ISSN 0029-3865

CBPF-NF-041/85

RAPID LOCATION OF THE PREFERRED INTERACTION SITES BETWEEN SMALL POLAR MOLECULES AND MACROMOLECULES. II. BINDING OF WATER TO A MODEL SEGMENT OF B-DNA*

by

Marçal de Oliveira Neto

Centro Brasileiro de Pesquisas Físicas - CNPq/CBPF
Rua Dr. Xavier Sigaud, 150
22290 - Rio de Janeiro, RJ - Brasil

*In part this work was carried out at the Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique, associé au Centre National de la Recherche Scientifique, 13, Rue Pierre et Marie Curie, 75005, Paris, France.

ABSTRACT

The procedure developed in Part I of this series is applied to the homopolymeric sequences poly(dA).poly(dT) and poly(dG).poly(dC) on the double helical structure of B-DNA. Some aspects of the base sequence influence on the polymer's attraction for water molecule are described. The results are used to discuss the general hydration features of those systems in relation to recent experimental studies of DNA single crystals.

Key-words: B-DNA; Major groove; Minor groove; Interaction energy; Hydration.

I INTRODUCTION

Chemical or biological properties of macromolecules are often strongly influenced by their surroundings. Among the various environmental factors, the detailed nature of hydration certainly plays an important role. However, due to the size and complexity of macromolecules, a detailed picture is difficult to obtain either from the experimental or the theoretical viewpoint.

Nowadays, since the advent of fast digital computers, it has been possible to study theoretically the hydration of macromolecules such as DNA molecules or proteins[1-5]. These investigations have been carried out using the Monte Carlo method with different atom-atom pair potentials obtained from quantum mechanical calculations.

In part I of this series[6] a procedure was developed to take into account both the sterical and energetic aspects of the interaction between small polar molecules and macromolecules. The intermolecular interaction energies were computed through the overlap between the multipole expansions of the charge distributions of the interacting entities[7-9]. The macromolecule is built from its constituent fragments and the multipolar expansions (up to quadrupoles) of the charge distributions of each constituent are derived from *ab initio* SCF wave functions. This procedure is the same as that adopted in a large number of theoretical studies devoted to the reactivity of nucleic acids based on their molecular electrostatic potentials or fields[10-13]

As an example, the above procedure was applied, in part I, to the interaction surface between the component units of nucleic acids and water molecules. By investigating thoroughly

all orientations of a water molecule at each binding site, it was possible to characterize its lability.

Following the line of the previous work[6], an extension of this study to a model fragment of B-DNA is presented now. We show primarily the results regarding model segments of the homopolymers poly(dA).poly(dT) and poly(dG).poly(dC), for which it is assumed a double helical structure of B-DNA. These two homopolymeric sequences allows a description of some aspects of the base sequence influence on the polymer's affinity towards a water molecule.

Subsequently, a model is proposed for a relatively complete first hydration layer in the minor groove of the homopolymeric sequences. This attempt to go a step further is motivated by recent X-ray diffraction studies of single crystals of oligomeric DNAs; fragments of this molecule in different conformations, B[14-18], A[19-21] and Z[22,23] have been studied and in several cases the positions of cations, polycations and water molecules appearing in the crystals have been also located with precision. We further discuss the present theoretical results and those concerning hydration of single crystals of oligomeric B-DNA to establish the significant correlations between them.

The present study is restricted to the standard B-DNA conformation[24] of the oligomers.

II RESULTS

The results regarding water binding to the major groove, minor groove and phosphates groups of the B-DNA model are presented in Sections II.1, II.2 and II.3, respectively.

The model polymer consists of one turn of the double helix with 11 phosphates in each strand and homopolymeric base sequences $(GC)_n$ and $(AT)_n$. In section II.4 we describe the model adopted in order to obtain a profile of the second hydration layer in the minor groove of the oligomeric sequences.

Unless otherwise stated water molecules were bound to the central nucleoticles of the model double helices to minimize end effects. The binding of water molecule to bases electronegative heteroatoms, to amino group hydrogens (other than those involved in the hydrogen bonds of the nucleic acid base pair), to furanose and ester oxygens of the phosphodiester linkage and to phosphate group anionic oxygens are investigated. Standard notations are employed for the atoms of the various components of DNA, for example, $N_7(G)$ for the atom N_7 of guanine, $O_1'(S)$ for the sugar oxygen, $H_6(A)$ for the N_6-adenine free amino hydrogen and so on.

Table I shows the studied oligomer regions, DNA component atoms involved in the interaction site, the optimum energies and the number NC (see ref. 6) of water configurations having energies within 1 Kcal/mole of the optimum energy.

Figures 1 to 6 depicte the first layer water molecule interaction sites in poly(dA).poly(dT) and poly(dG).poly(dC) grooves for which we have employed only three nucleotides of each double helix strand. In order to simplify the visualization of these diagrams we indicate the water molecule at its optimum energy position; the oxygen atoms of the other water molecule configurations are indicated by a shaded zone. All figures in the present work were drawn with the help of a graphic plotter connected to a Apple II microcomputer.

II.1 MAJOR GROOVE

As indicated in table I, the strongest binding site in poly(dA).poly(dT) major groove lies between the N_7(A) and H_6(A) atoms. This site is virtually the same as that located for isolated adenine[6] but in the polymer its binding energy becomes more favourable by 3.2 Kcal/mole. However, for the second site located in the major groove, this extra stabilization is not found when a water molecule interacts with O_4(T); in this case the interaction energy falls by 1 Kcal/mole in comparison to that of the isolated base. In figure 1 the bound water habilities (that in a bridge position between N_7(A) and H_6(A) and the one bound to O_4(T) are denoted by W_1 and W_2, respectively) can be visualized: W_1 is not very labile whereas W_2 can be displaced over a wide region in the major groove of the polymer.

In poly(dG).poly(dC) major groove the strongest binding site giving rise to an interaction energy of -15.3 Kcal/mole is associated with a bound water molecule lying between the N_7(G) and O_6(G) atoms (denoted by W_3 in figure 2). This configuration is more stable by 4 Kcal/mole than the corresponding one in isolated guanine[6] (although in equivalent binding site). The second located site involves the N_4 free amino hydrogen of cytosine (W_4 in figure 3). We recall that for isolated cytosine the optimum interaction energy of a water molecule with this atom is -6.4 Kcal/mole; in the macromolecule this value increases to -11 Kcal/mole despite a relatively high water lability.

II.2 MINOR GROOVE

If we first considere binding to the poly(dA) strand of

poly(dA).poly(dT), energy optimization indicates that in preferred location the water molecule is shared by N_3(A) and $O_{1'}$ atom of the sugar bound to the consecutive adenine in the sequence $5'A_pA_{3'}$. The optimum energy in this case is -12.3 Kcal/mole which coincides with that found for water binding between N_3(A) and the N_9-hydrogen in isolated adenine.

As to bindings to the poly(dT) strand our calculations show that the resulting configuration bridges the O_2(T) and $O_{1'}$(S) atoms, with an interaction energy of -11.9 Kcal/mole. The resulting bond is oriented similarly to that obtained for the poly(dA) strand.

It may be remarked that in figure 3 the water molecules bound to poly(dA) and to poly(dT) strands (denoted by W_5 and W_6 respectively) presents labilities significantly smaller than the corresponding water molecule lability found for the association with $O_{1'}$(S) in the isolated sugar subunit[6]. For this latter case the interaction energy is -5.7 Kcal/mole; however, this water molecule gains an extra stabilization when embedded in the macromolecular system.

In the case of poly(dG).poly(dC) minor groove the stronger binding (-11.8 Kcal/mole) is associated with a water molecule (W_7 in figure 4) interacting simultaneously with three centers: one of these centers is the N_2(G) free amino hydrogen of the central guanine and the other two are the O_2(C) and $O_{1'}$(S) atoms both belonging to the nucleoticle on the 3' side of the central cytosine. As reported in table I the resulting interaction energy involving a water molecule and the N_3(G) atom in poly(dG).poly(dC) minor groove is less favourable than that mentioned above and the binding site (W_8 in figure 4) is located in the neighbourhood of the guanine amino group. Analyse of the

distance between the water molecules W_7 and W_8 shows that both sites cannot coexist due to steric hindrance. Hydration of $O_{1'}$(S) atom on poly(dG) strand yields a water configuration very similar to that found for N_3(G), the binding energy being, however, 2 Kcal/mole less favourable. According to these results it would seem that only one binding configuration can be found in this region of poly(dG).poly(dC) minor groove. We shall return to this point later.

Figure 5 and 6 depicte a view of the bound water configurations in the major and minor groove of poly(dA).poly(dT), and poly(dG).poly(dC), respectively, on a plane perpendicular to their helical axis.

II.3 PHOSPHATE GROUPS

The strongest binding energies within the two studied sequence models are associated with phosphate groups (see table I). The binding of water to anionic oxygens is favoured both from sterical and energetic considerations. For instante, O_3, atom is sterically highly hindered preventing it from directly attacking a water molecule. The resulting configurations show high water labilities associated with the anionic oxygens. A water molecule may bridge the O_1 and O_2 atoms and also the O_2 and O_5, atoms through a hydrogen bond, although the latter is less favoured. Figure 6 shows, as an example, the poly(dG).poly(dC) sequence in which the water molecule W_9 is situated between the two anionic oxygens of the phosphate group. As indicated in table I there are about a hundred possible configurations for W_9; such configurations spread over a wide region (also depicted in figure 6) around the anionic oxygens.

II.4 A Model to Investigate a Second Hydration Layer in Minor Groove

Such a model consists in the preliminary construction of a relatively complete first hydration layer. In the case of poly(dA).poly(dT) minor groove we have located the two bound water configurations W_5 and W_6. These two molecules are distant enough from each other, as the minimal distance between two of their atoms is approximatly 5.5 Å ; consequently it may be assumed these binding associations do not interfere with subsequent hydration. In this way our model for a first hydration layer in minor groove has been built with the following approximation: in view of the homopolymeric poly(dA).poly(dT) sequence we inserted into the polymer at each base pair water molecules identical to W_5 and W_6, i.e., the configurations at the minimal energy position found for the central base pair. Figure 7 shows the resulting scheme: W_5', W_5 and W_5'' represent the binding configurations on poly(dA) strand ; W_6', W_6 and W_6'' the corresponding ones on poly(dT) strand. Now, the water molecules are integrated to the macromolecular system and their multipoles are also taken into account for the interaction energies computations. In order to obtain a profile of the second hydration layer we then studied a subsequent binding with the two central W_5 and W_6 water molecules. The resulting hydration scheme indicates that it is constituted by bridging the binding configurations on poly(dA) strand and those on poly(dT) strand. However, this bridge does not link water molecules associated with one base pair, but instead a water molecule bound to adenine and the following one in the 5'-3' direction bound to thymine. The interaction energy for this binding (-17.8 Kcal/mole) is markedly larger than those obtained

for water molecules of first hydration layer. In figure 7 it is shown the resulting configurations W_{10} (bridging W_5 and W_6'') and W_{11} (bridging W_5' and W_6).

As to poly(dG).poly(dC), we recall that apparently only the water molecule W_7 may associate with the central nucleotides in minor groove. In order to investigate the structure of water molecules in this region we have adopted a similar procedure to that employed for the poly(dA).poly(dT) case. Nevertheless, it is necessary to study the possibility of binding at the hydrophilic centers N_3(G) and $O_{1'}$(S) of the poly(dG) strand, taking into account the presence of other water molecules identical to W_7 in the minor groove of poly(dG).poly(dC). The results indicate that binding to $O_{1'}$ sugar oxygen (-9 Kcal/mole) is preferred to binding to N_3(G) atom by an amount of interaction energy about 3 Kcal/mole. In figure 8 W_7', W_7 and W_7'' denote the resulting configurations on poly(dC) strand; W_8'', W_8' and W_8''' those on poly(dG) strand.

Subsequent bindings to W_7 and W_8' molecules shows that in contrast to the poly(dA).poly(dT) minor groove, no cross strand bridging occurs and that, in fact, the binding configurations on poly(dC) strand are inaccessible to direct attaking of subsequent water molecules. Yet, the corresponding ones on poly(dG) strand can accept another strong hydrogen bonded water molecule (-20.5 Kcal/mole) locating in the neighbourhood of the backbone phosphate groups (W_{12} in figure 8).

A view of both poly(dA).poly(dT) and poly(dG).poly(dC) hydration scheme on a plane perpendicular to the helical axis is presented in figures 9 and 10, respectively.

III DISCUSSIONS AND CONCLUDING REMARKS

The above results indicate that when comparing hydration features in the minor and major grooves of the B-DNA model, as functions of its base sequence, relevant differences can be established. We remark that the major groove of both poly(dA).poly(dT) and poly(dG).poly(dC) have nearly the same general hydration features, the binding being associated with a single base. Moreover, the water molecule can bridge between two base atoms, which is the case for the site involving N_7(A) and H_6(A) or that involving N_7(G) and O_6(G). The highest labilities in the major groove are associated with the water molecule bound to O_4(T) and that bound to H_4(C) for which there appears to be no preferential neighbouring atom available to render these associations bidentates.

However, bidentate bound water configurations can be formed between atoms of different component units of the double helix and such states have been found in the two oligomer's minor grooves. This is the case for the bridge formed between base edge atoms O_2(T) or N_3(A) and O_1(S) in poly(dA).poly(dT) and that involving O_2(C), H_2(G) and O_1(S) in poly(dG).poly(dC) minor grooves.

A tentative investigation of a subsequent layer hydration in the minor groove of the B-DNA model reveals clear contrast in their hydration features and only for poly(dA).poly(dT) a cross-linking of the first hydration layer is observed. The presence of the guanine amino group in the poly(dG).poly(dC) minor groove appears to play on instrumental role in inhibiting such cross-linking.

Through the comparison of the two studied sequences we may deduce that not only the interaction energy of bindings in the grooves are strongly correlated with the base sequence involved, but also the geometrical arrangement of the water molecules; this effect is emphasized in the minor groove case.

As we have mentioned Dickerson and colaborators X-ray studies of a single crystal oligomeric DNA allowed to locate water molecules in different regions of the B-DNA self-complementary dodecamer CGCGAATTCGCG[16-18]. Here, it is important to remark that we have carried out the present investigations regarding oligomers with homopolymeric base sequence having the classical B-DNA conformation. In the case of Dickerson's dodecamer geometry the base sequence has a more complicated arrangement and presents noticeable local structural heterogeneities. Nevertheless a theoretical interpretation of its NMR spectrum suggests that, in solution, its conformation may be considerably closer to that of a regular B-DNA[26].

The results of the dodecamer crystal, indicate that one of the most stricking hydration feature observed is a regular spine of ordered water molecules binding between successive adenine N_3 and thymine O_2 atoms in the minor groove in the AT-rich center of the polymer. Moreover, the second hydration layer is found to bridge this first layer in an approximate tetrahedral bound configuration. This regularity is disrupted on passing from the AT center of the self-complementary dodecamer towards its GC ends. In contrast, hydration of major groove showed that most water molecules in the first layer interact with nitrogen or oxygen atoms on the edges of the base pair, and the majority of the binding observed are associated

with one single base. The hydration geometry in the major groove did not present a regularity comparable to the hydration in the minor groove.

Since the classical work of Falk *et al.*[26-28] the preferential hydration of the anionic oxygen atoms in the phosphate groups of DNA is well established. This feature was nevertheless invisible in the primitive Dickerson's X-ray results at room temperature and insufficiently resolved spectrum of the dodecamer crystal CGCGAATTCGCG, indicating thermal or static disorder of the backbone atoms[16]. A further investigation of the same dodecamer at 16°K[17] and of a bromo derivative d(CGCGAATT^{Br}CGCG) in a 60% MPD (2-methyl 2,4-pentanediol) solution[18] enabled a refinement of the initial conclusions, indicating the phosphates as the sites of strongest hydration. This correlates with our findings since strong binding may be observed favouring the anionic oxygens of the phosphate groups. Moreover the largest water labilities associated to these binding sites seems to be one of the most important factors rendering difficult the observation of water molecules in this regions of the dodecamer crystal.

Although the present theoretical investigation refers to classical B form of DNA it seems to account satisfactorilly for experimental observations such as those relevant to the structure of hydration in major and minor groove as well as those relevant to the affinity of this biomolecule for water molecules.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. A. Pullman and Professor B. Pullman for fruitful discussions during the course of this work. He also thanks M.S. de Giambiagi, M. Giambiagi and H.V. Ferreira for several helpful discussions.

This work was partially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazil).

FIGURE CAPTIONS

FIG. 1 - Preferred interaction sites of a water molecule in poly(dA).poly(dT) major groove.

FIG. 2 - Preferred interaction sites of a water molecule in poly(dG).poly(dC) major groove.

FIG. 3 - Preferred interaction sites of a water molecule in poly(dA).poly(dT) minor groove.

FIG. 4 - Preferred interaction sites of a water molecule in poly(dG).poly(dC) minor groove.

FIG. 5 - A view on a plane perpendicular to the helical axis of poly(dA).poly(dT). The three base pair planes are placed at the distances:
Z (———), Z-h (------), Z+h (———), o oxygen atoms, ● nitrogen atoms, ... hydrogen bonds involved.

FIG. 6 - A view on a plane perpendicular to the helical axis of poly(dG).poly(dC).

FIG. 7 - A schematic representation of the partial first and second hydration layer in poly(dA).poly(dT) minor groove.

FIG. 8 - A schematic representation of the partial first and second hydration layer in poly(dG).poly(dC) minor groove.

FIG. 9 - A view on a plane perpendicular to the helical axis of the scheme obtained in poly(dA).poly(dT) minor groove.

FIG. 10 - A view on a plane perpendicular to the helical axis of the scheme obtained in poly(dG).poly(dC).

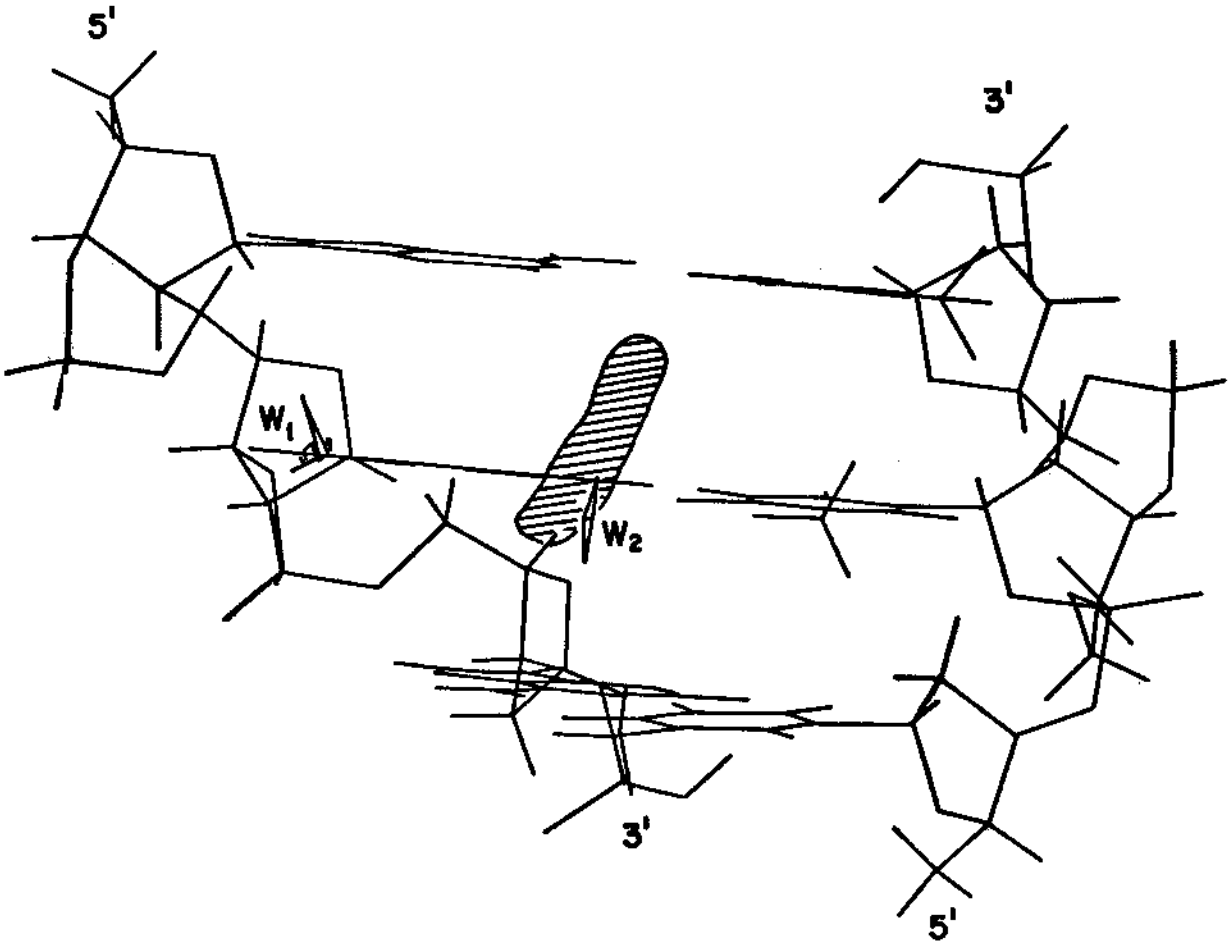

FIG. 1

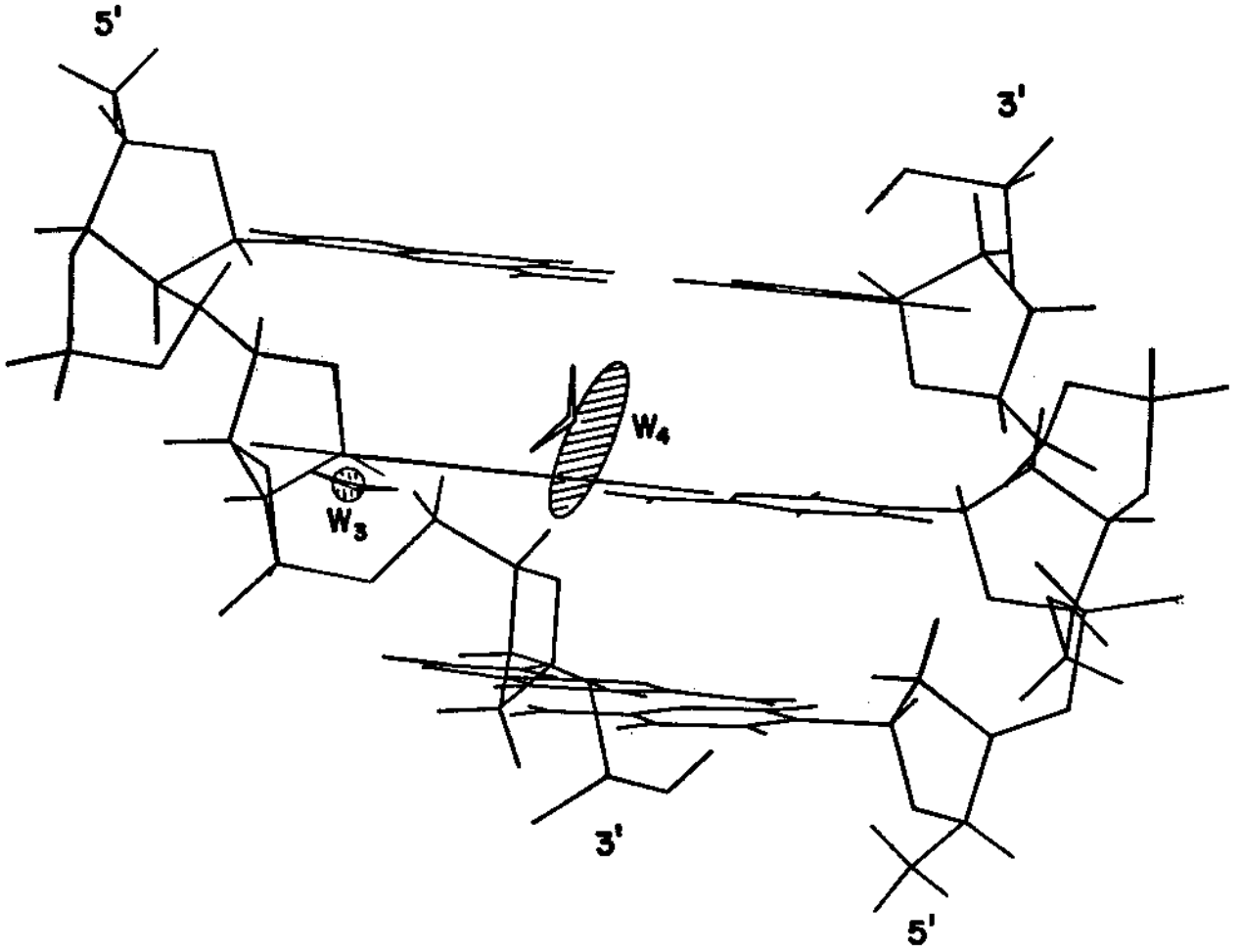

FIG. 2

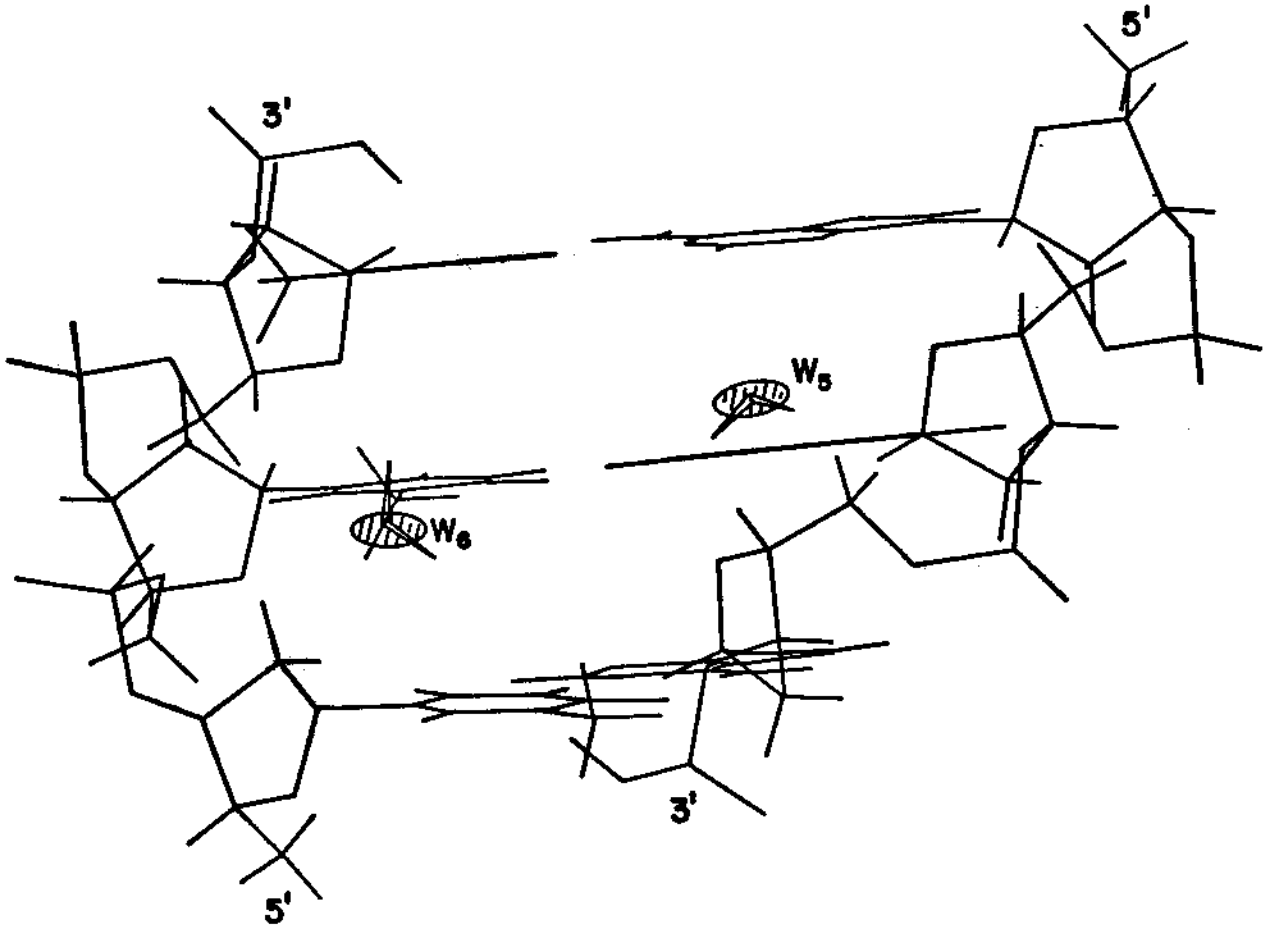

FIG. 3

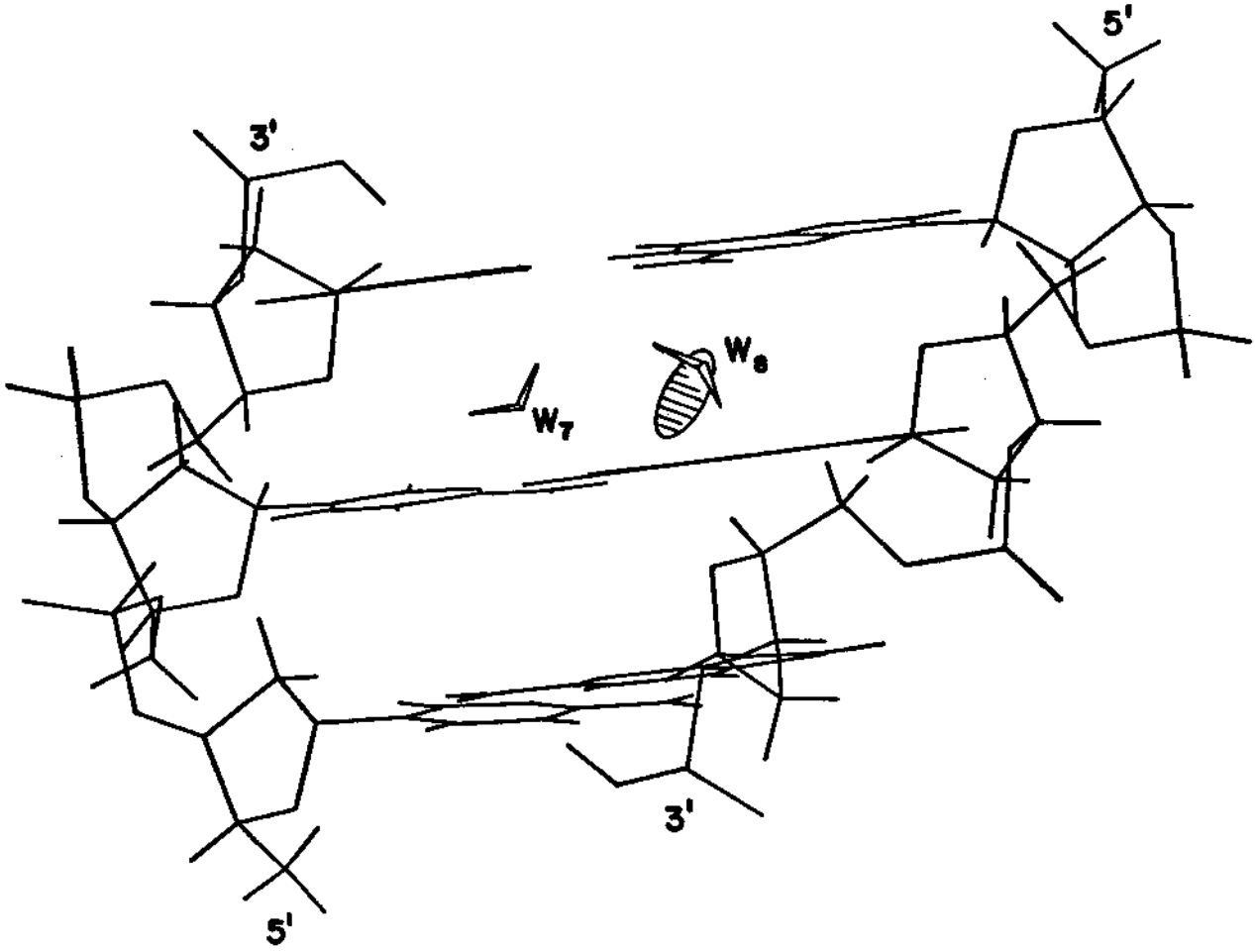

FIG. 4

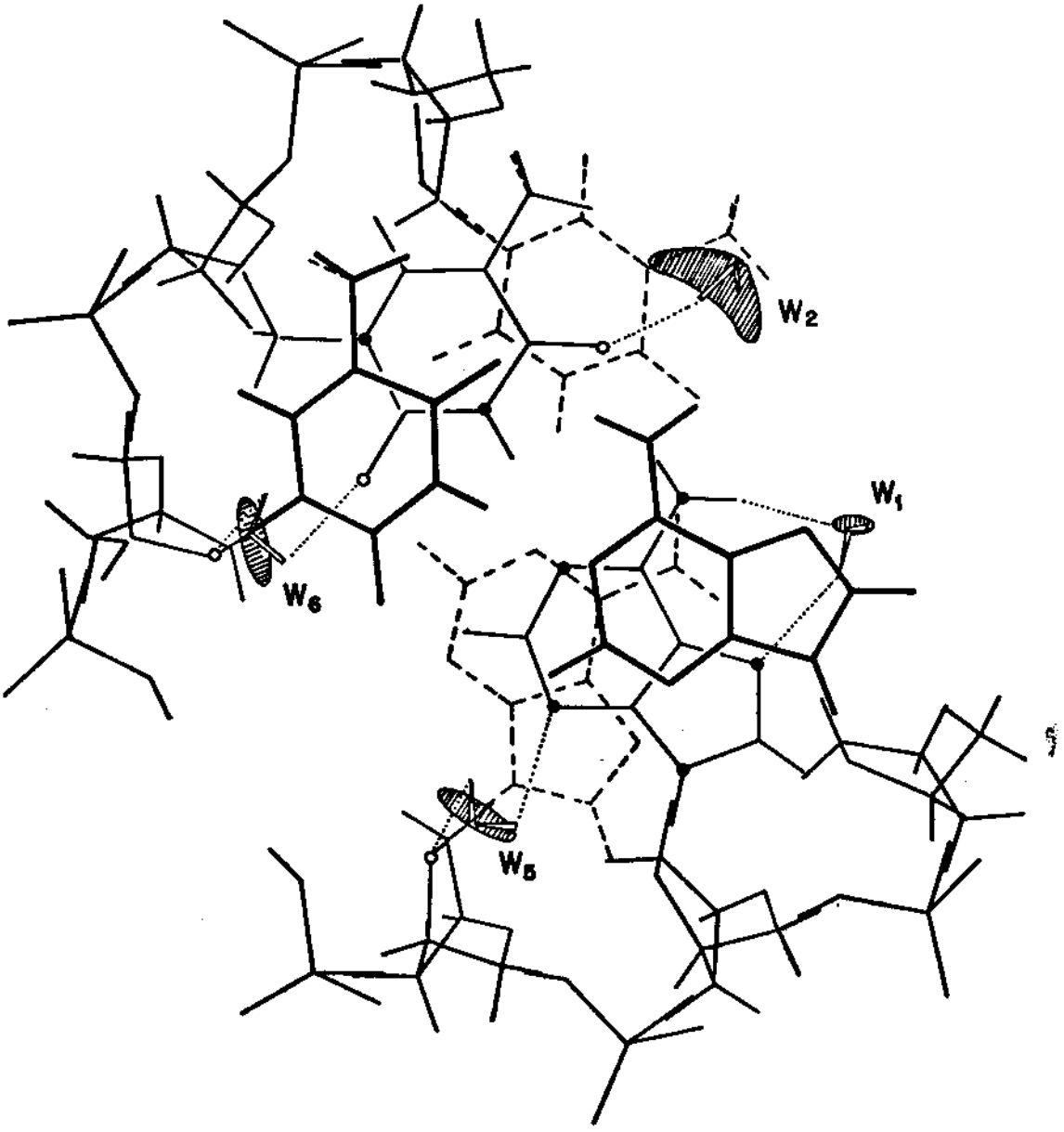

FIG. 5

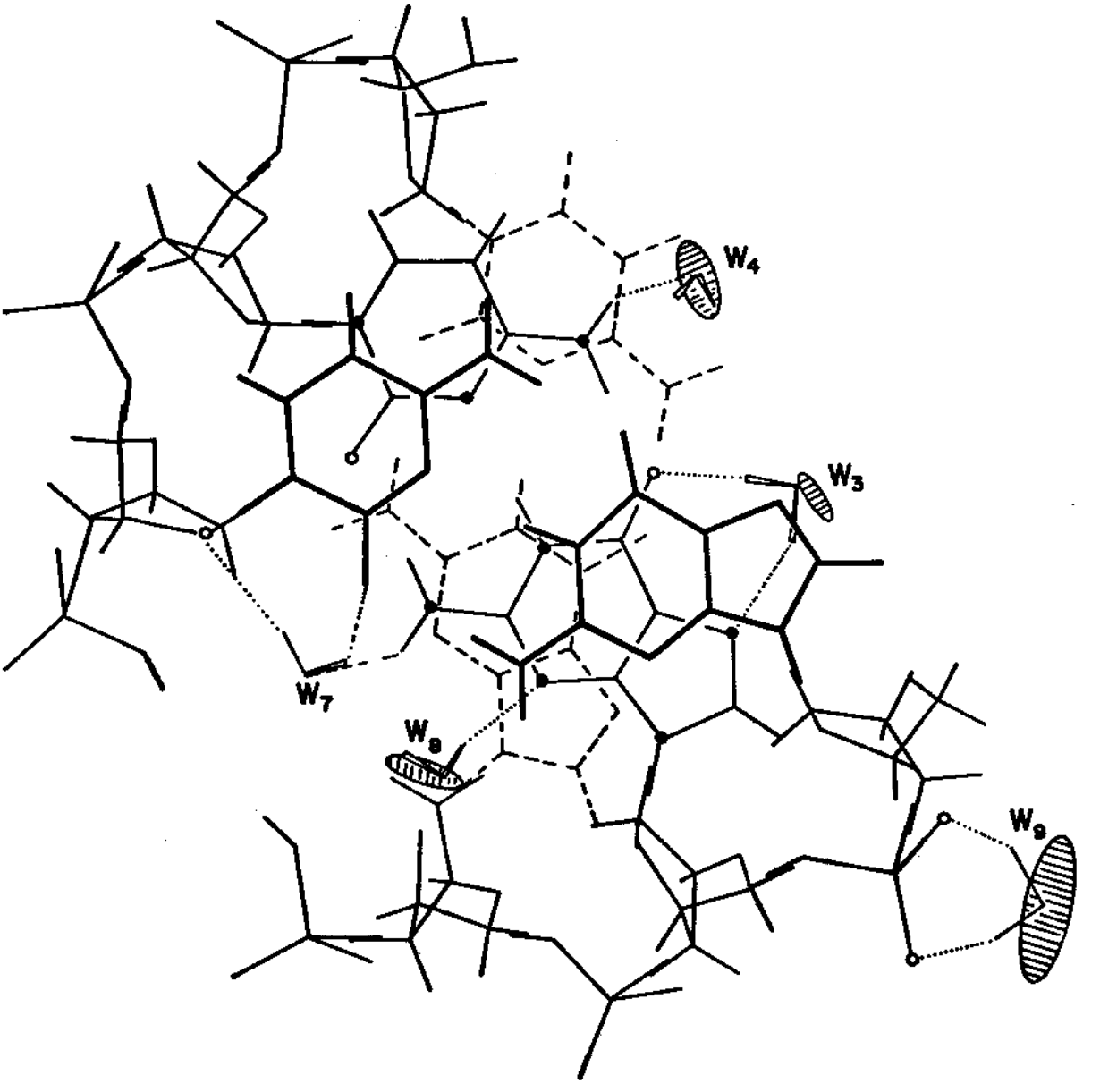

FIG. 6

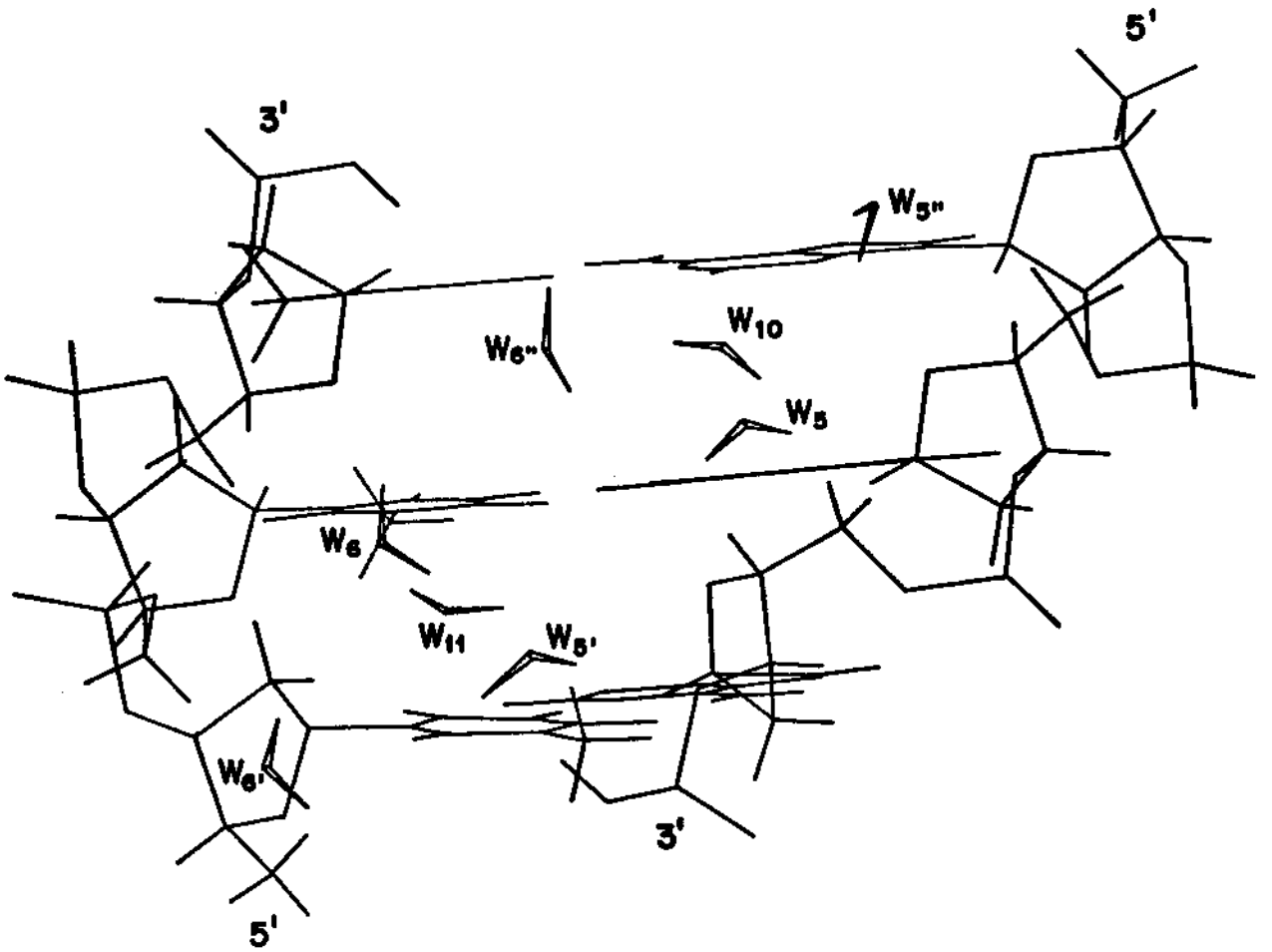

FIG. 7

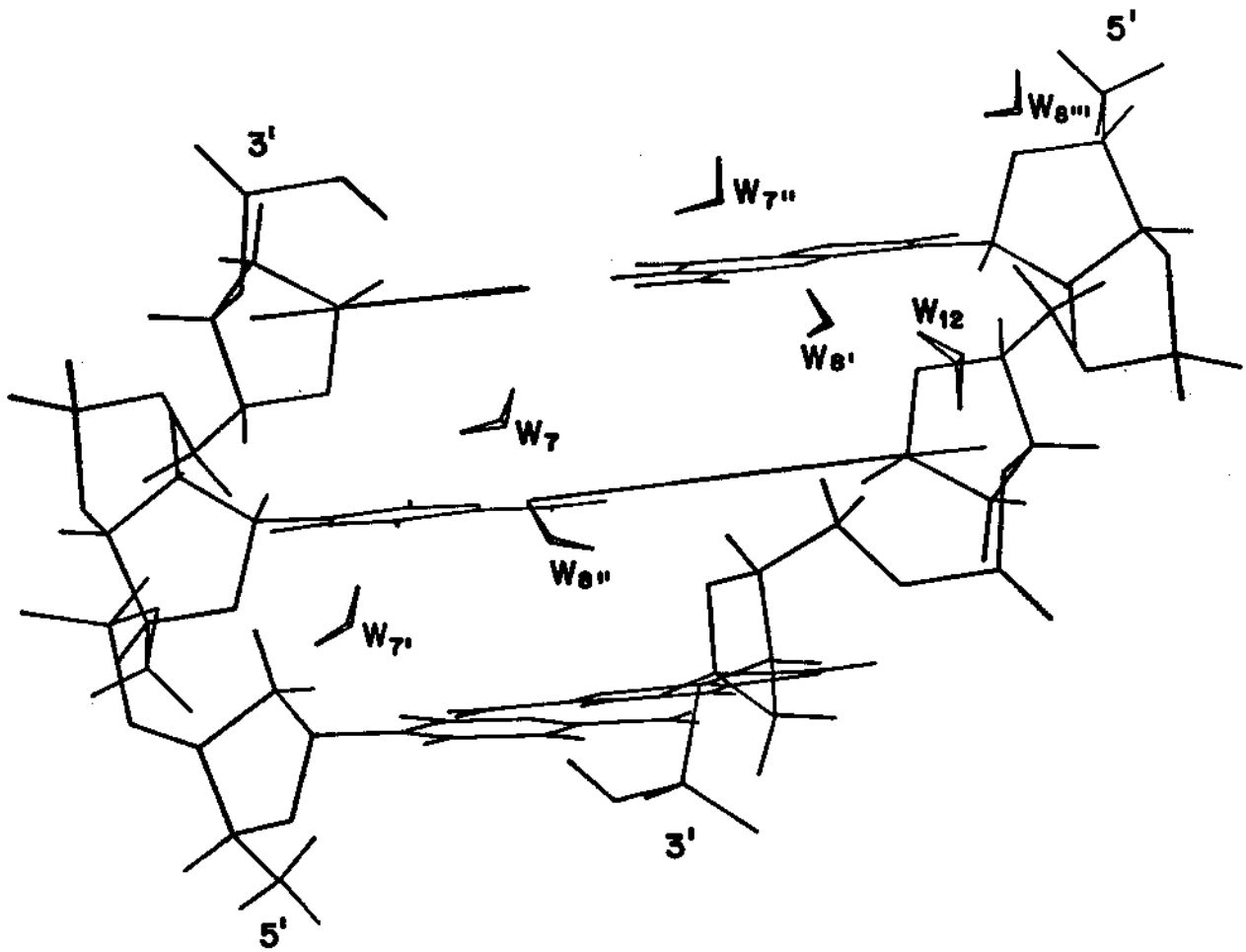

FIG. 8

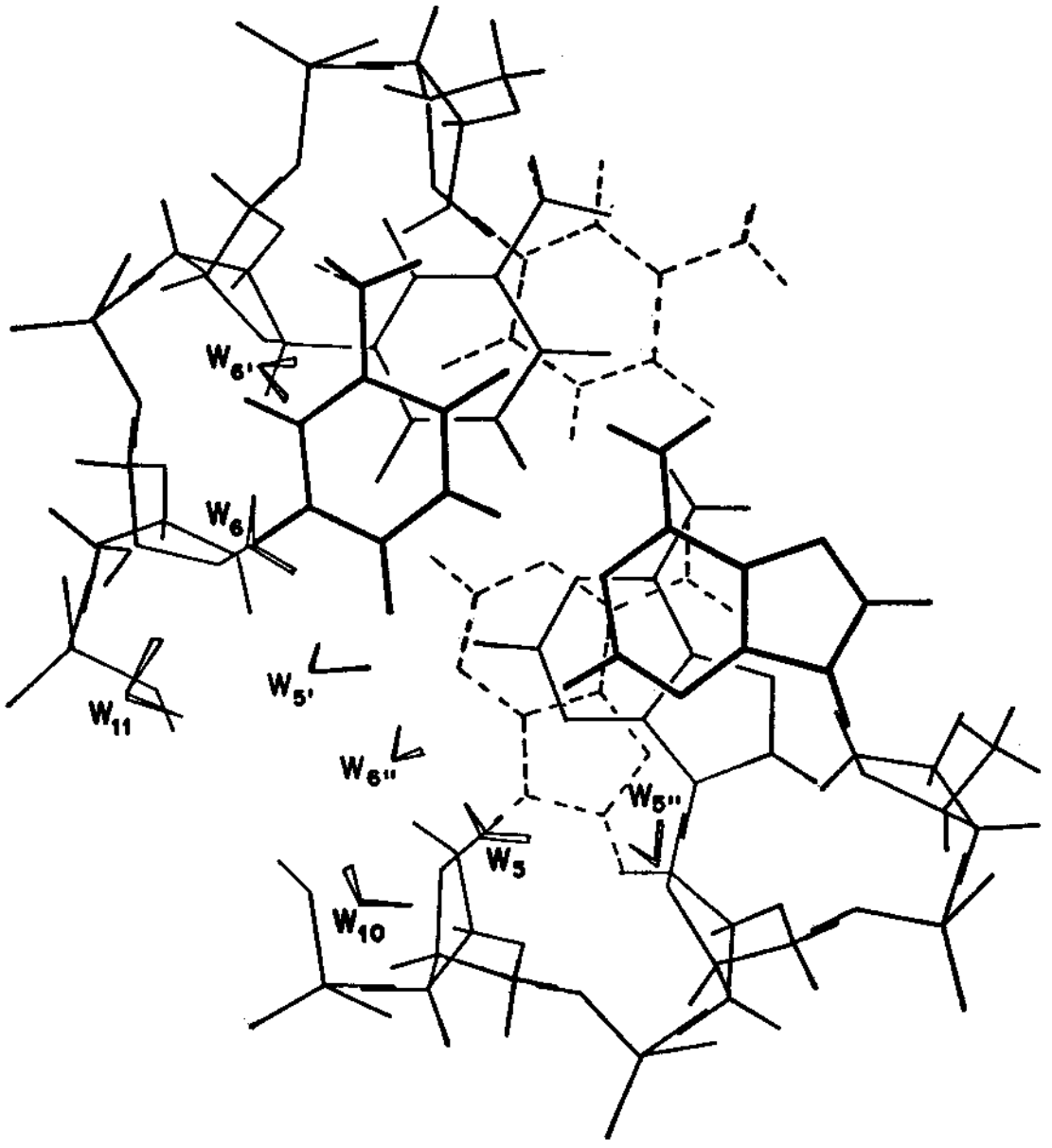

FIG. 9

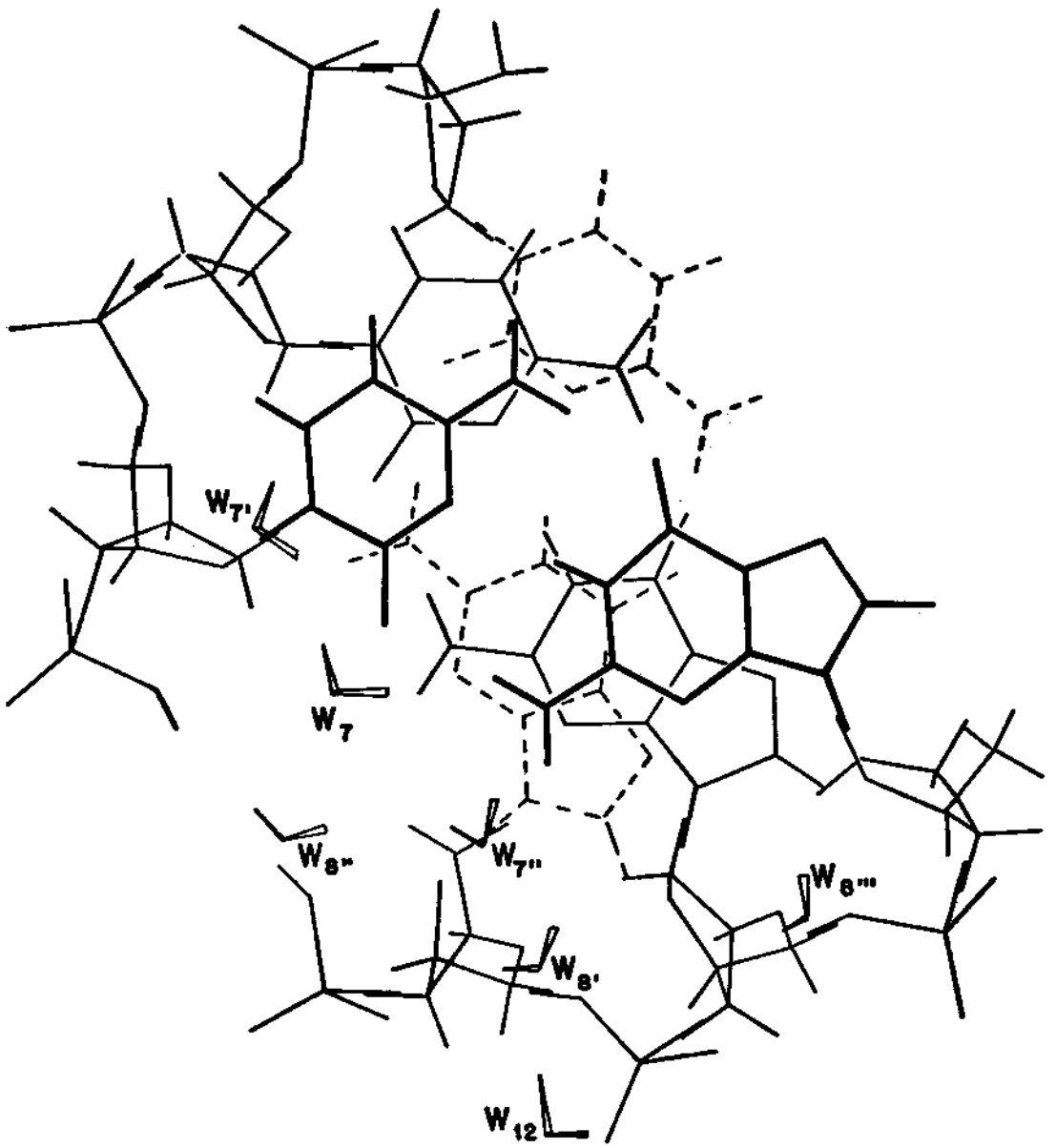

FIG. 10

Table 1

Preferred interaction sites between a water molecule and a model fragment of B-DNA. NC denote the number of a water molecule configurations at each binding site having energies within 1 Kcal/mole of the optimum energy.

Region	Base pair sequence	Receptor atoms involved	ΔE (Kcal/mole)	NC
Major Groove	AT	N_7(A), H_6(A)	- 14.6	9
		O_4(T)	- 8.4	24
	GC	N_7(G), O_6(G)	- 15.3	3
		H_4(C)	- 11.0	96
Minor Groove	AT	N_3(A), O_1,(S)	- 12.3	15
		O_2(T), O_1,(S)	- 11.9	20
	GC	O_2(C), H_2(G), O_1,(S)	- 11.8	5
		N_3(G)	- 11.0	4
Phosphate Group		O_1, O_2	- 25.5	103
		O_5,	- 20.8	2
		O_3,	—	—

REFERENCES

1. E. Clementi and G. Corongiu, Biopolymers, 18, 2431 (1979).

2. E. Clementi and G. Corongiu, Int. J. Quant. Chem., 16, 897 (1979).

3. E. Clementi and G. Corongiu, Chem. Phys. Lett., 60, 175(1979).

4. G. Corongiu and E. Clementi, Biopolymers, 20, 551(1981).

5. E. Clementi, G. Corongiu, M.Gratarola, P. Habitz, C. Lupo , P. Otto and D. Vercauteren, Int. J. Quant. Chem. Symp.,16 , 409(1982).

6. Submitted to publication in this Journal.

7. G.N.J. Port and A. Pullman, FEBS Lett., 31, 70(1973).

8. A. Pullman and D. Perahia, Theoret. Chim. Acta,48, 29(1978).

9. A. Goldblum, D. Perahia and A. Pullman, FEBS Lett., 91, 213 (1978).

10. R. Lavery, M. de Oliveira and B. Pullman, J. Comp. Chem., 1, 301(1980).

11. R. Lavery, A. Pullman, B. Pullman and M. de Oliveira, Nucleic Acid Res., 8, 5095 (1980).

12. A. Pullman and B. Pullman, Quart. Rev. Biophysics, 14, 289 (1981).

13. A. Pullman, B. Pullman and R. Lavery, in "Nucleic Acids:The Vectors of Life", (B. Pullman and J. Jortner, Eds.), p. 75, Reidel Publishing, Dordrecht, Netherlands (1983).

14. R. Wing, H.R. Drew, T. Takano, C. Broka, S. Tanaka,K.Itakura and R.E. Dickerson, Nature 287, 755 (1980).

15. H.R. Drew, R. Wing, T. Takano, C. Broka, S. Tanaka,K.Itakura and R.E. Dickerson, Proc. Natl. Acad. Sci. USA,78,2179 (1981).

16. H.R. Drew and R.E. Dickerson, J. Mol. Biol., 151,535(1981).

17. H.R. Drew, S. Samson and R.E. Dickerson, Proc. Natl. Acad. Sci. USA, 79, 4040 (1982).

18. M.L. Kopka, A.V. Fratini, H.R. Drew and R.E. Dickerson, J. Mol. Biol., 163, 129 (1983).

19. B.N. Connor, T. Takano, S. Tanaka, K. Itakura and R.E.Dickerson, Nature, 295, 294 (1972).

20. A.H.J. Wang, S. Fukii, J.H. van Boon and A. Rich, Proc. Natl. Acad. Sci. USA 79, 3968 (1982).

21. B.N. Conner, C. Yoon, J.L. Dickerson and R.E. Dickerson , J. Mol. Biol., 174, 663 (1984).

22. H.R. Drew, T. Takano, S. Tanaka, K. Itakura and R.E.Dickerson, Nature, 286, 567 (1980).

23. S. Arnott, R. Chandrasekaran, R. Birdsall, A.G.W. Leslie , and R.L. Ratliff, Nature, 283, 743 (1980).

24. S. Arnott and D.W.L. Hukins, Biochem. Biophys. Res. Comm., 47, 1504 (1972).

25. C. Giessner - Prettre and B. Pullman, Biochem. Biophys. Res. Comm., 107, 1539 (1982).

26. M. Falk, K.A. Hartman and R.C. Lord, J. Am. Chem. Soc.,84, 3843 (1962).

27. M. Falk, K.A. Hartman and R.C. Lord, J. Am. Chem. Soc.,85, 387 (1963).

28. M. Falk, K.A. Hartman and R.C. Lord, J. Am. Chem. Soc.,85, 391 (1963).