INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

Jayaram, Bhyravabhotla

MONTE CARLO COMPUTER SIMULATION STUDIES OF DIMETHYLPHOSPHATE ANION

City University of New York PH.D. 1987

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __✓__.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received ✓
16. Other _____

_____

_____

University
Microfilms
International

MONTE CARLO COMPUTER SIMULATION STUDIES OF DIMETHYLPHOSPHATE ANION

by

Bhyravabhotla Jayaram

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

1987

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

12/22/86
Date

David L. Beveridge
Chairman of the Examining Committee

1/7/87
Date

G. M. [illegible]
Executive Officer

Michelle S. Broido

[illegible]

[illegible]
Supervisory Committee

The City University of New York

Abstract

Monte Carlo Computer Simulation Studies of

Dimethylphosphate Anion

by

Bhyravabhotla Jayaram

Advisor: Professor David L. Beveridge

This dissertation is a collection of theoretical calculations on diverse aspects of the dimethylphosphate anion (DMP^-) in aqueous solutions and in free space based on Monte Carlo computer simulation methods. In addition, some extensions of the formal theory of environmental effects based on the dielectric continuum model were carried out. The individual topics considered are as follows.

Aqueous hydration of DMP^-, and Na^+DMP^- ion pair, with the two phosphodiester torsion angles in the gauche-gauche (gg), gauche-trans (gt) and trans-trans (tt) conformations were investigated using liquid state Monte Carlo computer simulation method in (T,V,N) ensemble at 25°C. The structural and energetic aspects of the hydration of each molecule were analyzed using the proximity criterion and were partitioned into ionic, hydrophilic and hydrophobic contributions. Free energy simulations were performed on $[DMP^-]_{aq}$ in the gg, gt and tt conformations using the probability ratio method. The performance of the simulations on the

thermodynamic cycle was also examined.

The intramolecular thermodynamics of the gg, gt and tt conformations of DMP^- in free space were determined by the Monte Carlo method, with configurational entropies estimated in the quasiharmonic approximation. Both inter and intramolecular thermodynamics favored the gg conformation of DMP^- relative to gt and tt forms.

The solvation shell model, modified by incorporating the coordination numbers and first shell radii obtained from the mean energy simulations was used to independently evaluate the relative free energies of hydration of the gg, gt and tt conformations of DMP^- and to provide a comparison with the results of Monte Carlo free energy studies.

In related studies from another point of view, the Tanford-Kirkwood theory for evaluating the electrostatic free energy of a discrete charge distribution in the presence of ion atmosphere was extended to concentric dielectric continua, and applied to study the conformational preferences of DMP^- and Na^+DMP^- in the absence and presence of ion atmosphere and at varying local dielectric constants. Theoretical extensions of coaxial cylindrical dielectric continua to evaluate environmental free energies were also carried out.

Preface

One of the basic tenets of biomolecular investigations is that form follows function and vice versa. We hope to understand function through a detailed knowledge of the structure. The present study is concerned with theoretical investigations of biopolymeric structural problems. Theoretical methods can be broadly classified into three categories based on the spatial and time scales involved in the problem at hand, namely (i) microscopic (ii) mesoscopic and (iii) macroscopic. The microscopic description involves mechanical models (H. L. Friedman, "A Course in Statistical Mechanics", Prentice-Hall, New Jersey (1985))(with a Hamiltonian specifying the interactions between the various particles constituting the system) either at Schroedinger level (where the particles are electrons and nuclei), or at Born-Oppenheimer level (where the particles are atoms and molecules and the interaction potentials are generally drived from quantum mechanical, S - level, calculations), or at Mcmillan-Meyer level (solvent is treated as a dielectric continuum and solutes are the interacting particles). Both Monte Carlo and molecular dynamics methods for instance, involve a microscopic description at B-O level, with the former method giving equilibrium (stationary) solutions to the problem, and the latter

following the temporal evolution of the system in the picosecond to nanosecond range. Statistical mechanics provides the necessary link between the microscopic results and macroscopic observables. At the mesoscopic level (R. Kubo, Science, 233, 330 (1986), and references therein) are the Kinetic theory (BBGKY hierarchy), the Brownian dynamics, the master equation method, the Fokker-Planck approach etc. to cite a few coarse grained descriptions. Equilibrium and nonequilibrium thermodynamics, electrohydrodynamics are some familiar examples for a macroscopic description all of which involve a continuum treatment of the macroscopic field variables. Some of the successful mean field theories at this level are the Van der Waals equation of state, the Debye-Huckel theory and the Flory's excluded volume approach. The present study is mainly concerned with a microscopic description of matter to investigate the equilibrium thermodynamic and structural properties.

The relevance of equilibrium studies of biopolymers cannot be overemphasized. Living systems feed on negentropy (E. Schroedinger, "What is Life?", Cambridge U.P. (1945)) and therefore it may appear that focus must be on nonequilibrium studies. Firstly, laws of thermodynamics are universal at and near equilibrium situations. Secondly a variational principle applies to

thermodynamic potentials at equilibrium. For an isolated system (E,V,N ensemble) entropy attains a maximum. Helmholtz free energy is at a minimum for a closed system (T,V,N ensemble), and Gibbs free energy takes a minimum value for the (T,P,N ensemble) or open system (T,V,u ensemble). It is not possible to identify a thermodynamic potential, for a system not at equilibrium, to which a variational principle applies (I. Prigogine and I. Stengers, "Order Out of Chaos", Bantam, New York (1984), ch-5.). Thirdly fluctuations are a source of order in living systems but Einsteinian formula for the probability distribution of fluctuations holds at equilibrium and near equilibrium and no such formula exists for nonequilibrium fluctuations. Fourthly the generalized second law (minimum entropy production principle) addresses the entropy production rate but the entropy flow term is system dependent and most complicated to grapple with. Above all, our knowledge of forces directing the behavior of biomolecules in varying environments, is incipient. Thus equilibrium studies of biopolymers is the place to begin for a systematic investigation.

Studies on biomolecules under physiological conditions must recognize their aqueous environment and this poses a challenging problem to theoretical investigations. Liquid state cannot be understood in

terms of idealizations of the system as an aggregate of independent subsystems such as a collection of harmonic oscillators as in the case of crystals, or as noninteracting or weakly interacting atoms and molecules as in the case of gases. The case of associated liquids is even more formidable since simple thumb rules such as "repulsive interactions govern the structure and attractive interactions control the density" break down. Computer experiments have made a major breakthrough in this area with the availability of alternate solute-solvent and solvent-solvent interaction potential functions, and methodologies to evaluate structural, equilibrium and transport properties. The rapidity with which these computations can be carried out today, provides an additional incentive. Moreover, given a model, computer simulations, uniquely provide an insight into the microscopic details of the structural aspects which are not directly accessible to experiment.

A major goal of the present study is to contribute to our knowledge on the forces stabilizing nucleic acids in aqueous solutions, using the techniques of computer simulations. Chapter I of this dissertation deals with Monte Carlo mean energy simulations on the aqueous hydration of dimethylphosphate anion. Chapter II describes the determinations of the

intramolecular thermodynamics. Chapter III is concerned with the modification of an empirical scheme to evaluate solvation free energies using the structural results obtained from the simulations. Chapter IV gives an account of liquid state free energy simulations as applied to the conformational problem in dimethylphosphate anion in aqueous solutions. Extensions of the discrete continuum approach, with a macroscopic description of the environment are discussed in Chapter V along with an application. Some relevant methodological explorations including a stochastic formulation of the dynamics starting from Monte Carlo results are given in Chapter VI. Scope of the thesis necessarily precludes discussion of many interesting topics. For instance, pair-wise additivity of the potential of intermolecular forces is at the heart of any computation conducted at molecular level, the justification of which can be found in Kirkwood's seminal paper (J. G. Kirkwood, J. Chem. Phys., 3, 300 (1935)). Although the present day simulations use effective pair-wise potentials, the influence of cooperativity merits further investigations. Periodic boundary conditions, system size and ergodicity are some of the important issues that are beyond the purview of this thesis. A recurring theme in the present work is the relevance of results to

biopolymers, comparison of the results with experiment, and merits and limitations of the methodologies. In science as in any other area of human endeavor there is so much more to do.

Acknowledgements

I consider it a special privilege to be a graduate student of Prof. D. L. Beveridge, who is both a scientist and a human being par excellence. The rigorous scientific training that he imparted, the role of focus and discipline in research that he repeatedly stressed on and the importance of efficient scientific communication that he taught me are some of my richest treasures today. I express my sincerest thanks to him.

Drs. M. Mezei and P. K. Mehrotra have taken me a long way from the alphabets of Monte Carlo algorithm to a specialist in computer simulations. I owe my gratitude to them, their patience and thoroughness. I gratefully acknowledge the many helpful discussions I have had with Dr. M. Mezei. I cannot miss this opportunity to thank Mr. T. R. Vasu for training me in methods of science particularly during the early stages of my Ph.D program. My association with Dr. G. Ravishanker has been very fruitful scientifically. I express my thanks to him. Special thanks go to Dr. P. V. Maye whose company I cherished all these years. I enjoyed many stimulating discussions with him, and with Mr. B. Gedulin and Mr. P. S. Subramanian.

My wife Lakshmi has been instrumental in seeing me through to the finish line of Ph.D. I would also like to express my gratitude to Prof. S. P. Narula, Dr. S.

Prahlad Rao, Mr. A. Gopinath and my parents who made my reentry into science easier.

I dedicate this thesis to all students of science.

Table of Contents

List of Tables

List of Figures

CHAPTER I

LIQUID STATE MONTE CARLO STUDY OF THE AQUEOUS HYDRATION OF DIMETHYLPHOSPHATE

I. A. INTRODUCTION

The structures of nucleic acids show considerable conformational flexibility and are known to be sensitive to hydration and ionic strength. The phosphodiester moiety, $-O-PO_2^-O-$, bears the anionic charge in each nucleotide unit and environmental effects are expectd to be very strong in this region. An understanding of the hydration scheme and conformational preferences of the diester torsions is essential to the investigations on the secondary structure of nucleic acids in aqueous solutions. We present herein a theoretical study of the aqueous hydration of the phosphodiester group, the effect of hydration and the influence of counterion on the conformational preferences of the phosphodiester torsion angles. The motivation for this study is two-fold: (1) to characterize the structural and energetic features of ionic, hydrophilic and hydrophobic hydrations in nucleic acid prototypes, and (2) to understand the relative conformational preferences of phosphodiester torsion angles in aqueous solutions.

Dimethylphosphate anion, DMP^-, prototypical of nucleic acid backbone is chosen for the current investigations. While the $C_5^{\prime}$ to $C_3^{\prime}$ polarity of nucleic acids is not well represented by DMP^-, and methyl groups do not mimic sugar rings, it is one of the

simplest model systems exhibiting the conformational problem for the phosphodiester torsion angles. Also DMP^- possesses all three types of chemically interesting functional groups: hydrophobic (methyl groups), hydrophilic (ester oxygens) and ionic (anionic oxygens). This study is based on (T,V,N) ensemble Monte Carlo mean energy computer simulations on the dilute aqueous solution of the dimethylphosphate anion, $[DMP^-]_{aq}$, and the sodium dimethylphosphate ion pair, $[Na^+DMP^-]_{aq}$, at a temperature of 25°C and experimental density.

I. B. BACKGROUND

The stability of B-DNA in aqueous solutions, the conformational transition to Z-DNA at increasing ionic strengths, and to A-DNA at decreasing humidity [1] are well documented but little understood at a molecular level. In the following we review some of the attempts to unravel the principles underlying the observed behavior of nucleic acids in the absence and presence of solvent and ionic environments.

Conformational analysis of the nucleic acid backbone is generally conducted in terms of six torsional angle variables [1], α, β, γ, δ, ϵ, and ζ shown in Figure 1, where α and ζ are the phosphodiester torsion angles. The magnitude of the problem involved in free space, is clearly depicted in Olson's [2] study based on sterically allowed rotational combinations. Each of the torsion angles α, β, γ and ζ takes three preferred values; g^+ (~60°), t (~180°) and g^- (~300°), δ assumes two (g^+,t) and ϵ two (t, g^-) giving a total of 324 possible rotational combinations for the six torsional angles along the backbone. This coupled with syn and anti orientations of the bases gives a multiplicative factor of 648 for each repeating nucleotide unit with sugar pucker not counting. Structural correlations however reduce the number of allowed combinations. Arnott and Hukins [3] and Sundaralingam [4], based on a

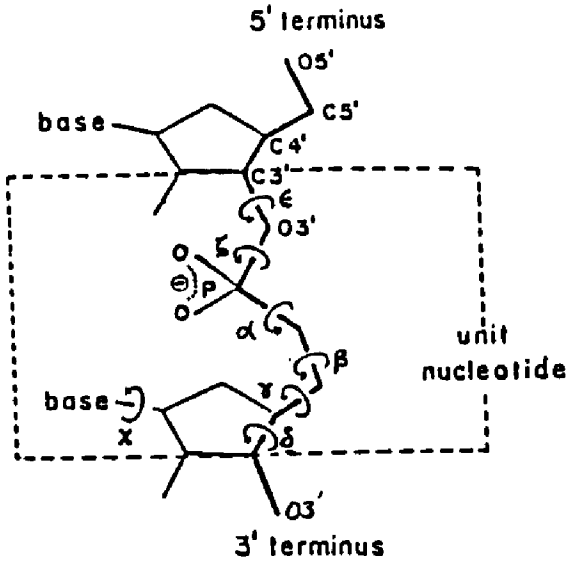

Figure 1. Definition of nucleic acid backbone torsion angles

statistical analysis of the known crystal structures of model compounds for nucleotides observed that nucleotides were more rigid than nucleosides, thus proposing the rigid nucleotide concept. The most preferred values for α , and ζ were found to be in the range of (g^-, g^-) for right handed helices. The other favored conformations (g^-, t) and (g^+, t) resulted in an extended backbone structure with unstacked bases and were conjectured to be of significance in solution and in the folding of polynucleotide chains. Berman et al. [5] proposed a categorisation of the crystal structures as falling into three conformational classes: (1) mini helices, (2) extended structures and (3) folded structures. The major distinguishing elements are the phosphodiester torsions (α,ζ) which fall into the ranges of (g^-, g^-) for class 1, (g^-, t) for class 2 and (g^+, g^+) for class 3. Thus to a large measure, the conformational flexibility in the oligonucleotides studied, was seen to occur with in the phosphodiester linkage.

The inherent tendency of the nucleotides in free space for helix formation ($\alpha = g^-, \zeta = g^-$) is attributed to anomeric effects [1]. There is an extensive literature on anomeric effects [6]. In the context of nucleic acids it refers to the polarization of the P-O ester bond by the lone pair of electrons on

the neighboring ester oxygen, resulting in the destabilization of the extended tt conformer relative to gt and gg forms with gg conformation emerging as the most stable form in free space. In aqueous solutions however, the picture is far more complicated by the screened phosphate-phosphate repulsions and by base stacking interactions. A logical conjecture stemming from the hydrophobic effect on the preferred conformations for phosphodiester torsions would be that nucleotides in aqueous solutions must also have an inherent tendency for helix formation with α , ζ adopting g, g values, since it is in this conformation that hydrophobic groups tend to cluster together along the backbone. However the significance of the extended conformations in aqueous solutions cannot be over emphasized, particularly in conformational transitions and in biological function. Thus it is essential to obtain quantitative thermodynamic and structural information on the hydration of phosphodiester torsions.

A partial understanding of the DNA hydration has essentially evolved in two stages, the advent of single crystal studies of model compounds drawing the demarcation line. Bloomfield et al. [7] summarized some of the early work on the primary shell hydration of DNA. Self diffusion coefficient studies gave a

coordination number of 6.5 waters per nucleotide, gravimetric studies yielded 8.4-10 waters at a relative humidity of 75-80%. Sedimentation coefficient studies indicated 7-13 waters, heat capacity studies suggested 9-10 waters and NMR at -35°C gave 11 waters as associated with the primary shell hydration of DNA.

Falk [8-10] and coworkers divided the hydration sites of DNA into three molecular subgroups which in decreasing order of priority were, the diesterified phosphate group, the deoxyribose and the heterocyclic bases. They found about 5-6 waters bound to the $PO_2^-Na^+$ group. Lewin [11] proposed four types of water bridges involving phosphate groups, namely (1) strong cation phosphate water bridges, (2) weak cation phosphate water bridges, (3) anchored phosphate amino water bridges and (4) nonanchored phosphate amino water bridges. In the classification of Hopfinger [12], out of the ~18 waters associated with the primary hydration of B-DNA, two waters were assigned to ionic phosphate representing the zone of highest affinity and the removal of these waters led to a disruption of the secondary structure of DNA. Next in binding strength were four waters lying at furanose oxygens and phosphodiester linkages. Dahlbord and Rupprecht [13] in their neutron scattering study of oriented Na-DNA found that water molecules were bound to H-bonding sites on

the double helix in a more or less regular way and concluded that it was not possible to assign a definite type of hydration structure.

Single crystal studies of nucleotides greatly facilitated an understanding of the hydration pattern in nucleic acids. Drew and Dickerson [14] concluded that at higher humidities B-form predominated to the exclusion of A and Z forms of DNA. Ordered water structures in the minor groove were proposed as an explanation for the stability of the B-form of DNA. Conner et al. [15] reported that minor groove was fully hydrated before the major groove in B-DNA. Kennard [16] examining the single crystals of oligonucleotides observed that in B-form primary hydration shell comprises 11-12 waters, that minor groove hydration was well ordered and that the water network in the major groove was discontinuous and much less regular. In the A-form first hydration shell constituted a ribbon of water molecules extending along the major groove, with no PO_2^- water bridges. Minor groove in A-DNA was less accessible and hydration was noticed to be less ordered. Based on these observations she proposed a mechanism for inducing B to A transition. On dehydrating the B-form, ordered water structure in the minor groove is disrupted and only highly polar phosphate oxygens remain hydrated. This induces a

conformational transition. A-DNA is stabilized by major groove hydration. Ordered water structures observed in crystal studies may well be due to the physical state of the system and hydration patterns concluded from crystal studies are not truly representative of the hydration of nucleic acids in aqueous solutions. Solution NMR data of Mai et al.[17] suggested a sequence for hydration. Up until five waters phosphate group was hydrated. As this was increased to 11 waters all exposed hydrophilic sites were hydrated. By about 25 waters groove hydration was also complete.

Pullman and coworkers carried out extensive theoretical studies [18-20] on DNA hydration. The molecular electrostatic potential was proposed to be a good approximation for the binding of cations, and the electrostatic field for the binding of neutral dipolar species, in particular for hydration. Deepest potentials were located in the grooves of the double helix, on the phosphate subunits, and greatest fields were concentrated at the phosphates. Counterion binding decreased the magnitude of the potentials but further increased the field on phosphates. They concluded that phosphates were the major sites of hydration, with hydration relatively sparse at esteric oxygens.

Clementi and coworkers [21-25] carried out a series of computer simulations on DNA-water clusters. They

identified three types of waters, those that were weakly H-bonded (R_c > 3.2 Å), those that were strongly H-bonded (R_c < 3.2 Å) and those which were both strongly H-bonded and optimally oriented for an H-bond, called very strongly H-bonded waters. Coordination numbers of strongly H-bonded type waters were found to be 5.9 for PO_4^-, Ø.3 for sugar rings, Ø.5 for sugar and bases and Ø.9 for bases. Initial saturation of the groove was reported to be coincident with the solvation of phosphate, sugar oxygens and of the base pairs corresponding to 11 waters. By increasing the extent of hydration to 2Ø waters, additional waters were packed into the groove and around the solvated sites. They concluded that waters not only enclosed PO_4^- group but also formed H-bonded filaments. Cluster calculations such as above disregard the liquid state densities completely and tend to over estimate solute-water interactions (also see section I. F) at the expense of water-water interactions, and are not very informative of the hydration in aqueous solutions at experimental densities.

Systematic studies on nucleic acid prototypes are relatively scarce. As a model compound for phosphate hydration and for phosphodiester conformational preferences dimethylphosphate anion has attracted considerable attention.

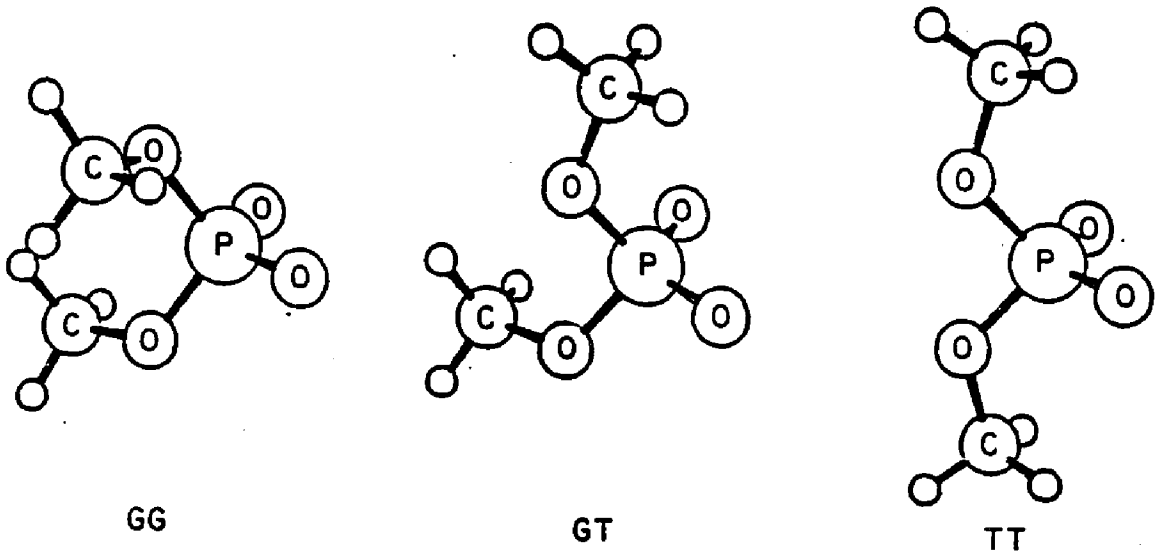

Figure 2. The structure of dimethylphosphate anion in gg, gt and tt conformations.

Dimethylphosphate at physiological pH is found predominantly in the anionic form. The conformations of DMP^- can be specified in terms of the torsional angles α(O-P-O-C) and ζ(C-O-P-O) (defined in Figure 1), following IUPAC notation. The angles α and ζ are identical to ω and ω' assigned to the phosphodiester torsions in some of the earlier literature. Both α and ζ follow roughly a threefold potential with minima in the regions of gauche$^+$ (g^+), trans (t) or gauche$^-$ (g^-). In this account, we contract the notation for the specification of conformation for DMP^- to simply "$\alpha\zeta$", i.e., gg, gt, or tt. These conformations are depicted in Figure 2. In the solid, DMP^- with ammonium counterion crystallizes with $\alpha = \pm 57.5^{\circ}$ and $\zeta = \pm 62.4^{\circ}$, a gg conformation [26]. Crystal structure data on phosphodiester torsion angles for various dinucleotides reveal a strong preference in these systems for gauche values of α and ζ, presumably stabilized by anomeric effects. A recent search of the Cambridge crystallographic data bank [27] turned up 45 structures with gg conformations versus only six structures exhibiting a combination of gauche and trans, i.e., gt, or tg values. The tt form seems to occur extremely rarely in phosphodiester torsion angles.

Information on the conformational preferences of DMP^- in solution comes from IR, Raman and NMR

spectroscopic studies. Shimanouchi, Tsuboi and Kyogoku [28] interpreted the IR spectrum of $Ba^{2+}(DMP^-)_2$ and the Raman spectrum of Na^+DMP^- in aqueous solution by means of normal coordinate calculations and concluded that the gg form was the likely conformation for DMP^- in both solid and solution. However, conformational analysis of DMP^- based on depolarized Rayleigh scattering by Garrigou-Lagrange et al. [29] suggested that extended rotational isomeric states of DMP^- in water were highly populated.

Proton NMR is insensitive to changes in the α, ζ torsions, but ^{31}P NMR proves to be a useful probe. Gorenstein et al. [30] investigated the ^{31}P chemical shifts of DMP^- and related compounds in aqueous solution as a function of temperature, in conjunction with studies of nucleic acid conformation. The ^{31}P resonance shifts downfield as a function of temperature. This was interpreted by an increase in the Boltzmann population of gt and possibly tt conformers with increasing temperature. By inference, the preponderance of gg conformation of DMP^- at lower temperatures, including ambient, was indicated. On the other hand, Lerner et al. [31] in their recent report on solvation effects on ^{31}P NMR chemical shifts and IR spectra of phosphate diesters in water and mixed organic solvent systems, interpreted changes in ^{31}P

chemical shifts with increasing concentration of nonaqueous solvents as due to changes in the phosphate hydration rather than conformational changes. They further adduced the "blue shift" of the antisymmetric stretching frequency of the anionic O-P-O group as corroborative evidence.

Extensive research has been reported on the calculation of conformational preferences in phosphodiester torsional angles in DMP^- and related compounds in the free space approximation. Some early studies [32, 33] with empirical potential functions favored the extended structures involving trans forms of α and ζ, while others [34] favored gauche forms. Investigations based on extended Huckel theory [35] and PCILO method [36] on model compounds favored trans and gauche forms, respectively. Quantum mechanical calculations by Newton [37] on the α, ζ conformational energy map showed gg< gt< tt, with tt destabilized by some 7 kcal/mole. Recently developed potential energy functions [38-41] effectively incorporate the interactions between the lone pairs on the ester oxygens, destabilizing the extended forms relative to gauche form.

Perahia, Pullman and Saran [42] showed the (α, ζ) energy surface to be sensitive to assumptions about molecular geometry. Subsequently, Perahia and

Pullman [43] and Gorenstein et al. [44-46] demonstrated a significant correlation between the phosphodiester torsional angles and the ester oxygen O-P-O bond angle, and found that a 4-5° reduction in O-P-O angle accompanied each phosphodiester torsional rotation from g to t. The conformational energy surface flattens when this is taken into account in calculations. All (α,β) conformations are then predicted to be thermally accessible to some extent, with a relative ordering in conformational energies of gg (Ø) < gt(+Ø.8-Ø.9 kcal/mole) < tt (+1.7-1.9 kcal/mole). This accentuates the possible role of hydration in deciding the relative conformational preferences for the phosphodiester torsion angles in solution.

Theoretical studies of the hydration of DMP^- by means of an exploration of the solute-water interaction energy hypersurface have been described in papers from several laboratories. Berthod and Pullman [47], and Perahia, Pullman and Berthod [48], using quantum mechanical calculations found extended and bridged structures for the DMP^- water monohydrate with stabilization energies of -2Ø to -28 kcal/mole depending upon whether or not d-functions were included in the atomic orbital basis set. PCILO calculations of Frischleder et al. [49] indicated that the strongest hydration occurs in the O-P-O plane bridging the

anionic oxygens with an interaction energy of -27 kcal/mole for the $DMP^-(H_2O)$ complex. Gay and Vanderkooi [50], performing CNDO/2 calculations, concluded that the lowest energy configuration (-20 kcal/mole) for $DMP^-(H_2O)$ was one with a linear hydrogen bond in which the atoms of H_2O were coplanar with the anionic oxygens of phosphate group. Alagona, Ghio and Kollman [51] in their quantum mechanical and molecular mechanical studies on the DMP^- monohydrate found the bridged structure for water to be more stable than the linear structure by about 5-6 kcal/mole. More extensive "solvation site" studies were reported by Pullman, Berthod and Gresh [52], who enumerated possible DMP^- monohydrate structures. From these calculations they predicted that six waters would be found in two "circular zones of attraction" near the anionic oxygens and perpendicular to the P-O bond.

Corongiu and Clementi [53] and subsequently Clementi, Corongiu and Lelj [54] studied the interaction of a single molecule with diethylphosphate anion and with the nucleic acid backbone model $MeCH_2$-C-O-PO_2-O-CH_2Me using quantum mechanical calculations. Contour plots revealed an energy minimum of -24 kcal/mole for DEP^- and -18 to -20 kcal/mole for the backbone fragment, in "Region A" which corresponds to the anionic oxygen zone of attraction of Pullman et al.

Another region called "B", in the ester O-P-O region, was also noted with binding energies apparently of a similar magnitude. The results of the calculations were used to construct an analytical potential energy function for these systems.

Clementi et al. used this analytical potential function in conjunction with a corresponding water-water potential to determine minimum energy structures for clusters of first four and then ten waters with DEP^- and with their nucleic acid backbone prototype and the results were mainly pertinent to phosphate hydration. They found their Region A to be preferentially populated and involved a single well-directed hydrogen bond pointing towards the two anionic oxygens, and that minimum energies were obtained by optimizing water-solute interaction at the expense of water-water interactions. However the detailed structure of the small cluster was found to depend strongly on the number of waters considered, 4 to 10 in these studies.

The question of polyhydration of DMP^- was taken up in a more elaborate study by Pullman, Berthod and Gresh [55]. The essential trihydrate structure around each anionic oxygen was maintained, however. An extensive study of polyhydration based on energy minimization was recently reported by Langlet,

Claverie, Pullman and Piazzola [56]. The $DMP^-(H_2O)_6$ hexahydrate complex was subjected to translational and rotational energy optimization from various starting configurations. The solvation site model with three waters per anionic oxygen in the circular zone of attraction was recovered with only slight modification, the principal refinement being that 6-7 waters were admitted to the first shell as defined on a binding energy criterion. The energies of the various conformers were found to remain close even with the inclusion of waters. Studies on the systems $DMP^-(H_2O)_{30}$ and $DMP^-(H_2O)_{60}$ permitted a detailed analysis of the hydration complex, and also the "radially oriented" structure in the vicinity of the anionic $>PO_2^-$ group and the "concentric structures", defined to describe the hydration of DMP^- methyl groups. The terms "radially oriented" and "concentric" structures seem to be equivalent to the terms "ionic hydration" and "hydrophobic hydration" widely used to discuss the structural chemistry of aqueous hydration. Langlet et al. found that solvent effects on conformational stability in DMP^- were likewise small in the higher order water clusters .

A treatment of the aqueous hydration of DMP^-, HDMP and H^+DMP^- has recently been carried out by Bleha, Mlynek and Tvaroska [57] using an Onsager continuum

model for the solvent. For DMP^-, the gg conformer was found to be preferentially stabilized in water as well as in free space, due primarily to the electrostatic dipolar contribution to the hydration energy.

The solvation shell model for the calculation of free energies of hydration, proposed by Scheraga and coworkers [58-60] and elucidated by Hopfinger [12, 61], assumes a model for hydration complexes of biological functional groups and subunits, which can be applied to DMP^-. The hydration numbers of solvation shell theory for DMP^- groups are: four for anionic $>PO_2^-$ group, two for each ester oxygen, and eight for each methyl group, leading to a value of 24 waters for the entire molecule in the extended tt form.

Monte Carlo simulation studies of the hydration of DMP^- have recently been reported in a preliminary form by Beveridge et al. [62], and more extensively in a recent article by Alagona, Ghio and Kollman [63]. Kollman and coworkers investigated the hydration of gg and gt conformations of DMP^- in the (T,P,N) ensemble. Partial atomic charges on the solute and the O-P-O valence bond angle were apparently treated as independent of conformation. Methyl groups were approximated by united atoms. The TIPS4P model was used to describe water-water interactions, and a 12-6-1 potential function for the solute-water interactions.

About 23 water molecules were assigned to the first shell of DMP^- on a geometric criterion and were classified as belonging to strongly polar, polar and apolar domains corresponding to ionic, hydrophilic and hydrophobic regions. The average coordination number of three waters per anionic oxygen, and the absence of bridged water structures in the ionic region emerging from their simulation studies were in accord with quantum mechanical predictions of Pullman and coworkers described above. Computed enthalpies of hydration favored gg conformer over gt by 28 kcal/mole and this was ascribed to more attractive water-water interactions in the case of gg. Solute-water interactions and in particular ionic hydration favored the gt conformer in their study. Detailed comparison of the results of Alagona et al. with those of the present study creates for the first time a perspective on the sensitivity of simulation results to choice of potential function, assumptions concerning solute geometry, thermodynamic ensemble and method of analysis, and statistical uncertainty in the calculated results.

The influence of counterion on the hydration and stability of model compounds for nucleotides was reported by Rich and coworkers [64-65] in their ApU and GpC crystal studies. In ApU, one sodium ion was found

to be bound to phosphate group while the other was located in the minor groove region, bound to uracil and screened from the phosphate group by its first shell waters. In GpC, the sodium counterion was bound to the ionized phosphate group, exhibiting an octahedral coordination and was considered to be a major organizing structural element. Phosphodiester torsions were in gg conformation in both ApU and GpC.

Diethylphosphate anion crystallized in gg conformation with barium counterion [66] while with silver cation [67] it was in gt form. DMP^- with ammonium counterion as pointed out earlier [26], crystallized in gg conformation. Glonek and Wazer [68] through their P-31 spin lattice relaxation studies on the aqueous solutions of several phosphate esters including DMP^- concluded that the anionic phosphate group in presence of sodium and potassium cations was associated with closely lying waters, but structural details were not accessible.

Several theoretical studies involving quantum calculations on metal-phosphate complexes have addressed the problem of geometry and stability of the complexes and few the conformational problem. Nanda and Govil [69] performed CNDO/2 calculations on metal-cation interactions with DMP^-. Both sodium and magnesium counterions favored two-centered (bridged)

interactions, with counterion in the plane of anionic $>PO_2^-$ group, equidistant from the anionic oxygens. Conformational trends were reversed (tt> gt> gg) in the presence of counterion in their study. Pullman, Gresh and Berthod [70] reexamined these trends with STO-3G ab initio calculations and found that the conformational trends were unperturbed (gg> gt> tt) by the presence of counterion. The bridge position for the counterion remained the stable form. Marynick and Schaefer [71] conducted ab initio calculations on a series of phosphate-counterion complexes including DMP^- (in tt geometry) with and without a water molecule attached to the metal ion. They inferred the contact interaction, with metal ion in C_{2v} symmetry with respect to the anionic oxygens of the phosphate group, as the most stable configuration for the metal-phosphate interactions. Pullman and Berthod [72] further investigated the effect of counterions on the molecular electrostatic potential of DMP^-. In their study, the counterion was placed on the bisector of PO_2^- at a distance of 2 Å from the anionic oxygens. The modified molecular potential indicated a strong decrease in the attractive nature of the PO_4^- group. Pack and coworkers [73] studied the geometric and charge transfer aspects of $M^+PO_4^-$ complexes. The preferred position for the sodium counterion was a

bridged structure in the $>PO_2^-$ plane, in agreement with the earlier theoretical calculations.

Berthod and Pullman [74] subsequently considered the competetivity in binding of Na^+ and water to the DMP^- anion, through their studies on $Na^+DMP^-(H_2O)_6$ system. They proposed two modes of binding. One involved a direct binding of Na^+ to DMP^- and the other, sodium cation binding to phosphate anion through an intermediate water molecule, shared by both anion and cation. Ab initio calculations of Pullman et al. [75] involving these two modes of binding showed that the interaction energies were comparable suggesting that both forms might contribute to the structure in solution.

Mlynek and Tvaroska [57] in their continuum study, referred to above, on H^+DMP^- intimate ion pair found tt to be the most stable conformation for the phosphodiester torsion angles.

Corongiu and Clementi [76] reported their cluster calculations on B-DNA fragment (12 base pairs, 24 sugar units and 22 phosphate groups) with one Na^+ placed fixed near the free oxygens of each PO_4^- group. They found an additional 1.5 water molecules per Na^+ in the first shell (R_c=3Å) of Na^+-B-DNA relative to B-DNA resulting in a highly dense and structured first shell termed an "ion induced compression effect", i.e.

electrostriction in the conventional physical chemistry literature. Sodium cation on the average maintained an octahedral coordination in their study.

Overall, it is clear from the literature on DMP^- that solvent interactions are potentially a significant influence on conformational stability, and the hydration of the phosphate group is an important area for study in nucleic acids research. The theoretical study of DMP^- in water via simulation procedures is the appropriate place to begin, but a number of difficulties are anticipated. The quality of the potential functions is a matter of continuous concern in obtaining a satisfactory account of basic thermodynamic indices, the principal point of comparison with experimental studies as well as other results not as amenable to experimental verification. Intrinsic limitations in the precision of calculated quantities due to statistical uncertainties in ensemble averages formed over finite segments of a potentially infinite numerical realization exist, and are manifest especially in estimates of conformational energy differences where the "small differences in large numbers" problem is encountered. In charged systems or systems with charge separation the periodic boundary conditions customarily assumed for simulations on aqueous system may be problematic. However, these

issues cannot be addressed without carrying out a series of well defined simulation studies, effectively computer experiments on the system, and gaining experience with the numerical problems and analyzing the results fully and critically. Thus the purpose of the calculations undertaken herein is twofold: to describe the system as well as possible and simultaneously to gain perspective on the methodology.

I. C. MONTE CARLO THEORY AND METHODOLOGY

A microscopic description of matter must necessarily start with a Hamiltonian for the system specifying the interactions between the particles (in the present study the particles are atoms and molecules) constituting the system. In the computer simulations [77], the equilibrium properties of the system are then evaluated either as time averages (dynamics simulations) or as ensemble averages (Monte Carlo method). Thus in the simulations, one needs an evolution equation for the system or alternatively a prescription for the evolution of the system in terms of transition probabilities to generate a series of configurations, given the Hamiltonian. (Theoretically equations of motion uniquely determine the transition probabilities and vice versa. While this program is clearly worked out in the quantum case - starting from transition probabilities obeying Markovian dynamics one can arrive at Schroedinger's equation via Stone's theorem [78] - the relationship is not well established in the classical case except at the level of analogies.) Classical equations of motion are the obvious choices for dynamical simulations. In the Monte Carlo method, proposed by Metropolis et al. [79], the evolution of the system is realized through a move algorithm specified in terms of time independent

transition probabilities. Thus the notion of time does not enter the Monte Carlo method, and the method is essentially useful for evaluating any quantity that can be expressed as an ensemble average. These aspects are detailed below.

The equilibrium value of a quantity F in the canonical (T,V,N) ensemble is expressed as [80]

$$\langle f \rangle = \frac{\int f \exp[-E(q_1,\ldots,q_{3N})/kT]\, d^{3N}q}{\int \exp[-E(q_1,\ldots,q_{3N})/kT]\, d^{3N}q} \qquad (1)$$

Since this is a multi-dimensional integral its evaluation is facilitated by the Monte Carlo method [81]. This is a numerical integration technique using random numbers. In the present context the method involves choosing a configuration 'i' randomly and weighting it with the Boltzmann probability. Thus the quantity f is evaluated as

$$\langle f \rangle = \sum_i f_i P_i \qquad (2)$$

where

$$P_i = \{\exp(-E_i/kT)\}/Q, \qquad (3)$$

Q is the configurational partition function, and E_i is the potential energy due to intermolecular interactions in state i. This is the so called "Crude Monte Carlo" method and is not practical for two reasons. (1) The normalized probabilities, P_i, are not known apriori

since this implies a knowledge of the partition function. (2) At liquid state densities configurations with low probabilities are chosen much more often resulting in an inefficient sampling.

Metropolis et al. [79], proposed an important sampling scheme as a modification of the crude Monte Carlo method. The essential difference lies in choosing configurations with Boltzmann probabilities and weighting them evenly as opposed to choosing configurations randomly and weighting them with Boltzmann probabilities. Thus the averages are expressed as

$$\langle f \rangle = (1/M) \sum_{i}^{M} f_i \qquad (4)$$

The summation extends over all states from 1 to M sampled, each state being defined as a point in configuration space. This is accomplished in the following manner. Choose a particle to move and generate a new state j from the old state i by changing its coordinates as follows.

$i \rightarrow j$

$x \rightarrow x + a z_1$

$y \rightarrow y + a z_2$

$z \rightarrow z + a z_3$

'a' is the maximum displacement allowed in one dimension. z_1, z_2 and z_3 are random numbers between -1

and +1. Then calculate $\Delta E_{ij} = (E_j - E_i)$. If $\Delta E_{ij} < 0$ accept the move. If $\Delta E_{ij} > 0$ generate another random number z_4. If $z_4 < \exp(-\Delta E_{ij}/kT)$ accept the move, otherwise reject the move and count the old configuration again.

To prove that the limiting distribution was canonical the authors argued that, (1) the method was ergodic since any particle was allowed to move to any point within a cube of side 2a with a finite probability, a large number of such moves would eventually let the particle reach any point in the complete cube, and that, (2) if $(n_i/n_j) > (\exp(-E_i/kT)/\exp(-E_j/kT))$, the algorithm was so constructed that more systems move from state i to state j and hence after many such moves the ensemble would reach Boltzmann distribution as the limiting distribution.

On the choice of 'a', it was pointed out that most moves would be forbidden if it was too large, and if it was too small the configuration would not change enough. In either case it would take longer to reach equilibrium. They chose 'a' such that about 50% of the moves were accepted.

In this Metropolis Monte Carlo the states i, j, .., M generated, constitute a realization of the stationary Markov chain [82], which is characterized by a one step transition probability matrix (P_{ij}) in which $P_{ij} \geq 0$ is

the conditional probability that if the system is in state 'i' at the nth step it would be in state 'j' at the (n+1)th step. The connection between the random walks in configuration space porposed by Metropolis et al., and finite stationary Markov chains is apparently a later development.

Wood [83] commenting on the necessary and sufficient conditions for the choice of P_{ij} stressed that the normalization condition

$$\sum_{j}^{M} P_{ij} = 1 \quad \text{for all } i, \tag{5}$$

together with the steady state condition

$$\sum_{i}^{M} P_i P_{ij} = P_j \quad \text{for all } j, \tag{6}$$

or the microscopic reversibility

$$P_i \; P_{ij} = P_j \; P_{ji}, \tag{7}$$

are adequate to ensure convergence of the Markov chain to a limiting stationary distribution. Wood further pointed out that conditions (5) and (6) or (7) are insufficient in number to completely determine the matrix (P_{ij}), accounting for the latitude in the choice of P_{ij}.

Hastings [84] gave a general formulation for the construction of the transition matrix (P_{ij}) starting with the assumption that

$$P_{ij} = Q_{ij}\, a_{ij} \quad (i \neq j) \tag{8}$$

and $$P_{ii} = 1 - \Sigma^{-}_{i \neq j} P_{ij} \tag{9}$$

$Q = (Q_{ij})$ is the transition matrix of an arbitrary Markov chain on the states $0,1,\ldots M$ and

$$a_{ij} = s_{ij}/(1+t_{ij}) \tag{10}$$

where

$$t_{ij} = P_i\, Q_{ij}/P_j\, Q_{ji} \tag{11}$$

and s_{ij} is a symmetric function of i and j chosen so that

$$0 \leq a_{ij} \leq 1 \text{ for all i and j.}$$

Q_{ij} determines the generation of a new state j from i and a_{ij} defines the acceptance criterion for the state so generated. The P_{ij}'s defined above satisfy the detailed balance requirement.

$$s_{ij}^{M} = \begin{cases} 1+t_{ij} & (\text{If } t_{ji} \geq 1) \\ 1+t_{ji} & (\text{If } t_{ij} \geq 1) \end{cases} \tag{12}$$

and

$$s_{ij}^{B} = 1. \tag{13}$$

For a symmetric Q (i.e. if $Q_{ij}=Q_{ji}$) $s_{ij}=s_{ij}^{M}$ corresponds to Metropolis choice and $s_{ij}=s_{ij}^{B}$ to Barker's [85] choice. More generally $s_{ij} = g\{\min(t_{ij},t_{ji})\}$ where g(x) is chosen such that

$$0 \leq g(x) \leq 1+x \text{ for } 0 \leq x \leq 1.$$

Peskun [86] through a consideration of the

asymptotic variance of the averages estimated through the Monte Carlo method showed that for a given transition matrix Q, the optimum symmetric function s_{ij} is the Metropolis choice, s_{ij}^{M}.

Valleau and Whittington [87] observed that since the Monte Carlo computations are based on an estimation of the products $\{f(X^N)\ P(X^N)\}$ and not the probabilities $P(X^N)$ themselves, there is considerable flexibility in choosing the distributions $\Pi(x^n)=W(x^N)\exp\{-E(x^N)/kT)\}$, which can vary depending on the quantity to be estimated for an efficient sampling.

Pangali, Rao and Berne [88] in an attempt to accelerate convergence in the Monte Carlo runs where the interaction potentials were strongly angle dependent proposed a force biased algorithm as an improvement over the Metropolis method. The modification lies in the Q_{ij} matrix. For Metropolis choice

$$Q_{ij} = (\Delta x_{\emptyset}\, \Delta y_{\emptyset}\, \Delta z_{\emptyset})^{-1} \quad \text{if } dR_j \in D$$
$$= \emptyset \quad \text{if } dR_j \notin D$$

In the force bias method

$$Q_{ij} = C \exp\{-(\nabla_{Rj}(V(R))\, dR_j/kT) \quad \text{for } \nabla\, dR \in D$$
$$= \emptyset \quad \text{for } \nabla\, dR \notin D,$$

where C is a normalization constant and ∇_{Rj} is the gradient operator with respect to positions and angles.

For monoatomic particles this is Q_{ij}= C exp($F_j dR_j/kT$) and for polyatomic molecules Q_{ij}=C exp(($F_j dR_j$+ $N_j dw_j$)/kT) where F_j, N_j, dR_j, and dw_j are the total force, torque, center of mass displacement and angular displacement of molecule j. In a subsequent publication [89] they incorporated an adjustable parameter λ in the exponential for Q_{ij} and argued that λ must be chosen as to maximize translational and rotational diffusion. Rossky et al. [90], suggested that a value of 1/2 for λ should substantially improve the acceptance probability.

Rossky, Doll and Friedman [90] derived a "Smart Monte Carlo" transition probability matrix from the Brownian dynamics algorithm, in which the particle displacements were made according to

$$\Delta r = A F/kT + R$$

where R is drawn from a Gaussian distribution, F represents the forces on the particle to be moved and A is related to the friction coefficient. The Q_{ij}'s are defined as

$$Q_{ij}= C \exp [-(\Delta r_i - A F_i/kT)^2/4A].$$

a_{ij} defining the acceptance criterion is the same as in Metropolis algorithm. If in the exponential, the linear terms in force alone are retained and if R is set to zero, this prescription corresponds to the force bias

algorithm mentioned above. For a given rms displacement per move of the particles the average acceptance ratio for smart Monte Carlo was expected to be larger than for the Metropolis Monte Carlo.

In our simulations we used the Q_{ij}'s of Pangali, Rao and Berne [88-89] and the a_{ij} of Metropolis et al. [79]. The excess thermodynamic internal energy U is computed as

$$U = < E(X^N) >$$

and the constant volume heat capacity as

$$C_V = \{ <E(X^N)^2> - <E(X^N)>^2 \}/k_B T^2.$$

The pressure is calculated as

$$P = (k_B T/V) (N - < \Sigma_{i=1}^{N} (R_i \, dE(X^N)/dR_i) >)$$

and the atom-atom spatial distribution function as

$$g_{ab} (R) = N_{ab} (R)/ d \, 4 \, \pi \, R^2 \, \Delta R$$

where R is the interatomic separation, N_{ab} is the average number of b neighbors of atom 'a' in a spherical shell between R and R + Δ R and d is the bulk density.

I. D. CALCULATIONS

Statistical thermodynamic (T,V,N) ensemble Monte Carlo simulations were carried out individually on dilute aqueous solutions of DMP^- anion and Na^+DMP^- ion pair in the gg, gt and tt conformations, using a modified Metropolis procedure [79] incorporating force bias method [89] and preferential sampling [91] for convergence acceleration. The system for study in each case consisted of 216 rigid particles, one DMP (anion/ ion pair), and 215 water molecules. The computer experiment was performed at 25°C and a density determined from the experimentally observed partial molar volume for water and derived partial molar volume (corresponding to infinite dilution) of 59.3 ml/mol for DMP^- [92], and 53.1 ml/mol for the ion pair [92-93]. The condensed phase environment of the system was simulated by face centered cubic boundary conditions which provides here in excess of two hydration shells for the solute. Convergence characteristics and statistical error bounds on each of the calculated quantities were monitored by the method of batch means [94]. Full details of the Monte Carlo procedure as applied here are described in the previous section and in a recent article by Mehrotra et al. [95].

The N-particle configurational energies of the system were calculated under the assumption of pairwise

additivity in intermolecular interactions using potential functions determined from ab initio quantum mechanical calculations. Water-water interactions were modelled by the MCY potential [96] developed by Matsuoka et al., and solute-water interaction by an analytical potential function developed by Clementi and coworkers. The performance of MCY water-water potential has been extensively documented in the recent literature [97], and is known to give good agreement with experimental radial distribution functions in simulations carried out at experimental density. The shortcomings arise in the neglect of cooperative effects, estimated to incur a 13% error in computed internal energies for $[H_2O]_1$ and an inordinately high calculated pressure indicative of deficiencies in the curvature of the potential. The latter problem is kept under control by working consistently with experimental densities in (T,V,N) ensemble simulations.

The DMP^--water interactions were computed using the 12-6-1 analytical potentials developed from quantum mechanical calculation by Clementi et al. [54]. The geometry for DMP^- adopted in our study, was that of Gorenstein and coworkers [44]. The (α, ζ) torsional angles were gg (60^o, 60^o), gt (60^o, 180^o) and tt (180^o, 180^o). The OPO valence angles for the gg, gt and tt conformations were taken to be 103.4^o, 97.5^o and 92^o

respectively, following Gorenstein et al. [44]. Net atomic charges for each of the conformers of DMP^- were computed using the Gaussian-80 system of programs [98] and the atomic orbital basis sets given by Matsuoka et al. [99], and consistent with the potential functions. We have looked at the behavior of this basis set for the conformational preferences of DMP^- in vacuum for both fixed [37] and optimized geometries [44]. By fixed geometry a constant value of 102.6^o is meant for the O-P-O valence angle in all three conformations and by optimized geometry, the optimized OPO valence angle for each conformation the values of which are quoted above. The energies relative to tt are -1.81 and -3.36 kcal/mol for gt and gg respectively in the fixed geometry, and 0.26 and 1.58 kcal/mol for the optimized geometry; the increased stability of the extended forms in vacuum upon optimization of the O-P-O valence angle is seen to be reproduced. A slice of the potential energy hypersurface for the DMP^-(g, g)H_2O interaction calculated from the 12-6-1 potential function in the $>PO_2^-$ plane is given in Figure 3. The energy minimum is in the plane of anionic oxygens and the interaction is worth ~-21.6 kcal/mole. This surface also indicates that an in-plane bridge structure is more stable than a coplanar sequential H-bond by about 2 kcal/mole. Next in interaction strength

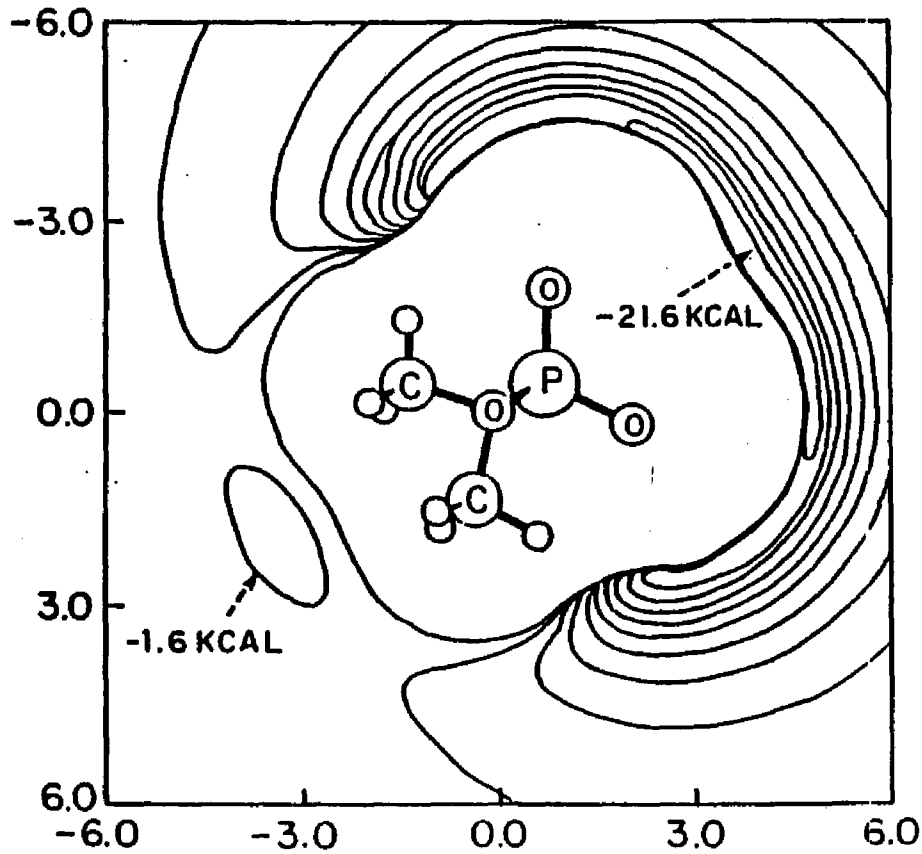

Figure 3. Isoenergy contour surface for DMP^- water dimer in the plane of anionic O-P-O^- group. Contour lines are separated by 2 kcal/mole; distances in Å.

are the planes of PO_1O_3 and PO_2O_4 with E ~ -16 kcal/mole, followed by the PO_3O_4 plane with E ~ -11 kcal/mole. Bridged structures are predicted to be the most stable structures for the gt and tt conformations as well. The stability of the bridged structures for the monohydrate complexes of DMP^-, as well as the interaction energies predicted by this potential function are in agreement with a majority of studies cited in the previous section.

In the case of Na^+DMP^- ion pair, sodium-water interactions were computed according to the "simple" model of Clementi and coworkers [100], the behavior of which was well characterized in an earlier publication [101] from this Laboratory. The interaction potential and the geometry for the anionic (DMP^-) part of the ion pair was identical to that described above. Sodium ion was held fixed in the anionic $>PO_2^-$ plane on the bisector of the OPO angle at a distance of 2.21 Å from the anionic oxygens, in conformity with the quantum mechanical calculations of Pack and coworkers [73] on metal-phosphate complexes. This geometry for sodium cation is also in accord with the other theoretical studies quoted in the background section. Partial atomic charges for each conformer of Na^+DMP^- were computed using Gaussian - 80 program [98] and basis sets of Clementi and coworkers [99,102].

In the computer simulations, solute-water interaction was treated under minimum image convention. Intra-solute and inter-solute interactions were not included. Since the solute modeled is never approached by an other solute molecule the simulated system essentially models infinite dilution. Water-water interactions were truncated at a spherical cutoff of 7.75 Å. Simulations on each conformer involved a total of ~3000 K configurations, preceded by 500K configurations of sampling which were treated as equilibration. The initial configuration was taken from a pre-equilibrated trial run on the appropriate conformation of $[DMP^-]_{aq}$. Ensemble averages for the mean energy are formed over the last 2000 K for each run separately.

We also carried out three additional sets of simulations on $[DMP^-]_{aq}$. The first set involved the anion in identical bond lengths, bond angles [37] and charges [54] in all the three conformations. This set, we call fixed geometry calculations, corresponds to the simulations described by Alagona et al.. The second set involved all the three conformations in optimized geometry [44] but with a larger cutoff of 10.5 Å for water-water interactions. The third set involved the anion in optimized geometry but without density constraints and corresponds to cluster calculations

widely published in the literature. Results of these three sets are quoted selectively.

These calculations were performed on the IBM 3081 machine and each of the twelve simulations on DMP^{-} with periodic boundary conditions took approximately 100 hours of CPU time for the production stage ($\sim$ 2000 K Monte Carlo moves) while the cluster simulations took approximately 50 hours for the same run length.

I. E. THE PROXIMITY CRITERION AND ANALYSIS

Monte Carlo simulations generate a set of configurations consistent with Boltzmann distribution. In order to glean the average microscopic environment around the solute one needs a unique analysis scheme starting from the configurations so generated. Proximity Criterion is one such scheme proposed [103-105], as an extension of the generalized molecular distribution function (GMDF) theory of Ben-Naim [106], to accomplish a compositional analysis of the statistical state of the system and gives a unique definition of the local environment of each identifiable substructure, atom or functional group of the solute. A brief description of the GMDF theory is given here followed by the proximity criterion.

Consider a system of N identical molecules. The supermolecular geometry of a given N-particle configuration is fully specified by the configurational coordinate $\underline{X}^N$

$$\underline{X}^N = \{\underline{X}_1, \underline{X}_2, \ \cdots \ \underline{X}_N\}$$

where $\underline{X}_i$ is the configurational coordinate of the particle i. For any molecular property Q which is a function of the coordinate $\underline{X}^N$, one can define a counting function

$$C_i^Q(\underline{X}^N, q) = d\ [q - Q_i(\underline{X}^N)]$$

where $Q_i(\underline{X}^N)$ gives the value of the property Q for particle i in the configuration $\underline{X}^N$ and d [ ] is the Dirac delta function. One then proceeds to define a Quasi Component Distribution Function (QCDF) for the property Q as

$$x_Q(q) = \frac{\int\cdots\int P(\underline{X}^N) \sum_{i=1}^{N} C_i^Q(\underline{X}^N,q) D_i^Q(\underline{X}^N) \, d\underline{X}^N}{\int\cdots\int P(\underline{X}^N) \sum_{i=1}^{N} D_i^Q(\underline{X}^N) \, d\underline{X}^N}$$

where $D_i^Q(\underline{X}^N)$ is a selector function whose value is either 0 or 1, which determines whether or not the particle i contributes to $x_Q(q)$. A similar definition of QCDF for properties which are functions of molecules i and j can be given as

$$x_Q(q) = \frac{\int\cdots\int P(\underline{X}^N) \sum_{i<j}^{N} C_{ij}^Q(\underline{X}^N,q) \, D(\underline{X}^N) \, d\underline{X}^N}{\int\cdots\int P(\underline{X}^N) \sum_{i<j}^{N} D_{ij}^Q(\underline{X}^N) \, d\underline{X}^N}$$

For discrete properties $x_Q(q)$ describes the configurational average of the fraction of molecules for which the property Q is exactly q. For continuous properties $x_Q(q)\,dq$ gives the configurationally averaged fraction of molecules with property Q in the interval [q, q+dq]. Configurational average of the property Q is obtained as

$Q = \int x_Q(q)\, dq$ for continuous properties

$Q = \sum_{\{q\}} x_Q(q) q$ for discrete properties.

To bring into QCDF, the volume element of the

configurational space sampled w.r.t the parameter q, another quantity called Quasi Component Correlation Function (QCCF) may be defined as

$$QCCF(q_Q(q)) = X_Q/v_Q(q).$$

The radial distribution function g(R) can be defined as the QCCF of distance. Take $Q_{ij}(\underline{X}^N)$ to be distance between particles i and j and use $D^R_{ij}(\underline{X}^N)=1$.

$$g_R(r) \equiv g(r) = X_R(r)/v_R(r) = X_R(r)/[4\pi r^2 N/(N-1)V]$$

The coordination number $C_i(\underline{X}^N)$ of i is defined as

$$C_i(\underline{X}^N) = \sum_{j=1}^{N} h(R_{ij}-R_c)$$

where $h(R_{ij}-R_c)$ is a unit step function equal to unity if the interparticle separation between i and j, R_{ij} is less than the radius of the coordination sphere R_c, which is chosen as the first minimum in the intermolecular center of mass g(R). The quantity $C_i(\underline{X}^N)$ gives the number of other particles j that fall within the first coordination sphere of particle i in configuration $\underline{X}^N$. The quantity running coordination number is simply the average coordination number $\bar{K}$ as a function of the cutoff radius R_c.

Binding energy of particle i in configuration $\underline{X}^N$ is defined as

$$B_i(\underline{X}^N) \equiv E(\underline{X}_1, \ldots \underline{X}_N) - E(\underline{X}_1, \ldots \underline{X}_{i-1}, \underline{X}_{i+1}, \ldots \underline{X}_N)$$

where E is the configurational energy of the system.

The thermodynamic configurational internal energy U is related to the average binding energy $\bar{V}$ by the expression $U = N\bar{V}/2$.

The pair energy between two particles i and j, P_{ij}, is the interaction energy between two particles. If only near neighbors are allowed to contribute the selector function is taken as

$$D^{P}_{ij}(\underline{X}^{N}) = h(R_{ij}-R_{c}). \text{ etc..}$$

The Proximity Criterion: Consider an infinitely dilute solution consisting of one solute molecule within a volume V together with N solvent molecules. The analysis is developed in terms of the coordinates of the N solvent molecules defined relative to the solute center of mass. In any given configuration each of the N solvent molecules is classified on the basis of the nearest solute atom A. The set of solvent molecules closer to A than to any other solute atom are designated as the total 1^{o} solvation of A. In geometrical terms, this is equivalent to saying that molecules that belong to the 1^{o} solvation fall into the Voronoi polyhedron of A, generated by the solute atoms and the boundary of the system. Higher orders of total solvation may also be similarly defined. The set of molecules for which A is the second nearest solute atom

gives the total 2^o solvation of A. Two normalization conditions for N_A that follow from the definition are

$$\Sigma_A^- N_A = N \quad \text{and}$$

$$\Sigma_k^- N_A^k = N \text{ for any A.}$$

N_A^k here is the total solvation number of A at order k.

For any given solvent molecule i, in an N-particle configuration of the system $\underline{X}^N$, collect as a set the solute atoms listed in order of k. The members of this set are the proximity indices for solvent molecule i, $S^k(\underline{X}^N)$. Consider this set as a generalized property of the system in the context of the GMDF theory.

$$S_i(\underline{X}^N) = \{S_i^{1o}(\underline{X}^N), S_i^{2o}(\underline{X}^N), \ldots\}$$

where $S_i^{1o}(R^N) = \{A/R_{Ai} = \min_M\{R_{Mi}\}\}$

i. e. the primary index of the solvent molecule is the solute atom A such that the distance R_{Ai} is the absolute minimum in the discrete set $\{R_{Mi}\}$ of all distances between the n solute atoms and the COM of the i th solvent molecule.

With the proximity indices thus defined for all solvent molecules one may develop an analysis of the solvation of a solute molecule atom by atom. For every QCDF $X_Q(q)$, one can define the k th order proximity QCDF $X_Q^k(q)$ by multiplying the selector function by

$d[A\text{-}S_i^k(R^N)]$. The extension of this analysis scheme to functional groups is straightforward. The selector function in this case is $\Sigma_{A \in \{F\}} d[A\text{-}S_i^k(R^N)]$ where {F} defines a functional group.

I. F. RESULTS

The calculated internal energies and related quantities for $[DMP^-]_{aq}$ and $[Na^+DMP^-]_{aq}$ are collected in Table I. The quantities entered here are the mean energy $\langle U_{sw} \rangle$ of the system (N_S=1, N_W=215), the energy $\langle U_W \rangle$ of 215 water molecules in $[H_2O]_l$ at 25°C, the corresponding energy $\langle U_W^{\prime} \rangle$ of solvent water in $[DMP]_{aq}$, the calculated partial molar internal energy of transfer for DMP into water, $\langle U_S \rangle$ and finally $\langle U_S^{\prime} \rangle$ and $\langle U_{rel} \rangle$, the solute-solvent and solvent-solvent contributions to $\langle U_S \rangle$. Each of these is formally defined in equations 1-12 and Figure 4 of a previous publication from this Laboratory by Swaminathan et al. [107]. The statistical noise levels upto a confidence limit of 95% (2 σ) on each of these quantities are also indicated in Table I underneath the corresponding row. The bottom row in Table I gives the calculated conformational energy differences relative to the trans extended form. The hydration energy is seen to favor the gt form in the free anion and the trans extended form in the ion pair. The statistical uncertainties in the total internal energies of hydration indicate that gg and gt conformers are not well differentiated in DMP.

Table I. Monte Carlo mean energy simulation results on $[DMP^-]_{aq}$ and $[Na^+DMP^-]_{aq}$

	DMP^-(gg)	DMP^-(gt)	DMP^-(tt)	Na^+DMP^-(gg)	Na^+DMP^-(gt)	Na^+DMP^-(tt)
<USW>	-1859.25	-1864.55	-1854.58	-1839.77	-1822.38	-1853.01
+/-2*SD	5.46	5.36	8.29	10.30	10.15	7.39
<US'>	-41.47	-31.74	-32.88	-16.31	-14.24	-34.15
+/-2*SD	3.22	1.90	2.45	4.08	2.24	1.87
<UW'>	-1817.78	-1832.81	-1821.70	-1823.46	-1818.14	-1818.86
+/-2*SD	4.41	5.01	7.92	9.47	9.90	7.15
<UW>	-1859.75	-1859.75	-1859.75	-1859.75	-1859.75	-1859.75
+/-2*SD	6.45	6.45	6.45	6.45	6.45	6.45
<Urel>	41.97	26.94	38.05	36.04	51.61	40.89
+/-2*SD	4.71	4.86	4.60	6.93	7.51	3.09
<US>	0.50	-4.88	5.17	19.73	37.37	6.74
+/-2*SD	5.71	4.48	5.21	8.04	7.84	3.61
Δ <US>	-4.67	-9.97	0.00	12.99	30.63	0.00
+/-2*SD	2.34	2.66		7.18	6.96	

The detailed analysis of the results is based on the Proximity Criterion described in the previous section. Tables II to VII summarize the results of the proximity analysis on $[DMP^-]_{aq}$ and $[Na^+DMP^-]_{aq}$ simulations. Column 1 of these tables lists the atom or functional group for which the analyses results are presented in the corresponding row. The sequence adopted is first methyl groups, followed by ester oxygens and then anionic $>PO_2^-$ group (and Na^+ for the ion pair) ending up with sums / averages for the entire molecule. Column 2 (RFS) gives the first shell radius in Å corresponding to the first minimum of the primary radial distribution functions of the relevant atom or functional group. The first shell proximity analyses of the simulations are based on these limits. In column 3 are indicated volumes (VFS) in $Å^3$ of the truncated spherical shell of the Voronoi polyhedron associated with the primary region of the solute analysis unit, be it solute atom or functional group. These volumes are generated by Monte Carlo method using a million random points. Column 4 gives the average coordination number <K>, computed as the average number of water molecules in the volume VFS. This is followed by local solvent densities <K/V> in Column 5 in gm/ml. The average first shell solute binding energies <SLTBE> are reported in column 6 in kcal/mole. These represent the net

Table II. Proximity analysis of the hydration of DMP^-(g,g)

AT NO	INDEX	TYPE		RFS	VFS	FIRST SHELL SOLUTE PROPERTIES <E>	<E/V>	<SLUTE>	<SLUPE>	TOTAL SLT PROPS <E>	<SLUTE>	WATER PROPERTIES RFS=1.38 RCP= 7.75 A <ENE>	<ENAPE>	<ERAPE>	
METHYL GROUPS															
1	6	53	C C1	5.3	2.82	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	7	61	H H1C1	4.2	95.52	2.91	0.91	-12.155	-4.284	22.106	-4.784	4.29	-2.937	-17.297	
3	8	61	H H2C1	4.2	67.98	1.78	0.75	1.685	0.945	12.964	8.289	4.24	-2.989	-17.188	
4	9	54	H H3C1	4.2	125.14	3.48	0.81	5.988	1.735	38.913	28.879	4.27	-3.828	-17.658	
TOTALS FOR FUNCTIONAL GROUP					291.47	8.01	8.82	-4.949	-0.618	65.983	24.385	4.18	-2.998	-17.428	
5	10	53	C C2	5.3	2.81	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
6	11	61	H H1C2	4.2	95.43	3.28	1.83	-14.347	-4.378	22.376	-7.789	4.28	-2.995	-16.859	
7	12	61	H H2C2	4.2	68.09	1.92	0.85	3.297	1.713	13.389	9.449	4.21	-3.868	-17.587	
8	13	54	H H3C2	4.2	125.48	3.47	0.83	3.683	1.862	38.488	18.583	4.18	-3.816	-17.244	
TOTALS FOR FUNCTIONAL GROUP					291.73	8.67	8.89	-7.367	-0.858	66.173	28.242	4.19	-3.818	-17.183	
AVERAGES OVER FUNCTIONAL GROUPS:—					291.60	8.34	8.86	-6.158	-0.734	66.078	22.274	4.25	-3.884	-17.386	-CH3
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.16	0.02	0.425	0.051	0.02	1.912	0.02	0.056	0.106	
ESTER OXYGENS															
9	4	52	O O1ES	3.2	29.58	1.25	1.27	-13.467	-10.747	6.161	-15.388	4.05	-2.921	-14.788	
10	5	52	O O4ES	3.2	29.54	0.75	0.76	-5.181	-6.838	5.155	-6.848	4.16	-2.989	-16.248	
AVERAGES OVER FUNCTIONAL GROUPS:—					29.56	1.88	1.81	-9.284	-8.788	5.658	-10.674	4.18	-2.915	-15.514	-O-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.06	0.05	1.848	1.764	0.01	3.329	0.06	0.182	0.329	
PO2- GROUP															
11	1	51	P PO2S	5.8	6.22	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
12	2	55	O O1AN	3.8	59.42	2.13	1.87	-31.934	-15.932	35.816	-34.323	4.33	-2.866	-16.387	
13	3	55	O O2AN	3.8	59.33	2.12	1.87	-32.233	-15.174	35.512	-30.349	4.38	-2.879	-16.728	
TOTALS FOR FUNCTIONAL GROUP					124.97	4.25	1.32	-66.168	-15.554	71.528	-64.672	4.32	-2.872	-16.557	>PO2-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.16	0.04	9.819	2.137	0.03	7.937	0.02	0.071	0.134	
DMP—															
MOLECULAR SUM/AVERAGE:					767.29	22.93	0.09	-97.852	-4.233	215.300	-41.472	4.26	-2.955	-16.957	DMP-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.27	0.01	5.699	0.250	0.04	3.218	0.01	0.043	0.280	

Table III. Proximity analysis of the hydration of DMP^-(g,t)

AT NO	INDEX	TYPE		RFS	FIRST SHELL SOLUTE PROPERTIES VFS	<N>	<N/V>	<SLTNB>	<SLTPE>	TOTAL SLT PROPS <N>	<SLTNB>	WATER PROPERTIES RFSW=3.30 RCD= 7.75 A <ND>	<WWLPE>	<BEWAT>	
METHYL GROUPS															
1	6	53	C C1	5.3	2.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	7	61	H H1C1	4.2	97.39	3.29	0.90	-6.060	-2.140	21.003	0.395	4.21	-2.992	-17.012	
3	8	61	H H2C1	4.2	96.26	2.87	0.89	-1.009	-0.351	22.638	6.019	4.23	-3.060	-17.544	
4	9	54	H H3C1	4.2	124.11	3.69	0.89	0.978	0.265	29.291	17.338	4.23	-2.999	-17.395	
TOTALS FOR FUNCTIONAL GROUP					320.47	9.76	0.91	-6.099	-0.707	73.011	23.751	4.23	-3.010	-17.330	
5	10	53	C C2	5.3	2.01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
6	11	61	H H1C2	4.2	85.35	2.44	0.85	-0.762	-0.312	15.600	6.320	4.23	-2.897	-17.015	
7	12	61	H H2C2	4.2	95.40	3.09	0.97	-7.260	-2.346	21.667	1.689	4.26	-3.033	-17.209	
8	13	54	H H3C2	4.2	129.27	4.07	0.94	3.354	0.824	33.495	16.060	4.33	-3.002	-17.960	
TOTALS FOR FUNCTIONAL GROUPS					312.90	9.60	0.92	-4.649	-0.484	70.762	24.069	4.28	-2.979	-17.540	
AVERAGES OVER FUNCTIONAL GROUPS:=					316.69	9.68	0.91	-5.774	-0.596	71.887	23.910	4.25	-2.999	-17.435	-CH3
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.10	0.02	0.182	0.017	0.03	1.493	0.02	0.032	0.086	
ESTER OXYGENS															
9	4	52	O O3ES	3.2	28.20	0.69	1.02	-5.649	-0.175	1.330	-5.771	3.75	-2.915	-13.160	
10	5	52	O O4ES	3.2	35.13	0.83	0.71	-7.293	-0.759	7.528	-3.930	4.20	-2.895	-16.449	
AVERAGES OVER FUNCTIONAL GROUPS:=					27.66	0.76	0.82	-6.471	-0.467	4.429	-4.851	3.98	-2.905	-14.805	-O-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.05	0.05	0.725	0.077	0.01	1.381	0.06	0.123	0.280	
PO2- GROUP															
11	1	51	P PHOS	5.0	6.26	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
12	2	55	O O1AN	3.0	52.21	2.42	1.38	-38.318	-15.059	26.171	-31.000	4.24	-2.880	-16.627	
13	3	55	O O2AN	3.0	60.76	2.46	1.21	-38.585	-15.655	36.197	-38.056	4.22	-2.919	-16.391	
TOTALS FOR FUNCTIONAL GROUP					119.23	4.88	1.22	-76.903	-15.756	62.368	-69.057	4.23	-2.906	-16.498	>PO2-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.18	0.04	4.005	0.912	0.03	6.921	0.03	0.046	0.120	
DMP-															
MOLECULAR SUM/AVERAGE:					807.93	25.76	0.95	-101.393	-3.936	215.000	-31.738	4.24	-2.968	-17.097	DMP-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.42	0.01	2.761	0.099	0.06	1.903	0.01	0.025	0.068	

Table IV. Proximity analysis of the hydration of $DMP^{-}(t,t)$

AT NO	INDEX	TYPE		RFS	FIRST SHELL SOLUTE PROPERTIES VPS	<D>	<K/V>	<SLTSD>	<SLTPE>	TOTAL SLT PROPS <D>	<SLTSD>	WATER PROPERTIES RFSA=3.38 RCB= 7.75 A <SD>	<SLTPE>	<SLTAE>	
METHYL GROUPS															
1	6	53	C C1	5.3	2.83	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	7	61	H H1C1	4.2	97.98	3.81	0.92	-9.863	-3.889	21.278	-2.877	4.25	-3.818	-17.181	
3	8	61	H H2C1	4.2	98.26	3.16	0.96	-8.268	-8.882	21.174	4.838	4.26	-2.987	-17.369	
4	9	54	H H3C1	4.2	129.59	4.87	0.94	1.988	0.487	35.788	17.388	4.25	-2.993	-17.318	
TOTALS FOR FUNCTIONAL GROUP					328.58	10.24	0.93	-7.343	-0.717	78.239	28.869	4.25	-2.998	-17.269	
5	10	53	C C2	5.3	2.83	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
6	11	61	H H1C2	4.2	98.26	2.87	0.87	-1.991	-0.695	21.914	5.872	4.25	-2.999	-17.338	
7	12	61	H H2C2	4.2	97.98	2.91	0.89	-5.398	-1.852	22.173	2.248	4.45	-2.975	-17.868	
8	13	54	H H3C2	4.2	129.59	4.02	0.93	-0.599	-0.149	34.833	15.749	4.21	-2.976	-17.133	
TOTALS FOR FUNCTIONAL GROUP					326.58	9.88	0.89	-7.985	-0.815	78.121	23.868	4.29	-2.982	-17.377	
AVERAGES OVER FUNCTIONAL GROUPS:—					328.58	10.82	0.91	-7.664	-0.766	78.180	21.569	4.27	-2.990	-17.323	-CH3
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.26	0.02	0.123	0.026	0.02	1.791	0.02	0.027	0.087	
ESTER OXYGENS															
9	4	52	O OES	3.2	38.67	0.81	0.79	-7.977	-0.795	3.379	-0.398	4.08	-2.914	-14.436	
10	5	52	O OES	3.2	38.67	0.70	0.68	-6.261	-0.951	3.311	-6.725	4.03	-2.858	-14.533	
AVERAGES OVER FUNCTIONAL GROUPS:—					38.67	0.76	0.74	-7.119	-0.373	3.345	-7.561	4.02	-2.886	-14.485	-O-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.07	0.07	0.415	1.179	0.08	3.087	0.07	0.127	0.352	
PO2- GROUP															
11	1	51	P PHOS	5.8	6.83	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
12	2	55	O O1AN	3.8	54.25	2.58	1.42	-29.928	-15.588	25.761	-36.185	4.21	-2.955	-16.427	
13	3	55	O O2AN	3.8	54.87	2.10	1.16	-32.863	-15.589	26.189	-24.785	4.31	-2.887	-16.417	
TOTALS FOR FUNCTIONAL GROUP					114.35	4.68	1.22	-72.723	-15.548	51.956	-68.898	4.26	-2.928	-16.422	>PO2-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.25	0.07	2.422	1.111	0.02	0.932	0.03	0.046	0.143	
DMP-															
MOLECULAR SUM/AVERAGE:					832.85	26.23	0.94	-102.298	-3.899	215.888	-32.875	4.26	-2.978	-17.815	DMP-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.59	0.02	1.437	0.118	0.04	2.458	0.01	0.023	0.073	

Table V. Proximity analysis of the hydration of $Na^{+}DMP^{-}$(g,g)

AT NO	INDEX	TYPE		RPS	VPS	FIRST SHELL SOLUTE PROPERTIES				TOTAL SLT PROPS		WATER PROPERTIES RFSH=3.30 RCD= 7.75 Å			
METHYL GROUPS															
1	6	53	C C1	5.3	2.88	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	7	61	H H1C1	4.2	92.83	2.97	0.96	-3.248	-1.892	22.668	6.781	4.27	-2.947	-17.235	
3	8	61	H H2C1	4.2	66.26	1.86	0.84	2.581	1.384	18.998	9.887	4.16	-3.897	-17.472	
4	9	54	H H3C1	4.2	121.42	3.53	0.87	5.983	1.672	28.395	16.885	4.18	-3.828	-17.378	
TOTALS FOR FUNCTIONAL GROUP					283.39	8.36	0.88	5.244	0.627	62.852	33.393	4.21	-3.818	-17.342	
5	10	53	C C2	5.3	2.85	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
6	11	61	H H1C2	4.2	92.92	2.75	0.89	8.285	0.875	28.454	13.647	4.27	-2.997	-17.677	
7	12	61	H H2C2	4.2	66.49	2.01	0.90	3.649	1.816	18.957	9.821	4.13	-3.852	-17.139	
8	13	54	H H3C2	4.2	121.84	3.45	0.85	6.081	1.762	30.383	22.378	4.28	-3.029	-17.722	
TOTALS FOR FUNCTIONAL GROUP					284.89	8.21	0.86	9.936	1.289	61.714	45.846	4.25	-3.818	-17.683	
AVERAGES OVER FUNCTIONAL GROUPS:—					283.74	8.29	0.87	7.590	0.918	61.883	39.219	4.23	-3.814	-17.473	-CH3
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.25	0.03	0.332	0.048	0.031	12.924	0.01	0.044	0.148	
ESTER OXYGENS															
9	4	52	O O3ES	3.2	28.96	0.95	0.98	-7.316	-7.688	4.448	-5.497	4.14	-2.781	-15.378	
10	5	52	O O4ES	3.2	28.95	0.90	0.92	-4.258	-4.757	5.376	-3.902	4.20	-2.879	-15.912	
AVERAGES OVER FUNCTIONAL GROUPS:—					28.95	0.92	0.95	-5.787	-6.223	4.912	-4.699	4.17	-2.838	-15.645	-O-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.18	0.18	0.897	1.163	0.889	5.497	0.04	0.146	0.469	
PO2- GROUP															
11	1	51	P PHOS	5.0	6.82	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
12	2	55	O O1AN	3.6	42.87	0.74	0.52	-4.977	-6.699	14.452	-4.982	4.38	-2.941	-16.889	
13	3	55	O O2AN	3.0	42.64	1.19	0.83	-9.799	-8.237	14.764	-8.985	4.46	-2.823	-16.888	
TOTALS FOR FUNCTIONAL GROUP					91.53	1.93	0.63	-14.776	-7.646	29.216	-13.967	4.38	-2.888	-16.849	>PO2-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.12	0.04	1.329	0.829	0.038	9.473	0.03	0.086	0.293	
NA+ CATION															
14	14	2	NA NA+	3.0	72.39	4.97	2.85	-86.957	-17.485	52.194	-71.383	4.39	-2.762	-16.211	NA+
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.23	0.10	5.858	1.418	0.048	36.223	0.02	0.062	0.211	
DMP-.NA+															
MOLECULAR SUM/AVERAGE:					789.30	25.33	0.96	-98.127	-3.874	215.088	-16.218	4.29	-2.925	-17.888	DMP-.NA+
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.57	0.02	3.253	0.155	0.082	4.078	0.01	0.032	0.189	

Table VI. Proximity analysis of the hydration of Na^+DMP^-(g,t)

AT NO	INDEX	TYPE		RPS	VPS	FIRST SHELL SOLUTE PROPERTIES <D>	<D/V>	<ELTRS>	<ELTPS>	TOTAL SLT PROPS <D>	<ELTRS>	WATER PROPERTIES RFSH=3.38 RCS= 7.75 Å <SD>	<SHAPE>	<SHAST>	
METHYL GROUPS															
1	6	53	C C1	5.3	2.77	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	7	61	H H1C1	4.2	93.22	2.79	0.89	-2.454	-0.881	28.428	7.816	4.23	-3.817	-17.264	
3	8	61	H H2C1	4.2	83.28	2.21	0.79	3.433	1.556	15.499	13.173	4.11	-3.879	-17.262	
4	9	54	H H3C1	4.2	125.66	4.17	0.99	5.823	1.286	32.182	20.493	4.23	-3.839	-17.215	
TOTALS FOR FUNCTIONAL GROUP					304.93	9.16	0.90	6.802	0.655	68.828	41.482	4.28	-3.841	-17.241	
5	10	53	C C2	5.3	2.79	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
6	11	61	H H1C2	4.2	89.69	3.07	1.02	0.347	0.113	17.955	0.711	4.28	-3.134	-17.408	
7	12	61	H H2C2	4.2	91.43	2.48	0.81	-3.424	-1.379	18.126	5.271	4.21	-3.885	-17.131	
8	13	54	H H3C2	4.2	120.58	3.56	0.88	6.889	1.712	26.111	19.418	4.37	-3.836	-17.952	
TOTALS FOR FUNCTIONAL GROUP					304.49	9.11	0.89	3.812	0.331	62.193	33.392	4.27	-3.855	-17.549	
AVERAGES OVER FUNCTIONAL GROUPS:—					304.71	9.13	0.90	4.587	0.493	65.110	37.437	4.24	-3.848	-17.395	-CH3
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.26	0.03	0.134	0.814	0.816	7.572	0.02	0.038	0.073	
ESTER OXYGENS															
9	4	52	O O3ES	3.2	28.04	0.23	0.35	-0.873	-1.726	0.784	-0.473	3.71	-3.181	-15.682	
10	5	52	O O4ES	3.2	34.12	0.95	0.83	-3.879	-4.888	7.762	-1.598	4.14	-2.988	-16.427	
AVERAGES OVER FUNCTIONAL GROUPS:—					27.88	0.59	0.65	-2.376	-3.983	4.273	-1.035	3.93	-3.048	-16.814	-O-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.07	0.08	0.275	0.422	0.804	0.817	0.07	0.149	0.262	
PO2- GROUP															
11	1	51	P PHOS	5.0	6.18	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
12	2	55	O O1AN	3.0	45.52	1.61	1.06	-13.484	-0.334	20.281	-7.699	4.30	-2.985	-16.735	
13	3	55	O O2AN	3.0	36.91	0.86	0.78	-7.463	-0.664	9.261	-6.788	4.41	-2.889	-16.543	
TOTALS FOR FUNCTIONAL GROUP					88.52	2.47	0.83	-20.947	-0.449	29.462	-14.487	4.33	-2.874	-16.674	>PO2-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.15	0.05	1.388	0.492	0.815	6.126	0.04	0.076	0.147	
NA+ CATION															
14	14	2	NA NA+	3.0	70.38	4.95	2.18	-81.887	-16.543	46.771	-72.633	4.36	-2.686	-15.681	NA+
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.24	0.10	4.049	0.764	0.819	24.512	0.03	0.056	0.110	
DMP-.NA+															
MOLECULAR SUM/AVERAGE:					822.48	26.87	0.98	-98.491	-3.666	215.080	-14.235	4.27	-2.941	-16.874	DMP-.NA+
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.60	0.02	2.272	0.079	0.042	2.241	0.01	0.029	0.055	

Table VII. Proximity analysis of the hydration of $Na^{+}DMP^{-}(t,t)$

AT NO	INDEX	TYPE		RFS	VFS	FIRST SHELL SOLUTE PROPERTIES <O>	<Q/V>	<ELTOT>	<ELTOT>	TOTAL SLT PROPS <O>	<ELTOT>	WATER PROPERTIES RFS=3.38 RCD= 7.75 Å <OLD>	<DELAPED>	<DELAY>	
METHYL GROUPS															
1	6	53	C C1	5.3	2.73	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
2	7	61	H H1C1	4.2	92.35	2.62	0.85	-3.342	-1.275	19.638	5.923	4.21	-2.986	-17.082	
3	8	61	H H2C1	4.2	92.16	2.66	0.86	-4.365	-1.641	21.128	4.740	4.28	-2.971	-17.412	
4	9	54	H H3C1	4.2	125.76	4.11	0.98	7.858	1.716	30.692	16.462	4.17	-3.812	-17.435	
TOTALS FOR FUNCTIONAL GROUP					312.99	9.40	0.90	-0.649	-0.869	71.458	27.125	4.21	-2.992	-17.331	
5	10	53	C C2	5.3	2.73	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
6	11	61	H H1C2	4.2	92.16	2.73	0.89	-6.853	-2.215	19.382	4.873	4.21	-3.836	-17.255	
7	12	61	H H2C2	4.2	92.35	2.73	0.89	-3.783	-1.355	18.444	6.465	4.21	-3.888	-17.317	
8	13	54	H H3C2	4.2	125.76	4.26	1.01	3.481	0.817	34.556	14.863	4.28	-3.878	-17.985	
TOTALS FOR FUNCTIONAL GROUP					312.99	9.72	0.93	-6.274	-0.645	72.382	26.201	4.25	-3.845	-17.578	
AVERAGES OVER FUNCTIONAL GROUPS:—					312.99	9.56	0.91	-3.462	-0.357	71.888	26.663	4.23	-3.819	-17.454	-CH3
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.22	0.02	0.128	0.014	0.026	1.789	0.02	0.027	0.088	
ESTER OXYGENS															
9	4	52	O O3ES	3.2	38.13	0.61	0.61	-3.238	-5.382	3.012	-1.671	3.98	-3.857	-16.134	
10	5	52	O O4ES	3.2	38.13	0.66	0.65	-3.148	-4.783	3.812	-1.267	3.87	-3.818	-15.925	
AVERAGES OVER FUNCTIONAL GROUPS:—					38.13	0.63	0.63	-3.193	-5.842	3.812	-1.469	3.88	-3.833	-16.029	-O-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.07	0.07	0.541	0.984	0.085	0.481	0.09	0.133	0.394	
PO2- GROUP															
11	1	51	P PHOS	5.0	5.96	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
12	2	55	O O1AN	3.0	39.66	1.64	1.24	-16.878	-9.782	13.362	-11.198	4.27	-2.842	-16.685	
13	3	55	O O2AN	3.0	39.52	0.96	0.72	-8.922	-9.333	12.355	-6.583	4.22	-2.917	-16.681	
TOTALS FOR FUNCTIONAL GROUP					85.14	2.60	0.91	-24.992	-9.617	25.716	-17.781	4.25	-2.878	-16.683	>PO2-
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.14	0.05	2.851	0.909	0.822	2.808	0.05	0.061	0.199	
NA+ CATION															
14	14	2	NA NA+	3.0	69.16	4.89	2.11	-88.738	-16.519	39.499	-66.838	4.34	-2.649	-15.588	NA+
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.22	0.09	5.347	1.259	0.827	8.556	0.04	0.845	0.149	
DMP-.NA+															
MOLECULAR SUM/AVERAGE:					840.54	27.87	0.99	-119.848	-4.271	215.888	-34.151	4.24	-2.932	-16.962	DMP-.NA+
STATISTICAL UNCERTAINTY (+/- 2*SD)						0.53	0.02	3.279	0.148	0.863	1.874	0.02	0.022	0.07	

interaction energies of each analysis unit with water molecules in their first shell. The next column gives the average first shell pair interaction energy <SLTPE> computed by dividing the elements in column 6 by those in column 4 for each row. In column 8 are indicated the total number of water molecules contained in the Voronoi polyhedron of each analysis unit. The molecular sum for this column adds up to 215, the total number of water molecules in the central cell. Column 9 gives the total solute binding energy in kcal/mole, which adds to $\langle U_S^i \rangle$ for the entire molecule. Columns 10 to 12 summarize water properties as modified by the presence of the solute. <KW> is the average coordination number of water molecules belonging to the primary region of the atom /functional group. The reference here is to be made to the MCY liquid water value of 4.34 [97]. <NNWWPE> is the average near neighbor pair interaction energy of water molecules in each primary region. The value to be compared with is -3.01 kcal/mole for MCY water. <BEWWT> is the total pair interaction energy in kcal/mole of all the waters in each Voronoi polyhedron. The corresponding MCY water value is -17.3 kcal/mole [97]. Column 13 gives functional groups for which the averages are reported in the corresponding row.

The RFS value adopted for carbon atoms is 5.3 Å and is 4.2 Å for methyl hydrogens. This is consistent with

earlier studies from this Laboratory on $[CH_4]_{aq}$ [107]. The first shell radial cutoff for ester oxygens is 3.2, slightly less than 3.3 for oxygens in liquid water, while that for anionic oxygens is 3.0. This contraction of the first shell is explicable in terms of the anionic nature or larger negative charge carried by the anionic oxygens. The value of 3.0 for Na^+ conforms to that adopted in the ion-water studies published from this Laboratory [101].These quantities are chosen from the position of the first minimum in the corresponding primary radial distribution function in the proximity analysis.

The volumes VFS can give a quick estimate of the average number of water molecules expected at liquid densities. Taking the average volume of water molecule in bulk water as 30 $Å^3$, the average coordination numbers expected for methyl groups is 10, one per ester oxygens and four for the PO_2^- for gg conformer of DMP^-, if the environment were bulk water like. The Voronoi volumes for the methyl groups in the extended conformations (gt & tt) increase relative to gg, while that of anionic oxygens decrease. The total first shell volumes for the entire molecule show a trend of gg < gt < tt. The Voronoi volumes for the ion pair for each conformer are larger than those for the corresponding anion. These volume elements are used in defining local

solvent densities and discussing the transferabilities of coordination numbers for a given atom or functional group in different molecules.

Structural features of DMP hydration : The calculated coordination numbers for DMP^- averaged over the two methyl groups from the detailed simulation analysis are 8.34, 9.68 and 10.02 for gg, gt and tt conformations respectively, showing a slight conformational trend. The corresponding averages for the ion pair are 8.29, 9.13 and 9.56 for gg, gt and tt conformations and parallel the conformational trend in the hydrated anion. Methyl group coordination is observed to be relatively unperturbed by the presence of counterion for the gg conformer. The extended forms however, show a slight decrease compared to their anionic counterparts. The average coordination number for the ester oxygens in gg form of DMP^- is 1.00 while that for gt and tt is 0.76, showing a decrease for tt not expected from the volumes of the primary regions. The corresponding averages for the ion pair follow similar pattern with 0.95 for gg, 0.65 for gt and 0.63 for tt. The average coordination number for PO_2^- group of gg conformation of DMP^- is significantly less (4.25) than for the extended conformations (4.88 for gt and 4.68 for tt). The corresponding coordination number for $>PO_2^-$ for the ion pair are of course much smaller due to

the presence of the sodium ion in the ion pair. The cation blocks some of the potential sites of hydration of the anionic oxygens. Here again the average for gg is (1.93) less than that of gt (2.47) and tt (2.60), indicating that statistical noise does not obscure the smaller coordination for $>PO_2^-$ of gg conformer. Coordination number for sodium cation is close to 5 and is seen to be relatively insensitive to conformational changes. Thus counting the two anionic oxygens together, sodium ion is heptacoordinated. This is in contrast to the octahedral coordination, with six waters in the first shell, found from the $[Na^+]_{aq}$ simulations [101, 108]. This appears to suggest that modifications are brought about in the solvent structure around the cation by the anionic $>PO_2^-$ group, but this may be sensitive to the assumed RFS values and the location of the sodium ion.

The total coordination numbers for DMP^- are 22.93 for gg, 25.76 for gt and 26.23 for tt. The corresponding figures for the ion pair are overall larger with 25.33 for gg, 26.87 for gt and 27.87 for tt. The reduction in $>PO_2^-$ group hydration is more than offset by the hydration complement of the cation to the whole molecule. Here again the conformational trends of gg< gt< tt are noticeable.

Local solvent densities show some interesting

trends. Hydrophobic hydration is characterized by a lower density (~10%) relative to the bulk water value in all six DMP (anion and ion pair) experiments. Ester oxygens show conformational sensitivity both for the anion and the ion pair, with a trend of gg> gt> tt. The anionic $>PO_2^-$ group indicates a considerable increase over the liquid water value resulting from electrostriction although the gg conformer of DMP^- is somewhat an exception. The solvent density of 2.05 for the sodium cation is double that of the bulk water value, clearly giving a structural criterion for electrostriction. The molecular averages for the anion are slighty less than that of liquid water indicating hydrophobic hydration dominates the overall first shell solvent densities. Values for the ion pair are closer to the bulk water value indicating that hydration of the counterion compensates the effect of hydrophobic hydration on the local solvent densities.

Energetics of DMP hydration : A large contribution to the first shell binding energies of the anion comes from the anionic ($>PO_2^-$) group hydration followed by hydrophilic (ester oxygens) and hydrophobic (methyl groups) hydration in all the three conformations. The first shell energetics of the ion pair likewise is dominated by the anionic and then the hydrophilic hydrations. The first shell binding

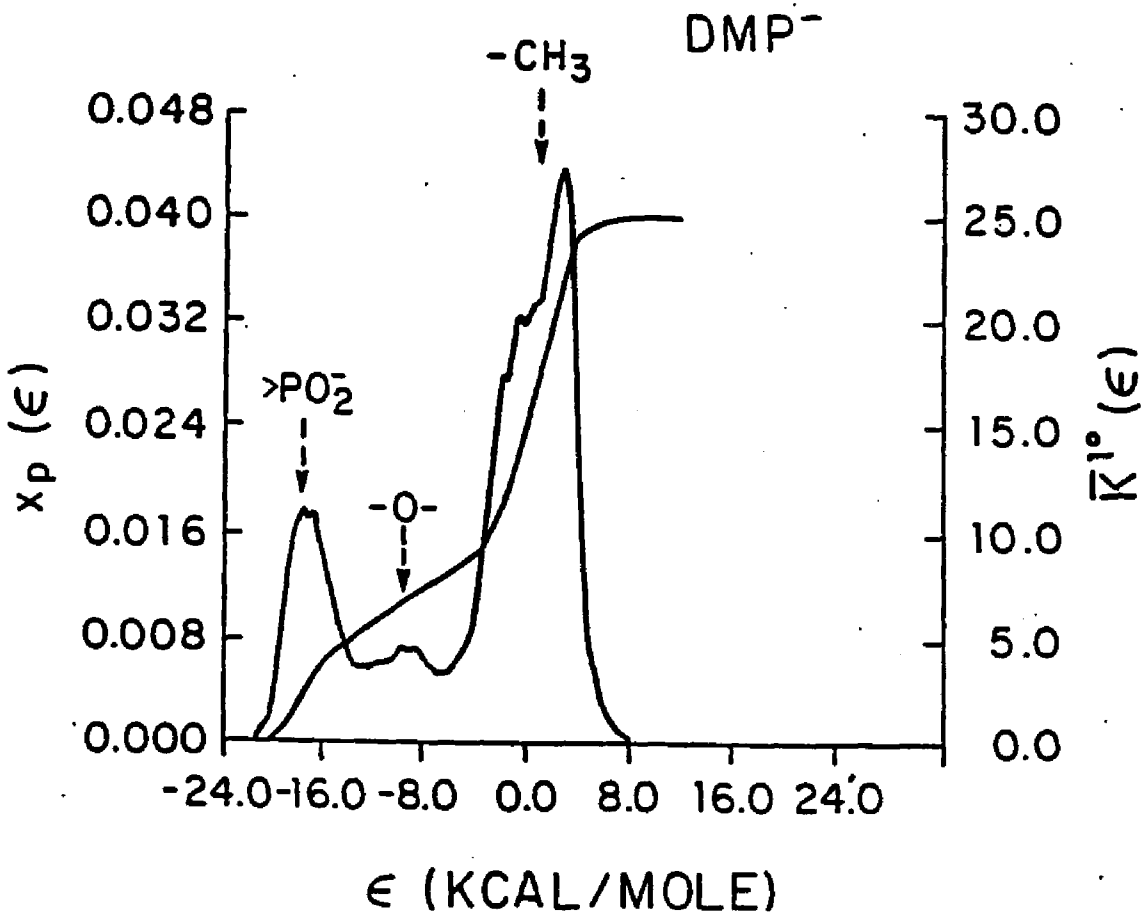

Figure 4. Calculated distribution $X_p(G)$ for solute-water pair energies versus energy G for $[DMP^-]_{aq}$.

energies for the hydrophobic hydration are positive, with the exception of tt conformer. Molecular sums for the first shell hydration are close to -100 kcal/mole for both DMP^- and Na^+DMP^- with tt form of the ion pair forming an exception.

The characteristic trends in pair interaction energies evaluated through simulations of aqueous solutions of multi-functional solutes are displayed in Figure 4. A trimodal distribution for the average near neighbor pair energies is typical of all $[DMP^-]_{aq}$ simulations, and shows the ionic (left most peak), hydrophilic (center peak) and hydrophobic hydration (right most peak) to be well resolved in this property.

The average first shell solute water pair interaction energies for the hydrophobic groups vary from -1 to +1 kcal/mole. This is consistent with the results of earlier simulations on hydrophobic hydration reported from this Laboratory and others [109-111]. The average pair energies for ester oxygens range from -8.5 to -9.4 kcal/mole for DMP^- but somewhat less (-3.9 to -6.2) for the ion pair. The average pair interaction energy for the PO_2^- group is about -15.6 kcal/mole. This is to be compared with -21.6 kcal/mole for the DMP^--Water, coming out of the contour surface calculations mentioned in section I. D. The value for sodium cation is also smaller (-16.5 to -17.5) compared

to the interaction energy of the isolated Na^+H_2O complex (-25.3 kcal/mole). The average pair interaction energies for the whole molecule vary from -3.7 to -4.3 kcal/mole.

The total solute binding energies of DMP^- are largely governed by the anionic and hydrophobic hydrations with interactions being strongly attractive for the $>PO_2^-$ group and strongly repulsive for the methyl groups. A large positive contribution to the total solute binding energy comes from the methyl group interactions with waters external to the first shell. This is either due to a disruption of water structure in the secondary shells of methyl groups caused by the ionic groups or due to inadequacies in the basis set used for computing partial atomic charges on the methyl groups or inaccuracies in the distance dependence of the methyl-water interaction potentials or all of the above. This repulsive contribution from methyl groups is larger for the ion pair. The calculated error bounds are much larger on all these atomic / functional group contributions to the total solute binding energies. The total solute-water energetics of PO_2^- group for DMP^- and $Na^+PO_2^-$ group for the ion pair are mostly dominated by the first shell contributions.

Columns 10 to 12 of Tables II to VII in general reflect water properties to be expected qualitatively

for ionic hydration. A value of -2.65 to -2.76 kcal/mole for example, for water-water average pair interactions in the proximity of sodium cation relative to the MCY water value of -3.Ø1 kcal/mole, gives an energetic criterion for ionic hydration. Solvent pair interaction energies in the vicinity of hydrophobic groups are close to the liquid water value while in the proximity of hydrophilic ester oxygens they are slightly smaller and are possibly dictated by the neighboring ionic groups. The computed error bounds for the solvent properties near the hydrophobic groups are much smaller relative to the ionic and hydrophilic groups.

Comparison with other DMP^- simulations : Finally, Table VIII summarizes some pertinent results from the other sets of simulations (columns 2^b to 4^d) on DMP^-. These are compared with the results from the main calculations (column 1^a) reported above. Energies are given in kcal/mole. Partial atomic charges (in atomic units) used in different simulations on DMP^- are collected in Table IX.

A larger cutoff (column 2^b) for water-water interactions results in a small but significant decrease in interactions originating both in water-water and solute-water terms. Solute binding energies show similar trends as in the main calculations

Table VIII. Energetics and coordination numbers for $[DMP^-]_{aq}$ from Monte Carlo simulations

$[DMP^-]_{aq}$		1[a]	2[b]	3[c]	4[d]
Total Energetics					
(g,g)	<USW>	-1859.3	-1824.6	-1898.4	-2256.8
	<US'>	-41.5	-33.8	-104.8	-142.7
	<UW'>	-1817.8	-1790.8	-1793.6	-2114.1
(g,t)	<USW>	-1864.6	-1821.9	-1870.5	-2228.9
	<US'>	-31.7	-21.2	-103.0	-147.3
	<UW'>	-1832.9	-1800.7	-1774.7	-2081.6
(t,t)	<USW>	-1854.6	-1818.8	-1895.7	---
	<US'>	-32.9	-21.9	-103.8	---
	<UW'>	-1821.7	-1796.1	-1792.7	---
First Shell Energetics					
(g,g)	<US'(FS)>	-97.1	-95.9	-116.5	-88.4
	(>PO2-)	-66.2	-77.3	-79.1	-63.8
	2(-O-)	-10.6	-19.2	-15.6	-8.7
	2(-CH3)	-12.3	+8.6	-21.8	-16.7
(g,t)	<US'(FS)>	-101.4	-95.2	-116.7	-95.8
	(>PO2-)	-76.9	-83.5	-93.2	-69.7
	2(-O-)	-12.9	-12.9	-9.3	-8.2
	2(-CH3)	-11.6	+1.2	-9.2	-17.9
(t,t)	<US'(FS)>	-102.3	-97.7	-124.1	---
	(>PO2-)	-72.7	-75.1	-78.1	---
	2(-O-)	-14.2	-12.4	-17.8	---
	2(-CH3)	-15.4	-10.2	-37.8	---
Coordination Numbers					
(g,g)	Total	22.93	25.42	25.85	22.96
	(>PO2-)	4.25	5.28	5.48	5.61
	2(-O-)	2.88	2.82	1.41	1.47
	2(-CH3)	16.68	18.28	19.84	15.88
(g,t)	Total	25.76	26.41	26.67	22.62
	(>PO2-)	4.88	5.45	5.23	6.88
	2(-O-)	1.52	1.48	1.41	1.53
	2(-CH3)	19.36	19.48	20.03	15.89
(t,t)	Total	26.24	26.26	27.51	---
	(>PO2-)	4.68	4.81	4.58	---
	2(-O-)	1.52	1.11	1.73	---
	2(-CH3)	20.04	20.34	21.19	---

a Main Calculations described in the text.
b Calculations with a larger cutoff for E(W-W)
c Calculations with identical charges and O-P-O valence angle for gg, gt & tt.
d From Ref.63

Table IX. Partial atomic charges on DMP^- used in the computer simulations

Atom	DMP^-(gg)	DMP^-(gt)	DMP^-(tt)	DMP^- [a]	DMP^- [b]
P	1.861	1.861	1.857	1.791	0.912
O1	-0.897	-0.905	-0.898	-0.878	-0.655
O2	-0.897	-0.892	-0.898	-0.878	-0.655
O3	-0.621	-0.619	-0.619	-0.631	-0.410
O4	-0.621	-0.631	-0.619	-0.631	-0.410
C1	-0.428	-0.424	-0.424	-0.211	---
H1C1	0.162	0.176	0.174	0.104	---
H2C1	0.165	0.178	0.174	0.104	---
H3C1	0.188	0.164	0.165	0.117	---
C2	-0.428	-0.426	-0.424	-0.211	---
H1C2	0.162	0.180	0.174	0.104	---
H2C2	0.165	0.175	0.174	0.104	---
H3C2	0.188	0.163	0.165	0.117	---
CH3(1)	0.088	0.095	0.089	0.114	0.109
CH3(2)	0.088	0.091	0.089	0.114	0.109
DMP^-	-1.000	-1.000	-1.000	-1.000	-1.000

a Calculations with identical charges and O-P-O valence angle for gg, gt & tt.
b From Ref. 63

reported above. The differences in the structural aspects are most pronounced for the anionic hydration in the gg form. Small changes are noticed for the extended forms. Hydration of methyl groups and ester oxygens and the total coordination numbers show conformational trends mostly matching with the main calculations.

Results on the fixed geometry and charge simulations (column 3^{c} of Table VIII) show a significant increase in the binding energies relative to the calculations previously described. The interactions are more negative by about 71 kcal/mole for the gg conformer and by ~83 and 81 kcal/mole for gt and tt conformations respectively. Given that the geometry for gg conformer is more or less similar in both studies, these differences are mainly attributable to the magnitude of charges, and point out the sensitivity of the calculated energetics to reasonable range of choices for these parameters. Optimization of the valence angle is seen to destabilize solute-water interactions over and above the charge effect, by about 10 kcal/mole. However the total internal energies of hydration favor the gg conformer relative to gt in the fixed charge and geometry calculations, while in the optimized geometry calculations the gg and gt conformers are energetically indistinguishable in consonance with the theoretical

studies on DMP^- in vacuum mentioned in the background section. Hydrophobic hydration in the first shell is slightly larger with coordination numbers per methyl group varying from 9.5 to 10.6. Hydrophobic hydration and the total coordination numbers show similar trends as in the other sets.

Column 4^d of Table VIII reproduces some pertinent results on $[DMP^-]_{aq}$ of Alagona et al.. A comparison of fixed geometry and charge calculations (column 3^c) and that of Alagona et al. shows that conformational trends match for the over all energetics and water-water interactions. However their results differ from the main calculations (column 1^a) presented above in that gt conformer was found to be more stable with regard to solute-water interactions, while the net conformational preference for gg in the total energetics in their calculations, originates in water-water terms which as they observed are more sensitive to statistical noise and convergence aspects of the runs. The total solute-water terms are more negative in their calculations. One major difference in the energy terms between the two sets arises in the water-water interactions which is -2164.4 +/- 2.3 kcal/mole for TIP4P water and -1859.75 +/-6.5 kcal/mole for MCY water for a system of 215 water molecules. A slightly smaller value for methyl group coordination is obtained in their

simulations, which is probably attributable to a smaller radial cutoff of 4.7 used for the first shell in their analysis. Also in general the united atom representation relative to a discrete representation is expected to give a smaller coordination number for the same cutoff as the accessible surface area is smaller in the former representation. That gg form should have a larger methyl coordination than gt is counter-intuitive. This is either due to differences in their analysis procedure or to the united atom representation for methyl groups used in their simulations. Over seven waters were assigned to PO_4^- group hydration, slightly more than that in the present study and this is possibly related to the potential function. The minimum in their potential function for phosphate water interaction (-16.Ø7 kcal/mole) is over 6 kcal/mole higher than that given by the potential function of Clementi and coworkers (-21.93 kcal/mole). A shallower potential energy surface appears to give a larger coordination number. Also a slightly larger cutoff of 3.2 Å was employed for the anionic oxygens in their analysis.

A visual description of the results on DMP^- (gt) cluster calculations is given in Figure 5. Conclusions from cluster calculations appear to be valid for ionic hydration and not for hydrophobic hydration, a

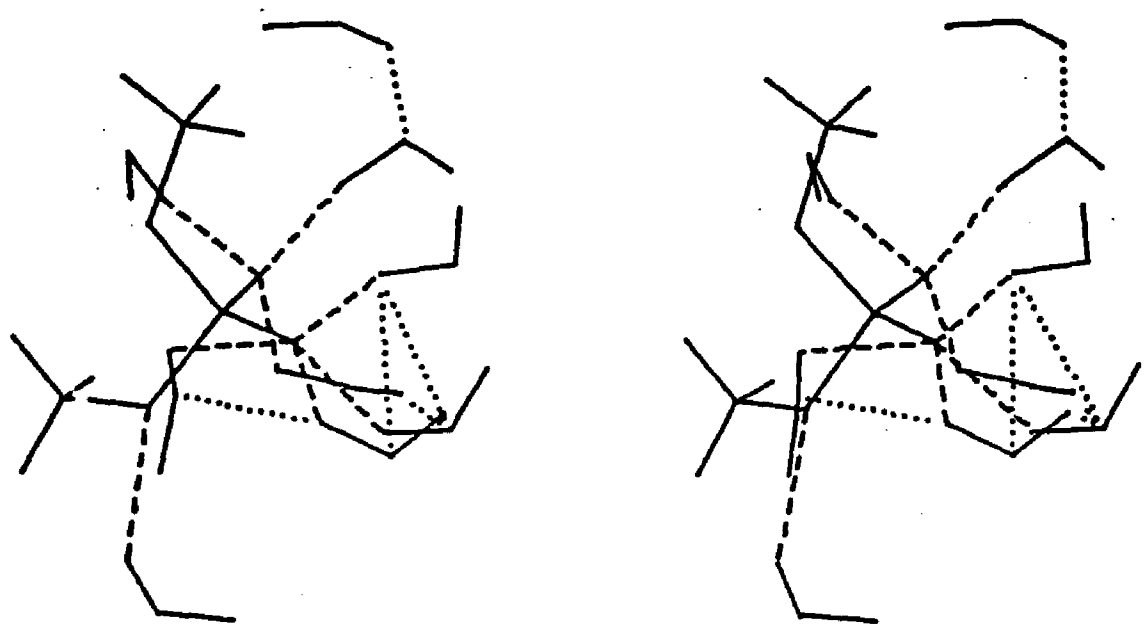

Figure 5. Representative hydration complex emerging from the simulation on DMP^{-}(g,t)-water cluster. Dashes and dots correspond to solute-water (< 3.2 Å) and water-water (< 3.3 Å) bonds.

reasonable expectation since the former is dominated by local solute-solvent interactions, whereas the latter depends on the solute volume and solvent-solvent interactions.

I. G. SYNTHESIS AND DISCUSSION

The essential structural features of the aqueous hydration of the dimethylphosphate anion as described by our simulations is a first hydration shell consisting of 23-26 water molecules, with ~six of the first shell waters identified with hydration of the phosphodiester (PO_4^-) moiety, and 17-20 with the hydrophobic hydration of the methyl groups. Conformational differences in the PO_4^- hydration show up mainly in a distribution of the six waters among the anionic and ester oxygens. Methyl groups in the gg form correspond to a contact pair (C to C distance ~ 3.6 Å) and are expected to be less solvated than in the extended conformations, and this is borne out by the simulation results. Counterion admits five additional waters into the first shell of the molecule while mostly decreasing the anionic hydration by about two waters in all conformations and of other functional groups as well to a minor extent. The hydrophobic hydration, taking into account the accessible volumes, is seen to be unperturbed by the presence of counterion.

Coordination numbers for methyl groups vary from 8-10 in all these simulations and show a high degree of transferability when these are normalized with the accessible first shell volumes. Local solvent densities

for the methyl groups in the fully extended form (tt) for instance, are 0.91 for the anion and 0.91 for the ion pair and for gt, they are 0.92 for the anion and 0.90 for the ion pair. This agrees with the results on N-methylacetamide [112] for hydrophobic hydration where methyl groups were assigned 8-9 waters and the results were observed to be independent of the potential functions. PO_4^- group hydration in the anion shows a small scatter while $Na^+PO_4^-$ group results indicate that the coordination numbers are transferable again.

Local solvent densities of ester oxygens clearly show a conformational effect, with ester oxygens in the gg form having a larger value. Figure 6 illustrates this effect. This is akin to the results reported by Pratt and Chandler (Figure 4) [113] on the hydration of butane where the interior methylene sites in the gauche form showed a larger density than the trans methylenes.

Bridged structures are a common theme in theoretical studies on monohydrate complexes as pointed out in the background section. We have looked at the statistical significance of such structures in $[DMP^-]_{aq}$ simulations. Statistical weight of configurations with a bridging water molecule between the two anionic oxygens of the phosphate group is defined as equivalent to the conditional probability for locating a water

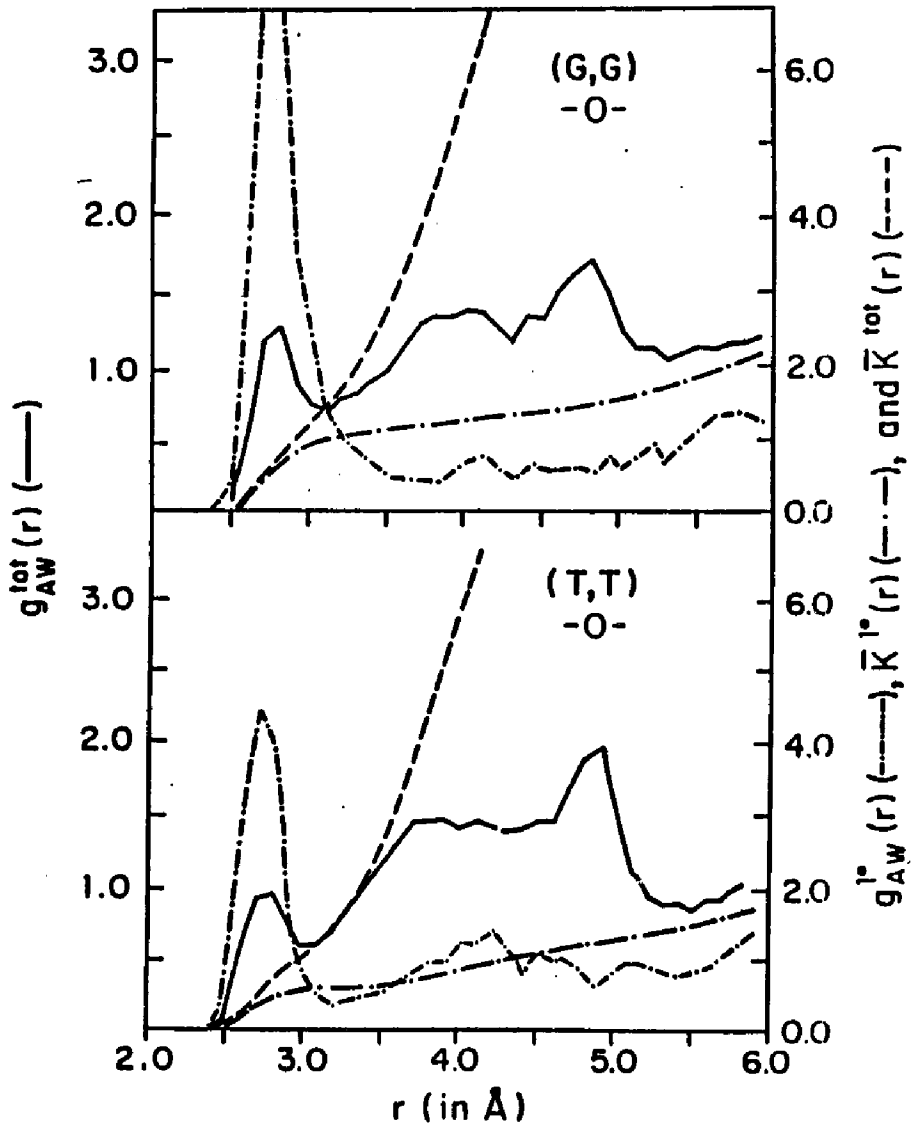

Figure 6. Primary and total radial distribution functions and running coordination numbers of gg and tt ester oxygens illustrating conformational effect.

molecule at a distance of 3.0 Å (first minimum in the radial distribution functions) from both the anionic oxygens. The value obtained for the statistical weight was 0.005, indicating a low probability for such structures in the statistical state of the systems. This observation matches with earlier studies on glycine zwitterion hydration reported from this Laboratory [114].

Representative structures illustrating the DMP^- hydration for the gg, gt and tt isomers are shown in Figures 7 to 9, respectively and for the ion pair in Figures 10 to 12. Of course, no single structure is representative of an entire simulation, but considered alongside the statistical description of the system given in the preceding section, general features can be considered. The anionic hydration of DMP^- involves primarily sequential, two-center hydrogen bonds in the structures shown. The most common type of interactions in the structures is bent hydrogen bonds. The general organization of the anionic hydration of DMP^- as determined from the simulation reflects clearly the idea of circular zones of hydration predicted in the studies of Pullman and coworkers described above [52, 55]. The hydrophilic hydration of DMP^-, essentially one water per ester oxygen shows an average separation of 2.7 Å, very near the liquid water value. The

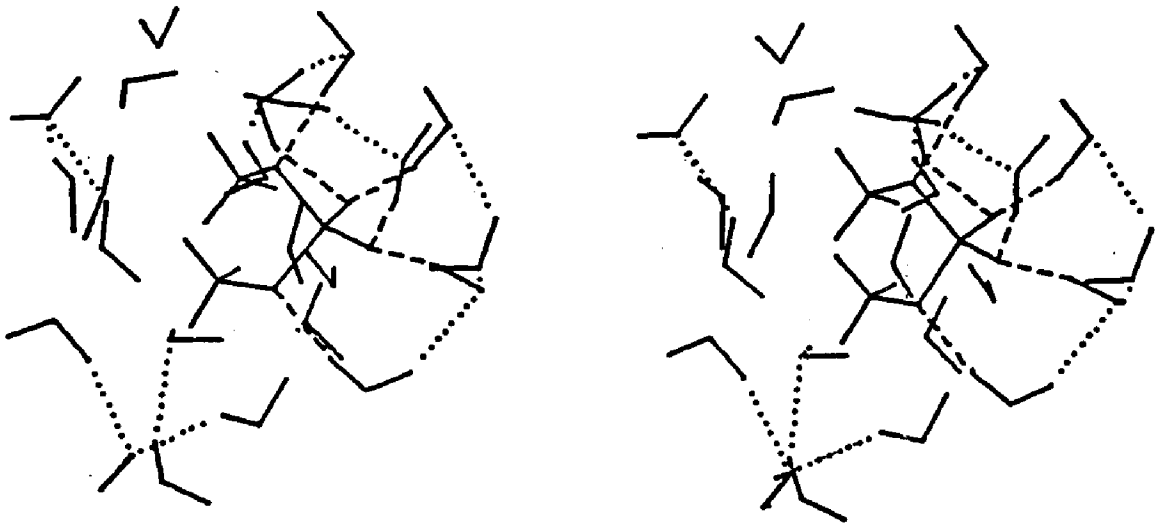

Figure 7. Stereo view of a representative hydration complex of $[DMP^{-}(g,g)]_{aq}$ sampled from the Monte Carlo run. Dashes and dots have the same significance as in Figure 5.

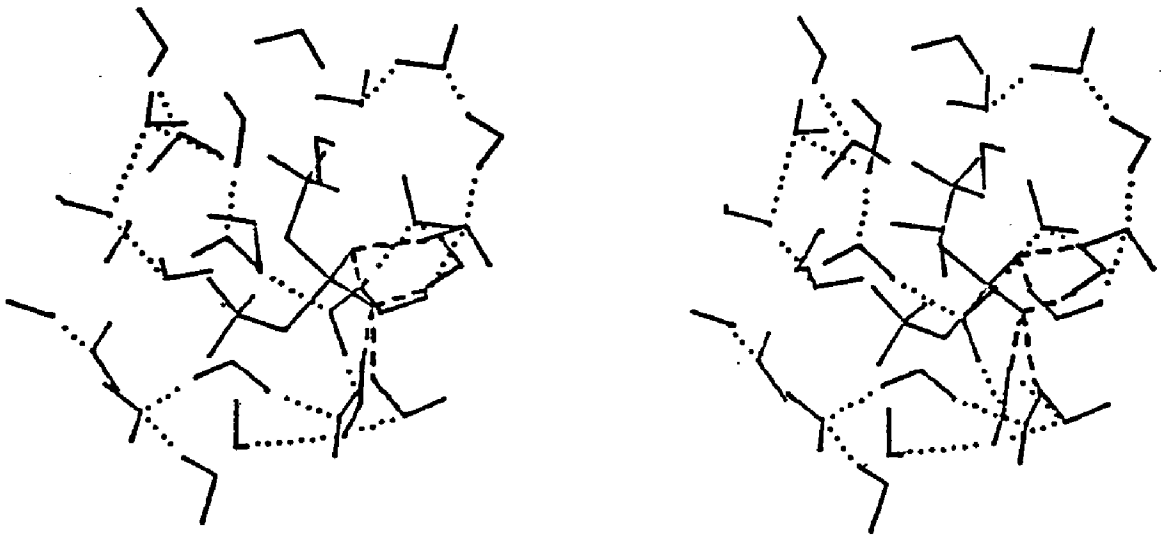

Figure 8. Stereo view of a representative hydration complex of $[DMP^{-}(g,t)]_{aq}$. Dashes and dots have the same significance as in Figure 5.

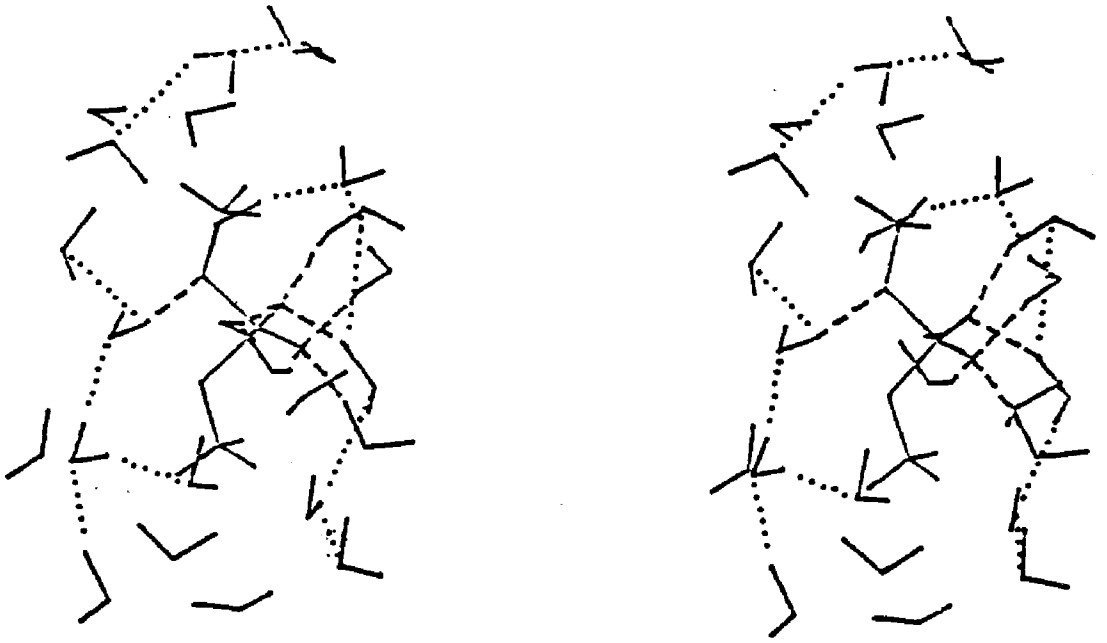

Figure 9. Stereo view of a representative hydration complex of $[DMP^{-}(t,t)]_{aq}$. Dashes and dots have the same significance as in Figure 5.

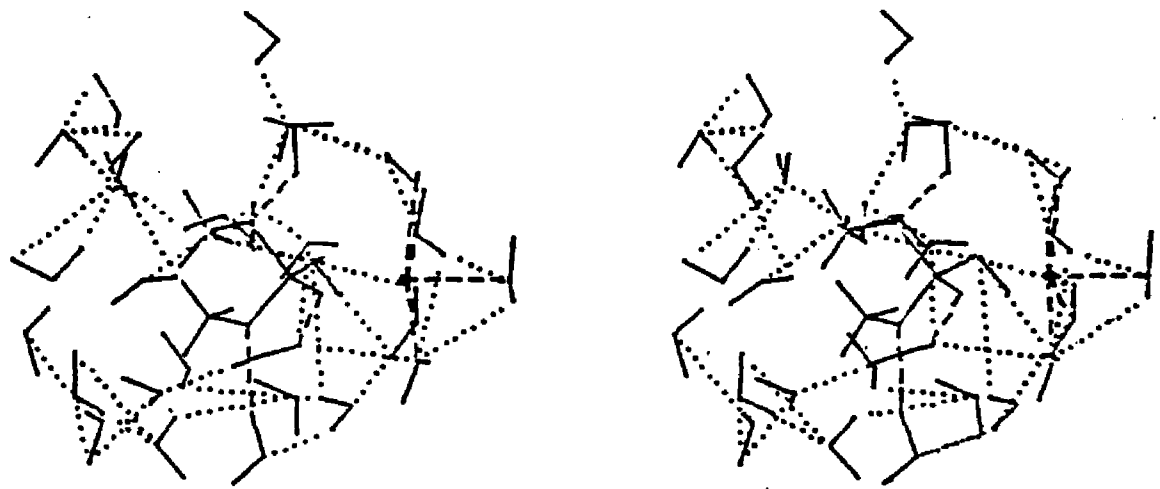

Figure 10. Stereo view of a representative hydration complex of $[Na^{+}DMP^{-}(g,g)]_{aq}$. Dashes and dots have the same significance as in Figure 5.

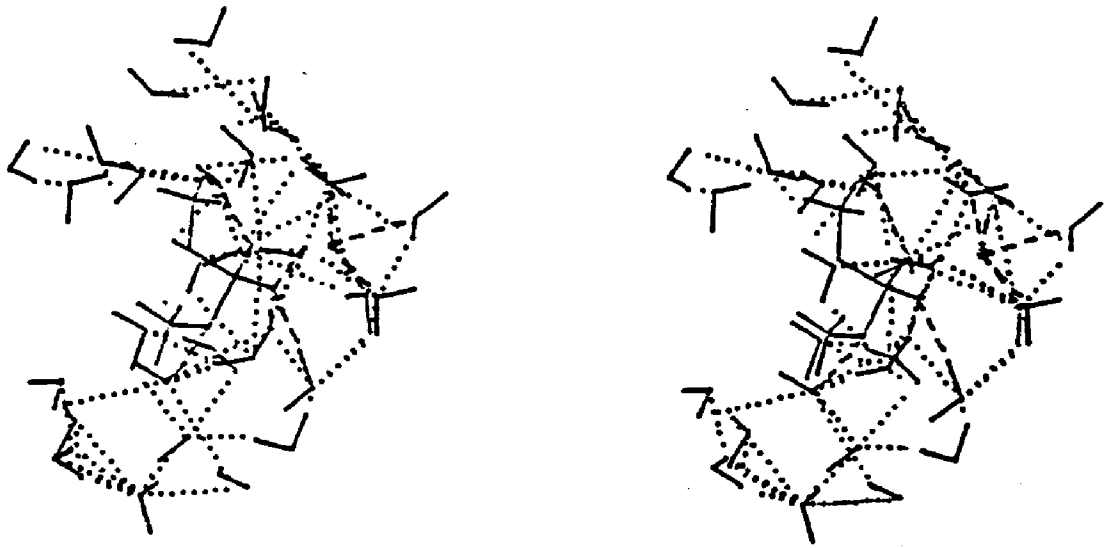

Figure 11. Stereo view of a representative hydration complex of $[Na^{+}DMP^{-}(g,t)]_{aq}$. Dashes and dots have the same significance as in Figure 5.

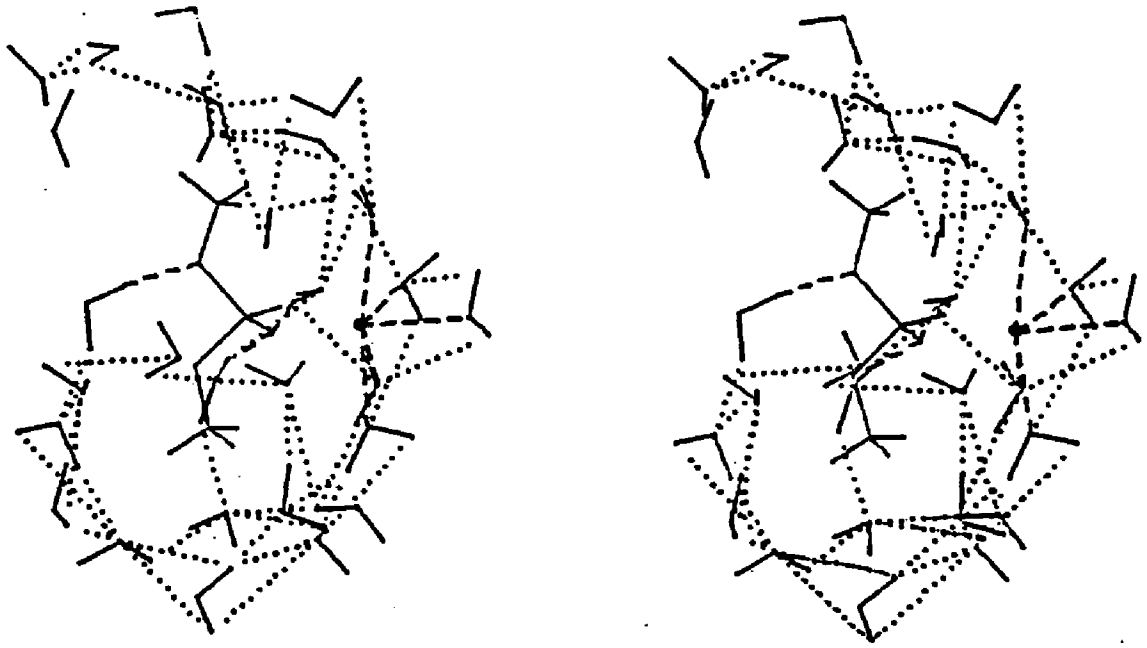

Figure 12. Stereo view of representative hydration complex of $[Na^{+}DMP^{-}(t,t)]_{aq}$. Dashes and dots have the same significance as in Figure 5.

hydrophobic hydration, quite extensive in DMP^-, is generally as expected; a structure of the hydrophobic hydration shells dominated by water-water than solute-water interactions is shown quite clearly in Figures 7 to 9. The presence of counterion is seen to increase the number of water-water hydrogen bonds in Figures 10 to 12. No water molecule is bound to both the sodium ion and the phosphate anionic oxygens simultaneously. Treating the midpoint of the anionic oxygens and the oxygens of water molecules as coordination sites, the symmetry of sodium hydration shell roughly corresponds to an octahedron in all the stereo pictures of ion pair hydration complexes. Attention here is to be drawn to the recent molecular dynamics simulation of Siebel, Singh and Kollman [115] on B-DNA in water with sodium counterions and without periodic boundary conditions, where they reported an average coordination number of 4.8 for sodium ions lying close to the phosphate anionic oxygens, in reasonable agreement with our observations.

An appraisal of the calculated energetics in the simulations must necessarily take into account two basic computational problems. Firstly, solute-solvent interactions are rather sensitive to the potential function adopted and the partial atomic charges. For ionic systems this is further compounded by the

approximations involved in periodic boundary conditions in modelling the system due to the long range interactions. Furthermore a non-negligible contribution for highly charged systems is expected to come from waters beyond the first shell which are numerous, and this leads to apprehensions about the small system size in the simulations. Secondly, solvent-solvent interactions are generally subject to considerable statistical uncertainty despite the long run lengths, more so for ionic systems. These together with concerns about intermolecular potentials leads us to only a provisional understanding of the relative stability of various conformations.

Trends in total solute-water interactions parallel conclusions from the continuum model calculations [57], with the gg conformer for the anion and tt for the ion pair favored over and above the calculated error bounds. The conformational differences in the total internal energies of hydration are not sufficiently significant for an unequivocal interpretation. However subject to the stated uncertainties, the tt conformer of the anion is seen to be destabilized by hydration relative to gg and gt, whereas it is stabilized by hydration when paired with sodium counterion.

Experimental data on the enthalpies of hydration of DMP is not available. Calculated transfer energies of

Table X. Calculated contributions to the transfer energies of DMP^-

	DMP^-(gg)	DMP^-(gg) [a]
<US>	0.5	-92.4
<Urel>	+42.0	+50.3
<US'>	-41.5	-142.7
<US'(FS)>	-97.1	-88.4
<US'(EXT)>	+55.6	-54.3
>PO2-(EXT)	+1.5	---
2(-O-)(EXT)	-3.8	---
2(-CH3)(EXT)	+56.9	---

a From Ref. 63

hydration are mostly endothermic. An analysis of the transfer energies for $DMP^-(g,g)$ is presented in Table X together with the results of (T,P,N) ensemble calculations of Alagona et al. [63]. First row of Table X gives the vacuum to water transfer energies $\langle U_S \rangle$. Second and third rows are the water-water $\langle U_{rel} \rangle$, and solute-water $\langle U_S^1 \rangle$ contributions to the transfer energies respectively. The $\langle U_{rel} \rangle$ results suggest that in both cases DMP^- disrupts the water-water interactions. A similar observation was made earlier with the halide anions and alkali metal cations [101]. The solute-water interactions however are much too small in magnitude in the present calculations relative to the results on ion-water simulations ($\langle U_S^1 \rangle \gtrsim -150$ kcal/mole) of Ref. 101 and of Alagona et al.. These are further analyzed in terms of first shell contributions $\langle U_S^1(FS) \rangle$ given in row 4 and interactions of solute with waters external to the first shell $\langle U_S^1(EXT) \rangle$ given in row 5 (and divided into the constituent functional group contributions in rows 6, 7 and 8). A contribution of +55.6 kcal/mole for solute-water binding energy from $\langle U_S^1(EXT) \rangle$, almost entirely coming from waters proximal to the methyl groups, is clearly unrealistic. Similar symptoms are observed with the gt and tt conformations of the anion and with the ion pair. Fixed charge and geometry calculations

indicate the sensitivity of the solute-water binding energies ($\langle U_S \rangle$= -104.8 kcal/mole given in column 3[c] of table VIII, as opposed to -41.5 kcal/mole in Table X discussed above) and hence the transfer energies to the choice of partial atomic charges adopted in the simulations as pointed out earlier. While a judicious choice of the solute-water and water-water potential functions is imperative to obtain acceptable transfer energies, another possible source for this discrepancy is modelling the ionic systems through periodic boundary conditions as mentioned earlier. These constitute some as yet unresolved methodological issues and further work is needed in this area. The $[DMP^-]_{aq}$ simulation results here (a perusal of Table VIII should suffice) appear to strongly suggest that structural and energetic aspects pertaining only to the first hydration shell of the solute are physically more meaningful.

A comparison of the performance of the solute-water potential functions of Clementi and coworkers employed in the present study with those of Kollman and coworkers [63], on the first shell energetics is undertaken in the following. (See columns 1[a] and 4[d] of Table VIII for details.) A total of -88.41 kcal/mole was assigned to the first shell solute-water interactions for the gg conformer in their simulations

with a division of -63.Ø3 kcal/mole for the ionic $>PO_2^-$ group, -8.72 kcal/mole for the ester oxygens and -16.66 kcal/mole for the methyl groups. This is to be contrasted with a total of -97.Ø5 kcal/mole apportioned as -66.17 kcal/mole belonging to the ionic group, -18.57 kcal/mole to the ester oxygens and -12.32 kcal/mole to the methyl groups in the simulation reported here. The average first shell solute-water pair interactions are even more revealing. They assign -11.24 kcal/mole for the ionic groups, -5.82 kcal/mole to the ester oxygens and -1.Ø5 kcal/mole for the methyl groups compared to -15.55 kcal/mole for the ionic group, -8.79 kcal/mole for the ester oxygens and -Ø.73 for the methyl groups in the present calculations. The first shell energetics of the hydrophobic hydration, taking into account the number of water molecules involved, appears to be relatively insensitive to the choice of the potential function. A smaller average first shell pair interaction energy and a larger coordination number for the ionic group in their calculations seem to be related to the choice of the potential function, while their product giving the total first shell binding energies cited above for the two potential functions agree well within the statistical uncertainties. Results on the gt conformer follow a similar pattern with the potential function of

Clementi and coworkers adopted in the present calculations turning in a slightly more negative term for the average pair interaction energies and a smaller coordination number for the PO_4^- first shell hydration compared to that of Kollman and coworkers.

Badger-Bauer rules [116] relate the red shift of the O-H stretching frequencies to the enthalpy of H-bonding formation. This empirical relation is extended to the first shell solute-water interactions obtained through the analysis of the simulation results. Lerner et al. [31] noted that the $>PO_2^-$ anti-symmetric stretching frequency (~1228 +/- 5 cm^{-1}) was insensitive to conformational changes and postulated that IR spectrum would not reflect conformational changes. Monte Carlo results on DMP are used to check this hypothesis through Badger-Bauer rule. The proportionality constant for PO_2^--H_2O frequency shifts, is obtained by calibrating the results on gg conformer. (An interaction energy of -66.168 kcal/mole is assumed to result in a frequency shift of 25 cm^{-1}.) The results on gg show that for a conformational change to gt or tt, a red shift of 2.5-4 cm^{-1} is expected for the anion and 2.3-3.8 cm^{-1} for the ion pair. A blue shift of 19.4 cm^{-1} is expected for ion pair formation (DMP^-(g,g) to Na^+DMP^-(g,g)). Monte Carlo results are thus seen to be in correspondence with the observations of Lerner et

al.. Thus the anticipated red shifts of the $>PO_2^-$ asymmetric stretch due to conformational changes are within the quoted experimental uncertainty, while solvent reorganization attendant upon ion pair formation is expected to result in a noticeable blue shift.

In another additional calculation based on the simulation results, ^{31}P spin lattice relaxation times of DMP^- in water are estimated following the methodology developed by Hertz and Raedle [117].

$$(1/T_1)_i=\{4/3\ \gamma_i^2\gamma_j^2\hbar^2S_j(S_j+1)\}x$$
$$\int\{\exp(-t/t_c)\,dt\int(g(r)/r^6)\,4\pi r^2\,dr.\}$$

The procedure essentially assumes that dipole-dipole (dd) relaxation is the dominant mechanism and involves a Markovian approximation for the relaxation process. The appropriate time correlation function [118] then has an exponential decay with rotational correlation time t_c of water as an input parameter (taken from Ref. 119). This is the well known rapid modulation limit where the solvent motions are on a much shorter time scale relative to the solute motions. The orientationally averaged spatial correlation function then involves evaluation of the ensemble average of $1/r^6$ where r is the distance between the relaxing solute nucleus and the solvent protons. This is evaluated using the radial distribution functions,

$g(r_{P-H})$, obtained from the Monte Carlo calculations. The calculated T_1 values for DMP^- are 216.2 seconds, 216.4 seconds and 217.5 seconds for gg, gt and tt conformations respectively. The experimentally estimated values [68] are close to 20 seconds. Thus either the procedure is limited in accuracy or dipole-dipole relaxation is not the dominant mechanism. Shindo [120] finds this as a usual situation met in the case of ^{31}P studies in macromolecules where chemical shift anisotropy provides an efficient alternative means, which may not be relevant for a small molecule like DMP^-. Relaxation scheme based on rapid modulation limit may not provide a valid model for strongly interacting systems such as DMP^- and water. Alternatively, with in the dd approximation one can utilize the experimental T_1 values to estimate the rotational correlation times of water in the vicinity of the solute and comment upon the strength of solute-water interactions computed via the given potential function. For DMP^- the t_c estimate is $\sim 2.5 \times 10^{-11}$ secs which is probably 3 to 4 times too large for an anion. Thus the solute-water potential function of Clementi and coworkers [53] employed here appears to over estimate the interactions, i.e. the first shell water molecules are predicted to be more tightly bound to the solute than is found experimentally.

I. H. SUMMARY AND CONCLUSIONS

To sum up, the phosphodiester group was characterized by a little over six water molecules in its first hydration shell. The total average internal energies of hydration and the relative transfer energies indicated that hydration destabilized the tt conformer of the anion, and this mostly originated in the $>PO_2^-$ group hydration. Solute-water interactions, in particular phosphate group hydration favored the gg form and water-water interactions favored the gt conformer of the anion. For the ion pair the total average internal energy of hydration was seen to stabilize the fully extended tt conformer. Convergence in water-water interactions was noted to be slow in all these simulations. The computed error bounds on the calculated energetics suggested that statistical noise was larger for the ion pair relative to the anion.

A comparison of the solute-water potential functions of Clementi and coworkers with those of Kollman and coworkers suggested that a meaningful comparison could be made only for the first shell properties of the solute wherein the results indicated a number of similarities. A larger cutoff for water-water interactions while preserving the conformational trends in coordination numbers resulted in an increase in the local solvent density of the anion in gg conformation.

Identical partial atomic charges in all conformations were observed to render the solute-water interactions conformationally indistinguishable. Also the transfer energies were noticed to be sensitive to the selection of partial atomic charges on the solute. The structural trends in hydration however, were in general agreement with other sets of simulations.

I. I. EXTENSIONS

The question of relative stability of the various conformations of DMP^- is of course a matter of free energy, and involves intramolecular energy and entropy contributions as well as intermolecular.

Determinations of the intramolecular thermodynamics of the gg, gt and tt conformations of DMP^- in free space, are described in Chapter II. A three-pronged approach (an empirical scheme, a computer simulation method and the dielectric continuum approach) was envisioned for the estimation of hydration free energies of the three conformations of DMP^-. These investigations are presented sequentially in Chapters III, IV and V.

Some relevant methodological explorations at different stages of completion are discussed in Chapter VII under suggestions for future work.

CHAPTER II

INTRAMOLECULAR THERMODYNAMICS OF DIMETHYLPHOSPHATE ANION IN FREE SPACE

II. A. INTRODUCTION

The statistical mechanical treatment of conformational analysis of macromolecules is traditionally based on the idea that conformational transitions are largely governed by the torsional degrees of freedom and vibrational modes around the torsional minima on the potential energy surface. Attempts were made in the past, with various constraints, on the solute geometry [121] to estimate the contribution of vibrational partition function to intramolecular thermodynamics of conformational changes. A method to estimate the configurational entropies through computer simulations based on Einstein's formula for the probability distribution of fluctuations was proposed by Karplus and Kushick [122], and this was further extended to compute the vibrational frequencies within the quasiharmonic approximation by Levy et al. [123]. The methodology for determining the intramolecular thermodynamics in the quasiharmonic approximation was illustrated by Ravishanker et al. [124]. Details of the calculations of internal energy and entropy contributions to the free energies of DMP^- in free space are described elsewhere [125] and are reviewed below.

II. B. THE QUASIHARMONIC APPROXIMATION

The intramolecular thermodynamics of the dimethylphosphate anion were determined by a Monte Carlo calculation on the internal degrees of freedom with entropies estimated in the quasiharmonic approximation. The configurational internal energy in the vicinity of a local minimum is computed by a numerical integration as the average value

$$U_i^{intra} = \langle E(q) \rangle \qquad q \in i$$

where E(q) is the intramolecular potential energy as a function of the q internal degrees of freedom and the brackets denote a Boltzmann configurational average in q space (c.f. section I. C.). The energy function E(q) is evaluated from an analytical force field [27].

$$\begin{aligned} E = {} & \Sigma^{-}_{bonds} K_r (r-r_{eq})^2 + \\ & \Sigma^{-}_{angles} K_\theta (\theta-\theta_{eq})^2 + \\ & \Sigma^{-}_{dihedrals} (V_n/2)[1 + \cos(n\phi - \gamma)] + \\ & \Sigma^{-}_{i<j} [(A_{ij}/R_{ij}^{12}) - (B_{ij}/R_{ij}^{6}) + (Q_iQ_j/\epsilon R_{ij})] + \\ & \Sigma^{-}_{H\ bonds} [(C_{ij}/R_{ij}^{12}) - (D_{ij}/R_{ij}^{10})] \end{aligned}$$

For entropies the procedure [122] is based on an effective potential of the form

$$E'(q) = (1/2)\ q \cdot F^{QH} \cdot q$$

with the elements of the force constant matrix F^{QH}

given by the expression

$$F_{kl} = k_B T[\sigma^{-1}]_{kl}$$

where σ is the covariance matrix of the internal coordinate fluctuations.

$$\sigma = \langle (q_k - \langle q_k \rangle)(q_l - \langle q_l \rangle) \rangle .$$

Assuming that the difference in the Jacobian evaluated at the two equilibrium conformations i and j are negligible, the quasiharmonic conformational entropy difference is given by

$$\Delta S^{int} = (1/2)\ k_B \ln(\sigma_i/\sigma_j)$$

where

$$\sigma_i = \det(\sigma) \qquad \sigma \in i$$

for displacements in the vicinity of the i th conformation. The relative internal energies and free energies follow straightforwardly from

$$\Delta U_{ij}^{int} = U_j^{int} - U_i^{int}$$

and

$$\Delta A_{ij}^{int} = \Delta U_{ij}^{int} - T \Delta S_{ij}^{int}.$$

II. C. CALCULATIONS

The initial geometries of the gg, gt and tt conformations of DMP^- are similar to those of Alagona et al. [27, 63] and identical to the values adopted by Newton [37] and are based on UpA crystal studies. Methyl groups were treated in the united atom approximation to be consistent with the force field employed [27]. The potential energy function used in the present calculations is that of Kollman and coworkers [27]. The intramolecular internal energies for the starting configurations of gg, gt conformations of DMP^- are seen to agree with the values reported in Table IV of Ref. 27. Monte Carlo simulations were conducted in cartesian space. Partial atomic charges [63] on the solute were held constant during the simulations. Since the sampling was confined to a local minimum, no gross distortions in the geometry and probably no major changes in the charges are to be expected. This besides, calculating charges at each Monte Carlo step is computationally prohibitively expensive. Also, representing the electronic charge distribution of the solute by a point charge model already involves an approximation. Full details of the calculations as carried out here are described in Ref. 124 and 125. About 2000 K Monte Carlo steps were run on each conformation.

Table XI. Calculated vibrational frequencies of DMP^- (in cm^{-1})

	tt	gt	gg	From Ref.29
$>PO_2^-$ as. str.	1213	1220	1231	1225
$>PO_2^-$ s. str.	1062	1082	1074	1086
C-O as. str.	1042	1057	1041	1055
C-O s. str.	1032	1022	1014	1040
$>PO_2$ as. str.	772	827	876	813
$>PO_2$ s. str.	747	738	707	750
Bends & Torsions	631	599	596	564
	554	560	551	523
	486	475	428	487
	431	410	408	390
	335	360	388	---
	324	326	311	260
	174	217	234	222
	122	122	162	---
	87	103	102	---

II. D. VIBRATIONAL SPECTRUM OF DMP^-

In a quasiharmonic molecular simulation, the force constant matrix can be used in a Wilson FG procedure for normal mode analysis. Results of the intramolecular thermodynamics can only be as good as the force field employed in the calculations. A check on the force field is performed by calculating the quasiharmonic force constant

$$F^{QH}/k_BT = [e-]^{-1}$$

and solving the secular equation

$$| GF^{QH} - \lambda I | = 0,$$

for the vibrational frequencies ($\lambda = 4\ \Pi\ y^2$). The calculated and the experimental [28, 29] vibrational frequencies are collected in Table XI. Results indicate reasonable agreement between theoretical predictions and experiment. Shimanouchi et al. through a normal mode analysis established the asymmetric stretchof the phosphodiester group ($>PO_2$) as a finger print zone for identifying the phosphodiester conformations. The gg conformation with C_2 symmetry was predicted to have a higher frequency for the asymmetric stretch ($>PO_2$ as. str.) relative to the symmetric stretch while the tt conformation with C_{2v} symmetry was predicted to give a higher frequency for the symmetric

stretch (>PO_2 s. str.). On the basis of this, and the Raman spectrum they concluded that the solution conformation was predominantly the gg form. The calculated frequencies here support their conclusions.

The agreement between the calculated and experimental frequencies is better in the high frequency region of the spectrum. The discrepancies in the calculated and experimental frequencies are to be traced to the quasiharmonic approximation, the force constants and other equilibrium parameters employed in the energy function. Also, the observed experimental frequencies are influenced by the presence of counterions while the present calculations do not take cognizance of the solute environment (free space approximation).

II. E. RESULTS AND DISCUSSION

The calculated thermodynamic quantities are collected in Table XII. The intramolecular free energies (ΔA^{int}) are observed to favor the gg conformation relative to gt and tt forms for the phosphodiester torsions of DMP^-. The internal energies (ΔU^{int}) are seen to dictate the trends. The calculatedtemperature weighted entropy contributions ($T \Delta S^{int}$) are small and counter the above trends.

The performance of the potential function appears to be satisfactory particularly in the high frequency region as evidenced by the agreement between the calculated and observed vibrational frequencies. The intramolecular thermodynamics is dominated by the internal energy contributions wherein gg form emerges as the most stable conformation of DMP^- in free space. The stability of the gg form in free space is generally attributed to the anomeric effects and this bias is already built into the force field through the dihedral term . The configurational averaging in the vicinity of a local minimum performed here has not altered the trends in the internal energies. The calculated entropies make no significant contribution to the conformational trends. It may be recalled that similar observations were made on configurational entropy contributions in the quasiharmonic approximation for

Table XII. Calculated intramolecular thermodynamic quantities for the gg, gt and tt conformations of DMP^- relative to tt.[a]

	tt	gt	gg
ΔU^{int}	0.0	-0.98 ± 0.07	-1.88 ± 0.13
$-T\Delta S^{int}$	0.0	0.11 ± 0.04	0.28 ± 0.28
ΔA^{int}	0.0	-0.87 ± 0.08	-1.60 ± 0.31

a) All entries are in kcal/mole. Error bounds are reported as $\pm 2\sigma$ as determined by the method of batch means.

alanine dipeptide [124]. The configurational entropies are extremely sensitive to the low frequency vibrational modes and moreover the multi-variate Gaussian distribution in the internals holds only when (a) the fluctuations in the internal coordinates are small and (b) they are statistically independent. It is uncertain if these conditions are satisfied particularly in the low frequency region. The thermal energy at room temperature corresponds to 200 cm^{-1}. Thus a classical treatment of all the vibrational degrees of freedom involves yet another approximation. Entropy changes due to diffusive modes which are probably important again in the low frequency region elude the present calculations. It must be mentioned however that the quasiharmonic approximation and the sampling scheme for the internals adopted here (limiting the sampling to a local minimum) are consistent with each other and while the objections and omissions listed above do matter to the absolute entropies, they are perhaps not very significant to the conformational differences. These are some possible areas for further investigations.

CHAPTER III

THE HYDRATION SHELL MODEL APPLIED TO DIMETHYLPHOSPHATE ANION

III. A. THE HYDRATION SHELL MODEL

Solvent effects play a significant role in biomolecular organization and hydration shell model is a simple scheme to incorporate solvent effects in exploring the free energy surfaces of biomolecules in aqueous media. Hydration shell model originally proposed by Scheraga and coworkers [58-60] and further developed by Hopfinger [12, 61] is an empirical scheme to evaluate solvation free energies and is particularly useful in studying conformational preferences of macromolecules.

The hydration free energy estimates are essentially based on P Δ V approximation for the Helmholtz free energy of hydration. A plausibility argument for the hydration shell model is presented below to make the underlying theory and approximations transparent.

$\Delta A = - P \Delta V$ (at constant temperature)

$= -(F/V) \Delta V$ (P = (F/V)= free energy density)

$= -(n f/V) \Delta V$ (n=Coordination number, f=(F/n))

$\Delta A^{hyd} = \sum_i \Delta A_i^{hyd}$ (Superposition approximation)

(ΔA_i is the free energy of hydration of atom i, and summation extends over all solute atoms/functional groups.)

$\Delta A_i = - (n_i f_i/V_i) \Delta V_i$

(Here ΔV_i is identified with the volume of exclusion

due to non-bonded neighbors.) The above equation in terms of the local solvent densities d_i can be expressed as

$\Delta A_i = - f_i d_i \Delta V_i$ (d_i is the local solvent density)

$\Delta A^{hyd} = - \sum_i f_i (n_i/V_i^H) \Delta V_i^{ex}$

where the f_i's are calibrated against experiment [59]. n_i's are evaluated from energy minimization studies [58]. V_i^H and ΔV_i^{ex} are computed analytically [60, 61].

1. The model thus assumes a linear relationship between free energy changes due to hydration and changes in accessible volumes due to the encroachment of non-bonded neighbors into the hydration shell.

2. Work done in removing n_i waters (scaled appropriately by the volume changes) from the primary hydration shell of atom "i" is related to the hydration free energy of atom "i".

3. Intra-solute spatial correlations in hydration are neglected (superposition approximation).

4. Coordination numbers or local solvent densities are assumed to be transferable.

5. Parameters are temperature independent and are valid only at 25° C.

6. Reference state from the formula is the fully stretched solute conformer with $\Delta V^{ex} = 0$, while from the interpretation of free energy as work done in

removing the waters of hydration, it is the solute in free space and from the calibration of the parameters with experiment reference state is the solute atom in a non-aqueous medium. This aspect needs a closer scrutiny.

Similar in spirit is the Gurney parameter approach of Friedman and coworkers to include solvent effects first applied to ionic solutions [126], and later extended to the study of hydrophobic interactions [127].

The interaction energy between species i and j is expressed as a sum of four terms.

$$U_{ij} = \text{Coulomb} + \text{Core} + \text{Cavity} + \text{Gurney}$$

Gurney term incorporates the solvent effects. Changes in free energy of hydration are linearly related to volumes of exclusion of the hydration spheres.

$$GUR_{ij} = (A_{ij})\ V^{ex}/\ V_w$$

Here too A_{ij}'s are calibrated against experiment and V^{ex} are computed analytically. The authors of this model point out that the errors due to the apparent inconsistency in treating the free energy term (as Gibbs free energy for the purposes of calibration while it in fact is Helmholtz free energy), are negligible. This argument applies to the hydration shell model as

well.

Ben-Naim following a different theoretical route in the context of hydrophobic interactions, finds a similar P Δ V approximation for free energies of solvation of macromolecules as justified [128]. It may be pointed out here that ΔA_i^{hyd} from the hydration shell model, as arising due to the non-bonded overlaps of hydrophobic groups is in direct correspondence with dA^{HI} used in the literature to characterize hydrophobic effect. In assesing the results of hydration shell model and similar empirical schemes, the breakdown of superposition approximation pointed out by Pratt and Chandler[129] are also to be kept in view.

Computer simulations within the scheme of hydration shell model come into play in handling the local solvent density term more accurately.

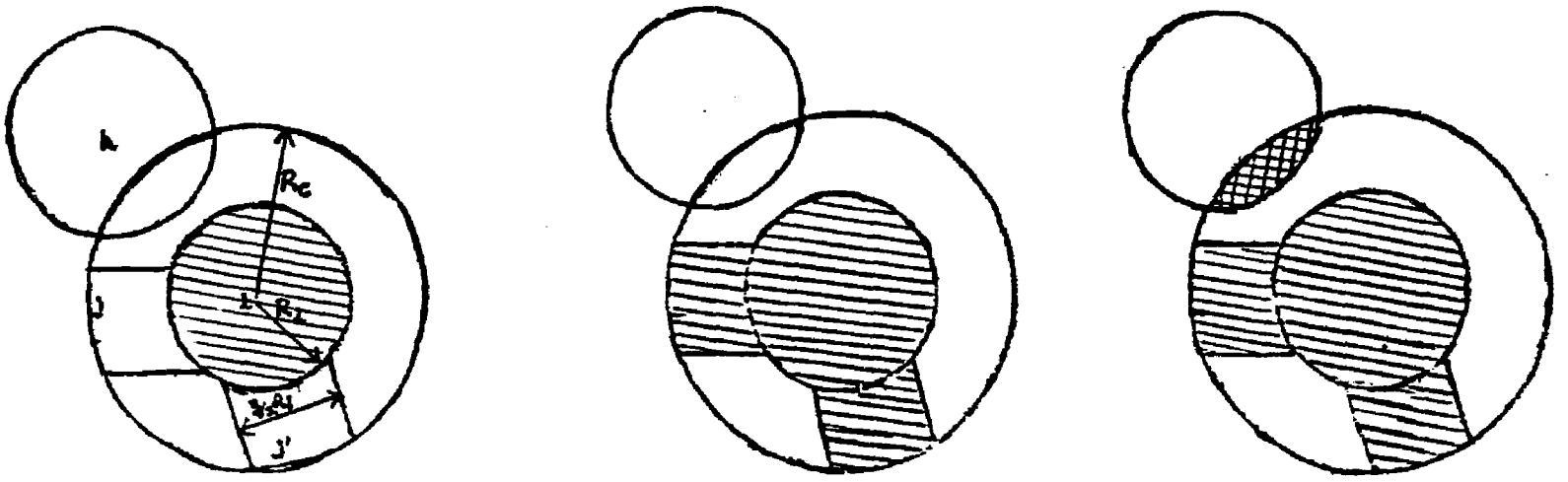

Figure 13. Illustration of the excluded volume calculations in hydration shell model. (A) R_i represents radius of the central atom i, and R_c radius of the hydration sphere of i. j and j' denote covalently bonded neighbors and k a nonbonded neighbor. (B) The shaded region is subtracted from the hydration sphere of atom i to compute maximally available hydration shell volume of atom i. $V_i^H = [(4/3)\pi(R_c^3 - R_i^3) - V_N]$ where V_N is a sum of volumes excluded because of bonded neighbors and is approximated by a cylinder of diameter $(3/2)R_i$. (C) The configuration dependent excluded volumes, shown as double hatched regions, due to atoms k, contribute to hydration free energy changes.

III. B. CALCULATIONS AND RESULTS

Hydration shell model calculations are performed on the gg, gt and tt conformations of dimethylphosphate anion (DMP^-). The excluded volumes are calculated as prescribed by Hopfinger [61] and by Hodes et al. [60]. This is illustrated in Figure 13. Contributions from specific hydration [60], (accounting for acceptor H-bonds of the anionic oxygens with water molecules), are equivalent for all the three conformers and hence cancel out in conformational differences in hydration free energies.

The geometry for DMP^- is identical to that of Newton [37] as mentioned in section II. C. Coordination numbers and radial cutoffs obtained from three separate Monte Carlo simulations on the gg, gt and tt forms of DMP^- in water, described in Chapter I (Table VIII, column 3^c), were utilized along with the free energy parameters of Hopfinger [12].

The hydration shell model results on DMP^- are collected in Table XIII. The hydration free energies are seen to favor the gg conformation relative to gt and tt forms of DMP^-. An analysis of the results on a functional group basis is reported in rows 2 to 4. Ionic hydration appears to play a major role in the destabilization of the extended tt form. Both ionic and hydrophobic hydrations favor the gg conformer.

Table XIII. Hydration free energies of DMP^- calculated through hydration shell model (kcal/mole)

	tt	gt	gg
ΔA^{hyd} ($>PO_2^-$)	0.0	-1.40	-1.77
ΔA^{hyd} 2(-O-)	0.0	0.84	1.01
ΔA^{hyd} 2(-CH3)	0.0	-0.17	-0.31
ΔA^{hyd} (DMP^-)	0.0	-0.73	-1.07

III. C. DISCUSSION AND CONCLUSIONS

The free energies of hydration emerging from the hydration shell model suggest a stability ordering of gg > gt > tt. The magnitude of the conformational differences is however small.

The conformational trends in hydration free energies predicted by the model are noticed to be sensitive to the coordination numbers adopted in the calculations. For instance the coordination number of ester oxygens is assigned as 2 [12] while in nucleic acids the hydrophilic hydration of ester oxygens seldom exceeds one water per ester oxygen. Thus if the parameter set given in Ref. 12 are used entirely, the hydration free energies relative to tt are 0.29 kcal/mole for gt and 0.59 kcal/mole for gg conformation. This trend is due to an unrealistically large emphasis given to the hydration of ester oxygens which stabilizes to a minor extent the fully extended tt form, the contributions from ionic hydration being equal for all the three conformations. Results were also seen to be sensitive to small variations in solute geometries and coordination numbers. If the coordination numbers given in column 1 of Table VIII corresponding to the optimized geometry simulations of $[DMP^-]_{aq}$ (Chapter I) are employed in the calculations, the observed stability trends were gt > gg > tt. Thus while the

model is computationally very attractive, a recalibration of the parameters for nucleic acid constituents is needed to improve the reliability of the hydration shell model conclusions.

CHAPTER IV

FREE ENERGY SIMULATIONS ON $[DMP^-]_{aq}$

IV. A. INTRODUCTION

Experimental and theoretical studies of phosphodiester torsion angles in model compounds in solution and in free space are reviewed in Chapter I and are summarized in Table XIV. The free energy of hydration of the backbone segment of nucleic acids exhibiting the phosphodiester torsional degrees of freedom is not a directly accessible quantity to experiment. All the theoretical investigations conducted hitherto focussed on the internal energy contributions to conformational differences. There is no information so far on the entropic contributions to the conformational free energies of hydration. As seen in Chapter III, empirical models while providing computationally attractive schemes to evaluate the hydration free energies directly, are fraught with uncertainties in terms of the parameters to be used and magnitudes and trends of the hydration free energies predicted. Thus theoretical studies must take recourse to alternative descriptions for reliable estimates of relative conformational free energies.

Liquid state computer simulations and Monte Carlo method in particular, have proven to be successful in evaluating structural and mechanical quantities such as radial distribution functions and internal energies. Computations of thermodynamic quantities such as

Table XIV. Summary of experimental and theoretical studies on the conformational prefernces of DMP^- (Energies in kcal/mole)

	gg	gt	tt
Giarda et al., Crystal, X-ray	x		
Shimanouchi et al., aq.soln., IR, Raman	x		
Garrigou-Lagrange et al., aq.soln., DPR			x
Gorenstein, aq.soln., ^{31}P NMR	x		
Newton, 1973, ΔU^{vac}, ab initio	0.0	3.0	7.0
Pullman et al., 1975, ΔU^{vac}, ab initio	0.0	3.4	8.0
Nanda et al., 1975, ΔU^{vac}, CNDO/2	0.0	1.5	5.0
Gorenstein et al., 1977, ΔU^{vac}, CNDO	0.0	0.1	0.9
Langlet et al., 1979, ΔU^{hyd}, Dis-Conti	0.0	-0.6	---
Alagona et al., 1985, ΔH^{hyd}, MC	0.0	28.0	---

free energy are difficult to obtain starting from a Hamiltonian model for the system, due to the special relationship between the partition function and free energy. Efforts in this direction have been extensively reviewed in a recent article from this Laboratory [130]. The problem is that high energy regions of the configurational space which are extremely important for the estimation of free energy, are either under-sampled or not sampled at all in any scheme designed to cull configurations with high Boltzmann probabilities. Non-Boltzmann sampling [87] (aimed at extracting ultimately Boltzmann averages of course) is a way to circumvent the above mentioned problem. A coupling parameter approach [131] was developed in this Laboratory along with the umbrella sampling technique [87] and the relative free energies were estimated via probability ratios. This method was applied to estimate the relative free energies of hydration of the conformations of alanine dipeptide in water [131]. The probability ratio method was subsequently extended to incorporate a self-consistent determination of the non-Boltzmann bias [132].

Free energy simulations on ionic systems are particularly problematic due to the strong and long ranged solute-water interactions. The studies reported so far are relatively few and and mostly confined to

simple spherical systems. Questions with regard to the practical applicability of the different free energy methodologies to biopolymers and their limitations such as the extent of uncertainties on the quantitative estimates, remain unanswered. An assessment of the probability ratio method, its application to aqueous solutions of nucleic acid constituents and specifically to evaluate relative conformational free energies of hydration of phosphodiester torsions in DMP^- are presented in the following. Special attention was given to the performance of the free energy simulations on the thermodynamic cycle.

IV. B. THEORY

To estimate the hydration free energy difference between any two conformations i and j ΔA_{ij}^{hyd}, through computer simulations, a coupling parameter (λ) is introduced to accomplish a smooth transition of the system from one state (λ=0) to another (λ=1) during the simulations. Thus any intermediate state k, between states i and j, along the one dimensional path defined by the coupling parameter is generated as

$$k = \lambda\, i + (1-\lambda)\, j$$

λ here is identifiable with a correlated conformational coordinate. For a conformational change this can be explicitly written as

$$\{q^k\} = \lambda\, \{q^i\} + (1-\lambda)\, \{q^j\}$$

A similar equation holds for the partial atomic charges. The free energy of hydration can be expressed as

$$A^{hyd}(\lambda) = -\, k_B\, T \ln P(\lambda) + \text{constant}$$

where $P(\lambda)$ is the unnormalized probability of occurrence of the state corresponding to the value of the coupling parameter λ. The problem then is reduced to developing $P(\lambda)$ where $\lambda \in [0, 1]$. A sequence of Monte Carlo runs with umbrella sampling with an appropriate weighting function then produce probabilities of state i and j and all other states along the path defined by the coupling parameter.

Whereas in the umbrella sampling technique the weighting function is conventionally chosen as a harmonic function of the form

$$E_w(\lambda) = c\ (\lambda - \lambda_o)^2,$$

the adaptive umbrella sampling scheme is based on the logic that the best weighting function is the free energy function itself.

$$E_w(\lambda) = k_B\ T\ \ln\ P(\lambda)$$

Since apriori $E_w(\lambda)$ is not known, an iterative approach is used where $P(\lambda)$ is first estimated on a smaller set of λ using the original potential energy function E and this estimate is used in the next step to enlarge the set of λ sampled. This refinement is continued until the adequate $E_w(\lambda)$ is found.

A ratio of the unbiased probabilities, properly matched, (matching is required to ensure that the different $P(\lambda)$ estimates along the path defined by λ have the same normalization factor), then yields relative free energies of hydration between any two states defined by the coupling parameter. The probabilities for the whole range of [0, 1] of λ in the present study are matched at the values of λ where the slopes of successive ln $P(\lambda)$ curves are equal. The relative free energy of hydration between the conformations i and j is obtained as

$$\Delta A_{ij}^{hyd} = A_j(\lambda=1) - A_i(\lambda=0) = k_B T\ \ln(P(\lambda=0)/P(\lambda=1)\ ).$$

IV. C. CALCULATIONS

Monte Carlo computer simulations in the (T,V,N) ensemble at a temperature of 25° C were performed to estimate relative free energies of hydration between (1) gg and tt and (2) gg and gt, conformations of DMP^- in water. A modified Metropolis procedure [79] incorporating force bias [89] and preferential sampling [91] was used. The geometries chosen for the solutes were those of Gorenstein et al. [44]. Partial atomic charges on each of the three conformations of DMP^- were computed using the Guassian-80 system of programs [98] with the atomic orbital basis sets of Matsuoka et al. [99]. The system in each case consisted of one DMP^- and 215 water molecules at experimental density. FCC (Face centered cubic) periodic boundary conditions were used to simulate the macroscopic system. The inter solute interactions were neglected, thus the system corresponds to one at infinite dilution. Solute-water interactions were computed using the potential functions of Clementi et al. [54], under minimum image convention [77]. Water-water interactions were modelled by the MCY representation [96], with a spherical cutoff of 7.75 Å. In the set gg to tt, the coupling parameter λ, carried the phosphodiester torsions from (60°, 60°) to (180°, 180°), the O-P-O valence angle from 103.4° to 92° and the partial atomic

charges (Table IX, Chapter I), as described in the previous section. Similarly in the set gg to gt the torsional angles were varied from (60°, 60°) to (60°, 180°), the O-P-O valence angle from 103.4° to 97.5° and along with the partial atomic charges.

Preliminary calculations using harmonic weighting functions for the non-Boltzmann bias, involved a total of six Monte Carlo simulations for gg to tt and five for gg to gt. The results suggested that gg was more stable than tt by 6.4±1.0 kcal/mole, and gt by 7.2±1.0 kcal/mole. The stability of the gg conformer relative to gt and tt predicted by these simulations was consistent with the preponderance of the gg form found in crystal studies of oligonucleotides. As a self-consistency check on the methodology and its quantitative predictions, six more Monte Carlo runs were conducted from gt to tt in an attempt to close the thermocycle. Computational details of these computer experiments are given in Table XV. As against a value of 0.8 kcal/mole in favor of the tt conformer to be expected from the previous two sets, the computed value for the free energy of hydration of gt relative to tt was -12.2 kcal/mole. The free energy cycle was off by 13 kcal. This raised serious questions about convergence of the Monte Carlo runs, the statistical noise in the estimated probabilities, the errors

XV. Characteristics of individual Monte Carlo runswith harmonic weightingfunctions

Characterization of the individual Monte Carlo runs (gt to tt)

Initial state	Final state	c	z k	zmin	zmax.	Run length
gt	tt	75.0	0.30	-0.24	0.88	1050K
gt	tt	75.0	0.45	-0.22	0.22	1150K
gt	tt	75.0	0.60	0.00	0.66	515K
gt	tt	75.0	0.75	0.54	1.10	250K
gt	tt	75.0	0.50	0.90	1.24	1120K
gt	tt	75.0	0.70	0.90	1.24	1050K
Δ A = 12.23 ± 1.00 kcal/mole						

Characterization of the individual Monte Carlo runs (gg to gt)

Initial state	Final state	c	z k	zmin	zmax	Run length
gg	gt	75.0	0.30	-0.24	0.14	450K
gg	gt	75.0	0.45	-0.06	0.70	800K
gg	gt	75.0	0.50	0.40	0.84	750K
gg	gt	75.0	0.65	0.54	1.10	1000K
gg	gt	75.0	0.85	0.96	1.24	600K
Δ A = 7.16 ± 1.00 kcal/mole						

Characterization of the individual Monte Carlo runs (gg to tt)

Initial state	Final state	c	z k	zmin	zmax	Run length
gg	tt	75.0	0.00	-0.24	0.00	600K
gg	tt	75.0	0.30	-0.10	0.20	500K
gg	tt	75.0	0.45	-0.06	0.40	600K
gg	tt	75.0	0.60	0.36	0.00	500K
gg	tt	75.0	0.85	0.74	1.12	500K
gg	tt	75.0	1.00	0.86	1.24	600K
Δ A = 6.36 ± 1.00 kcal/mole						

involved in matching and the path dependence of the computed quantities. A more rigorous convergence criterion was defined to eliminate one source of error. Runs which fluctuate around a given mean for the coupling parameter and do not show any drift during the simulation, were accepted as having converged. This provides one way of ensuring that the umbrella sampling scheme with harmonic weighting functions is properly implemented. Twenty more Monte Carlo simulations were performed meeting this criterion. Characteristics of these Monte Calo runs are given in Tables XVI and XVII and the corresponding matching curves are shown in Figure 14. The stability trend for the phosphodiester torsions in DMP^- in water was again observed to be gg > gt > tt as above. The first set from gg to tt gave a free energy of hydration of 6.1 ± 0.6 kcal/mole in favor of gg. The second set from tt to gg yielded -5.6 ± 0.6 kcal/mole indicating that each segment of the cycle is reproducible to within 1 kcal/mole. The thermocycle (gg -→ gt -→ tt -→ gg) however was still off by 8 kcal. While there is no obvious theoretical explanation for this anomalous result, the rotation of the solute during the simulations, particularly when the solute-water interactions arising in waters external to the first shell make significant contribution to solute binding energies, may lead to a very slow convergence

XVI. Characteristics of individual Monte Carlo runs for gg to tt & tt to gg simulations meeting convergence criterion

Characterization of the individual Monte Carlo runs (tt to gg)

Initial state	Final state	c	z k	zmin	zmax	Run length
tt'	gg	75.0	0.15	-0.10	0.20	500K
tt	gg	75.0	0.30	0.06	0.54	985K
tt	gg	75.0	0.45	0.26	0.78	755K
tt	gg	75.0	0.55	0.38	0.90	900K
tt	gg	75.0	0.75	0.64	1.38	750K
$\Delta A = -5.60 \pm 0.60$ kcal/mole						

Characterization of the individual Monte Carlo runs (gg to tt)

Initial state	Final state	c	z k	zmin	zmax	Run length
gg	tt	75.0	0.30	-0.10	0.28	500K
gg	tt	75.0	0.50	0.16	0.70	865K
gg	tt	75.0	0.65	0.16	0.86	745K
gg	tt	75.0	0.75	0.58	1.08	755K
gg	tt	75.0	0.80	0.58	1.24	915K
$\Delta A = 6.13 \pm 0.60$ kcal/mole						

XVII. Characteristics of individual Monte Carlo runs for gg to gt & gt to tt simulations meeting convergence criterion

Characterization of the individual Monte Carlo runs (gt to tt)

Initial state	Final state	c	z k	zmin	zmax	Run length
gt	tt	75.0	0.45	-0.22	0.22	1150K
gt	tt	200.0	0.20	-0.10	0.32	805K
gt	tt	75.0	0.30	0.06	0.54	920K
gt	tt	75.0	0.50	0.32	0.68	860K
gt	tt	300.0	0.60	0.44	0.74	1005K
gt	tt	300.0	0.75	0.66	0.92	770K
gt	tt	350.0	0.85	0.78	1.06	780K
gt	tt	75.0	0.60	0.90	1.20	600K
Δ A = 2.40 ± 1.10 kcal/mole						

Characterization of the individual Monte Carlo runs (gt to gg)

Initial state	Final state	c	z k	zmin	zmax	Run length
gt	gg	75.0	0.15	-0.24	0.14	695K
gt	gg	150.0	0.10	-0.20	0.34	1200K
gt	gg	75.0	0.10	0.06	0.64	1975K
gt	gg	75.0	0.25	0.40	0.68	1100K
gt	gg	75.0	0.50	0.48	0.84	750K
gt	gg	200.0	0.75	0.66	1.02	1045K
gt	gg	75.0	0.70	0.96	1.24	450K
Δ A = -12.42 ± 1.00 kcal/mole						

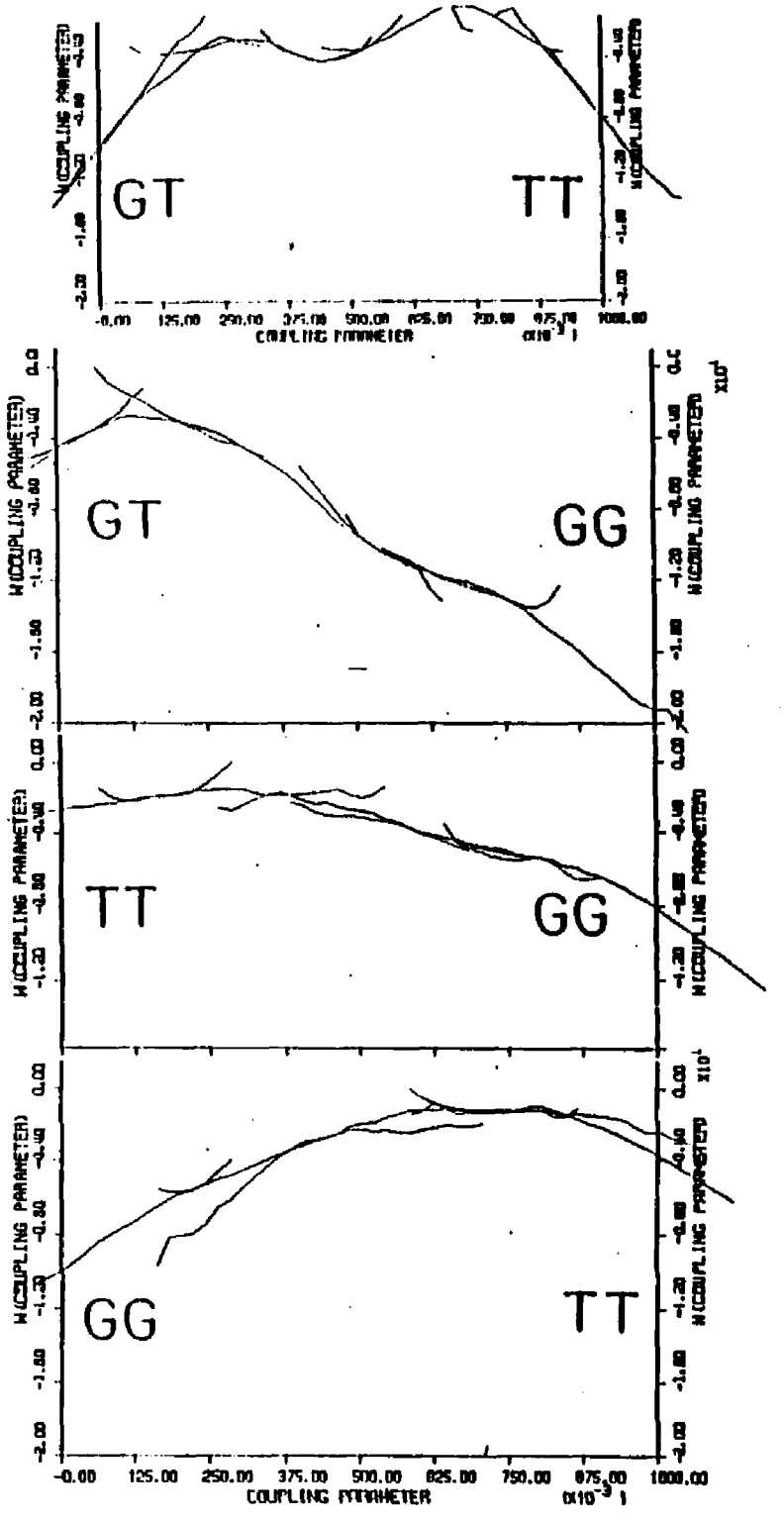

Figure 14. Matching curves for free energy simulations with harmonic weighting functions, meeting convergence criterion

of the estimated probabilities, since the "corners" of the simulation cell provide a slightly anisotropic environment. Secondly the influence of the selected harmonic weighting function and that of the initial configuration on the numerical estimates, are obseved to be nonvanishing in runs of length less than 1000K.

The hydration free energy calculations were repeated on the gg, gt and tt conformations of DMP^- with the probability ratio method using adaptive umbrella sampling technique involving a self consistent determination of the non-Boltzmann bias [132]. Solute was held rotationally fixed. Each of the two calculations (gg to tt and gg to gt) involved a total of three simulations, each of run length 2000K Monte Carlo steps. (To discourage sampling outside the targeted domains, the factor K in equation 20 of ref. 132 was set to 0.1 in the Monte Carlo runs and to encourage sampling in under sampled regions of the targeted domains the C factor in equation 19 of Ref. 132 was set equal to 0.5.) Results of these calculations are presented in the following section. The corresponding matching curves are shown in Figure 15. Calculations on a third set from gt to tt showed a double maxima for the $-kT\ln P(\lambda)$ curve and required five Monte Carlo simulations. This set together with the two above showed that the statistical

noise for the complete thermocycle simulation is ~ 0.6 kcal which is probably reasonable considering that a total of 11 simulations were involved in all. (The first law of thermodynamics stands verified after all by the free energy simulations on $[DMP^-]_{aq}$ to an accuracy of 0.6 kcal !)

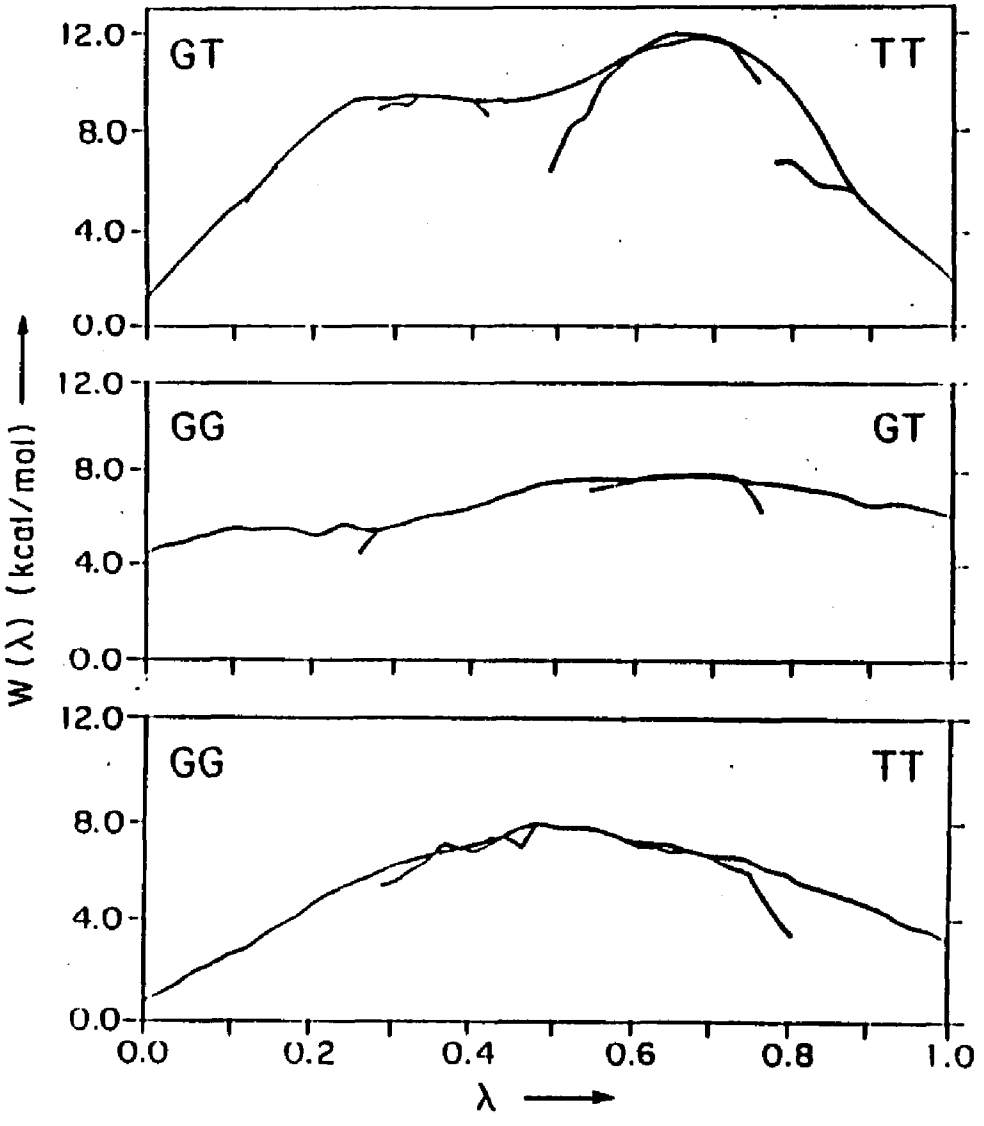

Figure 15. Matching curves for free energy simulations with adaptive umbrella sampling

IV. D. RESULTS AND DISCUSSION

The calculated free energies of hydration are collected in Table XVIII. The reported uncertainties in the estimated free energies are deduced from the differences in probability estimates of overlapping segments (matching). The stability ordering (row 1) for phosphodiester torsions in DMP^- in water is gg > gt > tt. This is consistent with the observed preponderance of gg conformer and lack of evidence for tt form in oligonucleotides and also with the IR and Raman inferences and ^{31}P NMR studies on DMP^- in aqueous solutions (section I. B. and Table XIV). Table XVIII gives a split of ΔA^{hyd} (row 1) into ΔU^{hyd} (row 2) and $T\Delta S^{hyd}$ (row 3) contributions. ΔU^{hyd} values are taken from the Monte Carlo mean energy simulations discussed earlier in Chapter I (Table I). The results show that while internal energies of hydration favor gt conformer, entropically gt turned out to be the least favored form. The stability of gg conformation for the phosphodiester torsional angles over the extended forms in free space is attributed to anomeric effects (Chapter II). The mean energy simulations on the aqueous solutions of DMP^- (Chapter I) showed that ionic hydration destabilizes the tt conformation relative to both gg and gt, while gt was favored over gg conformer. The hydration free energy simulations reported here

Table XVIII. Calculated intermolecular thermodynamic quantities for the gg, gt and tt conformations of DMP^- relative to gg. (in kcal/mole)

	gg	gt	tt
a ΔA^{hyd}	0.0	1.9 ± 0.1	2.3 ± 0.5
b ΔU^{hyd}	0.0	-5.3 ± 3.5	4.7 ± 2.3
$T\Delta S^{hyd}$	0.0	-7.2 ± 3.5	2.4 ± 2.4
c ΔA^{hyd}(HSM)	0.0	0.4	1.8
d ΔA^{hyd}(CDC)	0.0	2.6	3.3

a: The calculated relative free energies of hydration discussed in the text.

b: The mean internal energies of hydration. (Chapter I)

c: Relative free energies of hydration estimated through hydration shell model. (Chapter III)

d: Relative free energies of hydration evaluated through concentric dielectric continuum model. (Chapter V)

indicate that the relative preference of the phosphodiester torsions for gg conformation over gt form in aqueous solutions is to be traced to the $T\Delta S$ term. This is an interesting result and is supportive of the current view of the hydrophobic effect, given that the two methyl groups are closer to each other in the gg conformation. The overall picture presented by the conformational differences is that the hydration free energy surface is relatively flat.

The present study yields information for the first time on the relative conformational free energies of hydration of phosphodiester torsions in nucleic acid constituents and also highlights some of the problems to be encountered in complex systems. The series of computations here have shown that qualitative trends in relative hydration free energies evaluated through simulations are reproducible with consistency. Quantitative estimates however are more exacting. Thus caution is to be exercised in drawing inferences based on the assumption that the numerical estimates of hydration free energies of larger systems evaluated through free energy simulations honor the thermodynamic cycle. However thermocycle provides a stringent test on the methodology and an error estimate of 0.6 kcal obtained here is very encouraging.

IV. E. CONCLUSIONS

The free energy simulations reported here indicate that hydration stabilizes the gg conformation for the phosphodiester torsions in DMP^-. The anionic hydration through the solute-solvent interactions, and the hydrophobic hydration through the entropic contributions, are observed to be the major factors stabilizing the gg conformer of DMP^- in aqueous solutions. The conformational differences in free energies of hydration however suggested that the extended forms (gt and tt) are thermally accessible to some extent in aqueous solutions at ambient temperature. On the methodological front the performance of the free energy simulations is very satisfactory.

CHAPTER V

DIELECTRIC CONTINUUM MODELS FOR SOLVATION

V. A. CONCENTRIC DIELECTRIC CONTINUA

V. A. 1. BACKGROUND

The dielectric continuum as a description of environmental effects on the equilibrium properties of aqueous solutions and on the structure and conformation of biomolecules has received much attention over the last several decades due to its conceptual simplicity and computational convenience. Beveridge and Schnuelle in an earlier publication [133] described the evaluation of Helmholtz free energy of polarization of an arbitrary charge distribution imbedded in concentric dielectric continua, generalizing the models of Born [134, 135] for ion solvation and Onsager [136] for dipole solvation, within the framework of Kirkwood's [137] reaction potential [138] formalism. This can be useful for including electrostriction or bound water effects in a dielectric continuum treatment of solvent. In the following we present an extension of this treatment to incorporate ion atmosphere, thus combining the work on extended Debye - Huckel theory discussed by Tanford [139] with that in Ref. 133 to provide a unified theoretical scheme to apply to environmental effects. The theory is illustrated by investigating the conformational preferences of dimethylphosphate anion and sodium dimethylphosphate

ion pair in aqueous solutions with respect to phosphodiester torsional angles.

The continuum approach is motivated physically by a separation of spatial scales involved in the problem. If the dimensions of the solute are much larger compared to solvent and/or counterions, the environment (i.e. the solvent and counterions) may be treated as a homogeneous medium. The continuum formalism is characterised mathematically by an appropriate differential equation for the potential of interest. For the electrostatic potential, one starts with Laplace equation for treating solvent effects, and Poisson-Boltzmann equation for considering the effects due to ion atmosphere.

$$\nabla^2 \Phi = 0 \quad \text{(Laplace equation)} \tag{1}$$

$$(\nabla^2 - x^2)\, \Phi = 0 \quad \text{(Linearized P-B equation)} \tag{2}$$

Here Φ is the electrostatic potential and ∇^2 is the Laplacian operator. The ionic strength enters the theoretical treatment through x, the Debye's reciprocal length parameter. The general solution for equation (1) in polar coordinates [140], suitable for the symmetry of the problem of solute in a spherical cavity considered here, is

$$\Phi = \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \left(B_{nm}\, r^n + \frac{E_{nm}}{r^{n+1}} \right) P_n^m(\cos\,\theta)\, e^{im\phi} \tag{3}$$

and for equation (2) it is [137]

$$\bar{\phi} = \sum_{n=\emptyset}^{\infty} \sum_{m=-n}^{+n} \left(\frac{C_{nm}}{r^{n+1}} \right) e^{-xr} X_n(xr) P_n^m(\cos \theta) e^{im\phi} \quad (4)$$

where $P_n^m(\cos \theta)$ are the associated Legendre polynomials, B_{nm} and C_{nm} are constants, and E_{nm} is related to the charge distribution.

$$X_n(xr) = \sum_{s=\emptyset}^{n} \left(\frac{2^s \; n! \; (2n-s)!}{s! \; (2n)! \; (n-s)!} \right) (xr)^s \quad (5)$$

The form of equations (3) and (4) is due to an expansion of the solutions of equations (1) and (2) in terms of Legendre polynomials, which constitute a set of linearly independent functions and are particularly suitable here for the application of boundary conditions.

The problem as extended to concentric dielectric continua [133] considers three regions with solvent treated as a polarizable dielectric continuum; region A with a dielectric constant of ϵ_i is the cavity of radius a containing the solute represented as a discrete charge distribution, region B vicinal to the solute contains solvent of dielectric constant ϵ_{loc} in a spherical shell of thickness (b-a), and region C represents the bulk solvent of dielectric constant ϵ_o and extends radially from b to infinity. Thus for the

problem dealt with in Ref. 133 the potential inside the cavity is

$$\Phi_i = \epsilon_i^{-1} \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \{ (B_{nm}\, r^n + F_{nm}\, r^n + \frac{E_{nm}}{r^{n+1}}) \; (P_n^m(\cos\theta) e^{im\phi}) \} \quad (6)$$

where the $1/r^{n+1}$ terms originate in the multipolar expansion of the central charge distribution with n denoting the order of the electrical moments. E_{nm} contains the characteristics of the central charge distribution.

$$E_{nm} = \frac{(n-|m|)!}{(n+|m|)!} \sum_{k=1}^{n} q_k\, r_k^n\, P_n^m(\cos\theta_k)\, e^{-im\phi} \quad (7)$$

The terms B_{nm} and F_{nm} are due to the reaction field acting on the solute charge distribution, originating in the polarization of the local dielectric shell and the bulk dielectric continuum respectively. The reaction potential acting on the solute charge distribution is identifiable with

$$\Phi_R = \epsilon_i^{-1} \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} (B_{nm}\, r^n + F_{nm}\, r^n) P_n^m(\cos\theta) e^{im\phi} \quad (8)$$

The continuity of the potential and its first derivative across the boundaries between the three regions, were employed to specify the boundary

conditions. The coefficients B_{nm} and F_{nm} were determined in terms of E_{nm}. Helmholtz free energy of polarization was obtained as

$$A = \frac{1}{2} \sum_{k}^{-} q_k \Phi_R(r_k) \qquad (9)$$

$$= \frac{1}{2\epsilon_i} \sum_{n=0}^{\infty}{}^{-} \left\{ \left(\frac{(n+1)(1-\epsilon_a')}{(n+1)\epsilon_a' + n} \right) \frac{Q_n}{a^{2n+1}} + \left(\frac{(n+1)(1-\epsilon_b)}{(n+1)\epsilon_b + n} \right) \left(1 - \frac{n(1-\epsilon_a')}{(n+1)\epsilon_a' + n} \right) \frac{Q_n}{b^{2n+1}} \right\} \qquad (10)$$

where $Q_n = \sum_{k}^{-} \sum_{l}^{-} q_k q_l r_k^n r_l^n P_n(\cos \theta_{kl})$ (11)

and $\epsilon_a' = \epsilon_a / \left[1 + \frac{(n+1)(1-\epsilon_a)(1-\epsilon_b)}{((n+1)\epsilon_b + n)} \frac{a^{2n+1}}{b^{2n+1}} \right]$ (12)

Thus this model constitutes a step in advance of the earlier continuum models for solvation free energies in recognizing the special role of solvent close to the surface of the solute. A similar extension of Kirkwood and Tanford's treatment is presented below.

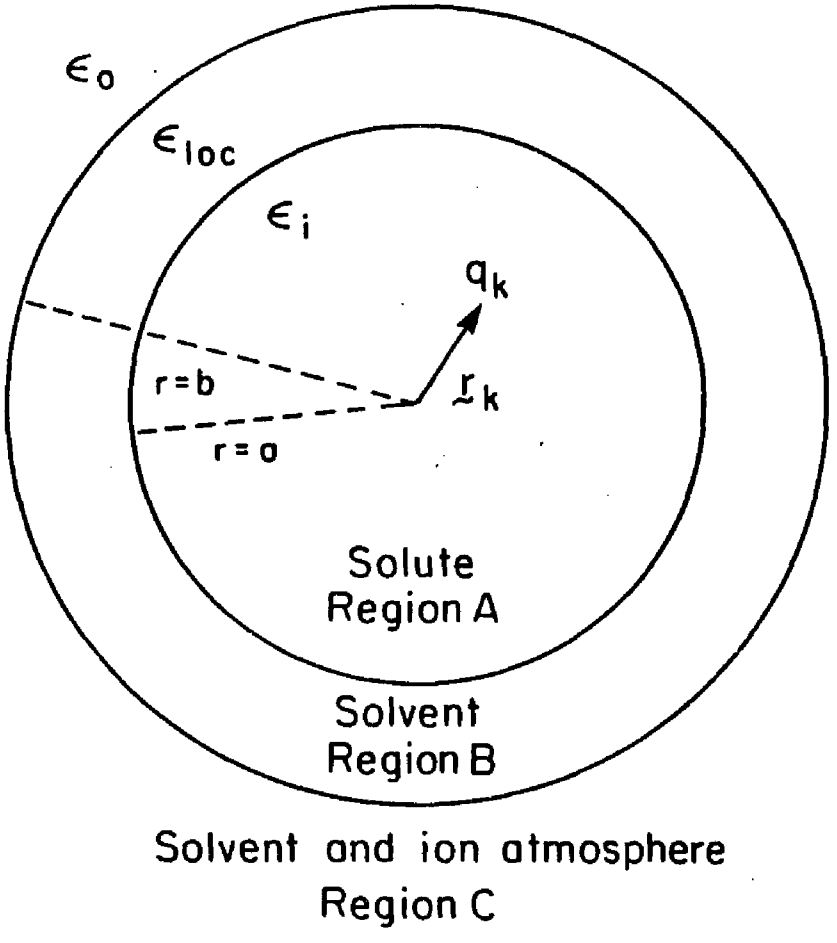

Figure 16. Definition of the parameters for the concentric dielectric continuum problem

V. A. 2. THEORY

The model considered in the present treatment consists of three regions as in Ref. 133 and depicted in Figure 16, but could of course be generalized further. Region A represents a cavity of radius r=a, and a dielectric constant of ϵ_i wherein the solute charges q_k are located at sites r_k. Region B extending from r=a to r=b represents the region of the solvent vicinal to the solute and contains solvent of dielectric constant ϵ_{loc}. Region C, ranging from r=b to r=∞ contains both the counter and co ions i.e. the ion atmosphere and the solvent of dielectric constant ϵ_o. This partitioning of the system, besides being physically realistic for spherical solutes at low ionic strengths, provides a unified scheme to combine all other models mentioned above. The objective here is to consider analytically the solute-solvent interactions as well as solute-ion atmosphere interactions at the continuum level. The appropriate equations to solve are the Laplace equation (eqn. (1)) in regions A and B and the linearized Poisson-Boltzmann equation (eqn.(2)) for region C. The notation adopted in the following is as similar as possible to that of Ref. 133.

The potential inside the cavity is given by equation (6) where now the F_{nm} terms include the reaction field due to the ion atmosphere as well. The

reaction potential at the solute charge site r_k is given by equation (8). The potential in the local region (region B) is

$$\Phi_{loc} = \epsilon_{loc}^{-1} \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \{(F_{nm}\, r^n + \frac{G_{nm}}{r^{n+1}}) (P_n^m(\cos\theta)\, e^{im\phi}) \} \quad (13)$$

The potential in the outer most region (region C) is

$$\Phi_o = \epsilon_o^{-1} \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} (\frac{C_{nm}}{r^{n+1}} e^{-xr}\, X_n(xr)) P_n^m(\cos\theta) e^{im\phi} \quad (14)$$

The boundary conditions for the model considered are

$$\Phi_{loc}(r=b) = \Phi_o(r=b), \quad (15)$$

$$\Phi_i(r=a) = \Phi_{loc}(r=a), \quad (16)$$

$$\epsilon_{loc}(d\Phi_{loc}/dr)_{r=b} = \epsilon_o(d\Phi_o/dr)_{r=b}, \quad (17)$$

and $$\epsilon_i(d\Phi_i/dr)_{r=a} = \epsilon_{loc}(d\Phi_{loc}/dr)_{r=a}. \quad (18)$$

Equations (13), (14) and (15) lead to

$$\epsilon_{loc}^{-1} (F_{nm} b^n + \frac{G_{nm}}{b^{n+1}}) = \epsilon_o^{-1} (\frac{C_{nm}}{b^{n+1}} e^{-xb} X_n(xb))$$

$$\text{or} \quad \epsilon_b (F_{nm}\, b^n + \frac{G_{nm}}{b^{n+1}}) = C_{nm}\, Y_b \quad (19)$$

where $\epsilon_b = \epsilon_o/\epsilon_{loc}$ (20)

and $$Y(r) = e^{-xr} X_n(xr)/r^{n+1};$$

$$Y_b = (Y(r))_{r=b};\ Y_b' = (dY(r)/dr)_{r=b} \qquad (21)$$

Equations (13), (14) and (17) lead to

$$n F_{nm} b^{n-1} - (n+1) \frac{G_{nm}}{b^{n+2}} = C_{nm} Y_b' \qquad (22)$$

Eliminating C_{nm} from equations (19) and (22)

$$F_{nm} = \left(\frac{(n+1)Y_b + \epsilon_b b Y_b'}{n Y_b - \epsilon_b b Y_b'}\right) \frac{G_{nm}}{b^{2n+1}} \qquad (23)$$

Similarly equations (6) and (8) together with (16) give

$$\epsilon_i^{-1} \left(B_{nm} a^n + F_{nm} a^n + \frac{E_{nm}}{a^{n+1}}\right) = \epsilon_{loc}^{-1}\left(F_{nm} a^n + \frac{G_{nm}}{a^{n+1}}\right)$$

$$\text{or } \epsilon_a \left(B_{nm} a^n + F_{nm} a^n + \frac{E_{nm}}{a^{n+1}}\right) = \left(F_{nm} a^n + \frac{G_{nm}}{a^{n+1}}\right) \qquad (24)$$

$$\text{where } \epsilon_a = \epsilon_{loc}/\epsilon_i \qquad (25)$$

and equations (6), (8) and (18) lead to

$$n B_{nm} a^{n-1} - (n+1) \frac{E_{nm}}{a^{n+2}} = -(n+1) \frac{G_{nm}}{a^{n+2}} \qquad (26)$$

Eliminating F_{nm} using (23) and (24) gives

$$G_{nm} = \epsilon_a' \left(B_{nm} a^{2n+1} + E_{nm}\right) \qquad (27)$$

$$\text{where} \quad \epsilon_a' = \epsilon_a / \left[1 + \left(\frac{a^{2n+1}}{b^{2n+1}}\right)(1-\epsilon_a)\left(\frac{(n+1)Y_b + \epsilon_b\, b\, Y_b'}{nY_b - \epsilon_b\, b\, Y_b'}\right)\right] \tag{28}$$

Combining (26) and (27) gives

$$B_{nm} = \left(\frac{(n+1)(1-\epsilon_a')}{(n+1)\epsilon_a' + n}\right)\frac{E_{nm}}{a^{2n+1}} \tag{29}$$

and from (27) and (29)

$$G_{nm} = \left(1 - \frac{n(1-\epsilon_a')}{(n+1)\epsilon_a' + n}\right) E_{nm} \tag{30}$$

and from (23) and (30)

$$F_{nm} = \left(\frac{(n+1)\ Y_b + \epsilon_b\, b\, Y_b'}{n\ Y_b - \epsilon_b\, b\, Y_b'}\right)\left(1 - \frac{n\ (1-\epsilon_a')}{(n+1)\epsilon_a' + n\ b^{2n+1}}\right)\frac{E_{nm}}{a^{2n+1}} \tag{31}$$

The reaction potential is

$$\Phi_R = \epsilon_i^{-1} \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \left\{ \left(\frac{(n+1)(1-\epsilon_a')}{(n+1)\epsilon_a' + n}\right)\frac{r^n}{a^{2n+1}} + \left(\frac{(n+1)Y_b + \epsilon_b\, b\, Y_b'}{nY_b - \epsilon_b\, b\, Y_b'}\right)\left(1 - \frac{n(1-\epsilon_a')}{(n+1)\epsilon_a' + n}\right)\frac{r^n}{b^{2n+1}}\right\} E_{nm} P_n^m(\cos\theta)\, e^{im\phi} \tag{32}$$

The electrostatic free energy is

$$A = \frac{1}{2\epsilon_i} \sum_{n=0}^{\infty} \left\{ \left(\frac{(n+1)(1-\epsilon_a')}{(n+1)\epsilon_a' + n}\right)\frac{Q_n}{a^{2n+1}} + \left(\frac{(n+1)Y_b + \epsilon_b\, b\, Y_b'}{nY_b - \epsilon_b\, b\, Y_b'}\right)\left(1 - \frac{n(1-\epsilon_a')}{(n+1)\epsilon_a' + n}\right)\frac{Q_n}{b^{2n+1}} \right\} \tag{33}$$

The above results transform into those of Ref. 133, in particular equation (33) reduces to (10) in the limit of vanishing ionic strength.

For $x=0$,

$$Y_b = 1/b^{n+1} \quad \text{and} \quad Y'_b = -(n+1)/b^{n+2} \tag{34}$$

Essentially the terms F_{nm} and ϵ'_a are redefined from those in Ref. 133 to include ion atmosphere through the Debye's reciprocal length parameter x.

For $n=0$, equation (33) reduces to

$$A_o = \frac{1}{2\epsilon_i}\left(\frac{(1-\epsilon'_a)}{\epsilon'_a}\frac{Q_o}{a} + \frac{1-\epsilon_b(1+xb)}{\epsilon_b(1+xb)}\frac{Q_o}{b}\right) \tag{35}$$

Substituting for ϵ'_a and setting $\epsilon_b=1$ results in

$$A_o = (A_{Born} + A_{DH}) = \frac{Q_o}{2\epsilon_i\,\epsilon_a\,a}\left((1-\epsilon_a) - \frac{xa}{(1+xb)}\right) \tag{36}$$

In comparing equation (36) with the Debye-Huckel result (see for instance equation 26-30 of Ref. 139) it must be borne in mind that the reference state here is the discrete charge assembly in free space as in Born's model for ion solvation. By letting $x=0$ and $\epsilon_i=1$ in the above equation, Born's charging energy is recovered.

$$A_o = \frac{Q_o}{2a}\left(\frac{1-\epsilon_a}{\epsilon_a}\right) \tag{37}$$

Similarly for n=1, $Q_1 = \mu^2$ and

$$A_1 = \frac{1}{2\epsilon_i}\left(\frac{2(1-\epsilon'_a)}{(2\epsilon'_a+1)}\frac{\mu^2}{a^3} + \left(\frac{2Y_b+\epsilon_b\, b\, Y'_b}{Y_b-\epsilon_b\, b\, Y'_b}\right)\left(1-\frac{(1-\epsilon'_a)}{(2\epsilon'_a+1)}\right)\frac{\mu^2}{b^3}\right) \quad (38)$$

$$\text{where } \epsilon'_a = \epsilon_a / \left[1 + \frac{a^3}{b^3}(1-\epsilon_a)\left(\frac{2Y_b+b\,\epsilon_b\, Y'_b}{Y_b-b\,\epsilon_b\, Y'_b}\right)\right] \quad (39)$$

$$Y_b = (e^{-xb}/b^2)(1+xb) \quad (40)$$

$$Y'_b = (-e^{-xb}/b^3)(2+2xb+x^2b^2) \quad (41)$$

Now by setting $\epsilon_b=1$, $\epsilon_i=1$ and $x=0$ in eqn. (38),

$$A_1 = \frac{(1-\epsilon_a)}{(1+2\epsilon_a)}\frac{\mu^2}{a^3} \quad (42)$$

This is the familiar Onsager's dipolar contribution to the solvation free energy.

V. A. 3. CALCULATIONS

Model calculations are performed on dimethylphosphate anion (DMP^-) and sodium dimethylphosphate ion pair (Na^+DMP^-), in the gauche-gauche (gg), gauche-trans (gt), and trans-trans (tt) conformations as systems of previous interest (Chapters I to IV). Literature on the conformational preferences of phosphodiester torsion angles, and on the hydration of DMP^- and Na^+DMP^-, is summarized in section I.B. and in Table XIV. Equation (33) in the limiting case was seen to reproduce the hydration free energies of monoatomic ions (Table 2.9 of Ref. 135), and here we consider a polyatomic conformationally flexible molecular ion. All calculations were performed at a temperature of 298 K and varied ionic strengths. Center of the first non-vanishing moment is chosen as the origin for the solute. The inner radius r=a was taken to be 3.2 Å and the outer radius r=b as 6.0 Å. These values are based on the solute-water radial distribution functions calculated from the Monte Carlo simulations described in chapter I. The inner radius as estimated from the partial molar volumes was ~ 2.87 Å but was observed to lead to divergences in the computed free energies since some solute charges lie outside the inner sphere. The choice for the outer radius also agrees with the value

of $b=a+2R_w$, where R_w is the radius of the water molecule. The geometry and charge distribution for DMP^- were identical to those of Alagona et al. [63], where the charges were chosen as fits to the quantum mechanically calculated electrostatic potentials and reproduce the moments of the electronic charge distribution well. The sodium ion for the ion pair was placed in the anionic POO^- plane on the bisector of the OPO angle at a distance of 2.21 Å from the anionic oxygens, with a unit positive charge, following Liebmann et al. [73]. The inner dielectric constant ϵ_i was taken to be 1. A choice close to 2 for the local dielectric constant (ϵ_{loc}) used in conjunction with the values of outer radii obtained from Monte Carlo simulations [101] was observed to reproduce the experimental free energies of hydration [141] of alkali metal cations and halide anions. Some of the estimated ϵ_{loc} values are 1.58 for Na^+, 1.72 for K^+, 2.25 for Cl^-, 2.16 for Br^- and 1.82 for I^- (b is taken as $a+2R_w$ for Br^- and I^-). This prompted a choice of 2 for ϵ_{loc} for the present study. A value of 4 for the local dielectric constant has some precedence in the work of Kollman et al. [142]. However, calculations were also performed with ϵ_{loc}=10, and 78.3 [135]. Different concentrations were employed to study the effect of ionic strength on the conformational preferences.

V. A. 4. RESULTS AND DISCUSSION

Results on the hydration free energies of the anion DMP^- are collected in Table XIX as a function of the number of terms considered in the multipole series expansion in equation (33). Conformational trends are seen to converge much faster than the absolute magnitudes. The calculations predict an ordering of gg > gt > tt for the relative conformational preferences of phosphodiester torsions in DMP^- in aqueous solutions. Table XX summarizes the influence of explicit counterion and the effects of variations in local dielectric constant on the conformational stabilities. Local dielectric constant provides an indirect handle on the ionic strength in the vicinity of the solute. The conformational differences are much smaller for the ion pair than for the anion, particularly for gg and gt conformations. A smaller value for the local dielectric constant brings the conformations free energetically even closer. Indeed, for the ion pair if the dipolar contributions (n=2, and ϵ_{loc}=2) alone are considered, the hydration free energies (in kcal/mole) are -11.57 for gg, -13.40 for gt and -15.21 for tt suggesting a trend of tt > gt > gg for the ion pair in concurrence with the studies of Bleha, Mlynek and Tvaroska [57] on H^+DMP^- and with the Monte Carlo studies (Chapter I). Changes in ionic

Table XIX. Convergence of hydration free energies of DMP^- calculated through concentric dielectric continuum model [a]

n[b]	gg	gt	tt
1	-39.43	-39.43	-39.43
2	-39.43	-39.43	-39.43
3	-39.85	-39.93	-40.27
4	-40.28	-40.45	-40.52
5	-40.71	-40.77	-40.71
10	-42.00	-41.36	-41.00
20	-43.38	-41.66	-41.05
30	-43.98	-41.70	-41.05
40	-44.24	-41.71	-41.05
45	-44.31	-41.71	-41.05

a. Energies are in kcal/mole, Temperature= 298 K, Ionic strength=0.0, a = 3.2 Å, b = 6.0 Å, ϵ_i=1.0, ϵ_{loc}=2.0, and ϵ_o=78.30.

b. Number of terms considered in the multipole expansion.

Table XX. Calculated hydration free energies of DMP^- and Na^+DMP^- shown as a function of local dielectric constant

	ϵ_{loc}	gg	gt	tt
DMP^-	78.30	-64.79	-57.32	-55.37
Na^+DMP^-	78.30	-58.20	-55.97	-49.85
DMP^-	10.00	-60.62	-54.32	-52.68
Na^+DMP^-	10.00	-50.87	-49.09	-44.03
DMP^-	4.00	-54.07	-49.41	-48.21
Na^+DMP^-	4.00	-39.99	-38.82	-35.25
DMP^-	2.00	-44.31	-41.71	-41.05
Na^+DMP^-	2.00	-25.28	-24.85	-23.07

Energies are in kcal/mole.

Table XXI. Environmental free energies of DMP^- and Na^+DMP^-

	C	gg	gt	tt
DMP^-	0.00	-44.31	-41.71	-41.05
Na^+DMP^-	0.00	-25.28	-24.85	-23.07
DMP^-	0.01	-44.37	-41.77	-41.11
Na^+DMP^-	0.01	-25.29	-24.85	-23.07
DMP^-	0.10	-44.45	-41.85	-41.18
Na^+DMP^-	0.10	-25.30	-24.86	-23.08
DMP^-	1.00	-44.55	-41.95	-41.28
Na^+DMP^-	1.00	-25.33	-24.89	-23.12

a. Ionic strengths are reported as concentration C (in molarities) of 1:1 electrolyte.

b. Energies are in kcal/mole.

strength in the outer region on the hydration free energies as presented in Table XXI areseen to have little or no influence on the conformational trends.

An attractive feature of the present model is its generality. It combines solvent effects and ion atmosphere and is particularly suited for incorporating a vicinal solvent zone and changes in the ionic strength. The applicability of the present model to biopolymers is to be judged in terms of the overall symmetry of the charge distribution, specifically the shape parameter $((\text{Area}^{1/2})/(\text{Volume}^{1/3}))$. For other shapes, a similar extension of the spheroidal model [143] is possible. Soumpasis [144] discussed the influence of ion atmosphere on a periodic charge distribution with cylindrical symmetry and this has been extended (section V. B.) further to discrete charge distributions incorporating a vicinal solvent zone. Size of the molecule needs to satisfy two mutually conflicting criteria. The solute molecule must be big enough for the bulk environment, and in particular the vicinal solvent to be treated as dielectric continua. The charges in the interior of the cavity must not be too close to the boundary, for the spherical symmetry of the potential assumed and in particular for the multipolar expansion of the potential to be valid [145].

Investigations on the conformational preferences of phosphodiester torsional angles in dimethylphosphate anion indicated that gauche-gauche conformation is preferred in aqueous solutions. This is in accord with the experimentally observed preponderance of the gg conformations for the diverse phosphodiester torsions (Chapter I). The conformational differences in the hydration free energies are smaller for the ion pair relative to the anion as expected. Variations in the ionic strength in the outer region are not energetically significant probably due to the large distance (6 Å) from the origin. Changes in the local dielectric constant have pronounced effect on the magnitudes of the hydration free energies but not on the relative conformational preferences.

V. B. Free Energy of a Discrete Charge Distribution in Coaxial Cylindrical Dielectric Continua

V. B. 1. BACKGROUND

Cylindrical charge distribution is the focus of constant research activity, primarily due to its relevance to biological and synthetic polyelectrolytes and especially to DNA. The stress in this area in recent years, is mostly on the condensation phenomenon and on the radial distribution of the counterions around the charged cylinder representing the polyion. The solvent effects are neglected. In the previous section (V. A.), a generalized treatment of solute-solvent and solute-ion atmosphere interactions at the continuum level for an arbitrary charge distribution embedded in concentric dielectric continua was presented. In the following a theoretical study of an arbitrary charge distribution enclosed in a coaxial cylindrical system containing solvent and counterions is undertaken.

In order to understand the polyelectrolyte behavior in solution at the continuum level the nonlinear Poisson-Boltzmann (PB) equation has to be solved. This however is not possible analytically. Thus either the nonlinear PB equation has to be integrated numerically, or Debye-Huckel's linearization has to be extended to

the cylindrical problem as well to obtain analytical solutions. There is an extensive literature on the polyelectrolyte theory and on the application of the PB equation to the cylindrical problem. Some of the earlier work in this area was reviewed by Morawetz [146], Tanford [139] and recently by Anderson and Record [147]. A critique on the application of the PB equation was advanced by Rice and Nagasawa [148].

A recollection of the spatial scales involved in the problem is pertinent here. The first is the average distance of axial charge separation projected along the backbone of the polyelectrolyte. This is not a such well defined quantity to be considered concentrated on a select few atoms along the backbone as any Mulliken population analysis would show. Thus 1.7 Å for native DNA for instance involves a coarse description of the charge separation and may be satisfactory for large distances from the helical axis. Second is the diameter of solvent molecule which for water is ~ 2.8 Å. Treating water as a continuum even in the vicinity of the polyionic charges involves a gross approximation. Third is the size of the mobile ion, isolated / solvated, and again the validity of the PB equation rests on the negligibility of the size. Fourth is the Debye's length, x^{-1}, which is inversely related to the square root of the ionic strength. Fifth is analogous

to Bjerrum's length for the polyion and may be defined as $Q_i\ Q_j/2\epsilon kT$, where Q_i and Q_j are any two neighboring charges along the backbone of the polyelectrolyte. The dimensionless linear charge density parameter Z often encountered in the literature on polyelectrolytes is defined as twice the ratio of two lengths, the fifth and the first.

Onsager observed [149] that if the charge density on a line charge exceeded a certain critical value corresponding to $Z > 1$ then statistical mechanics predicted that the system was unstable. This gave rise to the condensation theory of Manning.

Manning's, "Modern polyelectrolyte theory" [150], proposed that the residual charge on the cylindrical system was $1/NZ$ where N was the valency of the counterion and Z, a dimensionless structural parameter, related to the axial charge density of the polyion. A theoretical justification for this postulate was found by minimizing the total free energy of the system with contributions from the screened repulsions between charges on the linear array representing the polyion, and mixing of free cations, bound cations and solvent molecules. Thus the partial condensation of the counterions on the linear array was to be understood as a compromise between the repulsions which would favor complete binding and mixing free energy which would

lead to a complete dissociation. For simple electrolytes, it was pointed out, complete disssociation upon dilution was entropy guided and free energy changes due to dilution always remained finite. In contrast a complete dissociation of the counterions from the cylindrical polyion was shown to correspond to a physically unrealistic situation. Thus $Z \geq 1$ for cylindrical polyelectrolytes was found to be an instability condition. Linearized PB equation was asserted to be a good approximation for all values of $Z < 1$.

Dolar and Peterlin [151] conducted a numerical integration of the PB equation for a rod like model for polyelectrolyte solutions with mono and divalent counterions. Philip and Wooding [152] obtained a solution of the nonlinear PB equation about a cylindrical particle by a joining procedure for the sinh Φ function. Lawrence and Conway [153] studied the field and electrostriction at a linear array of charges. Katchalsky [154] formulated a cell model with the PB equation to evaluate the colligative properties of polyelecrolyte solutions and concluded, in consonance with Manning's thoery, that condensation of the counterions was expected if the charging parameter Z exceeded unity.

MacGillivray [155] studied the behavior of the PB

equation by estimating upper and lower bounds on the solutions, in the limit of infinite dilution. He showed that a clustering of counterions occurred about the cylindrical polyion when viewed from Debye length scale and that condensation assumption was not necessary for the case of $Z > 1$. The hypothesis that the uncondensed mobile ions might be treated in the DH approximation was found to be justified.

The electrostatic potential of a discretely charged cylinder in solution was studied by Bailey [156]. Sugai and Nitta [157] estimated the surface electric potential on a rodlike polyelectrolyte with periodic axial charge distribution and concluded that a nonuniform potential lowered the surface potential. The degree of the drop of the average surface potential due to non-uniformity was noted to be significant as the surface charge density decreased. Gregor and Gregor [158] evaluated the surface potentials of cylindrical polyelectrolytes with counterions of different sizes.

Stigter [159] gave an analysis of the ion condensation based on an interaction integral for the polyelectrolyte cylinder and the counterion and proposed a hierarchy of theoretical models with Manning's as the first in the series. If DH theory was assumed to be valid in the limit of infinite dilution,

the interaction integral diverged in the limit of xa=0; i.e., in the case of a line charge in a salt solution (a=radius of the cylinder=0, x = the reciprocal Debye length > 0) and also for a charged cylinder in a salt free solution (a>0 and x=0). Counterion condensation was shown to be a postulate to avoid this divergence. Stigter further pointed out that the logarithmic surface potential also diverged in the limit of infinite dilution.

Soumpasis [144] derived Coulomb and Debye-Huckel screened potentials for four different models of periodic charge distributions: a linear lattice, a cylindrical lattice, a helical lattice and a plectonemic double-helical lattice and found that colligative properties at infinite dilution were adequately described by a simple continuous line of charge. The range of validity of the linearized PB equation was reported to be confined to Z < 1.

Fixman [160] examined the validity of a generalized PB equation based on functional expansion techniques incorporating noncoulombic interactions into the potential. He noted that the PB equation was a good approximation for polyelectrolyte solutions upto a bulk concentration of 0.1 M.

Klein, Anderson and Record [161] carried out a numerical integration of the PB equation, extracted

colligative properties from a cell model and compared them with experiment with reasonable agreement. They also concluded that their results conformed over a range of salt concentrations to the limiting laws deduced by Manning.

Gueron and Weisbuch [162] discussed their results on the application of PB equation to different shapes in terms of counterion accumulation as opposed to Manning's counterion condensation. They found that their "CIV" (concentration of counterions in the immediate vicinity of the polyelectrolyte) was shape dependent only for a very low salt concentration (< 1mM) and that no importance could be attached to 1/NZ parameter.

Zimm and LeBret [163] studied the solutions of the PB equation at infinite dilution for a sphere, a cylinder and a plane, in a bid to assess the dependence of counterion condensation on system dimensionality. A spherical geometry for the system led to a total dissociation of the counterions, while a planar surface resulted in a double layer formation and total condensation. Cylindrical systems were observed to behave as a unique intermediate case with partial neutralization of the surface charge.

Ramanathan [164] investigated the properties of the PB equation for cylindrical polyelectrolytes in dilute

ionic solutions and proved that if the charge density on the polyion exceeded a critical value, the surrounding ionic solution separated into two phases. The inner phase predominantly consisted of counterions in verification of Onsager's observation and Manning's condensation theory.

Monte Carlo studies of Zimm and LeBret [165] demonstrated that the size of the mobile counterion was an important factor in determining the accuracy of the PB equation around a highly charged cylinder such as DNA. The PB equation led to very good results if the radius of the mobile ions was less than one tenth the radius of the cylinder. In correspondence with the condensation theory a striking accumulation of counterions in a layer of concentration greater than 1 M was noticed at the surface of the polyion even when the concentration of the counterions in the bulk was less than 0.01 M. A continuum versus a discrete charge representation for the polyion in their simulations was concluded to give similar results.

Pack and Klein [166] conducted a numerical integration of the PB equation to evaluate the radial distribution of counterions around B and Z forms of DNA and observed a higher concentration of counterions at the immediate surface of the Z conformer than in the vicinity of the B form, consistent with experimental

data on the B to Z transitions at higher ionic strengths.

It is thus apparent from the literature surveyed above that some aspects yet to be considered with in the frame work of cylindrical symmetry, are the solvation free energy of a charge distribution (this is not feasible for a polyion as a zero ionic strength term would result in the divergence of the logarithmic potential) and the influence of ion atmosphere on the solvation free energy of a neutral molecule with an approximate cylindrical symmetry. Another issue worthy of examination would be whether the range of validity of the extended Debye-Huckel theory would improve if the level of description of the solute was made finer and if regions with different dielectric constants were included taking into account both solute-solvent and solute-ion atmosphere interactions.

The motivation for the present work is two fold. One is to evaluate electrostatic free energy of a discrete charge distribution embedded in coaxial cylindrical dielectric continua, generalizing the existing models at a mean field level, to further an understanding of the conformational preferences of biopolymers in aqueous solutions. The second is to obtain theoretical expressions for the reaction potential to be used in conjunction with computer simulations on biopolymers,

accomplishing a reduction in the system to be represented and in computational time, while effectively incorporating the environmental effects.

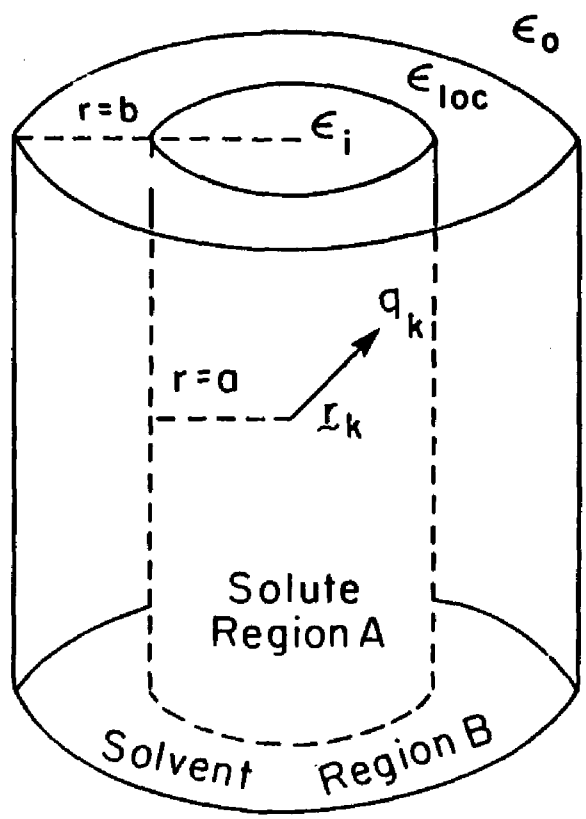

Figure 17. Definition of the parameters for the coaxial cylindrical dielectric continuum problem

V. B. 2. THEORY

The model considered in the present work, depicted in Figure 17, is so chosen as to be general enough to accomodate many of the existing approaches in the literature cited, besides being per se useful. The coaxial cylindrical system consists of three regions. Region A, the inner region, represents a cylindrical cavity of radius r=a, and a dielectric constant of ϵ_i wherein the solute charges q_k are located at sites r_k. Region B, the local region, extending radially from r=a to r=b represents the "first shell" of the solute and contains solvent of dielectric constant ϵ_{loc}. Region C, the outer region, ranging from r=b to r=co contains both solvent of dielectric constant ϵ_o and ion atmosphere, i. e. counterions and coions. The ion atmosphere enters the theoretical treatment through the Debye's inverse length parameter x which is proportional to the square root of ionic strength. The problem now is to solve the Laplace equation

$$\nabla^2 \bar{\phi} = 0 \qquad (1)$$

for regions A and B, and the linearized Poisson-Boltzmann equation

$$(\nabla^2 - x^2)\bar{\phi} = 0 \qquad (2)$$

for region C, where $\bar{\phi}$ is the electrostatic potential

and ∇^2 is the Laplacian operator. General solutions for regions A and B are

$$\Phi_{i \text{ and } loc} = A + \sum_{k=1}^{p} \frac{B_k}{|r-r_k|} \tag{3}$$

and for region C

$$\Phi_o = \sum_{k=1}^{p} \left(\frac{A_k \, e^{-x|r-r_k|}}{|r-r_k|} + \frac{B_k \, e^{x|r-r_k|}}{|r-r_k|} \right) \tag{4}$$

Φ_i and Φ_{loc} contain the coulomb potentials due to the discrete charge distribution and Φ_o contains the screened coulomb (Debye) potential. These general solutions are to be transformed into a coordinate system appropriate for the symmetry of the problem for the application of boundary conditions. Bessel functions are convenient to work with for cylindrical shapes. A discussion on the Laplace equation in cylindrical coordinates and the form of general solutions is given by Jackson [140]. An overview of the Bessel functions is given by Hildebrand [167] and Abramowitz and Stegun [168]. Soumpasis [144] gave an expansion for the Debye potential in Bessel functions.

The general solutions for the regions A, B and C are

$$\Phi_i = \epsilon_i^{-1} \sum_{k=1}^{p} \left(-\frac{2Q_k}{L} \ln w_k + B_k + F_k + \sum_{n=1}^{\infty} \left(B_{nk} \, I_0(l_n w_k) + F_{nk} \, I_0(l_n w_k) + \frac{4Q_k}{L} K_0(l_n w_k) \right) \cos l_n(z-z_k) \right) \tag{5}$$

$$\Phi_{loc}=\epsilon_{loc}^{-1}\sum_{k=1}^{p}(G_k \ln w_k + F_k + \sum_{n=1}^{\infty}(F_{nk} \; I_\emptyset(l_n w_k) + G_{nk} \; K_\emptyset(l_n w_k)) \cos l_n(z-z_k)) \quad (6)$$

$$\Phi_o=\epsilon_o^{-1}\sum_{k=1}^{p}(C_k \; K_\emptyset(x w_k)+ \sum_{n=1}^{\infty} C_{nk} \; K_\emptyset(l_n' w_k) \cos l_n(z-z_k)) \quad (7)$$

In the above equations, the functions $K_\emptyset(x')$ and $I_\emptyset(x')$ are the modified Bessel functions of zeroth order and $K_1(x')$ and $I_1(x')$ are of first order. The terms in B_{nk} originate in the polarization of the local dielectric. Similarly the terms in F_{nk} arise in the polarization of the bulk dielectric continuum. The constants B_k, B_{nk}, F_k, F_{nk}, G_k, G_{nk}, C_k and C_{nk} are to be evaluated by an application of the boundary conditions.

The reaction potential is identifiable with

$$\Phi_R=\epsilon_i^{-1}\sum_{k=1}^{p}(B_k+F_k+\sum_{n=1}^{\infty}(B_{nk}+F_{nk})I_\emptyset(l_n w_k)\cos l_n(z-z_k)) \quad (8)$$

The electrostatic Helmholtz free energy of the charge distribution is given by

$$A_{el} = \sum_{v=1}^{p} \frac{1}{2} Q_v \; \Phi_R(r_v) \quad (9)$$

where $\Phi_R(r_v)$ is the reaction potential at the charged site r_v.

The matching conditions for the problem are that the potential and its normal component across the

boundaries be continuous. This imposes on Φ_i, Φ_{loc} and Φ_o the following conditions.

$$(\Phi_i)_{\varrho=a} = (\Phi_{loc})_{\varrho=a} \tag{10}$$

$$(\Phi_{loc})_{\varrho=b} = (\Phi_o)_{\varrho=b} \tag{11}$$

$$\epsilon_i \left(\frac{d\Phi_i}{d\varrho}\right)_{\varrho=a} = \epsilon_{loc} \left(\frac{d\Phi_{loc}}{d\varrho}\right)_{\varrho=a} \tag{12}$$

$$\epsilon_{loc} \left(\frac{d\Phi_{loc}}{d\varrho}\right)_{\varrho=b} = \epsilon_o \left(\frac{d\Phi_o}{d\varrho}\right)_{\varrho=b} \tag{13}$$

Using these equations the constants B_k, F_k, B_{nk} and F_{nk} in the reaction potential expression, equation (8), are evaluated. These are

$$B_k = \frac{2Q_k}{L\epsilon_b} \frac{(1-\epsilon_a)}{\epsilon_a} \left(\frac{K_\emptyset(xw_k^b)}{xw_k^b K_1(xw_k^b)} + \epsilon_b \ln\left(\frac{w_k^b}{w_k^a}\right)\right) \tag{14}$$

$$F_k = \frac{2Q_k}{L\epsilon_b} \left(\frac{K_\emptyset(xw_k^b)}{xw_k^b K_1(xw_k^b)} + \epsilon_b \ln w_k^b\right) \tag{15}$$

$$B_{nk} = \frac{4Q_k}{L} (1-\epsilon_a') \frac{K_1(y_a)}{I_1(y_a)} \tag{16}$$

$$F_{nk} = \frac{4Q_k}{L} \epsilon_a' \Theta \tag{17}$$

where

$$\epsilon_a = \epsilon_{loc}/\epsilon_i \ , \ \epsilon_b = \epsilon_o/\epsilon_{loc}, \tag{18}$$

$$w_k^2 = (\varrho^2 + \varrho_k^2 - 2\varrho\varrho_k \cos(\phi-\phi_k)) \tag{19}$$

$$w_k^{a\,2} = (a^2 + \varrho_k^2 - 2a\varrho_k \cos(\phi-\phi_k)) \tag{20}$$

$$w_k^{b\,2} = (b^2 + \varrho_k^2 - 2b\varrho_k \cos(\phi-\phi_k)) \tag{21}$$

$$l_n = 2\pi n/L, \quad l_n' = (l_n^2+x^2)^{1/2} \tag{22}$$

$$y_a = \ln w_k^a,\ y_b = \ln w_k^b,\ y_b' = \ln' w_k^b \tag{23}$$

$$\theta = \left(\frac{\ln K_\emptyset(y_b')\ K_1(y_b) - \epsilon_b \ln' K_\emptyset(y_b)\ K_1(y_b')}{\ln K_\emptyset(y_b')\ I_1(y_b) + \epsilon_b \ln' I_\emptyset(y_b)\ K_1(y_b')}\right) \tag{24}$$

and

$$\epsilon_a' = -\frac{\epsilon_a}{y_a}\left[K_\emptyset(y_a)I_1(y_a) + \epsilon_a I_\emptyset(y_a)K_1(y_a) + (1-\epsilon_a)\theta I_\emptyset(y_a)I_1(y_a)\right]^{-1} \tag{25}$$

On substitution of the above expressions in equation (8) the reaction potential is given as

$$\Phi_R = \sum_{k=1}^{p} \frac{2Q_k}{L\epsilon_o}\left[\left(\frac{K_\emptyset(xw_k^b)}{xw_k^b K_1(xw_k^b)} + \frac{\epsilon_o}{\epsilon_{loc}} \ln\left(\frac{w_k^b}{w_k^a}\right) + \frac{\epsilon_o}{\epsilon_i} \ln w_k^a\right) + \frac{2\epsilon_o}{\epsilon_i} \sum_{n=1}^{\infty}\left((1-\epsilon_a')\frac{K_1(y_a)}{I_1(y_a)} + \epsilon_a'\theta\right) I_\emptyset(\ln w_k)\cos \ln(z-z_k)\right] \tag{26}$$

The electrostatic free energy is obtained as

$$A_{el} = \sum_{v=1}^{p}\sum_{k=1}^{p} \frac{Q_v Q_k}{L\epsilon_o}\left[(B'_k + F'_k) + \frac{2\epsilon_o}{\epsilon_i}\sum_{n=1}^{\infty}\{(B_{nk} + F_{nk})\ I_\emptyset(\ln w'_k)\cos\ln(z_v - z_k)\}\right] \tag{27}$$

where the constants are defined by equations (14) to (17) and by the following.

$B'_k = B_k\epsilon_a/(2Q_k/L)$

$F'_k = F_k\epsilon_a/(2Q_k/L)$

$B'_{nk} = B_{nk}/(4Q_k/L)$

$F'_{nk} = F_{nk}/(4Q_k/L)$

and $w_k' = w_k$, all evaluated at r_y.

Case (1): In the absence of counterions, the Helmholtz free energy is given by setting x to zero in equation (27) and using the asymptotic properties of modified Bessel functions.

For $x \ll 1$ $I_m(x) \sim (x^m/(2^m\, m!))$,

$K_m(x) \sim (2^{-m}\, (m-1)!\, x^{-m})$

$K_\emptyset(x) \sim -\log x$

For $x=\emptyset$, $K_\emptyset(x)$ diverges and it is also seen that both B_k' and F_k' in equation contain the $K_\emptyset(x)$ terms. Thus if the ionic strength is set equal to zero the Helmholtz free energy of polarization diverges. This result is in verification of the behavior of cylindrical charge systems surveyed in the background section (V. B. 1).

Case (2): Consider the linear lattice with lattice axis coiniciding with the z-axis. A further simplification is achieved by assuming that the magnitude of all charges along the lattice axis is equal.

$$Q_v = Q_k = Q$$

Then $\Sigma_{v=1}^{-p} Q_v Q_k / L\epsilon_o$ is replaced by $pQ^2/L\epsilon_o$

Also $\rho_k = \rho_v = \emptyset$, $\phi_k = \phi_v = \emptyset$ and $w_k^a = a$, $w_k^b = b$.

With these substitutions the electrostatic free energy is given as

$$A_{el}' = (p^2Q^2/L\epsilon_o)[B_\emptyset' + F_\emptyset' + (2\epsilon_o/\epsilon_i)\Sigma_{n=1}^{-\infty}(B_n' + F_n')]$$

where

$$B_\emptyset' = (1-\epsilon_a)[K_\emptyset(xb)/xb\ K_1(xb)] + \epsilon_b \ln (b/a)$$

$$F_\emptyset' = \epsilon_a[K_\emptyset(xb)/xb\ K_1(xb)] + \epsilon_a \epsilon_b \ln b$$

$$B_n' = (1-\epsilon_a')\ K_1(y_a)/\ I_1(y_a)$$

$$F_n' = \epsilon_a'\ \Theta$$

with y_a, ϵ_a' and Θ given by equations (23) to (25) after the above substitutions.

Case (3): For the case of linear lattice of Soumpasis [144] there is no solvent layer between ϱ = a and ϱ = b. Set a= b, $\epsilon_{loc} = \epsilon_i$ (i. e. $\epsilon_a = 1$) and the rest as in Case (2) above.

$$A_{el}'' = (p^2Q^2/L\epsilon_o)[F_\emptyset'' + (2\epsilon_o/\epsilon_i)\Sigma_{-n=1}^{-\infty} (B_n'' + F_n'')]$$

where $F_\emptyset'' = [(K_\emptyset(xb)/ xb\, K_1(xb)) + (\epsilon_o/\epsilon_i)\ \ln b)]$ with B_n'' and F_n'' identical to B_n' and F_n' of Case (2) respectively with a = b and $\epsilon_a = 1$

By noting a property of the modified Bessel functions $I_n(x)\, K_{n+1}(x) + I_{n+1}(x)\, K_n(x) = 1/x$ and using this in equation (25) it is seen that $\epsilon_a' = \epsilon_a = 1$ and $B_n'' = \emptyset$ $F_n'' = \Theta$ (defined by equation (24)).

These results agree with the conclusions in section III of Ref. 144.

Case (4): Hill's model [139] is obtained by taking the continuum limit (L --> ∞, with Q/L constant). n = $\emptyset$ terms remain. All higher order terms in 'n' vanish. $B_n'' = F_n'' = \emptyset$. Hill further assumes that $\epsilon_{loc} = \epsilon_o$ and $G_i = 1$. Given these, the expression for the electrostatic free energy simplifies to

$$A_{el} = (Q^2/L\epsilon_o)\ [K_\emptyset(xb)/xb\ K_1(xb) + \ln(b/a) + \epsilon_o \ln a]$$

The first two terms correspond to Debye-Huckel free energy for the cylindrical polyion. 'a' represents radius of exclusion for the ion atmosphere from the cylindrical axis. The third term is analogous to Born's work of discharging the polyion in vacuum (also see equation (36) on p-149).

To conclude, equation (27) gives the expression for electrostatic free energy of a discrete charge distribution (in the presence of ion atmosphere) imbedded in a coaxial cylindrical dielectric continuum.

CHAPTER VI

SUMMARY AND CONCLUSIONS

VI. A. A COLLECTIVE PERSPECTIVE

The phosphodiester torsions are found to favor the gauche-gauche (gg) conformation of dimethylphosphate anion relative to the extended forms, both in free space and aqueous solutions as concluded from the determined intra and intermolecular thermodynamics. Results are in general accord with experimental studies on this system. The intramolecular thermodynamics of DMP^- is seen to be dictated by anomeric effects stabilizing the gg conformation in free space. Both ionic and hydrophobic hydrations together are observed to be responsible for the stability of the gg form in aqueous solutions. This preference for the gg form over gt conformer emerging through the simulations is traced to hydrophobic effect. The conformational differences in the calculated thermodynamic indices suggested that the extended forms (gt and tt) may be thermally accessible at ambient temperature. If the anionic phosphodiester group is paired with an explicit counterion the internal energies of hydration are seen to favor the fully extended tt conformation of DMP^-.

Some salient features of hydration emerging out of the simulations are as follows. The anionic hydration of DMP^- is characterized by about 5 to 6 waters associated sequentially (through monodentate interactions with the anionic oxygens). Bridged water

structures connecting the two anionic oxygens of the phosphodiester group are seen to be preferred in the contour surface calculations (describing the interactions of the solute with a single water molecule at 0^{o} K, in free space) but not in the simulations performed at a temperature of 25^{o} C and experimental densities. The computed local solvent densities near ester oxygens of the three conformations of DMP^{-} studied suggested that this can be a finger print zone for identifying the conformational transitions (experimentally). The hydrophobic hydration of the methyl groups of DMP^{-} is characterised by about 8-10 waters in the first hydration shell with the average (solute-water) pair interaction energies close to zero. A comparison of the solute-water potential functions indicated that a shallower minimum on the contour surface can be correlated to a larger coordination number ($>PO_2^-$ group hydration). It also becomes evident now that cluster calculations are not very informative of the hydration at experimental densities, and particularly of the hydrophobic groups.

The free energy simulation results point out that the hydration shell model underestimates the conformational differences and probably needs a recalibration. The concentric dielectric continuum model overestimates the conformational differences

compared to the simulation results. A rationale for this result is to be found in the fact that continuum model underscores the attractive part of the solute-solvent interaction potential while the simulation results include contributions from both attractive and repulsive interactions. Presumably it is the latter that brings the gg and gt conformations free energetically closer in the case of DMP^- in water as observed in the simulations.

VI. B. SIGNIFICANCE OF RESULTS

In addition to providing a large amount of quantitative information on the hydration in multifunctional solutes, the present studies shed light on forces stabilizing the different phosphodiester conformations in free space and aqueous solutions. The phosphate group hydration as characterised here is of relevance to the hydration of nucleic acids and phospholipid head groups.

On the methodological front, the performance of the probability ratio method using adaptive umbrella sampling technique on thermodynamic cycle of DMP^- is a booster to the simulation methodology in its applications to biological problems. A continuum description of the environmental effects is made possible by the present work for spherical and cylindrical systems. The present investigations bring into perspective three different methodologies for estimating the effects of hydration on the conformational preferences of biological molecules.

Significance to Nucleic Acid Hydration

DMP is not DNA. All extrapolations here are to be taken as some possible suggestions and further work is called for on the hydration of a series of model compounds for nucleic acids for an emphatic statement.

The statistical weight (0.005) of bridged water structures calculated here, taken together with similar observations made on the hydration of glycine zwitterion hydration, is a strong pointer towards the insignificance of such structures in a statistical description of hydration and provides a strong argument against "ordered water filaments" postulated time to time to describe DNA hydration. Such ordered structures are not favored entropically as borne out by the simulations.

The hydrophobic hydration in all the DMP simulations is less dense than bulk water. This is significant to the minor groove hydration in B-DNA. The DMP^- results indicate that a conformational transition from gg to gt form can be induced by disrupting the hydrophobic hydration. This again is significant to the minor groove hydration in nucleic acids.

VI. C. RETROSPECTIVE

A systematic study of nucleic acid backbone hydration could have been undertaken starting from DMP^- and slowly building up to a duplex instead of methodolgical issues such as cutoffs, geometries and charges, but these are some unknowns as yet in the representation of hydration through simulation methodology and some more work is required on these lines. The free energy simulation methodology is not completely chartered yet. The determination of intramolecular thermodynamics in the quasiharmonic approximation is a relatively new scheme to comment decisively upon the magnitudes of the calculated entropies. Same is true for the calculated vibrational spectrum particularly in the low frequency region.

Some alternative solute-water potential functions became available and results on DMP^- were published, only after a majority of the studies mentioned in chapter I were completed. While this enabled a comparison of two independent simulation studies (see chapter I) on DMP^- hydration, lack of experimental data on the thermodynamics of hydration prevented a complete understanding of the relative merits of the potential functions in representing solute-water interactions. Similar comparative studies on larger systems is another possible area for future research.

VI. D. PROBLEMS AND PROSPECTS

As pointed out in the previous section and earlier in Chapter I, the quality of the solute-water potential functions to be used in the simulations is of continuous concern. The problem with endothermic transfer energies in the $[DMP^-]_{aq}$ simulations using the potential functions of Clementi and coworkers here, could not be pinned down to any one particular reason. A similar situation obtains with the choice of the parameters such as partial atomic charges to be used in the simulations, the radial cutoffs for truncating water-water interactions etc. and their implications on the computed hydration indices. Equally important are the solute geometries. The advent of supercomputers has opened up the possibility of acquisition of data on similar systems at a rapid pace and this must help further an understanding on several of the methodological issues in near future. Experimental studies on the thermodynamics of hydration of model compounds can be of immense help in this task.

Potential of mean force type calculations for considering a mobile counterion in the vicinity of the phosphate group are currently extremely time consuming and some methodological improvements are in order. Also, reliable solute-ion potential energy functions are required.

Free energy simulations with umbrella sampling scheme using harmonic weighting functions necessitate further investigations on the convergence of the estimated probabilites, criteria for convergence, the effect of the rotation of the solute in the simulations, and the size of the simulation cell. The closure of the thermodynamic cycle with adaptive umbrella sampling scheme is a promising feature for the free energy simulation methodology.

VI. E. SUGGESTIONS FOR FUTURE WORK

Several of the methodological issues which deserve to be addressed to improve the reliability of the predictions through computer simulations have already pointed out where appropriate so far in this chapter. Particular emphasis needs to be placed on solute-water potential functions and accumulation of theoretical and experimental data on model compounds for nucleic acids. Modifications in the hydration of phosphodiester group brought about by the presence of sugar ring and nucleic acid bases would be interesting to look at. Hydration of a mononucleoside triphosphate followed by an extension of these studies to ApU and GpC mini duplexes should be very informative on the issue of transferability of hydration to nucleic acids. Also mobile counterions in the simulations would bring the theoretical representation closer to reality. Another problem worthy of examination through simulation methodology is the study of barrier heights for conformational transitions of nucleic acids in aqueous solutions.

The role of hydrophobic effect in stabilizing helical forms (gg conformation for phosphodiester torsions in nucleic acids and ϕ, ψ torsions in proteins is an essential prerequisite for helices, and in both cases hydrophobic groups come closer along the

backbone) merits attention. Hydrophobic hydration in the minor groove of nucleic acids may be an important factor in conformational transitions.

During the course of the current investigations several related projects emerged as offshoots. One of them is to obtain dynamic information using the enormous microscopic data that the Monte Carlo simulations generate. Let Q be the dynamical variable and C(Q,t) be its correlation function. (See Ref. 169 and 118 for the experimental observable, the appropriate correlation function and the related dynamic variable). If the motivation is to obtain information on the relaxation processes, (or evaluation of line shapes), one may consider in the simplest treatment of the Monte Carlo data , two extreme cases depending on the type of interactions involved.

(A) Rapid modulation limit (Markovian limit).

$$C(Q,t) = \langle Q(t).Q(\emptyset)\rangle \simeq \langle Q(\emptyset)^2\rangle \exp(-t/t_c)$$

Here $\langle Q(\emptyset)^2\rangle$ is related to the static distribution function and t_c is the correlation time for the relaxation process under consideration. The power spectrum is the Fourier Transform of the correlation function C(Q,t). The exponential relaxation assumed here invariably yields a Lorentzian. A similar

treatment was used in the T_1 calculations discussed in Chapter I. If the solute-solvent interactions are strong, the static limit, is more appropriate.

(B) Static limit

Ergodic theorem can be exploited. The distribution of frequency shifts for each configuration can be evaluated and the shifts over all configurations accumulated. A plot of this directly gives the shape of the peak centered at gas phase value. This typically gives a Gaussian. Some relevant computational details are given in Ref. 170 and 171.

(C)Langevin approach

Both (A) and (B) are two extreme cases. A better description of the dynamics would be to treat the solvent as a heat bath, utilize the Monte Carlo data to evaluate the average forces on the solute due to the heat bath, set up the Brownian dynamics equations of motion and integrate them. The friction Kernel may be treated as a constant [172] or memory effects may be effectively built into it [173].

(D) Generalized Langevin approach.

Molecular time scale generalized Langevin approach (MTGLE) [174], is an improvement over a Brownian description of the heat bath in that it considers the solvent dynamics (although in a limited way). The four

computational steps involved are as follows.

(1) Calculation of the solute-solute and solvent-solute equilibrium pair distribution functions. (This can be done using the Monte carlo method.)

(2) Construction of the solute configuration dependent correlation functions.

(3) Construction of the chain equations from the correlation functions.

(4) Integration of the above equations to obtain solutions by the classical stochastic trajectory methods.

Details for steps (2) to (4) can be found in Ref. 174.

Monte Carlo Algorithms : Traditionally, Monte Carlo algorithms (reviewed in section I. C.) have been examined from the stand point of variance reduction techniques in a sampling scheme. A preliminary investigation of the random walks in phase space suggested three potentially useful approaches to the problem of specification of the transition probabilities to conduct the random walks, namely, (1) Group theoretical techniques (2) Measure theoretical and (3) Topological approaches. An outline of the program for these three approaches is given below.

Group theoretical approach :(1) Extend Gibbsian statistical groups to rings and give an algebraic structure to the groups and study the symmetry properties of the elements obeying this algebra. To elaborate, (a) define a set of pair-wise transition probabilities operating on the statistical group, permuting the elements. This constitutes a transition probability space. (b) If the definition of microscopic reversibility is incorporated into the definition of the above set, it forms a multiplicative group. (c) The above two together give a ring structure to the probability space. Viewing this as a vector space over the field of real numbers one can define an algebra for the probabilities. (d) Borrowing on the symmetry program in quantum mechanics, the Gibbsian ensembles

are isomorphic to a factor group of transition probabilities. (e) Enumerate the properties of the elements in the transition probability space.

Measure theoretical approach: Examine the definition of probability measure on the phase space. Extend the notion of measure to transition probabilities. Find a variational principle to obtain the optimum measure.

Topological approach: It is trivial to prove that if d is a metric (a) min {1,d} is a metric as well, as also (b) d/1+d. Now the parallel between the acceptance probabilities, a_{ij}, of Metropolis et al. and the above metric (a), and the relationship between the metric (b) and Barker's choice is a compelling source for considering an isomorphism between the Euclidean space and the transition probability space.

CHAPTER VII

REFERENCES

1. W. Saenger, "Principles of Nucleic Acid Structure", Springer Verlag, New York (1984).

2. W. Olson, "Topics in Nucleic Acid Structure", Pt. 2, editor: S. Neidle, Macmillan Publishers Ltd., London (1981).

3. S. Arnott and D. W. L. Hukins, Nature, London, 224, 886 (1969).

4. M. Sundaralingam, Biopolymers, 7, 821 (1969).

5. H. M. Berman and H-S. Shieh, "Topics in Nucleic Acid Structure", Pt. 1, editor: S. Neidle, Macmillan Publishers Ltd., London (1981).

6. N. S. Zefirov, Tetrahedron, 33, 3193 (1977).

7. V. A. Bloomfield, D. M. Crothers and I. Tinoco, "Physical Chemistry of Nucleic Acids", Harper and Row, New York (1974).

8. M. Falk, K. A. Hartman Jr. and R. C. Lord, J. Am. Chem. Soc., 84, 3843 (1962).

9. M. Falk, K. A. Hartman Jr. and R. C. Lord, J. Am. Chem. Soc., 85, 387 (1963).

10. M. Falk, A. G. Poole and C. G. Goymour, Can. J. Chem., 48, 1536 (1970).

11. S. Lewin, J. Theor. Biol., 17, 181 (1967).

12. A. J. Hopfinger, "Intermolecular Interactions and Biomolecular Organization", Wiley Interscience, New York (1977), p-324.

13. U. Dahlbord, A. Rupprecht, Biopolymers, 10, 849 (1971).

14. H. R. Drew and R. E. Dickerson, J. Mol. Biol., 151, 535 (1981).

15. B. N. Conner, T. Takano, S. Tanaka, K. Itakura, R. E. Dickerson, Nature, 295, 294 (1982).

16. O. Kennard, Pure and Appl. Chem., 56, 989 (1984).

17. M. T. Mai, D. E. Wemmer and O. Jardeztky, J. Am. Chem. Soc., 105, 7149 (1983).

18. D. Perahia, M. S. John, and B. Pullman, Biochim. Biophys. Acta, 474, 349 (1977).

19. B. Pullman, J. Biomol. Stereo Dynamics, I(3), 213 (1983).

20. E. Clementi and G. Corongiu, Biopolymers, 18, 2431 (1979).

21. E. Clementi and G. Corongiu, Int. J. Q. Chem., 16, 897 (1979).

22. E. Clementi and G. Corongiu, Biomol. Stereo Dynamics, 1, 209 (1983).

23. E. Clementi and G. Corongiu, Int. J. Q. Chem., 22, 595 (1982).

24. G. Corongiu and E. Clementi, Biopolymers, 20, 2427 (1981).

25. E. Clementi and G. Corongiu, Biopolymers, 21, 763, (1982).

26. L. Giarda, F. Garbassi and M. Calcaterra, Acta. Cryst., B 29, 1826 (1973).

27. S. J. Wiener, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta and P. Wiener, J. Am. Chem. Soc., 106, 765 (1984).

28. T. Shimanouchi, M. Tsuboi and Y. Kyogoku, Adv. Chem. Phys., 7, 435 (1964).

29. C. Garrigou-Lagrange, O. Bouloussa and C. Clement, Canadian Jl, Spectroscopy, 21, 75 (1976).

30. (a) D. G. Gorenstein in "Phosphorous-31 NMR, Principles and Applications", editor: D. G. Gorenstein, Acad. Press, New York (1984), ch-1.

30. (b) D. G. Gorenstein, J. B. Findlay, R. K. Momii, B. A. Luxon and D. Kar, Biochemistry, 15, 3796 (1976).

31. D. B. Lerner, W. J. Becktel, R. Everett, M. Goodman and D. R. Kearns, Biopolymers , 23, 2157 (1984).

32. W. K. Olson and P. J. Flory, Biopolymers, 11, 1 (1972).

33. N. Yathindra and M. Sundaralingam, Proc. Nat. Acad. Sci., USA, 71, 3325 (1974).

34. V. Sasisekharan and A. V. Lakshminarayanan, Biopolymers, 8, 505 (1969).

35. A. Saran and G. Govil, J. Theor. Biol., 33, 407 (1971).

36. A. Pullman, Biophys. Biochim. Acta, 269, 1 (1972).

37. M. D. Newton, J. Am. Chem. Soc., 95, 256 (1973).

38. G. Govil, Biopolymers, 15, 2303 (1976).

39. G. Lipari and C. Tosi, Theoret. Chim. Acta (Berl.), 50, 169 (1978).

40. A. R. Srinivasan, N. Yathindra, V. S. R. Rao and S. Prakash, Biopolymers, 19, 165 (1980).

41. E. Platt, B. Robson and I. H. Hillier, J. Theor. Biol., 88, 33 (1981).

42. D. Perahia, B. Pullman and A. Saran, Biochim. Biophys. Acta, 340, 299 (1974).

43. D. Perahia and B. Pullman, Biochim. Biophys. Acta, 435, 282 (1976).

44. D. G. Gorenstein, D. Kar, B. A. Luxon and R. K. Momii, J. Am. Chem. Soc., 98, 1668 (1976).

45. D. G. Gorenstein and D. Kar, J. Am. Chem. Soc., 99, 672 (1977).

46. D. G. Gorenstein, B.A. Luxon and J. B. Findlay, Biochim. Biophys. Acta. 475, 184 (1977).

47. H. Berthod and A. Pullman, Chem. Phys. Letters 32, 233 (1975).

48. D. Perahia, A. Pullman and H. Berthod, Theoret. Chim. Acta (Berl.), 40, 47 (1975).

49. H. Frischleder, S. Gleichmann and R. Krohl, Chem. Phys. Lipids, 19, 144 (1977).

50. R. Gay and G. Vanderkooi, J. Chem. Phys., 75, 2281 (1981).

51. G. Alagona, C. Ghio and P. A. Kollman, J. Am. Chem. Soc., 105, 5226 (1983).

52. A. Pullman, H. Berthod and N. Gresh, Chem. Phys. Lett., 33, 11 (1975).

53. G. Corongiu and E. Clementi, Gazz. Chim. Italiana, 108, 687 (1978).

54. E. Clementi, G. Corongiu and F. Lelj, J. Chem. Phys., 70, 3726 (1979).

55. B. Pullman, A. Pullman, H. Berthod and N. Gresh, Theoret. Chim Acta (Berl.), 40, 93 (1975).

56. J. Langlet, P. Claverie, B. Pullman and D. Piazzola, International J. Q. Chem. Quant. Biol. Symp., 6, 409 (1979).

57. T. Bleha, J. Mlynek and I. Tvaroska, Collection, Czech. Chem. Commun. 46, 1722 (1981).

58. K. D. Gibson and H. A. Scheraga, Proc. Nat. Acad. Sci., USA, 58, 420 (1967).

59. L. G. Dunfield, A. W. Burgess and H. A. Scheraga, J. Phys. Chem., 82, 2601 (1978).

60. Z. I. Hodes, G. Nemethy and H. A. Scheraga, Biopolymers, 18, 1565 (1979).

61. A. J. Hopfinger, "Conformational Properties of Macromolecules", Acad. Press, New York (1973), p-70.

62. D. L. Beveridge, P. V. Maye, B. Jayaram, G. Ravishanker and M. Mezei, J. Biomol. Structure and Dynamics, 2, 261 (1984).

63. G. Alagona, C. Ghio and P. A. Kollman, J. Am. Chem. Soc., 107, 2229 (1985).

64. N. C. Seeman, J. M. Rosenberg, F. L. Suddath, J. J. P. Kim and A. Rich, J. Mol. Biol., 104, 109 (1976).

65. J. M. Rosenberg, N. C. Seeman, R. O. Day and A. Rich, J. Mol. Biol., 104, 145 (1976).

66. Y. Kyogoku and Y. Iitaka, Acta Cryst., 21, 49 (1966).

67. J. P. Hazel, R. L. Collin, Acta Cryst., B28, 2951 (1972).

68. T. Glonek and J. R. Van Wazer, J. Phys. Chem., 80, 639 (1976).

69. R. K. Nanda and G. Govil, Theoret. Chim. Acta (Berl.), 38, 71 (1975).

70. B. Pullman, N. Gresh and H. Berthod, Theoret. Chim. Acta (Berl.), 40, 71 (1975).

71. D. S. Marynick and H. F. Schaefer III, Proc. Nat. Acad. Sci., USA, 72, 3794 (1975).

72. A. Pullman and H. Berthod, Chem. Phys. Letters, 41, 205 (1976).

73. P. Liebmann, G. Loew, A. D. Mc Lean and G. R. Pack, J. Am. Chem. Soc., 104, 691 (1982).

74. H. Berthod and A. Pullman, Chem. Phys. Letters, 46, 249 (1977).

75. A. Pullman, B. Pullman and H. Berthod, Theoret. Chim. Acta (Berl.), 47, 175 (1978).

76. G. Corongiu and E. Clementi, Biopolymers, 20, 2427 (1981).

77. J. P. Hansen and I. R. McDonald, "Theory of Simple Liquids", Acad. Press., New York (1976), ch-3.

78. E. G. Beltrametti and G. Cassinelli, "The Logic of Quantum Mechanics", Encyclopedia of Mathematics and its Applications, 15, Addison-Wesley, Massachusetts (1981), ch-6.

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

80. D. A. Mcquarrie, "Statistical Mechanics", Harper and Row, New York (1973).

81. D. Levesque, J. J. Weis and J. P. Hansen, "Monte Carlo Methods in Statistical Physics", editor: K. Binder, Springer-Verlag, NY, 47 (1979), ch-2.

82. J. G. Kemeny and J.L. Snell, "Finite Markov Chains", Van Nostrand Company Inc., Princeton, NJ (1960).

83. W. W. Wood, " Physics of Simple Liquids", editors: H. N. V. Temperly, G. S. Rushbrooke and J. S. Rowlinson, North Holand Publishing Co., Amsterdam (1968), ch-5.

84. W. K. Hastings, Biometrika, 57, 97 (1970).

85. A. A. Barker, Aust. J. Phys., 18, 119 (1965).

86. P. H. Peskun, Biometrika, 60, 607 (1973).

87. J. P. Valleau and S. G. Whittington, "Modern Theoretical Chemistry", editor: B. J. Berne, Plenum Press, New York, 5, 137 (1977), ch-4.

88. C. Pangali, M. Rao and B. J. Berne, Chem. Phys. Lett., 55, 413 (1978).

89. M. Rao, C. Pangali and B. J. Berne, Mol. Phys., 37, 1773 (1979).

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

91. (a). J. C. Owicki and H. A. Scheraga, Chem. Phys. Lett., 47, 600 (1979).

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

92. F. J. Millero, J. Phys. Chem., 82, 789 (1978).

93. F. J. Millero, in "Water and Aqueous solutions", editor: R. A. Horne, Wiley Interscience, New York (1972), p-519.

94. J. E. Erpenbeck and W. W. Wood, in "Modern Theoretical Chemistry", editor: B. J. Berne, Plenum, New York (1977), Vol. 6, Ch-2.

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

96. O. Matsuoka, E. Clementi and M. Yoshimine, J. Chem. Phys., 72, 3979 (1980).

97. D. L. Beveridge, M. Mezei, P. K. Mehrotra, F. T. Marchese, G. Ravishanker, T. R. Vasu and S. Swaminathan, in "Molecular Based Study of Fluids", editors: J. M. Haile and G. A. Mansoori, American Chemical Society, 1983.

98. Gaussian-80, QCPE, adapted for IBM by S. Topiol and R. Osman.

99. O. Matsuoka, C. Tosi and E. Clementi, Biopolymers, 17, 33 (1978).

100. H. Kistenmacher, H. Popkie and E. Clementi, J. Chem. Phys., 59, 5842 (1973).

101. M. Mezei and D. L. Beveridge, J. Chem. Phys., 74, 6902 (1981).

102. H. Kistenmacher, H. Popkie and E. Clementi, J. Chem. Phys., 58, 1689 (1973).

103. D. L. Beveridge, M. Mezei, P. K. Mehrotra, F. T. Marchese, T. R. Vasu and G. Ravishanker, "Liquid State Computer Simulations of Biomolecular Solvation Problem", prepared for conference on Quantum Chemistry in Biomedical Sciences, New York Acad. Sci., June 2-4 (1980).

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

105. M. Mezei and D. L. Beveridge, Methods in Enzymology, 127, 21 (1986).

106. A. Ben-Naim, "Water and Aqueous solutions", Plenum, New York (1974).

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

108. F. T. Marchese and D. L. Beveridge, J. Am. Chem. Soc., 106, 3713 (1984).

109. P. V. Maye, M. Mezei and D. L. Beveridge, in preparation.

110. W. L. Jorgensen, J. Chem. Phys., 77, 5757 (1982).

111. P. J. Rossky and M. Karplus, J. Am. Chem. Soc., 101, 1913 (1979).

112. S. W. Harrison, M. Mezei and D. L. Beveridge, submitted to Israel J. Chem. (1986).

113. L. R. Pratt and D. Chandler, J. Chem. Phys., 73, 3430 (1980).

114. M. Mezei, P. K. Mehrotra and D. L. Beveridge, J. Biomol. Structure and Dynamics, 2, 1 (1984).

115. G. L. Siebel, U. C. Singh and P. A. Kollman, Proc. Nat. Acad. Sci., USA, 82, 6537 (1985).

116. G. C. Piementel and A. C. McClellan, "Hydrogen Bonding , Freeman, San Francisco (1960).

117. H. G. Hertz and C. Raedle, Ber. Bunsenges. Phys. Chem., 77, 521 (1973).

118. R. G. Gordon, Adv. Magn. Reson., 3, 1 (1968).

119. H. G. Hertz, in "Water: A Comprehensive Treatise", editor: F. Franks, Plenum, New York, Vol.3, ch-7.

120. H. Shindo, Biopolymers, 19, 509 (1980).

121. N. Go and H. A. Scheraga, J. Chem. Phys., 51, 4751 (1969).

122. M. Karplus and J. Kuschick, Macromolecules, 14, 325 (1981).

123. R. M. Levy, O. Rojas and R. A. Freisner, J. Phys. Chem., 88, 4233 (1984).

124. G. Ravishanker, M. Mezei and D. L. Beveridge, J. Comp. Chem., 7, 345 (1986).

125. B. Jayaram, G. Ravishanker and D. L. Beveridge, in preparation.

126. P. S. Ramanathan and H. L. Friedman, J. Chem. Phys., 54, 1086 (1971).

127. H. L. Friedman and C. V. Krishnan, J. Soln. Chem., 2, 119 (1973).

128. A. Ben-Naim, Farad. Symp. Chem. Soc., 17, 121 (1982).

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

130. M. Mezei and D. L. Beveridge, "Free energy simulations", Prepared for the conference on, "Computer simulations and biomolecular systems", New York Acad. Sci., October (1985).

131. M. Mezei, P. K. Mehrotra and D. L. Beveridge, J. Am. Chem. Soc., 107, 2239 (1985).

132. M. Mezei, J. Comp. Phys., in print.

133. D. L. Beveridge and G. W. Schnuelle, J. Phys.Chem., 79, 2562 (1975).

134. M. Born, Z. Phys., 1, 45 (1920).

135. J. O'M. Bockris and A. K. N. Reddy, "Modern Electro Chemistry", Plenum, New York (1970), Vol. I, ch-2.

136. L. Onsager, J. Am. Chem. Soc., 58, 1486 (1936).

137. J. G. Kirkwood, J. Chem. Phys., 1, 351 (1934).

138. B. Linder, Adv. Chem. Phys., 12, 225 (1965).

139. C. Tanford, "Physical Chemistry of Macromolecules", John Wiley and Sons Inc., New York (1961), ch-7.

140. J. D. Jackson, "Classical Electrodynamics", Wiley, New York (1967).

141. M. H. Abraham and J. Liszi, J. Chem. Soc., Faraday Trans., I, 1604 (1978).

142. P. A. Kollman, P. K. Wiener and A. Dearing, Biopolymers, 20, 2583 (1981).

143. S. W. Harrison, H. J. Nolte and D. L. Beveridge, J. Phys. Chem., 80, 2580 (1976).

144. D. Soumpasis, J. Chem. Phys., 69, 3190 (1978).

145. H. Margenau and G. M. Murphy, "The Mathematics of Physics and Chemistry", Van Nostrand, Princeton, New Jersey (1956).

146. H. Morawetz, "Macromolecules in Solution", High Polymers, 21, Interscience, New York (1965).

147. C. F. Anderson and M. T. Record Jr., Ann. Rev. Phys. Chem., 33, 191 (1982).

148. S. A. Rice and M. Nagasawa, "Polyelectrolyte Solutions", Acad. Press, New York (1961), ch-7.

149. G. S. Manning, J. Chem. Phys., 51, 924 (1969).

150. G. S. Manning, Q. Rev. Biophys., 2, 179 (1978).

151. D. Dolar and A. Peterlin, J. Chem. Phys., 50, 3011 (1969).

152. J. R. Philip and R. A. Wooding, J. Chem. Phys., 52, 953 (1970).

153. J. Lawrence and B. E. Conway, J. Phys. Chem., 75, 2353 (1971).

154. A. Katchalsky, Rev. Pure. Appl. Chem., 26, 327 (1971).

155. A. D. MacGillivray, J. Chem. Phys., 56, 80 (1972), and 57, 4071 (1972).

156. J. M. Bailey, Biopolymers, 12, 559 (1973).

157. S. Sugai and K. Nitta, Biopolymers, 12, 1363 (1973).

158. H. Gregor and J. M. Gregor, J. Chem. Phys., 66, 1934 (1977).

159. D. Stigter, J. Phys. Chem., 82, 1603 (1978).

160. M. Fixman, J. Chem. Phys., 70, 4995 (1979).

161. B. K. Klein, C. F. Anderson and M. T. Record Jr., Biopolymers, 20, 2263 (1981).

162. M. Gueron and G. Weisbuch, Biopolymers, 19, 353 (1980).

163. B. H. Zimm and M. LeBret, J. Biomol. Structure and Dynamics, 1, 461 (1983).

164. G. V. Ramanathan, J. Chem. Phys., 78, 3223 (1983).

165. M. LeBret and B. H. Zimm, Biopolymers, 21, 271 (1982).

166. G. Pack and B. J. Klein, Biopolymers, 22, 233 (1983).

167. F. P. Hildebrand, "Advanced Calculus for Applications", Prentice Hall, New Jersey (1976).

168. M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions", Dover, New York (1965).

169. B. J. Berne and G. D. Harp, Adv. Chem. Phys., 17, 63 (1970).

170. D. W. Oxtoby, J. Phys. Chem., 87, 3028 (1984).

171. F. G. Dijkman and J. H. Van der Maas, J. Chem. Phys., 66, 3871 (1977).

172. D. L. Ermak and J. A. McCammon, J. Chem. Phys., 69, 1352 (1978).

173. J. A. McCammon and P. Wolynes, J. Chem. Phys., 66, 1452 (1977).

174. S. A. Adelman, Adv. Chem. Phys., 53, 61 (1983).