INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

Ravishanker, Ganesan

COMPUTER SIMULATION STUDIES OF THE HYDROPHOBIC EFFECT

City University of New York

PH.D. 1984

Computer Simulation Studies of the Hydrophobic Effect.

by

Ganesan Ravishanker

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

1984

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.

4/12/84
date

Chairman of Examining Committee

4/25/84
date

Executive Officer

Supervisory Committee

The City University of New York

ABSTRACT

The tendency of apolar solutes to aggregate spontaneously in aqueous solution is presently recognized as one of the primary organizing principles in structural biochemistry and biology, and is widely known as the hydrophobic effect. The major objective of this dissertation research is to characterize the hydrophobic effect at the molecular level using computer simulation techniques. It is essential to understand the hydration of a single apolar species in order to make conclusions on the hydrophobic effect. In this study, we performed large scale Monte Carlo computer simulation studies on methane, a representative aliphatic molecule,and benzene, a representative aromatic molecule, in water. The results emerging from this dissertation research supports the recent experimental evidence that solvent-separated hydrophobic interaction plays an important role in the description of hydrophobic interaction. The results on hydrophobic interaction between benzenes suggests interesting stacking mechanisms for pi-cloud systems.

PREFACE

It was a privelege to have had the opportunity to work with Dr Beveridge for the past 5 1/2 years. He was a major force in the completion of this dissertation research as well as in making me a good researcher. Our enthusiastic discussions on the subject coupled with the complete independence I enjoyed in my work proved very helpful. As a graduate student from a foreign country I felt depressed sometimes for varied reasons and the encouragement I received from Dr Beveridge during those difficult times is unforgettable. I am grateful to him for all the help and advice that I received.

I would like to thank the members of my doctoral committee, Dr Howell, Dr Diem and Dr Hoyer for helpful suggestions.

I greatly appreciate the support and help from the members of our group, Dr Mezei, Dr Mehrotra, Mr Vasu and Dr Marchese. Special thanks to Dr Mezei for his patiance with my initial difficulties in computer programming.

The encouragement I received from my wife Nalini was a major factor in my life as a graduate student. I would not have come this far in studies, if not for the help I received from my cousins Mr J. Sundaresan and Mr N.B. Rajendra. They were great enough to help me financially when I had no place to go for my college expenses. Therefore I dedicate this thesis for them.

This thesis is dedicated to

Mr and Mrs Sundaresan and Mr N.B. Rajendra

Who made this all possible by helping me get through college.

TABLE OF CONTENTS.

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I.

Introduction.

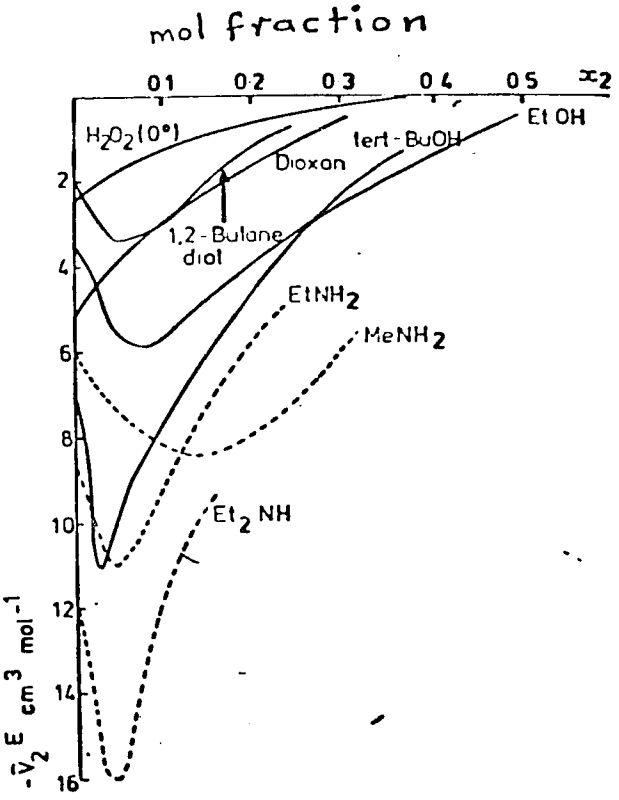
The tendency of apolar molecular species to associate spontaneously in aqueous solution is presently recognized as one of the primary organizing principles in structural biochemistry and biology. Kauzmann, in 1959, coined the term "hydrophobic bond" to describe this association process. Prevalent ideas about hydrophobic bonding at the molecular level involve contact interactions among the apolar groups, stabilized primarily by entropic effects originating in solvent water. The entropic nature of the process leads to an inverse temperature dependence for the association process, i.e. the tendency of apolar solutes to associate in water grows stronger with increasing temperature in the ambient range. However, another mode of hydrophobic interaction may operate over a longer range of distance than previously suspected, and may have important implications in structural biology. In this dissertation research, hydrophobic effects for prototypical alipahtic and aromatic molecules in water have been studied in detail at the molecular level using statistical mechanics. New knowledge of the nature of the hydrophobic interaction at both short and long range has been obtained.

The terminology in this research area requires precise specification. A controversy over the term "hydrophobic bond" arose in the 1960's from the recognition that the interactions involved are neither hydrophobic (the waters in the hydration complex make more favorable interactions with other waters than with the apolar solute,

rather than being repelled in any way by the solute), nor a bond (in the conventional enthalpic sense of the word, since the stabilization is entropic). However, in these systems a unique solvation process, quite distinct from other aqueous hydration processes such as hydrophilic and ionic, is undoubtedly involved and the issue has now become only a matter of terminology. The term "hydrophobic" remains with us now for historical reasons and also because of its mnemonic value in reminding us of the apparent result if not the correct nature of the process. These days, it is conventional to speak however of hydrophobic interactions (HI) rather than bonds, and to distinguish further the case described by Kauzmann as contact hydrophobic interaction (CHI), since the apolar species are presumed to be in direct spatial contact in this model. The detailed structural and energetical description of local solution environment of nonpolar molecules in the absence of solute-solute interactions is referred to as hydrophobic hydration.

During the mid 70's, Felix Franks and coworkers collected experimental evidence not explained by the classical model for HI proposed by Kauzmann. Partial molar thermodynamic indices, exemplified by volumetric measurements shown in Figure I.1 revealed that the indices of association are concentration dependent and indicate that more than one physicochemical association mechanism must be involved. This led Franks to propose a solvent modulated association in the

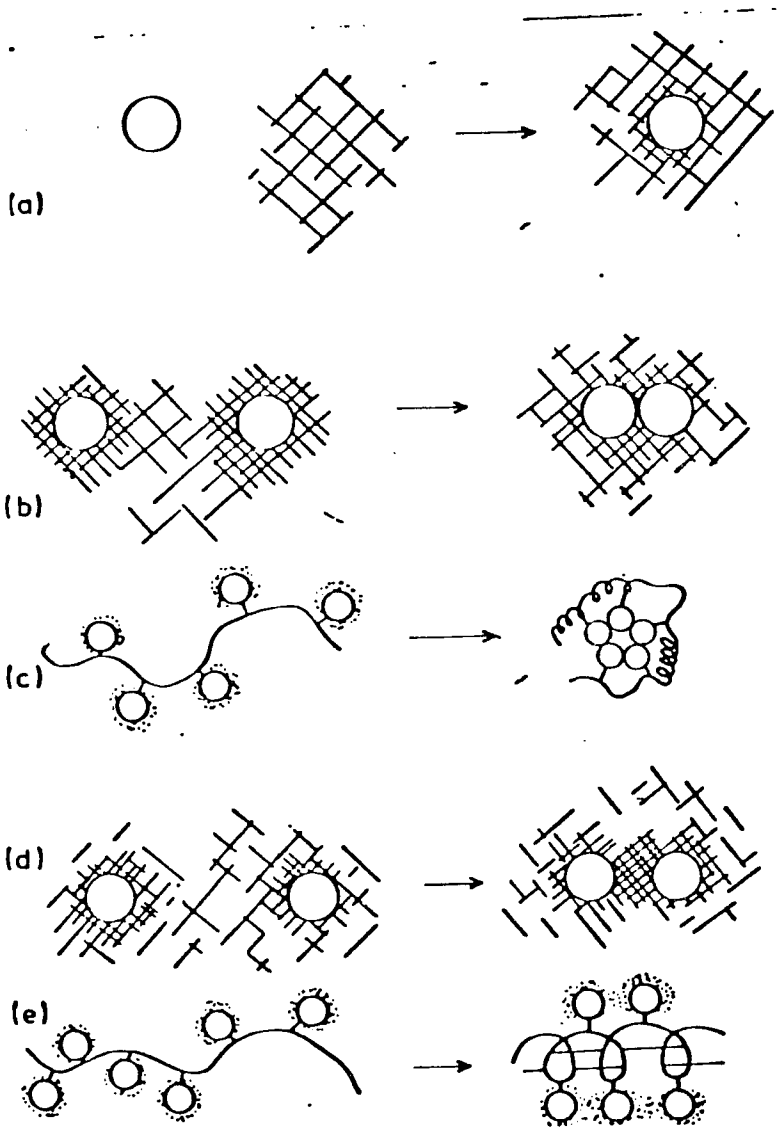
Figure I.1- Excess partial volume curves plotted against mol fraction of the solute for alcohols, amines, 1,2-butane diol and hydrogen peroxide at 25°C. The minima characteristic of essentially hydrophobic solutes are not observed for Dioxan or H_2O_2.[40]


mol fraction
0·1
0·2
0·3
0·4
0·5
x_2
H_2O_2 (0°)
Et OH
tert-BuOH
Dioxan
1,2-Butane diol
$EtNH_2$
$MeNH_2$
Et_2NH
2
4
6
8
10
12
14
16
$-\bar{V}_2^E$ cm^3 mol^{-1}

low concentration region of solute, hereafter referred to as a "solvent-separated hydrophobic interaction" (SSHI), and that the contact hydrophobic interaction takes over at higher concentrations of solute. The nature of the SSHI at the molecular level was not specified and no direct experimental technique has been used to test Franks's proposal. Computer simulation techniques can be used to test such a proposal. In this dissertation, Monte Carlo computer simulation techniques are used to probe the microscopic details of aqueous solutions of apolar molecules, and to study the relative importance of Kauzmann's model and Franks's proposal for HI. A pictorial representation of hydrophobic hydration and various hydrophobic interactions are given in Figure I.2.

The organization of susequent material in this thesis is as follows: In Chapter II, a detailed review of theoretical work on hydrophobic hydration and hydrophobic interaction is presented. For clarity, the literature is organized in chronological order and is divided into five subsections. Experimental work directly relevent to this study are also included. The theory and methodology used in our studies are described in Chapter III. The basic statistical thermodynamic setup framework for Monte Carlo computer simulation of solutions is discussed in this chapter, followed by the special techniques needed for modeling association processes. This chapter concludes with a brief discussion on intermolecular interactions.

Figure I.2- Diagrammatic representation of (a) hydrophobic hydration; (b) Kauzmann-Nemethy-Scheraga contact hydrophobic interactioni (CHI); (c) globular protein folding; (d) proposed solvent-separated HI (SSHI); (e) possible stabilization of helix by SSHI.


(a)
(b)
(c)
(d)
(e)

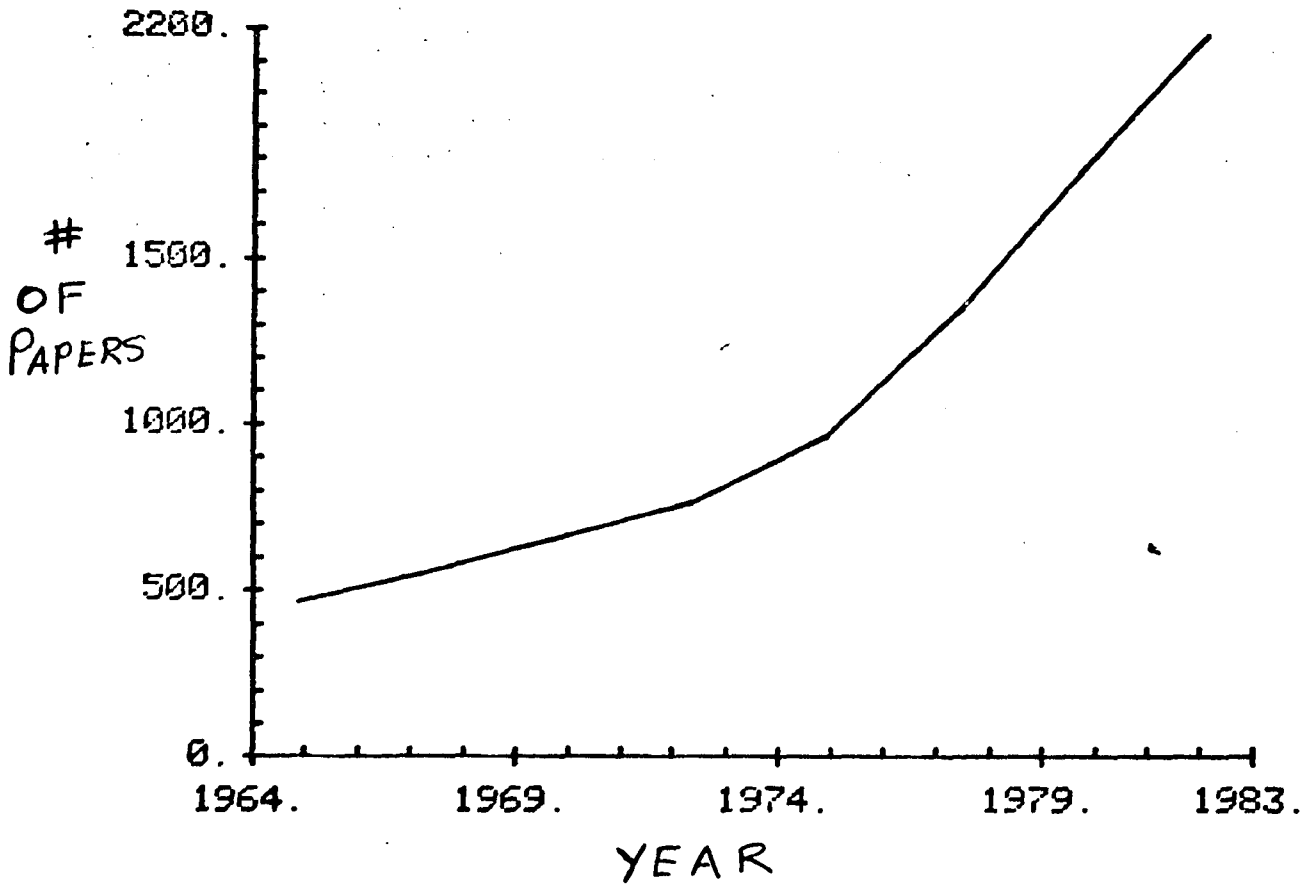
Chapters IV through VII discuss the results of computer simulations on dilute aqueous solutions of methane and benzene. Chapter IV deals with the hydrophobic hydration of methane. In chapter V, the simulations results on the association of methanes is given. Chapters VI and VII discuss the hydrophobic hydration of benzene, a prototypical aromatic system, and the association of benzenes along the C_6 axis.

CHAPTER II.

Background.

In this section a detailed review of previous theoretical studies on hydrophobic hydration and hydrophobic interaction is presented. The studies are reviewed in the chronological order they appeared. Included in the review are some experimental studies that are immediately relevent to the problems studied in this dissertation research. As shown in Figure II.1, the number of papers on hydrophobic interaction has grown so large and a complete review of such literature is beyond the scope of this section. This section is divided into five sections for convenience. The first section is devoted to the papers that appeared till 1960. Though this section is rather small, this is the most important era in the study of hydrophobic hydration and hydrophobic interaction, since the idea was formulated in that period. The four following sections review the works that appeared in the 60's, early 70's, late 70's and the early 80's.

Figure II.1-Number of papers containing the word "hydrophobic" in their title plotted against the year they appeared. This is taken from Science Citation Permuterm Index.


2200.
1500.
1000.
500.
0.
OF PAPERS
1964.
1969.
1974.
1979.
1983.
YEAR

1930-1960.

"Oil and water don't mix." The physical chemistry behind this generalization was studied in the first half of this century by measurements of the solubility of hydrocarbons in water by Butler[1], Eley[2], Lannung[3] and Valentiner[4]. Hydrocarbons were found to have very low solubilities in water and the water to vapor distribution coefficients of n-alkanes and aromatic molecules were found to be of the order of 10^3 and 10^2 respectively. The solubility was found to decrease with increasing chain length to the extent of 26% per $-CH_2-$ group. Butler, Eley and Frank and Evans[5] found the hydrocarbon dissolution process to be accompanied by a negative enthalpy and entropy changes and consideration of the process in terms of the free energy equation, $\Delta G=\Delta H-T\Delta S$, showed that the process is entropy controlled.

The behaviour of hydrocarbon/water systems was found to be a prototype for understanding important aspects of the three dimensional structures of proteins in solution. The structure of myoglobin (1959) and hemoglobin (1965) by Kendrew, Perutz and coworkers[6] showed that most of the non-polar groups cluster together in the protein interior and the polar groups were situated on the exterior making favorable contacts with solvent water and ions. This model remains today the textbook level description of

globular proteins. Perutz based on these observations called this the the oil drop model with the watch words non-polar in , polar out. However recent detailed studies[7] indicate the oil drop model to be oversimplified.

In 1945, Frank and Evans[5] proposed a theory to explain the hydrocarbon dissolution process at infinite dilution. The Frank-Evans theory is based on the idea that nonpolar molecules promote ordering of neighbouring water molecules leading to regions referred to as "icebergs", and characterized by increased hydrogen bonding among the neighbouring waters. Frank and Evans added the following note: "we shall use the word iceberg, without quotation marks or apology, to represent a microscopic region, either of pure water or surrounding a solute molecule or ion, in which water molecules are tied together in some quasi-solid structure. It is not implied that the structure is exactly ice-like, nor is it necessarily the same in every case where the word iceberg is used". Such a qualitative picture explained both the negative entropy change, arising from promotion of local water structure, and the negative enthalpy change, due to increased strength of hydrogen bonding between the local waters.

Kauzmann[8], in an important 1959 paper, advanced a theory for hydrocarbon association in water and introduced the controversial term "hydrophobic bonding". Kauzmann

adopted the Frank-Evans theory of hydrocarbon dissolution process to study the hydrocarbon association in water. The association brings together the hydrocarbons surrounded by Frank-Evans icebergs. The intersection of the solute hydrate cospheres causes some water molecules to be released from the ordered first solvation shells of hydrocarbons into the bulk, with a corresponding net increase in the entropy. Considering the enthalpy change for this process to be small, the free energy expression $\Delta G=\Delta H-T\Delta S$ with that negative ΔS, shows that the association process to have a negative free energy change. Thus, according to this view, the hydrocarbon association process in water is spontaneous and entropy driven. The entropic nature of the process leads to an inverse temperature dependence for the association process, i.e. the tendency of apolar solutes to associate in water grows stronger with increasing temperature in the ambient range.

1960-1969.

Nemethy and Scheraga[9], in a series of papers in 1962, treated hydrophobic hydration and hydrophobic interaction from a statistical mechanical point of view. They began with the flickering cluster model for water treated in an earlier paper and based on an idea by Frank and Wen[10], and extended it to encompass hydrophobic hydration and hydrophobic interactions. For each case, partition functions were written out explicitly in terms of the hydrogen bond interactions of various water clusters. The partition functions were used to calculate relevent thermodynamic properties for the model. Adjustable parameters in the Nemethy-Scheraga theory were parametrized against experimental data.

The details of Nemethy-Scheraga theory are as follows. Liquid water was considered to be an equilibrium mixture of clusters and monomers. Clusters are transient structures consisting of interconnected networks of water molecules. The networks arise from the hydrogen bonding of waters. Monomeric waters are those with no hydrogen bonds. Water molecules in the interior of clusters were assumed to have 4 hydrogen bonds, whereas the waters on the surface of clusters could have 1, 2, or 3 hydrogen bonds only. Water molecules are assigned different energies depending on the number of hydrogen bonds and their distribution among

possible energy states followed the Boltzmann Distribution Law. The main adjustable parameter in this theory was the water-water hydrogen bond energy, which was estimated from IR spectral data on liquid water[11]. The calculated values of free energy, enthalpy, and entropy of liquid water in the temperature range of Ø °C to 7Ø °C agreed with the experimental values within an error of less than 3%. The calculated temperature dependence of C_v was found to be too large compared to experiment. Using the mol fractions of various species present in water, Nemethy and Scheraga calculated the water-water radial distribution function, which compared well with the available x-ray data[12].

The Nemethy-Scheraga treatment of hydrophobic hydration incorporated the model for water described above and the idea that hydrocarbons the structure of water was increased in the vicinity of hydrocarbons. The probability of finding a water cluster in the first hydration shell of an apolar moiety was assumed to be greater than that in the bulk. Introduction of a hydrocarbon also introduces additional water-hydrocarbon interactions, favorable or unfavorable depending on how many hydrogen bonds are made or broken. Water-water interactions were assumed to be stronger than the water- hydrocarbon van der Walls interactions. The presence of a hydrocarbon stabilized a four-bonded water in the first shell, due to added Van der Walls interactions. On the other

hand, a hydrocarbon next to an unbonded water introduced unfavorable interactions. In addition, the size of the first shell depended on the size of the hydrocarbon and on temperature. A total partition function for this model of hydrophobic hydration was written as a product of partition functions of bulk water, of waters in the first shell of hydrocarbon, and of hydrocarbon. Parameters needed for the description of hydrocarbon partition functions were derived from experimental data[5] related to the transfer of hydrocarbon from pure liquid to water. Hydrocarbon-water interactions were assumed to be Lennard-Jones type with parameters evaluated from vapor phase data. Coordination numbers were derived from molecular models. Correlation between the calculated and experimental results on hydrophobic hydration were generally good. Standard enthalpies of solution and entropies of solution for aliphatic molecules agreed better with experimental values than for the aromatic compunds. Though the calculated trends in heat capacities reflected the experimental trends, their absolute values did not compare well with experiment, not surprising because the heat capacity for liquid water itself was in error. Also the model did not describe the temperature and pressure derivatives of the free energy for the transfer process.

The statistical thermodynamic treatment of hydrophobic interaction was set forth by Nemethy and Scheraga as an extension of their theory of liquid water and hydro-

phobic hydration. Statistical thermodynamic treatment of this process needed the addition of a description of the hydrocarbon pair in contact and the description of the concommitant release of some waters of the hydrophobic hydration into the bulk liquid. The solute-solute interactions were described by Lennard-Jones functions with parameters derived from vapor pressure measurements on hydrocarbons. Both the solutes were assumed to release an equal number of water molecules into the bulk when contact is established. The coordination number for the contact pair was necessary, which was evaluated from molecular models.

Nemethy and Scheraga studied several instances of the hydrophobic interaction using the approach described above. In studying the hydrophobic interaction between isolated chain molecules, it was shown that several levels of contacts are possible between the chains. Properties related to maximum and minimum hydrophobic interaction between aliphatic-aliphatic, aliphatic-aromatic and aromatic- aromatic side chains were evaluated and tabulated. The analysis was continued with various protein structures. In α-helical structures, geometry suggests that every 3rd or 4th member of the sequence will be neighbours, thus can form hydrophobic bonds, if possible. In a table they summarized the distances of closest approach of atoms of various side chains in α-helices of proteins as revealed in crystal structures. In the antipar-

allel β-pleated sheet it was found that extensive hydrophobic interactions were possible between α–α and between β-β carbons of apolar side chains. In the parallel chains, there is enough room between the residues for a sheet of water molecules parallel to the chain axis. Along the direction of the axis that is perpendicular to the chain axis, there is a possibility for extensive contact between α carbon atoms.

Solvent intervening hydrophobic interactions, subsequently referred to as solvent separated hydrophobic interactions(SSHI), were speculated as possibilities in this 1962 paper. Nemethy and Scheraga suggested that the solvent intervening hydrophobic pairs can arise mainly because of geometrical constraints which do not allow the apolar groups to come into contact, and presented results on such a system supporting their proposal. However, when there are no steric constraints preventing the hydrocarbon moieties from getting together, a contact hydrophobic pair was favored. Nemethy and Scheraga's work was is a milestone in the quantitative description of the hydrophobic interaction, but their speculation that the SSHI arises only due to geometrical constraints turns out to be questionable; study of this problem forms a major part of this dissertation research.

In 1962, Tanford[13] estimated the contribution from the HI to the stability of globular proteins. His calculations

are based on the following three assumptions about the native conformer of a globular protein: a) All charged groups in the side chain of amino acids appear only at the protein-water interface, b) most of the nonpolar side chains appear in the interior of the protien and c) polar groups may appear either in the interior or in the protein surface. The free energy of unfolding of a protein was expressed as

$$\Delta F = -T\Delta S^{o}_{conf} + \sum \Delta f_u$$

where ΔS^{o}_{conf} is the conformational entropy change due to unfolding and Δf_u is the free energy of transfer of small component groups of amino acids from their native environment to the unfolded environment. Tanford estimated ΔS^{o}_{conf} using some of Kauzmann's[8] data. ΔS^{o}_{conf} is given by xR ln Z, where x is the number of different conformations available for a protein and z is the number of isoenergy orientations of the protein for each such available conformation. Using Kauzmann's suggestions on the flexibility of random coil proteins, Tanford estimates $-T\Delta S^{o}_{conf}$ to be -1200 cal/residue. In order to calculate Δf_u, Tanford assumed that the major contribution was the buried nonpolar groups. In a previous study, Cohn and Edsall[14] reported the transfer free energies (Δf_t) for various amino acids from ethanol to water and found that Δf_t for various side groups are constants and are thus transferable. Tanford used the Δf_t values for various groups reported by Cohn and Edsall. A comparison of these quantities for other nonpolar solvents revealed them to be

independent of the nonpolar solvent system. Thus Tanford assumed that Δf_u was equal to Δf_t for the protein unfolding. Using the amino acid sequences for myoglobin, β-lactoglobin and ribonuclease, Tanford estimated the Δf_u values and $-T\Delta S^{o}_{conf}$, and found that the entropy increase during the unfolding process is mostly balanced by the decrease due to the loss of HI between the nonpolar solutes. Therefore Tanford suggests that the folding occurs mainly due to the hydrophobic effect which balances the entropy loss accompanying the folding. Tanford's model showed the importance of the HI in protein folding.

Scaled particle theory of liquids can be used to calculate the free energy of transfer of a hard sphere from vapor phase to a dense hard sphere fluid. Pierotti[15], in 1965, used the scaled particle theory of liquids to study the aqueous solutions of several hydrocarbons. He suggested that the cavity formation in liquids to be an important factor in the solution process and calculated the thermodynamics of cavity formation. Pierotti partitioned the solution process into two steps: a) the creation of a cavity in the solvent to accomodate the solute and b) the insertion of the solute in the cavity. The free energy of the solution process can therefore be written as,

$$G = \bar{G}_i + \bar{G}_C + \ln RT/V$$

where $\bar{G}_C$ is the partial molar free energy of cavity formation and $\bar{G}_i$ is the partial molar free energy of solute-solvent interaction resulting from the insertion of

the solute in the cavity. G_C was derived for a hard sphere fluid by Reiss et.al.[16] and is of the form,

$$G_C = K_\emptyset + K_1 a_{12} + K_2 a_{12}^2 + K_3 a_{12}^3$$

where K's are functions of density, temperature, and pressure. The quantity a_{12} is the diameter of the sphere which excludes any solvent centers.

Pierotti calculated G_C for several hydrocarbons using relation described above, and using hard sphere diameters for hydrocarbons from various sources and a σ_{H_2O} of 2.75 Å. He treated the solute-solvent interactions to be Lennard-Jones interaction and calculated G_i. ΔG values calculated for the solution process for hydrocarbons using this model were fairly close to the experimental values. Pierotti also evaluated the ΔH and ΔS of solution and found them to be in even better agreement with experiment. He extended this method to study in detail the solutions of argon and nitrogen in both water and benzene. The results revealed that the cavity formation in water is entropy controlled but it is enthalpy controlled in benzene. Pierotti therefore concluded that the structural changes in water resulting from the dissolution of hydrocarbon give rise to rather small enthalpy changes and large entropy changes, but since this is a thermodynamic study, specific details of the structural changes in water could not be derived. Pierotti, using the scaled particle theory of liquids, provided an elegant proof that there is a basic difference in the nature of cavity formation in water and

organic solvents.

In 1967, Kauzmann revisited the hydrophobic effect In a paper together with Kozak and Knight[17], lattice models and McMillan-Mayer[18] theory were used to investigate the hydrophobic interaction process. The heart of their arguments was the equation,

$$\ln \gamma_1 = Bx_2^2 + Cx_2^3 + \ldots\ldots\ldots$$

where γ_1 is the solvent activity coefficient, x_2 is the mol fraction of the solute and B and C were constants. The sign and magnitude of B and C play an important role in understanding the hydrocarbon solutions in water. Kozak et.al. discussed using theoretical models and how hydration, solute-solute association etc., influence these constants, and then trying to explain experimental observations based on these arguments.

Lattice models were used as a first approximation to the solution of hydrocarbon in water. Lattice theory for mixing two components already existed at that time, developed by Flory and Huggins[19a,b], and further advanced by Guggenheim[19c]. The solution was assumed to have N_1 solvents and N_2 solutes and the process of interest is their mixing. Each solute was allowed to be a polymer with some fixed solvent coordination number. Several cases were considered. First, the solvent molecules that were in the solvation shell of a given solute were not allowed to be shared by another hydrocarbon. Also, the enthalpy of

mixing was assumed to be zero. Flory and Huggins showed that calculation of B and C for such a system involved carefully evaluating the number of ways such a system can be arranged. Kozak et. al. calculated B and C using this approach, and found that B and C were functions of coordination numbers only, and were both negative. In the next step, the unique assignment of solvents to a solute was removed and a solvent was allowed to be in the first solvation shells of more than 2 solutes. This improvement yielded B and C which were still only functions of coordination numbers and both negative , but different in magnitude from the previous model. Kozak et.al. went a step further by removing the constraint that the enthalpy of mixing had to be zero. By applying Guggenheim's quasichemical approximation, they evaluated B and C. Here the sign of B depended on the relative magnitudes of solute-solute interaction and solute-solvent interaction. The sign of C was found to be always negative due to hydration. The sign of B was found to be positive only when the solute-solute interactions predominate the solute-solvent interactions.

Kozak et.al. also used McMillan-Mayer theory to understand the nature of the constants B and C and relate them to the potential of mean force(pmf) $w_{ss}(R)$ between the solutes. At that time no procedures were available to calculate the potential of mean force, so an approximate evaluation was carried out using an assumed model. The

pmf was decomposed into a repulsive and an attractive part,and relative magitudes of the terms was predicted. The decomposition itself is valid, but the model pmf that was used decreased monotonically with distance. The quantities B and C were found to depend on the solute size, and were contravariant; B becomes more negative on increasing the solute size, whereas C becomes more positive. Increased solute-solvent interactions tend to make both B and C more negative.

With all these considerations in hand, the next step was to compare them to the experiment. The analysis was done in two parts. All molecules except amino acids were analyzed in one group and the amino acids were analyzed separately. From a large database of experimental data on aliphatic alcohols, amides, carboxylic acids, amines, ketones as well as some aminoacids, the constants B and C were evaluated. Typically this involved measuring solvent activity data as a function of mol fraction of solute using freezing point methods. Then, from a plot of $\ln \gamma_1/x_2$ vs x_2 one can get the y intercept B and the slope C. Signs of temperature derivatives of both B and C were also calculated for some solutes.

The quantities B and C for most of the aliphatic compounds in the data base were found to be positive, which indicates that the solute-solute interaction predominate in hydration effects. Therefore, Kozak et.al. con-

cluded that the compounds in their database, mainly consisting of hydrophobic groups, favored the association. Further, B becomes more positive on increasing the solute size, showing that the association is favored as the hydrophobic group gets larger in size, and also on increasing temperature. The latter effect confirmed another well documented fact about hydrophobic effect, the inverse temperature dependence.

The sign of C and the conclusions from McMillan-Mayer theory pointed to similar conclusions. Amino acids showed a different behaviour, despite the fact they also have hydrophobic groups. Kozak et. al. explained this in terms of the strong electrostatic forces which dominate the relatively small hydrophobic effects in amino acids. Overall this study, featuring a totally different approach to the hydrophobic interaction problem,generally confirmed earlier predictions about hydrophobic interaction proposed by Kauzmann 1959.

In 1968, Hildebrand[20] wrote a small note in which he questioned the appropriateness of the phrase "hydrophobic bond" to describe the association of nonpolar solutes in water. He cited two examples to show that nonpolar molecules have no phobia for water. The adhesion between octane and water is 44 ergs cm^{-2}, which suggests considerable attraction between octane and water. Also, it takes more energy to evaporate a mole of butane from its aqueous

solution than from pure liquid butane, suggesting more attraction between butane and water than between butane molecules. Hildebrand added that the word 'bond' is inappropriate also, since the attraction between alkyl groups in water has none of the characteristics of a chemical bond.

Nemethy, Scheraga and Kauzmann[21] offered an explanation to Hildebrand's arguments in an adjoining reply. They agreed with Hildebrand that the nonpolar solute-water interactions are indeed attractive, but this is only the enthalpy term of the dissolution process. Experiments as well as quantum mechanical potentials show that the interaction between nonpolar molecules and water is attractive due to dipole-induced dipole interactions and dispersion forces. Nemethy et.al. pointed out that the word 'phobia' is not formulated based on enthalpy of interaction, rather on the free energy of solution. The latter quantity has a large unfavorable contribution in water from entropy, which makes the nonpolar molecules prefer nonpolar surroundings more than water. They also suggest that the word 'bond' is only a loose association and that it is understood by various workers in the field.

1970-1974.

Goldammer and Hertz[22], in 1970, used a novel approach to probe hydrophobic hydration and HI in aqueous solutions of effectively hydrophobic solutes. NMR data on time correlation functions was used to explain the life times of first hydration shells and the possibility of solute-solute aggregation. Goldammer and Hertz consider five different models for hydrophobic hydration and one for HI, and derive the relationship between solute and solvent correlation times. The first model of Goldammer and Hertz used a rigid solute with a long-lived first hydration shell, such that the solute and and the first shell waters translate and rotate together. Here two conditions must be satisfied: a) the self-diffusion coefficients of both the solutes and the first shell waters should be the same and b) the rotational correlation times between various atoms of the solute and the atoms of water must be the same. After comparison with experimental results, Goldammer and Hertz found no system in their databse to satisfy these rigorous conditions. and concluded that first hydration shells of hydrocarbons are not essentially long-lived. The results showed that hydrocarbon rotations are generally faster than that of first shell waters.

In their second model, Goldammer and Hertz still considered the solute to be rigid, but relaxed the

requirements on the rigidity of first shell waters. They considered that a distinct first hydration shell is possible as long as the distance between the center of mass of solute and of first shell waters remained the same over sufficient times. This model thus allows for the independent rotations of first shell waters about a fixed relative position from the solute. Under these assumptions, the self diffusion coefficients of the solute and first shell waters need to be the same, but the rotational correlations between the atoms of solutes and atoms of first shell waters need not be same. In the database of systems Goldammer and Hertz studied no entry satisfied the requirements of this model, proving that the first hydration shells of hydrocarbons are not long-lived even in a center-of-mass sense.

Having found no long-lived hydration shells for hydrocarbons, Goldammer and Hertz proposed their third model to check the existence of long-lived hydrocarbon-water pairs. This requires that the self-diffusion coefficients of solute and first shell waters be the same and also that relevent solute-water rotational correlations be the same at solute mol fractions of Ø.5. Experimental results showed that in the aqueous solutions of pyridine,ethanol, acetone and acetonitrile, long- lived hydrocarbon-water pairs are indeed possible. However, a deeper analysis showed that these pairs are due to the water hydrogen bonded to the polar groups such as -OH or

-CO rather than waters in the hydrophobic hydration sphere.

Goldammer and Hertz's next model treated the solute and the first hydration shell as molecular aggregates of fluctuating sizes. Under these assumptions, one can express the self-diffusion coefficients and rotational correlation times as functions of experimentally measurable quantities such as viscosity, and cluster size, an adjustable parameter. Goldammer and Hertz calculated cluster sizes for the molecules in their database based on molar volumes. Self-diffusion coefficients and the rotational correlation times calculated by this method did not compare well with the experimental values. This confirmed the conclusions from previous models, i.e. that the hydration shells are not long-lived enough to be considered as a molecular aggregate. Various pair correlation functions were used by Goldammer and Hertz to probe the nature of hydration shells of hydrocarbons. All solutes showed an increase in the water structure compared to the pure water, manifested by an increase in the amplitude and narrowing of the first peak of water-water pair correlation function. A further analysis showed that for some solutes, such as methanol, this structure promotion involved direct solute- solvent interactions. For weakly interacting solutes such as acetonitrile, the structure promotion did not involve any direct solute- water interactions.

In their sixth and last model, Goldammer and Hertz probed the possibility of solute-solute aggregation. The product of solute self-diffusion coefficient and normalized rotational correlation time of solute should be independent of the solute concentration if there is no aggregation. In addition, this product is inversely proportional to the cube of the closest distance the protons in different solutes can approach each other. An inverse dependence between the abovementioned product and concentration of solute would mean a solute aggregation. Experimentally the product could not be evaluated with proper accuracy at low concentrations of the solute. For some system the product showed inverse dependence and for some, a direct dependence, as a result of which no definite conclusions on this aspect of the problem were possible.

This work presented compelling evidence against the possibility of long-lived, distinct first hydration shells for hydrocarbons, and supports fluctuating structures. Goldammer and Hertz confirmed the increase of water structure near hydrocarbons using microscopic probes rather than such macroscopic probes as thermodynamic quantities. This work came up with a result that was not known before, and not given much attention: that the hydrocarbons rotate faster than the first shell water molecules. No conclusive picture for HI emerged from this work, as a result of the limitations that existed in the NMR meth-

odology in the early '70s.

Significant contributions to the understanding of HI came from a series of papers by Ben-Naim[23]. Using basic statistical mechanics, he derived a quantity,which he called δA^{HI} , as a measure of the stability of hydrophobic interaction. A system consisting of N water molecules and 2 solutes, fixed at a distance of R, in a (T,V,N) ensemble is considered. The Helmholtz free energy was derived to be

$$A(R) = A^{o} + U_{12}(R) + A^{HI}(R)$$

where A(R) represents the total free energy of the system, with the solutes separated at a distance R, and $U_{12}(R)$ is the solute- solute interaction energy when solutes. Here $A^{HI}(R)$ measures the work done in bringing the two solutes from an infinite distance to a distance R in water. The free energy change for that process in Ben-Naim's formalism reduces to

$$\Delta A(\sigma) = U_{12}(\sigma) + \delta A^{HI}(\sigma)$$

where σ is the distance of closest approach and $\delta A^{HI}(\sigma)$ is a measure of the solvent contribution to the free energy of HI process.

The quantity $\delta A^{HI}(\sigma)$ can be related to experimental quantities under certain assumptions, which essentially involve finding a representative molecule that closely resembles the two solutes at the distance of their closest approach. Ben-Naim treats the free energy

of HI for two methane molecules in water as an example. Two methane molecules at their closest distance of approach is represented by an ethane molecule, which reduces the $\delta_{A^{HI}}(\sigma)$ to,

$$\delta_{A^{HI}}(\sigma) = -RT \ln \gamma_E / \gamma^2 .$$

where γ represents the fugacities of the respective compounds. The quantity $\delta_{A^{HI}}(\sigma)$ was calculated for various solvent. All solvents except for water gave a value around -1.5 kcal/mol. The calculated $\delta_{A^{HI}}(\sigma)$ value for water was found to be -2.15 kcal/mol, which reflects an increased tendency for association, and points up to the uniqueness of water as a medium for nonpolar solute associations.

Ben-Naim argues that the quantity $\delta_{A^{HI}}(\sigma)$ is a better representative in measuring the HI than the total free energy change, because it removes the direct solute-solute effects and includes only the contribution from solvent modulated solute-solute interaction. His evaluation of $\delta_{A^{HI}}(\sigma)$ for various values of R of theoretical interest shows this to be a a monotonically increasing function of R. In an important derivation, Ben-Naim shows that the quantity $A^{HI}(R)$, which is the total free energy of the system is equal to $w_{ss}(R)$, the pmf between two solutes. Though the introduction of $\delta_{A^{HI}}(\sigma)$ helped establish the uniqueness of water in determining HI, furthur extensions were limited due to the requirement of a representative that closely resembles the two

solutes in contact. In addition, only contact hydrophobic interactions are treated by this theory.

Ben-Naim , Wilf and Yaacobi[24], in 1973, used experimental Ostwald absorption coefficients to calculate δA^{HI} for methane,ethane and benzene in both water and D_2O. The contact pairs for methane, ethane and benzene were assumed to be ethane, n-butane and biphenyl respectively. This study showed that the strength of HI is stronger in water than in D_2O for methane and ethane but it is opposite for benzene. The isotope effect on δH^{HI} and δS^{HI} were found to be rather insignificant. Ben-Naim et. al. reported the δA^{HI} values for benzene at various temperatures and it is -1.14 kcal/mol. This study provided an estimate of the strength of HI between benzene molecules in water. However, treating biphenyl as the contact pair for benzene is questionable, since there are other well packed dimers that are as important.

The first liquid state computer simulation study of the hydrophobic effect was reported by Dashevsky and Sarkisov[25] in 1973. They used Monte Carlo method to evaluate the configurational integrals for aqueous solutions ofmethane and for two methane molecules in water as a functions of intersolute separation. The configurational integral for a system of N molecules following Boltzmann statistics is,

$$Z_N = \int \ldots\ldots\ldots \int \exp(-E(X^N))\, dX^N$$

where $E(X^N)$ is the total energy of the system, when the system is in configuration X^N. The integral has to be evaluated over all configurations X^N. Once Z_N is known, it is straightforward to evaluate the thermodynamic properties of the system. The properties thus calculated can be compared with the experimental measurements to check the validity of the model.

Dashevsky and Sarkisov carried out MC simulations on a system of 64 waters, 1 methane and 63 waters and 2 methanes in 62 waters. Methanes were treated as spheres of radius 2 Å. All intermolecular interactions were described by Kitaygorodsky potentials[26]. Periodic boundary conditions were used and the simulations were carried out in (T,P,N) ensemble. Each simulation was 60K in length. The calculated free energy of solvation of methane evaluated from the MC simulation is positive and the entropy of solvation is negative. Dashevsky and Sarkisov showed that these compared well with experimental values after correcting for the same reference state. A total of eight different simulations were carried out on the system of 2 methanes in 62 waters. In each simulation, the methanes were kept at some fixed distance R. The free energy obtained from the configurational integrals of eight simulations yielded the potential-of-mean force, $w_{ss}(R)$ between the methanes as a function of R. A least square fitting yielded continuous $w_{ss}(R)$ with a single minimum at 4 Å. Dashevsky and Sarkisov interpreted this as the tendency

for methane molecules to be in contact at 4 Å in water.

Clear analysis reveals there are severe problems associated with the methodology used in Dashevsky and Sarkisov's calculations. Owicki and Scheraga[49a] discussed these in their paper of 1976. A careful look at the configurational integral shows that some configurations with unusually large $E(X^N)$ would contribute significantly towards Z_N. However, in a typical MC simulation, such high energy configurations are highly undersampled or never sampled, resulting in a highly unreliable Z_N. Since Z_N is unreliable, so are the thermodynamic properties derived from it. In light of this analysis, Dashevsky and Sarkisov's results have large errors associated with them and not reliable. The $w_{ss}(R)$ calculated by Dashevsky and Sarkisov for methane molecules is one of the non-oscillatory $w_{ss}(R)$ reported till now. This can easily be attributed to the defects associated with their methods.

O. Howarth[27] took an approach similar to that of Goldammer and Hertz in using NMR relaxation data to probe the structures of aqueous solutions of hydrocarbons. Whereas Goldammer and Hertz concentrated in studying the motional effects of water molecules in the first hydration shell of hydrocarbons, Howarth was interested in studying the motional effects of the hydrocarbon in water, compared to that in a non-polar slovent. He used FT NMR techniques to observe C^{13} relaxation times.

Howarth studied t-butanol and n-propanol in CCl_4 and in water. As the concentration of the alcohol increased in CCl_4, the tumbling of the alcohol became more and more restricted due to the increase in the intermolecular hydrogen bonding between the alcohols. In liquid water, the motions of the alcohols were independent of their concentrations, but their intrinsic magnitude suggested their motion to be more restricted in water than in CCl_4. Howarth proposed that the entropy reduction taking place during the transfer of a mole of hydrocarbon from pure liquid to water has two contributions. The first is due to the motional restrictions of the first hydration shell waters, as observed by Goldammer and Hertz, and the second is due to the motional restrictions the hydrocarbon itself undergoes. Howarth estimates the latter to contribute about 7.5 cal $K^{-1}mol^{-1}$. Thus Howarth demonstrated that a hydrocarbon solute undergoes motional restrictions upon dissolution in water, and the contributions from such restrictions to the total entropy reductions are indeed important.

In the early 70's R. B. Hermann[28] published some interesting work in which he combined basic statistical mechanics with significant structure theory of liquids to model the aqueous solutions of hydrocarbons. Hermann models hydrophobic hydration along the lines of Nemethy and Scheraga. He considers the solutes to be far from each other in water and noninteracting as a result.

The solutes are assumed to influence only those waters in their first hydration shell. He writes the partition function for the model exactly like Nemethy and Scheraga,but uses significant structure theory in writing the partition functions for bulk water and the waters in the first hydration shells of the solutes. This process introduces several adjustable parameters and assumptions into the model. Coordination numbers needed for the evaluation of partition functions are taken directly from Nemethy and Scheraga. Several experimental thermodynamic parameters are used to evaluate the parameters for the model. These parameters are then used to evaluate additional properties which compared well with the experiment. In a refinement to the above model, asymmetry in the electric fields of the water molecules in the first hydration shells of the solutes was introduced. This refined model is solvable only if one assumes that the hydrocarbon dissolution process is accompanied by a net increase in the volume. However, it is known from Masterton's[29] data this process involves a net decrease in volume.

In a following paper, Hermann[30] tries to relate quantitatively the surface areas of various hydrocarbons with their solubilities. He considers the process of bringing a hydrocarbon from pure liquid to water as involving creation of a cavity in liquid water to place the hydrocarbon. The free energy for the creation of cavity is therefore proportional to the surface area of the

hydrocarbon. Under these assumptions, the surface area of the hydrocarbon is proportional to -ln(solubility). Since the hydrocarbon is not rigid, and several conformations are accessible to the hydrocarbon, Hermann considers the surface area of a given molecule to be a weighted sum of surface areas of its conformers. The weight is unity for the molecule which has no gauche interactions and every additional gauche interaction introduced reduces the weight by exp(.8). A computer program is used to calculate the surface areas of most probable conformers of a given molecule and the weighted average of surface area was evaluated for a set of hydrocarbon molecules. The surface areas calculated by this method show an excellent correlation with -ln(experimental solubility). Therefore he suggests that weighted average surface areas of hydrocarbons is a good measure of the solubility. Traditionally one tries to correlate molar volumes of hydrocarbons to their solubilities. However no clear cut relation can be found in such a method, whereas, Hermann argues that his formalism has no exceptions, if applied carefully. McAuliffe's[31] solubility data shows that branching in hydrocarbons increases solubility. Simple molar volume- solubility relationships needed additional arguments to explain this observation. In Hermann's formalism, branching reduces the surface area of the hydrocarbon, which in turn increases the solubility.

Friedmann and Krishnan[32], in 1972, used the Gurney model[33] of solvation to study the HI in a set of alcohols. The solutes considered were methanol,ethanol, propanol, n-butanol and t- butanol. In the Gurney formalism, each solute in solution is surrounded by a region of solvent around the solute which is significantly perturbed by the presence of solute, called the cosphere. When two solutes approach each other, their cospheres overlap and some solvent molecules in the overlapping region are displaced into the bulk of the solution. This displacement has a free energy change associated with it, called $GUR_{12}(R)$, where R is the intersolute separation. The term $GUR_{12}(R)$ is proportional to the volume of overlap and A_{12}, the free energy per mol of solvent displaced. Usually A_{12} is treated as an adjustable parameter and experimental virial coefficients are used to evaluate them.

Friedmann and Krishnan suggest that the pmf between two solutes in water can be written as

$$w_{ss}(R) = COR_{12}(R) + GUR_{12}(R)$$

where $COR_{12}(R)$ is the work done against direct solute-solute interaction. above. Friedmann and Krishnan use a simple spherical repulsive potential of the form r^{-9} to describe the $COR_{12}(R)$ term and an effective spherical equivalent for non-spherical molecules. Volumes of overlap of cospheres of solutes are a function of the distance between the solutes and also a function of the conformation of the solute. An average over

significant conformations of given solutes was performed to calculate the average volumes of overlap. The cospheres are assumed to be only one water layer thick. Under these assumptions $w_{ss}(R)$ is a function of only the Gurney parameter, A_{12}. Friedmann and Krishnan calculated the solute- solute pair correlation function $g_{ss}(R)$ from $w_{ss}(R)$ and used hypernetted chain equations to calculate $g_{ss}(R)$ as a function of solute concentration. The $w_{ss}(R)$ and $g_{ss}(R)$ are then used to calculate some excess free energy and other thermodynamic quantities. A comparison of calculated results with experimental values from Kozak et.al.[17] yielded values for A_{12}.

The Gurney parameter, A_{12} turned out to be negative for all the solutes studied, and increased with the size ot solute. Thus the removal of waters from the overlapping regions of solute cospheres to the bulk is a spontaneous process, and increasingly so for larger solutes. Friedmann and Krishnan used the values of A_{12} to calculate some other thermodynamic quantities such as excess enthalpy which correlate well with the experiments.

Friedmann and Krishnan showed how the Gurney model can be adopted to explain hydrocarbon association processes. Their method however has some problems, such as using a single Gurney free energy parameter A_{12} as an aver-

age to represent several conformers of a given solute. The power of this method to yield the structural indices such as $w_{ss}(R)$ was later realized and used by Clark et.al.[38] in a calculation to be described later.

1975-1979.

In 1975, Ben-Naim[34] used equilibrium thermodynamics and power series expansion techniques to study the relationship between the structural changes in water on hydrocarbon association and the free energy of HI. In a previous work he observed that the δA^{HI} is markedly larger for water than for nonpolar solvents, which is often interpreted as arising from the structural changes in water. Such interpretations do not have direct experimental proof and are arrived at using some ad hoc model for liquid water. Ben-Naim worked with well defined molecular quantities for liquid water to investigate the role of structural changes in water to the free energy of HI.

The details of Ben-Naim study are as follows. Liquid water was assumed to consist of N_L molecules in some unspecified low structure form and N_H molecules in a high structure form. At equilibrium,

$$\mu_L = \mu_H$$

where μ_L and μ_H are the chemical potentials of the low and high structure forms of water. It is important to note that a representation of water as a two component mixture is not necessary for the following arguments. One can as well consider water to be a mixture of some n components based on any criterion nad carry out the following arguments without any problems. The two compo-

nent representation is chosen for clarity only. Under this representation, water structure at any instance can be specified by specifying N_L and N_H.

Ben-Naim then considers the free energy of HI process as consisting of two parts,

$$\Delta G^{HI} = \Delta G^* + \Delta G^r$$

where the first term ΔG^* is the work expended in bringing the solutes from an infinite distance to contact in water, keeping the water structure constant, viz.,

$$\Delta G^* = G(N_L, N_H, R=\sigma) - G(N_L, N_H, R=\infty)$$

Here R is the intersolute distance and σ is the contact distance and the numbers N_L and N_H represent the equilibrium numbers of the low and high structure form of water when the solutes are separated by an infinite distance. The second term, ΔG^r is the free energy associated with the relaxation of water structure to its equilibrium value when the solutes are at contact.

$$\Delta G^r = G(N_L', N_H', R=\sigma) - G(N_L, N_H, R=\sigma)$$

where N_L' and N_H' represent the number of low and high structure forms of water at equilibrium when solutes are at contact. ΔG^r therefore represents the contribution to G^{HI} from structural changes in water.

Ben-Naim proceeded to expand ΔG^r about ,

$$\Delta G^r = (\mu_L - \mu_H)\ dN_L + \frac{1}{2}\ (\mu_{LL} - 2\mu_{LH} + \mu_{HH})\ dN_L^2$$

where μ's represent the derivatives of G with respect to N_L. The second and higher derivatives of G can

be shown to be homogeneous functions of order minus one in N, the total number of water molecules, and they all vanish in the macroscopic limit of $N \to \infty$. However, the equilibrium considerations require that $\mu_L = \mu_H$, which means $\Delta G^r = 0$, or the free energy of structural changes in water during the HI process is zero. By similar arguments Ben-Naim shows that the corresponding enthalpy and entropy changes are nonzero.

Ben-Naim uses statistical thermodynamic arguments to substantiate the proof described above. The free energy change for the HI process can be derived for a (T,P,N) ensemble as,

$$\Delta G^{HI} = U(\sigma) + \int_{\infty}^{\sigma} \frac{\langle B(R) \rangle}{\langle R \rangle} \, dR$$

where $U(\sigma)$ is the solute-solute interaction energy at contact, and the second term represents the work done against an average force arising from the change in binding energy of the pair of solutes as they approach each other from to contact. Ben-Naim argues that there is no term in the above expression which depends on the structural changes in water, therefore contribution to ΔG^{HI} from structural changes in solvent is zero.

This work refuted the belief that the water structure plays an important role in the HI process, and has roused considerable controversy. In a latter paper Marcelja et. al.[47] questioned Ben-Naim's treatment on two grounds. They said that in definig the water structure in terms of N_L

and N_H, Ben-Naim omits the structural changes arising as a result of orientational differences. As a result, orientationally different structures having the same N_L and N_H are treated structurally same in Ben-Naim's theory. Secondly, Marcelja et. al. note that the expansion used for ΔG^r is incorrect. They argue that Ben-Naim ignores contributions from inhomogeneity near the solute-solvent interface to ΔG^r. In presence of such effects the free energy is known to be non-extensive in the thermodynamic limit and as a result ΔG^r is nonzero. Finally we are intrigued by this result for the following reason: if the structural changes in water are not responsible for HI, what is responsible and what makes water so unique with respect to HI?

Hermann[35a] used his earlier results on hydrocarbon surface areas to calculate the free energy of HI. He adopted Ben-Naim's ideas in defining the contact solute pairs for this study. Hermann embarks on a new route to study the hydrophobic hydration rather than use the significant structure theory of liquids. He uses Barker-Henderson perturbation theory for liquid mixtures. In this formalism, a reference liquid mixture consisting of hard spheres is perturbed by introducing long range forces between the molecules. Chemical potential for a liquid mixture consisting of N_1 solutes and N_2 solvents can then be written as a sum of two terms, one arising from hard sphere contribution , μ_{HS} , and a correction factor, μ_{CORR} , to account for the perturbation. Hermann uses

the μ_{HS} values calculated by Neff and McQuaire[35b]. A relation connecting experimental quantities and μ_{HS}, μ_{CORR} is as follows:

$kT \ln p_2/X_2 - kT \ln RT/V = \mu_{HS} + \mu_{CORR}$ where p_2 is the partial pressure of the hydrocarbon and X_2 is its mol fraction. Since μ_{HS} is known from Neff and McQuaire, μ_{CORR} can be calculated by using experimental partial pressures at various mol fractions.

Hermann then proposed a theoretical model to calculate μ_{CORR} which can be compared to the $_{CORR}$ calculated from experimental quantities. μ_{CORR} defined by Hermann is the same as G defined by Pioretti and consists of two terms, μ_C and μ_{IN}. μ_C is the work done in creating a cavity whose surface area is the same as that of the hydrocarbon solute in liquid water. Hermann uses the surface areas of hydrocarbons that he calculated before and evaluates μ_C. μ_{IN} is calculated by describing both thw water-water and water-solute interaction energies by Lennard-Jones interaction. Hermann calculates the free energy of HI as the difference in μ_{CORR} between hydrocarbons separated by a large distance in water and at contact distance. He adopts Ben-Naim's method of choosing the contact pair. Thus calculating the difference of μ_{CORR} between two methane molecules and one ethane molecule is an estimate of HI free energy for methane. He also tries to evaluate μ_{CORR} as a function of distance of separation between the solutes. Severe problems

arise in describing the regions in between the solutes when a water molecule can be placed there. Thus, he claims that his theory cannot be used to calculate HI free energy as a function of distance of separation between the solutes. This work showed how liquid state perturbation theory can be used to study the HI.

Wolfenden and Lewis[36], in 1976, presented a vapor phase analysis of HI in the further investigation of some earlier claims of Butler[1]. They argued that the HI can in principle arise from cohesive properties of water, from the affinity of nonpolar substances for each other, or from both. In 1937 Butler had proposed that the equilibrium transfer of a hydrocarbon from water to pure liquid is better understood by studying an equivalent two step process involving vapor phase equilibrium. The first step is the equilibrium transfer of the hydrocarbon from water to vapor phase and the second step is the equilibrium transfer of the hydrocarbon from vapor phase to pure liquid hydrocarbon. Assuming that the molecular interactions in the vapor phase are negligible, the first equilibrium measures the contribution of cohesive properties to HI, and the second measures the contributions from the attractive nonpolar interactions to the HI. Wolfenden and Lewis used hydrocarbon solubility data from McAuliffe[31] and Schlesinger[37] to calculate the free energy changes for the two vapor phase equilibria suggested by Butler, for the first eight

alkanes.

The data collected by Wolfenden and Lewis revealed some interesting information related to the HI. They found that the equilibrium distribution of hydrocarbon between water and pure liquid hydrocarbon decreased about 26% for every $-CH_2-$ group added. Wolfenden and Lewis estimate that the free energy transfer of a hydrocarbon from water to pure liquid hydrocarbon increased by 960 cal. for every additional $-CH_2-$ group. The partitioning as described above reveals that each added $-CH_2-$ group favors the vapor phase over water by only 140 cal., whereas each added $-CH_2-$ group makes the hydrocarbon prefer pure liquid over vapor by 840 cal. Wolfenden and Lewis therefore conclude that the nonpolar interactions indeed play an important role in hydrocarbon dissolution process, supporting an earlier speculation by Butler. They also point out that the cohesive properties of water contribute little compared to the attractive nonpolar interactions.

This note by Wolfenden and Lewis basically collects evidence in support of Butler's earlier claims on hydrocarbon solutions. Their technique powerfully demonstrates how the partitioning of overall equilibria can be of use in understanding contributions from subprocesses. Wolfenden and Lewis show that it is incorrect to treat HI as arising purely from the cohesive properties of water and that the hydrocarbon interactions can be of equal impor-

tance. A similar study involving hydrocarbon transfers from other common polar solvents to pure liquid hydrocarbon, in principle should reveal the uniqueness of water in relation to HI.

Clark, Franks, Pedley and Reid[38] in 1976, noted that the Gurney free energy calculated for HI by Friedmann and Krishnan[32] can be compared to the $\delta A^{HI}(R)$ formulated by Ben-Naim after suitable corrections. They were also interested in extracting the pmf between solutes using the Gurney model. Friedmann-Krishnan model for HI used a purely repulsive solute-solute potential to represent $COR_{12}(R)$, the direct solute-solute interaction term. This implies that either the solute-solute attractive interactions are negligible or that such interactions form a part of the Gurney free energy term, $GUR_{12}(R)$. If the former is true, then $GUR_{12}(R)$ is devoid of any direct solute-solute terms and is therefore identical to $\delta A^{HI}(R)$, whereas if the latter is true, then $GUR_{12}(R)$ contains a solute-solute term which has to be subtracted before comparison with $\delta A^{HI}(R)$. Clark et.al. suggest that the solute-solute attractions are not negligible, and by explicitely including such terms in $COR_{12}(R)$, one can calculate $GUR_{12}(R)$ directly comparable to $\delta A^{HI}(R)$. Clark et.al. studied the alcohols, methanol, ethanol, propanol, butanol and t-butanol. They used an atom-atom Lennard-Jones potential as an approximation to $COR_{12}(R)$. They represented $-CH_2-$ groups and

$-CH_3$ groups with effective atoms. Nonpolar molecules were represented by equivalent spheres for the purposes of calculation. A Monte Carlo technique was used to evaluate the orientationally averaged $COR_{12}(R)$ and the techniques suggested by Friedmann and Krishnan were followed to calculate $GUR_{12}(R)$. The Gurney free energy parameter, A_{12}, was calculated by using the experimental virial coefficient data on the alcohols. Clark et.al. also calculate the other related Gurney parameters E_{12},S_{12}, and V_{12} along the lines of Friedmann and Krishnan.

Calculated values for A_{12} by Clark et. al. turned out to be positive and increasing with the size of the solute, in contrast to the negative values obtained by Friedmann and Krishnan. This implies that the removel of a mole of water from cosphere overlap region is free energetically not favored. Therefore the solvent intervening solute pairs can be assumed to be free energetically favored. This result marks the beginning of a new speculation in the field of HI, viz. the SSHI. The resulting $w_{ss}(R)$'s have well developed minima which is in accordance with the accepted view of HI. A comparison of $w_{ss}(R)$ with $COR_{12}(R)$ reveals that the hydrocarbon association is predicted to be more favored in the vapor phase than in water, a conclusion which clearly contradicts the standard view of HI. This may be due to the nature of the Lennard-Jones potential used to represent $COR_{12}(R)$. Clark et.al. compare the calculated $GUR_{12}(R)$ with the corresponding $\delta A^{HI}(R)$

calculated by Ben-Naim[23] and find they differ in sign as well as magnitude.

Clark et.al. use some experimental results to explain an interesting observation made on the Gurney entropy parameter, S_{12}. By 1976, several experimental measurements on effectively hydrophobic solutes were available. Reliable excess volume measurements on these systems by Franks and coworkers showed that the mode of hydrocarbon association is concentration dependent. Approaching infinite dilution concentrations, the hydrocarbon association was found to be different from the standard association processes and remains so for a region of low concentration. After a particular critical concentration is reached, the hydrocarbon association follows the characteristics of standard association processes. Franks suggested that the behaviour in low concentration region may be due to a solvent separated HI. The values of S_{12} calculated by Clark et. al. for the association of nonpolar molecules were found to be less than what is expected from the complete expulsion of a water from cosphere to the bulk solvent, which suggests that the solute association still retains some of the cosphere waters in the overlap region. This is in agreement with the experimental findings of Franks and coworkers. The primitive form of the Gurney free energy used in this model does not allow Clark et.al. to probe any further on this speculation. The calculated $w_{ss}(R)$'s suggest pos-

sible second minima, which again could not be explored further due to the limitations in the model.

Gill and Wadso[39], in 1976, derived empirical relations between experimentally measured thermodynamic properties on aqueous solutions of hydrocarbons and number of hydrogen atoms in the hydrocarbon, n_H. They used enthalpies of solution for hydrocarbons measured in their laboratory, the hydrocarbon solubility data from McAuliffe[31] and calculated the rest of the thermodynamic properties. Near room temperature, the enthalpies of solution for most hydrocarbons were found to be almost zero. Also, the large heat capacity associated with the hydrocarbon dissolution was found to be independent of temperature. Using these observations, Gill and Wadso proposed the following thermodynamic equations,

$$\Delta H^o = \Delta C_p^o (T-T^*)$$

$$\Delta G^o = (T/T^*)\Delta G_T^{o*} - \Delta C_p^o [T\ln(T/T^*) + T^* - T]$$

where T^* is room temperature and ΔG_T^{o*} is the free energy change for the transfer of 1 mol of hydrocarbon from its pure liquid to water at room temperature. The rest of the thermodynamic properties refer to the same transfer process at a temperature T. Gill and Wadso use least square techniques to fit a line for ΔG_T^{o*} vs n_H and derive the relation,

$$\Delta G_T^{o*} = 6.4 + 1.85\ n_H$$

A similar fit for C_p^o vs n_H yields,

$$\Delta C_p^o = 33 n_H$$

A substitution of these relations in the expression for ΔG^{o} yielded the free energy of the system as a function of temperature and n_H. From this expression, the rest of the thermodynamic properties of the system can be easily derived as a function of n_H.

An interesting relation emerges from this analysis for ΔS^{o}, viz., $\Delta S^{o}=-21.5-6.24n_H$ which translates into a decrease of about 1.5 cal deg^{-1} mol^{-1} for each additional hydrogen atom. Gill and Wadso offer an explanation for this based on the following simple arguments. Number of orientations available for a water molecule in a tetrahedral network is 6, if it is fully hydrogen bonded. Replacing one such molecule from the tetrahedral network by a C-H group reduces the available configurations for the neighbouring waters to 3, since a hydrogen from water cannot face the C-H group to form a hydrogen bond. If this is assumed to control the entropy primarily, then ΔS^{o} should be $-R \ln(3/6)= 1.23$ cal deg $^{-1}$ mol^{-1}, in close agreement with the value expected from the empirical relation. Gill and Wadso argue that the large heat capacity of dissolution can be explained by considering the effect of temperature on the number of available configurations for water in the vicinity of hydrocarbon. The emprical relations derived by Gill and Wadso are extremely helpful tools. They provide a good first approximation to the thermodynamics of HI. Given that these relations are consistent with the available experimental results, one can extend these relations to cal-

culate the thermodynamic properties of hydrocarbons of low solubility.

Franks[40], in 1976, used experimental excess volumes of mixing for the four isomeric butanols to establish that the molecular nature of the hydrocarbon association is concentration dependent. This is a followup on Clark et.al.'s work[38], but Franks does not invoke any models to prove his claims. The excess volume of mixing for an aqueous solution of hydrocarbon can be written as,

$$V^E(m) = m(\Phi_V(m) - \nabla^o) = \nu_{xx} m^2 + \nu_{xxx} m^3$$

where Φ_V is the apparent molal volume and ∇^o is the limiting partial molar volume. Here ν_{xx} and ν_{xxx} are the excess volume virial coefficients. Since $\Phi_V(m)$ and ∇^o are experimentally known a linear fit of $(\Phi_V(m) - \nabla^o)/m$ vs m can give ν_{xx} and ν_{xxx} Franks finds that ν_{xxx} values thus calculated are not reliable and the ν_{xx} values for all substances he studied are negative. Using the data from Nemethy and Scheraga[9], Franks estimates that the complete removal of an apolar group like $-CH_3$ from water at infinite dilution involves a positive volume change of .74 cm^3(mol water)$^{-1}$. This means ν_{xx} is positive contrary to the experimental results. Franks therefore suggests that the nonpolar molecular association at low hydrocarbon concentrations is not simply a partial reversal of solution process as suggested by Nemethy and Scheraga. The behaviour of V^E at higher concen-

trations does indicate that associations at those concentrations follow the model suggested by Nemethy and Scheraga. Franks therefore concludes that there is a definite difference between the mode of association at low hydrocarbon concentration involving pairwise HI, and that at higher hydrocarbon concentrations involving multiple aggregation. He speculates that the association at low concentrations of hydrocarbon may involve a solute pair with intervening water molecules. This view by Franks is in a sense an extension of arguments made by Clark et.al, but there is no assumed model for HI in the latter interpretations which indicates SSHI at low hydrocarbon concentrations.

The low hydrocarbon solubilities are often attributed to large unfavorable entropy arising from iceberg formation around the solute. Shinoda[41], in 1976, questioned this view by showing that the enthalpy of iceberg formation is indeed larger than the corresponding entropy effect, but is cancelled by a corresponding large enthalpy of mixing. Shinoda suggests that in the hydrocarbon dissolution process the destruction of hydrogen bonding interactions and the formation of icebergs are the dominating processes. The enthalpy of solution and entropy of solution can be written as,

$$\Delta H = \Delta H(\text{H-bonds}) + \Delta H(\text{iceberg})$$

$$\Delta S = -R \ln x_2 + \Delta S(\text{H-bonds}) + \Delta S(\text{iceberg})$$

where $\Delta H(\text{H-bonds})$ is the enthalpy change due to the des-

truction of H-bonds and is assumed to be the major contributor to the enthalpy of mixing. $-R \ln x_2$ is the entropy of mixing and ΔS(H-bonds) is the entropy change due to the destruction of H-bonds. ΔH(iceberg) and ΔS(iceberg) are the enthalpy and entropy changes due to iceberg formation. Shinoda argues that since around 160 °C there is no iceberg formation, the ΔH(solution) and ΔS(solution) can be written as,

$$\Delta H(\text{solution}) = \Delta H(\text{H-bonds})$$

$$\Delta S(\text{solution}) = \Delta S(\text{H-bonds}) - R \ln x_2$$

Since ΔH(solution), ΔS(solution), R and x_2 are known from experiments, ΔH(H-bonds) and ΔS(H-bonds) can be calculated. Using these values and the temperature dependence of solubility, one can derive ΔH(iceberg) and ΔS(iceberg).

Shinoda uses the technique mentioned above to study the aqueous solutions of benzene, toluene and ethylbenzene. He finds that both ΔH(iceberg) and ΔS(iceberg) are negative and both ΔH(H-bonds) and ΔS(H-bonds) are positive. Shinoda points out that since ΔH(H-bonds) and ΔH(iceberg) cancel each other, we experimentally observe a small ΔH(solution). Then he argues that ΔH(iceberg) and $T\Delta S$(iceberg) also cancel each other, thus, the overall process is controlled by ΔH(H-bonds), or the enthalpy of mixing. To substantiate his arguments, he evaluates the solubilities for hypothetical system of aqueous solutions of hydrocarbons where

there is no iceberg formation. The solubilities thus calculated are found to be lower than the experimental values, suggesting that the iceberg formation increases solubility and does not decrease the solubility as suggested by Nemethy and Scheraga.

Shinoda's arguments can be viewed in the following way.

$$\Delta G(\text{solution}) = \Delta H(\text{solution}) - T\Delta S(\text{solution})$$

$$\Delta H(\text{solution}) = \Delta H(\text{H-bonds}) + \Delta H(\text{iceberg})$$

and $$\Delta S(\text{solution}) = \Delta S(\text{iceberg})$$

Since experimentally $\Delta H(\text{solution})$ is zero and $\Delta S(\text{solution})$ large and negative, one considers the hydrocarbon dissolution process as entropy controlled. What Shinoda claims is, since $\Delta H(\text{iceberg}) = \Delta S(\text{iceberg})$, $\Delta G(\text{solution}) = \Delta H(\text{H-bonds})$ and therefore the hydrocarbon dissolution process is enthalpy controlled. This argument is obviously very sensitive to the nature of the partitioning of $\Delta H(\text{solution})$ into $\Delta H(\text{iceberg})$ and $\Delta H(\text{H-bonds})$. In this model $\Delta H(\text{H-bonds})$ is invariant with respect to temperature and therefore Shinoda's model cannot explain the temperature dependence of $\Delta G(\text{solution})$, which is as much a function of temperature as $\Delta H(\text{iceberg})$.

In 1976, an interesting paper appeared on the effect of solvent-solvent attraction on solute-solute interactions. Ronis, Martina and Deutsch[42] used hypernetted chain(HNC) approximation to study a system of infinitely dilute solution of hard sphere solutes in Yukawa solvent.

Yukawa potential can be described as,

$$U_{11}(r) = \begin{cases} \infty & r/\sigma < 1 \\ \varphi e^{-r/r^*} & r/\sigma \geq 1 \end{cases}$$

where σ is the hard sphere diameter of the solvent and $r^* = r/\sigma$, φ is the strength of attraction between the solvents and is treated as an adjustable parameter. Such simple potential functions are chosen in order to control the HNC approximation methods, which otherwise become too complex and sometimes unsolvable.

The results obtained by Ronis et.al. using such simple methods and systems are extremely interesting. In order to check their methods, they initially set φ to 0, which reduces the system to a mixture of hard spheres. The calculated pmf and pair correlation functions agreed well with a previous calculation by Klein et.al. Ronis et.al. then calculated $w_{ss}(R)$ as a function of φ and also as a function of reduced density of solvent. They found that on increasing the reduced density ρ of solvent, the $w_{ss}(R)$ became oscillatory with well developed second minima. A careful analysis of the results led Ronis et.al. to two important conclusions: a) the first minimum in $w_{ss}(R)$ becomes deeper as φ decreases for a given solvent density. This suggests that the more solvent-solvent attraction is, the larger is the tendency for the solutes to aggregate. b) The second minimum in $w_{ss}(R)$ begins to disappear as is decreased for a given solvent density, suggesting that the increased solvent-solvent attraction tends to

exclude solvents from the region in between the solutes. The solute-solvent pmf shows that increased solvent-solvent attractions progressively increased solute-solvent repulsions.

An interesting explanation is offered by Ronis et.al. for the monotonic behaviour of $w_{ss}(R)$ obtained by Dashevsky and Sarkisov. Their calculations show that the $w_{ss}(R)$ becomes monotonous as the solvent reduced densities are lowered, and suggests that Dashevsky and Sarkisov probably used a low solvent reduced density. Though this is possible, Dashevsky and Sarkisov's[25] method has been shown to be numerically defective as well.

In a 1977 paper, R. Cramer[43] presented the discrepancies between experimental results and the predictions of various theories of HI. His arguments are based on partitioning the hydrocarbon transfer from water to a nonpolar solvent into vapor phase similar to Wolfenden and Lewis. Experimental water ⟶ octanol ΔG values , which is a classic measure of the strength of HI, has an inverse dependence on the molar volume of the solute for nonpolar molecules. The estimated ΔG(water⟶ octanol) for every additional $-CH_2-$ group is -.72 kcal/mol. The partitioning reveals that the ΔG(water ⟶ vapor) has no resemblence at all to the ΔG(water ⟶ octanol) and therefore Cramer argues that it is incorrect to assume that water plays an important role in HI. In addition he observed that the ΔG(water ⟶ vapor)

and the ΔG(octanol ——> vapor) are very similar to each other. This suggested that the solvation of nonpolar solutes by water as well as octanol should be similar, contradicting the predictions of the theories of HI where water is assumed to play a unique role in nonpolar molecule solvation. Cramer also noted that the S of solvation for an increment of $-CH_2-$ group in water is not much different from the values for some other typical nonpolar solvents.

Cramer observes several other features in thermodynamics of the vapor phase equilibria that contradict the predictions of the commonly accepted theory of HI. He observes the ΔG between water and a $-CH_2-$ group to be repulsive and only .18 kcal/mole, whereas it is attractive and .54 kcal/mole between water and a $-CH_2-$ group. This is a quantitative proof that the nonpolar solvent plays a more important role than the water in the HI process. The solubility of rare gases in water was found to increase with the size of the solute, in direct opposition with the HI theories. To add to the list of discrepancies, Cramer notes that the cyclic alkanes are more soluble than a straight chain hydrocarbon of the same molar volume, whereas HI theories predict them to be equally soluble. Cramer concludes by saying that the purpose of his article was to call attention to the discrepancies mentioned above and not to offer any alternate theory for HI. Most of Cramer's arguments are based on the thermodynamics of incremental $-CH_2-$ groups. Though that approach can explain part of the HI it is incomplete all by

itself. It is clear from the data presented by Cramer that the ΔG(water ——> vapor) for nonpolar solutes are twice as large as ΔG(octanol ——> vapor) showing the major role played by water. Cramer's observation about the increased solubility of cyclic alkanes has been explained by Hermann based on the surface volumes.

Pratt and Chandler[44] presented an extensive study of the HI using Weeks-Chandler-Anderson[45] (WCA) theory of dense liquids. WCA theory proposes the intermolecular structure in dense liquids to be dominated by short-range repulsive interactions, since they are the only quickly varying functions of intermolecular distance. Under this assumption the pair correlation function $g_{MM}(R)$ for any liquid is the same as the $g_{MM}(R)$ calculated for a hypothetical fluid with all attractions switched off. Pratt and Chandler use these ideas of WCA theory to model the solute-water and solute-solute interactions. Water does not belong to the class of liquids in the domain of the WCA theory. Pratt and Chandler did not want to use any apriori models for water structure, so they incorporated the experimental $g_{OO}(R)$ to describe the solvent.

In order to calculate the solute-solute properties, Pratt and Chandler assumed that the introduction of hydrocarbon in water is the same as that of a hard sphere. Therefore,

$$y_{SS}(R) = y_{HS}(R)$$

where $y_{HS}(R)$ is the hard sphere cavity distribution func-

tion. The quantity $y_{ss}(R)$ is rich in structural information pertinent to the solutes and is calculated by coupled Ornstein-Zernicke equations, which relate the solute-solute and solute-solvent pair correlation functions to the solvent-solvent g(R). Using the experimental $g_{OO}(R)$ for water from Narten and Levy[46], Pratt and Chandler calculated $g_{ss}(R)$ and $y_{ss}(R)$. The calculated $y_{ss}(R)$ is related to the free energy of solution via a perturbation term containing the attractive part of the solute-water potential which was neglected in the WCA approximation. Assuming the perturbation to be the attractive part of a Lennard-Jones interaction, Pratt and Chandler derived the thermodynamic properties of hydrocarbon solutions which compared well with the experiment.

As stated earlier, $y_{ss}(R)$ is related to the $\delta A^{HI}(\sigma)$ of Ben-Naim[23] by,

$$\delta A^{HI}(\sigma) = -kT \ln y_{12}(\sigma)$$

or $$\delta A^{HI}(\sigma) = -kT \ln y^{1}_{12}(\sigma) - kT \ln y^{o}_{12}(\sigma)$$

where the $y^{1}_{12}(\sigma)$ and $y^{o}_{12}(\sigma)$ are the contributions from the attractive and repulsive solute-solvent interactions respectively. Since the attractive perturbation was assumed to be Lennard-Jones type, and since $y^{o}_{12}(\sigma) = y_{HS}(\sigma)$,

$$\delta A^{HI}(\sigma) = \epsilon - kT \ln y_{HS}(\sigma)$$

where ϵ is the Lennard-Jones parameter. The δA^{HI} calculated by Pratt and Chandler reproduced the inverse temperature dependence of HI due to the kT term in the above

expression, but conclude that the inverse temperature dependence seen for HI does not arise from any microscopic properties of water structure. This contradicts the frequent speculation that the inverse temperature dependence arises from water structure.

The potential of mean force $w_{ss}(R)$ was calculated for hard spheres of various diameters using the $y_{ss}(R)$. Pratt and Chandler noted that all the $w_{ss}(R)$ were oscillatory with a distinct second minimum corresponding to the solvent separated solute pair. They also found that the effective solute interaction increased with increasin solute size, consistent with expectations. However the $w_{ss}(R)$ showed no inverse temperature dependence: Solute association becomes less favorable on increasing temperature. Pratt and Chandler were the first to report oscillatory $w_{ss}(R)$ from calculations on aqueous solutions.

Marcelja, Mitchell, Ninham and Sculley[47], in 1977, used a generalized Pople model[48] for water and Landau theory[49] for free energy density to study the aqueous solutions of hydrocarbons. The Pople model for water assumes liquid water to be an interconnected network of hydrogen bonding waters, and that the hydrogen bonds bend independently with associated energy, $F(\theta)=g\ \mathrm{Cos}\theta$, where g is the hydrogen bond bending constant and θ is the deviation of either the OH bond or the lone pair of a water from the O-O line. Marcelja et. al. modify this model to include

cooperativity and call the modifed model as the generalized Pople model. In the new model, $F(\theta)$ includes the deviations of both the OH bond and the lone pair from the O-O line of neighbouring waters.

Marcelja et. al. calculated the free energy of solvation for hydrocarbons by considering them to be spheres. The excess free energy of solution arising from the introduction of a solute in water is derived using Landau theory and takes the form,

$$\Delta G = a \int [\epsilon^2 + \xi_\epsilon^2 (\nabla \epsilon)^2] dR$$

where ϵ refers to the difference in the order parameters of the first shell and bulk waters, and ξ_ϵ is the correlation length of the decay of perturbation in water structure resulting from the insertion of the solute. Marcelja et. al. use a value of 1.9 Å for ξ_ϵ. Here a is proportional to the free energy density induced at the solute-solvent interface. Marcelja et. al. suggest that by knowing the value of ϵ at the solute-solvent interface, called ϵ_o, the spatial dependence of ϵ can be calculated by minimization of ΔG given above.

The free energy of solvation calculated by Marcelja et. al. using the method above is of the form,

$$\Delta G^{sphere} = -\frac{a\epsilon_o^2}{k} 4 R^2 (1 + -\frac{1}{kR})$$

where k is ξ_ϵ^{-1}. The calculated free enrgy of solvation for methane, ethane and propane agreed well with the experimental values. Marcelja et. al. estimated the free energy

of solvation per $-CH_2-$ group to be 832 cal/mol. Their calculation on hydrophobic hydration suggested that the free energy of solvation not only depends on the surface area of the solute but also on the curvature of the surface area. A careful examination of ΔG^{sphere} shows the nonadditivity of the free energy of solvation for separate groups.

Marcelja et. al. calculated the ΔG^{HI} by calculating the difference in the total free energy of the system when the solutes are at a distance R and when they are at an infinite distance. The free energy thus calculated as a function of R corresponds to $w_{ss}(R)$, and it was found to be monotonic, with the minimum at contact distances of the order of 1.5 kcal/mol. For a methane like sphere, the calculated ΔG^{HI} of 1 kcal/mol corresponds well with with the value of 1.4 kcal/mol reported by Dashevsky and Sarkisov[25]. This work by Marcelja et. al. showed the nonadditivity of free energies of solvation but failed to produce an oscillatory $w_{ss}(R)$.

Increased access to computers during the mid 70's made large scale computer simulations of liquids and solutions possible. Both Monte Carlo(MC) and Molecular dynamics(MD) techniques were used to study the aqueous solutions of hydrocarbons in detail. Owicki and Scheraga[49b] in 1977 reported MC computer simulation results on liquid water and $[CH_4]_{aq}$ in the (T,P,N) ensemble. The general details of MC computer simulation techniques and related topics are discussed in detail in the methodology section. The system

consisted of one methane molecule in 100 waters at 25°C and 1 atm. pressure. Face centered periodic boundary conditions were used to simulate a condensed phase environment. Throughout the simulation methane was kept fixed at the origin of the cubic cell and was not perturbed. Preferential sampling, also discussed in the methodology section, was used to move water molecules close to the methane more often than the bulk waters. Owicki and Scheraga used the quantum mechanical CI potential developed for water by Matsuoka et. al.[51] to describe the water-water pair interactions. They used a subset of the ab-initio quantum mechanical energies for methane-water calculated by Ungemach and Schaefer and fitted them to a 12-6-1 form. The empirical potential thus derived is used to describe methane-water pair interactions in the simulation. Several hundreds of simulation steps were discarded for equilibration and the last 674K steps (1K=1000 steps) were used in taking ensemble averages of various properties.

Owicki and Scheraga calculated some thermodynamic properties and structural indices for $[CH_4]_{aq}$ from the simulation results. Partial molar energy for $[CH_4]_{aq}$ was calculated to be -11 ± 15 kcal/mol which compared fairly well with the experimental value -2.6 kcal/mole; large errors on the calculated partial molar energy arises from the fact that it is the difference between two large quantities. The calculated partial molar volume was 25 ± 34 cm^3/mol^{-1}, which also compared fairly well with the experimental value

of 37 cm^3/mol^{-1}. Owicki and Scheraga presented various atom-atom pair correlation functions, $g_{AB}(R)$, where A refers to an atom in the solute and B refers to the atom in water. $g_{CO}(R)$ has a distinct first peak at 4.0 Å followed by a rather flat minimum from 5.5 to 6.0 Åand a second peak at 7 Å. Integrating $g_{CO}(R)$ upto 6 Å suggests that the first hydration shell of methane contains about 23 water molecules, consistent with the clathrates of size 20 and 24 known for methane. The peaks in $g_{CO}(R)$ and $g_{CH}(R)$ indicated a preference for water hydrogens to be closer to methane than water oxygens. Water-water pair energy distributions showed that the waters in the first hydration shell of methane interacted .68 kcal/mole stronger than the bulk waters. This calculation illustrates the power of computer simulation techniques in yielding the structural details of solutions. The standard errors on the calculated pair correlation functions are of the order of 10% and are fairly reliable. The functions are extremely hard to obtain experimentally due to the low solubilities of hydrocarbons in water and the limited applicability of diffraction experiments. Thus computer experiments can be used to calculate them and interpret the structure of first hydration shell of nonpolar molecules.

Chan, Mitchell, Ninham and Pailthorpe[52] in 1978, applied integral equation techniques to study the dependence of $w_{ss}(R)$ on solute size and on solute-solvent interactions. The system used in their study consisted of an infi-

nted various

(R), where A

s to the atom in

t 4.0 Å fol-

.0 Åand a sec-

Å suggests

ontains about 23

rates of size 20

n $g_{CO}(R)$ and

ydrogens to be

ater-water pair

s in the first

al/mole stronger

illustrates the

in yielding the

rd errors on the

of the order of

s are extremely

solubilities of

capability of dif-

ents can be used

re of first hyd-

e[52] in 1978,

the dependence

lvent interac-

sted of an infi-

nitely dilute

R_1 in sticky

potential u_{ij}

by,

$$\exp(-U_{ij}(R)$$

where $R_{ij}=\frac{1}{2}(R$

the strength

al. calculat

ick(PY) and

They found t

and $f_{22} \rightarrow \infty$

tory as the s

the solute-so

as the solve

et. al. al

$w_{ss}(R)$ becom

However the s

ficant. Thi

into the e

$w_{ss}(R)$. The

the $w_{ss}(R)$

realistic pot

analyses prov

not be used t

ity of the po

nitely dilute solution of sticky sphere solutes of radius R_1 in sticky sphere solvents of radius R_2. The pair potential $u_{ij}(R)$ between any two species i and j is given by,

$$\exp(-U_{ij}(R)/kt)-1= \begin{cases} -1+R_{ij}/12\tau_{ij} \quad (r-R_{ij}) & r<=R_{ij} \\ 0 & r>R_{ij} \end{cases}$$

where $R_{ij}=\frac{1}{2}(R_i+R_j)$ and τ_{ij} is a measure of the strength of adhesion between species i and j. Chan et. al. calculated $w_{ss}(R)$ for this system using Percus-Yevick(PY) and hypernetted chain(HNC) approximation methods. They found that for hard sphere solutes,(when both τ_{12} and $\tau_{22} \longrightarrow \infty$) $w_{ss}(R)$ is monotonic but becomes oscillatory as the solute-solvent attraction increases. In addition the solute-solute attractions were found to become stronger as the solvent-solvent attractions became stronger. Chan et. al. also found that for a given solvent size the $w_{ss}(R)$ becomes oscillatory on decreasing the solute size. However the solute-size dependence of $w_{ss}(R)$ is not significant. This interesting analysis provides more insight into the effect of solute-solvent potentials on the $w_{ss}(R)$. The results from this study can be used to test the $w_{ss}(R)$ calculated by other methods involving more realistic potential functions. As mentioned before, these analyses provide valuable informations on $w_{ss}(R)$ but cannot be used to model an aqueous solution due to the complexity of the potential functions involved to describe water.

In 1978 Swaminathan, Harrison and Beveridge[53] presented the results of a (T,V,N) ensemble Monte Carlo computer simulation on $[CH_4]_{aq}$ at 25°C. The simulation involved one methane in 124 waters in a simple cubic box and the condensed phase environment was modelled by periodic boundary conditions. They used MCY-CI potential to model water-water pair interactions and three different potentials, referred to as MO,MP and HM, to describe methane-water interactions. The MO potential was derived by fitting 225 ab-initio quantum mechanical energy points to a 12-3-1 empirical form; MP potential was essentially the same as MO, but electron correlation effects were used in the calculation of the quantum mechanical energies. HM is the MO potential with the attractive part set to zero everywhere. Three simulations were carried out using the three methane-water potential functions. The length of the simulation ranged between 575K for MO potential to 900K for HM potential.

The structural properties calculated in this simulation closely resembled the results of Owicki and Scheraga. The partial molar internal energy for methane calculated by Swaminathan et. al. is in the range of -15 to -26 kcal/mol, considerably different from the experimental value of -2.6 kcal/mol. The qualitative features of the methane-water pair correlation function $g_{MW}(R)$ calculated in the three simulations appeared the same with slight differences in the amplitudes. Swaminathan et. al. calculated the coordination number for methane to be 19.35 and a struct-ural

analysis showed that the first hydration shell resembled a pentagonal dodecahedron. This is slightly different from the coordination number of 23 suggested by Owicki and Scheraga, which corresponds to a tetrakaidecahedron, but the ideas emerging from these independent studies are similar. A comparison of the results from the simulations using MO and MP potentials suggested that the correlation effects in the methane-water interactions shifted the coordination number distribution towards higher values and the distribution of solute binding energies to lower values. Similar comparison between the the results of simulations using the MO and HM potentials showed that the soft part of the methane-water interactions shifts the distribution of coordination number by about 15% higher and the distribution of binding energies by 3.2 kcal/mol lower. This work by Swaminathan et. al. added further evidence that computer simulationis an excellent tool for studying solutionsand it also showed that the effect of solute-water potential on structural properties such as $g_{MW}(R)$ is within the statistical uncertainities.

Pangali, Rao and Berne[54], in 1978 reported the mean force $\langle F_Z(R) \rangle$ between two xenons in water along the line connecting the centers of mass of the xenons. The mean force between the solutes is the derivative of $w_{ss}(R)$ with respect to R. Their technique involved performing several Molecular dynamics simulations keeping the solutes fixed at various distances. In principle, a large number of such

simulations should give a continuous $\langle F_Z(R) \rangle$. However Pangali et. al. calculated $\langle F_Z(R) \rangle$ for only five intersolute separations ranging from 5.35 Å to 9.18 Å. Their system consisted of two Lennard-Jones spheres of σ=3.43 Å dissolved in 214 ST2 waters. The time step for the MD computer simulation was of the order of 10^{-16}sec and a total of 5000 time steps were carried out for each of the five simulations. The first 2000 time steps were discarded as equilibration and the last 3000K were used to calculate $\langle F_Z(R) \rangle$ between the solutes. Pangali et. al. reported the $\langle F_Z(R) \rangle$ thus calculated along with the direct solute-solute contribution and the indirect contributions from the solvent effect. The error bars on the $\langle F_Z(R) \rangle$ were rather large to make any quantitative conclusions. However, the mean force showed a definite oscillation, suggesting that at xenon-xenon separations greater than contact distance, a water molecule is likely to get in between the xenons. This work by Pangali et. al. added support to the possibility of SSHI.

The first application of Molecular Dynamic(MD) computer simulation technique to the study of aqueous solutions appeared in 1979. Geiger, Rahman and Stillinger[55] studied a system of two neon-like Lennard-Jones atoms in 214 water molecules using MD computer simulation techniques. In principle, the presence of two solute atoms makes it possible to study the hydrophobic hydration and HI. However, due to the methodological difficulties discussed in detail in Chapter

III, the latter cannot be explored in detail with the methodology by Geiger et.al. They presented an in-depth analysis of the hydrophobic hydration of neon, but could only discuss the HI between neon atoms on a qualitative level.

A brief discussion of a Molecular Dynamics technique can be found in Chapter III. Geiger et. al. used ST2 potential developed by Stillinger[56] to describe the water-water pair interactions. Both neon-water and neon-neon were taken as Lennard-Jones interactions. All the pair interactions beyond the distance of 7.Ø5 Å were treated insignificant for efficiency of computing. The time step used for numerical integration of the dynamical equations is $2.1X10^{-16}$ sec. A total of 35ØØØ time steps were carried out, of which the initial 1ØØØØ were discarded as equilibration. In a MD simulation the temperature fluctuates and the average temperature for this system was calculated to be 32.3°C.

Some interesting observations were made with regard to the neon pair motion as the simulation progressed. During the first 2.5 picoseconds, the two neon atoms remained at contact. Then they began to separate and setteled down in individual solvent cages with a layer of water in between them. A detailed analysis showed that one tight bond between two waters persisted for a rather long time and was situated between the neon pair. The structural analysis using pair correlation functions and orientational correlation functions showed that this bond forms a common edge of the

two adjacent hydration cages. This is an extremely important observation since it supports the possibility of solvent-separated HI between the neons. Geiger et. al. could not quantitatively describe any solute-solute properties due to insufficient statistics on this aspect in their calculation.

Analysis of the hydrophobic hydration of neons revealed intricate details of the structure of the first hydration shell of neon. Geiger et. al. calculated that the first hydration shell of neon contains about 14 water molecules. Orientational distributions were used to show that the first shell itself contains two subshells. The first subshell consists of 8 waters orienting in way that only one of the water hydrogens points directly away from the neon. The second subshell has the remaining 6 waters orienting in such a way that both their hydrogens point directly away from the solute. A comparison of $g_{OO}(R)$ and energetics between the first shell waters and the bulk water confirmed that the first shell waters are more structured than the bulk. The calculated heat capacity difference between the first shell waters and bulk is 68.6 cal/deg.mol, which compared fairly well with the experimental values of 30 to 65 cal/deg.mol for noble gases in water. Geiger et. al. used the results of the simulation to calculate some dynamic properties. The self diffusion coefficients indicated that the first shell waters diffuse about 20% slower than the bulk. The reorientational correlations for shell waters were found to be lar-

ger larger than for bulk factor by a factor of 1.2 to 1.7.

This important calculation provided further insight into the nature of nonpolar solute hydration. The recognition of various orientational preferences by first shell waters is new and interesting. Though this calculation could not conclude that SSHI is favored over CHI, it certainly points in that direction.

Tucker and Christian[57] presented very reliable Henry Law constants for aqueous solution of benzene in 1979. They found that the aqueous solution of benzene showed a definite negative deviation from Henry Law and attributed this deviation to associated benzene dimers. The formation constants calculated for benzene under this assumption is 47 and the free energy of dimer formation was derived to be -99 cal/mol. After converting the standard states to correspond to mol fraction scale, ΔG^{HI} for benzene is -2.3 kcal/mol., considerably less negative than δA^{HI} for benzene calculated by Ben-Naim[24].

This study is highly reliable due to the techniques used, and demonstrated the association of benzene in water. The formation constants derived for benzene dimer are very useful, and it can be related to $w_{ss}(R)$ through the second virial coefficient. The ΔG^{HI} calculated by Tucker and Christian cannot be directly compared with the δA^{HI} of Ben-Naim since ΔG^{HI} includes direct solute-solute interactions whereas δA^{HI} includes only the indirect contribu-

tions, and also that the choice of biphenyl as the benzene dimer in the calculation of δA^{HI} is not applicable.

In 1979 Elkoshi and Ben-Naim[58] extended the one dimensional model for water proposed by Bell[59] to study the HI. Bell's model exhibits the well known maximum in the temperature dependence of density for water. This is known to arise from peculiar packing nature of Bell's model. Elkoshi and Ben-Naim suggests that a similar model developed for HI may explain the inverse temperature dependence of HI.

Bell's model for water is a one dimensional lattice of equidistantly placed sites. Each site may be occupied by a "water" or else empty. Two waters that are second neighbours can only form hydrogen bonds. This set up leads to three possible states for each site: a) the site may be occupied by a water. b) It may be empty and situated between two hydrogen bonding waters. c) It may be empty but not flanked by hydrogen bonding waters. There are nine different possibilities for each adjacent pair of sites and the interaction energy for all the nine possibilities can be represented by a 3X3 matrix, called the A matrix. The grand partition function is related to the largest eigenvalue of the A matrix in the thermodynamic limit. It is also a function of the hydrogen bond energy of water. For reasonable hydrogen bond energies this model produces thermodynamic properties in agreement with experiment.

Elkoshi and Ben-Naim extended this model to allow "solutes" to occupy the empty sites in the lattice. This introduces two more possible states for each site: a) the site may be occupied by a solute or b) the solute may occupy a site inbetween two hydrogen bonding waters. Grand partition function for this system interms of A matrix is written under the assumption that the solute cannot occupy a site in between the hydrogen bonding water. It is used to derive the cavity distribution function for the solutes, $y_{ss}(R)$. $y_{ss}(R)$ is related to the $w_{ss}(R)$ by,

$$w_{ss}(R) = -kT \ln y_{ss}(R) - U_{12}(R)$$

where $U_{12}(R)$ is the direct solute-solute pair potential. In this model, the distances are discrete and $y_{12}(1)$ corresponds to the contact solute pair.

Elkoshi and Ben-Naim's analysis of $y_{ss}(R)$ revealed several interesting results. At modest pressures and temperatures, $y_{12}(1)$ is larger for water than for normal solvents. In this model, water refers to a solvent with maximum hydrogen bonding energy and a normal solvent refers to a solvent with no hydrogen bonding. He also found that the solute-solute association increased as a result of increasing the solvent-solvent interaction, a result in agreement with the calulations of Ronis et. al.[42] and Chan et. al.[52] Elkoshi and Ben-Naim calculated the temperature dependence of $y_{ss}(R)$ in the limit of moderate temperatures. He concluded that the water-water pair energy ϵ should be at least twice more attractive than the solute-

solvent pair energy η for the system to exhibit inverse temperature dependence. Since this condition is satisfied by hydrocarbons in general,this model explains the inverse temperature dependence of the HI. There is no such region of energetics for normal solvents that show inverse temperature dependence. Since $-kT \ln y_{ss}(R)$ measures the indirect contributions from the solvent effects to the HI, δA^{HI}, δH^{HI} and δS^{HI} can be calculated. δA^{HI} was found to be $-\epsilon+2\eta$, and since this is <0 in general for hydrocarbons, δA^{HI} is also negative. Since this result refers to $y_{12}(1)$ and a contact solute pair, these result confirm the free energy favored formation of solute contact pairs in water. Elkoshi and Ben-Naim thus used a very simple model to explain the HI, and showed that the inverse temperature dependence of HI results from the enthalpic term, viz. the solvent interaction energy compared to the solute-solvent interaction energy.

Okazaki, Nakanishi, Touhara and Adachi[60] in 1979 published results of a (T,V,N) ensemble Metropolis Monte Carlo simulations on dilute aqueous solutions of methane and isobutane. They also used the umbrella sampling technique to evaluate the excess free energies of the solutions over pure water. Okazaki et. al. used a system of 63 waters interacting via ST2 pair potential and one solute. The solute is Lennard-Jones sphere of radius corresponding to either meth-

ane or isobutane. A total of 400K to 500K configurations were sampled of which the first 130K to 150K were discarded. The weighting function used in the umbrella sampling procedure was derived by trial and error method and was a function of the total energy of the system.

The hydrophobic hydration presented by Okazaki et. al. agreed well with the previous theoretical predictions. The g(R) for both the solutes showed well developed peaks suggesting shell structures. The $g_{OO}(R)$ calculated for just the shell waters was compared with the bulk water $g_{OO}(R)$ to show increased structuration near the solute. The Helmholtz free energy change between $[CH_4]_{aq}$ and $[H_2O]_l$ was calculated to be 3.2 kcal/mol. Considering the internal energy difference between the two to be the enthalpy difference Okazaki et.al. calcuated the excess entropy for $[CH_4]_{aq}$ over $[H_2O]_l$ to be 300 cal/mol.

This study is a rather short Monte Carlo simulation and the results are subjected to fairly large error bars. Secondly, the number of water molecules used in this study is rather small for a solute like isobutane and may result in the boundary condition influence on results. The weighting function used by them takes values in the range of 1 to 10^{64}, which may not be sufficient to give a correct free energy. The numerical results of this study must be viewed with caution.

In 1979, Swaminathan and Beveridge[61] reported the mean force between two methanes in water, along the line joining the centers of the solutes. Their method was exactly the same as that used by Pangali et. al. except that Swaminathan et. al. used Monte Carlo simulation. A total of six Monte Carlo simulations were performed with the methanes kept fixed at distances ranging from 3.5 to 7.2 ^{o}A. The details of the potential functions etc. related to the simulation is similar to their calculation on $[CH_4]_{aq}$. Ideally, the average force along the lines perpendicular to the line joining the centers of the solutes should be zero. Due to the errors in the simulation methodology they were not found to be zero, and they are used to estimate the errors on $\langle F_Z(R) \rangle$. The function $\langle F_Z(R) \rangle$ showed oscillation and suggested a possible SSHI. The errors in the calculated $\langle F_Z(R) \rangle$ were too high to make any quantitative predictions.

In 1979, Pangali, Rao and Berne[62] reported the $w_{ss}(R)$ between two argon atoms in water using a modified Monte Carlo method. Since this work mainly concerned itself with the calculation of $w_{ss}(R)$, the Monte Carlo method had to be modified. The commonly used Metropolis Monte Carlo method is expected to produce statistically insignificant $w_{ss}(R)$ due to problems discussed in great depth in Chapter III. Pangali et.al. therefore augmented the Metropolis Monte Carlo method with the Umbrella sampling algorithm, also discussed in Chapter III. They presented the hydrophobic hydra-

tion of argon resulting from this study in an adjoining paper.

The details of the Pangali et.al. calculations are as follows. Their system consisted of two argons in 214 water molecules. The temperature of the simulation was 298 ^{o}K. The pair interactions between the water molecules were modeled using the ST2 potential. Both the argon-water and argon-argon interactions were treated as Lennard-Jones interactions. All the pair interactions were truncated at an intermolecular distance of 8.46 $\mathring{A}$. Force bias technique was used to aid faster convergence of the simulation results. Umbrella sampling is used to modify the standard Boltzmann factor for the sampling. The new Boltzmann factor contained an additional harmonic potential $U_H(R_{12})$ between the two solute atoms, and a correction for this modification is done in the ensemble averaging process. $U_H(R_{12})$ restricts the motion of the solutes to a small value around a preselected equilibrium intersolute separation, R_{12}^{o}. Pangali et. al. performed four simulations, each referred to as a window, at R_{12}^{o}values of 3.88,5.33,6.Ø8 and 6.6 $\mathring{A}$. In each of these simulations the solutes moved about 1.4 $\mathring{A}$ about the corresponding R_{12}^{o}values. A total of 1188K configurations were sampled in each window and the initial 1Ø8K was discarded for equilibration. The solute-solute pair correlation function, $g_{ss}(R)$, calculated in each window were matched to produce a single $g_{ss}(R)$. Then, using the relation

$w_{ss}(R) = -kT \ln g_{ss}(R)$, $w_{ss}(R)$ was derived.

Very interesting results emerged from this simulation. In order to test the validity of the method, Pangali et. al. calculated the $w_{ss}(R)$ between two "marked" argon atoms in liquid argon. This was compared with the $w_{ss}(R)$ calculated for argons from a long simulation by Kalos et. al.[63] where $w_{ss}(R)$ was averaged over all pairs of argons. The comparison showed that the umbrella sampling produced a $w_{ss}(R)$ that was within 4% of the Kalos et.al. $w_{ss}(R)$. The $w_{ss}(R)$ calculated by Pangali et. al. is oscillatory. Due to normalization problems, this is shifted from the true value by an unknown constant. Pangali et. al. corrected their $w_{ss}(R)$ by moving the first minimum in $w_{ss}(R)$ to correspond to the first minimum in the $w_{ss}(R)$ calculated for an equivalent hard sphere pair using Pratt and Chandler's method described above. The resulting $w_{ss}(R)$ has two distinct minima, the first at the solute pair contact distance and the second at a distance which corresponds to the solvent separated solute pair. The calculated barrier between the contact and solvent separated pairs is of the order of kT. Pangali et. al. also estimated the statistical weights of the contact and solvent separated pairs and concluded that the latter is favored by a factor of 3.

This study by Pangali et. al. presented further evidence for the importance of SSHI in understanding the HI. The $w_{ss}(R)$ calculated here still has considerable error asso-

ciated with it, but the conclusions reached about the preference of SSHI over CHI are within the statistical error limits. The small difference between $w_{ss}(R)$ for two marked argons and the all-argon-pair averaged $w_{ss}(R)$ is rather surprising. In a latter study Mehrotra et. al.[64] calculated such marked water pair correlation functions and showed that for most water pairs the g(R) was significantly different from that of the all-water-pair averaged g(R). They also noted that there were a few pairs for which the two g(R)'s were significantly similar. In view of this we suspect that the comparison given by Pangali et, al, is only a coincidence, and thus the convergence is not as good as originally expected.

Pangali et. al. present the analysis of hydrophobic hydration of argons for all four windows used in the above study. The argon-water pair correlation function, $g_{AO}(R)$, was studied by partitioning the space around each solute into two regions: a) an exterior region defined to be the hemisphere around a solute that points away from the other solute, and b) an interior region defined to be the hemisphere which points towards the other solute. There were no noticeable changes in g_{AO}^{int} and g_{AO}^{ext} except in window 1. In window 1 the solutes are at contact and as a result the interior region is drastically different from the exterior. Pangali et. al. calculated the coordination number to be about 20 for argon consistent with a dodecahedral clathrate. Water-water binding energy distributions showed

that the distribution for the first shell waters is shifted towards more attractive energies than the bulk water distribution. Pair energy distribution between various pairs of water molecules, viz. shell-shell, shell-bulk and bulk-bulk, revealed that the interactions between the shell waters is about 2 kcal/mol stronger than the same for bulk waters. This analysis further supports the iceberg model for hydrophobic hydration.

1980-1983.

In 1980, Pratt and Chandler[65] extended their earlier study of HI to investigate the effects of solute-water attractive forces on HI. In their previous study, such attractive forces were neglected according to the WCA theory. The inclusion of solute-solute and solute-solvent attractive forces demand modified integral equations. These were taken from the EXP theory of Anderson et. al. Pratt and Chandler recalculated the $w_{ss}(R)$ for Lennard-Jones solutes used by Pangali et. al. and found a drastic change in the $w_{ss}(R)$ from their previous results. They found the second minimum in the $w_{ss}(R)$ to be lower than the first minimum. They offer an explanation for this in terms of solute-solvent and solute-solute interaction energies, viz. ϵ_{sw} and ϵ_{ss}. In going from contact dimer to solvent separated dimer, two new solute-solvent interactions are introduced at the expense of one solute-solute interactions, assuming the rest of the first shell remains roughly the same. Therefore this amounts to an energy change of $2\epsilon_{sw}-\epsilon_{ss}$. Since this quantity in their calculation is negative, the SSHI is energetically favored. They also found that the solute-water properties are not affected significantly due to the inclusion of the attractive forces. In this work Pratt and Chandler showed the extreme sensitivity of $w_{ss}(R)$ towards the choice of solute-solvent and solute-solute interaction. Their explanation for the calculated $w_{ss}(R)$ is in terms of energetics and and thus of

enthalpic origin.

Rossky and Friedmann[66] used the Gurney model to study the HI between the benzene molecules in water. The application of Gurney model to the HI is discussed earlier in this section. The method basically involves calculating the Gurney free energy parameter, A_{12}, for an assumed benzene-benzene interaction potential. Using the Henry law constants reported by Tucker and Christian for aquous solutions of benzene, Rossky and Friedmann calculated A_{12} parameter for benzene association. If the benzene-benzene interaction is assumed to be exponential, A_{12} is calculated to be -175 cal/mol and if it is assumed to be of the r^{-9} form, then the A_{12} value comes out to be -200 cal/mol. Rossky and Friedmann noted that the A_{12} values calculated for other hydrophobic groups of the same size as benzene are very close to the values derived in this study. Therefore, they concluded that A_{12} values depend mostly on the size of the hydrophobic group and are thus transferable. Rossky and Friedmann added that there is no special force involved in the benzene association process and benzene association occurs due to the same forces that bring about any other nonpolar solute association. They clarify that the Gurney model proposed by them cannot be used to model an oscillatory $w_{ss}(R)$ and that it does not allow any distinction between the SSHI and CHI. The reason is that the solvent effects on the solute interaction responsible for the oscillatory $w_{ss}(R)$ is presented as in an average manner through A_{12}.

The important conclusion from this work of Rossky and Friedmann is that the A_{12} parameters are transferable. This transferability allows the study of various hydrocarbons that are otherwise not possible due to low solubilities of hydrocarbons. Their hypothesis that benzene association is the same as any hydrocarbon association is surprising, since the polarizability of the pi-cloud of benzene not present in some hydrocarbons may play an important role in the association.

In 1981, Bolis and Clementi[67] presented the results of a Monte Carlo simulation on $[CH_4]_{aq}$ at 300°C. The system for this study consists of one methane and 202 waters. The quantum mechanical CI potential developed by Matsuoka et. al.[51] is used to model the water-water interactions. Two new methane-water empirical potential functions were developed by Bolis and Clementi by fitting ab-initio quantum mechanical energies to assumed functional forms. The first function referred to as a four term(4T) potential is of 12-6-4-1 form and the second, called three term (3T) potential, is of 12-6-1 form. All through the simulation methane was kept fixed at the center of the spherical cell. The Monte Carlo simulation was augmented with the force-bias technique to accelerate the convergence. A total of 5800K and 3400K configurations were generated using 3T and 4T potentials respectively. The methane-water structural properties were in good agreement with the previously reported simulation results of Owicki and Scheraga and

Swaminathan et.al. However, there were considerable differences in the solute-water energetics. Bolis and Clemeti noted that even after 3000 to 400K run, the solute-water pair energy did not converge. The calculated enthalpy of transfer is -6.41±2.99 kcal/mol which is closer to the experimental value of -2.6 kcal/mol than any of the previous results. Bolis and Clementi attribute this to the increased length of the MC run. This study by Clementi demonstrated a known fact that the solute-water properties converge much slower than solvent-solvent properties due to the reduced statistics. They also demonstrated that the structural properties converge faster than the energetic properties. It is interesting to note that, despite an increase in the length of the MC run by a factor of 7 to 10, the energetics improved very little, whereas the structural properties were same compared to the short runs.

The $w_{ss}(R)$ calculated using the umbrella sampling techniques as in Pangali et. al.'s calculations, does not exactly correspond to infinitely dilute solution and is shifted by a constant due to normalization effects. Saul Goldmann[68], in 1981 proposed a method to calculate the absolute pmf from computer simulations. He suggested that several simulations must be performed at varying concentrations of solutes and the potential of mean force calculated for each concentration can then be used to calculate the $w_{ss}(R)$ at infinite dilution by extrapolation.

Goldmann performed five Monte Carlo simulations on a system containing 108 particles. In one simulation there was no solute and the system corresponded to pure solvent. In other simulations, 8, 11, 16 and 22 solutes were placed in the system. The water-water interactions were modeled by a generalized Stockmeyer potential and both solute-water and solute-solute interactions were Lennard-Jones interactions. The simulation length varied from 3000K to 500K. During the course of the simulations the solutes did not show any increased tendency to go apart as noted in the simulation by Geiger et. al. Goldmann attributes this to the increased number of solutes in his study. The calculated $w_{ss}(R)$ for various solute concentrations were found to be smooth and the errors were found to increase at large R values. The relation between $g_{ss}(R)$ and $w_{ss}(R)$ at finite solute concentrations is given by,

$$w_{ss}(R) = -kt \ln g_{ss}(R) + kt \ln\left(1 - \frac{1}{N_2}\right)$$

where N_2 is the number of solute particles. According to Goldmann the entire $w_{ss}(R)$ can be plotted as a function of concentration and then an extrapolation should yield the $w_{ss}(R)$ for infinite dilution. However, the errors on the $w_{ss}(R)$ values beyond the first minima are prohibitively large for such an extrapolation procedure. So he uses only the $w_{ss}(R)$ values at the first minima. The extrapolated value for the system he studied was of the order of .6kT. This calculation showed a method for evaluating the absolute pmf. The linear extrapolation used in this study needs a

careful look. Usually the relation between free energies and concentrations are logarthmic rather than linear and in view of this a logarthimic extrapolation may be proper.

In a 1982 paper, Rapaport and Scheraga[69] presented the results of a very long Molecular Dynamics simulation study on a system of four nonpolar solutes in 329 water molecules. MCY-CI potential was used to model water-water pair interactions and a shifted and truncated Lennard-Jones potential to describe both the solute-water and solute-solute interactions. This Lennard-Jones potential has the form,

$$V(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^{6} + \tfrac{1}{4}] \qquad r \leqslant r_c$$

$$= \emptyset \qquad r > r_c$$

where $r_c = 2^{1/6}\sigma$. Rapaport and Scheraga carried out two simulations for solutes of different size, viz. σ=3.1 Å and 3.3 Å. The time step for solving the dynamical equations was about 10^{-15} sec and the duration of each simulation was about 70 psec. They reported that such a long simulation was made possible by high speed array processors.

Rapaport and Scheraga analysed the solute-water structural properties by partitioning the region around each solute along the lines of Pangali et. al. The solute-water g(R) agreed qualitatively well with those previously reported from simulations on aqueous solutions of nonpolar molecules. The presence of more solutes and the increased length of the simulation aided in the convergence of solute-water properties and the errors on these properties

were of the order of 5%, smaller than all previous simulation studies. A detailed analysis of the first shell agreed with the clathrate picture for hydrophobic hydration emerging from other studies with about 15 waters in the first shell and the hydrogens of these waters pointing away from the solute.

Rapaport and Scheraga found that the solutes were highly mobile and showed no tendency to aggreagate. They found only one of the six pairs approached the contact distance but remained at contact for an insignificant time. The solutes were found to be separated by water molecules during most of the simulation time. This is interpreted as the SSHI. Though the simulation was long, the statistics was not good enough to compute solute-solute properties such as $w_{ss}(R)$. Therefore no quantitative conclusions about the SSHI could be made from this study. Rapaport and Scheraga also calculated the self diffusion coefficients for solute, first shell waters and bulk waters. They found considerable exchange between the waters in the first shell and bulk. The calculated self diffusion coefficients were corrected for this exchange and showed no statistically significant difference between the motions of the first shell and bulk waters. This study by Rapaport and Scheraga showed no tendency for the nonpolar solutes to aggregate in water. It also pointed out that a very long simulation may not be enough by itself to study the solute-solute $w_{ss}(R)$. One needs special sampling techniques to study such quantities.

CHAPTER III.

Theory and Methodology.

A theoretical description of liquid state properties and compositional characteristics of a system at a temperature T and density N/V (number of particles N per unit volume V) follows from the semi-classical canonical ensemble partition function

$$Q(T,V,N) = (q^N/(8\pi^2)\,\Lambda^{3N}\,N!)\;Z(T,V,N) \quad (1)$$

$$Z(T,V,N) = \int\cdots\int \exp[-E(X^N)/kT]\,dX^N \;. \quad (2)$$

Here q is the partition function for internal degrees of freedom and Λ is the one-dimensional translational partition function for each particle. The quantity $E(X^N)$ is the configurational energy of the system, discussed later in this section under intermolecular interactions. The integration ranges over the configurational coordinates X^N of the N particles of the system. The formalism is developed here in the context of the (T,V,N) ensemble. Parallel developments can be given for the (T,P,N) and (T,V,μ) ensembles[70,71].

The thermodynamical properties of the system follow from the statistical thermodynamic definition of the Helmholtz free energy

$$A = -kT\ \ln\ Q(T,V,N) \quad (3)$$

and its derivatives; the thermodynamic internal energy much discussed herein is given by $-T(\delta A/\delta T)_{V,N}$. On substitution, differentiation and management of terms, can be partitioned into "ideal" and "excess" contributions, with

the latter expressed as the configurational average of $E(X^N)$,

$$U = \int \ldots \int E(X^N)\ P(X^N)\ dX^N = \langle E(X^N) \rangle \qquad (4)$$

where $P(X^N)$ is the Boltzmann probability for the system to be found in configuration X^N,

$$P(X^N) = \exp[-E(X^N)/kT]/Z(T,V,N) \quad . \qquad (5)$$

Other thermodynamic properties can be expressed in an analogous manner. Of particular interest here will be the constant volume excess heat capacity,

$$C_v = (\langle E(X^N)^2 \rangle - \langle E(X^N) \rangle^2)/kT^2 \qquad (6)$$

pressure,

$$P=(kT/V)\ (N-\langle \sum_{i=1}^{N} (R_i .\ E(X^N)/\ R_i) \rangle) \qquad (7)$$

and atom-atom spatial distribution functions

$$g_{\alpha\beta}(R) = N_{\alpha\beta}(R)/(\rho 4\pi R^2 \Delta R) \qquad (8)$$

where R is the interatomic separation, $N_{\alpha\beta}(R)$ is the average number of β neighbours of an α atom in a spherical shell between R and $R+\Delta R$ and ρ is the bulk density.

An alternative expression for the Helmholtz excess free energy advantageous for computer simulation is

$$A = \int_0^1 U(\xi)\ d\xi \quad . \qquad (9)$$

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

average

$$U(\xi) = \int \cdots \int E(X^N) P(X^N/\xi) dX^N \qquad (10)$$

where $P(X^N/\xi)$ is the probability of observing the system in configuration X^N, conditional upon the value of the parameter ,

$$P(X^N/\xi) = \exp[-E(X^N/\xi)/kT] / \int \cdots \int \exp[-E(X^N/\xi)/kT] dX^N. \qquad (11)$$

When the system is coupled by the auxiliary parameter through the expression

$$E(X^N/\xi) = \xi E(X^N) \qquad (12)$$

the free energy is defined with respect to an ideal gas reference state of liquid density[72]. Mezei, following Torrie and Valleau[73], has studied the use of other reference state for tractability and computational efficiency[74], but methods for calculating free energy in computer simulation remains the subject of active research[75] rather than application. With the internal energy and free energy available, the excess entropy can be simply obtained from the expression

$$S = (A-U)/T \quad . \qquad (13)$$

The additional equilibrium properties of the system accessible to calculation in the (T,P,N) ensemble simulation are the constant-pressure heat capacity:

$$C_p = (\langle H^2\rangle - \langle H\rangle^2)/kT^2 \qquad (14)$$

isothermal compressibility:

$$K = (\langle V^2\rangle - \langle V\rangle^2)/(kT^2\langle V\rangle) \qquad (15)$$

and coefficient of thermal expansion:

$$\alpha = (\langle VH\rangle - \langle V\rangle\langle H\rangle)/(kT^2\langle V\rangle) \quad , \qquad (16)$$

where H is the enthalpy of the system.

The theoretical study of the equilibrium properties and diffusionally averaged structure of a fluid can be approached by either Monte Carlo or molecular dynamics. Molecular dynamics gives dynamical as well as equilibrium properties of the system but the calculations are usually done in the microcanonical ensemble, with the determination of temperature an empirical problem. Anderson[76] has recently devised an approach for extending molecular dynamics to other ensembles. The Monte Carlo method is defined on convenient statistical thermodynamic ensembles and allows for various extensions necessary for the study of solvent effects on inter- and intramolecular interactions and for the calculation of the free energy, but does not yield access to dynamical properties. Results on equilibrium properties of liquid water from Monte Carlo and molecular dynamics have been compared and found to agree closely[77,78,79]. The efficiency of the two methods in computing equilibrium properties was also compared and was

found to be comparable[78]. Thus for equilibrium properties and structure, either method of study is acceptable. The research described in this dissertation incorporates the Monte Carlo method.

Monte Carlo computer simulation is based on the simultaneous numerical integration of integrands analogous to the internal energy expression, Eq.4. This integral is well-known to be ill-conditioned for direct numerical integration, since the integrand $E(X^N)\ P(X^N)$ is large in only a very restricted region of configuration space. The integration can be successfully carried out as suggested by Metropolis et. al.[80] by sampling from the Boltzmann distribution in order to automatically concentrate the effort in the important regions of configuration space. In the Metropolis method, the N-molecule configurations of the problem are taken to be the states of an irreducible Markov chain. The one-step transitions between any two states k and l of the chain have the probability

$$p_{kl} = Pr[X^N_{t+1} = l \mid X^N_t = k] \qquad (17)$$

The p_{kl} can be collected in array form as a stochastic transition probability matrix, $P=[p_{kl}]$. Integration in the Metropolis method is carried out by means of a stochastic walk through configuration space generating a realization of an irreducible Markov chain whose unique limiting stationary distribution π is the Boltzmann distribution, i.e.

$$\pi = \pi P \qquad (18)$$

where $\pi_i = P(X_i^N)$. In a realization of this process, the configurations X^N are then sampled with a frequency proportional to $P(X^N)$, and the determination of average properties reduces to a simple summation over the energy of the individual configurations X^N:

$$E(X^N) = (1/M) \sum_{t=1}^{M} E(X_t^N) \qquad (19)$$

Provided the system is ergodic, as $M \longrightarrow \infty$, $E(X^N) \dashrightarrow \langle E(X^N) \rangle$, the cumulative average energy becomes an increasingly good estimator of the energy expectation value. The computation of heat capacity and other configurational properties of the system take analogous form.

To implement this procedure in practice, a particle is selected, a trial move is attempted from configuration X_k^N to X_l^N. The change in energy of the configurational energy of the system is used to calculate the quantity

$$R = (q_{lk}/q_{kl}) \exp[-(E(X_l^N) - E(X_k^N))/kT] \qquad (20)$$

If R is greater or equal to 1, the move is accepted. If R is less than one, then the move is accepted with the probabilty R and rejected with the probability 1 - R. To do this, one compares R with a random number uniformly distributed on the interval (0,1). If the R is greater than the

random number the move is accepted, otherwise the move is rejected. Repeated application of this process forms a sequence of configurations that are a realization of the desired Markov process, and any configurational average property of the system may in principle be calculated by averaging over these configurations. Optimum sampling for a property other than energy may require sampling from a modified Boltzmann or even non-Boltzmann distribution[73,81].

The convergence and statistical error bounds of Metropolis Monte Carlo calculations are generally monitored according to the method of block averages (also known as the method of batch means)[82,83]. Here the Monte Carlo realization is partitioned into several non-overlapping blocks of equal lengths, and the averages of the property under consideration (e.g. mean energy) are computed over each block. Let $\bar{f}_i$ denote the average of the property f computed over the block i. Under the assumptions that the f_i's are independent and normally distributed and that the Markov chain is ergodic, the error bounds for the property f at a 95% confidence level are 2σ, where

$$\sigma^2 = (1/B(B-1)) \sum_{i=1}^{B} [\bar{f}_i^2 - (\bar{f}_i)^2] \qquad (21)$$

and the summation runs over the B blocks. In computer simulations of small lengths, the above assumptions are honored more in the breach than in observance, and thus computed error bounds by the method of batch means are to be taken

with caution. To ensure the validity of the estimate by the batch means method, the block size has to be increased until reliable statistical tests show that the batch means are indeed independent. Other methods have also been proposed to estimate the confidence intervals of Monte Carlo estimates of this type. Good reviews can be found in Refs. 84-86.

The details of the Metropolis method and subsequent elaborations thereof can be specified in the following general notation: The elements of the one-step transition probability matrix of the Markov chain, p_{kl}, are written as a product of two terms,

$$p_{kl} = q_{kl} \alpha_{kl}$$

The first factor q_{kl} is dependent on the method of generating the state l from the state k in a single step transition. The second factor α depends on the way in which state l is accepted to ensure that the microscopic reversibilty conditions are satisfied. The Metropolis choice,

$$\alpha_{kl} = \min(1, P(X_l^N)\, q_{lk} / P(X_k^N)\, q_{kl}) \qquad (22)$$

has been shown by Peskun[87] to be asymptotically optimum. The elements of the one-step transition probability matrix for the Markov chain can be rewritten into a more popular notation,

$$p_{kl} = q_{kl} \min(1, p_l q_{lk} / p_k q_{kl}),\ k \neq l \qquad (23)$$

and

$$p_{kk} = 1 - \sum_{k \neq l} p_{kl} \qquad (24)$$

The various sampling methods discussed herein differ essentially in the definition of q_{kl}, i.e the way in which state l is generated from state k in a single step transition. In principle, all the sampling schemes allow for more than 1-particle moves. However, in practice, for convergence efficencies, the moves are restricted to a single particle. Thus, the configurational coordinates of the state l are related to the configurational coordinates of state k,

$$X_l^N = X_k^N + \delta^N \qquad (25)$$

where

$$\delta^N = \{ \emptyset, \emptyset, \ldots \delta(X_m), \emptyset, \ldots \} \qquad (26)$$

and $\delta(X_m)$ is a displacement vector for the molecule m selected for the move. For rigid polyatomic molecules,

$$\delta(X_m) = \{ \delta x_{cm}, \delta y_{cm}, \delta z_{cm}, \delta w, \eta \} \qquad (27)$$

where δx_{cm}, δy_{cm}, δz_{cm} are the displacements for the center of mass and ow is the rotation around a chosen axis η, passing through the center of mass of the molecule m. The magnitudes for the center of mass displacement and for the rotation angle are further restricted by certain step-

size parameters r and w , which are optimized in the initial stages of the simulation. In Metropolis sampling, the components of the displacement vector $\delta(X_m)$ are obtained by uniformly sampling from the domain D, centered at the coordinates of the molecule m in the state k, and defined by the step-size parameters r and w. The elements q_{kl} of the transition probability matrix Q are then

$$q_{kl} = \text{constant} \tag{29}$$

$$q_{kl} = \emptyset \tag{30}$$

Moreover, Q is a symmetric matrix .

In a typical Monte Carlo computer simulation on a molecular liquid in the (T,V,N) ensemble, the system consists of a simulation cell containing N molecules in a volume V determined by N/ρ, where ρ is the experimental density at the system temperature T. The configurational energy of the system is computed by means of analytical potential functions. The system is presented with a condensed phase environment by means of periodic boundary conditions, with the central cell surrounded at each face, edge and vertex by a self-image. Calculations from this Laboratory use mainly simple cubic or face centered cubic boundary conditions. To reduce the effect of the periodic images, most calculations include only interactions between the nearest images of each pair (minimum-image cutoff). Quite often, an additional cutoff criterion is applied to the nearest pair to decrease the computational effort (spherical cutoff). Calculations

on liquid water reported from this Laboratory used the spherical cutoff criterion. The initial segment of the calculation is an equilibration phase, and is discarded in the formation of ensemble averages.

Several methods have recently been proposed for acceleration of convergence in Monte Carlo calculations, with development and testing carried out on the liquid water system. These procedures involve appending additional importance sampling criterion to the Metropolis method. The principal procedures currently under consideration are the force-bias procedure developed for liquid water by Pangali, Rao and Berne[78] and an alternative force-bias scheme based on Brownian dynamics proposed by Rossky, Doll and Friedman[88]. In the force-bias sampling, the particle moves are biased in the direction of forces and torques on the molecule selected for the move. The elements q of the transition probability matrix for force-bias sampling are given by the expressions

$$q_{kl} = N(X_k^N) \exp\{ (F_m(X_k^N).\delta r + N_m(X_k^N).\delta w)/kt\} \qquad (31)$$
$$\delta(X_m) \in D$$

and

$$q_{kl}=\emptyset \qquad \delta(X_m) \notin D \ . \qquad (32)$$

Here $\delta r=\{\delta x_{cm},\delta y_{cm},\delta z_{cm}\}$, $\delta w = \eta \delta w$, $N(X_k^N)$ is a normalization constant and η is a parameter to be optimized during the sampling. The quantities $F_m(X_k^N)$ and

$N_m(X_k^N)$ are the forces and torques in the state k on the particle m to be moved. Note that $q_{kl} \neq q_{lk}$. In our force-bias calculations, is set equal to Ø.5, following Rossky et. al.[88].

Many of the interesting properties of the dilute aqueous solutions are primarily determined by the solute-solvent and solvent-solvent interactions near the solute. This suggests that the computational efficiency for the convergence of many solution properties could be expedited with sampling concentrated primarily in the neighbourhood of the solute. Clearly the solvent molecules far away from the solute also should be sampled, but not necessarily as frequently as those near the solute. This strategy is the basis for the method of preferential sampling first proposed by Owicki and Scheraga[5Ø] for Monte Carlo simulations on $[CH_4]_{aq}$. Owicki[89] subsequently presented a generalized version of the preferential sampling methodology. The two preferential sampling schemes have recently been applied by Bigot and Jorgensen[9Ø] to study the conformational equilibria of a n-butane in CCl_4 liquid.

Owicki's sampling scheme can be described as follows. Let w(R) be a weighting function which is nonnegative and decreasing function of R, where R is dependent on solute-solvent distances. For each N-particle configuration X_k^N define a probability distribution function $W(X_k^N)$,

$$W(X_k^N)=\{\ W_2(X_k^N), W_3(X_k^N) \ldots . W_N(X_k^N)\}, \qquad (33)$$

where the subscript refers to the molecule number and the solute is labelled as the molecule number 1. Furthermore,

$$W_m(X_k^N)= w(R_m)/ \sum_{i=1}^{N} w(R_i) \qquad (34)$$

In preferential sampling, the solvent molecules are selected for the moves by sampling from the probability distribution $W(X_k^N)$. If m is the solvent molecule thus selected in a single-particle move, then the elements of q_{kl} are,

$$q_{kl}=\text{const.}* W_m(X_k^N). \qquad (35)$$

Clearly, q_{kl} is not equal to q_{lk}, and

$$q_{lk}=\text{const.}* W_m(X_l^N). \qquad (36)$$

The selection of the solute for the move follows a different treatment. The preferential sampling scheme, in principle, alows that the solute can be picked with an arbitrary frequency, thus,

$$q_{kl}=q_{lk}= \text{constant}. \qquad (37)$$

In the limits, the solute is either never moved or is always selected to move. As Owicki points out, there are definite statistical advantages if the solute is also moved, because a solute move leads to perturbations in N-1 solute-solvent

pairs, whereas a solvent move causes only one solute-solvent perturbation. On the other hand, if the solute moves are attempted too frequently there is little room for the solvent molecules to properly relax. Clearly, a compromised approach is desired. Instead of carrying out a complete optimization requiring several long test runs, the solute molecule, in present studies, is perturbed with the same frequency as its neighbouring solvent molecules.

The preferential sampling can be combined with the Metropolis sampling as well as with the force bias sampling. Then the elements q_{kl} are the product of the preferential sampling q_{kl} with the Metropolis or the force bias q_{kl} as the case may be. For the simple weighting function used in our studies, the computational overhead due to the preferential sampling is negligible. In all the studies presented in this dissertation, we use a weighting function $1/R$, where R is the distance between the solute and water. Mehrotra et. al.[64] studied liquid water with and without convergence acceleration procedures and found that the latter increases the convergence by a factor of 3 to 4. Their method of attack consisted of marking a water in liquid water as solute and comparing the calculated g(R) for the marked water with the g(R) calculated by averaging over all the water molecules. The g(R) for marked water differs considerably from the total g(R) for liquid water if Metropolis method alone is used, and the differences diminished on adding the force-bias and preferential sampling techniques.

This study also included results from the $[CH_4]_{aq}$ described in the next chapter.

Application of Monte Carlo computer simulation to molecular associations in solution presents a special problem. The Metropolis method samples configurations according to Boltzmann distribution and thus the sampling of intersolute coordinate is likely to be very restricted. Patey and Valleau[81] encountered this problem in the Monte Carlo simulation of an ion pair in water and suggested a non-Boltzmann sampling scheme, referred to as umbrella sampling, to study the association process in solution.

The details of umbrella sampling are as follows. Ensemble average of any system property x in a (T,V,N) ensemble may be written as,

$$\langle x \rangle = \frac{\int \cdots \int x(X^N)\ P(X^N)\ dX^N}{\int \cdots \int P(X^N)\ dX^N} \qquad (38)$$

as a direct extension of Eq.4. Patey and Valleau suggested that the Eq.38 can be modified as,

$$\langle x \rangle = \frac{\int \cdots \int (x(X^N)/w(r_{12}))\ w(r_{12})\ P(X^N)\ dX^N}{\int \cdots \int (1/w(r_{12}))\ w(r_{12})\ P(X^N)\ dX^N} \qquad (39)$$

or,

$$\langle x \rangle = \frac{\int \cdots \int (x(X^N)/w(r_{12}))\ P'(X^N)\ dX^N}{\int \cdots \int (1/w(r_{12}))\ P'(X^N)\ dX^N} \qquad (40)$$

so that, finally,

$$\langle x \rangle = \frac{\langle x/w \rangle_{P'}}{\langle 1/w \rangle_{P'}} \qquad (41)$$

where $\langle\rangle_{P'}$ refers to averaging according to the modified probability distribution $P'=w(r_{12})*P(X^N)$. Here r_{12} refers to the intersolute distance in configuration X^N and w is called the "weighting function". The weighting function is adjusted so that any desired range of r_{12} is sampled in a given simulation.

Theoretically, it is possible to find a $w(r_{12})$ such that the entire range of r_{12} of interest is sampled in a given simulation, but this is nearly impossible in practice. Patey and Valleau therefore suggested performing several independent simulations, each referred to as a window, such that different $w(r_{12})$ are used to sample overlapping regions of r_{12}. The system properties calculated for a given r_{12} from two overlapping windows are not comparable because of the differences in normalization with respect to r_{12}. For eg., the potential of mean force for the solutes on a given window in the simulation can be given by

$$w_{ss}(R) = -kT \ln g_{ss}(R) + C \qquad (42)$$

where $g_{ss}(R)$ is the solute-solute pair correlation function calculated from the simulation and C is a constant

arising as a result of unknown concentration of the system. This additive constant C is indeterminate and differs from window to window in an unknown manner. By requiring that the values of any system property at a given r_{12} must be the same, independent of the window, the system properties covering the full range of r_{12} can be generated. This procedure is called matching. Since matching generates properties covering the full range of r_{12}, like the covering by umbrella, this is called umbrella sampling. The two important aspects of umbrella sampling, weighting function and matching are discussed below.

The implementation of umbrella sampling procedure is simple. The computational scheme remains the same as Metropolis method with or without the convergence acceleration procedures. However, the limiting probability distribution used for sampling is now changed from Boltzmann distribution to $w(r_{12})*P(X^N)$. Patey and Valleau used umbrella sampling procedures to calculate a reliable $w_{ss}(R)$ for ions in water.

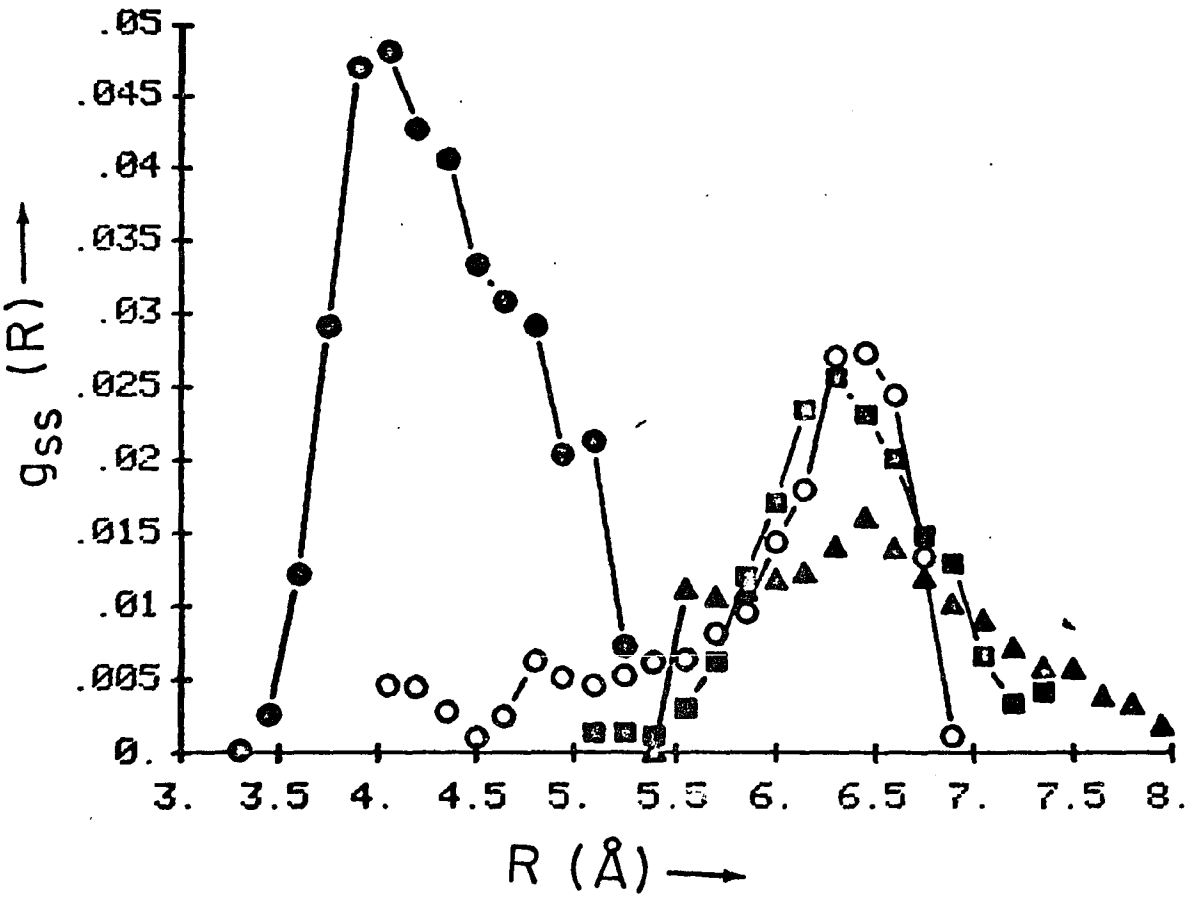
Pangali et. al. used a simple harmonic constraining potential as the weighting function in the umbrella sampling procedure used to calculate $w_{ss}(R)$ between two apolar solutes in water. The form of the function $w(r_{12})$ used by them is,

$$w(r_{12}) = 1/2\ k_o\ (r_{12} - r_{12}^o)^2 \qquad (43)$$

where k_o is the force constant and r_{12}^o is the equilibrium intersolute separation r_{12} around which the solutes are restricted to move in the simulation. A given r_{12}^o and k_o defines a window. A large positive value for k_o would restrict the solutes to move only a small range around r_{12}^o whereas a small positive k_o would allow the solutes to move rather freely. A value of 1.5 kcal/$^oA^2$ for nonpolar solutes was found satisfactory for our study. The quantities r_{12}^o are chosen depending on the region of r_{12} values of interest. Typically two windows, one with r_{12}^o= solute and another with $r_{12}^o = \sigma_{solute} + \sigma_{solvent}$ are necessary. One needs additional windows in between these two values for matching. In this study the simulations involving umbrella sampling used three to four windows.

The details of matching are as follows. In a typical umbrella sampling procedure, the solute-solute properties are of most interest. Therefore here we describe the matching procedure for solute-solute properties. In Fig. 1, the $g_{ss}(R)$ calculated for different overlapping windows for two methanes in water is presented. Harmonic constraining potentials centered at 4.Ø, 5.3, 6.Ø7 and 6.8^oA and k_o=1.5 kcal/$^oA^2$ define the four windows. $g^i(R)$ where i=1 to 4 refers to $g_{ss}(R)$ calculated for window i. It is evident from Fig. 1 that the amplitudes as well as the shapes of $g^i(R)$ in the overlapping regions are different, as expected, due to normalization problems. Matching bet-

Figure 1- Calculated $g_{ss}(R)$ for $[(CH_4)_2]_{aq}$ plotted against intersolute separation R at 25°C. ●, $g^1(R)$; o, $g^2(R)$; ■, $g^3(R)$, and ▲, $g^4(R)$.


.05
.045
.04
.035
.03
.025
.02
.015
.01
.005
0.
g_{ss} (R) ⟶
3.
3.5
4.
4.5
5.
5.5
6.
6.5
7.
7.5
8.
R (Å) ⟶

ween two adjacent $g^i(R)$ can in principle be done at any R value in the overlapping region. However the tail region of $g^i(R)$ have large errors due to poor sampling and therefore those regions should be avoided for matching. Pangali et. al. suggest the points that have reasonable error bars in the overlapping regions are the best points for matching. We adopt their suggestion in all our calculations. One can easily check to see if any other significant matchings change the conclusions.

Matching between the two overlapping windows involves selecting a point R_M in the overlapping region such that the points to the right of R_m in $g^i(R)$ and points to the left of R_M in $g^{i+1}(R)$ are discarded as insignificant and $g^{i+1}(R)$ is scaled by $g^i(R_M)/g^{i+1}(R_M)$ for $R \geq R_M$. In words, this scaling simply means that the points to the right of R_M in g^{i+1} are scaled so that at R_M, the pair correlation functions from both the windows have the same value. This matching reduces the number of curves from 4 to 3, and the result is given in Fig. 2. The point at which the matching was done is also marked. Similar matching is carried out between the $g^{i'}(R)$ and $g^{i+2}(R)$ and $g^{i''}(R)$ and $g^{i+3}(R)$, where $g^{i'}(R)$ and $g^{i''}(R)$ are the results of first and second matching. Curves resulting from second and third matching are given in Figs. 3 and 4. Since each matching propagates error through the scaling, one should be careful about choosing the correct points for matching to ensure minimal error in the final $g_{ss}(R)$. The example

Figure 2- The set of pair correlation functions resulting after the first matching between $g^1(R)$ and $g^2(R)$. ●, $g^{1'}(R)$; ■, $g^3(R)$; and ▲, $g^4(R)$.

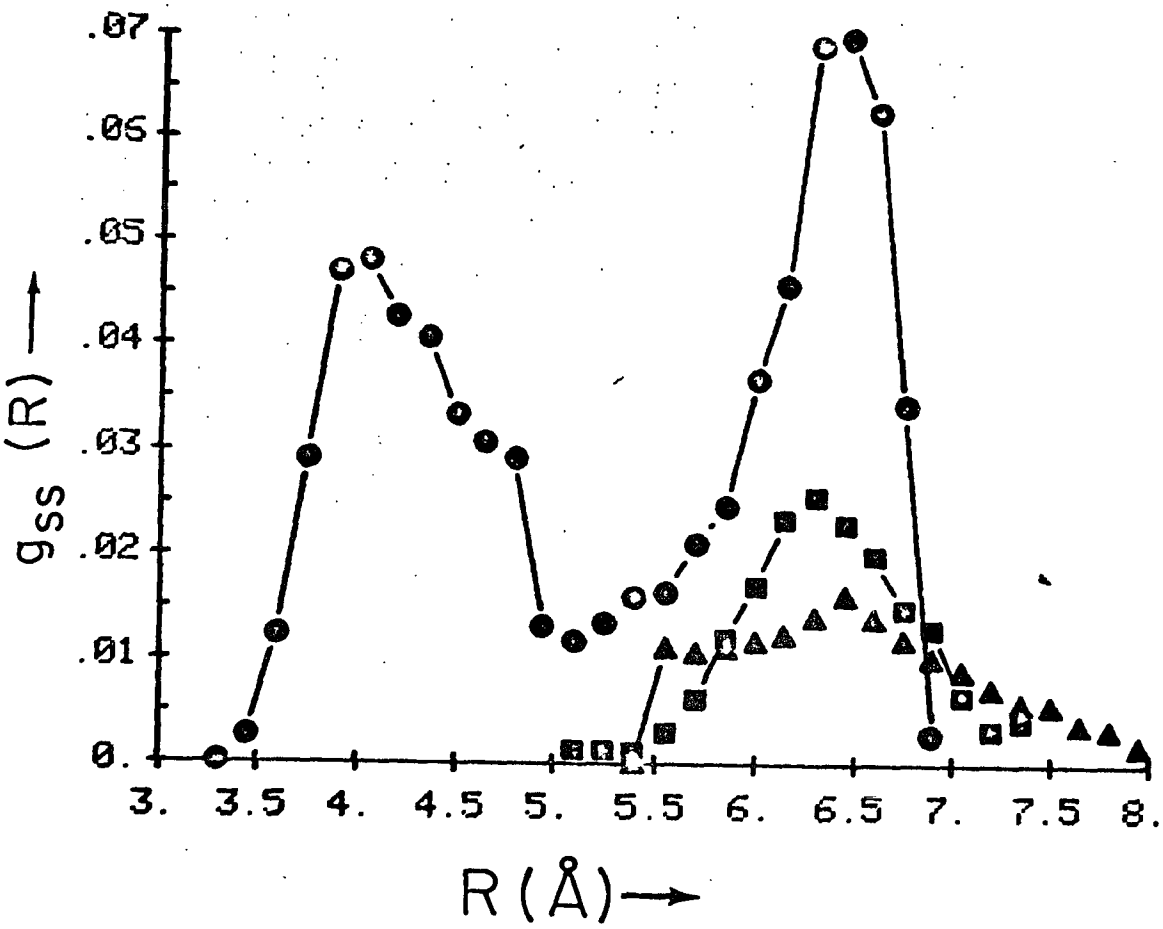

.07
.06
.05
.04
.03
.02
.01
0.
g_{ss} (R) ⟶
3.
3.5
4.
4.5
5.
5.5
6.
6.5
7.
7.5
8.
R (Å) ⟶

Figure 3- The pair correlation functions after the matching between $g^{1'}(R)$ and $g^{3}(R)$. ●, $g^{1''}(R)$ and▲, $g^{4}(R)$.

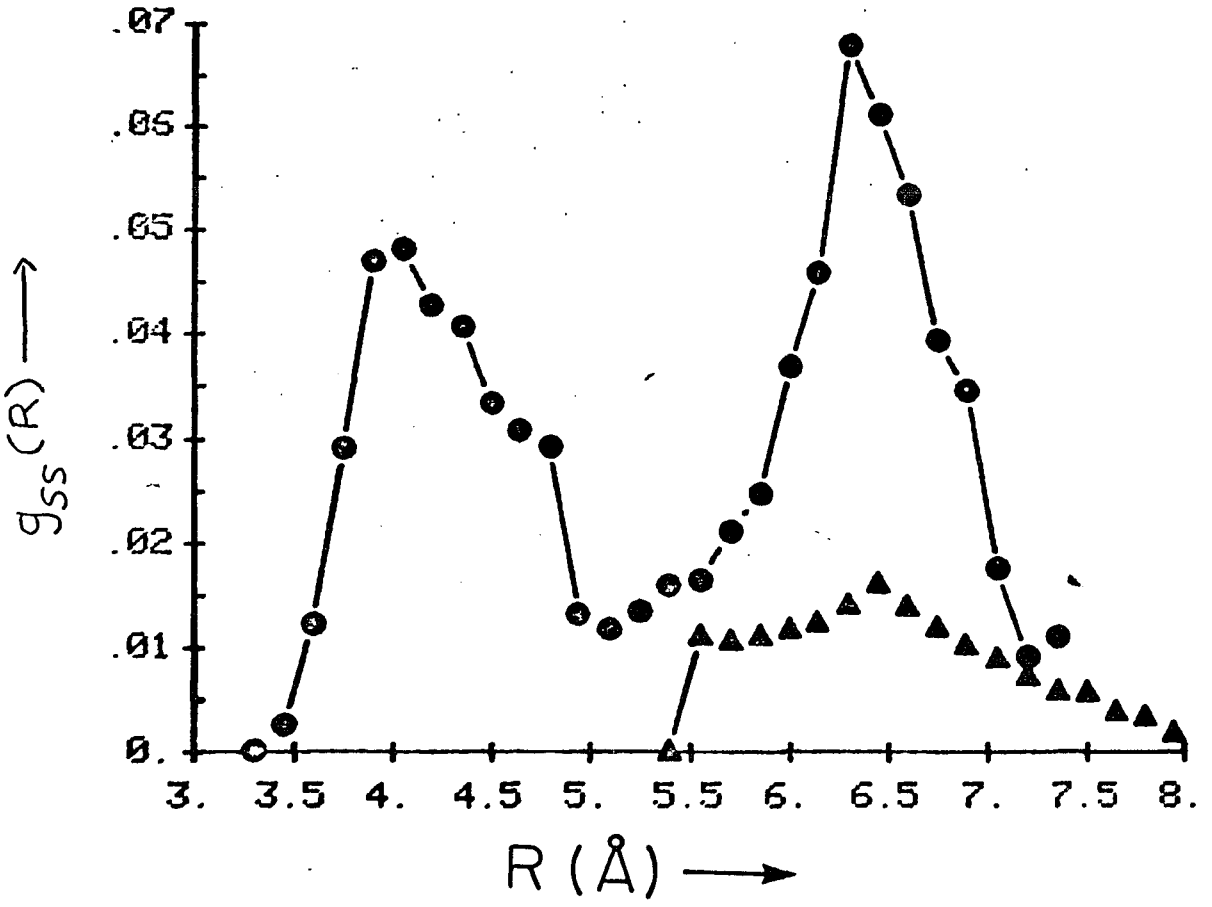
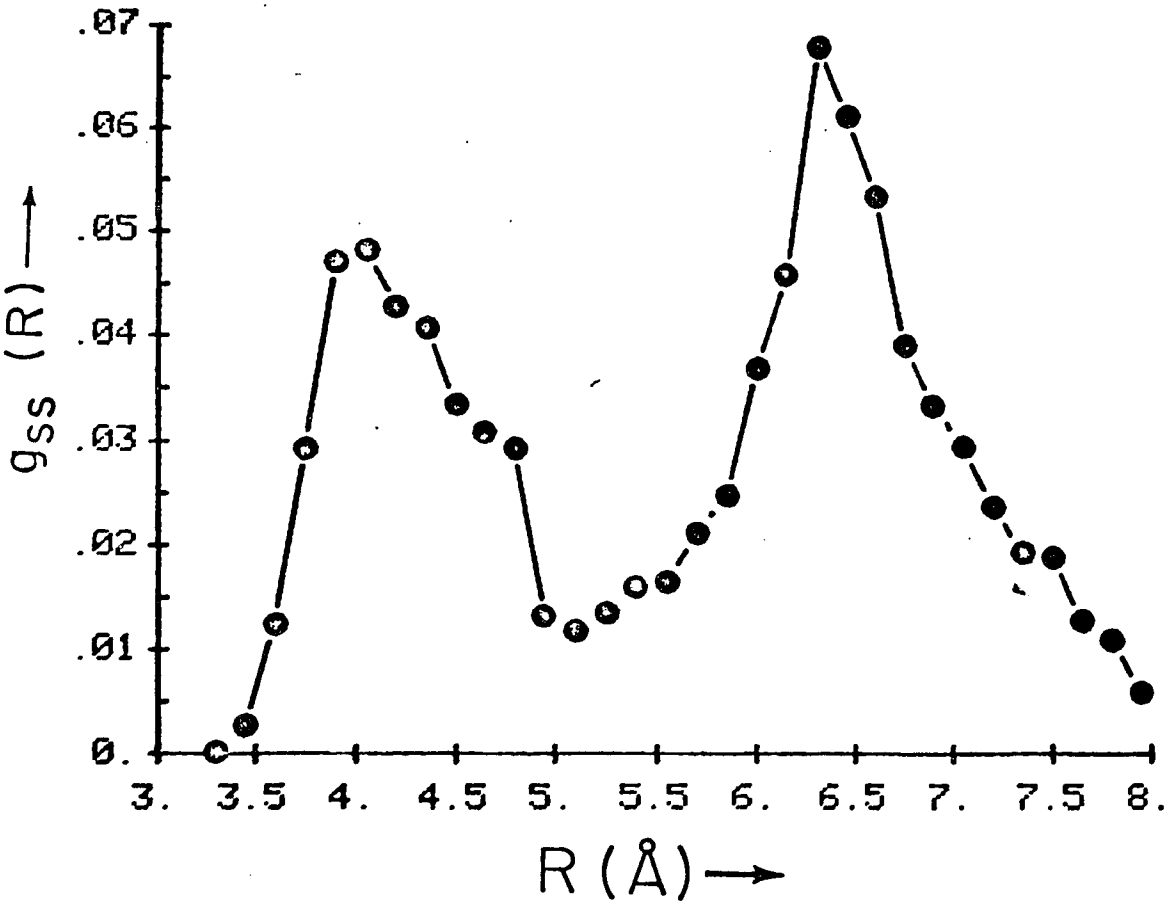

.07
.06
.05
.04
.03
.02
.01
0.
$g_{ss}(R)$ ⟶
3.
3.5
4.
4.5
5.
5.5
6.
6.5
7.
7.5
8.
R (Å) ⟶

Figure 4- Pair correlation function after the final matching between $g^{1''}(R)$ and $g^{4}(R)$.


.07
.06
.05
.04
.03
.02
.01
0.
g_{ss} (R) ⟶
3.
3.5
4.
4.5
5.
5.5
6.
6.5
7.
7.5
8.
R (Å) ⟶

given here carried out matching from left to right. Matching can also be done from right to left as in the case of Pangali et. al. A left to right matching propagates errors to second and larger minima in $w_{ss}(R)$ and a right to left matching propagates errors to first minimum in $w_{ss}(R)$.

In this study two types of simulations were carried out. The first involves one solute in water, simulating hydrophobic hydration, and the second involves a pair of solutes in water, simulating the hydrophobic interaction. All hydrophobic hydration studies involved lengthy Monte Carlo computer simulations using Metropolis method augmented with convergence acceleration techniques. Hydrophobic interaction studies involve the calculation of $w_{ss}(R)$ and used the umbrella sampling scheme of Patey and Valleau with the harmonic weighting functions procedure similar to that suggested by Pangali et. al.

The remaining aspect of the calculations to be specified is the choice of intermolecular potential functions. The individual characteristics of a molecular system are introduced into a computer simulation via the configurational potential energy. To formally define this quantity, let us specify an N-particle configuration of a molecular assembly by the configurational coordinate vector X ,

$$X^N = \{X_1, X_2, X_3, \ldots, X_N\} \qquad (44)$$

where each X is a product of position R and orientation Ω_i

$$X_i = \{R_i, \Omega_i\} \qquad (45)$$

The configurational potential energy in this notation is $E(X^N)$, and represents the energy of an N-molecule system in configuration X^N relative to the energy of N isolated molecules. The configurational energy may be expanded in terms of successive orders of interaction as

$$E(X^N) = \sum_{i<j}^{N} E\,(X_i,X_j) + \sum_{i<j<k}^{N} E\,(X_i,X_j,X_k) + \ldots\ldots (46)$$

where E is the energy of dimerization,

$$E_2(X_i,X_j) = E(X_i,X_j) - 2E(X_i) \qquad (47)$$

and E_3 is a correction term for three-body effects,

$$E_3(X_i,X_j,X_k) = E(X_i,X_j,X_k) - 3E(X_i) - [E_2(X_i,X_j)+E_2(X_i,X_k),+E_2(X_j,X_k)] \qquad (48)$$

Analogous terms E_4, E_5, etc. can be developed to represent even higher order effects. The terms E_n for $n>2$ introduce cooperative effects into the configurational potential. When all these terms are neglected, the configurational energy is expressed as a sum of interaction energies for molecular pairs, an assumption referred to as "pairwise additivity".

Computer simulation requires rapid evaluation of the configurational energy for a large number of N-molecule complexions of the system. This task is accomplished by means of potential energy functions, simple analytical expressions for interaction energy as a function of configurational coordinates and a set of disposable parameters. Potential

functions for intermolecular interaction energies can be grouped for purposes of discussion into three classes: model, empirical, and quantum mechanical. Typical model potential functions are the hard sphere and Lennard-Jones potentials, both studied extensively in the formal development of liquid state theory. Empirical potentials result when the disposable parameters of a function are selected based on experimental data. In quantum mechanical potentials, disposable parameters are determined on a best fit criterion from a discrete data base of quantum mechanically calculated interaction energies.

Both empirical and quantum mechanical approaches have been used for the determination of potential functions describing the interaction energy of water molecules, and there are advantages and disadvantages to both. The construction of functions based on experimental data has the decided advantage of building all possible observed information about the intermolecular interactions into the function. In addition, the experimental nature of the data partially compensates for the assumption of pairwise additivity in the functional form, leading to so-called "effective" pair potentials which include higher order effects in some averaged form. There are extensive experimental data to draw upon in this approach such as electric moments, vibrational frequencies, lattice constants, etc. However, the available experimental information corresponds only to certain limited regions of intermolecular configuration space,

and a function determined only from experimental data is not necessarily accurate in regions not represented in the data base i.e. the behaviour of the interaction energy in the configuration space is considerably underdetermined by the available data. In practice, a sensible functional form partially compensates for this problem.

On the quantum mechanical side, the non-empirical calculation of intermolecular interactions for small and modestly sized systems using molecular quantum mechanics is now feasible. The main advantages here are that interaction energies can be determined at rigorously defined levels of approximation and that any possible geometrical arrangement of molecules can be considered. Various reasonable approaches to sampling configuration space have been suggested and used effectively, and fitting functional forms to data can be accomplished with reasonable precision. However, the task of generating the data base of quantum mechanically calculated interaction energies becomes prohibitively expensive as the size of the system under consideration increases, when larger basis sets are necessary, or when electron correlation must be included. Compromises in the quality of the quantum mechanical calculations used for the data base remain of course inherent in the resulting intermolecular potential function. A particular problem in quantum mechanical calculations of intermolecular interaction energies is the basis set superposition error, whereby small basis sets result in spuriously inflated interaction

energies and commensurate errors in other properties, particularily a foreshortening of the calculated equilibrium intermolecular separation.

The development of quantum mechanical potentials describing the interaction of water molecules has been pursued mainly by Clementi and coworkers[91]. Functions describing pairwise interactions over all configuration space were successfully developed from data bases of O(1ØØ) quantum mechanical calculations, and produced functions representative of molecular orbital calculations near the Hartree-Fock (HF) limit and for several levels of electron correlation via electronic configuration interaction. These models place positive charges on the hydrogen atoms and negative charge at a point on the bisector of the HOH angle. The repulsive core is represented by exponential terms centered on the individual atoms, making the repulsion anisotropic. The quantum-mechanical potential due to Matsuoka, Clementi and Yoshimine (MCY-CI)[51] based on intermolecular CI calculations, is presently the most widely used of the quantum mechanical potentials in computer simulation studies of liquid water.

In all the computer simulations carried out in this study we use MCY-CI potential to describe the water-water pair interactions. The charactristics of this potential function used in the Monte Carlo simulations for the description of liquid water have been treated in several earlier papers; in

summary this function produces a $g_{OO}(R)$ in agreement with that observed from diffraction experiments and gives most of the thermodynamic variables in fairly good agreement with the experimental values. The most serious problem with this potential function in computer simulations is the inflated value of computed pressure of the liquid, indicating problems in describing the curvature of the potential in the equilibrium region.

The potential functions used to describe solute-water and solute-solute in our study varies from system to system and are described in each chapter.

CHAPTER IV.

Aqueous Hydration of Methane.

In this chapter we present the results from Monte Carlo simulations on a prototype system for hydrophobic hydration, $[CH_4]_{aq}$, and discuss the results in context of results from other simulations on this system[50,53,67]. In addition, the effects of convergence acceleration procedures on the simulation results are discussed. The present study forms the necessary first step to understand the hydrophobic interaction between methane discussed in the next chapter, and differs from previous studies on $[CH_4]_{aq}$ mainly in the potential function used for the description of methane-water interactions.

Calculations.

Statistical thermodynamic (T,V,N) ensemble Monte Carlo calculations were carried out on $[CH_4]_{aq}$ using a Metropolis procedure[80] (referred to as MMC) and a modified Metropolis procedure incorporating the force-bias method[78] and preferential sampling[89] (referred to as FBPS) for convergence The system for study in both the cases comprised of 216 rigid particles, one methane molecule and 215 waters. The simulations were performed at 25°C and at a density determined from experimental molar volumes of methane and water.[91] The condensed phase environment was provided by face-centered cubic boundary conditions, which ensures more than two complete hydration shells for methane. Convergence characteristics were monitored by control functions and statistical error bounds were calculated by the batch mean method using batches of size 50K.

We use MCY-CI[51] potential to describe the water-water pair interactions. Methane-water interactions were modeled by QPEN potentials developed by Marchese et. al.[92] QPEN potential is a transferable empirical potential constructed from quantum mecahnical energies. In the QPEN formalism the interaction energy between methane and water consistes of an electrostatic term and a non-bonded term. The former arises from the interaction of atoms in methane with those of water, and the latter from the interaction of bonding electron pair(for each C-H bond) of methane with the bonding

electron pair(for each O-H bond) and the lone pairs in water. The transferable parameters for the QPEN methane-water potential are listed in TableIV.1. These parameters correspond to a QPEN potential function fit to 225 methane-water quantum mechanical energies calculated including second order Moller-Plesset correction. The details of the potential function can be found in Ref.92 The potential energy hypersurface shows a global minimum of -2.17 kcal/mol, and corresponds to the water situated such that the plane of water is parallel to HCH plane of methane. In this configuration, the lone pairs of water point towards the methane. The methane-water interaction energy derived from QPEN potential function is more attractive than the previously reported values for this quantity. We discuss the effect of this on the calculated properties in the next section.

A total of 2500K configurations were sampled in both MMC and FBPS simulations. The first 1000K were discarded as equilibration and the ensemble average was done on the last 1500K configurations. The convergence profile for the calculations is shown in Figure IV.1 and Figure IV.2 respectively for MMC and FBPS simulations.

Table IV.1 -Transferable QPEN parameters for methane and water. BP refers to the bond pair and LP refers to lone pair.

	1 CHARGES.	2 A	3 B	4 C
1. C (METHANE)	4.			
2. H (METHANE)	1.			
3. BP (C-H METHANE)	-2.	103636.9	5.0762	20.9184
4. O (WATER)	6.			
5. H (WATER)	1.			
6. BP (O-H WATER)	-2.	3145.54	4.222	51.5344
7. LP (WATER)	-2.	242144.9	6.4694	8.87587

Figure IV.1- Convergence profile for the MMC simulation on $[CH_4]_{aq}$. Mean energy is denoted as $\bar{U}$ and mean energy for batches of size 50K is denoted as $\bar{U}_{50}$.

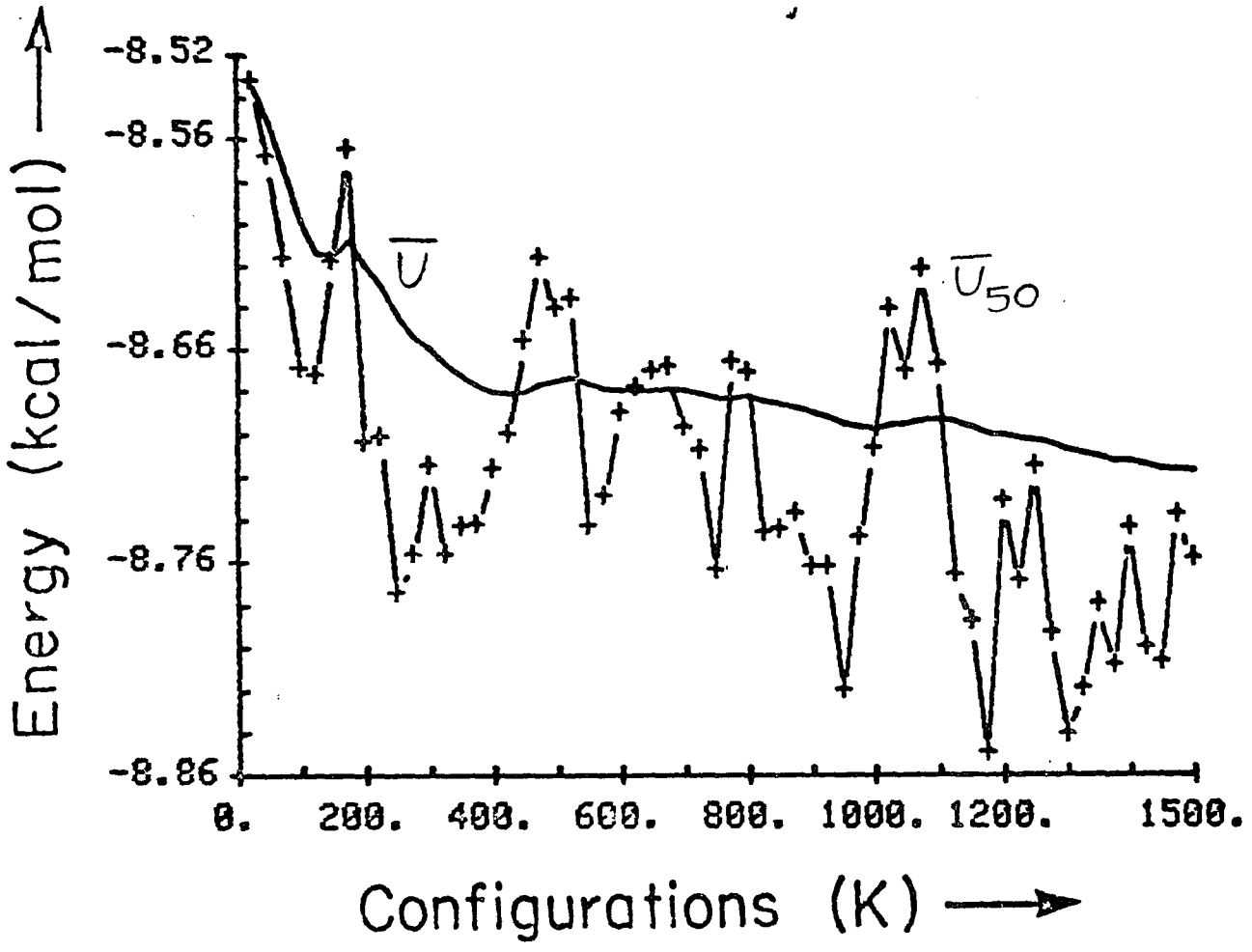
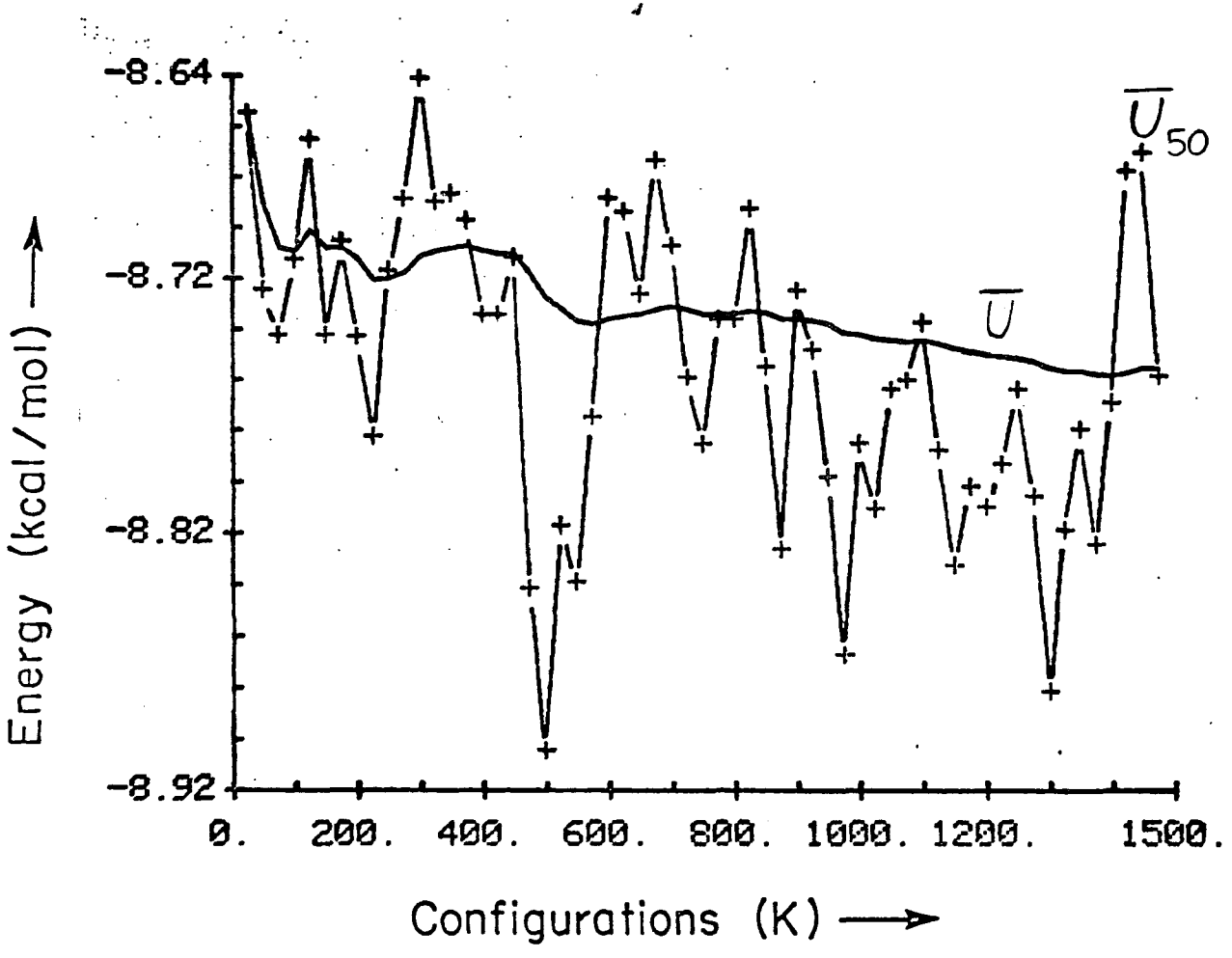

-8.52
-8.56
-8.66
-8.76
-8.86
Energy (kcal/mol)
$\overline{U}$
$\overline{U}_{50}$
0.
200.
400.
600.
800.
1000.
1200.
1500.
Configurations (K)

Figure IV.2- Convergence profile for the FBPS simulation on $[CH_4]_{aq}$. Mean energy is denoted as $\bar{U}$ and mean energy for batches of size 50K is denoted as $\bar{U}_{50}$.


-8.64
-8.72
-8.82
-8.92
Energy (kcal/mol)
$\overline{U}_{50}$
$\overline{U}$
0.
200.
400.
600.
800.
1000.
1200.
1500.
Configurations (K)

Results.

The calculated internal energies and related quantities for $[CH_4]_{aq}$ are collected in Table IV.2 The results from MMC simulation appears as column 1 and those from FBPS simulation appears as column 2. The quantities entered here are the mean energy U of the system (N_S=1, N_W=215), the energy U_W of 215 water molecules in $[H_2O]_1$ at 25°C, $U_{W'}$, the corresponding energy of solvent water in $[CH_4]_{aq}$, $\bar{U}_S$, the calculated partial molar internal energy of transfer for $[CH_4]_{aq}$ into water, and finally $\bar{U}_S$, and U_{rel} the solute-solvent and solvent-solvent contributions to $\bar{U}_S$. Each of these quantities is illustrated in Figure IV.3. The calculated molecular distribution functions and analysis thereof for $[CH_4]_{aq}$ at 25°C follow.

In Table IV.3 certain setup characteristics and results from computer simulation studies on $[CH_4]_{aq}$ are collected. The methane-water potential functions used in these studies are widely different and give rise to a range of pair energy minima. Our studies, MMC and FBPS, uses the largest number of water molecules and the studies referred to as 3T and 4T, by Bolis and Clementi[87], are the longest simulations. A comparison of structural indices such as methane-water pair correlation function $g_{MW}(R)$ shows that the general picture of hydrophobic hydration emerging from these studies agree with each other, considering the influ-

Table IV.2 Calculated Internal Energies for the Dilute Aqueous Solution of Methane at 25°C in kcal/mol.

	MMC	FBPS
U_{SW} (N_W=215,N_S=1)	-1882.84	-1891.29
U_W (N_W=215)	-1859.75	-1859.75
$U_{W'}$ (N_W=215)	-1874.00	-1882.40
$\bar{U}_{S'}$	-8.78	-8.97
$\bar{U}_{rel}$	-14.30	-22.70
$\bar{U}_S$	-23.10	-31.60

Figure IV.3 - Thermocycle illustrating the energetic quantities produced in a Monte Carlo simulation of a dilute solution of solute S, in water W.

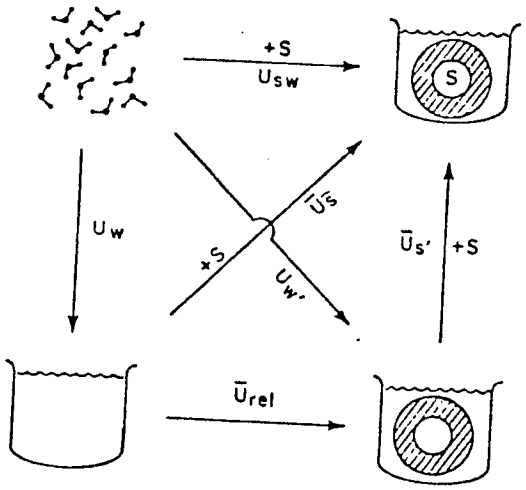

+S
U_{sw}
S
U_w
$\bar{U}_{s'}$
+S
+S
$U_{w'}$
$\bar{U}_{rel}$

Table IV.3- Comparison of computer simulation setup characteristics and results from studies on $[CH_4]_{aq}$ at 25°C. RDF refers to radial distribution function. OS- Owicki and Scheraga[50]; MO,MP- Swaminathan et. al.[53]; and 3T,4T- Bolis and Clementi[67]. The last two columns refer to MMC and FBPS simulations of this study.

	1 OS	2 MP	3 MO	4 3T	5 4T	6 QPEN	7 QPEN FB AND PS.
1. PAIR ENERGY MIN	-.42	-1.17	-.75	-.7	-.65	-2.13	-2.13
2. N WATER	100	124	124	202	202	215	215
3. ENSEMBLE	NPT	NVT	NVT	NVT	NVT	NVT	NVT
4. CONFIG.S	670K	750K	575K	5800K	3400K	1500K	1500K
5. RDF C-O MAX.	4.2	3.4	3.5	3.9	3.8	3.4	3.5
6. RDF C-O INTEN.	2.5	2.6	2.2	2.14	1.97	3.05	2.86
7. RDF C-O MIN.	6.0	5.3	5.6	5.4	5.6	5.15	4.95
8. RDF C-O INTEN.	.7	.6	.6	.51	.34	.538	.596
9. USLT	-.87	1.71	1.73	-1.02	-.03	-8.8	-8.9
10. U(S)	-11	-23.3	-21.8	-11.8	-6.41	-23.1	-31.6
11. COORD. NO.	23	19.35	20	20	20	19.1	20.1

ence of methane-water interaction potentials. The enthalpy of transfer of methane differs widely among the studies and is discussed in the following section.

The methane-water radial distribution function calculated from MMC and FBPS simulations are given in Figures IV.4 and IV.5. The curves flanking the g(R) in these figures are the upper and lower error bars on g(R). The figures also contain the running coordination numbers. We wish to designate the calculated properties with superscripts MMC and FBPS referring to the simulation from which they were calculated, for easy referencing; thus, $g_{MW}^{MMC}(R)$ refers to methane water radial distribution function calculated from MMC simulation. $g_{MW}^{MMC}(R)$ shows two distinct peaks, one at 3.4 Å and the other at 6 Å. The first minimum occurs at about 5°A and is broad. There are 19 water molecules in the first hydration shell of methane based on $g_{MW}^{MMC}(R)$. $g_{MW}^{FBPS}(R)$ resembles $g_{MW}^{MMC}(R)$ very much but is much smoother. The errors on $g_{MW}^{FBPS}(R)$ are much smaller than the errors on $g_{MW}^{MMC}(R)$. The calculated first shell coordination number for methane based on $g_{MW}^{FBPS}(R)$ is about 20.

The calculated quasi-component distribution (QCDF) of coordination numbers, $X_C^{MMC}(K)$ and $X_C^{FBPS}(K)$, are given in Figures IV.6 and IV.7 respectively. $X_C^{MMC}(K)$ is symmetric and ranges

Figure IV.4- Methane-water pair correlation function calculated using MMC simulation plotted against center of mass separation R. The curves flanking g(R) are the 95% confidence curves. The dotted line is the running coordination number.

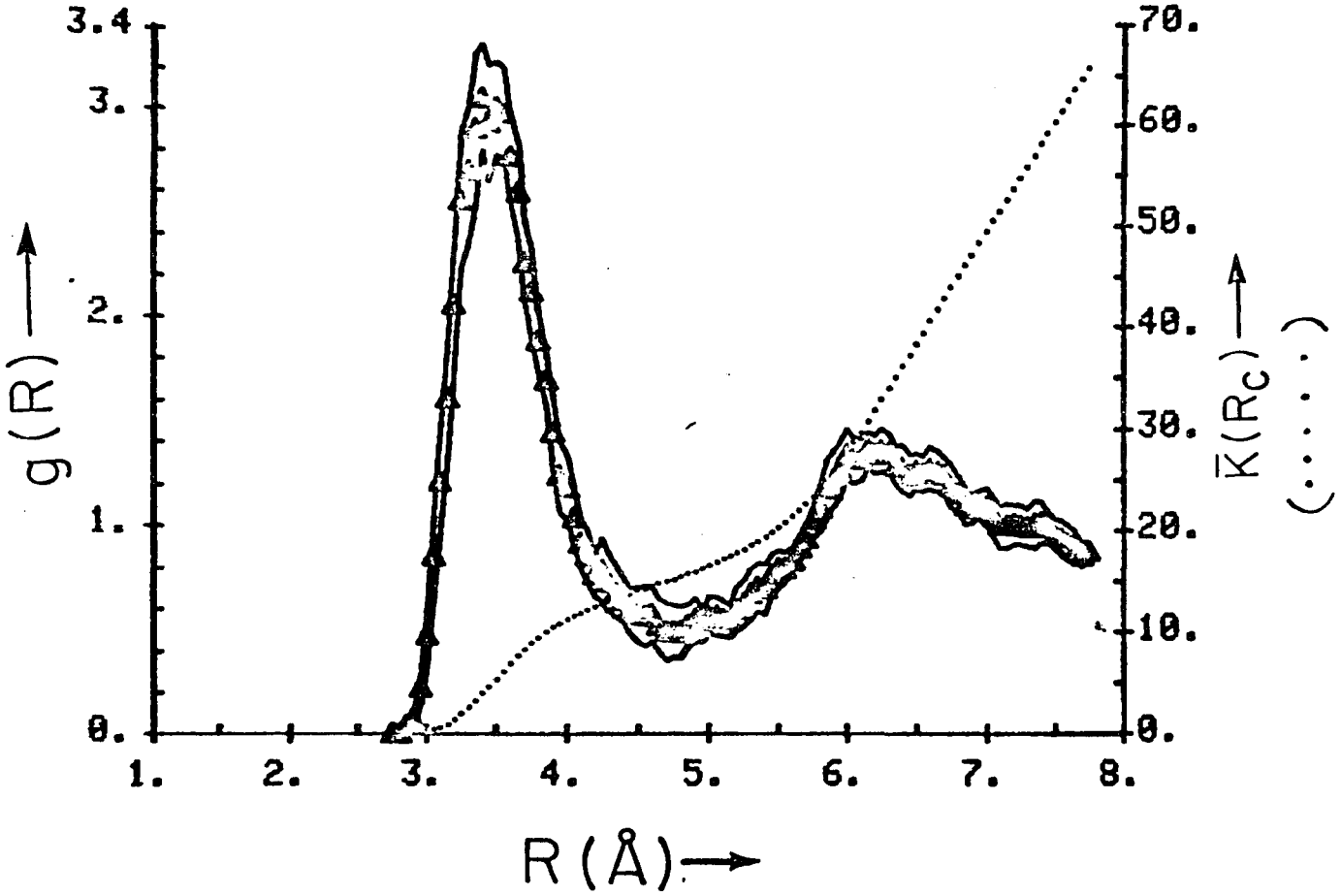

3.4
3.
2.
1.
0.
g(R)
70.
60.
50.
40.
30.
20.
10.
0.
$\bar{K}(R_c)$ (.....)
1.
2.
3.
4.
5.
6.
7.
8.
R (Å)

Figure IV.5- Methane-water pair correlation function calculated using FBPS simulation plotted against center of mass separation R. The curves flanking g(R) are the 95% confidence curves. The dotted line is the running coordination number.

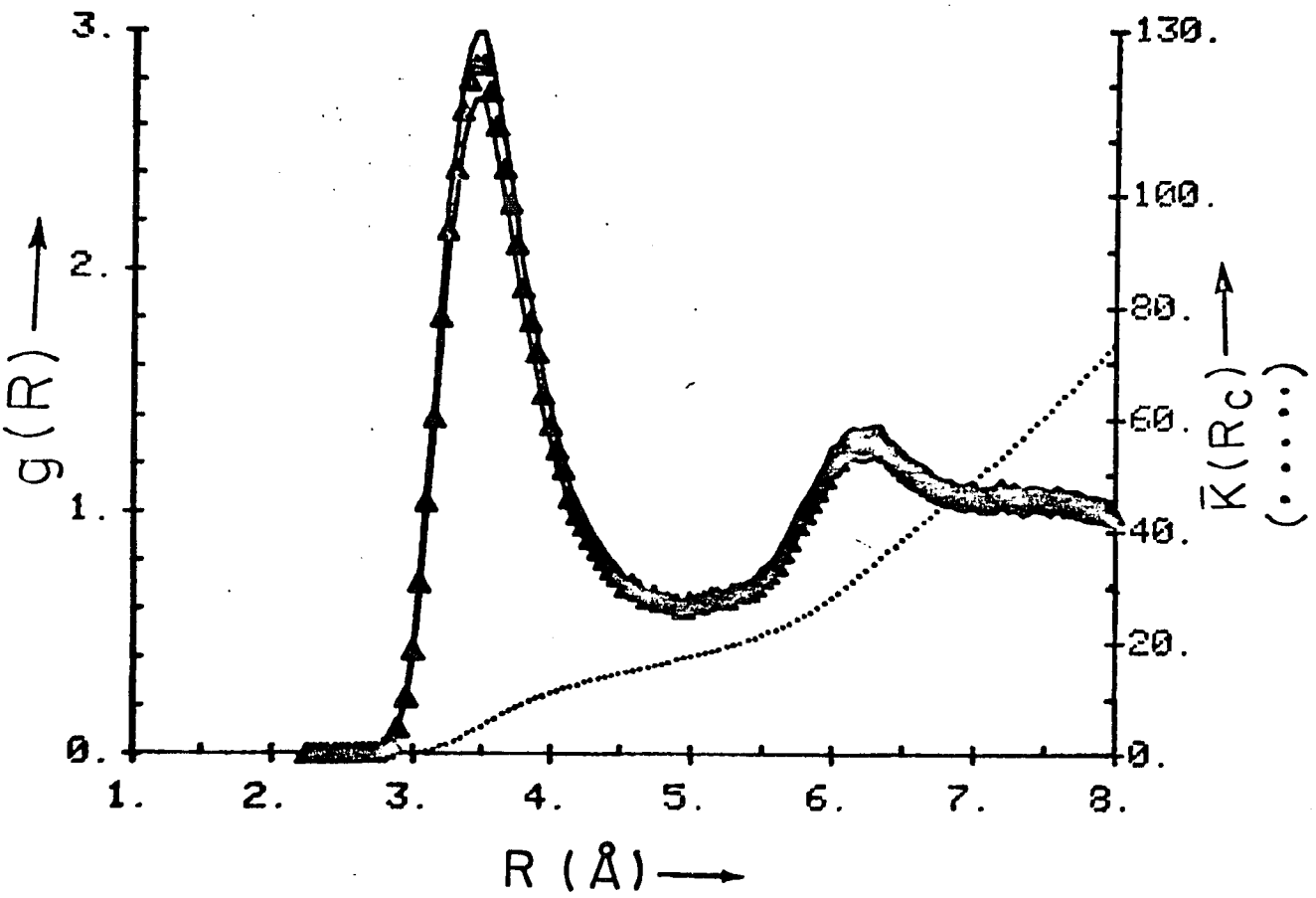

3.
2.
1.
0.
g(R)
130.
100.
80.
60.
40.
20.
0.
K(Rc)
(.....)
1.
2.
3.
4.
5.
6.
7.
8.
R (Å)

Figure IV.6- Calculated QCDF of coordination number from MMC simulation on $[CH_4]_{aq}$ at 25°C.

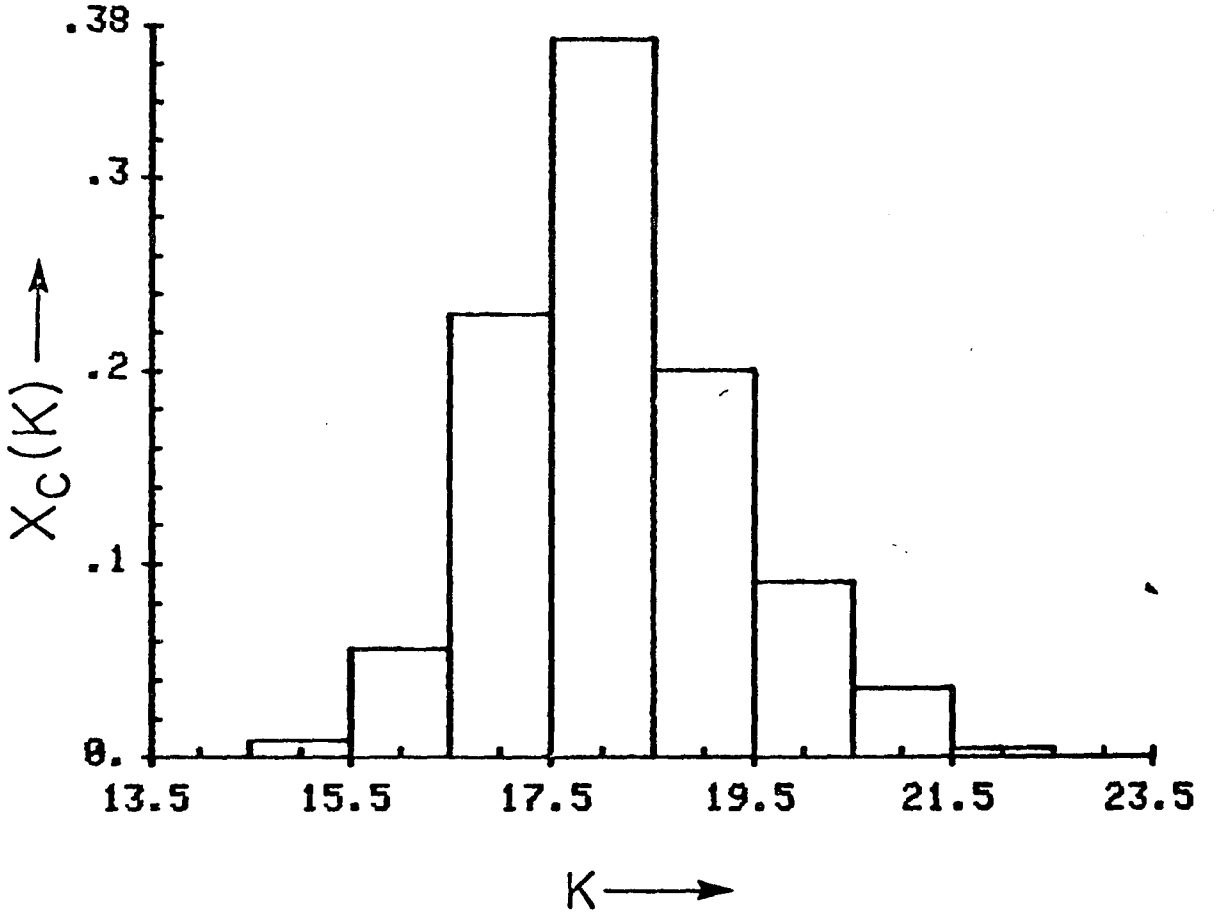
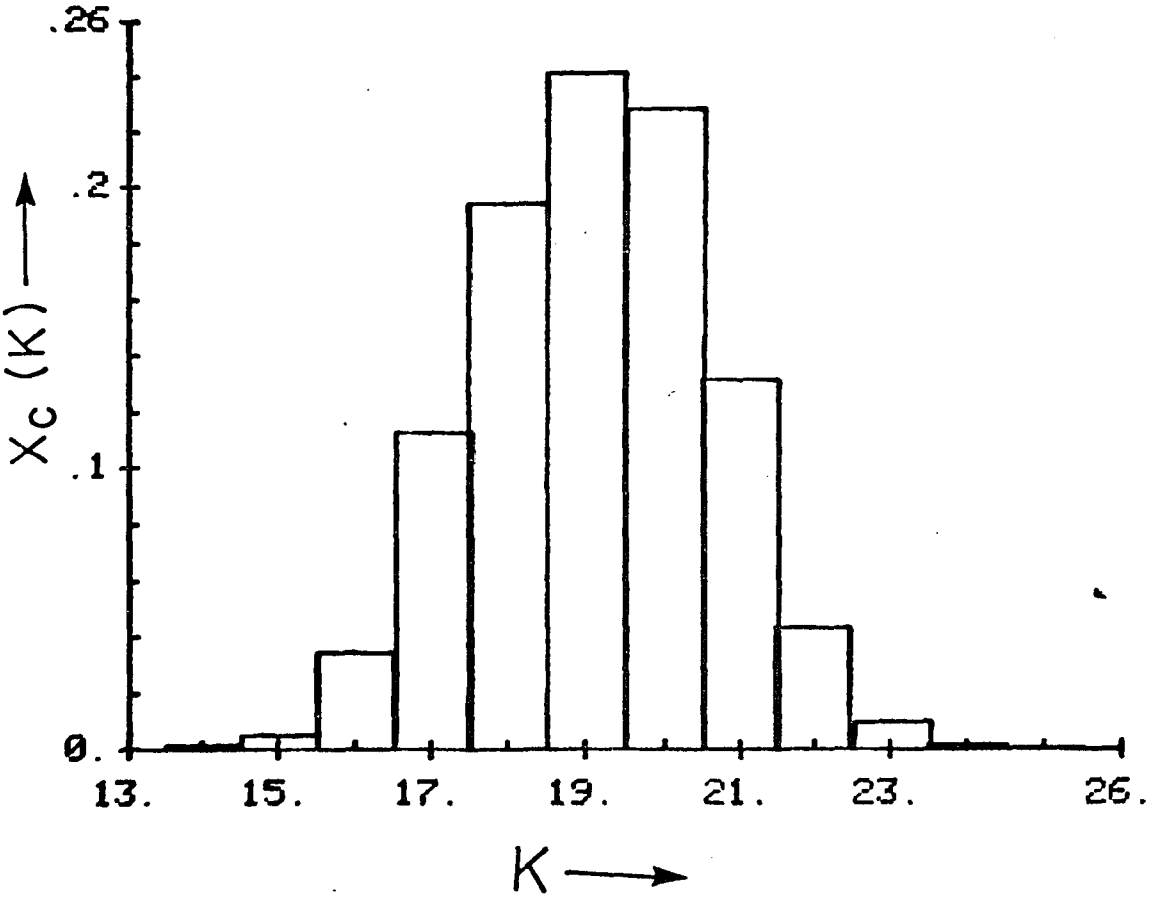

.38
.3
.2
.1
0.
$X_C(K)$
13.5
15.5
17.5
19.5
21.5
23.5
K

Figure IV.7- Calculated QCDF of coordination number from FBPS simulation on $[CH_4]_{aq}$ at 25°C.


.26
.2
.1
0.
X_c (K)
13.
15.
17.
19.
21.
23.
26.
K

from K= 15 to 22. with the peak at K=18, and an average $\bar{K}$=19.1. $X_C^{FBPS}(K)$ looks significantly different from $X_C^{MMC}(K)$ with the distribution shifted towards higher coordination numbers. $X_C^{FBPS}(K)$ ranges from 15 to 23 with an average $\bar{K}$=20.1. The peaks in $X_C^{FBPS}(K)$ at K=19 and 20 have the same amplitude. The QCDF of coordination numbers for both MMC and FBPS simulations were based on a R_C value of 5.3 Å. We chose this value to facilitate the comparison of our results with those of Swaminathan et. al.[53]

The QCDF of binding energies, $X_B^{MMC}(\mathcal{V})$ and $X_B^{FBPS}(\mathcal{V})$, of water molecules in $[CH_4]_{aq}$ are presented in Figures IV.8 and IV.9. The distribution $X_B^{MMC}(\mathcal{V})$ ranges from -10.8 to -5.3 kcal/mol with $\bar{\mathcal{V}}$=-8.8 kcal/mol and $X_B^{FBPS}(\mathcal{V})$ ranges from -11.3 to -5.8 kcal/mol with $\bar{\mathcal{V}}$=-8.9 kcal/mol.

The computed functions $X_P^{MMC}(\epsilon)$ and $X_P^{FBPS}(\epsilon)$ are given in Figures IV.10 and IV.11 respectively. Both the distributions range from -1.75 to .25 kcal/mol with the maximum at about 0 kcal/mol. These distributions are not restricted to first shell waters alone, and the majority of water molecules that are far away contribute to 0 kcal/mol.

A typical structure representative of the hydrophobic hydration of methane extracted from the FBPS simulation is given in Figure IV.12. Such a structure extracted from MMC resembles Figure IV.12 closely.

Figure IV.8- Calculated QCDF of binding energy from MMC simulation on $[CH_4]_{aq}$ at $25^{o}C$.

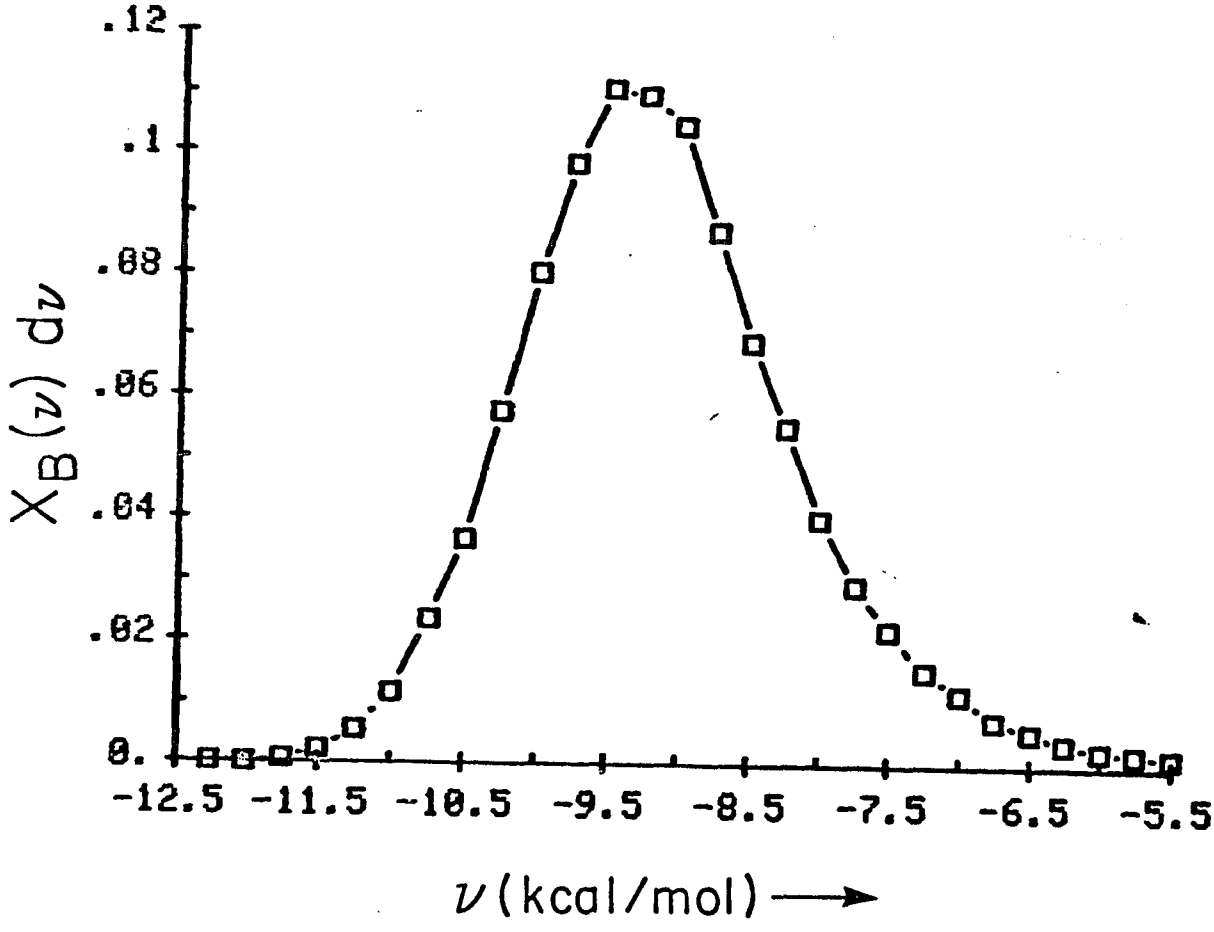

.12
.1
.08
.06
.04
.02
0.
$X_B(\nu)\, d\nu$
-12.5
-11.5
-10.5
-9.5
-8.5
-7.5
-6.5
-5.5
ν (kcal/mol) ⟶

Figure IV.9- Calculated QCDF of binding energy from FBPS simulation on $[CH_4]_{aq}$ at $25^{o}C$.

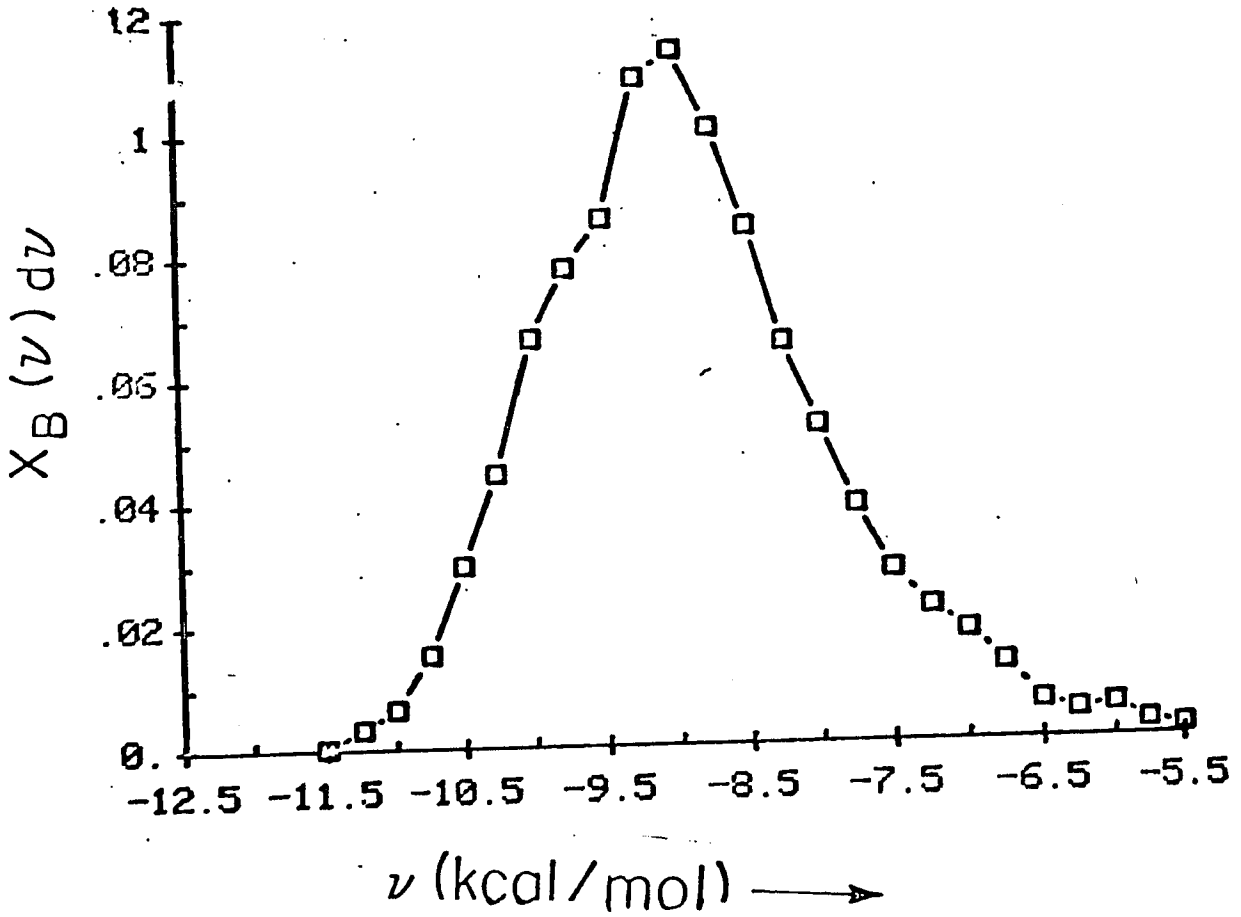

12
1
.08
.06
.04
.02
0.
$X_B(\nu)\,d\nu$
-12.5
-11.5
-10.5
-9.5
-8.5
-7.5
-6.5
-5.5
ν (kcal/mol) ⟶

Figure IV.1Ø- Calculated QCDF of pair energy from MMC simulation on $[CH_4]_{aq}$ at 25°C.

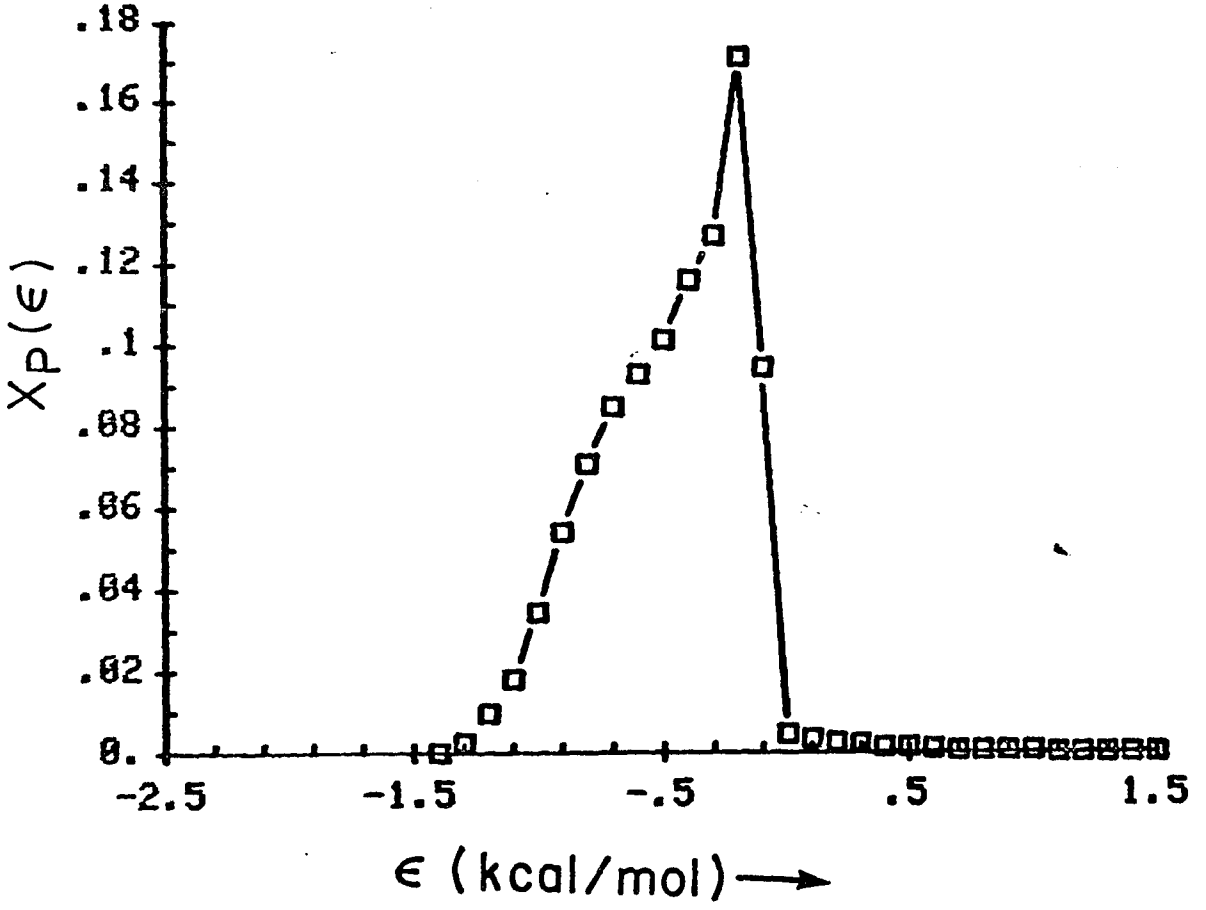
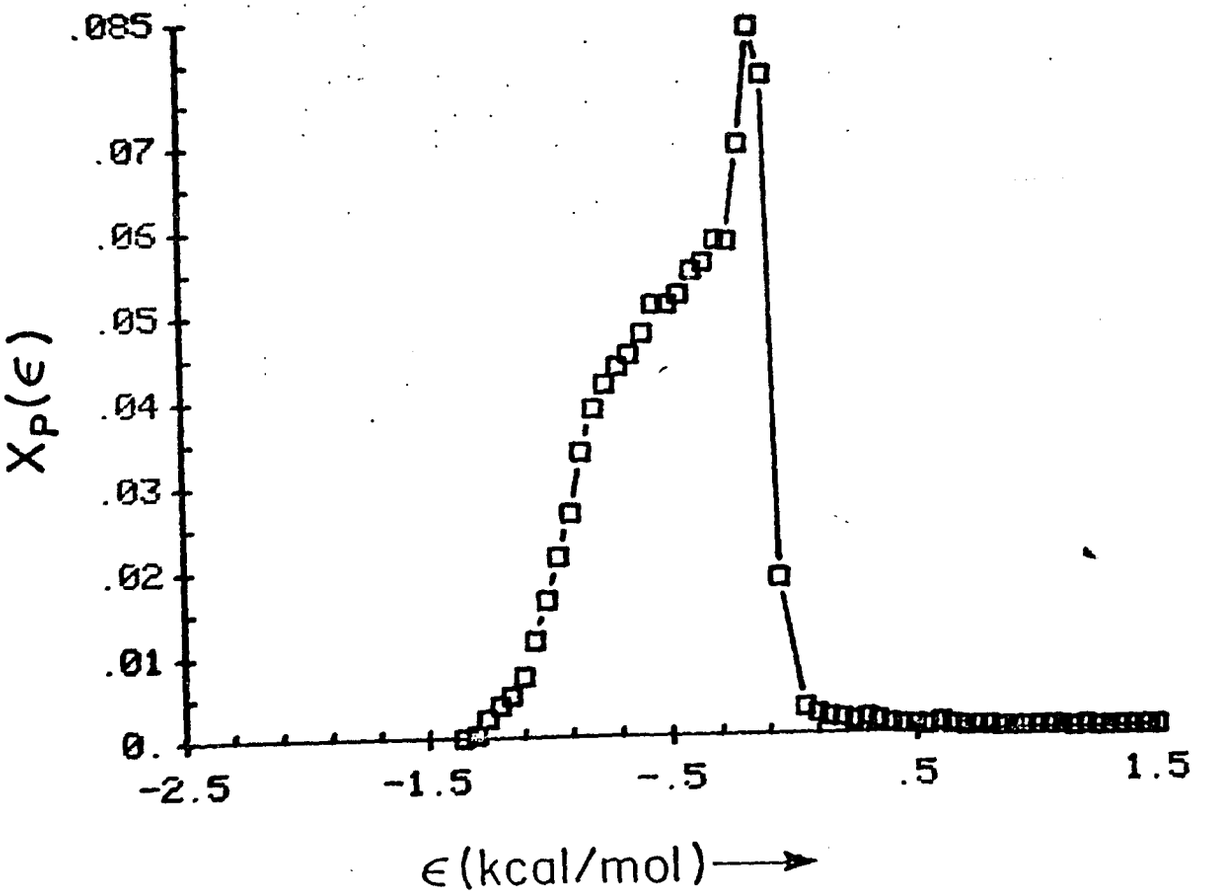

.18
.16
.14
.12
.1
.08
.06
.04
.02
0.
$X_p(\epsilon)$
-2.5
-1.5
-.5
.5
1.5
ϵ (kcal/mol) ⟶

Figure IV.11- Calculated QCDF of pair energy from FBPS simulation on $[CH_4]_{aq}$ at 25°C.

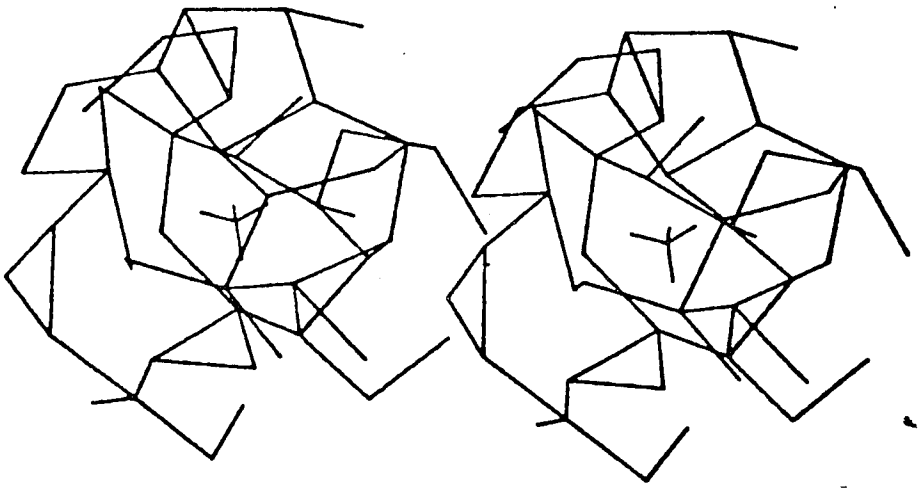

.085
.07
.06
.05
.04
.03
.02
.01
0.
$X_P(\epsilon)$
-2.5
-1.5
-.5
.5
1.5
ϵ(kcal/mol)⟶

Discussion.

The essential structural features emerging from the simulation results is first hydration shell of methane consisting of about 20 water molecules. A stereo view of the computer generated Dreiding model of a methane-water hydration complex is shown in Figure IV.12; the "bonds" in the figure connect oxygens of water molecules within the hydrogen bonding distance. The cage-like nature of the methane-water hydration complex is quite evident from this figure. Quite a few puckered pentagonal forms of water can be discerned but also contributions from lower and higher order polygonal forms. The coordination number of 20, combined with the above observation suggests possible arrangement of first shell waters in a pentagonal dodecahedral form. This result is also consistent with the predictions of Glew[93].

The average structure of methane-water hydration complex emerging from this study is essentially the same as that predicted from previous simulations, despite significant differences in methane-water interaction potentials. This shows that altering the methane-water interaction potentials does not alter the structural indices of $[CH_4]_{aq}$ significantly. Similar conclusions were derived by Pratt and Chandler[65] using integral equations techniques. They found that changing the solute-water interactions from hard sphere to Lennard-Jones type, changes $g_{SW}(R)$ by about 10. This is roughly the error associated with the $g_{SW}(R)$ cal-

Figure IV.12- Stereographic view of the Dreiding models of the first hydration shell of methane taken from the FBPS simulation described herein on $[CH_4]_{aq}$ at 25°C. Water oxygens within 3.2°C are bonded.



culated from computer simulations.

It is interesting to note the effect of convergence acceleration techniques on our results. The control functions presented in Figures IV.1 and IV.2 show that the FBPS simulation converged faster than the MMC simulation. The properties calculated from FBPS simulation were found to have much smaller error bars than those calculated from MMC simulation. Another way to test the convergence of results is to divide the simulation into 3 or 4 blocks and to calculate system properties in each block. The system properties in each block should remain significantly close to each other, if the convergence is good. Mehrotra et. al.[64] used the results from this study to show that $g_{MW}(R)$ calculated for blocks of size 500K remained significantly close to each other in FBPS simulation but not so in MMC simulation. This is further proof that the convergence acceleration procedures make the results converge faster. We therefore augment the Metropolis procedure with force-bias and preferential sampling techniques in all simulations reported in the following chapters.

The enthalpy of transfer of methane from vapor phase to water calculated from this study is very large compared to the experimental value of -2.6 kcal/mol[94]. It is also very different from the values calculated from other simulation studies. This quantity is calculated as the difference in two large quantities and has a large error associated with

it. Since the enthalpy of transfer depends on the solute-water interaction energy, the wide range for this quantity reported from simulations is understandable. A comparison of U(S) and the pair energy minimum in Table I shows that U(S) becomes more negative as the pair energy becomes more negative.

Calculations were extended to study the effect of temperature on the local solution environment of methane in $[CH_4]_{aq}$. Simulations with and without convergence acceleration were performed on $[CH_4]_{aq}$ at 25, 37 and 50°C. The calculated solute-water structural indices in this temperature range did not change significantly. This suggests that the fluctuations in the local structure of methane in $[CH_4]_{aq}$ in the temperature range 25 to 50°C may be too small and may be within the statistical error. The inflated value of pressure for the water model used in this study may also play an important role in reducing the effects of temperature. However the calculated $g_{OO}(R)$ for bulk water in this study shows a temperature dependence consistent with the experiments[95] in the temperature range of 25 to 50°C. thus making the latter explanation weak.

The calculations described herein showed that the loacl solution environment of methane in $[CH_4]_{aq}$ is described well using QPEN potential function for methane-water interactions. The convergence acceleration procedures were shown to aid in reducing the errors in calculated system proper-

ties. In addition, we found that temperature effects on solute-water properties cannot be detected even using convergence acceleration in reasonable temperature ranges.

CHAPTER V.

Monte Carlo Computer Simulation study of the Hydrophobic Effect.

Potential of Mean Force for $[(CH_4)_2]_{aq}$ at 25 and $50^{o}C$.

In this chapter we present the results of Monte Carlo simulations on a prototype system for study of hydrophobic interactions, $[(CH_4)_2]_{aq}$, and discuss our results in context of other recent theoretical and experimental work on this topic[25,44,47,62]. The major objective of this study was to calculate $w_{ss}(R)$ for two apolar molecules using potential functions derived from quantum-mechanical calculations and to analyze the structural details in the system. Umbrella sampling techniques[81] were incorporated in the Metropolis Monte Carlo method[80] for the calculation of $w_{ss}(R)$ and convergence acceleration procedures[78,89] were used in acheiving a better convergence in the results.

Calculations.

The calculation of $w_{ss}(R)$ for the interaction between two methanes in water described herein proceeds via the calculation of $g_{ss}(R)$ using Monte Carlo simulations augmented with umbrella sampling procedures. The theory and methodology section describes in detail the umbrella sampling. The umbrella sampling weighting functions used in this study are harmonic restoring potentials similar to those suggested by Pangali et. al.[62] in their study of pmf between two spherical apolar molecules.

The calculations described herein are intended as a direct extension of the earlier studies from this laboratory on hydrophobic hydration to the problem of hydrophobic interaction. The potential functions used for the evaluation of water-water pair interactions in the simulation was that of Matsuoka et. al.[50] based on quantum mechanical calculations. Methane-water interactions are treated by means of QPEN potential functions developed by Marchese et. al.[92] for this study. The details of this potential function can be found in the previous chapter on the hydrophobic hydration of methane. The parameters for methane-methane interactions are taken from the transferable methane-water QPEN potential function.

The calculations described herein are aimed at producing $g_{ss}(R)$ and $w_{ss}(R)$ for the intersolute coordinate in the system $[(CH_4)_2]_{aq}$. The system for study consists of

two methane molecules and 214 water molecules configured under face-centered cubic boundary conditions with a spherical cutoff treatment of the potential. Four points in origin on the intersolute coordinate were chosen, defined as 3.9,5.3,6.Ø7 and 6.8 Å. Harmonic constraining functions with a force constant of 1.5 kcal $Å^{-2}$ were used to define the four windows used in the umbrella sampling procedure. The density of the system was computed from the experimental measurements of partial molar volumes of methane and water[91]. Calculations for all four windows were carried out at 25°C and, for reasons described below, at 5Ø°C for the window centered at 5.3 Å. For each window that was studied, a total of $3 \times 1Ø^6$ configurations were sampled, of which the first $1.5 \times 1Ø^6$ configurations were treated as equilibration. All the properties reported are the ensemble averages from the second $1.5 \times 1Ø^6$ configurations.

Results and Discussion.

The $g_{ss}(R)$ for the four windows studied at 25°C are shown in Figure V.1. For window 1, the point of origin was chosen at 3.9Å to coincide with the solute-solute contact distance estimated from the position of the first peak in $g_{MW}(R)$. In this realization, the intersolute region from 3.2 to 5.2 Å is sampled. The maximum falls at 3.9Å, with a bias in the distribution in favour of longer distances.

Window 2 has the point of origin at 5.3 Å and was observed to sample a quite broad range of the intersolute coordinate from R=4.Ø Å to almost 7 Å. The maximum in the probability distribution for this window was found at 6.4 Å. The remaining two windows, R=6.Ø7 and 6.8 Å, both sample ca.3.5 Å of the intersolute coordinate and find their maxima in the region of 6.2 Å.

The probabilty distributions for the four windows were matched on the basis that the points in overlapping regions should coincide and the solute-solute radial distribution function $g_{ss}(R)$ was generated. The corresponding potential of mean force $w_{ss}(R)$ was generated and is plotted in Figure V.2. The oscillatory behaviour in $w_{ss}(R)$ noted previously by Pratt and Chandler[44] and Pangali et. al.[62] is clearly evident but differs from the monotonic $w_{ss}(R)$ reported by Marcelja et. al.[47] and Dashevsky and Sarkisov[25]. Two distinct minima, one at R=3.9 Å and the

Figure V.1.-Calculated radial distribution function $g_{ss}(R)$ plotted against intersolute separation R for each of the four windows of umbrella sampling for $[(CH_4)_2]_{aq}$ at 25°C.
●, Window 1; o, window 2; ■, window 3; ▲, window 4.

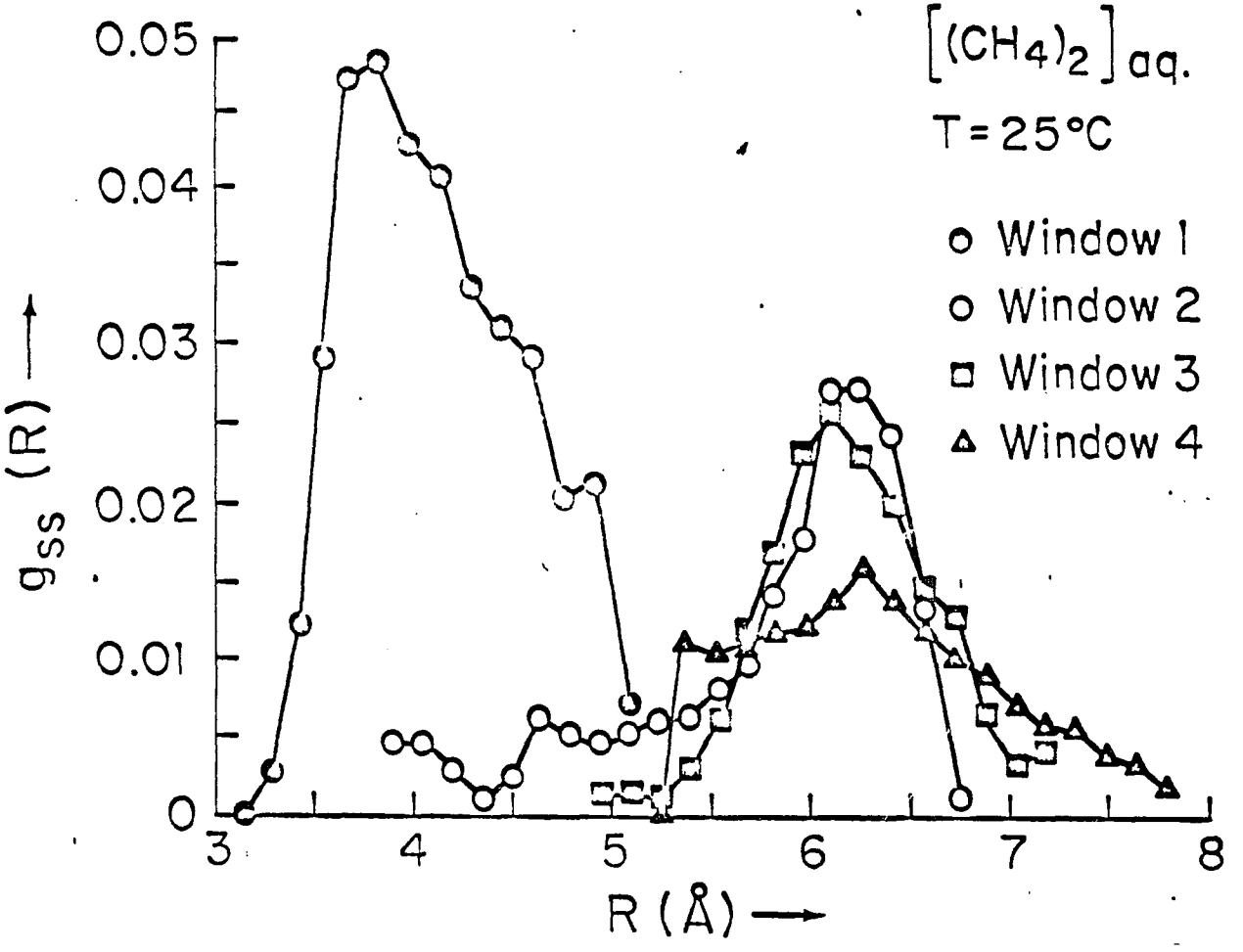
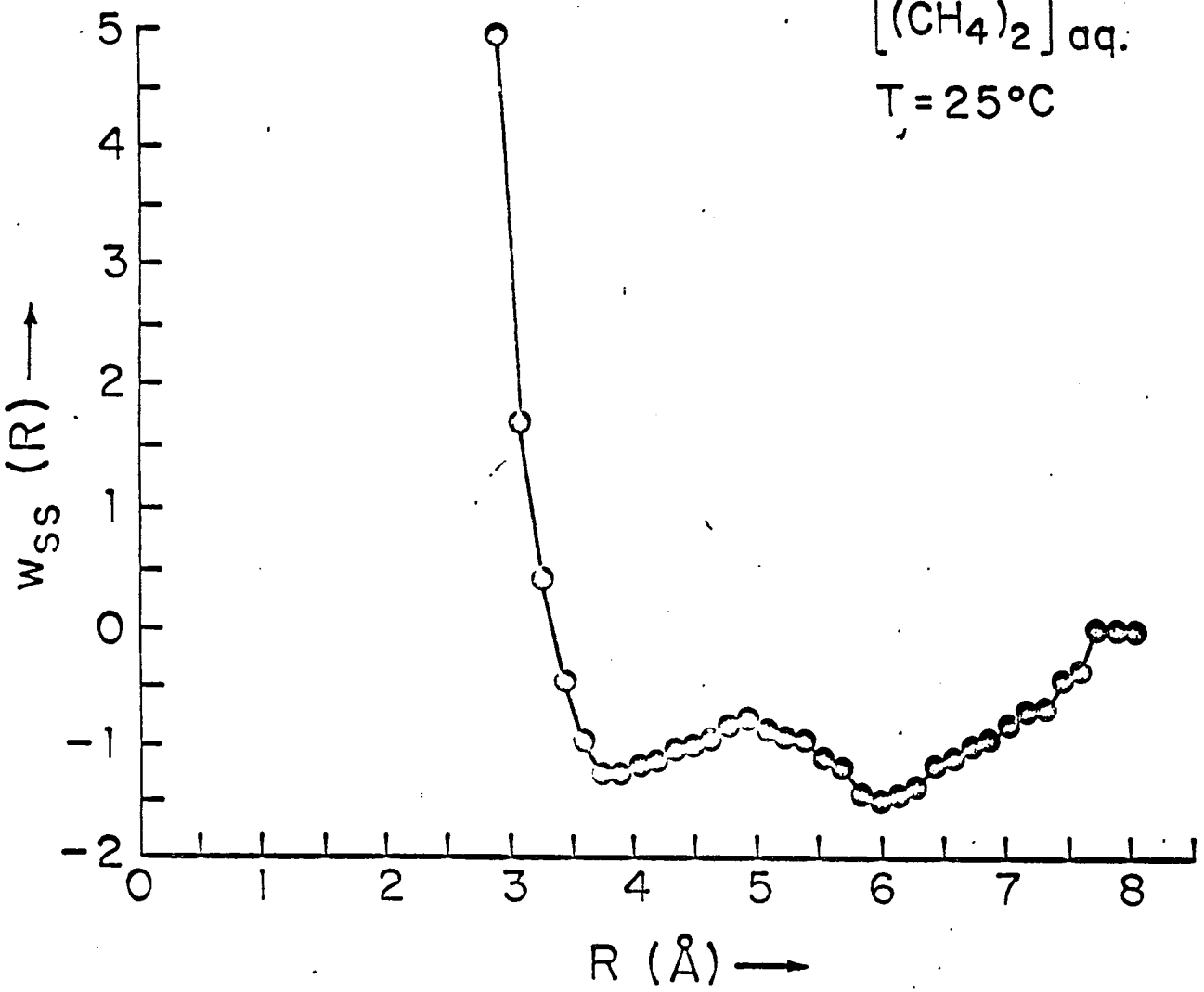

[(CH4)2]aq.
T=25°C
Window 1
Window 2
Window 3
Window 4
gSS (R)
0.05
0.04
0.03
0.02
0.01
0
3
4
5
6
7
8
R (Å)

Figure V.2.-Calculated potential of mean force $w_{SS}(R)$ as a function of R, after matching, for $[(CH_4)_2]_{aq}$ at 25°C.


$[(CH_4)_2]$ aq.
T = 25°C
W_{SS} (R) ⟶
R (Å) ⟶
5
4
3
2
1
0
−1
−2
0
1
2
3
4
5
6
7
8

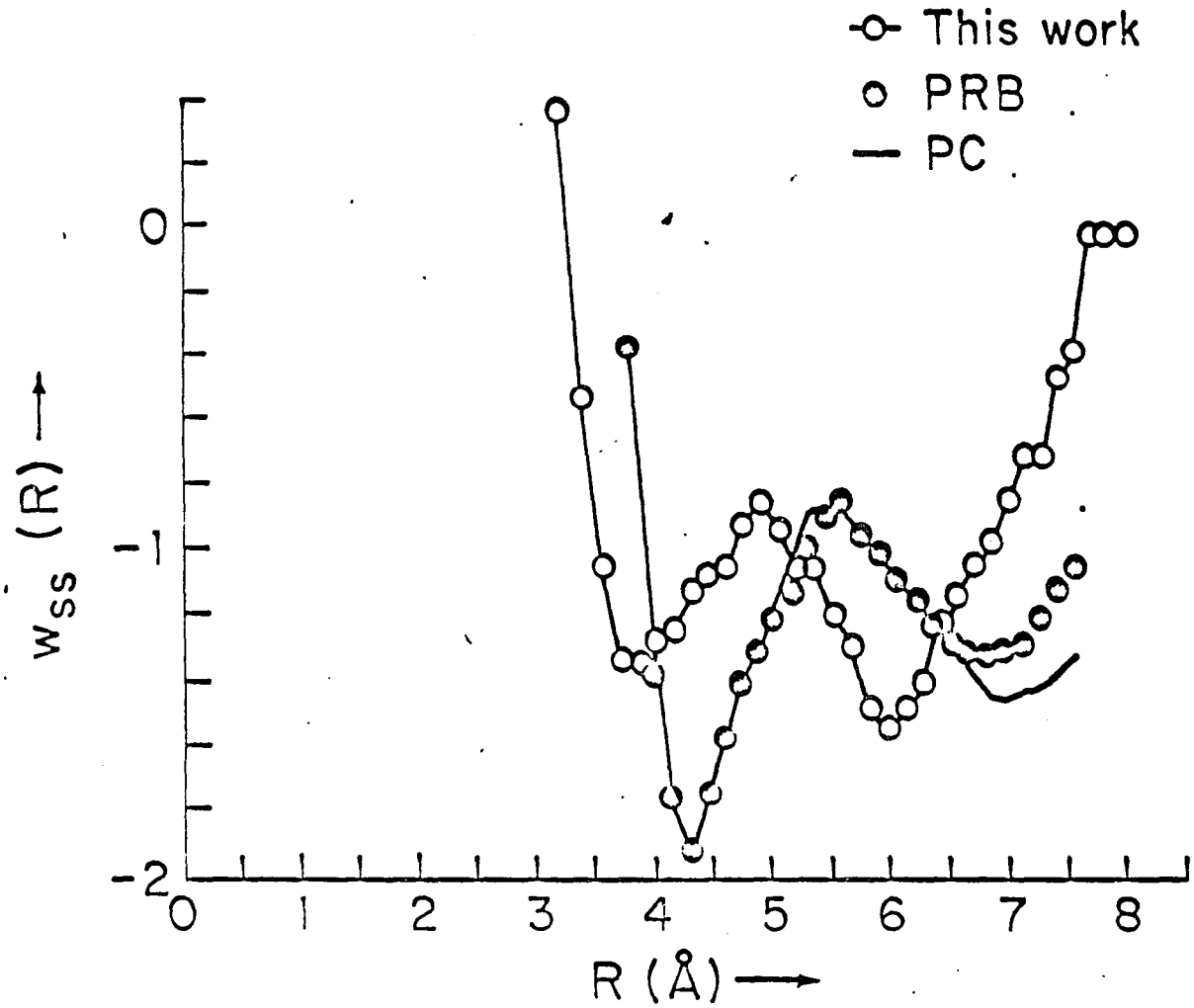
other at R=6.0 Å, are found. The former corresponds to the contact hydrophobic interaction and the latter to a solvent-separated structure.

The results of Pratt and Chandler(PC) and Pangali et. al.(PRB) are included in Figure V.3 for comparison and contrast. Two essential points of difference are noted: (a) the positions of the minima in our study occur at slightly shorter intersolute separations and (b) the solvent-separated minima in our study are deeper than the minima for the contact interaction, whereas in the previous studies the contact minima were deeper.

The difference in position of the minima in $w_{ss}(R)$ is due primarily to a difference in the solute-water potential. The PC and PRB studies both assume a Lennard-Jones form for the solute-water potential, which lacks the granularity expected when the molecules interact with one another. The quantum-mechanical potential used herein contains a more detailed description of this interaction and this is reflected in the positions of the calculated minima in the $w_{ss}(R)$.

The discrepancies in relative depth of the contact and solvent separated minima in the $w_{ss}(R)$ can also be traced back to differences in the potential function. The methane-water binding energy is negligible in the PC and PRB studies, whereas our function exhibits an equilibrium binding energy of -2.17 kcal mol^{-1}. The description of water-wa-

Figure V.3-Comparison of the calculated $w_{ss}(R)$ in this work(-- --) with that computed by Pangali et. al.(●) and Pratt and Chandler(--).


This work
PRB
PC
0
-1
-2
w_{ss} (R) ⟶
0
1
2
3
4
5
6
7
8
R (Å) ⟶

ter interactions varies among the studies as well, but since there is such close agreement between PC[using the experimental $g_{OO}(R)$] and PRB[using ST2 water, which overemphasizes considerably the tetrahedral nature of water interactions], we feel that the origin of this discrepancy probably resides in the different solute-water binding energy. This effect was anticipated by Pratt and Chandler[65]. The interpretation of the results is significantly affected at this point. The PC and PRB results produce an equilibrium constant for the CHI $\underset{<---}{\overset{--->}{}}$ SSHI equilibrium, of the order of unity, indicating significant contributions from both the contact and solvent-separated structures to the statistical state of the system at 25°C. Our results predict a strong preference for the solvent-separated form. There is at present no means of deciding which result is correct, but sensitivity of this crucial result to the choice of potential functions is notable. All studies concur in indicating that the solvent-separated form should be seriously considered in the structural biochemistry of the hydrophobic effect.

A preliminary study of the temperature dependence of the hydrophobic interaction was carried out by means of a Monte Carlo simulation on $[(CH_4)_2]_{aq}$ at 50°C. This study was carried out for window 2 only, which was observed to span the contact and solvent-separated structures on the intersolute coordinate. The calculated $g_{ss}(R)$ and $w_{ss}(R)$ for 50°C are shown in Figures V.4 and V.5, respectively.

Figure V.4-Calculated $g_{SS}(R)$ as a function of R for window 2 for $[(CH_4)_2]_{aq}$ at o, 25 and ●, 50°C.

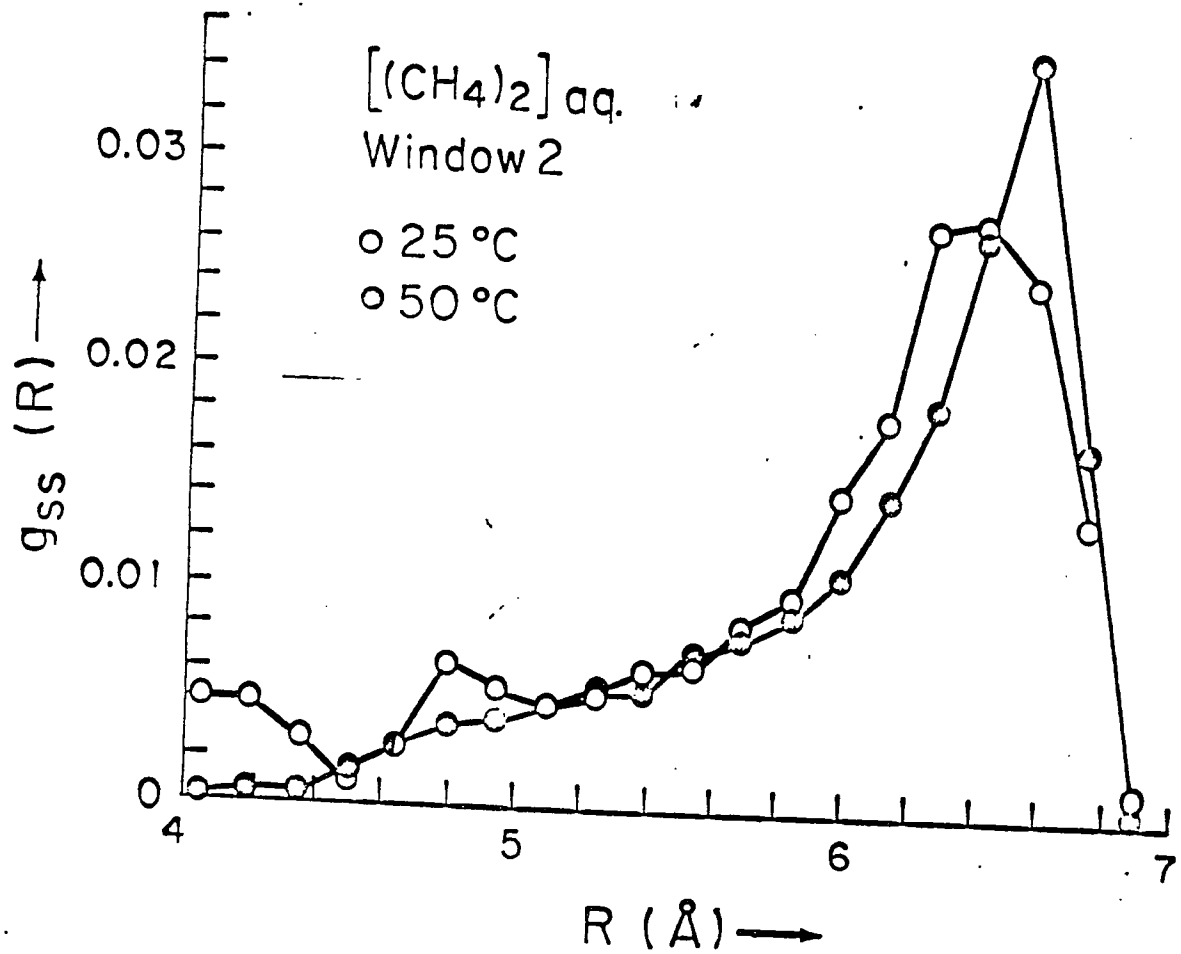
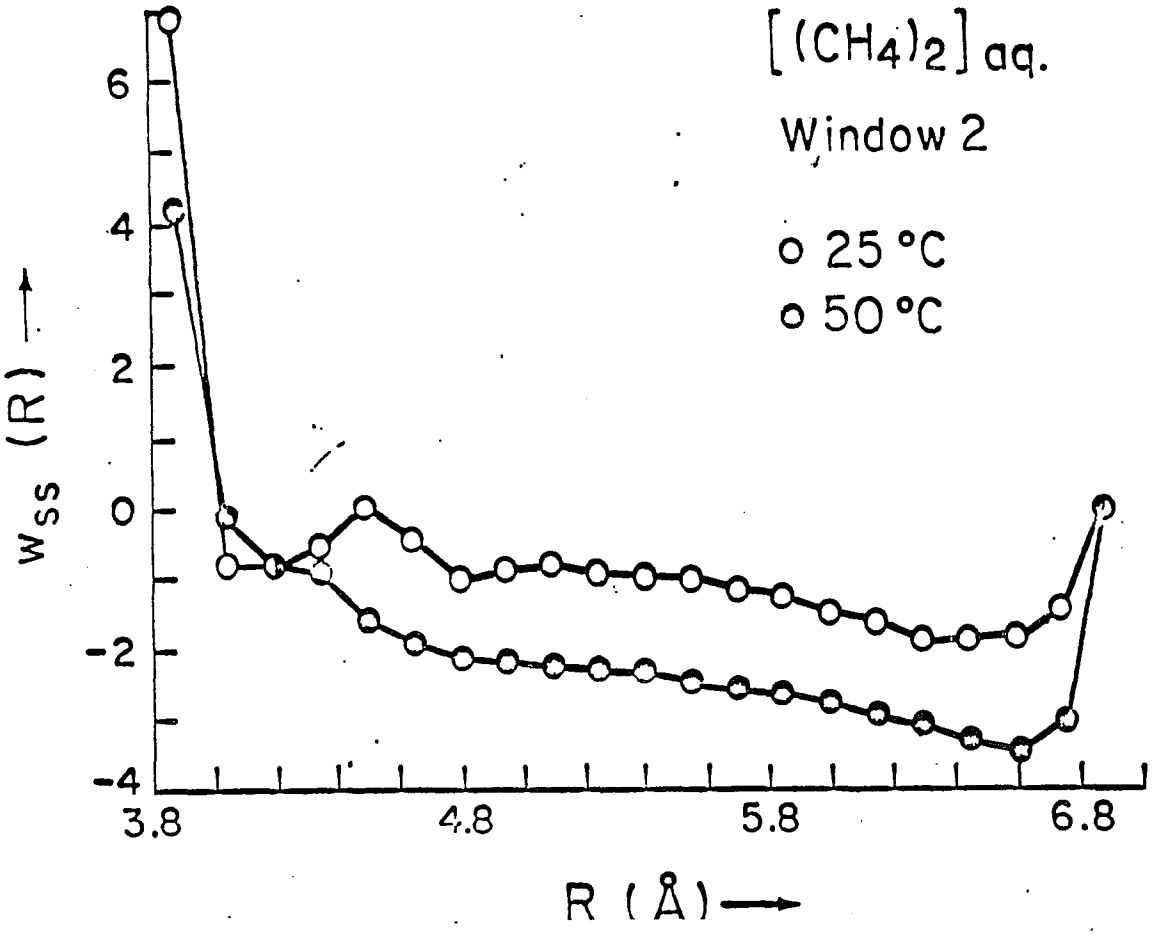

[(CH4)2] aq.
Window 2
o 25 °C
o 50 °C
0.03
0.02
0.01
0
gss (R)
4
5
6
7
R (Å)

Figure V.5-Calculated $w_{ss}(R)$ plotted against R for window 2 for $[(CH_4)_2]_{aq}$ at o, 25 and ●, 50°C.


[(CH4)2] aq.
Window 2
o 25 °C
o 50 °C
6
4
2
0
-2
-4
Wss (R) ⟶
3.8
4.8
5.8
6.8
R (Å) ⟶

The minimum in $w_{ss}(R)$ corresponding to the contact interactions is no longer present and the solvent separated structure is preferentially stabilized by the increase in temperature. Thus the inverse temperature dependence of the hydrophobic effect is accounted for by computer simulation and linked to the preferential stabilization of the solvent-separated hydrophobic interaction.

Analysis of Results.

The results of the preceding section strongly implicate the solvent-separated structures in the hydrophobic interaction. In this section we inquire into the molecular nature of these structures. Two procedures are employed for this analysis: (a) examination of the distribution functions $g_{\bullet O}(R)$ and $g_{\bullet H}(R)$, describing the organization of solvent water with respect to the center of mass(•) of the SS complex, and (b) consideration of stereographic pictures of supermolecular structures from configurations that contribute significantly to the statistical state of the system. The $g_{\bullet O}(R)$ for window 1,Figure V.6, shows successive peaks at 2.8,4.1 and 5 Å. Assignment of these peaks to waters in the first hydration shell follows straightforwardly from a consideration of the excluded volume of methane and water molecules, as shown in the inset of Figure V.6. The distribution $g_{\bullet H}(R)$, Figure V.7, is more complex because of the greater orientational possibilities,but the onset of probability at 1 Å is clearly consistent with a contact interaction between the methanes. Structures representative of this type of configuration were extracted from the simulation, and typical example is shown in Figure V.8.

For window 2, the distribution functions $g_{\bullet O}(R)$, Figure V.9, shows a well developed peak at 1 Å. Examination of $g_{\bullet H}(R)$, Figure V.1Ø, shows dominant contributions from structures in which a single hydrogen atom of a water molec-

Analysis of Results.

The results of the preceding section strongly implicate the solvent-separated structures in the hydrophobic interaction. In this section we inquire into the molecular nature of these structures. Two procedures are employed for this analysis: (a) examination of the distribution functions $g_{\bullet O}(R)$ and $g_{\bullet H(R)}$, describing the organization of solvent water with respect to the center of mass(•) of the SS complex, and (b) consideration of stereographic pictures of supermolecular structures from configurations that contribute significantly to the statistical state of the system. The $g_{\bullet O}(R)$ for window 1,Figure V.6, shows successive peaks at 2.8,4.1 and 5 Å. Assignment of these peaks to waters in the first hydration shell follows straightforwardly from a consideration of the excluded volume of methane and water molecules, as shown in the inset of Figure V.6. The distribution $g_{\bullet H(R)}$, Figure V.7, is more complex because of the greater orientational possibilities,but the onset of probability at 1 Å is clearly consistent with a contact interaction between the methanes. Structures representative of this type of configuration were extracted from the simulation, and typical example is shown in Figure V.8.

For window 2, the distribution functions $g_{\bullet O}(R)$, Figure V.9, shows a well developed peak at 1 Å. Examination of $g_{\bullet H(R)}$, Figure V.1Ø, shows dominant contributions from structures in which a single hydrogen atom of a water molec-

Figure V.6-Calculated $g_{\bullet O}(R)$ for $[(CH_4)_2]_{aq}$ for window 1.

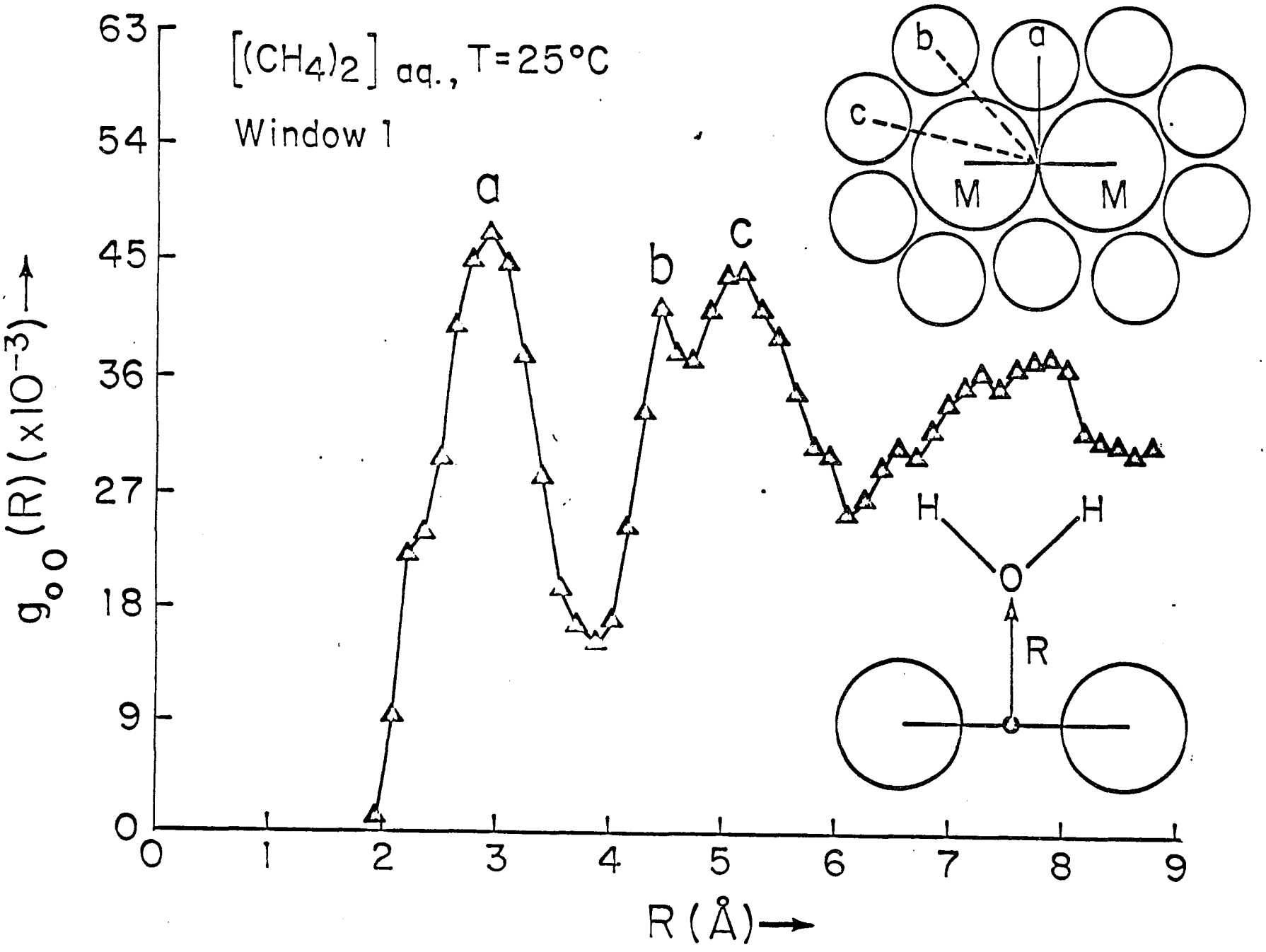

[(CH4)2] aq., T=25°C
Window 1
a
b
c
M
M
H
H
O
R
63
54
45
36
27
18
9
0
0
1
2
3
4
5
6
7
8
9
goo (R) (x10-3)
R (Å)

Figure V.7-Calculated $g_{\bullet H}(R)$ for $[(CH_4)_2]_{aq}$ window 1.

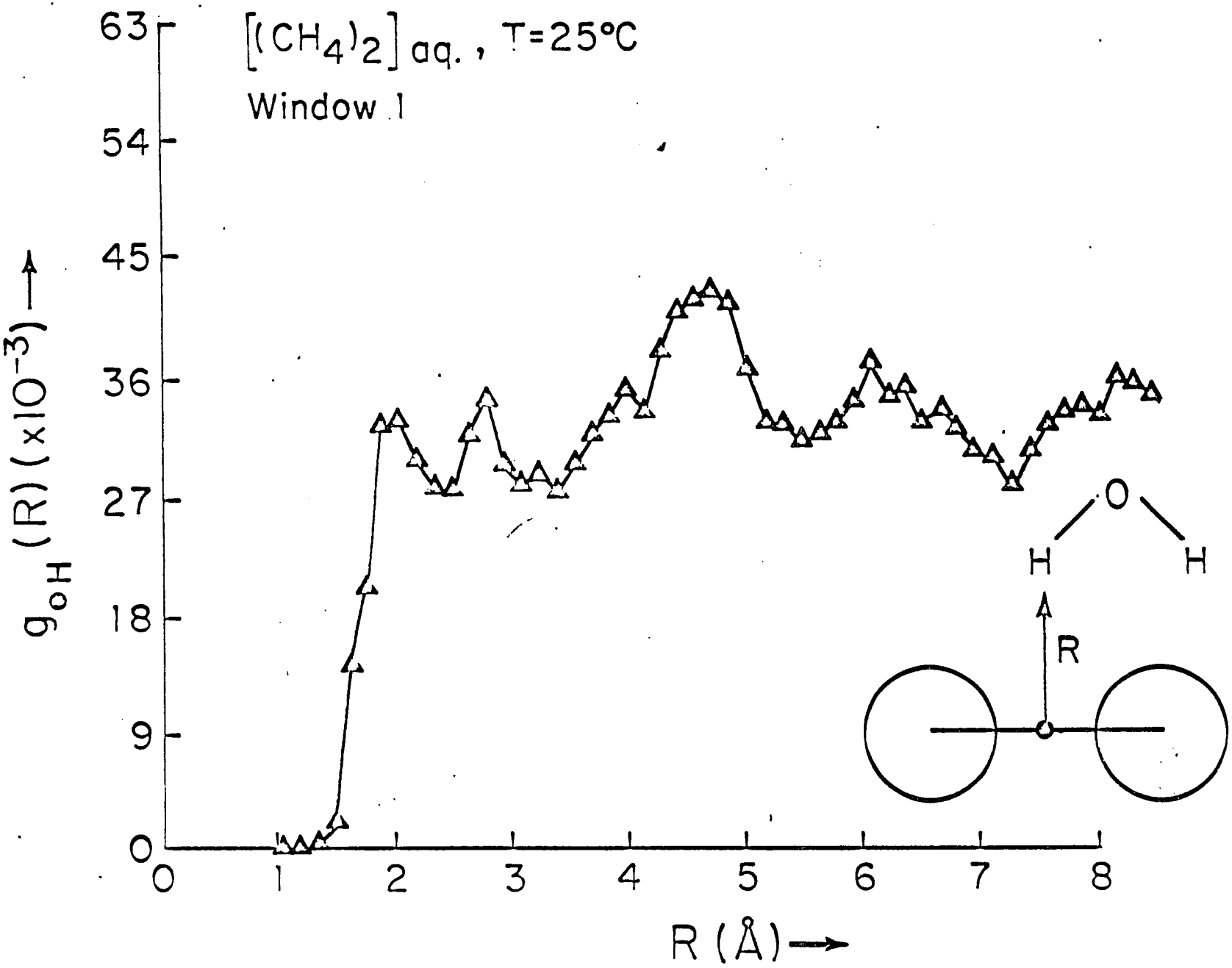

[(CH4)2] aq., T=25°C
Window 1
63
54
45
36
27
18
9
0
0
1
2
3
4
5
6
7
8
g_oH (R) (x10^-3) →
R (Å) →
O
H
H
R

Figure V.8-Typical structure corresponding to the most probable configuration in window 1.

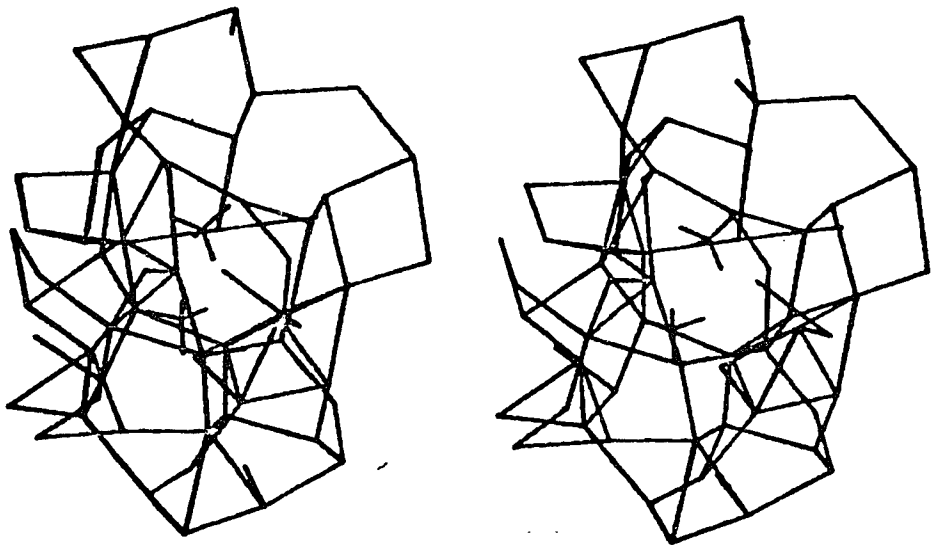


Figure V.9-Calculated $g_{O}(R)$ for $[(CH_4)_2]_{aq}$ for window 2.

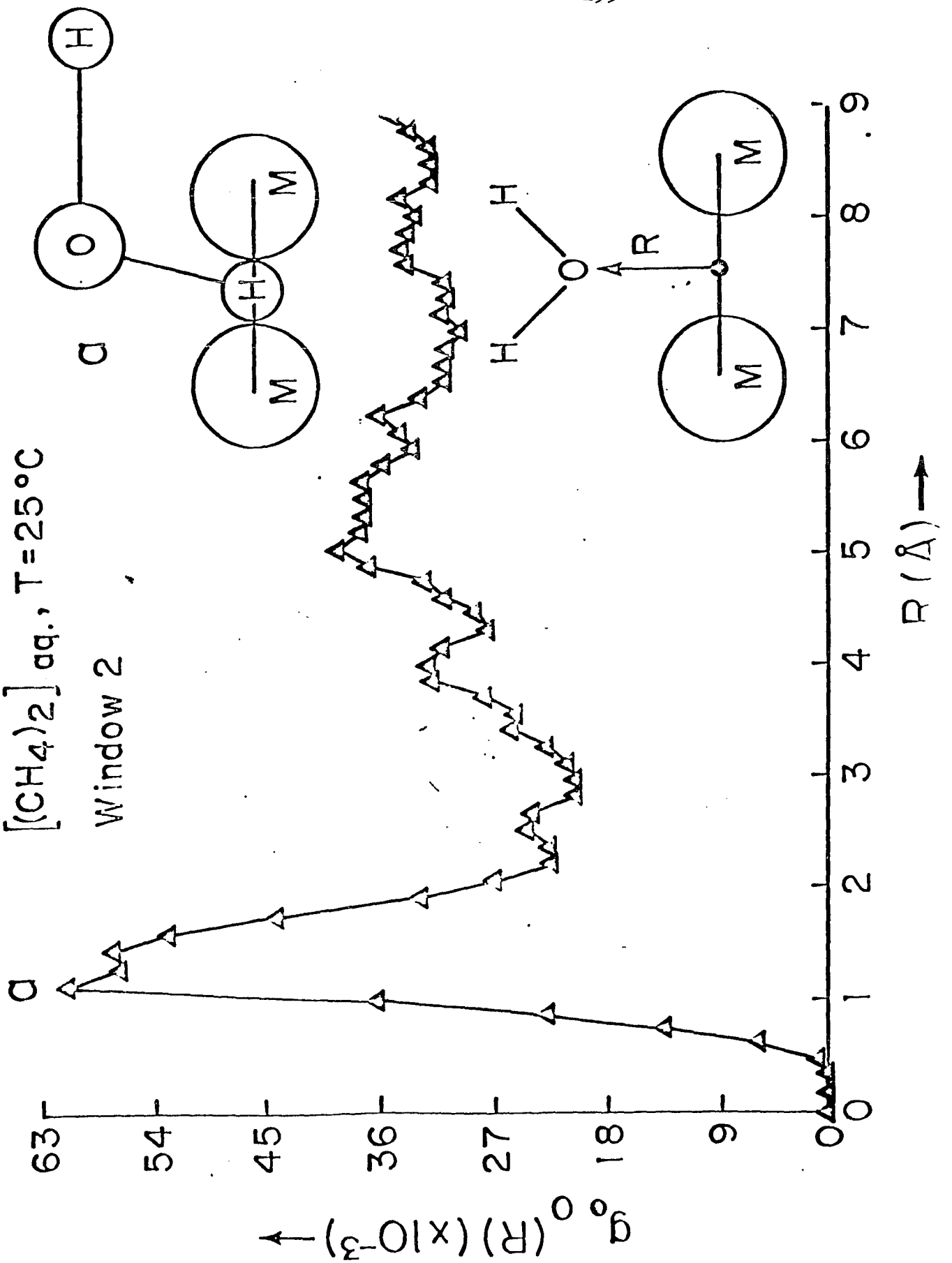

[(CH4)2] aq., T=25°C
Window 2
a
a
H
O
H
M
M
H
O
H
R
M
M
63
54
45
36
27
18
9
0
0
1
2
3
4
5
6
7
8
9
R (Å)
$g_{o\,o}(R)$ (x10^{-3})

Figure V.1Ø-Calculated $g_{\bullet H}(R)$ for $[(CH_4)_2]_{aq}$ window 2.

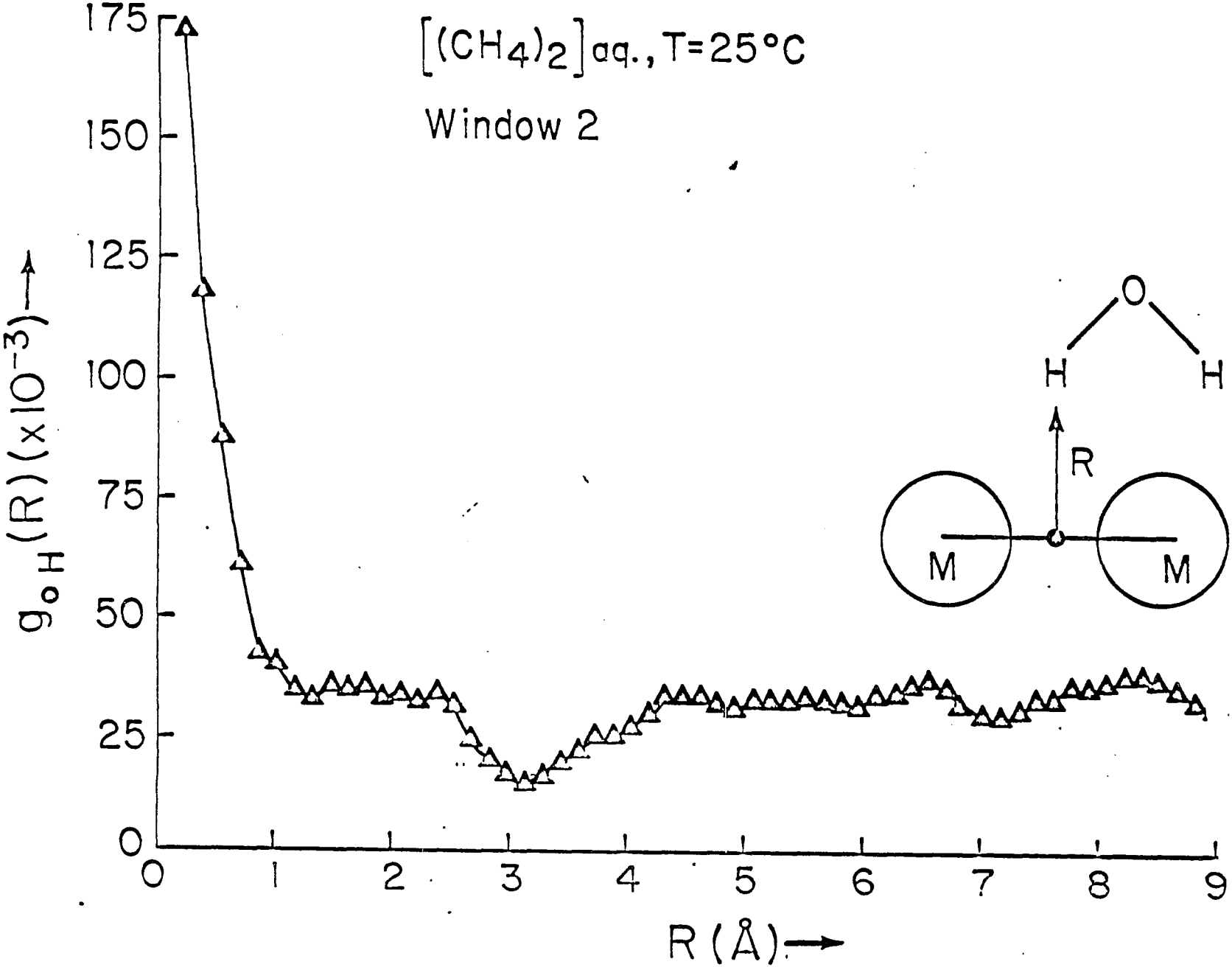
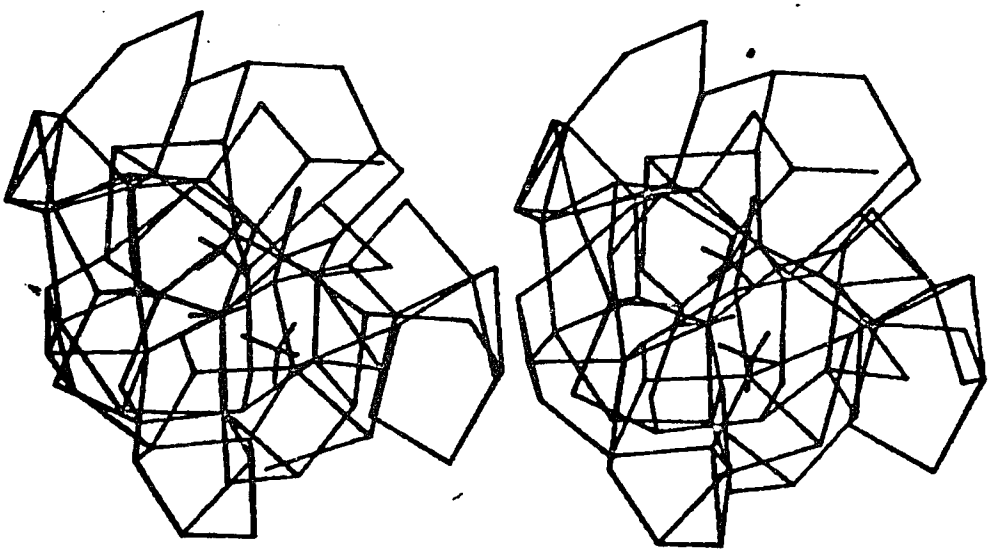

[(CH4)2]aq., T=25°C
Window 2
175
150
125
100
75
50
25
0
0
1
2
3
4
5
6
7
8
9
g°H(R) (x10⁻³)
R (Å)
O
H
H
R
M
M

Figure V.11-Typical structure corresponding to the most probable configuration in window 2.



ule intervenes between the methanes, and the peak in $g_{\bullet O}(R)$ is clearly consistent with significant contributions from this structure. A structure corresponding to this configuration is shown in Figure V.11.

Windows 3 and 4 are centered on solvent-separated values of the intersolute coordinate. The $g_{\bullet O}(R)$ and $g_{\bullet H(R)}$ for these two windows are very similar in appearence and we discuss here only the results from window 4. The $g_{\bullet O}(R)$, Figure V.12, shows a strong peak at R=Ø, corresponding to solvent-separated structures in which the water oxygen atom intervenes, and a smaller peak at R=1 Å corresponding to structures in which a water hydrogen intervenes. A solvent-separated structure extracted from the configuration is shown in Figure V.14.

In conclusion, some comments on the molecular structures in Figures V.8, V.11 and V.14 are in order. In viewing these structures, keep in mind that no one structure in a simulation is necessarily representative of the statistical state of the system, and structures arbitrarily chosen may be misleading. The urge to look at these structures is nevertheless irresistable, and we present then here along with a strong caveat emptor. The structures presented here are given as computer Dreiding models, with oxygens closer than 3.2 Å shown as bonded. Hydrogen atoms are included only for the methanes ans the solvent-separated water molecule. Only those water molecules within 6.5 Å of methane molec-

Figure V.12-Calculated $g_{\bullet O}(R)$ for $[(CH_4)_2]_{aq}$ for window 4.

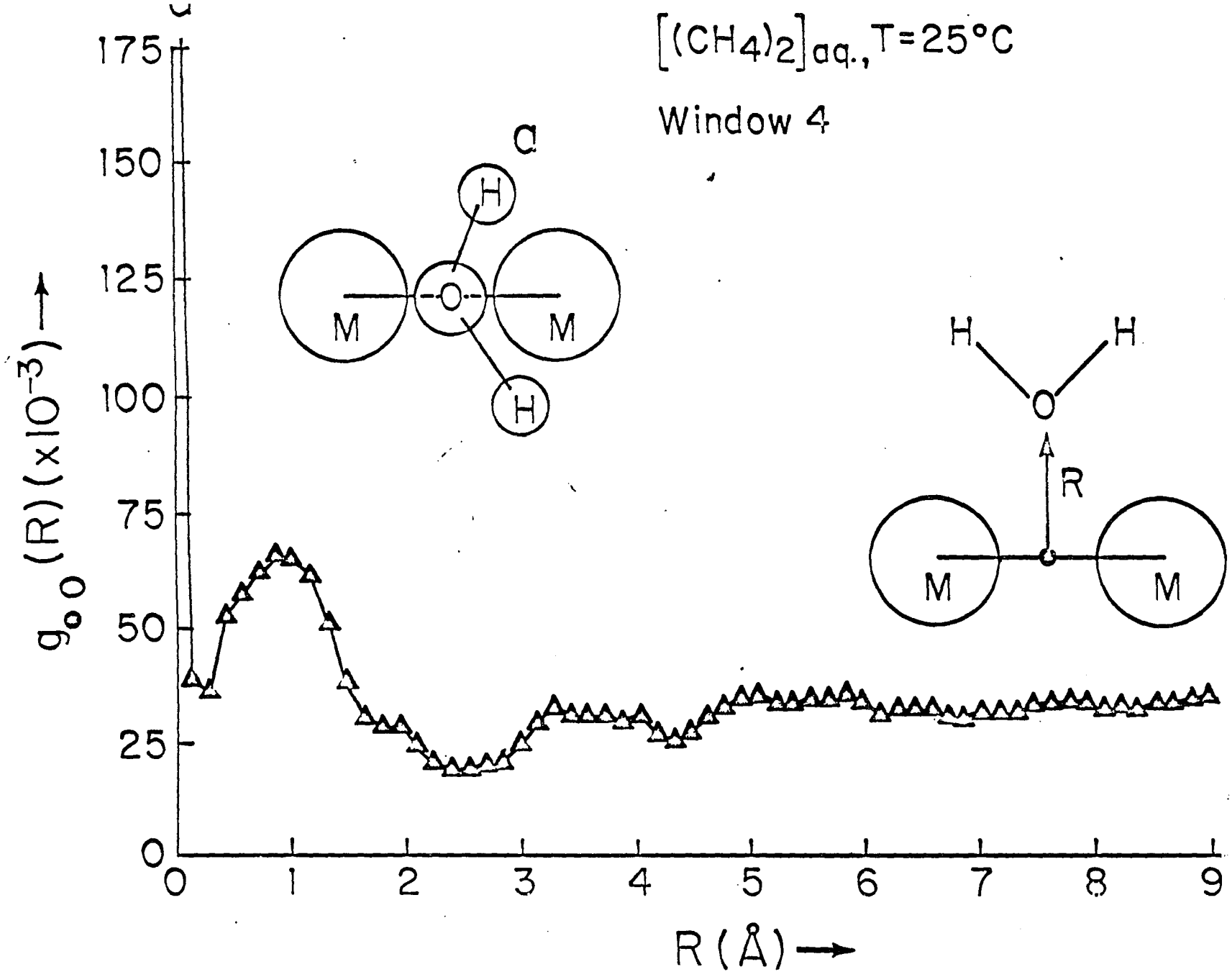
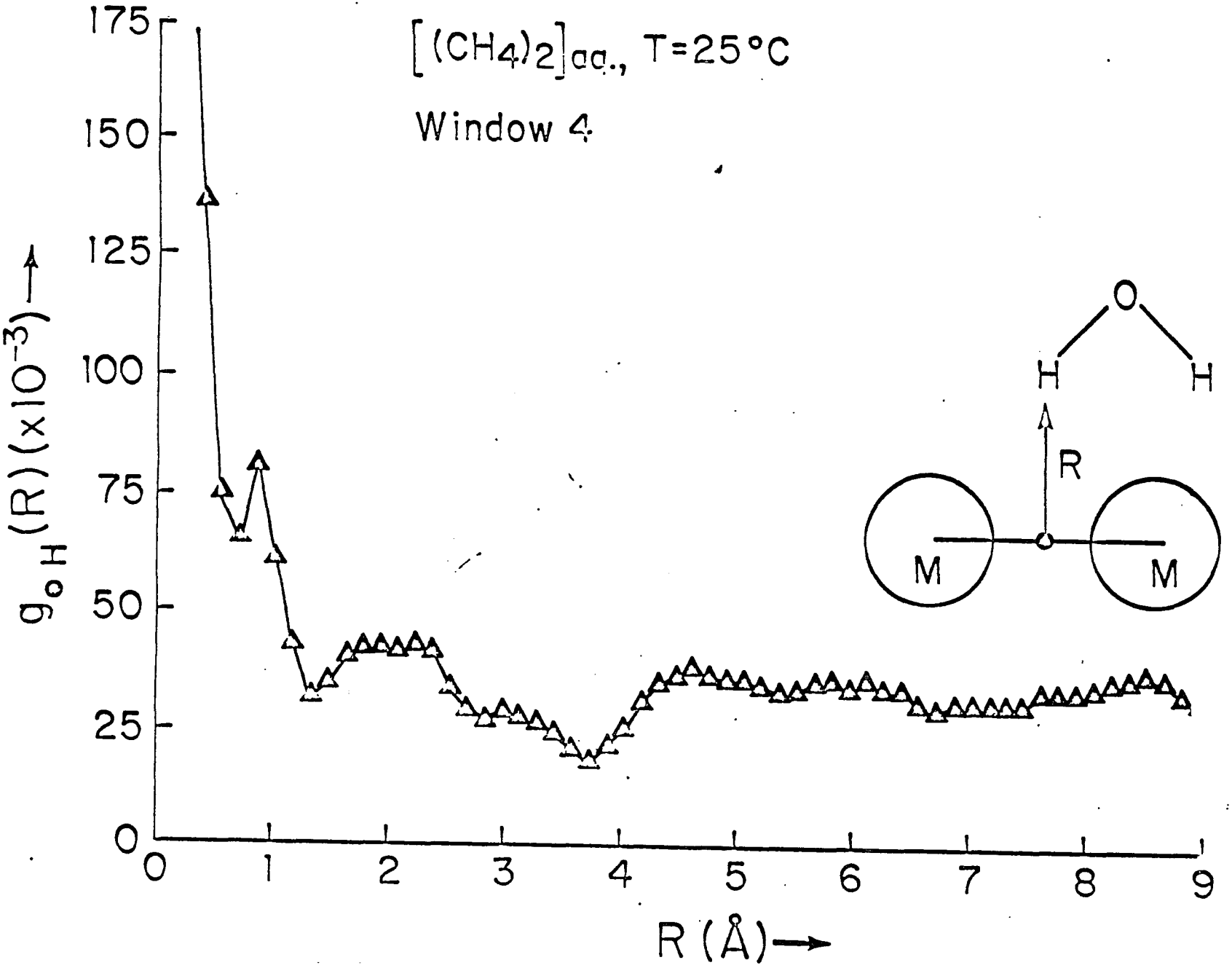

[(CH4)2]aq., T=25°C
Window 4
a
H
M
O
M
H
H
H
O
R
M
M
175
150
125
100
75
50
25
0
0
1
2
3
4
5
6
7
8
9
gOO(R) (x10-3)
R (Å)

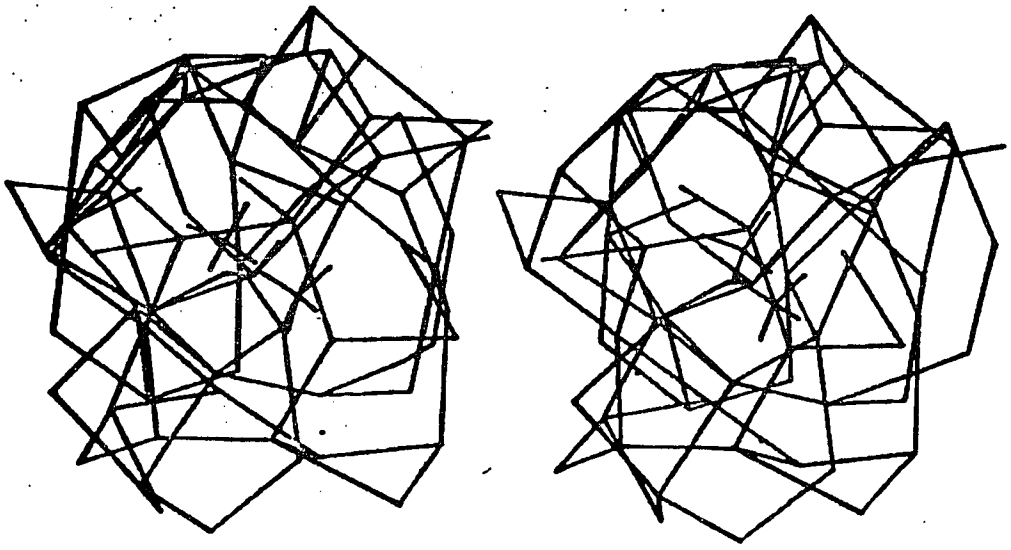
Figure V.13-Calculated $g_{\bullet H}(R)$ for $[(CH_4)_2]_{aq}$ window 4.


[(CH4)2]aq., T=25°C
Window 4
175
150
125
100
75
50
25
0
0
1
2
3
4
5
6
7
8
9
$g_{o\,H}(R)\,(x10^{-3})$
R (Å)
O
H
H
R
M
M

ules are included for clarity, which amounts to ca. 70.

Extensive networking is found in all three cases displayed. Three-,four- and five-coordinate waters are readily seen in all structures as expected. While the statistical indices of the calculations are consistent with the existence of clathrate-like contributions, the individual structures do not feature much which can be interpreted as regular clathrate polyhedra, although quite good individual pentagons can be found in all structures. The intervening water molecule in the solvent-separated structures in Figure V.11 and V.14 and the interaction of this molecule with the other waters is clearly evident.

Figure V.14-Typical structure corresponding to the most probable configuration in window 4.



CHAPTER VI.

Aqueous Hydration of Benzene.

Previous chapters have described how liquid state computer simulations can be used to study the structure and energetics of dilute aqueous solutions of CH_4, a prototype for hydrophobic hydration, and the potential of mean force between CH_4 molecules in aqueous solution, a prototype system for hydrophobic interaction. From the results of these studies, the essential clathrate-like nature of the hydrophobic hydration complex, expected from the early Frank-Evans "iceberg" ideas[5], was further delineated. Relatively high first shell coordination numbers were found to be a common characteristic of hydrophobic hydration, and were identified with water cage effects. Such contributions were also recognizable in the aqueous hydration of CH_3, CH_2 and CH groups in polyfunctional solutes studied by computer simulation[96]. To the extent these results are transferable, one can infer the structure of hydration complexes for amino acid residues in alanine, valine, leucine, isoleucine, and to some extent, proline.

Other amino acid residues such as phenylalanine and tryptophane as well as the nucleotide bases adenine, guanine, thyamine,cytosine and uracil involve unsaturated organic ring structure. These moieties are typically classified as hydrophobic, although the situation is possibly more complex than for alkyl group hydration since the polarizability of the pi-electron cloud admits the possibility of hydrophilic hydration above and below the molecular plane as well as in interactions involving the heteroatoms. An important proto-

type aqueous hydration problem for this class of molecules is the dilute aqueous solution of benzene, $[C_6H_6]_{aq}$. Benzene is of course only sparingly soluble in water. The thermodynamics of transfer have been determined and are ΔG=+5.33 kcal/mol and ΔS=-15.7 cal/deg.mol, as quoted by Edsall and Mackenzie[97]. Evidence for specific benzene-water interactions comes from studies by Backx and Goldman[98], who observed non-classical rotational behaviour of D_2O in water/benzene solutions and anticipated a weak hydrogen bond between H_2O and the benzene pi-electron cloud. Though there has been experimental work on the hydrophobic interaction of benzene molecules[57,99], there is little known at the molecular level about the nature of benzene hydration. To explore the aqueous hydration of the phenyl group and to determine the nature of the hydrophobic and hydrophilic effects in this system, we have extended our computer simulation studies to $[C_6H_6]_{aq}$ at 25°C.

While this work was in progress, Karlstrom et.al.[100] reported a new theoretical determination of an intermolecular potential function for the benzene-water interactions using quantum mechanical calculations. A novel means of correcting for the basis set superposition error was introduced. Comparison of the Karlstrom et.al. results with those of the potential function used for our study give an indication of the sensitivity of calculated interaction energies to basic assumptions in the theoretical methodology.

Calculations.

Statistical thermodynamic (T,V,N) ensemble Monte Carlo calculations were carried out on $[C_6H_6]_{aq}$ using a modified Metropolis procedure[80] incorporating the force bias method[78] and preferential sampling[89] for convergence acceleration. The system for study was comprised of 216 rigid particles, one benzene molecule and 215 water molecules. The simulation was performed at a temperature of 25°C and a density determined from the experimentally observed partial molar volume of water and benzene[99]. The condensed phase environment of the system was provided by means of face centered cubic periodic boundary conditions, which provide in excess of two complete hydration shells for the solute. Convergence characteristics and statistical error bounds on each of the calculated quantities was monitored by control functions based on the method of batch means. Full details of the Monte Carlo methodology are given in Chapter III.

The N-particle configurational energies of the system were calculated under the assumption of pairwise additivity in intermolecular interactions using potential functions determined from ab initio quantum mechanical calculations. For the water-water interactions we continue to use the MCY-CI(2) potential developed by Matsuoka et. al.[50], and representative of moderately large configuration interaction calculations on the water dimer. For the calculation of ben-

zene-water interactions, a potential function was constructed from the 12-6-1 functional form and the transferable parameters determined from minimal basis set LCAO-SCF-MO calculations by Clementi et. al.[101]. The net charge on C and H were taken to be -.2 and .2 respectively, obtained from molecular orbital calculations using Clementi's basis set in the Gaussian-80 system of programs.

For the benzene-water potential, no direct test of quality is available. Slices of the potential energy hypersurface for the benzene-water interaction computed from the potential used in this study is shown Figures VI.1 and VI.2. Benzene-water interaction energies in the molecular plane (Figure VI.1) and perpendicular to the molecular plane (Figure VI.2) are found to be worth about 3 kcal/mol stabilization energy. The function recently developed by Karlstrom et. al.[100] shows a similar placement of in-plane and out-of-plane energy minima in the benzene-water surface, with corresponding energy minima of -1.9 and -2.9 kcal/mol respectively. The discrepancy in the out-of-plane interaction energy as computed by the two approches is quite small. The difference in the in-plane binding energy, about 1 kcal/mol, is not expected to influence the results on solution structure from the simulation, but the effect of such a difference on energetics is more complicated to predict and will be discussed below. The benzene-water interaction energies in both the cases are slightly greater than the about 1 kcal/mol found for methane-water interac-

Figure VI.1-Isoenergy contour surface for benzene-water dimer in the molecular plane of benzene. Lowest energy is labelled A, and corresponds to -2.9 kcal/mol. Each successive contour increases in energy by 1.Ø kcal/mol.

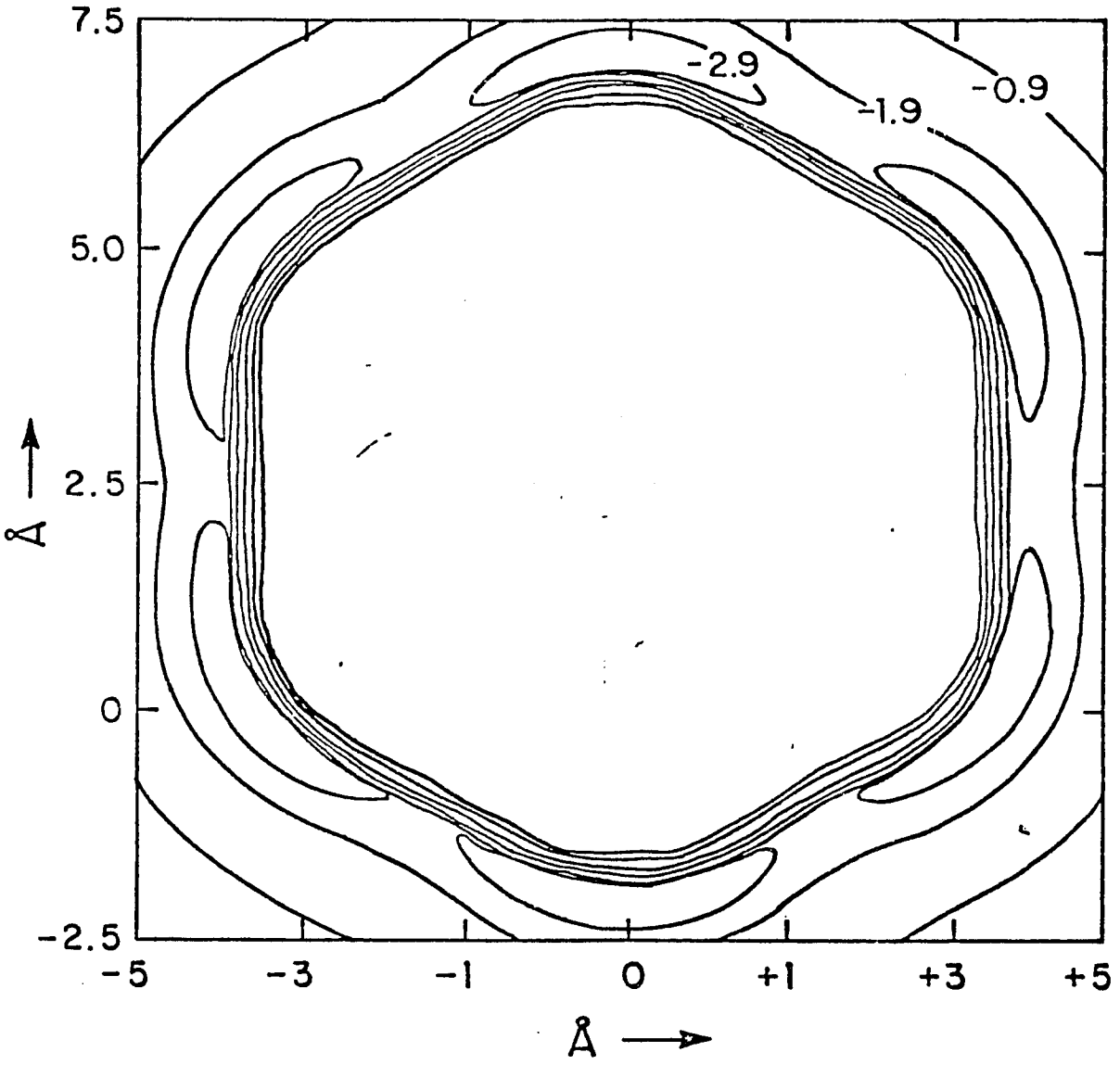
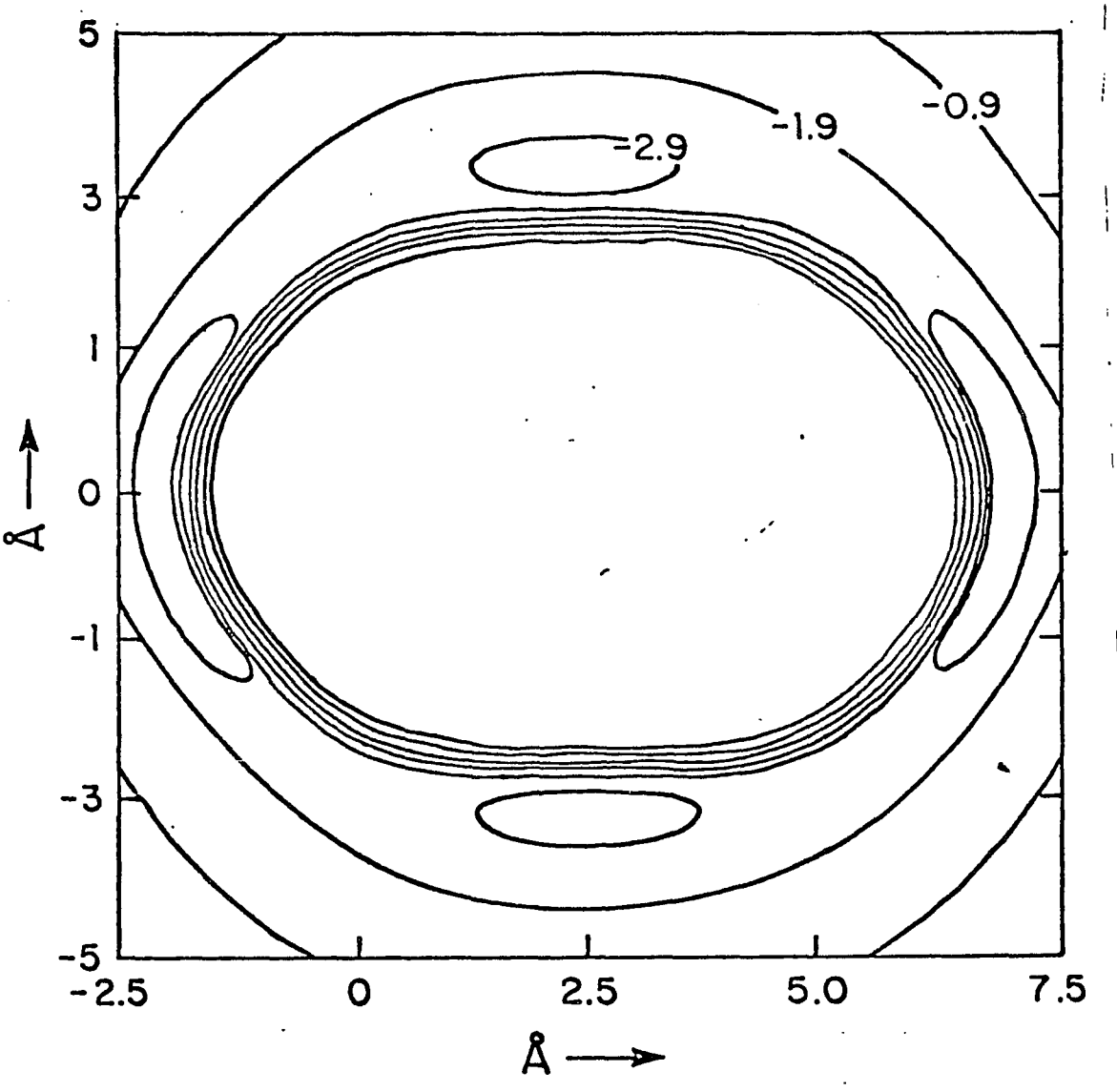

7.5
5.0
2.5
0
-2.5
Å
-5
-3
-1
0
+1
+3
+5
Å
-2.9
-1.9
-0.9

Figure VI.2-Isoenergy contour surface for benzene-water dimer in the plane perpendicular to the molecular plane of benzene, bisecting the molecule. Lowest energy is labelled A, and corresponds to -2.9 kcal/mol. Each successive contour increases in energy by 1.Ø kcal/mol.

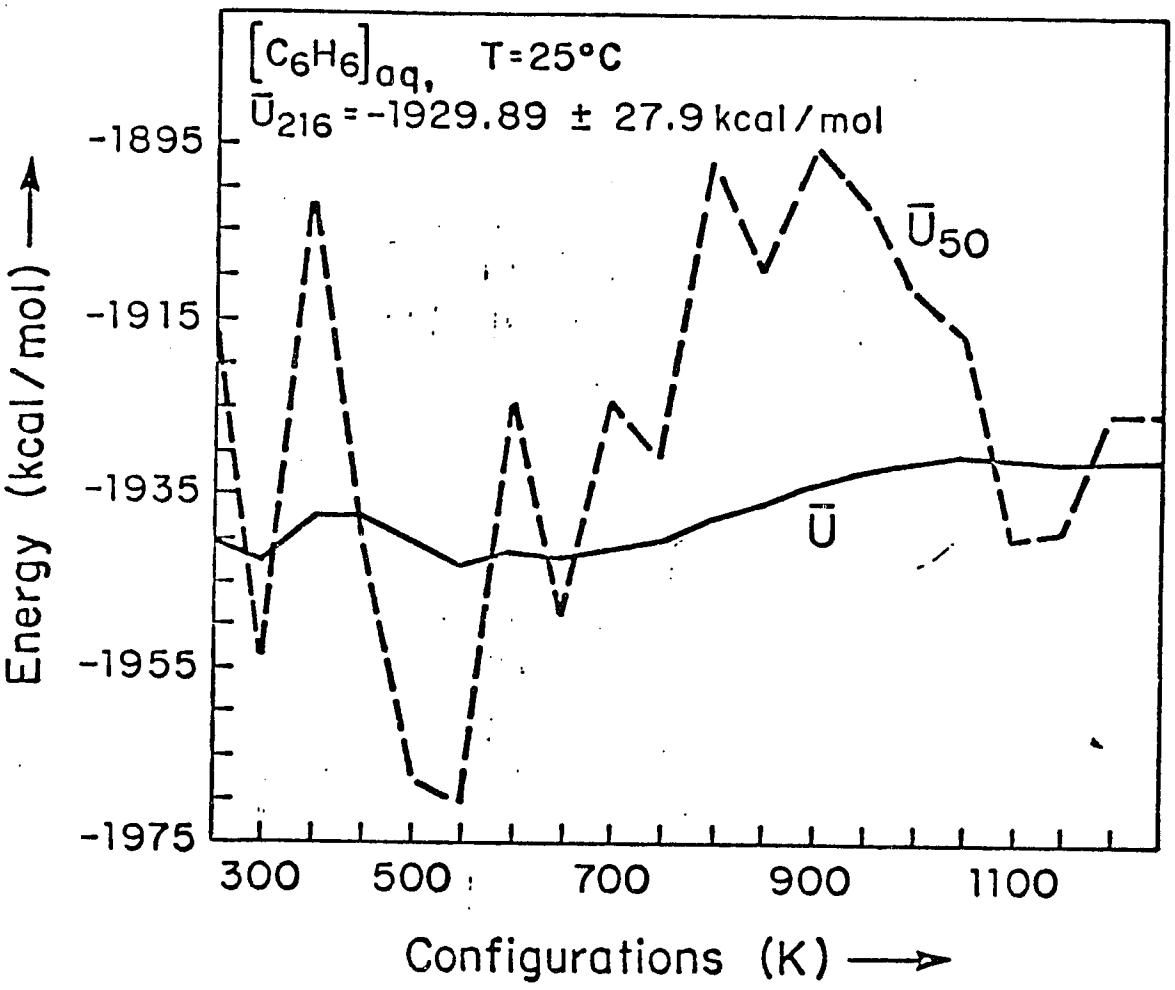

-0.9
-1.9
-2.9
5
3
1
0
-1
-3
-5
Å
-2.5
0
2.5
5.0
7.5
Å

tions and less than the 5 kcal/mol intermolecular hydrogen bond between water molecules. The benzene-water stabilization energy is therefore in a range reasonable for a hydrophobic-like dipole-induced dipole interaction or a weakly hydrophilic hydrogen bonding interaction.

In the computer simulation, all potential functions for water-water interactions were truncated at a spherical cutoff of 7.75 Å, whereas the benzene-water interaction energies were treated under the minimum image convention. No solute-solute interactions are included and the simulated system then corresponds to a dilute aqueous solution.

The complete simulation involved a total of 1700K configurations. The initial configuration was a random distribution of non-overlapping particles. The initial 700K configurations of the sampling was treated as equilibration and was discarded, and ensemble averages were formed over the remaining 1000K configurations. The convergence profile for the calculation is shown in Figure VI.3.

Figure VI.3-Convergence profile for the force-bias, preferential sampling augmented Monte Carlo simulation on $[C_6H_6]_{aq}$. Mean energy is denoted as Θ and mean energy for batches of size 50K is denoted as $\overline{\Theta}_{50}$.


$[C_6H_6]_{aq}$, T=25°C
$\bar{U}_{216}$ = -1929.89 ± 27.9 kcal/mol
$\bar{U}_{50}$
$\bar{U}$
-1895
-1915
-1935
-1955
-1975
Energy (kcal/mol)
300
500
700
900
1100
Configurations (K)

Results.

The calculated internal energies and related quantities for $[C_6H_6]_{aq}$ are collected in TableVI.1 The quantities entered here are the mean energy U of the system (N_S=1, N_W=215), the energy U_W of 215 water molecules in $[H_2O]_1$ at 25°C, $U_{W,,}$ the corresponding energy of solvent water in $[C_6H_6]_{aq}$, $\bar{U}_S$, the calculated partial molar internal energy of transfer for C_6H_6 into water, and finally $\bar{U}_S$, and U_{rel} the solute-solvent and solvent-solvent contributions to $\bar{U}_S$. Each of these is formally defined in Equations 1-12 and Figure VI.4 of a previous paper from this laboratory by Swaminathan et.al.[53] The calculated molecular distribution functions and analysis thereof for $[C_6H_6]_{aq}$ at 25°C follow. The analysis formalism follows that described by Mehrotra and Beveridge[102], except where noted. The results are displayed first on a solute atom-by-atom basis, then developed in terms of C-H and C_6 fragments, and finally extended to indices referred to the entire C_6H_6 solute molecule. The interpretation and implications of the results are discussed in the following section.

The solute atom- solvent water radial distribution functions are described in the following paragraphs. Two forms of these functions are presented for each atom: a) a "total" solute stom-water radial distribution function $g_{AW}^{tot}(R)$, conventionally defined, and b) a solute

Table VI-1 Calculated Internal Energies for the Dilute Aqueous Solution of Benzene at 25°C in kcal/mol.

	A	B
U_{SW} (N_W=215,N_S=1)	-1929.89	-1861.59
U_W (N_W=215)	-1859.75	-1859.75
U_W, (N_W=215)	-1869.29	-1864.42
$\bar{U}_S$,	-60.60	2.83
$\bar{U}_{rel}$	-9.55	-4.77
$\bar{U}_S$	-70.15	-1.86

A- Results from simulation as described in Calculations section. B- Results from simulation based on benzene-water potential with attractive region set to zero everywhere.

atom- water $g^{1^o}_{AW}(R)$ describing only those solvent water molecules designated "primary" to the solute atom based on the proximity criterion, i.e. those waters closer to that atom than to any other. This g(R) is renormalized to the volume element of the truncated spherical shell of the Voronoi polyhedron associated with the primary region of the solute atom. The $g^{tot}_{AW}(R)$ and $g^{1^o}_{AW}(R)$ are collected for each solute atom on a single graph, together with the corresponding running coordination numbers.

The calculated solute-water radial distribution functions of the benzene carbon and hydrogen atom in $[C_6H_6]_{aq}$ are shown in Figure VI.4. All solute atom-water radial distribution functions refer to the center-of-mass of water molecules unless otherwise noted. The total $g_{CW}(R)$ and the $g^{1^o}_{CW}(R)$ for carbon, symmetry averaged over the six carbon atoms, and the corresponding running coordination numbers are given in Figure VI.4a. The total g(R) shows two peaks, each relatively broad, since both in-plane and out-of-plane solvent molecules are included together. The $g_{CW}(R)$ for carbon by virtue of the proximity criterion, describes mainly those water molecules above and below the C_6 hexagon in benzene. A well defined first shell with a maximum value at 3.2 $\mathring{A}$ is evident. Integrating this shell up to R_C=4.5 $\mathring{A}$ gives a value of .37 water molecules per carbon atom in this region.

The total and primary radial distribution functions and

running coordination number for the benzene hydrogen atom, also symmetry averaged, are shown in Figure VI.4b. The $g^{1}_{HW}(R)$ is of most interest, and shows a maximum at 2.2 Å. The peak decays slowly to minimum at4.Ø Å. The area under the curve indicates 3.42 waters are included in the first shell for each hydrogen. Here a composite of in-plane and out-of-plane solvation is reflected even in the $g^{1}_{HW}(R)$.

We turn now to an atom by atom analysis of the primary solvation of the carbon and hydrogen atoms of benzene in aqueous solution by means of quasicomponent distribution functions. This analysis is based on the $g^{1}_{AW}(R)$ and the R_C values discussed above. The distribution $X_C(K)$ of primary solvent coordination numbers K for benzene carbon and hydrogen atoms is shown in Figure VI.5. For the carbon atom, Figure VI.5a, the distribution ranges from Ø to 2 with an average K = Ø.39. For hydrogen, Figure VI.5b, the distribution ranges from 2 to 5, with 3 and 4 being the maximum contributors, and $\bar{K}$=3.23.

The distribution $X_B(\nu)$ of binding energies for water molecules primary to the benzene atoms in $[C_6H_6]_{aq}$ is shown in Figure VI.6. For the carbon atom, Figure VI.6a, the distribution ranges from -4.5 kcal/mol to Ø.5 kcal/mol with $\bar{\nu}$=1.Ø7 kcal/mol. For hydrogen the range is from -14.Ø kcal/mol to -4.Ø kcal/mol with $\bar{\nu}$=-9.46 kcal/mol.

Figure VI.4-Calculated total(----) and primary(- - -) solute-solvent radial distribution functions and the corresponding running coordination numbers on an atom by atom basis in $[C_6H_6]_{aq}$. These distributions are symmetry averaged over six carbon atoms(a) and six hydrogen atoms(b) of benzene.

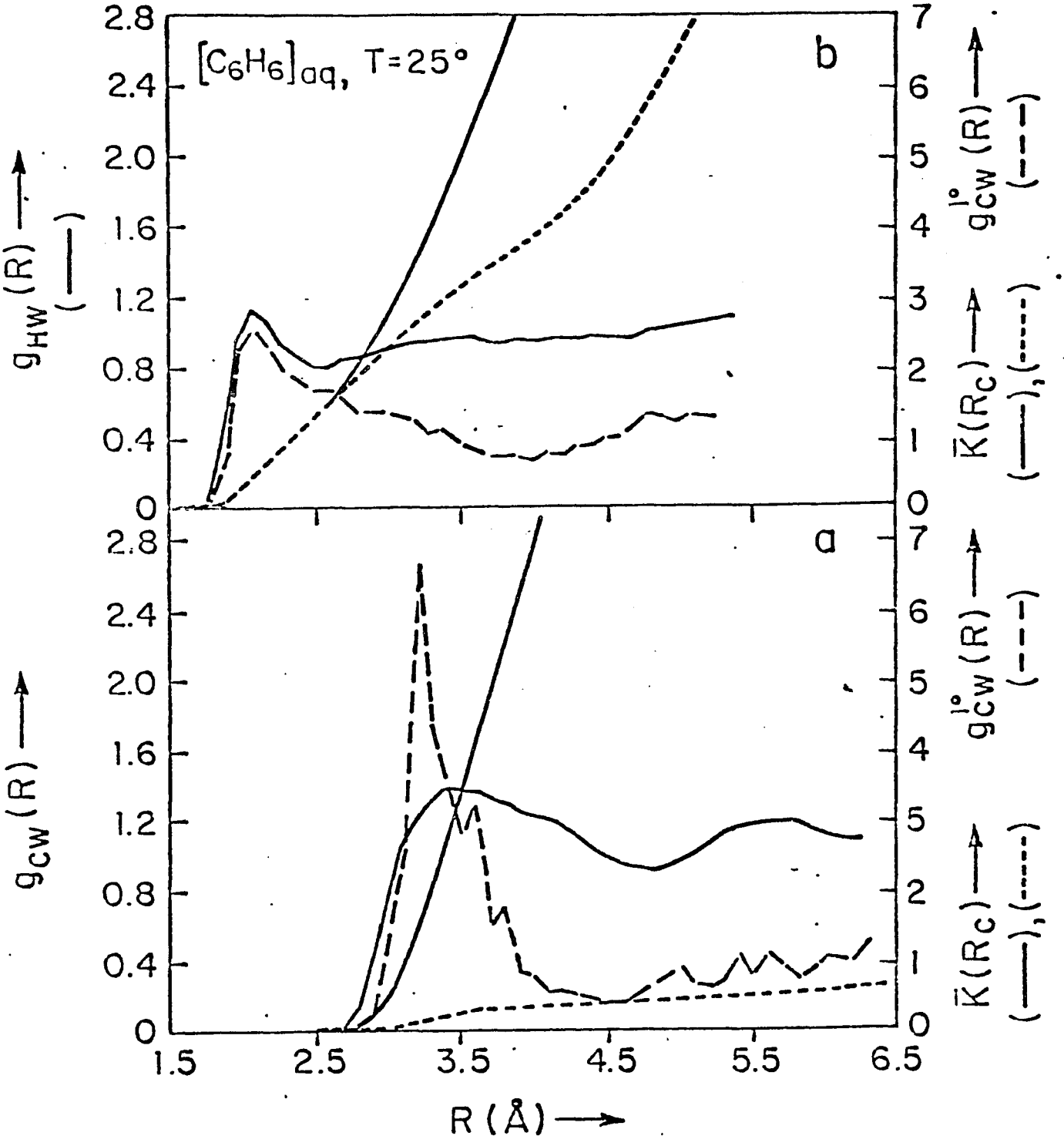

[C6H6]aq, T=25°
b
a
gHW(R) (——)
gCW(R)
g°CW(R) (– – –)
K̄(Rc) (——), (– – –)
R (Å)
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
7
6
5
4
3
2
1
1.5
2.5
3.5
4.5
5.5
6.5

Figure VI.5-Calculated symmetry averaged QCDF for primary solute-solvent coordination number on an atom by atom basis in $[C_6H_6]_{aq}$.

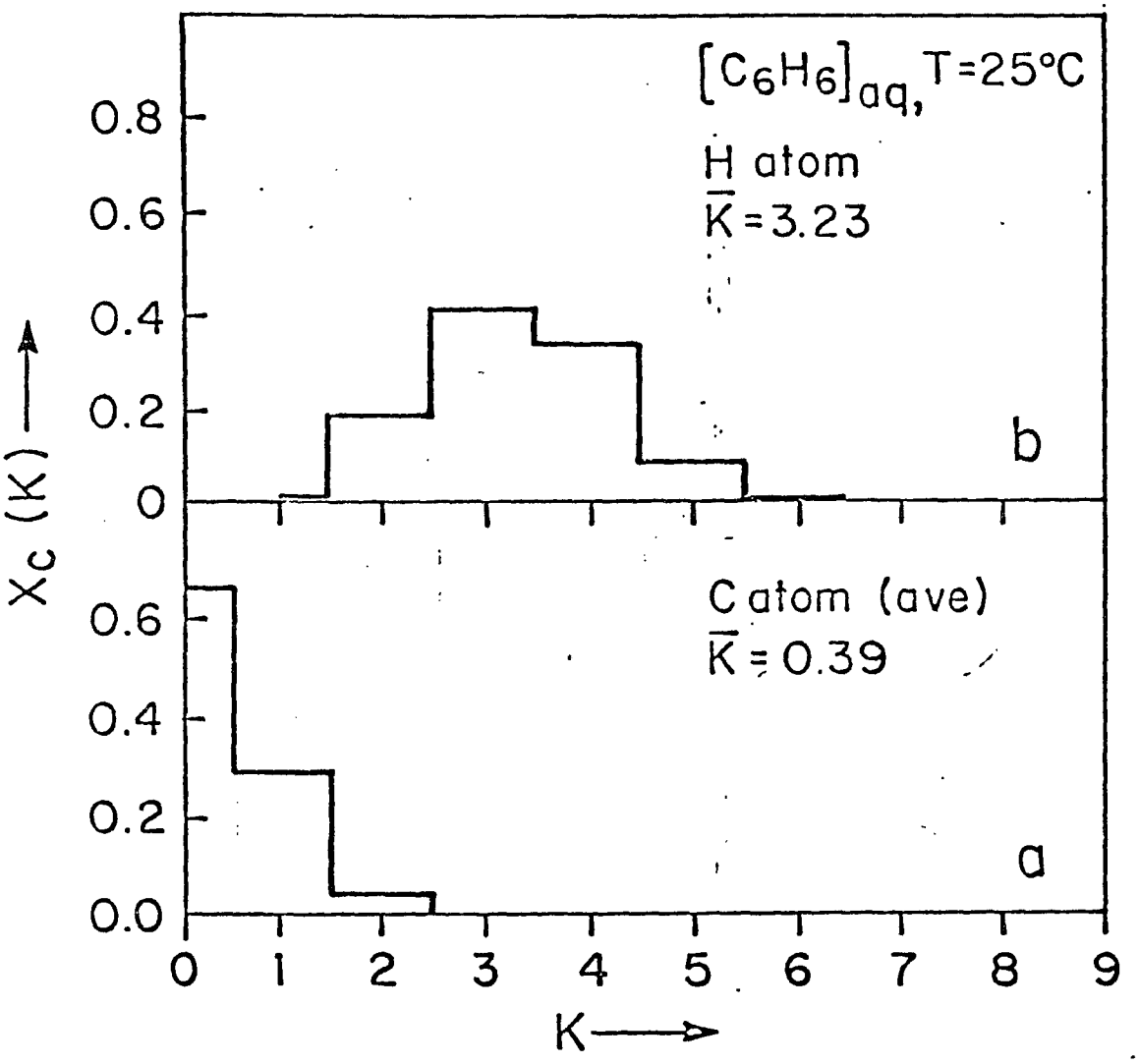
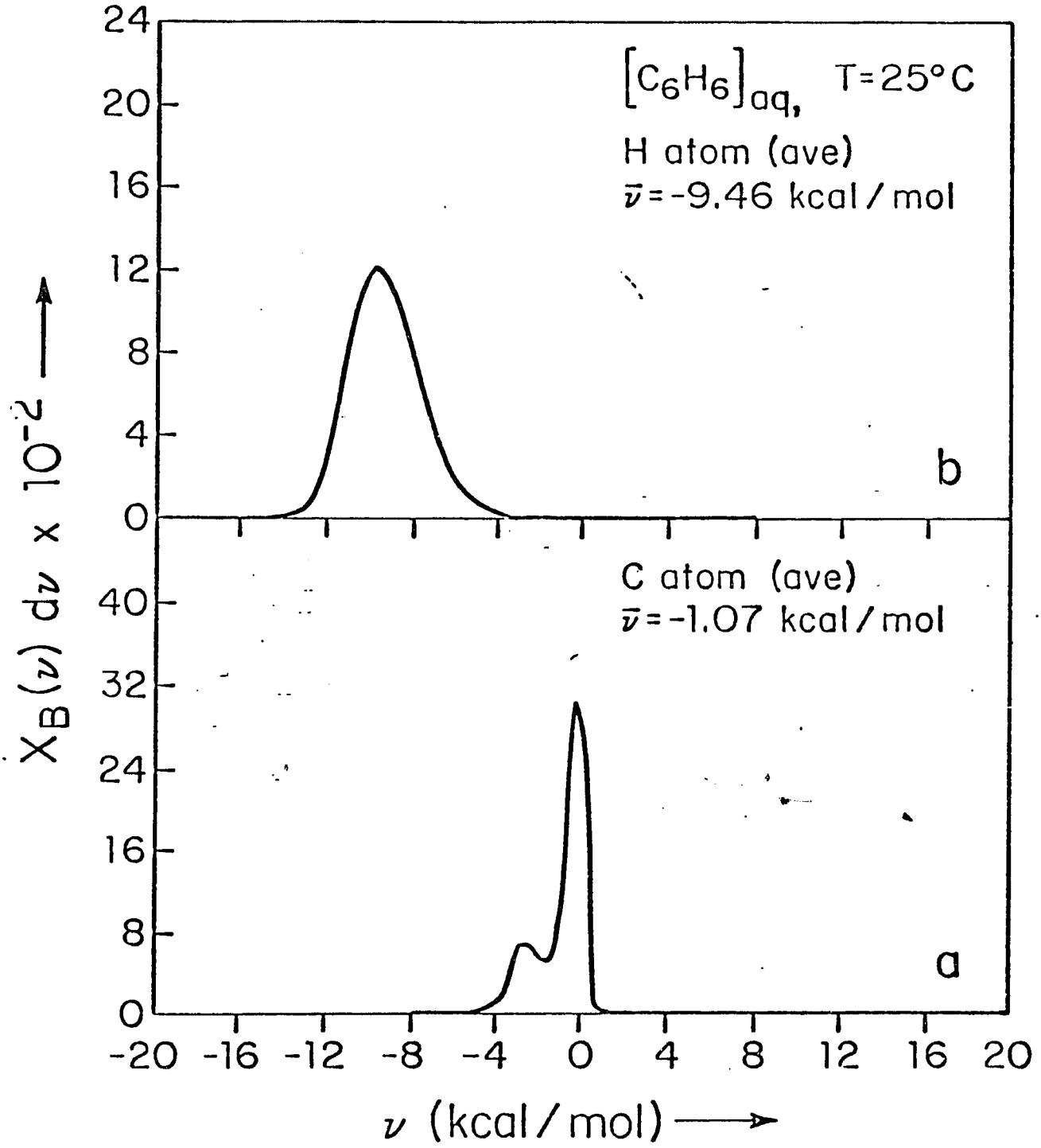

[C6H6]aq, T=25°C
H atom
K̄=3.23
b
C atom (ave)
K̄=0.39
a
Xc (K)
K
0.8
0.6
0.4
0.2
0
0.6
0.4
0.2
0.0
0 1 2 3 4 5 6 7 8 9

Figure VI.6-Calculated symmetry averaged QCDF for primary solute-solvent binding energy on an atom by atom basis in $[C_6H_6]_{aq}$.


$[C_6H_6]_{aq}$, T=25°C
H atom (ave)
$\bar{\nu}$ = -9.46 kcal / mol
b
C atom (ave)
$\bar{\nu}$ = -1.07 kcal / mol
a
$X_B(\nu)\, d\nu \times 10^{-2}$
ν (kcal / mol)
24
20
16
12
8
4
0
40
32
24
16
8
0
-20 -16 -12 -8 -4 0 4 8 12 16 20

The computed distribution functions $X_p(\epsilon)$ for solute-water pair interaction energy ϵ are given in Figure VI.7. For carbon, Figure VI.7a, the distribution ranges from -3.6 kcal/mol to 1.8 kcal/mol with the most probable ϵ value being -2.5 kcal/mol and with $\bar{\epsilon}$=-1.64 kcal/mol. For hydrogen, $X_p(\epsilon)$ ranges from -4 to 2.5 kcal/mol, with the most probable ϵ being -1.2 kcal/mol and $\bar{\epsilon}$=-1.Ø2 kcal/mol. The out-of-plane benzene-water interactions are slightly stronger energetically than the in-plane interactions.

The above results can be combined to produce a description of the local solution environment of benzene in $[C_6H_6]_{aq}$ in terms of the C-H group and the C_6 fragment and also the entire C_6H_6 molecule. The groupwise distributions for coordination numbers, binding energies and pair interaction energies are shown in Figures VI.8 ,VI.9 and VI.1Ø respectively. Of particular interest is the C_6 coordination number distribution, with contributions from K=Ø,1,2 and 3 and $\bar{K}$=2.34. This shows that essentially one water molecule above the plane and one below comprise the first hydration shell of the pi cloud of benzene in $[C_6H_6]_{aq}$, with a corresponding pair interaction energy placed on the average at about -3. kcal/mol and a distribution favoring bound values.

The distribution of the various analysis quantities referred to the entire molecule is shown in Fig-

Fig. 7-Calculated symmetry averaged QCDF for primary solute-solvent pair energy on an atom by atom basis in $[C_6H_6]_{aq}$.

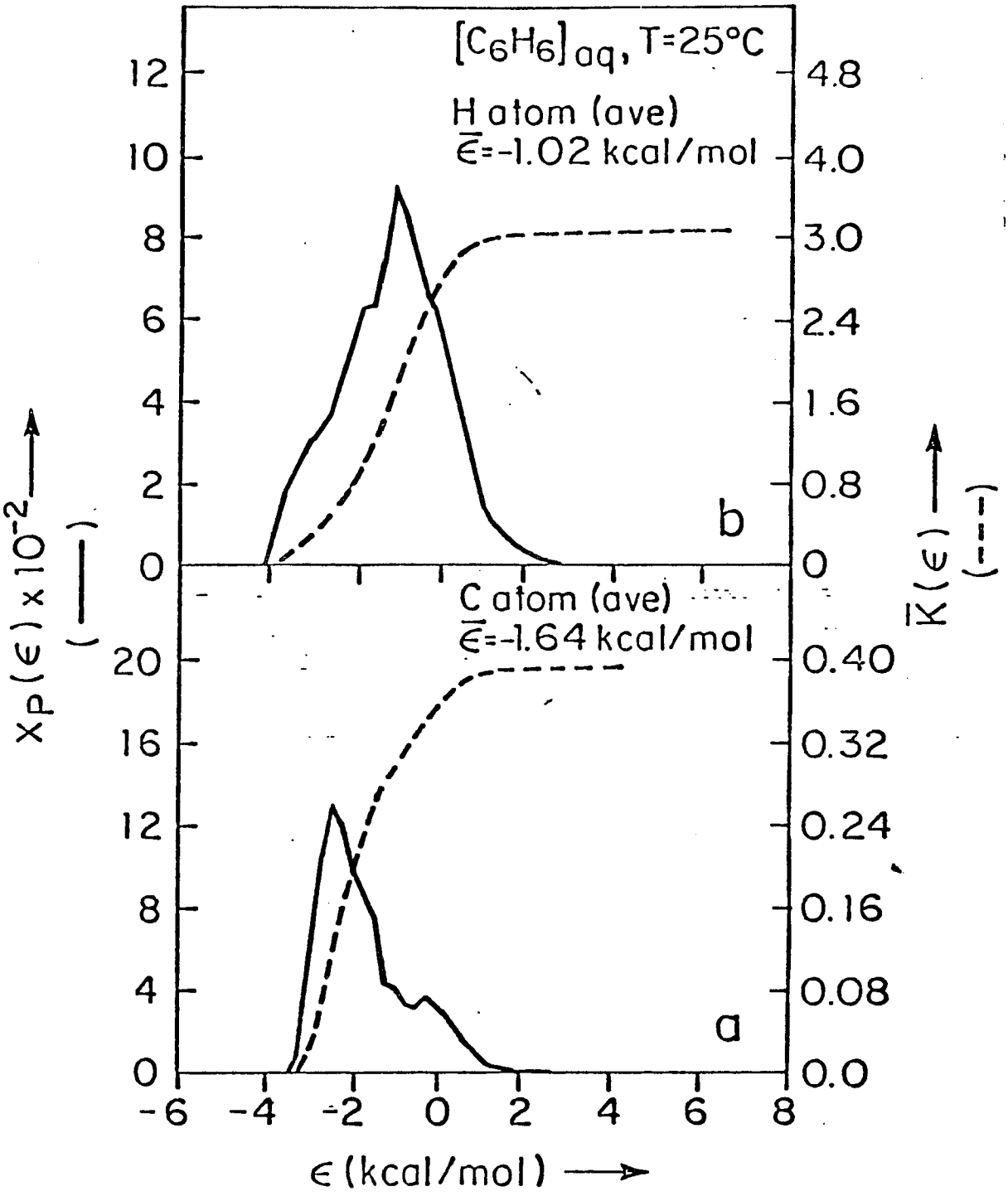

[C6H6]aq, T=25°C
H atom (ave)
ε̄=-1.02 kcal/mol
b
C atom (ave)
ε̄=-1.64 kcal/mol
a
Xp(ε) x 10^-2 (——)
K(ε) (- - -)
ε (kcal/mol) →
12
10
8
6
4
2
0
20
16
12
8
4
0
4.8
4.0
3.0
2.4
1.6
0.8
0
0.40
0.32
0.24
0.16
0.08
0.0
-6
-4
-2
0
2
4
6
8

Figure VI.8-Calculated symmetry averaged QCDF for primary solute-solvent coordination number on a functional group basis in $[C_6H_6]_{aq}$.

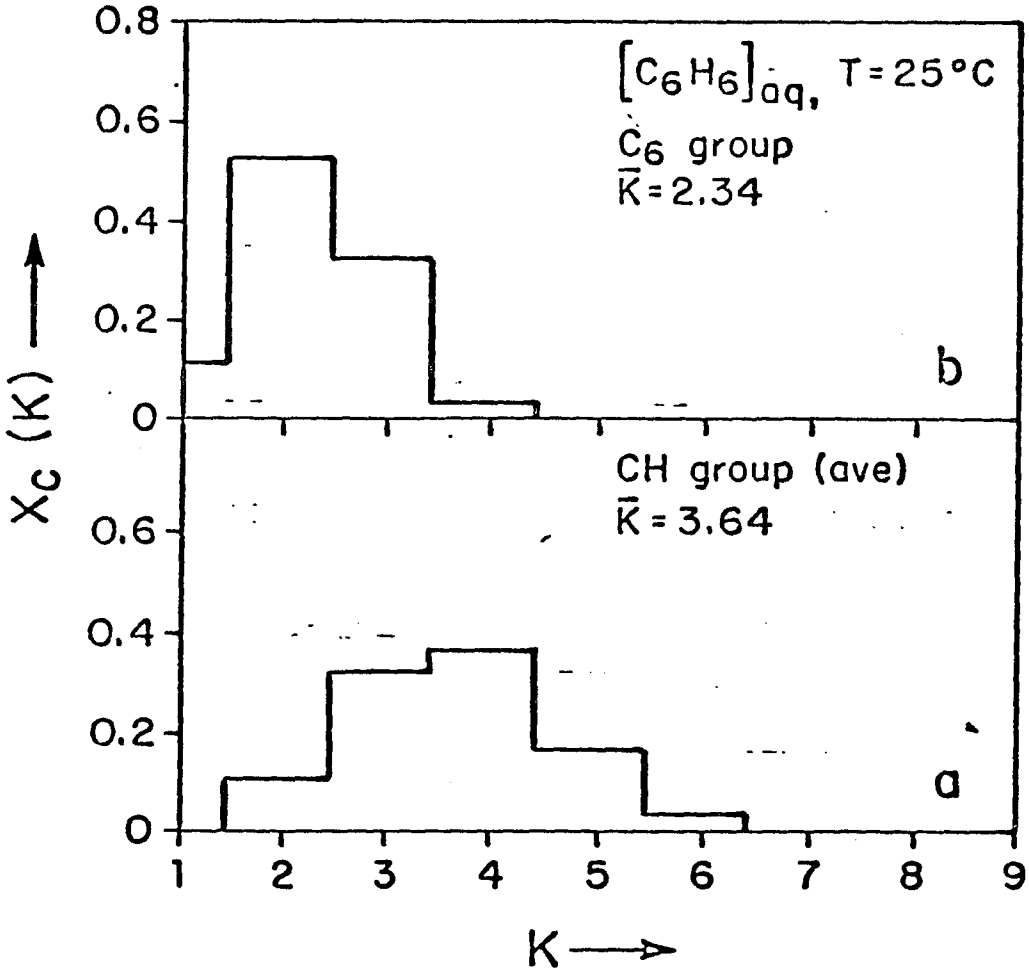

$[C_6H_6]_{aq}$, T = 25°C
$\bar{C}_6$ group
$\bar{K}$ = 2.34
b
CH group (ave)
$\bar{K}$ = 3.64
a
X_c (K)
K
0.8
0.6
0.4
0.2
0
1
2
3
4
5
6
7
8
9

Figure VI.9-Calculated symmetry averaged QCDF for primary solute-solvent binding energy on a functional group basis in $[C_6H_6]_{aq}$.

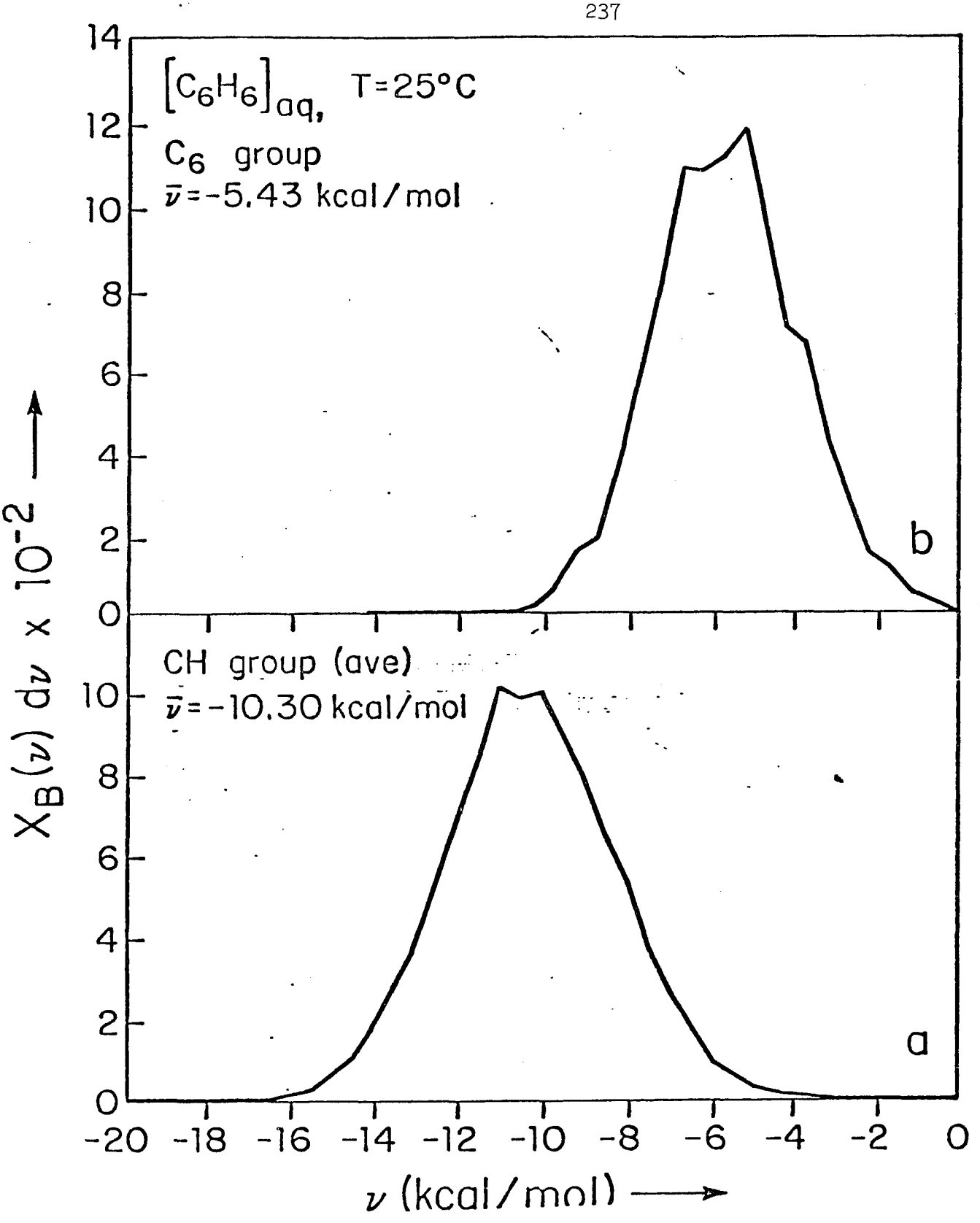

$[C_6H_6]_{aq}$, T=25°C
C_6 group
$\bar{\nu}$=-5.43 kcal/mol
b
CH group (ave)
$\bar{\nu}$=-10.30 kcal/mol
a
$X_B(\nu)\, d\nu \times 10^{-2}$
ν (kcal/mol)
14
12
10
8
6
4
2
0
-20
-18
-16
-14
-12
-10
-8
-6
-4
-2
0

Fig. 10-Calculated symmetry averaged QCDF for primary solute-solvent pair energy on a functional group basis in $[C_6H_6]_{aq}$.

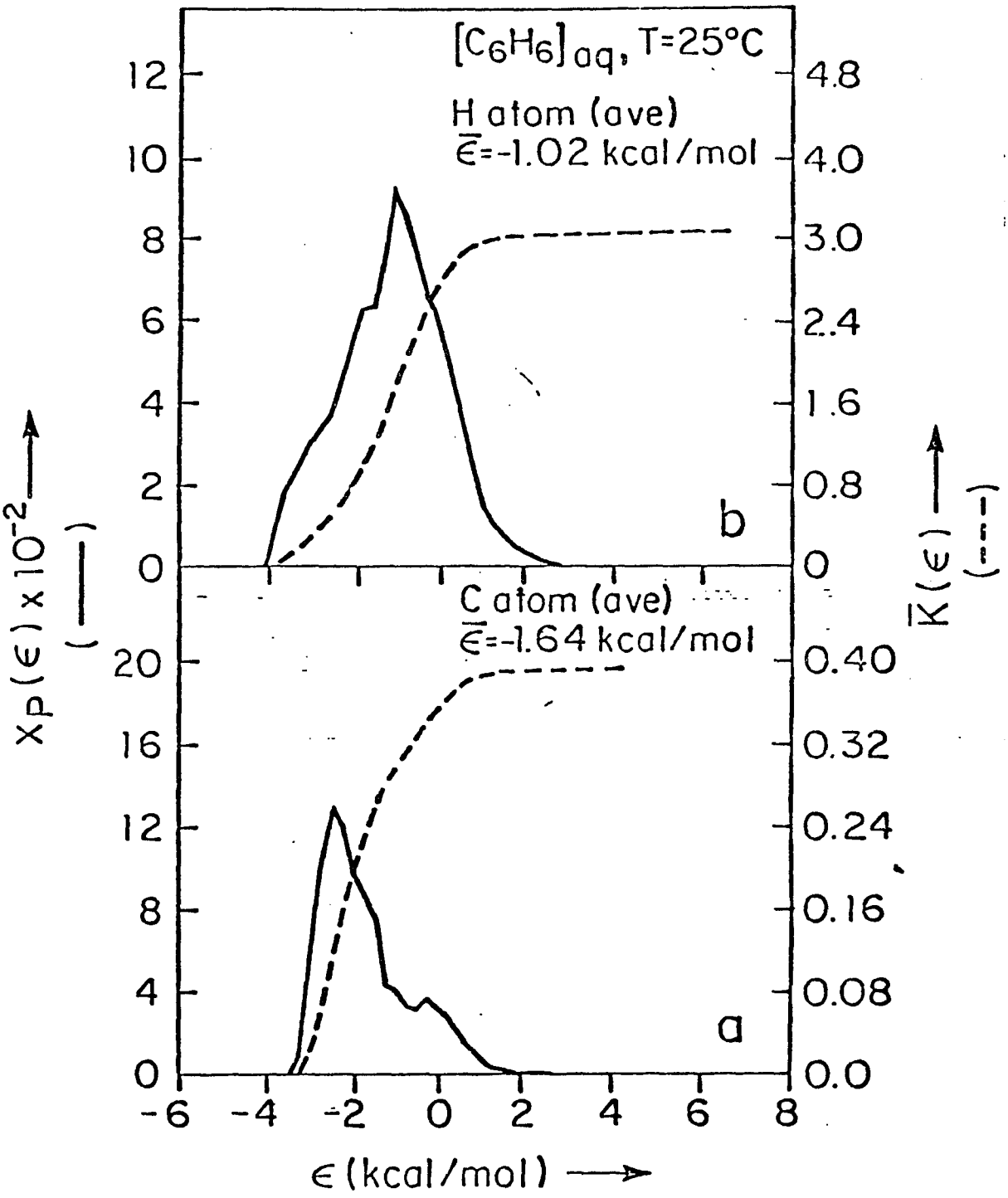

[C6H6]aq, T=25°C
H atom (ave)
ε̄=-1.02 kcal/mol
b
C atom (ave)
ε̄=-1.64 kcal/mol
a
Xp(ε) x 10^-2 (——)
K̄(ε) (- - -)
ε (kcal/mol) ⟶
12
10
8
6
4
2
0
20
16
12
8
4
0
4.8
4.0
3.0
2.4
1.6
0.8
0
0.40
0.32
0.24
0.16
0.08
0.0
-6
-4
-2
0
2
4
6
8

Figure VI.11-Calculated radial distribution function and the corresponding running coordination number on a molecular basis in $[C_6H_6]_{aq}$.

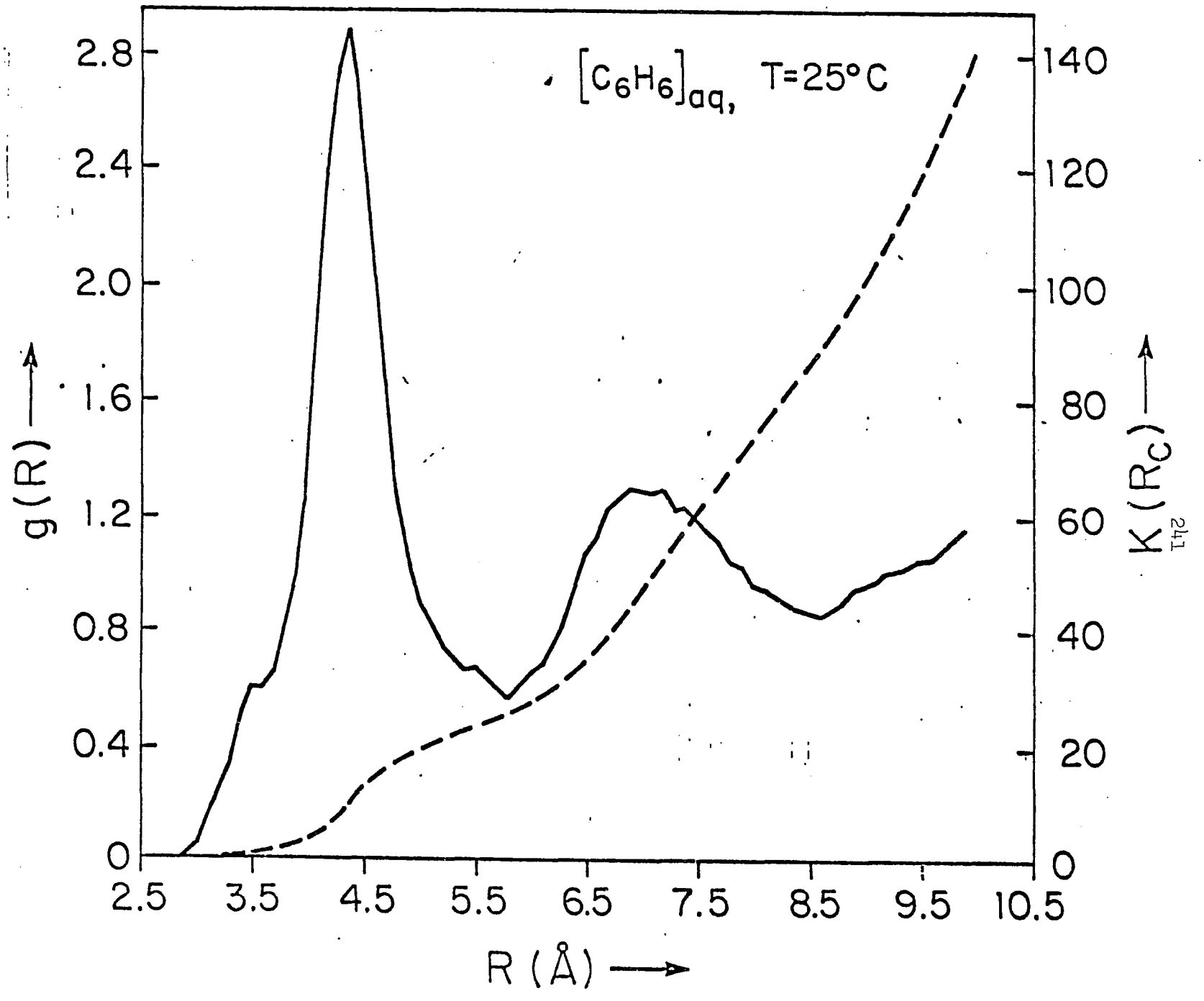

$[C_6H_6]_{aq}$, T=25°C
g(R)
$K(R_C)$
R (Å)
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0
140
120
100
80
60
40
20
0
2.5
3.5
4.5
5.5
6.5
7.5
8.5
9.5
10.5

Figure VI.12-Calculated QCDF for primary solute-solvent coordination number on a molecular basis in $[C_6H_6]_{aq}$.

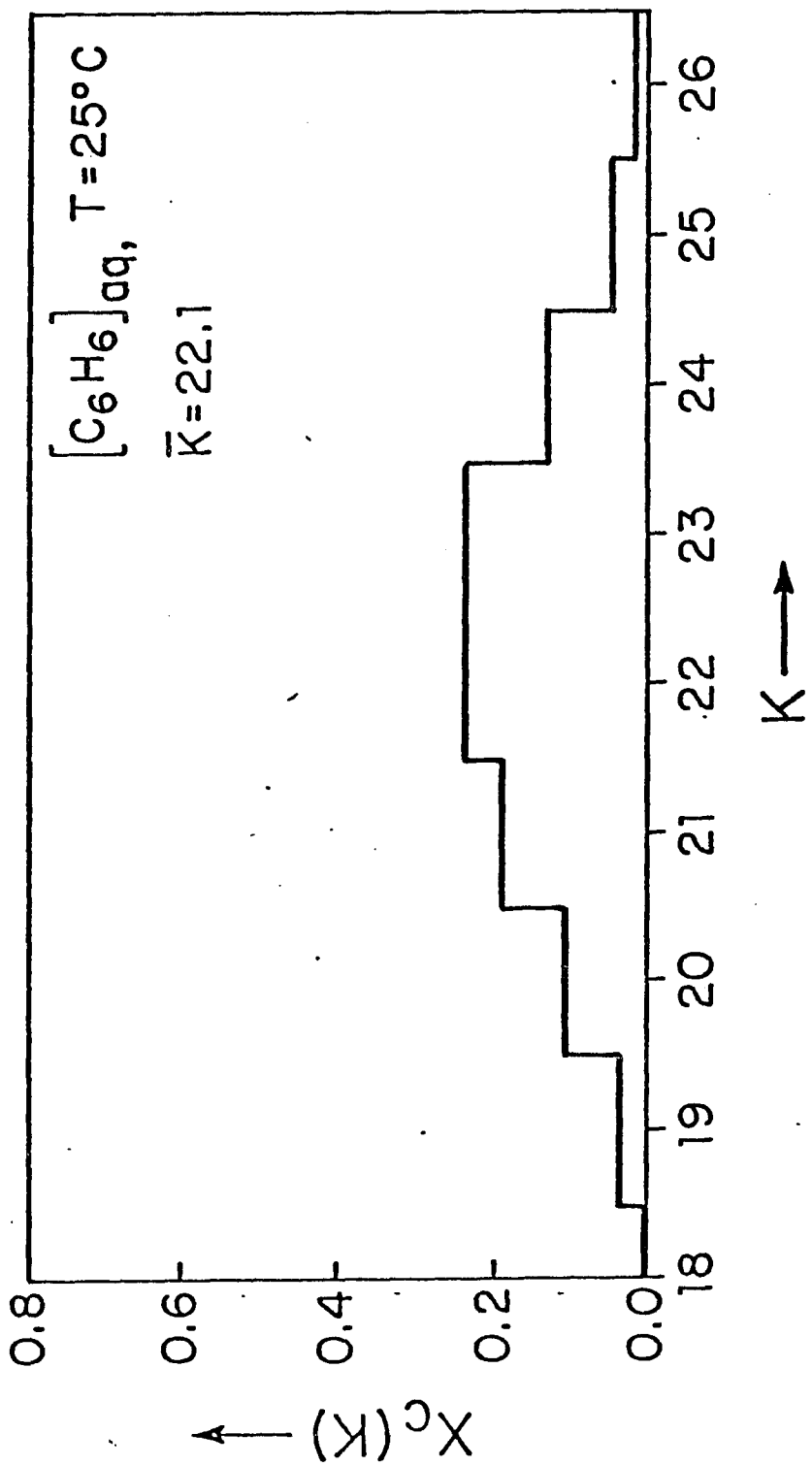

$[C_6H_6]_{aq}$, T=25°C
$\bar{K}$=22.1
$X_C(K)$
0.0
0.2
0.4
0.6
0.8
18
19
20
21
22
23
24
25
26
K

Figure VI.13-Calculated QCDF for primary solute-solvent binding energy on a molecular basis in $[C_6H_6]_{aq}$.

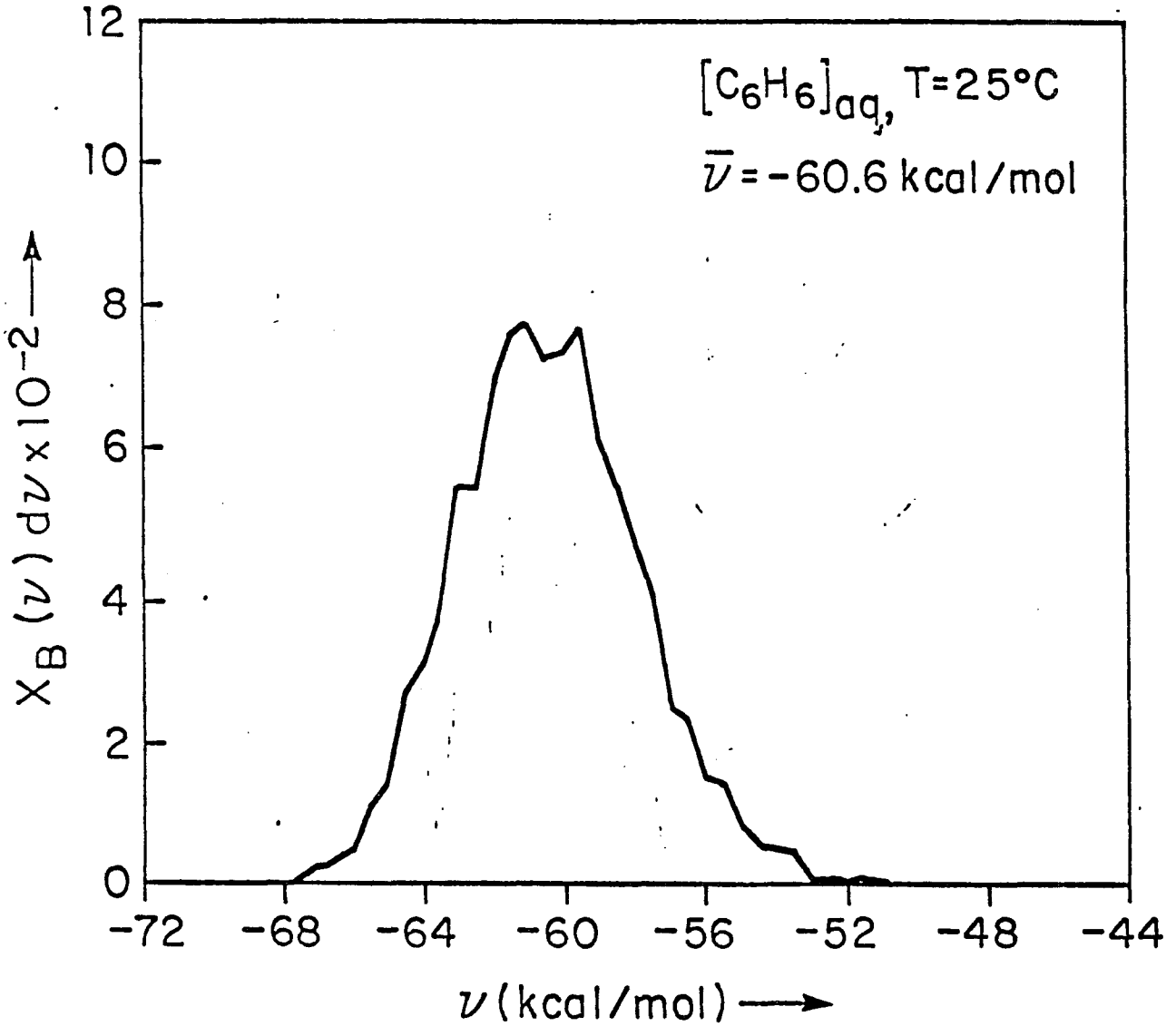
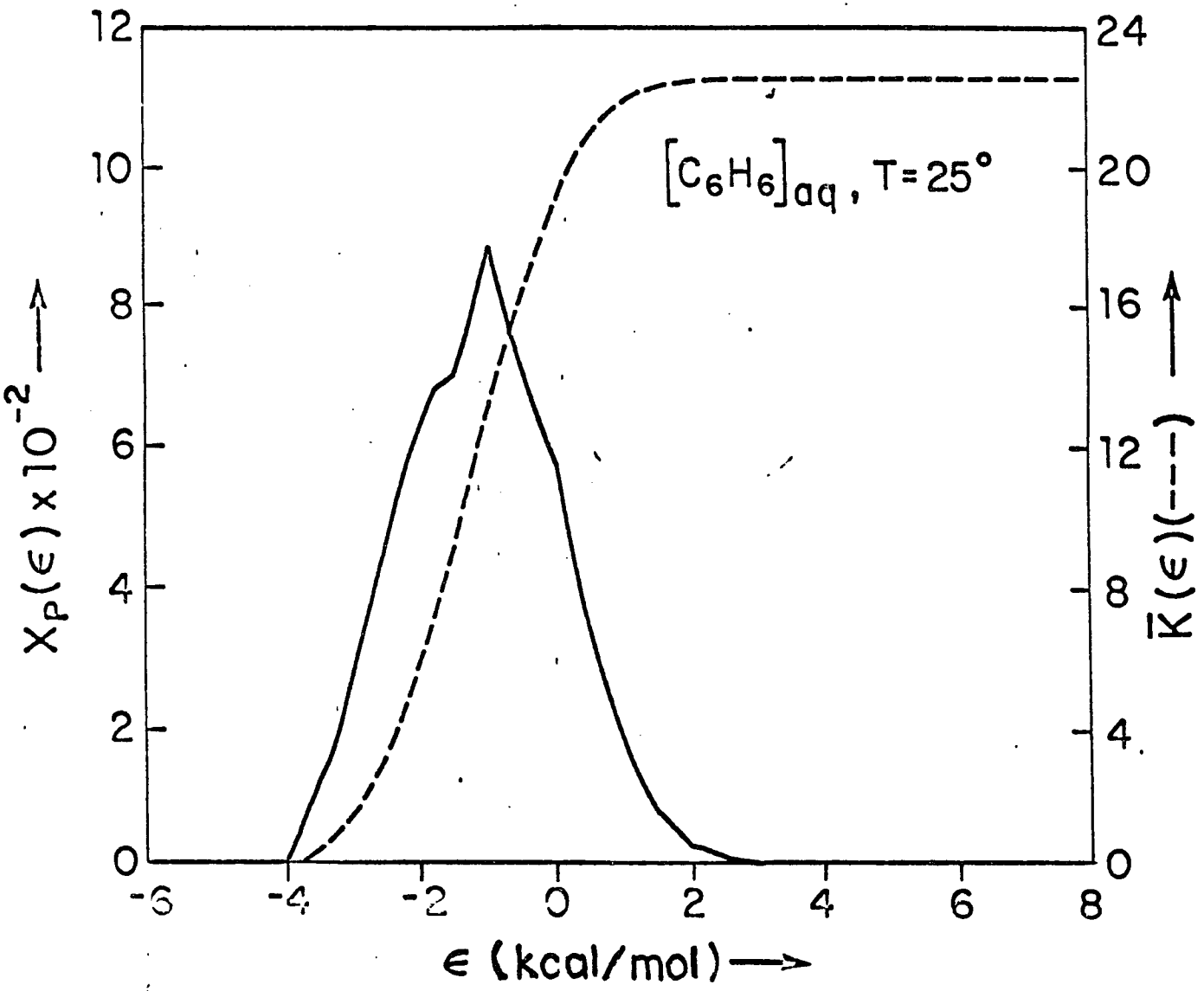

$[C_6H_6]_{aq}$, T=25°C
$\bar{\nu}$ = -60.6 kcal/mol
$X_B(\nu)\,d\nu \times 10^{-2}$
12
10
8
6
4
2
0
-72
-68
-64
-60
-56
-52
-48
-44
ν (kcal/mol)

Figure VI.14-Calculated QCDF for primary solute-solvent pair energy on a molecular basis in $[C_6H_6]_{aq}$.


12
10
8
6
4
2
0
24
20
16
12
8
4
0
-6
-4
-2
0
2
4
6
8
$[C_6H_6]_{aq}$, T=25°
$X_p(\epsilon) \times 10^{-2}$ ⟶
$\bar{K}(\epsilon)$ (---) ⟶
ϵ (kcal/mol) ⟶

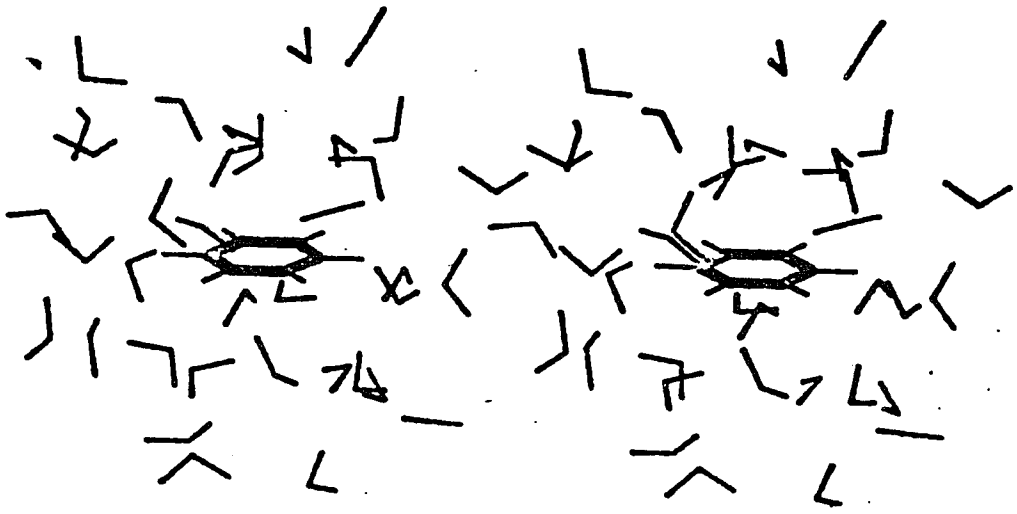
ures VI.11-VI.14. The first hydration shell of benzene is seen to involve from 19 up to 26 water molecules, with $\bar{K}$=22.1. The average total binding energy for water molecules is -60.6 kcal/mol. The pair interaction energy extends from -3 to 0 kcal/mol, with the contribution at $\bar{\epsilon}$=0 coming from all the distant waters. The average pair energy is -1.08 kcal/mol.

Discussion.

The essential structural features of the aqueous hydration of benzene emerging from the simulation results is a first hydration shell consisting of 23 water molecules,shown in Figure VI.15. 21 of the first shell waters can be associated primarily with H-region hydration and 2 are associated with hydration of the benzene ring above and below the carbon skeleton. Further insight into the nature of the calculated hydration can be obtained by examining details of the local hydration of the benzene ring C-H groups and pi electron cloud in individual structures contributing to the simulation. A stereo view of the computer generated Dreiding model of the benzene hydration complex is shown in Figure VI.16; the "bonds" in the figure connect oxygen atoms of water molecules that are within hydrogen bonding distance. This figure reveals the cage-like features of the benzene hydration complex. Quite a few puckered pentagonal forms can be discerned, but also contributions from higher and lower order polygonal forms. The irregularity of the polygons is a natural consequence of thermal disorder in the system at ambient temperatures. For the in-plane interactions, the number of waters and the spatial extent of the hydration shell are consistent with previous examples of hydrophobic hydration found in simulations on dilute aqueous solutions of alkyl group containing molecules. The average pair interaction energy of water molecules primary to the CH groups at -1.Ø8 kcal/mol turns out

Figure VI.15-Stereographic view of a significant molecular structure contributing to the statistical state of $[C_6H_6]_{aq}$.

Figure VI.16-Stereographic view of Dreiding models of the first hydration shell of benzene taken from the Monte Carlo simulation described herein on $[C_6H_6]_{aq}$ at 25°C. Water oxygens within 3.2 Å are bonded.

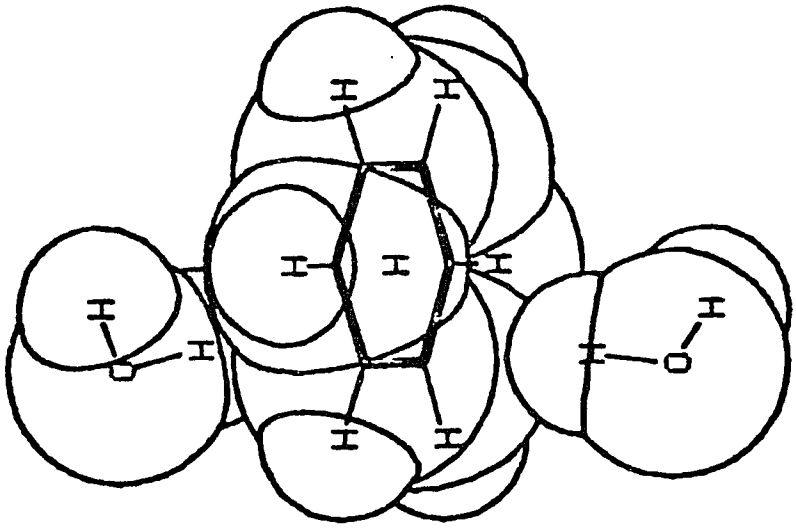


to be significantly closer to corresponding value computed for methane water interactions in $[CH_4]_{aq}$ suggesting that the apolar solute-water interactions are quite similar in both the cases, and essentially hydrophobic.

The hydration complex above and below the molecular plane, Figure VI.17, features two water molecules, one on each side of the benzene ring, located one above and one below the center of the pi-electron cloud. A hydrogen atom on each water molecule extends into the pi-cloud, towards the center of the molecule. and a mean pair energy of -1.64 kcal/mol is associated with this structure. The single water molecules interact with a set of second shell waters which extend over the carbon atoms, and articulate with the H-region waters completes the benzene hydration. It is interesting to note that the in-plane and out-of-plane potential minima in the pairwise interaction energy surface, both about 3 kcal/mol, give rise to quite different hydration structures in the simulation. The pi-cloud hydration structure is favoured by both a weak hydrogen bonding interaction and by steric factors, since the hydrogen atom of water can be accomodated better than the oxygen in the pi-cloud. The relative importance of these two effects are discussed below.

Partial molar internal energy of transfer for benzene calculated from simulation results comes out to be -70.15 kcal/mol with error bounds estimated to be ±30.0 kcal/

Figure VI.17-Space filling model of benzene and two water molecules primarily belonging to the pi-cloud.



mol. This calculation used a value of -8.65 kcal/mol for the energy of water as calculated in a previous study[64]. A recent, 3000K long run using the force-biased sampling scheme predicted a value of -8.75 kcal/mol for the energy of water[103]. Using this value, the partial molar internal energy of transfer of benzene is calculated to be -48.6 kcal/mol, with the estimated error bound still about ±30.0 kcal/mol. The calculated transfer energy also has a large error associated with it due to the fact that it is a small quantity derived from the difference of two large numbers known only with a considerable degree of statistical uncertainity. This number can also be expected to be quite sensitive to the well depth of the benzene-water interaction energy in the in-plane region. The energy difference of 1 kcal/mol between the function used herein and that of Karlstrom et. al. propagated over 20 in-plane interactions could change the calculated transfer energy by 30%. Experience with similar problems in liquid water system[104] indicates that the calculated structural characteristics of the system are not highly sensitive to small changes in energetics, and thus the description of the essential nature of the benzene hydration complex set forth herein is expected to remain valid.

Finally, we pursued the question of sensitivity of results to choice of potential function, with an additional simulation, identical to that previously described, except that the attractive part of the benzene-water potential was

everywhere set to zero. The complete simulation also involved 1700K configurations with ensemble averages formed over the last 1000K. The computed energetics are given in column II of Table 1. Here the transfer energy is reduced to -1.86 kcal/mol, still with large error bounds. Thus, the attractive part of the benzene-water potential influences the transfer energy significantly, with the experimental vaue bracketed by the two simulation results reported herein. The structural indices turned out to be essentially insensitive to this change in benzene-water potential, which indicates the steric contribution to the structure of the pi-cloud to be quite significant. Although a water hydrogen is proximal to the pi-cloud in this model, the nature of the interactions is not exclusively hydrophilic.

>

CHAPTER VII.

Monte Carlo Computer Simulation Study of the Hydrophobic Effect.

Potential of Mean Force for the Stacking of Two Benzenes.

Interactions between the nucleic acid bases are a significant feature of the three dimensional structure of the nucleic acids[105]. These interactions are stabilized by hydrogen bonding between the bases in the molecular plane and by the out-of-plane stacking. Thermodynamic studies[106] on the aqueous solutions of the nucleic acid bases show that the vertical stacking of bases has characteristics of hydrophobic bonding. Stacking interactions may also be stabilized by dispersion forces between the pi-clouds of the molecules. A prototypical system for the study of the stacking of molecules with delocalized pi-electron systems is $[(C_6H_6)_2]_{aq}$. In this chapter we describe the calculation of the potential-of-mean force $w_{ss}(R)$, between benzene molecules in water using Monte Carlo simulation augmented with umbrella sampling and convergence acceleration techniques. General aspects of benzene-benzene interactions are considered, but the calculations reported herein are confined to interactions along the vertical stacking coordinate.

Hydrophobic interactions between benzene molecules in water has been demonstrated by thermodynamic measurements from Tucker and Christian[57] and Dutta-Choudhury et. al.[99], with the negative deviations from Henry Law constants obtained in these studies attributed to benzene association. Tucker and Christian also noted that the G^{HI} for benzenes in water was comparable to $G^{stacking}$ found for purines and pyrimidines. The only theoretical study on

the hydrophobic interaction of benzene molecules comes from Rossky and Friedmann[66], who applied Gurney model to calculate the thermodynamics of benzene association. The calculations supported the results of Tucker and Christian, but could not propose the microscopic nature of the benzene association in water due to limitations inherent in their methods.

Calculations.

The calculation of $w_{ss}(R)$ for two benzenes in water proceeded in a manner similar to that used for two methanes in water, i.e. umbrella sampling techniques. Here $g_{ss}(R)$ were obtained for three small overlapping windows on the intersolute coordinate R and were subsequently matched to produce $g_{ss}(R)$ over the range of interest. The weighting function used to constrain the movement of the benzenes is a harmonic potential, the effect of which was removed by the standard procedures in umbrella sampling theory.

The calculations decribed herein are intended as a direct extension of studies on hydrophobic hydration of benzene and hydrophobic interaction between methanes. The potential function used to describe water-water pair interactions was that of Matsuoka et.al. based on quantum mechanical calculations. The nature of this potential function has been discussed in previous chapters. Both benzene-water and benzene-benzene interactions were modeled using empirical potential functions developed by Karlstrom et. al.[100] from quantum mechanical calculations. The functional form of this empirical potential function is

$$V(r) = \sum [A_i r^{-1} + B_i r^{-4} + C_i r^{-6} + D_i r^{-9} + E_i r^{-12}]$$

where the summation is over the interatomic pairs C-O, C-H(water), H(benzene)-O and H(benzene)-H(water) for benzene-water interaction and similarly over C-C, C-H and H-H over benzene-benzene interaction. The coefficients

A_i,B_i,C_i,D_i and E_i for both benzene-water interactions and benzene-benzene interactions are given in Tables VII.1 and VII.2. The potential energy hypersurface for benzene-water interactions show two noticeable minima. The first is worth -2.9 kcal/mol and corresponds to the interaction of water molecule with the pi-cloud of benzene, and the second worth -1.9 kcal/mol arises from the in-plane benzene-water interactions. Potential surfaces for these interactions were given in the preceding chapter. The benzene-benzene interactions are more complex due to the variety of possible dimer configurations. The global minimum in the benzene-benzene potential energy surface is worth -2.6 kcal/mol and corresponds to the benzenes perpendicular to each other. In this configuration one of the C-H bonds of the benzene points directly into the pi-cloud of the other benzene. The interaction between the stacked benzenes, which is the one of most interest in this study, is rather weak and has a minimum of -.3 kcal/mol. There is no direct experimental test of quality for benzene-water and benzene-benzene potentials.

In order to make a direct comparison between the results of this calculations and those of hydrophobic hydration of benzene, it is essential to use the same potential functions in both the study. In the study of hydrophobic hydration of benzene we used the potential functions of Clementi et. al.[101] to describe the benzene-water interactions whereas, here we use Karlstrom's potential for the same. This change

Table VII.1The benzene-water parameters for Karlstrom potential. The first entry in the first column refers to the atom in benzene and the second refers to the atom of water.

0 ROWNAME	1 A	2 B	3 C	4 D	5 E
1. C-O	33.658	163.883	-2586.96	44883.5	109937.
2. C-H	-16.829	-94.375	858.81	-6506.8	27667.
3. H-O	-33.658	-176.163	1144.37	-5840.7	13935.
4. H-H	16.829	100.514	-642.56	2238.8	3119.7

Table VII.2 The benzene-benzene parameters for the Karlstrom potential.

0 ROWNAME	1 A	2 B	3 C	4 D	5 E
1. C-C	7.3365	-29.468	-325.44	16005.8	194320.
2. C-H	-7.3365	27.889	-286.6	1641.5	27667.
3. H-H	7.3365	-26.31	114.73	-590.	3119.7

was done because Karlstrom potential is available for both benzene-water and benzene-benzene interactions in the same form. The potential of Clementi et. al. was not extended to benzene-benzene interactions. The hydrophobic hydration of benzene emerging from the work of Karlstrom et. al. using the Karlstrom potential is not much different from that emerging from our study using the Clementi et. al. potential, and a direct comparison is possible.

Monte Carlo simulations in this study were carried out using the Metropolis method augmented with Force-bias method and Preferential sampling method to accelerate the convergence of the results. The convergence acceleration methods are discussed in the methodology section. The system for this study consisted of two benzenes in 510 waters, configured under face centered boundary conditions with a spherical cutoff treatment of the potential. Since our interest was to study the $w_{ss}(R)$ for the stacking of benzenes, the benzenes were restricted to move only along their C_6 axis. They were allowed to rotate, resulting in the sampling of eclipsed and staggered configurations of benzene dimer. However in all configurations, the two benzene rings are parallel to each other.

Three points on the intersolute coordinate were chosen, 4.5,5.0 and 5.5 $\mathring{A}$, were chosen as the origin of umbrella sampling windows. The first two simulations were run using R_o values of 4.5 and 5.5 $\mathring{A}$ and failed to produce over-

lapping segments of $g_{ss}(R)$. Then an intermediate window centered at 5.0°A was introduced into the calculation to ensure proper overlapping. The region of intersolute coordinate spanned in each window is considerably less than those in previous studies on spherical apolar molecules such as methane and argon. Harmonic restoring potentials with a force constant of 1.5 kcal°A^2 were used to define the three windows. The density of the system was calculated from the experimental molar volumes of water and benzene. All calculations in this study correspond to 25°C. For each window $2.5X10^6$ configurations were sampled and the initial $1.X10^6$ configurations were discarded for equilibration. All properties reported in this study were calculated from the last $1.5X10^6$ configurations.

Results and Discussion.

The $g_{ss}(R)$ calculated for the three windows are given in Figure VII.1. For window 1 the point of origin chosen at 4.5 Å corresponds to the solute-solute contact distance of 2.2 Å along the C_6 axis of benzenes. This distance was estimated from the peak in $g_{BW}(R)$ corresponding to the hydration of pi-cloud of benzene. In this realization, the intersolute region from 3.6 to 4.8 Å is sampled. Maximum in $g_{ss}(R)$ for the first window falls at 4.3 Å and the distribution is rather smooth.

Window 2 has the point of origin at 5.0 A and was found to sample the intersolute coordinate from 4.1 to 5.8 Å. The maximum in $g_{ss}(R)$ for the second window falls at 5 Å. The third window, with the point of origin at 5.5 Å spans the intersolute region from 4.6 to 6.4 Å with the maximum at 5.2 Å.

The $g_{ss}(R)$ for the three windows were matched on the basis that the points in the overlapping region should coincide and a single $g_{ss}(R)$ was generated. The $w_{ss}(R)$ corresponding to this $g_{ss}(R)$ is given in Figure VII.2. The oscillatory behaviour in $w_{ss}(R)$, noted in previous studies is clearly evident. Two distinct minima at 4.3 and 5.1 Å are found. The $w_{ss}(R)$ calculated from the simulation is arbitrary. The additive constant can be determined by assuming that the second virial coefficient[57] of benzene in water arises from the stacking interactions of benzene.

Figure VII.1-Calculated radial distribution function $g_{ss}(R)$ plotted against intersolute separation R for each of the three windows of umbrella sampling for $[(C_6H_6)_2]_{aq}$ at 25°C. ●, Window 1; o, window 2; ■, window 3.

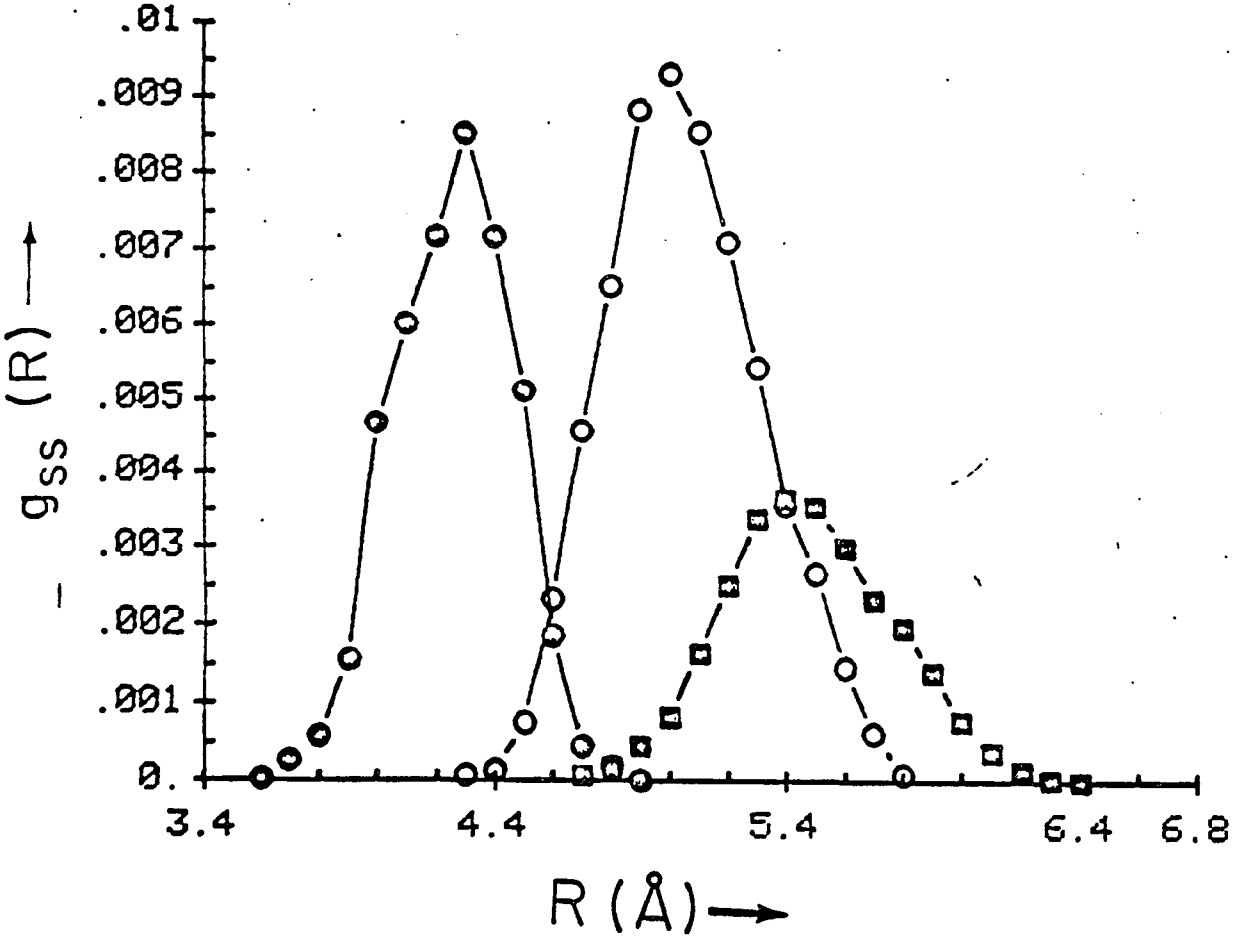

.01
.009
.008
.007
.006
.005
.004
.003
.002
.001
0.
g_{ss} (R)
3.4
4.4
5.4
6.4
6.8
R (Å)

Figure VII.2-Calculated potential of mean force $w_{ss}(R)$ as a function of R, after matching, for $[(C_6H_6)_2]_{aq}$ at 25°C.

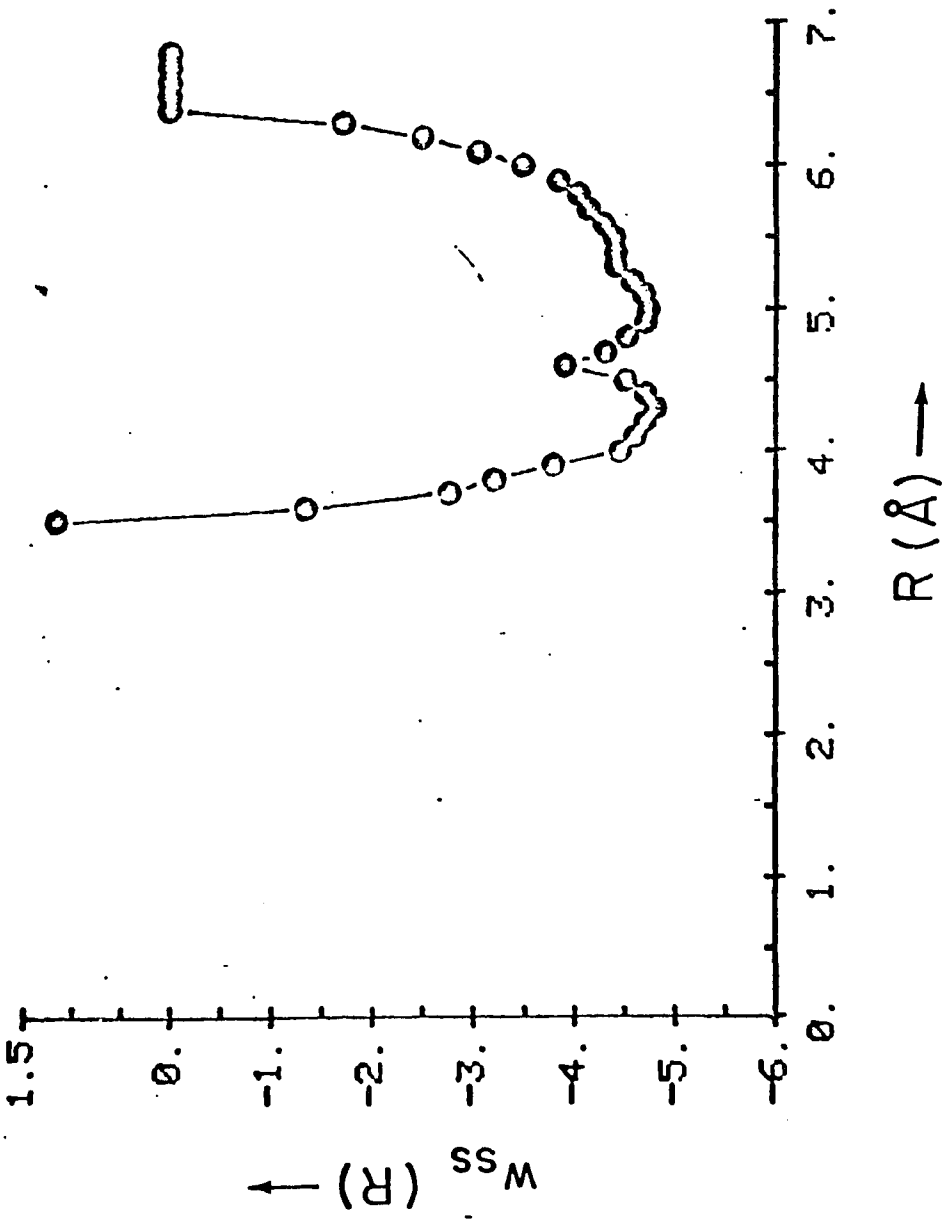
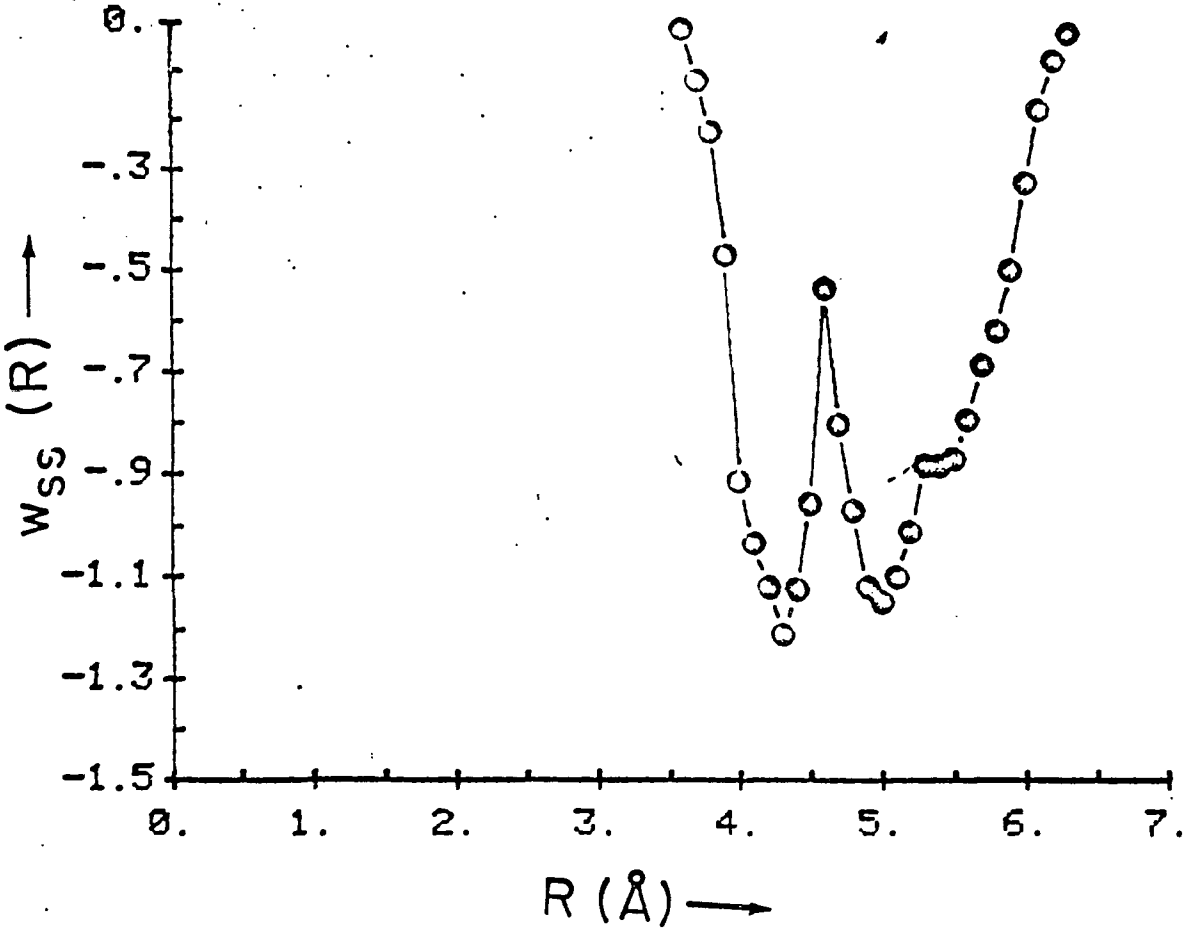

1.5
0.
-1.
-2.
-3.
-4.
-5.
-6.
W_{ss} (R) ⟵
0.
1.
2.
3.
4.
5.
6.
7.
R (Å) ⟶

Figure VII.3-Calculated potential of mean force $w_{ss}(R)$ as a function of R, after shifting the $w_{ss}(R)$ in Figure VII.2 to correspond to the experimental virial coefficient, for $[(C_6H_6)_2]_{aq}$ at 25°C.


0.
-.3
-.5
-.7
-.9
-1.1
-1.3
-1.5
W_{SS} (R)
0.
1.
2.
3.
4.
5.
6.
7.
R (Å)

This approximation produces the $w_{ss}(R)$ in Figure VII.3, where the values of $w_{ss}(R)$ look reasonable. The approximation used in calculating the absolute $w_{ss}(R)$ is not complete for several reasons. Experimental virial coefficients correspond to $w_{ss}(R)$ calculated by allowing benzenes to move freely, whereas in our study the $w_{ss}(R)$ corresponds to only the stacking. Therefore the absolute $w_{ss}(R)$ reported in Figure VII.3 should not be taken as quantitatively correct. We are however suggesting a method to relate the experimental virial coefficient to the arbitrary $w_{ss}(R)$ calculated from the umbrella sampling methods, and furthur studies along this line seems worthwhile.

The first minimum in $w_{ss}(R)$ corresponds to the contact benzene pair and the second minimum to a solvent-separated pair. The solvent-separated pair is not one where a full water molecule separates the benzenes; the nature of the arrangement of water in this configuration is discussed in the following section. The distance between the two minima is rather small compared to previous studies. The two minima have essentially the same depth and the calculated statistical weights suggest that the solvent-separated form is favored strongly over the contact form. It has to be noted that Pratt and Chandler predicted that the $w_{ss}(R)$ to have lower second minimum if the solute-water interactions are stronger than the solute-solute interactions. In our study the benzene-benzene interactions are weaker than the benzene-water interactions and thus the form of calculated

$w_{ss}(R)$ may be explained by Pratt and Chandler's proposition. The system of two benzenes in water differs considerably from the system of two methanes studied previously, as discussed below.

Analysis of results.

The calculated statistical weights for the contact and solvent-separated benzene pairs show a strong preference for the latter. In this section we present an analysis of the molecular nature of these structures in terms of $g_{\bullet O}(R)$ and $g_{\bullet H}(R)$, where $\bullet$ refers to the center of mass of the benzene dimer. The $g_{\bullet O}(R)$ for window 1, Figure VII.4, shows well defined shell structure around the benzene dimer. The first shell with respect to the center of mass of benzene dimer is defined by the peak in $g_{\bullet O}(R)$ at 4 Å and by the minimum at 5 Å. This shell contains a layer of six water molecules within 5 Å of the center of mass of benzene dimer and these waters were found to be situated in between the two benzenes. The peak in $g_{\bullet O}(R)$ at 5.6 Å can be assigned by referring to the hydrophobic hydration of benzene in the preceding chapter. The $g_{BW}(R)$ calculated for $[C_6H_6]_{aq}$ has a peak at 3.5 Å referring to the water molecules assigned to the pi-cloud of benzene and another peak at 4.5 Å referring to the waters proximal to the C-H bonds in benzene. In window 1, the benzenes are separated by an average distance of 4.3 Å and the waters proximal to the pi-clouds of benzenes from the direction away from the contact region and the waters proximal to the C-H bonds fall at a distance of about 5.6 Å from the center of mass of benzene dimer. The peak in $g_{\bullet O}(R)$ at 8.5 Å arises from the long range shell structures of C-H bond hydration. The distribution $g_{\bullet H}(R)$, Figure VII.5, has

Figure VII.4-Calculated $g_{\bullet O}(R)$ for $[(C_6H_6)_2]_{aq}$ for window 1.

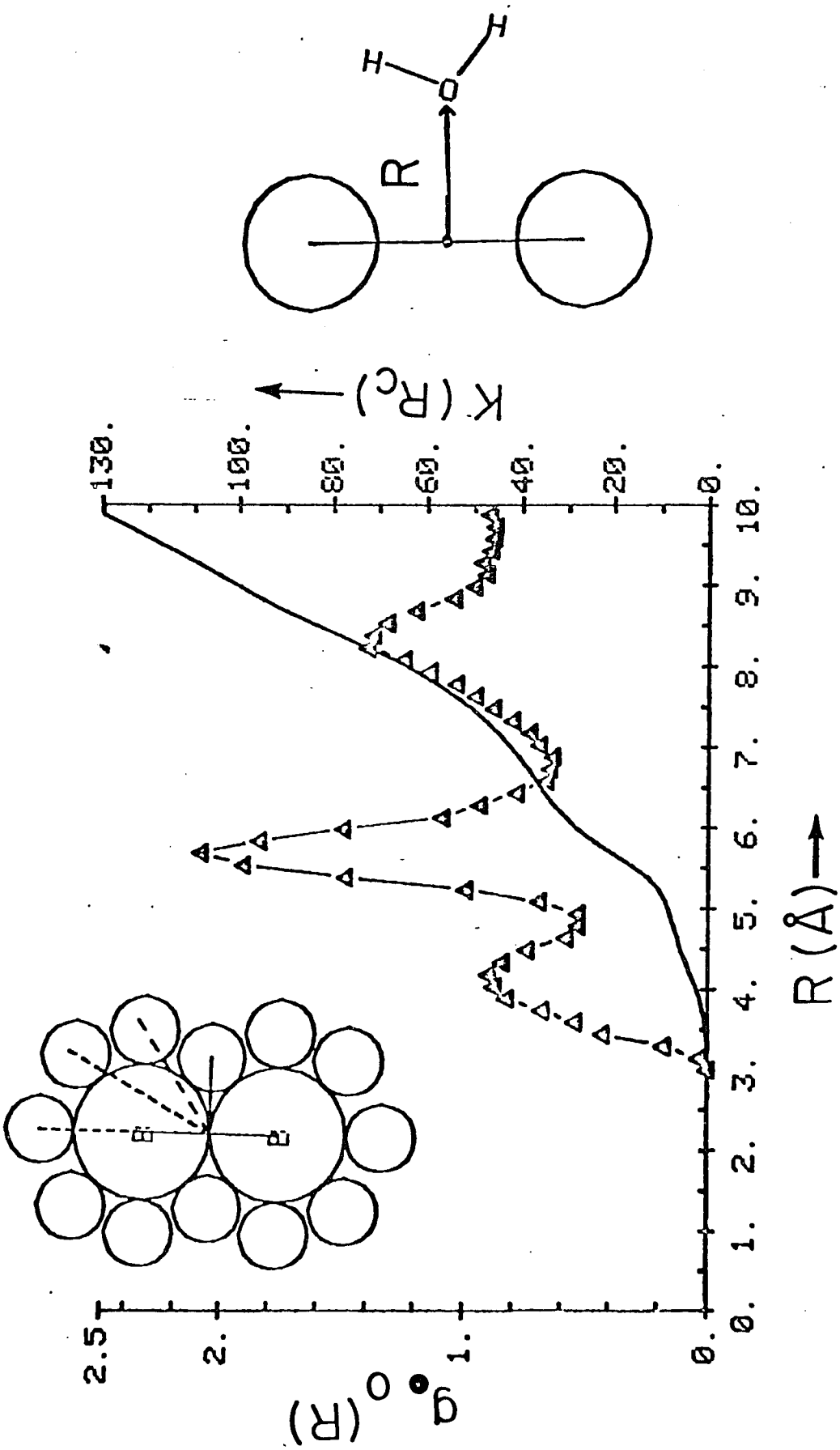
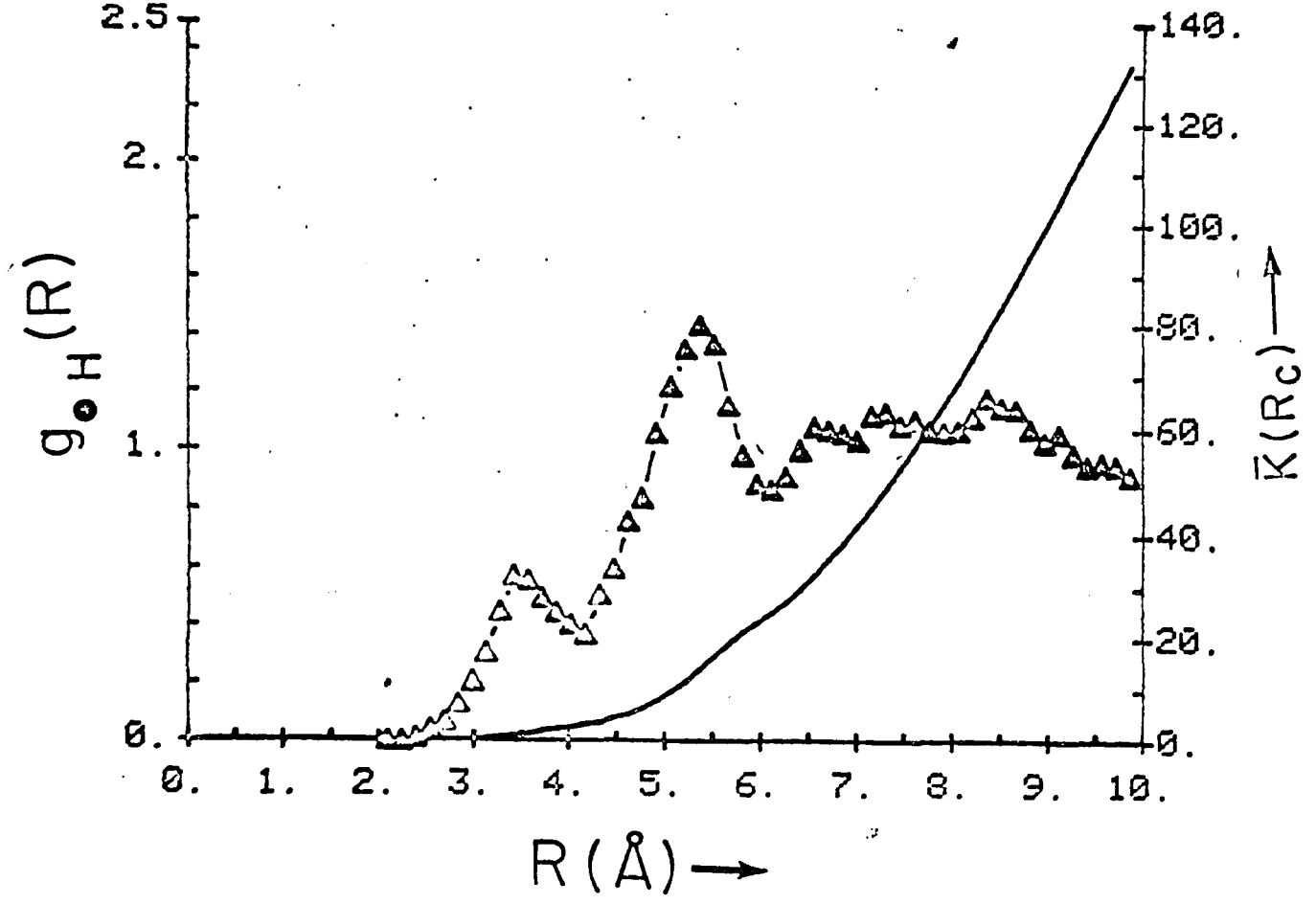
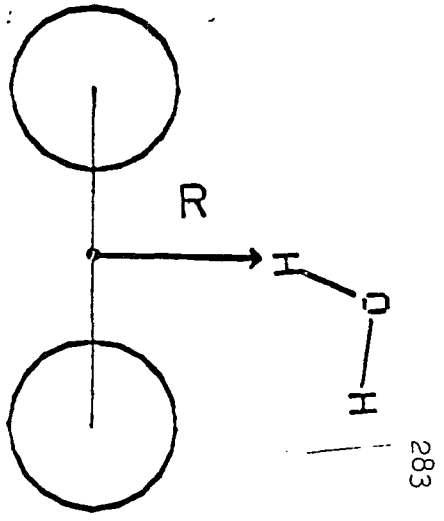

H
H
O
R
K (Rc)
130.
100.
80.
60.
40.
20.
0.
10.
9.
8.
7.
6.
5.
4.
3.
2.
1.
0.
R (Å)
2.5
2.
1.
0.
gO O(R)

Figure VII.5-Calculated $g_{\bullet H}(R)$ for $[(C_6H_6)_2]_{aq}$ for window 1.

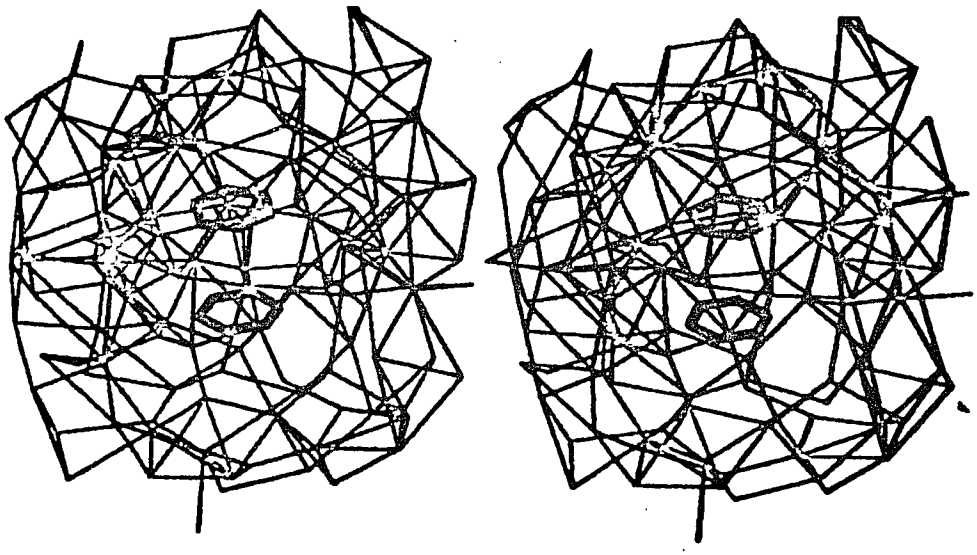

2.5
2.
1.
0.
$g_{O_H}(R)$
140.
120.
100.
80.
60.
40.
20.
0.
$K(R_c)$
0. 1. 2. 3. 4. 5. 6. 7. 8. 9. 10.
R(Å)


R
H
O
H

essentially all the features of $g_{\bullet O}(R)$ and is shifted by about 1 Å in the direction of lower R values. The first peak at 3.4 Å refers to the first shell of six water molecules discussed above. The observation that the first peak in $g_{\bullet H}(R)$ for window 1 occurs at a distance .6 Å shorter than the corresponding peak in $g_{\bullet O}(R)$ suggests that on the average the hydrogens tend to be closer to the region of contact between the benzenes than the oxygens. The region of contact in window 1 refers to the contact region of the pi-clouds and we note that there is already a tendency for the hydrogens to approach the region in between the pi-clouds. This tendency of hydrogen to point towards the pi-cloud was noticed in $[C_6H_6]_{aq}$ and was found to arise from the favorable interaction between the hydrogen and pi-cloud and also because a hydrogen fits inside the hole in the pi-cloud better.

A typical structure representing the contact benzene pair hydration is given in Figure VII.6. The layer of water molecules considered as the first shell for the benzene dimer is specially marked. A careful analysis of these waters shows that there are at least two water molecules with one of their hydrogens pointing at the pi-cloud. The rest of the waters just fill the space and engage in mutual hydrogen bonding. The separation between the benzenes is not large enough to let hydrogens penetrate in between the benzenes.

Figure VII.6-Typical structure corresponding to most probable configuration in window 1.



The hydration of the pi-cloud of benzenes from the direction away from the contact region of benzenes is also seen in these structures. The hydrophobic hydration of the C-H bonds of benzenes can also be seen in Figure VII.6.

The distribution functions $g_{\circledcirc O}(R)$ and $g_{\circledcirc H}(R)$ for window 2 are presented in Figures VII.7 and VII.8. In this window the average separation between the benzene rings is 5.1 A. The first peak in $g_{\circledcirc O}(R)$ is shifted by 1 Å closer to the center of mass of benzenes compared to the $g_{\bullet O}(R)$ for window 1. This is certainly the result of more available space between the benzene molecules. The first shell still contains only six water molecules and a comparison of $g_{\circledcirc O}(R)$ and $g_{\circledcirc H}(R)$ for window 2 shows again the hydrogens to be closer to the contact region of benzenes than oxygens. The second peak in $g_{\circledcirc O}(R)$ has shifted to slightly larger distance and is not as smooth as the $g_{\bullet O}(R)$ for window 1. This can also be explained by reference to the results from the study on $[C_6H_6]_{aq}$. With the benzene-benzene average separations of 5.1 Å, the water molecules proximal to the pi-cloud of benzene are at a different distance from the waters proximal to the C-H bonds of benzene. This distorts the second peak in $g_{\circledcirc O}(R)$ and also makes it wide. The layer of six water molecules defining the first shell of the benzene dimer has moved closer to the contact region by 1 Å compared to window 1. A typical structure representing the benzene dimers in this window is provided in Figure VII.9. An examination of sev-

Figure VII.7-Calculated $g_{\bullet O}(R)$ for $[(C_6H_6)_2]_{aq}$ for window 2.

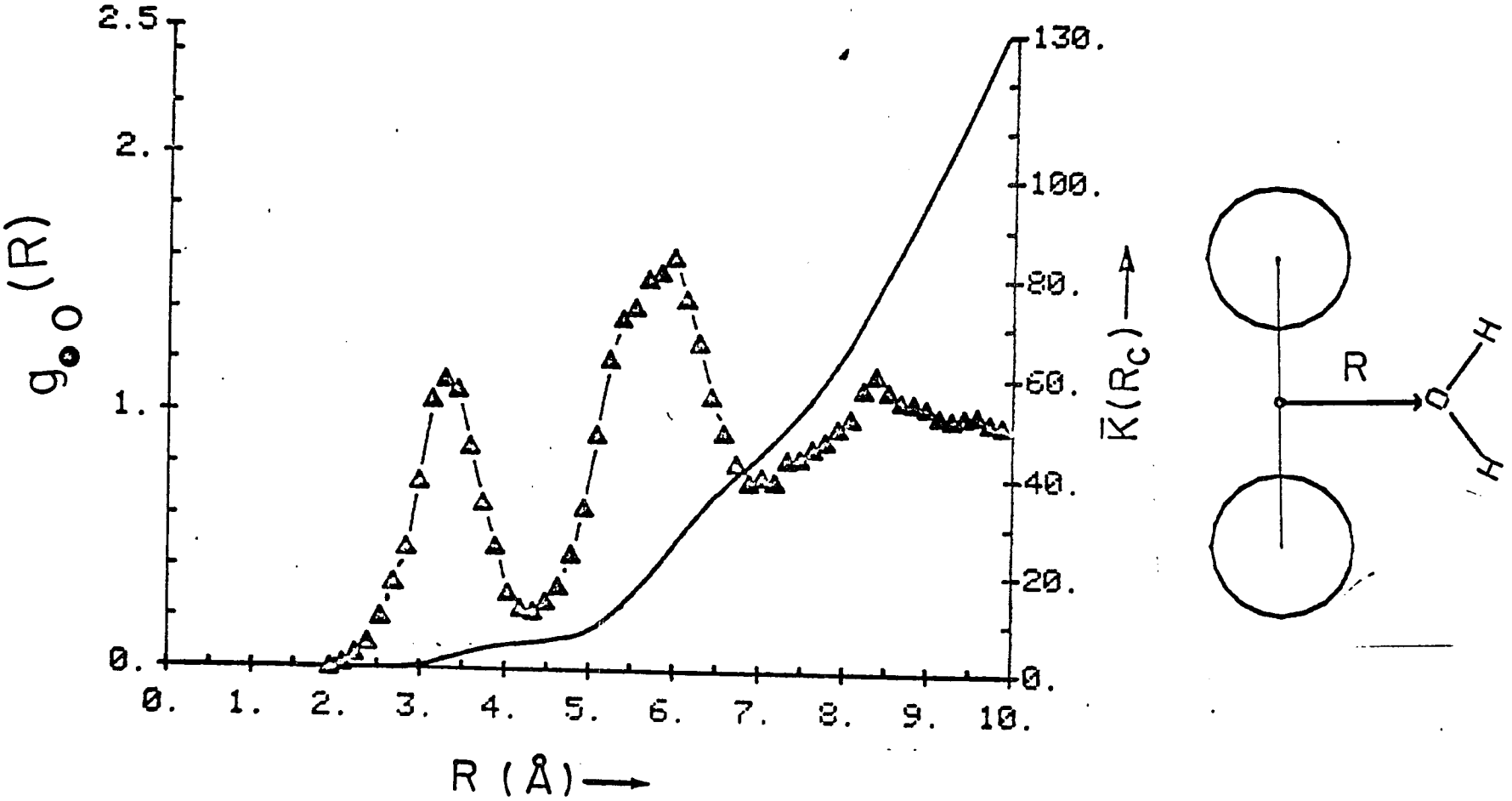

2.5
2.
1.
0.
$g_{OO}(R)$
130.
100.
80.
60.
40.
20.
0.
$\bar{K}(R_c)$
0. 1. 2. 3. 4. 5. 6. 7. 8. 9. 10.
R (Å)
R
O
H
H

Figure VII.8-Calculated $g_{\bullet H}(R)$ for $[(C_6H_6)_2]_{aq}$ for window 2.

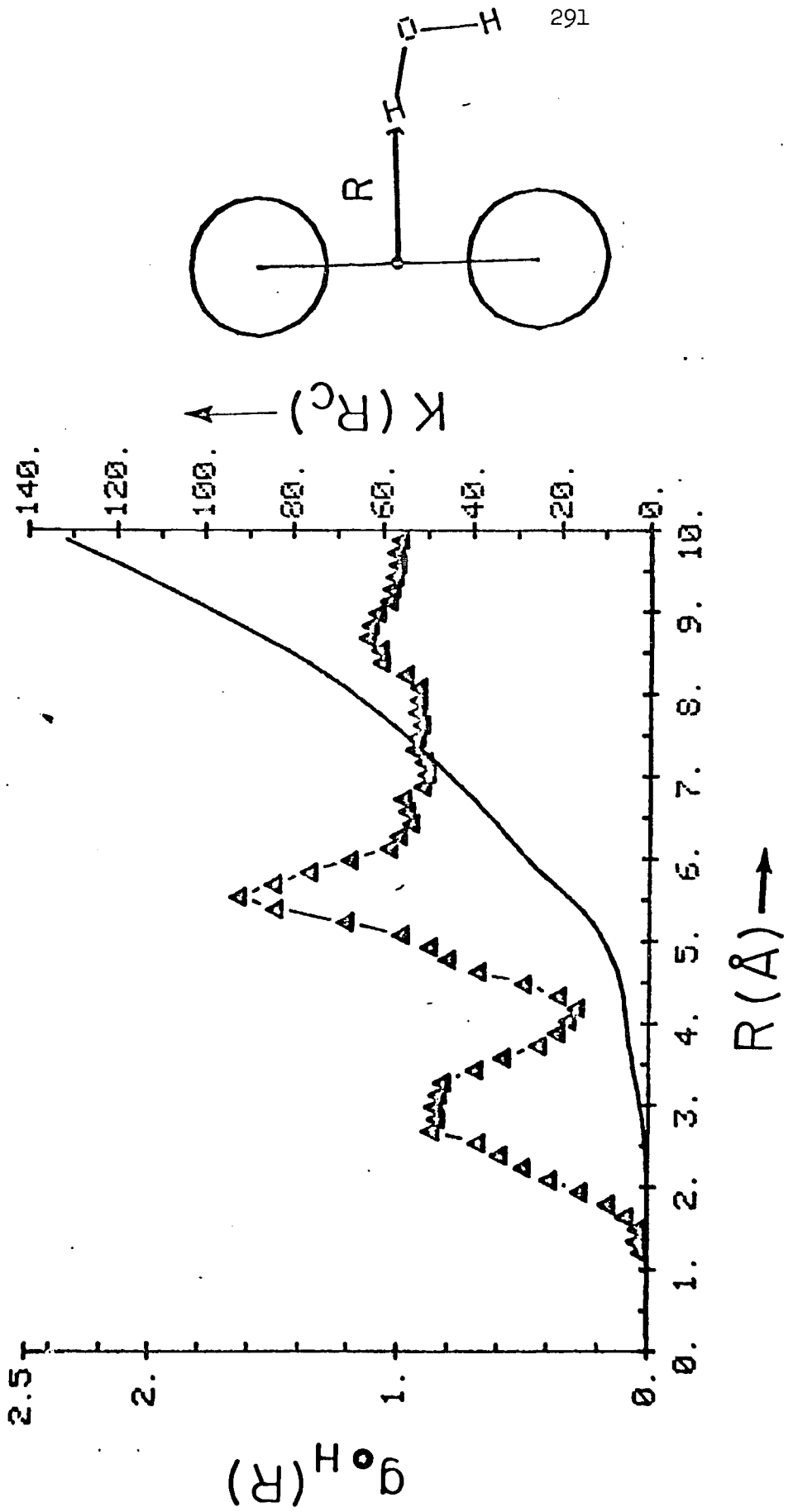
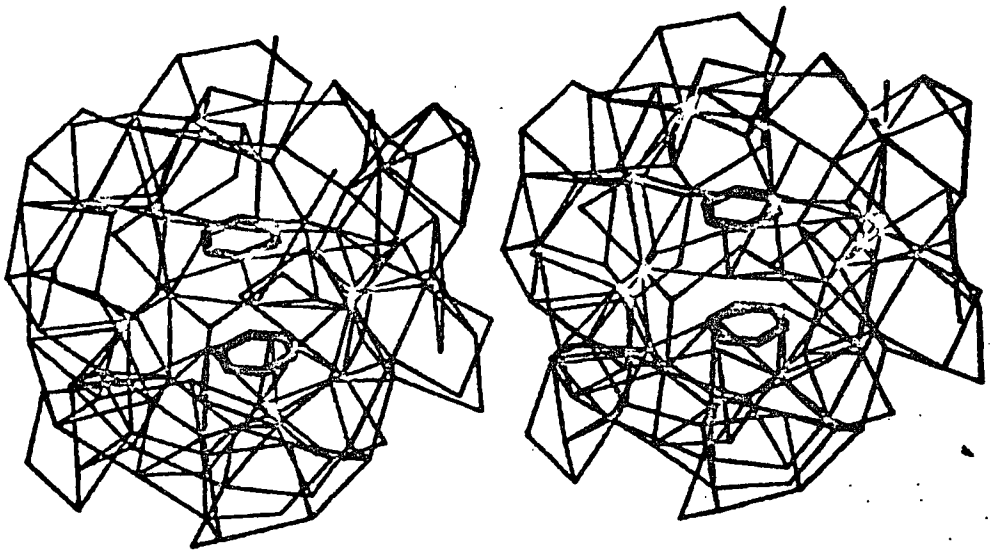

O—H
H
R
K (Rc)
140.
120.
100.
80.
60.
40.
20.
0.
0. 1. 2. 3. 4. 5. 6. 7. 8. 9. 10.
R (Å)
2.5
2.
1.
0.
gOH(R)

Figure VII.9-Typical structure corresponding to most probable configuration in window 2.



eral structures show that the hydrogens of water molecules in the first shell are better accomodated in the region between the benzenes than before. There are very few structures where both the hydrogens are pointing directly at the pi-clouds of benzenes. However, several structures showed that two water molecules in the first shell situated themselves such that one hydrogen from one of them pointed directly at the pi-cloud of one of the benzenes and the hydrogen from the other water pointed directly at the pi-cloud of the other benzene. Such hydrogens were found to be closer to the pi-clouds of benzenes than found in window 1. The rest of the waters in the first shell were found to fill the space in between the benzenes and formed hydrgen bonds with the rest of the waters as in window 1. This analysis, combined with the statistical weight of the configurations in window 2 suggests that the separation of about 5.1 Å between the benzenes gives enough room for a layer of six water molecule to intervene between the benzenes and stabilize such structures over contact benzene pairs.

The distribution functions $g_{\bullet O}(R)$ and $g_{\bullet H}(R)$, Figures VII.1Ø and VII.11, for window 3 is very similar to those for window 2, though the benzene-benzene separation in this window is about .4 Å larger. This shows that despite a larger benzene-benzene separation, the six-water first shell does not get any closer to the region in between the benzenes than when the separation was only 5.1 Å. A typical structure representing this window is given in Fig-

Figure VII.1Ø-Calculated $g_{\bullet O}(R)$ for $[(C_6H_6)_2]_{aq}$ for window 3.

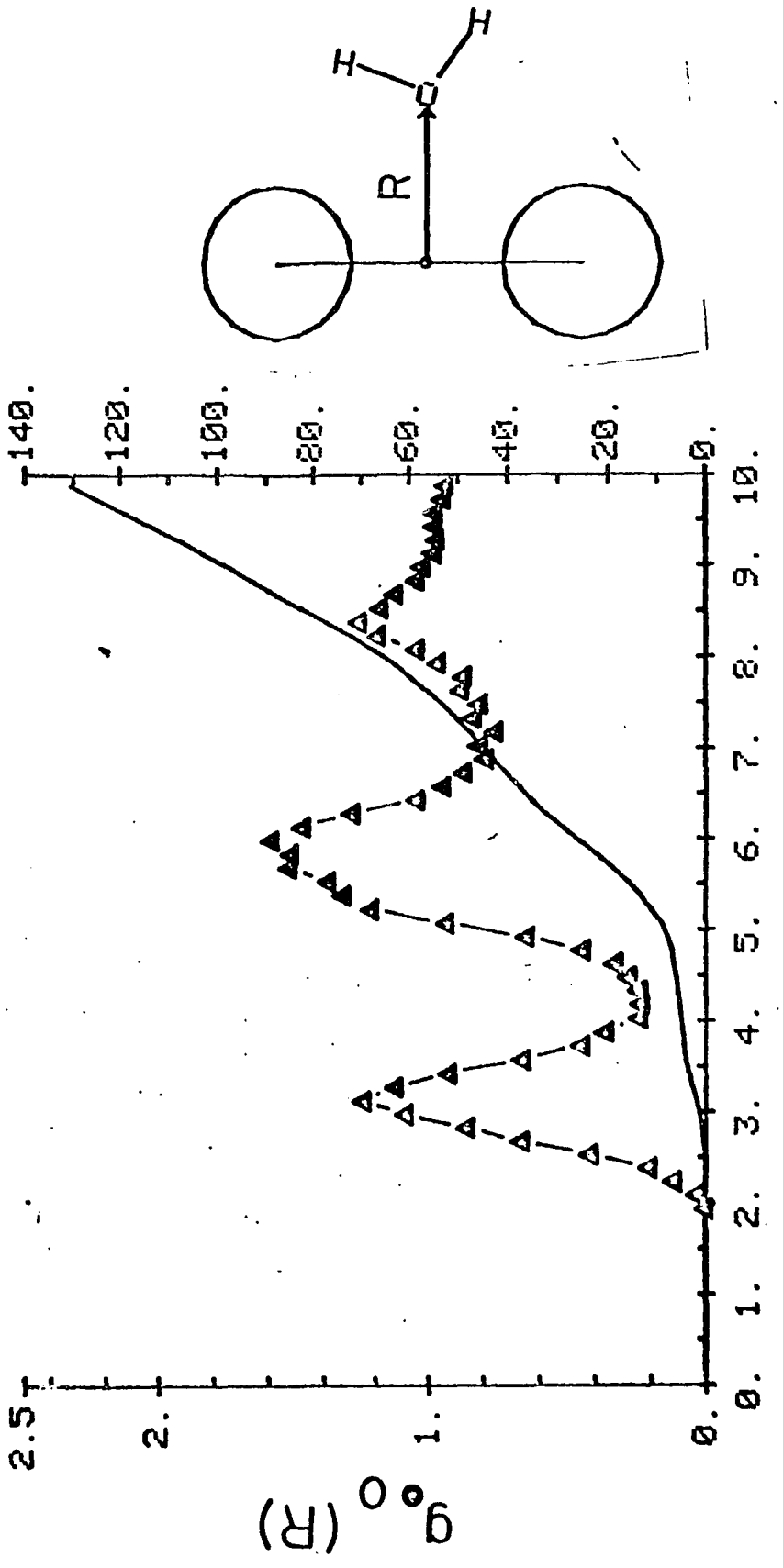

H
H
O
R
140.
120.
100.
80.
60.
40.
20.
0.
0.
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
2.5
2.
1.
0.
$g_{oo}(R)$

Figure VII.11-Calculated $g_{\bullet H}(R)$ for $[(C_6H_6)_2]_{aq}$ for window 3.

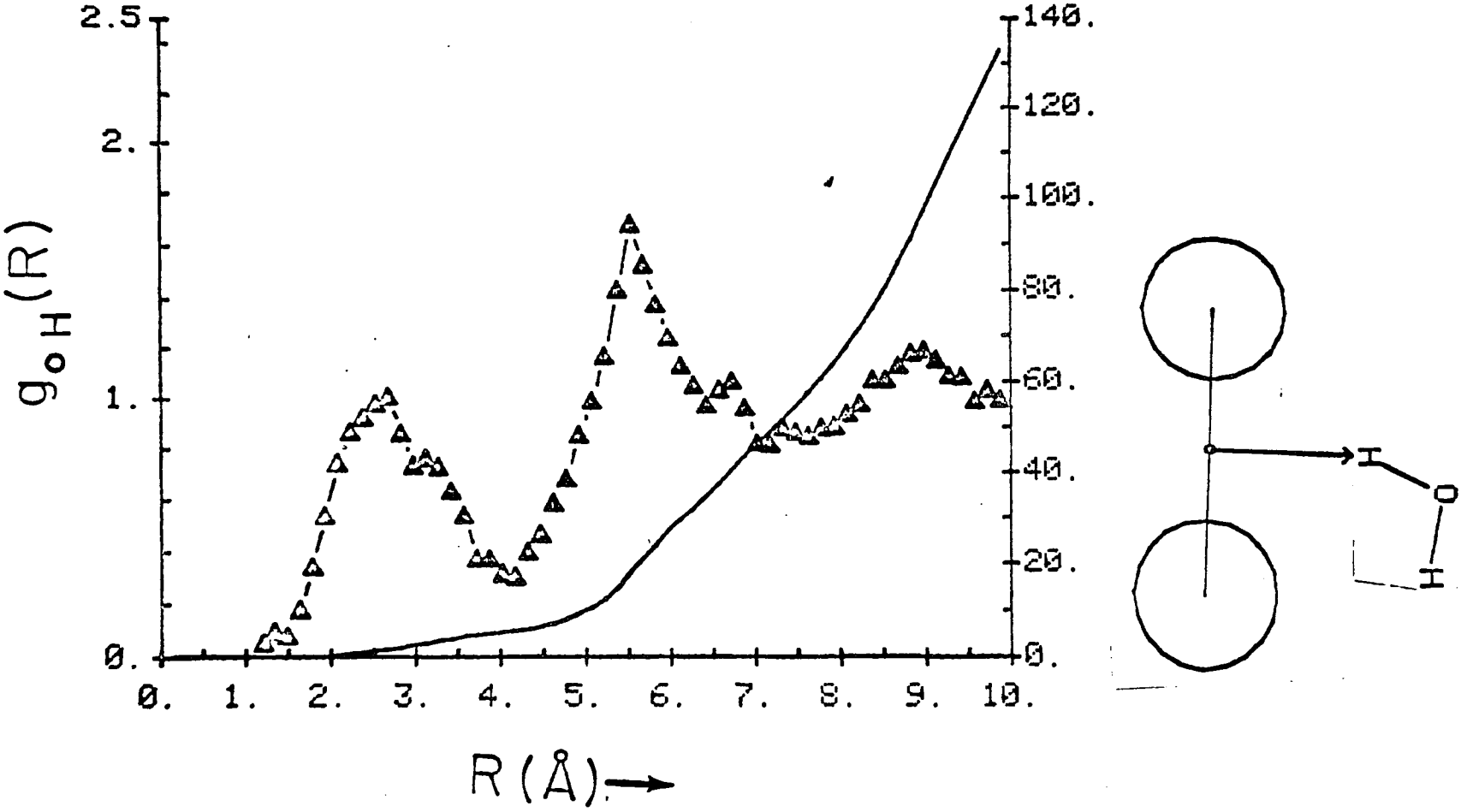
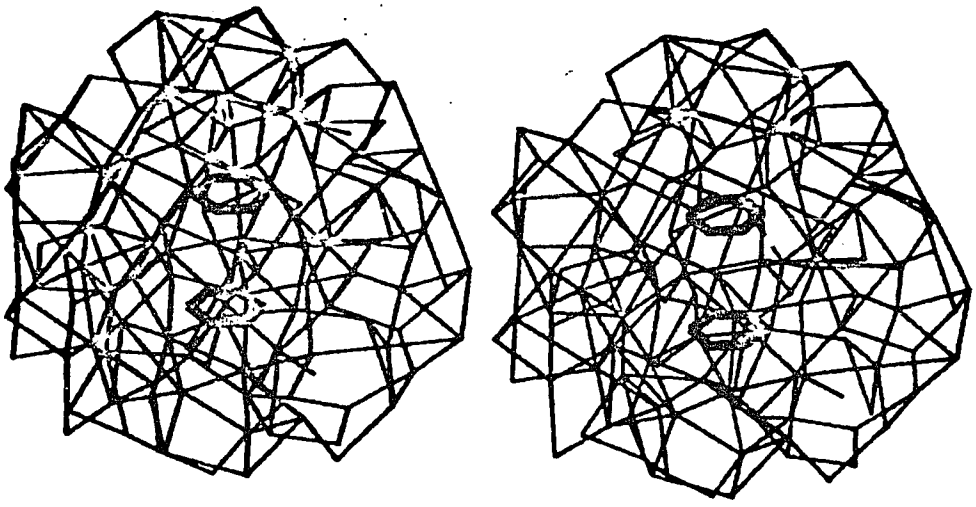

2.5
2.
1.
0.
$g_{OH}(R)$
140.
120.
100.
80.
60.
40.
20.
0.
0. 1. 2. 3. 4. 5. 6. 7. 8. 9. 10.
R (Å)
H
O
H

Figure VII.12-Typical structure corresponding to most probable configuration in window 3.



ure VII.12 and resembles Figure VII.8 very much.

In Figure VII.13, we present a structure where only the first shell waters of the benzene dimer are shown. This space filling model represents a statistically significant configuration in window 3 and includes the waters within 4 Å of the center of mass of the benzene dimer and in between the benzenes. Figure VII.14 shows the space filling model of pi-cloud hydration reproduced in this study.

Further analysis was done on some interesting orientational correlations. These results indicated that the benzene dimer prefers the eclipsed configuration over staggered configuration in both the contact and solvent-separated forms. The distribution of the angle between the center of mass of benzene dimer, the closest hydrogen of waters in the first shell and oxygen showed slight prefernce for the angles 60° and 120°. The configurations of water molecules corresponding to these angles and correspond to the hydration of the pi-cloud of benzene in the region between the benzenes.

The above analysis predicts that even when benzenes are at contact there is a layer of six water molecules in between the planes of the benzene rings, showing a considerable tendency to get in between the benzene rings. On increasing the benzene-benzene distance to 5.1 Å these waters tend to push themselves into the region between the benzenes 1 Å closer and further increase of benzene-benzene distances

Figure VII.13-Space filling model of benzene and water molecules within 4 Å of the center of mass of benzene dimer and in between the planes of the two benzene molecules.

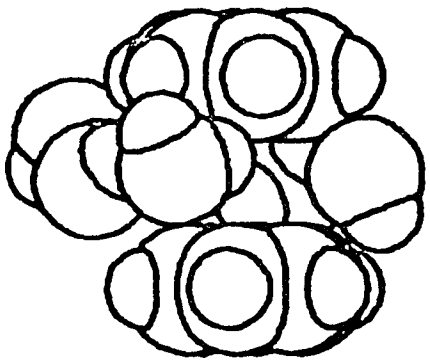
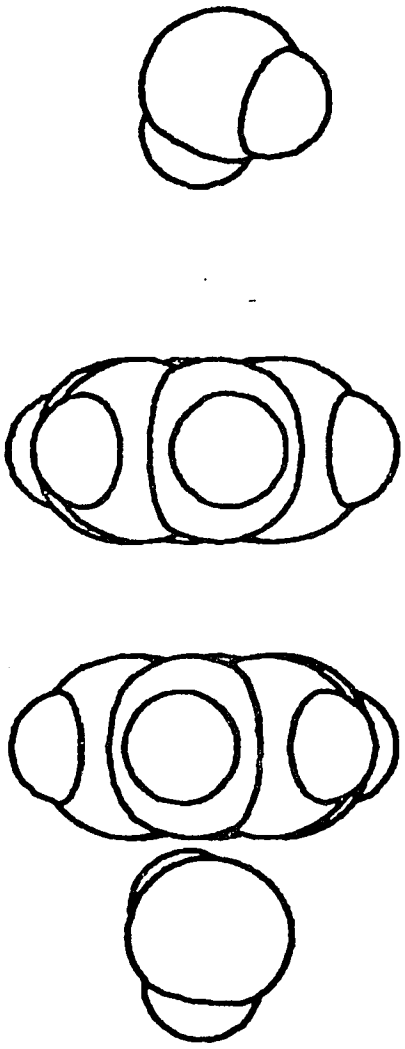


Figure VII.14-Space filling model of the benzene pi-cloud hydration.



upto 5.6 Å does not change the hydration picture significantly. We therefore propose that the HI between the benzenes involves a considerable contribution from solvent-separated structures where the benzenes are separated by a distance of 5.1 Å. The solvent-separated structures here do not imply a full water molecule in between the benzenes. The nature of this solvent separated form is described adequately in the above paragraphs. The shoulder in $w_{ss}(R)$ at 5.6 Å suggests that there are certinly other solvent-separated minima in the $w_{ss}(R)$ possible. What we have demonstrated in this study is that there is even a short-range solvent-separated HI between benzenes that has a higher statistical weight than the contact benzene pairs. The picture emerging from our study may be extended to the stacked nucleic acids in biomolecules. The distance between the molecules containing pi-clouds,in the direction of stacking need not have large space to accomodate a water in between to realize a solvent separated pair. As demonstrated by this study there are interesting intermediate short-range solvent-modulated interactions that may be free energetically more favorable over the direct contact pairs.

This study differs from the previous simulation studies on molecular association in that the solutes are restricted to move only along the vertical stacking direction. Such a restriction was added because our primary aim is to simulate a prototype for the base stacking found in nucleic acid structures. The calculated $w_{ss}(R)$ shows two distinct

minima, indicating a contact benzene pair and a short-range solvent-separated benzene pair, occuring at a distance considerably less than $\sigma_{benzene}+\sigma_{water}$. The shape of $w_{ss}(R)$ for stacking of benzenes suggests the possibility of several long-range solvent separated benzene pairs, but in this study only the short-range solvent-separated benzene pair is characterized. In this connection, we note that Parthasarathy et. al.[105b] have found intercalated waters in the crystal structures of purines and pyrimidines. These correspond to long-range solvent-separated pairs and it would be interesting to extend the calculations reported herein to study such pairs.

References.

1. J.A.V. Butler, Trans. Faraday Soc., 33, 29(1937).

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

3. A. Lannung, J.Amer.Chem.Soc., 52, 68(1930).

4. S. Valentiner, Zeits.f.Physik., 42, 253(1927).

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

6. (a) G. Bodo, H.M. Dintzis, J.C. Kendrew and Wyckoff, Proc.Roy.Soc. A, 253, 70(1959); (b) M.F. Perutz, J.C. Kendrew and H.C. Watson, J.Mol.Biol., 13,669(1965).

7. B. Lee and F.M. Richards, J.Mol.Biol., 55, 379-400(1971).

8. W. Kauzmann, Advances in Protein Chem., 14,1,(1959).

9. (a) G. Nemethy and H.A. Scheraga, J.Chem.Phys.,36, 3382(1962); (b) G. Nemethy and H.A. Scheraga, ibid.,36,3401(1962); (c) G. Nemethy and H.A. Scheraga, J.Phys.Chem.,66,1773(1962).

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

11. (a) P.A. Giguere and K.B. Harvey, Can.J.Chem., 34, 798(1956); (b) J.W. Schultz, PhD thesis, Brown University, June 1957.

12. J. Morgan and B.E. Warren, J.Chem.Phys., 6, 666(1938).

13. C. Tanford, J.Amer.Chem.Soc., 84, 4240(1962).

14. E.J. Cohn and J.T. Edsall, in "Proteins, Amino acids and Peptides", Reinhold Publishing Corp., New York, 1943. (Chapter 9).

15. R.A. Pierotti, J.Phys.Chem., 69, 281(1965).

16. H. Reiss, H.L. Frisch and J.L. Lebowitz, J.Chem.Phys., 31, 369(1959).

17. J.J. Kozak, W.S. Knight and W. Kauzmann, J.Chem.Phys.,48, 675(1968).

18. W. McMillan and J. Mayer, J.Chem.Phys., 13, 176(1945).

19. (a) P.J. Flory, J.Chem.Phys., 9, 660(1941). (b) M.L. Huggins, J.Chem.Phys., 9, 440(1941). (c) E.A. Guggenheim and M.L. McGlashan, Proc.Roy.Soc. A, 203, 435(1950).

20. J.H. Hildebrand, J.Phys.Chem.,72,1841(1968).

21. G. Nemethy, H.A. Scheraga and W. Kauzmann, J.Phys.Chem., 72, 1842(1968).

22. E.v. Goldammer and H.G. Hertz, J.Phys.Chem., 74, 3734(1970).

23. (a) A. Ben-Naim, J.Chem.Phys., 54, 1387(1971); (b) A. Ben-Naim, in "Hydrophobic Interactions", Plenum Press, New York, 1980.

24. A. Ben-Naim, J. Wilf and M. Yaacobi, J.Phys.Chem., 77, 95(1973).

25. V.G. Dashevsky and G.N. Sarkisov, Mol.Phys., 27, 1271(1974).

26. A.I. Kitaygorodsky, K.V. Mirskaya and V.V. Nauchitel, Kristallographiya(in Russian), 14, 900(1969).

27. O.W. Howarth, J.Chem.Soc.(London) Faraday Trans. I, 71, 2303(1975).

28. R.B. Hermann, J.Phys.Chem., 75, 363(1971).

29. W.L. Masterson, J.Chem.Phys., 22, 1830(1954).

30. R.B. Hermann, ibid., 76, 2754(1972).

31. C. McAuliffe, J.Phys.Chem., 70, 1267(1966).

32. H.L. Friedmann and C.V. Krishnan, J.Solution.Chem.,2, 119(1973).

33. R.W. Gurney, In "Ionic Processes in Solution", Dover Publications, New York, 1962.

34. A. Ben-Naim, Biopolymers, 14, 1337(1975).

35. (a) R.B. Hermann, J.Phys.Chem., 79, 163(1975).

(b) R.O. Neff, and D.A. McQuairre, J.Phys.Chem., 77, 413(1973).

36. R. Wolfenden, C.A. Lewis Jr., J.Thoer.Biol., 59, 231(1976).

37. G.G. Schlessinger, in "Handbook of Chemistry and Physics", pp. D151-D170, Chemical Rubber Co., Cleveland, 1972.

38. A.H. Clark, F. Franks, M.D. Pedley and D.S. Reid, J.Chem.Soc.(London), Faraday I, 73, 290(1977).

39. S.J. Gill and I. Wadso, Proc.Natl.Acad.Sci., 73, 2955(1976).

40. (a) F. Franks, J.Chem.Soc.(London) Faraday Trans. I, 73, 830(1977); (b) F. Franks, in "Water- A Comprehensive Treatise", F. Franks, Ed., Plenum Press, New York, 1975,Vol. 4, Chapter 1. (c) F. Franks and D.G.J. Ives, Quart.Rev.Chem.Soc., 20, 1(1966).

41. K. Shinoda, J.Phys.Chem., 81, 1300(1977).

42. D. Ronis, E. Martina and J.M. Deutsch, Chem.Phys.Lett., 46, 53(1977).

43. R.D. Cramer, J.Amer.Chem.Soc., 99, 5409(1977).

44. L.R. Pratt and D. Chandler, J.Chem.Phys., 67, 3683(1977).

45. J.D. Weeks, D. Chandler and H.C. Anderson,

J.Chem.Phys., 54, 5237(1971).

46. A.H. Narten and H.A. Levy, J.Chem.Phys., 55, 2263(1971).

47. S. Marcelja, D.J. Mitchell, B.W. Ninham and M.J. Sculley, J.Chem.Soc.Farad.Trans. II, 73, 630(1977).

48. J.A. Pople, Proc.Roy.Soc. A, 205, 163(1951).

49. B.V. Deryaguin and L.D. Landau, Acta Physico-Chem., URSS, 14, 633(1941).

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

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

52. D.Y.C. Chan, D.J. Mitchell, B.W. Ninham and B.A. Pailthorpe, Chem.Phys.Lett., 56,533(1978).

53. S. Swaminathan, S.W. Harrison and D.L. Beveridge, J.Amer.Chem.Soc., 100, 5705(1978).

54. C.S. Pangali, M. Rao and B.J. Berne in "Computer Modeling and Matter", P. Lykos, Ed., American Chemical Society, Washington D.C., 1978.

55. A. Geiger, A. Rahman and F.H. Stillinger, J.Chem.Phys., 70, 263(1979).

56. F.H. Stillinger and A. Rahman, J.Chem.Phys., 60, 1545(1974).

57. E.E. Tucker and S.D. Christian, J.Phys.Chem., 83,426(1979).

58. Z. Elkoshi and A. Ben-Naim, J.Chem.Phys., 70, 1552(1979).

59. G.M. Bell, J. Math.Phys., 10, 1753(1969).

60. S. Okazaki, K. Nakanishi and H. Touhara, J.Chem.Phys., 71, 2421(1979).

61. S. Swaminathan and D.L. Beveridge, J.Amer.Chem.Soc., 101, 5832(1979).

62. (a) C. Pangali, M. Rao and B.J. Berne, J.Chem.Phys., 71, 2975(1979); (b) C. Pangali, M. Rao and B.J. Berne, ibid., 71, 2982(1979).

63. M.H. Kalos, J.K. Percus and M.Rao, J.Stat.Phys., 17, 111(1977).

64. P.K. Mehrotra, M. Mezei and D.L. Beveridge, J.Chem.Phys., 78, 3156(1983).

65. L.R. Pratt, and D. Chandler, J.Chem.Phys., 73, 3434(1980).

66. P.J. Rossky and H.L. Friedmann, J.Phys. Chem., 84, 587(1980).

67. G. Bolis and E. Clementi, Chem.Phys.Lett., 82, 147(1981).

68. S. Goldmann, J.Chem.Phys., 74, 5851(1981).

69. Rapaport and H.A. Scheraga, J.Phys.Chem., 86, 873(1982).

70. A. Ben-Naim, "Water and Aqueous Solutions"; Wiley: New York, 1972.

71. J.A. Barker and D. Henderson, Rev.Mod.Phys., 48, 587(1976).

72. (a) J.G. Kirkwood In "Theory of Liquids", Ed. B.J. Alder, Gordon and Breach, New York, 1968. (b) M. Mezei, S. Swaminathan and D.L. Beveridge, J.Amer.Chem.Soc., 100, 3255(1978).

73. G. Torrie and J.P. Valleau, J.Comp.Phys., 23, 187(1977).

74. M. Mezei, Mol.Phys., 47, 1307(1982).

75. (a) S. Romano and K. Singer, Mol.Phys., 37, 1765(1979). (b) G. Jacucci and N. Quirke, Mol.Phys., 40, 1005(1980). (c) N. Quirke and G. Jacucci, Mol.Phys., 45, 823(1982).

76. H.C. Anderson, J.Chem.Phys., 72, 2384(1980).

77. M. Mezei, S. Swaminathan and D.L. Beveridge,

J.Chem.Phys., 71, 3366(1979).

78. (a) C.Pangali, M. Rao and B.J. Berne, Chem.Phys.Lett., 55, 413(1979). (b) M. Rao, C. Pangali and B.J. Berne, Mol.Phys., 37, 1773(1979).

79. M. Mezei, Chem.Phys.Lett., 74, 105(1980).

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

81. G.N. Patey and J.P. Valleau, J.Chem.Phys., 61, 534(1974).

82. J.G. Kirkwood In "Theory of Liquids", Ed. B.J. Alder, Gordon and Breach, New York, 1968.

83. M. Mezei, S. Swaminathan and D.L. Beveridge, J.Amer.Chem.Soc., 100, 3255(1978).

84. G.S. Fishman, In "Principles of Discrete Event Simulation"; Wiley-Interscience, New York, 1978.

85. A.M. Law and W.D. Kelton, Tech.Rep.No. 78-56, Univ. WI, 1978.

86. P. Heidelberger and P.D. Welch, Comm. of ACM, 24, 233(1981).

87. P.H. Peskun, Biometrika, 60, 3(1973).

88. P.J. Rossky, J.D. Doll and H.L. Friedman, J.Chem.Phys., 69, 4628(1978).

89. J.C. Owicki, In "Computer Modelling of Matter", Ed. P.G. Lykos, American Chemical Society, Washington D.C., 1978.

90. B. Bigot and W.L. Jorgensen, J.Chem.Phys., 75, 1944(1981).

91. A. Ben-Naim In "Water and Aqueous Solutions", Plenum Press, New York, 1974.

92. F.T. Marchese, P.K. Mehrotra and D.L. Beveridge, J.Phys.Chem., 86, 2592(1982).

93. D.N. Glew, J.Phys.Chem., 66, 605(1962).

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

95. A.H. Narten, J.Chem.Phys., 55, 2268(1971).

96. a) F.T.Marchese, P.K.Mehrotra, and D.L.Beveridge, in "Biophysics of Water", F.Franks and S.Mathias, eds., John Wiley, New York, 1982. b) F.T.Marchese, P.K.Mehrotra, and D.L.Beveridge, J.Phys.Chem., submitted for publication. c) M.Mezei, P.K.Mehrotra, and D.L.Beveridge, MS in preparation.

97. J.T. Edsall and H.A. McKenzie, Adv.Biophys., 10, 137(1978).

98. Peter Backx, and Saul Goldman, J.Phys.Chem., 85, 2975(1981).

99. Malay K.Dutta-Choudhury, Nada Mlljevic, and W.Alexander Van Hook, J.Phys.Chem., 86, 1711(1982).

100. G. Karlstrom, P. Linse, A. Wallqvist, and B. Jonsson, J.Amer.Chem.Soc., 105, 3777(1983).

101. E. Clementi, F. Cavallone, and R. Scordamaglia, J.Amer.Chem.Soc., 99, 5531(1977).

102. P.K.Mehrotra, and D.L.Beveridge, J.Amer.Chem.Soc., 102, 4287(1980).

103. M. Mezei, S. Swaminathan, and D.L. Beveridge, J.Chem.Phys., 71, 3366(1979).

104. D.L.Beveridge, M.Mezei, P.K.Mehrotra, F.T.Marchese, G.Ravishanker, T.R.Vasu, and S.Swaminathan, in "Molecular-Based Study and Prediction of Fluid Properties", J.M.Haile and G.Mansoori, eds., Advances in Chemistry Series, American Chemical Society, in press.

105.(a) W. Saenger in "Principles of Nucleic Acid Structure", Springer-Verlag, New York, 1984. Chapter 6. (b) R. Parthasarathy, T. Srikrishnan and S.L. Ginell, Proceedings of the second SUNYA conversation in the Discipline of Biomolecular Stereodynamics, Ed. R.H. Sarma, 1, 261(1981).

106. P.O.P. Ts'o, in "Basic Principles in Nucleic Acid

Chemistry", Ed. P.O.P. Ts'o; 1, Academic Press, New York, 1974.