850 cm^{-1}. Anal. Calcd for $C_6H_{11}O_6P$: C, 34.29; H, 5.28. Found: C, 34.22; H, 5.39.

Registry No. 1a, 61356-97-6; 1b, 77989-23-2; 1c, 77989-24-3; 10, 88229-79-2; 15, 88229-80-5.

References and Notes

(1) For a review, see: Kobayashi, S.; Saegusa, T. *Pure Appl. Chem.* **1981**, *53*, 1663.
(2) Saegusa, T.; Yokoyama, T.; Kobayashi, S. *Macromolecules* **1977**, *10*, 791.
(3) Kobayashi, S.; Hashimoto, T.; Saegusa, T. *Macromolecules* **1980**, *13*, 1650.
(4) Kobayashi, S.; Yokoyama, T.; Saegusa, T. *Polym. Bull.* **1980**, *3*, 505.
(5) Kobayashi, S.; Chow, T. Y.; Saegusa, T. *Macromolecules* **1982**, *15*, 202.
(6) Kobayashi, S.; Huang, M. Y.; Saegusa, T. *Polym. Bull.* **1982**, *6*, 389.
(7) Saegusa, T.; Yokoyama, T.; Kobayashi, S. *Polym. Bull.* **1978**, *1*, 55.
(8) Saegusa, T.; Kimura, Y.; Ishikawa, N.; Kobayashi, S. *Macromolecules* **1976**, *9*, 724.
(9) Kobayashi, S.; Narukawa, Y.; Hashimoto, T.; Saegusa, T. *Chem. Lett.* **1980**, 1599.
(10) Kobayashi, S.; Narukawa, Y.; Shinkai, H.; Saegusa, T. Presented at the 47th Annual Meeting of the Chemical Society, Japan, 3I35, 1983.
(11) Kobayashi, S.; Narukawa, Y.; Mori, S.; Hashimoto, T.; Saegusa, T. *Macromolecules* **1983**, *16*, 858.
(12) (a) Lynen, F. *Ber. Dtsch. Chem. Ges.* **1940**, *73*, 367. (b) Cope, A. C.; Jackson, H. L. U.S. Patent 2748153, 1956; *Chem. Abstr.* **1957**, *51*, 498*f*.
(13) Our preliminary results referred to the polymer structure as **2a**.[1] This was probably because the structure was examined with a polymer sample obtained after long reaction time.
(14) (a) Khan, S. A.; Kirby, A. J. *J. Chem. Soc. B* **1970**, 1172. (b) Cox, J. R., Jr.; Wall, R. E.; Westheimer, F. H. *Chem. Ind. (London)* **1959**, 929.
(15) Saegusa, T.; Kobayashi, S.; Kimura, Y.; Yokoyama, T. *J. Am. Chem. Soc.* **1976**, *98*, 7843.
(16) Gerrard, W.; Richmond, M. J. *J. Chem. Soc.* **1945**, 853.
(17) Lucas, H. J.; Mitchell, F. W., Jr.; Scully, C. N. *J. Am. Chem. Soc.* **1950**, *72*, 5491.
(18) Cox, J. R., Jr.; Westheimer, F. H. *J. Am. Chem. Soc.* **1958**, *80*, 5441.

Conformational Study of the Dinucleotide dGpdCp–Tetrapeptide Ala_4 Complex†

Young Kee Kang‡ and Mu Shik Jhon*

Department of Chemistry, Korea Advanced Institute of Science and Technology, Seoul 131, South Korea. *Received February 23, 1983*

ABSTRACT: Conformational analysis of dinucleotide dGpdCp and tetrapeptide β-alanine (Ala_4), to the complex has been carried out with an empirical potential function by varying all the independent degrees of freedom of the nucleotide and peptide backbones. To understand the factors involved in the change of one conformation to another, we have estimated approximately the configurational entropies of the conformer and water molecules bound to it by using a potential surface analysis method. The total free energy changes for each conformational transition between dGpdCp, Ala_4, and their complex at room temperature were calculated and compared with each other. The free energy change of the free dGpdCp and Ala_4 to the free complex is −70.0 kcal/mol, in which the interaction energy change is −49.5 kcal/mol and the entropy change is 68.9 eu. Through the hydration, there have been large changes of free energies: −664.0 kcal/mol for dGpdCp, −72.0 kcal/mol for Ala_4, and −682.3 kcal/mol for the complex. The entropy changes corresponding to them are −28.1, −48.3, and −178.5 eu, respectively. The free energy change of the hydrated dGpdCp and Ala_4 to the hydrated complex is −16.2 kcal/mol, in which the energy change is −26.1 kcal/mol and the entropy change is −33.2 eu. It is found that this entropy change is mainly caused by the conformational entropy change of dGpdCp and Ala_4 through the complex formation and that the major contribution to the total interaction energy is ascribed to the hydrogen bond between the conformer and water molecules bound to it, whereas the hydration effect of counterions and bound water molecules is proved to play an important role in determining the conformational stabilities of dGpdCp, Ala_4, and their complex in the hydrated state.

I. Introduction

The recognition between nucleic acids and proteins is one of the most fundamental processes and plays a central role in molecular communication in all living cells.[1-4] Examples of such interactions include complexes between histone and DNA, between repressor and DNA, between restriction endonuclease and DNA, and between transfer RNA and cognate synthetase.

The "structural recognition" between nucleic acids and proteins can be classified into two categories: nonspecific interaction and specific interaction. In the former case, protein recognizes some common structural features of nucleic acids and in the latter case, protein recognizes a particular base sequence of nucleic acids. Although ample experimental results[5-9] and model approaches[10-14] are available for these interactions, we know very little about the detailed mechanism of such interactions at the atomic or structural level. Carter and Kraut[10] proposed a model for a double-stranded RNA (ds RNA) and an antiparallel two-stranded β-sheet (β-ribbon), in which the 2′-hydroxyl of the ribose in RNA forms a hydrogen bond to the free carbonyl oxygen of the peptide backbone and the free NH group forms two hydrogen bonds through a water molecule with the ring oxygen and the 2′-hydroxyl oxygen of the next nucleotide on the narrow groove of ds RNA. They also pointed out that there is no room in the antiparallel β-ribbon for residues other than those with very small side-chain groups since the narrow groove of ds RNA is too shallow and ruled out the possibility of a similar model for ds DNA because DNA lacks a 2′-hydroxyl group and

† From the thesis submitted by Young Kee Kang in fulfillment of the requirements for the degree of Doctor of Philosophy, Korea Advanced Institute of Science and Technology, Seoul, South Korea, 1983.

‡ Present address: Department of Chemistry, College of Natural Sciences, Chungbuk National University, Cheongju 310, South Korea.

0024-9297/84/2217-0138$01.50/0

the minor groove of ds DNA in the B form is narrower than that of ds RNA. Kim et al.,[12] however, proposed a model for structural recognition between ds DNA and an antiparallel two-stranded β-sheet (β-ribbon), in which not only all the symmetry elements (two kind of pseudo twofold axes) of DNA and β-ribbon coincide as in the model for ds RNA but also the channel formed between the β-ribbon and the minor groove is large enough to allow the minor groove of ds DNA to be recognized by various amino acid side chains on the inner surface of the antiparallel β-ribbon. The unique portion of the complex is pointed out to be composed of one nucleotide and two peptides, in which the 3′-oxygens of the DNA backbone form hydrogen bonds to every alternating amido NH group of the peptide backbone and alternating carbonyl oxygens of the peptide backbone are involved in hydrogen bonds to the opposite strand of the antiparallel β-ribbon or with nearby water molecules or amino acid side chains.

As a further step in understanding the mechanism of the structural recognition between DNA and protein as well as the participation of water molecules in this interaction, a study has been undertaken for a model complex between a single-stranded dinucleotide dGpdCp and a single-stranded tetrapeptide β-alanine (Ala$_4$) by calculating the conformational energy using an empirical potential function. To recognize the factors involved in the stability of a given conformation and in the change of one conformation to another, it is necessary to evaluate their configurational entropy. There has been considerable effort in developing methods for its evaluation.[15-18] Recently, Karplus et al.[19] proposed a method of calculation of configurational entropy using a normalized multivariate Gaussian distribution function. They expressed the configurational entropy for each conformation in terms of the covariance matrix obtained from the harmonic analysis results.[20] However, it is not applicable to biomolecules such as DNAs or proteins since it is not easy to obtain their harmonic analysis results at the local minima. Hence an approximate method for the evaluation of the covariance matrix with the analysis of potential surface is proposed and extended to determine the configurational entropies of the biomolecule as well as the water molecules bound to it.

II. Methods and Definitions

A. Potential Functions and Parameters. Conformational energy calculations were carried out with the empirical potential energy functions and energy parameters described by Kang and Jhon.[21-24,26] The total conformational energy E_{tot} is the sum of the electrostatic energy E_{el}, polarization energy E_{pol}, nonbonded energy E_{nb}, and torsional energy E_{tor}. The hydrogen bond energy E_{hb} is also included explicitly in the total energy. The total energy is given by the expression

$$E_{tot} = \sum_{\text{nb pairs}} (E_{el} + E_{pol} + E_{nb}) + \sum_{\text{torsions}} E_{tor} + \sum_{\text{hb pairs}} E_{hb} \quad (1)$$

The nonbonded energy was approximated by using the familiar Lennard-Jones 6–12 potential, in which the dispersion coefficient was calculated by using the London formula and the equilibrium distance was assigned to be a distance approximately 0.2 Å greater than the sum of the van der Waals radii of the interacting atoms.[25] The atomic static polarizabilities used in the polarization and nonbonded terms were taken from the work of Kang and Jhon.[26] The torsional potential was of the form of the usual periodic function, and especially, for the phosphate group of DNA backbone, two additional potentials due to the anomeric effect of the polar P–O bond and the virtual bond between lone-pair orbitals of phosphodiester oxygens were also considered.[23]

Atomic partial charges used for the calculation of the electrostatic and polarization terms were obtained from the so-called "σ–π method",[18] in which a set of parameters for π-charge of the phosphate group of DNA backbone was refined. The results of ab initio calculation with including d orbitals for dimethyl phosphate[27] were used for the refinement of those parameters. In the case of the free (vacuum) state, a dielectric constant of 1.0 was used. For the hydrated state, a distance-dependent dielectric constant, $\epsilon(r_{ij})$, was used. It was assumed that at the contact distance r_{con}, $\epsilon(r_{ij}) = 1$ and for some particular interaction distance r_{eff}, the dielectric constant $\epsilon(r_{ij}) = \epsilon_0$, the effective bulk dielectric constant.[18] In the region $r_{con} < r_{ij} < r_{eff}$, $\epsilon(r_{ij})$ was considered to be a function of r_{ij} in the form of

$$\epsilon(r_{ij}) = (\epsilon_0 - 1)[(r_{ij} - r_{con})/(r_{eff} - r_{con})]^n + 1 \quad (2)$$

Boundary conditions such as $\epsilon(3\text{ Å}) = 1.0$ and $\epsilon(7\text{ Å}) = 4.0$ were taken from the work of Hopfinger[18] since there are no available experimental data for them. When $n = 1$, eq 2 is identical with the equation of Hopfinger. $n = 2$ is optimally chosen to yield good results of hydration energies of K^+ and Na^+ in agreement with experimental data.[28] Potential parameters of hydrogen bond were obtained optimally fitted to the results of ab initio computations. The detailed formulation and potential parameters used here are presented elsewhere.[21-24,26]

B. Conformational Entropy. 1. Conformational Entropy of Biomolecules. Because of the large number of atoms present in many biomolecules of interest, the set of torsion angles is commonly chosen as the coordinate set to be included in evaluation of the configurational integral.[15-17] Although a normal-mode analysis or a molecular dynamics simulation of the local fluctuations on the torsional motions of biomolecules may be feasible in principle, it is not applicable in practice because of the magnitude of the system. Karplus et al.[19] pointed out from the conformational study of decaglycine that the only internal coordinates that may be safely neglected are the bond lengths and that all other degrees of freedom contribute significantly to the entropy; however, our choice of the torsion angles for the set of internal coordinates for the nucleotide–peptide complex may be valid since the conformations of nucleotide and peptide are not changed significantly through the formation of the complex. If the joint probability distribution function for a set of torsion angles satisfies a normalized multivariate Gaussian distribution function, the conformational entropy at the ith minimum is given by[19]

$$S_q^{(i)} = \tfrac{1}{2}nR + \tfrac{1}{2}R \ln [(2\pi)^n \sigma^{(i)}] \quad (3)$$

Here, n is the number of degrees of freedom in a set of torsion angles and $\sigma^{(i)}$ is the determinant of the covariance matrix at the ith minimum. Hence the entroy difference associated with the conformational change of a biomolecule is given by

$$\Delta S_q(i \rightarrow j) = \tfrac{1}{2}R \ln (\sigma^{(j)}/\sigma^{(i)}) \quad (4)$$

where $\sigma^{(i)}$ and $\sigma^{(j)}$ are the determinants of the covariance matrices at the ith and jth minima, respectively.

The elements of the covariance matrix are readily available from numerical simulations such as molecular dynamics or Monte Carlo simulations and from harmonic analysis results. However, these methods are not feasible in evaluating the elements for each conformation at the local energy minima of large biomolecules because of the

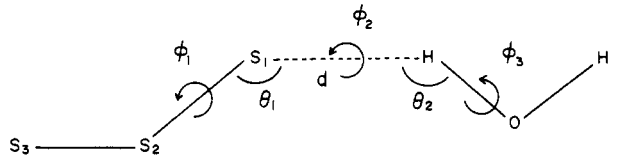

Figure 1. Definition of external variables required to generate the geometry of water molecule, where S_1, S_2, and S_3 are three atoms of the substrate.

magnitude of the system. Thus, we propose in this work an approximate method to obtain the elements by using the potential energy surface. The principal assumption made in the method is that the local fluctuations of biomolecules depend on the torsional motions and that the probability distribution function for each torsion angle satisfies a normal distribution function. Hence the element of the covariance matrix at the ith minimum is[29]

$$\sigma_{kl}^{(i)} = \rho_{kl}^{(i)}\sigma_k^{(i)}\sigma_l^{(i)} \tag{5}$$

where $\sigma_k^{(i)}$ and $\sigma_l^{(i)}$ are standard deviations of the kth and lth torsion angles, respectively, and $\rho_{kl}^{(i)}$ is the correlation coefficient between them. If there are small fluctuations in the torsion angles about the energy minimum, each standard deviation can be easily calculated since the probability distribution function for each torsion angle satisfies a normal distribution function.[29] In addition, the joint probability distribution function for the kth and lth torsion angles satisfies a bivariate Gaussian distribution function, and the correlation coefficient $\rho_{kl}^{(i)}$ can be obtained from the relation[29]

$$q_{k/l}^{(i)} = q_{k0}^{(i)} + \rho_{kl}^{(i)}\sigma_k^{(i)} \tag{6}$$

where $q_{k/l}^{(i)}$ is the conditional mean of the kth torsion angle with the lth torsion angle $q_l^{(i)} = q_{l0}^{(i)} + \sigma_l^{(i)}$ at the ith minimum and $q_{k0}^{(i)}$ is the value of $q_k^{(i)}$ at the energy minimum.

For a stable conformation of a bipolymer chain with small conformational fluctuations in the torsion angles about the energy minimum, Scheraga et al.[17] expressed the conformational entropy as $-^1/_2R \ln F$, where F is the determinant of the matrix of second derivatives at the energy minimum. Hence the conformational entropy change from the ith minimum to the jth one becomes

$$\Delta S(i \rightarrow j) = {}^1/_2R \ln (F^{(i)}/F^{(j)}) \tag{7}$$

where $F^{(i)}$ and $F^{(j)}$ are the determinants of the matrices of second derivatives at the ith and jth energy minima, respectively. However, eq 7 is identical with eq 4 since $1/\sigma^{(i)} = F^{(i)}/RT$, which is derived in Appendix A. Our method may be an alternative to the method to evaluate the second-derivative matrix of the potential energy surface.

2. Configurational Entropy of Bound Water Molecules. The coordinates of a water molecule bound to the biomolecule can be expressed with six external variables (one distance, two bond angles, and three torsion angles) if the coordinates of the three atoms to which a rigid water molecule is attached are known.[30,31] In Figure 1, the external variables are defined. If it is assumed that the joint probability distribution function for the external variables of a water molecule follows the multivariate Gaussian distribution function,[19] the configurational entropy of a water molecule is approximately given by

$$S_w = {}^1/_2nR + {}^1/_2R \ln [(2\pi)^n\sigma_w^*] \tag{8}$$

where n is the external degrees of freedom of the water molecule and σ_w^* is the determinant of the reduced covariance matrix of the water molecule. To make the determinant of the covariance matrix dimensionless, the standard deviation of each external variable must be reduced appropriately.

The potential barriers of the external variables can be approximately classified into the three cases:[32] the very high, low, and intermediate barriers. For very high barriers, the external motion is confined to small harmonic oscillation about the potential minimum. Then the expression for the entropy of the one-dimensional harmonic oscillation is given by[33]

$$S_{ho,i} = R(x_i \exp(-x_i)/(1 - \exp(-x_i))) - R \ln (1 - \exp(-x_i)) \tag{9}$$

where R is the gas constant and x_i is a parameter given by $x_i = h\nu_i/kT$ for the ith external variable. If the probability distribution function satisfies a normal distribution, the force constant can be expressed in terms of its standard deviation, and the parameter x_i becomes

$$x_i = h/(2\pi(\mu kT)^{1/2}\sigma_i) \tag{10}$$

where h is the Planck constant, k the Boltzmann constant, T the absolute temperature, μ the molecular weight (or the moment of inertia) of water molecule, and σ_i the standard deviation of the ith variable. Since the one-dimensional configurational entropy is described by

$$S_i = {}^1/_2R + {}^1/_2R \ln (2\pi\sigma_i^*) \tag{11}$$

the reduced standard deviation σ_i^* for the ith variable can be calculated by equating eq 9 and 11. This approximation is applied to the evaluation of the reduced standard deviations for the distance d, the bond angles ϕ_1 and ϕ_2, and the torsion angle ϕ_1, ϕ_2, or ϕ_3 with the higher potential barrier (the definition of the geometrical variables is shown in Figure 1). In practical calculation, the boundary potential barrier between the harmonic oscillator and the hindered rotation was taken to be 12 kcal/mol.

The torsional motion with the intermediate potential barrier are treated as a hindered rotation. The entropy of the hindered rotation is given by[33]

$$S_{hr,i} = S_{fr,i} + S_{hp,i} - \ln 2\pi \tag{12}$$

where $S_{fr,i}$ and $S_{hp,i}$ are the one-dimensional configurational entropies for the free rotation and the hindered perturbation due to the potential barrier, respectively. By equating eq 11 and 12, one can obtain σ_i^*. The simple expression for the entropy of the free rotation is given by[33]

$$S_{fr,i} = R({}^1/_2 \ln I_i + {}^1/_2 \ln T - \ln n_s - 0.522) \tag{13}$$

where I_i is the moment of inertia, T the absolute temperature, and n_s the symmetry number. The entropy associated with the hindered perturbation is written in terms of σ_i by[21]

$$S_{hp,i} = R\left[\int_0^b t^2 \exp(-t^2)\,dt \Big/ \int_0^b \exp(-t^2)\,dt + \ln\left(\int_0^b \exp(-t^2)\,dt\right) + \ln (8^{1/2}n_s\sigma_i)\right] \tag{14}$$

where

$$b = \pi/(2^{1/2}n_s\sigma_i) \tag{15}$$

The detailed derivation of eq 14 is given in Appendix B and the integrations in eq 14 are performed numerically. This approximation is applied to the evaluation of the reduced standard deviations for the torsion angles ϕ_1, ϕ_2, and ϕ_3 with the intermediate potential barriers. For the case of the low potential barrier $V_i < RT/2$, an additional approximation is also considered in Appendix B.

C. Model Compounds. As a model compound of the DNA–protein complex, the complex of a single-stranded dinucleotide dGpdCp and a single-stranded tetrapeptide

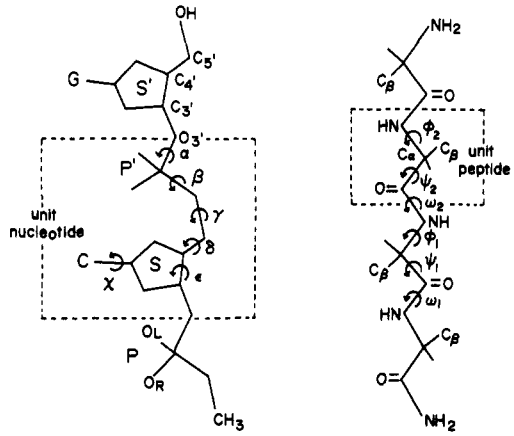

Figure 2. Nomenclature of torsion angles and atoms for nucleotide and peptide backbones.

β-alanine (Ala$_4$) was taken, where each end of the complex was terminated with hydrogen atom. Although it was found that there are two ways in which the anntiparallel β-ribbon can be fitted into the narrow groove of ds DNA parallel or antiparallel to each other,[12] the antiparallel type was ruled out and only the parallel type of the complex was considered in this work since the latter has more favorable stereochemistry under the constraint of keeping DNA as close to the B form as possible and since the total interaction energy of the antiparallel type was found to have a very large positive value.[21] To consider simply the effects of the side chain of the peptide backbone in the formation of the complex, β-alanine was chosen. In addition, for the hydrated dGpdCp, Na^+ counterions were placed along the phosphate anion axis of the nucleotide backbone. The notation for torsion angles of the model compound is illustrated in Figure 2.

D. Procedures of Computation. The total conformational energy change for each given conformation of dGpdCp, Ala$_4$, or their complex was calculated as the sum

$$\Delta E_{tot} = \Delta E_{con} + \Delta E_{int} \quad (16)$$

where ΔE_{con} is the internal conformational energy change of dGpdCp or Ala$_4$, and ΔE_{int} is the intermolecular interaction energy.

The torsion angles of the backbones of dGpdCp and Ala$_4$ were allowed to move during minimization by using a quasi-Newton method developed by Fletcher[34] with a convergence criterion of 0.005 kcal/mol and with a step length of 2° for all the torsion angles. The number of iterations was limited to 100 cycles. In order to generate the coordinates of the complex, 14 and 12 independent torsion angles are needed for dinucleotide and tetrapeptide, respectively. However, they were reduced to 7 (α, β, γ, δ, ϵ, ζ, and χ) and 6 (ϕ_1, ψ_1, ω_1, ϕ_2, ψ_2, and ω_2) variables during the energy minimization. For example, a torsion angle ($C_{5'}$-$C_{4'}$-$C_{3'}$-$O_{3'}$) of a sugar S' follows that of a sugar S, ϵ, etc. (see Figure 2 for the notation of torsion angles and atoms). Crystallographic results of DNA[35] and protein[36] and a refined model of their complex[12] were used as the preliminary geometries of minimization. The starting points for minimization of the total energy of each conformation were then determined by the refinement of those geometries varying each torsion angle of the backbones one after another. In the case of the complex, the geometry of the hydrogen bond was initially taken from that of the model building[12] and then the torsion angles of each conformer were allowed to move during the energy minimization.

For the hydrated systems, the hydration schemes were obtained by the optimization of water molecules bound to the systems,[30,31] and then the total interaction energies of the conformers were minimized by varying all the torsion angles once more. The numbers of water molecules bound to each nucleotide and peptide were initially decided from the results of infrared spectra of Falk et al.[37] and from the theoretical study of Scheraga et al.,[38] respectively. This procedure was repeated until the difference of each torsion angle between the previous and later iterations remains within 1°. Each water molecule bound to the system was optimized by varying the six external variables defined in Figure 1 by the Fletcher procedure,[34] in which the optimization step lengths were 0.01 Å for distance and 2° for all the angles. the maximum number of iterations was 60 cycles and a convergence criterion was 0.005 kcal/mol. In addition, Na^+ counterions placed along the phosphate anion axis of the nucleotide backbone were allowed to move during the energy minimization.

The total entropy change of each conformer through the hydration was estimated as the sum

$$\Delta S_{tot} = \Delta S_{con} + \Delta S_w \quad (17)$$

where ΔS_{con} is the conformational entropy change of the conformer given by eq 4 and ΔS_w is the configurational entropy change of water molecules bound to the conformer. It is assumed that ΔS_w is the difference between the entropy of liquid water at room temperature and the configurational entropy of the optimized water molecules. Hence ΔS_w is given by

$$\Delta S_w = S_{sw} - nS_w^\circ \quad (18)$$

where S_{sw} is the sum of the configurational entropy S_w given by eq 8, n the number of bound water molecules, and S_w° the entropy of liquid water at room temperature (i.e., 16.72 eu at 298.15 K).[39] The Helmholtz free energy change ΔA at each energy minimum is given by $\Delta A = \Delta E_{tot} - T\Delta S_{tot}$ at a temperature T. Since these calculations were carried out at constant pressure and volume, the Gibbs free energy change ΔG is approximately equal to ΔA.

III. Results and Discussion

A. Minimum-Energy Conformations. The optimized torsion angles for the model compounds of dGpdCp, Ala$_4$, and their complex are listed in Table I. The notation for torsion angles of dGpdCp and Ala$_4$ is illustrated in Figure 2. Three kinds of torsion angles for each model compound are compared with each other: the first column corresponds to the calculated results with a dielectric constant $\epsilon = 1.0$, the second to those of the hydrated compound with a distance-dependent dielectric constant described previously, and the third to the available experimental values. From the calculated results it is found that the stereochemistry of dGpdCp, Ala$_4$, or their complex with considering the hydration effect is more close to that of crystallographic results.

B. Dinucleotide dGpdCp. To investigate the relative stability of dGpdCp during the formation of the complex, it is necessary to calculate the internal conformational energy of dGpdCp itself as well as the configurational entropy of dGpdCp and water molecules bound to it. Since the water molecules in the first hydration shell are mainly responsible for the hydration structure of solute,[40,41] only the directly bound water molecules to dinucleotide dGpdCp were considered in the calculation, whereas 23 water molecules were optimized around dGpdCp, which is in good agreement with the experimental[37] and theoretical[42] results. Four water molecules are bound to each Na^+ counterion placed at the phosphate anions, five to guanine base, and three to cytosine base. The remaining water molecules are attached to the sugar and backbone

Table I
Backbone Torsion Angles in Degrees in an Asymmetric Unit

angle	Ala_4			dGpdCp			complex		
Peptide									
ϕ_1	−150.3[a]	−144.1[b]	−138.6[c]				−158.9[a]	−160.0[b]	−152.6[e]
ψ_1	157.3	162.8	134.7				142.6	142.2	130.5
ω_1	−179.0	−173.5	−178.5				−182.4	−185.5	−175.5
ϕ_2	−149.2	−142.6	−138.6				−142.0	−142.0	−144.3
ψ_2	157.5	145.7	134.7				179.5	181.6	181.5
ω_2	−180.0	−178.0	−178.5				170.0	168.6	175.8
Nucleotide									
α				−94.6[a]	−91.6[b]	−95.2[d]	−97.4	−94.4	−85.0
β				−57.1	−56.1	−46.9	−48.9	−47.9	−47.3
γ				−149.5	−148.5	−146.0	−149.5	−148.5	−143.3
δ				40.4	38.4	36.4	29.0	29.0	17.3
ϵ				149.5	146.5	156.4	149.8	149.8	152.7
ζ				163.7	163.7	155.0	169.9	169.9	165.0
χ				79.5	80.5	82.3	74.6	76.6	85.3

[a] Calculated values with a dielectric constant $\epsilon = 1$. [b] Calculated values with a distance-dependent dielectric constant due to the hydration. [c] X-ray crystallographic data from ref 36. [d] X-ray crystallographic data from ref 35. [e] Model building data from ref 12.

Table II
Optimum Geometry, Interaction Energy, and Configurational Entropy of Water Molecules Bound to Dinucleotide dGpdCp[a]

water molecule	ref atoms[b] S_1	S_2	S_3	optimum geometry d	θ_1	θ_2	ϕ_1	ϕ_2	ϕ_3	ΔE	S_w	ΔA
W1	$Na(P')O_L(P')O_R(P')$			3.05	76	27	261	133	220	−12.25	10.71	−15.44
W2	$Na(P')O_L(P')O_R(P')$			2.77	145	155	27	116	161	−12.67	16.67	−17.64
W3	$Na(P')O_L(P')O_R(P')$			2.81	62	48	190	207	130	−8.21	16.99	−13.28
W4	$Na(P')O_L(P')O_R(P')$			2.95	92	26	94	276	157	−10.38	15.93	−15.13
W5	$O_{1'}(S')C_{4'}(S')C_{3'}(S')$			2.00	129	176	203	223	14	−15.67	21.96	−22.22
W6	$O_{1'}(S')C_{4'}(S')C_{3'}(S')$			2.05	93	178	103	272	272	−14.78	23.99	−21.93
W7	$O_{3'}(P')P'(P')O_{5'}(P')$			1.95	100	161	139	99	194	−14.07	16.67	−19.04
W8	$O_R(P')P'(P')O_{5'}(P')$			1.48	208	178	214	214	206	−24.88	10.36	−27.97
W9	$N_7(G)\,C_8(G)\,N_9(G)$			1.97	98	186	166	190	0	−8.64	19.65	−14.50
W10	$O_6(G)\,C_6(G)\,C_5(G)$			1.72	126	175	191	344	168	−10.31	20.86	−16.53
W11	$H_2(G)\,N_2(G)\,C_2(G)$			2.47	176	23	242	161	145	−10.13	9.39	−12.93
W12	$H_{2'}(G)\,N_2(G)\,C_2(G)$			2.20	121	48	267	206	204	−6.42	9.76	−9.33
W13	$N_3(G)\,C_2(G)\,N_1(G)$			1.93	111	183	169	282	180	−7.98	10.59	−11.14
W14	$Na(P)O_L(P)O_R(P)$			3.03	79	26	259	91	240	−15.55	9.07	−18.25
W15	$Na(P)O_L(P)O_R(P)$			2.76	133	109	16	100	222	−8.67	16.02	−13.45
W16	$Na(P)O_L(P)O_R(P)$			2.81	62	47	171	203	111	−11.68	12.47	−15.40
W17	$Na(P)O_L(P)O_R(P)$			3.13	81	23	91	193	125	−16.26	12.27	−19.92
W18	$O_{1'}(S)C_{4'}(S)C_{3'}(S)$			2.16	84	146	92	314	357	−20.76	18.06	−26.14
W19	$O_{3'}(P)P(P)O_{5'}(P)$			1.95	107	165	134	104	204	−12.76	20.12	−18.76
W20	$O_L(P)P(P)O_{5'}(P)$			1.51	127	181	316	27	209	−15.69	12.16	−19.32
W21	$H_4(C)N_4(C)C_4(C)$			2.39	160	34	193	170	177	−8.46	14.24	−12.71
W22	$H_{4'}(C)N_4(C)C_4(C)$			2.23	163	40	1	170	221	−5.87	14.88	−10.31
W23	$N_3(C)C_4(C)C_5(C)$			2.06	141	186	145	6	179	−7.98	13.09	−11.88

[a] Definition of reference atoms and geometry is given in Figure 1; ΔE, S_w, and ΔA are interaction energy (kcal/mol), configurational entropy (eu), and Helmholtz free energy (kcal/mol) at room temperature, respectively. [b] Each atom is designated with parentheses to which it belongs.

of the nucleotide. The optimized structure of the hydrated dGpdCp is shown in Figure 3. It is found that there are a few hydrogen bonds between the bound water molecules: W1–W7, W2–W8, W11–W23, W12–W18, and W17–W15. The average distance between Na^+ counterion and water molecules coordinated to it is 2.49 Å, which is in good agreement with the average cation–oxygen bond lengths observed in crystals containing polycoordinated sodium cation.[43-45]

The optimized geometry, interaction energy, configurational entropy, and free energy of each water molecule bound to dGpdCp are shown in Table II. The interaction energies of water molecules bound to Na^+ counterions lie within −8.2 to −16.3 kcal/mol, those to phosphate and sugar groups are −12.8 to −24.9 kcal/mol, and those to bases amount to −5.9 to −10.3 kcal/mol. The configurational entropies of water molecules are 9.1–17.0, 10.4–24.0, and 9.4–20.9 eu for those groups, respectively. Hence it is found from the comparison of free energies with each other that the water molecules are more strongly bound to phosphate and sugar groups than the other water molecules and their values lie within −18.8 to −28.0 kcal/mol. The free energies of water molecules on Na^+ counterions and bases are −13.3 to −19.9 and −9.3 to −16.5 kcal/mol, respectively.

The interaction energies for the hydrated dGpdCp are summarized in Table III. The internal conformational energy change of dGpdCp through the hydration is about −19.4 kcal/mol, to which the electrostatic energy is the major contribution of energy. The total value of intermolecular interaction energy is −653.0 kcal/mol, to which the polarization and hydrogen bond energies contribute dominantly. Hence it is clear that the total energy change of dGpdCp due to the hydration amounts −672.4 kcal/mol and that the hydration effect including counterions exerts an important effect upon the relative stability of dGpdCp. The configurational entropy change of dGpdCp through the hydration of 10.5 eu (see Table IV), which is due to

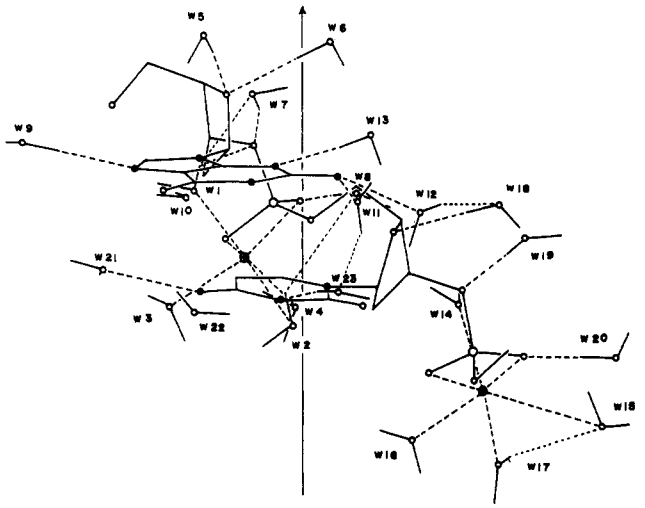

Figure 3. Structure of the hydrated dGpdCp along the helical axis. For clarity, hydrogen atoms are not shown. Hydrogen bonds and coordination of water molecules to counterion are shown in broken lines. Each atom is designated as follows: (•) carbon; (o) oxygen; (●) nitrogen; (O) phosphorus; (⊕) sodium.

Table III
Interaction Energies of Hydrated Dinucleotide dGpdCp and Tetrapeptide Ala$_4$ (kcal/mol)[a]

components of interaction energies	dGpdCp		Ala$_4$	
ΔE_{el}	−26.08	−5.88	−4.51	180.89
ΔE_{pol}	7.72	−146.56	0.09	−13.41
ΔE_{nb}	−1.37	43.34	−0.32	−3.92
ΔE_{tor}	0.35		1.71	
ΔE_{hb}		−543.92		−264.89
ΔE_{tot}	−19.38	−653.02	−3.03	−83.83

[a] Values in the first and second columns are the internal conformational energy changes and the intramolecular interaction energies of dGpdCp and Ala$_4$ through the hydration, respectively.

Table IV
Conformational Entropy Change of dGpdCp, Ala$_4$, and Their Complex (eu)[a]

state	dGpdCp	Ala$_4$
free complex	30.60	38.28
hydrated	10.48	−6.58
hydrated complex	−1.84	−41.26

[a] Values of conformational entropy changes of dGpdCp and Ala$_4$ are the relative values to those of the isolated dGpdCp and Ala$_4$ in the free state, respectively.

the increase of correlation coefficients between phosphate–backbone torsion angles of dGpdCp even though there is the reduction of standard deviation for each torsion angle through the hydration.

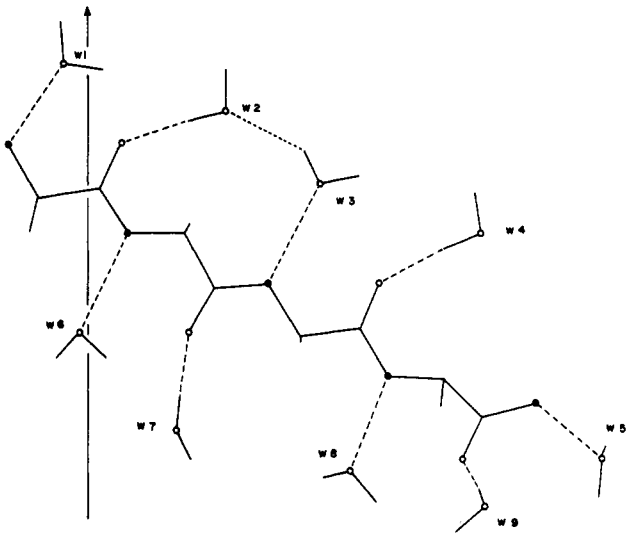

Figure 4. Structure of the hydrated Ala$_4$ along the z axis of rectangular coordinates.

By using the values in Table II and the experimental data,[39] the total entropy change of water molecules at room temperature can be approximately calculated as −38.6 eu. Hence the total entropy change and the total free energy change through the hydration of dGpdCp are about −28.1 eu and −664.0 kcal/mol, respectively, and the relative contribution of entropy to free energy change amounts to about 1.3%.

C. Tetrapeptide Ala$_4$. For tetrapeptide β-alanine in the free and hydrated states the internal conformational energy and the configurational entropy were also calculated. For the hydrated Ala$_4$, nine water molecules were optimized around the conformer. Their geometries, interaction energies, and configurational entropies are listed in Table V. In Figure 4, the optimum structure of Ala$_4$ and water molecules directly bound to it is shown. It is clear that each water molecule is attached to each amino or carbonyl group and there is one water molecule hydrogen bonded to an adjacent water molecule.

The interaction energies of water molecules bound to NH groups lie within −11.0 to −14.6 kcal/mol and those to carbonyl groups within −6.6 to −9.0 kcal/mol. The larger value of those to NH groups is due to the interaction with adjacent carbonyl group. Especially, water molecule W3 has the largest interaction energy through the hydrogen bond to an adjacent water molecule W2. The configurational entropies of water molecules bound to NH group lie within 4.2–10.2 eu and those to carbonyl group within 14.5–19.9 eu. It is found that the free energies of water molecules on Ala$_4$ are within −12.5 to −16.1 kcal/mol and that water molecules are attached to amido groups stronger

Table V
Optimum Geometry, Interaction Energy, and Conformational Entropy of Water Molecules Bound to Tetrapeptide Ala$_4$[a]

water molecule	ref atoms S_1 S_2 S_3	optimum geometry d	θ_1	θ_2	ϕ_1	ϕ_2	ϕ_3	ΔE	S_w	ΔA
W1	H_1 N_1 C_1^α	2.42	164	26	182	185	197	−11.59	8.54	−14.14
W2	O_1 C_1' N_2	1.77	136	177	21	357	181	−8.75	19.08	−14.44
W3	H_3 N_3 C_3^α	2.35	153	36	141	178	164	−14.63	4.87	−16.08
W4	O_3 C_3' N_4	1.77	140	178	22	350	182	−9.04	14.46	−13.35
W5	H_5' N_5 H_5	2.21	158	40	146	171	224	−10.96	10.23	−14.01
W6	H_2 N_2 C_2^α	2.46	164	27	229	188	225	−12.05	4.20	−13.30
W7	O_2 C_2' N_3	1.79	139	182	321	294	178	−7.07	18.58	−12.61
W8	H_4 N_4 C_4^α	2.36	164	33	182	191	212	−12.14	8.96	−14.81
W9	O_4 C_4' N_5	1.76	115	177	323	13	170	−6.61	19.82	−12.52

[a] Refer to footnotes in Table II.

Table VI
Interaction Energies of Complex between Dinucleotide dGpdCp and Tetrapeptide Ala_4 (kcal/mol)[a]

components of interaction energies	free complex			hydrated complex		
ΔE_{el}	2.14	1.53	-12.35	-22.81	-2.58	100.48
ΔE_{pol}	0.18	-0.60	-19.05	8.08	-0.76	-175.54
ΔE_{nb}	-4.58	0.27	0.87	-3.30	0.38	47.13
ΔE_{tor}	3.48	4.32		2.84	5.18	
ΔE_{hb}			-25.66			-744.04
ΔE_{tot}	1.22	5.52	-56.19	-15.19	2.22	-771.97

[a] Values in the first and the second columns correspond to the internal conformational energy changes of dGpdCp and Ala_4 relative to the isolated dGpdCp and Ala_4 in the free state, respectively; values in the third column correspond to the intramolecular interaction energy of complex between dGpdCp and Ala_4.

than to carbonyl groups by about 2–3 kcal/mol.

The optimized geometries of Ala_4 in the free and hydrated states are compared with the experimental value[36] in Table I. The structure of hydrated Ala_4 is known to be closer to that of experimental results. The interaction energy changes for Ala_4 through the hydration are also summarized in Table III. It is seen that the internal conformational energy change of Ala_4 is -3.0 kcal/mol and that the intermolecular interaction energy of the hydrated Ala_4 is -83.4 kcal/mol. The total energy change of Ala_4 through the hydration is thus -86.4 kcal/mol, in which the interaction energy between β-alanine and water molecules bound to it is the decisive factor. There is a decrease of configurational entropy about 6.6 eu for Ala_4 through the hydration (see Table IV). The entropy change of water molecules amounts to -41.7 eu. Thus, total entropy change due to the hydration is approximately -48.3 eu and the total free energy change at room temperature is about -72.0 kcal/mol. This value corresponds to about one-ninth of that for the hydrated dGpdCp.

D. Complex of dGpdCp and Ala_4. To understand the factors involved in stabilizing the dGpdCp–Ala_4 complex, the studies of the formation of the complex are carried out by considering the hydration effect.

The optimized geometry of the complex is listed in Table I. The hydrogen bond lengths of the complex are 3.13 and 3.17 Å for the free complex and 3.05 and 3.17 Å for the hydrated complex. There is no large structural difference between the two complexes, but the stereochemistry of the hydrated complex is closer to that of the model building study. The major differences between the isolated dGpdCp and the complex are in conformational angles β and δ of the nucleotide backbone, but those between Ala_4 and the complex are in all the torsion angles.

In case of the free complex, the optimized structure is shown in Figure 5 and the interaction energy changes are summarized in Table VI. The internal conformational energy changes are 1.2 and 5.5 kcal/mol for dGpdCp and Ala_4 due to the formation of the complex from the isolated dGpdCp and Ala_4, respectively. Although there is a decrease of internal energy through the formation of the complex, the total interaction energy change is -49.5 kcal/mol, in which the interaction energy between dGpdCp and Ala_4 is -56.2 kcal/mol. It is clear from analyzing the interaction energy that the major contribution to the relative stability of the complex comes from the hydrogen bond and the minor contribution is the nonbonded interaction. There is a configurational entropy change of 68.9 eu upon the complex formation (refer to the values of configurational entropy changes of conformers shown in Table IV), which is caused by the increase of correlation coefficients between backbone torsion angles of dGpdCp and Ala_4 while there is the reduction of mean fluctuation for each torsion angle through the formation of the complex. Hence the configurational free energy change of the

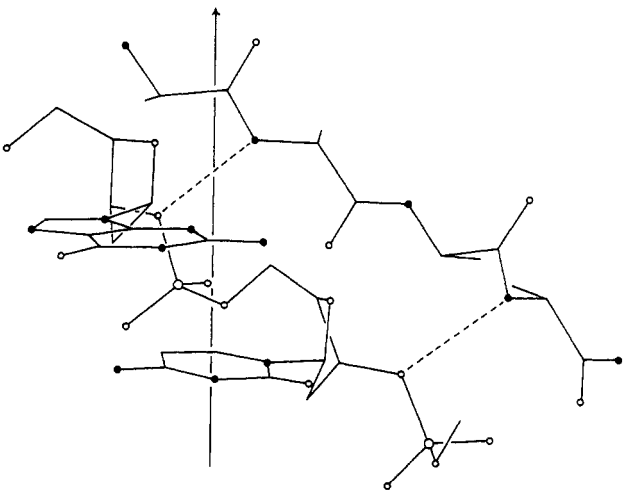

Figure 5. Structure of the free complex between dGpdCp and Ala_4 along the helical axis of dGpdCp. Hydrogen bonds are shown in broken lines.

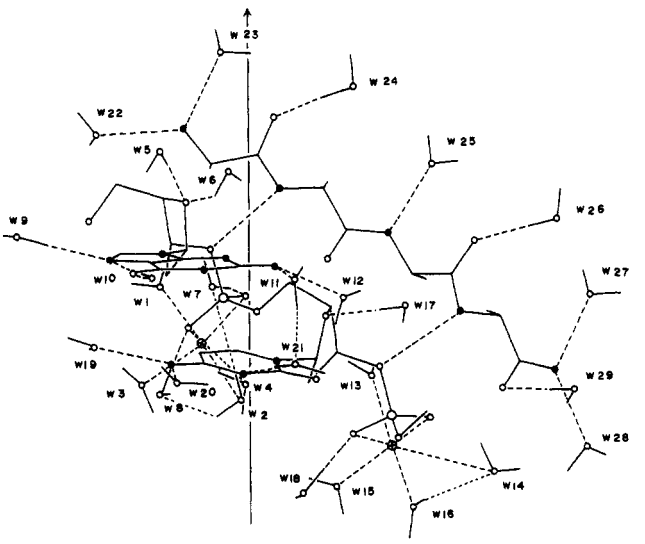

Figure 6. Structure of the hydrated complex between dGpdCp and Ala_4 along the helical axis of dGpdCp. Hydrogen bonds and coordination of water molecules to counterion are shown in broken lines. The other notation refers to the caption of Figure 3.

complex from the free dGpdCp and Ala_4 is -70.0 kcal/mol and the contribution of configurational entropy to it corresponds to about -20 kcal/mol.

In Figure 6, the structure of the hydrated complex is shown, in which the total 29 water molecules were optimized around the complex. The results of optimized geometry, interaction energy, and configurational entropy of each water molecule are also listed in Table VII. The major difference of hydration scheme between the complex and the isolated molecules comes from the complex for-

Table VII
Optimum Geometry, Interaction Energy, and Configurational Entropy of Water Molecules Bound to Complex between Dinucleotide dGpdCp and Tetrapeptide Ala_4[a]

water molecule	ref atoms S_1	S_2	S_3	optimum geometry d	θ_1	θ_2	ϕ_1	ϕ_2	ϕ_3	ΔE	S_w	ΔA
W1	$Na(P')$	$O_L(P')$	$O_R(P')$	3.04	75	31	268	138	247	−6.45	15.02	−10.93
W2	$Na(P')$	$O_L(P')$	$O_R(P')$	2.90	144	102	29	125	158	−13.85	16.31	−18.71
W3	$Na(P')$	$O_L(P')$	$O_R(P')$	2.77	57	50	194	201	123	−8.83	13.60	−12.88
W4	$Na(P')$	$O_L(P')$	$O_R(P')$	3.11	78	16	100	291	153	−16.26	10.59	−19.42
W5	$O_{1'}(S')$	$C_{4'}(S')$	$C_{3'}(S')$	1.95	134	177	208	222	14	−14.04	21.76	−20.53
W6	$O_{1'}(S')$	$C_{4'}(S')$	$C_{3'}(S')$	1.91	114	183	102	300	271	−12.78	15.43	−17.38
W7	$O_{3'}(P')$	$P'(P')$	$O_{5'}(P')$	1.47	170	178	276	158	213	−23.44	8.25	−25.90
W8	$O_L(P')$	$P'(P')$	$O_{3'}(P')$	1.46	145	177	242	182	213	−20.55	13.56	−24.59
W9	$N_7(G)$	$C_5(G)$	$N_9(G)$	1.97	100	186	170	196	0	−7.71	20.12	−13.71
W10	$O_6(G)$	$C_6(G)$	$C_5(G)$	1.77	116	173	197	345	163	−10.98	16.91	−16.02
W11	$H_2(G)$	$N_2(G)$	$C_2(G)$	2.48	175	24	246	161	145	−11.01	8.79	−13.63
W12	$H_{2'}(G)$	$N_2(G)$	$C_2(G)$	2.05	148	47	325	173	189	−12.12	16.77	−17.12
W13	$Na(P)$	$O_L(P)$	$O_R(P)$	3.07	79	27	264	93	240	−22.61	8.40	−25.11
W14	$Na(P)$	$O_L(P)$	$O_R(P)$	2.86	114	109	17	110	212	−10.14	16.89	−15.18
W15	$Na(P)$	$O_L(P)$	$O_R(P)$	2.81	62	49	167	199	108	−11.56	15.53	−16.19
W16	$Na(P)$	$O_L(P)$	$O_R(P)$	3.04	73	31	86	187	106	−16.22	12.33	−19.90
W17	$O_{1'}(S)$	$C_{4'}(S)$	$C_{3'}(S)$	2.07	89	182	140	121	1	−14.45	13.67	−18.53
W18	$O_R(P)$	$P(P)$	$O_{3'}(P)$	1.46	139	177	238	181	212	−23.43	12.66	−27.20
W19	$H_4(C)$	$N_4(C)$	$C_4(C)$	2.37	158	36	95	188	233	−10.45	10.84	−13.68
W20	$H_{4'}(C)$	$N_4(C)$	$C_4(C)$	2.26	156	39	323	170	226	−5.66	15.72	−10.35
W21	$N_3(C)$	$C_4(C)$	$C_5(C)$	1.99	140	186	147	7	178	−6.86	13.09	−10.76
W22	$H_{1'}$	N_1	C_1^{α}	2.38	156	36	207	152	237	−10.64	14.59	−14.99
W23	H_1	N_1	C_1^{α}	2.42	162	26	213	169	190	−12.50	11.35	−15.88
W24	O_1	$C_{1'}$	N_2	1.76	135	181	57	308	182	−7.96	19.18	−13.68
W25	H_3	N_3	C_3^{α}	2.24	157	35	166	173	161	−11.69	13.27	−15.65
W26	O_3	$C_{3'}$	N_4	1.76	139	175	37	334	129	−7.23	19.31	−12.99
W27	H_5	N_5	$H_{5'}$	2.27	157	43	86	114	223	−10.15	12.97	−14.02
W28	$H_{5'}$	N_5	H_5	2.24	141	41	162	167	230	−10.16	14.65	−14.53
W29	O_4	$C_{4'}$	N_5	1.88	97	173	297	39	177	−7.50	16.70	−12.48

[a] Refer to footnotes in Table II.

mation, in which water molecules bound to the vicinity of oxygen atoms of nucleotide backbone and of amido or carbonyl group of β-alanine are squeezed out and the 3′-oxygens of the nucleotide backbone form the hydrogen bonds to every alternating amido NH group of the peptide backbone, whereas the major difference between the hydrated complex and the hydrated dGpdCp and Ala_4 comes from the distribution change of water molecules bound to regions between dGpdCp and Ala_4. There are three hydrogen bonds between bound water molecules: W2–W7, W11–W21, and W16–W14.

Table VI contains the interaction energy of the hydrated complex between dGpdCp and Ala_4. The internal conformational energy changes of the complex through the hydration are −16.4 and −3.3 kcal/mol for dGpdCp and Ala_4, respectively. Since the intermolecular interaction energies of the complexes in the free and hydrated states amount to −56.2 and −772.0 kcal/mol, respectively, the energy change of intermolecular interaction due to the hydration is −715.8 kcal/mol, to which the major contribution of energy comes from the hydrogen bond and polarization energies. It is clear that Na^+ counterion plays an important role in stabilizing the conformation of the hydrated complex as well as the hydrated dGpdCp. These results are in good agreement with the experimental results of Aslanian et al.[46] Although the interaction between dGpdCp and Ala_4 is very essential in the formation of the complex, its contribution to the total interaction is found to be of less consequence. The total interaction energy change of the complex through the hydration is thus −735.5 kcal/mol.

In Table IV, the configurational entropy changes of dGpdCp and Ala_4 in the hydrated complex are shown and the values of those for dGpdCp and Ala_4 are −1.8 and −41.3 eu, respectively. By using data in Table IV, the configurational entropy change of the complex through the hydration can be estimated and amounts to −32.4 and −79.5 eu for dGpdCp and Ala_4, respectively. The configurational entropy change of water molecules can be calculated by using eq 18, and its value corresponds to −66.5 eu. Thus the total free energy change of the free complex between dGpdCp and Ala_4 through the hydration at room temperature is −682.3 kcal/mol and the entropic contribution to the total free energy is about 53 kcal/mol.

Although there are extensive experimental results[6] that the formation of the complex between DNA and protein is very favorable, there is no experimental evidence as yet to support the model implying the binding of a β-sheet polypeptide structure to the minor groove of DNA. To investigate this kind of interaction, the free energy change of the hydrated dGpdCp and Ala_4 to their hydrated complex is estimated. The interaction energy change through the formation of the hydrated complex from the hydrated dGpdCp and Ala_4 is −26.1 kcal/mol, which can be calculated by using the values in Tables III and VI. This value is about the same magnitude as that of the contribution of the dGpdCp–Ala_4 interaction to total interaction energy of the hydrated complex.

The conformational entropy changes of the hydrated dGpdCp and Ala_4 to the hydrated complex are −12.3 and −34.7 eu, respectively (see Table IV). The entropy change due to water molecules is 13.8 eu. Hence the total entropy change is −33.2 eu and the total free energy change at room temperature is −16.2 kcal/mol. It is clear that the formation of the complex with the hydration is favorable and the entropy change is mainly caused by the conformational entropy change of dGpdCp and Ala_4 upon the complex formation. If the entropy change due to the structural change of water molecules outside the directly bound water molecules during the formation of the complex, more exact value of free energy change may be obtained. The overall free energy changes for the various conformational tran-

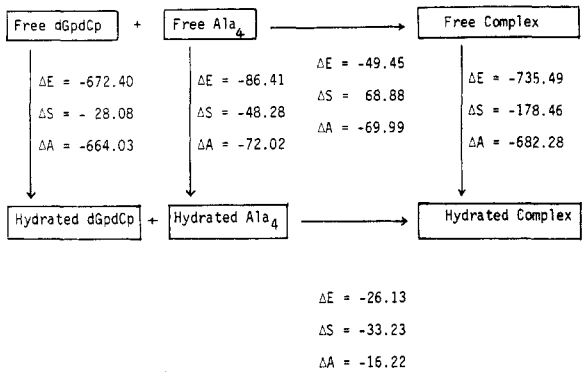

Figure 7. Overall free energy changes among dinucleotide dGpdCp, tetrapeptide Ala_4, and their complex in the free and hydrated states; ΔE, ΔS, and ΔA are the changes of the interaction energy (kcal/mol), the conformational entropy (eu), and the Helmholtz free energy (kcal/mol) at room temperature, respectively.

sitions among dGpdCp, Ala_4, and their complex in the free and hydrated states are shown in Figure 7.

IV. Conclusions

The conformational analysis of dinucleotide dGpdCp and tetrapeptide β-alanine to their complex has been carried out by using an empirical potential function varying all the independent degrees of freedom of the nucleotide and peptide backbones. To recognize the factors involved in the stability of a given conformation and in the change of the conformation to another, an approximate method based on the small fluctuation of torsional motion is suggested to estimate the confugurational entropy of the conformer and water molecules bound to it.

For dinucleotide dGpdCp, tetrapeptide β-alanine, and their complex in the free and hydrated states, the interaction energies and the configurational entropies are calculated. The conformational changes of dGpdCp and Ala_4 due to the complex formation are discussed with the total free energy change and with the distribution of water molecules around the conformers. It is found that the entropy change is mainly caused by the configurational entropy change of dGpdCp and Ala_4 upon the complex formation and that the interaction energy change comes from the interaction between dGpdCp and Ala_4 but the major contribution to the total interaction energy is the hydrogen bond energy due to the hydration. From analyzing the structural data of torsion angles, it is clear that there are not large differences between the hydrated and free complex although the complex formation yields a somewhat large conformational change of the nucleotide and peptide backbones. However, the water structure induced in the vicinity of phosphate group of the nucleotide and of amido or carbonyl group of β-alanine is significantly changed through the formation of hydrogen bond between nucleotide and peptide backbones. This result is in good agreement with experimental result.

Although further study on the evaluation of entropy change due to the configurational changes of biomolecules as well as water molecules outside the directly bound water molecules during the conformational change of biomolecules is needed, we obtained somewhat qualitative results upon the complex formation between dinucleotide dGpdCp and tetrapeptide Ala_4.

Acknowledgment. We are grateful to Professor Sung-Hou Kim at the University of California, Berkeley, for his kind discussion on the structure of complex between nucleic acid and protein. This work is supported in part by the Korea Science and Engineering Foundation.

Appendix A. Derivation of $1/\sigma = F/RT$

The ratio of statistical weight at $\mathbf{q}$ and $\mathbf{q}_0$, where $\mathbf{q} = (q_1, q_2, ..., q_n)$ and $\mathbf{q}_0 = (q_{10}, q_{20}, ..., q_{n0})$, is

$$P(\mathbf{q})/P(\mathbf{q}_0) = \exp(-(V(\mathbf{q}) - V(\mathbf{q}_0))/RT) \quad \text{(A-1)}$$

For a stable conformation with small conformational fluctuations about the energy minimum

$$V(\mathbf{q}) = V(\mathbf{q}_0) + \tfrac{1}{2}(\mathbf{q} - \mathbf{q}_0)'\mathbf{F}(\mathbf{q} - \mathbf{q}_0) \quad \text{(A-2)}$$

Hence the ratio $P(\mathbf{q})/P(\mathbf{q}_0)$ becomes

$$P(\mathbf{q})/P(\mathbf{q}_0) = \exp(-\tfrac{1}{2}(RT)^{-1}(\mathbf{q} - \mathbf{q}_0)'\mathbf{F}(\mathbf{q} - \mathbf{q}_0)) \quad \text{(A-3)}$$

If the probability distribution function $P(\mathbf{q})$ is a normalized multivariate Gaussian distribution function, the ratio $P(\mathbf{q})/P(\mathbf{q}_0)$ becomes

$$P(\mathbf{q})/P(\mathbf{q}_0) = \exp(-\tfrac{1}{2}(\mathbf{q} - \mathbf{q}_0)'\sigma^{-1}(\mathbf{q} - \mathbf{q}_0)) \quad \text{(A-4)}$$

From eq A-3 and A-4, we obtain the relation

$$\sigma^{-1} = (RT)^{-1}\mathbf{F} \quad \text{(A-5)}$$

or

$$1/\sigma = F/RT \quad \text{(A-6)}$$

where F and σ are the determinants of matrices $\mathbf{F}$ and σ, respectively.

Appendix B. Derivation of Eq 14

The configurational integral for a hindered rotation associated with a torsion angle ϕ is given by

$$Q_{hr} = \tfrac{1}{2}\pi q_{fr} q_{hp} \quad \text{(B-1)}$$

where

$$q_{fr} = (8\pi^2 I_\phi kT/h^2)^{1/2} \quad \text{(B-2)}$$

and

$$q_{hp} = \int_0^{2\pi} \exp(-V_\phi/RT)\, d\phi \quad \text{(B-3)}$$

If the vibrational motion of a torsion angle ϕ is assumed to be governed by a harmonic motion, the potential function V_ϕ becomes

$$V_\phi = V_\phi^0 + \tfrac{1}{2} f_\phi(\phi - \phi_0)^2 \quad \text{(B-4)}$$

where f_ϕ is a force constant of torsion angle ϕ. The configurational integral q_{hp} becomes

$$q_{hp} = \exp(-V_\phi^0/RT)G(T) \quad \text{(B-5)}$$

where

$$G(T) = \int_{-\pi}^{\pi} \exp(-f_\phi^2/2RT)\, d\phi \quad \text{(B-6)}$$

The function $G(T)$ by the transformation of variable becomes

$$G(T) = 2^{3/2} n_s \sigma_\phi \int_0^b \exp(-t^2)\, dt \quad \text{(B-7)}$$

with

$$t = (f_\phi/2RT)^{1/2} \quad \text{(B-8)}$$

$$f_\phi = RT/\sigma_\phi^2 \quad \text{(B-9)}$$

and

$$b = \pi/(2^{1/2}\sigma_\phi n_s) \quad \text{(B-10)}$$

Hence the final expression for the entropy due to the

hindered perturbation is given by

$$S_{hp} = R\left[\int_0^b t^2 \exp(-t^2)\, dt / \int_0^b \exp(-t^2)\, dt + \ln\left(\int_0^b \exp(-t^2)\, dt\right) + \ln\left(8^{1/2} n_s \sigma_\phi\right)\right] \quad \text{(B-11)}$$

For the case of $V_\phi < RT/2$, the following approximation is adopted:

$$G(T) = 2\int_0^\pi \exp(-f_\phi^2/2RT)\, d\phi \simeq \int_0^\alpha \exp(-\beta\phi)\, d\phi \quad \text{(B-12)}$$

with

$$\alpha = \phi_{max} - \phi_0 \quad \text{(B-13)}$$

$$\beta = V_\phi^0/\alpha \quad \text{(B-14)}$$

where ϕ_{max} and ϕ_0 are the values of ϕ at the maximum and minimum energies, respectively.

Hence the function $G(T)$ becomes

$$G(T) = 2(1 - \exp(-\beta\alpha))/\beta \quad \text{(B-15)}$$

and the entropy due to the hindered perturbation with the very low potential barrier can be approximately given by

$$S_{hp} = 2RTF(\alpha)/\beta + R \ln (2/\beta) + R \ln F(\alpha) \quad \text{(B-16)}$$

with

$$F(\alpha) = 1 - \exp(-\beta\alpha) \quad \text{(B-17)}$$

Registry No. (Ala_4), 926-79-4; dGpdCp, 6818-28-6; (Ala_4)·dGpdCp complex, 88229-00-9; water, 7732-18-5.

References and Notes

(1) J. J. Champoux, *Annu. Rev. Biochem.*, **47**, 449 (1978).
(2) P. Modrich, *Q. Rev. Biophys.*, **12**, 315 (1979).
(3) K. Geider and H. Hoffmann-Berling, *Annu. Rev. Biochem.*, **50**, 233 (1981).
(4) M. T. Record, Jr., S. J. Mazur, P. Melançon, J.-H. Roe, S. L. Shaner, and L. Unger, *Annu. Rev. Biochem.*, **50**, 997 (1981).
(5) N. V. Beaudette and N. Langerman, *CRC Crit. Rev. Biochem.*, **9**, 145 (1980).
(6) C. Helene and J.-C. Maurizot, *CRC Crit. Rev. Biochem.*, **10**, 213 (1981).
(7) R. V. Hosur, N. V. Kumar, and G. Govil, *Int. J. Quantum Chem.*, **20**, 23 (1981).
(8) R. W. Warrant and S. H. Kim, *Nature (London)*, **271**, 130 (1978).
(9) T. S. Young, S. H. Kim, P. Modrich, A. Beth, and E. Jay, *J. Mol. Biol.*, **145**, 607 (1981).
(10) C. W. Carter, Jr., and J. Kraut, *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 283 (1974).
(11) N. C. Seeman, J. M. Rosenberg, and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.*, **73**, 804 (1976).
(12) G. M. Church, J. L. Sussman, and S. H. Kim, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 1458 (1977).
(13) R. Day, F. Martino, and J. Ladik, *J. Theor. Biol.*, **84**, 651 (1980).
(14) R. V. Hosur and A. Pohorille, *Int. J. Quantum Chem.*, **20**, 33 (1981)..
(15) P. J. Flory, "Statistical Mechanics of Chain Molecules", Wiley, New York, 1969.
(16) N. Gō and H. A. Scheraga, *J. Chem. Phys.*, **51**, 4751 (1969).
(17) N. Gō and H. A. Scheraga, *Macromolecules*, **9**, 535 (1976).
(18) A. J. Hopfinger, "Conformational Properties of Macromolecules", Academic Press, New York, 1973.
(19) M. Karplus and J. N. Kushick, *Macromolecules*, **14**, 325 (1981).
(20) R. M. Levy and M. Karplus, *Biopolymers*, **18**, 2465 (1979).
(21) Y. K. Kang, Ph. D., Thesis, Korea Advanced Institute of Science and Technology, Seoul, Korea, 1983.
(22) Y. J. Park, Y. K. Kang, B. J. Yoon, and M. S. Jhon, *Bull. Korean Chem. Soc.*, **3**, 50 (1982).
(23) C. N. Yoon, Y. K. Kang, and M. S. Jhon, *Biopolymers*, in press.
(24) M. J. Moon, Y. K. Kang, and M. S. Jhon, *Bull. Korean Chem. Soc.*, **4**, 133 (1983).
(25) D. A. Brant, *Annu. Rev. Biophys. Bioeng.*, **1**, 369 (1972).
(26) Y. K. Kang and M. S. Jhon, *Theor. Chim. Acta*, **61**, 41 (1982).
(27) D. Perahia, A. Pullman, and H. Berthod, *Theor. Chim. Acta*, **40**, 47 (1975).
(28) I. Džidič and P. Kebarle, *J. Phys. Chem.*, **74**, 1466 (1970).
(29) R. V. Hogg and A. T. Craig, "Introduction to Mathematical Statistics", 4th ed., Macmillan, New York, 1978.
(30) Y. K. Kang and M. S. Jhon, *Bull. Korean Chem. Soc.*, **2**, 24 (1981).
(31) D. Perahia, M. S. Jhon, and B. Pullman, *Biochim. Biophys. Acta*, **474**, 349 (1977).
(32) W. J. Orville-Thomas, Ed., "Internal Rotation in Molecules", Wiley, New York, 1974.
(33) J. H. Knox, "Molecular Thermodynamics" (revised edition), Wiley, New York, 1978.
(34) R. Fletcher, "Fortran Subroutines for Minimization by Quasi-Newton Methods", AERE Report R7125, 1972.
(35) S. Arnott and D. W. L. Hukins, *Biochem. Biophys. Res. Commun.*, **47**, 1504 (1972).
(36) S. Arnott, S. D. Dover, and A. Elliott, *J. Mol. Biol.*, **30**, 201 (1967).
(37) M. Falk, K. A. Hartman, Jr., and R. C. Lord, *J. Am. Chem. Soc.*, **84**, 3843 (1962); **85**, 387, 391 (1963).
(38) Y. Paterson, G. Némethy, and H. A. Scheraga, *Ann. N.Y. Acad. Sci.*, **367**, 132 (1981).
(39) R. C. Weast, Ed., "Handbook of Chemistry and Physics", 56th ed., CRC Press, Cleveland, 1975.
(40) H. L. Friedman and C. Y. Krishnan, "Water: A Comprehensive Treatise", F. Franks, Ed., Plenum Press, New York, 1973, Vol. 3, p 1.
(41) A. Pullman, "The New World of Quantum Chemistry", B. Pullman and R. Parr, Eds., Reidel, Dordrecht, 1976, p 149.
(42) E. Clementi and G. Corongiu, *Ann. N.Y. Acad. Sci.*, **367**, 83 (1981).
(43) N. C. Seeman, J. M. Rosenberg, F. L. Suddath, J. J. P. Kim, and A. Rich, *J. Mol. Biol.*, **104**, 109 (1976).
(44) J. M. Rosenberg, N. C. Seeman, R. O. Day, and A. Rich, *J. Mol. Biol.*, **104**, 145 (1976).
(45) N. Camerman and J. K. Fawcett, *J. Mol. Biol.*, **107**, 601 (1976).
(46) V. M. Aslanian, Yu. S. Babaian, and P. P. Yasem, *Stud. Biophys.*, **87**, 101 (1982).