Computational studies of polynucleotide flexibility

Wilma K. Olson

Department of Chemistry, Rutgers, The State University of New Jersey, New Brunswick, NJ 08903, USA

Received 16 September 1981; Accepted 9 December 1981

ABSTRACT

Details of polynucleotide flexibility may be probed through a combination of semiempirical potential energy calculations and statistical mechanical analyses. The pseudorotational motions of the furanose and the long-range correlated rotations of the chain backbone are described briefly here.

INTRODUCTION

The physical and biological properties of the polynucleotides reflect the three-dimensional spatial arrangements, or configurations, that the atoms comprising these chain molecules can assume. The chain configurations depend, in turn, upon the structural parameters (e.g., bond lengths and valence bond angles) defined by the chemical architecture of the system and the angles of internal rotation (ψ' , ϕ' , ω' , ω , ϕ , ψ , and χ in Figure 1) described about the single bonds of the chain skeleton. Although subject to minor fluctuations, the structural parameters usually remain fixed in computational studies. The experimentally observed variations in bond lengths and valence angles occur more or less symmetrically about mean values and the effects of those of opposite signs tend to cancel one another.¹⁻³ The rotations about the skeletal bonds, which are subject to much wider latitudes of variation, thus constitute the principal determinants of polynucleotide configuration. Those chain structures generated solely by rotational variations constitute a special category of molecular configurations, generally referred to as conformations.¹

The conformational flexibility of the polynucleotide chain is frequently described in terms of the mean-square end-to-end dimensions $\langle r^2 \rangle_0$ of the ideal unperturbed chain. This parameter may be related quantitatively to the structural geometry and the potential energies governing the local hindrances to internal rotations in the chain backbone through a simple sequence of matrix operations.⁴ The facility with which cyclic and looped polynucleo-

Figure 1

Computer generated representation of a pdApdAp fragment of a polydeoxynucleotide chain showing chemical bonds and internal rotation angles.

tide structures are formed, through enzymatic reaction and hydrogen bonding associations, respectively, from the acyclic chain is related to the statistical distribution of the two ends of the molecule relative to one another and thus is also dependent upon the conformational character of the system. The spatial density distribution functions $W(\mathbf{r})d\mathbf{r}$ that describe the complete array of conformations accessible to the flexible polynucleotide may be estimated, for short chains, by direct Monte Carlo simulations.⁵⁻⁷ The spatial distributions of longer chains are approximated by a three-dimensional Hermite series expansion of the Gaussian in terms of average tensor moments of the chain.^{7,8} In addition to the dependence upon $W(\mathbf{r})$, ring and loop closure is also governed by the distribution of angular orientations $\Gamma(\theta)d\theta$ of terminal bonds in the acyclic chain fragment.⁹⁻¹¹

CONFORMATIONAL PREFERENCES AND INTERDEPENDENCIES

Much of the information describing the conformational character of the polynucleotides has originated from X-ray crystallographic analyses of low molecular weight nucleic acid analogs — principally the subunit nucleosides and nucleotides with a few recently refined larger oligonucleotide fragments.¹² A more direct probe of polynucleotide conformation in solution is

obtained from high resolution nuclear magnetic resonance (NMR) three-bond spin-spin coupling constants. The observed splitting J_{X-Y} (which is only possible between atoms such as ^1H , ^{13}C , ^{31}P , etc. with unpaired nuclear spin) is dependent upon the intervening dihedral angle ϕ according to simple empirical (Karplus)¹³ relationships (i.e., $J_{X-Y} = J_0 \cos^2 \phi + J_1 \cos \phi + J_2$ where J_0 , J_1 , and J_2 are constants dependent upon the nature of substituents X and Y and the local chemical environment). According to such measurements,¹⁴ the furanose ring is subject to rapid pseudorotational variations between two distinct puckered forms — the C3'-endo pucker where ψ' assumes a gauche⁺ (g^+) state ($-60 \pm 30^\circ$) and the C2'-endo pucker where ψ' adopts a trans (t) arrangement ($\sim 180 \pm 30^\circ$). The exocyclic C-C rotation ψ is similarly found to adopt three different conformers, the g^+ state with atom O5' in staggered gauche arrangements with respect to both O1' and C3' being most favored. Compared to the C-C bonds of the polynucleotide backbone, the flexibility about the two C-O rotations (ϕ' and ϕ) is more restricted. According to measured $^1\text{H} \dots ^{31}\text{P}$ and $^{13}\text{C} \dots ^{31}\text{P}$ coupling constants, each of these angles is confined almost exclusively to its sterically unencumbered t range. The ϕ' angle, however, is found in the g^- range in the Z form of DNA.¹⁵ The major uncertainty in the analysis of polynucleotide flexibility then is the conformational nature of the P-O torsions (ω' and ω). Unfortunately, there is as yet no direct and reliable experimental probe of these rotations in model nucleotide systems. While the observed ^{31}P chemical shifts have been useful in monitoring the helix-to-coil transitions of short oligonucleotides,¹⁶⁻¹⁸ these data are not able to describe the blend of ω' and ω rotational conformers with certainty. Furthermore, various theoretical predictions of the phosphodiester motions are widely discrepant.¹⁹ NMR¹⁴ and theoretical²⁰ studies, however, are in agreement over the predominance of anti glycosyl (χ) conformers, which position the more unwieldy portions of the heterocyclic bases (i.e., the 2'-keto group of a pyrimidine or the six-membered ring of a purine) away from the sugar-phosphate backbone.

Recent experimental^{15,21,22} and theoretical^{4,23-25} studies suggest that the polynucleotide is subject to unique long-range rotational correlations. The strong stacking interactions between adjacent bases favor sequences of crankshaft-like motions of nonadjacent rotations that maintain close base-base contacts. Variations of the sugar pucker (ψ') between C3'-endo and C2'-endo forms introduce simultaneous conformational changes of the ω' rotation angle from the g^- to t range. Changes in the ψ rotation

from g^+ to t to g^- states similarly lead to variations of ω from g^- to t to g^+ arrangements. Fluctuations of ϕ' toward g^- conformers also correlate with rotations in ϕ toward g^+ states. As illustrated for the $\omega\psi$ pair in Figure 2, such angle correlations introduce considerable mobility even in highly ordered (stacked) polynucleotide chains. The shaded portion of this figure depicts all $\omega\psi$ combinations that maintain the angular ($\Lambda \leq 45^\circ$) and distance ($3 \text{ \AA} \leq Z \leq 4 \text{ \AA}$) requirements of base stacking in a B-DNA double helix. While the long-range rotational interdependencies in the polynucleotide backbone complicate theoretical treatment of the chain, these concerted motions also provide an indirect probe of the P-O torsional motions. Using the $\psi'\omega'$ and $\omega\psi$ conformational correlations, the P-O rotational tendencies follow at once from the experimental observations of the C-C angles.⁴

FURANOSE PSEUDOROTATION

Comprehensive treatment of correlated motions in the polynucleotide backbone requires knowledge of the energetic preferences of the many conformations of the chain backbone. A semiempirical potential function has thus been developed to estimate the pseudorotational motions of simple

Figure 2

Composite contour diagram of the base stacking angles Λ (dashed curves) and distances Z (solid curves) as functions of the $\omega\psi$ torsion angles relative to the B-DNA reference helix (noted at $\times$). The allowed base stacking pathway, where $\Lambda \leq 45^\circ$ and $3 \text{ \AA} \leq Z \leq 4 \text{ \AA}$, is shaded.

ribose and deoxyribose sugars.²⁶ Previous potential energy calculations²⁷ were not designed to match the puckering preferences or to mimic the known geometric (valence angle) changes that accompany furanose pseudorotation. Theoretical predictions based upon such potential functions must be regarded with caution. According to both X-ray¹² and NMR¹⁴ measurements, the C3'-endo and C2'-endo forms are almost equally favored in ribose systems; the C2'-endo pucker, however, is predominant in 2'-deoxyribose and the C3'-endo in 3'-deoxyribose. The five valence angles of the furanose ring are also known to fluctuate (by 2-4°) in a sinusoidal fashion over the complete cycle of pseudorotational changes.^{28,29}

The potential energies of the differently puckered furanose molecules studied here reflect the combined contributions of nonbonded interactions (V_{NB}), valence angle strain (V_{STR}), intrinsic torsional barriers (V_{TOR}) and gauche effects (V_G). Each conformer additionally satisfies a pseudorotational constraint energy.

$$V_{PSEU} = \sum_{j=0}^4 1000(1 - \cos(\tau_j - \tau_j^0)) \quad (1)$$

that ensures that the five torsion angles ($\tau_0 = \angle C4'-O1'-C1'-C2'$, $\tau_1 = \angle O1'-C1'-C2'-C3'$, etc.) adopt the ideal values predicted by Altona and Sundaralingam³⁰ for the particular puckering. Because this term remains virtually constant for all puckered forms, it does not enter into the computed energy totals.

The potential energy V to furanose pseudorotation is expressed here by the simple summation:

$$V = V_{NB} + V_{STR} + V_{TOR} + V_G \quad (2)$$

The V_{NB} term includes the London attractions ($-c/r^6$), van der Waals' repulsions (d/r^{12}) and Coulombic interactions ($e\delta_k\delta_l/r$) between all pairs of nonbonded atoms or heavy atom groups (e.g., CH_3 , NH_2 , and OH) separated by at least three intervening chemical bonds. Pairs of atoms separated by three intervening bonds of the furanose, however, do not contribute to V_{NB} . Instead, we consider the interactions between ring atoms in V_{STR} . Because bond lengths are kept fixed, no bond stretching energies enter these computations.

The constants c describing the pairwise London attraction of atoms k and l are evaluated from atomic polarizabilities α (in $\text{\AA}^3$) and the effective number of valence electrons N using the Slater-Kirkwood equation:³¹

$$c_{k1} = 365\alpha_k\alpha_1/[(\alpha_k/N_k)^{1/2} + (\alpha_1/N_1)^{1/2}] \quad (3)$$

Values of α and N as well as r^0 (van der Waals radii) and δ (atomic charge) are listed in Table I. The van der Waals constant d is chosen so that the London term plus the van der Waals term display a minimum at the distance $R_{k1} = r_k^0 + r_1^0 + 0.2\text{\AA}$. This additional distance is used to correct for the effects of attractions imposed by other atoms in the molecule on the two-body force.³² The parameter d_{k1} is then given as

$$d_{k1} = R_{k1}^6 c_{k1}^{12} \quad (4)$$

The dielectric constant is set at 4.0 so that the numerical constant e required to yield Coulombic energies in kcal/mole is 83.

The endocyclic and exocyclic valence angle strain is accounted for by a sum of harmonic angle bending terms of the form:

$$V_{STR} = K_\theta (\theta - \theta^0)^2 \quad (5)$$

The rest angles θ^0 are taken to be tetrahedral (1.91 rad.) and the K_θ are estimated to be 40, 34, and 30 kcal/mole-rad² for C-O-C, C-C-O, and C-C-C sequences, respectively.

The V_{TOR} term is included to take account of the more subtle contributions from bond orbitals associated with the atoms attached to a given bond,³³⁻³⁵ including the effects of distortion of these orbitals by rotation.³⁶ The potential is taken to be threefold for the five torsions of

Table I. Parameters for Nonbonded Interactions

Atom or Group	$\alpha, \text{\AA}^3$	N	$r^0, \text{\AA}$	δ, esu
H	0.42	0.9	1.3	.051 to .053 [†]
C	0.93	5	1.8	-.041 to .127 [†]
O	0.64	7	1.6	-.271
CH ₃	1.77	7	2.0	-.003
NH ₂	1.87	8	2.0	-.053
OH	1.06	8	1.7	-.157

[†]See reference 26 for further details.

the furanose ring and is represented by

$$V_{\text{TOR}} = \sum_{j=0}^4 (V_3/2)(1 + \cos 3\tau_j) \quad (6)$$

Barrier heights V_3 of 2.8 and 1.8 kcal/mole are assigned to rotations centered about C-C and C-O bonds, respectively.

The intrinsic tendency of O-C-C-C and O-C-C-O sequences to favor gauche in favor of trans conformations is modeled by a phenomenological term

$$V_G = \sum_{i=0}^4 \sum_{j=1}^3 (V_2/2)(1 + \cos 2(\tau_i + \Delta_j)) \quad (7)$$

Barrier heights V_2 of 0.2 and 1.0 kcal/mole are introduced to reproduce the known t/g energy differences of O-C-C-C and O-C-C-O bond fragments,³⁷ respectively. The parameter Δ is a phase angle that relates the rotation of a given fragment to the torsion angle τ_i sharing a common central bond.

The puckering preferences predicted on the basis of the potential energies are compared with the frequencies of X-ray observations in Table II. The theoretical populations are obtained from the relative contributions of the Boltzmann factor of the potential energy over the four major quadrants of pseudorotation space — the favored C3'-endo and C2'-endo regions as well as the intermediate O1'-endo and O1'-exo domains. The intermediate conformers are disfavored by a combination of unfavorable steric interactions and gauche

Table II. Comparative Population of Model Furanoses

Furanose	X-ray [†]				Theory			
	C3'-endo	O1'-endo	C2'-endo	O1'-exo	C3'-endo	O1'-endo	C2'-endo	O1'-exo
Ribose (108 X-ray structures)	0.44	0.01	0.55	0.0	0.48	0.01	0.51	0.0
2'-Deoxyribose (27 X-ray structures)	0.23	0.10	0.67	0.0	0.15	0.11	0.74	0.0
3'-Deoxyribose (1 X-ray structure)	1.0	0.0	0.0	0.0	0.69	0.11	0.20	0.0

[†]Numerical data refer to fractional populations of each puckering category.

contributions. The equivalent populations of C3'-endo and C2'-endo ribose are a reflection of the (pseudo) geometrical symmetry of the ring. The biased conformational preferences in the deoxyribose systems can be attributed principally to the intrinsic tendency of O-C-C-O bond sequences to adopt *gauche* conformations. The computed and experimental populations are in close accord. No previous potential was able to account for the puckering preferences of the deoxyribose structures.

As evident from Figure 3, the valence angle fluctuations computed on the basis of this potential (dashed lines) are also in good agreement with experimental observations (solid curves). The correspondence of theoretical and observed proton coupling constants in Table III is equally satisfactory.

Figure 3

Comparative pseudorotational variations of endocyclic valence angles θ^X of furanose rings associated with energy optimization in this work (dashed curves) and observed in X-ray crystallographic regression (RGN) analyses (solid lines).^{28,29} The four quadrants (n, e, s, w) of pseudorotation space associated with the major categories of sugar puckering (C3'-endo, O1'-endo, C2'-endo, O1'-exo) are noted above the figure.

POLYNUCLEOTIDE PROPERTIES

A sequence of preliminary calculations of $\langle r^2 \rangle_0$ and $W(r)dr$ indicate that correlated motions of $\psi'\omega'$ or $\omega\psi$ do not significantly affect the unperturbed dimensions and loop closure tendencies of single-stranded polynucleotide helices. A theoretical model that allows free correlated changes in ω and ψ as described above matches the limiting characteristic ratio³⁸ of helical poly rA at -12°C ($\lim_{n \rightarrow \infty} \langle r^2 \rangle_0 / n\bar{l}^2 \sim 80$ where n is the number of chemical bonds and $\bar{l}^2$ the mean-square bond length) just as well as a model that restricts these angles to their predominant g^- and g^+ domains, respectively. The ribose ring adopts a rigid C3'-endo puckering in these calculations and the ϕ' and ϕ C-O angles assume single rotational isomeric states centered

Table III. Comparison of Theoretical Proton Coupling Constants in Hz with Experimental Measurements.

Coupling Constant	Theory	Experiment ^{14,26,27}
<u>Ribose</u>		
$J_{1'2'}$	4.9	4.5 ± 1.5
$J_{2'3'}$	5.2	5.2 ± 0.2
$J_{3'4'}$	4.8	5.0 ± 1.5
<u>2'-Deoxyribose</u>		
$J_{1'2'}$	9.0	7.4 ± 0.5
$J_{1'2''}$	6.4	6.5 ± 0.5
$J_{2'3'}$	7.0	6.5 ± 0.5
$J_{2''3'}$	2.9	3.3 ± 0.5
$J_{3'4'}$	2.3	3.1 ± 0.5
<u>3'-Deoxyribose</u>		
$J_{1'2'}$	2.6	2.1 ± 0.4
$J_{2'3'}$	7.0	5.8 ± 0.3
$J_{2'3''}$	3.5	2.7 ± 0.4
$J_{3'4'}$	8.7	8.8 ± 1.1
$J_{3''4'}$	6.5	5.9 ± 0.1

in their τ ranges. The short-range rotational interdependence of the ω' angles also excludes consecutive g^-g^+ rotational combinations that introduce sharp U-turns in the chain backbone.

A hard core analysis^{39,40} of the base stacking that accompanies correlated $\omega\psi$ rotational changes reveals a general opening of the helix as the angles vary from their preferred g^-g^+ combinations to the tt state. Rotational changes to the g^+g^- arrangement, on the other hand, reduce the base separation distance to less than normal van der Waals' distances and consequently raise the potential energy. While the increased separation distance between parallel bases of the tt conformer reduces stacking interactions and also raises the potential energy, this arrangement allows planar aromatic moieties to intercalate between adjacent base pairs. In the absence of intercalating species, the $\omega\psi = tt$ combination enhances the syn-anti transition of the heterocyclic bases along the chain backbone.^{39,40} Such concerted rotational changes may possibly account for the transition of the familiar right-handed DNA-B helix⁴¹ to the recently described left-handed Z-type backbone.¹⁵

ACKNOWLEDGEMENTS

This research was sponsored by the U.S. Public Health Service under grants CA 25981 and GM 20861 and the donors of the Petroleum Research Foundation to grant AC 11586. Computer time was supplied by the Rutgers University Center for Computer and Information Services. W.K.O. is also the recipient of a U.S.P.H.S. Research Career Development Award (GM 155).

REFERENCES

1. P.J. Flory, Statistical Mechanics of Chain Molecules, Interscience, New York, 1969, pp. 12-15.
2. W.K. Olson, Ph.D. thesis, Stanford University, Stanford, California, 1970, pp. 102-105.
3. W.K. Olson and P.J. Flory, Biopolymers, 11, 25-56 (1972).
4. W.K. Olson, Macromolecules, 13, 721-728 (1980).
5. C. Delisi and D.M. Crothers, Biopolymers, 10, 1809-1827 (1971).
6. R. Tewari, R.K. Nanda, and G. Govil, Biopolymers, 13, 2015-2035 (1974).
7. R. Yevich and W.K. Olson, Biopolymers, 18, 113-145 (1979).
8. D.Y. Yoon and P.J. Flory, J. Chem. Phys., 61, 5358-5365 (1974).
9. P.J. Flory, U.W. Suter, and M. Mutter, J. Am. Chem. Soc., 98, 5733-5739 (1976).
10. W.K. Olson, in Stereodynamics of Molecular Systems, R.H. Sarma, Ed., Pergamon Press, New York, 1979, pp. 297-314.
11. N.L. Marky and W.K. Olson, Biopolymers, submitted.
12. See, for example, N. Seeman, in Nucleic Acid Geometry and Dynamics, R.H. Sarma, Ed., Pergamon Press, New York, 1980, pp. 109-142.
13. M. Karplus, J. Chem. Phys., 30, 11-15 (1959).

14. R.H. Sarma, in Nucleic Acid Geometry and Dynamics, R.H. Sarma, Ed., Pergamon Press, New York, 1980 pp. 143-184.
15. A.H.-J. Wang, G.J. Quigley, F.J. Kolpak, G. van der Marel, J.H. van Boom, and A. Rich, Science, 211, 171-176 (1981).
16. D.J. Patel, Biopolymers, 15, 533-558 (1976).
17. D.G. Gorenstein, J.B. Findlay, R.K. Momii, B.A. Luxon, and D. Kar, Biochemistry, 15, 3796-3803 (1976).
18. C.A.G. Haasnoot and C. Altona, Nuc. Acids Res., 6, 1135-1149 (1979).
19. W.K. Olson, Biopolymers, 14, 1775-1795 (1975).
20. W.K. Olson, Biopolymers, 12, 1787-1814 (1973).
21. M.A. Viswamitra, O. Kennard, P.G. Jones, G.M. Sheldrick, S. Salisbury, L. Falvello, and Z. Shakked, Nature, 273, 687-688 (1978).
22. A.R. Srinivasan and W.K. Olson, Nuc. Acids Res., 8, 2307-2329 (1980).
23. N. Yathindra and 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, Maryland, 1975, pp. 649-676.
24. H. Broch and D. Vasilescu, Biopolymers, 18, 909-930 (1979).
25. S. Broyde and B. Hingerty, Nuc. Acids Res., 6, 2165-2178 (1979).
26. W.K. Olson, J. Am. Chem. Soc., in press.
27. For a complete bibliography and comparison see, W.K. Olson and J.L. Sussman, J. Am. Chem. Soc., in press.
28. P. Murray-Rust and S. Motherwell, Acta Cryst., B34, 2534-2546 (1978).
29. E. Westhof and M. Sundaralingam, J. Am. Chem. Soc., 102, 1493-1500 (1980).
30. C. Altona and M. Sundaralingam, J. Am. Chem. Soc., 94, 8205-8212 (1972).
31. K.S. Pitzer in Advances in Chemical Physics, Vol. 2, I. Prigogine, Ed., Interscience Publishers, Inc., New York, 1959, pp. 59-83.
32. D.A. Brant, W.G. Miller, and P.J. Flory, J. Mol. Biol., 23, 47-65 (1967).
33. D.A. Brant and P.J. Flory, J. Am. Chem. Soc., 87, 2791-2800 (1965).
34. R.A. Scott and H.A. Scheraga, J. Chem. Phys., 42, 2209-2215 (1965).
35. I.R. Epstein and W.N. Lipscomb, J. Am. Chem. Soc., 92, 6094-6095 (1970).
36. W.L. Jorgensen and L.C. Allen, J. Am. Chem. Soc., 93, 567-574 (1971).
37. A. Abe and J.E. Mark, J. Am. Chem. Soc., 98, 6468-6472 (1976).
38. B. Stannard and G. Felsenfeld, Biopolymers, 14, 299-307 (1975).
39. L. Ciancia and W.K. Olson, unpublished data.
40. W.K. Olson, in Biomolecular Stereodynamics, Vol. 1, R.H. Sarma, Ed., Adenine Press, New York, in press.
41. R. Chandrasekaran, S. Arnott, A. Banerjee, S. Campbell-Smith, A.G.W. Leslie, and L. Puigjaner, in Fiber Diffraction Methods, A.D. French and K.H. Gardner, Eds., ACS Symposium Series, Vol. 141, American Chemical Society, Washington, D.C., 1980, pp. 483-502.