Vol. 89, No. 2, 1979 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 27, 1979 Pages 374-385

THE POSSIBLE CHAIN BENDINGS IN POLYDEOXYNUCLEOTIDES

P.K. PONNUSWAMY and P. THIYAGARAJAN

Department of Physics, Autonomous Postgraduate Centre,
University of Madras, Tiruchirapalli 620 020,
Tamilnadu, INDIA

Received May 22,1979

SUMMARY: Energy minimization studies were carried out on the trimeric subunit of DNA, dApdApdA by treating the thirteen relevant dihedral angles as simultaneous variables in the total energy function. The four important sugar pucker states ^{3}E-^{3}E-^{3}E, ^{2}E-^{2}E-^{2}E, ^{3}E-^{2}E-^{3}E and ^{2}E-^{3}E-^{2}E were investigated. From the characteristic chain diverting properties of the low energy conformations four possible types of bend families were identified and their probable sites of occurrence in the polydeoxynucleotides predicted.

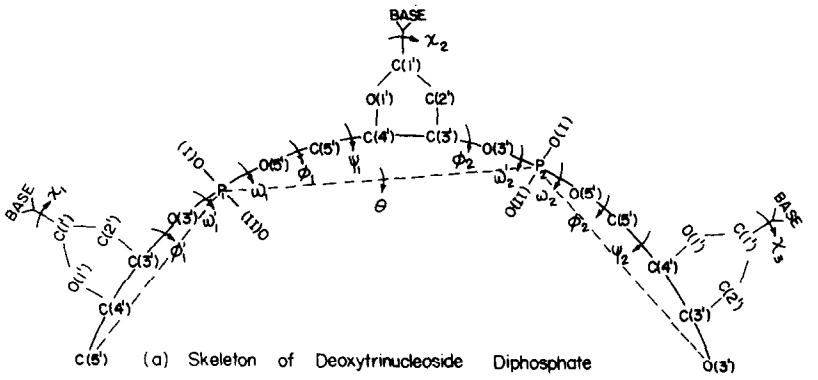
It is now firmly established that at specific sites a polypeptide chain typically bends to effect globularity in protein molecules (1,2). Systematic studies have further revealed that these bend segments are mainly associated with a few specific amino acid residues (3,4). However, it is not yet clearly known whether such recognizable well defined bend segments are also present in nucleic acids with base/sugar specificity. There exists considerable evidence to suggest the involvement of such bends in chromatin systems (5,6) as well as in tRNA molecules (7-10) and recently attempts are being made to understand the folding pattern of nucleic acid systems (11-15). Although references to the loop and bend structures are often made in the realm of both theoretical and experimental studies (16,17) on nucleic acids, information is still lacking

0006-291X/79/140374-12$01.00/0

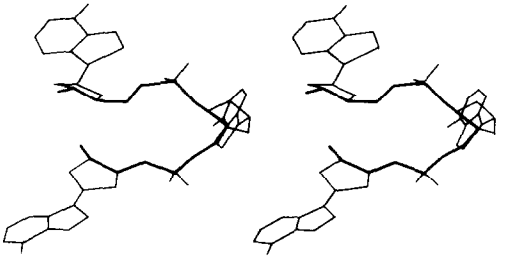
on the precise folding pattern of the nucleic acid chain. Insight into this aspect with reference to the relationship between base/sugar specificity and backbone bending with stability may be gained if it were possible to investigate the conformational characteristics of specific oligonucleotides. As a part of our ongoing program of theoretical investigations on polynucleotides (18-22), we have carried out an extensive analysis of the preferred conformations of a number of dimeric subunits of DNA and RNA and a typical trimeric subunit of DNA by means of potential energy calculations. The results of the trimeric unit dApdApdA were found to give immense information as to the most probable bend structures in polydeoxynucleotides (21). This unit, consisting of three nucleosides, two phosphate linkages and thirteen variable dihedral angles (see diagram a in Fig.1) is of sufficiently large but of the required minimum size to define bend structures in polynucleotides. The present study enables us to identify four families of backbone conformations that could characteristically alter the chain direction introducing typical bend sections which may be of considerable interest in the problem of polynucleotide folding.

METHODS

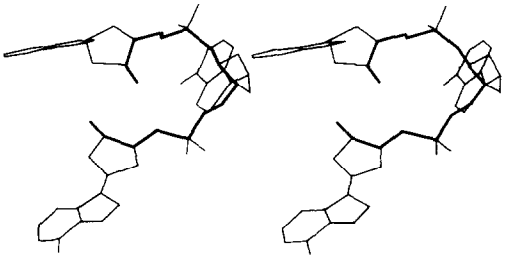
The total potential energy comprising of nonbonded, electrostatic, hydrogen bonding, torsional and ring distortion contributions was computed by the use of appropriate expressions(22) and minimized by simultaneously varying the ten backbone and three glycosyl dihedral angles through the use of a powerful minimization technique (23,24). The three sugar units were allowed to adopt the four important pucker sequences $^{3}E\text{-}^{3}E\text{-}^{3}E$, $^{3}E\text{-}^{2}E\text{-}^{3}E$, $^{2}E\text{-}^{2}E\text{-}^{2}E$ and $^{2}E\text{-}^{3}E\text{-}^{2}E$.



(a) Skeleton of Deoxytrinucleoside Diphosphate



(b) U - Bend

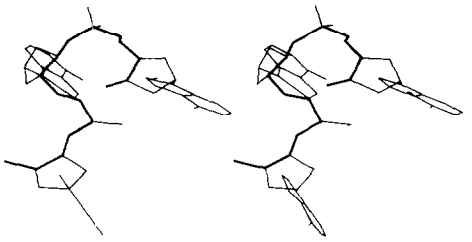


(c) Lr - Bend, Type 1

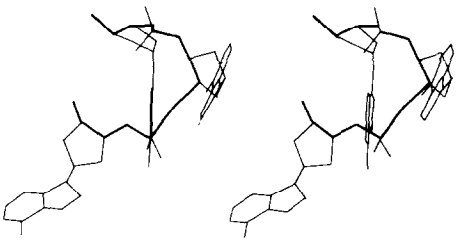
Fig.1. : a) Skeleton and conformational parameters for trinucleotide diphosphate: see ref.18 for the convention of measuring these parameters.

b-i) Stereo pictures for the four families of bends.

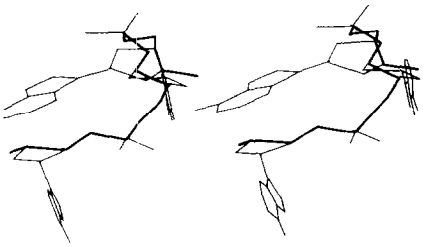
To start with ninety six probable conformations picked up by combining all the low energy domains for the various dihedral angles for the dimeric subunit dApdA were considered and their stabilities determined after minimiza-



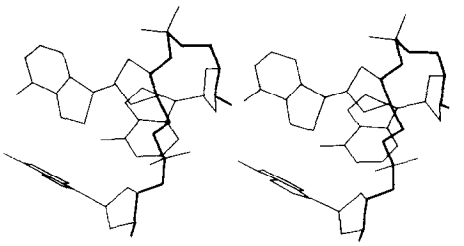
(d) Lr - Bend, Type 2



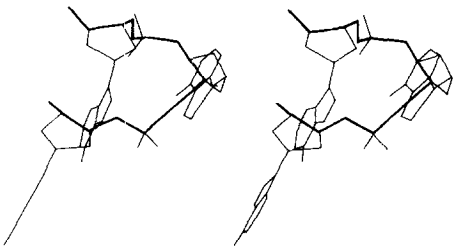
(g) L_ℓ - Bend, Type 2



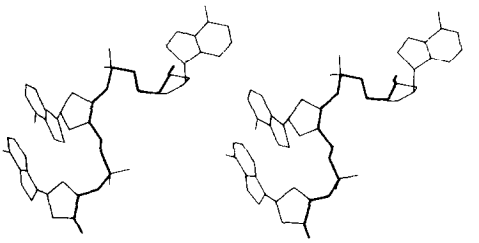
(e) Lr - Bend, Type 3



(h) Z - Bend, Type I



(f) L_ℓ - Bend, Type I



(i) Z - Bend, Type 2

Fig. 1 (continued)

tion. This resulted in seventeen conformations within an energy limit of 5.0 kcal/mole with respect to the lowest energy conformation (25). Making use of the starting conformations corresponding to these low energy cases for the dimeric subunit dApdA, seventy probable conformations were selected for the energy minimization study on dApdApdA unit.

RESULTS AND DISCUSSION

The conformations which after the minimization process, ended up within an energy level of 5.0 kcal/mole above

TABLE I

Low Energy Conformations for d-(ApApA)

No.	Sugar sequence	Conformational parameters (°)												Relative Energy	Comment
		χ_1	ϕ_1'	ω_1'	ω_1	ϕ_1	ψ_1	χ_2	ϕ_2'	ω_2'	ω_2	ϕ_2	ψ_2	χ_3	(kcal/mole)	
1	^3E-^3E-^3E	60	-178	142	-76	160	57	13	-164	101	59	-151	82	40	0.00	Nonhelical
2	^3E-^3E-^3E	101	-178	154	97	170	68	-3	-170	110	72	174	63	21	0.30	Nonhelical
3	^3E-^3E-^3E	38	-133	-69	-77	-176	63	10	-172	106	72	-154	70	22	1.51	Nonhelical
4	^2E-^2E-^2E	27	-171	-114	-57	178	72	62	-174	-101	-73	-176	61	85	1.57	DNA-B
5	^3E-^3E-^3E	19	-165	-78	163	133	169	11	-166	-70	174	164	170	3	2.98	Watson-Crick DNA
6	^3E-^2E-^3E	45	-136	179	69	-175	75	27	-133	59	82	-165	69	27	3.43	Nonhelical
7	^3E-^3E-^3E	93	-172	159	77	-175	60	1	172	130	-60	179	71	13	3.79	Nonhelical
8	^3E-^3E-^3E	9	-143	-68	-73	170	73	10	-179	119	-64	174	64	12	4.25	Nonhelical
9	^3E-^3E-^3E	52	-172	167	72	179	71	-8	-113	-42	-84	-178	69	22	4.35	Nonhelical
10	^3E-^3E-^3E	101	-175	155	71	171	67	-7	-120	-54	160	-171	161	6	4.49	Nonhelical
11	^3E-^3E-^3E	11	-158	-64	-75	175	72	5	-164	-55	-85	-169	73	28	4.88	DNA-A

the lowest energy found are listed in Table I. Conformations close to B-DNA (26) Watson-Crick-DNA (27) and A-DNA (28) type helical forms (Nos. 4,5 and 11 in Table I) are noted to lie, respectively, at energy barriers of 1.57, 2.98 and 4.88 kcal/mole above the lowest energy found.

Our main interest here is the lowest energy conformations found for dApdApdA which on repetition cannot promote a regular helical structure in the backbone of the polynucleotide chain (conformations other than 4, 5 and 11 in Table I). A detailed examination of these nonhelical low energy conformations of the trimeric unit were made by constructing wire models and stereopictures to extract information on the best possible bend promoting conformations. We identified four families of conformations which could characteristically introduce bends in the backbone and they are given in Table II along with parameters defining these bend structures.

The main feature that distinguishes the four families of bends is the disposition of the bonds at the two phosphate linkages of the trimeric unit, which is eventually manifested in the mutual orientation of the two arms containing the end residues about the central 'rigid nucleotide unit' within the two phosphate atoms. The 'U-bend' is the one in which the two end arms represented, respectively, by the virtual bonds $C(5')...P_1$ and $O(3')...P_2$ and the middle rigid nucleotide unit represented by the virtual bond $P_1...P_2$ lie more or less in plane and <u>cis</u> configuration . This 180° chain reversal has resulted due to the adoption of similar conformational domains (t) by ω_1' and ω_2' and of opposite domains (g^+ and g^-) by ω_1 and ω_2 angles. In this bend the two end atoms O(3') and C(5') approach

TABLE II

The Possible Kinds of Bends in Polydeoxynucleotides and Their Characteristic Parameters

Kind of bend	Sugar sequence of the chain section	Conformational domains of two consecutive phosphate linkages ω_1'	ω_1	ω_2'	ω_2	Distance between the end atoms * d(Å)	Dihedral angles between the bending arms about $P_1 \ldots P_2$ direction * θ (in degrees)
U-bend	$^3E\text{-}^3E\text{-}^3E$	t	g^+	t	g^-	4.3	8 (0 ± 40)
L_r-bend:							
Type 1	$^3E\text{-}^3E\text{-}^3E$	t	g^+	g^+	g^+	4.6	53 (90 ± 40)
Type 2	$^3E\text{-}^3E\text{-}^3E$	t	g^-	g^+	g^+	7.8	110 (90 ± 40)
Type 3	$^3E\text{-}^2E\text{-}^3E$	t	g^+	g^+	g^+	8.5	120 (90 ± 40)
L_l-bend:							
Type 1	$^3E\text{-}^3E\text{-}^3E$	t	g^+	g^-	g^-	6.3	-59 (-90 ± 40)
Type 2	$^3E\text{-}^3E\text{-}^3E$	t	g^+	g^-	t	6.4	-63 (-90 ± 40)
Z-bend:							
Type 1	$^3E\text{-}^3E\text{-}^3E$	g^-	g^-	g^+	g^+	8.2	152 (180 ± 40)
Type 2	$^3E\text{-}^3E\text{-}^3E$	g^-	g^-	t	g^-	11.9	149 (180 ± 40)

* The d and Θ values correspond to structures defined by dihedral angles noted in sequence Nos. 7,2,1,6,9,10,3 and 8 in Table I. See text also. The bracketed value in the Θ column denotes the expected range for the respective bend structure.

to the closest distance (d) possible. The dihedral angles defining the energetically best U-bend correspond to the values of conformation No. 7 in Table I, which is displayed in diagram b of Fig.1. Conformations orienting the $C(5')...P_1$ and $O(3')...P_2$ virtual bonds within the limit of $\theta = 0^\circ \pm 40^\circ$ could be taken, as the U-bend family.

The 'L_r-bend' is characterised by a right handed twist, $\theta \simeq 90^\circ$ between the $C(5')...P_1$ and $O(3')...P_2$ arms about the $P_1...P_2$ virtual bond which gives an 'L-shape' to the molecular unit. It was found from the low energy conformations listed in Table I, three of them have this characteristics: these three L_r-type bends adopt similar conformational domains (g^+) for ω_2' and ω_2 : Type 1 is characterised by ^3E-^3E-^3E sugar sequence and (ω_1', ω_1) = (t, g^+); Type 2 differs from Type 1 in adopting $\omega_1 = g^-$ instead of g^+ ; Type 3 differs from Type 1 in adopting the ^2E puckering for the central sugar unit instead of the ^3E puckering. It is very interesting to note that a change in the puckering of the middle sugar or a change (g^- to g^+) in the ω_1 domain produces more or less the same backbone course with a clockwise twist of θ around 110°-120° between the bending arms of the chain. The major difference between Type 1 and Type 2 or Type 3 bends is the distance between the end groups; Type 2 and Type 3 bends open up the chain to a distance almost double the value found for Type 1 bend as a consequence of the increased twist between the chain arms. The dihedral angles corresponding to the energetically best L_r-type bends are those given as conformation Nos. 1,2 and 6 in Table I, which are shown in diagrams c to e of Fig.1.

The 'L_l-bend' is characterised by a left-handed twist of $\theta \simeq -90^\circ$ between the end residue virtual bonds about the

central $P_1 \ldots P_2$ virtual bond. We notice two types of L_ℓ -bends, with identical chain diverting features but with a change in the ω_2 domain. The end group distances in the L_ℓ -bend types are in midway between U-bend and L_r-bend values. Conformation Nos. 9 and 10 in Table I represent the energetically best Type 1 and Type 2 - L_ℓ -bends (see diagrams f and g in Fig.1).

The final class of bend family, the 'Z-bend' is characterised by $\theta \simeq 180^\circ$, with a Zig-Zag 'Z-shape' backbone course. Although this kind of backbone course eventually keeps the chain direction unaltered, it introduces a 'kink' at the central residue. Two types of Z-bends are noted; while both the types have similar sugar sequences and similar ω_1' and ω_1 domains, Type 1 has $(\omega_2', \omega_2) = (g^+, g^+)$ whereas Type 2 has $(\omega_2', \omega_2) = (t, g^-)$ domains. Both the types of bends have similar θ values, but Type 2 is highly extended compared to Type 1 bend. Conformation Nos. 3 and 8 in Table I (diagrams h and i in Fig.1), correspond to the energetically best Z-type 1 and Type 2 bends. The Z-type 1 sharp bend conformation has been observed in the ApApA crystal (29) and also in the nucleotide sequence N74-N76 of the $tRNA^{Phe}$ crystal (12). The Type 2 bend conformation (No. 8 in Table I) is observed in the same nucleotide sequence in the model for the $tRNA^{Phe}$ crystal due to Quigley et al (13).

It is to be noted that all the four families of bends adopt ψ_1 and ψ_2 in the g^+ domain, with only one exception of L_ℓ-type 2 bend wherein we have $\psi_2 = t$. Another notable similarity in almost all the bend conformations is the absence of base stacking feature. The only exception here is the Z-type 2 bend. Because of the absence of base-stacking in the

U, L_r and L_ζ-bend families it is likely that these bends are equally probable in other types of base sequenced trimer units, however, with slightly altered stabilities. It is likely that there may be a few additional bends in trimer units with pyrimidine bases. Due to the enormous computational cost involved, we have not investigated these cases.

As to the possible sites of occurrence of the bends listed in Table II, the following may be stated from the knowledge of their chain diverting characters. The occurrence of the U-bend in the double helical portions is unlikely; L_r-bend Type 1 and Type 2 are very likely at the 5'-end of the polynucleotides; Type 1 - L_ζ-bend is likely to occur within the polynucleotide chain at a site where a conformational change takes place from A-DNA course to any other kind. Type 2 - L_ζ-bend is similarly possible at a site where a conformational change from Watson-Crick type structure to any other kind is needed. The Z-bends are possible both at the 3'-end of the polynucleotide chain as well as at the interior where an opening up of the chain occurs by way of a conformational change from A-DNA structure to any other kind.

REFERENCES

1. Venkatachalam, C.M. (1968) Biopolymers 6, 1425-1436.
2. Lewis, P.N., Momany, F.A. and Scheraga, H.A. (1973) Biochim. Biophys. Acta. 303, 211-229.
3. Crawford, J.L., Lipscomb, W.N. and Schellman, C.G. (1973) Proc. Natl. Acad. Sci. USA 78, 538-542.
4. Burgess, A.W., Ponnuswamy, P.K. and Scheraga, H.A. (1974) Israel Jnl. Chem. 12, 239-286.
5. Kovacic, R.T. and Van Holde, K.E. (1977) Biochemistry, 16, 1490-1498.

6. Finch, J.T., Lutter, L.C., Rhodes, D., Brown, R.J., Rushton, B., Levitt, M. and Klug, A. (1977) Nature 269, 29-36.

7. Stout, C.D., Mizuno, H., Rubin, J., Brennen, T., Rao, S.T. and Sundaralingam, M. (1976) Nuc. Acids Res. 3, 1111-1124.

8. Jack, A., Ladner, J.E. and Klug, A. (1976) J. Mol. Biol. 108, 619-649.

9. Quigley, G.J., Seeman, N.C., Wang, A.H.J., Suddath, F.L. and Rich, A. (1975) Nuc. Acids Res. 2, 2329-2339.

10. Sussman, J.L. and Kim, S.H. (1976) Biochem. Biophys. Res. Commun. 68, 89-96.

11. Crick, F.H.C. and Klug, A. (1975) Nature 255, 530-533.

12. Sobell, H.M., Tsai, C.C., Gilbert, S.G., Jain, S.C. and Sakore, T.D. (1976) Proc. Natl. Acad. Sci. USA. 73, 3068-3072.

13. Levitt, M. (1978) Proc. Natl. Acad. Sci. USA 75, 640-644.

14. Sussman, J.L. and Trifonov, E. (1978) Proc. Natl. Acad. Sci. USA 75, 103-107.

15. Namaradze, N.Z., Goryunov, A.N. and Birshtein, T.M. (1977) Biophys. Chem. 7, 59-70.

16. Yathindra, N. and Sundaralingam, M. (1974) Proc. Natl. Acad. Sci. USA 71, 3325-3328.

17. Kim, S.H. and Sussman, J.L. (1976) Nature 260, 645-646.

18. Thiyagarajan, P. and Ponnuswamy, P.K. (1978) Biopolymers 17, 533-553.

19. Thiyagarajan, P. and Ponnuswamy, P.K. (1978) Biopolymers 17, 2143-2158.

20. Ponnuswamy, P.K. and Thiyagarajan, P. (1978) Biopolymers 17, 2503-2518.

21. Thiyagarajan, P. and Ponnuswamy, P.K. (1979) Biopolymers 18, (to appear).

22. Ponnuswamy, P.K. and Thiyagarajan, P. (Biophysical Journal, Submitted).

23. Davidon, W.C. (1959) AEC. Res. and Development Report ANL-5990.

24. Fletcher, R. and Powell, M.J.D. (1963) Computer J. 6, 163-168.

25. Thiyagarajan, P. and Ponnuswamy, P.K. (Biophysical Journal, Submitted).

26. Arnott, S., Dover, S.D. and Wonacott, J. (1969) Acta Crystallogr. B25, 2192-2206.

27. Crick, F.H.C. and Watson, J.D. (1954) Proc. Roy. Soc. (Lond.) A223, 80-96.

28. Arnott, S. and Hukins, D.W.L. (1973) J. Mol. Biol. 81, 93-105.

29. Suck, D., Manor, P., Germain, G., Schwalbe, C., Weiman,G. and Saenger, W. (1973) Nature New Biol. 246, 161-165.