ON THE ALTERNATIVE STRUCTURE OF DNA: ROLE OF *SYN* CONFORMATION OF THE BASES

V. SASISEKHARAN AND GOUTAM GUPTA

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, *India*

ABSTRACT

Anti conformation about the glycosydic bond is equally favoured by purines and pyrimidines in nucleic acids. Using *anti* conformation for all the bases, earlier an alternative model of DNA designated as Type II structure was proposed. However, purines (and not pyrimidines) prefer some- time s *syn* conformation. Thus, polynucleotide duplexes with *syn* conformation throughout the structure, are not energetically favourable. In this paper, we have investigated the role of *syn* conformation for the purines in polynucleotide duplexes. For this purpose, a polynucleotide duplex with alternate purine-pyrimidine sequences was chosen with trinucleoside diphosphate as the repeating unit. A few examples of molecula1 conformations of trinucleoside diphosphate which led to *uniform* and *zig-zag* helices are discussed. Alternating right and left double helical segments of *uniform* as well as *zig-zag* helices can be joined to obtain the RL models of DNA which are only minor variants of Type II structure proposed earlier by us.

INTRODUCTION

Recently an alternative structure of B-DNA was proposed from this laboratory[1-4]. This structure, which will be referred to as the RL model, is not a regular double helix but has alternating right and left helical segments, each approximately five base pairs in length, in a repeat of 10 base pairs. The structure was arrived at by exploiting the conformational flexibility inherent in the DNA molecule. Further studies of different polymorphous forms of DNA showed that both right and left handed duplexes are compatible with the allowed stereochemistry and the observed X-ray data [5-6]. These studies also indicated that for a given polymorphous form of DNA, there can be a variety of right handed and left handed duplexes with dinucleoside monophosphate as the repeating unit. A right handed segment can be joined to a left handed segment to yield a RL structure and the model so obtained retains the essential features of what we have referred to earlier as the type Il structure[1-4]. In this paper, we first consider the implications of using dinucleoside monophosphate as the repeating unit and later indicate the role of the *syn* conformation of the bases when trinucleoside diphosphate is the basic unit.

RESULTS OF MODEL BUILDING WITH DINUCLEOSIDE MONOPHOSPHATE AS THE REPEATING UNIT

In the dinucleoside monophosphate the sugar puckering is kept the same at the 5' and 3' ends. So, mononucleotide is the exact repeating unit. However, as a dinucleoside monophosphate embodies two essential features of polymeric DNA, torsion around two P-0 bonds and base stacking which are absent in a mononucleotide, the former was chosen as a typical repeating unit for formulating stereochemical guideline of molecular model building.

With dinucleoside monophosphate as the repeating unit, we could demonstrate that the (β-γ) space can be divided into different helical domains as shown in Fig. 1. For the most frequently observed conformation[7] (*gg*) about the bond C4'- C5', we obtained helical domains I and II with (C3' - *endo,* g^-g^-) and (C2' - *endo,* tg^-) conformations respectively. In each domain, (a) both right and left helical duplexes are stereochemically possible, and (b) in a given domain, one can go over from one handedness to the other by small changes in the backbone torsion angles. However, the glycosyl torsion of the left handed duplex is about 60° lower than that of the right handed duplex. For example, glycosyl torsion is in the low *anti* region ($10° \leq \chi \leq 40°$) for the right handed duplex in domain l, while that for the left handed one is near the *syn* region ($310° \leq \chi \leq 340°$). For the right handed duplex in domain II, χ is 55°-75° while χ for the left handed one is 0°- 30°, both being in the *anti* region. Using the less commonly observed conformation (gt) about the bond C4' -C5', the helical region shifts to domain III for the C3' -*endo* puckering (Fig. l). Both right and left handed duplexes in domain III invariably have large radii and therefore cannot be packed in the unit cell of B-DNA. The values of χ for the left and right handed duplexes in this domain are similar to those obtained for the models in domain I. The results of model building using dinucleoside monophosphate as the repeating unit are summarized in Table I. This table shows that dinucleoside monophosphate as the repeating unit never leads to a right handed duplex with all the bases in *syn* conformation. Similarly, the left -duplexes in domain II cannot have all the bases in *syn* conformation. However, near-*syn* conformations are possible for left duplexes in domains I and III. But in view of the fact that the pyrimidine bases in *syn* conformation are not energetically favourable[8], the left helical duplexes with all the bases in *syn* confor-

mation are not very likely to occur in DNA. However, energy·calculations indicate *syn* conformation for purines (especially for guanine) is as favourable as the anti conformation[2-3]. Also, a few single crystal structures of nucleosides and nucleotides of guanine [10-11] indicate *syn* conformation for this base. This prompted us to look into the possibility of double helical structures with all purines in *syn* conformation. For this purpose, we have chosen a poly-nucleotide duplex with altemating purine and pyrimidine sequence as the model system. A trinucleoside diphosphate then turns out to be a typical repeating unit instead of dinucleoside monophosphate. Note that the exact repeating unit in such a case is a dinucleotide.

Model Building with Trinucleoside Diphosphate as the Repeating Unit

Use of a trinucleoside diphosphate as the repeating unit, leads to two topologically distinct types of duplexes, the *uniform* and *zig-zag* helices. In the *uniform*

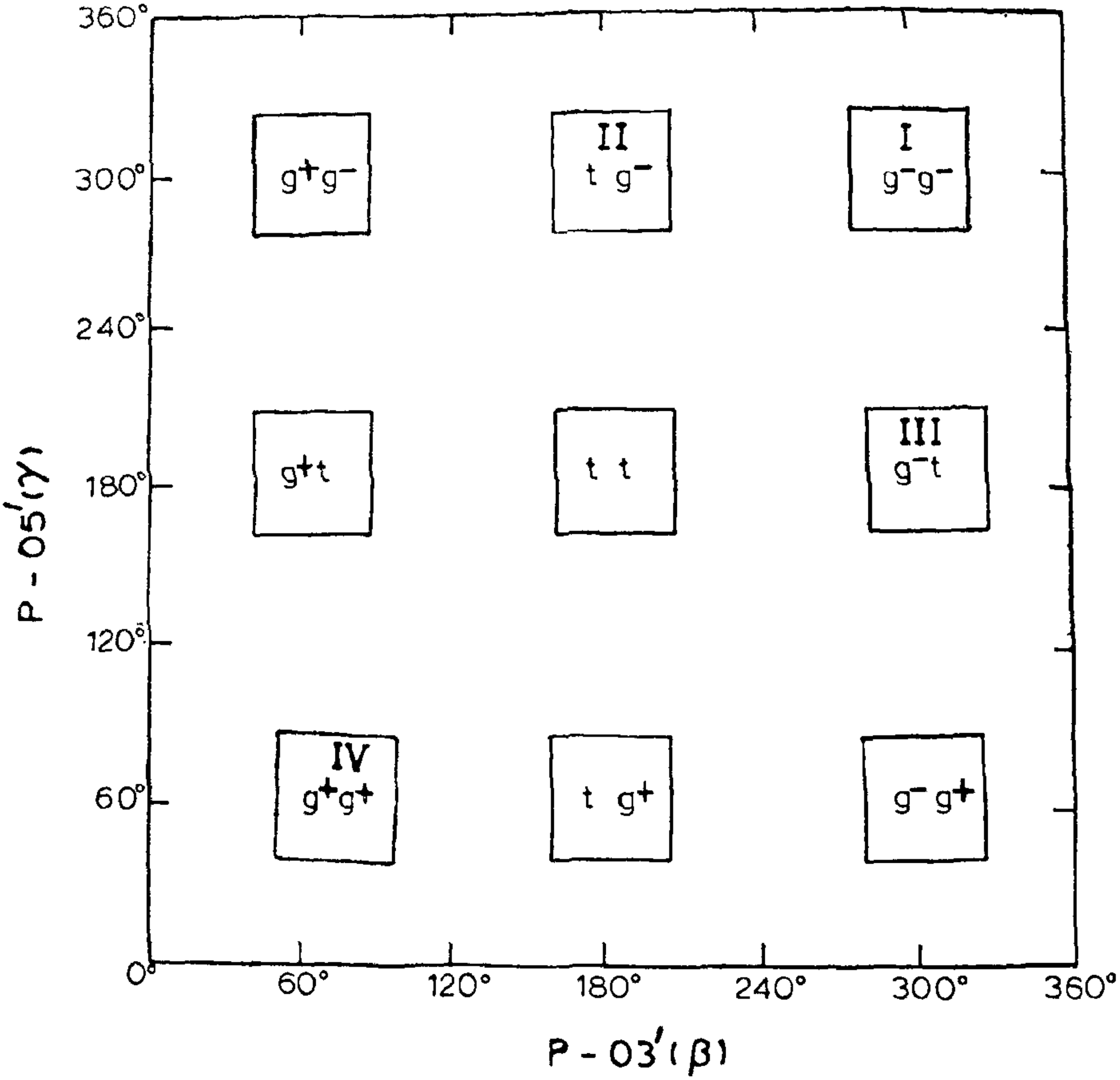


Fig. 1. Schematic representation of the helical domains in the (β-γ) space. The conformational features of the various domains are summarized in Table I. The helical domains are (g^-g^- and tg^-), (g^-t and tt) and (gg^+ and tg^+) for *gg*, *gt* and *tg* conformations respectively about the C4'-C5' bond. The remaining three are non-helical domains. We have only investigated the helical domains I, II and III and the non-helical domain IV as shown. We have found that the (C3'-*endo*, g^-g^+) domain is stereochemically unsatisfactory. The helical domain (C2'-endo, tt) and (C2'-endo, tg^+) and the non-helical domains g^+g^- and g^+t are not considered here, as these conformations have not been observed so far, for nucleotides and higher oligomers.

TABLE I

Results of the conformational domains investigated in (β-γ) space with dinucleoside monophosphate as repeating unit

Domain	Torsion (C4′–C5′) ϵ	Puckering (C4′–C3′) ζ	03′–P–05′ torsions (β, γ)	Glycosyl torsion χ	
				Right handed duplex	Left handed duplex
I	gg ($55° \leqslant \epsilon \leqslant 75°$)	C3′-*endo* ($75° \leqslant \zeta \leqslant 100°$)	g^-g^- ($270° \leqslant \beta, \gamma \leqslant 305°$)	low *anti* ($10° \leqslant \chi \leqslant 40°$)	near *syn* ($310° \leqslant \chi \leqslant 340°$)
II	gg ($40° \leqslant \epsilon \leqslant 65°$)	C2′-*endo* ($135° \leqslant \zeta \leqslant 155°$)	tg^- ($195° \leqslant \beta \leqslant 220°$) ($280° \leqslant \gamma \leqslant 315°$)	*anti* ($55° \leqslant \chi \leqslant 75°$)	low *anti* ($0° \leqslant \chi \leqslant 30°$)
III	gt ($150° \leqslant \epsilon \leqslant 195°$)	C3′-*endo* ($75° \leqslant \zeta \leqslant 100°$)	g^-t ($280° \leqslant \beta \leqslant 310°$) ($155° \leqslant \gamma \leqslant 195°$)	low *anti* ($10° \leqslant \chi \leqslant 40°$)	near *syn* ($310° \leqslant \chi \leqslant 340°$)
IV	gg ($50° \leqslant \epsilon \leqslant 75°$)	C3′-*endo* C2′-*endo*	g^+g^+ ($50° \leqslant \beta, \gamma \leqslant 90°$)	Structuie not possible	Structure not possible
	gt ($150° \leqslant \epsilon \leqslant 195°$)	C3′-*endo* C2′-*endo*	g^+g^+ ($50° \leqslant \beta, \gamma \leqslant 90°$)	Structure not possible	Structure not possible

The alphabetical nomenclature of the torsion angles are adopted from Seeman *et.al.*[14]. The molecular models were generated using modified LALS method wherein flexibility in the furanose ring was incorporated.

helices, the helical twist and the vertical displacement between successive phosphate groups are approximately the same. In the *zig-zag* helices the phosphate groups go around the helix axis in a non-uniform *(zig- zag)* fashion. In what follows, we describe the conformational features of the *uniform* and the *zig-zag* helices. We also show that one can join alternate right and left handed segments of the *uniform* helix to form a RL model of DNA, so also for the *zig-zag* helix. The common feature of these two kinds of RL models is that the left variety of both of them has either near-*syn* or pure *syn* conformation for all the purine bases.

THE UNIFORM HELIX

It is stereochemically possible to join two dinucleo- side monophosphates with conformations in in two helical domains and the resulting trinucleoside diphosphate can be used as a repeating unit to generate *uniform* helices. For example, we could link up alternately (C3'-*endo*, g^-g^-) and (C2'-*endo*, tg^-) to obtain both right and left helical duplexes for B-DNA. In such structures, all the purines are attached to sugars with C3'-*endo* sugar puckering while the pyrimidines are attached to sugars with C2'-*endo* puckering. For the right handed duplexes the baes (both purines and pyrimidines) are in *anti* conformation while for the left handed duplexes the purines arc in near *syn* conformation (*see* Fig. 2a). Such right, and left handed double helical segments (each 5 base-pairs in length) could be combined to arrive at a RL model of B-DNA. A space filling model so constructed is shown in Fig. 2*b*. In this model although the back-bone conformation is almost identical in the left and right helical segments, the purines in the right helical segment are in low *anti* conformations while in the left segment they are in the near-*syn* conformation.

THE ZIG-ZAG HELIX

Zig-zag helices were generated when two dinucleoside monophosphates, one with conformation in a helical domain and the other with conformation in a non-helical domain or both with conformations in the non-helical domains were joined. We have chosen (C3'-*endo*, g^+g^+) or (C2'-*endo*, g^+g^+) conformation as representative of a non-helical domain (domain IV in Fig.1.) g^+g^+ conformations were found in the single crystal structures of ApApA and UpA [12-13]. Therefore, the role of g^+g^+ conformation around P-O bonds in a polynucleotide duplex was investigated in detail. For example, a tri-nucleoside diphosphate with (C3'-*endo*, g^-t-C2'-*endo*, g^+g^+ C3'-*endo*) conformation (see Fig.3a) led to both right and left helical duplexes. For the left handed duplex all the purines are attached to sugars with C3'-*endo* puckering all of them have anti conformation. For the right handed duplex, although the purines are attached to sugars with C-endo puckering all of them have *anti* conformation. The depositions

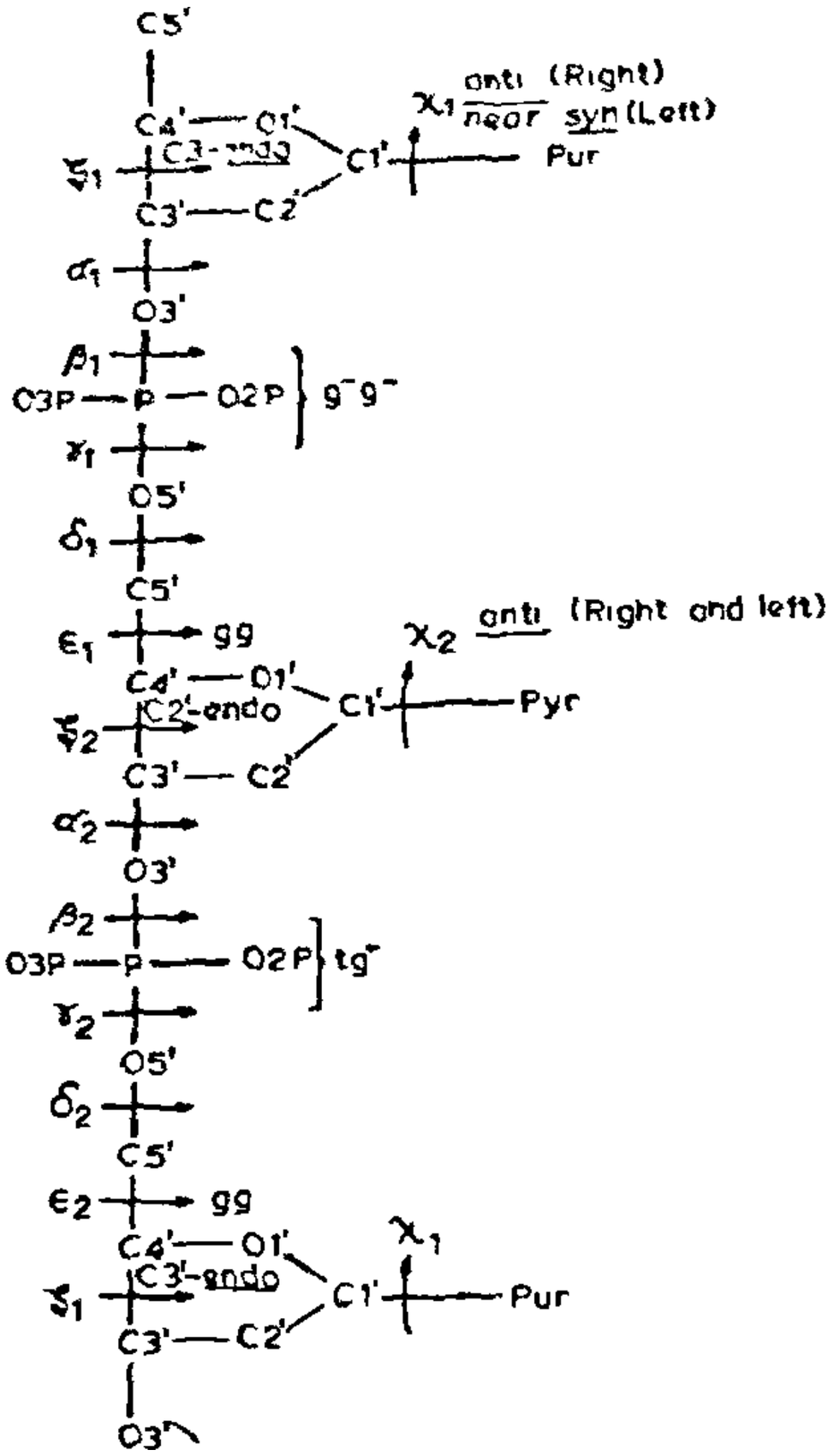


Fig. 2 *a*. Trinucleoside diphosphate as the repeating unit which leads to *uniform* helix. The glycosyl torsion regions are indicated for right and left handed duplexes.

of the phosphate groups for such structures are schematically shown in Fig. 3*b* and 3*c*. It is seen thataround each phosphate group, the two neighbouring ones are not symmetrically situated; one of them is horizontal while the other is vertically down. Such a left handed duplex can be joined smoothly with a right handed counterpart, within a repeat of 10 base-pairs of B-DNA. Such space filling model for B-DNA is shown in Fjg. 4. Here all the purines in the left handed helical segments have pure *syn* conformations.

In a similar fashion, right handed and left handed duplexes with *zig-zag* progression of the phosphate groups were arrived at when the (C3'-*endo,* g^+g^+-C2'-*endo,* g^+g^+-C3'-*endo)* conformation for the trinucleoside diphosphate is adopted. For the right handed duplex, both the sugars have *gt* conformation about the bond C4'-C5' (see Fig. 3) while the left handed duplex has *gt* conformation only for the C3'-*endo* sugar. The left handed *zig-zag* duplex so constructed has greater chain separation than the models with (C3'- *endo*, *gt*-C2'-*endo*,

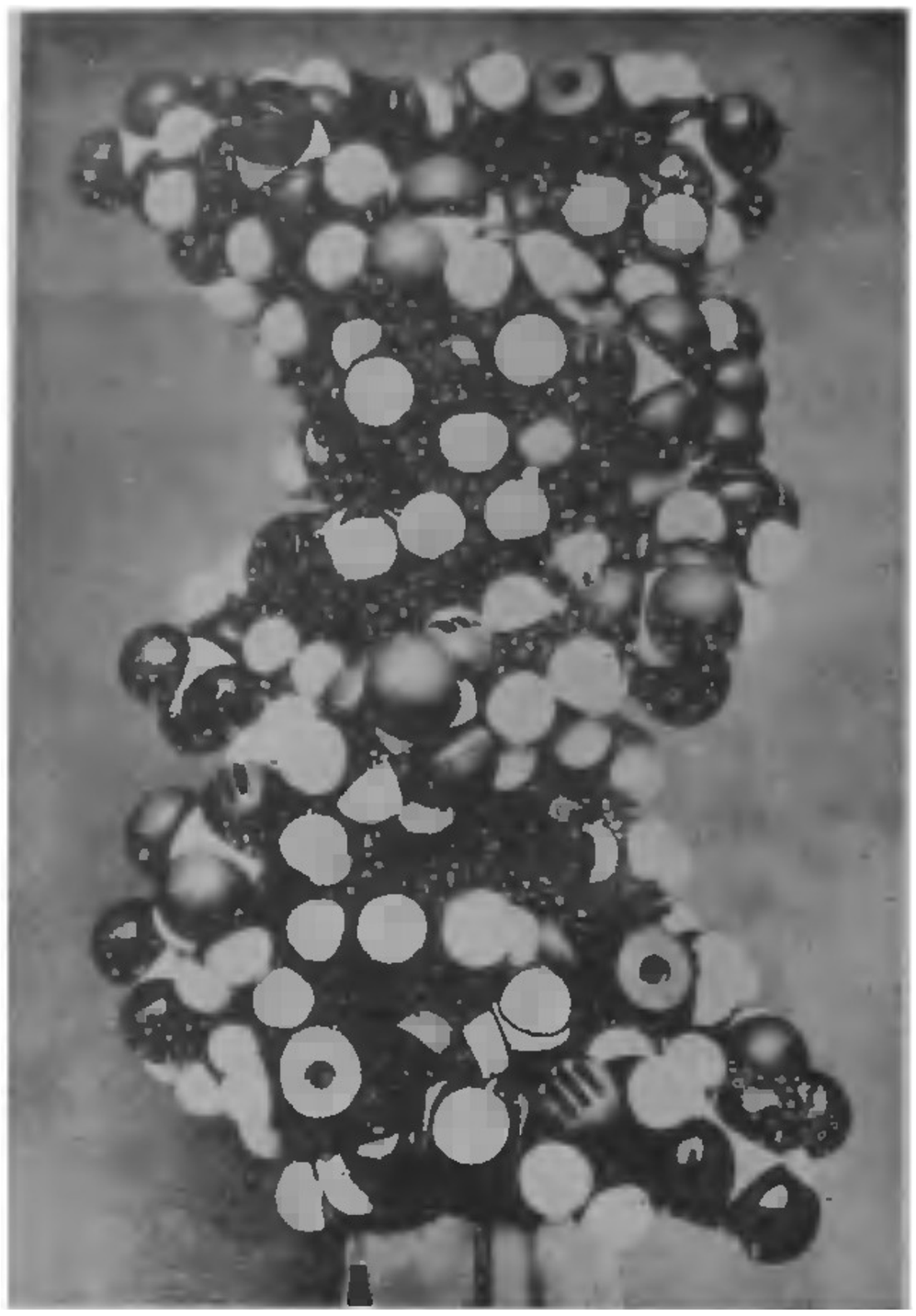


Fig. 2 *b*. Space-filling (CPK) model of a RL model obtained by joining alternately right and left helical segments of the *uniform* helix.

g^+g^+-C3'-*endo)* conformation. Here again all the purines in the left handed duplex have pure *syn* conformation, while in the right hand duplex they are all in *anti* conformation. However, in the case of the left handed duplex, adjacent sugars in the same chain point in opposite directions while the sugars attached to a given base pair point in the same direction. This is in striking contrast to right handed structures. The right and left helical segments, can be joined to obtain a RL model of B-DNA, *a* space filling representation, of which is shown in Fig.5. A left handed segment with either (C3'-*endo, g⁻t*-C2'-*endo,* g^+g^+-C3'-*endo*) or (C3'-*endo*), g^+g^+ -C2'-*endo*, g^+g^+- C2'-*endo*) conformation can easily be joined to any right handed segment generated from (C3'-*endo,* g^-g^-) or (C2'-*endo,* tg^-) conformation. ln such an arrangement, the phosphate groups in the left handed segment have a *zig-zag* progression while those in the right handed segment are uniformly wrapped around the helix surface.

CONCLUSIONS

These studies clearly indicate that *syn* conformation of the bases is possible only for the left handed duplexes. In such cases, the purines will have only the *syn* conformation and the·sugar attached to them should necessarily have C3'-*endo* and not C2'-*endo* puckering. As a result, if left stacking of the bases is preferred to

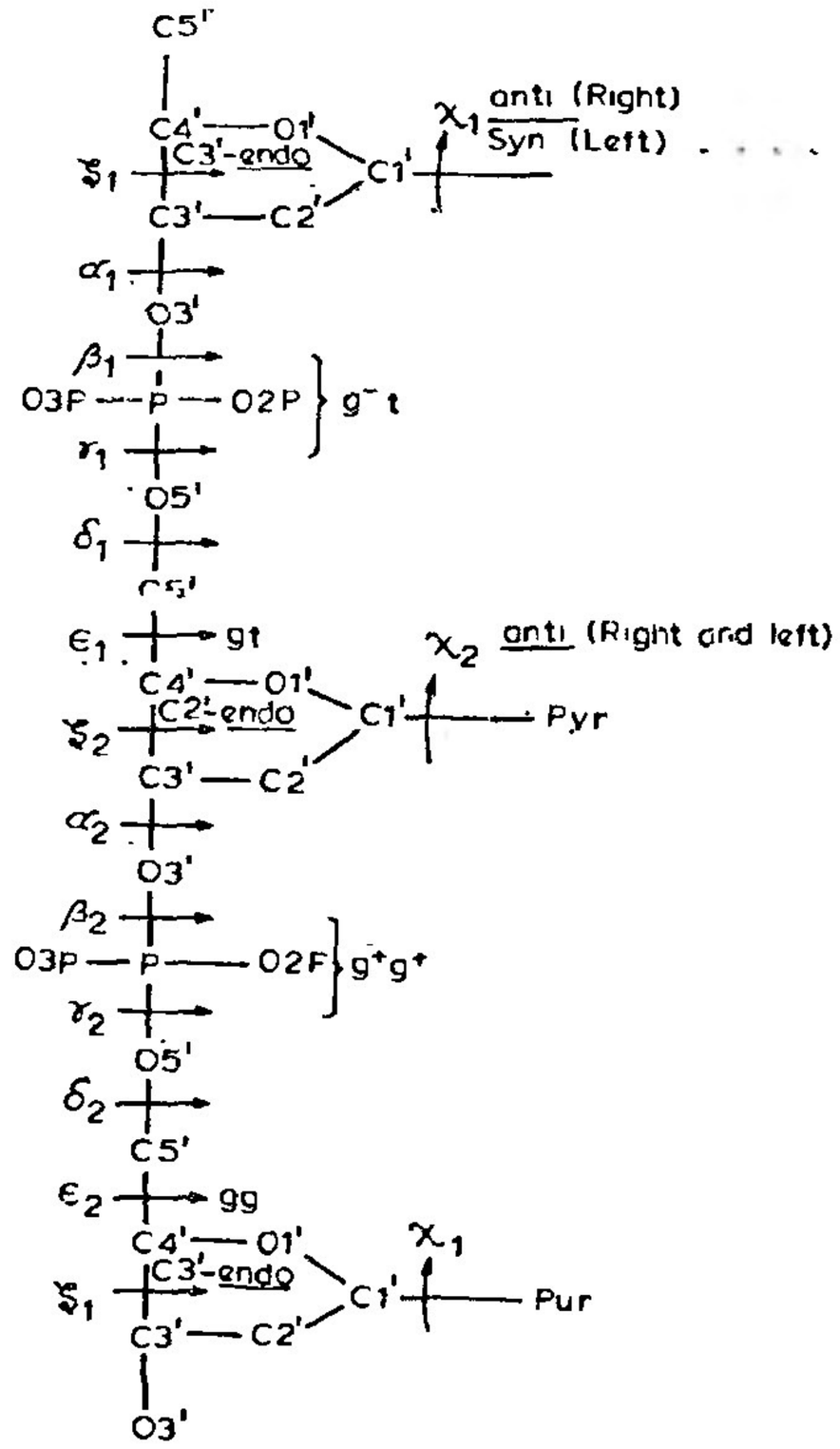
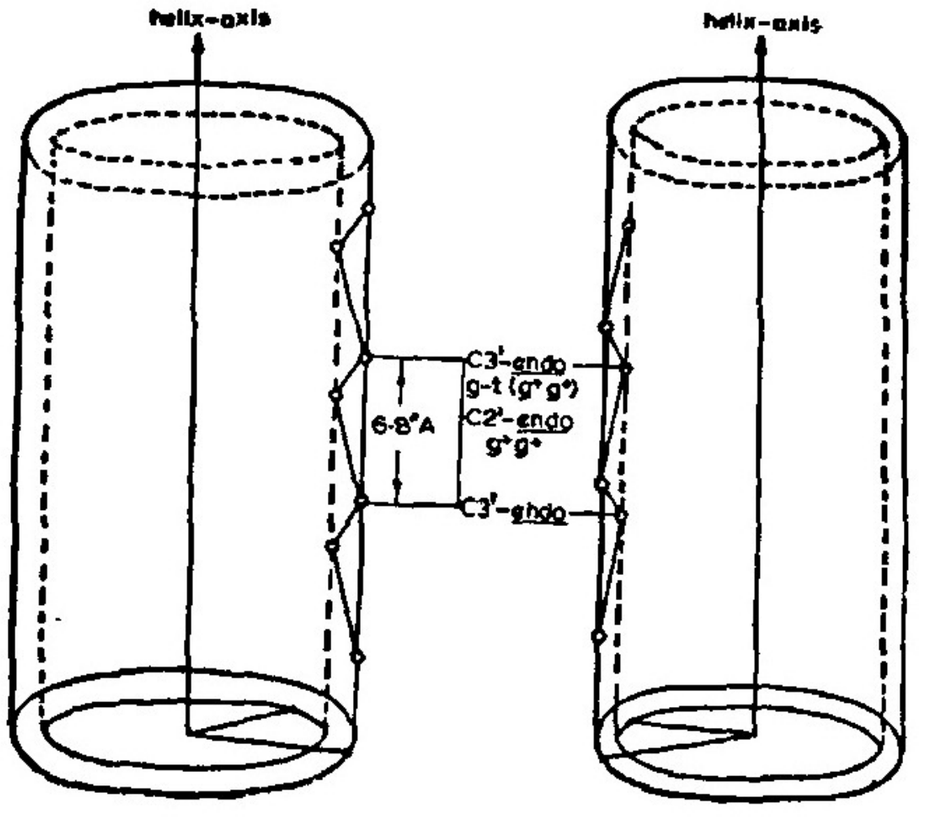


FIG. 3 *a.* Trinucleoside diphosphate which leads to a *zig-zag* helix.* The glycosyl torsion regions are indicated for right and left minded duplexes.

*It was noted that gt conformation around C4'-C5' bond can be interchanged between sugar with C2'- *endo* puckering and the one with C3'-*endo* puckering. It was also found that (C3'-*endo, g⁻g⁻*-C2'-*endo, g⁺g⁺*-C3'-*endo*) and (C3'-*endo, tg⁻*-C2'-*endo, g⁺g⁺*-C3'-*endo*) conformations led to similar *zig-zag* helices. Note that the phosphodiester conformation refers to the phosphate group attached to the 3'-end of the sugar.



3 *b* 3 *c*

Figs. 3 *b-c. b.* Schematic representation of a right handed *zig-zag* helix showing the progression of the phosphate groups. *c.* Schematic representation of a left handed *zig-zag*helix.

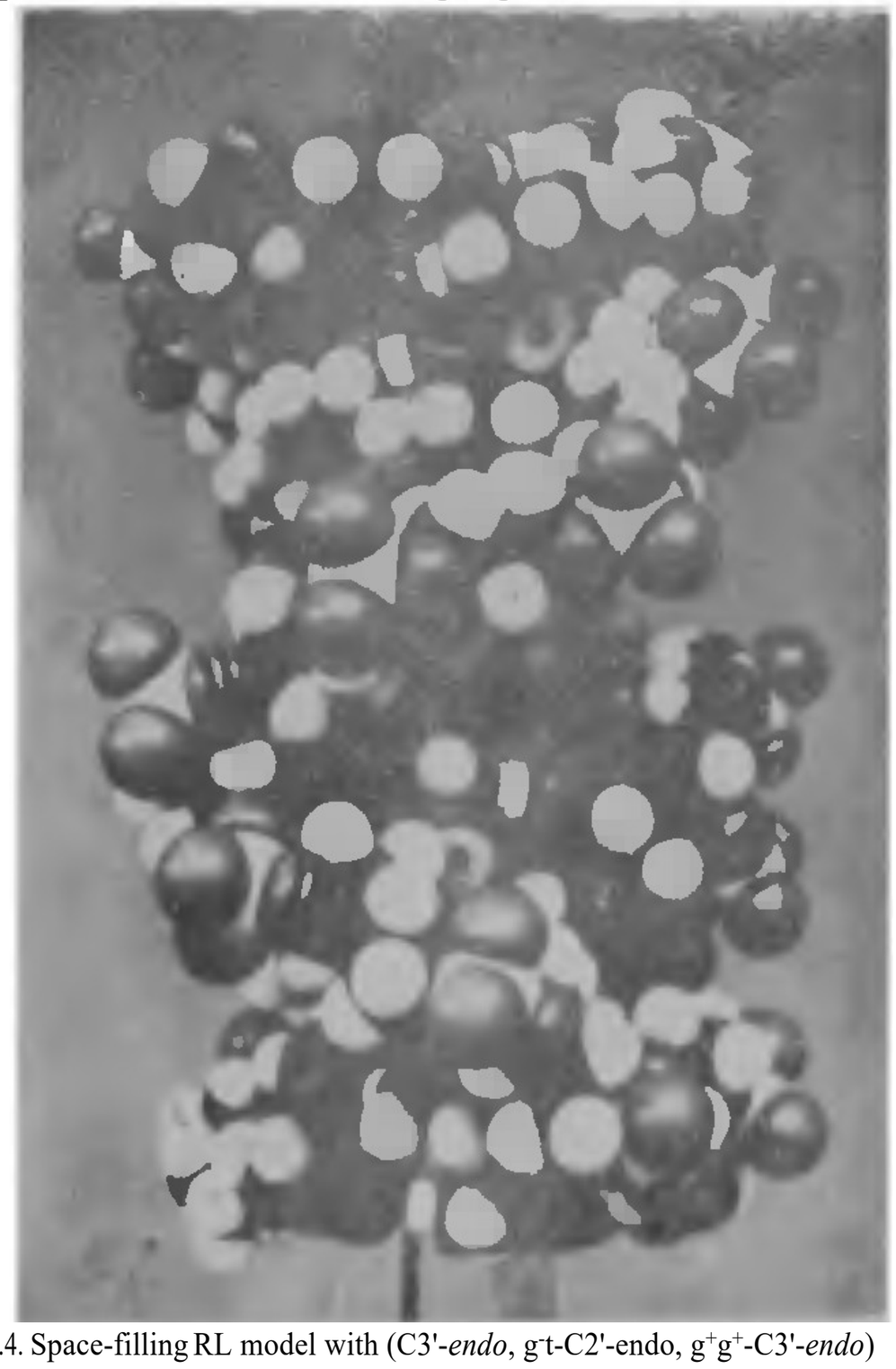


Fig.4. Space-filling RL model with (C3'-*endo*, g⁻t-C2'-endo, g⁺g⁺-C3'-*endo*) conformation.

right stacking in duplexes with alternate purine-pyrimidine sequences, the *syn* conformation for the purines becomes inevitable. After this work was completed, it has come to our notice that cxperiments by Wang *et al.*[15], bear out this prediction: In the (dC-dG)8 crystal structure internal G's are in syn conformation and sugars attached to them have C3'-*endo* puckering. The resulting structure, because of the reasons cited above, is a left handed *zig-zag* helix.

In all the RL models discussed above, the bases are turned over (or flipped over) each other at the bend region where the right and left helical segments join together. This inverted stacking arrangement, at the bend region, is a characteristic feature of the type II model published earlier [1-4]. The alternative

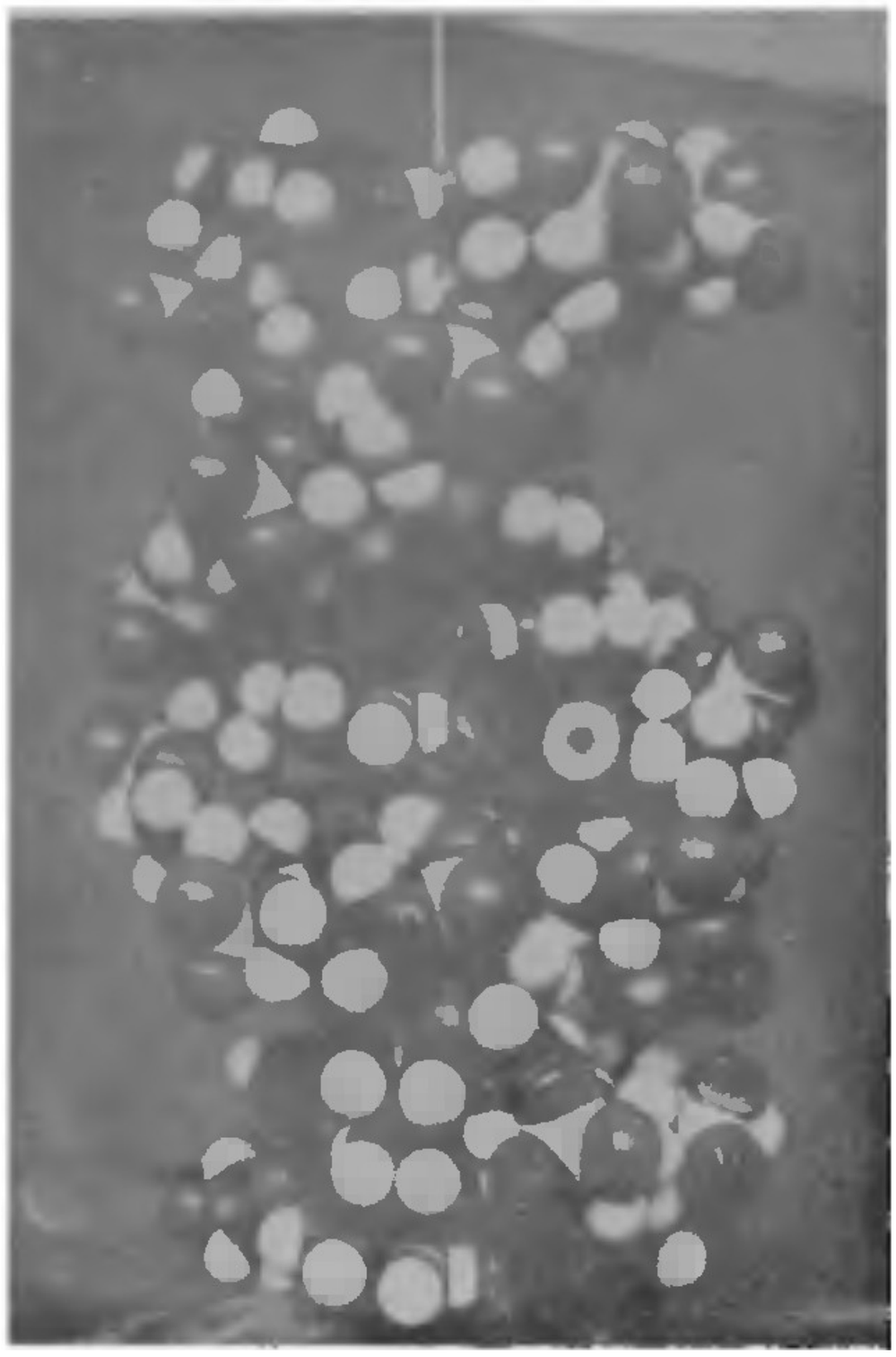


Fig. 5. Space-filling RL model with (C3'-*endo*, g^+g^+ C2'-*endo*, g^+g^+-C3'-*endo*) conformation.

models of DNA considered in this article are minor modifications of the type II structure. It is interesting to note that in the model of Wang *et al.* [15], a segment of left handed *zig-zag* DNA is combined with right banded B-DNA. The resulting structure is again a variant of the typeII structure proposed by us [1-4]

ACKNOWLEDGEMENT

This work was supported by the Department of Science and Technology, New Delhi, and also NIH Scheme PL-480-USPHS-grant 01-126-N. GG wishes to thank U.G.C. for a fellowship.

1. Sasisekharan, V. and Pattabiraman, N., *Curr. Sci.,* 1976, **45**, 779.
2. -, - and Goutarn Gupta. *Ibid.,* 1977 **46**, 763.
3. -, - and -, *Proc. Nat. Acad. Sci.* USA., 1978, **75**, 4092.
4. - and -, *Nature,* 1978, **275,** 159.
5. Goutam Gupta, Manju Bansal and Sasisekharan, V., Sent for publication.
6. -, - and -, Sent for publication.
7. For notations of conformations around different -single bonds *(i.e.,* gg, g^-g^-, etc.) refer to Sundaralingam, M., *Structure and Conformation of Nucleic Acid and Protein-nucleic Acid Interaction,* edited by Sundaralingam, M. and Rao, S. T., 1975, p. 487.
8. Berthod, H. and Pullman, B., *FEBS Letters,* 1973, 30, 231.
9 Yathindra, N. and Sundaralingam, M., *Bio-polymers* 1973, 12., 287, *2015.*
10. Brennan, T., Weeks, C., Shefter, Rao, S. T. and Sundaralingam, M., *J. Am. Chem. Soc.,* 1972, **94**, *8548.*
11.. Haschemeyer, A. E. and Sobell, H. M., *Acta Crystallogr.,* 1965, **19**, 125.
12. Suck, D., Manor, P. and Saenger, W., *Ibid.,* 1976, **B32**, 1927.
13. Sussman, J. L., Seeman, N. C., Kim, S. H. and Berman, H. M., *J. Molec. Biol.,* 1972, **66**, 403.
14. Seeman, N. C., Rosenberg, J.M., Suddath, F L., Kim, J. J. P. and Rich, A., *Ibid*., 1976, **104,** 109.
15. Wang, A. H. J., Kolpak, F. J., Quigley, G. J., Cranford, J. L., Van Boom, H. H., Van der Marel, G. and Rich, A., *Nature,* 1979, **282**, 680.