Molecular Dynamics Simulations of 1,2-Disubstituted (Hydroxy- and Amino-) Ethanes: Pure Molecular Liquids and Their Aqueous Solutions

by

Anna V. Gubskaya

M.Sc., Kharkov State University
Ph.D., Institute for Low Temperature Physics and Engineering,
Ukrainian National Academy of Science
M.Sc., Dalhousie University

Submitted in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

at

Dalhousie University

Halifax, Nova Scotia

August 2003

Your file Votre référence
ISBN: 0-612-83719-X
Our file Notre référence
ISBN: 0-612-83719-X

DALHOUSIE UNIVERSITY

DEPARTMENT OF CHEMISTRY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "Molecular Dynamics Simulations of 1,2-Disubstituted (Hydroxy- and Amino-) Ethanes: Pure Molecular Liquids and Their Aqueous Solutions" by Anna V. Gubskaya in partial fulfillment for the degree of Doctor of Philosophy.

Dated: August 18, 2003

External Examiner: [redacted]

Research Supervisor: [redacted]

Examining Committee: [redacted]

[redacted]

[redacted]

Departmental Representative: [redacted]

DALHOUSIE UNIVERSITY

DATE: August 2003

AUTHOR: Anna V. Gubskaya

TITLE: "Molecular Dynamics Simulations of 1,2-Disubstituted (Hydroxy- and Amino-) Ethanes: Pure Molecular Liquids and Their Aqueous Solutions"

DEPERTMENT OR SCHOOL: Chemistry

DEGREE: Ph.D. CONVOCATION: October YEAR: 2003

Permission is herewith granted to Dalhousie University to circulate and to have copied for non-commercial purposes, as its discretion, the above title upon the request of individuals and institutions.

Signature of Author

DEDICATION

To all I love.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF ILLUSTRATIONS

Page

ABSTRACT

This thesis is a comparative computational study of the local structure of three widely used representatives of 1,2-disubstituted ethanes, namely ethylene glycol (EG), ethylenediamine (ED) and 2-aminoethanol (AE), in liquid state and their mixtures with water. Classical molecular dynamics combined with three-dimensional atomic density maps, known as spatial distribution functions (SDF's), are the computational tools used in this study.

The present work consists of three parts. In the first part, twelve molecular models were designed and gas-phase simulations were carried out for each. The results obtained were compared with the most reliable experimental estimates in order to test different force fields and molecular representations. In the second part liquid-phase simulations were performed on the most successful (AMBER/OPLS-based) models. The heats of vaporization and self-diffusion coefficients were used as criteria for the final selection of molecular models to be employed in the subsequent simulations of aqueous solutions.

In the third part, a detailed structural analysis was performed. As an essential part of this analysis the dihedral angle distributions were calculated and relative populations of conformers with respect to the central dihedral angle were determined for pure EG, ED and AE, as well as their mixtures with water, where four compositions of each compound were considered. It has been confirmed that in the liquid phase the *gauche* conformation accounts for the major population of rotational isomers for EG and AE, while ED exhibits a significant population of *trans* conformers. Additionally, the first theoretical estimates of the compositional dependence of self-diffusion coefficients for the aqueous solutions of EG, ED and AE were obtained. The analysis of radial distribution functions in conjunction with calculated numbers of nearest neighbors around oxygen and nitrogen atoms of the main functional groups provided some structural insights into the H-bonding pattern of the systems studied. The number of strongly H-bonded neighboring groups was determined and their possible positions were located by means of SDF's. The possibility of four-membered H-bond arrangements (comprised of two strong and two weak H-bonds) found around oxygen and nitrogen atoms leads to the conclusion that in the liquid phase the generalized H-bonding pattern for EG, ED and AE can be described as a three-dimensional, branched network.

LIST OF SYMBOLS USED

LATIN SYMBOLS

$\vec{a}$ - acceleration;

$A(t)$, $B(t)$ - some dynamic quantities;

a, b - coefficients;

ac - plane;

b - scattering length;

$\vec{b}$ - third time derivative of position vector;

c_0, c_1, c_2 - Gear coefficients in equations of motion;

$C(t)$ - normalized velocity autocorrelation function;

$C_{AB}(t)$ - time correlation function between properties A and B;

C_{2h} - symmetry group;

C_{2h}^5 - crystallographic space group;

C_v - constant volume heat capacity;

2-D, 3-D - two, three dimensional;

D - self-diffusion coefficient;

d - equilibrium bond length;

<E>- average potential energy;

ΔE_{T-G} - energy difference between *trans* and *gauche* conformers;

E_{HB}- energy of hydrogen bond;

$E_{inter}(g)$, $E_{intra}(g)$ - inter- and intramolecular energy for the gas;

$E_{inter}(l)$, $E_{intra}(l)$ - inter- and intramolecular energy for the liquid;

$f(k)$ - atomic form factor;

$f(\theta)$ - intensity (amplitude) of the scattered component;

f_i - force on atom i;

g - number of degrees of freedom of the physical system;

$g(r)$- radial distribution function;

$g^{(2)}(\vec{r}^2,\vec{\omega}^2)$, $g(\vec{r}_{12},\vec{\omega}_2)$ - pair distribution functions;

$g_{\alpha\beta}(r)$ - site-site distribution function;

$g^D(r)$, $g^D_{\alpha\beta}(r)$ - site-site pair correlation function obtained from diffraction data;

ΔG_{hyd} - free energy of hydration;

$\vec{g}_i$ - constraint force on atom i;

$g^{(N)}$- N-body distribution function;

$g^{ref}_{\alpha\beta}(r)$- reference value of site-site pair correlation function;

$\langle g(\bar{r}_{12},\bar{w}_2)\rangle_{\vec{w}_2}$ - spatial distribution function;

$g_{OO}(r,\Omega)$, $g_{OC}(r,\Omega)$ - spatial distribution functions;

$H(\vec{p},\vec{q},t)$ - Hamiltonian function;

$h(l_1l_2l;n_1n_2;r)$ - trial coefficients in Fourier-Bessel expansion;

i, j, k - labels for atoms;

$j_l(kr)$ - spherical Bessel function of order l;

k - wavenumber;

<K> - average kinetic energy;

k_B - Boltzmann's constant;

$K_{b,a,t,\chi}$ - force constants;

$L(\vec{q},\dot{\vec{q}})$ - Lagrangian function;

M - number of constraints;

m_i - mass of atom i;

m_t - multiple factor for torsion angle;

N - number of molecules, atoms, nuclei;

P - pressure;

$P2_1$, $P2_1/n$, $P2_1/c$, $P2_12_12_1$ - crystallographic space groups;

p_k - generalized momentum;

Q - thermal inertia parameter;

q_i, q_j - charge on the atoms i and j;

q_k - generalized coordinate;

R - gas constant;

$\vec{r}$ - position vector;

$\vec{r}_{12}$ - the separation vector between molecules 1 and 2;

r_{eq} - equilibrium value for the bond length;

r_i - nuclear coordinates;

r_{ij} - distance between the atoms i and j;

r_i^n - coordinates of the n-th electron of the i-th atom;

$S(k)$ - structure factor;

$S_{\alpha\beta}(k)$ - partial structure factor;

$S(l_1 l_2 l; n_1 n_2; k)$ - Fourier-Bessel coefficients for the partial structure factors;

t - time;

T - temperature in degrees kelvin;

T_A - actual temperature in degrees kelvin;

T_D - desired temperatures in degrees kelvin;

U - potential energy of the configuration;

<U> - total average configurational energy;

$<U_a>$, $<U_d>$ - average bond and dihedral angles energies;

$u(r)$ - atomic potential;

$U^0_{\alpha\beta}(r)$, $U^{ref}_{\alpha\beta}(r)$, $U^N_{\alpha\beta}(r)$ - initial, reference and new potential, between sites α and β;

$<U_Q>$, $<U_{LJ}>$ - average Coulombic and Lennard-Jones energies;

V - volume, force field potential;

$\vec{v}$ - velocity;

$V(\vec{r})$ - scattering potential;

x_i - group electronegativity;

XZ - plane;

Z - atomic number.

GREEK SYMBOLS

α, β - label for atoms (sites);

χ - out-of-plane (improper dihedral) angle;

χ^2 - statistical factor;

$\Delta(l_1 l_2 l; n_1 n_2; r)$ - perturbation coefficients in SHR technique;

ε - Lennard-Jones parameter;

$\varepsilon(r)$ - experimental error in RMC technique;

ϕ - dihedral angle, angle in the spherical-polar representation;

$\lambda_\alpha, \gamma_\alpha$ - undetermined multipliers associated with constraints;

μ - chemical potential;

ν - frequency of normal mode;

θ - polar angle;

θ_{eq} - equilibrium values for the bond angles;

ρ - density;

ρ_n - number density;

σ - Lennar-Jones parameter;

σ^* - antibonding orbital;

$\sigma(\vec{r})$ - constraint;

$\vec{\omega}$ - orientation vector;

Ω - scattering angle;

ξ - friction coefficient;

$\psi_{\alpha\beta}(r)$ - potential of mean force between atoms α and β;

$\psi^{D}_{\alpha\beta}(r)$ - potential of mean force from diffraction data.

LIST OF ABBREVIATIONS USED

AE - aminoethanol, $NH_2(CH_2)_2OH$;

AM1-SM1a, AM1-SM2 - Austin Model 1 Solvation Models;

AMBER - Assisted Model Building with Energy Refinement;

B3LYP - Becke-Lee-Yang-Parr functional;

cc-VDZ- correlation consistent polarized Valence Double Zeta basis set;

cc-VTZ - correlation consistent polarized Valence Triple Zeta basis set;

CHARMM - Chemistry at Harvard Macromolecular Mechanics;

CN - Coordination Number;

DD - Domain Decomposition method;

DFT - Density Functional Theory;

DMSO - dimethylsulfoxide;

DNA - deoxyribonucleic acid;

DSDF - Difference Spatial Distribution Function;

DX - Data Explorer;

ED - ethylene diamine, $(CH_2NH_2)_2$;

EG - ethylene glycol, $(CH_2OH)_2$;

EPEN - Empirical Potential using Electrons and Nuclei;

EPMC - Empirical Potential Monte Carlo;

EXAFS - Extended X-ray Absorption Fine Structure method;

GROMOS - Groningen Molecular Simulations;

H1, H2 - molecular models based on Haughney's force field;

HA - Hydrogen Acceptor;

HD - Hydrogen Donor;

HF - Hartree-Fock;

HH - Hydrophobic Hydration region;

HTN - Hayashi, Tanaka and Nakanishi;

IR - Infra Red;

IUPAC - International Union of Pure and Applied Chemistry;

J, J1, J2, JMOD - molecular Models based on W.L. Jorgensen's force field;

LCAO-SCF - Linear Combination of Atomic Orbitals - Self-Consistent Field method;

LJ - Lennard-Jones;

MC - Monte Carlo simulations;

MC/FEP - Monte Carlo simulations with Free Energy Perturbation Calculation;

MD - Molecular Dynamics;

ME - methanol;

MM2, MM3 - Molecular Mechanics potentials;

MMC - Metropolis Monte Carlo;

MNDO - Modified Neglect of Diatomic Overlap;

MP - Møller-Plesset perturbation theory;

MPI - Message Passing Interface;

NBO - Natural Bond Orbital theory;

NMR - Nuclear Magnetic Resonance;

OPLS - Optimized Potential for Liquid Simulations;

PEG - polyethylene glycol;

PM3-SM3 - Parameterized Model 3 Solvation Models;

PSF - Partial Structure Factor;

RD - Replicated Data method;

RDF - Radial Distribution Function;

RISM-SCF - Reference Interaction Site - Self-Consistent Field method;

RMC - Reverse Monte Carlo;

SDF - Spatial Distribution Function;

SHR - Spherical Harmonic Reconstruction;

SPC/E - Single Point Charge, Extended water model;

TCF - Time Correlation Function;

TIP4P - Transferable Intermolecular Potential applied to 4-Points (sites) model of water;

TIPS - Transferable Intermolecular Potential for fluid Simulations;

TRANAL - TRAajectory ANALysis;

WP - Widmalm and Pastor.

ACKNOWLEDGMENTS

I would like to thank my supervisor Dr. Peter G. Kusalik for his friendship, guidance, encouragement and financial assistance.

I thank Dr. A. Laaksonen and Dr. A. Lyubartsev for giving me an opportunity to use M.DynaMix simulation package and numerous helpful tips they provided with regard to utilization of this software.

I express my sincere gratitude to my family for their endless love, understanding, patience, and permanent moral support.

I am also thankful to my friend Charlotte Harper for her ability always to offer a hand when it is most wanted.

Finally, I greatly appreciate all the help from members of the Chemistry Department during the period of my study at Dalhousie University.

CHAPTER I - INTRODUCTION

This work builds on a series of investigations which have focused on the spatial solvation structure within selected strongly associated, hydrogen-bonded liquids and solutions. One of the primary tools in this investigation is three-dimensional atomic density maps, known as the spatial distribution functions (SDF's). The results obtained for the liquids and solutions studied so far (see Chapter III) confirmed that the SDF's provide a very detailed picture of the solvation structure. They make it possible to define much more clearly and quantitatively a "hydrogen bond" and allow the investigation of the average local environment around the molecule within the liquid. The present work is a comparative investigation of the local structure in pure liquids and the solvation structure in aqueous solutions for a set of "simple" molecules with two functional groups, each of which can simultaneously act as donors and acceptors of hydrogen bonds. A detailed analysis of such relatively small molecular systems should help us to understand better the role of cooperative effects in the hydration structure of these particular molecules as well as to provide insights into the hydration structure and behavior of larger molecules containing similar functional groups.

1,2-Disubstituted ethanes are the class of saturated molecules with the general formula XCH_2CH_2Y where the functional groups X and Y are, for example, F, Cl, NH_2, OH [1]. In the present study the common and widely used representatives of this class containing amino and hydroxyl groups, in particularly, $(CH_2NH_2)_2$ (ethylenediamine), $NH_2(CH_2)_2OH$ (2-aminoethanol) and $(CH_2OH)_2$ (ethylene glycol), henceforth ED, AE and EG, respectively, will be considered.

Ethylenediamine is one of a variety of molecules containing two amino groups which is used extensively in both organic and inorganic (coordination) chemistry. ED and similar larger molecules (1,3-diaminopropane and 1,4-diaminobutane) are widely used as chelating agents for some metal ions and are precursors of biologically relevant polyamines (e.g. spermidine and spermine) that play an essential role in cell growth and differentiation in eukaryotic organisms. Moreover, some metal complexes (e.g. Pt^{2+}) formed with these amines are known to display anticancer properties through DNA-binding [2]. The exact nature of these biochemical mechanisms still remains unknown and requires additional intensive studies.

2-Aminoethanol is a polar organic molecule that contains both an alcohol and amine group. Alkanolamines are mainly used as aqueous solutions in the petroleum industry to eliminate H_2S and CO_2 from natural gas streams. Precise knowledge of the structural and thermodynamical properties of alkanolamine solutions is important for designing and operation of refining and purification processes [3].

Ethylene glycol is the simplest example of a saturated carbon compound with two vicinal hydroxyl groups and represents the monomeric entity of the polyethylene glycol (PEG) polymer. EG as a solvent has many applications especially as a cryogenic liquid, like other members of the family of organic molecules with multiple hydroxyl groups. Aqueous solutions of EG are widely used as antifreezes and cryoprotectants. What makes this substance superior to other polyols is that lower concentrations of EG are needed to obtain the desired results. This lowers the toxicity, which is of crucial importance in cryopreservation of living cells and tissues. Another aspect of this solvent relevant to its use in biology relates to the prevention of denaturation of proteins, which adopt a more ordered structure in polyol solvents or aqueous polyol solutions [4].

One more important issue concerning these three compounds is that they can be used in developing parameters for macromolecular modeling. They can serve as small prototypical systems that contain amino and hydroxyl groups and exhibit both intra- and inter- H-bonded molecular interactions. For example, an accurate depiction of the intramolecular H-bonding patterns of vicinal hydroxyl groups (in EG) is crucial to the modeling of sugars and their polymers [5]. EG is particularly interesting as a model compound due to its amphiphilic character (it possesses at the same time polar and non-polar groups), its importance as an organic solvent, and the occurrence of hydroxyl groups in the side chains for serine, threonine, tyrosine and several less common amino acids [6]. Amphiphilic molecules in aqueous solutions can self-associate into aggregates called micelles [7,8] where the contact between water and hydrocarbon is greatly reduced while the polar groups maintain their hydration [9]. In addition EG is one of the simplest organic molecules which may be considered as a water analog. EG has one of the lowest ratios of weakly polar methylene or methyl groups vs. polar hydroxyl groups in alcohols which means that the properties of this system will be strongly influenced by hydrogen bonding (HB) interactions. Previous studies have shown that likely both EG and water have four nearest H-bonded neighbors on average and both present three-dimensional HB structures [4]. In contrast, the structure of short monohydric alcohols consist of mainly linear winding chains of H-bonded molecules [10-14].

One distinctive feature of pure molecular liquids of ED, AE and EG is the possible competition between inter- and intramolecular interactions (H-bonds) which is not present in water. This also raises an interesting question with respect to aqueous solutions of these compounds: are the intramolecular H-bonds found between amino and hydroxyl groups in the gas phase disrupted in

aqueous solution in order to permit additional H-bonding with water molecules? The present study will help to elucidate both of these issues.

This thesis is organized as follows. A literature review of previous computational results for 1,2-disubstututed ethanes (and analogous systems containing hydroxyl and amino groups) in the gas phase and in the liquid phase is presented in Chapter II. The results of conformational analysis of EG, ED and AE performed using *ab initio* methods, as well as evaluations of the rotational flexibility of these molecules through examination of the potential barriers and contributions from the gauche effect and H-bonded interactions in the lowest energy conformers, will be discussed in detail. Only a few studies of 1,2-disubstituted ethanes (in particular EG and AE) in the liquid phase were found in the literature and hence the trends observed on small analogous molecular systems (e.g. methanol and methylamine) were examined in order to provide grounds for understanding the more complicated flexible molecules in a liquid environment.

Knowledge of the structure specific for the compounds of interest becomes of utmost importance. In the first part of Chapter III the structural results obtained for EG, ED and AE in gas, liquid and solid states from experimental (mainly spectroscopic) and theoretical investigations will be presented and discussed. A separate section is devoted to the analysis of previous X-ray studies taking into account the possibility that some structural features typical for these species in the solid state could survive in a liquid environment. Conventional and novel theoretical approaches explored in the structure determination of molecular liquids as well as their link with experimentally determined parameters will be presented in the second part of this chapter.

A brief description of molecular dynamics simulation, which served as the main simulation technique in the present work, will be given in Chapter IV. In

particular, the most common approaches used in computational practice for solving the equations of motion, in conjunction with some specific procedures (e.g. implementations of constraints, thermostats, velocity scaling) implemented in the computational package utilized, will be discussed in some detail. The distinctive features of the standard interaction potentials (force fields) and the potential functions specifically designed for simulations of liquid alcohols, amines and alkanolamines will be outlined. In the final (methodological) section the selected force field parameters employed in the present study and simulation specifications are given as well as descriptions and definitions for the properties of interest and characteristics used in structural analysis.

The results of simulations of pure EG, ED and AE and their aqueous solutions followed by elucidation of the local structure in these systems are presented in Chapters V and VI, respectively. The analysis starts with a comparison of simulated gas-phase geometries and conformations for different models created in order to help identify the best potential representation for each of the compounds. After elimination of the least successful models, further criteria are proposed and on their basis some models are chosen for liquid-phase simulations. During the simulations performed for pure compounds, the different force field regimes (with and without scaling of 1-4 nonbonded interactions) are also examined in order to achieve the best agreement with available experimental data for each of the compounds. The detailed structural analysis of the pure systems and their aqueous mixtures was performed by means of radial and spatial distribution functions; the results of this structural elucidation are presented and discussed in a comparative manner. Finally in the Chapter VII concluding remarks are made as well as proposals with regard to the most suitable strategy for future computational investigations of this class of molecules.

CHAPTER II - SURVEY OF PREVIOUS COMPUTATIONAL RESULTS

A detailed review of previous theoretical studies of 1,2-disubstituted ethanes and some relative compounds are presented in this chapter. In the gas phase the conventional conformational analysis of EG, ED and AE, which includes determination of the most stable unique conformers and their relative stabilities, has been performed by means of *ab initio* techniques at the HF and MP2 levels of theory. There have also been some interesting attempts to evaluate the rotational flexibility of these molecules through examination of the heights of potential barriers for the most important conformational interconversions. The relative contributions to the stability of the lowest energy conformers due to the H-bonded interactions and the *gauche effect* were also examined in a comparative manner for the compounds of interest. A discussion devoted to the theoretical studies of liquids and aqueous solutions of simple non-electrolytes starts by reviewing previous results obtained for methanol and methylamine. The trends observed for these small analogous molecular systems will aid in furthering our understanding of the behavior specific to more complicated flexible molecules in a liquid environment represented by molecules of the same kind or by molecules of a polar solvent (water). Only a few investigations of EG and AE in liquid state have been published and they will be described in some detail; it appears that there are no computational studies reported to date on liquid ED and its aqueous solutions.

2.1 THEORETICAL STUDIES OF THE GAS-PHASE SYSTEMS

2.1.1 Conformational analysis

Several conformational studies based on theoretical calculations are reported in the literature for ED, AE and EG [1,2,5,16-26]. Conformational

isomers of the set of saturated molecules XCH_2CH_2Y differ by the specification of the three possible rotameric dihedral angles, one about C-C bond and one about each of the C-X and C-Y bonds. Assuming typical 3-fold torsion potentials with minima for angles 60°, 180° and 300° (-60°), this gives rise to 27 (i.e. 3^3) conformational minima. Symmetry makes many of these degenerate; 10 unique possible conformers can be generated for ED and EG and 14 for AE. The notation for identifying rotational isomers of the XCH_2CH_2Y is given below.

1. The uppercase symbols T (*trans*) and G (*gauche*) are used to describe the arrangement of X and Y about the central C-C bond [1].
2. Orientations about the C-X and C-Y bonds (where X or Y are NH_2 or OH) are defined by CCOH and CCN(L) dihedral angles for OH and NH_2 groups, respectively (where (L) is used for the lone pair direction) [1].
3. The lowercase symbols t, g and g' refer to the positions of OH or NH_2 groups. The prime is used to distinguish the positive direction of the rotation from the negative one [1,15]. In Ref. 1 the appropriate dihedral angles are 180, +60 and -60° for t, g and g', respectively (see Fig. 1).
4. The rotational angle is defined from either side of the axis. The positive direction is defined such that the front bond is twisted anticlockwise against the rear bond (the same as IUPAC definition).
5. In the present study the range of the rotational angle ϕ is $-30^\circ < \phi < 30^\circ$ for the *cis*; $-90^\circ < \phi < -30^\circ$ or $30^\circ < \phi < 90^\circ$ for the *gauche*; and $-150^\circ < \phi < 150^\circ$ for the *trans* conformation [15].

It is well known that 1,2-disubstituted ethanes (XCH_2CH_2Y) possess internal rotation about the central C-C bond which leads to nonequivalent staggered arrangements corresponding to a *trans* and a pair of *gauche* structures [1]. If the internal rotation potential function has minima near these

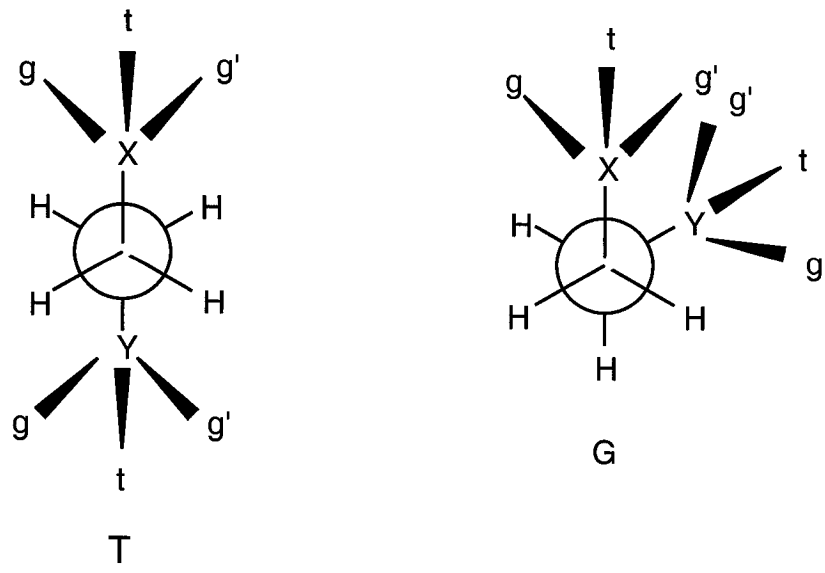


Figure 1. Notation used to specify rotational isomers [1].

staggered arrangements, the molecule will have distinct rotational isomers (rotamers). Radom *et al.* [1] used standard LCAO-SCF orbital theory with the extended 4-31G basis set to obtain a complete set of total and relative energies for staggered conformations of XCH_2CH_2Y molecules. They also presented calculated and experimental values of the energy change for the formal reaction

$$XCH_2CH_2Y + CH_3CH_3 \rightarrow CH_3CH_2X + CH_3CH_2Y \ . \qquad (2.1)$$

These energies measure 1,3 interactions between bonds. It has been shown that for the ED molecule the most stable conformations (among 10 examined) are tGg' and gGg' (see Fig. 2, Chapter 3), which are stabilized by intramolecular H-bonds (N-H···N). The 1,3-interaction energies for these rotamers are stabilizing while all remaining conformers have negative interaction energies [1]. The semiempirical MNDO/HB method was used in the work of Rodnikova *et al.* [16] to evaluate the relative stability of the *cis*, *gauche* and *trans* conformers of ED. The

calculations were performed with optimization of all geometric parameters except the bond lengths and bond angles, which were set accordingly to the experimental values. The results for the heats of formation and the dipole moments of the *cis*, *gauche* and *trans* conformations obtained by rotation of one NH_2 group around the C-C bond showed that the tGg' conformation with the optimal value of the NCCN dihedral angle of 65° has the greatest stability [16]. A different conformer, gGg', was identified as the most stable in the work of Chang *et al.* [17], where the global conformational potential was obtained at the MP2/6-311+G(2d,p) level of theory. Kudoh *et al.* [18] reported results of DFT calculations performed to optimize the geometrical structures with a 6-31++G** basis set and the B3LYP density functional. Again, in contrast to the results obtained in other work [1,16] these authors reached the conclusion that the most stable conformer is gGg' and the second most stable conformer is tGg', although their energy difference was only 0.63 kJ/mol. This result is in excellent agreement with data reported by Bultinck *et al.* [19] where the Hartree-Fock (HF) level of theory with the same basis set was used to calculate the relative energies of ED. In addition it has been shown [18] that tTt is the most stable among the *trans* (around the C-C axis) conformers, however, all the *trans* conformers are less stable than gGg', tGg' and gGg.

A majority of the theoretical as well as experimental studies indicate that a free molecule of EG exists only in the *gauche* arrangement with respect to the central bond. Radom *et al.* [1] predicted the tGg' form (see Fig. 3, Chapter 3) with an intramolecular O-H···O H-bond to be the most stable and the only low energy conformation. However, electron correlation effects, which are especially important in systems with strong non-bonded interactions, were not included in their calculations. Chang and co-authors [17,20] found the same global minimum for EG in gas-phase corresponding to the tGg' conformer; their calculations were

performed at the MP2 level of theory. Teppen *et al.* [5] optimized the geometries of 10 conformers of EG at the HF/4-21G, HF/6-311G** and MP2/6-311G** level in order to better understand intramolecular H-bonding and to develop more reliable energy and geometric parameters for future molecular modeling. Absolute and relative energies demonstrated that correlation effects split the conformers into two broad energy groups, those within 7.5 kJ/mol of the global minimum (the tGg' conformer) and those more than 11.7 kJ/mol above the minimum. The first group is stabilized by intramolecular H-bonds that are enhanced by the delocalizing effects of electron correlation. Compared with HF energies, the MP2 energies of the second group rise relative to the most stable conformer only by an average of 1.7 kJ/mol. These authors also showed that the inclusion of electron correlation markedly influences the absolute values of the geometric parameters. For instance, the MP2-optimized structures have C-O and O-H bonds up to 0.02 Å longer, OCC and HOC angles 1 and 4° smaller, respectively, and torsional angles deviating from ideal trans or gauche values by an additional 5-8° compared with HF/6-311G** results. The net effect of these changes is to greatly enhance intramolecular hydrogen bonding. It was concluded that electron correlation selectively stabilizes the conformers with more intramolecular H-bonds, but decreases the energy differences among conformers with the same number of H-bonds [5].

Following the same aim as Teppen *et al.* [5], Cramer and Truhlar [21] performed correlated *ab initio* calculations on 10 rotational isomers of EG at both the HF and MP2 levels (with cc-pVDZ and cc-pVTZ basis sets). They also have found some interesting trends. First, isomeric relative energies differ considerably with improving basis set. For identical geometries on going from the cc-pVDZ to the cc-pVTZ basis set, energies relative to the global minimum (tGg' rotamer) change by up to 3.4 kJ/mol at the both HF and MP2 levels. The larger basis set

appears to stabilize conformers with t-hydroxyl orientations relative to the global minimum, which itself has a single t-hydroxyl. The largest stabilizations were observed for the tTt and tGt conformers where both hydroxyl torsions are *trans*. The larger basis set also decreased the energetic separation between the two internally H-bonded conformers and the eight others and this effect is reduced at correlated levels compared to the HF level. In addition it was noted that the difference in potential energy between the lowest energy intramolecular H-bonded conformer and the lowest energy conformer without such a H-bond is 4.184 kJ/mol and also the difference in energy between the lowest energy *gauche* C-C conformer and the lowest energy *trans* conformer is 10.9 kJ/mol. Finally it was predicted that EG adopts a gas-phase population of conformers at 298 K comprised of 98% *gauche* (where 83% of rotamers possess intramolecular H-bonds) and 2% *trans* rotamers about the C-C bond [21].

For AE there are several conformations in which intramolecular hydrogen bonding is possible and these have been found in computational studies to have the lowest energies. Radom *at al.* [1] reported g'Gg' (see Fig. 4, Chapter 3) with an N···H-O H-bond as the most stable form. The gGt and tGt conformers both have N-H···O type H-bonds and they are, respectively, 1.17 and 2.88 kJ/mol higher in energy than g'Gg'. The intramolecular H-bonds in these three conformations are reflected in positive 1,3-interaction energies. The calculated interaction energies suggest that the N···H-O H-bond is stronger than the N-H···O bond [1]. These estimates are in very good agreement with the previous experimental findings [22,23] (as discussed in Chapter 3). A similar conclusion about the relative strengths of the two possible H-bonds in AE was made by Kelterrer and Ramek [24] who performed *ab initio* self-consistent field (SCF) calculations with a 4-31G basis set. The relative energies of ten symmetry-unique equilibrium geometries of AE indicated that the only conformation with the

N···H-O interaction (g'Gg') corresponds to the global minimum. In this conformation the molecule forms an envelope-like type five-membered ring with an N···H distance of 2.336 Å. The authors concluded that the N···H-O type of H-bond is much stronger than N-H···O ones and its presence has a more significant influence on bond lengths, vibrational frequencies and energetics [24].

All the possible rotamers of AE were more recently studied at a higher *ab initio* level of theory, namely MP2/6-31G**, with correlation energy inclusion and zero-point energy evaluation by G. Buemi [25]. When optimizations of AE were performed in this work 13 out of the 14 conformers expected on symmetry considerations were found. The gGg' form was missing because during the minimization procedure it evolved toward the g'Gg' conformer (the related energy barriers being essentially nonexistent). In agreement with previous literature reports [1,24], the g'Gg' conformer was again found to be the most stable. Its absolute stability with respect to the other forms was not confirmed because (taking into account the very low energy barriers between some conformers) the energy values involved depend on the basis sets used in the calculations and on the inclusion of the correlation energy [25]. Buemi deduced approximate values for the H-bond strength, E_{HB}, as the difference between the energy of the conformer under consideration and the energy of a reference structure showing no H-bond. For the O-H···N bridge of the g'Gg' conformer (assuming tTt or gGt as reference structures) the calculated E_{HB} value was about 17 kJ/mol, which is in reasonable agreement with the value of 13.4 kJ/mol obtained by Penn and Curl [23] using simple point charge models. E_{HB} for the O-H···N bridge of the gGt conformer was found to be more than twice (~8 kJ/mol) that calculated for the N-H···O bridge [25]. Performing calculations at the same level of theory but with an extended basis set (6-311+G(2d,p)) Chang *et al.* [17] confirmed that g'Gg' is the prevailing conformation in AE. In their study they reported relative

energies for 13 rotamers and evaluated the conformational energy decomposition due to the so-called "intrinsic *gauche effect*" and H-bonding. The former issue will be discussed in detail below.

Recently Vorobyov *et al.* [26] performed geometry optimizations of AE conformers using DFT with the B3LYP hybrid functional and the 6–311++G(2d,2p) basis set. In this study the emphasis was placed on the nature of the intramolecular hydrogen bonds between the OH and NH_2 groups in monomers as well as intermolecular H-bonds in AE dimers. Natural bond orbital (NBO) theory [26] was used to analyze H-bonded conformers in terms of orbital interactions. It was found that for the stable conformers of AE the formation of O-H···N or N-H···O intramolecular H-bonds enhances the stability of those conformers, causes elongation of the O-H or N-H bonds, contraction of the COH or CNH bond angles, and red shifts and increases the IR intensity of the $\nu(OH)$ or $\nu(NH_2)$ normal modes. NBO analysis explains these changes as a consequence of remote $n \rightarrow \sigma^*$ electron delocalizations, which bring electron density from the lone pair of a hydrogen bond acceptor into the antibonding orbital of the donor. These energetically favorable charge-transfer processes are the strongest for the global minimum, which corresponds to the O-H···N hydrogen-bonded g'Gg' conformer. Hydrogen bonds of the N-H···O type appear to be much weaker and the stabilizing effect of remote delocalizations is mostly canceled by the opposing stereoelectronic effects [26]. The geometric parameters obtained demonstrate that both intramolecular H-bond types are markedly bent, which results in the increase in steric exchange repulsion and consequent weakening of these interactions. It was further demonstrated in this computational study with sets of AE dimers that intramolecular H-bonds can be easily overcome by stronger and more linear intermolecular H-bonds. The formation of intermolecular H-bonds is the main stabilizing factor in the

dimerization of AE monomers. To investigate possible cooperative effects a few representative structures were considered, in particular, dimers composed of two AE monomers in the gas-phase global minimum g'Gg' conformation and in the most abundant liquid-phase gGt conformation [27] with an N-H···O intramolecular H-bond. Six AE dimers, two of them cyclic, each one with two equivalent intramolecular H-bonds of the O-H···N or N-H···O type, and four open-chain structures with four possible types of intermolecular H-bonds were studied. The trends in interaction energy show that among cyclic structures the dimer with two O-H···N intermolecular H-bonds (gGt/gGt) has the largest magnitude of interaction energy while for the open-chain dimers the magnitude of the interaction energy decreases in the order (gGt/gGt, O-H···N)>(gGt/g'Gg', O-H···O)>(g'Gg'/gGt, N-H···N)>(g'Gg'/g'Gg', N-H···O). Vorobyov *et al.* [26] predicted that it is very likely for open-chain structures composed of gGt conformers with the strong O-H···N intermolecular interactions to be formed in polar condensed media.

2.1.2 Rotational barriers and stabilization factors

In conformational analysis relative values of conformational energy are commonly used as a measure of the stability of one conformation with respect to another. However these data by themselves in fact do not provide sufficient information needed to explain the origin of their apparent high or low stability. In spite of the fact that conformational energy landscapes can be efficiently described in terms of the height of potential barriers, this parameter appears not to be often reported in the relevant literature. It has been estimated theoretically that chain molecules (in particular, n-mer alkanes) possess many high dihedral barriers in the range of 12-25 kJ/mol, which makes their potential surfaces quite complex [28]. Clearly, the highest rotational barriers for these molecules are

associated with rotation around C-C bonds and, therefore, one might expect that an analogous trend could be observed in the case of 1,2-disubstituted ethanes. Unfortunately, there are only a few sources reporting heights of rotational barriers available in the literature for EG, ED and AE and the data presented are not consistent with one another [17,25,29,30].

Nagy *et al.* [29] in their study of the gauche-trans equilibrium in EG looked at the potential energy change as a function of the HOCC and OCCO torsions and concluded that there are high barriers along the minimum energy paths probably hindering certain rotations. They found (assuming that their values are somewhat overestimated by the neglect of geometry optimization) that the C-C rotation has a barrier 33.5-37.5 kJ/mol, where the rotation of the internally bound hydroxyl group requires more than 21 kJ/mol. Buemi [25] estimated the rotational barriers between the most important conformations of AE. The highest barriers are those due to the rotation around the C-C and C-N bonds and, in particular, the interconversions starting from the most stable g'Gg' conformation. For instance, the barrier height for the g'Gg'→g'Tg' transformation (the rotation around C-C) is 32.66 kJ/mol and for gGg'→tGg' (the rotation around C-N) is 32.10 kJ/mol. Surprisingly, extremely low barriers were found for some interconversions involving rotation around the C-O bond (i.e. 0.68 kJ/mol for g'Gt→g'Gg', 1.25 kJ/mol for tGg'→tGt). According to data reported in Buemi's study [25] the average energy penalties for rotation about the C-C and C-N bonds in AE are 23.00 and 16.12 kJ/mol, respectively, while it is 4.25 kJ/mol for rotation around the C-O bond. Burgess *et al.* [30] performed empirical conformational energy calculations of EG, ED and AE exploring their so-called "empirical potential using electrons and nuclei" (EPEN). They have reported values 31.38, 28.7 and 26.98 kJ/mol for ED, EG and AE, respectively, for barriers (averaged when appropriate) to rotation in going from conformer G to conformer

T (i.e. rotation about C-C bond). Thus, the barriers heights decrease in the order ED>EG>AE. It is worth noting that this estimate for the rotational barrier in AE appears to be in good agreement with that calculated by Buemi [25] for the corresponding type of rotation. However, as has been mentioned by Burgess *et al.* [30], the procedure adopted in their work differs significantly from the usual procedures for calculations of conformational minima and rotational barriers and, therefore, any comparisons should done with a caution. This can be illustrated by looking at the values extracted from the minimum energy paths for the rotation around the C-C bond shown in the study of Chang *et al.* [17]. The barrier heights for the interconversions from the global minima are 21.33 (g'Gg→g'Tg), 22.38 (tGg'→tTg'), 23.43 (gGg'→gTg) kJ/mol for ED, EG and AE, respectively. In this case one can immediately notice the trend is opposite to that predicted in Ref. [30]. The results shown above lead one to the conclusion that the absolute numerical values for the rotational barriers can not be directly compared with one another because they appear to be rather sensitive to the computational procedure used, in particular, type of potential, the means of scanning conformational space, the level of theory and basis set employed in calculations. However, some common trends can be revealed. It is clear that in 1,2-disubstituted ethanes the average height of the rotational barrier around the C-C bond lies in the range of 20 to 30 kJ/mol, which is in good agreement with predictions made for n-mer alkanes. There is also a definite trend of decreasing barriers (up to 5 times compared with those for C-C bonds) for the rotation around the C-N and especially the C-O bond. It is very likely (relying on the "quality" of the interaction potential used) that for the rotation involving the central dihedral angle the rotational barriers of ED, EG and AE follow the trend observed in Ref. [17].

Another closely related question arises; namely, what kind of interactions are responsible for the strong stabilizing effects observed and what are their relative contributions for the compounds of interest? The current responses of researchers to these questions have included contributions from two stabilizing factors: H-bonding and the so-called *"gauche effect"*. The *gauche effect* can be defined as the preference of the *gauche* to the *trans* conformation in the molecular structures X-C-C-Y in which X and Y are groups with large electronegativities [17]. Originally the term "*gauche effect*" was applied to 1,2-difluoroethane and related compounds which demonstrated a much greater stability of their *gauche* forms than one might have predicted from steric and electrostatic effects. On the basis of conformational data for hydrocarbons, in which the energy (or enthalpy) of the *gauche* conformation was found to be by approximately 3.3 kJ/mol less stable than that of *trans* conformation, one might expect the 1,2-difluoroethane systems to behave in a similar fashion. This energy difference might be expected to be even smaller for fluorines where the C-F bonds contain sizable dipoles causing larger repulsion in the *gauche* conformation than in the *trans*. This expectation was also in good agreement with the very first *ab initio* calculations performed for these compounds at the HF level using relatively small basis set. These calculations did not reveal any abnormality: the *trans* conformation was found to be more stable. Nevertheless, reliable experiments as well as more recent *ab initio* calculations, which have included electron correlation, definitely show that the *gauche* conformation is actually more stable (by about 4.2 kJ/mol) than the *trans* in the difluoride [31]. The effect has been rationalized as being a consequence of an interaction between the adjacent electron pairs and/or polar bonds of the molecular fragments as the conformation varies; the bond bending caused by large electronegativity of susbstituents destabilizes the *trans* form more than the

gauche [32]. This anomaly, referred in the literature as "the *gauche effect*", was later observed in other 1,2-disubstituted ethane derivatives such as EG [33,34], AE [23], 2-fluoroethanol [35] and ED [36]. Oxygen and nitrogen are less electronegative than fluorine and so one might expect that these atoms cause a similar but less pronounced effect than that caused by fluorine. For EG, AE and ED the situation becomes somewhat more complicated because of the intramolecular H-bonding present in these compounds, which leads to a situation where the two stabilizing effects (*gauche effect* and H-bonding) overlap.

Chang *et al.* [17], who developed *ab initio* global conformational potentials for ED, AE and EG, have analyzed the different types of conformation interactions in these compounds in detail. They have made one of the first attempts to separate and measure quantitatively the *gauche effect* contribution. In the proposed procedure the full conformation potential was decomposed according to the nature of three types of conformational interactions: (1) H-bonding between the functional X...Y groups, (2) the interactions between the functional groups and their geminal CH_2 groups and (3) the *gauche* interactions. Interactions of type (1) were represented by dipole-dipole, dipole-quadrupole and quadrupole-quadrupole terms, (2) by the net multipole (electrostatic) term and type (3) by the decoupled-rotor potential (T-G) [17]. The energy difference between the G and T conformers in the decoupled-rotor potential is defined as the stabilization energy due to the intrinsic *gauche effect*.

For ED the decoupled-rotor and $NH_2 \cdots CH_2$ interaction potentials were combined into one because of the interdependency of these two potentials in the fitting procedure. The energy difference for this combined potential between the G and T forms is 1.84 kJ/mol. In addition it was calculated that through intramolecular H-bonding the G form is stabilized with respect to the T form by 3.5 kJ/mol. The three most stable conformers, gGg', tGg', and gGg, are all

stabilized by intramolecular H-bonding with different multipole interaction weights. It was confirmed that the most stable conformer, gGg', is stabilized mainly by the dipole-quadrupole term (-2.0 kJ/mol), while the next two stable conformers, tGg' and gGg, are stabilized primarily by the dipole-dipole term (-3.6 and -3.2 kJ/mol, respectively).

The energy decomposition analysis of Chang *et al.* [17] shows that for EG the intrinsic *gauche effect* of the two hydroxyl groups contributes 3.5 kJ/mol to the extra stability of the G form compared with the T form. The most stable conformer (tGg') is further stabilized by the dipole-dipole interactions between the OH groups. For this conformer the dipole-dipole, dipole-quadrupole and quadrupole-quadrupole potentials of OH···OH and the net electrostatic contribution of the $OH\cdots CH_2$ interaction are -6.7, 0.92, -0.38, -2.5 kJ/mol, respectively. These results suggest that tGg' is mainly stabilized by the dipole-dipole interaction and secondarily by the $OH\cdots CH_2$ electrostatic interactions.

The decoupled-rotor potential of AE yields an intrinsic gauche energy of 1.8 kJ/mol. The angular dependence of the decomposed conformational energies shows that the most stable conformer, g'Gg', is further stabilized mainly by the $NH_2\cdots OH$ dipole-dipole interaction as part of the O-H···N H-bond. Their analysis of the interaction contributions for the three most stable conformers reveals that g'Gg' and gGg' are stabilized by a O-H···N type and gGt conformer is stabilized by a N-H···O type of H-bond. The intramolecular H-bond of the gGg' conformer is much weaker than that of g'Gg' although they both belong to the same category. The H-bonding strengths of O-H···N (g'Gg') and N-H···O (gGt) are 15.35 and 3.22 kJ/mol, respectively [17]. These results are in agreement with the energy ordering of these two types of H-bonds reported in the previous literature. The multipole interaction weights for these and some other representative stable conformers of AE can be found in Ref. [17].

To complete the study of the *gauche effect* in molecules with intramolecular H-bonds, in particular, to predict an intrinsic gauche energy of ED, Chang *et al.* [17] also calculated the energy differences, ΔE_{T-G}, between *gauche* and *trans* conformers of 1,2-dihaloethanes (XCH_2-CH_2Y, where X,Y=F,Cl,Br). These authors suggested that the intrinsic gauche energies obtained using the decoupled-rotor potentials are consistent with the conventional gauche energies for the simpler 1,2-dihaloethane systems. Their idea was based on the correlation between the gas-phase energy differences of *gauche* and *trans* conformers and the sum of electronegativities of the halogen atoms of 1,2-dihaloethanes observed by Phillips and Wray [37]. Using the intrinsic gauche energies determined for the set of 1,2-dihaloethanes along with the intrinsic gauche energy of EG (3.5 kJ/mol) and AE (1.8 kJ/mol), the relationship of these energies with various group electronegativity scales was examined by the equation

$$\Delta E_{T-G} = a + b\sum x_i, \tag{2.2}$$

where $\sum x_i$ is the sum of the two group electronegativities. Four electronegativity scales were considered by Chang *et al.* and the value of the gauche energy for ED was predicted to be -2.0 kJ/mol [17].

The final conclusion to be drawn from this conformational analysis is that the *gauche effect* and the intramolecular H-bonding are both important for a complete characterization of the unusual stability of the *gauche* conformers of EG, ED, and AE in the gas phase. The intramolecular H-bonding plays the major role in the stabilization of *gauche* conformers in comparison with *trans* conformers. The *gauche effect* further stabilizes the most stable *gauche* conformers of ED and AE but destabilizes that of ED [17]. On the basis of this conclusion and previous theoretical estimates [21], Chang *et al.* [17] made

predictions with regard to possible behavior of these molecular systems in the condensed phase. In particular, since the strength of the H-bond decreases in the order AE≅EG>ED, it would be expected that in the condensed phase this trend will be preserved i.e. AE should behave conformationally similarly, while ED would behave differently, from EG [17].

2.2 LIQUIDS AND AQUEOUS SOLUTIONS

2.2.1 Liquid methanol, methylamine and their mixtures with water

Among the variety of representatives of hydrogen-bonded liquids the smaller alcohols are of particular interest. They are economical to simulate and they have been the objects of much experimental effort. Methanol (ME) is one of the most widely studied molecular liquids [6,38-45]. The pattern of H-bonding in ME and ME-like systems seems to be very different from that found in water. For example, X-ray studies of liquid ME [38] have shown that instead of forming tetrahedral networks each monomer engages in approximately two H-bonds which is consistent with a chain-like structure similar to that found in the solid state [39].

Haughney *et al.* [40] performed MD simulation of liquid ME. They reported results of detailed studies of four different intermolecular potential models. All four were three-site models with the centers of interaction positioned on the carbon (C) and oxygen (O) nuclei and the hydroxyl proton (H_O), and the conventional representation of Columbic and Lennard-Jones terms for each site-site potential. Two of the models (J1 [41] and J2 [6]) were proposed by Jorgensen where model J2 was a revised version of J1. Models H1 and H2 were developed by the authors [40] in order to improve J1; H1 differs from J1 in the choice of charge distribution and H2 differs also in the values used for the Lennard-Jones parameters. Conventional (constant volume and energy) MD

calculations were carried out with the molecules treated as rigid objects. The adequacy of these four models was tested and results were reported for thermodynamic (internal energy and pressure) and structural (site-site distribution functions - $g_{\alpha\beta}(r)$) properties, self-diffusion coefficients and reorientational correlation times. The static characteristics of H-bond formation (i.e. the mean number of H-bonds per molecule, the mean energy of H-bonded pairs) in the four models at three different temperatures were examined as well as the dynamics of the H-bonded network (H-bond lifetimes). Two of the models H1 and J2 gave results in good agreement with a wide variety of experimental data while two other models (J1 and H2) were substantially less realistic. The results for models H1 and J2 suggested that at room temperature the dominant structural units in liquid ME are chains consisting of fewer than ten monomers and the lifetime of individual H-bonds is of order 1 to 2 ps, which is several times larger than the corresponding time for liquid water [40].

Some years later Svishchev and Kusalik [42] reported results from MD simulations of liquid ME at room temperature performed with the H1 model of Haughney *et al.* They determined spatial distribution functions (SDF) for the oxygen-oxygen ($g_{OO}(r,\Omega)$) and oxygen-carbon ($g_{OC}(r,\Omega)$) correlations and have used these to elaborate further the equilibrium structure in this liquid. These authors showed that a structural analysis based on SDF's (rather than on radial distribution functions) greatly improves our understanding of the local structure in molecular liquids. The integration of the first peak in $g_{OO}(r,\Omega)$due to nearest neighbors yields an average coordination number of 1.92, which is in a very good agreement with the previous results from chain length analysis [40]. The examination of SDF's has also provided insights into the packing pattern of chains in the bulk liquid. It appears that H-bonded chains in liquid ME prefer a local parallel arrangement, similar to that found in the solid state. The authors

suggested that O···CH_3 interactions are important when considering H-bonded chains in liquid ME [42].

Methanol, as the simplest (smallest) alcohol, can also serve as a suitable model molecule for studies of structural aspects of solvation in aqueous mixtures. In dilute aqueous solutions the CH_3 group is thought to reinforce the water structure while the OH group makes the whole molecule soluble in water [43]. In accordance with previous theoretical suggestions ME would simply substitute water in its H-bonding network in dilute solutions. Jorgensen and Madura [44] carried out Monte Carlo (MC) simulations for a dilute solution of ME in water at room temperature. The water-ME interaction potential (TIPS) employed, which explicitly incorporated methyl hydrogens, allowed the internal rotation of the methyl group to be included in the calculations. It was shown that in aqueous solutions ME prefers to be staggered (*gauche*, ϕ =60°) rather than eclipsed (ϕ =0°) just as it does in gas phase. Structural results obtained from a detailed analysis of the solute-solvent radial distribution functions confirmed that the water molecules form a cage around the methyl group, while two or three water molecules are H-bonded to the hydroxyl end of the methanol. These findings are in reasonable agreement with the more recent neutron diffraction studies by Soper and Finney [45] which suggest that water molecules form a distorted cage around the methanol molecule. According to Soper and Finney [45], the cage around the methyl group consist of approximately 20 water molecules and is believed to be comparable to the water clathrate structure around a methane molecule. Laaksonen *et al.* [43] performed MD simulations at room temperature for five water-ME liquid mixtures scanning the entire range of compositions. Both oxygen-oxygen radial distribution functions (RDF) and oxygen-oxygen spatial distribution functions were employed in their work to estimate the coordination numbers (around oxygen atoms) and to analyze the three-dimensional local

structure in water-ME solutions. It was shown from the RDF analysis that water has higher H-bond coordination numbers than ME; the total coordination around water is consistently about one molecule larger than for methanol, over the entire composition range. In the water-rich mixtures the total coordination around a water molecule drops from 4.5 (in pure water) to 4.0. For water in methanol-rich solutions and for methanol in water-rich mixtures their numbers of H-bonded neighbors approach 3. However, these authors [43] demonstrated that RDF's provide an incomplete and sometimes even misleading picture with respect to the local order in these liquid mixtures. Detailed analysis based on oxygen-oxygen SDF's in water-rich solutions confirmed a high degree of ordering characterized by a strong preference of tetrahedral arrangements with water molecules most highly localized around the hydroxyl group of the methanol solute. In methanol-rich solutions the water-water correlations are evident even at longer range while those for ME resemble its pure liquid structure [43].

Usually the study of small functionalized organic molecules is an essential first step in modeling of large biomolecular systems in their natural aqueous environment. For instance, N-H···O and NH^+···O H-bonds are important factors determining conformational and crystals structures of biomolecules. Understanding of the hydration of a small molecule such as methylamine could provide insights into the hydration structure of larger molecules containing amino groups. Previously reported investigations of aqueous amine solutions include MC simulation studies of low concentration water-methylamine mixtures [46-48], a hybrid RISM-SCF study of methylamines examining their basicity anomaly in aqueous solution [49], several simulation studies focused upon solvation free energies of simple amines [50-52], and investigations of the solvation structure of methylamine in aqueous solutions [53,54]. Kusalik *et al.* [54] carried out MD simulations of liquid methylamine and 10 and 30 wt% aqueous solutions of

methylamine. A three-site simple point charge model (SPC/E) [55] was used for the water interaction potentials and the four-site model of Impey *et al.* [53] with the methyl group represented as a united-CH_3 atom was used for methylamine. Radial distribution and spatial distribution functions were employed in their structural analysis. In pure liquid methylamine the nitrogen-nitrogen RDF revealed roughly four nearest neighboring nitrogen atoms while the nitrogen-hydrogen RDF showed that each methylamine molecule makes only two strong H-bonds. The SDF's provided the necessary information required to determine that the remaining nitrogen coordination is due to either weak or non-H-bonding first neighbors or closely lying second neighbors. SDF's also clearly showed that the methylamine molecules prefer arrangements with a few amino groups clustered together in small compact ring or ring-like structures unlike the chain-like arrangements found in liquid methanol [43]. The authors found that the information provided by RDF's is not sufficient to resolve entirely the complex local structure in methylamine aqueous solutions. Furthermore, aspects of the determined hydration structure were rather unexpected. Two primary types of H-bonded water molecules were found. Two molecules are involved in making strong H-bonds, one as an accepting neighbor and one as a donating neighbor. The second H-bond forming site of the amino group is not vacant but rather is occupied by a more distant nearest-neighbor water that forms only a weak H-bond to the amino hydrogen [54]. The water molecules around the methyl group tend to lay flat against its hydrophobic surface. In addition there were no long range correlations of methylamine molecules observed as was the case in dilute methanol solution [43]. These authors remarked that it may be particularly interesting to observe the local structures in aqueous systems with both alcohol and amino groups present.

2.2.2 Liquid phase 1,2-disubstituted ethanes

A limited number of theoretical studies of liquid 1,2-disubstituted ethanes and their aqueous solutions have been published within the last decade [4,6,11,21,29,56-57,59,63-67], while EG was investigated the most extensively. It is also somewhat surprising that no computational studies of liquid ED or its mixtures with water were found in the available literature.

Padro, Saiz and Guardia [11] performed a series of MD simulations to investigate H-bonding in liquid methanol, ethanol, ethylene glycol and glycerol at 298K. H-bonding statistics as well as the mean lifetime of the H-bonds were analyzed and results were compared with those obtained for liquid water. It was shown that the H-bonded structure within these systems differs markedly. Methanol and ethanol are characterized by H-bonding patterns that may be described as linear winding chains while EG and glycerol appear to show three-dimensional H-bonding patterns, which are not well characterized [11]. The same group of authors [4] later examined the structure of liquid EG at room temperature by carrying out MD simulation studies for four different force fields, namely, J [6], JMOD [4], WP [56] and HTN [57]. The J model is a generalization of the OPLS model proposed by Jorgensen [6] for simulations of liquid alcohols. In WP and JMOD models several aspects (e.g. the torsion potential) of the OPLS model were modified. Models J, JMOD and WP, all being OPLS-based models, have identical parameters for the intermolecular interactions [6]. For the HTN model an entirely new force field was developed by Hayashi *et al.* [57], where the EG molecule consisted of six interaction sites (i.e. was modeled using an united atom approximation) and bond lengths and bond angles were kept fixed during the simulations. The molecular shape of EG was studied through an analysis of the distribution of dihedral angles for motions about the C-C and C-O bonds. For example with respect to C-C bond, it was found that for the WP model the

percentage of *gauche* conformers (45.2%) is slightly lower than the percentage of *trans* conformers. In the case of the J model, most of the molecules adopt a *gauche* conformation and only 5% of the molecules remain in the *trans* form. For the HTN model the percentage of *trans* conformers (58%) is similar to that found for the WP model; eclipsed and *gauche* forms equally share the rest of the population. The *gauche* maxima are located at about $\pm 60^{o}$ for the J model, at $\pm 75^{o}$ for the WP model, and at $\pm 80^{o}$ and $\pm 50^{o}$ for the JMOD and HTN models, respectively. A description of the local structure in the liquid was obtained on the basis of RDF's. It was found that the structure is dominated by a three-dimensional network of H-bonded molecules with a mean number of H-bonds per molecule slightly lower than four [11]. Interestingly, the structural results obtained were similar for each model and appeared to be independent of the mean conformation of the molecule. On the contrary, the data obtained for the dynamics indicated that the four models give quite different results. The authors found good agreement for the self-diffusion coefficients between the OPLS-based models (7.4×10^{-7} and 11.7×10^{-7} cm^2/s for J and JMOD, respectively) and the experimental value of 9.0×10^{-7} cm^2/s[58].

In the force field proposed by Hayashi *et al.* [57] to simulate EG in the liquid state a newly developed torsional potential was combined with the standard OPLS intramolecular potential [6] for non-bonded interactions. The simulations were carried out on 25 mol% mixtures of EG with water and EG with Xe in order to examine the solvent effect. Comparison of the results for the aqueous solution of EG with Xe solution revealed that there is a preference for the *trans* form (up to 97%) in the aqueous solution in spite of the occurrence of some *gauche* conformations possessing intramolecular H-bonds [57]. These results contradict the findings reported by other authors [21,29,59] possibly

because in this work two interaction potentials of different origin were combined together without appropriate re-parameterization.

Aqueous solvation effects on the EG molecule were studied theoretically with explicit solvent inclusion by Nagy *et al.* [29] and Hooft *et al.* [59]. The former authors performed MC simulations using the OPLS force field [6] and the TIP4P water model [60]. Changes in free energies were monitored while individual dihedral angles were driven from one conformer to another and four conformers (tTt, tGg', gGg', g'Gg') were considered throughout. The latter group employed MD methods using the GROMOS force field [61] and the SPC [62] water model. In their work all conformers (with only bond distances frozen) were sampled during the simulation run assuming that the simulation time was long enough for this to occur in practice. Both studies found that the tTt rotamer is better solvated than tGg' with a free energy difference of 5 kJ/mol. Considering that there is only one *trans* conformer (tTt), Nagy *et al.* [29] predicted a 99:1 gauche/trans ratio for the rotation about the C-C bond. In contrast, Hooft *et al.* [59] considering all possible conformers, predicted about 67:33 gauche/trans ratio for the same equilibrium in aqueous solution. Most of the difference in these estimates is due to the gas-phase energy values to which the differential solvation free energies were added; Nagy *et al.* [29] used gas-phase free energies calculated at MP2/6-31G*//HF/6-31G* level, while Hooft *et al.* [59] obtained their gas-phase energies using the GROMOS force field. Hooft *et al.* [59] also concluded that the gas-phase preference for internally H-bonded solute molecules disappears completely in dilute aqueous solution.

An alternative method, in which the solvent was represented by a continuum dielectric, was used by Alagona and Ghio [63] and Varnek *et al.* [64] to analyze the effects of aqueous solvation on EG. The polarized continuum model of Tomasi and co-workers [65] was used to this end. Alagona and Ghio

[63] considered differences only in the electrostatic portions of the solvation free energy at the HF/4-31G and HF/6-31G* levels for the four most stable conformers. At both of these levels they found that the tTt rotamer is better solvated than the tGg' rotamer by only 0.83 kJ/mol, but gGg' and g'Gg' are better solvated than tGg' by 2.5-4.6 kJ/mol. Varnek *et al.* [64] taking into account both the solvation and the intrinsic energy term reported the order of relative stabilities among *cis*, *gauche* and *trans* conformers of EG in water. They found that the most stable form is *gauche* (with OCCO dihedral angle of 54º) < *cis* (ΔE=18.8 kJ/mol) < *trans* (ΔE=45.2 kJ/mol) [64]. Their prediction of the *gauche* form being the most stable is consistent with the results of Nagy et al. [29] with explicit water included in the simulations.

In view of the large portion assigned to the *trans* population by Hooft *et al.* [59], Cramer and Truhlar [21] questioned whether this dramatic predicted solvation effect is valid. They presented a third approach for analyzing the conformational equilibrium in aqueous solutions of EG, employing quantum statistical solvation models, in particular, the Austin Model 1 (AM1-SM1a, AM1-SM2) and Parameterized Model 3 (PM3-SM3) solvation models, which are discussed in detail in Ref. [21]. These models divide the free energy of solvation into two parts. The first accounts for the mutual polarization of solute and solvent (like in a continuum model) and the second includes local effects specific to the first solvation shell (like in an explicit solvent simulation). Their results for aqueous solution together with gas-phase calculations suggest that on passing from the gas phase into aqueous solution there is a relative increase in the population of *trans* conformers by a factor of 3-5 (or to about 6-10%). For instance, for the relaxed SM3 model the gauche/trans ratio is 92:8 vs 98:2 in the case of gas phase. This difference comes from the increase in the population of the *trans* conformers tTg and gTg' at the expense of conformer g'Gg' [21].

Overall these estimates for the gauche/trans ratio lie somewhere in between those obtained by Nagy *et al.* [29] and Hooft *et al.* [59]; unfortunately a more precise comparison with results from explicit water simulations is difficult.

The number of theoretical studies of liquid AE is small and presently none have focussed on its thermodynamics and structural behavior in aqueous solutions. In 1996 Button *et al.* [66] used MD simulations to study hydrogen bonding in liquid AE. The authors employed a force field which is a combination of those reported for alcohols [6] and amines [66]. The results were not compared with experimental data and the only major conclusion of Button *et al.* [66] was that "different behavior was observed for the alcohol and amine functional groups". Later Alejandre *et al.* [67] developed a special force field for AE as a precursor of a more general modeling potential for simulations of alkanolamines and their aqueous solutions. The geometry of AE, the bending and torsional energies, and the charges were first obtained by means of *ab initio* calculations. The charges obtained give a dipole moment of 3.044 D, which is in excellent agreement with the experimental value of 3.05 D [67]. The torsion angle potential chosen reproduced the observed experimental behavior for the most stable conformation [21] of AE (with 180°, 80° and 50° for NHCC, NCCO, CCOH torsional angles, respectively) quite well. Canonical MD simulations were carried out for the liquid phase and for the liquid-vapor equilibrium state. To explore the formation of H-bonds the simulations were performed at six temperatures between 298 and 530 K. From RDF analysis it was found that the H-bond is favored when a hydrogen of the hydroxyl group is at a distance of ~1.8 Å from an oxygen or nitrogen of another molecule. The hydrogen of the amino group does not favor H-bonding with the oxygen or the nitrogen from neighboring molecules. The proposed force field satisfactorily reproduces the experimental liquid and vapor densities and the surface tension.

CHAPTER III - ANALYSIS OF STRUCTURAL CHARACTERISTICS

In this chapter emphasis is placed on the structure specific to 1,2-disubstituted ethanes, as obtained from experimental and theoretical investigations. All three compounds of interest have been extensively studied in the gas phase using gas electron diffraction and microwave spectroscopy, while the elucidation of their structures in the liquid and in the solid state have been mainly performed using spectroscopic (Raman and IR) methods. Unfortunately, spectroscopic methods do not provide a precise picture of the local structural arrangements of molecules within the liquid. However, one might reasonably expect the local structural pattern in the liquid phase to resemble somewhat that found in the solid state, and hence a separate section will be devoted to results obtained on solid samples using X-ray diffraction. A summary of the most abundant conformations in the gas, liquid and solid state is given at the end of this section. Conventional theoretical approaches used in the determination of the structure of molecular liquids are primarily based on analysis of radial distribution functions. The relationship between radial distribution functions and structure factors obtained from scattering experiments, as well as the extraction of pair distribution functions from experimental data using such numerical techniques as spherical harmonic reconstruction, reverse Monte Carlo and empirical potential structure refinement will be discussed in detail. Finally a novel structural probe, namely the spatial distribution function will be presented. The function can be obtained computationally and then visualized to give a more complete representation of the average local structure within a liquid system and particularly its hydrogen-bonding pattern.

3.1 EXPERIMENTAL STUDIES OF 1,2-DISUBSTITUTED ETHANES

Information on the structures, relative energies and the potential barriers between the separate rotamers of ED, AE and EG in the gas, liquid and solid states has been obtained by a wide variety of experimental methods [2,22-23,36,68-70,74,79]. Structural properties of the ED molecule have been studied most extensively in the gas phase. Yokozeki and Kuchitsu [69,70] investigated the structure and the rotational isomerism of ED by means of gas electron diffraction. One of the characteristic aspects revealed by this study was that this molecule has a large-amplitude torsion about the *cis* conformation. Preliminary results [69] suggested a dominant *gauche* conformer. Further analysis took into consideration the hydrogen positions of the amino groups. The structural parameters were estimated from an analysis of the molecular intensity on the assumption that the central dihedral angle was 60°. The presence of one (*gauche*) conformer with the NCCN angle of 64±4° measured from the *cis* position was revealed in the vapor phase (at 50-120° C) [70]. For a quantitative examination of the fractions of the *gauche* conformer and other possible candidates, an analysis of the scattering intensity observed in the small-angle region was undertaken [70]. The *gauche* conformer was found to be by far the most stable in the gas phase and the fraction of any other rotamer, if present, was estimated to be less than 5% (conformational notation is shown on Fig. 1). The microwave spectrum of ED has been investigated in the 12.4-39.5 GHz spectral region by Marstokk and Mollendal [36]. The two NCCN *gauche* conformations denoted I (tGg') and II (gGg') were identified (see Fig. 2). Both rotamers have NCCN angles of 63±2° and posses intramolecular H-bonds with the two amino groups acting, respectively, as proton donor and acceptor within each conformation. The roles of the amino group of each conformation may be interchanged by appropriate rotation about both C-N bonds. The existence of two

identical forms of each of the conformations I and II leads to separate double minima in the potential for the appropriate transformation motion. The tunneling frequencies of the transformation motion for each conformation were identified and a moderate barrier height in the 17-25 kJ/mol range was estimated for each conformation. It was found [36] that the gas phase is composed roughly of 2/3 of conformation I and 1/3 of II where conformation I is favored by 1.26±0.8 kJ/mol relative to conformation II. However, it appears that the *gauche* form of the ED molecule is not preferred in the liquid and crystalline state [2].

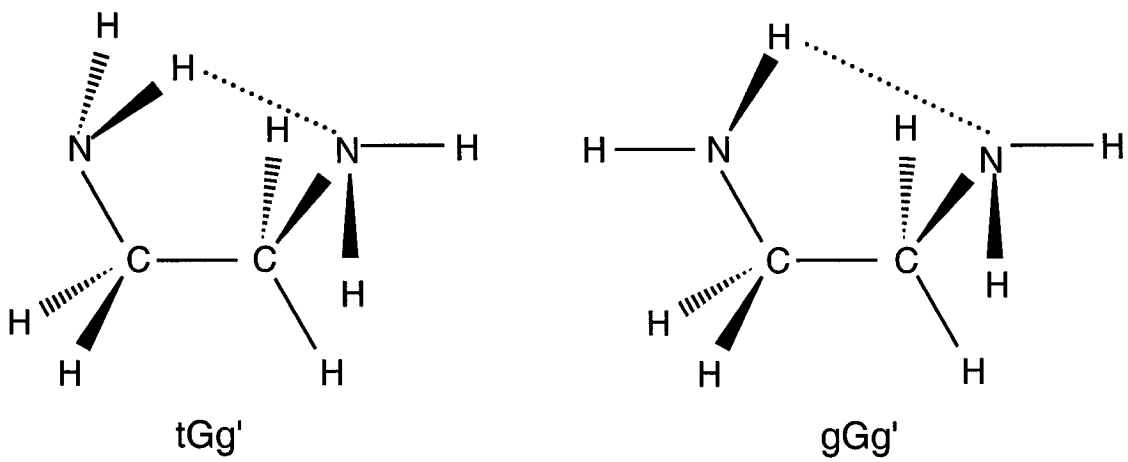


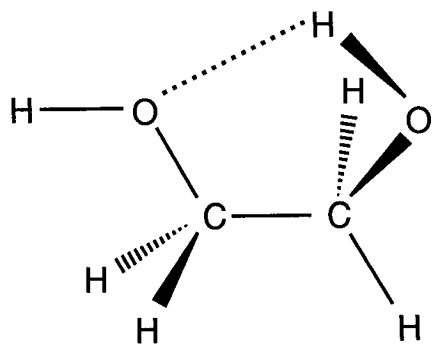
Figure 2. The most stable conformers of ED.

Batista de Carvalho *et al.* [2] performed conformational analysis of ED using Raman spectroscopy of the pure liquid, the aqueous solution and the solid. A complete assignment of the vibrational spectra was made for both the non-deuterated and N-deuterated molecules. In the liquid phase it was possible to assign the Raman band observed at 473 cm^{-1} to the tTt conformer and the bands at 341 and 512 cm^{-1} to the conformations displaying a "skeletal *gauche* arrangement" [2]. In the solid sample the band at 470 cm^{-1} confirmed the

existence of only the tTt conformer. In aqueous solution this band was also clearly detected while the bands corresponding to *gauche* arrangements seemed to disappear. The authors concluded that for the pure liquid, the aqueous solution and the solid sample the conformational preferences of ED are determined by the relative importance of intra- versus intermolecular H-bonds [2]. Further evidence that in the solid phase the molecule adopts the NCCN *trans* conformation is available from X-ray data [71] and infrared and Raman spectra [72]. The infrared and Raman spectra of crystalline ED have been obtained at 200 and 78 K and interpreted by Righni and Califano [72] on the basis of the earlier measured crystal structure [71]. In an earlier investigation of the solid ED [73] the existence of two crystalline modifications (hereafter forms I and II) with a phase transition temperature at about 203 K was discovered. Since only the crystal structure of the high temperature phase (form I) is known [71] it was possible to perform a meaningful calculation of the vibrational spectrum only for this phase [72]. It was found that form I is a monoclinic system, space group C_{2h}^5, with two molecules in the unit cell and the ED molecule in the crystal has a *trans*-planar conformation of C_{2h} symmetry. On the basis of spectroscopic observations the authors also concluded that the phase transition between forms I and II involves a conformational change in the molecule such that the center of symmetry is preserved. The only possible conformational change is then a rotation of the NH_2 groups around C-N bonds in opposite directions. Thus in form II the NCCN backbone remains *trans*-planar whereas the NH_2 groups assume a *gauche* conformation with respect to the CH_2 groups [72].

The conformation and structure of EG have been the subject of several experimental investigations [34,74-79]. The structure of EG in the gas phase has interested chemists for years: Bastiansen [34] reported an electron diffraction study of this compound in 1949. Only the heavy-atom *gauche* conformation was

detected in this work and its stability was ascribed to intramolecular H-bonding [34]. In 1981 Caminati and Corbelli [74] measured the microwave spectrum of the two possible O-monodeuterated species so that the four possible H/D hydroxyl species were present with a 25% concentration each. It had been noted previously that the microwave spectrum of EG is complicated by the tunneling motion between the double minima in the potential due to the simultaneous rotation of the two hydroxyl hydrogens. The isotopic substitution combined with an application of an ordinary semi-rigid rotor model, which excludes tunneling effects [74], allowed these authors to reach a definitive conclusion about EG's conformation. Again only one conformer was detected: one oxygen is *gauche* with respect to the other; one hydroxyl hydrogen participates in the H-bond while the second one is *trans* with respect to C-C bond [74] (tGg', see Fig. 3). The microwave spectrum of EG was assigned for the first time using microwave-microwave double resonance and molecular beam Fourier transform microwave spectroscopies [75]. These techniques were explored because they select the



tGg'

Figure 3. The lowest energy conformer of EG.

desired lines in the crowded spectrum, which does not follow a standard rigid rotor pattern because of the large-amplitude motion displayed by the molecule. As a result of this analysis the existence of the tGg' conformer of EG was again verified [75]. However, Kristiansen *et al.* [76] analyzing the microwave spectrum of an asymmetrical isotopically substituted species (e.g. $HOCH_2CD_2OH$) have found two possible heavy-atom *gauche* conformations possessing internal H-bonds: tGg' and gGg'. It was shown that the gGg' conformation (with a OCCO angle of 53.9 ± 6^o) is 1.4±0.4 kJ/mol less stable than tGg' (with a OCCO angle of 53.4 ± 6^o). The fact that the OCCO dihedral angles of both conformations are approximately 6^o less than a completely staggered arrangement indicates a strong interaction between the two hydroxyl groups in this molecule. It was suggested that in EG the internal H-bond alone cannot account for this stability and that a *gauche effect* [32] must augment the H-bonding and be of considerable magnitude [76].

The equilibrium configurations of EG in liquid phase have been studied by Crupi and co-authors using Fourier transform IR spectroscopy [77,78]. The extracted activation and binding energies were attributed to the transition from a *gauche* configuration stabilized by an intramolecular H-bond to a simple *gauche* configuration. It was shown that the existence of an intramolecular H-bond makes the corresponding *gauche* conformation the most energetically favored [77]. Finally, the equilibrium configuration of EG in the vapor, liquid and solid phases was investigated by IR spectroscopy of the two isotopically substituted molecules $(CD_2OH)_2$ and $(CH_2OD)_2$ by Buckley and Giguere [79]. The assignments of most of the fundamental frequencies, the phase and temperature behavior of the spectra confirmed that in all cases EG exists only as *gauche* conformers, which are more stable than the *trans* as a result of intramolecular H-bonds. Interestingly, the positions of OH stretching and especially bending

vibrations, which are strongly shifted towards higher frequencies, indicates the presence of highly bent (i.e. weak) H-bonds. However, the EG molecule retains the *gauche* conformation up to fairly high temperatures (398 K) [79] and is interpreted by some as evidence that an effect in addition to the internal H-bonding must be operative [76]. In addition to the analysis of hydrogen bonds in EG, some molecular energetics were also examined by Buckley and Giguere [79]. For instance, the average rotational barrier heights were estimated to be of the order of 42 kJ/mol for the rotation about the C-C bond, and 12.5-16.5 kJ/mol for the rotation about C-O bonds. These results seem to be markedly overestimated in comparison with theoretical predictions for the analogous parameters, in particular, the barrier heights for the rotation about the C-C bond in EG were found to be 28.7 and 23.43 kJ/mol by Burgess *et al.* [30] and Chang *et al.* [17], respectively.

In contrast to ED and EG, for AE only a few sources of experimental structural information are available. Kreuger and Mettee have carried out an infrared spectroscopic study [22] and showed that in dilute tetrachloroethylene solutions AE exists in three forms. The most stable one (g'Gg', see Fig. 4) is a

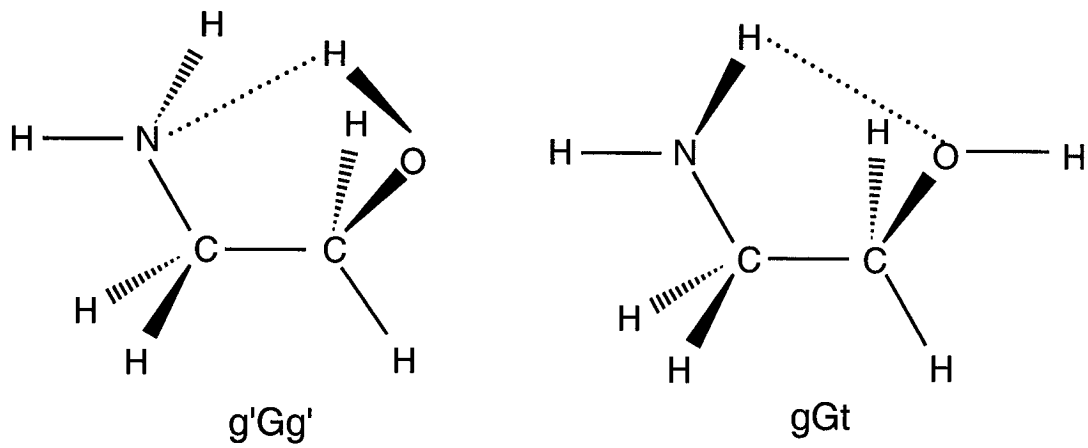


Figure 4. The most stable conformers of AE.

H-bonded form in which the hydroxyl group is the proton donor and the amino group is the proton acceptor (O-H···N). A *trans* (or non-H-bonded form) is less stable than the O-H···N form by 2.93 kJ/mol. There is another H-bonded form (gGt, Fig. 4) in which the amino group is the proton donor and the hydroxyl group is the acceptor (N-H···O). This form is 5.86 kJ/mol less stable than the O-H···N form. The microwave spectra of six singly substituted isotopic species of AE have been measured and assigned by Penn and Curl [23]. They found that the molecule exists in a *gauche* form with an intramolecular H-bond of O-H···N type. The internal rotation dihedral angles assume values that minimize the distance between the hydroxyl proton and the lone-pair electrons of the nitrogen. This corresponds to the O-H bond halfway between the eclipsed and staggered conformation with respect to the C-C bond [23]. The results of an infrared study of AE isolated in matrices at low temperature [80] demonstrate the presence of at least five different conformers in the samples: the most stable O-H···N intramolecularly H-bonded g'Gg' form and the conformers gGt, tGt, tGg and gGg. The latter four conformers possess an N-H···O intramolecular H-bond and have energies within 10 kJ/mol of the most stable conformer. The molecular structures, relative energies, dipole moments and vibrational spectra (infrared and Raman) of the possible conformers of AE in liquid phase were reported by Silva *et al.* [27]. Their data are consistent with the absence of the intramolecularly H-bonded g'Gg' (or gGg') form and the presence of gGt and tGt conformers as the most abundant species in the liquid phase. It was also pointed out that no evidence of *trans* conformers (about the NCCO axis) was found in the liquid phase spectra [27].

A summary of experimental structural data used as the input for MD simulations performed for ED, EG and AE in the present work (see Chapter IV) is given in Table 1.

Table 1. Structural parameters for ED, EG and AE used in the present work.

	AE [23]	ED [36]	EG [74]	EG [79]
Bonds (Å)				
CC	1.526	1.546	1.526	1.54
CO	1.396	-	1.423	1.43
CN	1.475	1.469	-	-
OH	1.139	-	0.956	1.00
NH	1.017	1.017	-	-
CH	-	1.093	1.090	1.08
Angles (deg.)				
CCO	112.1	-	109.5	110.0
CCN	108.1	111.50	-	-
CCH	-	109.48	109.0	-
CNH	110.4	109.48	-	-
	111.3			
COH	103.7	-	108.0	105.0
HNH	109.86	109.48	-	-
HCH	-	109.48	-	-
Dihedrals (deg.)				
OCCN	55.4	-	-	-
NCCN	-	63.2		
OCCO	-	-	60.0	74.0
CCOH	28.3	-	-28.0	-
			198.0	
CCNH	-78.2			
	159.5			

3.2 SOLID STATE STRUCTURAL DETERMINATION FOR ETHYLENE GLYCOL ETHYLENEDIAMINE AND 2-AMINOETHANOL

As has been shown in the previous section, the structural information available for the compounds of interest in liquid state is rather limited and it becomes especially noticeable where the secondary (or intermolecular) structural

arrangements are concerned. In this case the structural details obtained from the X-ray scattering experiments may serve as important complementary material in the elucidation of liquid state structure.

A structural study of the alkanediols HO-$(CH_2)_n$-OH (n=2-10) and alkanediamines H_2N-$(CH_2)_n$-NH_2 (n=2-8), referred to as diols and diamines hereafter, was undertaken by Thalladi *et al.* [81] to investigate interference between hydrogen bonding and hydrophobic interactions. The X-ray data were collected at the same temperature (130 K) for all the compounds to allow a comparison of calculated densities. The density gives a measure of compactness in the packing and it was shown that the odd-numbered members are more poorly packed than the even-numbered members in both homologous series. The even diols with $n \geq 4$ form layer-like networks (space groups $P2_1/n$ and $P2_1/c$) and the odd members form three-dimensional structures ($P2_12_12_1$ space group) [81].

The end groups in the diols and diamines form H-bonds and the alkyl chains participate in hydrophobic interactions. The common packing pattern in both even and odd n-alkanes is a column-like structure with the molecules stacked such that CH_2 groups of successive molecules are intergrooved (such a close packing of CH_2 groups occurs at an intermolecular separation of 4.7 Å [82]).

Since both OH and NH_2 groups act as H-bond donors and acceptors simultaneously, they form several types of structural patterns, the most common of which is a chain pattern [83]. For the diols formation of O-H···O chains at both ends of the molecule can lead to a layered structure consisting of molecules in columns, where successive molecules in a column are separated by 5.1 Å, a distance that is too long for the hydrophobic packing observed in n-alkanes. The

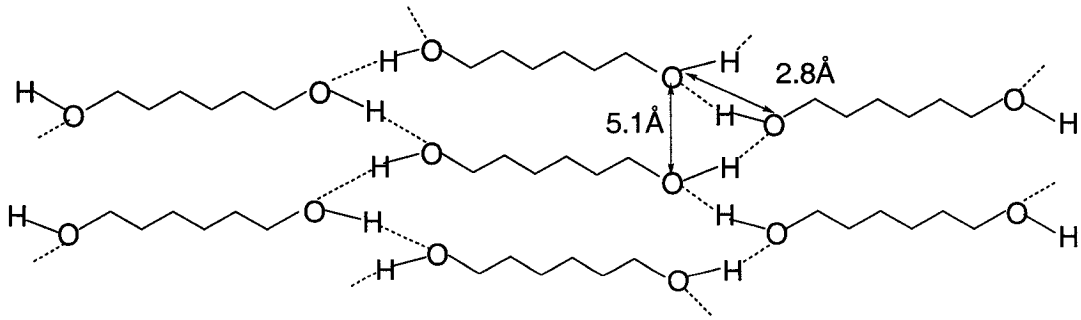


Figure 5. Hypothetical layered structure formed by O-H···O hydrogen-bonded chains in even diols [81].

corresponding anticipated structure is given in Fig. 5. However, it was observed that the layer structure of the even diols differs slightly from that depicted in Fig. 5. The molecules are offset along their length in such a fashion that the α-CH_2 groups fit into supramolecular grooves, and adjacent columns are inclined to preserve the O-H···O chains (see Fig 6a). This arrangement becomes possible because of the all-*trans* conformation observed in even diols, which leads to an antiparallel projection of the two C-O groups. In odd diols these groups are projected in the same direction precluding the formation of O-H···O chains at the opposite ends of the molecule (Fig. 6b). In addition in the odd diols one of the C-O groups adopts a *gauche* conformation, leading to the formation of a three-dimensional network with the packing of the hydrophobic moiety not as effective as in the even diols [81].

The layered structure in even diamines is similar to that found in even diols except that the N···N separation (3.2 Å) is generally longer than O···O separation (2.8 Å). Adjacent molecules are now spaced at 5.2 Å and the hydrocarbon chain therefore swivels around the main molecular axis by 45º to fill the corresponding gaps. Odd diamines, unlike odd diols, form a layer structure with N-H···N chains

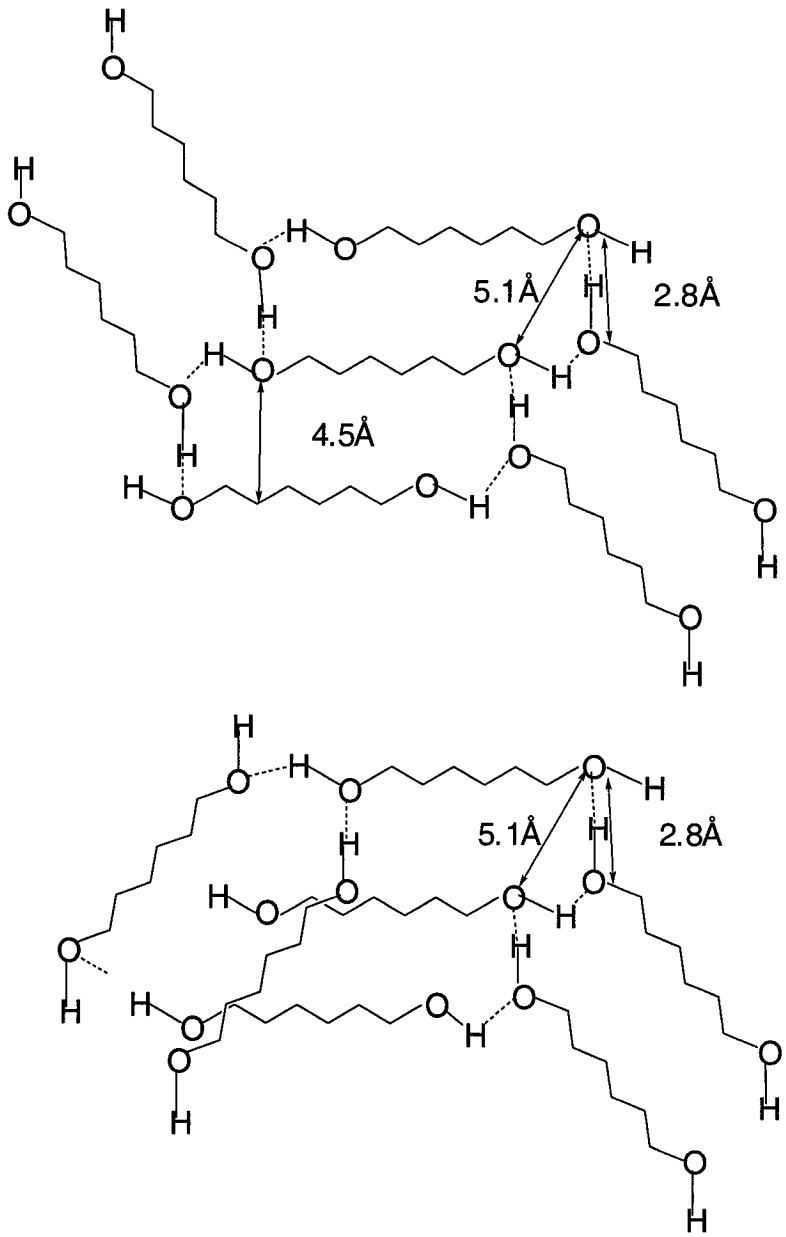


Figure 6. (a) Offset of even diol molecules within a column due to the intergrooving pattern of hydrophobic interactions. (b) Geometrical hindrance of molecular offset on odd diols; the H-bonded chains are formed only at one of the ends [81].

but without the offset of the molecules within a column (adjacent molecules are now separated by 5.7 Å). This structure is analogous to the hypothetical layer shown in Fig. 5. In spite of the swiveling of the hydrocarbon chain the odd diamines appear to be less densely packed than the even ones. The interlayer packing in even and odd is similar with the second NH group forming a longer N-H···N hydrogen bond between the layers [81].

Unfortunately, the only basic structural information relevant to EG and ED was reported in Ref. 81 (summarized in Table A1). Thalladi *et al.* [81] only briefly mentioned that "the packing patterns in 1,2-ethanediol (EG) and 1,2-ethanediamine (ED) are different from those described in this work because in these lower members the H-bonds overrule the hydrophobic interactions". Two years earlier Boese and Weiss [84], who are the co-authors of Ref. 81, published a paper devoted exclusively to structural elucidation of EG. The crystallization was performed on the diffractometer at a temperature of 245 K with a miniature zone melting procedure using a focused IR-laser beam. They reported that in the solid state EG adopts a *gauche* conformation with the torsion OCCO angle of -64.4°. The packing consist of a "catamer-like" [85] two dimensional network with four hydrogen bonds per molecule parallel to the [100] plane (O···O distances are 2.711 and 2.730 Å, and O-H···O angles are 176 and 171°) [84]. Comparison of these results with the general trend observed for higher members of the diol series leads to the conclusion that EG rather resembles the structural specificity of the odd representatives of this series. In particular, it has a similar space group ($P2_12_12_1$) and conformation (*gauche*), but differs in the type of packing pattern, forming 2-D chains instead of a 3-D network.

The crystal structure of pure ED was determined by Jamet-Delcroix [71] in 1973 from three-dimensional single-crystal X-ray data obtained at 213 K. The

basic structural parameters agree well with those reported by Thalladi *et al.* [81] including the fact that ED (an even diamine) adopts the *trans* conformation. It is also in agreement with the findings presented for a series of diamines [84], where it was reported that the NH_2 groups of ED are responsible for the formation of two H-bonded chains and the formation of molecular layers parallel to the bc (XZ) plane [71].

The crystal structure of 2-aminoethanol (AE) was determined by Mootz *et al.* [86] by low-temperature single-crystal X-ray diffraction. The measurements were performed at a temperature of 263 K on a Syntex $P2_1$ diffractometer. A miniature zone-melting technique using focused heat radiation was applied. The authors found that the crystal structure of AE is governed by the geometry of the molecule (see Table A1) and extensive intermolecular H-bonding. Molecules in crystalline AE are linked by strong (short-2.05 Å) N-H···O hydrogen bonds into angular chains parallel to the longer diagonal of the ac plane (direction $[10\bar{1}]$). Markedly weaker N-H···O and O-H···N hydrogen bonds (2.23 and 2.36 Å, respectively) were observed between these chains [86]. This structural arrangement leads to the formation of a three-dimensional network with molecules hydrogen-bonded to six nearest neighbors (see Fig. 7). It has to be noted that in contrast to the hydroxyl oxygen of EG, which acts as a single donor and a single acceptor of H-bonds, the oxygen of AE appears to be a single donor but a double acceptor forming simultaneously one strong and two weak H-bonds.

As a brief conclusion on experimental results presented and discussed in Sections 3.1 and 3.2, Table 2 summarizes the experimental findings related with the conformations of ED, EG and AE.

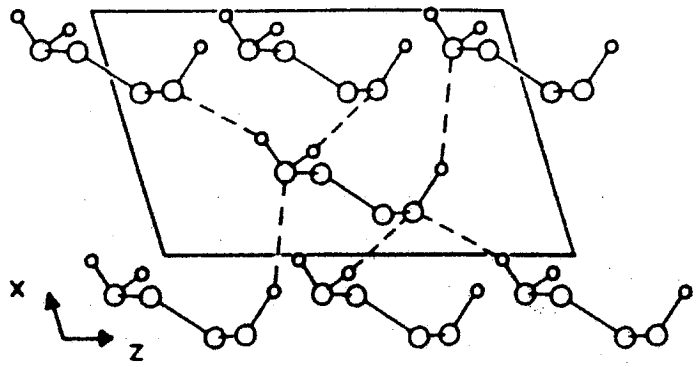


Figure 7. Projection of the structure of AE. All atoms are of arbitrary size and the H atoms of the CH_2 groups are not shown for clarity [86].

Table 2. The most abundant conformations with respect to the central dihedral angle for ED, EG, AE in the gas, liquid and solid state.

	ED	EG	AE
Gas	*gauche* [36,69,70]	*gauche* [34,74-76]	*gauche* [22,23,80]
Liquid	*trans* [2]	*gauche* [77,78]	*gauche* [27]
Solid	*trans* [71-73]	*gauche* [79]	*trans* [86]

3.3 STRUCTURE OF MOLECULAR LIQUIDS: FORMAL DESCRIPTIONS

3.3.1 Radial distribution function (RDF)

The structure in molecular liquids can be completely described by its N-body distribution function:

$$g^{(N)}(\vec{r}_1,\ldots,\vec{r}_N;\vec{\omega}_1,\ldots,\vec{\omega}_N) \equiv g^{(N)}(\vec{r}^N,\vec{\omega}^N), \tag{3.1}$$

which is a function of the positions, $\vec{r}$, and orientations, $\vec{\omega}$, of all the molecules in the system [87]. This function represents the relative probability density of finding N molecules of the system in any particular configuration. Unfortunately, the complexity of such a function makes it virtually useless in understanding the structure within a liquid. The most important distribution function in liquid state theory is the pair distribution function, $g^{(2)}(\vec{r}^2,\vec{\omega}^2)$, which represents the probability of finding any two molecules at any two points in space with any two orientations. For a molecular liquid (i.e. an isotropic homogeneous fluid) this function becomes invariant to the choice of the position and orientation of the first molecule. This $g^{(2)}(\vec{r}^2,\vec{\omega}^2)$ can be expressed as $g(\vec{r}_{12},\vec{\omega}_2)$, where $\vec{r}_{12}$ (the separation vector between molecules 1 and 2) and $\vec{\omega}_2$, respectively, defines the position and orientation of the second molecule relative to that of the first. However, the practical considerations of presenting and then analyzing this full 6-dimensional function have made it virtually inaccessible [87].

For atomic liquids represented by only spherically symmetric interactions, the pair distribution function has no angular dependence. Therefore the structure of such systems will be completely described by the radial distribution function, $g(r)$, where $r=|\vec{r}|$ is simply the magnitude of the separation vector. Knowledge of the radial distribution function is sufficient to determine most thermodynamic quantities of atomic systems. For instance, energy and pressure can be calculated using $g(r)$ as

$$\frac{\langle E\rangle}{N}=\frac{3}{2}k_BT+\frac{1}{2}\rho\int g(r)u(r)d\vec{r}, \tag{3.2}$$

and

$$\frac{\beta P}{\rho}=1-\left(\frac{\beta\rho}{6}\right)\int g(r)\,r\frac{du(r)}{dr}d\vec{r}\ , \tag{3.3}$$

respectively, where $\beta = \left(k_B T\right)^{-1}$, ρ is the density and $u(r)$ represents the atomic potential [88].

The isotropic nature of a liquid allows one to examine its structure through its structure factor, $S(k)$, obtained from an X-ray or neutron scattering experiment on the molecules within the liquid. In the case of a homogeneous fluid the function, $S(k)$ can be written as the Fourier transform of $g(r)$[89]:

$$S(\vec{k}) = 1 + \rho \int \exp(-i\vec{k} \cdot \vec{r}) g(\vec{r}) d\vec{r} \tag{3.4}$$

and conversely, $g(r)$ is the Fourier transform of $S(k)$

$$g(\vec{r}) = 1/\rho(2\pi)^3 \int \exp(i\vec{k} \cdot \vec{r})[S(\vec{k}) - 1] d\vec{k}, \tag{3.5}$$

where $\vec{k}$ is the wavevector. For an isotropic system where $k = \left|\vec{k}\right|$ eqn 3.4 becomes

$$S(k) = 1 + 4\pi\rho \int r^2 g(r) \frac{\sin kr}{kr} dr. \tag{3.6}$$

The structure factor of a fluid can be obtained experimentally from measurements of the cross-section for scattering of neutrons or X-rays as a function of scattering angle.

In the case of neutron scattering the differential cross-section, $d\sigma/d\Omega$, for scattering into a solid angle, $d\Omega$, in the direction θ, ϕ is given by

$$\frac{d\sigma}{d\Omega} = \left|f(\theta)\right|^2, \tag{3.7}$$

where $f(\theta)$ is the intensity (amplitude) of the scattered component. The amplitude of the scattered component is usually given by the Fourier transform of the scattering potential, $V(\vec{r})$. Also $f(\theta)$ can be expressed as a matrix element of the interaction $V(\vec{r})$ between initial and final plane wave of the neutron (see

Ref. 89 for the detailed derivation). The final expression for $d\sigma/d\Omega$ in terms of $S(k)$ becomes

$$\frac{d\sigma}{d\Omega} = Nb_{inc}^2 + Nb_{coh}^2 S(\vec{k}), \tag{3.8}$$

where the subscripts refer to incoherent and coherent scattering, respectively, b is the scattering length and the amplitude of the wave scattered by a single fixed nucleus is $f(\theta) = -b$ [89]. The structural information for the fluid can be obtained entirely from the coherent contribution to the cross-section; there is no incoherent contribution from a sample consisting of a single isotopic species of zero nuclear spin. The expression for the cross-section in the absence of incoherent scattering then becomes

$$\frac{d\sigma}{d\Omega} = Nb^2 S(\vec{k}), \tag{3.9}$$

where Nb^2 is the scattering from N independent nuclei and $S(k)$ represents the effect of correlations.

For the scattering of X-rays there is no separation into coherent and incoherent contributions but the expression for the differential cross-section has the same general form as eqn. 3.9. The main difference is that X-rays are scattered by the atomic electrons and instead of b the atomic form factor, $f(k)$, has to be used. Unlike b, the atomic form factor is a function of k and it is defined as

$$f(\vec{k}) = \left\langle \sum_{n=1}^{Z} \exp[-i\vec{k} \cdot (\vec{r}_i^{\,n} - \vec{r}_i)) \right\rangle_Q , \tag{3.10}$$

where the subscript Q indicates a quantum-mechanical expectation value, r_i^n represents the coordinates of the n-th electron of the i-th atom (with nuclear coordinates r_i), and Z is the atomic number [89].

The radial distribution function can be obtained by substituting the appropriate experimental data in eqn. 3.5. This functional relationship can also be extended to systems of more then one component. In this case a partial structure factor has to be defined and then related to the partial radial distribution function. The analogous procedure can be performed to establish the connection with experimental data for a site-site radial distribution function within a one component homogeneous fluid. For example, site-site partial structure factors, and therefore site-site radial distribution functions, $g_{\alpha\beta}(r)$, can be extracted using the recently developed technique of isotropic substitution [90-92] in neutron diffraction experiments. Unfortunately, even a complete set of site-site radial distribution functions cannot properly reconstruct the full molecular pair distribution function because $g_{\alpha\beta}(r)$ represents integrals (convolutions) over the full pair distribution function. In order to allow a more complete description based on the pair distribution function to be extracted from sets of $g_{\alpha\beta}(r)$ provided by scattering experiments, the numerical techniques of spherical harmonic reconstruction [93], reverse Monte Carlo (RMC) [94] and empirical potential structure refinement [95] have been developed.

The spherical harmonic reconstruction (SHR) technique was first developed for diatomic molecules [92] as an attempt to generate a three-dimensional picture of the local orientations in a molecular liquid directly from a set of measured partial structure factors. It has been shown by Soper [93] that the structure of the molecular liquid is fully characterized by the orientational pair correlation function between neighboring molecules which plays the same role as that played by the pair correlation function for simple liquids. The problem is that the site-site partial structure factors (PSF) that can be measured for a molecular liquid involve averages over molecular orientations and therefore represent only indirect information about the true molecular structure at the microscopic level.

The SHR technique, generalized by Soper to the case of molecules of arbitrary shape, appeared to be a method capable of extracting sufficient orientational information from PSF data [93]. Soper concluded that in order to make such an extraction possible any estimated pair correlation function should be smooth but also consistent with the data. The algorithm proposed in Ref. [93] is based on the relationships of Gray and Gubbins [96] where both the orientational pair correlation function and the site-site partial structure factors are expanded as a series of spherical harmonic functions. They are related to one another by a straightforward Fourier-Bessel transform,

$$S(l_1 l_2 l; n_1 n_2; k) = 4\pi\rho \int_0^\infty r^2 h(l_1 l_2 l; n_1 n_2; r) j_l(kr) dr, \tag{3.11}$$

where the coefficients for each expansion are labeled as $h(l_1 l_2 l; n_1 n_2; r)$ for $h(\vec{r}, \omega_1, \omega_2)[= g(\vec{r}, \omega_1, \omega_2) - 1]$ and as $S(l_1 l_2 l; n_1 n_2; k)$ for the partial structure factors $S_{\alpha\beta}(k)$, and $j_l(kr)$ is a spherical Bessel function of order l [96]. The basic aim is to minimize a statistic factor, which measures how well the trial structure factors fit the measured data, and the noise function, which measures how much variation there is in the trial $h(l_1 l_2 l; n_1 n_2; r)$ coefficients. The lower the noise value then the smoother the radial dependence of $h(l_1 l_2 l; n_1 n_2; r)$. The goal to make this dependence as smooth as possible but still remain a sensible fit to the data can be achieved by using the maximum number of coefficients that can be accommodated in a reasonable computing time scale. The main stage of the reconstruction is to generate the set of perturbation coefficients, $\Delta(l_1 l_2 l; n_1 n_2; r)$, which are obtained as the difference between the current and the previously generated reference coefficients. The $\Delta(l_1 l_2 l; n_1 n_2; r)$ are then used in place of $h(l_1 l_2 l; n_1 n_2; r)$ to reconstruct the orientational pair correlation function for the specified molecular orientations [93].

The SHR method was applied to the partial structure factors of water to extract a map of the 3-D structure in water and a picture of the local orientation of neighboring molecules. The results show that in addition to the conventional view of water as a tetrahedrally coordinated liquid there is a possibility of other local structural arrangements. In particular, trigonal and pentagonal structures are also compatible with the data [93].

RMC modeling [97] is a variation of the standard Metropolis Monte Carlo (MMC) method. Since 1988 it has become a powerful and widely used tool for the structural analysis in disordered materials [94]. The aim of RMC is to produce a structural model (usually called configuration) which is consistent with one or more sets of experimental (X-ray or neutron diffraction) data within their errors and subject to a set of constraints [95]. The basic procedure of a RMC simulation is to set up distributions of particles, which are consistent with measured quantities such as the site-site pair correlation function. It is normally assumed that short-range interactions exist between the particles to prevent atomic overlap. The simulation follows that of a standard Monte Carlo simulation algorithm with random moves. The site-site pair correlation function is calculated (providing a move does not violate atomic overlap) and then compared with the data using a χ^2statistic factor. This factor measures how well the simulated correlation function fits the data:

$$\chi^2 = \sum_r \left(\frac{g(r) - g^D(r)}{\varepsilon(r)} \right)^2 , \tag{3.12}$$

where $g(r)$ is the simulated site-site pair correlation function, $g^D(r)$ is the correlation function obtained from diffraction data and $\varepsilon(r)$ represents the experimental error. If the change in χ^2 is negative then the move is accepted

and the "new" configuration becomes the "old" configuration. If χ^2 is positive the move is accepted only with a Boltzmann-type probability distribution. Otherwise it is rejected. As the number of accepted moves increases χ^2 will initially decrease until it reaches an equilibrium value about which it will then oscillate. The resulting configuration will yield a pair correlation function that is consistent with the experimental data within its errors [94,95].

Comparison of the RMC method with conventional MMC reveals that some analogies can be made. In MMC modeling the quantity sampled is $(U_n - U_0)/kT$, where U is the potential energy of the configuration, T is the temperature and k is Boltzmann's constant. In the example of the RMC procedure given above the correlation function plays the role of the energy and $\varepsilon(r)$ plays the role of the temperature. The convenience of this analogy determined the functional form of χ^2. The RMC algorithm can be easily modified for simultaneous fitting to different sets of data. The types of data used include neutron diffraction (including isotopic substitution), X-ray diffraction (including anomalous scattering), electron diffraction, extended X-ray absorption fine structure (EXAFS) and NMR (magic angle spinning and second moment) [94]. However, in computer simulations of the site-site pair correlation function using the RMC method certain problems may arise. For instance, χ^2 is unable to distinguish between two configurations of particles, one of which has a large statistical uncertainty but lies within the supplied data, and the other, with lower statistical uncertainty located out of the peaks. Another problem is that for any computer simulation the pair correlation function will have its own statistical uncertainty, which for a moderate sized system was estimated [95] to be 10% at r= 3 Å. On the other hand, modern diffraction data usually have statistical uncertainties much better that 1%. Therefore any attempt to fit $g(r)$ better than

its intrinsic uncertainty by changing the size of $\varepsilon(r)$, may result in the simulation "trapped" in a local minimum without proceeding on a true random walk.

Soper extended the RMC algorithm so that it establishes empirically an effective site-site pair potential (or set of potentials in the case of multicomponent systems) which reproduces quantitatively the site-site pair correlation function obtained from diffraction data [95]. Comparison of conventional RMC with this empirical potential Monte Carlo method (EPMC) reveals that the *energy* rather than χ^2 is the equilibrated parameter. Therefore it is less likely for the simulation to become locked in a local configuration and the simulated pair correlation functions have correct statistical uncertainties.

The starting point of the EPMC algorithm is the potential of mean force between atoms α and β in the fluid,

$$\psi_{\alpha\beta}(r) = -kT\ln[g_{\alpha\beta}(r)], \tag{3.13}$$

where $g_{\alpha\beta}(r)$ is the simulated site-site pair correlation function. There is also an equivalent potential of mean force $\psi^D_{\alpha\beta}(r)$ for the corresponding correlation function obtained from diffraction data, $g^D_{\alpha\beta}(r)$. The procedure continues by introducing an assumed reference potential, $U^{ref}_{\alpha\beta}(r)$, between sites α and β. The potential of mean force (eqn. 3.13) is then used to generate a new potential energy function, $U^N_{\alpha\beta}(r)$, as a perturbation to the initial potential:

$$U^N_{\alpha\beta}(r) = U^0_{\alpha\beta}(r) + [\psi^D_{\alpha\beta}(r) - \psi_{\alpha\beta}(r)] = U^0_{\alpha\beta}(r) + kT\{\ln[g_{\alpha\beta}(r)/g^D_{\alpha\beta}(r)]\}, \tag{3.14}$$

where the initial potential $U^0_{\alpha\beta}(r)$ and the site-site correlation function $g_{\alpha\beta}(r)$ equal their reference values $U^{ref}_{\alpha\beta}(r)$ and $g^{ref}_{\alpha\beta}(r)$, respectively. The new potential, $U^N_{\alpha\beta}(r)$, replaces $U^0_{\alpha\beta}(r)$ in the simulation which now proceeds with the revised potential. This process is repeated until convergence is reached, i.e.

$$U^{0}_{\alpha\beta}(r) \approx U^{N}_{\alpha\beta}(r) \text{ and } g_{\alpha\beta}(r) \approx g^{D}_{\alpha\beta}(r). \quad (3.15)$$

As a result an empirical potential energy function is determined that is able to reproduce the observed site-site pair correlation function of the material of interest [95]. EPMC method has been tested on the SPC/E model and on experimental water and it gave excellent results in both cases. It revealed the strong similarities between the H-bonding in SPC/E water compared to experimental water as well as highlighted the similarities and differences between the model and the empirical potential. Soper also concluded that for complex systems and complex-shaped molecules EPMC analysis appears to be preferable in comparison with alternative methods such as SHR [93].

3.3.2 Spatial distribution function (SDF)

In molecular systems (e.g. water) with a very strong directionality in the interaction potential, the local arrangements of neighboring particles is expected to be quite anisotropic. Yet, this angular dependence will be simply lost in the angle averaged radial distribution function. It is clear that a complete picture of the local packing pattern around each molecule is essential for understanding the anisotropic local structural order in such molecular systems. There are two ways to solve this problem. In the first approach one could try to extract structural information exploring the angular dependence of the pair distribution function (e.g. using spherical harmonic expansions or geometrical constructions such as Voronoi polyhedra) [87, 98]. The second approach utilizes a quantity termed the spatial distribution function (SDF).

The position of a neighboring particle can be localized (within the local frame of the central molecule) by means of the function given by

$$g(\vec{r}) \equiv g(\vec{r}_{12}) = \langle g(\vec{r}_{12}, \vec{\omega}_2) \rangle_{\vec{\omega}_2}, \tag{3.16}$$

where $\langle \ldots \rangle_{\vec{\omega}_2}$ denotes the average over all orientations of the second molecule. While not reproducing the full pair distribution function, the SDF does completely describe the local packing structure [87]. As a three-dimensional function, SDF can be visualized and its accumulation becomes manageable with modern computer technology.

There are some special issues, which have to be taken into account in the representation, accumulation and visualization of SDF's. These issues arise because of their three-dimensional nature. In comparison with a RDF, a SDF requires a considerably longer simulation run in order to obtain a reasonably well averaged function because of its higher dimensionality. For pure molecular liquids a smooth SDF can be obtained by averaging a system of a few hundred molecules over a trajectory of 100-200 ps, while in mixtures (especially if one of the components is present at low composition) a run length of at least 0.5-1.0 ns may be necessary.

Another important issue is an appropriate choice of the local frame. The local frame for rigid molecules, once chosen, can always be clearly specified. However, if the molecule is very flexible, care should be taken to prevent an interchange of axes in the resulting local frame. The local frame for a flexible molecule may experience the influence of frequently occurring conformational transitions as well as large fluctuations at ambient temperatures. In this case the calculated SDF's could become highly distorted because of the geometrical changes undergone by the central molecule. When the solvation structure near specific areas of large flexible (organic) molecules is of primary interest, it is usually recommended to define a local frame on specific parts of the molecule like a residue, functional group or any other reasonably rigid unit in the area of focus. The accumulated SDF will then provide detailed information about the

local structure around the particular molecular residue or functional group. The technique of assigning the local frame proceeds according to the following steps. If two axes are fixed between any two pairs of atoms, then a third perpendicular axis can be defined by their vector (cross-) product; after that the interatomic site-site distance vector can be transformed into this temporary local frame.

Once a local frame is defined one has to choose a coordinate system to express the separation vectors. The spherical-polar coordinate system, in which $\vec{r}$ has the components (r, θ, ϕ) can be a good choice, especially for small rigid molecules. Unfortunately, the spherical-polar representation of an SDF also has some disadvantageous. The major one is that it samples the different regions of the local space non-uniformly introducing difficulties in visualization. In this case Cartesian coordinates represent an attractive alternative. The data structure obtained is uniform over the local space and can be easily accessed, normalized and visualized. Use of the Cartesian representation appears to be a necessity for complex or very flexible molecules [87].

The visualization of SDF's requires additional consideration. There are two basic choices, either to view the variation of the SDF in two-dimensional slices through the full three-dimensional data set, or to examine the structure present in the data at particular threshold values. Both approaches offer different benefits. The first approach provides an immediate sense of the three-dimensional structures. This can be done by analyzing many different two-dimensional slices and noting how the observed features change with slicing plane. The second gives the better sense of the specificity of structural features. This approach is based on observations of how the nature of three-dimensional isosurfaces changes with threshold value. In either case the main goal of such an analysis is to obtain information on the variations of the SDF over the entire local three-dimensional space. The generation and viewing of SDF isosurfaces can be

performed by means of available graphics software packages developed for visualization of three-dimensional data sets.

Two examples of oxygen-oxygen SDF's for liquid methanol [42] and water- methanol mixtures [43] are shown in Figs. 8a and 8b, respectively. From Fig. 8a one can see the two features due to H-bonded neighbors. It has been noted [42] that the feature due to H-bond donor neighbors appears similar in shape to that found for water [99], although it contains only a single neighbor and has its maximum at the dipole rather than tetrahedral positions. The H-bond acceptor feature in methanol, when compared with that of water, is more highly localized indicating the strong association in this liquid. For a 3:1 water-methanol solution the neighboring water-oxygen density around methanol is shown in Fig. 8b. Two features due to nearest H-bond neighbors can be clearly seen. The third and the fourth features are represented by the upper edge of the cap over the methyl group and the upper part of the second neighbor ring surrounding the H-bond acceptor cap. The latter feature becomes resolved as three distinct tetrahedral second neighbors indicating the strong tetrahedral character of the local structure in water-rich solution.

It is worth noting that the evaluation of solvent effects by considering the local structure of water around a solute molecule on the basis of spatial distribution functions is a relatively new approach. It has been developed by Svishchev and Kusalik and first published in 1993 [99]. Since then it has been successfully applied to many different molecular systems where much of this work has been carried out by the research groups of P.G. Kusalik and A. Laaksonen. It is very encouraging that recently Hata and Ono [100] carried out Monte Carlo (MC) simulations for an infinitely dilute aqueous solution of EG at 298K with the major goal to establish the optimal analysis method for their SDF's, when applied to the hydration structure of a large molecular system. They

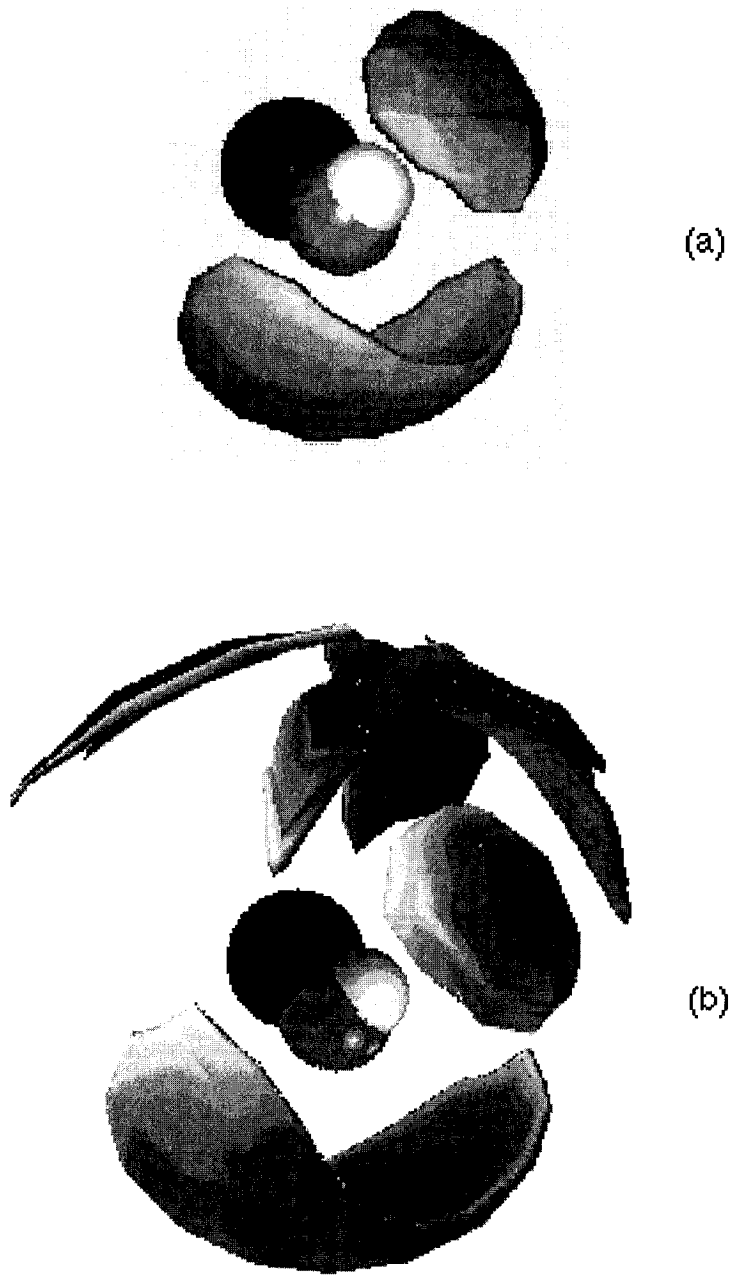


Figure 8. Oxygen-oxygen SDF's for (a) liquid methanol [42] and (b) a 3:1 water-methanol solution [43] at 298 K and isosurface threshold of 1.9.

intended to proceed by employing SDF's in the analysis of fundamental small molecular systems.

The authors examined the oxygen (solute) - oxygen (water) SDF's obtained for two stable conformers (gGg' and tGg') and three additional conformations in between for a EG molecule in liquid water. The oxygen-oxygen SDF was partitioned into three categories defining different high-density regions: hydrogen acceptor (HA), hydrogen donor (HD), MIX (for overlapped distributions of HA and HD), and hydrophobic hydration (HH) regions. The maximum distributions of the hydrogen and oxygen atoms of the hydration water molecules form a triple-layer structure in the HA region and the HA part of the MIX region; their spatial orientations are of a linear type. On the other hand, the maximum distributions of oxygen and hydrogen atoms in the HD region and the HD part of the MIX region form double-layer structures and their spatial orientations are also of a linear type. In addition, it was apparent that the spatial orientations of water molecules were of the linear type throughout the conformational change from the gGg' to tGg'. From the oxygen-oxygen difference SDF (DSDF) between the two conformers the authors concluded that the distribution of hydration water molecules in the HA and HD parts of the MIX region are governed by the competition of internal H-bonds between the hydrogen atom and the two lone-pairs on the oxygen atom of EG molecule [100].

CHAPTER IV - SIMULATION METHODOLOGY AND MODELS

This chapter starts with a brief description of a cornerstone of molecular dynamics simulation, the equations of motion and the most successful computational methods to achieve their solution. In particular, the Gear predictor-corrector and Verlet predictor integrating schemes in conjunction with the SHAKE algorithm, which introduces the concept of constraint into the equations of motion, is described in some detail. Important conceptual quantities, such as statistical ensemble, are introduced and the ensembles most commonly used in simulations are characterized. In addition to the procedure of scaling velocities, the implementation of thermostats within a simulation to generate the constant-temperature ensemble is discussed in the context of the present study. All simulations reported in this thesis were carried out using a molecular dynamics computational package; the general scheme, distinctive features and advantage of this particular program are presented. A separate section is devoted to the description and comparison of the standard interaction potentials (force fields) such as AMBER, CHARMM and OPLS implemented in the MD program used. Two OPLS-based potential functions for liquid alcohols and amines, as well as a special force field for alkanolamines, are also characterized and the selected parameters taken from these force fields and employed in the present work are summarized. The simulation specifications will be given for all stages of the calculations performed. The properties of interest both static and dynamic, such as average configurational energy, enthalpy of vaporization, and self-diffusion coefficient, as well as some characteristics specifically used in structural analysis (coordination numbers, torsional angle distributions, H-bonding) are defined in the final methodological section.

4.1 MOLECULAR DYNAMICS METHOD

4.1.1 **The equations of motion**

The term molecular dynamics (MD) is conventionally used to describe the numerical solution of the classical (Newtonian) equations of motion

$$m_i \ddot{\vec{r}}_i = \vec{f}_i, \tag{4.1}$$

for a set of molecules. In eqn. 4.1 m_i is the mass, r_i is the Cartesian coordinates, and f_i is the force on atom i. Within a MD computer simulation, the history of an initial arrangement of molecules is followed by calculating the trajectories of all the particles under the influence of the intermolecular potentials. Newtonian equations of motion can be used to predict where each particle will be after a short time interval (typically 1 fs for rigid molecules, which is significantly shorter than the time between collisions), and then the calculation is repeated for many (typically thousands or millions) such steps [101].

The equation 4.1 may also be represented in Lagrangian or Hamiltonian form. The most fundamental form is the Lagrangian equation of motion

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0, \tag{4.2}$$

where the Lagrangian function $L(\vec{q},\dot{\vec{q}})$ is defined in terms of kinetic and potential energies as $L = K - V$. It is considered to be a function of the generalized coordinates, q_k, and their time derivatives, $\dot{q}_k$. In other words, the Lagrangian formulation is given in terms of generalized coordinates and velocity with time as a parameter. In contrast, the Hamiltonian form for the equations of motion use generalized coordinates and generalized momenta as the independent variables [102]. The generalized momentum, p_k conjugate to q_k, is defined as

$$p_k = \frac{\partial L(q_k, \dot{q}_k, t)}{\partial \dot{q}_k}. \tag{4.3}$$

The change in independent variables can be performed by means of a Legendre transformation, which results in the definition of the Hamiltonian function,

$$H(\vec{p}, \vec{q}, t) = \sum_i \dot{q}_k p_k - L(q_k, \dot{q}_k, t). \tag{4.4}$$

Substituting the Hamiltonian obtained from eqn. 4.4 into equations referred to as Hamilton's canonical equations of motion [102]

$$\dot{q}_k = \frac{\partial H}{\partial p_k} \quad \text{and} \tag{4.5}$$

$$\dot{p}_k = -\frac{\partial H}{\partial q_k}, \tag{4.6}$$

yields first-order equations of motion. For Cartesian coordinates, Hamilton's equations become

$$\dot{\vec{r}}_i = \vec{p}_i / m_i, \tag{4.7}$$

$$\dot{\vec{p}}_i = -\nabla_{r_i} U = \vec{f}_i, \tag{4.8}$$

where p_i is the momentum of particle i. Therefore, the time evolution of a system of N molecules interacting via a potential U involves solving either a system of 3N second-order differential equations, eqn. 4.1, or an equivalent set of 6N first-order differential equations, eqns. 4.7 and 4.8.

4.1.2 Integrators for the equations of motion

A standard method of solving coupled ordinary differential equations, such as equations 4.7 and 4.8, is the finite difference approach. The general idea is to

obtain at time t new positions and velocities at a later time $t+\delta t$. Many different algorithms fall into the general finite difference pattern and they can be classified as either predictor or predictor-corrector methods. In the predictor methods the molecular coordinates are updated from quantities which are either calculated in the current step or known from previous steps. Examples of predictor algorithms are the Verlet algorithm [103] and its modifications [104,105]. In contrast, predictor-corrector algorithms involve predicting new molecular coordinates, using the predicted coordinates to calculate the value of some function and then using this value to correct the initial prediction. The most widely used in molecular dynamics is the Gear predictor-corrector algorithm [106]. Both these algorithms are most easily applied to atoms while for molecular systems (e.g. rigid molecules) special modifications are required.

The main principles of a predictor-corrector method may be described as following. For the continuous classical trajectory, an estimate of the positions, velocities etc. at time $t+\delta t$ may be obtained by Taylor expansion about time t:

$$\begin{aligned}
\vec{r}^{\,p}(t+\delta t) &= \vec{r}(t)+\delta t\vec{v}(t)+1/2\delta t^2\vec{a}(t)+1/6\delta t^3\vec{b}(t)+\ldots,\\
\vec{v}^{\,p}(t+\delta t) &= \vec{v}(t)+\delta t\vec{a}(t)+1/2\delta t^2\vec{b}(t)+\ldots,\\
\vec{a}^{\,p}(t+\delta t) &= \vec{a}(t)+\delta t\vec{b}(t)+\ldots,
\end{aligned} \tag{4.9}$$

where $\vec{r}$, $\vec{v}$ and $\vec{a}$ stand for the positions, velocities and accelerations, respectively, and $\vec{b}$ denotes the third time derivatives of $\vec{r}$. The superscript p marks these as "predicted" values. However, an equation like eqn. 4.9 will not generate exact trajectories as time advances; to help remedy this situation a correction step is added. From the new positions $\vec{r}^{\,p}$ one may calculate the forces at time $t+\delta t$ and hence the correct accelerations $\vec{a}^{\,c}(t+\delta t)$. These have to be compared with the predicted accelerations (from eqn. 4.9) to estimate the size of the error in the prediction step:

$$\Delta\vec{a}(t+\delta t)=\vec{a}^{c}(t+\delta t)-\vec{a}^{p}(t+\delta t). \tag{4.10}$$

This error as well as the results of the predictor step are included in a corrector step which reads as

$$\begin{aligned}
\vec{r}^{c}(t+\delta t)&=\vec{r}^{p}(t+\delta t)+c_0\Delta\vec{a}(t+\delta t),\\
\vec{v}^{c}(t+\delta t)&=\vec{v}^{p}(t+\delta t)+c_1\Delta\vec{a}(t+\delta t),\\
\vec{a}^{c}(t+\delta t)&=\vec{a}^{p}(t+\delta t)+c_2\Delta\vec{a}(t+\delta t),
\end{aligned} \tag{4.11}$$

where c_0, c_1 and c_2 are coefficients which depend upon the order of the differential equation being solved (here it is second order since the double time derivative of position is being compared with the accelerations computed from $\vec{r}$). Values for these coefficients were proposed by Gear [106] and they can also be found in Ref. [101]. The resulting values of $\vec{r}^{c}(t+\delta t)$, $\vec{v}^{c}(t+\delta t)$ etc., are better approximations to the true positions, velocities, and accelerations [102].

The general scheme of a step-wise simulation, based on a predictor-corrector algorithm (Gear predictor-corrector scheme), is summarized below [101]:

(a) predict the positions and velocities at a time $t+\delta t$, using the current values of these quantities and their derivatives;

(b) evaluate the forces using the new positions;

(c) correct the predicted positions and velocities and their derivatives, using the new forces and eqns. 4.7 and 4.8;

(d) calculate any variables of interest, such as the energy, order parameters etc., and carry out the accumulation of any time averages, before returning to (a) for the next step.

The simple predictor-corrector algorithm described above is a representative example of the many possibilities.

Among the most widely used methods of integrating the equations of motion is that initially proposed by Verlet [103]. This method gives a direct solution of the second-order equations (eqn. 4.1). It is based on positions $r(t)$, accelerations $a(t)$, and the positions $r(t-\delta t)$ from the previous step. In common with the corrector-predictor algorithm the starting point is the Taylor series expansions about $\vec{r}$:

$$\vec{r}(t+\delta t)=\vec{r}(t)+\delta t\vec{v}(t)+1/2\delta t^2\vec{a}(t)+..., \tag{4.12a}$$

$$\vec{r}(t-\delta t)=\vec{r}(t)-\delta t\vec{v}(t)+1/2\delta t^2\vec{a}(t)+... . \tag{4.12b}$$

When eqns. 4.12a and 4.12b are added then the equation for advancing the positions becomes

$$\vec{r}(t+\delta t)=2\vec{r}(t)-\vec{r}(t-\delta t)+\delta t^2\vec{a}(t). \tag{4.13}$$

Equation 4.13 is known as Verlet's algorithm [102,103]. It can be seen that the velocities do not appear in eqn. 4.13. Although the velocities are not needed to compute the trajectories, they are essential for estimating the kinetic and total energies of the system. They can be obtained as

$$\vec{v}(t)=[\vec{r}(t+\delta t)-\vec{r}(t-\delta t)]/2\delta t. \tag{4.14}$$

The Verlet method is compact, easy to implement, time reversible and has excellent energy conserving properties. The main disadvantages of this algorithm are that it does not handle velocities well and its numerical precision in not optimal [102].

The one of various modifications of the original Verlet algorithm was proposed by Hockney [104]. This so-called "leap-frog" scheme makes use of the velocity at half-time intervals:

$$\vec{r}(t+\delta t) = \vec{r}(t) + \delta t\vec{v}(t+\delta t/2), \tag{4.15a}$$

$$\vec{v}(t+\delta t/2) = v(t-\delta t/2) + \delta t\vec{a}(t), \tag{4.15b}$$

$$\vec{v}(t) = [\vec{v}(t+\delta t/2) + \vec{v}(t-\delta t/2)]/2. \tag{4.15c}$$

The algorithm follows the same general procedure as the Verlet algorithm; the difference is that the calculation of velocities now becomes an integral part of the algorithm. The calculated accelerations are used to determine the half-time velocities, which then are used to update the atomic positions and current velocities. The current position and the previous half-time velocity must be stored at each time step and updated following each change in coordinates [104]. The main advantage of the leap-frog algorithm is that the numerical imprecision is reduced because of the use of smaller differences.

4.1.3 Integrators for molecules: SHAKE algorithm

The integrators described above can be used directly to solve the equations of motion for atoms interacting via an appropriate potential (e.g. square-well, Lennard-Jones and other potentials). Unlike atoms, which are treated as spherical particles, the simulation of molecules is complicated by the presence of rotational motion, intramolecular bonds, etc. In particular, molecular bonds constrain the motion of the atoms that comprise the molecule and they have to be taken into account in the methods used for integrating the equations of motion.

Molecules are constrained by various intra-atomic separations and bond angles. For the case of a triatomic molecule comprised of atoms i, j and k there are at least two bonding constraints (angular constraints will be ignored for simplicity):

$$\left.\begin{aligned} \vec{r}_{ij}^{\,2} - d_{ij}^{2} &= 0 \\ \vec{r}_{jk}^{\,2} - d_{jk}^{2} &= 0 \end{aligned}\right\}, \tag{4.16}$$

where r_{ij} and r_{ik} are intra-atomic distances and the d values are the equilibrium bond separations. For a rigid bond constraint if one of the conditions (eqn. 4.16) is not satisfied a bond is broken and the molecule is fragmented. In general, a molecule composed of N atoms can have M constraints,

$$\left.\begin{aligned} \sigma_1(\vec{r}) &= 0 \\ &\cdot \\ &\cdot \\ \sigma_M(\vec{r}) &= 0 \end{aligned}\right\}, \tag{4.17}$$

where $\sigma(\vec{r})$ represents a constraint [107].

The concept of constraint can be introduced into Lagrangian and Hamiltonian mechanics to model the dynamics of large molecules. The resulting 3N Euler-Lagrange equations [102] for the molecule have a form:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\vec{r}}_i}\right) - \frac{\partial L}{\partial \vec{r}_i} = \vec{g}_i\,, \tag{4.18}$$

where L is the Lagrangian for the system of N atoms and $\vec{g}_i$ is the constraint force on atom i, which is defined as

$$\vec{g}_i = \sum_{\alpha=1}^{M} \lambda_\alpha \left(\frac{\partial \sigma_\alpha}{\partial \vec{r}_i}\right). \tag{4.19}$$

In eqn. 4.19 λ_α are the undetermined multipliers associated with the constraints. De Leeuw *et al.* [107] used eqns 4.18 to derive 6N Hamiltonian equations of motion:

$$\dot{\vec{r}}_i = \frac{\vec{p}_i}{m_i} - \frac{1}{m_i}\sum_{\alpha=1}^{M}\gamma_\alpha\left(\frac{\partial\sigma_\alpha}{\partial\vec{r}_i}\right), \tag{4.20}$$

$$\dot{\vec{p}}_i = -\frac{\partial u}{\partial\vec{r}_i} + \sum_{\alpha=1}^{M}\sum_{j=1}^{N}\gamma_\alpha\left(\frac{\partial^2\sigma_\alpha}{\partial\vec{r}_i\partial\vec{r}_j}\right)\dot{\vec{r}}_j, \tag{4.21}$$

where $\gamma_\alpha = \lambda_\alpha$. These first-order equations can be rewritten as 3N second-order expressions

$$\ddot{\vec{r}} = \vec{a}_i - \frac{1}{m_i}\sum_{\alpha}\lambda_\alpha\left(\frac{\partial\sigma_\alpha}{\partial\vec{r}_i}\right), \tag{4.22}$$

where $\vec{a}_i$ is the acceleration of the atom i.

The required values for the multipliers λ_α can be obtained (see Ref. 102 for details) and, therefore, the solution of eqn. 4.22 (or 4.20 and 4.21) can be found. However, a direct solution leads to considerable numerical inaccuracy and the error associated with it increases progressively with time. Ryckaert *et al.* [108] proposed a method to calculate the equations of motion for molecules which are subject to constraint. This, so called, SHAKE algorithm was developed specifically to be used in conjunction with the Verlet integrating scheme.

Substituting for the acceleration in the Verlet equation (eqn. 4.13) for unconstrained atoms one can write

$$\vec{r}_i'(t+\delta t) = 2\vec{r}_i(t) - \vec{r}_i(t-\delta t) + \frac{\delta t^2}{m_i}\vec{f}_i. \tag{4.23}$$

If constraints (e.g. bond separations) are applied then the true positions are

$$\vec{r}_i(t+\delta t) = 2\vec{r}_i(t) - \vec{r}_i(t-\delta t) + \frac{\delta t^2}{m_i}(\vec{f}_i - \vec{g}_i). \tag{4.24}$$

Substituting the constraining force (eqn. 4.19) into eqn. 4.24 and rearranging terms gives

$$\vec{r}_i(t+\delta t) = \vec{r}_i'(t+\delta t) - \frac{\delta t^2}{m_i} \sum_{\alpha=1}^{M} \lambda_\alpha \left(\frac{\partial \sigma_\alpha(\vec{r}(t))}{\partial \vec{r}_i} \right). \tag{4.25}$$

In addition the atomic coordinates at time $t+\delta t$ must obey the constraint conditions represented by eqn. 4.17, i.e.

$$\left.\begin{aligned} \sigma_1(r(t+\delta t)) &= 0 \\ &\cdot \\ &\cdot \\ \sigma_M(r(t+\delta t)) &= 0 \end{aligned}\right\}. \tag{4.26}$$

In the SHAKE algorithm the unconstrained atomic positions $r_i'(t+\delta t)$ are obtained (eqn.4.23) and then adjusted iteratively until the constraint equations (eqn.4.26) are satisfied to within a specified degree of accuracy. Each iteration goes through all the independent constraints where each given constraint usually involves only a sub-set of all atoms. At the I-th iterative loop, the forces of constraint act on these atoms to yield new positions that allow for the evaluation of the Lagrange multiplier. The multiplier calculated for the M-th constraint of the I-th iteration is then used to obtain a new guess for the atomic positions. After roughly ten such iterations, all the constraints are successfully satisfied and the new constrained positions are known [108,109].

4.1.4 Sampling from ensembles

MD simulation generates information at the microscopic level (atomic and molecular positions, velocities etc.) and the conversion of this information into macroscopic terms such as pressure, energy and many others is the field of statistical mechanics. The thermodynamic state of a one-component

macroscopic system is usually defined by a small set of parameters (the number of particles, N, the temperature, T, and the pressure, P). Other thermodynamic properties (density, ρ, chemical potential, μ, heat capacity, C_V,) are state functions, which may be derived through knowledge of the equation of state and the fundamental equations of statistical thermodynamics. Although these properties clearly depend on the microscopic structure and dynamics of the system, their values are completely dictated by the few variables (e.g. NPT) characterizing the thermodynamic state of the system, not by the numerous atomic positions and momenta that define the instantaneous mechanical state [101].

However, the positions and momenta can be considered as coordinates in the multidimensional phase space. The collection of points in the phase space described by a probability density is determined by the chosen set of macroscopic parameters (NPT, NVT etc.) and is referred to as an ensemble. The four ensembles of most common use are the microcanonical (constant-NVE) ensemble, the canonical (constant-NVT) ensemble, the isothermal-isobaric (constant-NPT) ensemble and the grand-canonical (constant-μVT) ensemble.

In computer simulations the choice of the ensemble determines which thermodynamic properties can be evaluated and it also governs the overall simulation algorithm. For instance, there are a few different ways to generate the NVT ensemble. The simplest method to obtain constant temperature is velocity scaling [102]. Exploring two expressions for the kinetic energy, the kinetic energy per particle for the ensemble

$$\langle K \rangle = \frac{1}{2N} \left\langle \sum_i m_i \vec{v}_i \cdot \vec{v}_i \right\rangle \tag{4.27}$$

and the expression from the kinetic theory of gases

$$\langle K \rangle = \frac{3}{2} k_B T \ , \tag{4.28}$$

one can obtain the following relationship for temperature

$$T = \frac{1}{3Nk_B} \left\langle \sum_i m_i \vec{v}_i \cdot \vec{v}_i \right\rangle . \tag{4.29}$$

From equation 4.29 the actual temperature (T_A) can be determined for the ensemble at any time. The velocities can then be scaled to achieve a desired temperature (T_D):

$$\vec{v}_i^{new} = \vec{v}_i \sqrt{\frac{T_D}{T_A}} . \tag{4.30}$$

In the microcanonical algorithm the kinetic energy is not a constant of motion, therefore, periodic re-scaling of velocities during the equilibration period is essential. In the velocity scaled algorithm for the canonical ensemble velocity scaling after the initialization phase is optional [102].

The NVT ensemble can also be generated by applying a thermostat, which requires modification of the equations of motion. Among the various strategies [102] the Nosé-Hoover algorithm [110,111], which includes a thermal reservoir as an integral part of the system, appears to be the most straightforward way for generating the NVT ensemble. The Nosé-Hoover equations of motion are

$$\dot{\vec{r}} = \frac{\vec{p}}{m_i}, \tag{4.31}$$

$$\dot{\vec{p}} = \vec{f}(\vec{r}) - \xi \vec{p}, \tag{4.32}$$

$$\xi = \frac{g}{Q}(kT_A - kT_D), \tag{4.33}$$

where g is the number of degrees of freedom of the physical system and Q is the thermal inertia parameter which couples the thermal reservoir to the real

system and affects temperature fluctuations. In the Nose-Hoover equations ξ is a friction coefficient, which evolves with time according to eqn. 4.33 and must be re-evaluated at each time step. The integration of the eqns. 4.31-4.33 can be performed by means of either Verlet or Gear algorithms [101,102].

4.2 M.DynaMix SIMULATION PACKAGE FOR MOLECULAR MIXTURES

The M.DynaMix - MD simulation package has been developed for simulations of arbitrary mixtures of both rigid and flexible molecules [112]. It employs modern simulation techniques for high-quality simulations and it is built around widely used conventional force fields (AMBER [113,114], CHARMM [115,116] and GROMOS [61]). The program also allows inclusion of the most common water models (rigid or flexible) in the simulations of aqueous solutions. Besides the NVE statistical ensemble, both the NVT and NPT ensembles can be sampled by the program by employing the appropriate Nosé-Hoover thermostats and barostats [112]. The following MD schemes and statistical mechanical environments are implemented in the package:

1. Standard Newtonian NVE molecular dynamics with the Verlet leap-frog algorithm [104]. If needed, the temperature can be controlled by scaling the velocities; this procedure is recommended in the case of initial states for simulations of dynamical properties.
2. Constant-temperature Nosé-Hoover molecular dynamics in a constant-volume cell [110,111].
3. Constant-temperature-constant-pressure molecular dynamics with isotropic cell fluctuations [117].
4. Constant-temperature-constant-pressure molecular dynamics with anisotropic cell fluctuations, which allows the simulation cell to fluctuate separately in the three coordinate directions [117].

In the case of flexible molecular models the double time-step algorithm of Tuckerman *et al.* [118] was implemented. Forces due to fast intramolecular motion and fast nearest short-range non-bonded interactions are recalculated at every short time-step. The forces due to long-range non-bonded interactions (the more slowly fluctuating internal forces) are recalculated at each long time-step. Typical values for the short and long steps are 0.2 and 2.0 fs, respectively. In the case of rigid molecular models or if the fast motion of bond stretching needs to be frozen, a constrained dynamics using the SHAKE algorithm [108] is applied (see Section 4.1.3 for details).

The issues related with the implementation or choice of a particular force field for MD simulations require special consideration. In the conventional MD method all atoms of the simulated system move according to classical equations of motion. The forces acting on the atoms are defined from the gradients of the potential energy (force field) and as functions of the distances between all interacting sites in the system (see eqn. 4.8). In the more popular conventional force fields such as AMBER, CHARMM, OPLS [119,120] and GROMOS the core of the interaction potential can be represented by the following functional form:

$$V = \sum_{bonds} K_b (r - r_{eq})^2 + \sum_{angles} K_a (\theta - \theta_{eq})^2$$

$$+ \sum_{torsions} 1/2 K_t (1 + \cos(m_t \phi - \gamma_t))$$

$$+ \sum_{i<j} \left\{ 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{r_{ij}} \right\}, \qquad (4.34)$$

where the first sum runs over all covalent bonds, the second runs over all bond angles, the third over all torsional angles and the fourth sum runs over all non-

bonded atom pairs. $K_{b,a,t}$ are the force constants, r_{eq} and θ_{eq} are equilibrium values for the bond lengths and bond angles, σ and ε are the Lennard-Jones parameters, and r_{ij} is the distance between the atoms i and j. Usually the cross terms (between atoms of different types) of the Lennard-Jones parameters are calculated by the Lorentz- Berthelot combination rules

$$\varepsilon_{ij} = (\varepsilon_i \varepsilon_j)^{1/2} \quad \text{and} \tag{4.35 a}$$

$$\sigma_{ij} = (\sigma_i + \sigma_j)/2. \tag{4.35 b}$$

One of the most prominent features of the M.DynaMix program is that in addition to the standard form of the interaction potential, its authors have implemented a variety of auxiliary terms for inter- and intramolecular contributions. Some of them are given below.

1. Morse potentials for covalent bonds [112]:

$$V_{bond} = D(1 - \exp(-\rho(r - r_{eq}))^2). \tag{4.36}$$

2. The MM3 force field potential for torsional angles [121]:

$$V_{tors} = K_1/2(1+\cos(\phi)) + K_2/2(1-\cos(2\phi)) + K_3/2(1+\cos(3\phi)). \tag{4.37}$$

3. Ryckaert-Bellemans potential for torsional angles [122]:

$$V_{tors} = \sum_{i=1}^{5} K_i(\cos(\phi - 180)^i). \tag{4.38}$$

Originally this potential was intended for hydrocarbon chains and polymers. However, it has recently been used successfully as a part of an unconventional force field by Alejandre *et al.* [67] to describe the rotational motion in AE.

4. Improper torsional angle potentials [113]:

$$V_{imp} = K_{imp}/2(\psi - \psi_{eq})^2. \tag{4.39}$$

5. Lennard-Jones 10-12 potentials [113] for explicit inclusion of H-bonds:

$$V_{H-bond} = \sum_{H-bonds} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right). \tag{4.40}$$

This type of potential allows for better control of H-bond distances.

M.DynaMix can be run on any single processor computer or on any parallel system with the MPI library installed. There are two conventional ways to parallelize MD programs. In the replicated data (RD) method [123] all the nodes keep the position of all the particles in the system, while the calculation of different contributions to the forces is divided between the nodes and is done in parallel. In the domain decomposition (DD) approach [124] the particles are distributed between the nodes with each of the nodes being responsible for the particles in the corresponding subcell. The RD method is normally preferable for a parallel simulation of complex biomolecular systems on computers with less than 100 processors and it has been implemented in the M.DynaMix program [112].

The general organization of calculations with the M.DynaMix package is presented in Fig. 9. The program reads the main input file through the standard input. This file specifies the system being studied and all the simulation parameters. For each molecular species a specific topology file must be created, which describes its molecular structure and parameters of the force field. These so-called .mol files contain information about relative coordinates of all atoms, q (charges), σ and ε (Lennard-Jones) parameters, and some additional optional force field parameters (see Section 4.3). These files can be prepared/edited manually or using a supplied utility (*makemol*) which was designed to simplify the procedure of writing .mol files for large molecules. This provides considerable

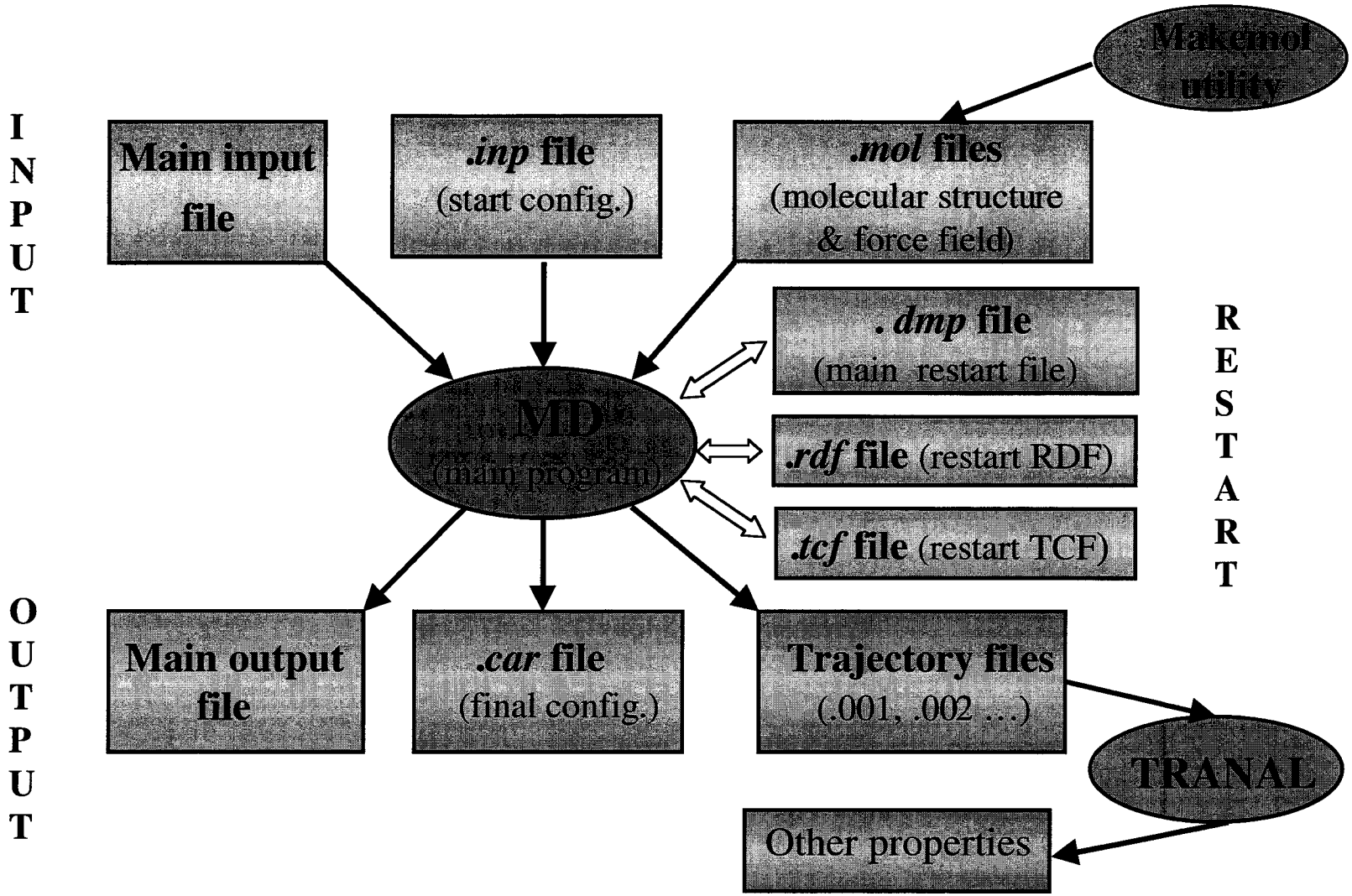


Figure 9. The general organization of calculations using the M.DynaMix package [112].

flexibility in changing force field parameters. Additionally, some features in the force field can be specified through the input file, for example, the scaling of the "1-4" intramolecular interactions. During the execution, the program periodically dumps a restart file containing the current configuration of the system and all averages calculated to that point. There are separate restart files for the radial distribution functions (RDF) and time correlation functions (TCF) if they are calculated during the MD run. Hence, the program can be interrupted at any time and the execution can be continued later from the restart file. The program produces an output file containing all the basic information and simulation results. The final configuration of the system can be dumped in appropriate format and then visualized by a molecular viewer (e.g. XMOL). During the simulation trajectories of the atomic coordinates can be produced and stored in sequentially numbered files. A set of files can be analyzed later using an analysis program named TRANAL. TRANAL allows one to extract the following additional properties of the simulated system:

- SDF's in a local coordinate system attached to a given molecule or residue. Output of SDF's can be later visualized using packages such as gOpenMol [125] or Data Explorer (DX);
- three-body correlations (i.e. a density distribution of a third particle around two particles separated by a specified fixed distance);
- RDF's;
- time dependence of the root-mean-square displacement of molecules and the self-diffusion coefficients;
- residence (life) times of specified atoms to be within a certain distance around another atom;
- population distribution of specified torsion angles.

The main block of the TRANAL module reads the trajectory files in binary format and then performs the required calculations for specified properties.

The M.DynaMix package has been successfully used in simulations of a wide variety of molecular systems such as aqueous electrolyte solutions, double-helix DNA fragments surrounded by water and ions, dicarboxylic acids in solution, *tert*-butyl alcohol in aqueous solution and disaccharides in water and water-DMSO mixtures [112].

4.3 FORCE FIELDS

4.3.1 General features and comparison of molecular force fields

A general description of the molecular modeling force fields in use today can be given in terms of a composite potential which incorporates the contributions of various possible intra- and intermolecular interactions. The four main energetic contributions are those associated with the deviation of bonds and angles away from their reference or equilibrium values, the rotation of bonds and the interaction between non-bonded parts of the system. More sophisticated force fields may have additional terms such as the out-of-plane bending potential, or/and the "cross terms" due to bond-angle or angle-angle interactions, but invariably contain these four components. The functional form for a force field that can be used to model single molecules or assemblies of atoms and/or molecules is given by eqn. 4.34.

In order to define a force field not only the functional form but also the parameters (the various constants such as $K_{b,a,t}$, σ and ε in eqn. 4.34) have to be specified; different force fields may have an identical functional form but yet have very different parameters. Some of the terms in a force field are essentially independent of all others (particularly the bond and angle terms), however a force field should be considered as a single entity. It is not correct to divide the

total energy into its individual components and to use some of the parameters from one force field in conjunction with parameters from another force field. The accuracy of the force field (intermolecular potential) can depend critically on the accuracy of the various parameters used. The bond and angle parameters can be obtained from X-ray, IR, NMR and microwave data. The non-bonded parameters can be obtained from experimental data for second virial coefficients and molecular beam experiments. The parameters can also be calculated using *ab-initio* and semi-empirical methods.

Force fields are generally designed to predict certain properties of certain molecular systems and therefore have been parameterized (optimized) accordingly. For instance, the Optimized Potential for Liquid Simulation (OPLS) [119] is specifically optimized to reproduce properties of liquids (which is reflected in its values for non-bonded parameters). The interaction sites of OPLS can be either nuclei or $(CH)_n$ groups modeled as a single pseudo particle (or united atoms). Other examples of potentials include the Chemistry at HARvard Macromolecular Mechanics (CHARMM) [115,126] and Assisted Model Building with Energy Refinement (AMBER) [113] potentials originally designed for simulations of proteins and nucleic acids, as well as molecular mechanics (MM2, MM3) potentials for modeling organic molecules in the gas phase [127,121]. These force fields contain the four components mentioned above with the non-bonded term represented by Lennard-Jones (for van der Waals interactions) and Coulomb (for point-charge electrostatic interactions) contributions. The CHARMM force field includes a contribution from out-of-plane bending (improper torsional potential) of the form

$$V(\chi) = \sum_{angles} K_\chi (\chi - \chi_0)^2, \tag{4.41}$$

where χ is the out-of-plane angle and K_χ is the force constant for out-of-plane bending. The CHARMM and AMBER force fields use the same representation of non-bonded forces; the only difference is that unlike CHARMM, the AMBER force field includes a hydrogen-bonding (the 10-12) term which is given by eqn. 4.40. The MM3 force field incorporates a more elaborate description of the various bonded and non-bonded intreraction potentials. For instance, a Morse potential is used in the bond stretching term and the accuracy of the bond angle bending and out-of-plane bending is improved by using higher order expansions. A distinctive feature of the MM3 potential is the inclusion of different cross-terms. Such a complex functional form is necessary for reproducing vibrational frequencies and subtleties of molecular geometries [102].

There are two important concepts that are common to most force fields. The first one is that of an atom type. The atom type is defined by such characteristics as atomic number, hybridization and details of the local environment. For example, for carbon atoms a distinction must be made between sp^3, sp^2, and sp hybridized carbons, which adopt tetrahedral, trigonal and linear geometries, respectively. The geometry of the atom type in turn affects the values of potential parameters. The local environment also affects these parameters. In the AMBER force field the carbon atom at the junction between a six and five-membered ring is assigned an atom type that is different from the carbon atom in an isolated five-membered ring [128].

The second common concept of the AMBER, CHARMM, MM3 and OPLS models is the assumption that the charge distribution associated with each atom does not depend on the conformation of the molecule. This concept gives rise to an important feature, namely transferability of the parameters of a force field. Transferability means that the same set of parameters chosen within a potential model can be used to model a series of related molecules. Unfortunately, in the

models discussed above the rearrangement of valence electrons found in molecular bonds is not taken into account and therefore site-site interactions are often not transferable to another molecule in the same group [102,128]. To overcome this problem it is sometimes desirable to define a new set of parameters specific to the molecular systems where particularly accurate work is required. Two examples of such a refinement are considered below.

4.3.2 Potential functions for liquid alcohols and amines

W. Jorgensen [6] developed intramolecular potential functions for liquid alcohols and other molecules with hydroxyl groups using standard procedures and Monte Carlo (MC) simulation techniques. These functions are based on earlier work for liquid hydrocarbons (the original OPLS potential model [119]) and required introduction of a few new parameters. In the process of parameterization Jorgensen followed the same guidelines as in earlier work [119] and the initial parameters derived in this way were then tested on H-bonded dimers of liquid methanol and methanol-water complexes. The density appeared to be a little low so some adjustments were made for the σ of oxygen. After this refinement simulations were performed at 298 K and 1 atm for five liquid alcohols: methanol, ethanol, 1-propanol, 2-propanol, and 2-methyl-2-propanol. It was found that the same parameters could be successfully used for the hydroxyl group in each case. It is worth noting that only four independent parameters were added to the OPLS set to describe alcohols, the charge, σ and ε for oxygen and the charge on hydrogen. The ε for oxygen was assigned an average value based on oxygen in water and amides. The charge on C(O) (carbon with attached oxygen) was determined from molecular neutrality and the Lennard-Jones parameters for hydroxyl hydrogen were taken to be zero [41]. The selected parameters for alcohols are given in Tables 3, 4 and 5 in Section 4.4.

Thermodynamic and structural results were obtained by Jorgensen [6] for several liquids and compared with experimental data. The computed and experimental data were found to be in excellent agreement with an average difference only 1.3%. The intramolecular energies were slightly lower in the liquids than in the gas phase. It was also demonstrated that the OPLS functions yield excellent densities for liquid alcohols with an average deviation from the experimental values of 1.8%. The results for heat capacities and compressibilities are also in good agreement with experimental data. The quality of the agreement between simulation and experiment is remarkable in view of the simple form and straightforward parameterization of the potential functions [6].

Early studies of free energies of hydration for both amines and amides revealed that standard force fields failed for these systems [129,130]. Rizzo and Jorgensen [52], using free energy perturbation calculations in Monte Carlo simulations (MC/FEP), developed force field parameters for amines primarily by fitting to experimental data for pure liquids and to hydrogen bond strengths from gas phase *ab initio* calculations. Initially standard bond-stretching and angle bending parameters (for the first two terms in eqn. 4.34) were assigned from the OPLS all-atom parameter set [120], and the main focus was placed on the development of the Fourier coefficients (eqn. 4.37), partial charges, and Lennard-Jones parameters. The details of parameterization as well as parameters obtained can be found in Ref. [52] (the selected set of parameters for amines is also presented in Tables 3, 4 and 5). Calculations were performed by Rizzo and Jorgensen [52] for ammonia, methylamine, dimethylamine and trimethylamine. The results obtained give that the most favorable ΔG_{hyd} occurs for methylamine in accord with the experimental data, but in contrast with previous computational reports. It was noted that the computed heats of vaporization are most affected by the choice of partial charges, while densities are particularly sensitive to the

Lennard-Jones radii. MC simulations of thirteen additional amines, not included in the original parameterization, demonstrated the transferability of this force field. The authors concluded that the prior problems with classical force fields for amines were a result of non-optimal parameterization rather than a lack of inclusion of explicit polarization [52].

Finally it is worth mentioning here the attempt of Alejandre *et al.* [67] to develop a force field for pure alkanolamines and their aqueous solutions. The authors chose AE to start their investigation in which they combined *ab initio* calculations with canonical MD simulations carried out for the liquid phase and for the liquid-vapor equilibrium. The potential includes intramolecular interactions given by harmonic type potentials for bond angles, and Ryckaert-Bellemans potentials for torsional angles (eqn. 4.38), and intermolecular interactions represented by the standard Lennard-Jones and Coulomb potentials. The potential parameters for AE are presented in Ref. [67]. The proposed force field gives a good description of the liquid-vapor coexistence in AE. The liquid density obtained at 298 K is 1.003 g/cm^3 which compares well with the experimental value of 1.012 g/cm^3. The estimated critical point is located at 584 K and 0.32 g/cm^3 in comparison with the experimental result of 614 K and 0.31 g/cm^3. The surface tension calculated at 323 K (43.2±2.5 mN/m) is in excellent agreement with the experimental value of 44.81 mN/m.

4.4 METHODOLOGICAL DETAILS

4.4.1 Molecular and potential models

In the present study each molecule was first modeled using a united atom approximation [131]. In this approximation the EG molecule consisted of six, the AE molecule of seven and the ED molecule of eight interaction sites. The methylene groups (CH_2) were considered as single interaction sites with their

centers located at the position of the carbon atoms. In the united atom approach, the hydrogen atoms of the methylene groups are not explicitly considered; they are taken into account implicitly in the parameters. Furthermore, the united atom description allows us to ignore some internal rotation within the molecule (i.e. the methyl rotation, which is believed to play a minor role in determining the hydration structure). All bond lengths were kept fixed by means of the SHAKE algorithm [108] during the major part of the present simulations. However, for the purposes of comparison, two additional models were constructed: a rigid all-atom model of EG represented by 10 interaction sites and a flexible united atom model of AE. Both these models were studied in the gas phase as well as in the liquid state. Only rigid, united atom molecular models of EG, ED and AE were used throughout the simulations of aqueous solutions.

Three AMBER/OPLS-based force fields by Jorgensen *et al.* [6,52,120] were considered in the present study. The generalizations of the OPLS model proposed for simulations of liquid alcohols [6] and amines [52] were used to describe intra- and intermolecular interactions in EG and ED. For AE force field parameters such as bond lengths, bond angles and non-bonded parameters specifically related to hydroxy and amino groups were taken from Refs. 6 and 52, respectively. These parameters are summarized in Tables 3, 4 and 5. Table 4 does not include parameters (i.e. Fourier coefficients) adopted for the NCCO and OHCC dihedral angles of AE; the appropriate force constants (for XCCX, where X stands for a non-hydrogen atom, and OHCC dihedral angles) can be found in Ref. 113. The all-atom characteristics for the methylene groups from Ref. 120 in conjunction with parameters for hydroxyl groups and the torsion potential from Ref. 6 were used in the all-atom molecular model of EG. In addition two auxiliary force fields, CHARMM and the one proposed by Alejandre *et al.* [67], were

Table 3. Geometrical parameters and force constants[a] used in simulations of pure molecular systems and their aqueous solutions.

Compound	Bond lengths, Å		K_b[d]	Bond angles, deg.		K_a[d]
EG[b]	C-C	1.526	1087.840	CCO	109.5	334.72
	C-O	1.423	1615.024	COH	108.0	230.12
	H-O	0.956	2313.752			
ED[c]	C-C	1.546	1087.840	CCN	111.5	334.72
	C-N	1.469	1535.528	CNH	109.5	146.44
	H-N	1.017	1815.856	HNH	106.4	182.45
AE[d]	C-C	1.526	1087.84	CCO	109.5	334.72
	C-O	1.425	1615.024	CCN	111.2	334.72
	C-N	1.471	1535.528	COH	108.5	230.12
	H-O	0.960	2313.752	CNH	109.5	146.44
	H-N	1.010	1815.856	HNH	106.4	182.42

a - units for K's are kJ/mol; b - [74]; c - [36]; d - [6,52].

Table 4. Fourier coefficients[a] for intramolecular rotational potential functions.

Compound	Dihedrals	K_0	K_1	K_2	K_3
Eg[b]	OCCO	0.000	2.9372	-0.8870	12.8031
	HOCC	0.000	3.4895	-0.4853	3.1254
Ed[c]	NCCN	0.000	10.0081	-2.8200	2.3012
ED, AE	HNCC	0.000	-0.7950	-1.7447	-1.7489

a - units for K's are kJ/mol; b - [6]; c - [52].

applied to AE for the sake of comparison of potential models. For all models 1-4 non-bonded interactions (for the sites separated by exactly three covalent bonds) were included in the calculations. For selected models 1-4 electrostatic and Lennard-Jones interactions were scaled by the factors of 1/1.2 [132] and 1/8 [133,134], respectively.

Table 5. Non-bonded parameters [6,52].

Atom or group	q	σ, Å	ε, kJ/mol
CH_2 (EG)	0.265	3.905	0.494
O	-0.7	3.07	0.711
H(O)	0.435	0.0	0.0
CH_2 (ED)	0.18	3.905	0.494
N	-0.9	3.3	0.711
H(N)	0.36	0.0	0.0
CH_2 (AE)	0.257	3.905	0.494
O	-0.644	3.07	0.711
N	-0.938	3.25	0.711
H(N,O)	0.374	0.0	0.0

In addition to the parameters listed in Tables 3-5, specific topology files (.mol files) created for the most successful molecular models of EG, ED and AE (see Chapter V for details) and explored in the present work are given in Appendix B.

4.4.2 Simulation specifications

The M.DynaMix program [112] has been used in the present work to perform calculations of EG, ED and AE in gas and liquid state as well as in aqueous solutions. The gas-phase geometries (conformations) of ED, AE and EG have been used as the starting geometries for all liquid phase simulations. MM3 type potentials (given by eqn. 4.40 [6,52,121]) have been used for torsional angles of ED, EG and AE. The SPC/E water model [55] has been employed in simulations of aqueous solutions. All calculations were carried out at room temperature (298 K) and experimental densities [135,136,137] (see Table C1). Both canonical (NVT) and isothermal-isobaric (NPT) conditions have been used throughout this work. The latter has been applied to a binary mixture of pure

compounds (EG and ED) as part of an investigation of cooperativity effects and to selected compositions of aqueous solutions when accurate experimental data were not available as an attempt to obtain uniform values for the self-diffusion coefficient.

To model gas phase conditions single molecule calculations of a total duration of 25 ps with a time step of 1 fs were carried out. For pure liquids and their aqueous solutions the total number of molecules was 256 per simulation cell. Liquid systems were simulated for 300 ps (with averages collected after 100 ps) for pure compounds and 1 ns (with averages collected after 200-300 ps) for most aqueous solutions. A detailed description of simulation specifications is given in Table 6.

Table 6. Simulation specifications for pure compounds and their aqueous solutions at 298 K.

	Run				Pure
	I	II	III	IV	liquid
Number of H_2O	248	230	179	51	-
Number of solute	8	26	77	205	256
X^a	0.03	0.1	0.3	0.8	1
Time step (fs)	0.5	0.5	0.5	0.5	1.0
Equilibration (ns)	0.3	0.3	0.2	0.2	0.1
Production (ns)	1.4	0.7	0.3	0.8	0.2

a - mole fraction of solute.

To prepare initial configurations of aqueous molecular systems the simulations are started at constant volume and the thermostat turned off. In this case simple velocity scaling regulates temperature. In addition the simulations are always started at low density (i.e. with a large simulation box). This

essentially produces gas phase conditions reducing the probability of overlapping atoms during the automatic generation of an initial configuration. Also any total force acting on any atom growing above a specified level is cut (truncated) to the threshold value. Then in a sequence of short runs (100-200 time steps) the density is gradually increased to the experimental value. During this procedure the molecules have enough time to adjust and organize themselves so as to avoid overlaps. Finally all the "emergency settings" (velocities, forces) are turned off at this point.

In addition to the main MD simulation package (M.DynaMix) other computational software has been explored in the present study. Molecular geometries were optimized by means of the GAUSSIAN 98 program [138]. It also has been used to produce starting coordinates for all molecular models under consideration. Visualization of current configurations and SDF's has been performed using XMol and gOpenMol [125] software.

4.4.3 Properties of interest

The positions, velocities and accelerations obtained as a result of solving Newton's equations of motion for a particular thermodynamic state of the system allow one to determine various thermodynamic, structural and dynamical properties.

The total average configurational energy, $\langle U \rangle$, which includes the average Coulombic and Lennard-Jones energies, $\langle U_Q \rangle$ and $\langle U_{LJ} \rangle$, respectively, is reported as part of the standard output of the M.DynaMix program. This quantity can be directly compared with experimentally determined heats of vaporization. Such a comparison appears to be the first (and main) criterion in conventional (e.g. AMBER, OPLS) force field developments. The (molar) heat of vaporization, can be determined as

$$\Delta H_{vap} = \Delta E + P\Delta V_m \text{ (at constant P)} \quad (4.42)$$

or

$$\Delta H_{vap} = (E_{inter}(g) - E_{inter}(l)) + (E_{intra}(g) - E_{intra}(l)) + RT, \quad (4.43)$$

where $E_{inter}(l)$ is the internal energy of the interacting molecules in the liquid (the intermolecular van der Waals and electrostatic contributions), $E_{intra}(g,l)$ the intramolecular energy (bond lengths, angles and torsions as well as intramolecular van der Waals and electrostatic terms) for the gas or liquid, R is the gas constant and T is the temperature (here RT accounts for the $P\Delta V_m$ term) [133]. If $E_{inter}(g)$ for the gas phase is approximated as zero (assuming an ideal gas) then eqn. 4.43 becomes

$$\Delta H_{vap} = E_{intra}(g) - E_{inter}(l) - E_{intra}(l) + RT. \quad (4.44)$$

To determine values of the heat of vaporization, the energy was substituted by the corresponding values of the appropriate energies in eqn. 4.44.

The discussion devoted to definitions and developments of structural quantities, such as radial and spatial distribution functions, can be found in the Chapter 3 of this thesis. By definition, the orientationally averaged radial distribution function, $g(r)$, represents the probability of finding a neighboring particle at a separation r from a chosen one. In spite of the fact that it usually provides only limited information on the local structure, it can be used to determine so-called the first coordination number, CN, defined in Ref. [139] as

$$CN = 4\pi\rho_n \int_0^{r_m} g(r)r^2 dr, \quad (4.45)$$

where r_m indicates the location of the first minimum in $g(r)$ and ρ_n is number density in N/Å^3 (see Table C2). In other words, numerical integration of the first peak of $g(r)$ may provide important quantitative information about the number of

nearest (possibly H-bonded) neighbors with respect to the chosen atomic site. This parameter can be especially useful in conjunction with results obtained from the analysis of SDF in elucidation of the local structure for simple liquids.

Since the main focus of the present work is a set of strongly associated liquids and the final goal is elucidation of their H-bonded structural patterns, some essential definitions have to be made in this regard. To date the term "hydrogen bond" has been applied to a wide range of interactions. Very strong H-bonds resemble covalent bonds while the energies of very weak hydrogen bonds are close to those of van der Waals forces. The majority of H-bond interactions are distributed between these two extremes [140]. In 1960 Pimentel and McClellan in the book "The hydrogen bond" [141] made a very general definition which appeared to satisfy all investigators at that time. With this phenomenological definition experimentalists could take different views as to what constitutes evidence of H-bonding. The definition of H-bond adopted in this study is closely related to the definition proposed by Reichardt for strongly associated liquids [142] while remaining in complete accord with that of Pimental and McClellan. In Ref. 142 it is stated that "Liquids possessing OH groups or other groups with a H atom bound to an electronegative atom X are strongly associated". Consequently for our particular case hydrogen bonding is "a special type of attraction that exists between the hydrogen atom in a polar bond (particularly an H-O or H-N bond) and an unshared electron pair on a nearby electronegative atom (O or N) on another or the same molecule" [140].

In polyatomic fluids molecules usually possess internal rotations that give rise to a variety of conformational possibilities which, in turn, dictate certain structural arrangements of molecules within the liquid. In this case, knowledge of the dominant molecular conformations is essential for recognizing the basic structural pattern. In principle, the average values of the torsional angles can be

obtained as part of the main output. Unfortunately, if the actual distribution of the torsion angle under consideration is not uniform, e.g. includes positive and negative values with approximately the same or entirely different probabilities, the final (average) value appears to be rather more confusing than helpful. To overcome this problem the module TRANAL has been used to obtain the average torsional angle distributions for all possible dihedral angles of EG, ED and AE. It has been noted, however, that the distributions with respect to peripheral dihedral angles (angles due to the rotation around C-O and/or C-N bonds) demonstrates sensitivity to the simulation conditions (i.e. run lengths) and, therefore, the main focus in the present analysis will be on the evolutions of the distributions with respect to central (Y-C-C-X) dihedral angle.

Usually the quantities based on the time-dependent behavior of the system (dynamical properties) require knowledge of time correlation functions. Time correlation functions measure how a value of some dynamic quantity $A(t)$ may be related to the value of some other quantity at some other time $B(t')$. The unnormalized cross-correlation function is defined as

$$C_{AB}(t) = < A(0) \cdot B(t) > , \tag{4.46}$$

where (0) denotes an arbitrary initial time and <...> the statistical ensemble average [143]. When A and B are the same quantity, C is called an autocorrelation function. For example, the normalized velocity autocorrelation function can be written as

$$C(t) = \frac{< \vec{v}_i(0) \cdot \vec{v}_i(t) >}{< v_i^2 >}, \tag{4.47}$$

where v_i is the center-of-mass velocity of a molecule i and $< v_i^2 >$ is its average square velocity. The velocity autocorrelation function can be utilized to compute the self-diffusion coefficient, D,

$$D = \frac{1}{3}\int_0^\infty < \vec{v}_i(0) \cdot \vec{v}_i(t) > dt. \tag{4.48}$$

The corresponding Einstein relation for the self-diffusion coefficient is

$$D = \lim_{t\to\infty} \frac{1}{6t}\left\langle \left|\vec{r}_i(t) - \vec{r}_i(0)\right|^2 \right\rangle, \tag{4.49}$$

where $\vec{r}_i(t)$ is the position of a molecule i [101]. Self-diffusion describes the random Brownian motion of the molecules in a homogeneous system. Therefore, the self-diffusion coefficient can serve as a measure of the mobility of each component in binary or multicomponent mixtures. In the present work the values for self-diffusion coefficients were taken from the standard output as a result of application of eqn. 4.49.

CHAPTER V - ELUCIDATION OF THE LOCAL STRUCTURE IN PURE COMPOUNDS

5.1 GAS-PHASE SYSTEMS

In the present study gas-phase (single molecule) simulations were performed as a first and essential stage in the investigation of structural behavior in liquid systems. It is well known [4] that in liquid phase simulations the quality of the final results depends strongly on the adequacy of the chosen force field and modeling conditions. In this context vacuum (gas-phase) simulations represent an attractive choice for preliminary testing of potential and molecular models. The fact that a significant amount of information related with geometric (conformational) and energetic characteristics of single molecules of AE, ED and EG is available from computational studies, provides an excellent opportunity for comparing results and making careful choices. Additionally, gas-phase simulations are absolutely crucial for this sort of computational work because they allow for comparison with the experimental literature. In numerous publications devoted to the development of model potentials it has become common practice to use the experimentally determined heat of vaporization as a main criterion for evaluating the quality of the parameterization. It is unfortunate that such a comparison appears to have been abandoned in the majority of publications on 1,2-disubstituted ethanes in the liquid state, independent of what kind of the force field (conventional or newly developed) was in use. In this work the computed heat of vaporization, which was defined in Chapter IV as a difference between total configurational energies in gas and liquid state (see eqn. 4.47), is compared with available experimental data to help select the best model representation for pure liquids of AE, ED and EG and their aqueous solutions.

Gas-phase simulation results for thermodynamic and structural properties (i.e. the average bond angle and dihedral angle energies, $<U_a>$ and $<U_d>$, respectively, the average Coulombic and Lennard-Jones energies, $<U_Q>$ and $<U_{LJ}>$, respectively, and the total average configurational energy, <U>, averaged and final values of dihedral angles) for different models of AE, ED and EG at 298 K are given in the Tables 7 and 8. In accord with experimental and theoretical findings (see Chapters II and III for details) the lowest energy conformations in the gas phase are g'Gg' (or gG'g), g'Gt and gGg' for AE, EG and ED, respectively. To allow comparison with these findings the final conformational arrangements for each model are also included in Tables 7 and 8.

The applicability of different force fields, in particular Jorgensen's AMBER/OPLS-based [52,120] and a "combined" model potential [6,52], CHARMM [115,116], and the special force field by Alejandre *et al.* [67], have been tested with the corresponding equilibrium geometries in single molecule simulations of AE (denoted as the AEeq, AEcmb, AEchm, and AEtst models, respectively, in Table 7). Comparing data from Tables 1 (for experimental geometries), 3 (for equilibrium geometric parameters) and 7 one can clearly see that both the AEeq and AEtst models give very good performance in terms of reproducing the corresponding gas-phase geometry, while the CHARMM potential causes noticeable distortion of the NCC and OCC bond angles, and as a result of this generates an eclipsed conformation with respect to the central torsion angle. AE was also used to examine the possible influence of molecular model representation (not shown in Table 7). No noticeable difference in the average molecular geometry was found between rigid and flexible molecular models, but there was a small (~1 kJ/mol) change in the average intramolecular configurational energy.

Table 7. Basic gas-phase simulation results for different models of AE at 298 K. Units for energy are kJ/mol and all angles are given in degrees.

	AEop	AEeq	AEcmb	AEtst	AEchm
$<U_a>$	4.78	4.34	2.40	1.71	19.11
$<U_d>$	3.14	2.94	7.90	-3.14	1.20
$<U_Q>$	60.59	60.86	72.97	165.05	128.36
$<U_{LJ}>$	3.04	2.87	0.94	0.14	14.30
$<U>^a$	71.55	71.01	84.21	163.76	162.98
Aver. Angles					
NCC	108.1	109.1	111.2	118.7	124.1
OCC	107.6	105.9	107.9	108.9	119.5
HNC	110.5	109.0	109.3	110.5	110.0
HOC	102.2	102.9	104.7	105.7	104.5
Aver. Dihedrals					
NCCO	40.91	41.18	55.57	76.85	2.37
HOCC	22.68	23.02	40.11	42.82	40.58
HNCC	170.17	171.78	169.42	175.85	174.38
Dihedrals[b]					
NCCO	-41.5	38.27	45.83	-83.4	0.02
HOCC	22.73	-16.40	-33.7	46.5	-47.00
HNCC	-179.7	-171.2	-173.8	179.8	-168.2
Conformation[b]	**gG'g'**	**c'Gg'**	**g'Gg'**	**gG'g**	**g'Cg'**

a - the largest error associated with total configurational energies is about 0.1 kJ/mol;
b - dihedral angles and conformations corresponding to the final configuration.

Table 8. Basic gas-phase simulation results for different models of EG and ED at 298 K. Units for energy are kJ/mol and all angles are given in degrees.

	EGex1	EGex2	EGeq	EGaa	EDex	EDop
$<U_a>$	1.14	1.52	1.17	18.97	2.54	2.49
$<U_d>$	9.00	10.22	9.20	10.92	-9.07	-9.27
$<U_Q>$	82.31	77.43	82.46	-196.63	57.08	58.59
$<U_{LJ}>$	0.49	0.54	0.36	154.37	1.10	0.98
$<U>^a$	92.95	89.72	93.20	-12.37	51.65	52.78
Aver. Angles						
OCC, NCC	109.7	108.9	109.8	104.7	111.8; 110.9	109.8; 114.6
HOC, HNC	107.8	101.2	108.3	109.7	103.8; 108.6	103.1; 109.2
HOC, HNC	104.7	104.8	105.2	103.8	108.9; 109.1	109.7; 110.9
Aver. Dihedrals						
OCCO, NCCN	57.43	55.33	57.43	54.62	50.79	51.87
HOCC, HNCC	177.29	172.88	175.19	173.80	174.04	172.16
HOCC, HNCC	50.85	47.66	50.98	43.66	44.00	44.93
Dihedrals[b]						
OCCO, NCCN	58.51	57.89	57.87	48.75	52.20	51.2
HOCC, HNCC	-52.00	-51.4	-55.6	-36.4	-40.80	-46.5
HOCC, HNCC	178.80	-168.6	-173.3	173.9	-174.4	-173.2
Conformation[b]	**g'Gt**	**g'Gt**	**g'Gt**	**g'Gt**	**gGg'**	**gGg'**

a - the largest error associated with total configurational energies is about 0.02 kJ/mol;
b - dihedral angles and conformations corresponding to the final configuration.

United and all-atom AMBER/OPLS molecular representation were tested on two models of EG (in Table 8 EGeq and EGaa, respectively). Only a very subtle discrepancy in the average geometrical parameters for both these models was noted while the configurational energy contributions were significantly rearranged. The rearrangement mainly involved the energy contributions from bond angle and intramolecular interaction (both Lennard-Jones and electrostatic) terms. The latter fact can be appropriately evaluated when gas-phase configurational energies are compared with corresponding values from liquid phase simulations.

In the test of the different model potentials the most important observation is related with the reproducibility of the lowest energy conformation for each compound of interest. There are two complementary ways to clarify this issue. Firstly, the most stable conformation (obtained experimentally or by *ab initio* methods) is used as an initial input for these classical vacuum simulations and, in the case of an appropriate choice of the force field, the final conformation is expected to be essentially (within acceptable error) unchanged. Secondly, when the most successful force field has been identified, the simulation run is repeated starting from an arbitrarily chosen conformational arrangement to prove the ability of the potential in again generating the lowest energy conformer. From Tables 7 and 8 one can see the results of the application of this procedure. Relying on average values for the dihedral angles all potentials (with the exception of CHARMM) reproduce the initial arrangement quite well. There are, however, slight variations in the final configuration. In the case when the special AE force field was employed (AEtst), the optical isomer of the most stable conformation of AE with desirable values for the dihedral angles was produced, while for the AMBER/OPLS-based models (AEop and AEeq) the value for the HOCC angle was definitely lowered (in comparison with the conventional value

for a *gauche* position). Model AEcmb was designed to test a "combined" torsion potential where the Fourier coefficients for the NCCO and HOCC torsion angles were taken from a force field for liquid alcohols [6] and those for the HNCC angle came from a model potential for amines [52]. At the present level of analysis it appears that AEcmb gives a fairly good representation of an AE molecule in the gas phase. In particular, it is the only model among all those considered for AE, which exactly reproduces its most stable conformation (g'Gg') with a reasonable average value (55°) for the NCCO torsion angle.

An analysis of the gas-phase simulation results obtained from models developed using different geometric parameters also revealed some interesting trends. The following three groups of geometry-dependent models were considered: those based exclusively on experimental parameters such as EGex1 (gas-phase data by Caminatti and Corbelli [74]), EGex2 (liquid-phase structure by Buckley and Giguere [79]), EDex and AEex (gas-phase measurements by Marstokk and Mollendal [36] and Penn and Curl [23], respectively), models based on optimized structures such as AEop and EDop, and models based on AMBER/OPLS equilibrium parameters such as AEeq and EGeq. In all cases these model molecules were treated as rigid bodies using a united atom representation and the AMBER/OPLS force field. The final structural characteristics for the EGex1, EGex2 and EDex models nicely correlate with the original experimental data (see Table 1). Both the EGex1 and EGex2 models predict very similar geometries as well as thermodynamic properties (see Table 8). Entirely analogous behavior is observed for EDex and EDop characteristics.

In contrast to EG and ED, the experimental-geometry-based model for AE (not shown in Table 7) appeared to be a special case. At an early stage of the simulation process the AEex model exhibited a totally distorted structure and calculations were automatically interrupted. Following optimization of this

experimental structure with a modest basis set (6-31+G(d)) at the MP2 level of theory, this situation was significantly improved; simulations performed for AEop exited normally and final geometric parameters were in good agreement with data obtained from the tests for AEeq and AEcmb. Overall structural parameters for both AMBER/OPLS-based models appear to be lowered in comparison with those for AEcmb as is apparent in the magnitudes of bond and dihedral angle energy contributions. However in the absence of reliable experimental data it is difficult to make a definite conclusion about relative quality of calculated geometries for an isolated molecule of AE.

It was expected that equilibrium-geometry-based models would give results very similar to those observed for those models based on high quality experimental data; all the characteristics for EGeq, when compared with EGex1 and EGex2, confirm this suggestion. It is remarkable that all the models presented in Table 8 reproduce fairly well the most stable conformations of EG and ED, indicating a consistently adequate combination of the force field and model geometry.

One might argue that employing the experimental gas-phase structural parameters as input for the liquid phase simulations is, in principle, not an optimum choice. Unfortunately, liquid phase experimental geometric parameters for these compounds are either incomplete and augmented by unreliable theoretical assumptions (e.g. the experimental structure of AE) or simply unavailable. Moreover, as mentioned earlier, there is some experimental evidence for a trend among the compounds investigated to prefer the same conformation in the liquid state as in the gaseous state.

5.2 LIQUID-PHASE SIMULATIONS

The analysis of gas-phase simulations results has shown that the information obtained is essential but not sufficient for a critical assessment of the quality of the molecular models and interaction potentials. On the basis of gas-phase results only two models, namely AEchm (showing an entirely wrong conformation with respect to the central dihedral angle) and EDop (giving essentially the same results as the model EDex) were excluded from calculations of liquid AE, ED and EG. At this point it became clear that for future calculations of aqueous mixtures there was still a need to impose additional criteria which would allow the elimination of at least half of the models in the present set.

5.2.1 Criteria for selecting liquid-phase models

The criteria chosen for selecting liquid-phase models were the liquid density and heats of vaporization. The density, ρ, is available from a constant-temperature, constant-pressure MD simulation as the total mass of the N molecules divided by the average volume of the periodic box [133]:

$$\rho = \frac{mass_N}{< volume >_{box}}. \qquad (5.1)$$

ED has been chosen as a representative density test for all AMBER/OPLS-based models. Its liquid density (0.9443 g/cm^3) obtained from a constant-pressure MD run for the EDex model, for example, was in reasonable agreement (within ~5%) with the experimental value of 0.8990 g/cm^3 [135] confirming the applicability of the AMBER/OPLS force field to the compounds of interest.

Another criterion, the (molar) heat of vaporization, can be determined using eqn. 4.44. The different contributions to the total configurational energies for the selected liquid-phase models are given in Table 9, as well as the

calculated (by substituting thermodynamic data from Tables 7, 8 and 9 into eqn. 4.44) and experimentally determined heats of vaporization. Special note should be made with respect to the experimental values. In particular, in the contemporary scientific literature one can find at least three different values namely 57.07 [144], 58.71 [145] and even 65.6 [146] kJ/mol, for ΔH_{vap} of EG at 298 K. The latter value is the most recent (1988) estimate deduced from direct calorimetric measurements of alkanediols [146-148]. For AE the only experimental value of ΔH_{vap} was measured at its boiling temperature (T_b=444 K) [149] making a direct comparison with data obtained at 298 K difficult; however it is clear that this value at 444 K should be somewhat lower than the value at 298 K.

Table 9. Coulombic and Lennard-Jones contributions to the total configurational energy and heats of vaporization of liquid EG, ED and AE at 298 K.

	EGex2	EGaa	EDex	AEcmb	AEtst
Intra.					
$<U_Q>$	88.17	-194.30	61.29	81.11	169.49
$<U_{LJ}>$	0.89	161.65	1.68	0.87	-0.16
Inter.					
$<U_Q>$	-60.38	-34.35	-35.54	-48.16	-36.80
$<U_{LJ}>$	-8.65	-16.15	-15.98	-13.62	-16.37
Total					
<U>	29.21±0.04	-60.79±0.02	19.32±0.03	34.67±0.03	128.36±0.03
ΔH_{vap}	63.04±0.07	50.92±0.03	34.83±0.05	52.04±0.05	37.90±0.03
Experim.					
ΔH_{vap}	57.0-65.6[a]	57.0-65.6[a]	44.98[b]	49.83[c]	49.83[c]

(Units for energy and ΔH_{vap} are kJ/mol; a - [144-146]; b - [145]; c - [149]).

From Table 9 one can see that the heat of vaporization calculated for the AEtst model is in rather poor agreement with the experimental value while AEcmb gives excellent performance. It can also be noted that there is a marked discrepancy between the experimental and calculated results for the EDex and EGaa models. The EDex model has the lowest magnitude ΔH_{vap} among the all models presented in Table 9.

An analysis based mainly on values for the heat of vaporization allowed for highlighting and elimination of some of the least successful models for each compound. In particular, model EGaa (with its rather large error in ΔH_{vap}) and EGex2 (with a similar magnitude error in ΔH_{vap} as for model EGex1 but with a very undesirable self-diffusion coefficient, results for which will be discussed below) were removed from further consideration. Hence, on this stage only one model (EGex1) remained for EG; for ED the EDex model was chosen (in spite of its large error in ΔH_{vap}) since no obvious better alternative was available. For AE three models were chosen for a further consideration: both the AEeq and AEcmb models which have almost equally good basic static and dynamic properties and AEtst as the test-representative of the specifically designed model potential.

In force field development [132,134] one of the ways for possibly improving the agreement between calculated and experimental data is by adjusting parameters in the Coulomb and Lennard-Jones energy terms for the so called 1-4 interactions (see Chapter IV). This can be technically achieved by using appropriate scaling factors. All the data reported in Tables 7, 8 and 9 (with the exception of the AEtst model) were obtained with scaling factors of 0.833 and 0.125 for the electrostatic and Lennard-Jones (LJ) non-bonded interactions, respectively. It has been previously shown that the 1-4 LJ scale factor of 0.125 should be used in conjunction with the AMBER/OPLS force field [134] and in the

study of simple alcohols, amines and EG [4,132] it was found that a 1-4 electrostatic scale factor of 0.833 gives the best model representation. However, for the present systems the heat of vaporization appears not to be the only parameter which should be taken into consideration when scaling is applied. In the case of 1,2-disubstituted ethanes, 1-4 interactions strongly influence the molecular conformation. Hence the major population of rotational isomers and the self-diffusion coefficient, D, in liquid-phase simulations are rather sensitive to the values used to scale these non-bonded parameters. In this thesis five different pairs of values for 1-4 scaling were tested, in particular, (1) 0.25 and 0.25, (2) 0.5 and 0.5, (3) 0.833 and 0.125, (4) 0.125 and 0.833 (as discussed above), as well as (5) 1 and 1 (i.e. no scaling used) for van der Waals and electrostatic terms, respectively. The tests for the first three pairs were performed on the EGex1 model and the poor results obtained confirmed their inapplicability. For all three pairs the values for the self-diffusion coefficient appeared to be the major problem. For example, the values of D for pairs (2) and (3) were four and two times, respectively, lower than the experimental estimate and for pair (1) D was overestimated by almost two times while the average value of the OCCO torsion angle were dramatically shifted towards the *cis* conformation. Therefore, only two pairs of scale factors, namely (4) and (5), were chosen for further analysis in the most promising models.

5.2.2 Dihedral angle distributions

At this stage the conformational characteristics (dihedral angle distributions) will be examined for the models EGex1, EDex, AEeq (henceforth EG, ED and AE, respectively) considered as the basic models, as well as AEcmb and AEtst as test models. For each of the models EG, ED and AE two separate liquid-phase simulation runs were carried out with the standard scale factors

(0.125 for Lennard-Jones and 0.833 for electrostatic) and no scaling (1 and 1, respectively) used. Only one (the former) scaling scheme was used for AEcmb and the latter for the AEtst model.

The dihedral angle distributions obtained for all models are shown in Figs. 10-13. It can be seen from Fig. 10a that for EG when 1-4 scaling is used the major population of rotamers (99%) is in the *gauche* conformation with respect to the OCCO dihedral angle with the maximum probability of the distribution centered at 57.5°. Distributions for both HOCC torsions exhibits three probability maxima, centered at -82.5, 77.5 and 187.5° for one angle while the other torsion has maxima at -82.5, -32.5 (these two peaks overlap in Fig. 10a) and 57.5°. In contrast, for the EG model simulated without scaling of the 1-4 interactions (see Fig.10b) all molecules are in the *trans* conformation (182.5°) with respect to the central dihedral angle and both distributions for the HOCC angle demonstrate a similar pattern with maxima located at about -80 and 185°.

The ED model simulated with 1-4 scaling (Fig. 11a) has two peaks in the NCCN dihedral angle distribution located in the G (83.4% of conformers) and G' (16.6% of conformers) regions. As it was expected one pair of HNCC torsions (from one end of the molecule) has similar distributions with three maxima corresponding to two *gauche* (about -70 for g' and 77.5 and 47.5° for g) and one *trans* (about 185°) conformation. Another pair of HNCC torsions has distributions with two maxima centered at -62.5 (for both HNCC angles) and at 52.5 or 172.5°. Again a similar trend of generating a significant population (64.2%) of *trans* rotamers with respect to the central torsion angle was observed for ED simulated without 1-4 scaling (see Fig. 11b). In addition, two types (g and g') of *gauche* conformers were present on this distribution with an overall population of 35.8%. Each distribution for HNCC angles in Fig. 11b has two maxima located in *gauche* (g and/or g') and/or *trans* positions.

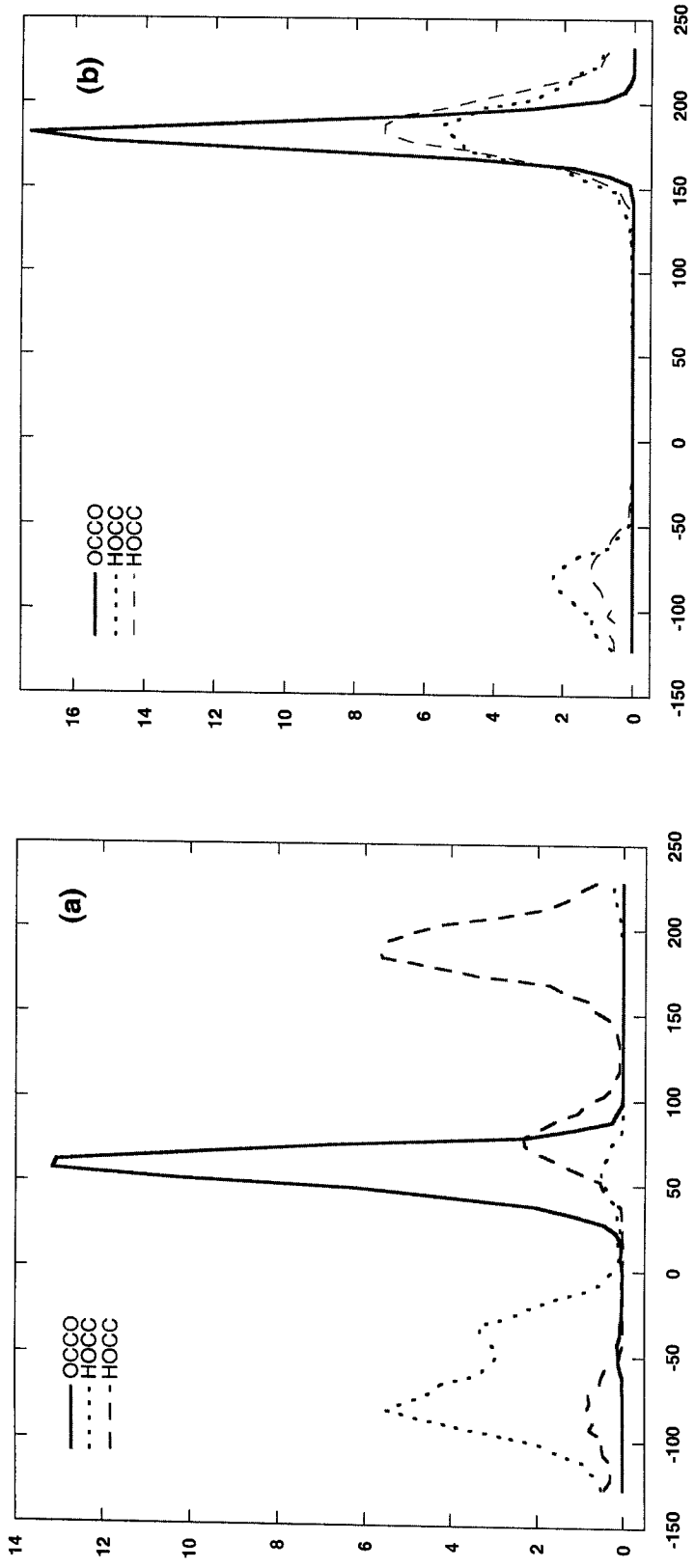


Figure 10. Dihedral angle distributions for EG: (a) with scaling factors for 1-4 interactions of 0.125 and 0.833 for the Coulomb and Lennard-Jones terms, respectively; (b) with no scaling factors used.

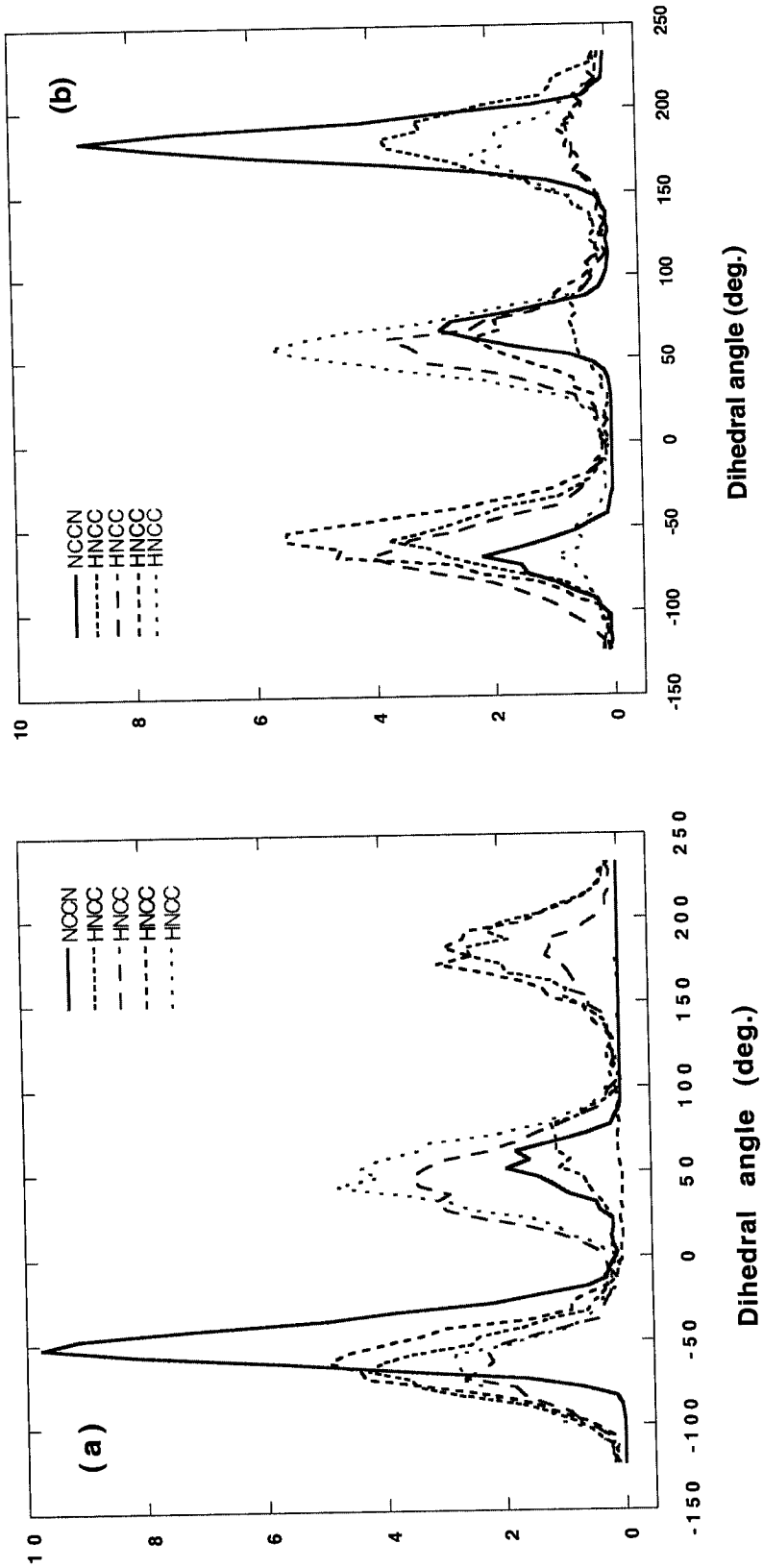


Figure 11. Dihedral angle distributions for ED: (a) with scaling factors for 1-4 interactions of 0.125 and 0.833 for the Coulomb and Lennard-Jones terms, respectively; (b) with no scaling factors used.

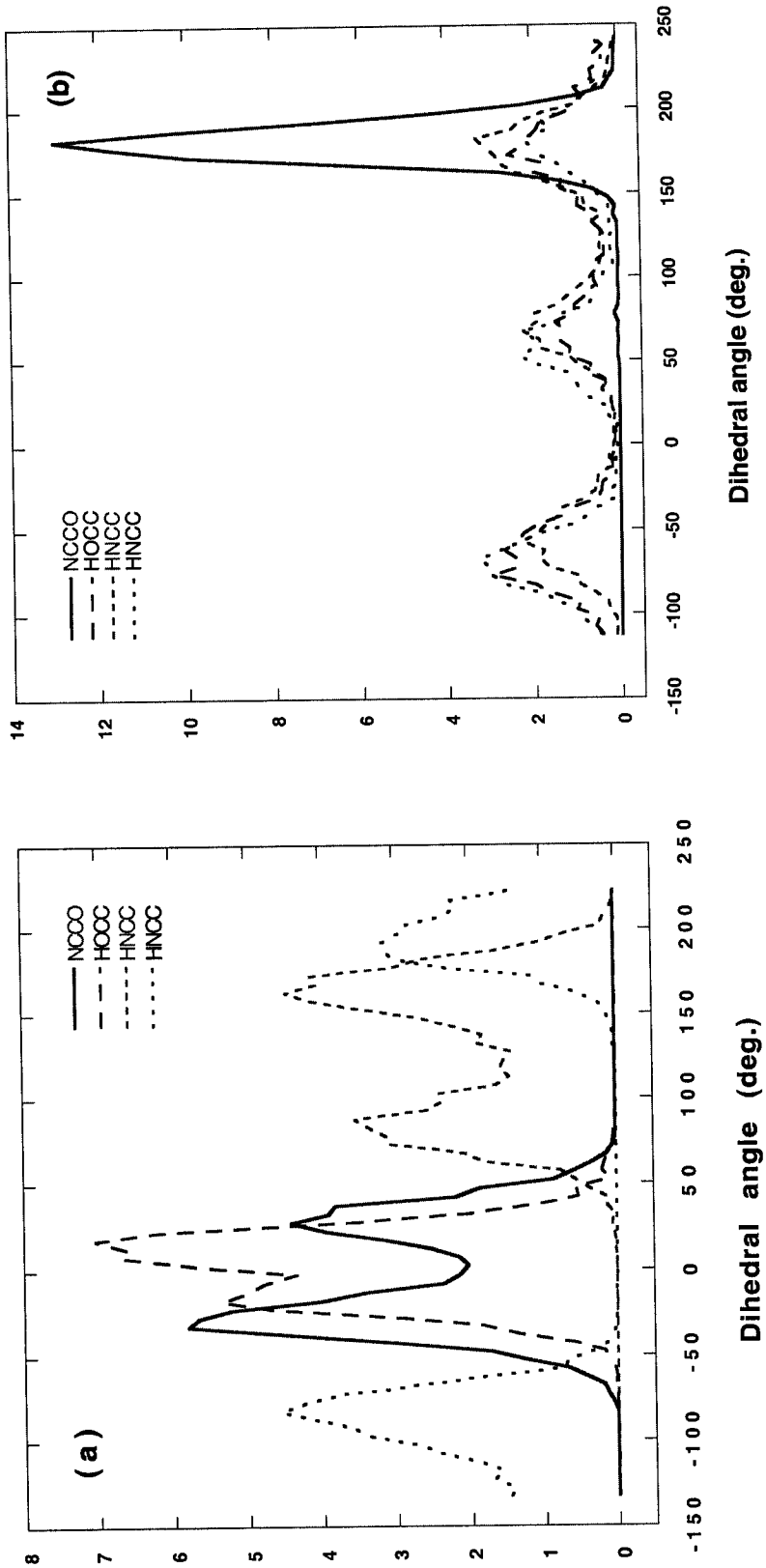


Figure 12. Dihedral angle distributions for AE: (a) with scaling factors for 1-4 interactions of 0.125 and 0.833 for the Coulomb and Lennard-Jones terms, respectively; (b) with no scaling factors used.

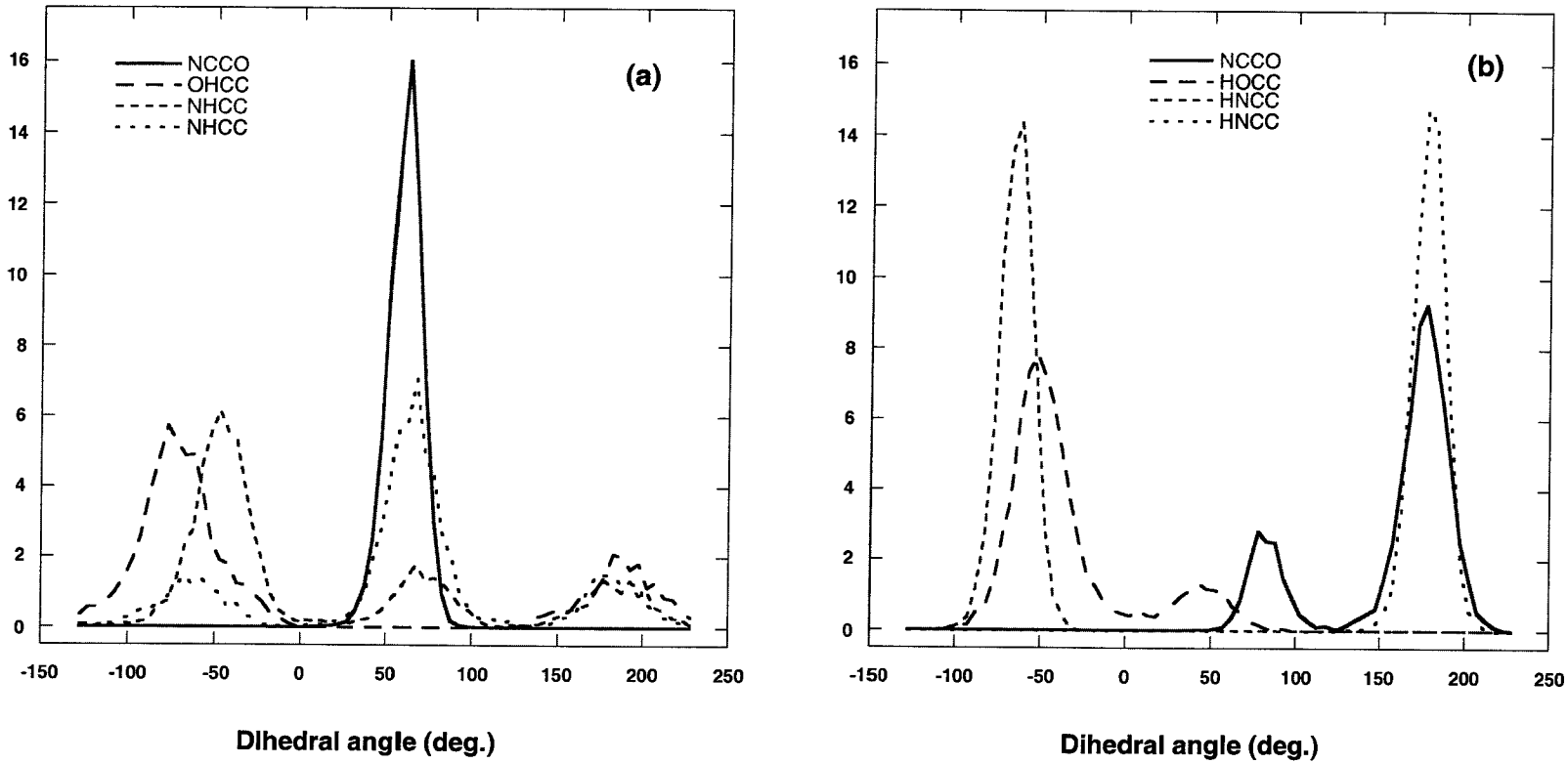


Figure 13. Dihedral angle distributions for AE: (a) model AEcmb (scaling factors are 0.125 and 0.833); (b) model AEtst (no scaling factors used).

Distributions for all three dihedral angles of AE obtained using appropriate scale factors exhibit the most complicated behavior (Fig. 12a). The distribution of the NCCO dihedral angle has almost equal population (54.4 and 45.6%) of conformers located at -32.5 and 27.5°, respectively, which indicates (analogously to the situation for the gas-phase) a shift towards the *cis* position. In addition it should be noted that these two peaks are significantly overlaped. This trend is also clearly seen in the distributions for the OHCC dihedral angle (virtually one broad peak corresponding to a *cis* conformation with maxima at -17.5 and 17.5°) and one of the HNCC angles (visible overlap of two peaks located at 87.5 and 162.5°). In contrast, all peaks on the distributions shown in Fig. 12b (no scaling used) are nicely distinguishable and appear at the expected locations corresponding to two types of *gauche* (at about ± 60°) and a *trans* (170-180°) populations of conformers for this compound.

Finally the torsion angle distributions for two test models of AE were examined in order to understand better the specificity of the observed conformational behavior. From the probability distribution of the NCCO dihedral angle for the AEcmd model shown in Fig. 13a one can see a single intense peak corresponding to a 100% population of *gauche* rotamers. In contrast the same distribution for the AEtst model indicates the presence of both *gauche* (20%) and *trans* (80%) conformers (see Fig. 13b). In both models the distribution of OHCC dihedral angles exhibits two maxima, for AEcmd at -77.5 and 182.5° and for the AEtst model at -52.5 and 42.5°. The corresponding distributions of the HNCC torsion angles also differ markedly. For the AEcmb model there are three maxima at about -50.0, -70.0 and 180.0°, while the distributions for AEtst has a peak at -62.5 and another peak at 177.5°. Such a diversity of conformational patterns of AE leads one to the conclusion that this compound will require special consideration during the further analysis.

5.2.3 Results and final choice of models

The results of simulations carried out with and without scaling of the 1-4 non-bonded interactions on the basic liquid-phase models of EG, ED and AE are summarized in Table 10. One can clearly see that for virtually all compounds the conformation of molecules with respect to the central dihedral angle changes from *gauche* to *trans* (a mixture of both kinds of rotamers are present in the case of ED) upon removal of scaling. These results can be interpreted by taking into account the information given in Table 2. It has been shown experimentally that *gauche* (G) is the most abundant conformation for EG [77,78] and AE [27] in the liquid state, which corresponds to results obtained when 1-4 scaling is used. For ED the experimentally determined conformation (G+T or T) [2] is in excellent agreement with our observations from the scaling-free run. The populations of conformers of these models were calculated and these are presented in Table 11.

As the next step, the experimental values for the heat of vaporization are compared with calculated parameters. From Table 10 one can see a trend of increasing magnitudes of ΔH_{vap} for all three compounds comparing conventional and scaling-free regimes. Interestingly, for both AE and EG the latter values of 57.99 and 67.4 kJ/mol appear to be overestimates when compared with the experimental results of 49.83 [149] and 58.71 kJ/mol [145], respectively. In contrast, for ED the value obtained from the scaling-free run (39.96 kJ/mol) is closer to the experimental value (44.98 kJ/mol) [145].

Comparison of estimates of the self-diffusion coefficient, D, reveals a clear trend of decreasing magnitude in this quantity for the scaling-free simulation runs. For AE, D becomes approximately 3.5 times smaller with no 1-4 scaling (unfortunately there is no experimental estimate available); for EG the decrease is even more pronounced where the calculated value of 0.012×10^{-5} cm^2/s is

Table 10. Selected liquid phase simulation results for AE, EG and ED at 298 K.

	AE		EG		ED	
Scale factor	f=0.833;125	f=1; 1	f=0.833;125	f=1; 1	f=0.833;125	f=1; 1
Conformation	G	T	G	T	G	G+T
Experiment	G		G		G+T/T	
ΔH_{vap}	45.05±0.21	57.99±0.03	53.37±0.04	67.40±0.08	34.83±0.05	40.01±0.03
Experiment	49.83		57.0-65.6		44.98	
D	0.52	0.15	0.08	0.012	1.16	1.0
Experiment	-		0.09[a]		0.98[b]	

(Units for dihedral angles are degrees, for D - $10^{-5}cm^2/s$, for ΔH_{vap} - kJ/mol; a - [58]; b - [150].)

Table 11. Conformational characteristics of pure liquids.

	EG		ED			AE	
Conformation	G'	G	T	G'	G	G'	G
(%)	0.9	99.1	64.2	16.2	19.6	54.4	45.6
Dihedral	-42.5	57.5	182.5	-67.5	67.5	-32.5	27.5

(Units for dihedral angles are degrees and all values for dihedral angles correspond to the maximum probability of dihedral angle distributions).

7.5 times lower than the experimental one (see Table 10). Somewhat surprisingly, the diffusion coefficient of ED decreases only slightly and achieves excellent agreement with experimental findings [150] with no scaling applied.

The results obtained for two test models of AE (not shown in Table 10) can be summarized as it follows. The AEtst model gives rather poor performance for all quantities considered, in particular, the wrong conformation (see also Section 5.2.2 for details) and a very low magnitude for the heat of vaporization (37.88 kJ/mol). The self-diffusion coefficient for AEtst was found to be 0.35×10^{-5} cm^2/s. The results for AEcmd can be classified as more satisfactory. This model exhibits the *gauche* conformation for the NCCO dihedral angle, has a heat of vaporization of 51.98 kJ/mol and D of 0.22×10^{-5} cm^2/s, these results being comparable to those obtained for the basic AE model. The absence of an experimental result for D for AE and the fact that the only experimental heat of vaporization available corresponds to the boiling temperature of AE does not allow one to make a completely justified choice between these two model representatives.

As a result of the present analysis the basic EG model simulated using the standard scaling scheme [132,134] was chosen as the most realistic one. It is also clear that all characteristics of the basic ED model obtained with scaling applied were dramatically improved in the scaling-free simulation run, which makes the latter model the superior choice for the further investigation. In the case of AE the final choice between the basic AE and AEcmd models was made on the basis of accumulated experience with the modeling of this class of compounds. The basic AE model simulated with standard scaling was chosen for the further structural investigation and for the simulations of aqueous solutions. The major thermodynamic characteristics, in particular, Coulombic and Lennard-

Jones contributions to intra- and intermolecular energies, as well as the total configurational energies for the models of choice are given in Table 12.

Table 12. Thermodynamic characteristics of liquid EG, ED and AE at 298 K.

	AE	EG	ED	Max. error
$<U_a>$	14.99	6.02	11.78	-
$<U_d>$	5.25	11.69	-7.17	-
Intra.				
$<U_Q>$	59.46	90.71	148.67	0.20
$<U_{LJ}>$	6.22	0.53	-0.31	0.05
Inter.				
$<U_Q>$	-41.65	-48.80	-36.94	0.57
$<U_{LJ}>$	-15.60	-12.24	-16.28	0.03
<U>	**28.46**	**42.08**	**102.33**	**0.12**

(Units for energy are kJ/mol.)

5.2.4 Structural analysis of pure liquids

The structural analysis of pure EG, ED and AE (and their aqueous solutions) was performed using both radial and spatial distribution functions (RDF's and SDF's, respectively). In spite of the fact that RDF's are functions of interparticle separation only and therefore ignore the orientation of the local frame (see discussion in Chapter III), they are still widely used because of their ability to provide some insights into the immediate environments of those atoms interacting strongly with other molecules (atoms). Hence in strongly associated liquids the most interesting and informative RDF's are usually those due to the atoms involved in H-bond formation. The RDF's between several site pairs for pure liquid EG, ED and AE are shown in Figs. 14,15 and 16, respectively. The corresponding coordination numbers were calculated using eqn. 4.45 and results

are given in Table 13. The values in Table 13 for EG are for *inter*molecular coordinated neighboring atoms only; results for ED and AE include intramolecular coordination.

Table 13. Coordination numbers for pure EG, EG and AE.

EG		AE	
O-O	2.0	O-O	0.7
O-H	0.5	N-N	1.3
		N-O	2.5
ED		O-H(N)	1.0
N-N	3.6	O-H(O)	0.3
N-H	1.0	N-H(N)	0.4
		N-H(O)	0.7

The oxygen-oxygen (O-O) RDF for EG (Fig. 14) shows a sharp first peak at 2.8 Å which drops into the narrow minimum followed by a second broad peak at about 5.5 Å, which is somewhat reminiscent of the corresponding RDF of liquid methanol [42]. The O-H RDF has two peaks one at 1.8 Å, due presumably to H-bonding and a second less well defined maximum at 3.4 Å. These results for the RDF's of EG are in very good agreement with those of Saiz *et al.* reported for two OPLS-based models of EG [4]. Integration of the first peak of $g(r_{OH})$ gives a coordination number of 0.5 (due to intermolecular hydrogens) indicating that each oxygen atom of EG accepts on average only one strong H-bond, where there is an equal likelihood that this coordination is due to intra- and intermolecular neighboring hydrogens. In addition, the coordination number from the first peak in O-O RDF suggests the presence of two intermolecular neighboring oxygens around each oxygen atom. Clearly, one of these oxygen atoms does not appear to participate in a (strong) H-bond with the central atom.

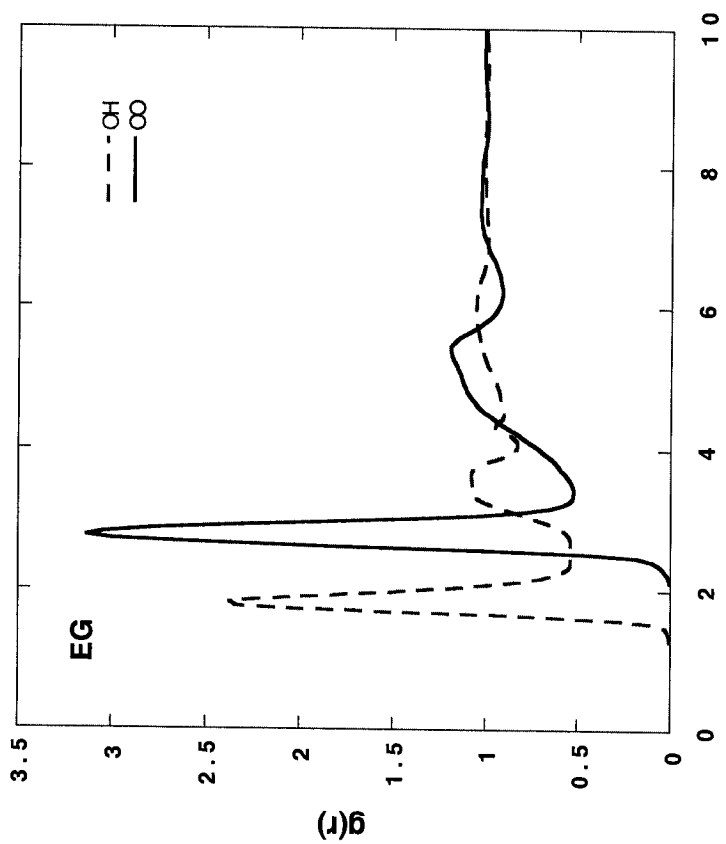


Fig. 14.
Radial distribution functions for pure EG at 298 K. The solid and dashed lines represent the oxygen-oxygen and oxygen-hydrogen site-site RDF's, respectively.

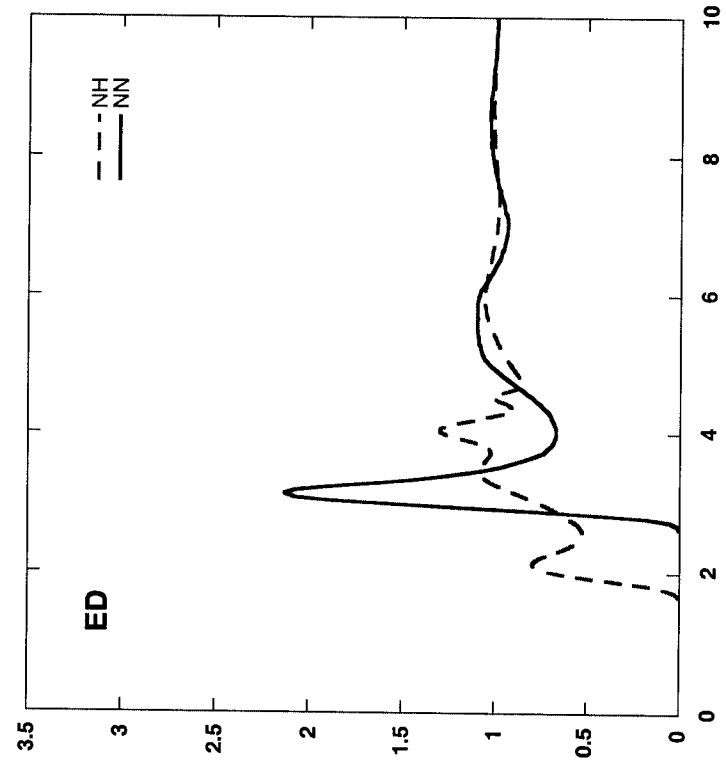


Fig. 15.
Radial distribution functions for pure ED at 298 K. The solid and dashed lines represent the nitrogen-nitrogen and nitrogen-hydrogen site-site RDF's, respectively.

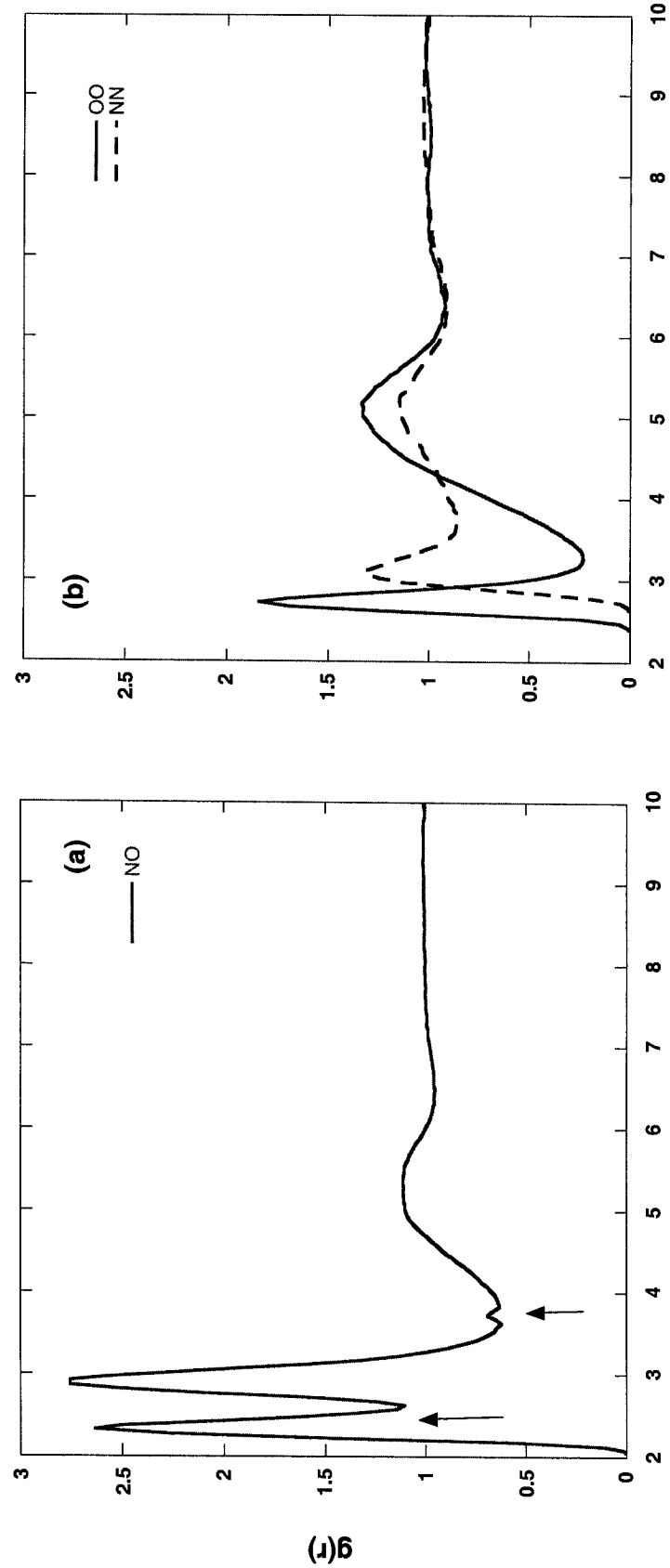


Fig. 16. Radial distribution functions of pure AE at 298 K: (a) nitrogen-oxygen (arrows indicate peaks due to intramolecular features); (b) oxygen-oxygen and nitrogen-nitrogen RDF's (solid and dashed lines, respectively).

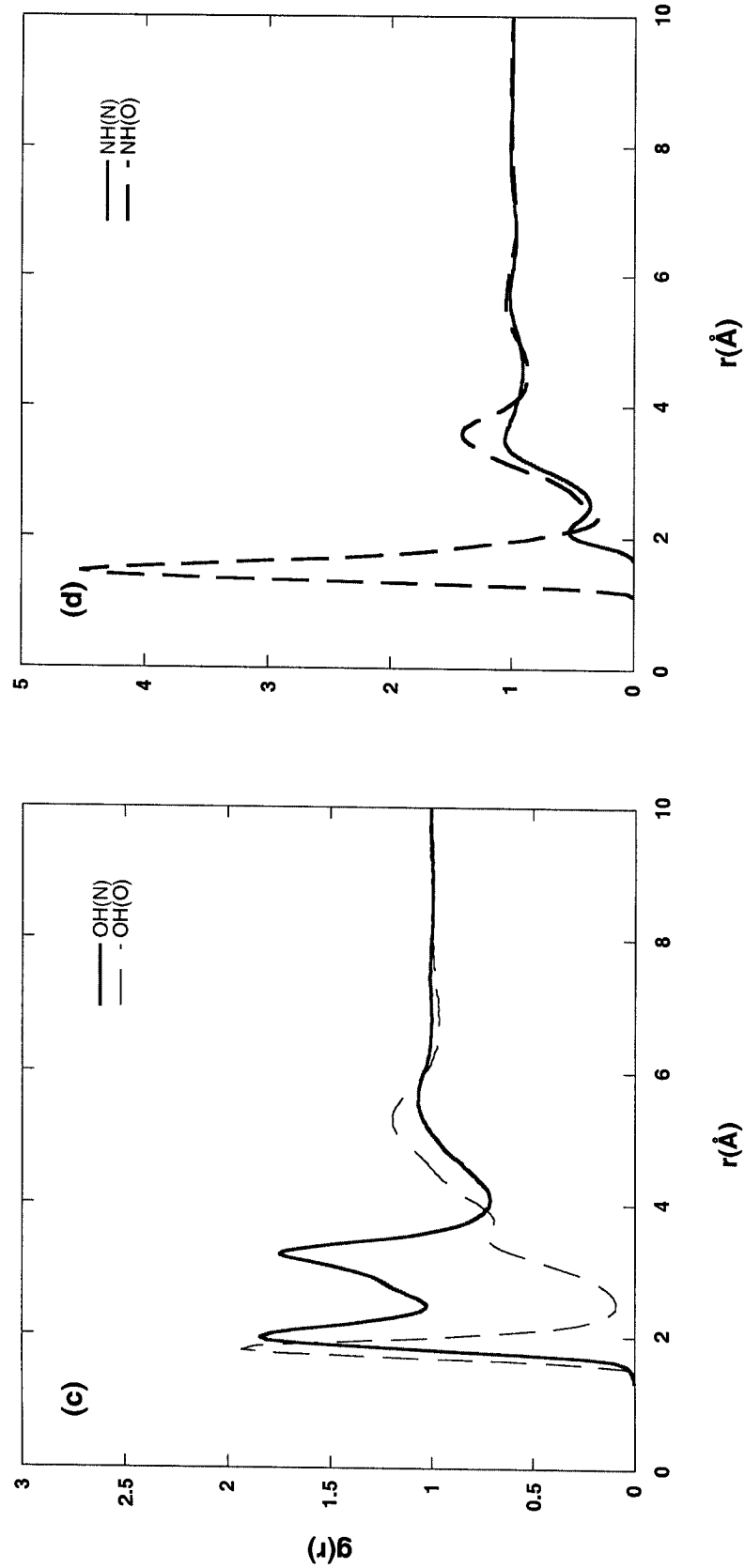


Fig. 16 continued. (c) Oxygen-hydrogen and (d) nitrogen-hydrogen site-site RDF's.

From the present RDF analysis one can see that although EG possesses a total of six H-bonding sites per molecule, the constraint of H-bond balance (the number of donors must equal the number of acceptors) dictates that two of the acceptor sites remain unutilized in spite of the apparent presence of two (extra) near neighboring oxygens. The SDF analysis below will be used to provide more insight into the nature of the near-neighbor coordination.

The RDF's of ED for N-N and N-H site pairs is shown in Fig. 15. The N-N RDF exhibits a still well defined but not intense first peak at 3.1 Å followed by a shallow minimum at 4 Å and a broad second peak at about 5. 5 Å. The N-H RDF has a small first peak at 2.1 Å due to relatively weaker (in comparison to EG) H-bonded neighbors, as well as a set of quite complicated peaks corresponding to secondary structure (one can distinguish a maximum at 4.0 Å located between two shoulders at 3.5 and 4.4 Å). Unfortunately, neither theoretical nor experimental RDF's of ED were found for the comparison in the available literature.

Analysis of the coordination numbers for liquid ED requires that we recall that approximately 65% of molecules adopt a *trans* conformation and hence only one third of all rotamers may possess an intramolecular H-bond. Integration of the first peaks in $g(r_{NH})$ and $g(r_{NN})$ gives coordination numbers of 1.0 and 3.6, respectively. One can reasonably assume that one third of hydrogens H-bonded to nitrogen are within the same molecule. Taking into account that ED has four H-bond donor and two H-bond acceptor sites per molecule one can suggest that each molecule makes on average four (relatively strong) H-bonds while two of the donating sites appear not to participate (on average) in strong H-bonding. As complementary evidence for this the N-N coordination number indicates the presence of at least two non H-bonding (or weakly H-bonding) nitrogens located in the first coordination shell of the ED molecule.

Seven selected RDF's between oxygen, nitrogen and their associated hydrogen atoms for pure AE are presented in Fig. 16. The first peak in the nitrogen-oxygen RDF (Fig. 16a) can be identified as one due to the intramolecular H-bonded oxygen and nitrogen (with an interatomic distance of 2.3 Å); the second distinct peak (at 2.9 Å) is due to intermolecular nearest neighbors and a small shoulder at 3.7 Å is again an intramolecular feature corresponding to the trace amount of *trans* conformer. The oxygen-oxygen and nitrogen-nitrogen functions are compared in Fig. 16b. The O-O RDF exhibits a well defined, sharp peak at 2.7 Å which drops into the deep minimum followed by broad but distinct peak at 5.2 Å due to second neighbors. In contrast the N-N RDF has a short and broadened first peak at 3.1 Å followed by similarly broad second maximum (5.2 Å). Four RDF's of the O-H and N-H types (with H attached to N and H attached to O) are shown in Figs. 16c and 16d. Both O-H(O) and N-H(O) RDF's have large and well defined first peaks at 1.8 and 1.5 Å, respectively, indicating the presence of strong H-bonding between the corresponding sites (contributions due to intra and intermolecular near neighbors do overlap in this case). The shortened interatomic distance (H-bond) of 1.5 Å is presumably a result of the small separation (almost *cis* conformation) between nitrogen and oxygen atoms also present in an isolated AE molecule. In the O-H(O) function the second peak is less well defined and has a little shoulder at 3.5 Å while the second peak in N-H(O) has a distinct maximum at 3.6 Å. The O-H(N) has two broader peaks at 1.9 Å (indicating the presence of a weaker H-bond) and at 3.2 Å (whose identity is not resolvable from the RDF). The less pronounced (appearing more as a shoulder) H-bonded peak observed at 2.1 Å in the N-H(N) RDF suggests that the nitrogen donor sites trend to form rather weak H-bonds with other N atoms. Finally, it is worth mentioning that the four O-H and N-H RDF's (with some exceptions for O-H(N) which can be accounted for by the

difference in the model potentials used) are all in good qualitative agreement with the results reported by Alejandre *et al.* [67].

The coordination numbers of AE listed in Table 13 reveal some interesting trends. For instance, the total coordination number of two (0.7 and 1.3) for both nitrogen and oxygen atoms (when they are considered as separate sites) is smaller than the 2.5 value for $g(r_{NO})$ which indicates a preference for unlike-atom coordination. Analysis of coordination numbers from N-O together with the O-H(N) and N-H(O) RDF's suggests that the AE molecule (which has three donor and three acceptor sites) makes on average four strong H-bonds and has in its nearest surrounding presumably 3 weakly H-bonded (or non bonded) neighbors. In addition coordination numbers for O-H(N) and N-H(O) (and the corresponding RDF's) clearly indicate the preference for N-H···O versus O-H···N hydrogen bonds which leads to the conclusion that in the liquid phase AE should trend to adopt the gGt conformation. This result can be confirmed by previous experimental [27] and theoretical [26] findings.

In order to get more detailed insights into the three-dimensional local structure around EG, ED and AE in their pure liquids spatial distribution functions (SDF's) [99] were calculated and visualized (Figs. 17-20). For the sake of clarity in Figs. 17-20 only the fragment of the central molecule defining the local frame for each compound is represented. This fragment is always composed of the appropriate functional group (i.e. OH for EG, NH_2 for ED and both for AE) and its attached carbon atom.

In Figs. 17a and 17b the oxygen-oxygen SDF's for EG are shown for thresholds 3.0 and 1.8 times of the bulk density, respectively. One can see in Fig. 17a four features of different types due to nearest neighbors. Two of them (below and above the hydroxyl group) are the principal features corresponding to donors and acceptors of strong intermolecular H-bonds. The wide ring around the

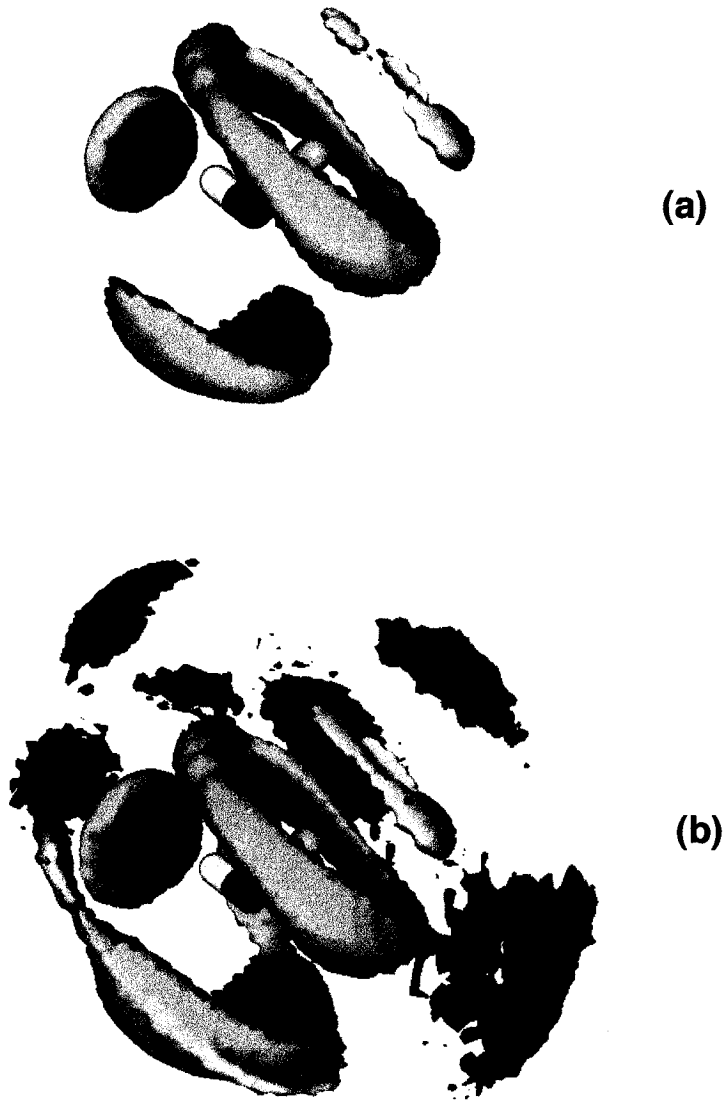


Fig. 17. Oxygen-oxygen spatial distribution functions for pure EG at isosurface thresholds of (a) 3.0 and (b) 1.8. The surfaces are colored by separation, from dark blue to red.

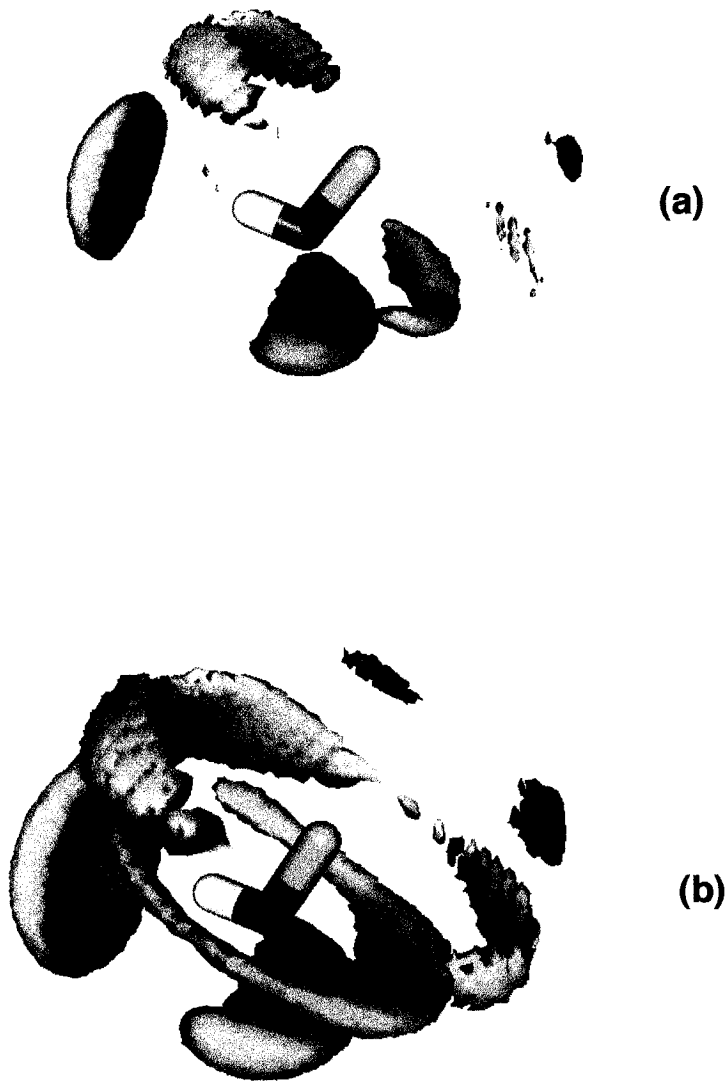


Fig. 18 Oxygen-hydrogen spatial distribution functions for pure EG at thresholds of (a) 5.0 and (b) 2.5.

carbon atom is due to the intramolecular oxygen (the second oxygen on the EG molecule) when it is in the *gauche* position with respect to the central dihedral angle and at the same time performs rotation around the O-C bond (i.e. rotation with respect to the HOCC dihedral angle). The second small ring (which is not well defined at this threshold) also indicates the same kind of rotation of the second oxygen when the central molecule adopts the *trans* conformation with respect to the OCCO dihedral angle.

At the lower isosurface threshold shown in Fig. 17b the SDF becomes more complex. The red distant features correspond to the secondary structure. These are the round-shaped feature right behind the principal H-bond acceptor, the group of features around the central ring (those are first nearest neighbors with respect to the second intramolecular oxygen) and the feature above the small ring due the first neighbor to the oxygen in the *trans* position. It is especially interesting to note the appearance of a rim on the first H-bond donating feature, which becomes larger and develops two "wings" on both sides extending upwards towards the cap due to the H-bond acceptor. It is likely that the rim indicates the presence of a rather weak H-bond donor, which is trying to occupy the second accepting site on the central oxygen. As has been shown previously, the "wings" that develop at larger separations on the edge of the principal H-bond donor feature are evidence of the presence of such "bifurcated" H-bond arrangements [54].

The oxygen-hydrogen SDF's for EG (Fig. 18) provide an additional confirmation of formation of such coordination. In Fig. 18a one can see at a threshold of 5.0 only strong H-bonded features survive at this relatively high threshold from a central molecule, in particular, the feature due to hydrogen attached to the accepting oxygen (in front of H on the hydroxyl group) and hydrogen directly H-bonded to the central oxygen. The four other features

apparent in Fig. 18a are due to intramolecular hydrogens (analogous to the oxygen-oxygen SDF) and are not clearly defined at this threshold, although they become more pronounced at a threshold of 2.5 (see Fig. 18b) and can be identified as follows. Two small more distant (red) features arise due to intramolecular hydrogens directly attached, as well as those H-bonded, to the oxygen in the *trans* position. The large ring around the carbon atom is due to hydrogens related to the intramolecular oxygen; the hydrogen covalently bonded to this oxygen contributes to the upper part of the ring while the H-bonded intermolecular hydrogen is responsible for the presence of its lower part. The next nearest feature to the central oxygen consist of a smeared cap and two long "wings" broadened at the ends. Its lower part (the cap) is also due to the hydrogen covalently bonded to the intramolecular oxygen when the EG molecule adopts an inverse conformation and the "wings" are another indication of a weak H-bond donor. In spite of the fact that the corresponding weak H-bond acceptor could not be definitely localized, it becomes apparent that EG trends to participate in two weak (possibly birfurcated) H-bonds in addition to two strong ones which is in agreement with results of the present radial and CN analysis.

The local structure around ED has been explored through the nitrogen-hydrogen and nitrogen-nitrogen SDF's (see Fig. 19). The nitrogen-hydrogen SDF at a threshold of 1.5 shown in Fig. 19a exhibit two principal features due to strong H-bonds. The single feature below the nitrogen corresponds to a single neighbor while each of the two caps connected by a wide "bridge" and located above the amine hydrogens combine to account for a second strong and a third weakly H-bonded nearest neighbor [54]. The large cap below the single (H-bond donating) feature is due to the second hydrogen of the amino group donating the H-bond to the central nitrogen. The small (red) distant feature appears to be a rather weakly H-bonded nearest neighbor, which unexpectedly approaches the amino

Fig. 19. Spatial distribution functions for pure ED: (a) nitrogen-hydrogen at a threshold of 1.5; (b) nitrogen-hydrogen at a threshold of 1.6 overlayed with nitrogen-nitrogen at a threshold of 2.3 (light shading).

group from "above". It is likely that the presence of the large (about 2/3) population of *trans* conformations in pure ED make such arrangements quite possible. Interestingly, no features associated with intramolecular hydrogen density (i.e. from the second amino group in the *trans* position) were noted. This can be seen as being a manifestation of the greater flexibility of the ED model (recall that no scaling was applied for 1-4 non-bonded interactions in ED).

From Fig. 19b, where the nitrogen-hydrogen SDF was superimposed with the corresponding nitrogen-nitrogen SDF, the local structural arrangement in liquid ED is further clarified. First one can see the appearance of the large ring due to internal rotation of the *gauche* nitrogen (pure ED has 36% of *gauche* rotamers) and the nitrogen density above the feature due to the weakly H-bonded hydrogen. The principal H bonded features become multi-layered (doubly for the H-bond accepting and triply for the H-bond donating features) because of the presence of both H and N atoms in the strongly associated amino groups. One can again conclude that within liquid ED (somewhat similar to the situation for EG) the nitrogen atom of ED participates in two strong and two weak H-bonds.

The SDF's for the two functional groups of AE (see Fig. 20) somewhat unexpectedly demonstrate structure rather similar to that observed for their counterparts in EG and ED. The nitrogen-nitrogen (shown from the "front side" of amino group) and the nitrogen-oxygen (shown from the "back side" of amino group) SDF's are presented at a threshold of 2.2 in Figs. 20a and 20b, respectively. It was necessary to change the orientation of the central molecule in order to show clearly all the features of these two functions. In both SDF's the regular well defined features due to the two H-bonded nearest neighbors are evident. The H-bond donating feature due to the neighboring oxygen (see Fig.

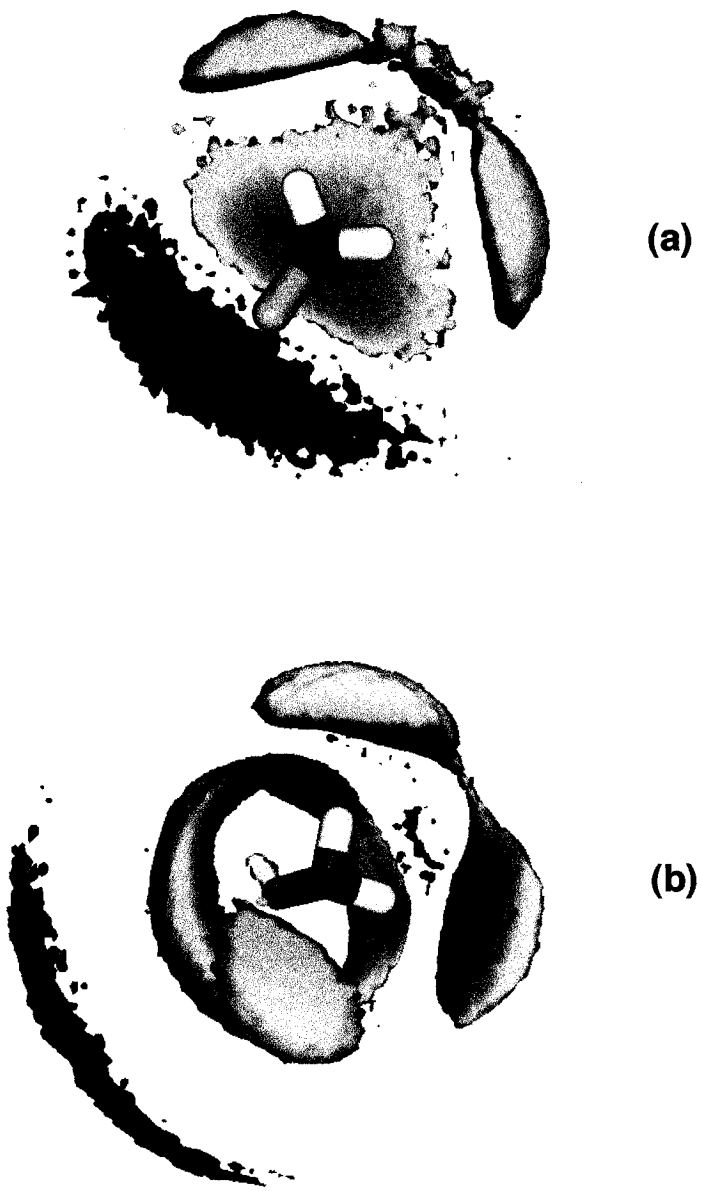


Fig. 20. Spatial distribution functions for pure AE: (a) nitrogen-nitrogen and (b) nitrogen-oxygen SDF's at a threshold of 2.2.

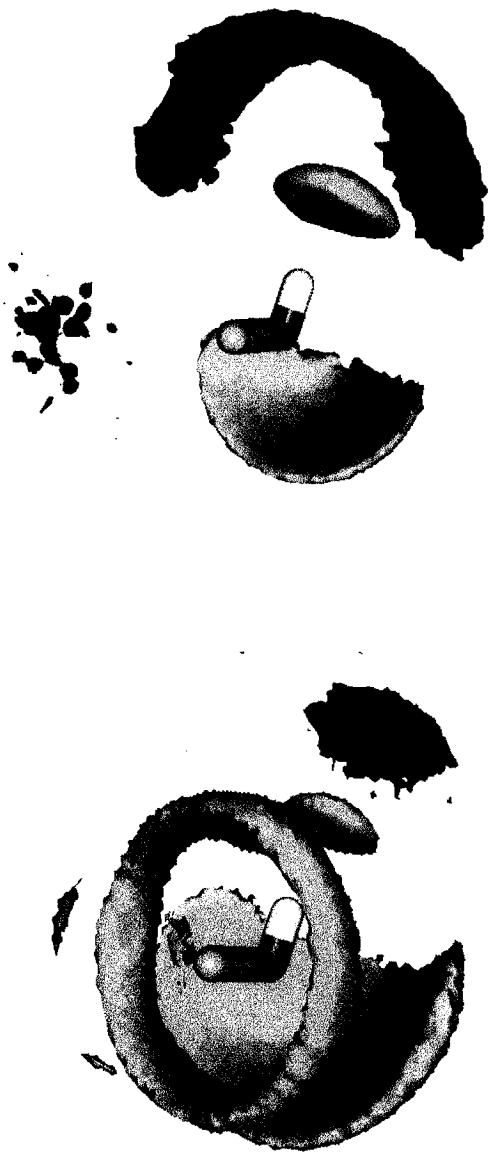


Fig. 20 continued. (c) Oxygen-oxygen and (d) oxygen-nitrogen SDF's at a threshold of 2.5.

20b) appears to be more localized than that due to neighboring nitrogen (see Fig. 20a). The N-N SDF has a large extended cap (appearing red at approximately 4.0 Å separation from the central nitrogen) corresponding to a secondary N; it appears to be a nearest H-bonded neighbor for the intramolecular oxygen in *gauche* position. The broad ring typical for this intramolecular oxygen can be clearly seen in Fig. 20b as well as the distant elongated secondary feature, which can be again identified as oxygen density nearest (and H-bonded) to the oxygen of the intramolecular ring.

The H-bonding pattern observed in oxygen-oxygen and oxygen-nitrogen SDF's (see 20c and 20d, respectively) has both familiar and unfamiliar aspects. The large secondary feature which forms a broad semicircular cap around the H-bond accepting feature in O-O SDF (see Fig. 20c) appears to be sharply contrast to the highly linear and localized H-bonding arrangement recorded in the corresponding O-N SDF (see Fig. 20d). In addition, the presence of a very small amount of atomic density due to intramolecular oxygen and nitrogen in *trans* position can be recognized in the diffuse pattern in front of the carbon atom (see Fig. 20c) and in the three small features outlining a second ring in Fig. 20d.

Together the results displayed in Fig. 20 indicate that both nitrogen and oxygen atoms of AE participate in two strong H-bonds. Unfortunately it was difficult to identify completely the locations of the three (as it follows from RDF and CN analysis) perhaps remaining weakly H-bonded nearest neighbor atoms. However, it is worth noting that such features can be expected to be rather sensitive to possible inadequacy of the model potential used for this particular compound.

CHAPTER VI - STRUCTURAL ORDERING IN AQUEOUS SOLUTIONS

In this chapter the theoretical investigation of EG, ED and AE is extended to aqueous mixtures of these compounds. On the basis of the analysis performed to estimate the quality of the models previously designed (see Chapter V), three models, namely basic EG, ED and AE models, were chosen for further simulations. For each of these models, four mixtures representing various typical cases along the concentration axis for binary aqueous solutions were prepared. Two dilute and two concentrated solutions with respect to each of the components were simulated (see Table 6, Chapter IV, for the mole fractions of the solute, X, and computational details) and the corresponding static and dynamic characteristics were determined. The discussion of conformational possibilities of the compounds of interest in aqueous surroundings will be followed by a detailed structural analysis for all compositions.

6.1 SIMULATION RESULTS

6.1.1 Energetic and dynamic properties of binary mixtures

Simulation results for the energies of the four compositions of EG, ED and AE in water are given in Tables 14, 15 and 16 (the intermolecular energies are given per mole of solution). The average total configurational energy, <U>, is comprised of the average potential energies due to solute-solute, solute-solvent and solvent-solvent interactions. These energies in turn can be divided into short-range (Lennard-Jones) and long-range (Coulombic) contributions to investigate the significance of specific interactions to changes in the local structure. From Tables 14-16 one can see increases of 14-20 times in the magnitudes of $<U_Q>$ for the solute-solute intermolecular energies moving from dilute to concentrated

Table 14. Energies for aqueous mixtures of EG.

	Composition (X)				Max.
	0.03	0.1	0.3	0.8	error
$<U_a>$	5.91	5.94	5.99	6.01	-
$<U_d>$	11.06	10.84	11.18	11.20	-
Intra. (EG)					
$<U_Q>$	94.54	93.99	92.91	91.59	0.17
$<U_{LJ}>$	0.41	0.40	0.45	0.48	0.006
Inter. (EG-EG)					
$<U_Q>$	-2.44	-6.99	-19.51	-42.86	0.52
$<U_{LJ}>$	-1.42	-4.27	-8.92	-12.11	0.02
Inter. (EG-H_2O)					
$<U_Q>$	-121.64	-107.84	-75.49	-77.20	1.80
$<U_{LJ}>$	-3.63	-0.73	3.03	8.38	0.08
Inter. (H_2O-H_2O)					
$<U_Q>$	-56.50	-51.43	-37.75	-9.83	0.23
$<U_{LJ}>$	8.93	8.57	7.08	2.14	0.03
<U>	**-43.94**	**-37.82**	**-20.25**	**24.08**	**0.02**

(Units for energy are kJ/mol.)

solutions. This trend is most pronounced in the case of ED. The corresponding magnitudes of the solute-solute $<U_{LJ}>$ behave in similar fashion but increasing less dramatically (by about 8-9 times) for all compounds. One can clearly see that the major contribution to the solute-solute intermolecular energy is electrostatic for all compounds and compositions with only exception of the two dilute solutions of ED and AE at X=0.1.

It is also useful to examine solute-solute and water-water contributions to the total average energies normalized per molecule of that species, which can be obtained by summation of corresponding $<U_Q>$ and $<U_{LJ}>$ terms followed by division by the respective mole fractions (not shown in Tables 14-16). The

Table 15. Energies for aqueous mixtures of ED.

	Composition (X)				Max.
	0.03	0.1	0.3	0.8	error
$<U_a>$	12.23	12.17	12.17	11.89	-
$<U_d>$	-6.89	-6.85	-6.94	-7.12	-
Intra.(ED)					
$<U_Q>$	148.76	148.66	148.63	148.68	0.03
$<U_{LJ}>$	-0.24	-0.35	-0.32	-0.29	0.02
Inter. (ED-ED)					
$<U_Q>$	-1.50	-4.22	-10.79	-30.20	0.88
$<U_{LJ}>$	-1.85	-5.27	-11.13	-15.69	0.04
Inter. (ED-H_2O)					
$<U_Q>$	-110.36	-95.29	-69.54	-72.20	2.05
$<U_{LJ}>$	-8.39	-4.69	0.48	4.39	0.07
Inter. (H_2O-H_2O)					
$<U_Q>$	-56.93	-51.74	-38.32	-11.71	0.34
$<U_{LJ}>$	8.99	8.69	7.22	2.29	0.03
<U>	**-42.36**	**-32.58**	**-4.21**	**71.79**	**0.01**

(Units for energy are kJ/mol.)

values calculated appear to be very similar in magnitude for all types of solute when the same compositions are compared. The magnitudes of energies in the case of solute-solute interactions decrease from -133.3, -128.7 and -111.7 kJ/mol at X=0.03 for AE, EG and ED, respectively, to about -55 to -60 kJ/mol for X=0.8. This trend reflects the apparent increase in the mutual ordering for the solute molecules in dilute aqueous solutions. Somewhat similar behavior was also observed in the case of water-water interactions (e.g. -49.5 for all water-rich compositions and -47.1, -43.9 and -38.5 kJ/mol for water-poor compositions of ED, AE and EG, respectively). The fact that the present water-water energy for the water-rich solutions (-49.0-49.7 kJ/mol) is larger than the value obtained for

Table 16. Energies for aqueous mixtures of AE.

	Composition (X)				Max.
	0.03	0.1	0.3	0.8	error
$<U_a>$	13.88	13.81	14.46	14.59	-
$<U_d>$	3.76	3.65	4.39	4.67	-
Intra. (AE)					
$<U_Q>$	64.51	64.98	62.28	61.43	0.42
$<U_{LJ}>$	5.19	5.08	5.68	5.82	0.10
Inter. (AE-AE)					
$<U_Q>$	-2.39	-4.64	-14.70	-34.43	0.42
$<U_{LJ}>$	-1.61	-5.04	-10.60	-15.13	0.03
Inter. (AE-H_2O)					
$<U_Q>$	-113.11	-97.83	-74.92	-77.73	1.90
$<U_{LJ}>$	-9.08	-4.96	0.50	6.79	0.10
Inter. (H_2O-H_2O)					
$<U_Q>$	-57.24	-52.32	-39.09	-11.32	0.26
$<U_{LJ}>$	9.03	8.77	7.48	2.54	0.08
<U>	**-44.46**	**-39.37**	**-24.90**	**13.14**	**0.05**

(Units for energy are kJ/mol.)

pure SPC/E water (-46.73 kJ/mol) suggests an increase in water-water correlations (structure-making) in these dilute solutions (similar to that found for water-ME mixtures [43]). Finally, a comparison of total configurational energies reveals a consistent decrease in their magnitudes from X=0.03 to 0.3 for all compounds; when the solute content is further increased from X=0.3 to 0.8 the total energy for all compounds, and especially for ED, continued to increase and actually changed sign.

The mean self-diffusion coefficient, D, was selected as a representative dynamical property for the systems of interest. The simulated values of D are

compared with the composition dependence of experimental estimates in Fig. 21. Self-diffusion coefficients for aqueous solutions of EG and ED were measured at 298 K by spin-echo NMR [9,150]; again it is very unfortunate that there appears to be no experimental results available for binary mixtures of water and AE. All raw experimental data as well as simulated values of D obtained in the present study are listed in Table D1.

In Fig. 21a one can see that the composition dependence for ED is quite well reproduced, in particular the decreasing trend of the mean D up to X=0.3 (0.33 in accord with the experimental data). At this concentration Val'kovskaya *et al.* [150] observed a minimum in ED mobility, and suggested that it corresponds to the formation of cyclic ED·2H_2O complexes [151]. Earlier the same research group (Rodnikova and co-authors [152]) studied the dielectric properties of water-ED mixtures and concluded that a molecule of ED is incorporated into the H-bonded network of water without its alteration up to a concentration of 10 mol%.

For aqueous solutions of EG the experimental D gradually decreases becoming essentially constant in EG rich solutions (i.e. X>0.8 in Fig. 21b). The calculated dependence nicely correlates with the experimental curve with a slight inconsistency in the constant (EG-rich) region. Similar trends (see also Fig. 21b) have been observed for AE in the water-AE mixtures considered in the present study. The somewhat unusual increase in D at X=0.8 can be explained by a high sensitivity of constant-volume simulations to the quality of the experimental density; small changes (error) in this parameter can cause noticeable deviation of D from its correct value. In principle this problem can be partially resolved by running simulations at constant-pressure allowing consistent densities (and therefore D) to be determined during the simulation run.

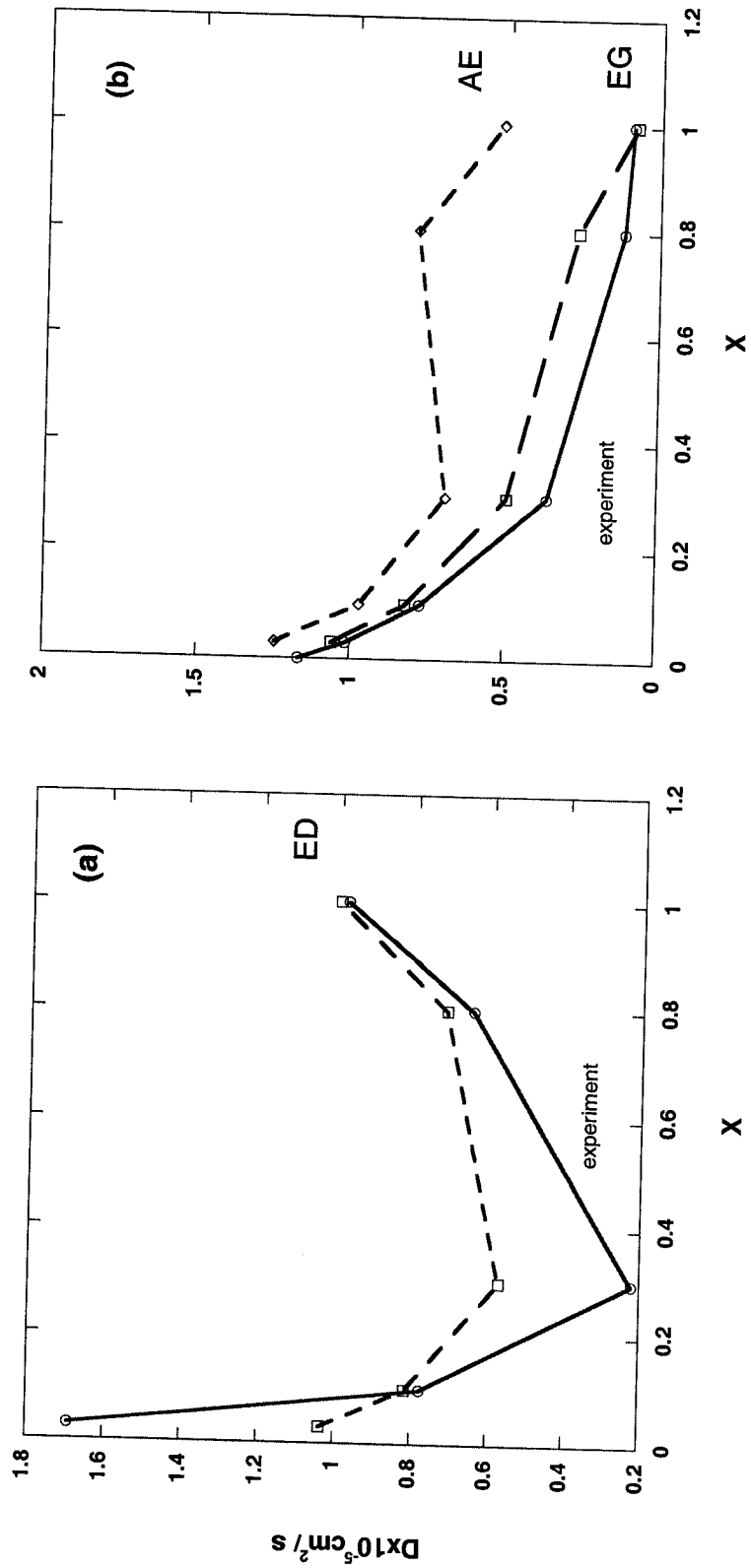


Fig. 21 Composition dependence of the self-diffusion coefficient for aqueous solutions of ED, AE and EG at 298 K: (a) ED; (b) AE and EG. The solid and dashed lines correspond to the experimental and calculated results, respectively.

6.1.2 Dihedral angle distributions

Conformational characteristics of aqueous mixtures of EG, ED and AE were investigated in a manner similar to those for pure liquids. The dihedral angle distributions were obtained for each concentration of all species and they are shown in Figs. 22-24. The corresponding populations of conformers (for the central torsion angle) were also calculated and the results are given in Table D2.

The most prominent feature of the OCCO dihedral angle distributions of EG (see Fig. 22) is the presence of a significant population (56%) of T conformers which is observed only for the lowest composition. For this system the remaining population is comprised of G (35.3%) and a small amount (8.7%) of G' conformers. For the three other (higher) concentrations only *gauche* conformers (consistently located at -62.5 and 62.5°) were registered analogously to pure EG. All HOCC dihedral angle distributions for EG-water mixtures exhibit three (two *gauche* and one *trans*) stable maxima at about -62.5, 62.5 and 177.5°, respectively.

The conformational difference observed in dilute EG solution indicates that in the presence of a large number of water molecules the free energies of the G and T conformers become roughly equal. This issue will be discussed later in terms of possible preferences for H-bond formation. It is worth noting that these findings are in qualitative accord with the *ab initio* results of Nagy *et al.* [29], Hooft *et al.* [59], and Cramer and Truhlar [21] where the presence of both T and G conformers in infinitely dilute solutions of EG was clearly shown. In addition our roughly 50:50 gauche/trans ratio appears to be in reasonable quantitative agreement with the 67:33 estimate predicted by Hooft *et al.* [59].

In the case of ED solutions the conformational content is quite uniform with respect to the NCCN torsion angle (see Fig. 23). *Trans* conformers predominate at all compositions with a population of about 70% for low and 90%

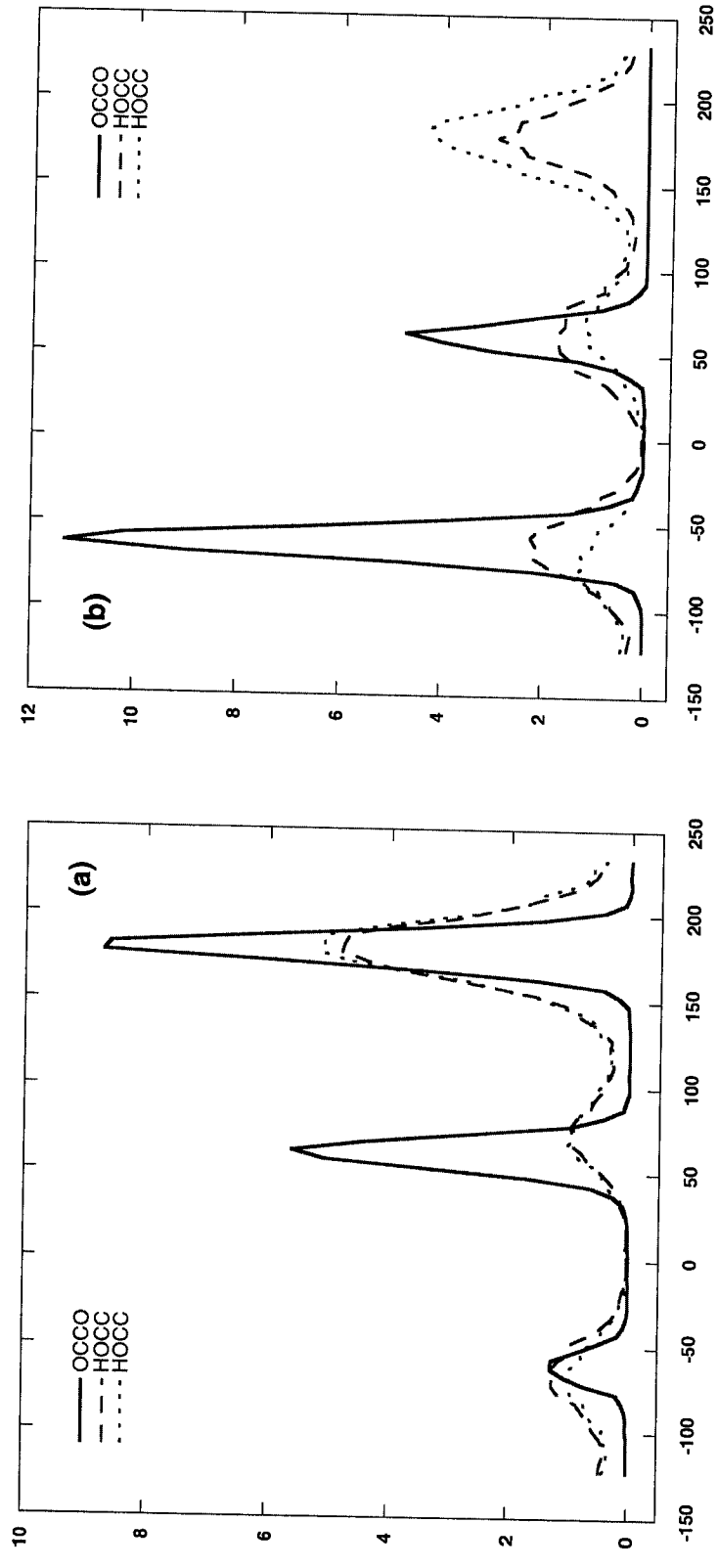


Fig. 22. Dihedral angle distributions for aqueuos solutions of EG: (a) mole fraction of 0.03; (b) mole fraction of 0.1.

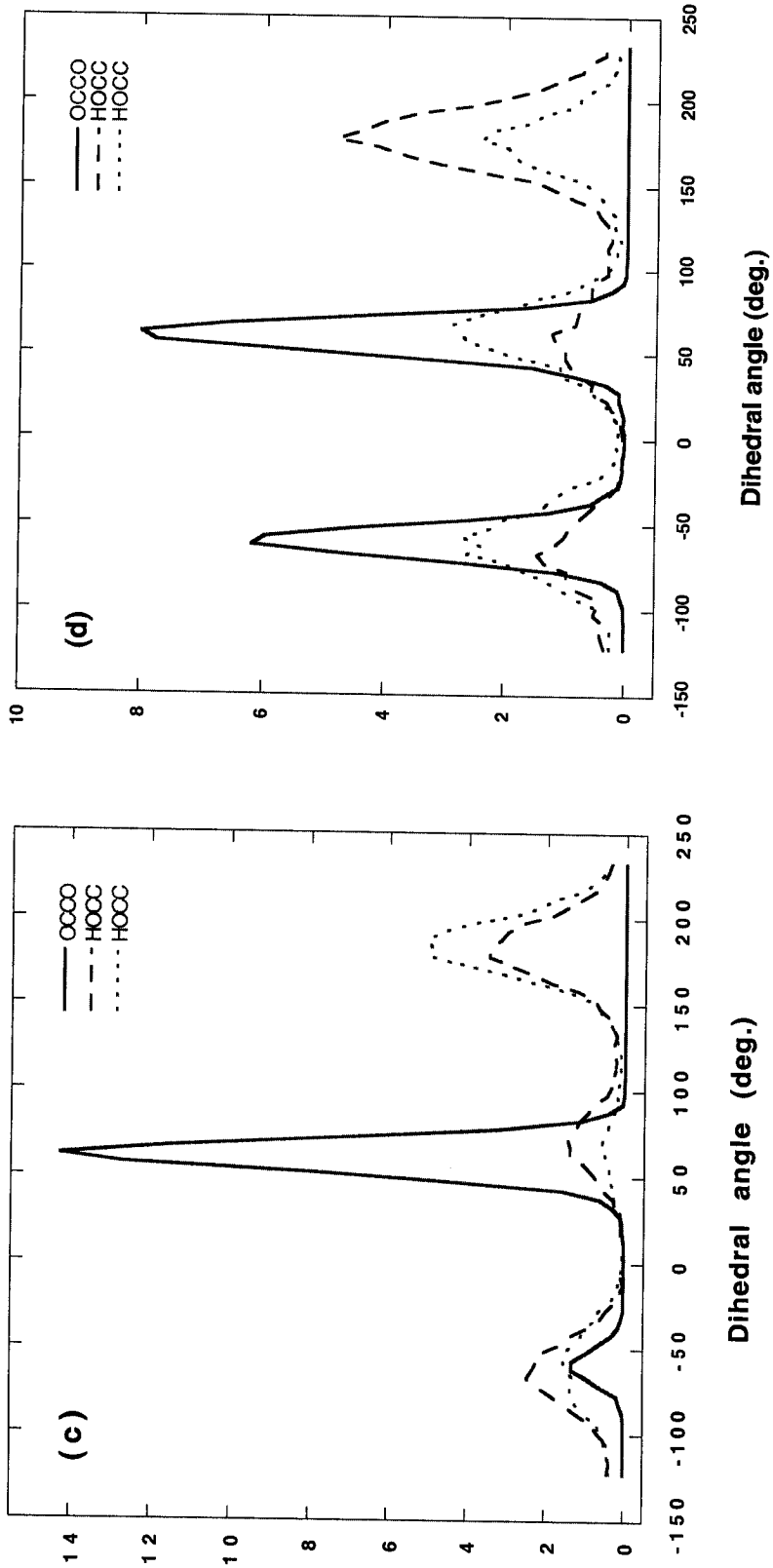


Fig. 22 continued. (c) Mole fraction of 0.3; (d) mole fraction of 0.8.

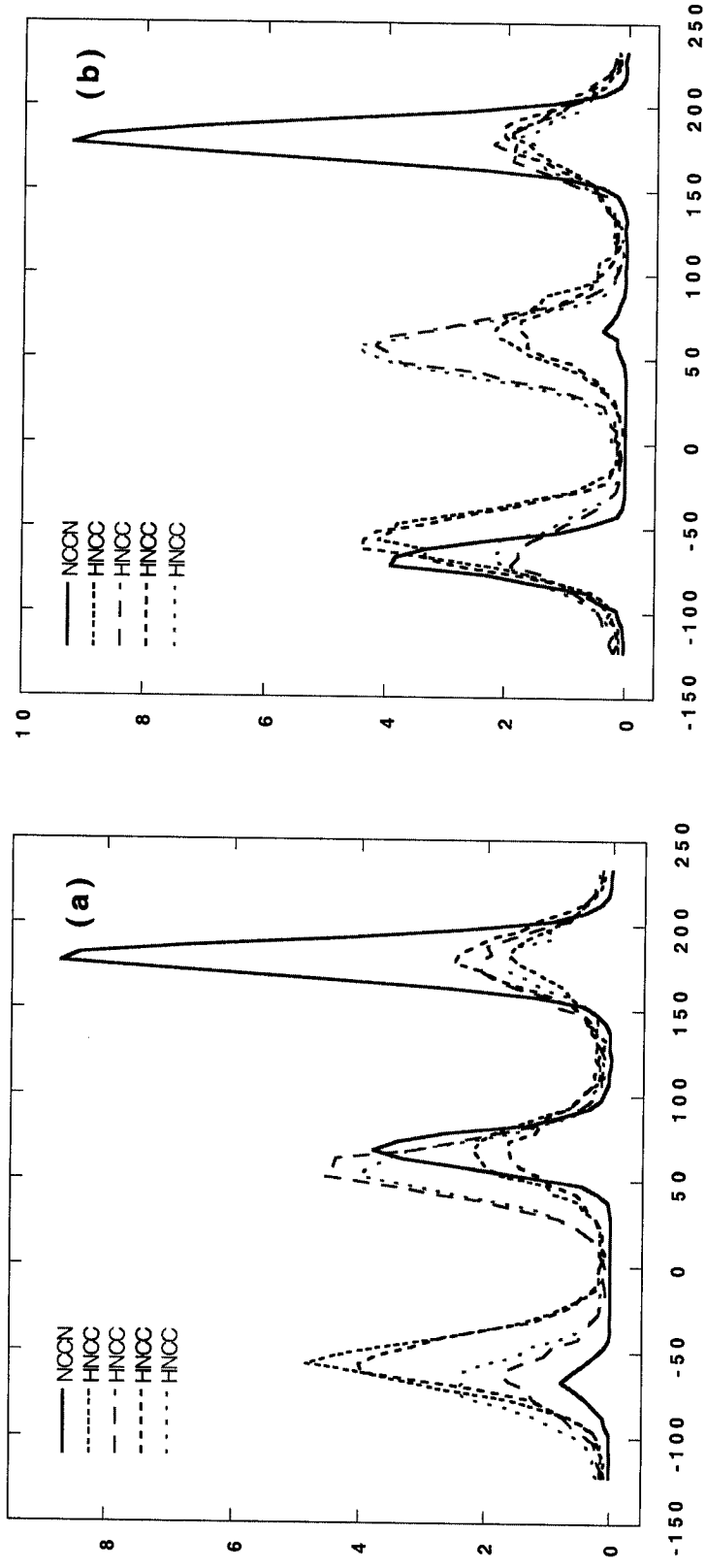


Fig. 23. Dihedral angle distributions for aqueuos solutions of ED: (a) mole fraction of 0.03; (b) mole fraction of 0.1.

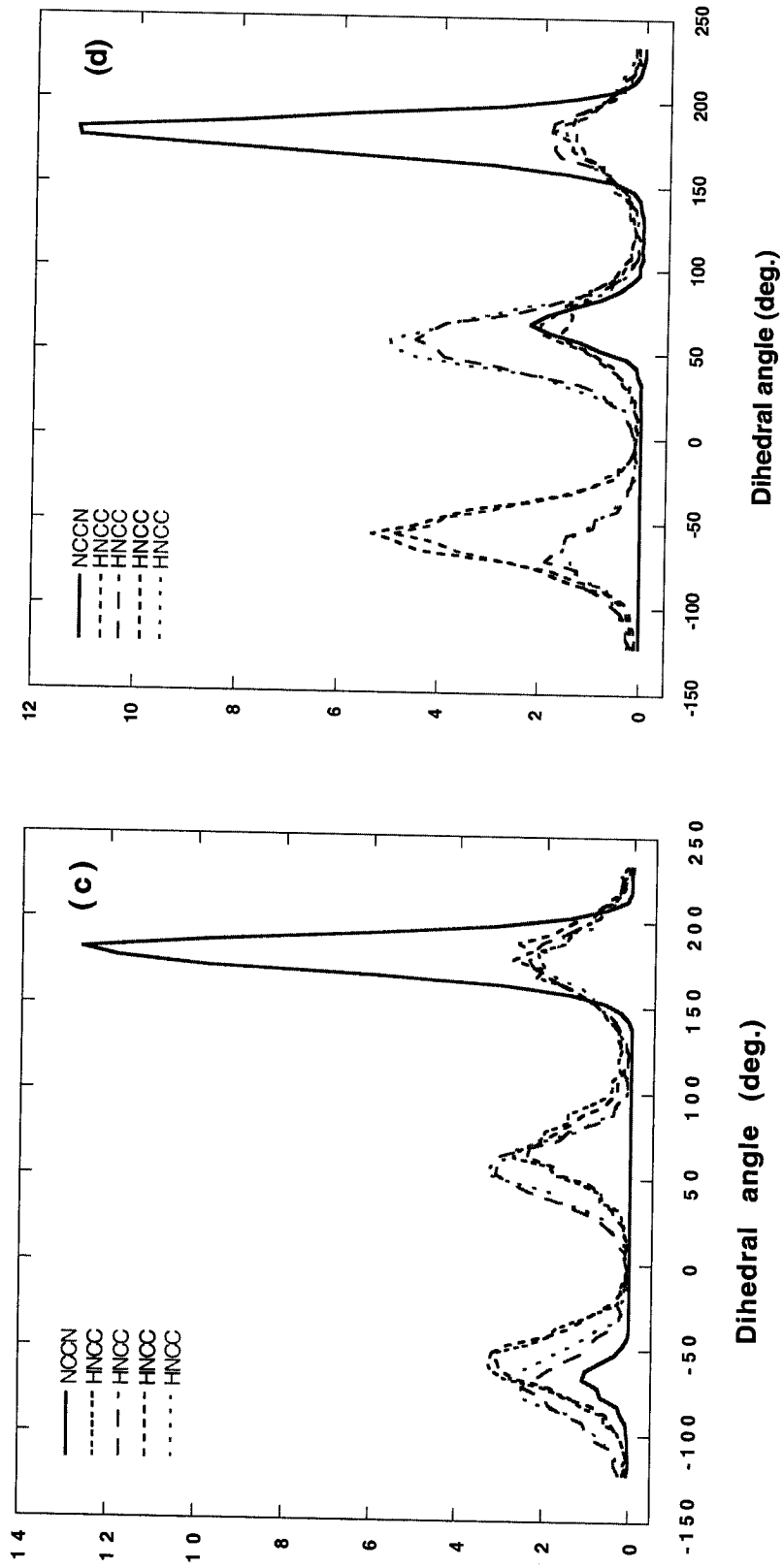


Fig. 23 continued. (c) Mole fraction of 0.3; (d) mole fraction of 0.8.

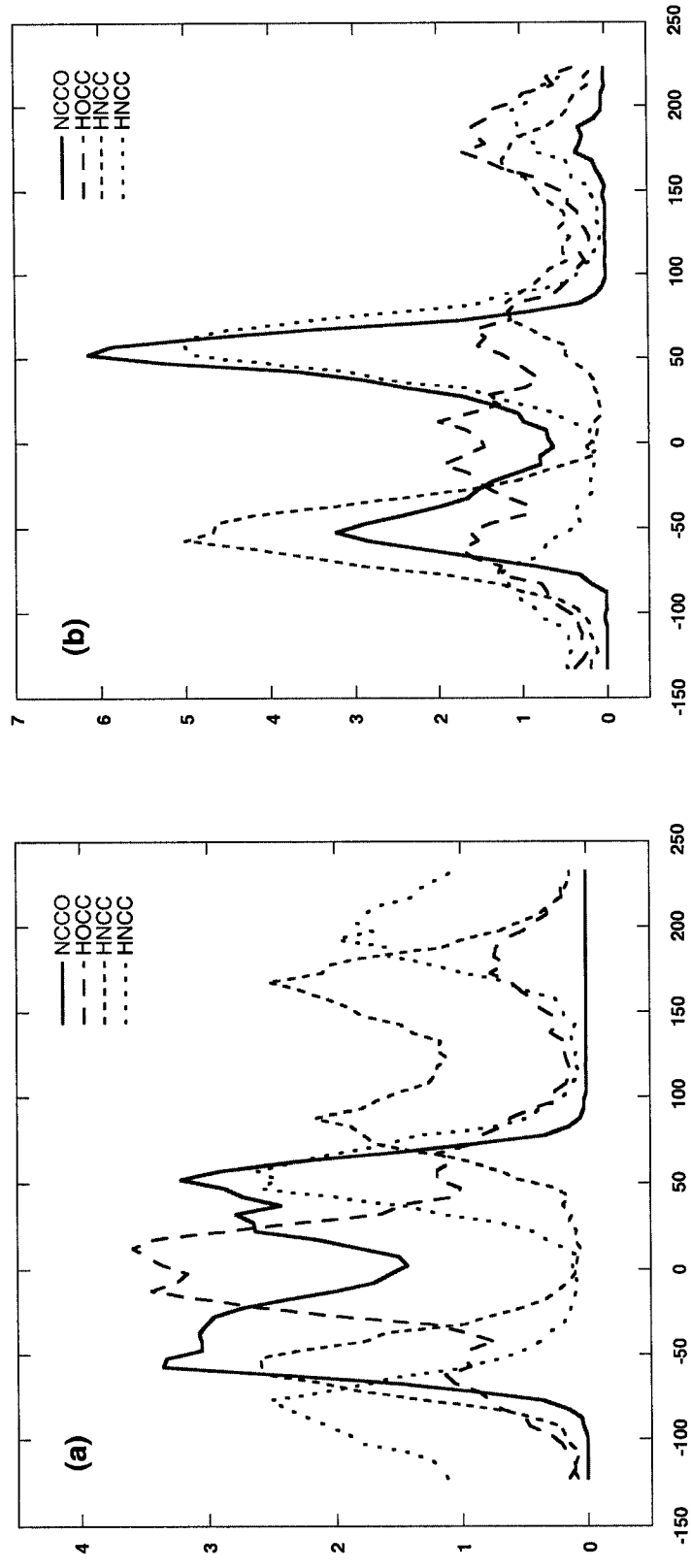


Fig. 24. Dihedral angle distributions for aqueuos solutions of AE: (a) mole fraction of 0.03; (b) mole fraction of 0.1.

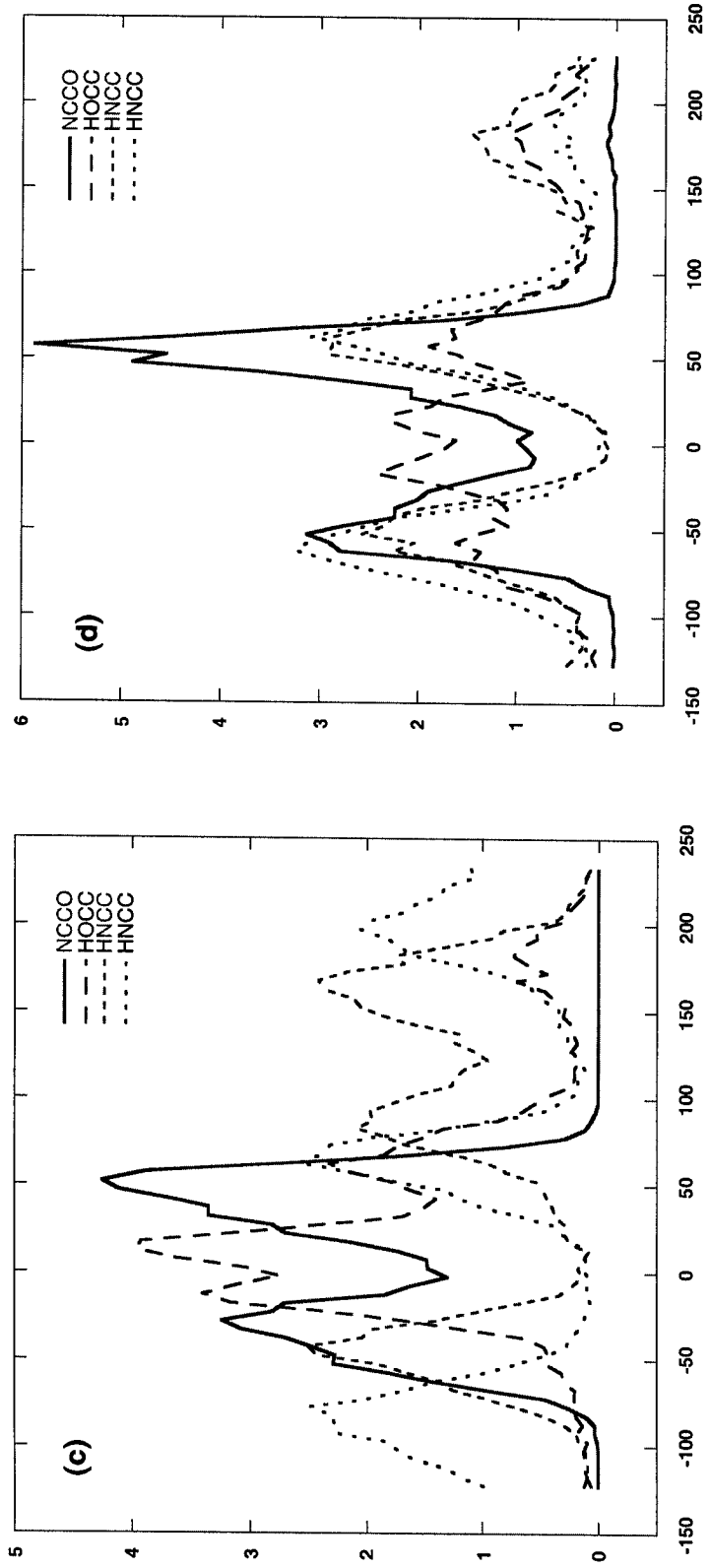


Fig. 24 continued. (c) Mole fraction of 0.3; (d) mole fraction of 0.8.

for high concentrations. It can be noted that, similar to EG, the locations of the probability maxima for the central dihedral angle remain virtually unchanged when composition changes. The HNCC distributions have three maxima (g', g and t) whose positions and intensities vary slightly depending on composition.

The dihedral angle distributions for AE-water mixtures (see Fig. 24) exhibit the most complicated behavior, resembling to a great extent that observed for the pure system. In particular, both *gauche* peaks on the NCCO distributions markedly overlap, but in contrast to pure AE, are centered much closer to ideal *gauche* positions. Furthermore the HOCC torsion has a broad peak with a double maximum in the *cis* region (see Figs. 24a and 24c). Interestingly, these similarities become more pronounced for X=0.03 and 0.3 solutions. The conformational pattern for the two other concentrations (X=0.1 and 0.8) appears to be quite different, with the NCCO distributions exhibiting trace amounts (2.5 and 0.6%, respectively) of *trans* population and the major *gauche* peaks becoming very well resolved (see Figs. 24b and 24d). Even larger changes in conformational behavior are visible in the HNCC and HOCC dihedrals. For the latter the distinct peaks seen in the pure liquid become virtually indistinguishable, which suggests that during these runs HOCC scans almost all conformational space. In contrast, peaks in the HNCC distributions appear to be better resolved and located at about ±50 and 180°.

6.2 STRUCTURAL ANALYSIS

To our knowledge the structural analysis of 1,2-disubstituted ethanes within binary aqueous mixtures has not been previously performed by means of RDF's. The fact that averaging over the angular coordinates of the pair distribution functions can often result in the cancellation of contributions from regions of low and high probability at the same distance but composing different

parts of the local structure in solution, implies that RDF's can provide a complete structural picture only for liquids of spherical particles. However, the aqueous binary mixtures of some associating liquids (e.g. methanol, methylamine) have been examined at the radial distribution function level followed by structural analysis going beyond radial averages [43,54]. The analogous approach has been performed in the present study with respect to the aqueous mixtures of EG, ED and AE.

6.2.1 Solution structure from radial distribution functions

The selected radial distribution functions for pure water and aqueous solutions of EG, ED and AE for four different concentrations are given in Figs. 25, 26, 28 and 30, respectively. Composition dependence of total coordination numbers calculated with respect to the selected atomic sites for each solution are given in Figs. 27, 29 and 31 (complete data tables for the original values of CN are available in Appendix E).

Beginning with EG solution, first compare water-water correlations are compared with those of pure water. From Figs. 25 and 26a one can see that the first maximum at 2.8 Å in $g(r_{OO})$ increases monotonically with EG concentration. In water-rich mixtures (X=0.03 and 0.1) the structural features (for both primary and secondary coordination) are similar to that found for pure SPC/E water (see Fig. 25) which indicates that the presence of EG molecules in this concentration range does not appear to influence water-water coordination. In water-poor mixtures (X=0.8) the shift and lowering of $g(r_{OO})$ at and beyond the first minimum suggests that there is a tendency for the association of a small number of water molecules (e.g. as dimers) in this system. This phenomenon has been previously observed in water-methanol mixtures [43].

The first maxima in the EG-EG $g(r_{OO})$, which occurs at 2.8 Å similar to

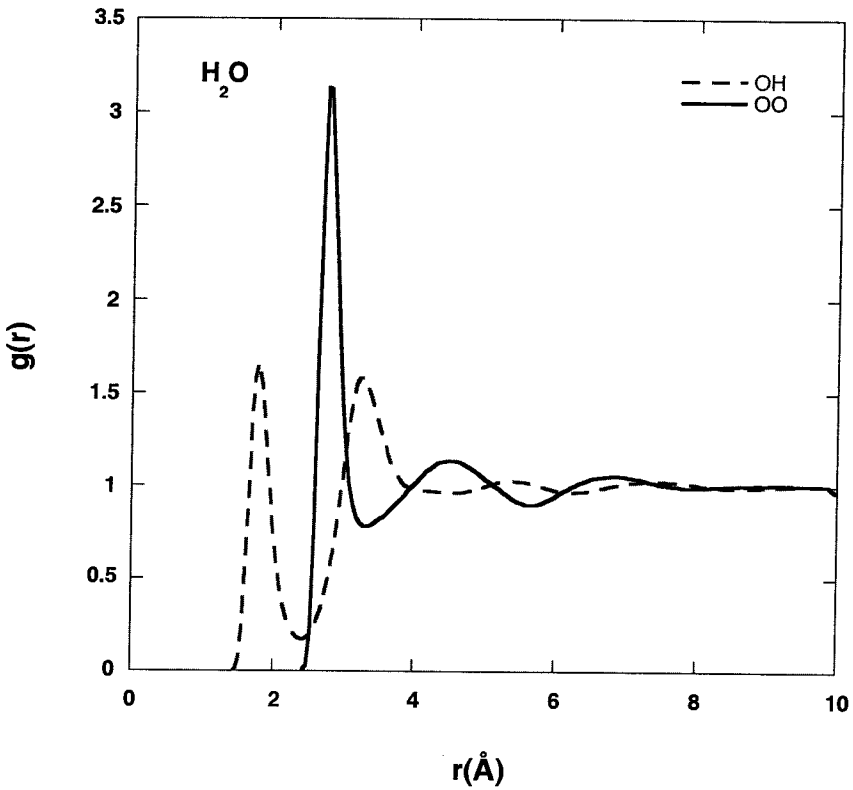


Fig. 25.
Radial distribution functions for pure water at 298 K. The solid and dashed lines represent the oxygen-oxygen and oxygen-hydrogen site-site RDF's, respectively.

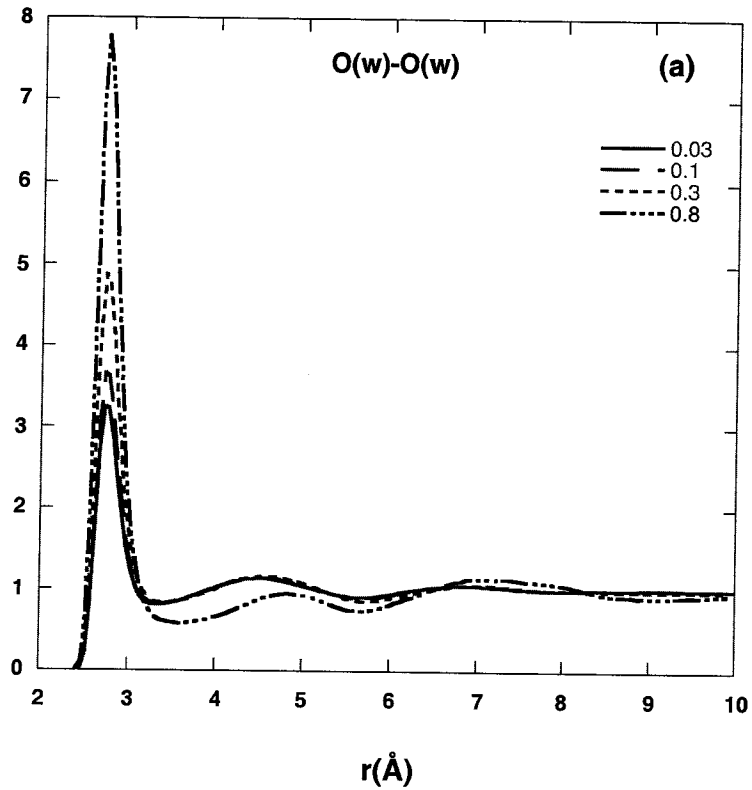


Fig. 26.
Radial distribution functions for water-EG solution at 298 K. The solid, long dashed, short dashed and dot-dashed lines represent X=0.03, 0.1, 0.3 and 0.8, respectively. (a) Water-water RDF's for oxygen atoms.

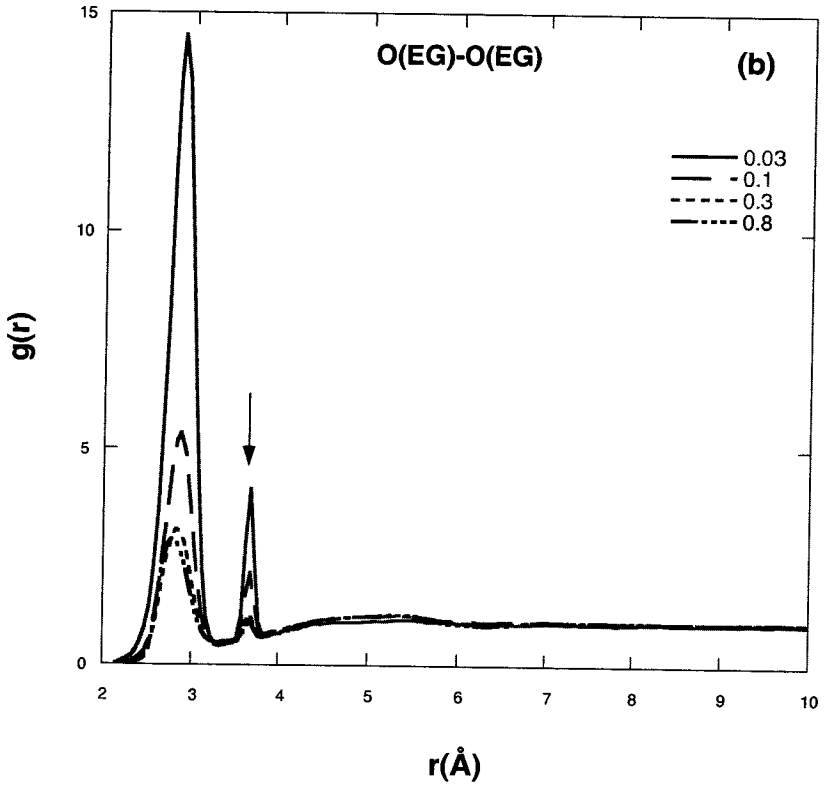


Fig. 26 continued.
(b) EG-EG RDF's for oxygen atoms. Arrow indicates peaks due to intramolecular features.

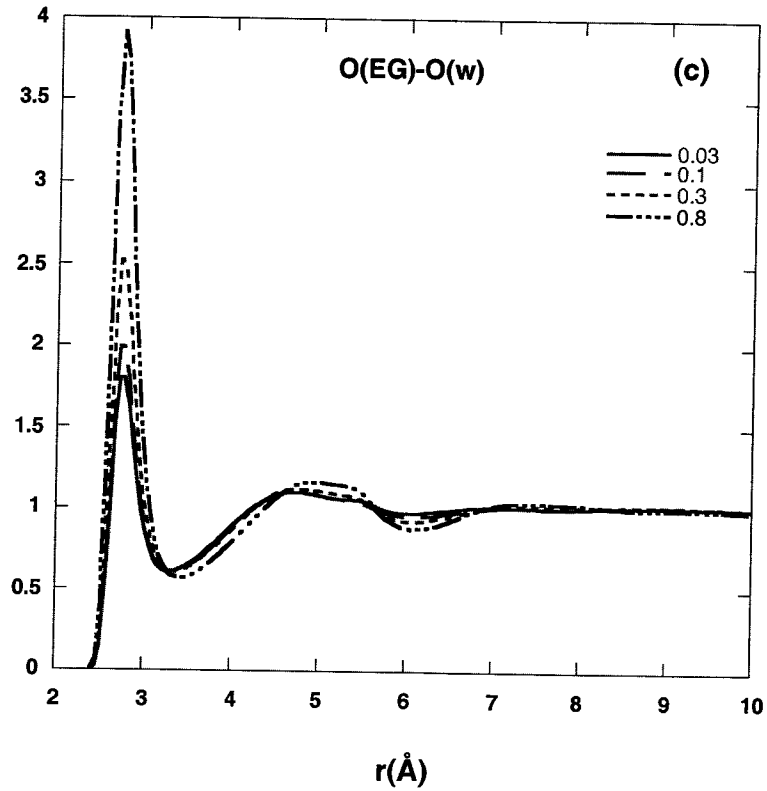


Fig. 26 continued.
(c) EG-water RDF's for oxygen atoms.

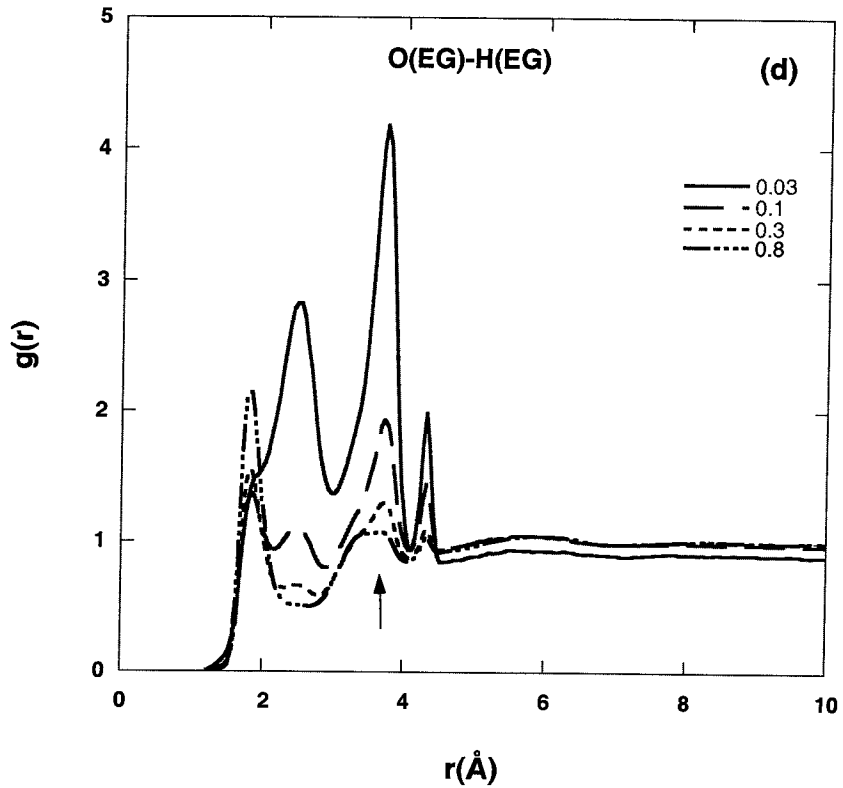


Fig. 26 continued.
(d) RDF's between oxygen and hydrogen of EG.
Arrow indicates peaks due to intramolecular features.

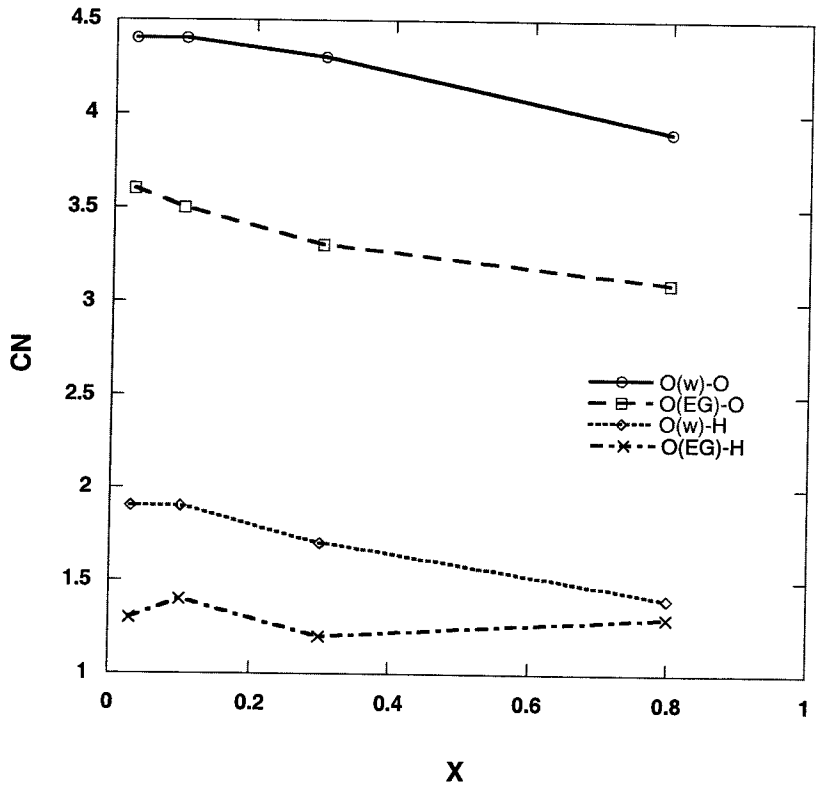


Fig. 27.
Composition dependence of coordination numbers for aqueous solutions of EG at 298 K.

that in water-water $g(r_{OO})$ indicating the presence of strong H-bonding between the corresponding sites, increases with decreasing EG concentration as can be seen from Fig. 26b. This suggests that EG-EG H-bonds persist even as this component becomes more dilute. The peaks observed in $g(r_{OO})$ at 3.7 Å (with the highest maximum for X=0.03) can be identified with the intramolecular oxygen in the *trans* position. We recall that in the dilute solution the number of molecules in the *trans* conformation increases to more than 50%.

Curves for $g(r_{OO})$ for water-EG pairs are shown in Fig. 26c. Overall the features in these functions become sharper (i.e. maxima higher, minima deeper) indicating strong water-EG correlations as the EG content of these solutions increased. It is also apparent from Fig. 26c that the position of the second maximum shifts to slightly larger separations with increasing X, indicating a slight shift in the specific structure of the second coordination shell in these systems.

The oxygen-hydrogen RDF's of EG in its aqueous mixtures (see Fig. 26d) is the richest in features and demonstrates the strongest dependence on the concentration. The first peak at 1.8 Å (due to relatively strong intermolecular H-bonding) followed by a less well defined peak at 3.7 Å (due presumably to the intramolecular hydrogens from *trans* conformers) can be observed for X=0.8 and 0.3. As the concentration is lowered further the peak gradually diminishes and finally (for the most dilute solution) transforms into a small shoulder on the large peak at 2.5 Å, which has been increasing after having first appeared at X=0.3. As expected the highest peak due to intramolecular hydrogens in *trans* conformers is observed in the lowest concentration solution. The distinct peaks recorded at 4.3 Å, which also increase with decreasing concentration, cannot be readily explained on the basis of RDF analysis.

Several features in the estimated total coordination numbers (around oxygen atoms) given in Fig. 27 are noteworthy. It might be expected that the

coordination numbers around these oxygen atoms would indicate the number of nearest neighbors H-bonded to the central molecule. From Fig. 27 one can see that water maintains a coordination number of about 4.4 (the same found for pure SPC/E water) until X=0.3 after which point it drops to 3.9 for the highest concentration of EG. Naturally the coordination numbers due to neighboring (H-bonding) hydrogens behave similarly; they approach two in water-rich solutions (indicating that the two acceptor sites on a water molecule are fully occupied) and then drop to 1.4 for EG-rich mixtures. This behavior can be interpreted as reflecting the predominant nature of EG structure in EG-rich aqueous solution when water is unable to establish tetrahedral coordination but rather is "forced" to adopt a structural motif more typical of EG. The number of neighboring hydrogens around EG oxygens remains relatively unchanged over the entire composition range indicating that the H-bonding tendency of EG in aqueous solutions does not appear concentration sensitive, in contrast to the behavior found for methanol [43]. In general water has a higher total or species specific (see Table E1) coordination numbers than EG. The total oxygen coordination around water is consistently about one larger than for an EG oxygen over the entire composition range.

RDF's for the water-ED solutions are shown in Fig. 28. In comparison with EG solution, water-water correlations did not reveal any unusual structural features. The first peak in $g(r_{OO})$ (not shown) indicates the presence of strong H-bonding at a separation of 2.7 Å and exhibits a similar trend of increasing intensity in ED-rich mixtures. The nitrogen-oxygen RDF's for ED-water pairs (see Fig. 28a) have their first peak at the usual position of 2.8 Å, but in contrast to the analogous functions of EG the first minimum on the curve corresponding to the highest concentration is markedly lowered and shifted towards larger separations. The first maxima in the corresponding nitrogen-hydrogen RDF's

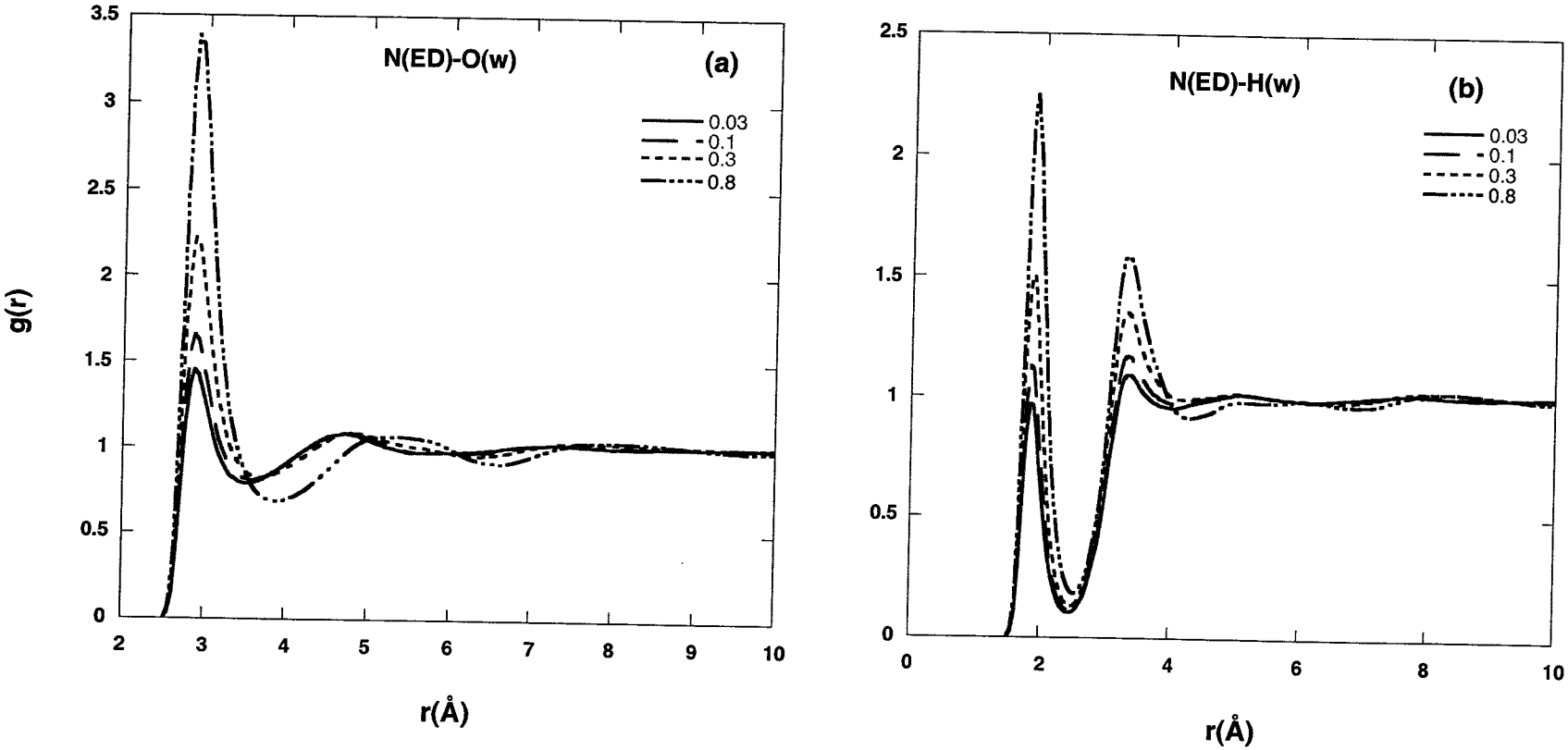


Fig. 28. Radial distribution functions for water-ED solutions at 298 K. The solid, long dashed, short dashed and dot-dashed lines represent X=0.03, 0.1, 0.3 and 0.8, respectively. ED-water RDF's for (a) nitrogen-oxygen and (b) nitrogen-hydrogen atoms.

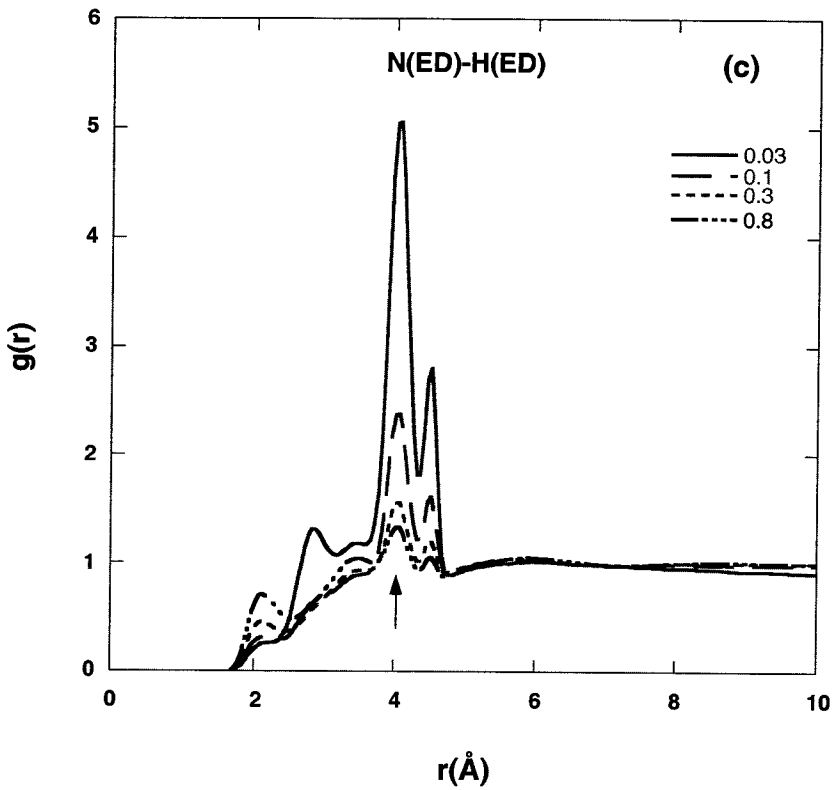


Fig. 28 continued.
(c) ED-ED RDF's for nitrogen and hydrogen atoms.
Arrow indicates peaks due to intramolecular features.

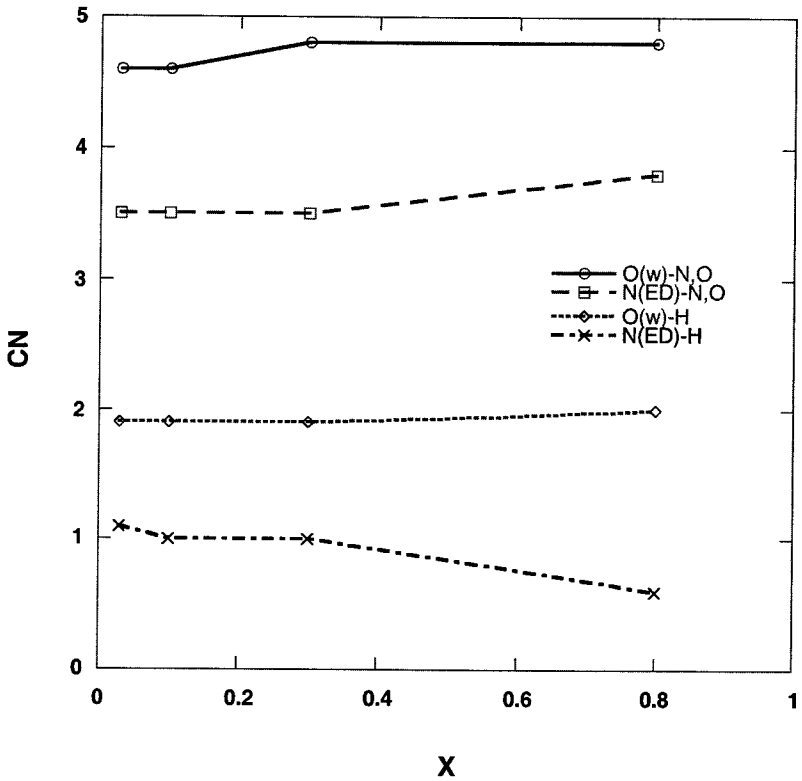


Fig. 29.
Composition dependence of coordination numbers for aqueous solutions of ED at 298 K.

(due to apparently strong O-H···N hydrogen bonds) are very well defined and are followed by slightly broadened but also well defined secondary maxima at 3.3 Å, after which the functions become essentially featureless. The first minima in this function reaches a value of 0.1 indicating a rather stable solute-solvent H-bond when ED acts as the acceptor.

The most complicated RDF obtained for ED-ED correlations was for nitrogen-hydrogen atomic pairs. Interestingly, the same features and trends apparent in the analogous oxygen-hydrogen RDF's of EG solution (see Fig. 26d) can also be found in $g(r_{NH})$ for ED mixtures (see Fig. 28c). For both ED and EG solutions these corresponding RDF curves have four distinct maxima where in the case of ED they are shifted to somewhat larger separations. The first (H-bonding) peak appears at 2.1 Å (for only X=0.8 and 0.3) indicating the presence of weaker N···H-N hydrogen bonds. At two low concentrations the disappearance of this first peak is accompanied by the appearance of a broad peak at 2.9 Å which grows with decreasing ED concentration. Again, analogously to dilute solutions of EG, the largest maximum at 4.1 Å is believed to arise from intramolecular hydrogens of *trans* conformers. The intensity of this peak correlates well with the population of *trans* conformers present in ED-water mixtures. One might also suggest that this population is also "responsible" for the last sharp peak noted at about 4.5 Å for both EG and ED solutions.

Examination of the coordination numbers provided by the RDF's in Fig. 29 (also see Table E2) reveals further interesting trends. The water oxygen total coordination numbers (due to the presence of both oxygens from water and nitrogens from ED) for the two water-rich solutions are 4.6 (i.e. only slightly higher than the 4.4 recorded for pure water) indicating a structural "shift" opposite to that observed in EG solutions. In the ED-rich solutions the coordination number increases yet further and approaches a value of 4.8. One possible

explanation for this behavior suggests the formation of "weak" H-bond arrangements around water oxygen (acceptor sites) in order to maintain a total H-bond balance in these systems; we recall that in pure ED nitrogen trends to participate in two strong and two weak H-bonds in order to achieve a balance. This trend is consistent with the fact that total coordination numbers of the water oxygen due to the neighboring hydrogens in the first hydration shell remain virtually unchanged (~1.9) indicating that the oxygen acceptor sites are fully occupied. The total coordination around water oxygen is consistently about one more than that for the ED nitrogen over the entire composition range. This can be explained by noting that both oxygen (in water) and nitrogen (in ED) are double H-bond donors whereas nitrogen can accept only a single (strong) H-bond. Finally, as expected the total hydrogen coordination around ED nitrogen remains relatively unchanged (at about 1.0) for X=0.03, 0.1 and 0.3; the slight decrease below unity for X=0.8 may be attributed to a problem with the experimental density for this particular composition.

One might reasonably suggest that the local structural arrangements around OH and NH_2 groups in AE aqueous mixtures would be somewhat reminiscent to those observed around these same groups in EG and ED aqueous solutions. The extent to which this is the case can be considered by examining the RDF's shown in Fig. 30 for AE solution. Analysis of water-water correlations (see Fig. 30a) reveals a first maximum in $g(r_{OO})$ at 2.8 Å (typical for all three solutions studied) which again increases when water concentration decreases. Similar to EG solution, the first minimum in the RDF for AE-rich mixtures is lowered and shifted to slightly larger separations (with respect to the functions for other compositions). As it was mentioned above, a tendency for the association of small numbers of water molecules (preferably as dimers [43]) can explain such a structural feature specific to concentrated EG and AE solutions. It

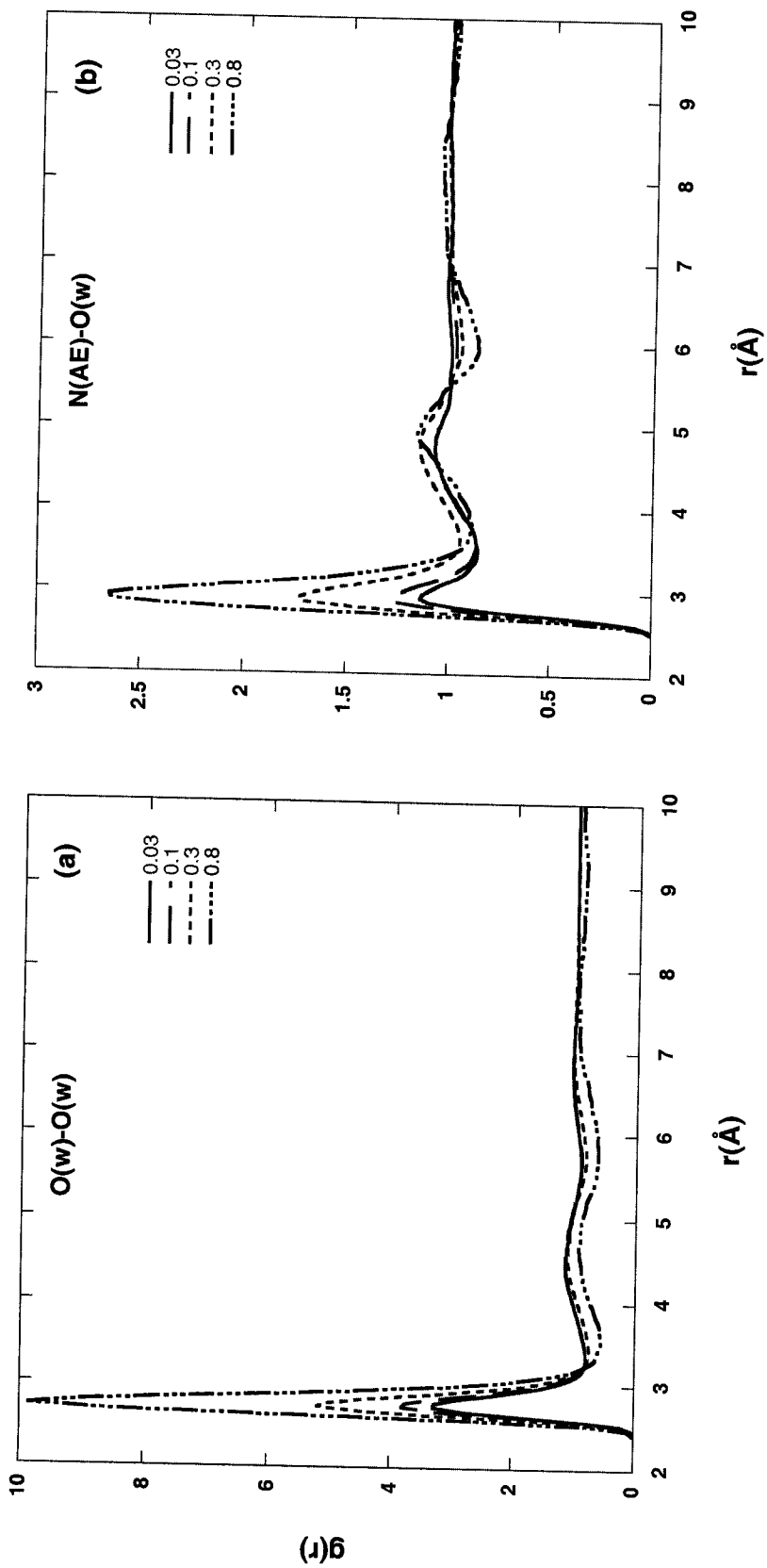


Fig. 30. Radial distribution functions for water-AE solutions at 298 K. The solid, long dashed, short dashed and dot-dashed lines represent X=0.03, 0.1, 0.3 and 0.8, respectively. (a) Water-water RDF's for oxygen atoms. (b) AE-water RDF's for nitrogen and oxygen atoms.

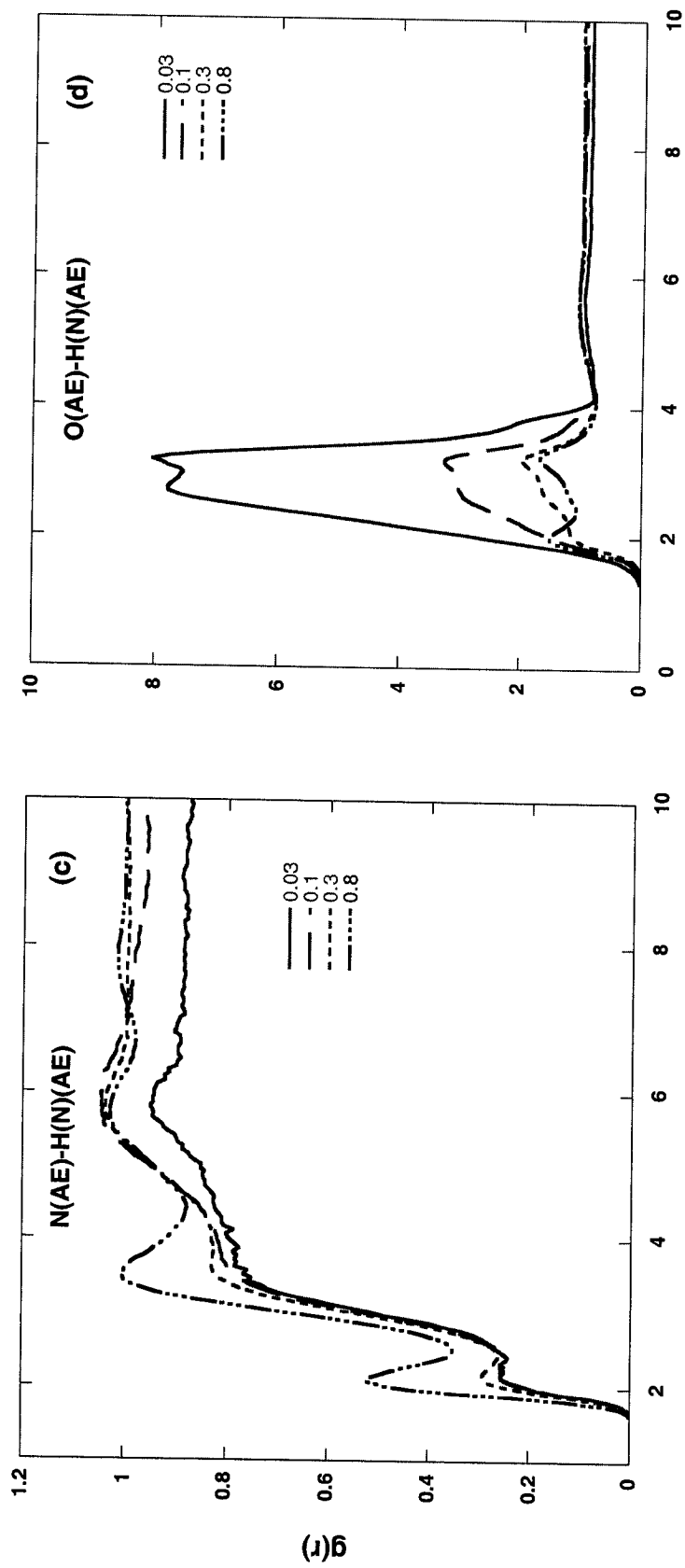


Fig. 30 continued. AE-AE RDF's for (c) nitrogen and hydrogen and for (d) oxygen and hydrogen atoms.

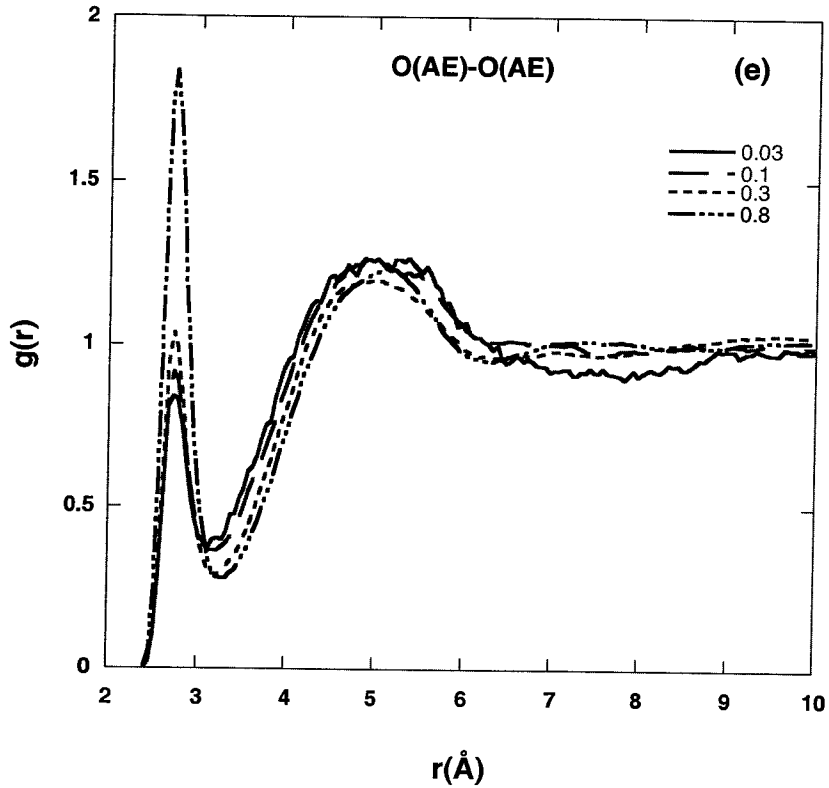


Fig. 30 continued.
(e) AE-AE RDF's for oxygen atoms.

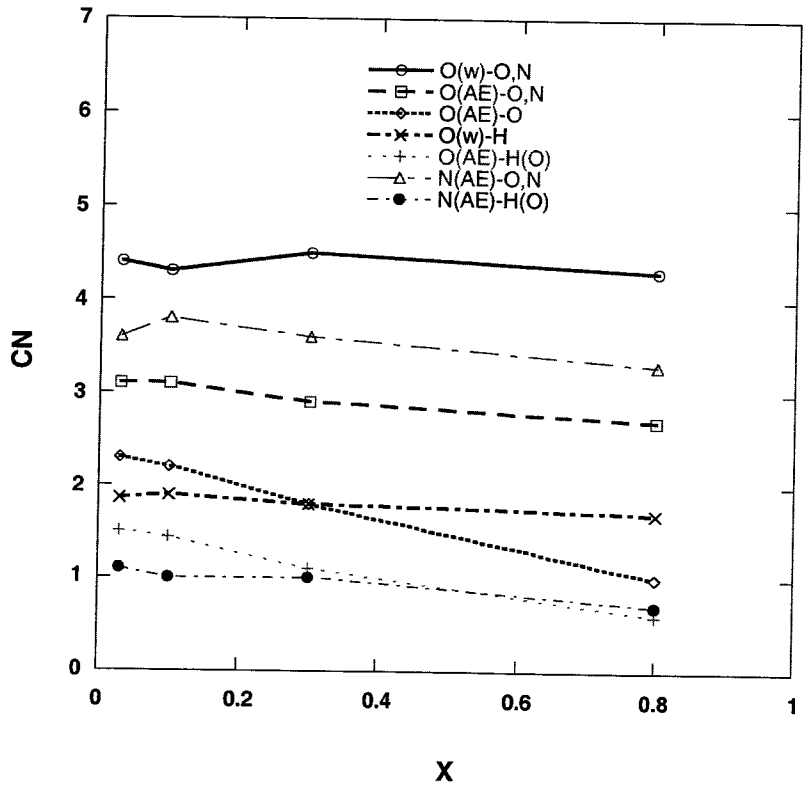


Fig. 31.
Composition dependence of coordination numbers for aqueous solutions of AE at 298 K.

should also be noted that $g(r_{OO})$ of EG-rich solution (see Fig. 26a) indicates the presence of specific features due to water molecules in the second coordination shell of this system which do not appear in the corresponding RDF of AE (see Fig. 30a). A trend of increasing peak height with increasing AE composition somewhat similar to that seen in Fig. 26c, has been observed for the RDF for nitrogen-oxygen AE-water pairs (Fig. 30b). This function exhibits a first peak at 2.9 Å, which corresponds to strongly H-bonded nearest neighbors, but the growth of secondary features (beyond the first peak) can again be clearly seen for AE-rich solutions.

Two further sets of AE-AE RDF's exhibit interesting structural features. The nitrogen-amine hydrogen RDF's (see Fig. 30c) has a small first peak at 2.1 Å corresponding to a relatively weak H-bond which is well defined only for AE-rich solution; it rapidly vanishes and transforms into the poorly defined shoulder as the concentration decreases. This RDF also has a second broad peak at 3.4 Å presumably due to the second hydrogens of NH_2 groups. For the three other less concentrated solutions $g(r_{NH(N)})$ becomes rather featureless maintaining only a broad peak at 5.7 Å. From the oxygen-amine hydrogen RDF shown in Fig. 30d one can notice that, in contrast to the nitrogen-amine hydrogen RDF's of AE but similar to $g(r_{OH})$ of EG solution, the apparent structure decreases with increasing AE concentration. In AE-rich solution the first peak at 2 Å due to H-bonding (followed by a second peak at 3.2 Å) is replaced by a new poorly defined peak at 2.7 Å as concentration decreases while the position of the second peak remains unchanged. In water-rich solution all features overlap into one broad peak with two barely distinguishable maxima at 2.7 and 3.2 Å that suggests a change in average relative orientation of NH_2 to O.

Finally, for the sake of comparison with the analogous RDF's of EG solution, the AE-AE $g(r_{OO})$ is shown in Fig. 30e. The trend observed for the

height of the first peak is opposite to that noted for EG mixtures, i.e. a decrease of intensity with decreasing concentration can be observed as well as a lowering and shift of the first minima from 3.1 (for X=0.03 and 0.1) to 3.3 Å (for X=0.3 and 0.8). This behavior indicates that AE molecules become well hydrated in dilute solution.

The composition dependence of the coordination numbers for AE mixtures is shown in Fig. 31 (detailed numerical results can be found in Table E3). Comparison with the appropriate coordination numbers for EG and ED solutions reveals more apparent similarities than differences. For instance, similar to EG, the total number of oxygens and nitrogens around the water oxygen remains virtually unchanged over the entire composition range (recall that it was also unchanged for EG and increased slightly for ED mixtures). There is no apparent trend of "structure making" in water with the initial addition of AE, as the total coordination around a water molecule in water-rich compositions is 4.4 similar to the coordination in pure water and the corresponding EG mixtures. The total number of hydrogens around water oxygen also remains virtually unchanged (1.9 for water-rich compositions) and drops only slightly (to 1.7) for AE-rich solution. This trend appears to be opposite to that reported for methanol [43] or *tert*-butyl alcohol [154] aqueous solutions where both solutes were found to be strong structure makers in water.

In contrast to water, the coordination numbers around AE oxygen (total or oxygen specific) decrease slightly with concentration (see Fig. 31). An estimate for oxygen-oxygen CN in AE-rich mixtures correlates well with the appropriate coordination number for pure AE. A similar trend can be observed for the total number of hydroxyl hydrogens around AE oxygen; the coordination number drops from 1.5 to 0.6 in going from water-rich to and AE-rich compositions. The analogous coordination around AE nitrogen remains almost unchanged over the

entire composition range. From Fig. 31 one can also see that for water-rich solutions the total coordination around nitrogen in AE is approximately 3.6-3.8, which is about one larger than that around oxygen (3.1); for AE-rich solutions this difference decreases to 0.6. Together these observations imply that the local structure of AE in its aqueous mixtures is somewhat affected by the presence of water molecules. A structural analysis using SDF's will be performed to add more details to the qualitative picture described above.

6.2.2 Solution structure revealed by spatial distribution functions

In the present SDF analysis the site-specific local structure and its dependence on concentration were examined for aqueous solutions of EG, ED and AE. The sets of spatial distribution functions representative of the local water structure around (1) a water molecule and (2) a solute molecule in the solutions of interest, as well as the local solute structure around (3) water and (4) the corresponding solute were visualized and systematically studied. A representative sampling of the most typical (or, in contrast, usual) SDF's selected from these sets are presented in Figs. 32-39.

The compositional behavior of the local water structure around a water molecule as measured by the oxygen-oxygen SDF's for EG, ED and AE solutions reveals more similarities than differences. Two water-water SDF's for X=0.03 and 0.3 solutions of AE are shown in Fig. 32. At an isosurface threshold of 1.6 times that of the bulk one can see three principal features very reminiscent of those observed in pure water [155]: the two distinct caps centered directly over the hydrogens of the central molecule due to its H-bond accepting neighbors and the single large feature below the central molecule due to its two H-bond donating neighbors. The two more distant features in Fig. 32a indicate the presence of additional ("interstitial") coordination also typical for pure liquid [155].

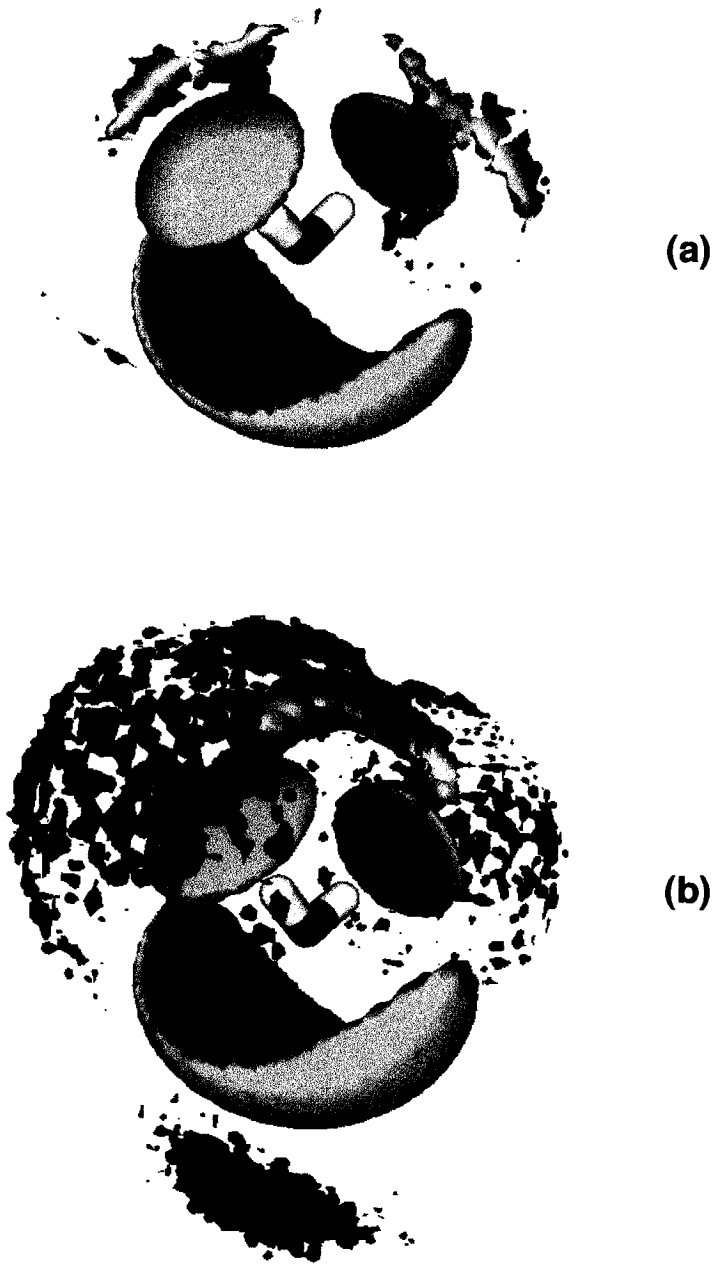


Fig. 32. Local water structure around a water molecule in AE solution as measured by the oxygen-oxygen SDF: (a) mole fraction of 0.03; (b) mole fraction of 0.3. The isosurfaces threshold is 1.6 times that of the bulk.

Results (at the same threshold) for dilute EG and ED solutions (not shown) appear almost identical to Fig. 32a. At higher concentrations (see Fig. 32b) one can see the appearance of secondary structure as two large hemispherical caps surrounding the principal H-bond accepting features and a single feature below the principal H-bond donating neighbors. Interestingly, the distance from the central oxygen to the lower part of the "bridge" structure located between the two caps in the region of the bisecting plane is only about 2.6-2.8 Å; this indicates the presence of a classical "bifurcated" H-bond arrangement [54]. We note that this unusual feature has been observed only for the X=0.3 solution of AE.

Spatial density maps of EG oxygen and ED nitrogen around water oxygen in corresponding EG and ED solutions are shown in Figs. 33 and 34, respectively, at a threshold of 1.6 for low and high compositions. From Fig. 33a one can see that the principal donating feature is splitting into two shallow caps indicating that in the case of dilute EG solution water molecules prefer to accept a H-bond from EG at tetrahedral rather than dipolar (axial) positions. This preference for tetrahedral coordination becomes especially pronounced in Fig. 34a (i.e. in the dilute ED solution) where it manifests itself in two rather distorted (triangular) caps below each of the "lone pair" sites. This behavior can be explained by again considering the balance of H-bond donor and acceptor sites. With an excess of donor sites in ED solutions, it would seem that water, when it bonds to the nitrogen, strongly prefers to utilize both its acceptor sites. One more similarity can be noted in Figs. 33a and 34a. These are two equatorial out-of-plane features, which represent the tendency of a water molecule hydrating the CH_2 groups of EG and ED molecules to lie flat to these hydrophobic surfaces. At the threshold (1.6) used in Figs. 33a and 34a these features are just beginning to appear.

At high concentration (see Figs. 33b and 34b) the hydration picture

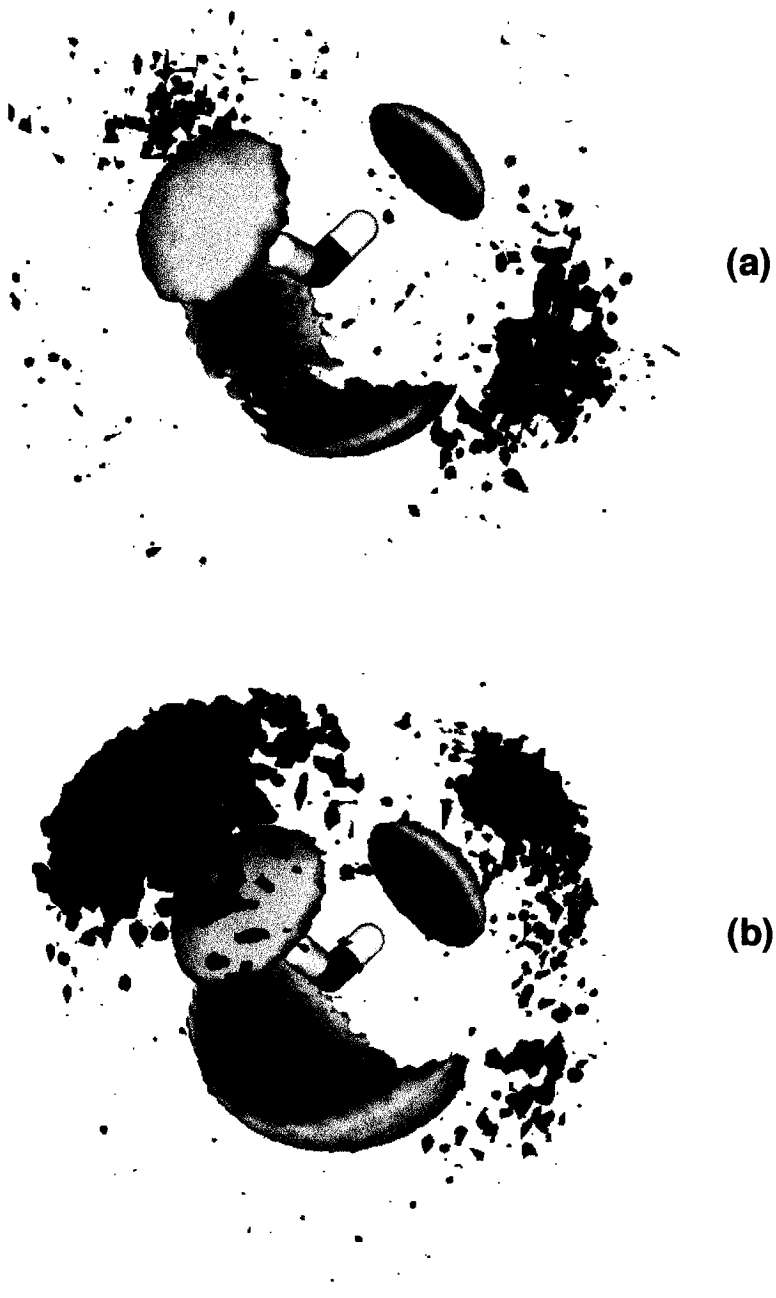


Fig. 33. Spatial density of EG oxygen around water oxygen in EG solutions: (a) mole fraction of 0.03; (b) mole fraction of 0.8. The isosurface threshold is 1.6 times that of the bulk.

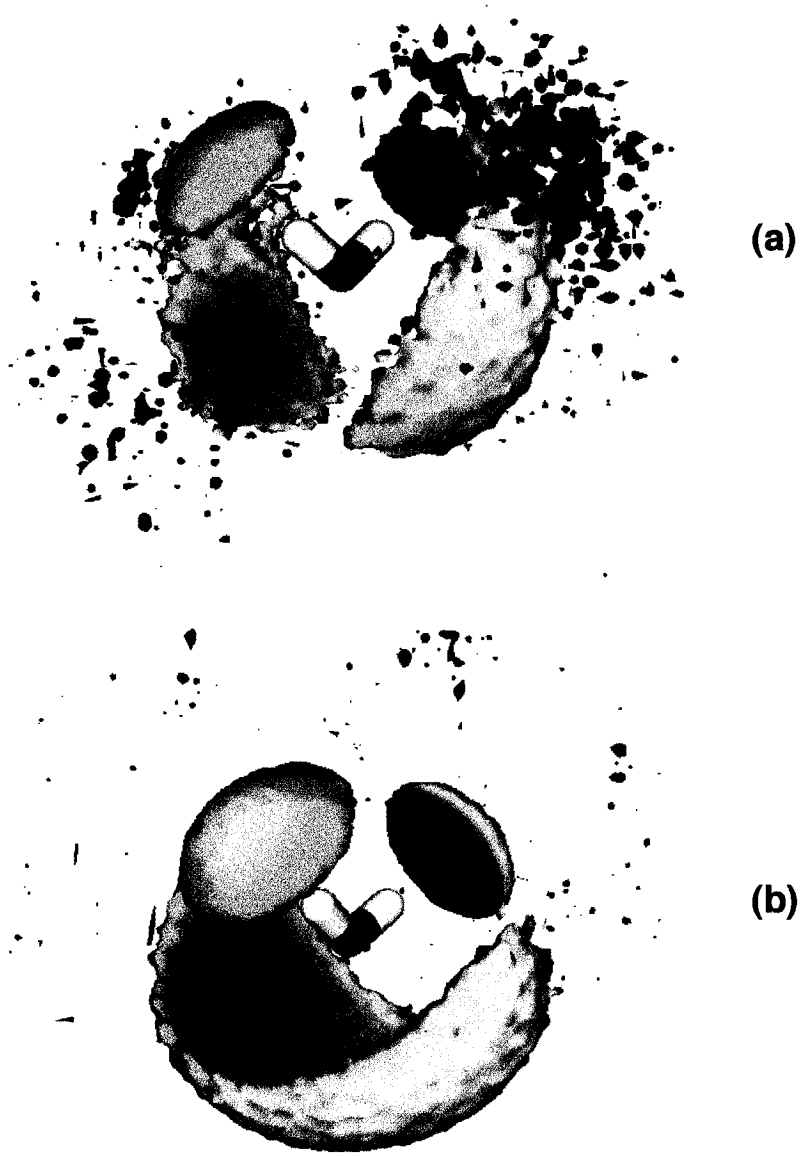


Fig. 34. Spatial density of ED nitrogen around water oxygen in ED solutions: (a) mole fraction of 0.03; (b) mole fraction of 0.8. The isosurface threshold is 1.6 times that of the bulk.

changes dramatically: the tendency for tetrahedral coordination appears to be completely lost for both EG and ED as well as the out-of-plane (hydrophobic hydration) features, especially in the case of ED. However, for EG solution the appearance of secondary caps above the principal H-bond donating features in Fig. 33b indicates a clear trend to hydrophilic hydration at high concentration while in dilute solution hydrophobic hydration becomes prevalent.

Note that the water-solute coordination in AE solution (not shown) is very reminiscent to that observed for the corresponding functional groups in EG and ED solutions. However, there are some subtle differences. For instance, the oxygen-oxygen SDF measured at the highest concentration (X=0.8) still reveals the presence of some hydrophobic structural arrangements which is not seen in Fig. 33b for EG solution.

The local water structure around solute molecules, in particular around the hydroxyl groups of EG and AE and the amino groups of ED and AE, is displayed in Figs. 35 and 36. In Fig. 35a the oxygen-oxygen SDF for EG solution at X=0.1 (which can still be considered as low concentration) exhibits a primary H-bonding pattern very similar to that observed for pure EG (see Fig. 17b), note that the same isosurface threshold allows direct comparison. The rim on the lower edge of principal donating feature as well as two apparent "wings" (which appear as small separate features in equatorial position with respect to the central hydroxyl) are the most prominent additional features in Fig. 35a. They clearly indicate the presence of weakly H-bonded nearest water oxygens and serve as evidence that even in dilute aqueous solution the EG oxygen exhibits a H-bonding pattern similar to that found in the pure system. In contrast to EG, for the AE hydroxyl (see Fig. 35b) only principal H-bonded features are apparent; it is not possible to localize weakly bonded nearest neighbors (which was also the case for the pure liquid).

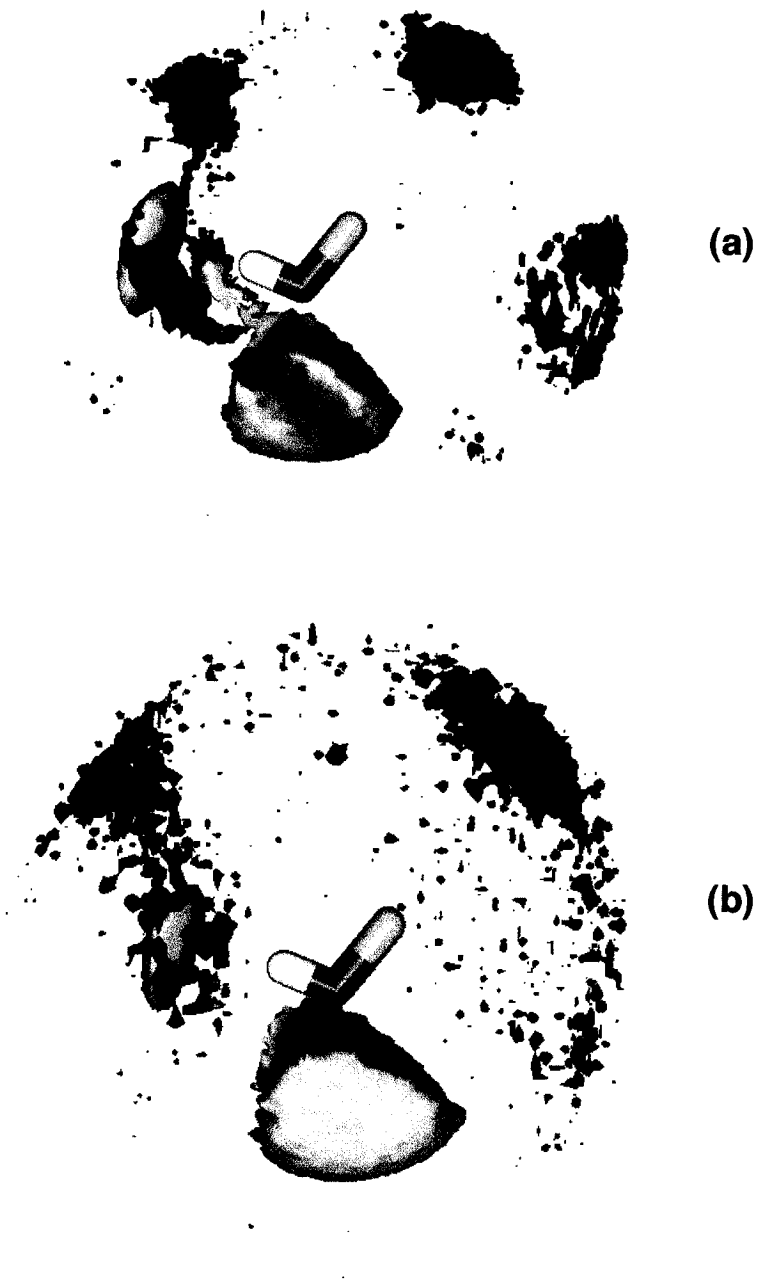


Fig. 35. Local water structure around hydroxyl group in dilute solutions (X=0.1) of (a) EG and (b) AE as measured by the oxygen-oxygen SDF's. The isosurface threshold is 1.8 times that of the bulk.

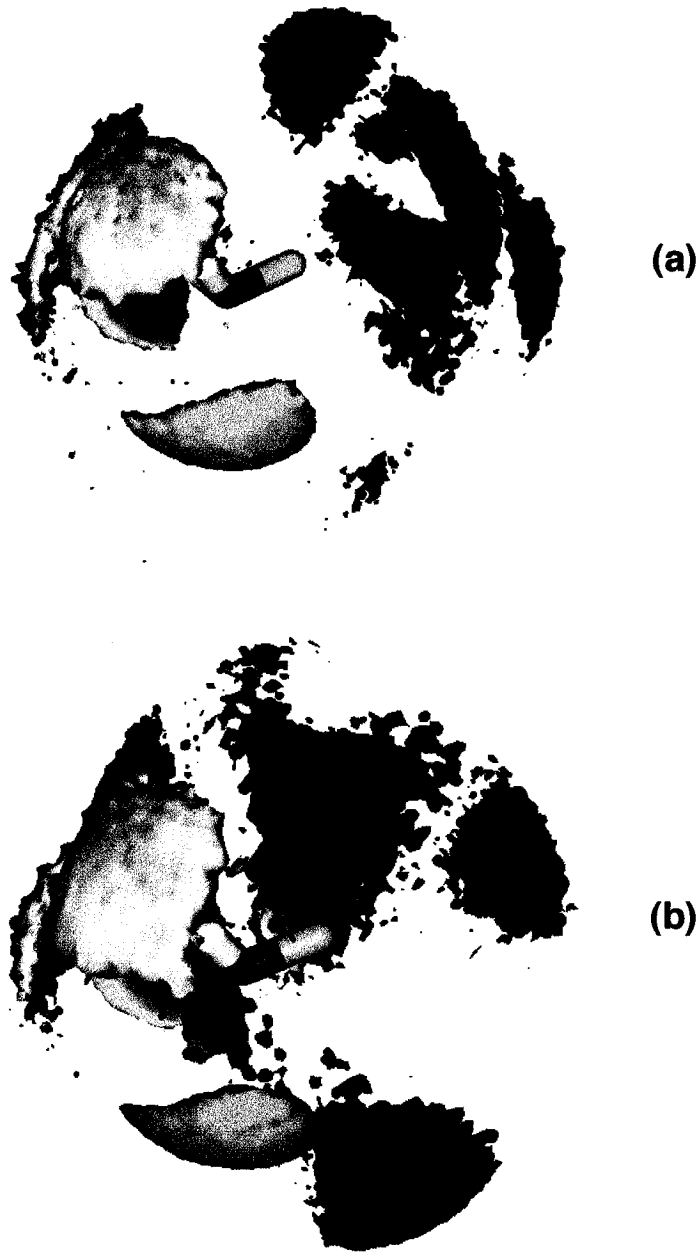


Fig. 36. Local water structure around the amino group in X=0.3 solutions of (a) ED and (b) AE as measured by the nitrogen-oxygen SDF's. The isosurface threshold is 1.8 times that of the bulk.

The local water structure around the amino group was examined for X=0.3 solutions of ED and AE by means of the nitrogen-oxygen SDF at a threshold of 1.8 in Figs. 36a and 36b, respectively. In Fig. 36a the primary H-bonded features are clearly evident as caps over each of two hydrogens and as a slightly elongated feature immediately below the amino group. The ridge extending between the two amino hydrogen caps is particularly interesting because it represents a bridging second neighbor water. It can be expected that this water forms (likely strong) H-bonds with the two neighbors occupying the primary H-bond acceptor positions over the amino group. An analogous H-bonding arrangement around an amino group was observed previously in aqueous solutions of methylamine [54]. It is worth noting that this SDF also allows one to see clearly all the basic H-bonded features due to the hydration of the second intramolecular nitrogen when an ED molecule adopts a classical *trans* conformation. The principal hydration pattern around the amino group in AE solution (see Fig. 36b) is very similar to that shown in Fig. 36a with exception of the secondary features associated with the intramolecular oxygen in *gauche* position. However one can note the presence of two small features located on both sides of the central nitrogen between the single donating feature and the accepting caps. They could be indicative of a first (weak) H-bond donor water neighbor necessary to maintain proper H-bond balance and forced to occupy a rather unexpected position because of the internal geometry of the AE molecule (the water molecule attempting "to bridge" the hydroxyl and amino groups of a molecule in the *gauche* conformation).

Finally, spatial solute-solute correlations for dilute and concentrated solutions of EG, ED and AE are shown in Figs. 37-39. In fact all features (due to both first and second neighboring atoms) which are present in Fig. 37 can be immediately recognized when compared with those identified in Fig. 17 for pure

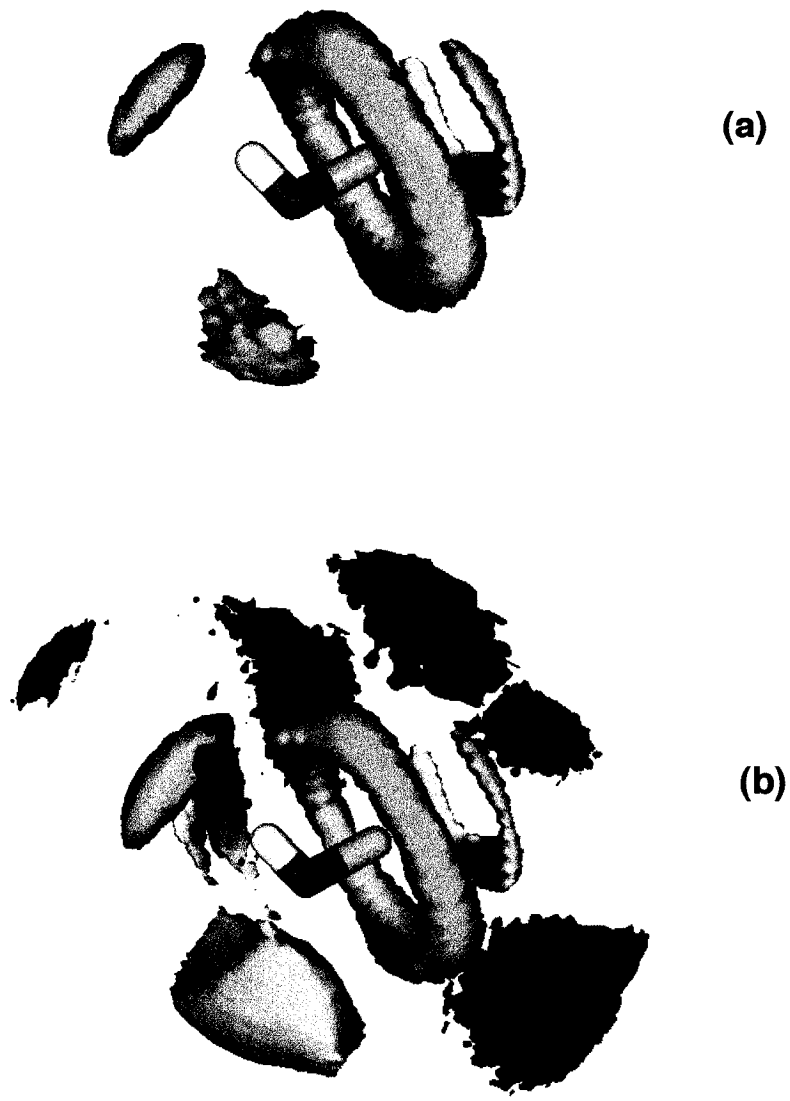


Fig. 37. Oxygen-oxygen spatial distribution functions for EG in (a) dilute (X=0.1) and (b) concentrated (X=0.8) EG solutions at isosurface thresholds of 3.0 and 1.6, respectively.

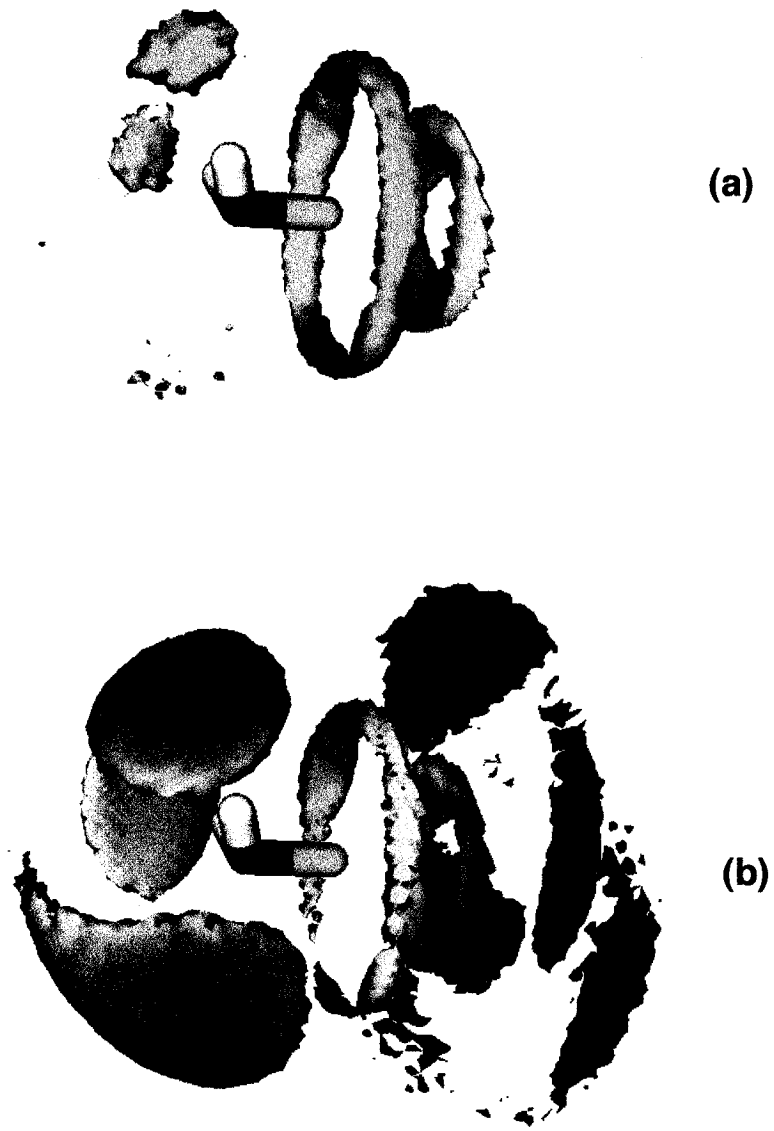


Fig. 38. Nitrogen-nitrogen spatial distribution functions for ED in (a) dilute (X=0.1) and (b) concentrated (X=0.8) ED solutions at isosurface thresholds of 3.0 and 1.6, respectively.

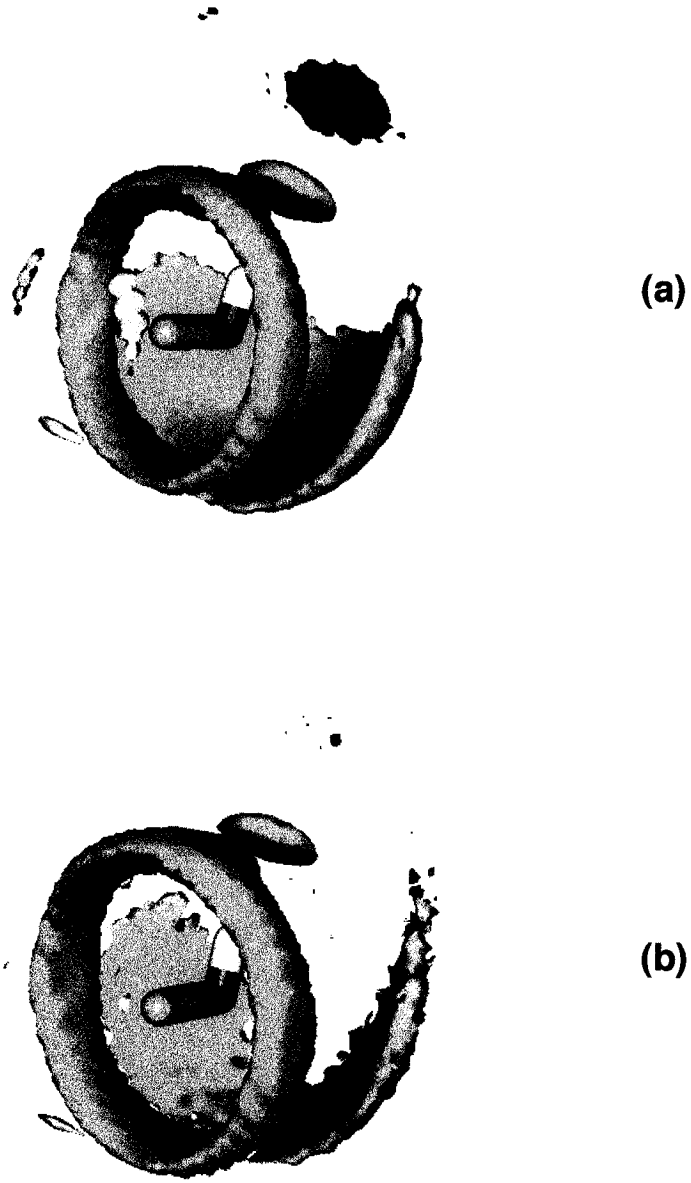


Fig. 39. Oxygen-nitrogen spatial distribution functions for AE in AE aqueous solutions: (a) mole fraction of 0.8 and (b) mole fraction of 0.3 at a threshold of 2.0.

EG. Thus, the presence of water appears to have little qualitative effect on the structural preferences of EG around another EG molecule. At the highest concentration (X=0.8), the nitrogen-nitrogen SDF for ED (see Fig. 38b) and all examined SDF's for AE (only O-N is shown in Fig. 39a) essentially reproduce the main structural patterns found in the respective pure liquids (again with an exception for the coordination of the second amino group of the *trans* conformers in ED solution). However, further analysis of intensities of the main SDF maxima for all compositions reveals a consistent weakening of solute-solute correlations (even in the presence of a small amount of water molecules).

Somewhat surprisingly, a rather dramatic change can be noted in the oxygen-nitrogen SDF of AE when the concentration of the solute decreases to X=0.3 (or below). Comparing Figs. 39a and 39b one can see that the principal donating H-bond feature below its amino group splits into two large caps upon dilution, indicating that the hydroxyl group is better able to utilize both its H-bond acceptor sites, while the well defined feature due to the second neighbor above the hydroxyl (see Fig. 39a) disappears. These observations suggest that the aqueous media is a more accommodating (H-bonding) environment for the AE molecule than is its pure liquid.

CHAPTER VII - CONCLUSIONS

In this thesis a computational investigation of the local structure of three widely used representatives of 1,2-disubstituted ethanes, namely ethylene glycol, ethylenediamine and 2-aminoethanol, in the liquid state and their mixtures with water has been performed, and the results obtained were presented and discussed in a comparative manner. Classical molecular dynamics simulations combined with a novel technique to produce three-dimensional atomic density maps (spatial distribution functions) for the systems of interest were the main computational tools used in the present study.

Previous and current experimental and theoretical studies of 1,2-disubstituted ethanes in gaseous, liquid and solid state were reviewed and analyzed. It has been shown that in the gaseous state EG, ED and AE prefer to adopt *gauche* conformations with respect to the central dihedral angle, which are stabilized mainly by the intramolecular hydrogen bond. In the liquid phase the *gauche* conformation still accounts for the major population of conformers for EG and AE while ED exhibits a significant population of *trans* rotamers. To date the detailed local structure for all three compounds in their pure liquids and aqueous solutions had remained relatively unknown, with EG being the focus of some previous efforts. Therefore, the purpose of this investigate was to determine quantitatively the populations of rotamers (i.e. to clarify the conformational picture) and to obtain a detailed picture of the H-bonding patterns and local structure in pure liquid systems and in aqueous mixtures (at different compositions) for the three compounds of interest.

Several necessary and important steps were performed to achieve this goal. First, twelve molecular models were designed in order to test different

potentials (force fields) and molecular representations. Gas-phase (isolated single molecule) simulations were carried out and simulated geometries and conformations were compared with the most reliable experimental estimates. As a result of this step five of the most successful models where chosen, each of whose molecular descriptions included the AMBER/OPLS-based parameterization for alcohols and amines [6,52].

In the next step liquid-phase simulations for pure systems were performed. The influence of inclusion and of the values of scaling coefficients for the Lennard-Jones and Coulombic contributions to 1-4 non-bonded interactions were examined. The experimental heats of vaporization and self-diffusion coefficients were used as criteria for the final selection of molecular models. It was shown that for EG and AE the best agreement of simulated properties with experimental estimates is achieved when scaling factors of 0.125 and 0.833 were used for the Lennard-Jones and Coulombic terms, respectively; no scaling was necessary in the case of ED. It is worth noting that there had been no previous computational results reported in the literature for the heat of vaporization and self-diffusion coefficient for pure AE.

As an essential part of the structural analysis the dihedral angle distributions were calculated for the peripheral and the central (X-C-C-Y) torsional angles of pure EG, ED and AE and their aqueous mixtures, where four compositions of each compound were considered. Focusing on the central dihedral angle, the integration of the distribution curves allowed us to determine the relative populations of conformers in each of the pure liquids. These results were found to be in very good qualitative agreement with experimental findings. The absence of experimental conformational data for aqueous solutions made the corresponding comparison impossible, however, for a dilute solution of EG excellent agreement with a previously reported theoretical result [59] has been

achieved. The quality of the molecular models used in the present simulations of aqueous solutions has been additionally confirmed by the results obtained for the self-diffusion coefficient. The compositional dependence of the self-diffusion coefficients for EG and ED agreed well with experimentally determined curves (again, no experimental data were available for AE). To our knowledge these are the first theoretical estimates of this characteristic for the aqueous solutions studied.

The investigation of the local structure around EG, ED and AE in pure liquids and aqueous solutions included analysis of radial distribution functions and coordination numbers, as well as production and interpretation of spatial distribution functions. The radial analysis, in conjunction with calculated numbers of nearest neighbors around oxygen and nitrogen atoms of the main functional groups, provided some structural insights into the H-bonding pattern within the pure systems. The number of strongly H-bonded neighboring groups (among all recorded nearest neighbors) was determined and then their possible positions with respect to the reference (central) molecule have been located by means of SDF's. It was found that in EG and ED both nitrogen and oxygen atoms trend to make on average two strong and two weak H-bonds, for which the corresponding neighbor atom densities were successfully identified during the spatial analysis. A similar picture appears to be valid for AE, however the weakly bound nearest neighbors were not clearly identifiable on the corresponding atomic density maps. The possibility of four-membered H-bond arrangements (around a central H-bonding group) leads one to the conclusion that in liquid EG, ED and AE the generalized H-bonding pattern can be described as a three-dimensional, branched network.

On the basis of similar analysis performed for the aqueous solutions of these compounds, an analogous (i.e. three-dimensional) picture of hydration has

been revealed. Furthermore, some interesting trends were noted associated with behavior specific to water acting in this case as an accommodating solvation media. For instance, there is a tendency for the association of a small number of water molecules in water-poor mixtures of EG and AE, the tendency of water molecules to lie flat to the hydrophobic surface of the CH_2 groups (particularly in dilute solution), as well as water ability to maintain its tetrahedral coordination in water-rich solutions. The water-solute structural arrangements revealed a general trend of hydrophilic hydration at high concentrations, while in the dilute solutions hydrophobic hydration becomes clearly apparent for all three types of solute. The solute-solute correlations are mainly characterized by weakening of the intensity of the principal SDF's maxima as the composition is decreased, which indicates a clear preference for each of the solute molecules to be completely hydrated rather than self-associated as it has been observed in aqueous solutions of *tert*-butyl alcohol [154].

Comparison of the present results with those reported in the literature for relatively simple liquid nonelectrolyles and their aqueous solutions [40,43,54,153, 154] revealed some similarities (e.g. nitrogen coordination in the amino groups of methylamine, ED and AE) and differences, the latter presumably due to the degree of cooperativity observed in the systems studied. As in previous work it was found that H-bond balance appears to play a key role in determining the preferred local structures around H-bonding molecules. However, unlike the chaining motif found in the local structures of liquid methanol or small ring-like structures observed for methylamine, the chosen representatives of 1,2-disubstituted ethanes rather accommodate a three-dimensional, branched local environment. This behavior is also apparent in their aqueous solutions.

Some concluding remarks should also be made with respect to applicability of the force fields used in the present work, as well as the utility of

spatial density analysis to this particular class of compounds. The ability of SDF's to provide essential detailed insights into the local structure of strongly associated molecular liquids has again been demonstrated. However, the conformational flexibility (i.e. the presence of three dihedral angles) of EG, ED and AE sometimes does not allow localization and identification of all the neighboring atoms. This makes precise structural elucidation difficult. This problem can be partially resolved in future studies either by producing SDF's for certain (fixed) conformers or, more generally, by generating the atomic density maps which allow inclusion of three-body correlations (i.e. that would fix two functional groups not necessary in the same molecule).

The applicability of the current force fields developed for simple alcohols and amines to molecules containing more than one amino or hydroxyl group (or the combination of both) appears to be a questionable issue. In my opinion the improvement of model potentials should determine the main thrust of future work. It is clear that the general features of the local structure determined in this study with the present models are correct in a qualitative sense. However, since it has been shown that the internal geometry of these molecules can strongly impact their local structure, one can reasonably assume that a better description of the dihedral angles present in these particular compounds would provide more "realistic" molecular models. It would then be possible to capture more subtle features and to better characterize the more important structural details of EG, ED and (especially!) AE.

APPENDIX A

Table A1. Basic structural information for EG [81], ED [81] and AE [86].

Compound	Space group	Z	Lattice parameters (Å, deg.)				C_k*	Lattice energy (kJ/mol)	
			a	b	b	β		H-bond	v.d.Waals
EG	$P2_12_12_1$	4	5.0130	6.9149	9.2710	90	0.693	32.112	23.995
ED	$P2_1/c$	2	5.0467	7.1552	5.4746	115.363	0.717	14.974	35.233
AE	Cc	4	5.0260	8.8180	8.3190	107.440	-	-	-

* The packing factor, C_k.

APPENDIX B

Molecular description file for the EG model

Number of sites

6

	X	Y	Z	Mass	q	σ	ε
C	.0000	.0000	.0000	14.0268	.2650	3.9050	.4940
C	.0000	.0000	1.5260	14.0268	.2650	3.9050	.4940
O	1.3414	.0000	2.0010	15.9994	-.7000	3.0700	.7110
O	.6707	-1.1617	-.4750	15.9994	-.7000	3.0700	.7110
H	1.8878	-.4268	1.3429	1.0079	.4350	.0000	.0000
H	.9089	-1.0124	-1.3887	1.0079	.4350	.0000	.0000

Semi-rigid EG model (with CH_2 as united atoms). Gas-phase geometry by Caminati *at al.* [74], force field for alcohols by W. Jorgensen [6].

Number of bonds

5

ID	N1	N2	Reqv	Kb	D	Rho (=**-1)
0	1	2	1.5260	1087.8400	.0000	.0000
0	2	3	1.4230	1615.0240	.0000	.0000
0	1	4	1.4230	1615.0240	.0000	.0000
0	3	5	.9560	2313.7520	.0000	.0000
0	4	6	.9560	2313.7520	.0000	.0000

Number of angles

4

N1	N2	N3	Aeqv	Ka
1	2	3	109.5000	334.7200
2	1	4	109.5000	334.7200
2	3	5	108.0000	230.1200
1	4	6	108.0000	230.1200

Number of dihedrals

0

Supported types

tors1

3

N1	N2	N3	N4	K1	K2	K3
4	1	2	3	2.9372	-.8870	12.8030
5	3	2	1	3.4895	-.4853	3.1254
6	4	1	2	3.4895	-.4853	3.1254

Molecular description file for the ED model

Number of sites

8

	X	Y	Z	Mass	q	σ	ε
C	.0000	.0000	.0000	14.0268	.1800	3.9050	.4940
C	.0000	.0000	1.5460	14.0268	.1800	3.9050	.4940
N	1.3668	.0000	2.0844	14.0067	-.9000	3.3000	.7110
N	.6205	-1.2178	-.5384	14.0067	-.9000	3.3000	.7110
H	1.5834	-1.2786	-.2167	1.0079	.3600	.0000	.0000
H	.1037	-2.0325	-.2167	1.0079	.3600	.0000	.0000
H	1.8580	.8303	1.7627	1.0079	.3600	.0000	.0000
H	1.3309	.0000	3.1008	1.0079	.3600	.0000	.0000

Semi-rigid ED model (with CH_2 as united atoms). Gas-phase geometry by K.-M. Marstokk, H. Mollendal [36], OPLS force field for amines by W. Jorgensen [52].

Number of bonds

7

ID	N1	N2	Reqv	Kb	D	Rho (=**-1)
0	1	2	1.5460	1087.8400	.0000	.0000
0	2	3	1.4690	1535.5280	.0000	.0000
0	1	4	1.4690	1535.5280	.0000	.0000
0	4	5	1.0170	1815.8560	.0000	.0000
0	4	6	1.0170	1815.8560	.0000	.0000
0	3	7	1.0170	1815.8560	.0000	.0000
0	3	8	1.0170	1815.8560	.0000	.0000

Number of angles

8

N1	N2	N3	Aeqv	Ka
1	2	3	111.5000	334.7200
2	1	4	111.5000	334.7200
1	4	5	109.4800	146.4400
1	4	6	109.4800	146.4400
5	4	6	106.4000	182.4224
2	3	7	109.4800	146.4400
2	3	8	109.4800	146.4400
7	3	8	106.4000	182.4224

Number of dihedrals

0

Supported types

tors1

5

N1	N2	N3	N4	K1	K2	K3
4	1	2	3	10.0081	-2.8200	2.3012
5	4	1	2	-.7950	-1.7447	1.7489
6	4	1	2	-.7950	-1.7447	1.7489
7	3	2	1	-.7950	-1.7447	1.7489
8	3	2	1	-.7950	-1.7447	1.7489

Molecular description file for the AE model

Number of sites

7

	X	Y	Z	Mass	q	σ	ε
C	.0000	.0000	.0000	14.0268	.2225	3.9050	.4940
C	.0000	.0000	1.5260	14.0268	.2225	3.9050	.4940
O	1.2934	.0000	2.0512	15.9994	-.7000	3.0700	.7110
N	.7961	-1.1540	-.4582	14.0067	-.9000	3.3000	.7110
H	1.9099	.5246	1.2499	1.0079	.4350	.0000	.0000
H	.5658	-1.4045	-1.4166	1.0079	.3600	.0000	.0000
H	1.7899	-.9514	-.3831	1.0079	.3600	.0000	.0000

Semi-rigid AE model (with CH_2 as united atoms). Equilibrium geometry and force fields by J. Weiner [113], and by W. Jorgensen (for amines) [62].

Number of bonds

6

ID	N1	N2	Reqv	Kb	D	Rho (=**-1)
0	1	2	1.5260	1087.8400	.0000	.0000
0	2	3	1.4250	1615.0240	.0000	.0000
0	1	4	1.4710	1535.5280	.0000	.0000
0	3	5	.9600	2313.7520	.0000	.0000
0	4	6	1.0100	1815.8560	.0000	.0000
0	4	7	1.0100	1815.8560	.0000	.0000

Number of angles

6

N1	N2	N3	Aeqv	Ka
1	2	3	109.5000	334.7200
2	1	4	111.2000	334.7200
2	3	5	108.5000	230.1200
1	4	6	109.5000	146.4400
1	4	7	109.5000	146.4400
7	4	6	106.4000	182.4224

Number of dihedrals

2

N1	N2	N3	N4	Del	Vn/2	n
4	1	2	3	.0000	8.3680	3
5	3	2	1	.0000	2.0920	3

Supported types

tors1

2

N1	N2	N3	N4	K1	K2	K3
6	4	1	2	-.7950	-1.7447	1.7489
7	4	1	2	-.7950	-1.7447	1.7489

Molecular description file for the AEcmb model

Number of sites

7

	X	Y	Z	Mass	q	σ	ε
C	.0000	.0000	.0000	14.0268	.2225	3.9050	.4940
C	.0000	.0000	1.5260	14.0268	.2225	3.9050	.4940
O	1.2934	.0000	2.0512	15.9994	-.7000	3.0700	.7110
N	.7961	-1.1540	-.4582	14.0067	-.9000	3.3000	.7110
H	1.9099	.5246	1.2499	1.0079	.4350	.0000	.0000
H	.5658	-1.4045	-1.4166	1.0079	.3600	.0000	.0000
H	1.7899	-.9514	-.3831	1.0079	.3600	.0000	.0000

Semi-rigid AE model (with CH_2 as united atoms). Equilibrium geometry and force field (for alcohols and amines) by W. Jorgensen [6,52].

Number of bonds

6

ID	N1	N2	Reqv	Kb	D	Rho (=**-1)
0	1	2	1.5260	1087.8400	.0000	.0000
0	2	3	1.4250	1615.0240	.0000	.0000
0	1	4	1.4710	1535.5280	.0000	.0000
0	3	5	.9600	2313.7520	.0000	.0000
0	4	6	1.0100	1815.8560	.0000	.0000
0	4	7	1.0100	1815.8560	.0000	.0000

Number of angles

6

N1	N2	N3	Aeqv	Ka
1	2	3	109.5000	334.7200
2	1	4	111.2000	334.7200
2	3	5	108.5000	230.1200
1	4	6	109.5000	146.4400
1	4	7	109.5000	146.4400
7	4	6	106.4000	182.4224

Number of dihedrals

0

Supported types

tors1

4

N1	N2	N3	N4	K1	K2	K3
4	1	2	3	2.9372	-.8870	12.8031
5	3	2	1	3.4895	-.4853	3.1254
6	4	1	2	-.7950	-1.7447	1.7489
7	4	1	2	-.7950	-1.7447	1.7489

Molecular description file for the AEtst model

Number of sites

7

	X	Y	Z	Mass	q	σ	ε
C	.0000	.0000	.0000	14.0268	.2570	3.9050	.4940
C	.0000	.0000	1.5400	14.0268	.2570	3.9050	.4940
O	1.3558	.0000	1.9623	15.9994	-.6440	3.0700	.7110
N	.5772	-1.1500	-.6466	14.0067	-.9380	3.2500	.7110
H	1.7021	.8773	1.7834	1.0079	.3740	.0000	.0000
H	.5279	-1.0513	-1.6405	1.0079	.3470	.0000	.0000
H	1.5355	-1.2541	-.3805	1.0079	.3470	.0000	.0000

Semi-rigid AEtst model (with CH_2 as united atoms). Equilibrium geometry and force field by J. Alejandre [67].

Number of bonds

6

ID	N1	N2	Reqv	Kb	D	Rho (=**-1)
0	1	2	1.5400	1087.8400	.0000	.0000
0	2	3	1.4200	1615.0240	.0000	.0000
0	1	4	1.4400	1535.5280	.0000	.0000
0	3	5	.9600	2313.7520	.0000	.0000
0	4	6	1.0000	1815.8560	.0000	.0000
0	4	7	1.0000	1815.8560	.0000	.0000

Number of angles

6

N1	N2	N3	Aeqv	Ka
1	2	3	107.3000	546.8000
2	1	4	116.6800	506.1200
2	3	5	106.8000	298.7300
1	4	6	110.3500	316.5800
1	4	7	110.3500	316.5800
7	4	6	108.4800	269.3300

Number of dihedrals

0

Supported types

tors5

3

N1	N2	N3	N4	K1	K2	K3	K4	K5
7	4	1	2	11.05	-2.37	-25.29	0.43	2.01
4	1	2	3	17.78	23.60	-36.80	-16.63	13.32
1	2	3	5	19.14	-2.79	-13.61	-0.12	-1.55

APPENDIX C

Table C1. Experimental densities, ρ (g/cm^3), used in simulations of pure EG, ED and AE and their mixtures with water [135,136,137].

Comp. (X)	EG	ED	AE
0.03	1.0070	0.9942	1.0007
0.1	1.0285	0.9943	1.0092
0.3	1.0638	0.9904	1.0242
0.8	1.0950	0.9168	1.0178
1.0	1.1000	0.8990	1.0118

Table C2. Number density, ρ_n (N/Å^3), used to calculate the coordination numbers (CN) for aqueous solutions.

	Solute-solvent			Solvent-solute		
Comp. (X)	EG	ED	AE	EG	ED	AE
0.03	0.0009415	0.0009324	0.0009371	0.03044	0.03015	0.03030
0.1	0.002764	0.002696	0.002741	0.02488	0.02426	0.02452
0.3	0.006156	0.005841	0.005984	0.01436	0.01363	0.01396
0.8	0.009906	0.008547	0.009346	0.002476	0.002138	0.002336

APPENDIX D

Table D1. Self-diffusion coefficients, D ($10^{-5}cm^2/s$), for aqueous solutions of ED, EG and AE.

Comp. (X)	ED		EG		AE
	experimental[a]	calculated	experimental[b]	calculated	calculated
0.03	1.690	1.036	1.0135	1.0590	1.2450
0.1	0.775	0.8146	0.7720	0.8207	0.9706
0.3	0.222	0.5699	0.3590	0.4924	0.6932
0.8	0.646	0.7172[c]	0.1175	0.2669	0.7950
1.0	0.980	1.000	0.0900	0.0800	0.5188

a - [150]; b - [9]; c – value obtained from constant-pressure simulation run.

Table D2. Conformational characteristics of AE, EG and ED in aqueous solutions. All values for dihedral angles correspond to the maximum probability on dihedral angle distributions.

Comp. (X)		EG			ED			AE		
	Conformation	G'	G	T	G'	G	T	G'	G	T
0.03	(%)	8.7	35.3	56.0	5.1	28.3	66.6	51.1	48.9	-
	Dihedral (deg.)	-62.5	62.5	177.5	-67.5	67.5	177.5	-57.5	52.5	-
0.1	(%)	71.3	28.7	-	28.3	1.9	69.8	36.0	61.5	2.5
	Dihedral (deg.)	-62.5	62.5	-	-72.5	67.5	177.5	-52.5	52.5	172.5
0.3	(%)	9.0	91.0	-	7.3	-	92.7	43.7	56.3	-
	Dihedral (deg.)	-62.5	62.5	-	-67.5	-	182.5	-27.5	52.5	-

APPENDIX E

Table E1. Coordination numbers for EG aqueous solutions. The first site listed identifies the central atom while the second is the coordinating atom.

	Composition (X)			
	0.03	0.1	0.3	0.8
$O(H_2O)$-$O(H_2O)$	4.2	3.9	2.9	0.7
O(EG)-$O(H_2O)$	2.6	2.3	1.6	0.4
$O(H_2O)$-O(EG)	0.2	0.5	1.4	3.2
O(EG)-O(EG)	1.0	1.2	1.7	2.7
$O(H_2O)$-H(EG)	0.1	0.2	0.5	1.1
H(EG)-$O(H_2O)$	0.9	0.8	0.6	0.1
$O(H_2O)$-$H(H_2O)$	1.8	1.7	1.2	0.3
O(EG)-$H(H_2O)$	1.3	1.2	0.8	0.2
$H(H_2O)$-O(EG)	0.04	0.1	0.3	0.8
O(EG)-H(EG)	-	0.2	0.4	1.1

Table E2. Coordination numbers for ED aqueous solutions. Labels are as in Table E1.

	Composition (X)			
	0.03	0.1	0.3	0.8
$O(H_2O)$-$O(H_2O)$	4.4	3.9	2.8	0.8
N(ED)-$O(H_2O)$	3.2	3.0	2.3	0.5
$O(H_2O)$-N(ED)	0.2	0.7	2.0	4.0
N(ED)-N(ED)	0.3	0.5	1.2	3.3
$O(H_2O)$-H(ED)	0.1	0.2	0.7	1.6
H(ED)-$O(H_2O)$	0.7	0.6	0.4	0.1
$O(H_2O)$-$H(H_2O)$	1.8	1.7	1.2	0.4
N(ED)-$H(H_2O)$	1.1	1.0	0.8	0.2
$H(H_2O)$-N(ED)	0.04	0.1	0.3	0.7
N(ED)-H(ED)	-	0.04	0.2	0.4
H(ED)-N(ED)	-	0.1	0.3	0.8

Table E3. Coordination numbers for AE aqueous solutions. Labels are as in Table E1.

	Composition (X)			
	0.03	0.1	0.3	0.8
$O(H_2O)$-$O(H_2O)$	4.2	3.8	2.9	0.8
O(AE)-$O(H_2O)$	2.3	2.1	1.5	0.4
$O(H_2O)$-O(AE)	0.1	0.2	0.7	1.6
O(AE)-O(AE)	0.03	0.1	0.3	0.6
$O(H_2O)$-H(AE)	0.06	0.2	0.5	1.3
H(AE)-$O(H_2O)$	0.5	0.5	0.4	0.1
$O(H_2O)$-$H(H_2O)$	1.8	1.7	1.3	0.4
O(AE)-$H(H_2O)$	1.5	1.4	1.0	0.3
$H(H_2O)$-O(AE)	0.02	0.1	0.2	0.6
O(AE)-H(O)(AE)	0.01	0.04	0.1	0.3
O(AE)-H(N)(AE)	-	-	-	0.6
N(AE)-N(AE)	0.09	0.2	0.4	1.1
N(AE)-H(N)(AE)	-	-	0.1	0.3
N(AE)-H(O)(AE)	0.4	0.4	0.5	0.6
N(AE)-$H(H_2O)$	0.7	0.6	0.5	0.1
$H(H_2O)$-N(AE)	0.01	0.04	0.1	0.2
N(AE)-O(AE)	0.8	0.9	1.1	1.7
N(AE)-$O(H_2O)$	2.7	2.7	2.1	0.5
$O(H_2O)$-N(AE)	0.1	0.3	0.9	1.9

REFERENCES

1. L. Radom, W.A. Lathan, W.J. Hehre, J.A. Pople, *J. Am. Chem. Soc.*, **95**, 693 (1973).
2. L.A.E. Batista de Carvalho, L.E. Lourenco, M.P.M. Marques, *J. Mol. Struct.*, **482-483**, 639 (1999).
3. Kirk-Othmer, *Encyclopedia of Chemical Technology,* 3rd ed., **1**, John Wiley and Sons, New York (1978).
4. L. Saiz, J.A. Padro, E. Guardia, *J. Chem. Phys.*, **114**, 3187 (2001).
5. B.J. Teppen, M. Cao, R.F. Frey, Ch. van Alsenoy, D.M. Miller, L. Schafer, *J. Mol. Struct. (Theochem)*, **314**, 169 (1994).
6. W.L. Jorgensen, *J. Phys. Chem.*, **90**, 1276 (1986).
7. R.G. Laughlin, in *Advances in Liquid Crystals,* ed. G.H. Brown, Academic Press, New York, 41 (1978).
8. R.G. Laughlin, in *Advances in Liquid Crystals*, ed. G.H. Brown, Academic Press, New York, 99 (1978).
9. L. Ambrosone, G. D'Errico, R. Sartorio, L. Costantino, *J. Chem. Soc. Faraday Trans.*, **93**, 3961 (1997).
10. B.M. Ladanyi, M.S. Skaf, *Annu. Rev. Phys. Chem.*, **44**, 335 (1993).
11. J.A. Padro, L. Saiz, E. Guardia, *J. Mol. Struct.*, **416**, 243 (1997).
12. L. Saiz, J.A. Padro, E. Guardia, *J. Phys. Chem.*, B **101**, 78 (1997).
13. L. Saiz, J.A. Padro, E. Guardia, *Mol. Phys.*, **97**, 897 (1999).
14. L. Saiz, E. Guardia, J.A. Padro, *J. Chem. Phys.*, **113**, 2814 (2000).
15. Y. Hamada, M. Tsuboi, M. Nakata, M. Tasumi, *J. Mol. Spectrosc.*, **106**, 164 (1984).
16. M.N. Rodnikova, A.N. Isaev, S.A. Zasypkin, *Koordinatsionnaya Khimiya*, **17**, 1467 (1991). Journal written in Russian.
17. Y.-P. Chang, T.-M. Su, T.-W. Li, I. Chao, *J. Phys. Chem.*, **101**, 6107 (1997).
18. S. Kudoh, M. Takayanagi, M. Nakata, T. Ishibashi, M. Tasumi, *J. Mol. Struct.*, **479**, 41 (1999).

19. P. Bultinck, A. Goeminne, D.V. Vondel, *J. Mol. Struct.*, **339**, 1 (1995).

20. T.-S. Yeh, Y.-P. Chang, T.M. Su, *J. Phys. Chem.*, **98**, 8921 (1994).

21. C.J. Cramer, D.G. Truhlar, *J. Am. Chem. Soc.*, **116**, 3892 (1994).

22. P.J. Kreuger, H.D. Mettee, *Can. J. Chem.*, **43**, 2970 (1965).

23. R.E. Penn, R.F. Curl, *J. Chem.Phys.*, **55**, 651 (1971).

24. A.-M. Kelterer, M. Ramek, *J. Mol. Struct. (Theochem)*, **232**, 189 (1991).

25. G. Buemi, *Int. J. Quantum Chem.*, **59**, 227 (1996).

26. I. Vorobyov, M. C. Yappert, D.B. DuPre, *J. Phys. Chem.*, **106**, 668 (2002).

27. C.F.P. Silva, M.L.T.S. Duarte, R. Fausto, *J. Mol. Struct.*, **482-483**, 591 (1999).

28. Z. Zhu, M. E. Tuckerman, Sh. O. Samuelson, G. J. Martyna, *Phys. Rev. Lett.*, **88**, 100201-1 (2002).

29. P.I. Nagy, W.J. Dunn III, G. Alagona, C. Ghio, *J. Am. Chem. Soc.*, **113**, 6719 (1991).

30. A.W. Burgess, L.L. Shipman, R.A. Nemenoff, H.A. Scheraga, *J. Am. Chem. Soc.*, **98**, 23 (1976).

31. J.-H. Lii, K.-H. Chen, T.B. Grindley, N.L. Allinger, *J. Comp. Chem.*, **24**,1490 (2003).

32. S. Wolfe, *Acc. Chem. Res.*, **5**, 102 (1972).

33. K.-M. Marstokk, H. Molendall, *J. Mol. Struct.*, **22**, 301 (1974).

34. O. Bastiansen, *Acta Chem. Scand.*, **3**, 415 (1949).

35. K. Hagen, K. Hedberg, *J. Am. Chem. Soc.*, **95**, 8263 (1973).

36. K.-M. Marstokk, H. Mollendal, *J. Mol. Struct.*, **49**, 221 (1978).

37. L. Phillips, V. Wray, *J. Chem. Soc. Chem. Commun.*, **90**, 1 (1973).

38. A.H. Narten, A. Habenschuss, *J. Chem. Phys.*, **80**, 3387 (1984).

39. W.C. Hamilton, J.A. Ibers, *Hydrogen bonding in solids*, Benjamin, New York (1968).

40. M. Haughney, M. Ferrario, I.R. McDonald, *J. Phys. Chem.*, **91**, 4934 (1987).

41. W.L. Jorgensen, *J. Am. Chem. Soc.*, **103**, 341 (1981).

42. I.M. Svishchev, P.G. Kusalik, *J. Chem. Phys.*, **100**, 5165 (1994).

43. A. Laaksonen, P.G. Kusalik, I.M. Svishchev, *J. Phys. Chem.*, A **101**, 5910 (1997).

44. W.L. Jorgensen, J.D. Madura, *J. Am. Chem. Soc.*, **105**, 1407 (1983).

45. A.K. Soper, J.L. Finney, *Phys. Rev. Lett.*, **71**, 4346 (1993).

46. W.J. Dunn, P.I. Nagy, *J. Phys. Chem.*, **94**, 2099 (1990).

47. W.J. Dunn, P.I. Nagy, E. Collantes, *J. Am. Chem. Soc.*, **113**, 7898 (1991).

48. P.I. Nagy, *Acta Chim.*, Hung. **129**, 429 (1992).

49. M. Kawata, S. Ten-no, S. Kato, F. Hirata, *Chem. Phys.*, **203**, 53 (1996).

50. P.Y. Morgantini, P.A. Kollman, *J. Am. Chem. Soc.*, **117**, 6057 (1995).

51. T.I. Spector, P.A. Kollman, *J. Phys. Chem.*, B **102**, 4004 (1998).

52. C.R. Rizzo, W.L. Jorgensen, *J. Am. Chem. Soc.*, **121**, 4827 (1999).

53. R. Impey, M. Sprik, M. Klein, *J. Am. Chem. Soc.*,109, 5900 (1987).

54. P.G. Kusalik, D. Bergman, A. Laaksonen, *J. Chem. Phys.*, **113**, 1 (2000).

55. H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, *J. Phys. Chem.*, **91**, 6269 (1987).

56. G. Widmalm, R.W. Pastor, *J. Chem. Soc. Faraday Trans.*, **88**, 1747 (1992).

57. H. Hayashi, H. Tanaka, K. Nakanishi, *Fluid Phase Equilibria*, **104**, 421 (1995).

58. N. Chandrasekhar, P. Krebs, *J. Chem. Phys.*, **112**, 5910 (2000).

59. R.W.W. Hooft, B.P. van Eijck, J. Kroon, *J. Chem. Phys.*, **97**, 3639 (1992).

60. W.L. Jorgensen, J.D. Madura, *Mol. Phys.*, **56**, 1381 (1985).

61. W.F. van Gunsteren, H.J.C. Berendsen, *Groningen Molecular Simulation* (GROMOS), Library Manual, Groningen (1987).

62. H.J.C. Berendsen, J.P.M. Posta, W.F. van Gunsteren, J.Hermans, *Intermolecular Forces*, Reidel, Dordrecht (1981).

63. G. Alagona, C. Ghio, *J. Mol. Struct. (Teochem)*, **254**, 287 (1992).

64. A.A. Varnek, G. Wipff, A.S. Glebov, D. Feil, *J. Comp. Chem.*, **16**, 1 (1995).

65. S. Miertus, E. Scrocco, J. Tomasi, *Chem. Phys.*, **55**, 117 (1981).

66. J.K. Button, K.A. Gubbins, H.Tanaka, K. Nakanishi, *Fluid Phase Equilibria*, **116**, 320 (1996).

67. J. Alejandre, J.L. Rivera, M.A. Mora, V. de la Garza, *J. Phys. Chem.*, B **104**, 1332 (2000).

68. L.V. Lanshina, M.N. Rodnikova, K.T. Dudnikova, *J. Struct. Chem.*, **30**, 684 (1989).

69. A. Yokozeki, K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **43**, 2664 (1970).

70. A. Yokozeki, K. Kuchitsu, *Bull. Chem. Soc. Jpn.*, **44**, 2926 (1971).

71. S. Jamet-Delcroix, *Acta Cryst.*, B **29**, 977 (1973).

72. R. Righini, S. Califano, *Chem. Phys.*, **17**, 45 (1976).

73. M.Jouan, N.Q. Dao, *C.R. Acad. Sci. Paris,* C **279**, 1987 (1972).

74. W. Caminati, G. Corbelli, *J. Mol. Spectrosc.*, **90**, 572 (1981).

75. D. Christen, L.H. Coudert, R.D. Suenram, F.J. Lovas, *J. Mol. Spectrosc.*, **172**, 57 (1995).

76. P.-E. Kristiansen, K.-M. Marstokk, H. Mollendal, *Acta Chem. Scand.*, A **41**, 403 (1987).

77. V. Crupi, M.P. Jannelli, S. Magazu, G. Maisano, D. Majolino, P. Migliardo, D. Sirna, *Mol. Phys.*, **84**, 645 (1995).

78. V. Crupi, G. Maisano, D. Majolino, P. Migliardo, V. Venuti, *J. Phys. Chem.*, A **104**, 3933 (2000).

79. P. Buckley, P.A. Giguere, *Can. J. Chem.*, **45**, 397 (1967).

80. M. Rasanen, A. Aspiala, J. Homanen, J. Murto, *J. Mol. Struct.*, **96**, 81 (1982).

81. V.R. Thalladi, R. Boese, H.-C. Weiss, *Angew. Chem. Int. Ed.*, **39**, 918 (2000).

82. R. Boese, H.-C. Weiss, D. Blaser, *Angew. Chem.*, **111**, 1042 (1999); *Angew. Chem. Int. Ed.*, **38**, 988 (1999).

83. C.P. Brock, L.L. Duncan, *Chem. Mater.*, **6**, 1307 (1994).

84. R. Boese, H.-C. Weiss, *Acta Cryst.*, C **54**, IUC 9800024 (1998).

85. J.L.G. de Paz, J. Elguero, C. Foces-Foces, A.L. Llamas-Saiz, F. Aguilar-Parrila, O. Klein, H.-H. Limbach, *J. Soc. Perkin Trans.*, **2**, 101 (1997).

86. D. Mootz, D. Brodalla, M. Weibcke, *Acta Cryst.*, C **45**, 754 (1989).

87. P.B. Balbuena, J.M. Seminario (Editors), Theoretical and Computational Chemistry, V **7**, *Molecular Dynamics: From Classical to Quantum Methods*, Elsevier, Amsterdam (1999).

88. D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, New York (1986).

89. J.P. Hansen, I.R. McDonald, *Theory of Simple Liquids*, Academic Press, London (1986).

90. J.E. Enderby, *Ann. Rev. Phys. Chem.*, **34**, 155 (1983).

91. A.K. Soper, *Discuss. Faraday Soc.*, **103**, 41 (1996).

92. A.K. Soper, C.Andreani, M. Nardone, *Phys. Rev.*, **47**, 2598 (1993).

93. A.K. Soper, *J. Chem. Phys.*, **101**, 6888 (1994).

94. R.L. McGreevy, *J. Phys.: Condens. Matter*, **13**, R877 (2001).

95. A.K. Soper, *Chem. Phys.*, **202**, 295 (1996).

96. C.G. Gray, K.E. Gubbins, *Theory of Molecular Liquids. Vol. I: Fundamentals*, Oxford University Press, New York (1984).

97. R.L. McGreevy, L. Pusztai, *Mol. Sim.*, **1**, 359 (1988).

98. J.P. Shih, S.Y. Sheu, C.Y. Mou, *J. Chem. Phys.*, **100**, 2202 (1994).

99. I.M. Svishchev, P.G. Kusalik, *J. Chem. Phys.*, **99**, 3049 (1993).

100. T. Hata, Y. Ono, *Chem. Pharm. Bull.*, **48**, 1660 (2000).

101. M.P. Allen, D.J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford (1987).

102. R.J. Sadus, *Molecular Simulation of Fluids. Theory, Algorithms and Object-Orientation*, Elsevier Science B.V. , Amsterdam (1999).

103. L. Verlet, *Phys Rev.*, **159**, 98 (1967).

104. R.W. Hockney, *Methods Comput. Phys.*, **9**, 136 (1970).

105. W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson, *J. Chem. Phys.*, **76**, 637 (1982).

106. C.W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, New Jersey (1971).

107. S.W. de Leeuw, J.W. Perram, H.G. Petersen, *J. Stat. Phys.*, **61**, 1203 (1990).

108. J.-P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, *J. Comput. Phys.*, **23**, 327 (1977).

109. G. Ciccotti, . J.-P. Ryckaert, *Comp. Phys. Rep.*, **4**, 345 (1986).

110. S. Nosé, *J. Chem. Phys.*, **81**, 511 (1984).

111. W.G. Hoover, *Phys. Rev.*, A **31**, 1695 (1985).

112. A.P. Lyubartsev, A. Laaksonen, *Comput. Phys. Commun.*, **128**, 565 (2000).

113. S.J. Weiner, P.A. Kollman, D.A. Case, U.Ch. Singh, C. Ghio, G. Alagona, S. Profeta, Jr., P. Weiner, *J. Am. Chem. Soc.*, **106**, 765 (1984).

114. S.J. Weiner, P.A. Kollman, D.T. Nguyen, D.A. Case, *J. Comp. Chem.*, **7**, 230 (1986).

115. B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, *J. Comp. Chem.*, **4**, 187 (1983).

116. A.D. MacKerell, Jr., J. Wiorkiewicz-Kuczera, M. Karplus, *J. Am. Chem. Soc.*, **117**, 11946 (1995).

117. G.J. Martyna, D.J.Tobias, M.L. Klein, *J. Chem. Phys.*, **101**, 4177 (1994).

118. M. Tuckerman, B.J. Berne, G.J. Martyna, *J. Chem. Phys.*, **97**, 1990 (1992).

119. W.L. Jorgensen, J. Tirado-Rives, *J. Am. Chem. Soc.*, **110**, 1657 (1988).

120. W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.*, **118**, 11225 (1996).

121. a. N.L. Allinger, Y.H. Yuh, J.-H. Lii, *J. Am. Chem. Soc.*, **111**, 8551 (1989); b. J.-H. Lii, N.L. Allinger, *Ibid.* 8566; c. . J.-H. Lii, N.L. Allinger, *Ibid.* 8576.

122. J.-P. Ryckaert, A. Bellemans, *Chem. Phys. Lett.*, **30**, 123 (1985).

123. W. Smith, *Comput. Phys. Commun.*, **62**, 229 (1991).

124. D. Fincham, *Mol. Simul.*, **1**, 1 (1987).

125. http://laaksonen.csc.fi/gopenmol/gopenmol. html.

126. J.C. Smith, M. Karplus, *J. Am. Chem. Soc.*, **114**, 801 (1992).

127. N.L. Allinger, *J. Am. Chem. Soc.*, **99**, 8127 (1977).

128. A.R. Leach, *Molecular Modelling. Principles and Applications*, Longman (1996).

129. P.-Y. Morgantini, P.A. Kollman, *J. Am. Chem. Soc.*, 117, 6057 (1995).

130. Y.B. Ding, D.N. Bernardo, K. Krogh-Jespersen, R.M. Levy, *J. Phys. Chem.*, **99**, 11575 (1995).

131. W.L. Jorgensen, J.D. Madura, C.J. Swenson, *J. Am. Chem. Soc.*, **106**, 6638 (1984).

132. W.D. Cornell, P. Cieplak, Ch.I. Bayly, P.A. Kollman, *J. Am. Chem. Soc.*, **115**, 9620 (1993).

133. S.E. DeBolt, P.A. Kollman, *J. Am. Chem. Soc.*, **117**, 5316 (1995).

134. J. Tirado-Rives, W.L. Jorgensen, *J. Am. Chem. Soc.*, **112**, 2773 (1990).

135. E.F. Ivanova, S.M. Kiiko, *Vestnik Kharkovskogo Universiteta*, **378**, 92, (1993). Journal written in Russian.

136. T.M. Aminabhavi, B. Gopalakrishna, *J. Chem. Eng. Data.*, **40**, 856 (1995).

137. Y. Maham, T. T. Teng, L. G. Hepler, A. E. Mather, *J. Solution Chem.*, **23**, 195 (1994).

138. M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.S. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T.A. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, L.J. Andres, M. Head-Gordon,

E.S. Replogle, J.A. Pople, *Gaussian 98,* Gaussian, Inc., Pittsburgh, PA (1998).

139. A. Ben-Naim, *Water and Aqueous Solutions. Introduction to a Molecular Theory*, Plenum Press, New York (1974).

140. G.A. Jeffrey, *An Introduction to Hydrogen Bonding*, Oxford University Press, New York (1997).

141. G.C. Pimentel, A.L. McClellan, *The Hydrogen Bond*, W.H. Freeman, San Francisco (1960).

142. Ch. Reichardt, *Solvents and solvent effects in organic chemistry*, VCH, New York (1988).

143. J.M. Haile, *Molecular Dynamics Simulation. Elementary Methods.*, John Wiley and Sons, New York (1992).

144. A.F. Gallaugher, H. Hibbert, *J. Am. Chem. Soc.*, **59**, 2521 (1937).

145. R.C. Weast (Ed.), *CRC Handbook of Chemistry and Physics*, 64th edn., CRC Press, W. Palm Beach, Florida (1983).

146. P. Knauth, R. Subbuh, *Bull. Soc. Chim. Fr.*, 834 (1988).

147. P. Knauth, R. Subbuh, *J. Chem. Thermodyn.*, **21**, 203 (1989).

148. P. Knauth, R. Subbuh, *Struct. Chem.*, **1**, 43 (1990).

149. V. Majer, V. Svoboda, *Enthalpies of Vaporization of Organic Compounds*, Blackwell Scientific Publications, Oxford (1985).

150. F.M. Samigullin, M.N. Rodnikova, T.M. Val'kovskaya, *Zhurnal Neorganicheskoi Khimii*, **42**, 1049 (1997). Journal written in Russian.

151. T.M. Val'kovskaya, M.N. Rodnikova, F.M. Samigullin, G.V. Spivak, *Zhurnal Fizicheskoi Khimii*, **72**, 616 (1998). Journal written in Russian.

152. M.N. Rodnikova, P.C. Jastremskij, V.S. Kharkin, K.T. Dudnikova, *Izv. AN SSSR (Cer. Khimija)*, N5, 1157 (1987). Journal written in Russian.

153. P.G. Kusalik, A.P. Lyubartsev, D.L. Bergman, A. Laaksonen, *J. Phys. Chem.*, B **104**, 9526 (2000).

154. P.G. Kusalik, A.P. Lyubartsev, D.L. Bergman, A. Laaksonen, *J. Phys. Chem.*, B **104**, 9533 (2000).

155. P.G. Kusalik, I.M. Svishchev, *Science*, **265**, 1219 (1994).