Systematic Study of Nuclear Overhauser Effects vis-à-vis Local Helical Parameters, Sugar Puckers, and Glycosidic Torsions in B DNA: Insensitivity of NOE to Local Transitions in B DNA Oligonucleotides Due to Internal Structural Compensations†

Nikolai B. Ulyanov,‡,§ Andrei A. Gorin,§ Victor B. Zhurkin,§ Ban-Chin Chen,‡ Mukti H. Sarma,‡ and Ramaswamy H. Sarma*,‡

Institute of Biomolecular Stereodynamics, Department of Chemistry, State University of New York, Albany, New York 12222, and The Engelhardt Institute of Molecular Biology, Academy of Sciences of the USSR, 117984 Moscow B-334, USSR

Received September 24, 1991; Revised Manuscript Received February 20, 1992

ABSTRACT: A method has been developed to solve structures of DNA oligomers in solution from the experimental NOE data. The method is a combination of two approaches: (1) full matrix NOESY simulations and (2) conformational calculations of DNA double helix based on generalized helical parameters. The process of the refinement of a solution structure does not involve NMR-derived interproton distance constraints; rather it consists of a direct fitting of a structure to the experimental NOE data, a weighted sum of energy, and *R* factor being under minimization. A helical parameters-based generation of DNA forms makes it possible to organize the search for the optimal structure more effectively, systematically varying starting conformations. The method has been used to calculate a structure for the self-complementary DNA hexamer GGATCC, which is consistent with the available experimental data. The structure belongs to the B family of forms, although the local structural heterogeneity is very strong. Sugar puckers vary from O4′-exo to C3′-exo; helical steps are open with different magnitudes toward the minor groove. Next, we have addressed the question of how uniquely the structure is defined by the existing NMR data. Different structural parameters have been systematically varied, and their effect on individual NOE's and the *R* factor has been studied. Two energetically conjugated parameters, sugar puckers and glycosidic angles, can be determined very reliably, because of the strong dependences of the intraresidue H6/H8 to H2′/H2″/H3′ NOE's. In contrast, the local helical conformation of DNA and the geometry of base pairs proved to be underdetermined by the existing NOE information, because the effect of any helical parameter on interproton distances can be compensated by the concerted changes in other parameters.

A great deal of effort has been devoted lately to the investigation of the sequence-dependent structural microheterogeneity of DNA[1] in solution. This direction of research is stimulated by several factors. From the several available crystal structures of DNA oligonucleotides, it is clear that the local conformation of B DNA does depend to a large extent on the nucleotide sequence [for a review, see Yanagi et al. (1991)]. Besides, there is a general understanding that in certain cases such conformational variations may have a biological function, e.g., it has been documented that the sequence-dependent bending not only is involved in such relatively "passive" processes as packaging of DNA in chromatin but also participates in such "active" events as transcription, replication, recombination, etc. [for reviews, see Trifonov (1985), Steitz (1990), and Hagerman (1990)].

Two-dimensional NOESY spectroscopy has become a traditional tool for elucidation of solution structure of DNA oligomers. Different approaches are being used to extract and utilize the information about relative positions of protons from the NOESY spectra. In order to avoid systematic errors due to multiple pathways of magnetization transfer, a number of methods have been developed to simulate NOESY spectra for a given structural model (Keepers & James, 1984; Olejniczak et al., 1986; Lefevre et al., 1987). These methods permit the derivation of a set of unbiased distance constraints, which are then incorporated into a subsequent structural refinement (Banks et al., 1989; Metzler et al., 1990; Schmitz et al., 1991). In more direct approaches, a direct fitting of a structure to the experimental NOE data is used (Gupta et al., 1988; Lane, 1990; Baleja et al., 1990).

However, a question arises as to how precisely a local geometry of DNA is defined by NOESY spectroscopy. In other words, how unique is a structure which is consistent with experimental NOE's. Recently, this problem was addressed by Metzler et al. (1990). They independently refined several structures for two DNA decamers, starting from different initial conformations which were generated by a distance geometry algorithm. All obtained structures were in a good agreement with the experimental data as decided by a visual comparison with simulated NOESY spectra. The variation of local geometries in these few structural models was so significant that the authors have concluded that it is not possible to determine reliably the dependence of helical parameters on the sequence of DNA in solution from the existing NMR data (Metzler et al., 1990).

† This work was supported by a grant from the National Institutes of Health (GM29787) and by a contract from the National Foundation of Cancer Research. The high-field NMR experiments were performed at the NMR facility for Biomolecular Research located at the Francis Bitter National Magnet Laboratory, MIT. The NMR facility is supported by Grant RR00995 from the Division of Research Resources of the NIH and by the National Science Foundation under Contract C-670. We thank the Advanced Scientific Computing Laboratory, NCI Frederick Cancer Research Facility for access to the Cray X-mp/24.

‡ State University of New York.

§ Institute of Molecular Biology, Moscow.

[1] Abbreviations: DNA, deoxyribonucleic acid; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect.

0006-2960/92/0431-3918$03.00/0

The objective of the present study was to investigate rigorously and systematically how simulated NOE's depend on various helical parameters and to determine the ranges of parameters which are in agreement with the experimental NOESY spectra. For this purpose, an approach was developed incorporating two methods which have been in use previously in our laboratories. One of them is a full matrix NOESY simulation for a given structure of DNA (Keepers & James, 1984). The second one is a method to generate a double-helical DNA conformation (Zhurkin et al., 1978). The method uses generalized helical parameters as independent variables; the present modification of the method (Ulyanov et al., 1989) takes into account all possible degrees of freedom in base pairs and between neighboring pairs as defined by the Cambridge convention (Dickerson et al., 1989). The combination of these two methods yields a program which permits calculation of DNA structure from the helical parameters, calculation of energy of the structure, simulation of NOESY spectra for the structure, calculation of an agreement R factor between simulated and experimental NOESY spectra, and minimization of a weighted sum of energy and R factor. An essential feature of our approach is a direct fitting of a structure to the experimental NOE data, without a preliminary determination of distance constraints. The use of the generalized helical parameters rather than Cartesian coordinates as independent variables reduces greatly the number of degrees of freedom in the system. But it also enables us to generate structures with helical parameters that are set to any predetermined values and to conduct a "conditional refinement" where any parameter can be held fixed, thus making it possible to calculate the dependences of R factor as well as individual NOE's on various helical parameters.

The self-complementary hexadeoxyribonucleotide d-$(GGATCC)_2$ was chosen for the present investigation. It contains a site for the *Bam*HI restriction enzyme and is referred to further as the *Bam*HI hexamer. It has NOESY spectra with relatively well-separated cross peaks (Sarma et al., 1985), and this is the main reason for selecting this sequence for the present study. The proposed method was used to refine the solution structure and to estimate the range of helical parameters which are consistent with the experimental NOESY spectra for the *Bam*HI hexamer.

Materials and Methods

The *Bam*HI hexamer $d(GGATCC)_2$ was synthesized from blocked dinucleotides using phase phosphotriester techniques. The numbering scheme for the *Bam*HI hexamer is

5′-G1-G2-A3-T4-C5-C6
C6-C5-T4-A3-G2-G1-5′

NMR Spectroscopy. The DNA concentration was 1.9 mM in duplex, the salt concentration was 100 mM in NaCl (pH 7.0 in 10 mM sodium phosphate buffer with 1 mM EDTA). The 2D NOESY spectra at 7 °C in D_2O were recorded at 500 MHz by the phase-sensitive methods (States et al., 1982) with a relaxation delay 1.5 s and a data matrix ($t_2 = 1024 \times t_1 = 350$); 64 transients were acquired per t_1 experiment. Mixing times were 10, 25, 50, and 150 ms. The residual HDO signal was presaturated. All spectra were Fourier transformed into a matrix ($t_2 = 2048 \times t_1 = 1024$) with zero filling.

Measurements of NOE's. Individual NOE's were measured by integrating areas of cross peaks in appropriate slices, only the slices with well-resolved diagonal peaks being used. In the case of overlapped cross peaks, the sums of several NOE's were measured and used for the calculation of the R factor. NOE's were quantified by normalizing the magnetization of a cross peak (its area on a slice) to the total magnetization of the corresponding slice, including the autopeak and all cross peaks (Broido et al., 1985). All (normalized) NOE's are expressed in percents (when we say that some NOE of 10% was increased by 10%, throughout all the paper we mean the absolute values, not the relative change, i.e., the new NOE is 20%, not 11%).

NOESY Simulation. NOESY simulation was performed as earlier (Gupta et al., 1988; Sarma et al., 1988, 1990; Umemoto et al., 1990a,b), by calculating an exponent of a relaxation matrix. Expressions for the relaxation rates and spectral densities were used as they are given by Keepers and James (1984) for the molecules with a single isotropic overall motion. A single effective correlation time was used for the whole molecule, since Reid et al. (1989) have shown recently that bases and sugars have a similar effective correlation time in DNA oligonucleotides. The protons of the methyl group of thymine T4 were treated explicitly. For the sake of simplicity, no rotational averaging has been made for the methyl group, although it introduces some systematic error in the calculation of NOE's which involve CH_3 (Koning et al., 1990).

R Factor. The R factor of agreement between experimental and simulated NOE's was defined for each mixing time t_m:

$$R_{t_m} = \frac{\sum |a_{ij}^{exp}(t_m) - a_{ij}^{cal}(t_m)|}{\sum a_{ij}^{exp}(t_m)} \times 100 \qquad (1)$$

Here the superscripts "exp" and "cal" stand for experimental and calculated NOE's a_{ij}, respectively. The summation is being performed over all nondiagonal experimentally measured cross peaks (ij) plus all theoretical NOE's which are greater than 0.5%. This definition of the R factor is commonly used in crystallographic studies [see, e.g., Arnott et al. (1969) and Fratini et al. (1982)]. Only those cross peaks were considered during the computation of R factors which belong to the slices through resonances of protons with nonoverlapped diagonal peaks. Namely, we used slices through H6/H8 and H1′ of A3, T4, C5, and C6. Slices through protons of G1 and G2 were not used because of a complete overlap of H8 and H1′ resonances in G1 and G2. As in the case of NOE's, we express R factors in percents; when we say that the R factor was decreased by 1%, it always means the absolute decrease (e.g., if the old value was 20%, then the new one is 19%). Also, by R factor without an indication of mixing time we mean the sum $R_{50} + R_{150}$, as this sum was used in the refinement of the structure (see below).

Generation of Structures: Independent Variables. The main idea of our approach is to define the positions of bases by independent variables (generalized helical parameters) and then to find the conformations of a sugar–phosphate backbone by means of a chain closure procedure (Zhurkin et al., 1978; Ulyanov et al., 1989). The notations for and definitions of helical parameters as well as definitions of frames of references are in agreement with the accepted recommendations (Dickerson et al., 1989; for the definition of frame of reference, which is bound to a base pair, see also Figure 10 later in the present paper). The exact analytical definitions of step (wedge) parameters (Ω, τ, ρ, D_x, D_y, D_z) and pair parameters (ω, κ, σ, S_x, S_y, S_z) are given by Gorin et al. (1990). We call these parameters "active", as they are used to generate the structure of double helix. In addition, the sugar conformations [which are defined by a single parameter, pseudorotation P (Ulyanov & Zhurkin, 1982)] and the glycosidic angles χ were varied independently for each residue.

Description of Structures. The geometry of the base pair step (wedge) is defined by six parameters, three rotations and

three translations. Our FORTRAN program to generate the DNA structures is written in such a way that the three consecutive rotations around the three orthogonal axes are performed in the following order: roll, tilt, and then helical twist (Ulyanov & Zhurkin, 1982, 1984; Gorin et al., 1990). However, since different authors use different orders of rotations [e.g., Srinivasan et al. (1987) perform the rotations in the order tilt–roll–twist, whereas Levene and Crothers (1983) use the same order as we do], we devised the other ("descriptive" or "passive") set of parameters, which do not depend on this order. Our descriptive angles twist, roll, and tilt are the local parameters, which are independent of the geometry of neighboring residues. We especially checked that they satisfy the following requirements: (1) tilt (τ) and shift (D_x) equal zero for the structures with the dyad symmetry like ApT; (2) parameters for the transition from the base pair 1 to 2 are the same as for the reverse transition, from 2 to 1. The formal definition of these parameters is given in the Supplementary Material. The descriptive set of step (wedge) parameters is used throughout this paper. In the case of pair parameters, a single set of parameters was used both for generation and description of base pair geometries (Gorin et al., 1990).

Dyad Symmetry. A dyad symmetry was imposed on the structure of the *Bam*HI hexamer. Only the tetramer G1-G2-A3-T4:A3-T4-C5-C6 was treated explicitly in the computations. Two AT pairs were considered as having an identical internal geometry, i.e., they had the same values of ω, σ, S_y, and S_z; parameters buckle κ and shear S_x were of opposite signs in the two neighboring AT pairs (e.g., see Figure 10). Also, the two deoxyadenosines as well as the two thymidines had identical sugar conformations and glycosidic angles. No additional restrictions were imposed on the geometry of step ApT:ApT. The remaining part of the molecule was calculated by a symmetry transformation in those cases when we dealt with the whole hexamer (e.g., Figure 1).

Energy of the Structure. The energy of the structure was calculated in the same way as before (Gorin et al., 1990; Zhurkin et al., 1991) with the use of atom-atomic potential functions of Zhurkin et al. (1981) and Poltev and Shulyupina (1986). The distance-dependent dielectric constant $\epsilon(r) = r$ was used in the electrostatic energy calculations.

Refinement. The target function, a weighted sum of energy and R factor, was under optimization during the refinement:

$$F = \text{energy} + C(R_{50} + R_{150}) \qquad (2)$$

Here the energy is expressed in kilocalories per mole and the R factor (1) in percent. The R factor was taken with a rather high weight: $C = 2$ kcal/mol; 1% increase in $(R_{50} + R_{150})$ contributes the same amount into the target function F as a 2 kcal/mol increase in energy. The reason why we chose such a high weight of the R factor is that the objective of the present investigation was to establish the limits imposed by the experimental data on the structural parameters of DNA. On the other hand, the minimization of the R factor alone, without a term of energy, proved to be ineffective because structures obtained in such a way have an unrealistically strained sugar–phosphate backbone and/or disrupted base pairs. The data at 10 and 25 ms were not included in the refinement because of the worse signal to noise ratio at these mixing times.

Scanning Procedure. The following procedure was used to investigate the dependences of energy, R factor, individual NOE's, etc., on conformational parameters. A parameter was set to some definite value, and the structure was optimized with respect to the target function by a variation of all other independent variables. Then the parameter that the scanning procedure was applied to was changed by a small increment, usually 1° for rotations and 0.1 Å for translations. Again, the structure was optimized with a new value of a fixed parameter. As a result, we obtain a variation of a structure vis-a-vis a specified conformational parameter. Previously, this procedure was used by us in the studies of torsional and bending rigidities and of interconversions of sugar rings in B DNA (Zhurkin et al., 1982; Ulyanov & Zhurkin, 1984; Gorin et al., 1990). In the present version of program (Ulyanov et al., 1989), the procedure is flexible and fully automated, the sequence of all operations being specified in the input file by the user.

At first, the scanning procedure was applied to the search for the optimal structure of the *Bam*HI hexamer; such systematic variation of structural degrees of freedom helps avoid the "trap" due to an arbitrary choice of starting conformations. Then this procedure was repeated starting from the optimal structure, in order to calculate the variations of R factor, individual NOE's, distances, etc. versus conformational parameters. We performed more than 60 scans of the conformational space; on the average, 10–15 conformations were "conditionally optimized" (i.e., with one of the parameters being held fixed) in each scan. Typically, a few hundreds of structures were tested during the optimization of each conformation. Totally, it took about 40 h of CPU time on Cray XMP in order to fulfill this project, the most time-consuming part being NOESY simulations during the optimization process.

Results and Discussion

Assignment. Assignment of all nonexchangeable base and sugar protons was done by traditional methods using 2D NOESY spectra at mixing times of 50 and 150 ms. The table of assignment and a 2D plot of the portion of spectra with H6/H8–H1′ and H6/H8–H3′ sequential connectivities are given in the Supplementary Material.

Effective Correlation Time. At first we estimated the effective correlation time t_c from NOE's H6–H5 in the nonterminal cytosine C5 with the use of the data at 7 °C with mixing times of 10, 25, and 50 ms. The two-spin approximation was used for this estimation, 6.6 ± 0.8 ns, following the procedure of Clore and Gronenborn (1984). Then we performed several trial NOESY simulations for different conformations of DNA within the A and B families of forms. In order to achieve a better agreement with the experimental H6–H5 NOE's, we had to increase the effective correlation time up to 7.5 ns. Such correlation time is apparently too high for a double-stranded hexamer of DNA. Assuming the spherical shape of a hexamer with a radius $r = 13$ Å (Clore & Gronenborn, 1984), the correlation time is expected to be 4.6 ns at 7 °C. The higher correlation time is indicative of possible intermolecular contacts for the *Bam*HI hexamer. Indeed, we have found several specific end-to-end contacts, for example, the H1′G1–H2′C6 contact, which is clearly seen even at a relatively low mixing time, 50 ms (see Supplementary Material).

The correlation time of 7.5 ns was used in all subsequent NOESY simulations for the *Bam*HI hexamer. The experimental H6–H5 NOE's were reproduced in the nonterminal cytosine in the optimal structure with precisions 0.4%, 0.3%, 1.1%, and 0.7% at mixing times of 10, 25, 50, and 150 ms, respectively. However, it is worth noting that this NOE can vary as much as 0.4% at 50 ms and 2.0% at 150 ms with variation of conformation of the *Bam*HI hexamer, even though the corresponding distance is constant and rather short, 2.45 Å.

Optimal Structure. The optimal structure for the *Bam*HI hexamer is shown in Figure 1, and its parameters are listed

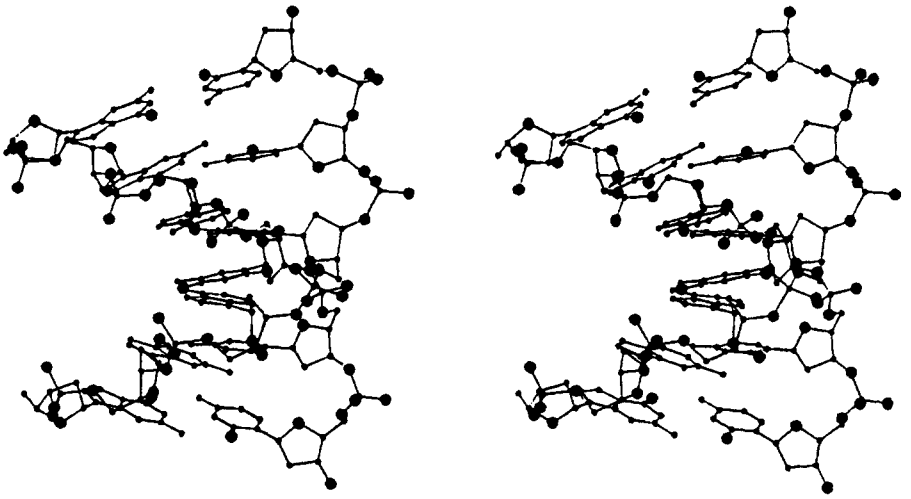

FIGURE 1: Refined structure for the *Bam*HI hexamer in stereo. An overall bend is seen in the direction of the major groove.

Table I: Conformational Parameters of the Optimal Structure[a,b]

	Ω	τ	ρ	D_x	D_y	D_z
GpG	33.3	1.1	8.8	−0.41	−0.04	3.38
GpA	36.3 (20 to 38)	−2.5 (−7 to 10)	11.0 (4 to 23)	−1.24 (−2.1 to −0.4)	0.47 (−0.9 to 0.8)	3.51
ApT	31.1 (20 to 38)	0.0	5.4 (1 to 7)	−0.02	−0.59 (−0.9 to 0.0)	3.29

	PrTw	buckle	opening	S_x	S_y	S_z
G1:C6	1.4	2.8	0.1	0.03	0.02	0.07
G2:C5	−7.0 (−15 to 3)	5.6 (−5 to 14)	−1.8	0.12	0.01	−0.07
A3:T4	−9.3 (−12 to −3)	1.6 (0 to 2)	0.6	0.15	0.00	0.10

	γ C4′–C5′	β C5′–O5′	α O5′–P	ζ P–O3′	ε O3′–C3′
GpG	40.3	200.2	273.7	274.3	195.1
GpA	53.7	179.8	297.1	226.3	181.6
ApT	56.3	178.7	285.6	255.2	183.5
TpC	59.2	167.2	291.8	266.2	195.3
CpC	52.8	176.1	281.2	269.4	190.8

[a] Rotations are in degrees; translations are in angstroms. [b] Step parameters Ω, ρ, τ, D_x, D_y, and D_z are "descriptive" variables as defined under Materials and Methods; pair parameters propeller, buckle, opening, S_x, S_y, and S_z are independent variables. Torsion angles α, β, γ, ε, and ζ were calculated after the chain closures were found. Meanings and signs of step and pair parameters are in agreement with Cambridge convention. Note that pair parameters define the rotation/translation of each complementary base relative to its position in the standard Watson–Crick pair in a symmetric manner, e.g., the full propeller twist in the pair A3:T4 is −18.6°, twice the value shown in the table. The spreads of parameters which are consistent with the experimental data are shown in round brackets; they were determined as values of parameters that increase $R_{50} + R_{150}$ within the limits of 1% relative to the optimal structure. Those spreads are underlined which coincide with the upper or lower limits of our investigation (see details in the text).

in Tables I and II. This structure is a compromise between a low energy and a low *R* factor. The energy of the structure can be significantly improved at the expense of the *R* factor, e.g., during the scanning procedure a structure was found with an energy 10 kcal/mol lower and an *R* factor 22% higher than in the optimal structure. Similarly, we can somewhat improve the *R* factor (not dramatically, by 1%) with a substantial increase in energy.

The printouts of simulated NOESY spectra at mixing times 10, 25, 50, and 150 ms for the optimal structure are given in the Supplementary Material. In the optimal structure, R_{50} = 16.6% and R_{150} = 19.2%. R_{10} and R_{25}, which have not been under optimization, are 28.5 and 23.1%, respectively. Despite the increased values of R_{10} and R_{25}, the general agreement between the simulated and experimental spectra at mixing times 10 and 25 ms is quite satisfactory: the deviations between individual simulated and experimental cross peaks do not exceed 1.5% (see Supplementary Material).

The optimal structure for the *Bam*HI hexamer has an almost perfect dyad symmetry. The descriptive tilt angle τ and shift D_x are very close to zero in the step ApT:ApT (Table I); they are supposed to be exactly zero in symmetric dinucleotides (see Materials and Methods). Equivalent torsions differ by no more than 3° in the two ApT strands. The rest of the molecule (including symmetrically related base pairs and steps) has the exact symmetric conformation by definition

Table II: Range of Pseudorotation Parameters (deg)

residue	P_o[a]	variation[b]	ΔP[c]	sugar pucker	χ_o[d]
G1[e]	203			C3′-exo	115
G2[e]	168			C2′-endo	156
A3	178	175–200	25	2′-endo/C3′-exo	143
T4	127	110–135	25	C1′-exo	128
C5	109	95–125	30	O1′-endo/C1′-exo	124
C6	79	50–90	40	O4′-exo/O1′-endo	118

[a] Pseudorotation parameter in the optimal structure. [b] Determined as a variation within a 1% increment of $(R_{50} + R_{150})$ relative to the optimal structure. [c] Spread of variation. [d] Glycosidic angle in the optimal structure, which is defined as dihedral angle C2′–C1′–N9–C4 for Pur and C2′–C1′–N1–C2 for Pyr. [e] Spread of variation of pseudorotation was not determined.

(see Materials and Methods).

The refined structure belongs to the B family of forms, but the local conformational heterogeneity is very strong. The sugar puckers vary from C3′-exo for G1 to O1′-endo/C4′-exo for C6, pseudorotations of sugars in pyrimidines being systematically less than in purines (Table II). The helical winding angle Ω varies from 31.1° in the ApT:ApT dimer to 36.3° in GpA:TpC (Table I), the values of Ω in all three complementary dinucleotides being in a very good agreement with the predictions based on the analysis of different crystal and solution data (Kabsch et al., 1982). The preference of purine–pyrimidine steps to have a decreased helical twist was also found in our conformational calculations and Monte Carlo simulations of DNA fragments (Zhurkin et al., 1990). All nonterminal base pairs have a pronounced propeller twist (Table I), the full propeller being 14.0° and 18.6° for nonterminal GC and AT pairs, respectively. Another feature of base pairs is the buckle angle. Buckle is *positive* in the *Bam*HI hexamer for all pairs where a purine belongs to the "strand I", i.e., the pairs with purines in the left strand are buckled *down* when they are viewed from the minor groove (Figure 10). By definition, buckle is of opposite sign for symmetrically related base pairs, so that the step ApT:ApT has a "rhombic" structure, in agreement with the results of Monte Carlo simulations (Zhurkin et al., 1990). However, the magnitude of buckle angle is significantly reduced for the junction ApT:ApT as compared to the middle of the oligopurine block (Table I), again in agreement with Monte Carlo simulations (Zhurkin et al., unpublished results).

All complementary dinucleotides of the *Bam*HI hexamer are locally bent into the major groove (see positive roll angles, ρ, in Table I), which gives the structure some resemblance to the A form of DNA. Globally, the structure is also apparently bent into the major groove in the direction of the dyad axis (Figure 1), the maximum bends being situated in symmetrically related GpA:TpC steps (Table I). According to our conformational calculations of the GGAA:TTCC tetramer of DNA, it is easier to bend the GpA:TpC dimer into the major groove as compared to the minor groove (Zhurkin et al., 1988). In contrast, conformational calculations of AATT:AATT revealed some preference of bending the ApT:ApT step into the minor groove (Ulyanov & Zhurkin, 1984), while it is bent into the major groove in the NMR-derived structure of *Bam*HI (Table I). However, the magnitude of bending is the smallest for ApT:ApT among all complementary dinucleotides of the *Bam*HI hexamer (Table I), so in both cases the ApT:ApT dimer reveals qualitatively the same tendency. The rest of this paper is devoted to the question of how uniquely the structure for the *Bam*HI hexamer is defined by the existing experimental data.

Glycosidic Angles and Conformations of Deoxyriboses. We calculated a series of structures for the *Bam*HI hexamer where the phase angle of pseudorotation *P* in the nucleotides A3, T4, C5, and C6 was varied in the range of 20–200° with steps of 10°. For this purpose, the scanning procedure was employed as described under Materials and Methods. Earlier we used this method to study the energetics of C2′-endo–C3′-endo transitions of sugars in DNA fragments d(AAA):d(TTT) and d(GGG):d(CCC). It has been shown that the general framework of B form affords variation of any single deoxyribose in the range from C2′-endo to C3′-endo conformations, although with different ease for purines and pyrimidines; pathways for such transitions have been found (Gorin et al., 1990; Zhurkin et al., 1990).

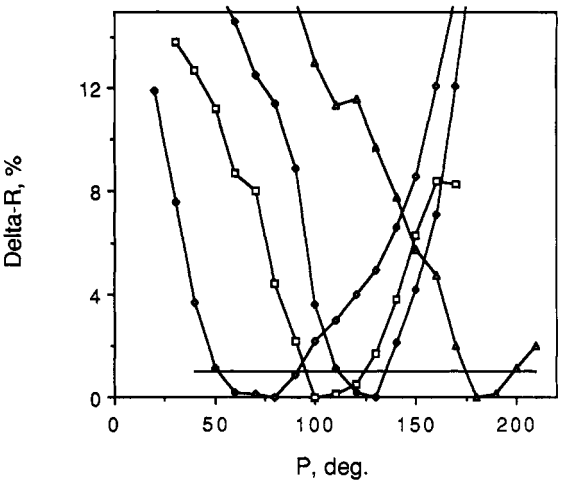

FIGURE 2: Dependences of the increments of the *R* factor, $R_{50} + R_{150}$, on sugar pseudorotations, *P*'s, for the residues A3 (triangles), T4 (filled diamonds), C5 (squares), and C6 (open diamonds). The increments were calculated relative to the *R* factor of the optimal structure. Each point represents a refined structure with a fixed value of pseudorotation in the corresponding residue; a weighted sum of energy and *R* factor was under the minimization with the use of a scanning procedures (see details in the text). The horizontal line shows the 1% level of the *R* factor's increment.

The emphasis of the present study is to investigate the dependences of individual NOE's and interproton distances on conformational parameters of *Bam*HI hexamer. The results for sugar conformations are presented in the following sections.

Estimation of the Range of Pseudorotation Which Is Consistent with the Experimental Data. Below we deal, if not stated otherwise, with the structures optimized with respect to the target function (2). In Figure 2 the increments of the *R* factor ($R_{50} + R_{150}$) are shown as functions of sugar pseudorotation *P* for the residues A3, T4, C5, and C6; the increments are calculated relative to the optimal structure. Every point represents an optimized structure with a fixed value of *P*. Each residue has a well-defined area with a low *R* factor corresponding to a range of pseudorotations which is consistent with the experimental data. Those ranges within a 1% increment of ($R_{50} + R_{150}$) are given in Table II. No attempts were made to estimate similar variations of pseudorotation for G1 and G2 because of the strong spectral overlap for these residues.

It is interesting to note that the variation of sugar pseudorotation *P* for the nonterminal nucleotides, 25–30° (Table II), is very close to the thermal fluctuation of that parameter calculated for complementary trimers of DNA (Gorin et al., 1990). But of course, this coincidence is quite accidental: due to a big weight of *R* factor used in the calculation of pseudoenergy (2), the estimated variation of *P* does not necessarily reflects the conformational mobility of DNA in solution; rather it must be considered as a measure of uncertainty in deriving a structure from the experimental data. As it will be shown in the following sections, the situation is drastically different for some other conformational parameters, where the uncertainty of their estimation exceeds their thermal fluctuation.

Intraresidue NOE's from Base to Sugar Protons. The intraresidue distances H6/H8–H2′/H2″/H3′ are determined completely by sugar and glycosidic conformations of a nucleoside; see, e.g., a comprehensive survey by Wüthrich (1986). It is also well known that the glycosidic angle χ is correlated to a high degree with the sugar pseudorotation in DNA double helices [Figure 3; see also Fratini et al. (1982)]. In such a case the situation is essentially one-dimensional: if χ can be expressed as a function of *P*, then the intraresidue base to sugar

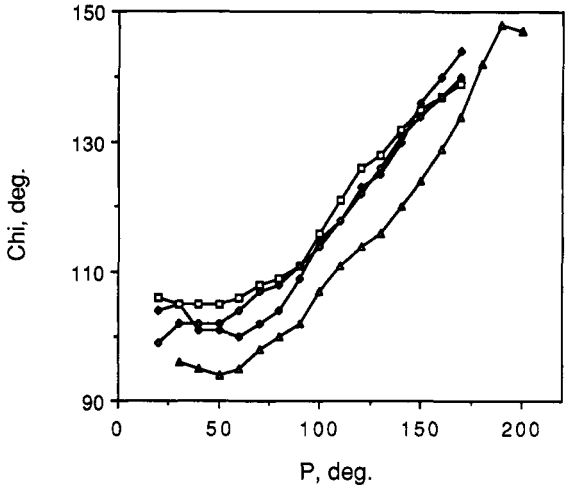

FIGURE 3: Dependences of glycosidic angle χ on pseudorotation P for the individual residues. Data points represent structures obtained similar to those in Figure 2, but with a minimization of energy alone without the R factor; notations for the residues are the same as given in the legend to Figure 2.

distances are the functions of mainly just a single variable, of pseudorotation P. The calculated dependences for the H6/H8–H2′ and H6/H8–H3′ distances and NOE's are shown in Figure 4. The data shown in Figures 3 and 4 were obtained by optimization of energy alone, without a term of R factor. The curves are almost identical for all pyrimidines (T4, C5, and C6). This coincidence is quite natural. Changes in the intraresidue base to sugar interproton distances can be compensated only by the glycosidic angle χ, but the glycosidic angle itself depends on the pseudorotation parameter in the energy-optimized structures (Figure 3). The difference between the single representative of purines (A3) and all pyrimidines is also quite natural; it results from the different geometry of aromatic rings of purines and pyrimidines.

It is well known that C2′-endo conformations of sugars in the B form of DNA are distinguished by a short H6/H8–H2′ intraresidue distance, while this distance is increased for C3′-endo conformations in the A form; conversely, the H6/H8–H3′ distance is short in the A form and increased in the B form (Sarma et al., 1986; Wüthrich, 1986). In Figure 4, we show how these distances change due to a variation of sugar conformation within the B form of DNA. The H6/H8–H2′ distance changes almost linearly with pseudorotation P, while the H6/H8–H3′ distance has a plateau when the pseudorotation is above 100° (Figure 4c,f). Accordingly, there is a very strong NOE between a base proton and H2′ in the C2′-endo region of sugar conformations, which drops dramatically with the decrease in P. The shape of curves for NOE's between H6/H8 and H2″ is very similar to that for H6/H8–H2′ (not shown) due to indirect transfer of magnetization H6/H8 → H2′ → H2″. In contrast, the H6/H8–H3′ NOE is a nonmonotonic function of the pseudorotation P. Similar to the A form, this cross peak is very strong for C3′-endo/C4′-exo conformations of deoxiriboses within the B form and then decreases very fast and almost levels off for O1′-endo sugar puckers. Then, despite the almost constant distance H6/H8–H3′, the NOE steadily increases when P changes toward C2′-endo/C3′-exo conformations. This increase is due to spin diffusion: the distance H6/H8–H2′ becomes shorter at pseudorotations above 100°, and the pathway of magnetization transfer H6/H8 → H2′ → H3′ becomes very effective. This process is very strong for a mixing time of 150 ms, especially for pyrimidines (Figure 4e), but it is still significant even at lower mixing times (Figure 4d).

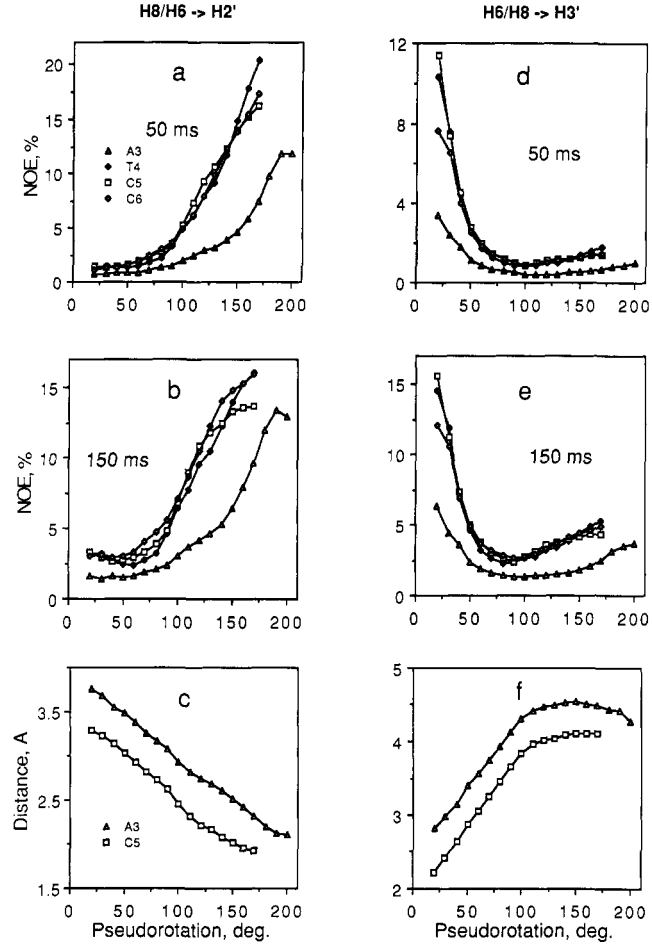

FIGURE 4: Dependences of intraresidue base-to-sugar interproton distances and NOE's on sugar pseudorotation. Data points represent the same structures as in Figure 3; for notations see the legend to Figure 2. (a–c) H6/H8–H2′ NOE's and distances; (d–f) H6/H8-H3′ NOE's and distances; (a and d) NOE's at 50 ms; (b and e) NOE's at 150 ms; (c and f) interproton distances.

Correspondence between Individual Experimentally Measured and Simulated Cross Peaks. Proton resonances in the spectral region of H2′/H2″ are often strongly overlapped. We were able to measure the following cross peaks for the *Bam*HI hexamer: H8A3–(H2′A3 + H2′G2), H8A3–H2″A3, H6T4–H2′T4, H6T4–H2″T4, H6C5–(H2′T4 + H2′C5), H6C5–(H2″T4 + H2″C5), and H6C6–(H2′C6 + H2″C6), and also H8A3–(H3′A3 + H3′G2), H6T4–H3′T4, H6C5–(H3′C5 + H3′T4), and H6C6–H3′C6. The sugar pucker dependences of the corresponding simulated cross peaks are shown in Figure 5 for cross peaks involving intraresidue base proton to H2′ NOE's and in Figure 6 for base to H3′ NOE's. In all cases horizontal lines with 1% error bars show the experimental data; they span the intervals of pseudorotation which are consistent with the experimental data as determined from the 1% increase of the R factor (Figure 2). Curves with open squares represent structures obtained by the minimization of energy together with the R factor; results of the minimization of energy alone are shown for comparison (curves with filled squares). If individual experimental NOE's match those in simulated spectra for the optimal structure, the corresponding curve must intersect the horizontal line. It is clearly seen that there is generally a good agreement between the simulated H6/H8 to H2′ NOE's and the experimental data (Figure 5).

The situation is essentially different for the cross peaks between H6/H8 and H3′ (Figure 6). The simulated NOE's

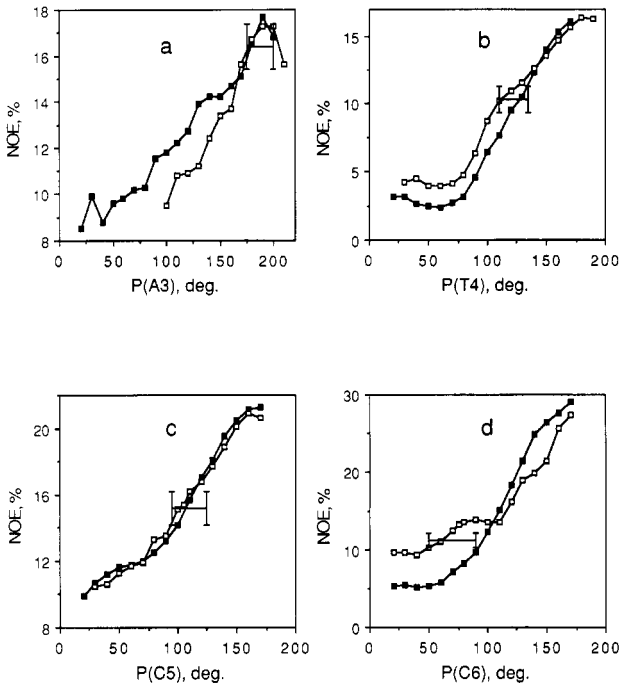

FIGURE 5: Dependences of H6/H8 to H2′ NOE's at a mixing time of 150 ms on sugar pseudorotation of individual residues. Shown are the single NOE's for the resolved cross peaks and the sums of NOE's for the overlapped ones, in correspondence with experimentally measured quantities. The experimental values are shown by horizontal lines which span the uncertainty of definition of pseudorotation P, as determined from Figure 2 (see details in the text). Vertical bars represent 1% errors of measurements. Open squares are for the structures with refined weighted sum of energy and R factor (same as in Figure 2); filled squares stand for the structures where the energy alone was under optimization (as in Figures 3 and 4). (a) Sum of NOE's H8A3–H2′G2 and H8A3–H2′A3 in dependence on sugar pseudorotation of A3, P(A3); (b) NOE H6T4–H2′T4 in dependence on P(T4); (c) sum of NOE's H6C5–H2′T4 and H6C5–H2′C5 in dependence on P(C5); (d) sum of NOE's H6C6–H2′C6 and H6C6–H2″C6 in dependence on P(C6).

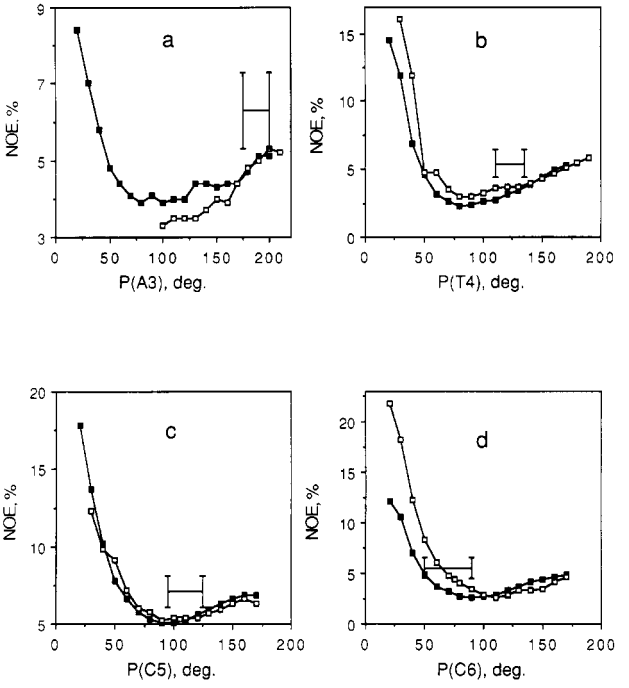

FIGURE 6: Dependences of NOE's from H6/H8 to H3′ at a mixing time of 150 ms on sugar pseudorotations in individual residues; see the legend to Figure 5. (a) Sum of NOE's H8A3–H3′A3 and H8A3–H3′G2 in dependence on sugar pseudorotation of A3, P(A3); (b) NOE H6T4–H3′T4 in dependence on P(T4); (c) sum of NOE's H6C5–H3′C5 and H6C5–H3′T4 in dependence on P(C5); (d) NOE H6C6–H3′C6 in dependence on P(C6).

are *systematically* less than experimental values, especially at higher mixing time, despite taking into account the process of spin diffusion. Although the deviations are not very big (they vary from 1.2% to 1.7% at a mixing time of 150 ms), they are presumably more than the experimental error. The reasons for such a systematic discrepancy are not quite clear. One possible explanation involves the *internal motion of deoxyriboses*. The problem of internal mobility of sugar rings in DNA has a history. Raman spectroscopy (Wartell & Harrel, 1986), analysis of scalar couplings (Rinkel et al., 1987; Zhou et al., 1988; Celda et al., 1989; Lane & Forster, 1989), and conformational calculations (Gorin et al., 1990; Zhurkin et al., 1990) have demonstrated that the repuckering of deoxyriboses easily takes place in the B form of DNA. Such repuckering can explain the higher values of observed intraresidue NOE's between H6/H8 and H3′. Even a small fraction of N-conformers of deoxyriboses would effectively increase the base proton to H3′ NOE's, since these NOE's grow rapidly with decreasing pseudorotation below $P = 50$–$60°$ (Figure 6). Unfortunately, the correct quantitative averaging of NOE's for repuckering sugars requires a significant number of unknown parameters [see discussion of that subject by Borgias and James (1988)]. Nevertheless, an attempt of such analysis has been made recently by Lane (1990a,b), who interpreted the NOE data in terms of an equilibrium between C2′-endo and C3′-endo conformers, assuming that the sugar repuckering is slow on the Larmor time scale. Populations of sugar conformers emerging from such an analysis are comparable with those obtained by analysis of coupling constants for the DNA hexamer CGTACG (Lane, 1990b; Lane & Forster, 1989). However, one must be cautious with results of such an analysis, as it requires several approximations. There is no doubt that a consideration of S–N equilibrium would improve the agreement between theoretical and experimental intraresidue H6/H8–H3′ NOE's, but the resulting populations of conformers must be very sensitive to the choice of equilibrium point for N-conformers, because of the strong dependence of these NOE's on pseudorotation in the range of $P < 50$–$60°$ (Figures 4 and 6). Usually, the geometry of N-conformers is assumed to be constant when fitting the data; P_N was 18° during the analysis of NOE data by Lane (1990b). Such values of P_N may not necessarily be correct in the case of B DNA, especially for pyrimidines [for a discussion of this subject, see Gorin et al. (1990)]. Anyway, in the case of the *Bam*HI hexamer, the discrepancies between the experimental H6/H8–H3′ NOE's and those calculated in the single-conformation approximation, although being of a systematic character, are on the margin of the experimental error. This effect is probably too small to make any definite conclusion.

Helical Parameters. With the use of the "scanning procedure" (see Materials and Methods), we have systematically varied "step" (or "wedge") parameters in complementary dinucleotides GpA:TpC and ApT:ApT and also propeller twist and buckle in AT and nonterminal GC pairs. The uncertainties in the magnitudes of these parameters, calculated similar to those of sugar pseudorotations, are listed in Table I. The majority of these parameters are defined much less precisely by the experimental data than sugar puckers. The reason for that is the correlated changes in different helical parameters, which compensate the effect of any particular parameter on interproton distances. Below we show in several examples the mechanisms of these compensational correlated changes.

Table III: Variation of Representative Distances and NOE's with Helical Winding Angle in Step GpA:TpC, Ω_{GA}[a]

	distance (Å)		NOE (%)					
			τ_m = 50 ms			τ_m = 150 ms		
proton pair	Ω_{GA} = 41°	Ω_{GA} = 20°	Ω_{GA} = 41°	Ω_{GA} = 20°	exp.	Ω_{GA} = 41°	Ω_{GA} = 20°	exp.
H8A3–H8G2	5.52	4.17	0.0	0.5	0.0	1.3	1.9	1.2
H6T4–H6C5	4.91	4.11	0.4	0.6	0.6	1.9[b]	1.7	1.8
H8A3–H1′G2[c]	3.28	2.73	1.7	3.9	3.9	4.2	6.9	7.7
H6C5–H1′T4	4.36	4.15	0.7	0.6	1.1	2.8	1.9	3.5
H1′C5–H2A3[d]	4.33	5.13	0.3	0.1	0.6	0.7	0.3	0.6
H1′T4–H5′C5	5.30	3.74	0.2	1.0		0.8	3.3	
H1′T4–H5″C5	4.61	2.86	0.6	2.7		1.1	5.0	
H1′T4–H5′/5″C5[e]			3.1	5.9	3.5	9.3	15.8	15.8

[a] Simulated data are given in the columns entitled "Ω_{GA} = 41°" and "Ω_{GA} = 20°", and experimental data are in the columns "exp". [b] Maximum of NOE is 2.1% at Ω_{GA} = 31°. [c] Minimum of distance (2.71 Å) and maximum of NOE (4.2% at 50 ms and 7.7% at 150 ms) are at Ω_{GA} = 27°. [d] Interresidue interstrand contact across the minor groove. [e] Strongly overlapped cross peaks between H1′T4 and H4′T4 + H4′/5′/5″C5 at a mixing time of 50 ms and H4′/5′/5″T4 + H4′/5′/5″C5 at 150 ms.

Helical Winding Angles. Helical winding angles in the steps ApT:ApT and GpA:TpC (Ω_{AT} and Ω_{GA}) have been systematically varied by the scanning procedure in increments of 1° in the range 20–41°. It was found that helical winding angles are among the parameters that are determined very poorly by the NOESY spectra. The optimal structure of the *Bam*HI hexamer (see the corresponding section and Table I) has Ω_{AT} = 31.1°, Ω_{GA} = 36.6°, and *R* factor $R_{50} + R_{150}$ = 35.8%. In the range of Ω_{GA} from 20° to 38° the *R* value does not exceed 36.7%, so we consider all these structures as consistent with the experimental data. Essentially the same range was found for the parameter Ω_{AT} also. Usually, R_{50} slightly increases and R_{150} decreases during unwinding of steps ApT:ApT and GpA:TpC (data not shown), so that their sum is almost constant. In fact, structures of the *Bam*HI hexamer with a strongly underwound GpA:TpC step have even a smaller value of *R* factor than the optimal structure: $R_{50} + R_{150}$ = 34.8% for conformations with Ω_{GA} = 23 and 24°. However, the energy of these structures is about 20 kcal/mol higher than the energy of the optimal structure (Figure 7). 75% of this energy increase is due to the change in the hydrogen-bonded geometry of pairs and unfavorable interactions in sugar–phosphate chains in the underwound conformations.

The lack of sensitivity of the *R* factor to the helical winding angle indicates that all short-range distances vary insignificantly with winding/unwinding of DNA. The distances that change most of all with the variation of Ω_{GA} are listed in Table III. With some exceptions, all of them are of a medium- to long-range type. Though there are changes as big as 0.8–1.3 Å (notably, in the intrastrand H8A3–H8G2, H6T4–H6C5, and interstrand H1′C5–H2A3 contacts), the corresponding changes in NOE's are very small in that range of distances. It is interesting that despite the significant decrease in the H6T4–H6C5 distance, there is no corresponding increase in NOE at 150 ms between these two protons (Table III). Obviously, it results from a less effective spin diffusion in the underwound structures.

The variation of distances is even smaller during the unwinding of the ApT:ApT dimer. The distance between base protons H8A3 and H6T4 decreases from 4.8 to 4.6 Å, and the H6T4–H1′A3 distance varies in the range 3.3–3.4 Å when Ω_{AT} changes from 41° to 20°. The interresidue intrastrand distance H1′T4–H2A3 is very sensitive to Ω_{AT} (it increases from 4.0 to 5.2 Å when the step unwinds), but it is difficult to quantify this distance experimentally: the corresponding NOE changes from 0.5% to 0.1% at a mixing time of 50 ms and from 1.2% to 0.2% at 150 ms. If one believes absolutely in the experimentally measured NOE's (0.5% and 0.8% at mixing times of 50 and 150 ms, correspondingly), one is able to exclude the underwound structures. But if one uses the *R* factor as an agreement criterion, these structures are still consistent with the experimental data.

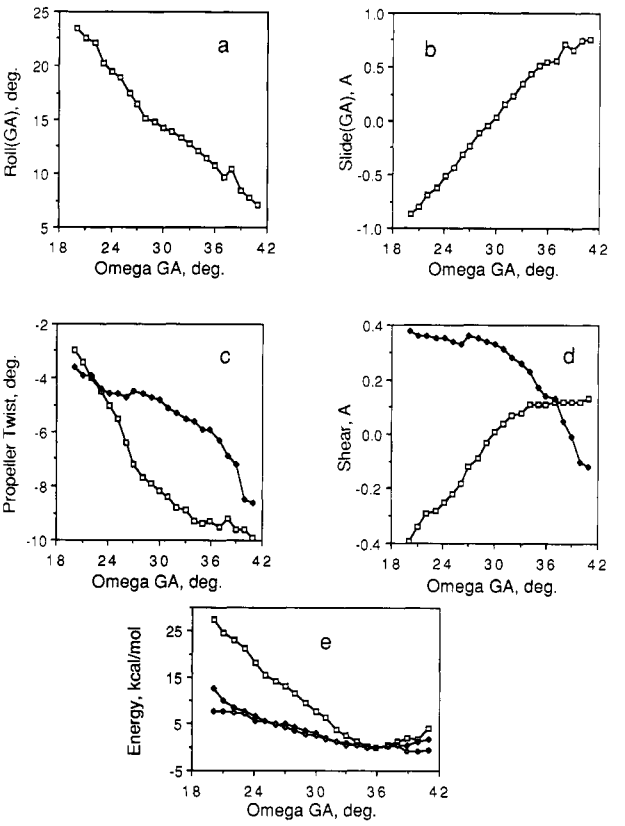

FIGURE 7: Selected conformational parameters versus helical twist of the GpA:TpC dimer, Ω_{GA}. Data points represent structures optimized with respect to a weighted sum of energy and *R* factor. (a) Roll angle in GpA:TpC dimer, ρ_{GA}; (b) slide $D_{y_{GA}}$ in the same dimer; (c) propeller twist ω in A3:T4 pair (open symbols) and in G2:C5 pair (filled symbols); (d) shear S_x in A3:T4 pair (open symbols) and in G2:C5 pair (filled symbols); (e) increment of energy relative to the optimal structure. The increments are shown for the total energy (open squares), energy of hydrogen-bonded interactions in pairs (filled diamonds), and energy of interactions in the sugar–phosphate backbone (open diamonds).

It is the inter- and intraresidue distances from a base proton to H2′/H2″ that mainly determine the value of the *R* factor, and all these distances and the corresponding NOE's are almost constant during unwinding of steps GpA:TpC and ApT:ApT from 40° to 20° (data not shown). This fact deserves a special consideration because, apparently, the interresidue intrastrand contacts must become shorter upon unwinding of DNA. To compensate for this decrease, the DNA double helix can employ several strategies. First, since the glycosidic bonds are directed toward the minor groove, a ro-

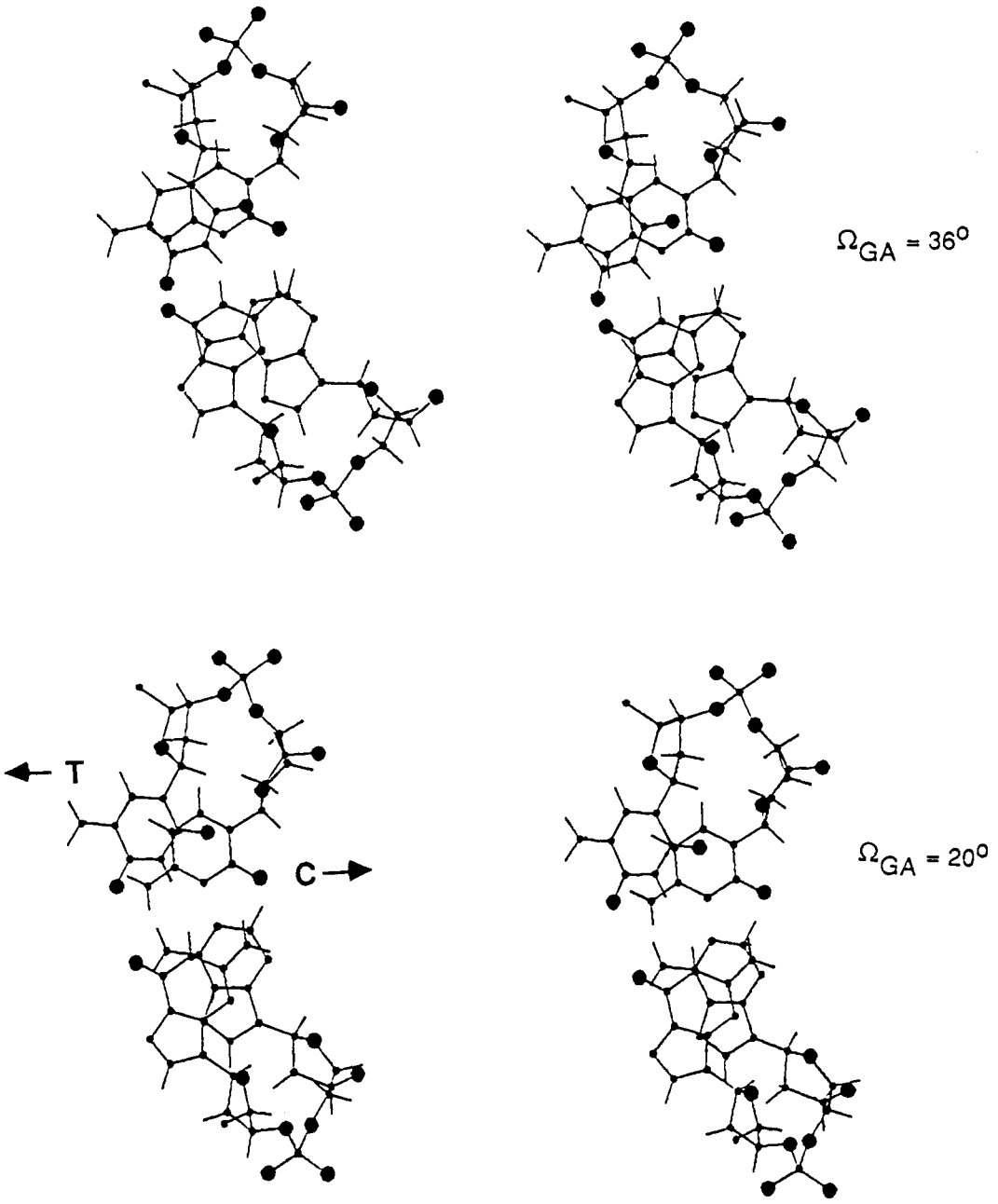

FIGURE 8: Stereoviews of dinucleotide GpA:TpC in the *Bam*HI hexamer optimized with fixed values of helical twist Ω_{GA}: 36° (top) and 20° (bottom). Note significant distortions in the geometries of both AT and GC pairs at Ω_{GA} = 20°: the cytosine is shifted into the minor groove relative to the guanine, while the thymine is shifted into the major groove relative to the adenine (cf. Figure 7d). Both structures are consistent with the experimental data.

tation of base pair along its long axis (in such a way that the dimer bends into the major groove) would increase the interresidue base-to-sugar interproton distances. Indeed, ρ_{GA} attains 23° when the GpA:TpC step unwinds down to 20° (Figure 7a). Second, a decrease in slide $D_{y_{GA}}$ from 0.8 to −0.9 Å (Figure 7b) also increases these interproton distances, moving the adenine A3 away from the deoxyribose of G2 and the deoxiribose of T4 away from the base C5. It is interesting that similar correlations between these helical parameters have been found as a result of energetical conformational calculations and Monte Carlo simulations of DNA fragments without experimental NMR restrictions (Ulyanov & Zhurkin, 1984; Zhurkin et al., unpublished results). Correlations between Ω and ρ, and Ω and D_y are among the strongest ones in the DNA double helix (Zhurkin et al., unpublished results).

Among other helical parameters let us mention propeller twist ω and shear S_x in both AT and GC pairs (Figure 7c,d). The diminution in absolute values of propeller twist in pairs A3:T4 and G2:C5 helps relieve steric clashes between bases in the major groove caused by bending of DNA (see Figure 8). Abrupt changes in parameters S_x (Figure 7d) are directed to move apart the neighboring residues in such a way that interresidue base-to-sugar interproton distances would increase and thus compensate their decrease due to unwinding of DNA. Namely, A3 and C5 are moving into the minor groove, while T4 and G2 are moving into the major groove with unwinding of step GpA:TpC (compare structures with Ω_{GA} of 20° and 36° in Figure 8). Naturally, such changes in the geometry of base pairs in underwound dimers are unfavorable in terms of energy (Figure 7e), but the structures are still consistent with the experimental data.

In the case of the ApT:ApT dimer, the compensational changes are similar to those in GpA:TpC, but their spreads of variation are different. Roll angle ρ_{AT}, shift $D_{y_{AT}}$, and

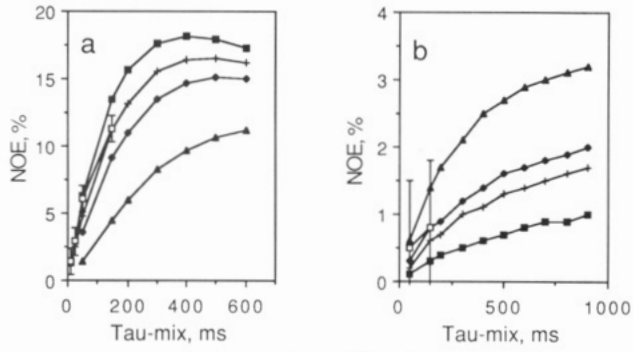

FIGURE 9: Dependence of NOE's on mixing time for several structures optimized with fixed values of roll angle ρ_{AT} in ApT:ApT step: 10° (filled squares), 4° (pluses), 0° (filled diamonds), and −10° (triangles). Experimental data are shown by open squares with 1% vertical error bars. (a) H8A3–MetT4; (d) H1′T4–H2A3.

propeller twist ω_{AT} vary in a lesser extent: ρ changes from 3° to 12°, D_y from −0.6 to −0.9 Å, and ω from −11° to −6° while Ω_{AT} decreases from 41° to 20°. At the same time, shear $S_{x_{AT}}$ varies to a greater degree: from −0.3 to +0.5 Å.

Roll Angles. Roll angles are parameters which are strongly correlated with the helical winding angle. The dependence of roll ρ_{GA} on Ω_{GA} is shown in Figure 7a for the step GpA:TpC. The range of Ω_{GA} 20–38° is consistent with the experimental data; the corresponding range for ρ_{GA} is from 23° to 10°. The compensatory conformational changes, which help keep the simulated NOE's constant during variation of Ω, were considered in the section on helical winding angle; these considerations are relevant to the case of variation of ρ as well, due to a strong correlation between Ω and ρ. The application of the scanning procedure specifically to the roll angle ρ_{GA} showed that structures that agree with the experimental data can be generated with ρ_{GA} as low as 4°. However, when ρ_{GA} was decreased further into the negative region (i.e., when the GpA:TpC step was bent into the minor groove), it caused a small but systematic increase of the R factor; $R_{50} + R_{150}$ = 38.8% at $\rho_{GA} = -8°$, which is 2% higher than for the optimal structure. Variations are not very high among interproton distances and NOE's, thus explaining the small increase of the R factor. The exceptions are the interresidue intrastrand distance H6T4–H5C5 and the interstrand contact H1′C5–H2A3. The first distance changes from 3.70 to 4.66 Å and the second one from 4.99 to 3.97 Å when ρ_{GA} decreases from 11° (in the optimal structure) to −8°. Their significant variation is mainly due to a change of the shift parameter $D_{x_{GA}}$ from −1.2 Å in the optimal form to −0.6 Å in the structure with $\rho_{GA} = -8°$. This movement shifts the AT pair into the major groove relative to the neighboring GC pair. Besides, the roll angle itself influences both distances. However, the corresponding changes of NOE's do not exceed 1% even at 150 ms.

The dinucleotide ApT:ApT has one additional contact compared to GpA:TpC, namely, H8A3–MetT4, which proved to be very sensitive to the roll angle ρ_{AT} (Figure 9a). The methyl group is directed into the major groove of DNA, so an increase in the roll angle (i.e., bending into the major groove) diminishes the distance between CH_3 and the 5′-neighboring H6/H8 proton (in our case H8A3). Conversely, bending into the minor groove makes this distance longer. It explains why after applying the scanning procedure to the parameter ρ_{AT} we found a relatively narrow range of those values which are in agreement with the experimental data (Table I). But as it was before with other helical parameters, the R factor increases very slowly with deviation of ρ_{AT} from its optimal value. Structures with $\rho_{AT} = -6°$ or +12° both have an R factor which is only 2.5% higher than that for the optimal structure. The reason for this is that all other contacts are not very sensitive to ρ_{AT}, or, more precisely, they can be kept constant by compensational conformational adjustments (see, e.g., Figure 9b).

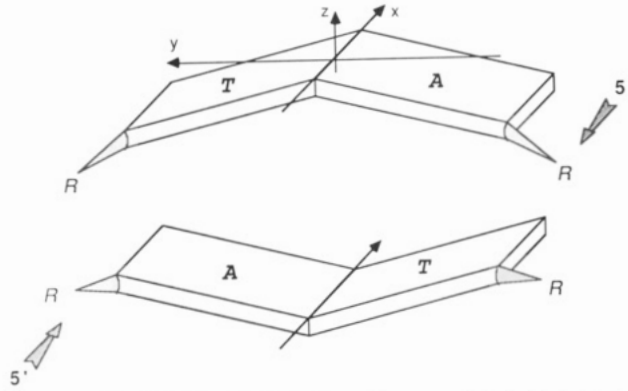

FIGURE 10: Schematic representation of a symmetric ApT:ApT step with buckled pairs, a view from the minor groove. The lower AT pair has a positive buckle angle; the upper one has identical internal geometry, but the buckle angle is negative due to a dyad symmetry of the step. Buckle angles are exaggerated. The frame of reference is shown which is bound to the upper base pair.

Parameters of Base Pairs. The scanning procedure was used to study the variation of propeller twist ω and buckle κ in the base pairs A3:T4 and G2:C5. Propeller twists ω_{AT} and ω_{GC} are among the parameters poorly determined by the experimental data (Table I). There are no measurable cross peaks in NOESY spectra in D_2O which directly depend on relative positions of bases in a pair. The variation of propeller twist requires a certain adjustment of double-helix conformation (data not shown), which could involve changes in some interproton distances. However, the adjustment occurs in such a way that all simulated NOE's, which correspond to experimentally measured cross peaks, change insignificantly. The same is true for buckle κ_{GC} in the pair G2:C7; for every tested value of κ_{GC}, we found a structure which is consistent with the experimental data (Table I).

But the situation is different for the buckle κ_{AT} in the pair A3:T4; only structures with κ_{AT} = 0–2° agree with the NOESY data. Beyond that range the R factor increases, attaining 38.9% at $\kappa_{AT} = -2°$ and 40.2% at $\kappa_{AT} = +9°$. This increase is due to variation of several interproton distances; the distance between H8A3 and the methyl group of T4 is one of the most variable, the corresponding simulated cross peak being weak at negative values of buckle and strong at high positive values. Also, there are some minor variations in the distances H8A3–H2′/2″A3, H8A3–H1′G2, H6T4–H2′/2″T4, and H1′A3–H5′/5″T4, which also account for the increase of R factor with deviation of κ_{AT} from its optimal value.

To explain this effect, we have to consider the nature of buckling motion in the AT pairs of the *Bam*HI hexamer. The ApT:ApT step was assumed to be symmetric (see Materials and Methods); therefore, the buckling motion occurs symmetrically in the two neighboring AT pairs. A positive buckle in the first (lower) pair A:T is coupled with a negative buckle in the next pair T:A, so the ApT:ApT structure looks like a rhombus (Figure 10). In the terminology of Dickerson and co-workers, it corresponds to a positive "cup" [see Figure 5 in Yanagi et al. (1991)]. In contrast, a negative κ_{AT} (buckle in the lower pair A:T) would lead to the "butterfly-like" structure (Zhurkin et al., 1990) or to a negative cup (Yanagi et al., 1991). If not compensated by other structural adjustments, the distance between the neighboring sugars in each

backbone is dramatically reduced in the "rhombic-like" structures (Figure 10), which prevents the correct closure of sugar–phosphate chains. One simple adjustment of the structure is increasing the rise parameter D_z, which helps relieve this compression of a backbone; another one is increasing the bend into the major groove, which pushes the neighboring sugars apart and thus also compensates the compression. Indeed, $D_{z_{AT}}$ is 4.46 Å and ρ_{AT} attains 7° in the ApT:ApT step of the optimized structure with $\kappa_{AT} = 9°$. In contrast, the rise $D_{z_{AT}}$ equals 2.93 Å and ρ_{AT} vanishes almost to zero in the butterfly-like structure with $\kappa_{AT} = -2°$. Here, decreasing $D_{z_{AT}}$ and ρ_{AT} restores the otherwise enlarged distance between neighboring deoxyriboses along the backbone. Of course, all these conformational motions affect the distance H8A3–MetT4. Since both atom H8 and the methyl group belong to the periphery of the double helix, the distance between them should be dramatically reduced in rhombic-like structures, in parallel with the reduction of the distance between sugars along the chain. A compensational increase of the $D_{z_{AT}}$ makes both distances longer. In contrast, the rolling motion affects these two distances differentially, because sugars are located in the minor groove and the methyl group is in the major groove of B DNA. A compensational bending into the major groove makes the distance H8A3–MetT4 shorter in rhombic-like dimers. As a net result of all these motions, the H8–CH_3 distance is shortened at high κ_{AT} (in rhombic steps) as compared to the optimal structure. At negative κ_{AT}, all the movements are of opposite direction, and the distance is enlarged.

Similar correlations exist also for the buckling motion of G2:C5 pair. But this pair does not possess a methyl group protruding from the plane of bases; the atom H5 of cytosine C5 is situated not too close to H6 of the neighboring T4, and the NOE H5C5–H6T4 is less sensitive to the variation of D_z in the step GpA:TpC. Besides, the pair G2:C5 does not belong to a symmetric step, and the buckling of this pair is relatively localized, not creating such big perturbations in the geometries of GpA:TpC and GpG:CpC steps as in the case of buckling of AT pairs in the symmetric ApT:ApT step.

Recently, Dickerson and co-workers performed a thorough correlational analysis of helical parameters in crystal structures of B DNA's solved up to date (Yanagi et al., 1991). They have found an exactly opposite correlation between buckle and rise: the rise D_z decreases when the cup is positive (i.e., in rhombic steps in our terminology). This apparent contradiction is a trivial consequence of different definitions of a rise parameter. D_z is being calculated from separations of deoxyriboses' atoms C1′ in Yanagi et al. (1991), whereas we determine D_z as a vertical shift between frames of reference which are attached to base pairs, not to sugars (see Materials and Methods).

Conclusions

The calculated uncertainties (see Table I for the helical parameters and Table II for sugar conformations) show that many of the helical parameters are determined by the experimental data strikingly poorly, especially those of step GpA:TpC and of the nonterminal GC pair. One of the least defined parameters of the *Bam*HI hexamer is the helical winding angle in both GpA:TpC and ApT:ApT steps. The only thing we can say about helical twist on the basis of the experimental data is that the double helix is not strongly overwound. But all values of Ω down to 20° (and probably even less) are consistent with the NOESY data. The probable reason for this result is that winding/unwinding of B DNA occurs in the planes of base pairs, and it can be relatively easily compensated by translations D_x, D_y, and S_x, also taking place in the planes of pairs (see Figure 7). The situation is getting more complicated here because we practically do not have reliable direct experimental restrictions on the parameters of base pairs, such as shear S_x. Although a strong variation of S_x is energetically unfavorable, it compensates the helical unwinding and keeps the interproton distances more or less constant (Figure 8). The above considerations could explain why some NMR-derived structures of DNA oligonucleotides in solution are so strongly underwound when the NOE information was the main driving force of refinement. For instance, the nonterminal step CpG has a helical twist as low as 6° in the NMR-derived structure of the dodecamer CGCGAATTCGCG (Nerdal et al., 1989). This structure is consistent with the experimental data (Nerdal et al., 1989); however, such a low winding angle contradicts the estimations of DNA twisting from conformational calculations (Zhurkin et al., 1978, 1982), and also it contradicts the following experimental data: the average helical twist of DNA in *solution* is 34–36° depending on the sequence, as measured by a band-shift method (Wang, 1979; Peck & Wang, 1981; Strauss et al., 1981). The variation of helical twist among individual dinucleotides is 28–40°, as derived from the analysis of solution data and crystal structures (Kabsch et al., 1982).

It is useful to compare the calculated uncertainties of helical parameters with their thermal fluctuations. The thermal fluctuation of helical winding angle was estimated from the experiments with cyclization of DNA (Frank-Kamenetskii et al., 1985; Levene & Crothers, 1986); also its fluctuation as well as fluctuations of other parameters were calculated in Monte Carlo simulation of DNA fragments with different sequences (Zhurkin et al., 1990, 1991; Zhurkin et al., unpublished results). Pseudorotations of deoxyriboses in the nonterminal residues of the *Bam*HI hexamer are defined quite well by the NMR data; half of their uncertainties, about 15° (Table II), is very close to their thermal fluctuations, as estimated by energy calculations. All wedge parameters of the GpA:TpC step (with the possible exception of rise D_z which has not been studied by the scanning procedure) are defined poorly from the NMR data; half of their uncertainties exceed their thermal fluctuation. The same holds true for propeller twist and buckle angle of the nonterminal GC pair and also for helical twist of the ApT:ApT dimer. Roll angle ρ_{AT}, slide $D_{y_{AT}}$, and the propeller twist of the AT pair are defined better; half of their uncertainties are comparable with the corresponding thermal fluctuations. Buckle angle κ_{AT} of the AT pair is defined much more precisely. Tilt angle τ and shift D_x in ApT:ApT were always close to zero, because this dimer has a dyad symmetry (see Materials and Methods).

When comparing our results with those obtained by Metzler et al. (1990), one has to take into account the differences in methods used by our groups. Metzler et al. determined the ranges of variation of conformational parameters in a set of seven refined structures for the DNA decamers CGCCTAATCG and CGTCACGCGC, the initial conformations being randomly generated by a distance geometry algorithm. We calculated the uncertainties of determination of parameters from the increments of the agreement R factor versus systematically varied parameters. Although our methodology enables us to generate structures with arbitrary values of conformational parameters (provided the sugar–phosphate chains can be closed), we restricted ourselves to the forms with a reasonable energy, the ranges of tested parameters being slightly more than their thermal fluctuations. In this situation, one cannot expect a quantitative agreement

between our estimates of the uncertainties in the determination of parameters and the ranges of their variation calculated by Metzler et al. (1990). Nevertheless, the main conclusions in this paper and that by Metzler et al. are similar: the conformations of deoxyriboses (and, therefore, the glycosidic angles also) are defined quite well, but the helical parameters are defined considerably worse by the NOESY data.

These results are in an apparent contradiction with the conclusions of Lane (1990a), who studied the dependences of NOE's on sugar conformations and helical parameters of DNA. He found that the intraresidue H6/H8–H2′/H2″ NOE's depend strongly on the glycosidic angle χ, whereas these NOE's change insignificantly with variation of the pseudorotation phase angle P. Those results were obtained by a variation of a single parameter (χ or P, correspondingly) while the rest of parameters were held fixed (Lane, 1990a). But in the double-helical DNA, P and χ cannot be varied independently, they are highly correlated (Fratini et al., 1982; Figure 3 of the present work). In order to obtain the correct dependences of NOE's on conformational parameters, one has to take into account all structural correlations which exist in the double helix. The above is relevant not only to conformations of residues but to helical conformations as well. When a helical parameter is varied and all other parameters are held fixed, it results in a strong dependence of internucleotide NOE H6/H8–H2″ [Figure 5 of Lane (1990a)]. However, if similar dependences are calculated by the scanning procedure, which optimizes the structure with respect to all other degrees of freedom, then the change in H6/H8–H2″ distance proves to be compensated by conformational adjustments, and the corresponding NOE is almost constant in most cases (see Results).

One may expect that future progress in NOESY methodology as well as consideration of additional NMR-derived information, like spin couplings for the sugar–phosphate backbone, will somewhat improve the degree of definition of DNA structure. However, as we showed in the course of the present investigation, there exists a great variety of helical conformations with very similar sets of short- and middle-range interproton distances. In this situation, one needs some additional source of information to study such effects as dependence of DNA structure on its sequence. From our point of view, such information can be obtained from the consideration of conformational energy of DNA. In our approach, theoretical energy calculations are conjugated with the algorithm in which the parameters of interest serve as independent variables for the generation of DNA structures. This makes it possible to investigate more effectively the essential part of the conformational space, taking advantage of scanning the energy of DNA along chosen coordinates. This method was proven to be capable of reproducing such characteristics as torsional and bending rigidities of B DNA [for a review, see Zhurkin et al. (1990)]. However, the well-known drawback of the method of conformational calculations is its inability to account properly for the solvent effects that could influence significantly the structure of small DNA fragments. We consider the combination of NOESY spectroscopy and conformational calculations as an approach that is capable of overcoming the drawbacks of each of these two methods.

Acknowledgments

We thank Drs. Goutam Gupta and Kimiko Umemoto for helpful discussions.

Supplementary Material Available

A formal definition of "descriptive" ("passive") set of helical parameters, a table of assignment, a table with experimentally measured and simulated for the optimal structure NOE's, a 2D plot of the H6/H8 vs H3′/H5/H1′ region of NOESY spectrum at 150 ms, and a 2D plot of the H1′ vs H2′/H2″ region of NOESY spectrum at 50 ms (16 pages). Ordering information is given on any current masthead page.

References

Arnott, S., Dover, S. D., & Wonacott, A. J. (1969) *Acta Crystallogr. B25*, 2192–2206.
Baleja, J. D., Pon, R. T., & Sykes, B. D. (1990) *Biochemistry 29*, 4828–4839.
Banks, K. M., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 6996–7010.
Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493–512.
Broido, M. S., James, T. L., Zon, G., & Keepers, J. W. (1985) *Eur. J. Biochem. 150*, 117–128.
Celda, B., Widmer, H., Chazin, W. J., Denny, W. A., & Wüthrich, K. (1989) *Biochemistry 28*, 1462–1471.
Clore, G. M., & Gronenborn, A. M. (1984) *FEBS Lett. 172*, 219–225.
Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O., Lavery, R., Nelson, H. C. M., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., von Kitzing, E., Wang, A. H.-J., & Zhurkin, V. B. (1989) *J. Biomol. Struct. Dyn. 6*, 627–634.
Frank-Kamenetskii, M. D., Lukashin, A. V., Anshelevich, V. V., & Vologodskii, A. V. (1985) *J. Biomol. Struct. Dyn. 2*, 1005–1012.
Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686–14707.
Gorin, A. A., Ulyanov, N. B., & Zhurkin, V. B. (1990) *Molek. Biol. (Engl. Trans.) 24*, 1036–1047.
Gupta, G., Sarma, M. H., & Sarma, R. H. (1988) *Biochemistry 27*, 7909–7919.
Hagerman, J. P. (1990) *Annu. Rev. Biochem. 59*, 755–781.
Kabsch, W., Sander, C., & Trifonov, E. N. (1982) *Nucleic Acids Res. 10*, 1097–1104.
Katahira, M., Sugeta, H., Kyogoku, Y., & Fujii, S. (1990) *Biochemistry 29*, 7214–7222.
Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–426.
Koning, T. M. G., Boelens, R., & Kaptein, R. (1990) *J. Magn. Reson. 90*, 111–123.
Lane, A. N. (1990a) *Biochim. Biophys. Acta 1049*, 189–204.
Lane, A. N. (1990b) *Biochim. Biophys. Acta 1049*, 205–212.
Lane, A. N., & Forster, M. J. (1989) *Eur. Biophys. J. 17*, 221–232.
Lefevre, J.-F., Lane, A. N., & Jardetsky, O. (1987) *Biochemistry 26*, 5076–5090.
Levene, S. D., & Crothers, D. M. (1983) *J. Biomol. Struct. Dyn. 1*, 429–435.
Levene, S. D., & Crothers, D. M. (1986) *J. Mol. Biol. 189*, 73–83.
Metzler, W. J., Wang, C., Kitchen, D. B., Levy, R. M., & Pardi, A. (1990) *J. Mol. Biol. 214*, 711–736.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.
Olejniczak, E. T., Gampe, R. T., & Fesik, S. W. (1986) *J. Magn. Reson. 67*, 28–41.
Patel, D. J., Shapiro, L., & Hare, D. (1987) *Q. Rev. Biophys. 20*, 35–112.
Peck, L. J., & Wang, J. C. (1981) *Nature 292*, 375–378.
Poltev, V. I., & Shulyupina, N. V. (1986) *J. Biomol. Struct. Dyn. 3*, 739–765.
Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 1–34.

Reid, B. R., Banks, K., Flynn, P., & Nerdal, W. (1989) *Biochemistry 28*, 10001-10007.
Rinkel, L. J., van der Marel, G. A., van Boom, J. H., & Altona, C. (1987) *Eur. J. Biochem. 163*, 275-286.
Sarma, M. H., Dhingra, M. M., Gupta, G., & Sarma, R. H. (1985) *Biochem. Biophys. Res. Commun. 131*, 269-276.
Sarma, R. H., Sarma, M. H., & Gupta, G. (1986) *Proceedings of the 4th Conversation in the Discipline of Biomolecular Stereodynamics* (Sarma, R. H., & Sarma, M. H., Eds.) pp 157-177, Adenine, New York.
Sarma, M. H., Gupta, G., & Sarma, R. H. (1988) *Biochemistry 27*, 3423-3432.
Sarma, M. H., Gupta, G., Garcia, A. E., Umemoto, K., & Sarma, R. H. (1990) *Biochemistry 29*, 4723-4734.
Schmitz, U., Pearlman, D. A., & James, T. L. (1991) *J. Mol. Biol. 221*, 271-292.
Srinivasan, A. R., Torres, R., Clark, W., & Olson, W. K. (1987) *J. Biomol. Struct. Dyn. 5*, 459-496.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.
Steitz, T. A. (1990) *Q. Rev. Biophys. 23*, 205-280.
Strauss, F., Gaillard, C., & Prunell, A. (1981) *Eur. J. Biochem. 118*, 215-222.
Trifonov, E. N. (1985) *CRC Crit. Rev. Biochem. 19*, 89-106.
Ulyanov, N. B., & Zhurkin, V. B. (1982) *Molek. Biol. (Engl. Trans.) 16*, 857-867.
Ulyanov, N. B., & Zhurkin, V. B. (1984) *J. Biomol. Struct. Dyn. 2*, 361-385.
Ulyanov, N. B., Gorin, A. A., & Zhurkin, V. B. (1989) *Proceedings of the International Conference on Supercomputing '89: Supercomputer Applications* (Kartashev, L. P., & Kartashev, S. I., Eds.) pp 368-370, International Supercomputing Institute, Inc., St. Petersburg, FL.
Umemoto, K., Sarma, M. H., Gupta, G., & Sarma, R. H. (1990a) *Biochemistry 29*, 4714-4722.
Umemoto, K., Sarma, M. H., Gupta, G., Luo, J., & Sarma, R. H. (1990b) *J. Am. Chem. Soc. 112*, 4539-4545.
Wang, J. C. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 200-203.
Wartell, R. M., & Harrel, J. T. (1986) *Biochemistry 25*, 2664-2671.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Yanagi, K., Prive, G. G., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 201-214.
Zhou, N., Bianucci, A. M., Pattabiraman, N., & James, T. L. (1987) *Biochemistry 26*, 7905-7913.
Zhurkin, V. B., Lysov, Yu. P., & Ivanov, V. I. (1978) *Biopolymers 17*, 377-412.
Zhurkin, V. B., Poltev, V. I., & Florentiev, V. L. (1981) *Molek. Biol. (Engl. Trans.) 14*, 1116-1130.
Zhurkin, V. B., Lysov, Yu. P., Florentiev, V. L., & Ivanov, V. I. (1982) *Nucleic Acids Res. 10*, 1811-1830.
Zhurkin, V. B., Ulyanov, N. B., & Ivanov, V. I. (1988) in *Structure & Expression* (Olson, W. K., Sarma, M. H., Sarma, R. H., & Sundaralingam, M., Eds.) Vol. 3, pp 169-190, Adenine, New York.
Zhurkin, V. B., Gorin, A. A., Charakhchyan, A. A., & Ulyanov, N. B. (1990) in *Theoretical Biochemistry & Molecular Biophysics* (Beveridge, D. L., & Lavery, R., Eds.) pp 411-431, Adenine, New York.
Zhurkin, V. B., Ulyanov, N. B., Gorin, A. A., & Jernigan, R. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7046-7050.

Structural Mapping of Cysteine-63 of the Chloroplast ATP Synthase β Subunit†

Kim K. Colvert,‡ Denise A. Mills, and Mark L. Richter*
Department of Biochemistry, The University of Kansas, Lawrence, Kansas 66045
Received January 17, 1991; Revised Manuscript Received January 16, 1992

ABSTRACT: The single sulfhydryl residue (cysteine-63) of the β subunit of the chloroplast ATP synthase F_1 (CF_1) was accessible to labeling reagents only after removal of the β subunit from the enzyme complex. This suggests that cysteine-63 may be located at an interface between the β and the α subunits of CF_1, although alternative explanations such as a conformational change in β brought about by its release from CF_1 cannot be ruled out. Cysteine-63 was specifically labeled with [(diethylamino)methylcoumarinyl]-maleimide, and the distance between this site and trinitrophenyl-ADP at the nucleotide binding site on β was mapped using fluorescence resonance energy transfer. Cysteine-63 is located in a hydrophobic pocket, 42 Å away from the nucleotide binding site on β.

The chloroplast coupling factor 1 (CF_1)[1] utilizes the energy of a transmembrane proton gradient to catalyze ATP synthesis. CF_1 is comprised of five different subunits designated α-ϵ in order of decreasing molecular weight. The probable subunit stoichiometry is 3α, 3β, 1γ, 1δ, and 1ϵ (McCarty & Moroney, 1985). The two larger subunits, α and β, are involved in nucleotide binding and catalysis. Attachment of photoaffinity nucleotide analogues located at known nucleotide binding/catalytic sites on CF_1 results primarily in labeling of the β

† This work was supported by grants from the National Science Foundation (DMB-8805048) and from The University of Kansas (GRF 3009).
* To whom correspondence should be addressed.
‡ Recipient of an NSF research opportunity award. Permanent address: Department of Physical Sciences, Ferris State University, Big Rapids, MI 49307.

[1] Abbreviations: CF_1, chloroplast coupling factor 1; TNP-ATP and TNP-ADP, 2′,3′-*O*-(2,4,6-trinitrophenyl) derivatives of ATP and ADP; Tricine, *N*-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; Tris, tris(hydroxymethyl)aminomethane; CPM, *N*-[4-[7-(diethylamino)-4-methylcoumarin-3-yl]]maleimide.

0006-2960/92/0431-3930$03.00/0