Volume 5 Number 1 January 1978 Nucleic Acids Research

Helix geometry of single stranded DNA 'A' and 'B' forms from minimum energy conformations of dimeric subunits

B. Hingerty† and S. Broyde*

†MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK and
*Biology Department, New York University, Washington Square, New York, NY 10003, USA

Received 30 September 1977

ABSTRACT

Low energy conformations with dihedral angles similar to those occurring in fibers of the 'A' and 'B' forms of DNAs have been calculated for the deoxydinucleoside phosphates dApdA, dCpdC, dTpdT, dGpdG and dGpdC (1-3). These conformers have been used as building blocks for generating larger single stranded polymers, whose helical parameters we have calculated. We find that single stranded 'A' and 'B' form helices tend to be narrower and more tightly wound than the duplexes obtained in fibers (4,5). This is consistent with experimental observations on single stranded fibers of poly (rC) (6). We also find that the different sequences have different helix geometries. In addition, it is observed that large variations in helix geometry for a given sequence are achievable at little energetic cost.

INTRODUCTION

The minimum energy conformations of the deoxydinucleoside phosphates dGpdC, dApdA, dCpdC, dGpdG and dTpdT have been calculated by classical potential energy methods in which the eight dihedral angles (see Figure 1 and Table I) and the deoxyribose puckering were the variables in a nine parameter minimization of the energy (1-3). Results on dGpdC with sugar pucker fixed were published earlier (7). In the present work we examine the properties of the helices generated from the 'A' and 'B' form low energy minima building blocks.

While the backbone torsion angles completely specify the geometrical properties of the helix, it is more illustrative to describe the helix in terms of its radius (r), the number of nucleotides per turn (n), and the rise per nucleotide residue along the helix axis (h). n is related to the turn angle or rotation per residue (t), where $n = 360^{o}/t$ $(0 \leq |t| \leq 180)$. We report characteristics n, h, r, and t of the helices generated from the 'A' and 'B' form minimum energy conformations of dGpdC, dApdA, dCpdC, dGpdG and dTpdT. We also calculate base stacking parameters of the deoxydinucleoside phosphates: Z, the mean perpendicular distance between adjacent base planes,

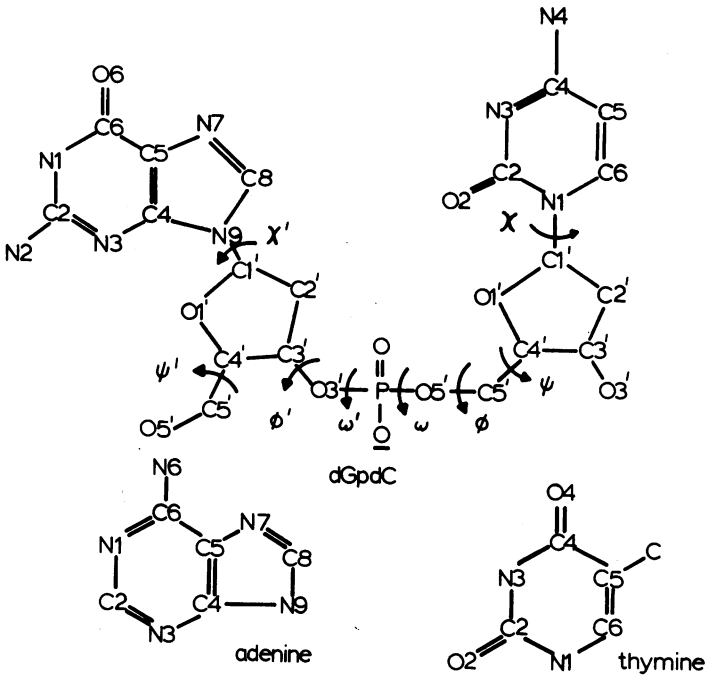

Figure 1. Structure, numbering scheme and conformational angle designations for dGpdC. Structure and numbering scheme for adenine and thymine.

TABLE 1. Dihedral Angle Definitions

Angle	
χ', χ	O1'-C1'-N9-C8 (pur)
	O1'-C1'-N1-C6 (pyr)
ψ', ψ	C3'-C4'-C5'-O5'
φ'	P-O3'-C3'-C4'
φ	C4'-C5'-O5'-P
ω'	O5'-P-O3'-C3'
ω	C5'-O5'-P-O_3'

All angles A-B-C-D are measured by a clockwise rotation of D with respect to A, looking down the B-C bond. A eclipsing D is 0°. Note that the angle ψ' is the exocyclic C4'-C5' bond (Fig. 1).

Λ, the mean dihedral angle between neighboring bases, and γ, the angle between the normal to the base plane and the helix axis.

The dihedral angles calculated for these single stranded 'A' and 'B' forms are quite similar to those obtained in duplex DNA fibers (4,5).

However, we find that these single stranded helices generally have a smaller number of nucleotides per turn and a narrower radius than the duplexes examined by fiber diffraction. Single stranded 'A' form poly (rC) fibers do have a similar low n and r (6). This agreement suggests that the calculated 'A' and 'B' form deoxydinucleoside phosphate conformations are representative building blocks for the ordered, single stranded polymers even though "end effects" occur, e.g. in the non-identical values of χ' and χ.

METHODS

Details on the calculation of the potential energies are to be found elsewhere (3).

The helical parameters were evaluated using the method developed by Mizushima and Shimanouchi (8a) for polypeptides, which was detailed for polynucleotides by Olsen (9). The base stacking parameters were also calculated as per Olson (9). In order to obtain Z, the mean perpendicular distance between base planes, we calculated Z_a and Z_b, the perpendicular distance from the center of one base to the other base, and averaged these quantities (Figure 2). We also calculated Z', the distance between base centers, which is a measure of the base overlap. In a helical dinucleoside phosphate the glycosidic torsion angles χ' and χ need not be identical, and the same holds true for ψ', the exocyclic C4'-C5' rotation and ψ, the backbone C4'-C5'. However, in a longer helical polymer the members of each pair

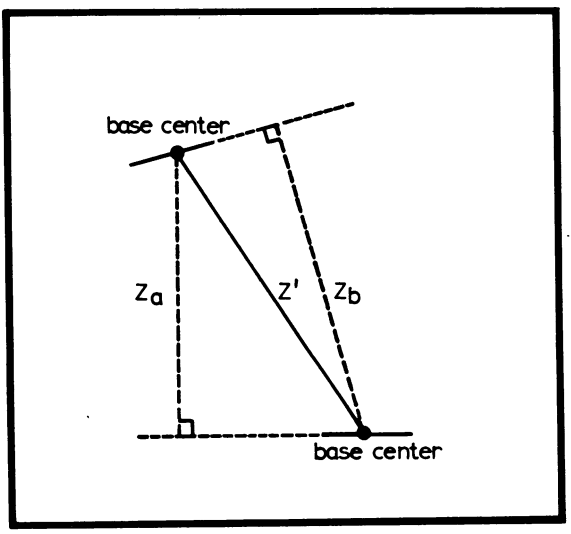

Figure 2. Definition of distances between base planes given in Table II. Z is the average of Z_a and Z_b.

are the same. Therefore, for the calculation of the helical parameters Ψ' was equated to Ψ. Base stacking parameters are reported for the deoxydinucleoside phosphates themselves. In most cases these parameters are only slightly different when the average of χ' and χ is employed. To calculate γ, χ' and χ had to be averaged, but the other helical parameters are independent of the mean χ.

The calculations for the helical parameters were done on the IBM/370 at the MRC Laboratory of Molecular Biology, Cambridge, England. The potential energy calculations, coordinate generation and the ORTEP drawings were made at the Georgia Institute of Technology on the CDC Cyber 6600 computer.

RESULTS AND DISCUSSION

Base Stacking

Figures 3 and 4 show the 'A' and 'B' forms calculated for each molecule. The 5' base is placed rigorously in the x-y plane, so this view illustrates the degree and kind of overlap between the bases. The differences between Z' and Z in Table II, which summarizes our results, are a quantitative measure of the extent of overlap. It is of interest that there is little over-

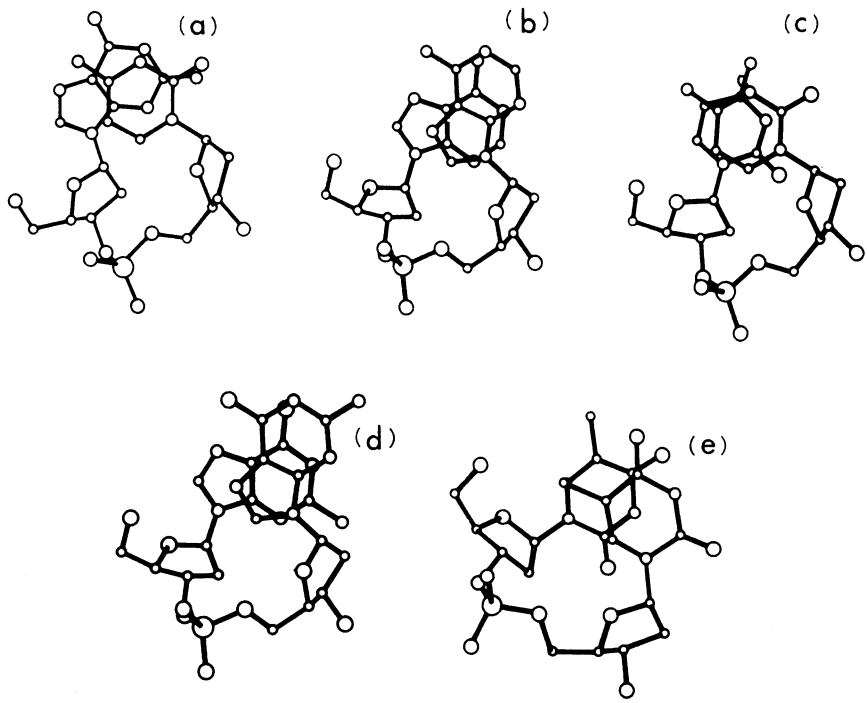

Figure 3. 'A' Forms. (a) dGpdC (b) dApdA (c) dCpdC (d) dGpdG (e) dTpdT

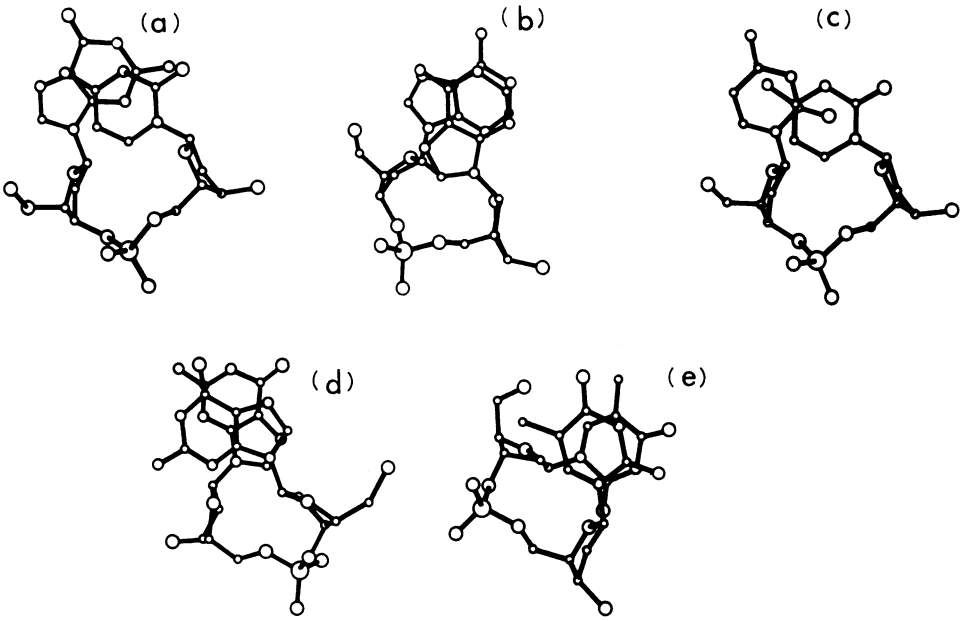

Figure 4. 'B' Forms. (a) dGpdC (b) dApdA (c) dCpdC (d) dGpdG (e) dTpdT

lap of base rings in the 'B' form of dCpdC, which is in agreement with experimental observations (10). However, stacking of heteroatoms above the aromatic rings is noted. There is considerably less intrastrand stacking in multiple strand structures than in the calculated single strands, due to the constraints of base pairing. This may be seen from a comparison of Z' and Z, calculated for some of the multiple strand structures obtained from fiber diffraction. The bases in our 'A' and 'B' form conformers are close to parallel. The deviation from planarity, measured by Λ, the mean dihedral angle between neighboring bases, is 21° at most, and generally in the range of $\sim 5° - 10°$. However, because the bases are not exactly parallel, Z_a and Z_b are not equal, and the difference between them is also a measure of the tilt. The calculated Z's, the averages of Z_a and Z_b, are in the range of 2.9 to 3.3 Å, except for the 'B' form of dTpdT, for which steric restrictions of the thymine methyls prevent the bases from approaching more closely than 4.7 Å. Thus, although the bases overlap, there is no stacking interaction at this distance. This might account, in part, for the little stacking observed for dTpdT in solution (11).

Helix Handedness

The ratio $t/h < 0$ defines a left-handed helix, while $t/h > 0$ defines a

TABLE II. Characteristics of 'A' and 'B' Form Helical Conformations

Molecule	Helix Type	χ′	Ø′	ψ′	ψ	Ø	Ψ	χ	P[a]	ΔE[b] kcal/mole	n	t	h Å	r Å	γ	Δ	Z_a Å	Z_b Å	Z Å	Z′ Å
		degrees										degrees			degrees	degrees				
dGpdC	'A'	9	197	302	283	187	49	32	17	0.	8.74	41.2	2.32	7.54	11.7	11.9	2.94	3.08	3.01	3.12
	'B'	65	172	257	302	191	46	65	169	0.	8.53	42.2	3.28	8.22	9.5	6.0	3.03	3.12	3.08	3.22
dApdA	'A'	5	208	312	280	183	47	21	7	0.2	5.74	62.7	2.43	4.93	8.9	4.9	2.86	2.98	2.92	3.26
	'B'	79	168	261	303	185	52	83	177	0.	6.51	53.3	3.10	6.65	5.8	7.3	3.10	3.03	3.07	3.12
dTpdT	'A'	10	201	316	278	183	51	24	19	0.	6.19	58.1	2.44	5.44	8.8	6.3	3.03	3.13	3.08	3.26
	'A'(alt)	44	192	302	283	177	55	43	13	0.1	8.77	41.0	2.15	8.21	19.8	18.9	3.23	3.16	3.20	3.25
	'B'	41	193	279	300	171	59	75	173	2.6	4.33	83.1	4.29	4.19	-16.7	19.6	4.70	4.67	4.69	5.18
dCpdC	'A'	5	207	324	273	182	50	22	14	2.2	4.97	72.5	2.37	4.41	10.1	8.2	2.84	2.99	2.92	3.11
	'A'(alt)	41	195	307	281	181	54	37	20	3.5	8.01	45.0	2.37	7.21	23.6	20.6	3.29	3.37	3.33	3.41
	'B'	66	173	255	299	184	51	66	164	2.3	8.27	43.5	2.79	8.46	8.9	4.3	2.83	2.95	2.89	3.36
dCpdG	'A'	6	207	311	281	183	47	24	7	6.7	5.92	60.8	2.46	5.08	8.8	7.3	2.79	3.00	2.90	3.36
	'B'	41	188	274	294	181	50	85	166	4.8	5.54	65.0	3.87	5.23	15.2	19.7	3.38	3.12	3.25	3.39
DNA(4,5)	'A'	23	178	313	275	208	45	23	14		11.	33.	2.56	8.86	20.2	11.4	3.41	2.97	3.19	4.05
	'B'	80	155	264	314	213	36	80	192		10.	36.	3.37	9.23	6.3	3.6	3.26	3.32	3.29	3.73
poly(dG-dC)·poly(dG-dC) (13)	'B'	8	141	259	298	208	69	8	192		8.	45.	3.03	7.86	-16.0	12.6	3.21	3.45	3.33	3.90
poly(dA)·poly(dT) (14)	'B'	82	145	273	308	224	39	82	192		10.	36.	3.29	8.78	8.0	4.9	3.13	3.26	3.20	3.63
poly(dT)·poly(dA)·poly(dT) (14)																				
poly(dA)	'A'	29	196	294	310	179	40	29	14		12.	36.	3.26	9.18	7.2	4.1	3.52	3.55	3.54	3.86
poly(dT_1)	'A'	26	192	299	297	186	47	26	14		12.	36.	3.26	9.36	9.1	3.7	3.50	3.53	3.52	4.04
poly(dT_2)	'A'	23	205	284	312	172	43	23	14		12.	36.	3.26	9.12	8.6	4.5	3.58	3.50	3.54	4.04
poly(rC)	'A'	19	231	295	303	173	47	19	19		6.	60.	3.11	6.7	21.4	21.3	3.57	3.66	3.62	3.86

a pseudorotation parameter of the sugar pucker (23)

b. energy difference between the 'A' or 'B' form local minimum and the lowest energy conformation in the same puckering domain

right-handed one (8b), ($0 \leq |t| \leq 180$). Of interest is the finding that t and h are positive in every instance, so that all the helices are right-handed. This is, of course, consistent with the situation obtained in all DNA fibers. However, it is remarkable in light of the fact that left and right handed helices in both the 'A' and 'B' form conformational domains are separated by only small differences in conformational angles (9,12).

Number of Nucleotides Per Turn

The familiar DNA double helices are usually identified as the 10-fold 'B' form (5), or the 11-fold 'A' form (4). However, Table II shows that particular sequences adopt other values of n in the duplex; for example, the 8-fold poly(dG-dC)·poly(dG-dC) (13) is a 'B' type known as 'D' DNA, and the 12-fold poly(dT)·poly(dA)·poly(dT) (14) is an 'A' type. The 'A' genus of helices are generally less variable in n than the 'B' (15), being 11 or 12-fold in the more recent structures. However, poly AH·poly AH^+ (16) was reported to be 8-fold. The 'B' genus exhibits values of n between 8 and 11.5 (15). Looking at our calculated minimum energy conformations, we see that the 'A' forms all have torsional angles very similar to those of duplex 'A' DNA (4). The ω', ω angle pair in the calculated conformations are within $\sim 10^{\circ}$ of those in the 'A' DNA fibers. Yet n is generally lower than in the double stranded fibers, taking on values between 5 and 9. A similar situation obtains with the 'B' forms. The calculated conformers are very similar to 'B' DNA. They have the higher values of χ and ω, and the lower values of ω' characteristic of double stranded 'B' DNA fibers as compared with the 'A' type (4,5). Nonetheless, n is between 4 and 9. It is clear that this parameter of the helix is very sensitive to seemingly small changes in conformation. The source of the flexibility lies primarily in the angles ϕ' and ϕ, which differ by as much as $\sim 25^{\circ}$-30° in the calculated conformations from those in the 'A' and 'B' DNA fibers. The n - h plots of Yathindra and Sundaralingam (12) show that flexibility in ω' over a range of $\sim 50^{\circ}$ in the 'A' or 'B' form region also permits such variability in n; however ω', in our calculated conformations does not differ greatly from the value in fibers. The model for single stranded poly rA (17), obtained from the conformation in the helical portion of the ApApA crystal (18) is 9-fold and also differs most from 11-fold 'A' RNA in ϕ. Another model of poly rA, based on potential energy calculations, predicts that this polymer is 11 or 12-fold (19). The only single stranded polymeric structure available from fiber diffraction analysis is the 'A' form poly (rC) (6). It is, most interestingly, 6-fold, and the high turn angle

is, in fact, due in large part to an increase in ϕ' as compared with double stranded 'A' forms. Thus, available experimental evidence supports these calculated results. Experimental results for fibers (21,22) as well as our calculations (3,7,20) find the ribo and deoxy nucleic acids in the 'A' form to have very similar conformations, so that a comparison between experimental results for ribopolymers and the present results are worthwhile.

Helix Radius

The helix radius, r, is correlated with n, wider radii being associated with a larger number of nucleotides per turn. Figure 5 shows n versus r for our calculated conformations, together with recent experimental results for fibers, as well as the poly rA model from the crystal structure (17). An approximately linear relationship is discerned. The narrower, more tightly wound (low n) helices are found in our calculated single stranded structures and in fibers of poly rC. Structures possessing more than one strand are wider and more loosely wound. This change in n and r, needed to form multistranded structures, can be achieved at little energetic cost by small changes in the torsion angles within their preferred ranges. Table II shows what the magnitude of these energies may be. Both dTpdT and

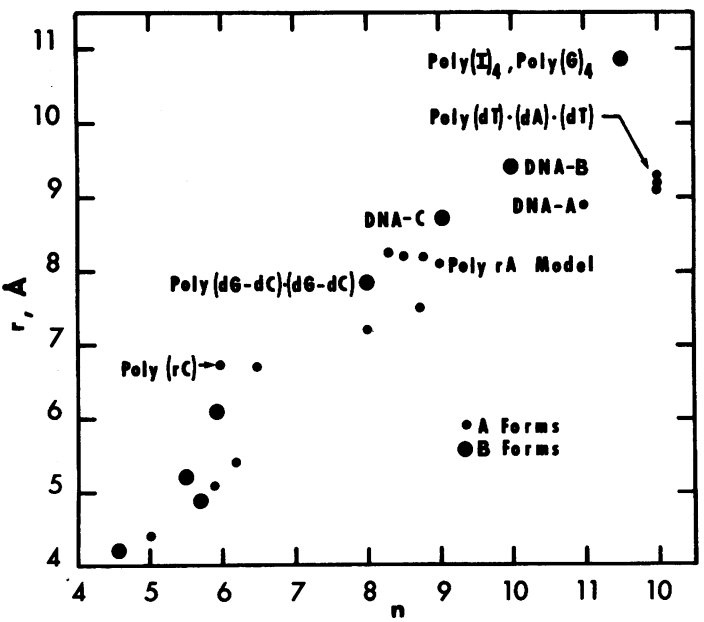

Figure 5. The number of nucleotides per turn, n, as a function of helix radius for calculated and observed conformations. Data for calculated conformations are from Table II. Observed conformations are labelled; data for these are from references 4, 5, 6, 13, 14 and 24.

dCpdC have two 'A' forms, the lower energy form being narrower and having fewer nucleotides per turn. Yet only 0.1 kcal./mole separates the six-fold from the nine-fold 'A' form in dTpdT, and 1.3 kcal./mole separates the five and the eight-fold helices in dCpdC. Cross strand stacking interactions and hydrogen bonds in the duplex could supply the energy needed for these rearrangements.

Axial Rise Per Residue

The quantity h, the axial rise per residue, has a lower value in fibers of duplex 'A' DNA than in 'B' DNA. In the former, h is 2.56 Å, while in the latter it is 3.37 Å (15). However, h is not consistently greater in all 'B' form fibers. Instead, the value of h is sequence dependent. For example, in the 'B'-type 'D' DNA form of poly(dG·dC)·(dG·dC) (13), h is 3.03 Å, while in the triple stranded poly(dT)·poly(dA)·poly(dT) (14), which is 'A' form, h is 3.26 Å. Nevertheless, our calculated 'A' form conformations for single strands all have lower values of h than the 'B' form of the same sequence. Possibly the lower rise per residue is, in fact, intrinsic to 'A' forms, reflecting the lesser local extension of the C(3')-endo sugars (25). Tertiary interactions in duplex fibers of certain sequences may overcome these inherent predilections.

Angle Between the Normal to the Base Planes and the Helix Axis

In double stranded structures h is related approximately linearly to the angle γ that the normal to the base planes makes with the helix axis, by the empirical relationship $\gamma = 69\text{-}19(h)$ (15). In 'A' DNA fibers a large tilt angle, γ, of 20° is observed while the tilt in 'B' DNA fibers is only 6°; however, the lower γ is not a persistent feature of 'B' forms, but follows the sequence dependence of h. In single stranded poly rC (6), the linear relationship between γ and h is not maintained. The same is true of the calculated single stranded conformers. The calculated 'A' forms do generally, but not invariably, have a higher absolute value of γ than the corresponding 'B' forms; however, the differences are small. Therefore concomitant with a lower h, a higher γ may also be preferred by single stranded 'A' forms, although the two properties are not linearly related in single strands.

CONCLUSION

The predicted dGpdC 'B' type single stranded helix is very similar to the 'B' type 'D' DNA observed in fibers of poly(dG-dC)·poly(dG-dC) (13). However, in general, our calculations on deoxydinucleoside phosphates show that single stranded 'A' and 'B' form helices tend to be narrower and more

tightly wound than the duplexes. This is in agreement with the fiber diffraction studies of single stranded poly (rC) (6) which is six-fold. Intrastrand strand stacking is generally greater in the calculated single strands than in multiple strand structures. Only small alterations in the conformational angles are needed to produce these marked differences between the overall geometric properties of the single stranded helix and those of the duplex. Furthermore, the changes which are needed to permit Watson-Crick base pairing require only small amounts of energy. It is conceivable that similar conformational alterations may occur spontaneously when the duplex "breathes". Such conformational changes could produce local variations in the helix radius and pitch (n x h) that are characteristic of specific sequences, since the geometric properties of the helices in the different sequences examined vary. These irregularities might then be recognizable to site specific proteins.

ACKNOWLEDGEMENTS

Grateful acknowledgement is made to the National Institute of Health for support of this work under Public Health Service Grant GM 22765-01, awarded to S.B. B.H. would like to thank the National Institutes of Health for a Postdoctoral Fellowship, 1F22CA02210-01. Special thanks are due to Prof. R.M. Wartell, School of Physics, Georgia Institute of Technology, for many interesting discussions, and particularly for suggesting that the geometries of nucleic acid helices might be base sequence dependent. We are also pleased to acknowledge the original inspiration and programming on dinucleoside phosphate energy minimization by Dr. Steven D. Stellman, now at American Health Foundation, New York, N.Y. 10019, which led to the present work.

A preliminary account of this work was presented at the 21st Annual Meeting of the Biophysical Society (2).

REFERENCES

1. Broyde, S., Wartell, R. and Stellman, S. (1976) Biophys. J. 16, 143a.
2. Broyde, S., Hingerty, B., Wartell, R. and Stellman, S. (1977) Biophys. J. 17, 113a.
3. Broyde, S., Wartell, R., Stellman, S. and Hingerty, B. (in press) Biopolymers.
4. Arnott, S. and Hukins, D. (1972) Biochem. Biophys. Res. Commun. 47, 1505-1509.
5. Arnott, S. and Selsing, E. (1975) J. Mol. Biol. 98, 265-269.
6. Arnott, S., Chandrasekaran, R. and Leslie, A. (1976) J. Mol. Biol. 106, 735-748.
7. Broyde, S., Stellman, S. and Wartell, R. (1975) Biopolymers 14, 2625-2639.

8a. Mizushima, S. and Shimanouchi, T. (1961) Adv. in Enzymology 23, 1-27.
8b. Sugeta, H. and Miyazawa, T. (1967) Biopolymers 5, 673-679.
9. Olson, W.K. (1976) Biopolymers 15, 859-878.
10. Brahms, J., Maurizot, J. and Pilet, J. (1969) Biochim. Biophys. Acta 186, 110-123.
11. Wood, D., Hruska, F. and Ogilvie, D. (1974) Can. J. Chem. 52, 3353-3366.
12. Yathindra, N. and Sundaralingam, M. (1976) Nucleic Acids Res. 3, 729-747.
13. Arnott, S., Chandrasekaran, R., Hukins, D., Smith, P. and Watts, L. (1974) J. Mol. Biol. 88, 523-533.
14. Arnott, S. and Selsing, E. (1974) J. Mol. Biol. 88, 509-521.
15. Arnott, S., Chandrasekaran, R., and Selsing, E., (1975) in Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions, M. Sundaralingam and S.T. Rao, Eds., pp. 577-596, University Park Press, Baltimore, Md.
16. Rich, A., Davies, D., Crick, F. and Watson, J. (1961) J. Mol. Biol. 3, 71-86.
17. Saenger, W., Riecke, J., and Suck, D. (1975) J. Mol. Biol. 93, 529-534.
18. Suck, D., Manor, P., Germain, G., Schwalbe, C., Weimann, G. and Saenger, W. (1973) Nature New Biol. 246.
19. Fujii, S. and Tomita, K. (1976) Nuc. Acids Res. 3, 1973-1984.
20. Broyde, S., Wartell, R., Stellman, S., Hingerty, B. and Langridge, R. (1975) Biopolymers 14, 1597-1613.
21. Milman, G., Chamberlain, M. and Langridge, R. (1967) Proc. Nat. Acad. Sci. U.S. 37, 1804-1810.
22. O'Brien, E. and MacEwan, A. (1970) J. Mol. Biol. 48, 243-261.
23. Altona, C. and Sundaralingam, M. (1973) J. Am. Chem. Soc. 95, 2333-2344.
24. Arnott, S. and Hukins, D. (1972) Biochem. J. 130, 453-465.
25. Sundaralingam, M. Op. cit. ref. 15, pp. 487-523.