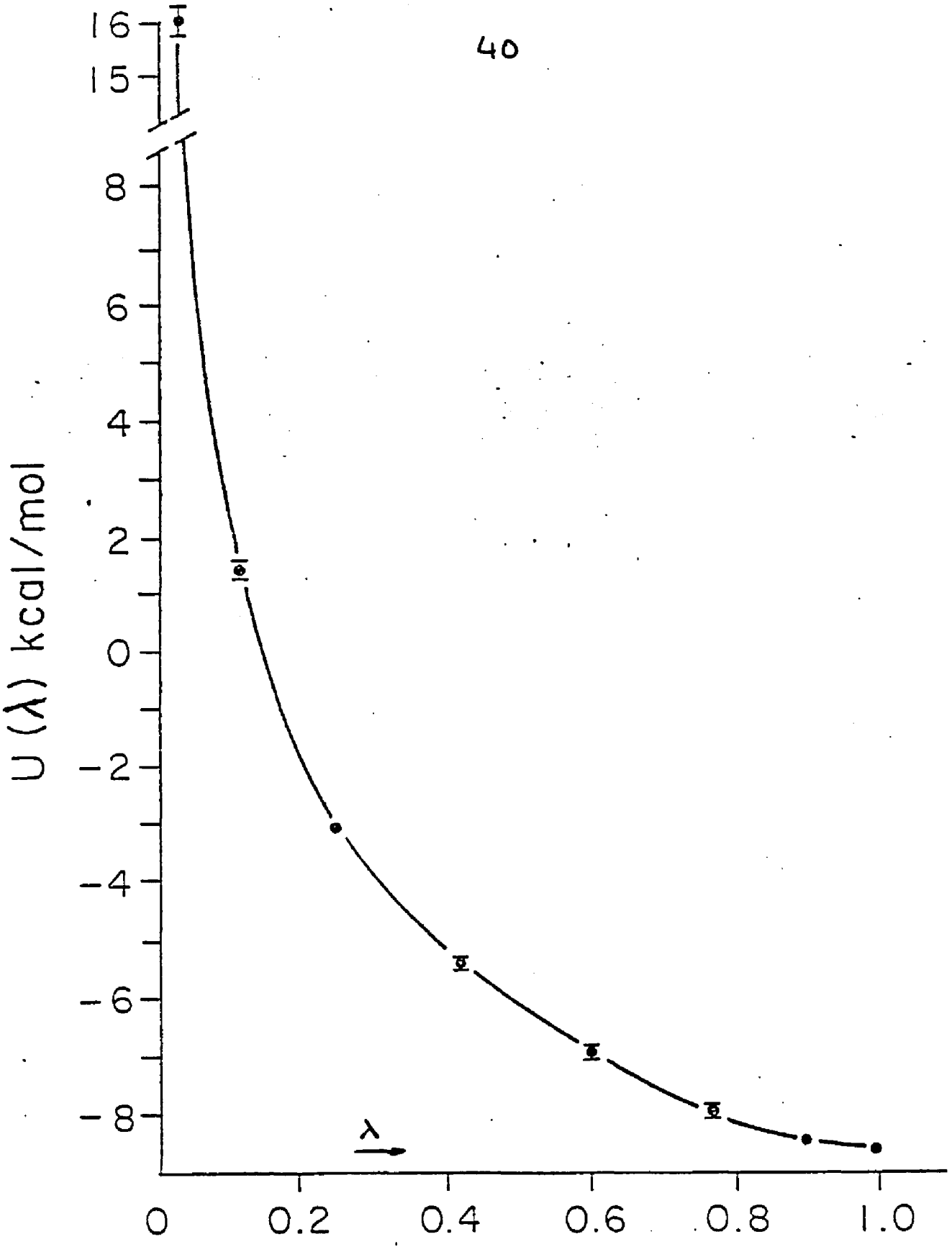


Figure 11. The ensemble average U(λ) vs. λ for all the eight points in the integration over λ to calculate the free energy of liquid water.

the integration over the Kirkwood coupling parameter[29]. The integration over λ was carried out by means of an 8-point Gaussian integration.[55] $U(\lambda)$ for all these eight points were determined by canonical ensemble Monte Carlo Metropolis computer simulation on 64 molecules under simple cubic periodic boundary conditions with a spherical cutoff equal to half the size of the cubic cell and using the CI potential described above. In previous studies it has been found that the results of the simulation based on 64 particles are essentially in agreement with O(125) particles. Convergence criteria and statistical error were established using control functions.

The Monte Carlo calculations of $U(\lambda)$ involved surprisingly few difficulties. The $U(\lambda)$ for all the eight points converged to within 1% within 500K Metropolis steps. A plot of $U(\lambda)$ vs. λ is shown in Fig. 11. The free energy was calculated directly from the integration over λ to be -4.31 $\pm$ 0.07 kcal/mol compared with an experimentally observed value of -5.74 kcal/mol. The entropy was calculated from the free energy and previously calculated internal energy to be -14.33 +0.09 cal/deg.mol compared with an observed value of -13.96 cal/deg.mol.

A comparison of the calculated and observed values shows the discrepancy between the calculated and observed free energy to be closer to that found for internal energy alone, and is thus ascribed to the same reasons enumerated above. There is a close accord between the calculated and

experimental values of entropy.

d) Discussion

Since the Monte Carlo simulation of liquid water using the CI potential produces a radial distibution function in close accord with experiment, we can look to the analysis of this calculation (Figs. 8, 9 and 10) for a theoretical view of the structure of liquid water. Considering first the idea of low density and high density contributions to water structure, the results in terms of $x_c(K)$ in Fig. 8 allows these contributions to be quantified. The mole fraction of low density component x_L is

$$x_L = \sum_{K=0}^{4} x_c(K) = 0.70$$

and the mole fraction of high density component x_H is

$$x_H = \sum_{K=5}^{\infty} x_c(X) = 0.30$$

for the statistical state of the system at 25°C

The predominant quasicomponent with respect coordination number at K=4 in Fig. 8 can be identified as part of the low density component considered in the ad hoc models. For an ice-like structure we expect four hydrogen bonds per molecule at -5.62 kcal/mol each ; a binding energy of -23 kcal/mol is near the lower limit of $x_{BC}(\nu,4)$ curve in Fig. 10. Note that the binding energy includes contribution from the first hydration shell as well as from the rest of the assembly. The

fact that the maximum in $x_{\beta c}(\gamma,4)$ falls at $\gtrsim$ -23 kcal/mol indicates that slightly distorted hydrogen bonds or other bound structures play a significant role in the system.

One can look to the plot of $x_{\beta}(\gamma)$ in Fig.9 for information on the relative viability of mixture models as opposed to continuum models. According to calculations on the model system the presence of distinct quasicomponents of a mixture model should be indicated by structure in the $x_{\beta}(\gamma)$ curve with separate distinct peak for each component, whereas a continuum model would be represented by a continuous distribution of individual binding energies. The results in Fig 9 favor the latter model [11]; similar conclusions emerged from molecular dynamics. This result should be considered in the perspective of the assumption of pairwise additivity, since the formation of certain types of quasicomponents may depend on cooperative effects.

The effect of electron correlation on the calculated structure of liquid water can be determined by comparing the results based on the HF potential function(Figs. 3-6) with those obtained from the CI potential function(Figs. 7-10). Comparing the results on radial distribution function(Figs. 2,6), the inclusion of electron correlation is found to bring the calculated g(R) into close accord with experiment. The g(R) calculated from a simulation using another potential of Matsuoka et. al. is similar to the one obeained using HF potential. This potential, representative of CI calculations on the water dimer displayed essentially the correlation

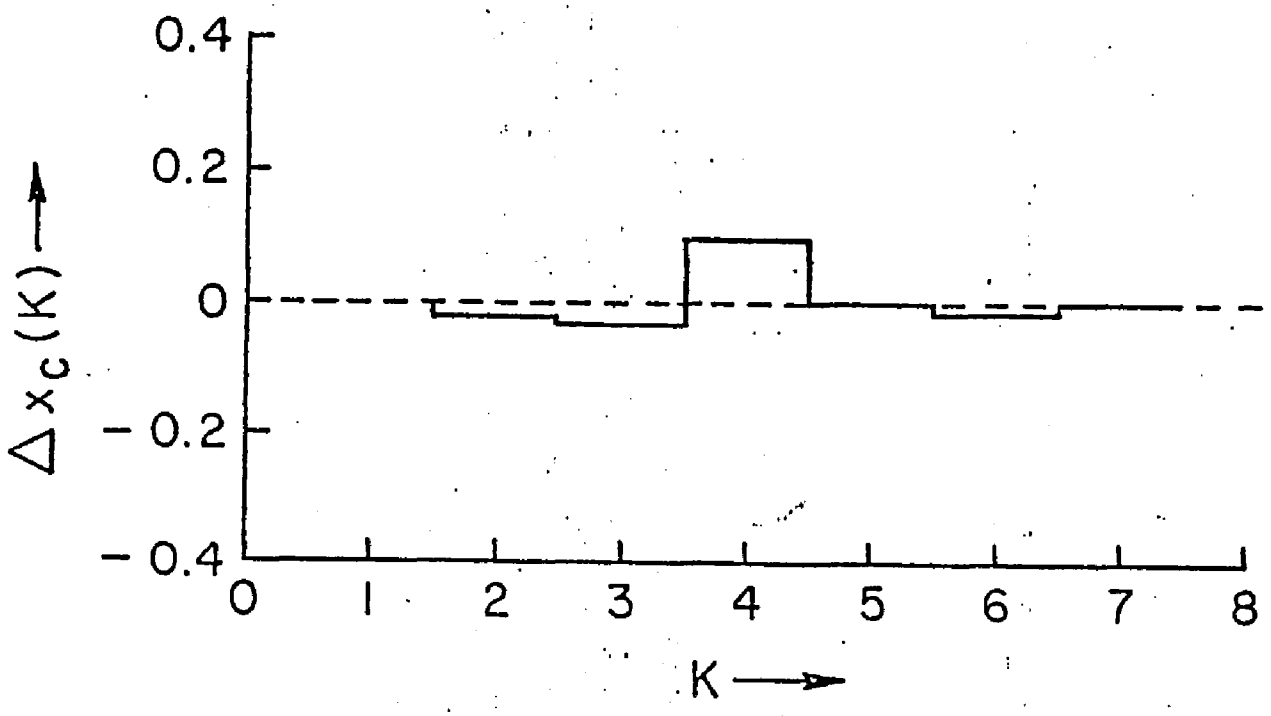


Figure 12. Caculated difference quasicomponent distribution function $\Delta x_c(K)$ for liquid water using CI potential and HF potential; ($x_c(K)$ for CI - $x_c(K)$ for HF)

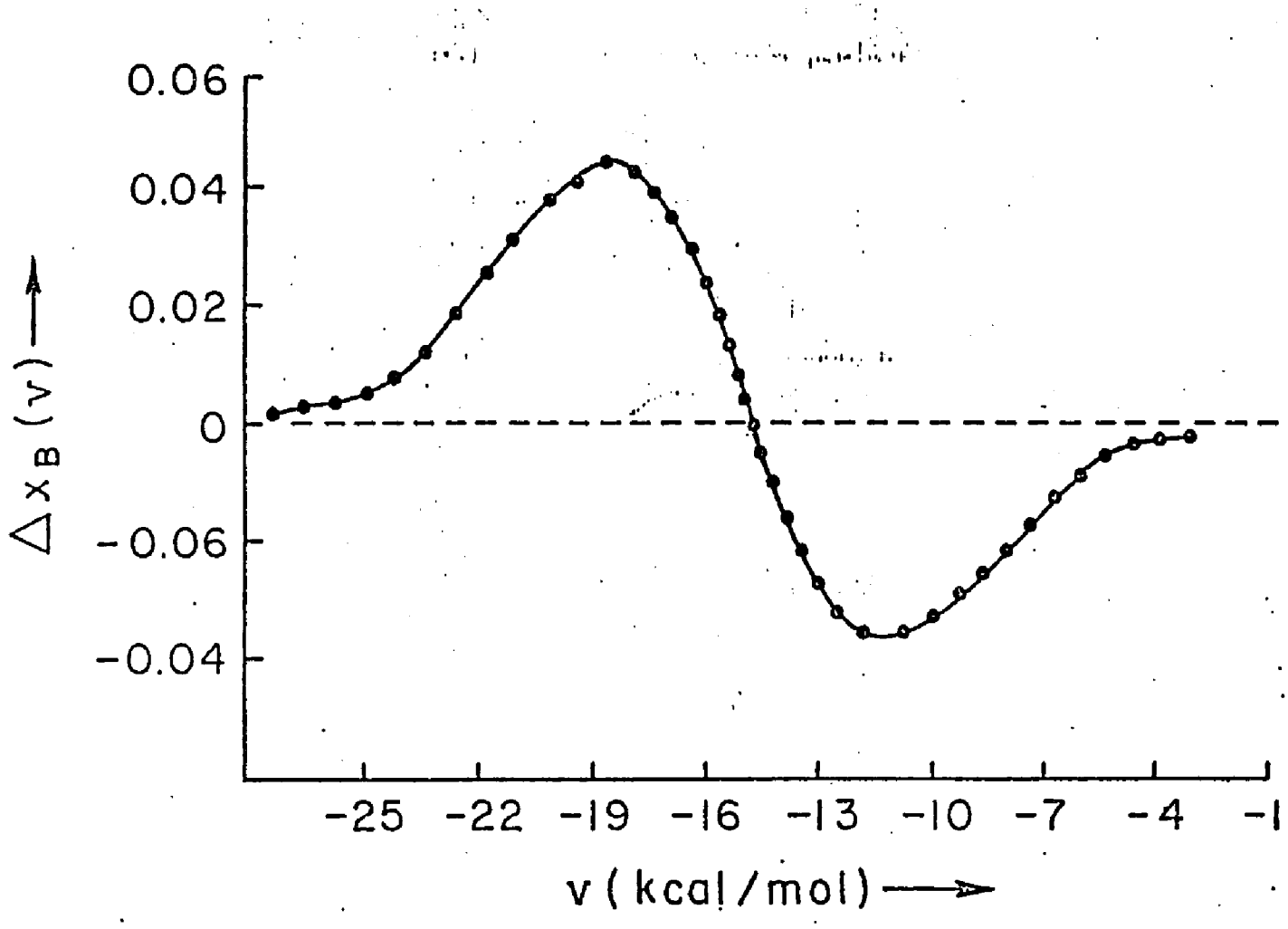


Figure 13. Caculated difference quasicomponent distribution function $\Delta x_{\beta}(\nu)$ for liquid water using CI potential and HF potential; ($x_{\beta}(\nu)$ for CI - $x_{\beta}(\nu)$ for HF)

effect on binding energy, but not on geometry. Therefore the correlation effect on geometry as well as energy is essential for obtaining agreement between calculated and observed radial distribution functions.

The effect of correlation on the calculated structure of liquid water can be quantified in terms of both coordination number and binding energy. This is most directly displayed in terms of difference(Δ) QCDFs,

$$\Delta x_c(K) = x_c^{CI}(K) - x_c^{HF}(K)$$

$$\Delta x_B(V) = x_B^{CI}(V) - x_B^{HF}(V)$$

The Δ QCDF for coordination number is shown in Fig. 12. Correlation effects are seen to increase the four coordinate species at the expense of low density components K=2 and K=3 and the higher density K=6 quasicomponent. The Δ QCDF for binding energy is shown in Fig 13 and displays essentially a simple displacement of binding energy curve towards lower values. Thus the effect of electron correlation on the calculated structure of liquid water is to increase the four-coordinate structure of the system by 10% and decrease the average binding energy of a water molecule by about 6 kcal/mol.

The structures contributing significantly to the statistical state of the system were looked at and two

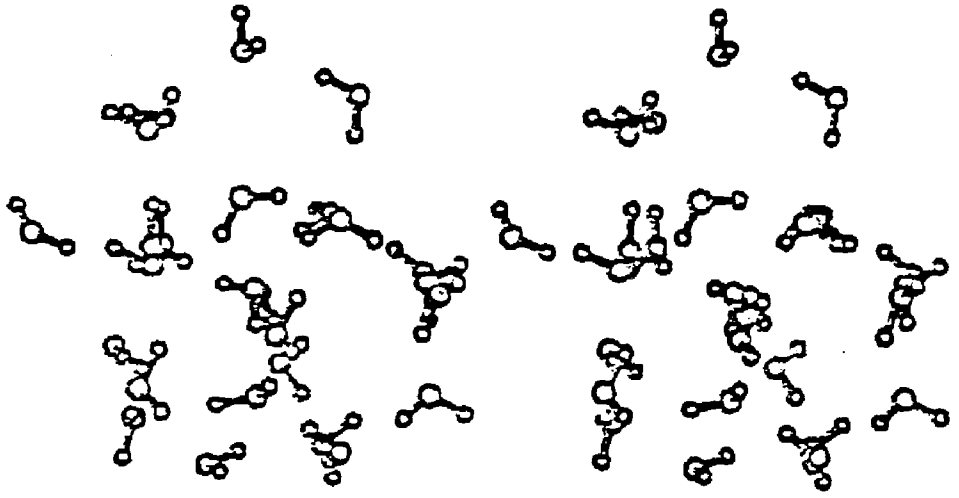


Figure 14. Stereographic view of a fragment of one of the molecular structure that contributes significantly to the statatistical state of liquid water.

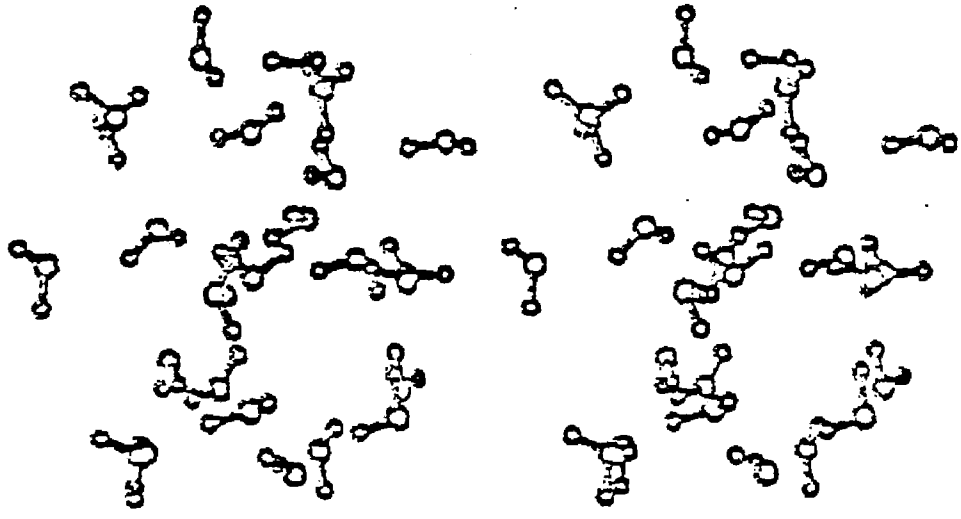


Figure 15. Stereographic view of a fragment of another molecular structures that contribute significantly to the statatistical state of liquid water.

representative samples have been chosen and displayed in Figs. 14 and 15 to visualise the hydrogen bonding patterns in in liquid water.

e) Summary and Conclusions

The analysis of results of Monte Carlo simulations on liquid water which are in satifactory accord with experimentally observed thermodynamic properties and radial distribution functions quantifies the distribution of coordination numbers and binding energies in the system. The predominance of four coordinate quasicomponent in the diffusionally averaged structure of the system is confirmed and a 70/30 ratio of low density to high density component is indicated. Analysis of the results in terms of the distribution of binding energies favors the idea of a continuous distribution structures rather than a mixture model of liquid water, and included in the continuum of structures are contributions from 3,5 and 6 coordinate species as well as the four coordinate component of "Pople-water"[42]. The use of quasicomponent distribution functions is seen to be a powerful means for the analysis of the structural chemistry of the statistical state of a molecular liquid.

The effect of electron correlation on the structure of water is to increase the four coordinate structure of the system and decrease the binding energy of the system. The free energy of liquid water has been successfully calculated using the CI potential. The good agreement between the calculated and observed entropy seems to indicate that the statistical

state of the system is being described well by the CI potential.

IV HEURISTIC POTENTIAL FUNCTIONS

a) Introduction

The development and application of methodology for determining potential functions representative of ab initio quantum mechanical calculations of intermolecular interaction energies is described here. The goal is to establish a systematic and well characterized general approach to obtaining the best quality function in the most economical way possible. To accomplish this, certain variables in the procedure will be permitted to evolve in the course of determining a potential function, leading to a heuristic method. The functions obtained are intended for use in computer simulation studies of molecular assemblies. The methodology was applied to two intermolecular interactions of significant interest in chemistry. They are the formaldehyde-water interaction and methane-water interaction. The calculations and results and discussion will be presented independently for each system.

The following section of this chapter provides background material for the project. The method of procedure is given in detail in section c including considerations on specification of variables and computational technology. The calculations are described in section d, followed in section e by a presentation of results and discussion thereof. Summary and conclusions are collected in section f.

b) Background

Theoretical studies of chemical systems using statistical mechanics and molecular dynamics require rapid accurate evaluation of the configuration energy of the system or the force at various positions in the system.[56] In numerical calculations at present this is accomplished using a potential function, a simple analytical expression for the energy of the system, as a function of configuration coordinates and disposable parameters. Forces can be determined from the gradient of the potential function. The quality of the numerical results obtained in computer simulation studies on chemical systems depends directly on the quality of the potential functions used.

The simplest way in which potential functions are used in numerical calculations is with the assumption of pairwise additivity, wherein the total configurational energy E of an assembly is written as a sum of pairwise interaction energies between the individual particles of the assembly,

$$E(\underset{\sim}{X}^N) = \sum_{A<B} V_{AB} \qquad (1)$$

where $\underset{\sim}{X}^N$ represents the configurational coordinates of N particles and V_{AB} is a pairwise potential function for molecules A and B. The most widely known pairwise potential function is the Lennard Jones 6-12 potential function,

$$V_{AB}(r) = 4\epsilon_{AB}\left[(\sigma_{AB}/r)^{12} - (\sigma_{AB}/r)^{6}\right] \qquad (2)$$

where r is the interparticule separation and ϵ_{AB} and σ_{AB} are disposable parameters characteristic of the interaction of molecules A and B. Numerous other functional forms for pairwise potential functions have been used in various applications. Generally the true intermolecular energy is unknown, so the functional form for the pairwise interaction is chosen with a physical model in mind, and the disposable parameters are determined empirically or semiempirically from independent considerations or experimental data. This type of potential function will be henceforth referred to as an empirical potential function (EPF).

Currently, good estimates of the true intermolecular pairwise interaction energy can be obtained using molecular quantum mechanics. Recent advances in computer technology and development of efficient computational algorithms have made ab initio quantum mechanical calculations tractable for large molecules and molecular complexes. However at present moment the computation time for individual pairwise interaction energies still precludes their direct use in statistical thermodynamics or molecular dynamics calculations.

One currently viable approach to the problem is to use quantum mechanics to generate a data base of interaction energies at various points in configuration space and choose the disposable parameters in an analytical function against the data base in a least squares sense. Thus one obtains an "analytical potential function" (APF) representative of ab

initio quantum mechanical pairwise interaction energies which can be used in computer simulations in the usual manner. Clementi and coworkers have recently pioneered this approach, and their potential functions representative of large basis set ab initio molecular orbital calculations for ion-water and water-water interactions [19-22], have been used in several recent statistical thermodynamic and molecular dynamic studies of water and solution structure, including studies of the thermodynamic indices of ion hydration.[57]

With the feasibility of generating APF's representative of quantum mechanical calculations thus established, there remains to consider the most economical way to proceed. A number of details must be specified in establishing a general procedure for determining APF'S, including a) quality of points in the data base, b) functional from for the APF, c) size of the data base, d) strategy for choosing configurational coordinates of points in the data base, e) weighting of points in the curve fitting procedure and, f) the predictive value of the function. Considering these factors, a methodology has been developed whereby decisions on variables are evolved in the course of the determination of the potential function for a given intermolecular interaction. Thus intrinsic in the procedure is a discovery of the optimum specification of the variables, thus involving a heuristic method leading to a "heuristic potential function" (HPF).

Before passing to details of this procedure it is necessary to comment further on pairwise additivity. While

this approximation holds fairly rigorously for simple non-polar particles, for a preponderent number of systems of chemical interest at least three-body terms will be significant, if not those of even higher order. The HPF procedure can in principle accommodate higher order effects. Initial studies however will be carried out for pairwise potential functions with close attention to the approximations involved in their application to chemical problems. In spite of the problems therein, computer simulation research on molecular systems is currently proceeding on the basis of pairwise or effective pairwise additivity. Evidence of the viability of this approach comes from the excellent agreement between observed oxygen-oxygen pair correlation function for liquid water and that from computer simulation by Clementi et.al. using a pairwise additive APF in a Monte Carlo procedure.

c) Methodology

This section includes a detailed description of the methodology followed by general considerations on specification of variables and computational details.

The HPF method: Consider the problem as the determination of a particular intermolecular potential function representative of a given level of quantum mechanical interaction energies as economically as possible. The other variables in the procedure are related to the data base and the curve fitting procedure. The data base consists of a set energies calculated at various points in the configuration

space of the problem and the points,

$$\underset{\sim}{B} = \left\{ \underset{\sim}{R}, \underset{\sim}{E} \right\} \qquad (3)$$

where $\underset{\sim}{R}$ is the set of entire geometrical parameters (coordinates or interatomic distances) corresponding to the energies $\underset{\sim}{E}$. The specification of data base variables involves determining the number and position of points to be included. The number and disposition of points should be large and diverse enough to be representative of $E(\underset{\sim}{R})$ but small enough to be economical. The curve fitting variables include the tolerance or standard deviation to be required of the potential function and any strategy involving preferential weighting of points in the data base. The tolerance should be low enough to enable, for example, calculation of internal energy at the order of

experimental error of thermodynamic measurements on chemical systems.

The HPF method involves a dynamic decision making process on specification of certain data base and curve fitting variables. The procedure can be illustrated with data base variables. We begin with an initial data base $\underset{\sim}{B}$ consisting of N_1 points, N_1 depending of course on the size of the molecules and the extend of configuration space to be spanned. These N_1 points in configuration space are chosen on the basis of some strategy ; various alternatives will be discussed below. The energies $\underset{\sim}{E}_1$ are calculated quantum mechanically for these

positions, giving $\underset{\sim}{B}_1$. An analytical functional form V is adopted (see also below) and fit to the initial data base by minimizing the sum of squares of the deviation Δ_1, Where

$$\Delta_1 = \frac{1}{N_1} \sum_{i}^{N_1} \left[E_i(\underset{\sim}{R}_i) - V_i(\underset{\sim}{R}_i, \underset{\sim}{a}, \underset{\sim}{b}, \cdots) \right]^2 \quad (4)$$

with $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... being symbols for sets of disposable parameters. The result is the analytical potential function

$$V_1(\underset{\sim}{R}, \underset{\sim}{a}, \underset{\sim}{b}, \cdots) \quad (5)$$

where $\underset{\sim}{a}$, $\underset{\sim}{b}$, ... give the best fit of the function to the data base in a least squares sense. The statistics on the curve fitting procedure are fully generated, including standard deviation σ_1, residuals, etc, and tell us how well the APF fits the data base. At this point we are also interested in a more difficult question : how well the data base itself represents true energy in the configuration space. To estimate this we generate a "test" data base $\underset{\sim}{B}_1' = \{\underset{\sim}{R}_1', \underset{\sim}{E}_1'\}$ consisting of N_1' points outside the initial data base. The predictive value of the function is tested by using the function to calculate the points on $\underset{\sim}{B}_1'$ and comparing them with their quantum mechanically clculated counterparts, $\underset{\sim}{E}_1'$. The resulting standard deviation is a test of the predictive value of the V_1.

The statistics σ_1 and σ_1' together indicate the quality of V_1. These can be compared against a prescribed desired

tolerance for the function. Initially the tolerance is unlikely to be satisfied and steps must be taken to improve the quality of the function. The most direct way is to increase the size of the data base. Since we have N_1' additional points already calculated for test purposes, the easiest way to do this is to concatenate the initial data base and the test data base. This gives a data base N_2 of a size $N_1 + N_1'$, for which we can redetermine the APF. The result of this step is $V_2(R, \underset{\sim}{a}, \underset{\sim}{b}, \ldots)$ with a standard deviation σ_2 and predictive value on a new test data base $\underset{\sim}{B}_2'$ of σ_1'. This procedure is repeated until an APF within tolerance is obtained, and for which the statistics are stable on successive iterations. Thus in the course of determining the potential function we have evolved a data base representative enough of energy in configuration space to produce a potential function within precribed tolerances. A straightforward extension of this procedure allows curve fitting vaiables as well as data base variables to evolve in the procedure.

Specification of variables: The totality of variables involved in the procedure can be conveniently grouped into "static" and "dynamic"variables, depending upon whether a variable is fixed or evolves in the course of an HPF determination. The static variables are the analytical form adopted for the potential function and the quality of quantum mechanically calculated points in the data base. The dynamic variables are certain data base variables and curve fitting variables.

The data base variables in the HPF procedure are a) the size of the initial data base N_1 b) the size of the test sets N_i and c) the position of the points in configuration space. Note that N_1 and N_i are static data base variables whereas the total number points in the data base, is a dynamic variable. N is given by the expression

$$N = N_1 + \sum_{i=1}^{m-1} N_i$$

where m is the number of cycles to obtain stability and statistics within tolerance in the procedure. The size of the initial data base will depend on the volume of configuration space under consideration. In considering alternative strategies for positioning initial data base points in configuration space, we can identify two extremes : random sampling and grid search. In random sampling, the points are chosen by a Monte Carlo technique using a random number generator. Grid search involves laying a multidimentional grid on configuration space and taking data base points at each intersection. Either of these two procedure could be modified by importance sampling, concentrating data base points in regions of special interest. The initial efforts as described herein are based on a random sampling strategy.

The curve fitting procedure also involves many aspects, some static and some dynamic with respect to the determination of the HPF. The curve fitting method itself is a multi-dimensional non-linear least squares procedure using an

algorithm based on Powell's method of isoconjugate directions.[66] The dynamical aspects of the curve fitting procedure are involved in the preferential weighting of certain points. This can be keyed on the requirements of the potential function for computer simulation. For example in the evaluation of ensemble averages, only values of energies below a certain threshold will make a significant contribution. This suggests tailoring the potential function to describe the low energy regions of the system well. This can be accmplished by concentrating data base points in the low energy regions(importance sampling) or by weighting the low energy points in the data base heavier in the curve fitting procedure.

The procedure adopted in the present studies is preferential weighting. The calcualtions involve a weighting function of the form "1 + αexp(- E/kT)" in the curve fitting procedure minimizing the mean square deviation

$$\Delta = \frac{1}{N} \sum_i \left[E_i \left(\underset{\sim}{R}_i \right) - V_i \left(\underset{\sim}{R}_i, a, b, \cdots \right) \right]^2 \left(1 + \alpha e^{-\Delta E_i / kT} \right) \qquad (7)$$

Here the weighting parameter may also evolve dynamically in the course of determining the HPF or be preset on the basis of exploratory calculations. The standard deviation below threshold $\sigma_<$ and the frequency of errors γ wherein a point above threshold is predicted to be below threshold is computed for each cycle. The parameter α can be adjusted in small increments to minimise σ and γ simultaneously in the HPF

procedure.

The remaining variables in the procedure to discuss %are the functional form of the potential functions generated and the quality of points in the data base. A functional form that is general for most intermolecular interaction and easily usable in the simulation is

$$V_{AB} = \sum_{i<j} \left[a_i a_j / r_{ij}^{12} + \sum_n \left(b_{in} b_{jn} / r_{ij}^n \right) \right], \quad \begin{matrix} i \in A \\ j \in B \end{matrix} \qquad (8)$$

where the i's and j's are atomic and pseudo-atomic centers on molecules A and B respectively and r_{ij} is the inter-center distance. The parameters a is essentially the "hard part" and b_n is the "soft part" of the function and they are determined in a least squares sense.

The quality of points in the data base is determined by level of quantum mechanical calculation of energies. The HPF procedure is of course independent of the method used for energy evaluation and thus the entire range of empirical, semi-empirical and non-empirical methods of molecular quantum mechanics are possible alternatives. The calculations described herein are at the non-empirical level.

Computational details: The HPF method described above can be automated to a higher degree using the sophisticated operating systems of third generation digital equipment. The principal programs involved are the quantum mechanical programs for energy calculation and a set of programs collectively referred to as "HPF" which manages data base

generation, curve fitting and the overall flow of computational events. The HPF program uses the IBM system utility ISDRVR together with dependent job processing to submit ordered jobs to the input stream. By linking the jobs submitted, HPF can transfer control to other programs for energy evaluation and then be automatically restarted after the energies are available.

C) Calculations

1) Formaldehyde-Water

As an initial application, the formaldehyde-water system has been chosen. The formaldehyde-water system is a prototype for the hydration of a carbonyl group and is thus of both chemical and biochemical interest. Previous theoretical studies of this system are due to Iwata and Morokuma [59], Johansson and Kollman [60], and Del Bene [61].

The data base for the HPF determination is most conveniently specified in terms of the geometry of water relative to formaldehyde with the internal coordinates of each moiety held fixed. All geometries were generated by random deployment of positional and orientational coordinates of water within 5.5 Å of the formaldehyde molecule. This distance represents twice the average diameter of a water molecule and is intended to produce a function capable of describing the formaldehyde-water interaction in the region of the first and second hydration shells. The choice of a functional form (see below) with proper long range limiting behavior effectively extends the range of the potential function to all

configuration space within good approximation. The initial data base for the calculation consisted of 100 points and all test sets involved 25 points each.

The formaldehyde-water(F-W) intermolecular interaction energies were calculated using matrix Hartree-Fock methods incorporating a 6-31G basis set[62] on each atom. The 6-31G basis was developed by Pople et. al.[62] and is an extended set with split valence shell functions. It was selected for calculations herein as the largest basis feasible within local economic constraints. The effects of basis set distortion were neglected. The intramolecular geometries for both moieties were maintained at their available experimental values in all calculations. The intramolecular geometrical parameters for formaldehyde were taken as CO=1.203 Å, CH=1.101 Å $\widehat{HCO}$=121.75° and $\widehat{HCH}$=116.5°. For water we use OH=0.958 Å and $\widehat{HOH}$=105.° The computer time required for obtaining each individual formal dehyde-water energy was approximatly 4 min. using the programme Gaussian-70 on the IBM 370-168.

The HPF procedure was carried out on this data base using two different functional forms, both adaptations of equation 8. In both cases parameters were identified with each of the atoms of the molecules. In addition two pseudo-atomic centres were introduced for each molecule, coordinated to the oxygen atoms. These types of parameters have been proved advantageous in previous empirical and analytical potential functions for water-water interaction. In the case of water the pseudo atomic centres were tetrahedrally coordinated to oxygen and in

the formaldehyde case, the centres were trigonally coordinated to oxygen at a distance of 0.1852 Å in both the cases. The explicit forms of the potential functions used are

$$V_{12-1} = \sum_{\substack{i<j \\ i \in F \\ j \in W}} \left(a_i a_j / r_{ij}^{12} + b_i b_j / r_{ij} \right) \quad (9)$$

henceforth referred to as the 12-1 potential, and

$$V_{12-1} = \sum_{\substack{i<j \\ i \in F \\ j \in W}} a_i a_j / r_{ij}^{12} + b_i b_j / r_{ij} + c_i c_j / r_{ij}^{3} \quad (10)$$

hereafter referred to as the (12-1-3) potential. The summations extend over all atoms and pseudoatoms of each molecule. A weighting procedure described in equation 9 was employed. Exploratory calculations showed best results were obtained with α=100 and this was the value of α used for all calculations.

2) Methane-water

A methane-water potential function is a prerequisite for theoretical studies of a dilute aqueous solution of methane, a prototype for a non-polar solute in liquid water. A detailed knowledge of the structure of methane-water solution at the molecular level can provide preliminary information on the more general problem of the interaction of water with dissolved hydrocarbon chains, and ultimately an understanding of the role of water in maintaining the three-dimensional

structural integrity of biological polymers in solution.

Ab initio self-consistent field calculations of the methane-water intermolecular interaction energies have been reported previously for a limited number of configurations.[63] Single configuration wave functions were used with both minimal and extended basis sets. The latter were double C,O(9s 5p/4s 2p) and H(4s/2s) contracted primitive sets. These studies yielded some detailed information about the H_2O-CH_4 potential energy surface within limits of neglect of dispersion forces inherent in the Hartree Fock(HF) model. It should be noted that correlation effects are expected to be extremely significant in the methane-water interaction. Methane-water;HF calculations do show a shallow minimum due to dipole-induced dipole and also higher order interactions, indicating that potential functions representative of this level of calculations, properly spanning the configuration space, could be useful for preliminary statistical mechanical studies of dilute solutions of simple hydrocarbon solutes.

The calculation of the individual methane-water interaction energy was carried out using matrix HF methods on 6-31G basis set.[62] The 6-31G basis is a split valence shell set nearly commensurate in quality with that used by Ungemach and Schaefer and was used in the HPF determination for the formaldehyde-water interaction discussed previously.

Individual entries in the data base for the HPF determination involve the interaction energy for specific methane-water geometries. The geometries are generated by

Table 1

Statistics and in the determination of (12-1-3) HPF for formaldehyde-water.

HPF Cycle	Points in Data Base	Initial Sample		Test Sample	
i	N	σ_ζ	ν	σ_ζ	ν
1	100	0.22	0.0	0.68	0.0
2	125	0.23	0.0	0.28	0.0
3	150	0.20	0.0	1.15	0.04
4	175	0.23	0i006	0.58	0.0
5	200	0.23	0.005	0.36	0.0

random positional and orientational deployment of the water molecule with respect to the methane with internal coordinates of each particule held fixed. The center of mass of the water molecule was maintained within 5.5 Å of the center of mass of methane molecule, allowing for direct sampling in the configuration space of the first and second hydration shells. The internal coordinates adopted were : for H_2O, r(OH) = 0.957 Å (HOH) = 104.52°; and for tetrahedral CH ,r(CH) = 1.094 Å.

As in the previous case studied, the initial data base consisted of 100 points and the subsequent test sets were comprised of 25 points each. The analytical function V used on the curve fitting part of the procedure is

$$V_{12-1-3} = \sum_{ij} a_i a_j / r_{ij}^{12} + b_i b_j / r_{ij} + c_i c_j / r_{ij}^3 \qquad i \in CH_4,\ j \in H_2O \qquad (11)$$

where the summation over i and j runs over atomic and pseudo-atomic centers in the molecule and the a's and b's are disposable parameters. Pseudo-atomic (PSA) centers are tetrahedrally coordinated to the oxygen atom in water at a distance of 0.1852 Å. Terms in r^{-12}, r^{-1} and r^{-3} were sufficient for our purposes, i.e. The (12-1-3) potential used previously.

d) Results and discussion

1) Formaldehyde-Water

The statistics of the HPF determination of a potential function for formaldehyde-water interaction based the (12-1-3) form are given in Table 1. For the initial data base of 100

Table 2

Parameters a, b and c for atomic and pseudo-atomic centers in the (12-1-3) HPF for formaldehyde-water.

Molecule	Center	$\underset{\sim}{a}$ $\hbar^{-12}$	$\underset{\sim}{b}$ $\hbar^{-1}$	$\underset{\sim}{c}$ $\hbar^{-3}$
	O	-6.33E-1	-2.06E00	9.71E-1
	H	-9.59E-4	3.93E-1	-1.15E-1
	PSA	+1.20E-1	6.73E-1	-3.50E-1
	C	-5.40E06	-2.01E-1	-1.35E01
	O	-1.05E06	3.75E00	1.93E01
	H	-3.09E05	2.54E-1	1.84E00
	PSA	-4.06E05	-2.08E00	-9.41E00

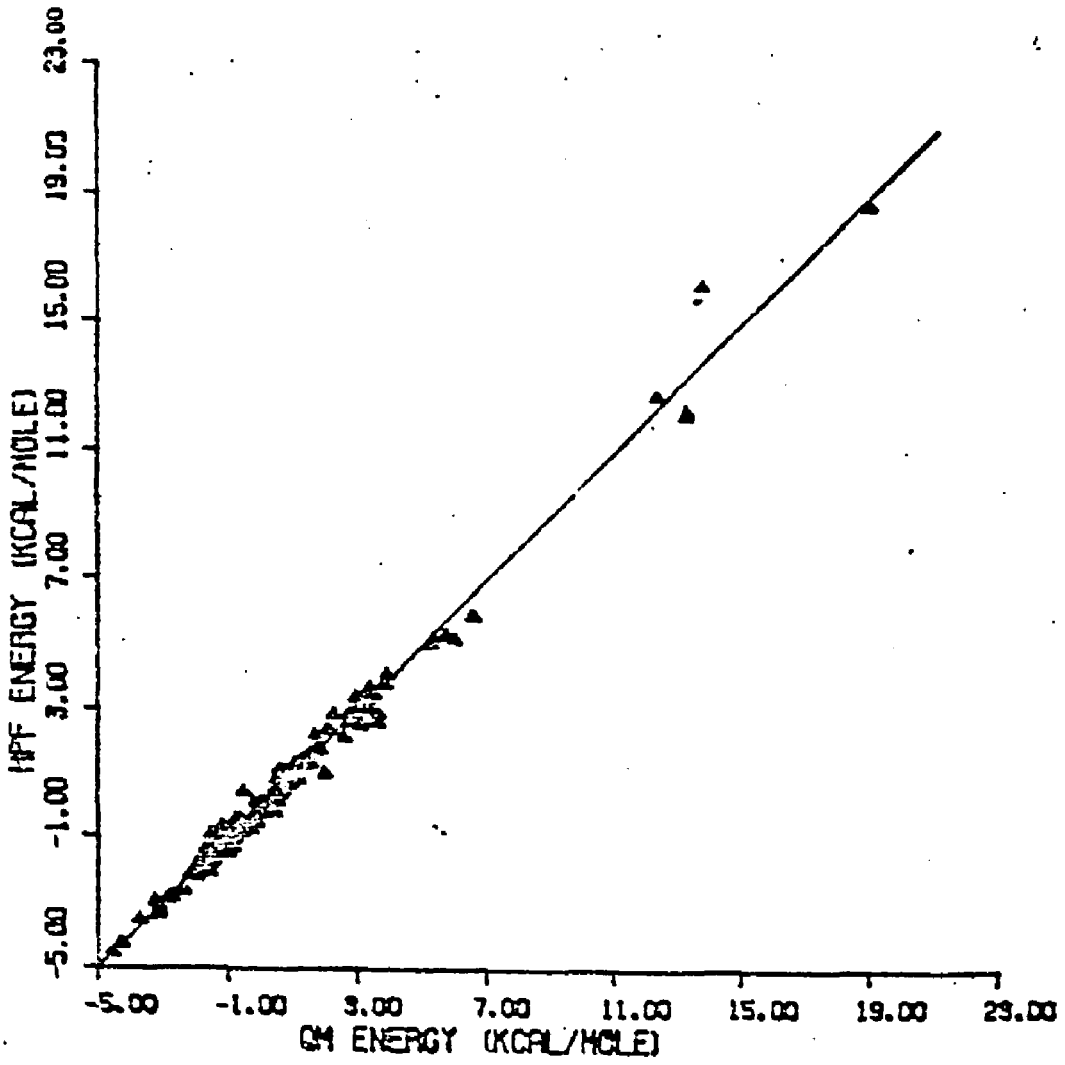


Figure 16. Formaldehyde-Water interaction energies calculated from (12-1-3) HPF vs. quantum mechanically calculated values.

points, the standard deviation was 0.22 kcal/mol. The test set of 25 points had a standard deviation of 0.68 kcal/mol in the first cycle. The calculation was continued for four cycles. The standard deviation of successive total sample remained well within acceptable limits. On the second cycle, the description of the test set degenerated slightly, indicating poor description of the configuration space sampled. On the following cycles the statistics improved, confirming the self-consistent nature of the HPF procedure. The final curve fitting of the 200 data points resulted in a standard deviation of 0.23 kcal/mol. The final values for the disposable parameters are given in Table 2.

A further indication of the quality of the function is found from a plot of the values predicted for each of the 225 energies in the data base vs. Their quantum mechanically calculated value. This plot for the (12-1-3) function is shown in Fig. 16. The standard deviation gives the composite measure of scatter, but the plot carries additional information on the function. The effect of preferential weighting of low energy points is clearly evident.

As a final test as well as an initial application of the function, the nature of the formaldehyde-water interaction was investigated by generating contour diagrams of energy as a function of intermolecular geometry. The search was limited to two regions which are of structural chemical interest: one in which the center of mass of water was constrained to be in the plane of the formaldehyde and the other in which the center of

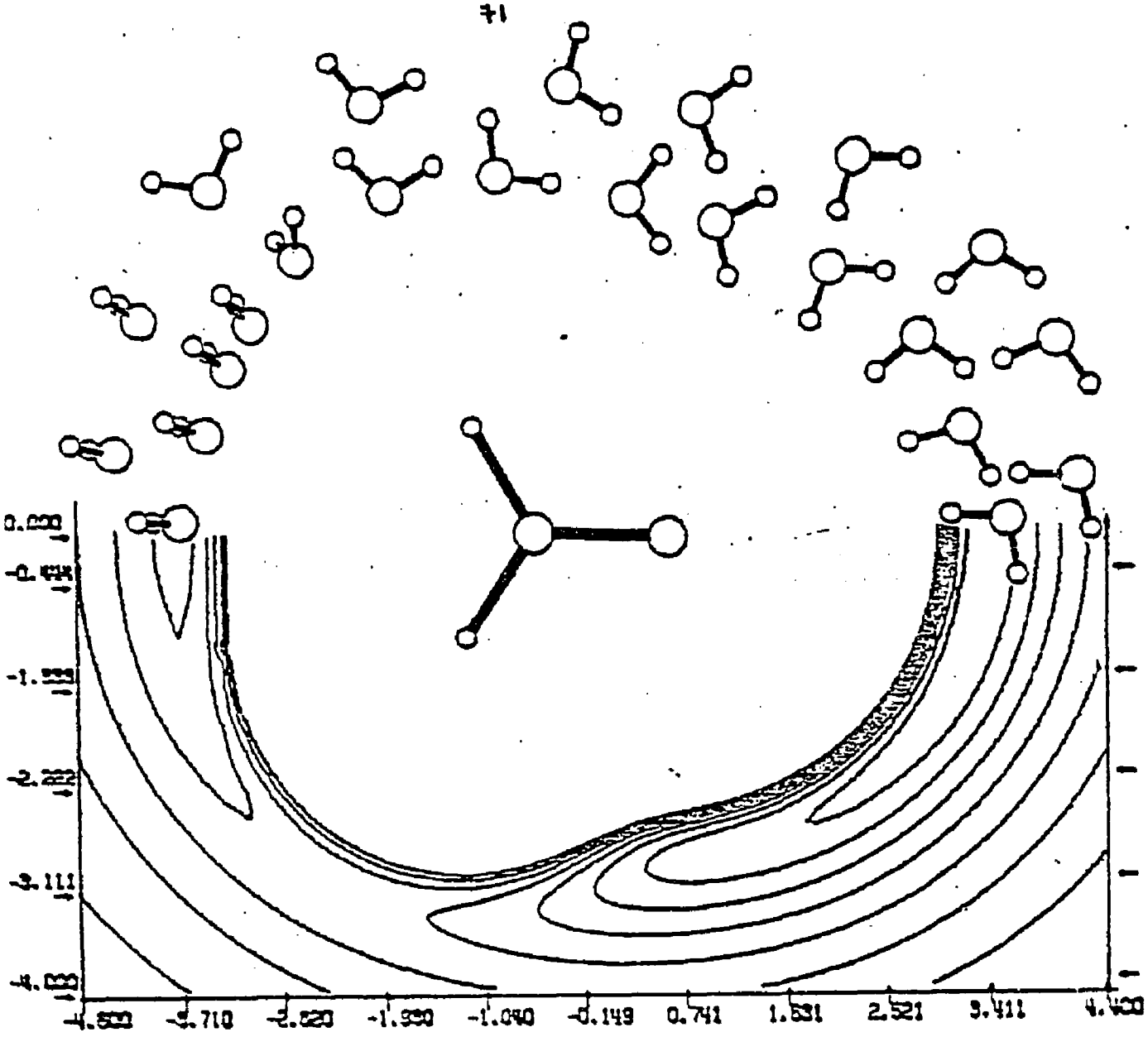


Figure 17. Isoenergy contour map of orientationally optimised formaldehyde-water pairwise interaction energies calculated from (12-1-3) HPF in the plane of the formaldehyde molecule. The distance coordinates refer to the separation between the centers of mass of formaldehyde and water in A. The molecular geometry corresponding to a given energy is depicted in a mirror image position in the top half of the figure. The relative size of the water molecule structures are scaled down for clarity.

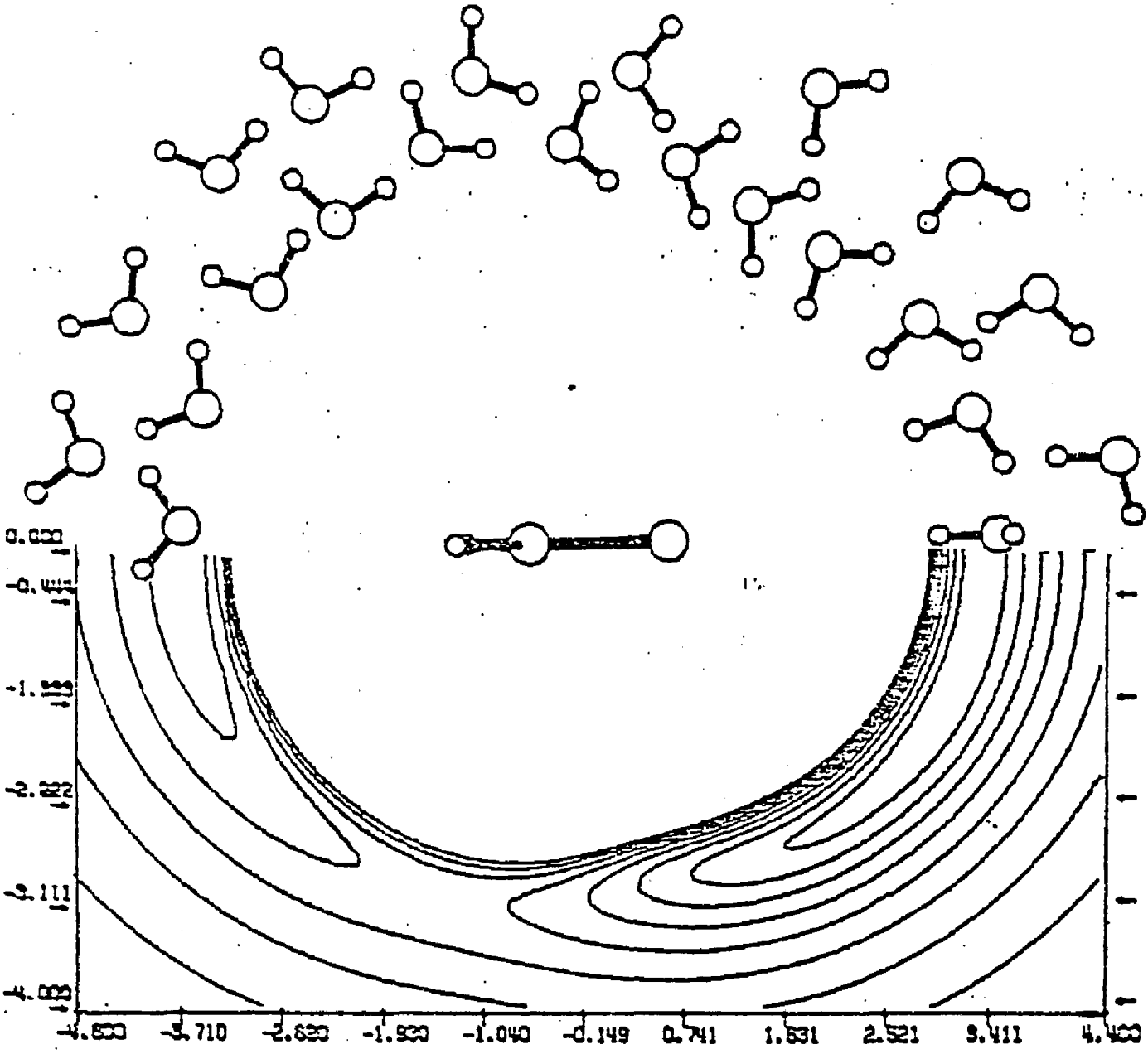


Figure 18. Isoenergy contour map of orientationally optimised formaldehyde-water pairwise interaction energies calculated from (12-1-3) HPF in the plane coincident with the C-O bond and perpendicular to the plane of the formaldehyde molecule. See caption for figure 18 for details.

Table 3

Statistics and in the determination of (12-1) HPF for formaldehyde-water.

HPF Cycle	Points in Data Base	Initial Sample		Test Sample	
i	N	σ_z	ν	σ_z	ν
1	100	0.47	0.0	0.41	0.0
2	125	0.47	0.0	0.43	0.0
3	150	0.37	0.0	0.41	0.04
4	175	0.37	0.006	1.34	0.0
5	200	0.44	0.005	0.45	0.0

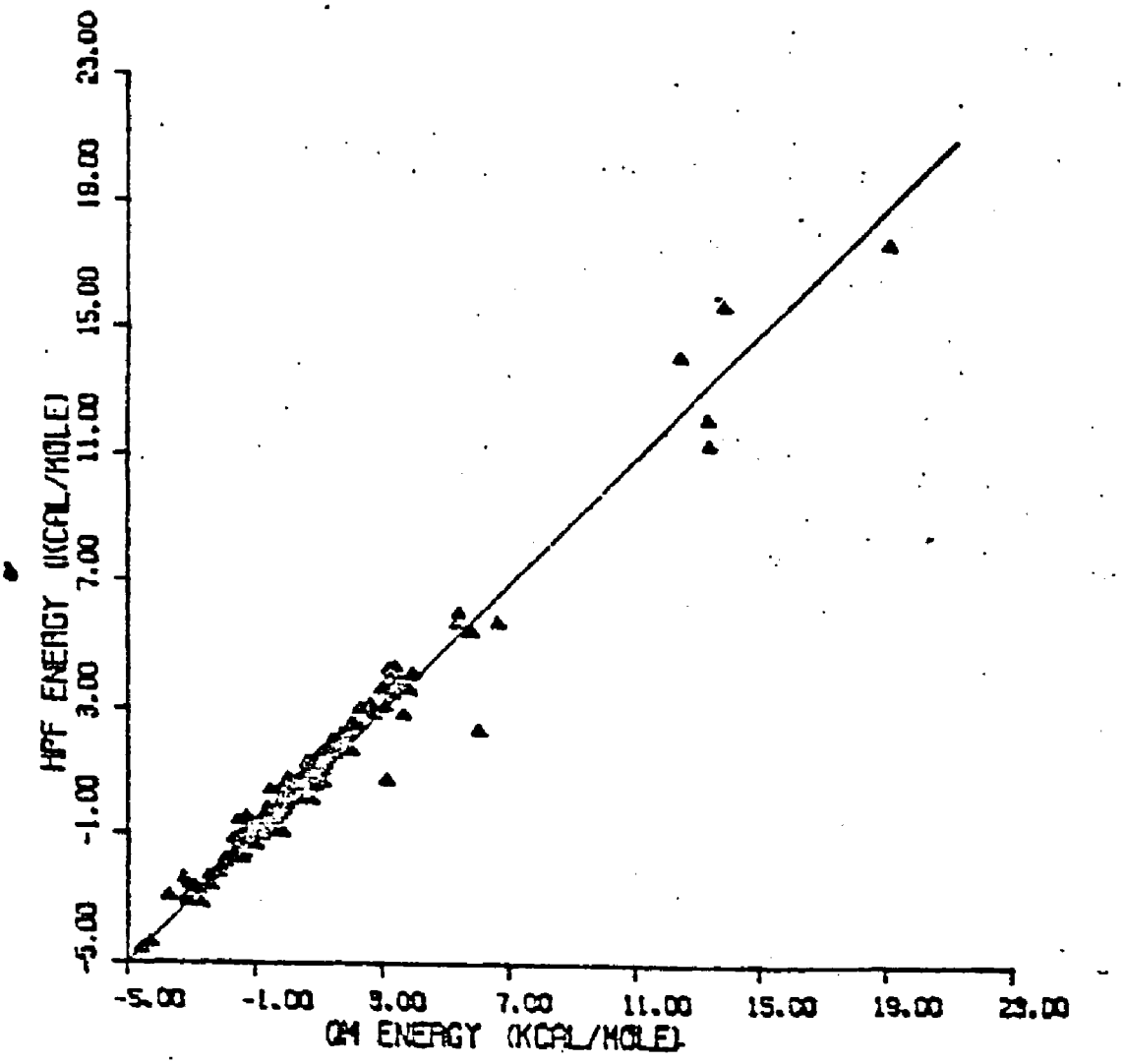


Figure 19. Formaldehyde-Water interaction energies calculated from (12-1) HPF vs. quantum mechanically calculated values.

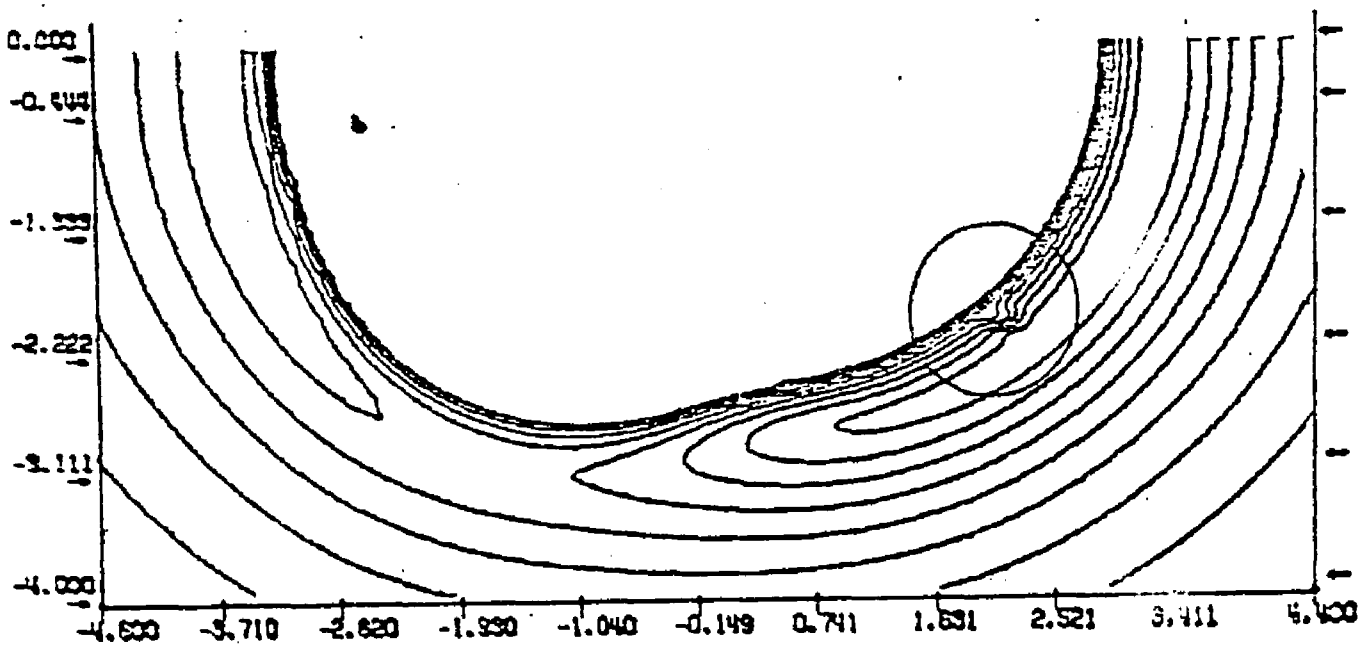


Figure 20. Isoenergy contour map of orientationally optimised formaldehyde-water pairwise interaction energies calculated from (12-1) HPF in the plane of the formaldehyde molecule. See caption for figure 18 for details.

Table 4

Standard deviations for the HPF cycles in the determination of the MO function for methane-water interaction.

HPF Cycle	Points in Data Base	Standard Deviations $\sigma_{\zeta 11}$		Standard Deviations $\sigma_{\zeta 2}$	
i	N	Initial	Test	Initial	Test
1	100	0.22	0.43	0.14	0.18
2	125	0.26	1.32	0.13	0.10
3	150	0.29	0.61	0.13	0.17
4	175	0.41	0.68	0.14	0.13
5	200	0.44	0.42	0.14	0.24

mass of water was perpendicular to the molecular plane and coincident with the plane of carbonyl group. The energy of each particular formaldehyde-water positional configuration was minimised with respect to the orientation of the water molecule. The resulting isoenergy contour maps are shown in the lower half of Figs. 17 and 18 and the orientations of the water molecule are shown in the upper half of the same figures. The plot demonstrates the behavior of the function.

The initial HPF determination was carried out using the (12-1) form of the potential function. The statistics of the HPF cycle are given in Table 3. and scatter of points in Fig. 19 The statistics do not indicate any serious deficiencies in the function.

Examination of the function on the basis of isoenergy contour map in the plane of formaldehyde shown in Fig. 20 reveals an unphysical region, indicating the inadequacy of the function in this region. This problem, while not significant enough to influence the results of the simulation was clearly rectified in the improved fit gained in the (12-1-3) function.

2) Methane-Water

The statistics of the HPF determination are collected in Table 4. The standard deviations $\sigma_{<11}$ and $\sigma_{<2}$ are reported, the former including all points below 11 kcal/mol below threshold and the latter giving detailed statistic on the binding region of the probelm. For the initial data base of 100 points σ below threshold was 0.22 kcal/mol. The values predicted for 25 randomly chosen points showed a standard deviation of 0.43

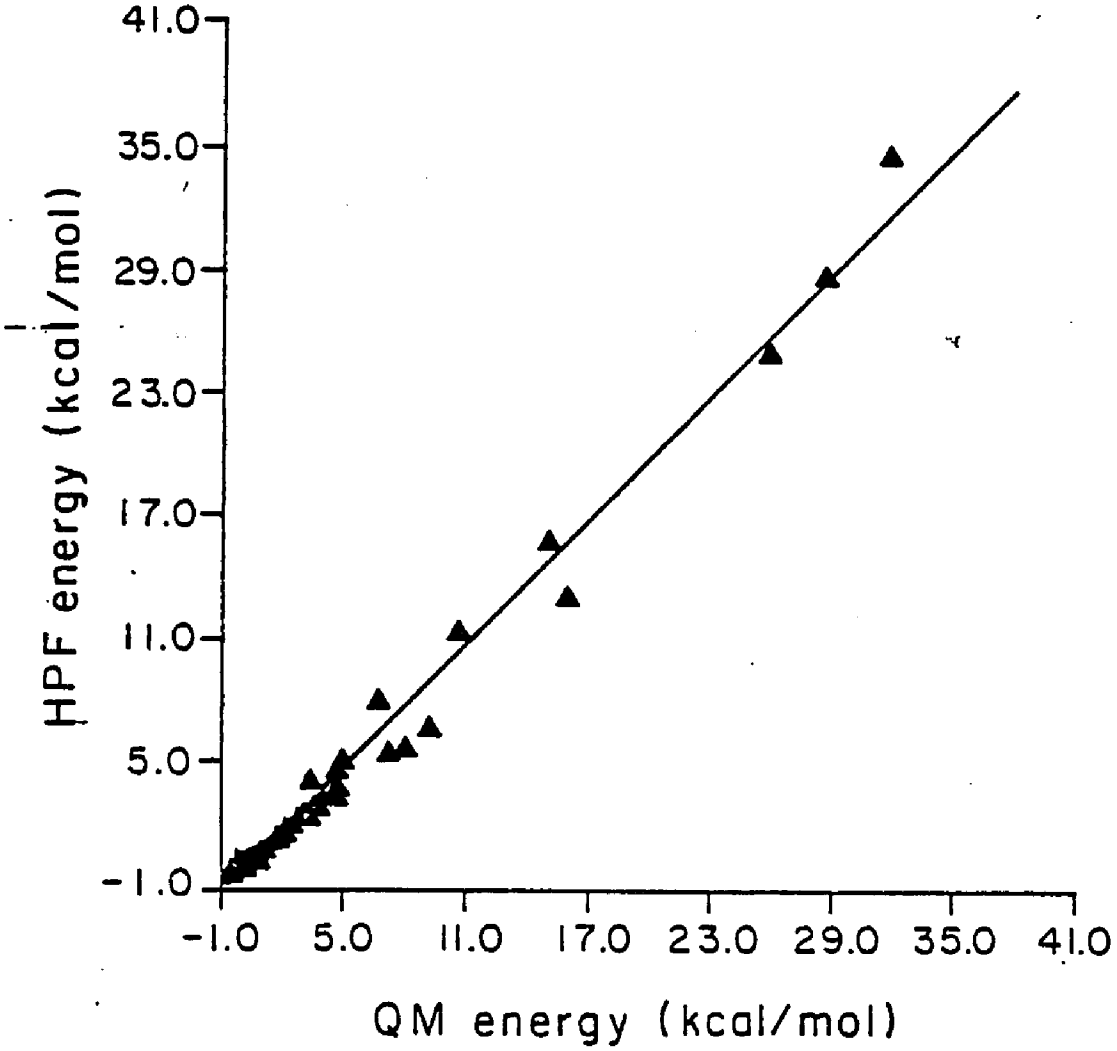


Figure 21. Methane-Water interaction energies (MO) calculated from (12-1-3) HPF vs. quantum mechanically calculated values.

Table 5

Parameters $\underset{\sim}{a}$, $\underset{\sim}{b}$ and $\underset{\sim}{c}$ for atomic and pseudo-atomic centers in the MO function for Methane-water.

Molecule	Center	$\underset{\sim}{a}$	$\underset{\sim}{b}$	$\underset{\sim}{c}$
		r^{-12}	r^{-1}	r^{-3}
	O	2.88E05	-1.37E02	4.98E-3
	H	1.03E04	1.14E01	1.24E-3
	PSA	-4.12E04	3.78E00	4.19E-4
	C	1.19E01	-9.75E-3	-9.07E02
	H	4.72E-2	2.41E-3	1.93E02

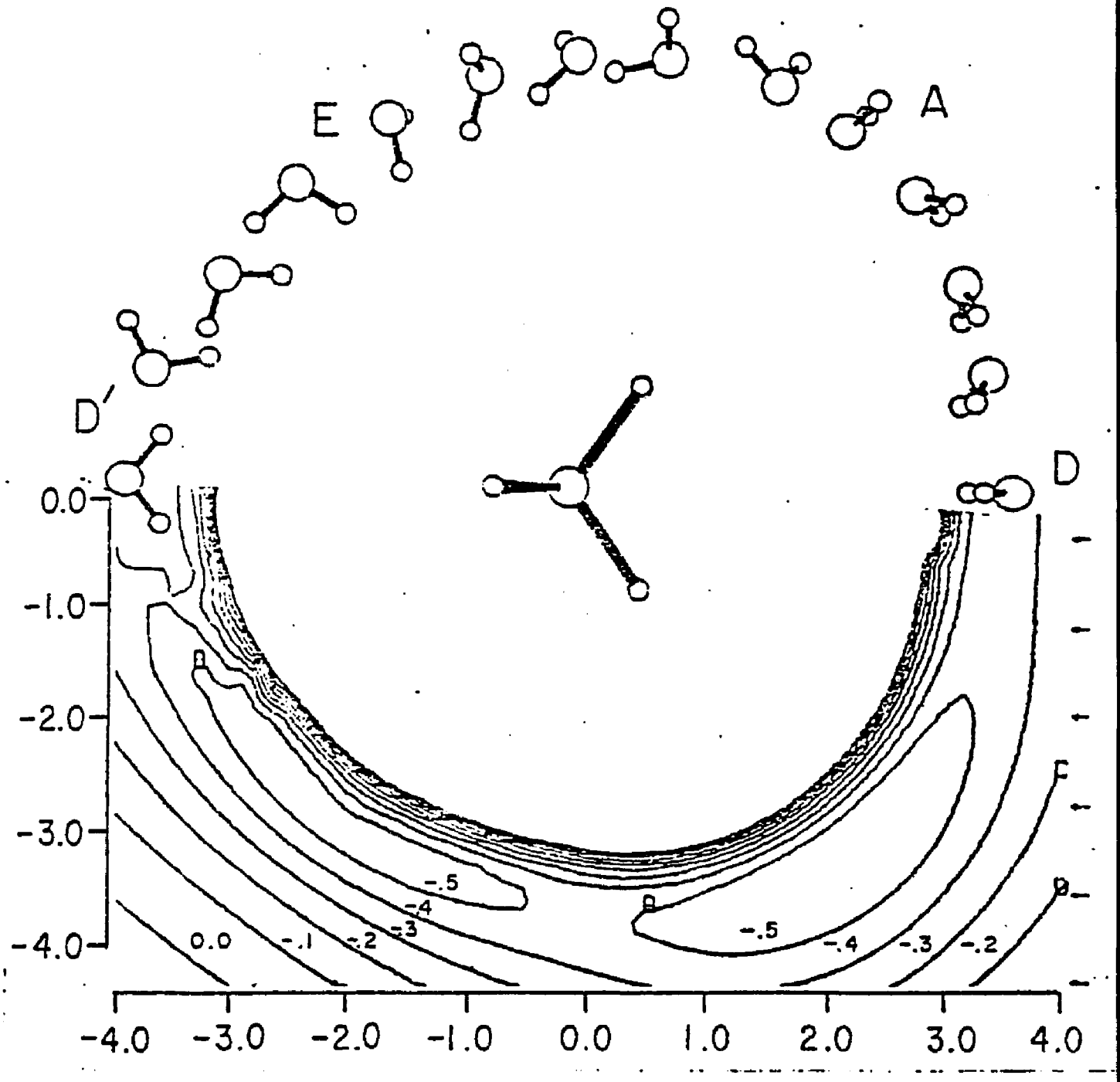


Figure 22. Isoenergy contour map of orientationally optimised methane-water pairwise interaction energies calculated from (12-1-3) HPF for STO 6-31G MO potential in the H-C-H plane. See caption for figure 18 for details.

kcal/mol.

On the second cycle the decription of the test set for all points below threshold degenerated. On the following cycles the predictive value of the function shows a commensurate overall, but montonic improvement. This confirms the self consistent nature of the procedure.

An overall indication of the quality of the function is conveyed also by Fig. 21, a plot of the calculated energy of data base points vs. The energy predicted by the function. This function is described in Table 5. Further details on the function obtained and basic information about the STO 6-31G molecular orbital calculation are presented in Fig. 22. The lower half of figure 22 is an isoenergy contour map of orientationally optimized methane-water interaction energies in a methane H-C-H plane, based on 0(1000) interaction energies determined using the methane-water HPF. The corresponding intermolecular interaction geometries for selected points are given in the top half of the figure.

The long range limiting behavior of the function is satisfactory, smoothly approaching zero. There are three distinct minima in evidence, labelled A, E and D (D') in figure 22, following the notation used by Ungemach and Schaefer. The minimum A corresponds to the lowest energy configuration of the system, stabilized by a C-H ... O interaction.

The structure of liquid water was previously observed to be quite sensitive to electron correlation effects in the

Table 6

Parameters $\underset{\sim}{a}$, $\underset{\sim}{b}$ and $\underset{\sim}{c}$ for atomic and pseudo-atomic centers in the MP function for Methane-water.

Molecule	Center	$\underset{\sim}{a}$	$\underset{\sim}{b}$	$\underset{\sim}{c}$
		R^{-12}	R^{-1}	R^{-3}
	O	-4.38E06	-5.65E02	+1.38E-2
	H	-2.62E05	6.08E01	-8.79E-4
	PSA	-5.93E05	1.88E02	-4.56E-3
	C	-4.07E-1	-9.78E-3	-1.28E03
	H	-1.61E-2	2.40E-3	2.05E02

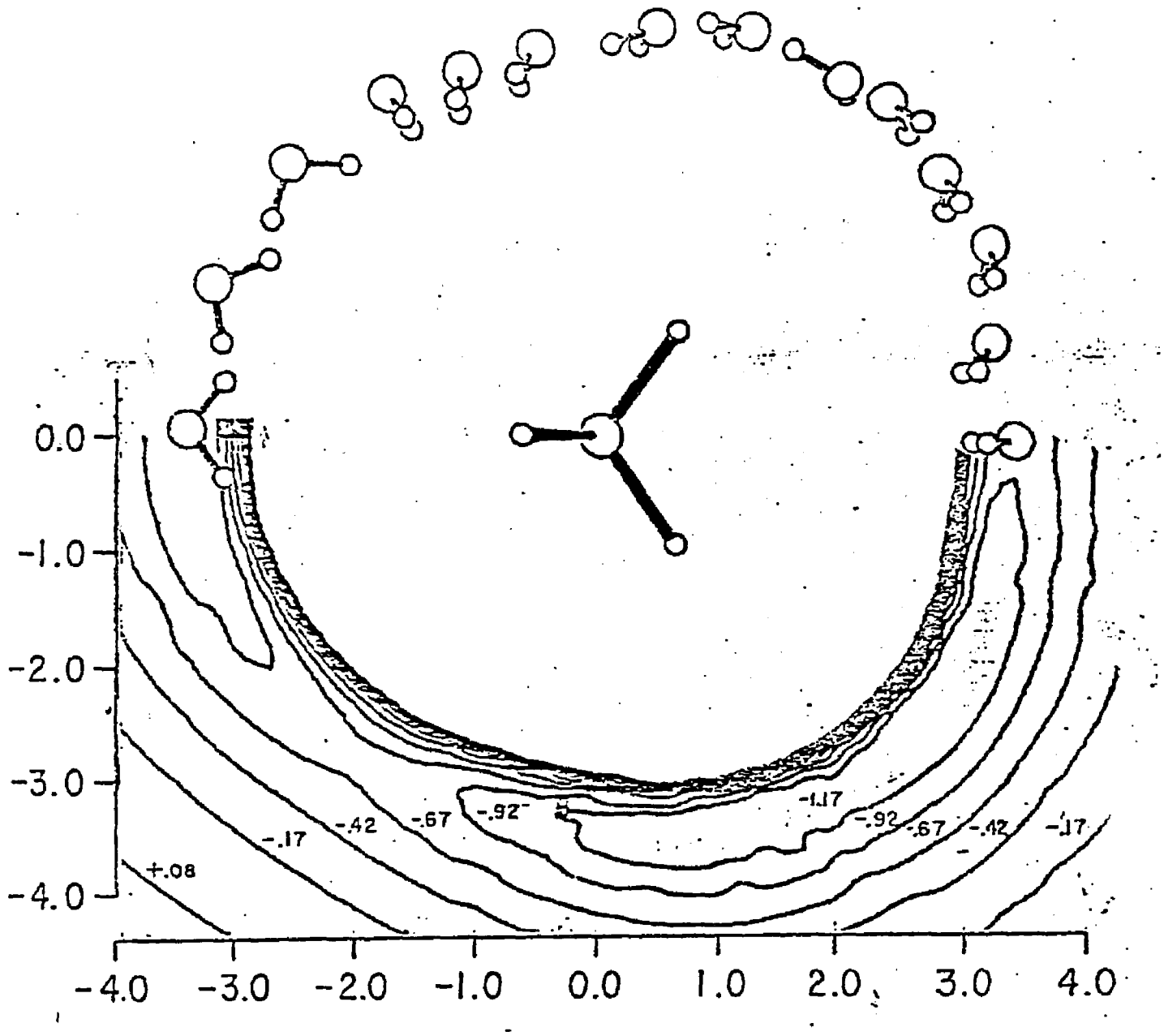


Figure 23. Isoenergy contour map of orientationally optimised methane-water pairwise interaction energies calculated from (12-1-3) Møller-Plesset (MP) potential in the H-C-H plane. See caption for figure 18 for details.

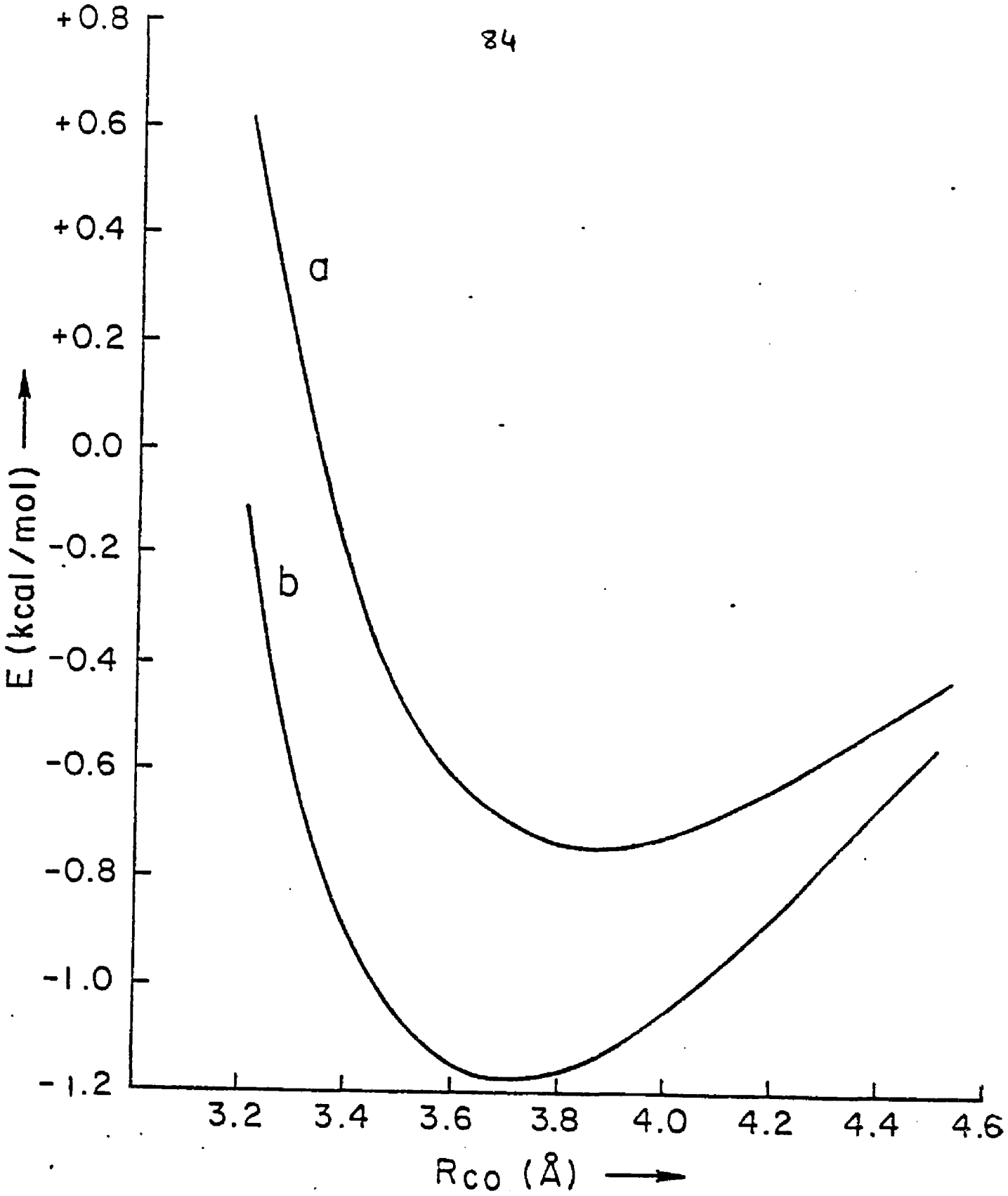


Figure 24. A comparison of the MO (a) and the MP (b) energies as a function of distance for geometry A of Ref 63.

19-22

water-water pairwise interaction. The pairwise interaction between methane and water is expected to be relatively weak (ca. 1-2 kT) and is also expected to be significantly dependent upon the correlation. For the purposes of the present study, the second order Møller-Plesset[64] (MP) correlation energy has been calculated for the methane-water interaction of each of 95 lowest energy points involved in the determination of the STO 6-31G MO function described in the previous paragraph, and developed an analytical potential function representative of MP calculations. The MP function is given in Table 6. The nature of the methane-water interaction as described by the MP function is shown in Fig. 23. The standard deviation of the function is σ_{fit} = 0.13 kcal/mole for all points below 2 kcal/mole. A comparison of the MO and MP energies for a slice of the methane-water potential energy hypersurface is given in Fig. 24. Electron correlation effects are seen to reduce the binding energy by %0.5 kcal/mole in low energy regions and decrease the equilibrium methane-water separation by $\simeq$0.15 Å.

e) Summary and Conclusions

A heuristic method has been described for the determination of intermolecular potential functions from quantum mechanical calculations and applied to the formaldehyde-water and methane-water systems. The stability of the statistical indices of the function and the reasonable computed energy grids indicate the procedure works satisfactorily and within acceptable tolerances. The fact that

that satisfactory numerical results were obtained using relatively small data base suggests that the heuristic potential function approach can be a viable means of generating intermolecular potential functions from quantum mechanical calculations.

V DILUTE AQUEOUS SOLUTION OF METHANE

a) Introduction

The dilute aqueous solution of methane is a system of prominent interest in molecular liquids as the prototype of a non-polar molecule dissolved in liquid water. Moreover, a detailed knowledge of the structure of the methane-water solution at the molecular level can provide leading information on the interaction of water with dissolved hydrocarbon chains in general, thereby contributing to the theoretical basis for understanding the role of water in maintaining the three- dimensional integrity of biological macromolecules in solution.

A set of new theoretical studies of the methane-water intermolecular interaction and Monte-Carlo computer simulations of the dilute solution of methane under canonical ensemble conditions with temperature T, volume V and number of molecules N specified and constant are reported in this chapter. All simulations are based on pairwise potential functions representative of ab initio quantum mechanical calculations of the intermolecular interactions.

The analysis of solution structure is developed in terms of quasicomponent distribution functions in a manner consistent with the analysis of liquid water structure discussed in chapter 3. Perturbation of water structure by the dissolved methane is developed in terms of difference

quasicomponent distribution functions.

b) Background

General background on non-polar solutions in water has been recently reviewed by Frank[47] and Ben-Naim.[24] Early important work on this system is due to Eley[65] and Frank and Evans.[66] Methane has been identified as a "structure maker" in aqueous solution in the language of Frank and Wen.[34] The nature of structural changes in solvent water by dissolved hydrocarbons has for sometime been discussed in terms of clathrate formation based on Work by Glew[67] and analogies drawn from a number of hydrate crystal structures of non-polar species, known to involve water clathrate cages of order 20 and 24.

Early computer simulation of the methane-water system were reported by Dashevsky and Sarkisov.[68] Recent important theoretical studies of the methane-water system are the ab initio molecular orbital calculations of the methane-water pairwise interaction energy by Ungemach and Schaefer[63] and the Monte Carlo computer simulation on the dilute aqueous solution in the isothermal-isobaric ensemble by Owicki and Scheraga(OS).[65] The OS simulation was based on the water-water CI potential discussed in the chapter on liquid water and a potential function representative of the Ungemach-Schaefer calculations for the methane-water interaction. Due to the limitations in the potential functions, the density of the system in these calculations turned out to be somewhat lower than that observed at 25°C; still the analysis of the results gave the best quantitative theoretical evidence of the

structuration of vicinal water in the solution to date and the calculated average methane-water coordination number of 23 is consistent with water clathrate contributions to the solution sturcture.

Theoretical studies of the methane-water solution require an accurate description of the structure of the pure solvent, liquid water, as a point of departure. The structure of liquid water has been discussed in chapter III. This chapter describes the extension of the (T,V,N) ensemble studies on liquid water to the dilute aqueous solution of methane, maintaining the use of potential functions representative of quantum mechanical calculations of the pairwise interaction energies. The analysis of solution structure involves the calculated distribution of coordination numbers and binding energies in the statistical state of the solution and identifies explicitly the supermolecular structures with high statistical weights. Additional questions taken up beyond previous theoretical studies include electron correlation effects of the methane-water potential on solution structure, methane excluded volume effects on solution structure, and quantitative considerations of solute perturbations on solvent structure as quantitatively described by difference quasicomponent distribution functions.

c) Calculations

The diffusionally averaged equilibrium structure of the dilute aqueous solution of methane in water using the Metropolis method for statistical thermodynamic Monte Carlo

computer simulation was studied. The specific formulation of the problem and notation relevant to this study are discussed in the second chapter. The Monte Carlo calculations involve one methane molecule and 124 water molecules at 25°C and 1 gm/cm^3; this is within 1% of the density computed from the observed partial molar volumes of methane and water. The condensed phase environment is modeled by conventional periodic boundary conditions with a spherical cutoff for the potential equal to half the box size. Standard deviations on each of the quantities calculated are determined using control functions in the manner set forth by Wood.[23]

The configurational energy of the system is developed under the assumption of pairwise additivity of intermolecular interactions using potential functions representative of ab initio quantum mechanical calculations of the water-water and methane-water interaction energies. The CI potential described in the third chapter is used for the water-water interaction. For the methane-water interaction energy, the potential functions are based on the calculations carried out in the previous chapter.

Monte Carlo computer simulations described herein are based on each of the three methane-water potential functions; The MP function, the MO function and a related function to be called "hard methane"(HM), obtained by setting the attractive part of the MO function to be zero in the manner of Barker and Henderson. A comparison of the results of the simulations based on the MP and MO potentials gives an idea of the effect

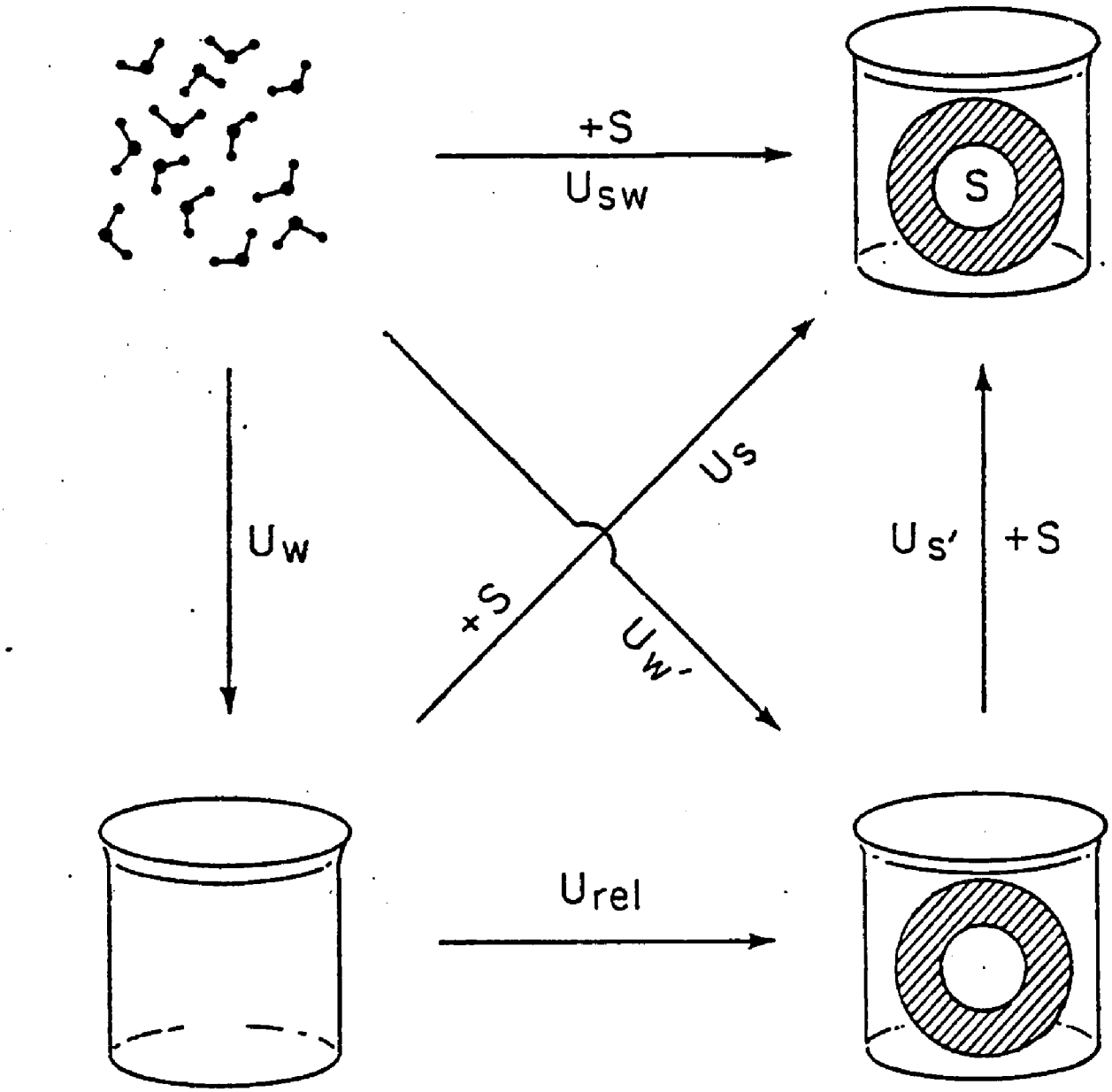


Figure 25. Thermocycle illustrating the energetic quantities produced in a Monte Carlo computer simulation of a solute, S, in water, W.

of electron correlation. The structure of the solution is expected to be largely due to methane excluded volume effects, and a comparison of simulation results based on MP and MO potentials with those obtained from the HM function permits a quantitative study of this particular feature of the system.

The principle point of comparison between calculation and experiment in this work are formally defined in Fig. 25 and the equation below. Following Ben-Naim[24] we have

$$u_{sw}(T,V,N_w,N_s) = u_w(T,V,N_w) + \left(\frac{\partial U}{\partial N_s}\right)_{T,V,N_w,} \cdot N_s + \cdots \quad (1)$$

where U_{sw} is the total internal energy of N_w water molecules and N_s solute molecules at temperature T and volume V and U_w is the total internal energy of N_w water molecules. In the addition of one solute molecule to water, $N_s = 1$ and the internal energy of transfer, neglecting higher order terms, is

$$\bar{u}_s = \left(\frac{\partial U}{\partial N_s}\right)_{T,V,N_w} = u_{sw}(T,V,N_w,N_{s=1}) - u_w(T,V,N_w) \quad (2)$$

The quantities U_{sw} and U_w are produced directly in computer simulations of the solution and pure solvent respectively as the configurational averages

$$u_{sw}(T,V,N_w,N_s) = \int\cdots\int E(\underset{\sim}{X}^s,\underset{\sim}{X}^w)\,P(\underset{\sim}{X}^s,\underset{\sim}{X}^w)\,d\underset{\sim}{X}^s\,d\underset{\sim}{X}^w \quad (3)$$

and

$$\mathcal{U}_w(T, V, N_w) = \int \cdots \int E(\underset{\sim}{X}^w)\, P(\underset{\sim}{X}^w)\, d\underset{\sim}{X}^w \quad (4)$$

where $E(\underset{\sim}{X}^s, \underset{\sim}{X}^w)$ is the configurational energy of the system with solute specified by the configurational coordinates $\underset{\sim}{X}^s$ and the N_w water molecules specified by the configurational coordinates $\underset{\sim}{X}^w$. The quantity $P(\underset{\sim}{X}^s, \underset{\sim}{X}^w)$ is the probability of observing the system in configuration $(\underset{\sim}{X}^s, \underset{\sim}{X}^w)$. The quantities $E(\underset{\sim}{X}^w)$ and $P(\underset{\sim}{X}^w)$ are analogous terms for pure water. Thermodynamic quantities expressed on a per particle basis are denoted by a bar, as in $\overline{U}_s$

Reference to Fig.25 shows an alternative formulation of $\overline{U}_s$ to be

$$\overline{\mathcal{U}}_s = \overline{\mathcal{U}}_{s'} + \overline{\mathcal{U}}_{rel} \quad (5)$$

The definition of the quantities on the right hand side of equation 5 follows from partitioning the configurational energy according to

$$E(\underset{\sim}{X}^s, \underset{\sim}{X}^w) = E_{sw}(\underset{\sim}{X}^s, \underset{\sim}{X}^w) + E_{ww}(\underset{\sim}{X}^w) \quad (6)$$

where E_{sw} and E_{ww} are solute-water and water-water contributions respectively, defined as

$$E_{sw}\left(\underset{\sim}{X}^{s}, \underset{\sim}{X}^{w}\right) = \sum_{i}^{Nw} E_i \left(\underset{\sim}{X}^{s}, \underset{\sim}{X}_i^{w}\right) \qquad (7)$$

and

$$E_{ww}\left(\underset{\sim}{X}^{w}\right) = \sum_{i<j}^{Nw} E_{ij}\left(X_i^{w}, X_j^{w}\right) \qquad (8)$$

where E_i and E_{ij} are the solute-water and water-water pairwise interaction energies. The transfer quantity $\overline{U}_s$ is then defined as

$$\overline{U}_{s'} = \int\cdots\int E_{sw}\left(\underset{\sim}{X}^{s}, \underset{\sim}{X}^{w}\right) P\left(\underset{\sim}{X}^{s}, \underset{\sim}{X}^{w}\right) d X^{s}\, d X^{w} \qquad (9)$$

and represents the direct solute-solvent contribution to the partial molar internal energy of transfer. The quantity can be written in terms of the QCDF for solute binding energy

$$\overline{U}_{s'} = \int_{-\infty}^{+\infty} \gamma\, x_B(\gamma)\, d\gamma \qquad (10)$$

where $x_B(\gamma)$ is the mole fraction of methane molecules with the binding energy γ. The second term on the right hand side of equation 5 represents a contribution from solvent reorganization on solute formation, and can be expressed in terms of a difference of a pure water U_w and solvent water term U_w'

$$\overline{U}_{rel} = U_w - U_{w'} \qquad (11)$$

where

$$U_{w'} = \int\cdots\int E_{ww}\left(\underset{\sim}{X}^{w}\right) P\left(\underset{\sim}{X}^{s}, \underset{\sim}{X}^{w}\right) d X^{s}\, d X^{w} \qquad (12)$$

Table 7

Calculated internal energies for the dilute aqueous solution of methane at 25 C based on the MP, MO and HM functions; in kcal/mol.

	MP	MO	HM
$U_{SW}(N_W=124, N_S=1)$	-1077.7±3.6	-1076.2±3.6	-1070.1±3.6
$U_{W'}(N_W=124)$	-1079.4±3.6	-1077.9±3.6	-1075.3±3.6
$U_W(N_W=124)$	-1054.4±5.5	-1054.4±5.5	-1054.4±5.5
$\bar{U}_{rel}$	-25.0±6.6	-23.5±6.6	-20.9±6.6
$\bar{U}_{S'}$	1.71±.29	1.73±.27	5.19±.27
$\bar{U}_S$	-23.3±6.6	-21.8±6.6	-15.7±6.6

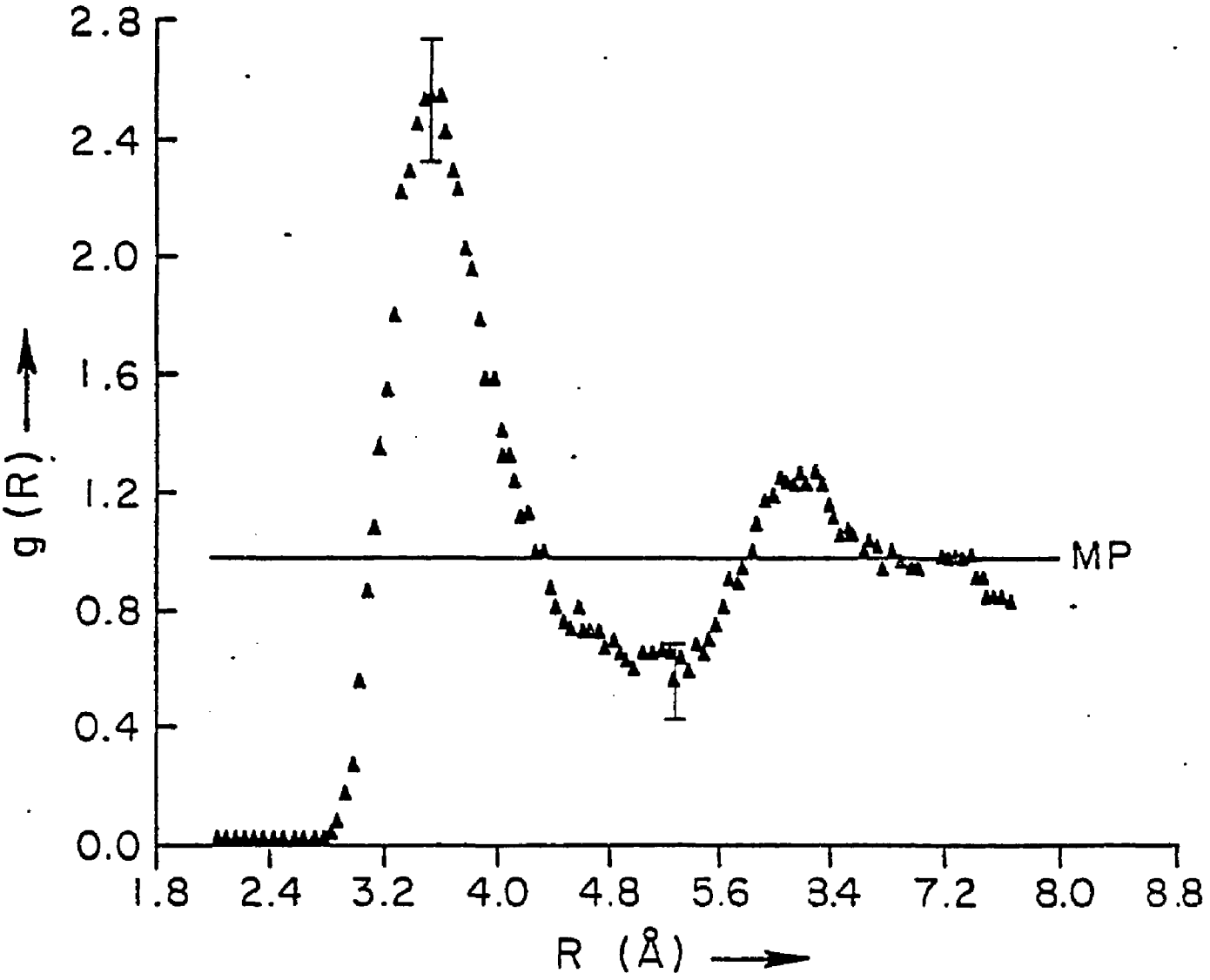


Figure 26. Calculated methane-water radial distribution function g(R) vs. center of mass separation R from Monte Carlo computer simulation based on the MP function.

Note $\bar{U}_{sol}$ is defined per solute particle.

d) Results

The simulation using the MP potential is based on a 750K stochastic walk with the first 100K discarded in forming the ensemble averages. Convergence criteria and error bounds were determined from control functions taken at 25K intervals in the calculations. The number of discards were determined from minimum error bounds. The partial molar internal energy of methane is -23.3±6.6 kcal/mol, compared with an experimental value of -2.6 kcal/mol.[70] The complete set of calculated energetic quantities for all simulations are reported in Table 7.

The calculated radial distribution function for the center of mass of water molecules with respect to the center of mass of the methane molecule for the MP function is shown in Fig.26. A broad unstructured first peak and a minimum in the region of 5.3 Å are found. Integrating the g(R) up to this point yields an average coordination number of 19.35. Experimental value for these quantities are not known, but taking into account differences in the (T,V,N) and (T,P,N) ensembles these are in satisfactory accord with the results of OS.

The simulation involving the MO potential function was based on 575K steps with 150K discarded. The calculated partial molar internal energy was found to be -21.8+6.6

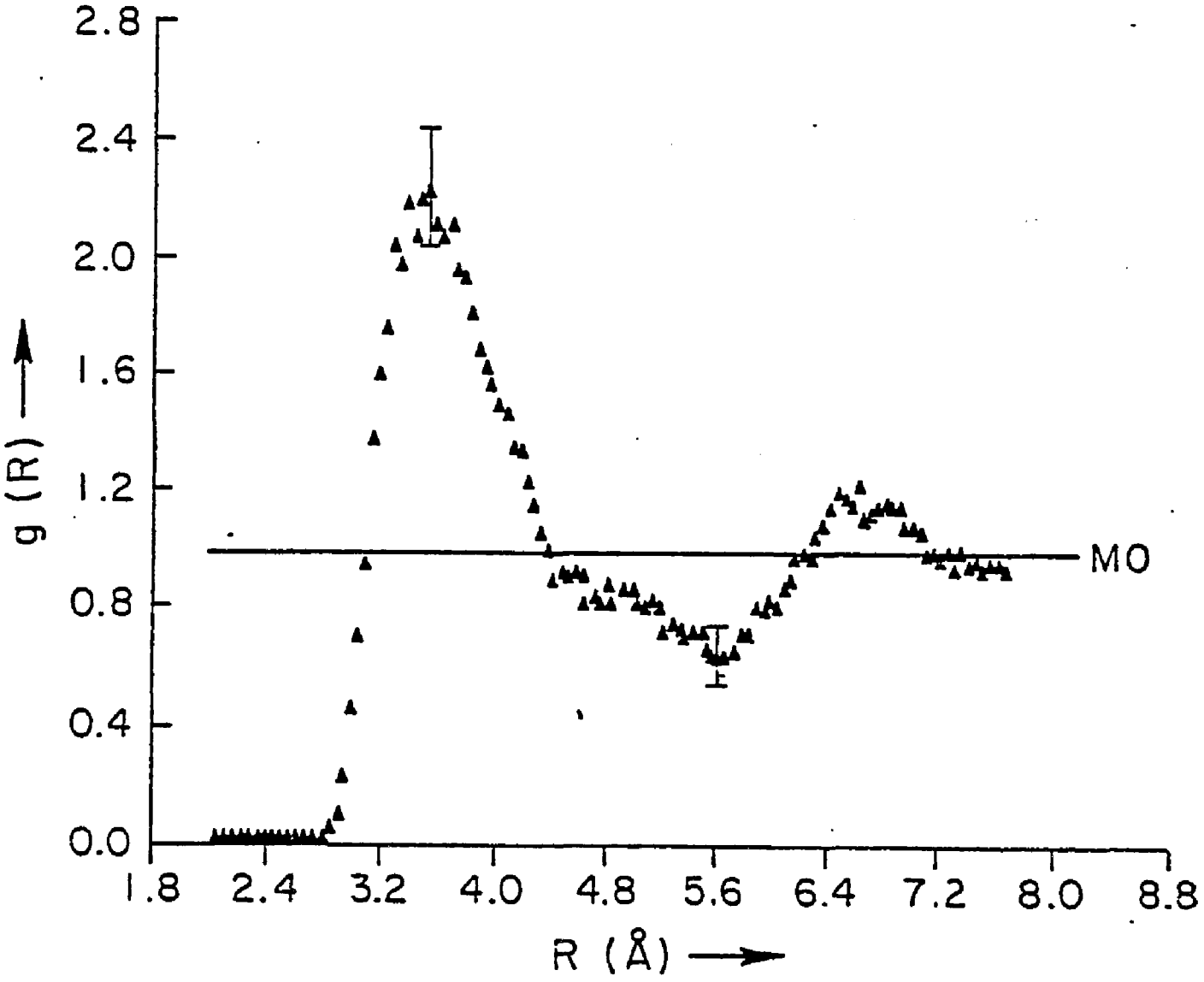


Figure 27. Calculated methane-water radial distribution function g(R) vs. center of mass separation R from Monte Carlo computer simulation based on the MO function.

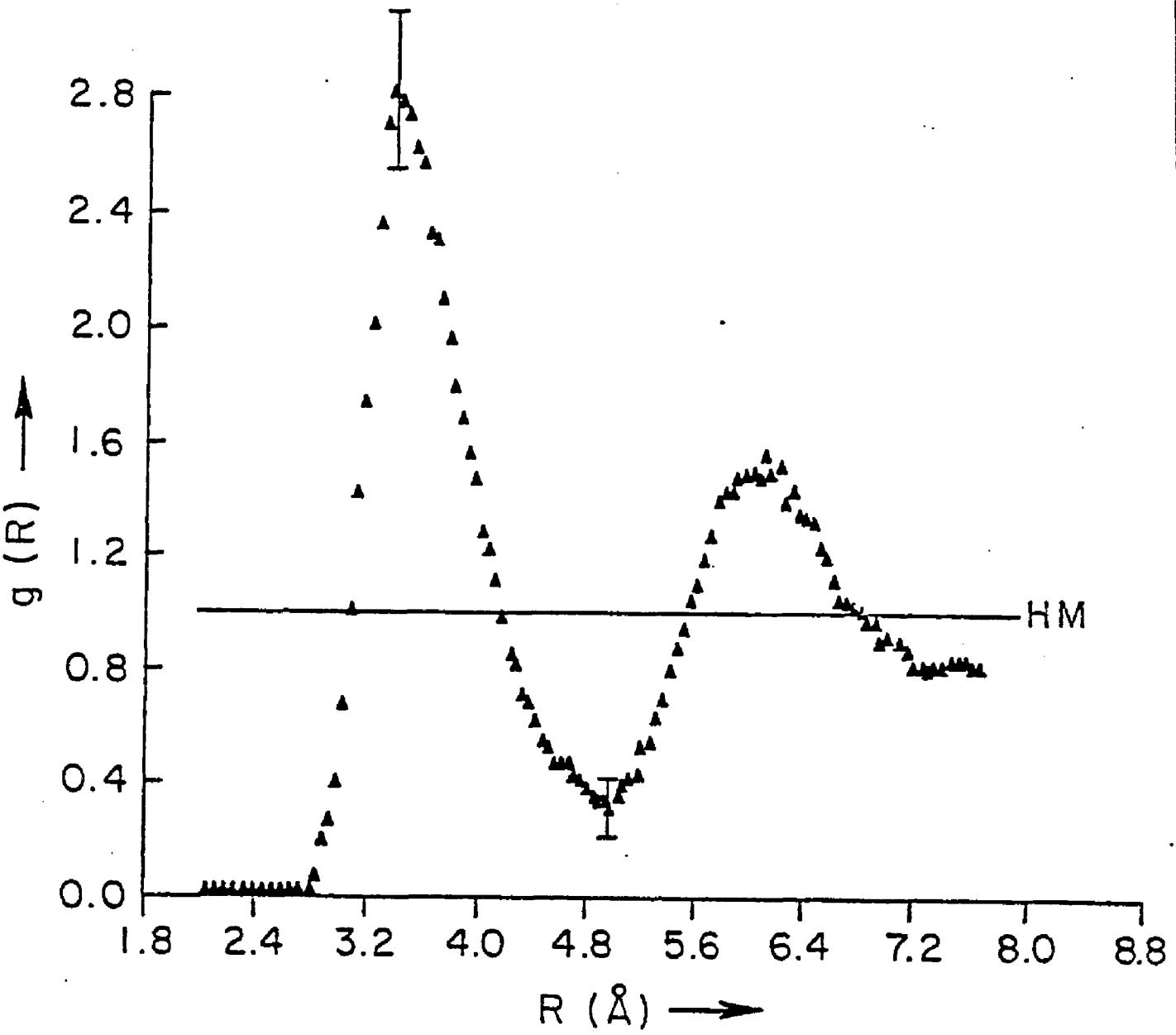


Figure 28. Calculated methane-water radial distribution function g(R) vs. center of mass separation R from Monte Carlo computer simulation based on the HM function.

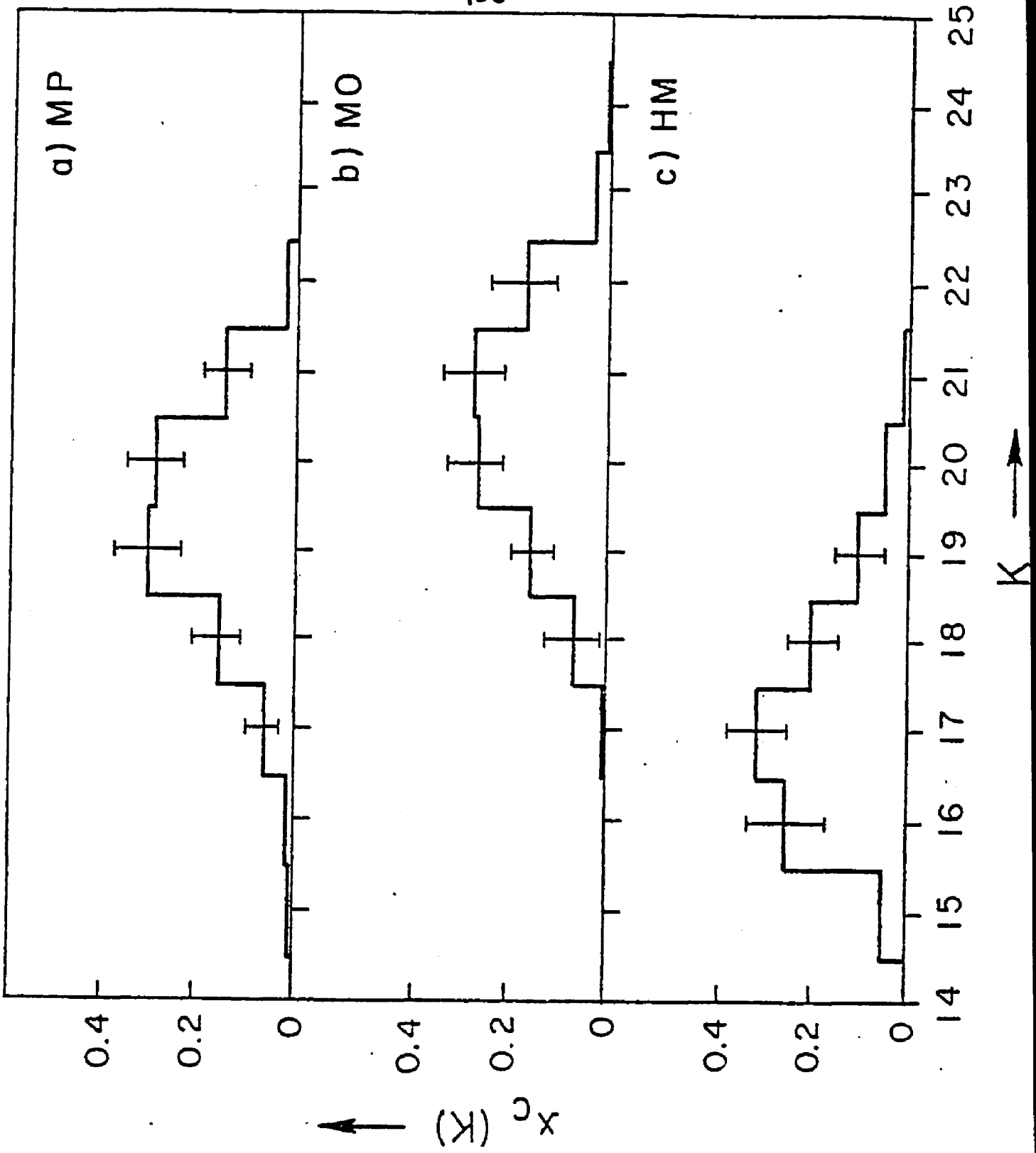


Figure 29. Calculated quasicomponent distribution function x_c (K) vs. coordination number K of methane in simulation based on the a) MP potential, b) MO potential and c) HM potential.

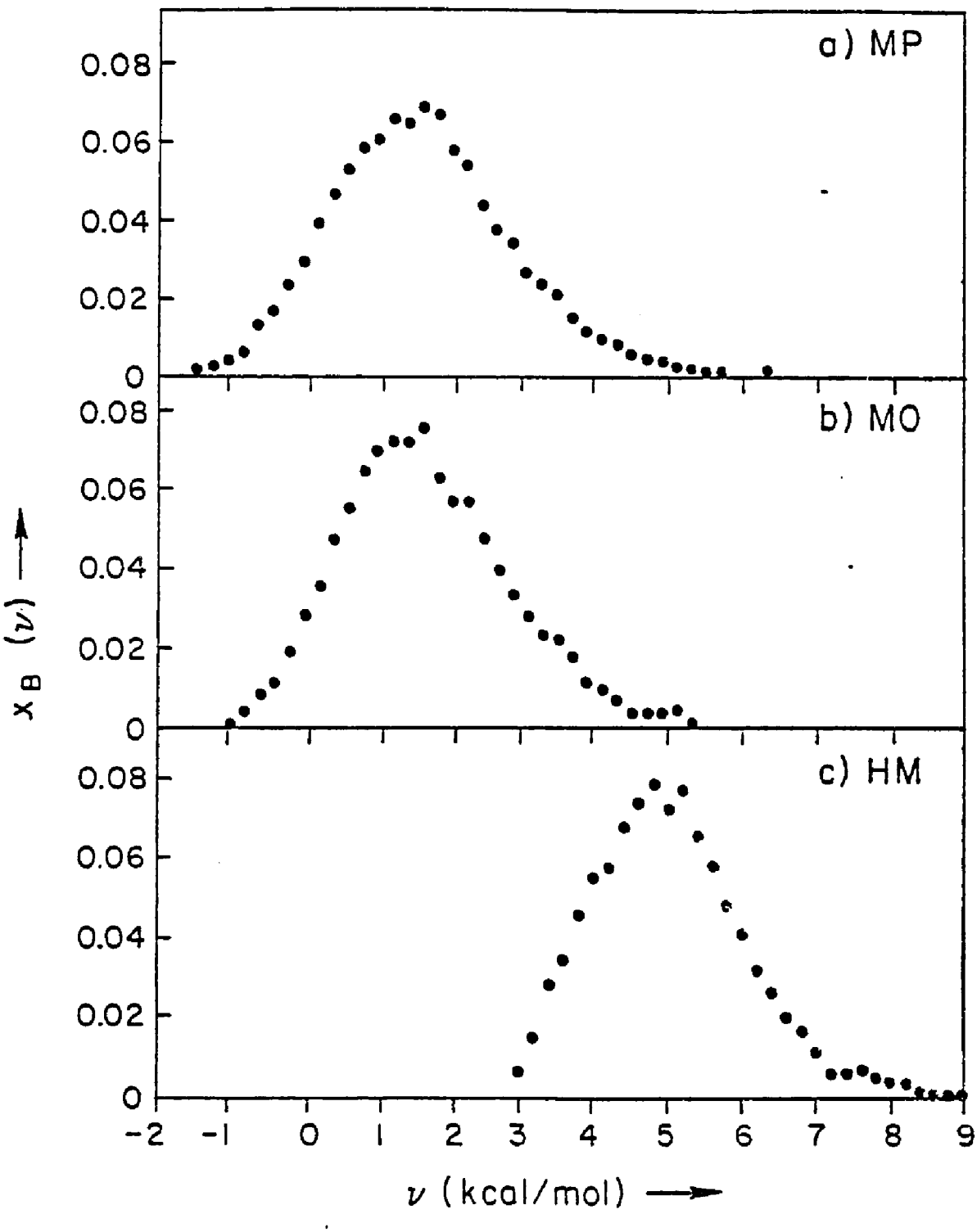


Figure 30. Calculated quasicomponent distribution function x_B (ν) vs. binding energy ν of methane in simulation based on the a) MP potential, b) MO potential and c) HM potential.

kcal/mol. The calculated methane-water radial distribution function is given in Fig.27.

The simulation on hard methane involved 900K steps with 275K discarded. The calculated partial molar internal energy was -15.7 ± 6.6 kcal/mol. The radial distribution function is given in Fig.28.

A theoretical analysis of the structure of the dilute aqueous solution of methane can be developed in terms of quasicomponent distribution functions for coordination number and binding energies as described by Ben-Naim in Reference 24. The distribution of coordination numbers found within the first hydration shell(R_c=5.3Å) of methane in the statistical state of the solution is shown for all three simulations in Fig.29. Here a plot of the mole fraction of particles $x_c(K)$ vs. coordination number K in histogram form is presented. The $x_c(K)$ from the MP is a broad unimodal distribution ranging from K=16 to K=22 with a maximum in the region of K=19 and K=20 biased slightly towards higher coordination numbers. The $x_c(K)$ for the MO simulation is slightly more symmetric, and the maximum is displaced to the region of K=17 and K=18. The HM simulation produces a slightly broader distribution of coordination numbers, unimodal with a maximum at K=17 and clearly biased towards higher K.

The calculated quasicomponent distribution functions for binding energies, the mole fraction of particles $x_B(\gamma)$ as a function of methane binding energy, is shown for all three simulations in Fig.30. All $x_B(\gamma)$ distributions are of a

similar shape but displaced toward higher energies on going from the results of MP simulation to those obtained from HM potential. There is some incipient structure in the curve but error bounds on the calculated values are too large to ascribe any significance.

d) Discussion

This section deals with the following problems: a) the comparison of calculated and observed partial molar internal energies for methane in aqueous solution, b) the microscopic nature of the solution environment of the solute methane, c) electron correlation effects on solution structure, d) the role of the soft part of the methane water potential on solution structure and e) structuration effects in solvent water due to the solute methane.

Comparing the calculated and observed partial molar internal energies for methane in water, the calculated values from each of the three simulations are seen to be negative but rather too low. The negative sign is an essential feature of the hydrophobic effect as currently understood and has been ascribed to solvent water stabilisation effects. The partitioning of $\overline{U}_s$ into solute effect $\overline{U}_{s'}$ and a solvent relaxation contribution $\overline{U}_{rel}$ is displayed in Table 7. The $\overline{U}_{s'}$ contribution is positive in all simulations and $\overline{U}_{rel}$ is negative, consistent with the current ideas on the hydrophobic solvation. The solvent stabilisation term is somewhat overestimated, however, and leads to a factor of 10 discrepancy between the calculated and observed values of $\overline{U}_s$,

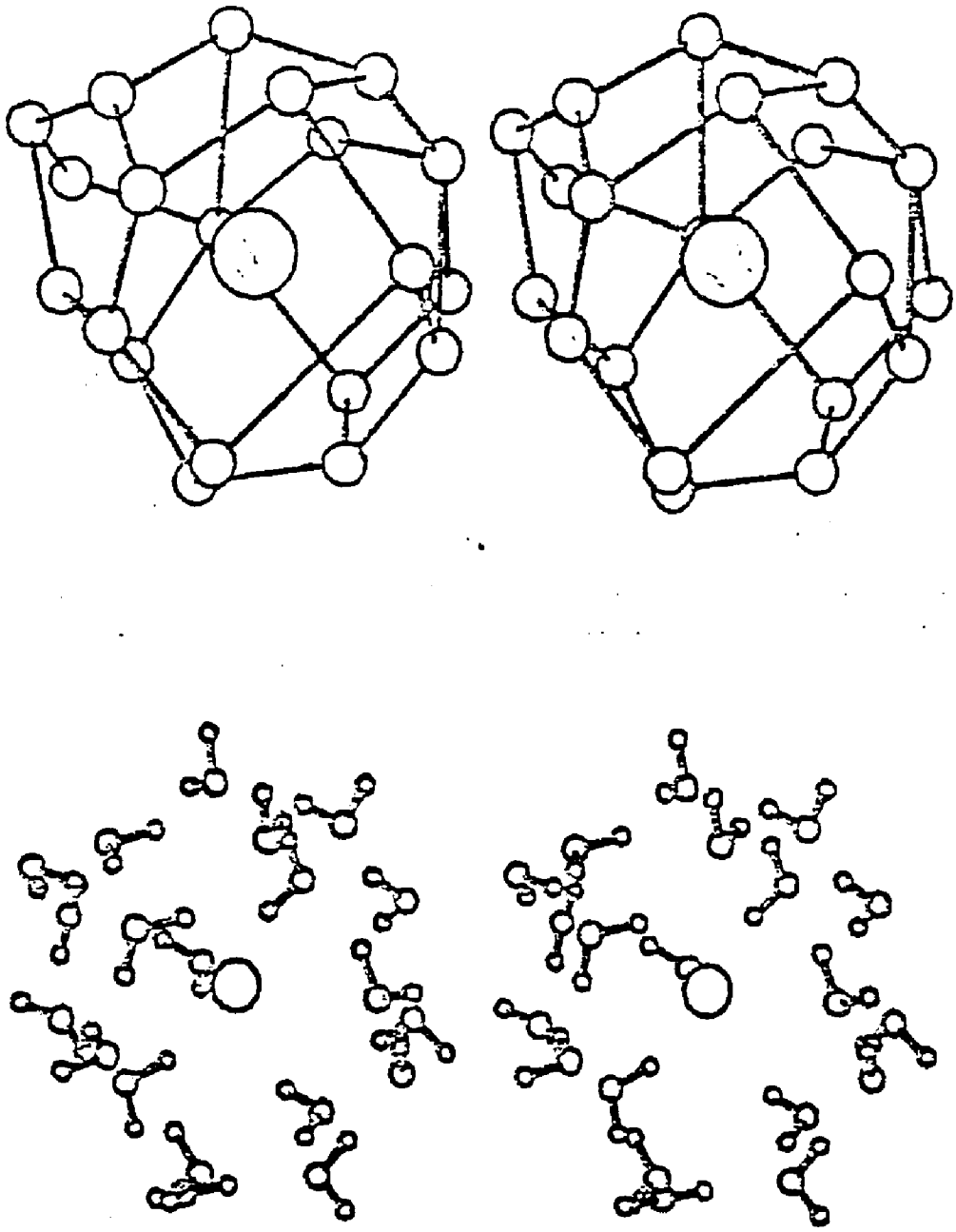


Figure 31. Stereographic view of methane and its first hydration shell for one of the structures with a high statistical weight in $x_c(20)$ of the MP simulation. a) disposition of centers of mass of water molecules about methane with the quasiclathrate delineated; b) disposition of water molecules about methane in the structure. Relative sizes of atoms scaled down and methane depicted as a sphere for greater legibility of the pictures.

a difference of the order of 20 kcal/mol. If we assume this stabilisation involves mainly those 20 water molecules found in the first hydration shell of methane, this is an error of 1 kcal/mol per water molecule or a fraction of a kcal/mol per pairwise interaction or hydrogen bond. The major assumptions inherent in the calculations are the neglect of three-body and higher order contributions to the configurational energy, truncation errors in the quantum mechanical calculations of the pairwise interaction energies, and statistical errors in the analytical potential function. The discrepancy between calcuated and observed values is thus reasonable in perspective of the capabilities and limitations of the configurational energy evaluation.

The distribution of coordination numbers for methane in the statistical state of the solution is consistent with water clathrate contributions. The values are closer to those expected for a pentagonal dodecahedaral cage whereas the the corresponding results obtained for the (T,P,N) simulation of OS were closer to those expected for tetrakaidechedral cage. In order to investigate futher the microscopic nature of the local solution environment of methane, a number of configurations of high statistical weight were extracted from the MP simulation and stereographic ORTEP plots of methane andits first hydration shell were made. The structure which closely resembled a pentagonal dodecahedral clathrate cage is shown in Fig.31. The resemblance seems to be a direct confirmation of clathrate contribution to solution structure.

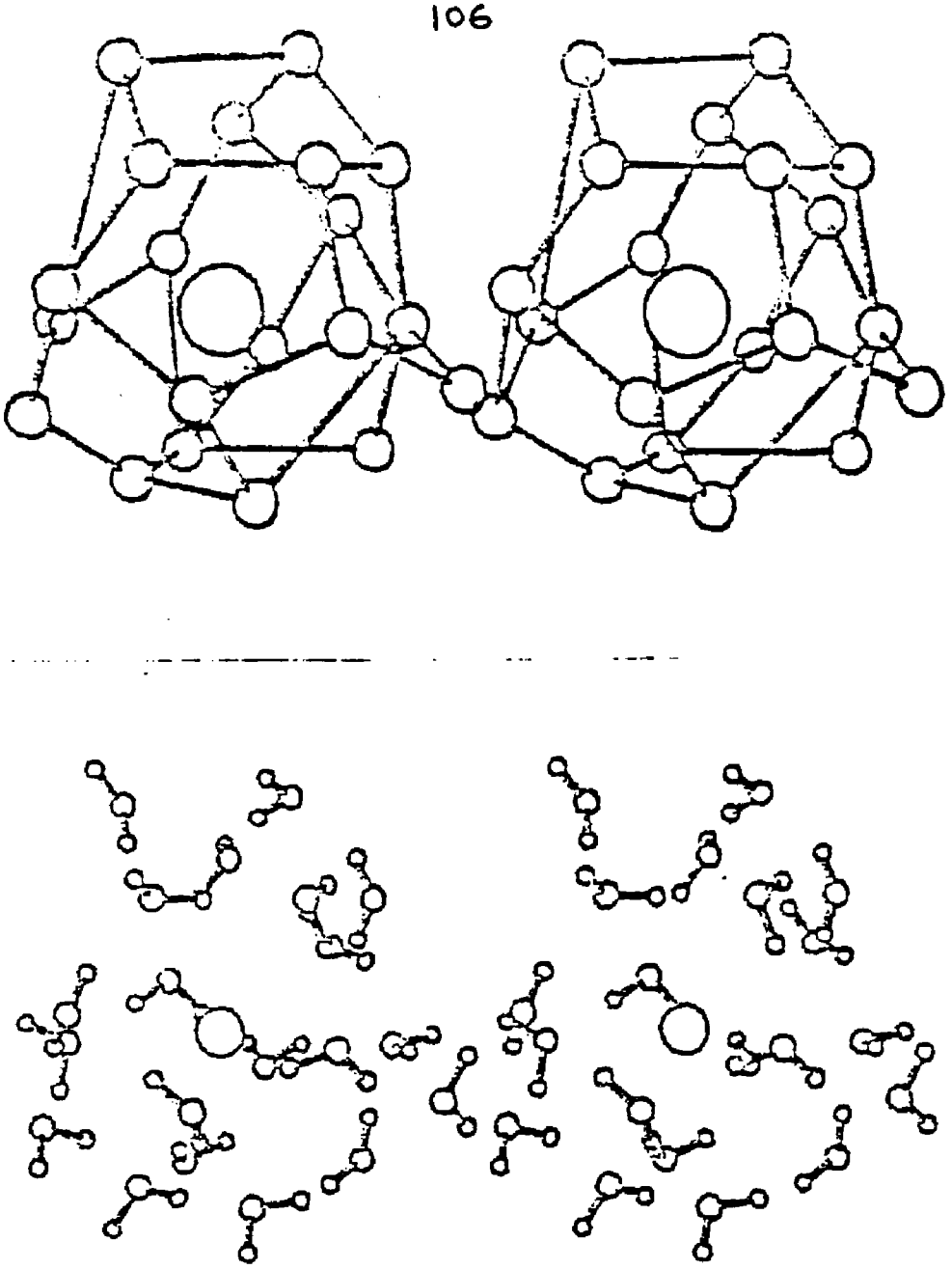


Figure 32. Stereographic view of methane and its first hydration shell for another structure with a high statistical weight in $x_c(20)$ of the MP simulation. a) disposition of centers of mass of water molecules about methane with the quasiclathrate delineated; b) disposition of water molecules about methane in the structure. Relative sizes of atoms scaled down and methane depicted as a sphere for greater legibility of the pictures.

In a number of other structures the clathrate-like regions of the first hydration shell could be found but defects and distortions were more prevalent than the structure in Fig.31. A structure that is representative of the sample set is presented in Fig.32. The quasicomponent distribution functions for binding energy, $x_B(\gamma)$ collected in Fig.35. show that the solute methane is presented with a continuous distribution of energetic environments in the solution. The emergent description of the dilute aqueous solution of methane is that of a slightly distorted and defective continuum clathrate structure.

A comparison of the computer simulation results based on the MP and MO potentials for the methane-water interaction gives an idea of the effects of electron correlation effects on the system. This may be taken only as leading information, since there are still deficiencies in the quality of the quantum mechanical data base due to basis set truncation effects of the electron correlation and the nature of the Moller-Plesset electron correlation. The calculated g(R) for the MP potential in Fig.26. and the MO potential in Fig.27. are similar in overall appearence with only minor visible discrepancies. The analysis of structure in terms of $x_C(K)$ and $x_B(\gamma)$ shows that electron correlation discernably leads to higher coordination number and lower binding energies. The displacements are of the order of 5-10% for both quantities.

Further comparison of the results with reference to the g(R), $x_C(K)$ and $x_B(\gamma)$ obtained on the computer simulation

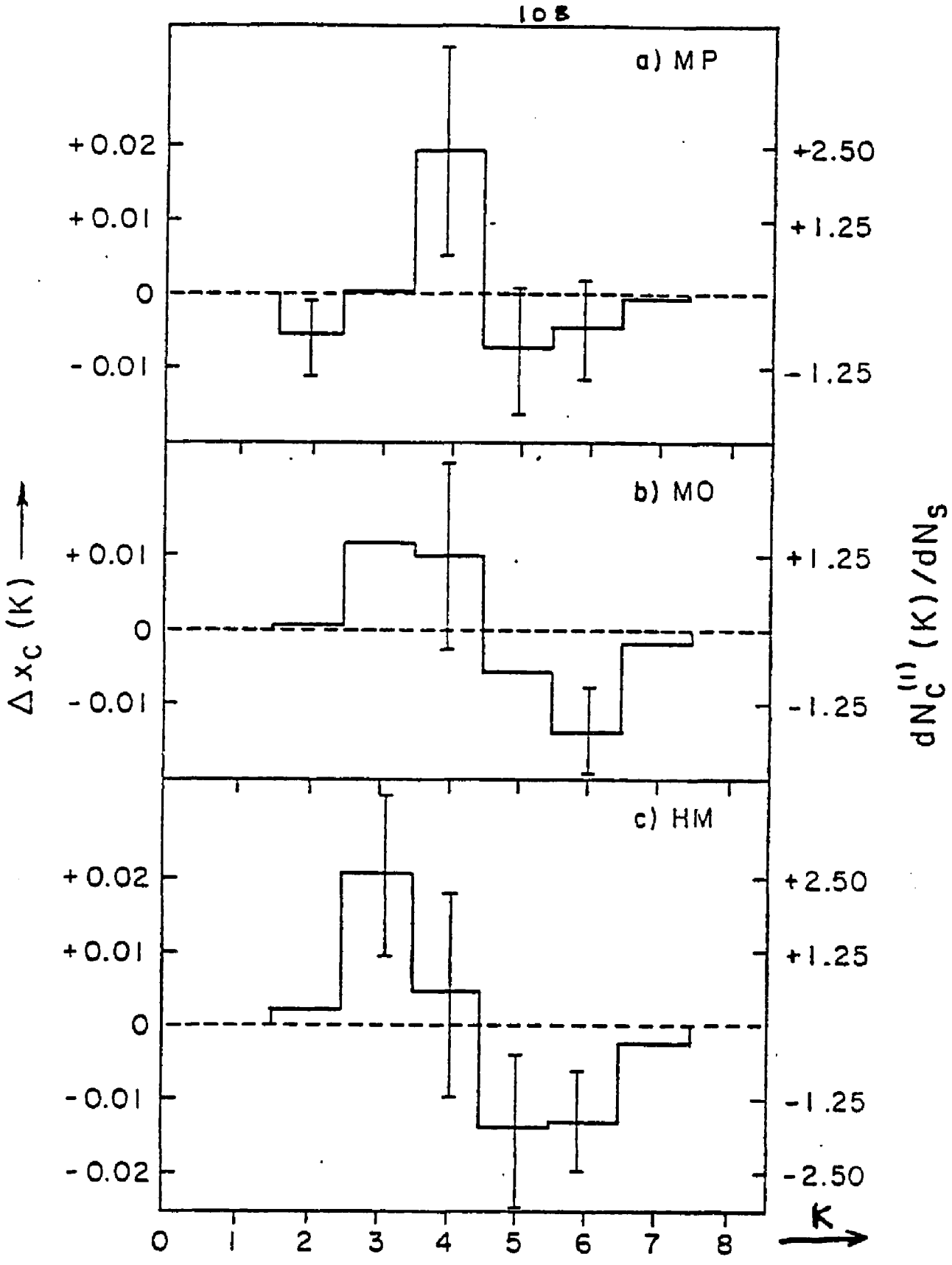


Figure 33. Caculated difference quasicomponent distribution function $\Delta x_c(K)$ for water in the simulations based on the a) MP potential, b) MO potential and c) HM potential.

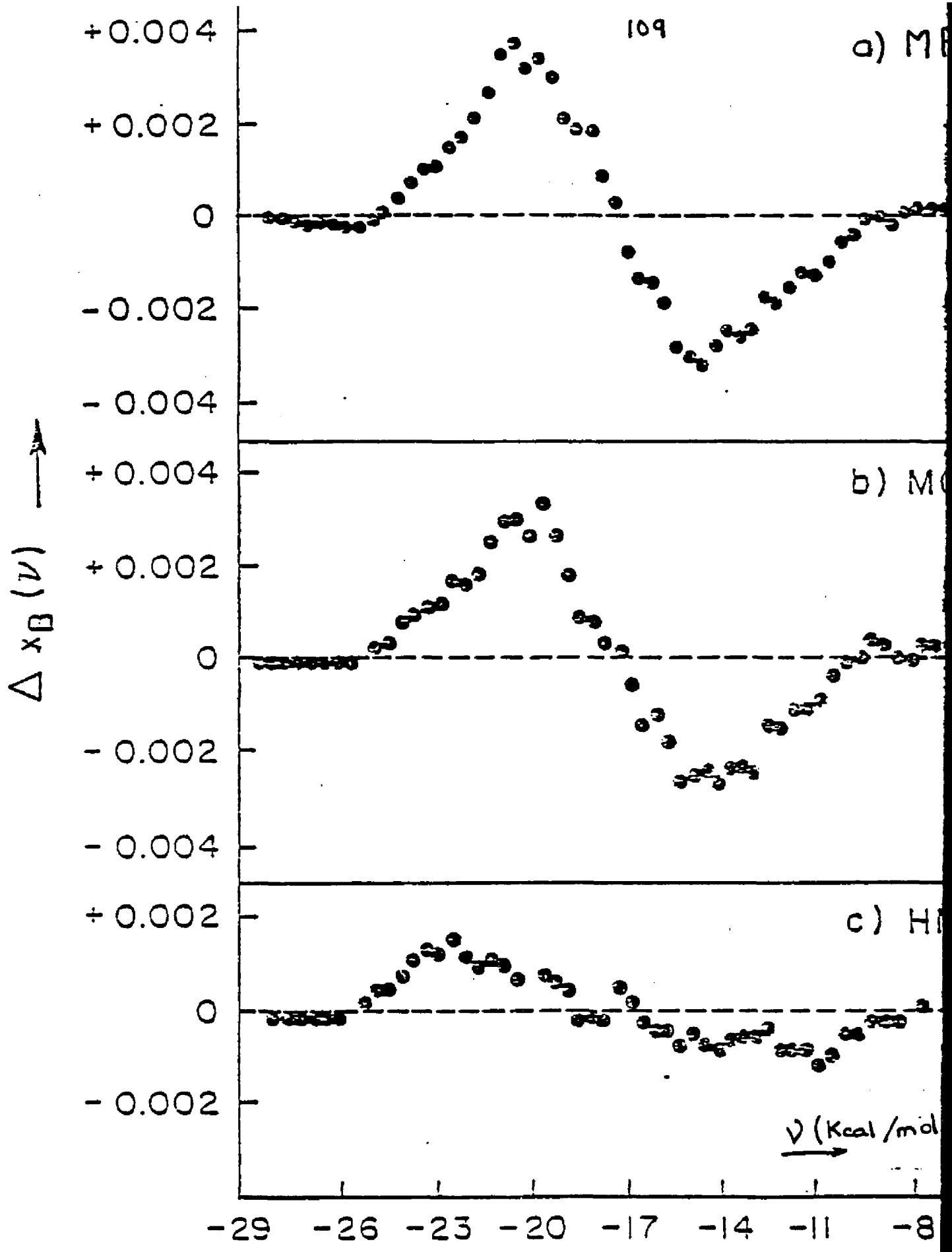
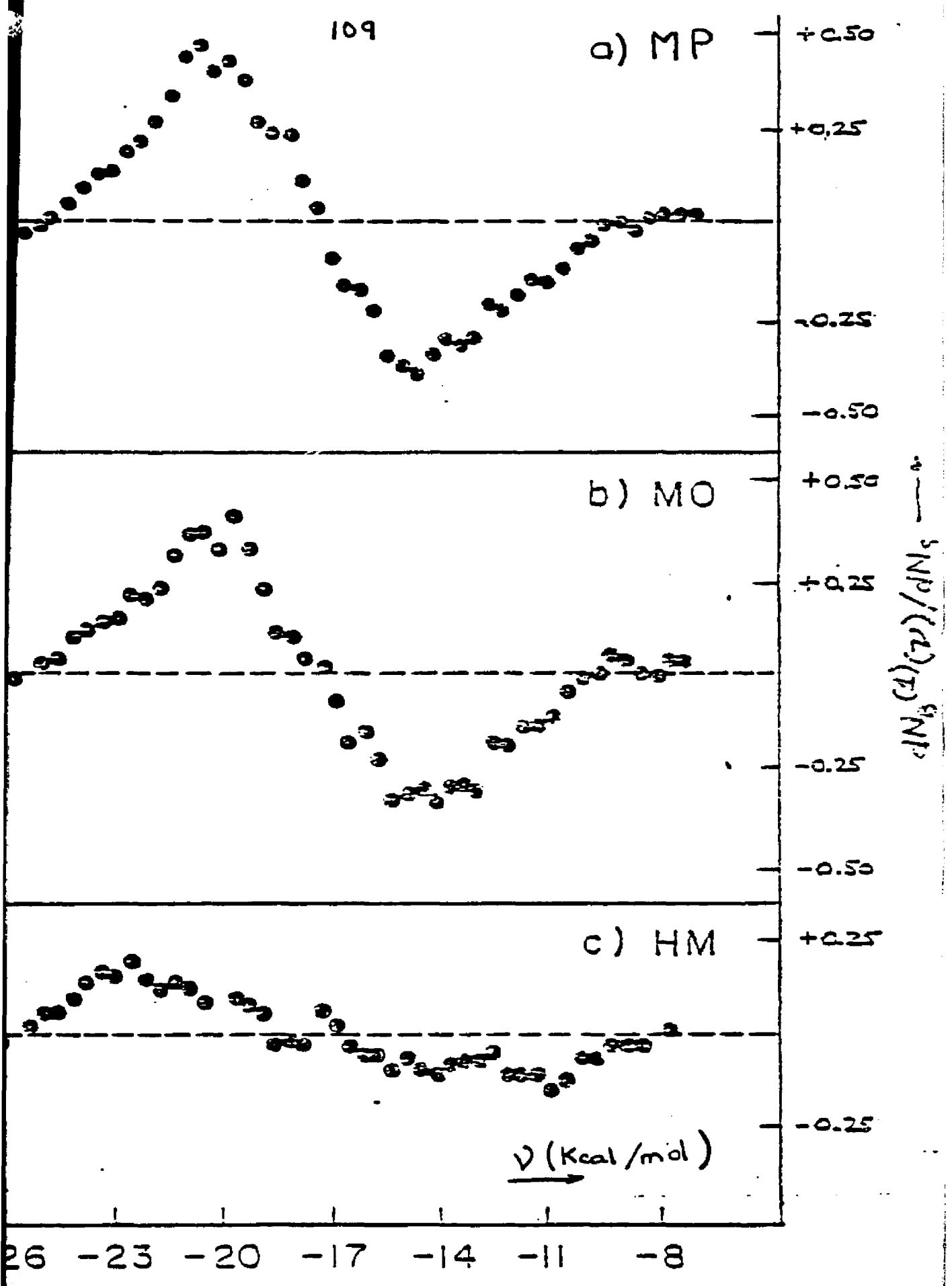


Figure 34. Caculated difference quasicomponent distri
function $\Delta x_\beta(\nu)$ for water in the simulations based on
MP potential, b) MO potential and c) HM potential.



culated difference quasicomponent distribution (ν) for water in the simulations based on the a) b) MO potential and c) HM potential.

using HM potential provides information the role of the soft part of the methane-water potential function on solution structure. First, the similarity in the calculated radial distribution functions in Figs. 26-28 shows the essential nature of the methane-water solution as an excluded volume effect as expected. The analysis of structure shows that the soft part of the potential function shifts the distribution of coordination numbers upwards by 15% and shifts the distribution of binding energies down by 3.2 kcal/mol.

A consideration of the effects of solute methane on water structure is now in order. It should be noted that the OS paper identified structuration effects in the solvent by studying explicitly the waters vicinal to the solute by a method involving partitioning of the configuration space. A similar analysis is being employed in a molecular dynamics study of a solution of a tripeptide in water by Rossky and Karplus.[71] The representation of the phenomenon is developed here in terms of difference quasicomponent distribution functions (QCDF). The idea is to describe the changes in water structure in terms of difference between $x_c(K)$ and $x_b(V)$ calculated for the solution and for the pure liquid. The effects of bulk water are removed by differencing allowing the direct display of structural changes in solution water without recourse to arbitrary partitioning of the configuration space.

The $\Delta x_c(K)$ and $\Delta x_b(V)$ for solvent water in the dilute aqueous solution of methane with respect to bulk water described in Chapter II are given in Figs. 33 and 34. In the

distribution of $x_c(K)$ there is a clear increase in the four coordinate species at the expense of coordination numbers 2,5 and 6. These results may be considered in the perspective of the error bounds on the Δ QCDF, but the increase in K=4 is definitely statistically significant. The $\Delta x_B(\gamma)$ shows slight but clearly discernable shifts toward stronger binding water molecules. The nature of the solvent structuration as seen from Figs. 31 and 32 involves mainly bent hydrogen bonds.

If one permits the identification of the increase in four coordinate species in water and increased binding with "structure making", the QCDF description of the structuration effects serves to quantify some of the longstanding terminology prevalent in the descriptive chemistry of solutions. The problem inherent in this connection are brought out in the critical review of the problem by Holzer and Emerson,[72] and the judgement on this point should be reserved until ΔQCDF for solvent water in a number of solutions of structure makers and structure breakers have been determined.

e) Summary and Conclusions

Monte Carlo computer simulation on the dilute aqueous solution of methane at 25° C using quantum mechanical potential functions reveal the local solution environment to be that of a distorted defective continuum pentagonal dodecahedral clathrate structure. Increased four coordination and stronger binding are induced in the solvent water in comparison with the bulk liquid.

REFERENCES

1. D. Eisenberg and W. Kauzman,"The structure and properties of Water", Oxford University Press (1969).
2. J.A. Barker and D. Henderson, Rev. Mod. Phys. 48, 587 (1976).
3. J. Yvon, Actualities Scientifiques et Industrielles, 203 (1935).
4. J.G. Kirkwood, J. Chem. Phys. 3, 300 (1935).
5. J.K. Percus and G.L. Yevick, Phys. Rev. 110, 1 (1958).
6. T. Morita, Prog. Theor. Phys. 20, 920 (1958).
7. R.W. Zwanzig, J. Chem. Phys. 22, 1420 (1954).
8. J.A. Barker and D. Henderson, J. Chem. Phys. 47, 4714 (1967).
9. J.D. Weeks, D.Chandler and H.C. Anderson, J. Chem. Phys. 54, 5237 (1971).
10. B.J. Alder and T.E. Wainwright, J. Chem. Phys. 27, 1208 (1957).
11. A. Rahman and F.H. Stillinger, J. Chem. Phys. 55, 3336 (1971).
12. I.R. MacDonald and M.L. Klein, J. Chem. Phys. 64, 4791 (1976).
13. B.J. Alder and W.G. Hoover,"Physics of Simple Liquids", North Holland, Amsterdam.(1968).
14. J. Kushick and B.J. Berne, "Statistical Mechanics Part B" 6, Modern Theoretical Chemistry, Plenum Press, New York.

15. J.M. Hammersly and D.D. Hanscomb, "Monte Carlo Methods", Methuen, London (1964).

16. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. 21, 1087 (1953).

17. J.A. Barker and R.O. Watts, Chem. Phys. Lett., 3, 144 (1969).

18. R.O. Watts, Mol. Phys. 28, 1069 (1974).

19. H. Popkie, H. Kistemacher and E. Clementi, J. Chem. Phys., 59, 1235 (1973).

20. G.C. Lie and E. Clementi, J. Chem. Phys., 62, 2195 (1975).

21. O. Matsuoka, E. Clementi and M. Yoshimine, J. Chem. Phys., 64, 1351 (1976).

22. G.C. Lie, E. Clementi and M. Yoshimine, J. Chem. Phys., 64, 2314 (1976).

23. W.W. Wood in "Physics of Simple Liquids", North Holland, Amsterdam. (1968).

24. A. Ben-Naim, "Water and Aqueous Solutions", Plenum Press, New York. (1976).

25. F. Kohler in "The Liquid State",Crane Russack Co. (1972).

26. D.H. Hankins, J.W. Moskowitz and F.H.Stillinger, J. Chem. Phys., 53, 4544 (1970).

27. B.R. Lentz and H.A. Scheraga, J. Chem. Phys. 58, 5296 (1973).

28. J.P. Valleau and G.M. Torrie, "Statistical Mechanics" , 6, Modern Theoretical Chemistry, Plenum Press, New York.

29. J.G. Kirkwood in "Theory of Liquids", Gordon and Preach, New York. (1968).

30. W.C. Röntgen, Ann. Phys., 45, 91 (1892).

31. J.D. Bernal and R.H. Fowler, J. Chem. Phys., 1, 515 (1933).

32. O.Ya. Samoilov, Zh. Fiz. Khim., 20, 1411 (1946).

33. V.A. Mikhailov, Zh. Strukt. Khim., 8, 189 (1967); 9, 397 (1968).

34. A.H. Narten, M.D.Danford and H.A. Levy, Disc. Faraday Soc., 43, 97 (1967).

35. L. Pauling, J. Am. Chem. Soc., 57, 2680 (1935); Science, 134, 15 (1961).

36. H.S. Frank and A.S. Quist, J. Chem. Phys. 34, 604 (1961).

37. C.M. Davis and T.A. Litovitz, J.Chem. Phys. 42, 2563 (1965).

38. V. Vand and W.A. Senior, J. Chem. Phys., 43, 1878 (1965).

39. H.S. Frank and W.Y. Wen, Disc. Faraday Soc., 24, 133 (1957).

40. A.T. Hagler, H.A. Scheraga and G.M.Nemethy, J. Phys. Chem., 76, 3229 (1972).

41. M.S. Jhon, J. Grosh, T. Ree, and H. Eyring, J. Chem. Phys., 44, 1465 (1966).

42. J.A. Pople, Proc. R. Soc. London, Ser. A, 205, 163 (1951).

43. C.M. Davis and J. Jarzynski, Adv. Molecular Relaxation Processes, 1, 155 (1967).

44. G.S. Kell in "Water and Aqueous Solutions", John Wiley, New York. (1972).

45. B. Kamb in "Structural Chemistry and Molecular Biology", W.H. Freeman, San Francisco (1968).

46. H.S. Frank in "Water-A comprehensive Treatise",Vol I, Plenum Press, New York (1972).

48. G.E. Walrafen in "Water-A comprehensive Treatise",Vol I, Plenum Press, New York (1972).

49. J.R. Scherer, J. Phys. Chem., 78, 1304 (1974).

50. A. Ben-Naim and F.H. Stillinger in "Water and Aqueous Solutions", John Wiley, New York. (1972).

51. H. Kistenmacher, H. Popkie, E. Clementi and R.O. Watts, J. Chem. Phys., 60, 4455 (1974).

52. G.H.F. Diercksen, W.P. Kraemer and B.O. Roos, Theor. Chim Acta, 36, 249 (1975).

53. G.N. Sariskov, V.G. Dashevsky and G.G. Malenkov, Mol. Phys., 27, 1271 (1974).

54. A.H. Narten, J. Chem. Phys. 56, 5681 (1972).

55. K.J. Nielsen, "Methods in Numerical Analysis", MacMillan Co., New York. (1964).

56. J.O. Hirshfelder, C.F. Curtiss and R.B. Bird, "Molecular Theory of Gases and Liquids", John Wiley, New York. (1954).

57. G.W. Schnuelle and D.L. Beveridge, J. Phys. Chem., 79, 2566, (1975).

58. M.J.D. Powell, Comp. J., 1, 303 (1965).

59. K. Morokuma, J. Chem. Phys., 55, 1327 (1971).

60. A. Johansson and P.A. Kollman, J. Am. Chem. Soc., 94, 6196, (1972).

61. J.E. Del Bene, J. Am. Chem. Soc., 95, 6517 (1973).

62. W.J. Hehre, R. Ditchfield and J.A. Pople, J. Chem. Phys., 56, 2257 (1972).

63. S.R. Ungemach and H.F. Schaefer III, J. Am. Chem. Soc., 96, 7898 (1974).

64. C. Møller and M.S. Plesset, Phys. Rev. 46, 618 (1934).

65. D.D. Eley, Trans. Faraday Soc. 35, 1281 (1939).

66. H.S. Frank and M.W. Evans, J. Chem. Phys., 13, 507 (1945).

67. D.N. Glew, J. Am. Chem. Soc. 66, 605 (1962)

68. V.G. Dashevsky and G.N. Sariskov, Mol. Phys., 27, 1271 (1974).

69. J.C. Owicki and H. A. Scheraga, J. Am. Chem. Soc., 99, 7413, (1977).

70. M. Yacobi and A. Ben-Naim, J. Phys. Chem., 78, 175 (1974).

71. A. Holzer and M.F. Emerson, J. Phys. Chem., 73, 26 (1969).