Volume 9 Number 18 1981 Nucleic Acids Research

The molecular electrostatic potential and steric accessibility of A-DNA

Richard Lavery and Bernard Pullman

Institut de Biologie Physico-Chimique, Laboratoire de Biochimie Théorique, associé au C.N.R.S., 13, rue Pierre et Marie Curie, 75005 Paris, France

Received 3 July 1981

ABSTRACT

The molecular electrostatic potential and steric accessibility of A-DNA are computed for base sequences $(dG.dC)_n$ and $(dA.dT)_n$. An interpretation of the results in terms of the structure of A-DNA is provided and differences with respect to other forms of DNA, namely B-DNA and Z-DNA, are discussed.

INTRODUCTION

Already in the early history of the resolution of DNA structure it was recognized that this biopolymer could exist in two distinct families of conformations termed A-DNA and B-DNA, the former being prevalent under conditions of low humidity and relatively low salt concentration (1). Detailed studies (2-4) have since brought to light the essential differences between these two forms, amongst which we may cite the changed sugar pucker (C3'-endo in A-DNA versus C2'-endo or C3'-exo in B-DNA), the altered rise per base pair (2.59 Å in A-DNA versus 3.38 Å in B-DNA), the greater number of phosphates per turn in A-DNA (11 versus 10 in B-DNA), the tilt of the base pairs (19° in A-DNA versus 6° in B-DNA).

It seems certainly interesting to investigate what may be the effect of these structural differences on the chemical and biochemical reactivity of these two conformational varieties. We present in this paper the results of a preliminary attempt at such a comparison based on the calculation of the molecular electrostatic potential and of the accessibilities to important reactive sites of the purine and pyrimidine bases for an A-DNA double helix. Previous studies from our laboratory carried out for B-DNA have shown (5-8) the significance of these properties for the understanding of many major aspects of the reactivity of this biologically essential form of DNA.

For the sake of simplification and comparison with the results obtained for B-DNA, the study of A-DNA was carried out, as for B-DNA, for the two base pair sequences poly (dG.dC) and poly (dA.dT). In fact, it is known that DNAs

having highly repeating sequences and few guanine-guanine base stacks do not adopt stable A form and thus, in particular, this form does not seem to exist for poly (dA.dT), while it is readily formed by poly (dG.dC) (9). Poly (dA.dT). poly (dA.dT) and poly (dG.dC).poly (dG.dC) do not form the A-helix either (10). On the other hand, it has been shown recently that, contrary to previous beliefs, A-T rich, and even very A-T rich (95%) DNAs can adopt this form (11, 12). Whatever be the significance of the precise sequencing of the bases for the stability of A-DNA, there can be no doubt that the utilization of the model sequences poly (dG.dC) and poly (dA.dT) for the study of the general aspects of the molecular electrostatic potential and base sites accessibility in A-DNA will reproduce correctly the main features associated with these two properties in this conformational form of DNA.

Finally, we shall also extend the comparison to the results which we have obtained (13, 14) for a third family of DNA conformations, the left-handed double helices termed Z-DNAs, which have recently appeared on the scene (14-17).

METHOD

The model A-DNA segments studied involve full turns of this double helix with 12 phosphates in each strand and repetitive base sequences dG.dC or dA.dT. The geometry employed for these helices is that due to Arnott and Hukins (4).

Molecular electrostatic potentials were calculated by the superposition of the base, sugar and phosphate potentials using the methodology described previously (6, 7, 18). The resulting potentials are presented in three ways : 1) the potential minima associated with nucleophilic sites (attractive to electrophiles) on the nucleic acid bases, 2) potential maps in planes perpendicular to the helical axis, at the center of the model double helices, 3) potentials on surface envelopes surrounding the double helices, these envelopes being formed (7) by the intersection of spheres, with radii proportional to the van der Waals atomic radii, surrounding each atom of the nucleic acid (as in our previous studies of B- and Z-DNA the proportionality factor of 1.7 was employed).

Atomic steric accessibilities were calculated by the methodology previously described for B- and Z-DNA (8, 13, 14). The results are presented as areas (in $\AA^2$) accessible on the van der Waals sphere of the atom concerned toward a test sphere of radius 1.2 $\AA$. (For related computations on accessibility see also Alden and Kim (19, 20)).

RESULTS AND DISCUSSION

We begin by considering the potentials associated with the reactive sites of the nucleic acid bases. The results for the A-DNA double helix are contained in Table I, guanine and cytosine potentials being derived from the central base pair of the poly (dG.dC) model and adenine and thymine potentials from the central base pair of the poly (dA.dT) model. The site potentials range between - 600 and - 800 kcal/mole, the large absolute values being due to the cumulative effect of the unscreened anionic phosphate groups of the double helix. (Possible screening schemes with countercations and their effect on the potentials will be considered in a forthcoming publication).

A graphical comparison of these results is given on the left-hand side of figure 1 where the site potentials are represented on a vertical scale, guanine and cytosine on its left and adenine and thymine on its right. The deepest potentials are seen to be associated with the purine ring nitrogens N7 of guanine and adenine (notation N7(G) etc.) and with the carbonyl oxygens O6(G) and O4(T). The next deepest potentials are those of C8(G), N4(C), N6(A) and C8(A). It is essential to note that all the sites enumerated above are associated with atoms lying in the major groove of the A-DNA double helix. In contrast, the sites on the opposing side of the base pairs, N3(G), O2(C),

TABLE 1. Potentials (kcal/mole) and steric accessibilities ($\AA^2$) of the bases in A-DNA.

BASE	SITES				
GUANINE	N2	N3	N7	C8	O6
Potential	- 623	- 666	- 787	- 699	- 761
Accessibility	0.3	0.9	3.4	0.8	2.3
CYTOSINE	N4	C5	O2		
Potential	- 692	- 656	- 638		
Accessibility	0.1	0.4	1.6		
ADENINE	N3	N6	N7	C8	
Potential	- 663	- 682	- 748	- 670	
Accessibility	2.1	0.1	2.7	0.7	
THYMINE	C5	O2	O4		
Potential	- 665	- 655	- 725		
Accessibility	0.0	2.8	1.3		

N3(A) and O2(T) are considerably weaker. We will return to this distinction shortly.

Table I also contains the atomic accessibilities for the base atoms associated with the potential sites discussed above. A graphic representation of these results is also given on the left-hand side of figure 2. The most accessible atoms are found to be the ring nitrogens and carbonyl oxygens intermingled in the resulting order :

N7(G) > O2(T) > N7(A) > O6(G) > N3(A) > O2(C) > O4(T)

With the exception of O2(T) the three most accessible atoms are thus located in the major groove of the helix. In distinction to N3(A), N3(G) has a rather low accessibility as do the purine C8 atoms. The lowest accessibilities are those of the pyrimidine C5 and the amino nitrogens.

Since accessibilities and potentials have been shown, through the examples of B- and Z-DNAs and of $tRNA^{Phe}$ (21), to play important roles in determining

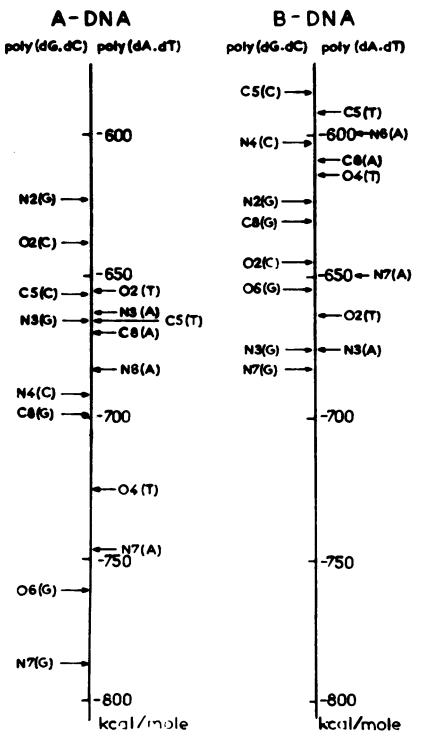


FIGURE 1. Comparison of A-DNA and B-DNA base site potentials (kcal/mole).

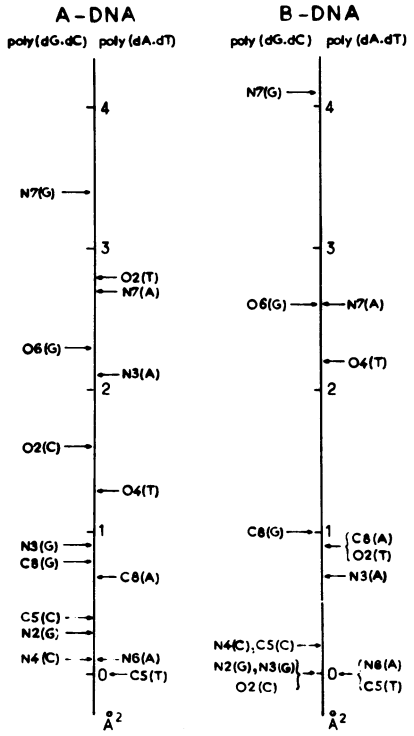


FIGURE 2. Comparison of A-DNA and B-DNA base atom accessibilities ($Å^2$).

the reactivity of the bases of these nucleic acids we can hope to predict similarly, be combining these two factors, particularly favourable sites in A-DNA. For this acid one obvious conclusion would be that electrophiles should preferentially attack N7(G), N7(A) and O6(G). Situated in the major groove these three sites have both the deepest potentials and great accessibilities. O2(T), O2(C) and N3(A), situated in the minor groove have large accessibilities but weaker potentials and O4(T) with a rather deep potential is disfavoured by a smaller accessibility. The remaining sites have both low accessibilities and relatively weaker potentials.

It may be interesting to compare the results found here for A-DNA with the previous ones (8, 13) for B-DNA. The comparison is presented in Fig. 1 and 2. As for A-DNA, N(7)G, N7(A) and O6(G) were also found to be associated with relatively deep potentials and good accessibilities in B-DNA. It may be observed from Fig. 1 that for these highly reactive positions the potentials are much deeper in A-DNA that in B-DNA. In fact, it can be seen from this figure that the limit of potential is deeper in A-DNA than in B-DNA. This is due, partly, to the greater number of phosphates in the models used for the A-DNA double helix (24 against 22 for B-DNA,in a helical turn) and, more importantly perhaps, to the tighter packing of phosphates in A-DNA, whose pitch is 28.2 Å compared to 33.8 Å for B-DNA. This overall situation does not preclude, however, the fact that the potentials of N3(G), O2(C), N3(A) and O2(T) (all in the minor groove) are weaker in A-DNA than in B-DNA.

Concerning the accessibilities, the most significant observations are the slight decrease of accessibility in A-DNA with respect to B-DNA for N7(G) and O4(T) (in the major groove) and an increase for N3(G), N3(A), O2(C) and O2(T) (in the minor groove).

These observations may be summarized by saying that, in passing from B-DNA to A-DNA the potentials associated with the major groove sites deepen while their accessibilities generally decrease (although they remain quite appreciable). the potentials associated with the minor groove sites become weaker and their accessibilities generally increase. The explanation for these phenomena appears to be in large part due to the displacement of the base pairs in A-DNA away from the helical axis toward the minor groove. The result is that this groove is extremely shallow in A-DNA, in fact, it represents a nearly flat face, with the bases exposed, while, in contrast, the major groove becomes narrow and very deep. The trends in accessibility thus become understandable, the "minor groove" sites in A-DNA being more exposed and the major groove sites more hindered.

The trends in the distribution of the potential become perhaps more understandable by looking on the isopotential maps in the planes perpendicular to the helical axis and passing through the central base pair of our model double helices. The results are shown in figure 3 for the poly (dG.dC) model of A-DNA and in figure 4 for its poly (dA.dT) one. In both figures the helical axis is indicated by a central cross, the phosphorus atoms of the backbones passing through the plane are marked by the letters P and the central base pair is indicated. From these figures the displacement of the base pairs from the helix axis is clear. As a result of this situation, the backbones bordering the deep and narrow major groove produce much more negative potentials in this region than those found on the opposing flat side of the double helix. This effect is accentuated by the orientation of the phosphate anionic oxygens which point across the entrance to the major groove, rather than to the outside of the double helix as in B-DNA. The presence of more negative site potentials in the major groove is thus comprehensible.

This situation is an interesting contrast to our findings for Z_I-DNA (13, 14) where, according to the geometry due to Wang et al. (15), the bases of this left-handed double helix are displaced towards the side of the "major groove", which becomes in fact a convex surface, leaving only a deep minor groove with a consequent deepening of the site potentials on the corresponding side of the nucleic acid bases, exactly the reverse of the situation in A-DNA.

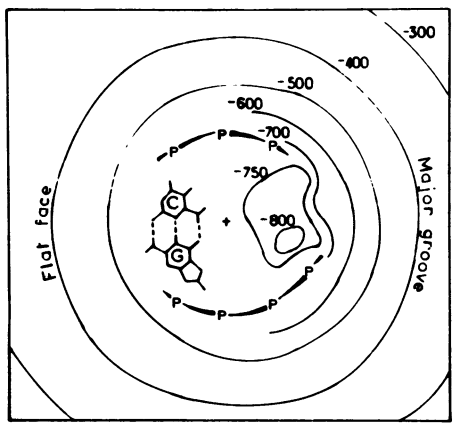


FIGURE 3. Midplane potential for A-DNA, poly (dG.dC) model (isopotentials in kcal/mole).

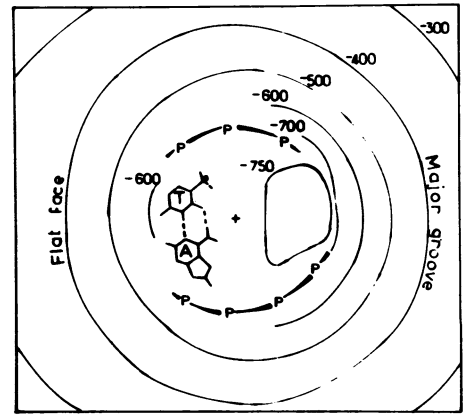


FIGURE 4. Midplane potential for A-DNA, poly (dA.dT) model (isopotentials in kcal/mole).

A final view of the distribution of the potential around A-DNA is given by the potentials on the surface envelopes surrounding the model double helices. A three-dimensional representation of the envelope for the poly (dG.dC) model of A-DNA is shown in figure 5, in which the deep, narrow major groove is clearly visible in the lower half of the figure. Above it may be seen the virtually flat opposing face. Figure 6 shows a graphic of this poly (dG.dC) double helix surrounded by the projected profile of the surface envelope and in figure 7 the surface potentials projected on a plane are indicated by different degrees of shading, darker shadings implying more negative potentials (see table 2 for details). Comparing figures 6 and 7, the most negative potentials are seen to be concentrated in the major groove of the double helix in the lower half of figure 7. The minimum potential of - 790 kcal/mole, (denoted by the letter M) lies on the surface of a phosphate anionic oxygen on the edge of the groove. On the opposing flat side of the helix (upper half of figures 6 and 7) there is almost no concentration of potential over the bases and the values are some 100 kcal/mole weaker than in the major groove (local minimum = - 700 kcal/mole).

The results for the poly (dA.dT) model of A-DNA are given in figures 8 (the molecular graphic of the double helix) and 9 (the surface envelope potential). A three-dimensional representation of the envelope is not given for this base sequence as it is virtually indistinguishable from that in figure 5. Once again a concentration of the most negative potentials can be observed in the major groove of the helix, in the lower half of figure 9. The minimum potential, denoted by the letter M, is - 784 kcal/mole. As for the poly (dG.dC) model the potentials on the opposing flat side of the helix are roughly 100 kcal/mole less negative (local minimum = - 696 kcal/mole).

Finally if the poly (dG.dC) sequence is compared with the poly (dA.dT)

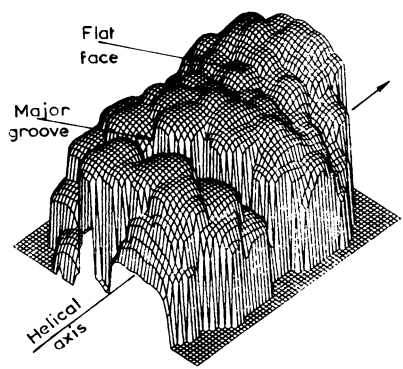
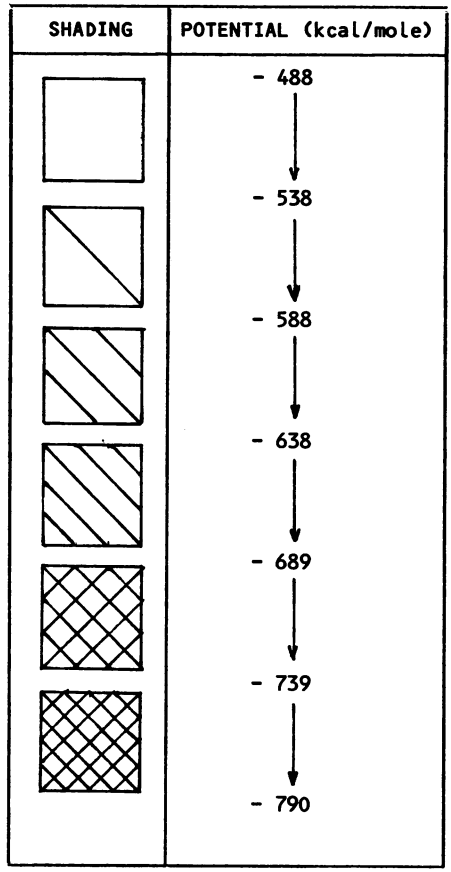


FIGURE 5. Surface envelope graphic of A-DNA, poly (dG.dC) model.

TABLE 2. Shading used for surface envelope potentials



sequence either in terms of the mid-plane potentials (figures 3 and 4) or in terms of the envelope potentials (figures 7 and 9) the same overall order of decreasingly negative potentials may be observed for the grooves, namely :

poly (dG.dC) major > poly (dA.dT) major > poly (dG.dC) minor > poly (dA.dT) minor

CONCLUSIONS

The calculation of the molecular electrostatic potential and steric accessibility for a model A-DNA double helix has drawn attention to the possible importance of the base pair displacement away from the helical axis in influencing the chemical and biochemical reactivity of this DNA conformation. The

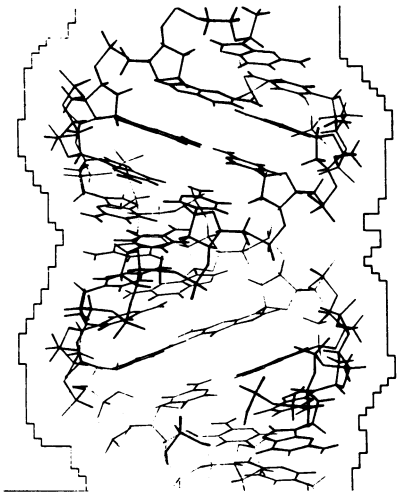


FIGURE 6. Molecular graphic of A-DNA, poly (dG.dC) model.

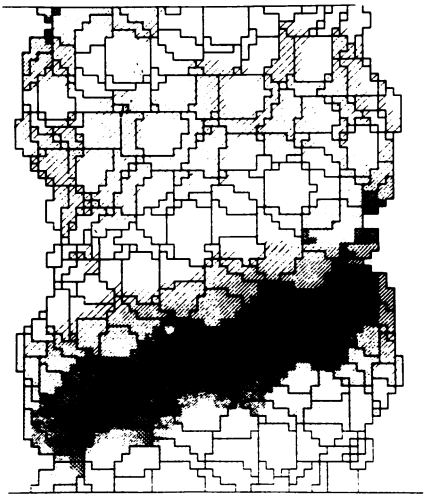


FIGURE 7. Surface potentials for A-DNA, poly (dG.dC) model (for significance of shading see table 2).

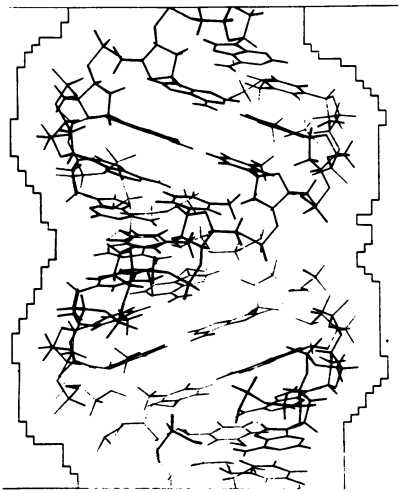


FIGURE 8. Molecular graphic of A-DNA, poly (dA.dT) model.

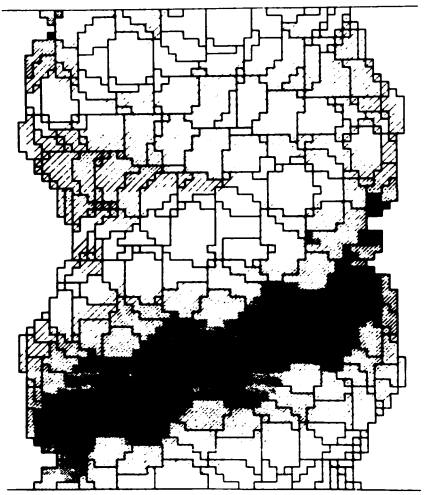


FIGURE 9. Surface potentials for A-DNA, poly (dA.dT) model (for significance of shading see table 2).

resulting inbalance between the two sides of the base-pairs with deeper potentials, but decreasing accessibilities, in the major groove and the flattening of the minor groove with variable effect on both properties, represent important changes with respect to B-DNA. In particular while in B-DNA the major groove represented the potentially priviledged region for electrophilic attacks in G-C rich segments and the minor groove such a region in A-T rich segments, the major groove will manifestly be predominant in A-DNA independently of its composition. This situation may perhaps be responsible in part for the fact that the recently much explored oligopeptide antibiotics netropsin and distamycin A, which interact without intercalation with nucleic acids and bind preferentially to the minor groove of A-T rich regions of helical B-DNA, bind much less, if at all, to A-DNA or to helical RNA (whose structure is similar to A-DNA) (22-24). Other, complementary reasons for this selectivity could be 1) the more negative values of the potentials at N3(A) and O2(T) in B-DNA than in A-DNA, these atoms being the ones utilized for hydrogen bonding with the amide hydrogens of the antibiotics and 2) the unsuitable arrangement of the anionic oxygens of the phosphates (pointing across the entrance to the major groove) in A-DNA : in B-DNA hydrogen bonds linking the positively charged propionamidine group(s) of the antibiotics with these oxygens stabilize the interaction.

The possible occurrence of local A-DNA characteristics is an overall B-DNA double helix, a possibility substantiated by the recent detailed study of the influence of base sequence on helix structure in the case of a B-DNA dodecamer (25), also implies the occurence of the associated features of electrostatic potential and accessibility. Such an occurence may therefore be significant for the manifestations of local biochemical reactivity. This aspect of the situation has been considered recently extensively with respect to Z-DNA and the possible significance of Z-DNA sequences in a B-DNA helix for the affinity toward chemical carcinogens (14, 15). It may be interesting to consider the same problem with respect to A-DNA fragments. As examples we may consider the NH_2 group and C8 atom of guanine which are targets for the carcinogenic polycyclic aromatic hydrocarbons and N-acetoxy-N-2-acetylaminofluorene, respectively. The data of table I and Fig. 1 indicate that the potential at N2(G) is practically the same in A- and B-DNA and its accessibility a little greater in A-DNA. There is therefore little probability that the presence of an A-DNA step in a B-DNA helix should produce a significant enhancement of the reaction with polycyclic aromatic carcinogens. On the other hand, while the accessibilities to C8(G) are very similar in A- and B-DNAs, the potential associated

with this site is significantly deeper in A-DNA than in B-DNA. Consequently, the presence of an A-DNA fragment in a B-DNA helix could produce an enhanced affinity for N-acetoxy-N-2-acetylaminofluorene.

Finally, we must remark that in spite of the obvious interest in evaluating the intrinsic molecular electrostatic potential of naked nucleic acids there is at least a similar interest in evaluating the influence of counter-cation binding. We have presented such studies for B-DNA (24, 25) and Z-DNA (13, 14) and shall present in a forthcoming publication the results for A-DNA.

ACKNOWLEDGMENT

The authors wish to thank Professor R.E. Dickerson for communication of manuscripts prior to publication and the National Foundation for Cancer Research for its support of this work.

REFERENCES

(1) R.E. Franklin and R.G. Gosling, Nature 171, 740-741 (1953).
(2) G. Malenkov, L. Minchenkova, E. Minyat, A. Schyolkina and V. Ivanov, FEBS Lett. S1, 38-42 (1975).
(3) S. Arnott, S.D. Dover and A.J. Wonacott, Acta Cryst. 25, 2192-2206 (1969).
(4) S. Arnott and D.W.L. Hukins, Biochem. Biophys. Res. Comm. 47, 1504-1509 (1972).
(5) B. Pullman, D. Perahia and D. Cauchy, Nucl. Acids Res. 6, 3821-3829 (1979).
(6) D. Perahia, A. Pullman and B. Pullman, Int. J. Quant. Chem., Quant. Biol. Symp. 6, 353-363 (1979).
(7) R. Lavery and B. Pullman, Int. J. Quant. Chem., Quant. Biol. Symp. 8, in press.
(8) R. Lavery, A. Pullman and B. Pullman, Int. J. Quant. Chem., Quant. Biol. Symp. 8, in press.
(9) S. Arnott and E. Selsing, J. Mol. Biol. 88, 551-552 (1974).
(10) S. Arnott, R. Chandrasekaran and E. Selsing in "Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions" Eds. M. Sundaralingam and S. Rao, University Park Press, Baltimore 577-596 (1975).
(11) E. Selsing and S. Arnott, Nucl. Acid Res. 3, 2443-2450 (1976).
(12) E. Selsing, A.G.W. Leslie, S. Arnott, J.G. Gall, D.M. Skinner, E.M. Southern, J.H. Speneer and K. Harbers, Nucl. Acid Res. 3, 2451-2458 (1976).
(13) K. Zakrzewska, R. Lavery, A. Pullman and B. Pullman, Nucl. Acids Res. 8, 3917-3932 (1980).
(14) K. Zakrzewska, R. Lavery and B. Pullman, in "Biomolecular Stereodynamics" Ed. R.H. Sarma, Pergamon Press, New York, in press.
(15) A. Wang, H-J. Quigley, F.J. Kolpak, G. van der Marel, J.H. van Boom and A. Rich, Science 211, 171-176 (1980).
(16) H.R. Drew, T.Tanako, S. Tanaka, K. Itakura and R.E. Dickerson, Nature 286, 567-573 (1980).
(17) S. Arnott, R. Chandrasekaran, D.C. Birdsall, A.G.W. Leslie and R.L. Ratliff, Nature 283, 743-745 (1980).
(18) A. Pullman and B. Pullman in "Molecular Electrostatic Potentials in Chemistry and Biochemistry" Eds. P. Politzer and D.G. Truhlar, Plenum, New York (1980) 381-405.
(19) C.J. Alden and S-H. Kim, J. Mol. Biol. 132, 411-434 (1979).

(20) C.J. Alden and S-H. Kim, in "Stereodynamics of Molecular Systems" Ed. R.H. Sarma, Pergamon Press, New York (1979) 331-349.
(21) R. Lavery, A. Pullman, B. Pullman and M. de Oliveira, Nucl. Acid Res. 8, 5095-5112 (1980).
(22) Ch. Zimmer, Progr. in Nucl. Acid Res. and Mol. Biol. 15, 285-317 (1975).
(23) G.V. Gursky, V.G. Tumanyan, A.S. Zasedatelev, A.L. Zhuze, S.L. Grohovsky and B.P. Gottickh in Nucleic Acid-Protein Interactions, H.J. Vogel Ed. Academic Press, New York (1977) 189-217.
(24) K.E. Reinert, D. Geller and E. Stutter, Nucl. Acid Res. 9, 2335-2349 (1981).
(25) R.E. Dickerson and H.R. Drew, J. Mol. Biol., in press.
(26) D. Cauchy, R. Lavery and B. Pullman, Theoret. Chim. Acta 57, 323-327 (1980).
(27) R. Lavery, D. Cauchy, O. de la Luz Rojas and A. Pullman, Int. J. Quant. Chem. Quant. Biol. Symp. 7, 323-330 (1980).