View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Publications of the IAS Fellows

Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci.,
Vol. 8, Nos 3 & 4, August 1985, pp. 603-614. © Printed in India.

Comparison of interproton distances in DNA models with nuclear Overhauser enhancement data

B. GOPALAKRISHNAN and MANJU BANSALs

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

Abstract. The conformational flexibility inherent in the polynucleotide chain plays an important role in deciding its three-dimensonal structure and enables it to undergo structural transitions in order to fulfil all its functions. Following certain stereochemical guidelines, both right and left handed double-helical models have been built in our laboratory and they are in reasonably good agreement with the fibre patterns for various polymorphous forms of DNA. Recently, nuclear magnetic resonance spectroscopy has become an important technique for studying the solution conformation and polymorphism of nucleic acids. Several workers have used 1_H nuclear magnetic resonance nuclear Overhauser enhancement measurements to estimate the interproton distances for the various DNA oligomers and compared them with the-interproton distances for particular models of A and B form DNA. In some cases the solution conformation does not seem to fit either of these models. We have been studying various models for DNA with a view to exploring the full conformational space allowed for nucleic acid polymers. In this paper, the interproton distances calculated for the different stereochemically feasible models of DNA are presented and they are compared and correlated against those obtained from ^{1}H nuclear magnetic resonance nuclear Overhauser enhancement measurements of various nucleic acid oligomers.

Keywords. DNA models; interproton distances; nOe data.

Introduction

Based on theoretical conformational energy calculations and taking into account the experimentally observed values for the six backbone torsion angles as well as the glycosidic torsion angle, Sasisekharan *et al.* (1971) and Sundaralingam's group (Sakurai and Sundaralingam, 1971) had first examined the possible combinations which can lead to double helical structures with mononucleotide repeating units. Recent detailed calculations (Gupta *et al.*, 1982) reveal that there are only 8 distinctly different combinations which lead to the formation of duplex structures of both right and left handedness, designated as RU and LU helices. Of these 8 conformations 1 is sterically disallowed for a polynucleotide, 5 correspond to the *trans* conformation about both the O5′–C5′ and C3′–O3′ bonds while in the other 2 the torsion angle about the C3′–O3′ bond is in the g^- conformation. In all the 7 sterically allowed helical domains, the glycosyl torsion angle χ falls in the anti region ($170° < \chi < 260°$) and hence both purines and pyrimidines can be accommodated in the RU and LU helices of these domains. Sasisekharan *et al.* (1981) have been able to obtain sterically acceptable models of both right and left handedness for the A, B and D forms of DNA. Recently Premilat and Albiser (1984), have proposed a left handed model for C–DNA which is very similar to the LU model for B–DNA, proposed in Sasisekharan *et al.* (1983). Models of both handedness give almost equally good agreement with the observed fibre diffraction pattern for their respective polymorphic forms and they have also been

found to agree equally well with the infra-red data. Single crystal structure analysis of short oligonucleotides has also conclusively shown that the backbone of a polyucleotide chain is inherently very flexible and can adopt many different conformations (Dickerson and Drew, 1981).

In recent days the proton-proton nuclear Overhauser enhancement (nOe) measureent has emerged as a potentially very powerful technique for determining the conformation of biomolecules in solution. It has been used to elucidate the general structural features of several oligonucleotides by comparing the observed proximity of some well identified protons with the coordinates of the A and B forms of DNA as given by Arnott and Hukins (1972). However, to date only one attempt has been made,by Gronenborn *et al.* (1984), to characterize in detail the solution conformation of two oligonucleotides, 5′d(C-G-T-A-C-G) and 5′d(A-C-G-C-G-C-G-T), based on a detailed calculation of the interproton distances and model building. Since in this case as well as in other studies the interproton distances observed in the solution state do not seem to agree well with either the A or B form, as proposed by Arnott's group we have undertaken a detailed model-building study to calculate the interproton distances for a variety of sterically allowed, conformationally favourable polynucleotides structures. As a first step in this study we have calculated the proton-proton distances for the various models proposed for the A and B forms of DNA and compared them with those obtained from the nOe measurements of nucleic acid oligomers.

Method and models

The conformational parameters for the seven models considered by us are listed in table 1. All these models have been refined against the available fibre diffraction data for the A and B forms of DNA. The models include two models for A-DNA, one right handed model (Arnott and Hukins, 1972) and one left handed model, obtained by us as being sterically acceptable, though it does not agree as well as the right handed model with the A- DNA fibre pattern. For the B form DNA five models have been considered, including one left handed and four right handed models. All the models have been refined against the B-form fibre diffraction data and represent the best models in their respective domains. The B-DNA model of Arnott and Hukins (1972) has some stereochemical problems, but has been included for completeness and also since most workers continue to use these coordinates as representative of the B-form DNA. The best right and left handed models for the B-form as obtained in our laboratory (Sasisekharan *et al.,* 1983) are stereochemically satisfactory, and also in excellent agreement with the available fibre-diffraction and infra-red data.

The RU models in the C2′-*endo* domain and C3′-*endo* domain are similar to the models of Arnott and Hukins (1972) and their revised model as given in Reid *et al.,* (1983). The LU model also has a C2′-*endo* sugar pucker, but has the unusual *gauche*$^-$ onformation about the C3′-O3′ bond. Following the stereochemical guidelines outlined in Gupta *et al.* (1982), the orientations about the phosphodiester bonds O3′–P and P–O5' are *trans* and *gauche*$^+$ respectively. This LU model is much superior to our earlier model (Gupta *et al.,* 1980) which had the more common *trans* conformation about the C3′–O3′ bond and the *trans, gauche*$^-$ conformations about the phos-

Table 1. Conformation angles (according to IUPAC-IUB Nomenclature, 1983) for the various DNA models considered in this study.

		Right handed models				Left handed models		
		B-Form				A-Form	B-Form	A-Form
Torsion angle	Bond involved	Arnott and Hukins (1972)	Sasisekharan *et al.* (1983) C2′-*endo* RU helix	Gupta *et al.* (1980) C3′-*endo* RU helix	Gupta *et al.* (1980) Alternating E^2–E^3 helix	Arnott and Hukins (1972)	Sasisekharan *et al.* (1983) C2′-*endo* LU helix	Sasisekharan *et al.* (unpublished results)
α	P–O5′	−46	−47	−91	−63/−74	−85	51	−55
β	O5′–C5′	−146	141	179	147/−179	−152	−132	136
γ	C5′–C4′	36	40	75	49/66	45	160	−45
δ	C4′–C3′	156	141	97	156/89	83	136	158
ε	C3′–O3′	155	−135	−176	−155/−177	178	−63	−72
ζ	O3′–P	−96	−152	−68	−150/−68	−47	178	156
χ	C1′–N9	−98	−110	−145	−122/−156	−154	−171	−159

phodiester bonds. The RU and LU models in the C2′-*endo* domain give equally good agreement with the fibre diffraction pattern as well as infra-red data (see Sasisekharan *et al.*, 1983, for details). The LU model for C-DNA (Premilat and Albiser, 1984) has the identical g– conformation about the C3′–O3′ bond and tg^+ orientations about the phosphodiester bonds O3′–P and P–O5', as our LU model for B-DNA. The other torsion angles are also in the same regions. The model with C3′-*endo* sugar pucker is also a stereochemically satisfactory structure but is not in as good agreement with the observed X-ray data as the above two models in the C2′-*endo* domain. We have also included one dinucleotide repeating model for B-DNA with alternate C2′-*endo* and C3′-*endo* sugar puckers which is similar to that proposed by Klug *et al.* (1979), and gives reasonably good agreement with the B-DNA fibre data (Gupta *et al.,* 1980).

The interproton distances for the different models of DNA were calculated after appropriately fixing the hydrogen atoms. The calculated interproton distances which are less than 6 Å and for which the nOe has been observed are listed in tables 2 and 3. They are compared with reported interproton distances obtained using ^{1}HNMR nOes in the hexamer 5"d(C-G-T-A-C-G), the octamer 5"d(A-C-G-C-G-C-G-T) (Gronenborn *et al.,* 1984) and a undecamer 5"d(A-A-G-T-G-T-G-A-C-A-T) (Clore and Gronenborn, 1984).

Results and discussion

A comparision of the calculated interproton distances for the various models clearly shows that the various models obtained by fitting to the X-ray fibre data can be grouped into two categories on the basis of the sugar puckers: *viz,* C2′-*endo* sugar puckered and C3′-*endo* sugar puckered structures. The C2′-*endo* sugar pucker domain contains structures of both handedness, the RU helices for B-DNA as well as the LU helices for the A and B-forms. However, the structures with different handedness can be distinguished on the basis of interproton distances between the base H8/H6 protons and the H2′/H2″ distances as well as the inter-base proton distances. Right handed structures with C2′-*endo* sugar pucker will show a large nOe between the H8/H6 atoms and H2′/H2″ since the H8/H6 proton is in close proximity and almost equidistant to the H2′ proton of its own sugar and the H2″ proton of the neighbouring sugar on the 5′-side as shown in figure 1. On the other hand, as shown in table 3, in the left handed models for both the A and B forms (figure 2) the H8/H6 protons are more than 3·4 Å from the H2′/H2″ protons in their own sugar as well as from the protons of the deoxyribose on the 5′-end. Hence the absence of any nOe in the H2′/H2″ region has been taken as implying the presence of a left handed structure, as reported by Gupta *et al.* (1983) for poly (dA-dT).

The major difference between the models with the C2′-*endo* sugar pucker and C3′-*endo* pucker (figure 3) is of course the fact that both the protons H2′ and H2″ on the sugar attached to the base are quite distant from the H 8 proton, though the interresidue H8 . . . H2′/H2″ distance is < 2·0 Å for all right handed models for the A and B-form DNA.

The alternating model for B-DNA shows features common to the two RU models with C2′-*endo* and C3′-*endo* sugar puckers with the interresidue distances being in

worse agreement with the observed ones than either of the other two models.

The interproton vectors characteristic of sugar pucker also show considerable variation as the pucker varies from C1′-*exo* to C4′-*exo*. The base-sugar intraresidue H8/H6 . . . H1′ distance does not vary much with the glycosidic torsion (χ) in the whole range for E^2 (C2′-*endo*) and E^3 (C3′-*endo*) sugar puckered models, which have χ varying from 170° to 100°. However, as mentioned above, the H8/H6 . . . H2′/H2″ distances differ significantly for the various models.

The calculated interproton distances of the non-exchangeable base and sugar protons, for the different models, therefore give a range ofcharacteristic values, some of which are within the range of observable nOe distances. However, most of the distances calculated from the nOe measurements appear to be shorter than those calculated for the various models which have been refined on the basis of X-ray fibre diffraction analysis.

A comparison of the calculated versus reported interproton distances is shown in tables 2 and 3. The interresidue interproton distances are within observable range only with the sugar protons on the 5′-side of the base proton, for all the models. Notably, some of the interresidue interproton distances are shorter than the corresponding intraresidue interproton vectors.

The r.m.s. differences between the calculated and the observed interproton distances have also been calculated for all the models, taking into account 14 intraresidue and 8 interresidue distances for the hexamer, 11 and 9 distances respectively for the octamer and 6 and 10 for the undecamer. These are also listed in table 2 for the right handed models.

An analysis of the r.m.s. differences calculated for the three deoxyoligonucleotides indicates no clear cut better agreement for any one of the models. However, in general the right handed models give a better agreement with the observed data as compared to the left handed models, which give an r.m.s. difference greater than 1·2 for all the three oligonucleotides. Among the right handed models, Sasisekharan *et al.*'s C2′-*endo* RU model gives a marginally better agreement than Arnott's models, for A and B-DNA, the C3′-*endo* model, or the alternating E^2-E^3 model. It is interesting to note that a comparison with the representative interproton distances for the crystal structure of the dodecamer d(C-G-C-G-A-A-T-T-C-G-C-G) which has an averaged C1′-*exo* structure, gives an almost equally good fit for the hexamer and the undecamer but a rather poorer fit for the octamer which has infact been interpreted as having a dinucleotide repeat (Clore and Gronenborn, 1983). Most of the variable torsion angles in this model have rather uncommon conformations. In particular, the orientation about the C3′–O3′ bond is predicted to be g_+ for pyrimidine nucleotides and g^+/t for purine nucleotides, which is a feature not observed for any oligo or polynucleotide structure.

Hence it may be concluded that the observed interproton distances show a similarity with the overall features of the right handed B-DNA models as obtained from fibre diffraction or single crystal studies. The major difference between the C2′-*endo* and C3′-*endo* models, particularly in the r.m.s. difference for the undecamer, (also noted by Clore and Gronenborn, 1984) arises from the differences in the intraresidue vectors H8/H6 . . . H2′ and H8/H6 . . . H2″. Since in addition, the interresidue H8/H6 . . . H2′/H2″ and the H5 . . . H2′/H2″ vectors for the B-DNA models are in

Table 2. Comparison of calculated and observed interproton distances for right handed models. The r.m.s. differences calculated for the various models in comparison with the observed nOe values are also given.

	Calculated distances						Observed distances (from NMR)		
Interproton vector	Arnott's B-Form*[a]	C2′-*endo* RU Helix[b]	C3′-*endo* RU Helix[c]	Alternating E^2–E^3[c]	Dodecamer[d]	Arnott's A-Form[a]	Hexamer[e]	Octamer[e]	Undecamer[f]
(i) Intraresidue									
H1′-H2′	3·0 (3·1)	3·0	2·7	3·0/2·7	3·1	2·7	2·5	2·3	2·3
H1′-H2″	2·3 (2·6)	2·4	2·3	2·4/2·3	2·4	2·3	2·2	2·1	2·1
H1′-H4′	3·7	3·2	3·4	3·4/3·3	3·0	3·4	2·9	—	—
H2′-H3′	2·4	2·4	2·4	2·4/2·4	2·4	2·4	2·2	2·2	—
H2″-H3′	2·7	2·7	3·1	2·7/3·0	2·8	3·0	2·3	2·2	—
H2″-H4′	3·9	4·0	3·0	4·0/2·9	4·0	2·8	3·0	—	—
H3′-H4′	2·7	2·8	3·0	2·8/3·0	2·8	3·0	2·6	2·4	—
H3′-H5″	2·8	2·8	2·6	2·4/3·0	2·5	3·1	2·4	—	—
H1′-H8	3·8 (3·8)	3·8	3·7	3·8	4·0	3·9	3·8	3·1	3·2
H2′-H8	2·2 (2·3)	2·1	3·8	3·9	2·3	3·9	2·1	2·1	2·0
H2″-H8	3·7 (3·5)	3·5	4·8	4·8	3·5	4·7	—	—	—
H3′-H8	4·2	4·2	3·4	3·3	4·6	2·9	3·2	2·6	—
H1′-H6	3·6 (3·7)	3·7	3·5	3·7	3·7	3·6	3·8	3·2	3·6
H2′-H6	2·4 (2·0)	2·0	3·6	2·4	2·3	3·7	2·2	2·0	2·0
H2″-H6	3·9 (3·3)	3·4	4·5	3·8	3·8	4·4	—	—	—
H3′-H6	2·6	4·1	3·1	4·3	3·6	2·6	2·9	2·9	—

(ii) Interresidue									
H3′-H8	4·9	4·0	3·3	3·9	4·6	3·3	—	—	—
H2′-H8	3·6 (3·9)	3·5	1·6	3·3	3·4	1·7	2·7	2·7	2·7
H2″-H8	2·2 (2·1)	1·7	2·8	1·7	2·5	3·3	2·6	2·3	2·3
H1′-H8	2·8 (3·6)	3·2	4·1	2·0	4·4	4·1	3·2	3·6	3·5
H3′-H6	5·0	4·0	3·4	3·4	4·8	3·5	—	—	—
H2′-H6	3·9 (4·0)	3·6	1·6	2·0	3·4	1·6	2·5	2·2	2·3
H2″-H6	2·4 (2·2)	1·8	2·6	3·4	2·1	3·1	2·5	2·1	2·3
H1′-H6	2·8 (3·5)	3·2	4·0	4·7	3·9	4·0	2·9	3·0	3·4
H2′-H5	3·2 (3·5)	3·2	2·6	3·4	3·2	3·0	—	2·4	—
H2″-H5	2·8 (3·0)	2·8	4·4	5·2	3·3	4·8	—	2·5	2·5
H1′-H5	3·7 (4·3)	3·8	4·9	5·6	4·7	4·8	—	—	3·6
H8-H5	3·8 (3·9)	4·0	4·2	4·7	4·4	3·7	3·5	3·5	3·8
H8-H6	5·2	5·1	5·2	4·6	4·4	4·6	3·6	—	4·1
H6-H8	4·7	4·8	4·8	5·2	5·2	4·5	3·7	—	—
R.M.S. Values:									
Intraresidue									
Hexamer	0·74	0·55	0·68	0·61	0·64	0·72			
Octamer	0·89	0·63	0·73	0·74	0·80	0·87			
Undecamer	0·44	0·40	1·02	0·78	0·51	1·09			
Interresidue									
Hexamer	0·87	0·88	0·96	1·16	0·82	0·83			
Octamer	0·77	0·68	0·83	1·40	0·76	1·07			
Undecamer	0·74	0·65	0·97	1·35	0·72	1·04			
Overall									
Hexamer	0·79	0·70	0·80	0·81	0·71	0·76			
Octamer	0·84	0·65	0·78	1·03	0·78	0·96			
Undecamer	0·65	0·57	0·99	0·87	0·65	1·06			

* The values within parentheses are listed by Reid *et al.* (1983) for the refined B-Form model proposed by Arnott and Chandrasekharan (private communication).

[a] Arnott and Hukins (1972); [b]Sasisekharan *et al.* (1983); [c]Gupta *et al.* (1980); [d]Dickerson and Drew (1981); [e]Gronenborn *et al.* (1984); [f] Clore and Gronenborn (1984).

Table 3. Comparison of calculated and observed interproton distances for left handed models.

Interproton vector	LU helix A-Form[a]	LU helix B-Form[b]	Hexamer	Octamer	Undecamer
(i) Intraresidue					
H1′-H2′	3·0	3·0	2·5	2·3	2·3
H1′-H2″	2·3	2·4	2·2	2·1	2·1
H1′-H4′	3·7	3·0	2·9	—	—
H2′-H3′	1·7	2·5	2·2	2·2	—
H2″-H3′	2·3	2·7	2·3	2·2	—
H2″-H4′	3·3	4·0	3·0	—	—
H3′-H4′	2·7	2·8	2·6	2·4	—
H3′-H5″	3·6	2·7	2·4	—	—
H1′-H8	3·8	3·5	3·8	3·1	3·2
H2′-H8	3·5	3·9	2·1	2·1	2·0
H2″-H8	4·6	4·7	—	—	—
H3′-H8	4·6	5·4	3·2	2·6	—
H1′-H6	3·6	3·3	3·8	3·2	3·6
H2′-H6	3·3	3·7	2·2	2·0	2·0
H2″-H6	4·4	4·5	—	—	—
H3′-H6	4·3	5·0	2·9	2·9	—
(ii) Interresidue					
H8-H3′	5·7	5.2	—	—	—
H8-H2′	5·5	5·3	2·7	2·7	2·7
H8-H2″	4·7	4·4	2·6	2·3	2·3
H8-H1′	3·5	2·5	3·2	3·6	3·5
H6-H3′	5·7	5·1	—	—	—
H6-H2′	5·6	5·4	2·5	2·2	2·3
H6-H2″	4·6	4·4	2·5	2·1	2·3
H6-H1′	3·6	2·5	2·9	3·0	3·4
H5-H2′	> 6·0	5·5	—	2·4	—
H5-H2″	> 6·0	5·1	—	2·5	—
H5-H1′	4·6	2·7	—	—	3·6
H5-H8	5·4	> 6·0	3·5	3·5	3·8
H6-H8	> 6·0	4·8	3·6	—	4·1
H8-H6	5·4	5·3	3·7	—	-

[a] Sasisekharan, V., Bansal, M. and Gupta, G. (unpublished results).
[b] Sasisekharan *et al.* (1983).

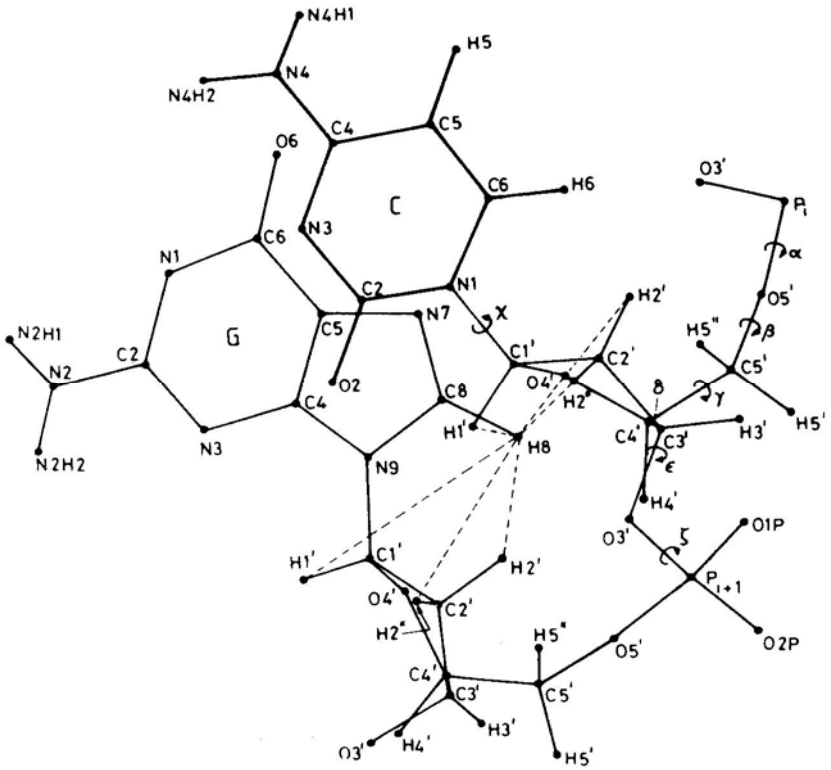

Figure 1. An x-y projection of 5 '-pCpG in the B-DNA RU model, showing the interproton base-sugar vectors, < 4 A, involving the H8 proton of the base.

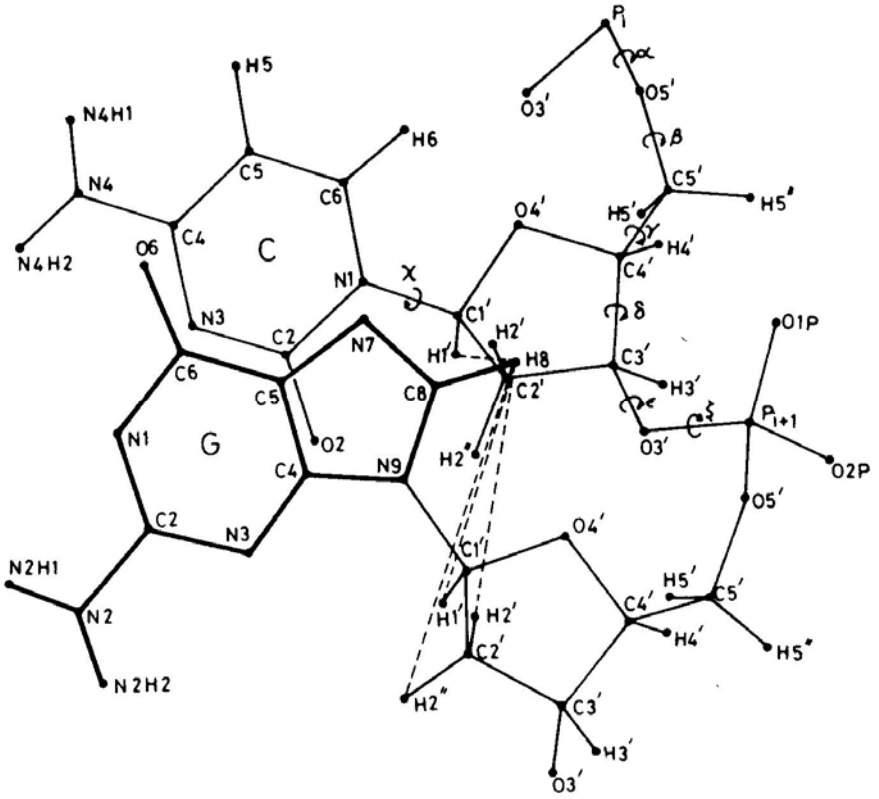

Figure 2. Same as in figure 1, but for the B-DNA left-handed model.

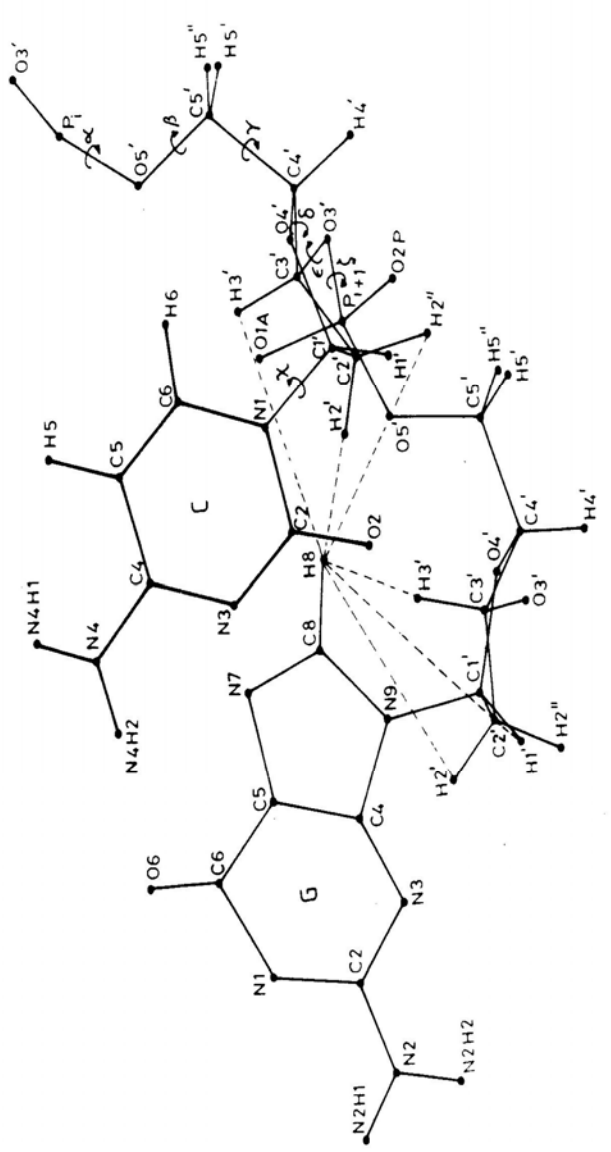

Figure 3. Same as in figure 1, for Arnott's right-handed A-DNA model.

better agreement with the observed distances, the B-DNA model naturally gives a much lower r.m.s. difference, than the C3' *endo* A-DNA model.

For the hexamer and the octamer the agreement between the calculated and observed distances is not really satisfactory for any of the models obtained from fibre diffraction analysis. Hence it will be interesting to investigate whether the solution structure is an averaged structure of several slightly varying conformations or is infact a structure, different from that existing in the solid state, such as that seen for a Li-DNA fibre, which gives a crystalline pattern. Much more experimental nOe data as well as model-building studies need to be analysed before any definite conclusion can be reached on the details of the solution structure of oligonucleotides.

However, it is interesting to note that the C2′-*endo* RU helix, with mononucleotide repeat, which has been optimized to fit the X-ray fibre diffraction pattern and the available infra-red data, also gives the best-fit to the solution structure of all the three oligonucleotides examined so far. This seems to indicate that while the oligonucleotides are inherently flexible, the overall structural features in solution are likely to be closer to those of DNA in fibres, as compared to oligonucleotides in single crystals.

References

Arnott, S. and Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun.*, **47,** 1504.
Clore, G. M. and Gronenborn, A. M. (1984) *FEBS Lett.*, **175,** 117.
Clore, G. M. and Gronenborn, A. M. (1983) *EMBO J.*, **2,** 2109.
Dickerson, R. E. and Drew, H. R. (1981) *J. Mol. Biol.*, **149,** 761.
Gronenborn, A. M., Clore, G. M. and Kimber, B. J. (1984) *Biochem. J.*, **221,** 723.
Gupta, G., Bansal, M. and Sasisekharan, V. (1980) *Int. J. Biol. Macromol.*, **2,** 368.
Gupta, G., Rao, S. N. and Sasisekharan, V. (1982) *FEBS Lett.*, **150,** 424.
Gupta, G., Sarma, M. H., Dhingra, M. M., Sarma, R. H., Rajagopalan, M. and Sasisekharan, V. (1983) *J. Biomol. Struct. Dynam.,* **1,** 395.
IUPAC-IUB Nomenclature for polynucleotide chains, (1983) *Eur. J. Biochem.*, **131,** 9.
Klug, A., Jack, A. Viswamitra, M. A., Kennard, O., Shakked, Z. and Steitz, T. A. (1979) *J. Mol. Biol.,* **131,** 669.
Premilat, S. and Albiser, G. (1984) *J. Biomol. Struct. Dynam.*, **2,** 607.
Reid, D. G., Salisbury, S. A. Bellard, S. Shakked, Z. and Williams, D. H. (1983) *Biochemistry,* **22,** 2019.
Sakurai, T, and Sundarlingam, M. (1971) *Am. Cryst. Assoc. Meeting Abs.*, p. 91.
Sasisekharan, V., Sato, T. and Langridge, R. (1971) *Fed. Proc. Abs.*, **30,** 1219.
Sasisekharan, V., Bansal, M., Brahmachari, S. K. and Gupta, G. (1981) in *Proceedings of the Second SUNYA Conversation in the Discipline Biomolecular Stereodynamics* (ed. R. H. Sarma) (New York: Adenine Press) Vol. 1, p. 123.
Sasisekharan, V., Bansal, M. and Gupta, G. (1983) in *Nucleic Acids: The Vectors of life* (eds B. Pullman and J. Jortner) (Dordrecht: D. Reidel Publishing Co.) p. 101.