large methyl groups rotating near the highly ionic O2 groups, although our derived analytical potential function (Figure 1) can be used if more clever approaches to get around the sampling problem are found. Such umbrella sampling approaches have worked nicely in studying internal rotation in *n*-butane, and it is likely that the presence of the strong ionic forces causes the difficulties in preventing a similar study of DMP.

Registry No. Dimethyl phosphate anion, 7351-83-9.

Monte Carlo Determination of the Free Energy and Internal Energy of Hydration for the Ala Dipeptide at 25 °C

M. Mezei, P. K. Mehrotra, and D. L. Beveridge*

Contribution from the Department of Chemistry, Hunter College of the City University of New York, New York, New York 10021. Received June 11, 1984

Abstract: The differences in the free energy and internal energy of hydration of the Ala dipeptide in the C_7, α_R, C_5, and P_{II} conformations were computed with the Monte Carlo method in the (T,V,N) ensemble at 25 °C. The free-energy differences were obtained by determining the relative probabilities of the conformations that lie on a line that connects two conformations in the (ψ,ϕ) torsion angle space. The determination of one free-energy difference required three to five separate Monte Carlo runs using non-Boltzman sampling. The results indicate that both the α_R and P_{II} conformations are preferentially stabilized by hydration. The major contributing factor for the stability of internal energy of hydration for these conformations can be traced to the hydration of the carbonyl group.

I. Introduction

The structure of the Ala dipeptide, *N*-acetylalanyl-*N*-methylamide (AcAlaNHMe), in aqueous solution is a matter of considerable current interest in structural biochemistry. The sterically allowed regions of conformation for AcAlaNHMe are prototypical of the polypeptide backbone with all amino acid residues except Gly and Pro, and thus knowledge of the conformational stability of this molecule is highly relevant to fundamental aspects of protein and enzyme structure.[1] The intramolecular energetics of AcAlaNHMe and related molecules have been extensively studied by means of empirical energy functions and molecular quantum mechanics. However, our knowledge of the effect of solvent, especially aqueous hydration, is less complete, and diverse experimental and theoretical results are at variance with one another.

We present herein a new theoretical determination of the thermodynamics of hydration for AcAlaNHMe at 25 °C, based on statistical thermodynamic liquid-state computer simulation. Direct calculations on the internal energy and the free energy of hydration are reported, and new methodological details for free-energy calculations are described. The results obtained are compared with those of previous theoretical calculations and available experimental data.

II. Background.

The conformation of AcAlaNHMe can be specified in terms of the angles of torsion ψ(N–C–C–N) and ϕ(C–N–C–C), as defined in ref 2. Structural studies to date indicate four conformations from the allowed regions of (ψ,ϕ) space for specific consideration:[2] C_7 (90°, –90°), C_5 (150°, –150°), α_R (–50°,

Table I. Calculated Intramolecular Energies (kcal/mol) of the C_5, C_7, α_R, and P_{II} Conformations of AcAlaNHMe Relative to C_7

authors	method	C_7	C_5	α_R	P_{II}
Brant and Flory[7]	EPF	0	0	1	0
Momany et al.[8]	CNDO	0	5	5	5
Hoffman and Imamura[9]	EHT	0	–2	0	0
Pullman et al.[1]	PCILO	0	2	3–4	4
Robson et al.[10,11]	3s2p	0	–3.2	7.8	
	4-31G	0	–2.1	7.5	
	CFF	0	1.6	7.7	
Karplus et al.[12]	EPF	0	6	8	6.5
Scheraga et al.[13]	ECEPP	0	0.38	1.13	1.0

–70°), and P_{II} (150°, –80°). The C_7 and C_5 structures, shown in Figures 1 and 2, respectively, are characterized by seven- and five-atom ring structures completed by an intramolecular hydrogen bond between a CO and NH group. The C_7 form can exist in both axial and equatorial forms; we consider herein the equatorial form. The α_R and P_{II} conformers, shown in Figures 3 and 4, respectively, have ψ and ϕ angles similar to those found in the right-handed polypeptide α-helix and the poly(L-proline)-II helix. However, the α_R and P_{II} forms of AcAlaNHMe are not fully representative of helical polypeptides, which are further stabilized by intramolecular hydrogen bonds between subunits.

The structure of AcAlaNHMe in CCl$_4$ was established by Avignon and Lascombe by infrared spectroscopy[3] to be predominantly C_7. The crystal structure, reported by Harada and Iitaka,[4] involved two conformers: P_{II} and the structurally similar poly-(L-proline)-I form. Avignon et al. subsequently extended their study to AcAlaNHMe in water[5] and proposed on the basis of an

(1) B. Pullman and B. Maigret in "Conformation of Biological Molecules and Polymers", E. Bergman and B. Pulman, Eds., Academic Press, New York, 1973.
(2) C. R. Cantor and P. R. Schimmel, "Biophysical Chemistry", Vol. 1, W. H. Freeman, San Francisco, 1980.

(3) M. Avignon and J. Lascombe in "Conformation of Biological Molecules and Polymers", E. Bergman and B. Pulman, Eds., Academic Press, New York, 1973.
(4) Y. Harada and Y. Iitaka, *Acta Crystallogr., Sect. B*, **30**, 1452 (1974).
(5) M. Avignon, C. Garrigou-Lagrange, and P. Bothorel, *Biopolymers*, **12**, 1651 (1973).

2240 *J. Am. Chem. Soc., Vol. 107, No. 8, 1985*

Mezei et al.
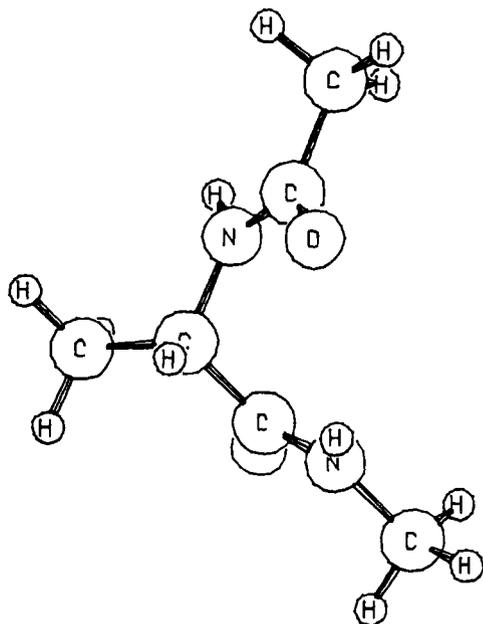
Figure 1. AcAlaNHMe in the C₇ conformation.
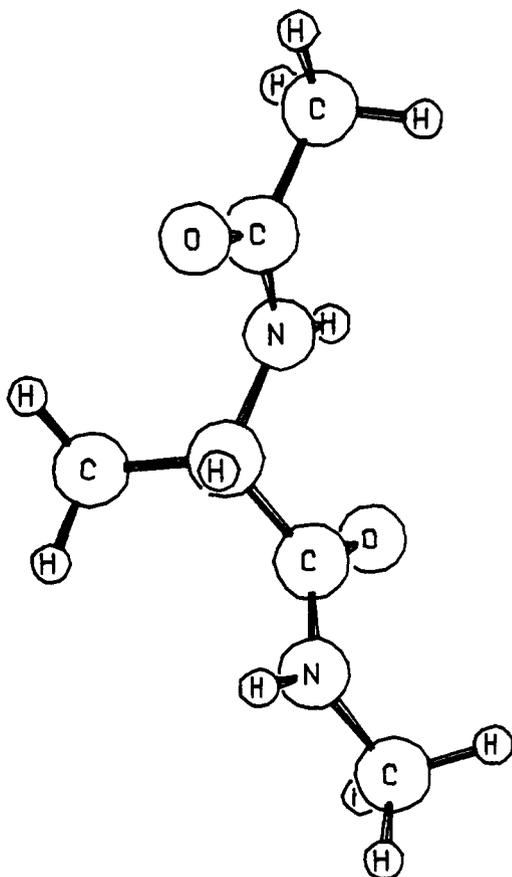
Figure 2. AcAlaNHMe in the C₅ conformation.

Figure 3. AcAlaNHMe in the α_R conformation.
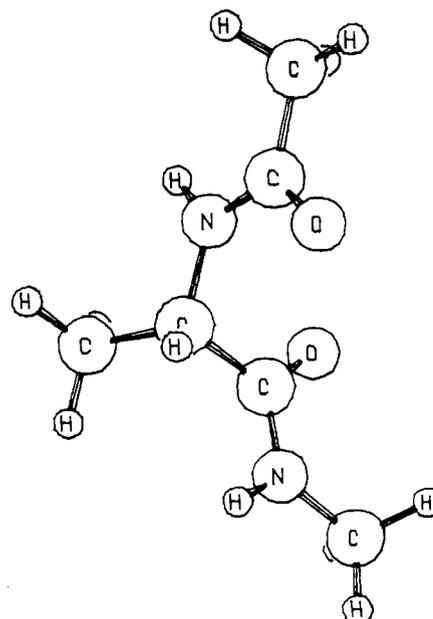
Figure 4. AcAlaNHMe in the P_{II} conformation.

The results of diverse theoretical calculations of the relative stability of the C₇, C₅, α_R, and P_{II} conformers in isolated Ac-AlaNHMe[7-13] are collected in Table I. Calculations using empirical energy functions (EPF) by Brandt and Flory[7] and by Scheraga and co-workers[13] show that C₇, C₅, and α_R structures correspond with energy minima of comparable depth, with P_{II} in a sterically allowed region likely to be thermally accessible. The consistent force field developed by Hagler et al.[14] applied to a related molecule, *N*-formylalanylamide, by Robson et al.[11] shows considerable destabilization of the α_R form; Karplus and Rossky[12] use a similar prescription for the intramolecular energy. Quantum

analysis of depolarized Rayleigh scattering that the C₇ conformer is maintained in water, stabilized by a water molecule bridging exterior CO and NH groups. Recently Madison and Kopple[6] studied the alanine dipeptide in CCl₄ and in H₂O using CD spectra and NMR spin coupling analysis. They confirmed earlier proposals of C₇ for nonpolar solvents, but found evidence for contributions from both α_R and P_{II} conformations of AcAlaNHMe in water.

(6) V. Madison and K. D. Kopple, *J. Am. Chem. Soc.*, **102**, 4855 (1980).

(7) D. A. Brandt and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).

(8) F. A. Momany, R. F. McGuire, J. F. Yan, and H. A. Scheraga, *J. Phys. Chem.*, **75**, 2286 (1971).

(9) R. Hoffman and I. Imamura, *Biopolymers*, **7**, 207 (1969).

(10) I. H. Hillier and B. Robson, *J. Theor. Biol.*, **76**, 83 (1979).

(11) B. Robson, I. H. Hillier, and M. F. Guest, *J. Chem. Soc., Faraday Trans. 2*, **74**, 1311 (1978).

(12) P. J. Rossky, M. Karplus, and A. Rahman, *Biopolymers*, **18**, 825 (1979).

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

(14) A. T. Hagler, G. Nemethy, and S. Lifson, *J. Am. Chem. Soc.*, **96**, 5319 (1974).

mechanical calculations[8-11] show sharper differences in conformational energies, but generally support C_7 as the energy minimum in the free space approximation. Inclusion of electron correlation, as in the PCILO calculation,[1] tends to flatten the surface.

Studies of solvent effects on the conformational stability of AcAlaNHMe were first reported by Venkatachalam and Krimm,[15] who computed a (ψ, ϕ) map with specific water molecules fixed in hydrogen-bonding positions to the interior CO and NH groups. Rein and co-workers[16] reported calculations on AcAlaNHMe in CCl_4 and in H_2O using an Onsager continuum model for the solvent. The minimum in CCl_4 was near C_7 in accord with experiment, and the minimum in water was found at $\psi = 120°$, $\phi = 60°$. Recently Scheraga and co-workers, using a new extended version of the hydration-shell approach, obtained C_7 as a minimum for AcAlaNHMe in CCl_4 and C_5 in H_2O.[13,17] Their calculated conformational map is, however, remarkably flat in the sterically allowed region, and thus other conformations are thermally accessible. In summary, one finds general agreement between experiment and theory on C_7 as the conformational preference of the alanine dipeptide in nonpolar solvents, but diverse results on the nature of the structure in water.

Liquid-state computer simulation has recently been used to study the molecular dynamics of $[AcAlaNHMe]_{aq}$ in the C_7 form by Rossky and Karplus.[18] Hagler[19] has reported Monte Carlo calculations of the internal energy and structure of C_7 and α_R forms of AcAlaNHMe and found a dipeptide–water energy difference of 5.6 kcal/mol favoring α_R. In related work, David[20] carried out Monte Carlo calculations on an aqueous droplet containing glycylglycine zwitterion, with torsion angles treated as a configurational variable. Very recently, Chandrashekar, Smith, and Jorgensen computed the free-energy change along the path of an S_N2 reaction in water.[21] The energetics and hydration of various other peptides have been studied by Madison.[22] Studies on this system by Brady and Karplus,[23] using the molecular dynamics, and by Pettitt and Karplus,[24] using the extended RISM method of Hirata, Pettitt, and Rossky,[25] are in progress. A preliminary report on the internal energy calculations reported here has also been published.[26]

III. Theory and Methodology

A conformationally flexible molecule in solution can in principle assume various structural forms, and the equilibrium state of the system must be described in terms of these structures and their corresponding statistical weights. The relative free energy of the various structures is simply related to their respective probability of occurrence in the ensemble average of structures representative of the statistical state of the system, considering all intramolecular and intermolecular degrees of freedom. A complete theoretical study of this type of problem is not feasible at present for AcA-

laNHMe in water, and thus various assumptions and approximations must be introduced in order to reduce the problem to tractable form.

First, we identify specific conformations of the dissolved molecule which are expected on the basis of other information, experimental or theoretical, to be of particular structural interest. We then direct our calculations to the difference in free energy and related quantities between the selected structures in solution. We assume for two specific conformations i and j with conformational free energies A_i and A_j, that

$$\Delta A_{ij} = \Delta A_{ij}^{int} + \Delta A_{ij}^{hyd} \tag{1}$$

where ΔA_{ij}^{int} is the difference in free energy due to intramolecular factors and ΔA_{ij}^{hyd} is the free-energy difference due to hydration. The coupling of hydration to internal modes and vice versa is neglected.

The quantities A_i and A_j include, in principle, averages over the respective vibrational manifolds of conformational states i and j. Preliminary estimates of ΔA_{ij}^{int} are typically constructed with the vibrational contributions to A_i and A_j assumed to cancel:

$$\Delta A_{ij}^{int} = \Delta E_{ij}^{int} \tag{2}$$

where ΔE_{ij}^{int} is the difference in intramolecular energy of states i and j. The AcAlaNHMe system, where structures with intramolecular hydrogen bonds are compared with open forms, is vulnerable in this approximation. Procedures for treating ΔA_{ij}^{int} more rigorously have become available, and detailed studies of the validity of the approximation in eq 2 are underway in this laboratory and elsewhere.[23] We focus herein on the determination of ΔA_{ij}^{hyd} and the related quantities ΔU_{ij}^{hyd} and ΔS_{ij}^{hyd}.

In a computer simulation on aqueous hydration[27] involving a dissolved molecule A and W water molecules, the total internal energy is determined by numerical integration of the expectation value

$$U^{hyd} = \langle E(X^A, X^W) \rangle \tag{3}$$

where $E(X^A, X^W)$ is the many-particle configurational energy for the system in configuration (X^A, X^W) and the brackets indicate Boltzmann averaging. The difference in the internal energy between two conformations i and j of A is determined from separate simulations, one involving A in conformation i the other A in j, each subject to statistical uncertainty (noise) in the simulation. In taking this difference, we neglect the effect of the partial molar volume of A on conformation. The calculated U_i^{hyd}, U_j^{hyd}, ΔU_{ij}^{hyd}, and the various molecular distribution functions of the system can be partitioned into various contributions for analysis and interpretation of results.[27,28]

The calculation of the free-energy difference between i and j is carried out in a separate series of simulations. Here we define a "conformational transition coordinate" ξ connecting i and j in conformation space by any desired path. The free energy as a function of this coordinate, a potential-of-mean force,[29] is

$$A^{hyd} = -kT \ln g(\xi) \tag{4}$$

$$A^{hyd} = -kT \ln P(\xi) + C \tag{5}$$

where $g(\xi)$ is a spatial correlation function for conformations along ξ in solution, $P(\xi)$ is the probability of occurrence of the conformation in solution, and C is a constant when only torsion angles are varied. Equation 5 follows from the definition of $g(\xi)$. If we now define ξ such that $\xi = 0$ corresponds to structure i and $\xi = 1$ corresponds to j, then

$$\Delta A_{ij}^{hyd} = A^{hyd}(\xi)_{\xi=1} - A^{hyd}(\xi)_{\xi=0} = kT \ln [P(\xi)_{\xi=0}/P(\xi)_{\xi=1}] \tag{6}$$

(15) C. M. Venkatachalam and S. Krimm in "Conformation of Biological Molecules and Polymers", E. Bergman and B. Pulman, Eds., Academic Press, New York, 1973.

(16) V. Renugopalakrishnam, S. Nir, and R. Rein, *Biochim. Biophys. Acta,* **434**, 164 (1976); V. Renugopalakrishnan and R. Rein in "Environmental Effects on Molecular Structure and Properties", B. Pullman, Ed., D. Reidel, Dodrecht-Holland, 1976.

(17) G. Nemethy, Z. I. Hodes, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.,* **75**, 5760 (1978).

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

(19) A. T. Hagler, D. J. Osguthorpe, and B. Robson, *Science,* **208**, 599 (1980).

(20) C. W. David, *J. Phys. Chem.* **85**, 2433 (1981).

(21) J. Chandrashekar, S. F. Smith, and W. L. Jorgensen, *J. Am. Chem. Soc.,* submitted for publication.

(22) V. Madison in "Peptides: Synthesis, Structure and Function", D. H. Rich and E. Gross, Eds., Pierce Chemical Co., Rockford, Ill., 1981.

(23) J. Brady and M. Karplus, private communication; Abstracts of Workshop on Molecular Dynamics and Protein Structure, J. Hermans and W. F. van Gunsteren, Eds., Chapel Hill, N.C., May 1984.

(24) M. Pettitt and M. Karplus, private communication; Abstracts of the American Chemical Society Conference on Theoretical Chemistry, Jackson Hole, Wy., June 1984.

(25) F. Hirata, B. M. Pettitt, and P. Rossky, *J. Chem. Phys.,* **78**, 4133 (1983), and references therein.

(26) P. K. Mehrotra, M. Mezei, and D. L. Beveridge, *Int. J. Quantum Chem.,* in press.

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

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

(29) D. A. McQuarrie, "Statistical Mechanics", Harper and Row, New York, 1976.

The determination of $\Delta A_{ij}{}^{hyd}$ from computer simulation is thus reduced to the determination of $P(\xi)$ for a conformational transition coordinate ξ. Generally, we cannot expect the entire range of from 0 to 1 to be adequately sampled in a standard computer simulation involving associated liquids, and one finds only a fraction of the ξ coordinate covered in a single simulation of reasonable, O(1000K), length. An approach to this problem was developed by Patey and Valleau[30] and Torrie and Valleau[31] and implemented for the study of molecular associations by Pangali et al.[32] and Ravishanker et al.[33] and for conformations of organic molecules in the neat liquid as reviewed by Jorgensen.[34] The idea, called "umbrella sampling", is to use separate simulations to develop $P(\xi)$ about points $\xi_1, \xi_2, \ldots$ such that the $P(\xi)$ for successive points overlap. Then $P(\xi)$ for the entire range of ξ desired can be determined by a straightforward matching procedure.

In a given simulation for point ξ_k, the coordinate ξ can be restricted to motion about ξ_k by adding to the configurational energy $E(X^A, X^W)$ a harmonic constraining function $U_H(\xi, \xi_k)$.[32,33] This is essentially a non-Boltzmann sampling procedure, but Valleau and co-workers[30,31] showed that unbiased estimates of $P(\xi)$ can be extracted by forming

$$P(\xi') = \mathcal{N}\langle\delta(\xi' - \xi)\exp(U_H(\xi)/kT)\rangle_H/\langle\exp(U_H(\xi)/kT)\rangle_H \tag{7}$$

where $\delta(\xi' - \xi)$ is the Dirac delta counting function for configurations with coupling parameter ξ', $\mathcal{N}$ is a normalization factor, and the subscript H denotes an ensemble average determined with the harmonic constraining function added to the conventional expression for the configurational energy. With $P(\xi)$ thus determined, $\Delta A_{ij}{}^{hyd}$ can be computed from eq 6, and in our problem relative free energies of hydration for the various conformations of AcAlaNHMe can be determined in computer simulation by umbrella sampling on conformational transition coordinates linking the various structures under consideration.

Various conformers of AcAlaNHMe differ in ψ and ϕ values, which suggest a two-dimensional version of umbrella sampling. We initially explored this possibility, but found that developing overlapping distributions for the entire (ψ,ϕ) space would require a minimum of 36 separate simulations. Also, special difficulties arise in matching multidimensional distributions as explained in the Appendix. We subsequently chose the one-dimensional procedure with the defining equation for ξ taken to be

$$(\psi,\phi) = \xi(\psi_j,\phi_j) + (1 - \xi)(\psi_i,\phi_i) \tag{8}$$

Thus ξ becomes a reduced, correlated conformational coordinate, analogous to the virtual bonds defined in complex conformational studies of nucleic acids.[36]

In many instances a correlated change as described here may at some points place the solute in unphysical, sterically forbidden conformations. However, in the calculation of the thermodynamics of hydration, the intramolecular energy is not included, i.e., assumed to be zero for this aspect of the calculation.[35] Therefore, intramolecular repulsions will not hinder the sampling of the system as a function of ξ. The required free energy, a state function, can be obtained from any path on which overlapping $P(\xi)$ can be constructed. Thus for the determination of differences in thermodynamic indexes of hydration, the most convenient and economical correlated conformational coordinate can be chosen. Clearly the correlated conformational coordinate approach can be generalized to any number and any types of internal coordi-

nates, and may be widely useful for problems in structural chemistry and biochemistry.

In this approach, the internal coordinates of the solute not explicitly considered become rigid constraints, and a metric correction as described by Fixman[37] is required. The correction is expected to be within the statistical uncertainty of energy expectation values. Also, in taking conformational energy differences constraint errors tend to cancel.[34] Thus we neglect herein errors due to constraints.

IV. Calculations

For the determination of internal energies, separate (T,V,N) ensemble Monte Carlo simulations were carried out on [AcAlaNHMe]$_{aq}$ in the conformations C_7, C_5, α_R, and P_{II}, respectively. A modified Metropolis procedure incorporating the force bias method and preferential sampling for convergence acceleration was used. The system for study in each case was comprised of 202 rigid molecules, 1 AcAlaNHMe, and 201 water molecules. The simulation was performed at a temperature of 25 °C. The volume of the simulation cell was determined based on the density measurement of Bose and Hudt.[38] The condensed-phase environment of the system was provided by means of periodic boundary conditions. Convergence was monitored by the method of batch means. Full details of the Monte Carlo methodology used in our programs are given in a recent article by Mehrotra et al.[39]

The N-particle configurational energy of the system was in each case calculated using potential functions determined from quantum-mechanical calculations. For the water–water interactions, the MCY–CI(2) potential developed by Matsuoka et al.[40] was used. For the solute–water interactions, a potential function constructed from the 12–6–1 functional form and transferable parameters of Clementi et al.[41] was obtained. Net atomic charges for atoms of AcAlaNHMe for the various conformations were determined by LCAO–SCF–MO calculations using GAUSSIAN 80[42] with the basis set described by Matsuoka et al.[43] and consistent with the potential function determination.

In the computer simulation, all potential functions for water–water interactions were truncated at a spherical cutoff of 7.75 Å, and solute–water interactions were treated under the minimum image convention. Solute–solute interactions are neglected, and thus the system represents a state of infinite dilution. Each complete simulation involved an equilibration period of at least 500K configurations with ensemble averages formed over a succeeding 1500K segment of the realization.

The free-energy determinations were carried out on correlated conformational coordinates between the C_7 and α_R conformations and the C_7 and P_{II} conformations of the alanine dipeptide. Details of the simulations were maintained fully consistent with the internal energy calculations described above. The atomic charges on solute atoms used in the solute–water interaction potential were varied accordingly with ξ:

$$q_n(\xi) = \xi q_{nj} + (1 - \xi)q_{ni} \tag{9}$$

where q_n denotes the net atomic charge of atom n and q_{ni} and q_{nj} are the corresponding quantities for states i and j. In the umbrella sampling, the configurational energy of the system was supplemented with the harmonic constraining function

$$U_H(\xi,\xi_k) = c(\xi - \xi_k)^2 \tag{10}$$

(30) G. N. Patey and J. P. Valleau, *Chem. Phys. Lett.* **21**, 297 (1973).
(31) G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, **28**, 187 (1977).
(32) C. Pangali, M. Rao, and B. J. Berne, *J. Chem. Phys.*, **71**, 2982 (1979).
(33) G. Ravishanker, M. Mezei, and D. L. Beveridge, *Faraday Symp. Chem. Soc.*, **17**, 79 (1982).
(34) W. L. Jorgensen, *J. Phys. Chem.*, **87**, 5304 (1983).
(35) The use of eq 1 and 2 to calculate ΔA_{ij} is formally equivalent to computing ΔA_{ij} with $E(X^A,X^W) = E(X^W) + E^{int}(X^A)$ and performing an umbrella sampling with $U_H' = U_H - E^{int}(X^A)$.
(36) W. Olsen in "Topics in Nucleic Acid Structure", Part II, S. Neidle, ed., MacMillan, London, 1982.

(37) M. Fixman, *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3050 (1974); D. Chandler and B. J. Berne, *J. Chem. Phys.* **70**, (~71) 5386 (1979).
(38) L. Bose and A. Hudt, *J. Chem. Thermodyn.*, **3**, 663 (1971).
(39) P. K. Mehrotra, M. Mezei, and D. L. Beveridge, *J. Chem. Phys.*, **78**, 3156 (1983).
(40) O. Matsuoka, E. Clementi, and M. Yoshemine, *J. Chem. Phys.*, **64**, 1351 (1976).
(41) E. Clementi, F. Cavallone, and R. Scordamaglia, *J. Am. Chem. Soc.*, **99**, 5531 (1977).
(42) J. S. Binkley, R. A. Whitehead, P. C. Hariharan, R. Seeger, J. A. Pople, W. J. Hehre, and M. D. Newton, GAUSSIAN 80, Quantum-Chemical Program Exchange, adapted for IBM by S. Topiol and R. Osmond.
(43) O. Matsuoka, C. Tosi, and E. Clementi, *Biopolymers*, **17**, 33 (1978).

Table II. Characterization of the Individual Monte Carlo Runs

initial state	final state	c	ξ_k	ξ_{min}	ξ_{max}	run length
C_7	α_R	100	0.10	-0.21	0.15	500K
C_7	α_R	100	0.13	-0.11	0.50	600K
C_7	α_R	100	0.35	0.09	0.67	750K
C_7	α_R	100	0.61	0.49	0.83	500K
C_7	α_R	100	0.85	0.73	1.16	500K
C_7	P_{II}	50	0.11	-0.09	0.44	500K
C_7	P_{II}	50	0.51	0.22	0.87	500K
C_7	P_{II}	50	0.81	0.57	1.18	500K

Table III. Calculated Thermodynamic Indexes of Hydration for the C_7, C_5, α_R, and P_{II} Conformations of AcAlaNHMe, Relative to C_7[a]

	C_7	C_5	α_R	P_{II}
ΔU_{ij}^{hyd}	0	8.5	-6.6	-5.5
ΔU_{SW}	0	-1	-5.1	-6.2
ΔU_{WW}	0	9.5	-1.5	0.7
ΔA_{ij}^{hyd}	0		-3.6	-3.2
ΔS_{ij}^{hyd}	0		-10.6	-7.7

[a] U_{SW} is the solute-water energy. U_{WW} is the water-water energy. Energies and free energies are in kcal/mol. Entropies are in cal/mol·°C.

Table IV. Calculated First-Shell Coordination Numbers for the C_7, C_5, α_R, and P_{II} Conformations of AcAlaNHMe, Relative to C_7

	C_7(abs)	C_7	C_5	α_R	P_{II}
total	33.5	0	-1.5	-1.4	0.3
CH_3	29	0	-2.3	-2	-1
CO	2.6	0	0.7	0.6	1.2
NH	1.9	0	0.1	0.0	0.1

Table V. Calculated Solute-Water Binding Energies for the C_7, C_5, α_R, and P_{II} Conformations of AcAlaNHMe, Relative to C_7

	C_7	C_5	α_R	P_{II}
CH_3	0	4.7	0.5	1.3
CO	0	-1.1	-8.2	-5.9
NH	0	-2.0	2.8	-2.2

where the parameter c and the positions ξ_k were determined from trial runs to sample ξ in overlapping segments. The determination of c and the ξ_k in U_H was performed in two stages. First, several 25–50K trial runs were carried out using different c values to determine the size of interval on ξ that could be successfully sampled per simulation. A set of 100K simultaneous runs were started that, based on initial trials, would sample the entire [0,1] interval on ξ. If the range sampled in a given run began to drift significantly, the simulation was restarted with an adjusted ξ_k. Long-range correlations complicated this process slightly, with regions where $P(\xi)$ was near a minimum giving the most problem. For example, two $\alpha_R \rightarrow C_7$ runs using $\xi_k = 0.77$ and $\xi_j = 0.90$ were strongly overlapping for about 300K but overlapped to a much smaller extent in the remainder of the run. With some trial and error, a set of points [ξ_k] which produced overlapping distributions for each transition was obtained. In the final calculations, the $C_7 \rightarrow \alpha_R$ path required five simulations and the $C_7 \rightarrow P_{II}$ route required three. The final set of parameters for U_H, the range of ξ covered, and other computational details of the various simulations are collected in Table II.

The matching of the probability distribution for ξ obtained from the different Monte Carlo runs can be performed, at least in principle, using any point in the distribution that overlaps with a neighboring distribution. In practice, the precision of the values in the distribution obtained varies, leading to slightly different ΔA_{ij}^{hyd} values with different points used for matching. The points to be used for matching can be determined by inspecting the ratios $P(\xi)/P'(\xi)$ for overlapping regions of two runs. Ideally, the ratio should be independent of ξ. While in our case this never was true for the whole range of overlap, we could select the range of ξ's where the ratio was nearly constant and use of one of these points.

V. Results and Discussion

The calculated internal energies of [AcAlaNHMe]$_{aq}$, total and partitioned into solute-water and water-water contributions, are given in Table III. In our results the calculated energetics can be expected to be quite sensitive to choice of intermolecular potential, an aspect to be dealt with in more detail in forthcoming work. Also, the statistical uncertainties in the relative energies are rather high since they are obtained as small differences between large numbers. At this stage of study, we recommend that our numerical results be taken only as a reasonable estimate of the effect of water on various conformations of AcAlaNHMe.

As described in section II, the C_7 conformation seems to be well established as the predominate form of AcAlaNHMe in nonpolar solvents. For ease of interpretation, we thus present our results for aqueous solutions of AcAlaNHMe relative to those for the C_7 conformation and consider those aspects of aqueous hydration

which lead to stabilization and destabilization. Overall, we find the internal energy of the C_5 conformation of AcAlaNHMe to be destabilized and the α_R and P_{II} conformers to be stabilized in water. The α_R and P_{II} conformations in solution are relatively close energetically, 1.1 kcal/mol, with a preference for α_R indicated. The partitionings of total internal energy into solute-water and water-water contributions indicate that the origin of the destabilization of C_5 lies in the water-water term, whereas the stabilization effects for α_R and P_{II} are found mainly in the solute-water term. The P_{II} conformation is, in fact, preferentially favored by solute-water interaction. The ultimate preference for α_R over P_{II} is due to the favorable contributions from the water-water solvent reorganization term. The calculated pressure values show a variation of 450 atm. Since this can give a pV correction of 0.2 kcal/mol only or less, the effect of keeping the partial molar volumes constant is well within the statistical noise in the simulation.

To explore in more detail the structural origins of these results, the calculated first-shell coordination numbers for C_7, C_5, α_R, and P_{II} are collected in Table IV. The average number of water molecules in the first-shell hydration complex for [AcAlaNHMe]$_{aq}$ is found to range between 32 and 33.8, with the variation relative to C_7 only 1.5 molecules. The coordination can be partitioned into contributions from the various functional groups by the proximity criterion.[28] Here we find 85% of the first-shell waters to be engaged in hydrophobic hydration of the methyl groups of AcAlaNHMe, with the remainder assigned to the hydrophilic hydration of the CO and NH groups. Overall there seems to be little change in the first-shell hydration numbers with conformation, and the nature of the hydration effects on structure must depend on the way in which waters interact with the solute and each other rather than the numbers involved.

Since the major stabilizing effect for the α_R and P_{II} conformations appears in the solute-water term, we partitioned to solute water binding energies according to functional groups also by means of the proximity criterion. The major factor contributing to the stabilization of the internal energy of hydration of α_R and P_{II} in aqueous solution is clearly due to energetic factors originating in the hydration of the carbonyl groups of the alanine dipeptide. Stereo views of the structures of a selected hydration complex for the α_R and P_{II} forms of the AcAlaNHMe are shown in Figures 5 and 6, respectively.

The original proposal by Avignon et al.[5] of the C_7 structure stabilized by a water bridge has not held up in subsequent experimental and theoretical studies. However, we have examined specific structures in the C_7 simulation to see the extent to which the characteristic bridging structure they proposed can be found. We observed a structure markedly similar to that of Figure 14.6 of their 1973 paper.[5] Thus, the structure seems to be involved, but does not confer a special stability relative to that of α_R and P_{II}.

From the free energy simulations, the ΔA_{ij}^{hyd} between the C_7 and α_R conformations was found to be -3.6 kcal/mol, and the ΔA_{ij}^{hyd} between C_7 and P_{II} conformations was -3.2 kcal/mol. Our assessment of the noise level in these results is as follows: observation of the ranges covered by successive 25K segments of
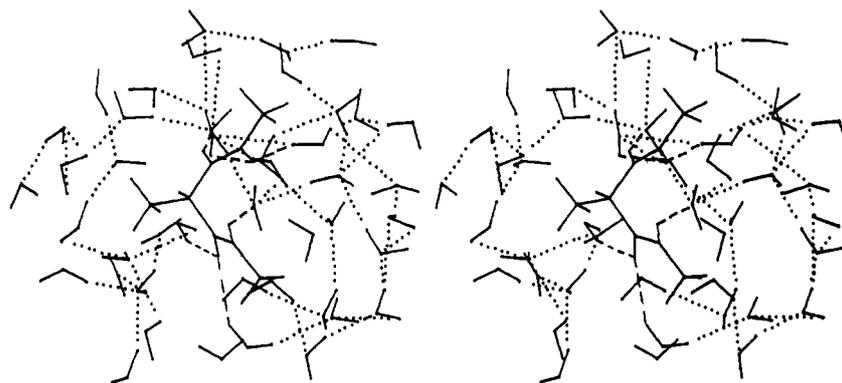
Figure 5. Stereoview of a hydration complex for AcAlaNHMe in the α_R conformation.
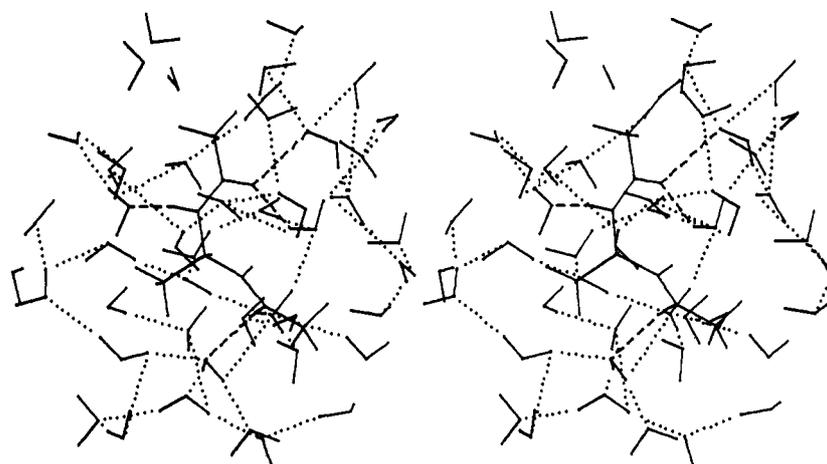
Figure 6. Stereoview of a hydration complex for AcAlaNHMe in the P_{II} conformation.

the runs showed a grand cycle of period 500K, similar to the long-range correlations in the energy found for liquid water.[44,45] Observations of the $P(\xi)/P(\xi')$ ratios on the longest run for ξ and ξ used in matching this run show it to be accurate within 10–20% after 500K. Depending on the number of runs required, this results in an uncertainty factor of 1.5–2.0 in the $P(0)/P(1)$ values, which translates to a statistical error of 0.3–0.6 kcal/mol in ΔA_{ij}^{hyd}. Errors due to assumptions about intermolecular potentials no doubt exceed this value.

Combining the results on the free energy and internal energy of hydration leads to estimates of the entropy difference between C_7 and α_R of $\Delta S_{ij}^{hyd} = -10.1$ cal/mol·°C; for C_7 and P_{II}, $\Delta S_{ij}^{hyd} = 7.72$ cal/mol·°C. Because of the stated error bounds for ΔA_{ij}^{hyd} and ΔU_{ij}^{hyd}, and especially the noise in ΔU_{ij}^{hyd}, these entropy differences are also subject to significant numerical uncertainty factors. However, it is reasonable to conclude that the entropy change for hydration is negative for both $C_7 \rightarrow \alpha_R$ and $C_7 \rightarrow P_{II}$ transitions, presumably owing to the ordering of water in the regions of hydrophilic hydration in the open forms of AcA-laNHMe. This could be compensated, however, by intramolecular entropy effects.

A comparison of the calculated results with experimental data for $[AcAlaNHMe]_{aq}$ requires knowledge of the total free-energy change for the various transitions, the intramolecular contribution plus that due to hydration. As seen from Table I, previous calculations of the intramolecular energy differences among the conformations under consideration here differ widely at the numerical level, and estimates of the intramolecular entropy change are not yet available for this system. Madison and Kopple[6] found evidence for both α_R and P_{II} conformations for $[AcAlaNHMe]_{aq}$ at 25 °C. Our calculations show these two conformations to be

preferentially stabilized by hydration, and quite close in their relative free energies of hydration. If the adiabatic (ψ,ϕ) intramolecular energy differences between sterically allowed conformations is small, as predicted in several of the studies cited, and most recently in the new version of energy functions by Scheraga and co-workers,[13] the calculations and experimental studies are in close accord and are mutually supportive. Further research on the nature of the intramolecular free-energy surface for AcAlaNHMe is required before a stronger position can be taken on this matter.

Appendix. Matching Probability Distributions in N Dimensions.

When an N-dimensional probability distribution is obtained by piecemeal matching of unnormalized distributions that were computed in N_r smaller overlapping domains, then the matching to produce a global normalized distribution is a nontrivial problem. The difficulty lies in the fact that there can only be N_r normalizing factors but there are many more neighboring domain pairs, each of which can itself determine a normalization factor whenever the unnormalized distributions are not exact (which is the case for distributions obtained by the Monte Carlo method). Thus the matching defines a minimization problem that can be formalized as follows.

Let $P_i(j,k)$ be the unnormalized probability distribution value for run i at the kth gridpoint of the domain i such that it overlaps with the domain of the run j. We assume here that the P_i are determined numerically on a grid. In case it is obtained in an analytical form, the sum over k in the equations below should be replaced by an integral over the overlapping regions of the domains

(44) M. Rao, C. S. Pangali, and B. Berne, *Mol. Phys.*, **37**, 1773 (1979).
(45) M. Mezei, S. Swaminathan, and D. L. Beveridge, *J. Chem. Phys.*, **71**, 3366 (1979).

of the individual runs. Let c_i be the normalization constant for run i. When there are only two domains to match, the matching involves finding a c such that

$$\sum_k w(1,2,k)*[P_1(2,k) - c*P_2(1,k)]^2 \rightarrow \text{minimum} \quad \text{(A1)}$$

where $w(1,2,k)$ is a weighting factor that should be designed to take into account the accuracy of the $P_1(2,k)$ and $P_2(1,k)$ values to ensure that the more accurate data will be dominant in determining the matching. In eq A1 and below, the sum over k is to be taken over all gridpoints that were sampled in both runs i and j. When there are more than two domains, then the scaling factors c_i should be determined by minimizing

$$\sum_{i=1}^{N_r} \sum_{j \in O(i)} \sum_k w(i,j,k)*[c_i*P_i(j,k) - c_j P_j(i,k)]^2 \quad \text{(A2)}$$

where $O(i)$ is the set of runs such that their domains overlap with

the domain of i. Taking the partial derivative of (A2) with respect to each c_i, we obtain N_r linear equations:

$$c_i^* \sum_{j \in O(i)} \sum_k P_i(j,k)^2 - \sum_{j \in O(i)} c_j \sum_k P_i(j,k)*P_j(i,k) = 0 \quad \text{(A3)}$$

Clearly, the equation system is homogeneous. As a result, it does not determine the solution fully. The remaining freedom in c_i, however, is needed at the end to determine the overall normalization factor that normalizes the global P to unity.

The necessity of solving a minimization problem introduces an added problem. There will be no unique value for the regions where two domains overlap. Furthermore, because of the many restrictions, there may even be significant breaks when the global P switches from one domain to another. This might necessitate the use of some smoothing, or restrict the use of the matching for more accurate distributions.

Metal–Metal Bonding and Antibonding Molecular Orbitals in Triangular Skeletons of the Type L₆M₃

Carlo Mealli

Contribution from the Istituto per lo Studio della Stereochimica ed Energetica dei Composti di Coordinazione, C.N.R., Via F. D. Guerrazzi 27, 50132 Florence, Italy. Received July 23, 1984

Abstract: The stabilization of triangular transition-metal frameworks of the type L₆M₃ is analyzed in terms of qualitative perturbation theory arguments and isolobal analogy concepts. Essentially each metal of the M₃ triangle is required to have two orbitals, one each of σ and π symmetry, which nest as three bonding (two tangential and one radial) and three antibonding (one tangential and two radial) MOs. Six electrons in the former and none in the latter set confer maximum stability to the system by analogy to the known electronic features of cyclopropane. The required energy level ordering is not attainable with models having three terminal L₂M fragments, unless supported by the presence of other coligands. The reason for this failure is that each L₂M fragment has a σ hybrid which is too destabilized, hence the intermetal radial bonding combination of these hybrids cannot descend below the tangential antibonding combination of π orbitals. Conversely, L₆M₃ clusters are stabilized when three of the six ligands function as bridges between metals. This structure, referred to as "unsupported-bridged", may be considered as formed by three planar L₃M fragments, and this allows the correct order of the direct bonding and antibonding M–M orbitals. Strategies for the extraction of the latter from the group of M–L bonding-antibonding orbitals are indicated in the presence of either π-acceptor or π-donor bridging ligands. In some instances one of the three M–M antibonding orbitals is also low enough in energy to be populated. Usually this lowers the calculated M–M bond order which is consistent with the experimental long M–M distances. Examples of "supported-unbridged" geometry are also known. The supporting ligands are capping one or both faces of the M₃ triangle formed by three terminal L₂M groupings. In other cases bridging hydrogen atoms provide the required support. It is shown that the main action of the supporting ligands is that of enriching the sp character of a particular radial bonding combination of metal d orbitals which thus stabilizes below a competing tangential antibonding M–M combination. The present analysis gives new weight to the contribution of metal d orbitals in the M₃ bonding network and allows the identification of the corresponding MOs.

It has been proposed that D_{3h} trinuclear metal clusters are the inorganic equivalent of cyclopropane.[1] The bonding network of cyclopropane can be constructed by means of two localized hybrids at each carbon atom containing one electron each, as shown in **1**.[2] In terms of MO theory one σ and one π orbital at each center
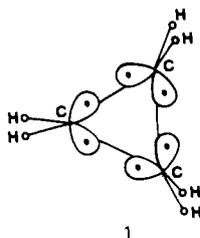
1

concur to form three radial (a′₁ and e′ᵣ) and three tangential (a′₂ and e′ₜ) combinations. Three C–C single bonds are then formed since the bonding combinations a′₁ and e′ₜ are fully occupied, as shown in **2**.

Similar schemes may be constructed for triangular metal clusters provided that each metal fragment has suitable σ and π orbitals. There is sufficient evidence that the L₆M₃ skeleton can exist by itself only if three L ligands function as bridges between two metals in what can be referred to as "unsupported-bridged" geometry. In other cases the required bonding-antibonding arrangement of the levels is attained because of the presence of other coligands, either capping one or both faces of the M₃ triangle or bridging three L₂M fragments approximately in the molecular

(1) Hoffmann, R. *Angew. Chem., Int. Ed. Engl.* **1982,** *21,* 711–724.

(2) (a) Jorgensen, W. L.; Salem, L. "The Organic Chemist's Book of Orbitals"; Academic Press: New York, 1973; pp 19–23. (b) Hoffmann, R. "Special Lectures at the 23rd International Congress of Pure Applied Chemistry"; Butterworths; London, 1971; Vol. 1, p 157.