(14) P. Andrews, *Biochemistry*, **91**, 222 (1964).
(15) U. K. Laemmli, *Nature (London)*, **227**, 680 (1970).
(16) S. D. Putney and P. R. Schimmel, in preparation.
(17) G. Davies and G. Stark, *Proc. Natl. Acad. Sci. U.S.A.*, **66**, 651 (1970).
(18) M. R. Kula, *FEBS Lett.*, **35**, 299 (1973).
(19) G. L. E. Koch, Y. Boulanger, and B. S. Hartley, *Nature (London)*, **249**, 316 (1974).
(20) R. M. Waterson and W. H. Konigsberg, *Proc. Natl. Acad. Sci. U.S.A.*, **71**, 376 (1974).
(21) C. J. Bruton, R. Jakes, and G. L. E. Koch, *FEBS Lett.*, **45**, 26 (1974).
(22) S. Robbe-Saul, F. Fasiolo, and Y. Boulanger, *FEBS Lett.*, **84**, 57 (1977).
(23) D. Nathans and H. O. Smith, *Annu. Rev. Biochem.*, **44**, 273 (1975).
(24) F. Bolivar, R. L. Rodriques, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow, *Gene*, **2**, 95 (1977).
(25) J. G. Sutcliffe, *Nucleic Acids Res.*, **5**, 2721 (1978).
(26) F. Sanger, S. Nicklen, and A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 5463 (1977).
(27) A. M. Maxam and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 560 (1970).
(28) K. Biemann, "Biochemical Application of Mass Spectrometry", Suppl. I, G. R. Waller, Ed., Wiley, New York, in press.
(29) T. Hashimoto and M. Sekiguchi, *J. Bacteriol.*, **127**, 1561 (1976).
(30) R. J. Roberts, *Gene*, **4**, 183 (1978).
(31) F. C. Neidhardt, J. Parker, and W. G. McKeever, *Annu. Rev. Microbiol.*, **29**, 215 (1975).
(32) J. E. Brenchley and L. S. Williams, *Annu. Rev. Microbiol.*, **29**, 251 (1975).
(33) P. R. Schimmel, *Acc. Chem. Res.*, **10**, 411 (1977).
(34) A. Rich and P. R. Schimmel, *Nucleic Acids Res.*, **4**, 1649 (1977).
(35) D. L. Meléndez and P. R. Schimmel, unpublished observation.
(36) D. Shortle and D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 2170 (1978).
(37) H. Hayatsu, *Prog. Nucleic Acid. Res. Mol. Biol.*, **16**, 75 (1976).
(38) P. J. Flory, "Statistical Mechanics of Chain Molecules", Chapter VII, Wiley, New York, 1969.
(39) H. A. Scheraga, *Pure Appl. Chem.*, **50**, 315 (1978).
(40) For a summary of recent literature see P. Y. Chou and G. D. Fasman, *Annu. Rev. Biochem.*, **47**, 251 (1978).
(41) J. H. Miller, C. Coulondre, M. Hofer, U. Schmeissner, H. Sommer, A. Schmitz, and P. Lu, *J. Mol. Biol.*, **131**, 191 (1979).
(42) J. H. Miller and U. Schmeissner, *J. Mol. Biol.*, **131**, 232 (1977).
(43) J. H. Miller, *J. Mol. Biol.*, **131**, 249 (1977).

Configurational Statistics of Polynucleotide Chains. An Updated Virtual Bond Model to Treat Effects of Base Stacking*

Wilma K. Olson

Department of Chemistry, Rutgers University, New Brunswick, New Jersey 08903.
Received February 13, 1980

ABSTRACT: An updated virtual bond scheme has been developed to include "long-range" effects of base stacking in the treatment of the spatial configurations of the polynucleotides. As a consequence of the relative rigidity of rotations about the C–O bonds (ϕ' and ϕ) of the sugar–phosphate backbone, it is possible to represent the six chemical bonds constituting each nucleotide repeating unit in terms of two virtual bonds of comparable magnitudes (spanning C–C–O–P bond fragments of the chain). The "long-range" (three-bond) interdependence of the alternate C–C and O–P bonds in the polynucleotide backbone somewhat complicates computations of chain averages compared to our previous treatments. Despite this complexity, the "long-range" interactions introduce a novel theoretical probe to deduce the solution conformation of the phosphodiester linkages ($\omega'\omega$) from the experimentally observed conformations of the C–C rotations ($\psi'\psi$). On the basis of hard-core conformational analysis and Karplus treatment of NMR coupling constants, we have modeled the helix-to-coil transition of poly(rA) over the temperature range −12 to 60 °C. At low temperatures the molecule is a flexible helix similar in conformational detail to the A-RNA family of structures. At higher temperatures as the nucleotide residues fluctuate over 12 distinct conformational domains, chain dimensions decrease in accordance with experimental data. In contrast to the conventional concept of unstacked, randomly coiling polynucleotides, these computations describe poly(rA) at high temperatures as a highly stacked system with 5 out of 12 of the nucleotides in such arrangements.

Recent applications of polymer chain statistics to the polynucleotides have provided a useful framework for relating subtle features of chemical architecture to the macroscopic properties of this system. As a consequence of geometric constraints imposed by the very complicated skeletal structure (cf. Figure 1), it is possible to represent the nucleotide repeating units comprising six chemical bonds by imaginary virtual bonds that span structural segments of fixed conformation. The virtual bond scheme first offered to treat the unperturbed dimensions of randomly coiling polynucleotides entails two such segments, one spanning four chemical bonds and the other two.[1] A later model applied to flexible nucleic acid helices involves a single virtual bond repeating unit that connects successive phosphorus atoms of the chain.[2-4] In both schemes the mutual orientation of any pair of successive virtual bonds is independent of the orientation of all other such pairs; statistical mechanical analyses of the spatial configurations of the polynucleotide as a whole are consequently facilitated.

Until now, the design of virtual bonds in polynucleotides and other biopolymers[5-8] has stemmed from the *local* conformational preferences of all pairs of *successive* rotation angles in the chain. In the two-virtual-bond scheme for polynucleotide random coils, the parameters originally determined to be flexible are the $\omega'\omega$ and $\phi\psi$ angle pairs (cf. Figure 1).[1,9] More recent studies,[10,14] however, suggest that ϕ is a "rigid" parameter confined almost exclusively to trans or t conformations and that only the ω', ω, ψ, and ψ' angles vary in polynucleotides. Because furanose puckering (or ψ') is fixed within the longer of the two

*In honor of P. J. Flory, a dear friend and teacher.

0024-9297/80/2213-0721$01.00/0

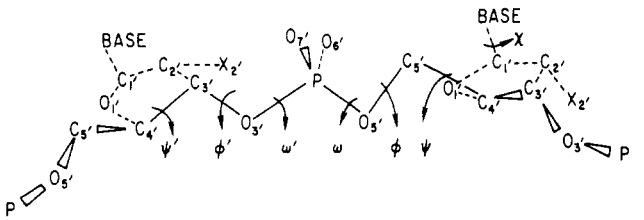

Figure 1. Section of an extended polynucleotide chain showing chain atoms and rotation angles. In the poly(rA) system treated here, $X_{2'}$ is hydroxyl and BASE is adenine.

virtual bonds, the unperturbed dimensions of the random coil must be interpreted in terms of a conformational copolymer of differently puckered nucleotides (the so-called $C_{3'}$-endo or ^3E and $C_{2'}$-endo or ^2E units).[9] In the single virtual bond treatment of flexible polynucleotide helices, on the other hand, only the ω' and ω rotations are allowed to vary. Chain dimensions of DNA double helices spanning a range of chain lengths between 2^8 and 2^{13} nucleotide units are reproduced by allowing only minor ($\sim$15°) variations in these two angles from a fixed reference conformation.[4]

Although the heterocyclic base is not itself a part of the skeletal backbone, the conformation of the polynucleotide chain may be markedly affected, under certain conditions, by the interactions of bases attached to adjacent furanose residues. At low temperatures that favor the stacking of adjacent bases, the unperturbed dimensions of the poly(riboadenylic acid) (poly(rA)) chain are increased dramatically compared to those of the largely unstacked system.[15,16] According to various X-ray crystallographic[12-14,17-19] and theoretical[21-26] studies detailed below, such strong interactions between bases appear to introduce "long-range" third-bond rotational correlations that may be important in the randomly coiling as well as in the helical state.

In this paper we develop a new virtual bond scheme that incorporates long-range third-bond effects of base stacking into the polynucleotide chain. The influences of base stacking are notably absent in our earlier studies of polynucleotide random coils based upon sequential conformational analysis of the furanose–phosphate backbone without attached bases.[1,9] Here we represent the six skeletal bonds of the nucleotide repeating unit by two virtual bonds that span the two C_4–C–O–P segments of the chain backbone. These two virtual bonds are similar in appearance to the virtual bonds that span the three chemical bond repeating units of polypeptides[5] and polysaccharides.[6-8] In accordance with the long-range effects, the conformations of the phosphodiester linkage (i.e., angles ω' and ω) depend upon the conformation of the nearest furanose unit (angles ψ' and ψ, respectively). Successive values of ψ and ψ', however, appear to be independent so that the configurational analysis of the polynucleotide as a whole reduces to the analysis of a single-nucleotide two-virtual-bond repeating segment of the chain. As we shall demonstrate, this new approach to rotational interdependence allows us to follow, for the first time, the conformational details of the noncooperative helix-to-coil transition in single-stranded poly(rA). We monitor the conformation of the phosphodiester rotations on the basis of the experimentally observed values of ψ' and ψ at different temperatures.[11] At very low temperatures (−12 °C) we describe poly(rA) as a flexible A-type helix with end-to-end dimensions consistent with those measured by Stannard and Felsenfeld[16] for a (close to) completely stacked system. As the temperature increases, the computed unperturbed dimensions additionally reproduce data found experimentally by Eisenberg and Felsenfeld[15] at 10, 25, and 60 °C for poly(rA) chains subject to different degrees of unstacking.

Virtual Bond Scheme

Long-Range Conformational Effects. Recent quantum mechanical (PCILO)[22,23] and semiempirical energy[24] analyses of various dinucleoside triphosphates support the unusual crystallographic observation in the $(pdApdT)_2$[17,18] molecule that pentose pucker (ψ') is correlated with the orientation of the 3′-phosphate group (ω'). In ^3E sugars where ψ' is gauche$^-$ (g^-) (or $-120 \pm 60°$ defined with respect to the trans = 0° orientation of atoms $C_{5'}$–$C_{4'}$–$C_{3'}$–$O_{3'}$), the ω' angle is found in the gauche$^+$ (g^+) range. In ^2E units where ψ' is t, the ω' parameter is shifted into the t range. In both the computations and the X-ray data, the intervening ϕ' angle is confined to a preferred t rotation. A similar correlation of the ψ' and ω' angles is seen additionally in the various fiber diffraction[19] and theoretical[3,25] models of A (^3E-type sugar puckering) and B (^2E-related sugar puckering) nucleic acid helices and in the composite data[21] of the three recently refined X-ray crystallographic models of yeast t-RNA^{Phe}.[12-14] In these latter X-ray studies, the ϕ' angle varies over a rather broad range of values (predominantly t) with apparently no effect upon the interdependence of ψ' and ω'.

PCILO computations suggest, in addition, a second long-range correlation between the ω and ψ angles of the polynucleotide chain.[22,23] Energetically favored stacked states occur for the combinations $\omega\psi = g^-g^+$, tt, and g^+g^- when the intervening ϕ angle is fixed in its t range. These three different stacking geometries appear as well in the three refined t-RNA models[12-14,21] while both the g^+g^- and tt combinations of ω and ψ occur in models of polynucleotide helices deduced by X-ray fiber diffraction data.[19,20] A low-energy valley connecting the three stacked $\omega\psi$ angle combinations also results from minimization studies of classical potential energy functions in several model dinucleoside monophosphates.[26] In contrast to the PCILO predictions of almost exclusive preference for stacked geometries, the X-ray data[12-14] and classical potential studies[26] predict, as well, the occurrence of several nonstacked $\omega\psi$ angle combinations in polynucleotide systems.

The major uncertainty in the analysis of polynucleotide conformation in solution is the phosphodiester linkage. The various theoretical predictions of potential energy in model phosphodiesters are widely discrepant.[27] Unfortunately, there is no direct and reliable experimental means to monitor the two P–O rotations. The observed ^{31}P chemical shifts are useful thermodynamic measures of helix–coil transitions[28,29] but are not able to describe the blend of random-coil conformations with certainty.[30] On the other hand, the proton coupling constants associated with the ψ' and ψ angles are fairly reliable measures of these two rotations.[31,32] Under the assumption that the above long-range interdependence of the $\psi'\omega'$ and $\omega\psi$ angle pairs holds in both polynucleotide helices and random coils, it is possible to monitor the P–O rotations indirectly from observed ψ' and ψ values and thus to model polynucleotide chains in detail under various experimental conditions.

Rotational Isomeric States. The preferred t conformations about the two C–O bonds of the polynucleotide chain backbone are represented here by the rotational isomeric states $\phi' = 30°$ and $\phi = 0°$. These values are consistent with the relatively narrow ranges of angles observed directly in model systems by NMR[11,31,32] and X-ray crystallographic[10,12-14] analyses and also predicted by various theoretical energy estimates.[9,27,33] The choices of fixed rotational states are both convenient and well

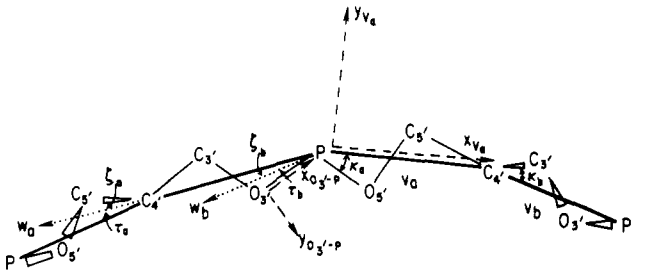

Figure 2. Section of a polynucleotide chain showing virtual bonds (heavy lines) and coordinate axes (dashed lines) of the chain backbone. The parameters w_a, w_b, ζ_a, ζ_b, τ_a, τ_b, κ_a, and κ_b are defined in the text.

founded for bonds such as these having distinct conformational minima separated by energy barriers that substantially exceed RT.

Assignment of ϕ' and ϕ to their rotational isomeric state values fixes the distances that separate $C_{4'}$ atoms from the following, as well as the preceding, P atoms in the chain. It is appropriate then to introduce two virtual bond vectors $\mathbf{v}_a$ and $\mathbf{v}_b$ joining these atoms as shown by the heavy solid lines in Figure 2. The lengths v_a and v_b of the two bonds depend upon the values assigned to the ϕ and ϕ' angles, respectively, as detailed below. With these two rotation angles fixed, the spatial configuration of the polynucleotide chain can be represented by the alternating succession of virtual bond vectors $\mathbf{v}_a$ and $\mathbf{v}_b$. These two bonds may thus replace the six real bonds of the repeat unit. Major simplification of the representation of the spatial configuration of the chain as a whole is achieved in this manner. As outlined below, it is advantageous with respect to both the helical conformations and to the random coil.

The orientation of virtual bond $\mathbf{v}_{ai}$ with respect to $\mathbf{v}_{bi-1}$ and with respect to the structure spanned by $\mathbf{v}_{bi-1}$ is determined by ω' and ω, as is apparent from Figures 1 and 2. Similarly, the orientation of vector $\mathbf{v}_{bi}$ with respect of $\mathbf{v}_{ai}$ and the portion of the chain spanned by $\mathbf{v}_{ai}$ is determined by the ψ and ψ' angles. In accordance with the long-range third-bond effects cited above, the conformations of the $\omega'\omega$ angle pair are dependent upon the values assigned the preceding ψ' and succeeding ψ rotations of the chain sequence. Simultaneous rotations of the $\omega'\omega$ angle pair are also known to exhibit mutual interdependence in that certain conformational combinations of the two angles ($g^{\pm}g^{\mp}$) are sterically forbidden. Successive rotations of the $\psi\psi'$ angle pairs centered at the $C_{4'}$ atoms in the chain, however, are apparently independent of one another. The almost identical conformational behavior of ^{3}E-^{3}E and ^{3}E-^{2}E dimers as well as of ^{2}E-^{2}E and ^{2}E-^{3}E dimers in various theoretical[23-25] and X-ray[17,18,34] studies is suggestive of approximate $\psi\psi'$ rotational independence. On the other hand, the slight correlation between NMR coupling consants that monitor the ψ and ψ' values in a series of ribo- and deoxyribonucleosides[35,36] is indicative of some conformational dependence between these two parameters. It is not clear whether the very small differences (0.4 Hz) in ψ-related coupling constants in the nucleoside series can be attributed to significant conformational variations or whether these are simply the consequence of electronegativity differences[32,36] between ribose (mixed ^{3}E and ^{2}E populations) and deoxyribose (predominantly ^{2}E puckering systems). The absence of the $\psi\psi' = g^+g^-$ conformational sequence in the available X-ray structures of nucleic acid analogues is also suggestive of $\psi\psi'$ rotational interdependence.[37] The $\psi = g^+$ rotamer, however, is highly disfavored by steric[38,39] and anomeric effects[40] regardless of the pucker of the adjacent pentose. The $\psi\psi'$ rotation angle pair is thus approximately independent because consecutive rotations of these two angles do not entail unfavorable second-order interactions in regions of conformation space permitted by first-order interactions of the separate angles. Statistical mechanical treatment of the spatial configuration of the polynucleotide chain is greatly simplified as a consequence of the mutual independence of ψ and ψ'.

Virtual Bond Parameters. The virtual bonds v_a and v_b may be related to the structural parameters of the polynucleotide chain by using methods of matrix algebra. For this purpose, a right-handed Cartesian coordinate system is defined for each chemical bond of the chain as follows. The x axis is chosen parallel to the direction of the chemical bond and the y axis is located in the plane defined by the chemical bond and its predecessor in the chain sequence, as illustrated in Figure 2 for chemical bond $O_{3'}$–P. The direction of the chain is defined from left to right in the figure. The direction of the y axis is chosen to make an acute angle with the preceding bond vector.

A chemical bond vector in the B–C coordinate system may be transformed into the coordinate system of its predecessor A–B upon multiplication with the orthogonal matrix $\mathbf{T}_{AB,BC}(\theta^B,\chi)$. This matrix is a function of θ^B, the supplement to the fixed valence bond angle between the two bonds, and of χ, the angle of rotation about A–B (defined with respect to trans = 0°). The matrix $\mathbf{T}_{AB,BC}$ is simply expressed as the matrix product

$$\mathbf{T}_{AB,BC}(\theta^B,\chi) = \mathbf{X}(\chi - 180°)\mathbf{Z}(-\theta^B) \tag{1}$$

where

$$\mathbf{X}(\chi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\chi & -\sin\chi \\ 0 & \sin\chi & \cos\chi \end{bmatrix} \tag{2}$$

and

$$\mathbf{Z}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

The matrix product $\mathbf{T}_{A-Z}$ of the form

$$\mathbf{T}_{A-Z} = \mathbf{T}_{AB,BC}\mathbf{T}_{BC,CD} \cdots \mathbf{T}_{XY,YZ} \tag{4}$$

then serves to transform a vector in the coordinate system of chemical bond Y–Z into the coordinate system of chemical bond A–B. The vectors $\mathbf{v}_a$ and $\mathbf{v}_b$ spanning the nucleotide repeating units are thus given in the coordinate systems of chemical bonds P–$O_{5'}$ and $C_{4'}$–$C_{3'}$, respectively, by the sums of sketetal bond vectors as

$$\mathbf{v}_a = \mathbf{l}_{P-O_{5'}} + \mathbf{T}_{P-C_{5'}}\mathbf{l}_{O_{5'}-C_{5'}} + \mathbf{T}_{P-C_{4'}}\mathbf{l}_{C_{5'}-C_{4'}} \tag{5}$$

and

$$\mathbf{v}_b = \mathbf{l}_{C_{4'}-C_{3'}} + \mathbf{T}_{C_{4'}-O_{3'}}\mathbf{l}_{C_{3'}-O_{3'}} + \mathbf{T}_{C_{4'}-P}\mathbf{l}_{O_{3'}-P} \tag{6}$$

Each bond vector $\mathbf{l}$ in these expressions is represented in its own coordinate system and the serial products are defined following eq 4. The magnitudes of the virtual bonds are obtained from the scalar products of $\mathbf{v}_a$ or $\mathbf{v}_b$ with itself. Values of the virtual bond lengths as well as values of orientational parameters that relate the virtual bonds to the chemical bonds of the polynucleotide are listed in Table I. In these computations chemical bond lengths and valence angles are maintained at the standard fixed values listed in Table II. As the pentose undergoes pseudorotation, the backbone valence angles at $C_{4'}$ and $C_{3'}$ are expected to fluctuate by approximately 2 and 1°, respectively, from these values (associated with the preferred ^{3}E and ^{2}E ring puckering).

Table I
Geometrical Parameters of the Virtual Bond Scheme

virtual bond	ϕ (or ϕ'), deg	υ, Å	κ, deg	Φ, deg	ζ, deg	τ, deg	Ψ, deg
a	0	3.95	13.5	0.0	90.0	90.0	0.0
b	30	3.88	28.6	21.5	16.0	27.5	42.0

Table II
Standard Structural Parameters of the Polynucleotide Backbone

bond lengths, Å		bond angle and supplement, deg	
C-C	1.52	O-P-O	$76.0 = \theta^{P}$
C-O	1.44	P-O-C	$58.5 = \theta^{O_{5'}} = \theta^{O_{3'}}$
O-P	1.60	O-C-C	$70.0 = \theta^{C_{5'}} = \theta^{C_{3'}}$
		C-C-C	$64.0 = \theta^{C_{4'}}$

Mean-Square Dimensions. The mean-square unperturbed end-to-end distance $\langle r^2\rangle_0$ of a polynucleotide chain comprising x nucleotide repeating units is expressed by the average scalar product

$$\langle r^2\rangle_0 = \sum_{i=1}^{x}\sum_{j=1}^{x}\langle(\mathbf{v}^{T}_{ai} + \mathbf{v}^{T}_{bi})\cdot(\mathbf{v}_{aj} + \mathbf{v}_{bj})\rangle = \sum_{i=1}^{x}\sum_{j=1}^{x}[\langle\mathbf{v}^{T}_{ai}\cdot\mathbf{v}_{aj}\rangle + \langle\mathbf{v}^{T}_{bi}\cdot\mathbf{v}_{bj}\rangle + 2\langle\mathbf{v}^{T}_{ai}\cdot\mathbf{v}_{bj}\rangle + 2\langle\mathbf{v}^{T}_{bi}\cdot\mathbf{v}_{aj}\rangle] \quad (7)$$

where $\mathbf{v}_{aj}$ and $\mathbf{v}_{bj}$ are the column representations of the two virtual bonds, and $\mathbf{v}^{T}_{ai}$ and $\mathbf{v}^{T}_{bi}$ are the transpose or row forms of the two vectors.

Evaluation of the average scalar products appearing in eq 7 requires that all virtual bonds in the polynucleotide chain be expressed in the same coordinate system. For this purpose let a right-handed Cartesian coordinate system be affixed to each virtual bond in the chain with its x axis in the direction of the given bond and its y axis in the plane defined by the virtual bond vector ($\mathbf{v}_a$ or $\mathbf{v}_b$) and the chemical bond vector (P–$O_{5'}$ or $C_{4'}$–$C_{3'}$) that projects from the same atom (P or $C_{4'}$) as the virtual bond. The positive direction of the y axis also makes an acute angle with the preceding chemical bond vector ($O_{3'}$–P or $C_{5'}$–$C_{4'}$). (See Figure 2.)

Angles employed in the various transformations outlined below are denoted in Figure 2. The parameters κ_a and κ_b are the angles formed by the respective virtual bonds with the chemical bond vectors that share the same origin. The angles ζ_a and τ_a are defined with respect to the dashed vector $\mathbf{w}_a$ that is normal to the planes defined by the positive z axes of virtual bond vector $\mathbf{v}_a$ and of the chemical bond vector along $C_{5'}$–$C_{4'}$ (i.e., $\mathbf{w}_a = (\mathbf{z}_{C_{5'}\text{-}C_{4'}} \times \mathbf{z}_{\upsilon_a})$) where the $\mathbf{z}$'s represent unit vectors along the z axes of the given bonds and the × denotes the vector product. The ζ_b and τ_b angles are analogously defined in terms of the hypothetical $\mathbf{w}_b$ vector equal to $\mathbf{z}_{O_{3'}\text{-P}} \times \mathbf{z}_{\upsilon_b}$.

The matrix

$$\mathbf{T}_{P\text{-}O_{5'},\upsilon_a}(\kappa_a,\omega - \Phi_a) = \mathbf{X}(\omega - \Phi_a)\mathbf{Z}(\kappa_a) \quad (8)$$

effects the transformation of a vector from the coordinate system of virtual bond υ_a into the coordinate system of chemical bond P–$O_{5'}$, while the matrix

$$\mathbf{T}_{C_{4'}\text{-}C_{3'},\upsilon_b}(\kappa_b,\psi' - \Phi_b) = \mathbf{X}(\psi' - \Phi_b)\mathbf{Z}(\kappa_b) \quad (9)$$

transforms a vector from the coordinate system of virtual bond υ_b to that of chemical bond $C_{4'}$–$C_{3'}$. The parameters Φ_a and Φ_b are the supplements to angles formed by the positive z axis of the designated virtual bond with the z axis of the chemical bond vector of common origin when the associated rotation angle (ω or ψ', respectively) is equal to 0°. Similarly, the matrix

$$\mathbf{T}_{\upsilon_a,C_{5'}\text{-}C_{4'}}(\zeta_a,\Psi_a,\tau_a) = \mathbf{Z}(\zeta_a)\mathbf{X}(\Psi_a)\mathbf{Z}(-\tau_a) \quad (10)$$

transforms a vector in the $C_{5'}$–$C_{4'}$ coordinate system into that of virtual bond υ_a and the matrix

$$\mathbf{T}_{\upsilon_b,O_{3'}\text{-P}}(\zeta_b,\Psi_b,\tau_b) = \mathbf{Z}(\zeta_b)\mathbf{X}(\Psi_b)\mathbf{Z}(-\tau_b) \quad (11)$$

effects a transformation of a vector in the $O_{3'}$–P coordinate system into that of virtual bond υ_b. The Ψ_a and Ψ_b parameters are the supplements to the angles defined by the z axes of the respective virtual bonds and the designated chemical bonds.

Transformation of a vector in the coordinate system of one virtual bond into that of the preceding virtual bond involves three successive transformations of vectors. The matrix $\mathbf{T}_{ab}$ for the transformation from the coordinate system of vector $\mathbf{v}_b$ to the coordinate system of vector $\mathbf{v}_a$ immediately preceding it is expressed by the product

$$\mathbf{T}_{ab} = \mathbf{T}_{\upsilon_a,C_{5'}\text{-}C_{4'}}(\zeta_a,\Psi_a,\tau_a)\mathbf{T}_{C_{5'}\text{-}C_{4'},C_{4'}\text{-}C_{3'}}(\theta^{C_{4'}},\psi) \times \mathbf{T}_{C_{4'}\text{-}C_{3'},\upsilon_b}(\kappa_b,\psi' - \Phi_b) \quad (12)$$

It is a function of the independent pair of rotation angles ψ and ψ' only. All other quantities occurring in eq 12 are parameters whose values are fixed. The fixed angles, however, vary with the specific values assigned the ϕ and ϕ' rotations.

The matrix representing the transformation from the coordinate system of virtual bond υ_a to that of the preceding virtual bond υ_b is given as the product

$$\mathbf{T}_{ba} = \mathbf{T}_{\upsilon_b,O_{3'}\text{-P}}(\zeta_b,\Psi_b,\tau_b)\mathbf{T}_{O_{3'}\text{-P},P\text{-}O_{5'}}(\theta^{P},\omega')\mathbf{T}_{P\text{-}O_{5'},\upsilon_a}(\kappa_a,\omega - \Phi_a) \quad (13)$$

This matrix depends upon the interdependent $\omega'\omega$ rotation angle pair; all other parameters in the above expression are fixed.

The scalar products appearing in eq 7 can now be formulated in terms of the transformation matrices $\mathbf{T}_{ab}$ and $\mathbf{T}_{ba}$. Separation of terms where $i = j$ together with combinations of terms where $j - i = k$ leads to the expression

$$\langle r^2\rangle_0 = x\upsilon_a^2 + x\upsilon_b^2 + 2\mathbf{v}^{T}_a\langle\mathbf{T}_{ab}\rangle\mathbf{v}_b + 2\sum_{k=1}^{x-1}[\mathbf{v}^{T}_a\langle(\mathbf{T}_{ab}\mathbf{T}_{ba})^k\rangle\mathbf{v}_a + \mathbf{v}^{T}_b\langle(\mathbf{T}_{ba}\mathbf{T}_{ab})^k\rangle\mathbf{v}_b + \mathbf{v}^{T}_a\langle(\mathbf{T}_{ab}\mathbf{T}_{ba})^k\mathbf{T}_{ab}\rangle\mathbf{v}_b + \mathbf{v}^{T}_b\langle(\mathbf{T}_{ba}\mathbf{T}_{ab})^{k-1}\mathbf{T}_{ba}\rangle\mathbf{v}_a](x - k) \quad (14)$$

As a consequence of the independence of successive nucleotide residues at the juncture of the ψ and ψ' rotation angles, the serial products of transformation matrices in eq 14 can be separated into factors, one for each independent unit of the chain. The average products $\langle(\mathbf{T}_{ab}\mathbf{T}_{ba})^k\rangle = \langle\mathbf{T}(\psi)\mathbf{T}(\psi')\mathbf{T}(\omega')\mathbf{T}(\omega))^k\rangle$ and $\langle(\mathbf{T}_{ba}\mathbf{T}_{ab})^k\rangle = \langle(\mathbf{T}(\omega')\mathbf{T}(\omega)\mathbf{T}(\psi)\mathbf{T}(\psi'))^k\rangle$ are thus reexpressed as $\langle\mathbf{T}(\psi)\rangle\langle\mathbf{T}(\psi')\mathbf{T}(\omega')\mathbf{T}(\omega)\mathbf{T}(\psi))\rangle^{k-1}\langle\mathbf{T}(\psi')\mathbf{T}(\omega')\mathbf{T}(\omega)\rangle$ and $\langle\mathbf{T}(\omega')\mathbf{T}(\omega)\mathbf{T}(\psi)\rangle\langle\mathbf{T}(\psi')\mathbf{T}(\omega')\mathbf{T}(\omega)\mathbf{T}(\psi))\rangle^{k-1}\langle\mathbf{T}(\psi')\rangle$, respectively. The expression in eq 14 then takes the form of a finite geometric series in the six averaged matrix products $\langle\mathbf{T}(\psi'\omega'\omega\psi)\rangle$, $\langle\mathbf{T}(\psi'\omega'\omega)\rangle$, $\langle\mathbf{T}(\omega'\omega\psi)\rangle$, $\langle\mathbf{T}(\omega'\omega)\rangle$, $\langle\mathbf{T}(\psi)\rangle$, and $\langle\mathbf{T}(\psi)\rangle$.

The summation in eq 14 for a chain of x repeating units is alternatively expressed in terms of a product of generator matrices

$$\langle r^2\rangle_0 = [1\ \ 0\ \ 0\ \ 0\ \ 0]\mathcal{G}_a(\mathcal{G}_{ab})^{x-1}\mathcal{G}_b\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix} \quad (15)$$

where the 5 × 5 matrices $\mathcal{G}_a$, $\mathcal{G}_{ab}$, and $\mathcal{G}_b$ are given respectively by

$$\mathcal{G}_a = \begin{bmatrix} 1 & \mathbf{v}^T{}_a\langle T(\psi)\rangle & v_a^2/2 \\ 0 & \langle T(\psi)\rangle & \mathbf{v}_a \\ 0 & 0 & 1 \end{bmatrix} \tag{16}$$

$$\mathcal{G}_{ab} = \begin{bmatrix} 1 & \mathbf{v}^T{}_a\langle T(\psi)\rangle + \mathbf{v}^T{}_b\langle T(\omega'\omega\psi)\rangle & (v_a^2+v_b^2)/2 + \mathbf{v}^T{}_b\langle T(\omega'\omega)\rangle \mathbf{v}_a \\ 0 & \langle T(\psi'\omega'\omega\psi)\rangle & \langle T(\psi'\omega'\omega)\rangle \mathbf{v}_a + \langle T(\psi')\rangle \mathbf{v}_b \\ 0 & 0 & 1 \end{bmatrix} \tag{17}$$

$$\mathcal{G}_b = \begin{bmatrix} 1 & \mathbf{v}^T{}_b\langle T(\omega'\omega)\rangle & v_b^2/2 \\ 0 & \langle T(\psi'\omega'\omega)\rangle & \langle T(\psi')\rangle \mathbf{v}_b \\ 0 & 0 & 1 \end{bmatrix} \tag{18}$$

Statistical Weight Matrices. For the purpose of evaluating the averaged transformation matrices in eq 16–18, we introduce three statistical weight matrices $\mathbf{U}_2(\psi',\omega')$, $\mathbf{U}_3(\omega',\omega)$, and $\mathbf{U}_4(\omega,\psi)$ that relate the rotational interdependence of these angle pairs (see below). We also define a premultiplication matrix $\mathbf{U}_1(\psi')$ and a postmultiplication matrix $\mathbf{U}_5(\psi)$ that reflect the rotational preferences of these two independent angles.

With statistical weight matrices $\mathbf{U}_1, \ldots, \mathbf{U}_5$ constructed in this manner, the partition function for each independent nucleotide unit is

$$\mathbf{z} = \mathbf{U}_1\mathbf{U}_2\mathbf{U}_3\mathbf{U}_4\mathbf{U}_5 = \mathbf{U}_1{}^{(5)} \tag{19}$$

where $\mathbf{U}_1{}^{(5)}$ is a short-hand notation for the serial product. The required average transformation matrices are given by

$$\langle \mathbf{T}_i{}^{(i'-i)} \rangle = \mathbf{z}^{-1}(\mathbf{U}_1{}^{(i)} \otimes \mathbf{E}_3)\|\mathbf{T}_i\|[(\mathbf{U} \otimes \mathbf{E}_3)\|\mathbf{T}\|]_{i+1}{}^{(i'-i-1)}(\mathbf{U}_{i'-i}{}^{(5)} \otimes \mathbf{E}_3) \tag{20}$$

in terms of the four transformation matrices $\mathbf{T}_1(\psi')$, $\mathbf{T}_2(\omega')$, $\mathbf{T}_3(\omega)$, and $\mathbf{T}_4(\psi)$ where $1 \le i < i' \le 5$; $\mathbf{E}_3$ is the identity matrix of order 3; ⊗ denotes the direct matrix product; and $\|\mathbf{T}_i\|$ is the pseudodiagonal matrix formed from the $\mathbf{T}_i$ matrices for the rotational isomeric states of bond i in the diagonal array. The subscript appended to the expression in square brackets applies to both U and T matrices enclosed therein.

Hard-Sphere Approximations. Each distinct conformational minimum of the ψ' and ψ rotations is approximated in the computations below by a single rotational isomeric state. According to numerous theoretical studies,[9,26,33,38,39] the minima associated with these two C–C bond rotations are separated by conformational energy barriers that substantially exceed RT. These data, however, have been challenged by Levitt and Warshel,[41] who claim to find a very low barrier to pentose pseudorotation in ribose and deoxyribose. Such low energy barriers are not consistent with available NMR and X-ray measures of ring puckering in various nucleic acid analogues.[40] The ψ' rotation is thus assigned here rotational isomeric states of 261 and 320° corresponding to the observed ^{2}E and ^{3}E ranges, respectively, of this angle. The ψ parameter is assigned three rotational states at 0 and ±120° corresponding to classical threefold rotations in the t and $g^\pm$ ranges, respectively. The $\omega'\omega$ phosphodiester rotations, in contrast, are sampled at smaller 30° intervals in view of both the theoretical uncertainty of these two angles[27] and the apparently low P–O rotational barrier suggested by broadly dispersed X-ray observations.[10,12–14] For simplicity, the glycosyl conformation angle χ is maintained in an anti (30°) orientation that positions the more bulky portion of the base away from the sugar–phosphate backbone; variations of this angle to the much less preferred syn range are expected to shift the ψ rotation from the usually favored g^- to the t range and hence to alter the conformation of the sugar–phosphate backbone.[38,39] According to a model by Govil et al.[42] of the unusual poly(8Br-rA) double helix, on the other hand, the syn–anti conformational transition is not expected to perturb the polynucleotide backbone.

The long-range interdependence of the $\psi'\omega'$ rotation angle pair is based upon the available X-ray and theoretical data cited above. When the ψ' angle is positioned in the g^- range associated with ^{3}E puckering, the ω' rotation is assigned to states between 30 and 120° in its g^+ range. If the ψ' rotation is in the t range associated with ^{2}E puckering, ω' is allowed to assume conformational states between −30 and 60° in the t range of this angle. The statistical weight matrix $\mathbf{U}_2(\psi',\omega')$ is thus given by the expression

$$\mathbf{U}_2(\psi',\omega') = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \tag{21}$$

where the rows index the two rotational values of ψ' and the columns the twelve (0–330°) states of ω'. For simplicity, all sterically allowed states are assigned a statistical weight of unity and all disallowed states a weight of zero in accordance with a hard-core potential. The 0's are null vectors of order 1 × 7.

The interdependence of the $\omega'\omega$ rotation angle pair is based upon our previous hard-core analysis of steric interactions in the dimethyl phosphate fragment of the polynucleotide backbone.[1] Of the 144 possible combinations of these angles sampled in a 30° analysis, 35 states centered about the (180°,180°) all-cis conformer are disallowed by second-order contacts between $C_{3'}$ and $C_{5'}$. The 12 × 12 statistical weight matrix $\mathbf{U}_3(\omega',\omega)$ is given by

$$\mathbf{U}_3(\omega',\omega) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{22}$$

where the 1's are unity matrices of orders required to conform.

The long-range interdependence of ω and ψ is described by a hard-core potential that reflects allowed rotational departures from the three low-energy $\omega\psi$ stacking geometries. When ψ is positioned in either the t or the g^+ rotational state, variations of the ω angle away from the t and g^- states, respectively, are free of steric contacts involving adjacent bases. If ψ is confined to the g^- range found in most nucleic acid helical structures, however, severe steric contacts are noted if ω is varied outside its preferred g^+ range. The 12 × 3 $\mathbf{U}_4(\omega,\psi)$ statistical weight matrix based upon these contacts is

$$\mathbf{U}_4(\omega,\psi) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \tag{23}$$

where the 1's are unity vectors and the 0's null vectors of order 3 × 1.

The $\mathbf{U}_1(\psi')$ matrix describing the sugar puckering through the ψ' rotation is given by the 1 × 2 row vector

$$\mathbf{U}_1(\psi') = [F_{3E} \quad F_{2E}] \tag{24}$$

where F_{3E} represents the fractional population of ^{3}E sugar puckering and $F_{2E} = 1 - F_{3E}$ the amount of ^{2}E puckering. The F_{3E} parameter is varied in the calculations reported below.

Finally, the $\mathbf{U}_5(\psi)$ matrix is given by the 3 × 1 column vector

$$\mathbf{U}_5(\psi) = \begin{bmatrix} F_{gt} \\ F_{tg} \\ F_{gg} \end{bmatrix} \tag{25}$$

F_{gg} refers to the fractional population of $\psi = g^-$ conformers where $O_{5'}$ is in the so-called double gauche (gg) conformation with respect to atoms $O_{1'}$ and $C_{3'}$. F_{gt} is the statistical weight of the $\psi = t$ state where $O_{5'}$ is gauche to $O_{1'}$ and trans to $C_{3'}$ and F_{tg} the weight for $\psi = g^+$ where $O_{5'}$ is trans to $O_{1'}$ and gauche to $C_{3'}$. F_{gg} is varied in the calculations below with F_{gt} and F_{tg} determined by the relationships $F_{tg} + F_{gt} + F_{gg} = 1$ and $F_{tg} = F_{gt}/3$.

The conformational details of the phosphodiester linkage are deduced indirectly from the rotational assignments of ψ' and ψ by using the above matrices. The predominant g^+g^+ phosphodiester conformer of A-RNA[12-14,19] is favored when ψ' is ^{3}E and ψ is g^- while the observed tg^+ phosphodiester state in B-DNA[19,20] is favored when ψ' is ^{2}E and ψ is g^-. The helices described by these rotational assignments, however, are flexible structures with limited variations in ω' and ω. If the ψ' and ψ populations are mixed among the various allowed conformational states, the phosphodiester motion is close to the free motion expected in a random coil. Among the conformers included in the array of random coil states are the A- and B-helical families, the so-called Watson–Crick helical structure,[20] various intercalated structures closely related to the helical geometries,[34] several sharp bends and reversals of the backbone, and even a backbone conformer close to the unusual zigzag state (Z-DNA) reported recently in the crystal structure of $(dCpdG)_3$.[43] Associated with the ^{2}E units of the random coil are extended $\omega' = t$ values of the phosphodiester linkage; these conformers were previously found to be important factors in accounting for the relatively large characteristic ratio of polynucleotide random coils.[27] Notably absent from the random coil states, however, is the $\omega'\omega = g^+t$, $\psi = g^-$ chain reversal conformer or π-bend observed in the TψC and anticodon loops of t-RNA[21,44] and predicted to occur in classical energy minimization studies of ApA.[26,45] This omission is not surprising since the π-bend also involves values of ϕ' and ϕ outside the t ranges considered here. Furthermore, the π-bends in t-RNA are found at pyrimidine–pyrimidine and pyrimidine–purine sequences, respectively, rather than in a purine–purine fragment of the chain. The theoretical π-bend in ApA is also predicted to be >3.0 kcal/mol in relative energy.[26]

Results of Numerical Calculations

Furanose Conformation. The conformation of the pentose in solution may be estimated by using so-called Karplus equations that relate proton–proton coupling constants J to the internal torsion angles τ of the structure.[31,32] The mean value of $J_{1'2'}$ expected for ^{3}E and closely related puckering states is 0.2 Hz and that associated with ^{2}E puckering and closely related states is 9.0 Hz.[32] The parameter F_{3E} used in these computations may thus be related to the experimentally observed $J_{1'2'}$ values by the relationships, $F_{3E} = (9 - J_{1'2'})/(9 - 0.2)$.

Table III
NMR Conformer Populations of ψ' and ψ

system	t, °C	$J_{1',2'}$-(ψ'), Hz[a]	F_{3E}	Σ(ψ), Hz[a]	F_{gg}
poly(rA)	22	2.0	0.79		
rA; rAMP	22			6.0	0.87
$(rA)_2$; $(rA)_3$	72	4.8	0.47	7.0	0.64
poly(rA)	86	5.9	0.22	7.3	0.57
approximate t dependence			F_{3E} = $-0.0065t + 0.93$		F_{gg} = $-0.0045t + 0.96$

[a] Data reported by Evans and Sarma.[11]

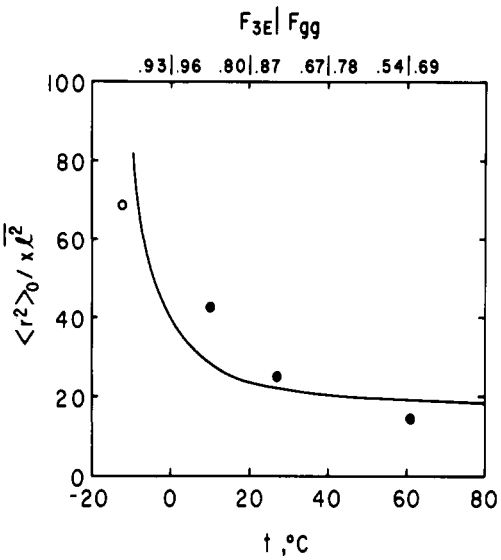

Figure 3. The calculated dependence of the limiting characteristic ratio upon temperature (lower abscissa) and $\psi'\psi$ conformer populations (upper abscissa). The open and closed circles denote experimental data explained in the text.

Similarly, the conformation of the ψ rotation in solution may be deduced from the sum of coupling constants Σ between the $H_{4'}$ and two $H_{5'}$ protons.[32] If the protons adopt a gauche orientation with respect to each other, J is predicted by Karplus analysis to assume an *average* value of 2.7 Hz. On the other hand, if the protons are trans, the mean J value is 7.1 Hz. These average J values are computed here by using several states within the respective ranges rather than for a single rotational state in each range. The sum of average coupling constants for the gg or $\psi = g^-$ conformer is 9.8 Hz while the Σ associated with either the tg or gt conformers is 5.4 Hz. The value of F_{gg} may then be related to the observed Σ values, following Sarma,[32] by the expression, $F_{gg} = (9.8 - \Sigma)/(9.8 - 5.4)$.

Values of $J_{1'2'}$ and Σ observed in NMR studies of poly(rA) and its low molecular weight analogues at several temperatures are listed in Table III.[11] The respective F_{3E} and F_{gg} values associated with these data are also given. The linear dependence of both F_{3E} and F_{gg} with temperature, as suggested by the data, is recorded in the table. These equations are utilized in the computations of chain dimensions reported below.

Polynucleotide Dimensions. Values of the characteristic ratio $C_n = \langle r^2\rangle_0/n\bar{l}^2$, where $n = 6x$ is the number of chemical bonds along the chain backbone and $\bar{l}^2$ is the mean-square bond length, were evaluated as a function of the experimental ψ' and ψ populations by using eq 15 above. The limiting values $C_\infty = \lim_{n\to\infty} C_n$ are plotted in Figure 3 as a function of temperature t. The approximate

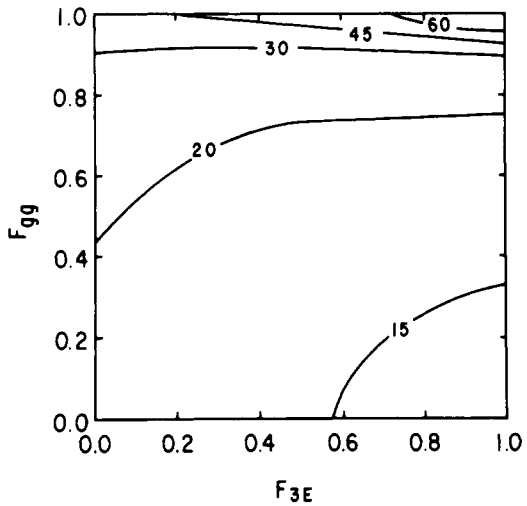

Figure 4. Contour diagram of the limiting characteristic ratio C_∞ as a function of the ψ' (F_{3E}) and ψ (F_{gg}) conformer populations. The flexible A-RNA helix is found at $F_{3E} = F_{gg} = 1$ while the flexible B-helix is located at $F_{3E} = 0$, $F_{gg} = 1$.

ψ' and ψ populations associated with various temperatures are reported in terms of F_{3E} and F_{gg} along with the upper abscissa. Also reported in the figure are the unperturbed dimensions of partially stacked poly(rA) (x = 1740 nucleotides) obtained by Eisenberg and Felsenfeld[15] at three different temperatures as well as the characteristic ratio found by Stannard and Felsenfeld[16] for a helical fragment of the molecule (x = 94) at −12 °C. Because limiting values of C_∞ are obtained within 90% for chains as short as 65 residues in these computations, the experimental data are compared to theoretical infinite chains.

The experimental points in Figure 3 are matched surprisingly well by this very approximate conformational analysis. The agreement may be improved when the dimensions are reassessed in terms of the potential energies of the $\psi'\omega'$, $\omega'\omega$, and $\omega\psi$ rotation angle pairs. Computations of C_∞ by us based upon the available PCILO potential energy surfaces[21,22] are found to exceed the dimensions of unstacked chains (calculated $C_\infty \sim 25$) and to fall considerably below those of the ordered helix (calculated $C_\infty \sim 35$). The $\psi'\omega'$ and $\omega\psi$ surfaces used in these latter computations must be approximated, as to be detailed elsewhere, from the several $\omega'\omega$ potential surfaces reported for all possible combinations of ψ' and ψ.

The limiting chain dimensions of Figure 3 are replotted as a function of F_{3E} and F_{gg} on the contour surface in Figure 4. In agreement with our earlier work on polynucleotide random coils,[9] C_∞ is found to depend upon the sugar pucker; here C_∞ varies from ~14 for all-^{3}E chains to ~20 for all-^{2}E chains in a highly random system, such as poly(rU), where $\Sigma \sim 8.0$ Hz or $F_{gg} \sim 0.35$. The increased value of ^{2}E random coil dimensions is due primarily to the more extended t backbone conformations of both ψ' and ω' in this type of unit. According to these data a flexible B-type (^{2}E) single-stranded helix is less extended than a similar A-type (^{3}E) helix. According to recent electric dichroism studies by Charney and Milstien,[46] a B-type double helix in solution is more extended than an A-type helix. It is possible for us to reproduce this experimental observation if we decrease slightly the allowed ranges of ω' and/or ω. The data is Figure 4, however, are seen to depend much more strongly upon F_{gg} than F_{3E}. The increase of C_∞ with F_{gg} is a consequence of the limited rotational flexibility in ω that is correlated through the $\mathbf{U}_4$ matrix from ψ in this range. Several $\omega\psi$ conformational combinations allowed at low values of F_{gg} are local bends of the polynucleotide backbone that depress chain dimensions. If the correlations between $\psi'\omega'$, $\omega'\omega$, and $\omega\psi$ are ignored and these four bonds are allowed to rotate freely over their respective first-order domains, C_∞ drops to 8.0.

Discussion

The above updated virtual bond scheme was designated to incorporate the strong influences of base stacking in the analysis of the spatial configurations of the polynucleotides (i.e., nucleic acids). Virtual bond schemes offered heretofore to treat this system[1-4] are based upon short-range nonbonded interactions within the sugar–phosphate repeating units of the chain backbone. According to several recent X-ray[17-20] and theoretical[21-26] studies, the polynucleotide is subject to "long-range" (three-bond) base stacking effects that involve rotations about the alternate C–C and O–P bonds in each C–C–O–P fragment of the sugar–phosphate backbone. Because rotations about the intervening C–O bonds are approximately rigid, each C–C–O–P chain fragment may be replaced by a hypothetical virtual bond. The six chemical bond nucleotide repeating unit is then simplified into two virtual bonds of comparable magnitude and also of similar size to the virtual bond repeating units of polypeptides[5] and polysaccharides.[6-8] In contrast to our previous analyses in terms of independent virtual bond units,[1-4] the orientations of the virtual bond pair constituting each $C_{4'}$-to-$C_{4'}$ nucleotide repeating residue are interdependent in this study. The chain, nonetheless, may be approximated as a sequence of independent nucleotide residues, each comprising two virtual bonds.

While the long-range interdependence of backbone rotations introduces a new level of complexity in the treatment of polynucleotides, it offers a novel theoretical probe to deduce the conformational details of the very flexible (and, to date, experimentally uncharacterized) phosphodiester rotations. We estimate these parameters for poly(rA) here from the experimentally determined conformations[11] of the two associated C–C backbone rotations. At low temperatures where the sugar pucker (ψ') is locked in the ^{3}E range and the acylic rotation about $C_{5'}$–$C_{4'}$ (ψ) is confined to the so-called gg conformer, the intervening phosphodiester linkage adopts the g^+g^+ conformation typical of A-RNA double helices. As the temperature increases and alternate conformations of ψ' and ψ become available, the phosphodiester rotations vary over more and more regions of conformation space. Under conditions where the pentose occurs with equal frequency in ^{2}E and ^{3}E puckered states and where the threefold $C_{5'}$–$C_{4'}$ backbone rotation is also free, the conformational domains of ω' and ω, as modeled here, include the standard tt, tg^+, tg^-, g^+t, and g^+g^+ angle combinations. The hard-core approximations that we invoke, however, exclude completely the g^-t, g^-g^-, g^-g^+, and g^+g^- phosphodiester combinations. In more refined estimates of intramolecular potential energy,[9,24,45] these four conformers occur with very low, although nonzero, probabilities.

The observed variations of ψ' and ψ with temperature thus provide a detailed conformational description of the entire transition of poly(rA) from a highly-stacked, single-stranded helix to a partially unstacked randomly coiling chain. The approximately linear dependence of coupling constants (and fractional ψ' and ψ populations) with temperature[11] reflects the well-known noncooperative nature of this transition.[47] Previously we approximated the poly(rA) transition by random occurrences of fixed helical nucleotide units within a randomly coiling chain backbone.[9] Long helices described by the earlier approach are

rigid rods in contrast to the flexible structures observed in solution at low temperature[46] and described by the present scheme. The earlier random coil backbone also covers more and slightly different regions of conformation space than presented here. The hard-core interactions described by eq 21–23 allow 12 different combinations of the flexible $\psi'\omega'\omega\psi$ quartet of angles per nucleotide of the random coil (***tttt***, ttg^+t, ttg^-t, $tttg^+$, ttg^+g^+, ***ttg*$^-$*t*$^+$**, ***ttg*$^+$*g***, ***g*$^-$*g*$^+$*tt***, $g^-g^+g^+t$, $g^-g^+tg^+$, $g^-g^+g^+g^+$, ***g*$^-$*g*$^+$*g*$^{+g-}$**). The five combinations of angles noted in boldface entail base stacking and describe loosely wound helical backbones;[3,25] moreover, four of these five stacked geometries fit known X-ray fiber diffraction patterns.[20] In contrast to the conventional concept of a completely unstacked random coil, approximately 5 out of 12, or 42%, of the nucleotides in the poly(rA) chain at high temperature assume stacked geometries typical of ordered systems. In other words, only 58% of the poly(rA) molecule unstacks in this helix–coil transition! Such a highly stacked random coil may account, in part, for the considerably lower enthalpy of base stacking determined by direct calorimetric measurements[48] in poly(rA) and its low molecular weight analogues compared to that deduced by van't Hoff analysis of various physical parameters. The computed calorimetric enthalpy assumes total unstacking of bases over the helix–coil transition while the van't Hoff values simply reflect the extent of reaction between two states of unspecified conformational detail. According to the above rough analysis, the computed calorimetric enthalpy should be corrected by the factor (1/0.58).

Finally, in accordance with the hard-core analysis offered here, the long-range conformational effects in the polynucleotide backbone depend upon the sequence of attached bases. The allowed combinations of backbone rotations, for example, will increase in pyrimidine systems which entail fewer steric contacts between bases.[45] The conformational freedom will also differ slightly between deoxyribo and ribo systems substituted at the $C_{2'}$ pentose atoms by protons and hydroxyl groups, respectively.[49] Hence, in heterogeneous DNA or RNA helices certain segments of the chain will be more subject to rotational fluctuations than other parts. Such conformational differences become important in site specific recognition of the nucleic acids by various proteins, drugs, and other outside agents.

Acknowledgment. The author is grateful to the National Institutes of Health (USPHS Grant GM 20861) and to the Charles and Johanna Busch Memorial Fund of Rutgers University for laboratory support; to Professor Joachim Seelig for his hospitality at the Biozentrum of the University of Basel, Switzerland, where this work was initiated; to the J. S. Guggenheim Memorial Foundation for a fellowship; to the University of Basel Computer Center and the Center for Computer and Information Services of Rutgers University for computer time; and to the USPHS for a Career Development Award (GM 00155).

References and Notes

(1) W. K. Olson and P. J. Flory, *Biopolymers*, 11, 1 (1972).
(2) W. K. Olson, Macromolecules, 8, 272 (1975).
(3) W. K. Olson, *Biopolymers*, **15**, 859 (1976).
(4) W. K. Olson *Biopolymers*, 18, 1213 (1979).
(5) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, 87, 2791 (1965).
(6) V. S. R. Rao, N. Yathindra, and P. R. Sundararajan, *Biopolymers*, 8, 325 (1969).
(7) C. V. Goebel, W. L. Dimpfl, and D. A. Brant, *Macromolecules*, 3, 644 (1970).
(8) S. G. Whittington, *Biopolymers*, **10**, 1481, 1617 (1971).
(9) W. K. Olson and P. J. Flory, *Biopolymers*, **11**, 25 (1972).
(10) M. Sundaralingam in "Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions", M. Sundaralingam and S. T. Rao, Eds., University Park Press, Baltimore, Md., 1975, p 487.
(11) F. E. Evans and R. H. Sarma, *Nature (London)*, **263**, 567 (1976).
(12) B. E. Hingerty, R. S. Brown, and A. Jack, *J. Mol. Biol.*, **124**, 58 (1978).
(13) S. R. Holdbrook, J. L. Sussman, R. W. Warrant, and S.-H. Kim, *J. Mol. Biol.*, **123**, 631 (1978).
(14) C. D. Stout, H. Mizuno, S. T. Rao, P. Swaminathan, J. Rubin, T. Brennan, and M. Sundaralingam, *Acta Crystallogr. Sect. B*, **34**, 1529 (1978).
(15) H. Eisenberg and G. Felsenfeld, *J. Mol. Biol.*, **30**, 17 (1967).
(16) B. Stannard and G. Felsenfeld, *Biopolymers*, **14**, 299 (1975).
(17) M. A. Viswamitra, O. Kennard, Z. Shakked, P. G. Jones, G. M. Sheldrick, S. Salisbury, and L. Falvello, *Curr. Sci.*, **47**, 289 (1978).
(18) M. A. Viswamitra, O. Kennard, P. G. Jones, G. M. Sheldrick, S. Salisbury, L. Falvello, and Z. Shakked, *Nature (London)*, **273**, 687 (1978).
(19) S. Arnott, R. Chandrasekaran, and E. Selsing, ref 10, p 577.
(20) S. Arnott, A. Banerjee, D. Birdsall, S. Campbell-Smith, R. Chandrasekaran, and L. Puigjaner, *ACS Symp. Ser.*, in press.
(21) A. R. Srinivasan and W. K. Olson, *Nucleic Acids Res.*, in press.
(22) D. Perahia, B. Pullman, D. Vasilescu, R. Cornillon, and H. Broch, *Biochim. Biophys. Acta*, **478**, 244 (1977).
(23) H. Broch and D. Vasilescu, *Biopolymers*, **18**, 909 (1979).
(24) N. Yathindra and M. Sundaralingam, ref 10, p 649.
(25) N. Yathindara in "Biomolecular Structure, Conformation, Function, and Evolution", R. Srinivasan, Ed., Pergamon Press, Oxford, in press.
(26) S. Broyde and B. Hingerty, *Nucleic Acids Res.*, **6**, 2165 (1979).
(27) W. K. Olson, *Biopolymers*, **14**, 1775, 1797 (1975).
(28) D. J. Patel, *Biopolymers*, **15**, 533 (1976).
(29) D. J. Patel, *Biopolymers*, **16**, 1635 (1977).
(30) D. G. Gorenstein, J. B. Findlay, R. K. Momii, B. A. Luxon, and D. Kar, *Biochemistry*, **15**, 3796 (1976).
(31) P. O. P. Ts'o in "Basic Principles in Nucleic Acid Chemistry", Vol. I, P. O. P. Ts'o, Ed., Academic Press, New York, 1974, p 454.
(32) M. M. Dhingra and R. H. Sarma in "Stereodynamics of Molecular Systems", R. H. Sarma, Ed., Pergamon Press, New York, 1979, p 3.
(33) R. Tewari, R. K. Nanda, and G. Govil, *Biopolymers*, **13**, 2015 (1974).
(34) H. M. Berman and S. Neidle, ref 32, p 367.
(35) F. E. Hruska in "Conformation of Biological Molecules and Polymers (Proceedings of the Fifth Jerusalem Symposium)", E. D. Bergmann and B. Pullman, Eds., Israel Academy of Sciences and Humanities, Jerusalem, 1973, p 345.
(36) D. B. Davies and S. S. Danyluk, *Biochemistry*, **13**, 4417 (1974).
(37) M. Sundaralingam, *Biopolymers*, 7, 821 (1969).
(38) W. K. Olson, *Biopolymers*, **12**, 1787 (1973).
(39) N. Yathindra and M. Sundaralingam, *Biopolymers*, **12**, 297 (1973).
(40) W. K. Olson, manuscript in preparation.
(41) M. Levitt and A. Warshel, *J. Am. Chem. Soc.*, **100**, 2607 (1978).
(42) G. Govil, C. Fisk, F. B. Howard, and H. T. Miles, *Nucleic Acids Res.*, **4**, 2573 (1977).
(43) A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, T. L. Crawford, J. H. van Boom, G. van der Marel, and A Rich, *Nature (London)*, **282**, 680 (1979).
(44) S.-H. Kim and J. L. Sussman, *Nature (London)*, **260**, 645 (1976).
(45) S. B. Broyde, R. M. Wartell, S. D. Stellman, B. Hingerty, and R. Langridge, *Biopolymers*, **14**, 1597 (1975).
(46) E. Charney and J. B. Milstien, *Biopolymers*, **17**, 1629 (1978).
(47) M. Leng and G. Felsenfeld, *J. Mol. Biol.*, **15**, 455 (1966).
(48) K. J. Breslauer and J. M. Sturtevant, *Biophys. Chem.*, **7**, 205 (1977).
(49) W. K. Olson and P. J. Flory, *Biopolymers*, **11**, 57 (1972).