

inactivate wild-type protein produced in the same cell, presumably by formation of stably folded but inactive heterodimers. Thus, the first 68 residues, which form helices A, B, and C in the intact protein, are sufficient for heterodimer formation. The following lines of evidence further support a dimeric structure for the nativelike N-fragment: mobility in native polyacrylamide gel electrophoresis (11); concentration dependence of sedimentation velocity; concentration dependence of the fluorescence spectrum and the fluorescence-detected urea denaturation midpoint; and glutaraldehyde cross-linking. In addition, N-fragment can form heterodimers with both intact and armless TrpR.

16. A. Bundi and K. Wüthrich, *Biopolymers* **18**, 285 (1979).
17. Other lines of evidence support the conclusion that reconstituted *trp* repressor has nativelike structure. First, the only Trp residues in TrpR are 19 and 99. Fluorescence spectra show nearly identical maximum wavelengths ( $\lambda_{\max}$ ) and total integrated intensities for intact TrpR, armless TrpR, intact Trp<sup>99</sup> → Phe mutant TrpR (C. Royer, personal communication), N-fragment, and the reconstituted mixture of N- and C-fragments. Time-resolved fluorescence spectra show a 4-ns lifetime component for both N-fragment and the Trp<sup>99</sup> → Phe mutant proteins. These results suggest that the fluorescence spectrum of intact protein is due almost entirely to Trp<sup>19</sup> and

that the environment of this residue in N-fragment is also nativelike. Second, we (in collaboration with C. L. Lawson) have obtained crystals of the reconstituted mixture in the presence of L-tryptophan under the same conditions that yield tetragonal crystals (4, 6) of intact *trp* holorepressor. In addition, reconstitution detected by CD shows that maximal increase in helicity occurs at a 1:1 ratio of N- and C-fragments at ~8  $\mu$ M each; addition of up to sixfold excess C-fragment causes no further increase in helicity.

18. Helix F represents ~32% of the residues of C-fragment (12 of 37 residues). The fraction of residues in helical conformation went down slightly after reconstitution of N-fragment with C-fragment, from ~56% to ~47%, which is consistent with formation of helical structure in only the helix F portion of C-fragment, with little or no change either in the remaining residues of C-fragment or in the N-fragment portion of the reconstituted molecule.
19. The intertwined topology of TrpR is strikingly similar to that of the all- $\alpha$ -helical interferon dimer (20), even though the two proteins share no apparent homology. Interferons occur in both monomeric and dimeric forms, and x-ray structures are known for each. Comparison of these two forms reveals a remarkable conservation of chain folds (20). The NH<sub>2</sub>-terminal helices are arranged almost identically in both monomeric and dimeric interferons, whereas

the position of the two COOH-terminal helices differs in the two forms: in the monomer, these helices are packed internally among the NH<sub>2</sub>-terminal helices, whereas in the dimer the distal segment of each polypeptide crosses the subunit interface to pack with the opposite subunit, as in *trp* repressor. These observations suggest the possibility that an early step in assembly of dimeric interferon involves folding of the NH<sub>2</sub>-terminal part of the protein into a nativelike structure common to both monomeric and dimeric forms. This fold provides a recognition surface for dimerization, whereas the distal chain segments are incorrectly or incompletely folded until a later stage of the folding process.

20. S. E. Ealick *et al.*, *Science* **252**, 698 (1991).
21. T. G. Oas and P. S. Kim, *Nature* **336**, 42 (1988); J. P. Staley and P. S. Kim, *ibid.* **344**, 685 (1990).
22. A. M. Lesk and G. D. Rose, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4304 (1981); J. S. Richardson, *Adv. Protein Chem.* **34**, 167 (1981).
23. R. S. Mulvey, R. J. Gualtieri, S. Beychok, *Biochemistry* **13**, 782 (1974).
24. Supported by grant GM43558 to J.C. We thank K. Valentine for assistance in acquiring some of the NMR spectra and R. Punch and F. Figueirido for assistance with computer modeling. Discussions with D. Barrick and P. Kim are acknowledged.

3 September 1991; accepted 25 November 1991

## Toward a Dynamical Structure of DNA: Comparison of Theoretical and Experimental NOE Intensities

JANE M. WITHKA, S. SWAMINATHAN,\* JAYASHREE SRINIVASAN, DAVID L. BEVERIDGE, PHILIP H. BOLTON

Comparisons of experimental and calculated interproton nuclear Overhauser effect (NOE) buildup curves for duplex d(CGCGAATTCGCG)<sub>2</sub> have been made. The calculated NOEs are based on molecular dynamics simulations including counterions and water and on the single-structure canonical A, B, and Z crystal forms. The calculated NOE effects include consideration of the motions of individual interproton vectors and the anisotropic tumbling of the DNA. The effects due to inclusion of anisotropic tumbling are much larger than those due to the local motion, and both improve the agreement between calculated and experimental results. The predictions based on the dynamical models agree significantly better with experiment than those based on either of the canonical forms or the crystal structure.

THE FIRST DETERMINATION OF THE structure of DNA was based on diffraction data on fibers (1); subsequently, higher resolution structural information was obtained by crystallography on oligonucleotides (2). A number of crystal structures in the A, B, and Z families have been reported (2–5). The analysis of these structures has revealed further evidence for sequence-dependent fine structure and bending. However, the applicability of these observations to the structure of DNA in solution has not been determined, because the solution structure per se is not accurately known.

Studies of DNA in solution based on

extensive molecular dynamics (MD) simulations including counterions and water (6) have recently been reported from this laboratory. A theoretical prediction on the structure of duplex d(CGCGAATTCGCG)<sub>2</sub> in solution (Fig. 1) exhibits qualitatively, if not quantitatively, the propeller twist in the base pairs and local axis deformations at or near the GC/AT interface seen in the crystal structure of Dickerson and Drew (2). However, the calculations involve approximations at the level of the intermolecular force field and assumptions in the simulation protocol (7–9), and the accuracy of the predictions based on MD simulation of DNA structure in solution has not yet been fully validated. A detailed comparison of theoretical and experimental results for DNA in solution is required to determine the quality of the MD results that are obtained in the absence of any artificial energy constraints.

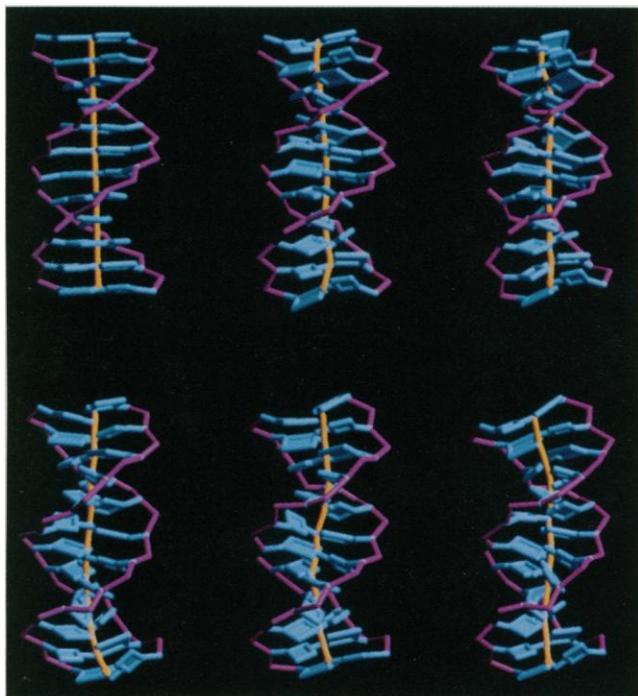
NOE data provide, in principle, extensive information on interproton distances in molecules (10–12). However, structures cannot be determined directly from the NOE data because of a number of factors, which include information on the short-range interproton distance ( $\leq 5$  Å), spin diffusion effects, and the internal motion and anisotropic tumbling of the DNA molecules (13–16). We have recently described a procedure for obtaining the NOE intensities directly from the trajectories of MD simulation (15). This method includes consideration of anisotropic tumbling of the duplex and the orientation of the <sup>1</sup>H–<sup>1</sup>H vectors to the symmetry axis of the duplex and the extent of local motion of the <sup>1</sup>H–<sup>1</sup>H vectors. These features along with the interproton distances can be used to determine the experimentally observable <sup>1</sup>H–<sup>1</sup>H NOEs (13–15).

The orientation of the <sup>1</sup>H–<sup>1</sup>H vectors to the helical symmetry axis is important because a duplex dodecamer DNA, an asymmetric molecule with a length about twice that of the diameter, undergoes anisotropic tumbling in solution (13–16). Therefore, a <sup>1</sup>H–<sup>1</sup>H vector pointing along the symmetry axis will have an effective correlation time longer than one perpendicular to the helical axis because the tumbling of the DNA helix in solution will be faster about the symmetry axis than about the axis perpendicular to the symmetry axis. The local motion of a <sup>1</sup>H–<sup>1</sup>H vector, of extent  $\theta$ , is about the average orientation,  $\beta$ , of the <sup>1</sup>H–<sup>1</sup>H vector to the helical axis (Fig. 2).

Kaptein and co-workers recently included some aspects of internal motion in conjunction with internuclear distance constraints in

Chemistry Department, Wesleyan University, Middletown, CT 06459.

\*Present address: Gilead Sciences, Foster City, CA 94404.



**Fig. 1.** A sequence of snapshots of the structure of  $d(\text{CGCGAATTCGCG})_2$  from the WC/B molecular dynamics trajectory. Structures in the top row (left to right) were obtained at 0, 20, and 50 ps, and those in the bottom row (left to right) at 80, 110, and 140 ps. The helical axis is depicted in yellow, the backbone of the duplex in pink, and the orientation of the bases in blue.

modeling DNA (17). The calculation of the experimental NOEs from the trajectories involves all of the magnetization transfer pathways, which is especially important in DNA because spin diffusion occurs even at the shortest accessible mixing times (13–15). Analysis of NOE buildup curves shows that, even at the shortest accessible mixing times, the slopes are not proportional to the inverse sixth power of the internuclear distance (15, 18).

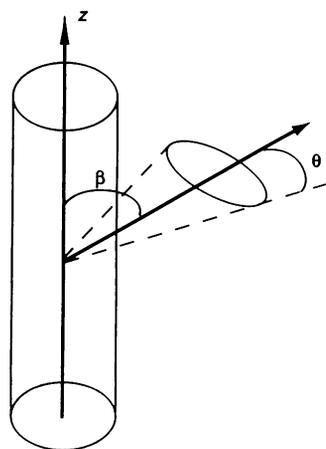
This report describes the determination of the time evolution of NOE intensities for  $d(\text{CGCGAATTCGCG})_2$  and a detailed comparison of the observed results with those calculated on the basis of MD simulations and the single-structure canonical and

crystallographic forms. Two-dimensional NOE data were obtained with a Varian Unity 500-MHz spectrometer for mixing times of 25, 50, 75, 100, 150, and 200 ms on a 3 mM sample of  $d(\text{CGCGAATTCGCG})_2$  in 0.1 M NaCl at pH 7.0 and a temperature of 30°C (19). The peak assignments in the dodecamer spectra were made by sequential methods (10, 11) and were found to be consistent with those previously reported (20, 21). A total of 217 resolvable  $^1\text{H}$ - $^1\text{H}$  NOE correlations were obtained for each strand of the symmetric duplex. The nuclear magnetic resonance (NMR) results show that the structure of the dodecamer in solution falls into the DNA B form family on the basis of the aromatic to ribose NOE

connectivities and intensities (11, 20, 21).

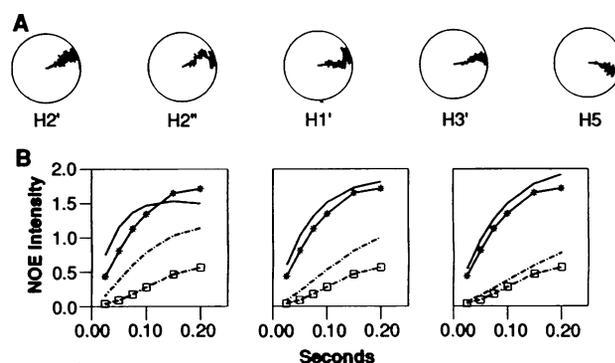
The MD simulations (6) involve the dodecamer duplex, 1927 water molecules, and 22  $\text{Na}^+$  counterions carried out in three different forms: (i) the ES/B model, with a canonical B form initial structure and the GROMOS 86 force field (8), (ii) the WC/B model, essentially the ES model including harmonic distance restraints between the atoms involved in base pair hydrogen bonding, and (iii) the WC/X model, utilizing the force field of the WC model described above starting with the x-ray crystal structure. The ES/B model leads to a disruption of Watson-Crick base pairing at the G4-C21 base pair that, on the basis of base-pair exchange rates, is considered to be unphysical (6, 22–24). Thus, the WC restraints were introduced to maintain base pairing in accord with experiment. Moreover, the comparison of the WC/B and WC/X simulations on the dodecamer provides information on the sensitivity of the results to the initial structure.

The MD simulations were carried out for 140 ps in the case of WC/B, 140 ps for ES/B, and 200 ps for WC/X. The dynamical description of the  $^1\text{H}$ - $^1\text{H}$  vectors from the simulation is then used in the determination of the theoretical NOE intensities. The cross- and autorelaxation rates are obtained by using the  $^1\text{H}$ - $^1\text{H}$  distances and motional parameters to predict the NOEs (13–15, 25). The inclusion of anisotropic tumbling and local motion in the description leads to differences in effective correlation times for each of the various  $^1\text{H}$ - $^1\text{H}$  vectors according to the procedures described elsewhere (13–15). These relaxation rates are used to construct the simultaneous differential equations that describe the evolution of magnetization as a function of mixing time. The differential equations are solved by numerical integration (15).

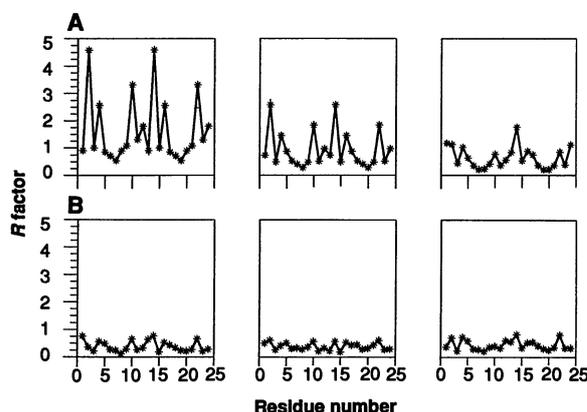


**Fig. 2.** Schematic depiction of duplex DNA as a cylinder. The orientation of a  $^1\text{H}$ - $^1\text{H}$  vector to the symmetry axis is described by the angle  $\beta$ . The extent of the internal motion of the  $^1\text{H}$ - $^1\text{H}$  vector is described by the angle  $\theta$ .

**Fig. 3.** (A) Dials that depict the orientation,  $\beta$ , of  $^1\text{H}$ - $^1\text{H}$  vectors relative to the helical axis as a function of the time of the trajectory. The start of the trajectory is at the center of the dial and the end of the trajectory at the circumference. The dials are shown for several proton pairs involving the H6 proton of cytosine 3 in the WC/B trajectory. The residue numbering is 5'( $\text{C}_1\text{G}_2\text{-C}_3\text{G}_4\text{A}_5\text{A}_6\text{T}_7\text{T}_8\text{C}_9\text{G}_{10}\text{C}_{11}\text{-G}_{12}$ )3' and for the complementary strand 5'( $\text{C}_{13}\text{G}_{14}\text{C}_{15}\text{G}_{16}\text{A}_{17}\text{A}_{18}\text{T}_{19}\text{T}_{20}\text{C}_{21}\text{G}_{22}\text{C}_{23}\text{G}_{24}$ )3'. (B) Theoretical (solid line) and experimental (\*) NOE buildup curves for the H6 to H2' pair of cytosine 3. Theoretical (dashed line) and experimental ( $\square$ ) NOE buildup curves for the H6 to H2'' pair of cytosine 3. The leftmost plots were calculated with neglect of orientation and local motion effects, the middle plots with inclusion of orientation effects, and the rightmost plots with inclusion of both orientation and local motion effects.



**Fig. 4.** (A)  $R$  values between the experimental and calculated NOE data for the single-structure models of A form, B form, and the crystal structure shown from left to right, respectively. The calculated NOE data include orientation effects. The  $R$  values were determined from the results at all six mixing times. (B)  $R$  values for the dynamical models of WC/B, ES/B, and WC/X are shown from left to right, respectively. The calculated NOE data include orientation and internal motion effects. Calculated  $R$  values for the dynamical structures were determined from the results at all six mixing times.



We use “dials” to show the motion of  $^1\text{H}$ - $^1\text{H}$  vectors during the time course of the MD trajectory for an illustrative case (Fig. 3A) (15, 26). Represented in each dial is the angle of a  $^1\text{H}$ - $^1\text{H}$  vector to the symmetry axis of the DNA duplex with the start of the trajectory at the center of the dial and the end at the circumference. The motion of the  $^1\text{H}$ - $^1\text{H}$  vectors is analyzed to determine the average angle to the symmetry axis,  $\beta$ . This orientation is important because a dodecamer duplex DNA is an anisotropic molecule. The internal motion is modeled as a damped libration, which is given by the root-mean-square fluctuation  $\theta$ , about the average orientation (13–16).

The calculated NOE buildup curves for an illustrative case are shown in Fig. 3B, which shows a comparison of the calculated and experimental intraresidue NOEs for selected  $^1\text{H}$ - $^1\text{H}$  pairs with the use of only the  $\langle r^3 \rangle$  average distances (left), of the average distances and the orientation of the  $^1\text{H}$ - $^1\text{H}$  vectors (middle), and of the extent of motion as well as the average distances and the orientation of the  $^1\text{H}$ - $^1\text{H}$  vectors (right). The inclusion of the overall anisotropic motion and the local motion effects in the model significantly improves the agreement of the theoretical and experimental results. The improvement in agreement due to inclusion of anisotropic effects tends to be much larger than gains due to inclusion of the rapid local motion both for NOEs arising primarily by direct magnetization transfer and for those connected by diffusion of magnetization.

One can determine the overall fit of the theoretical and experimental results by comparing the  $R$  values between calculated and observed NOE intensities (27) of each residue for the different models. Results are presented for the duplex DNA modeled as the canonical A form, the canonical B form, and the crystal form (single-structure models), as well as by the ensemble of structures comprising the dynamical model obtained from MD (Fig. 4). The comparisons for the

three single-structure models, including the effects of anisotropic motion, are shown in Fig. 4A. The crystal structure yields a better fit than the canonical B form structure, and both, as expected, are better than the results based on the canonical A form. However, none of the single-structure models gives a particularly good fit to the experimental data.

The corresponding results for the WC/B, ES/B, and WC/X dynamical models, including the effects of anisotropic and internal motion, are shown in Fig. 4B. Each of the simulations leads to a significantly better agreement than any of the single-structure models, canonical or crystallographic. The dynamical model intraresidue results are within experimental error, and the interresidue results tend to correspond to internuclear distances of no more than 0.2 Å less than the observed values. Therefore, the dynamical structure results agree significantly better with experiment than any of the single-structure models. With respect to  $R$  values, the three distinct MD models all have about the same overall quality of fit to the corresponding observed results, but detailed inspection reveals results that differ in the fit from site to site.

The improved agreement between the theoretical and experimental results including motional considerations demonstrates that both the anisotropic tumbling and local motions of the DNA need to be explicitly included for accurate modeling of DNA in solution. The improved agreement obtained from the MD results over that of single structures indicates that the dynamical models, particularly WC/B and WC/X, provide the best description of this dodecamer structure in solution quantitatively documented to date. The overall quality of fit of the experimental and theoretical data has been obtained without the introduction of interproton NOE constraints, which may lead to distorted structures and dynamics. The  $R$  values obtained for our free dynamical models are in fact comparable to those obtained

by the use of artificial internuclear constraints (17, 25).

#### REFERENCES AND NOTES

1. J. D. Watson and F. H. C. Crick, *Nature* **171**, 737 (1953).
2. R. E. Dickerson and H. R. Drew, *J. Mol. Biol.* **149**, 761 (1981).
3. R. E. Dickerson, in *Structure and Method*, vol. 3, *DNA and RNA*, R. H. Sarma and M. H. Sarma, Eds. (Adenine, New York, 1990).
4. W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag, New York, 1983).
5. R. V. Gessner, C. A. Fredrick, G. J. Quigley, A. Rich, A. H.-J. Wang, *J. Biol. Chem.* **264**, 7921 (1989).
6. S. Swaminathan, G. Ravishanker, D. L. Beveridge, *J. Am. Chem. Soc.* **113**, 5027 (1991).
7. A. J. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge Univ. Press, Cambridge, 1986).
8. W. F. van Gunsteren, H. J. C. Berendsen, R. G. Geurtsen, H. R. Zwinderman, *Ann. N.Y. Acad. Sci.* **482**, 287 (1986).
9. J. Srinivasan, J. M. Withka, D. L. Beveridge, *Biochem. J.* **58**, 533 (1990).
10. G. M. Clore and A. M. Gronenborn, *Crit. Rev. Biochem. Mol. Biol.* **24**, 479 (1989).
11. K. Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
12. P. H. Bolton, *Prog. Nucl. Magn. Reson. Spectrosc.* **22**, 423 (1990).
13. J. M. Withka, S. Swaminathan, P. H. Bolton, *J. Magn. Reson.* **89**, 386 (1990).
14. ———, in *Proceedings of the NATO Advanced Research Workshop, Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy*, J. Hoch, Ed. (Plenum, New York, 1991), p. 409.
15. J. M. Withka, S. Swaminathan, D. L. Beveridge, P. H. Bolton, *J. Am. Chem. Soc.* **113**, 5041 (1991).
16. W. Eimer, J. R. Williamson, S. G. Boxer, R. Pecora, *Biochemistry* **29**, 799 (1990).
17. T. M. G. Koning, R. Boelens, G. A. van der Marel, J. H. van Boom, R. Kaptein, *ibid.* **30**, 3787 (1991).
18. J. M. Withka, S. Swaminathan, J. Srinivasan, D. L. Beveridge, P. H. Bolton, in preparation.
19. J. M. Withka, J. A. Wilde, P. H. Bolton, A. Mazumder, J. A. Gerlt, *Biochemistry* **30**, 9931 (1991).
20. D. R. Hare, D. E. Wemmer, S. H. Chou, G. Drobny, B. R. Reid, *J. Mol. Biol.* **171**, 319 (1983).
21. W. Nerdal, D. R. Hare, B. R. Reid, *Biochemistry* **28**, 10008 (1989).
22. A. Pardi and I. Tinoco, Jr., *ibid.* **21**, 4686 (1982).
23. A. Pardi, K. M. Morden, D. J. Patel, I. Tinoco, Jr., *ibid.*, p. 6567.
24. J. G. Moe and I. Russu, *Nucleic Acids Res.* **18**, 821 (1990).
25. M. Gochin and T. L. James, *Biochemistry* **29**, 11172 (1990).
26. G. Ravishanker, S. Swaminathan, D. L. Beveridge, R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **6**, 669 (1989).
27. C. Gonzalez, J. A. C. Rullmann, A. M. J. J. Bonvin, R. Boelens, R. Kaptein, *J. Magn. Reson.* **91**, 659 (1991).
28. We thank G. Ravishanker for assistance with the preparation of Fig. 1, which was ray-traced with the use of the public domain program Raster3D and photographed by K. J. McConnell. V. V. Krishnamurthy's assistance in obtaining the NMR data and that of F. M. DiCapua in preparing figures are gratefully appreciated. Supported, in part, by grants from the American Cancer Society (NP-750, P.H.B.), from the National Institutes of Health (GM-37909, D.L.B.), from the Bristol-Myers Squibb Corporation via participation in a State of Connecticut Cooperative High Technology Research and Development Grant (P.H.B. and D.L.B.); J.M.W. is a recipient of a National Institutes of Health Traineeship in Molecular Biophysics (1T32 GM-08271, P.H.B. and D.L.B.). Computer time on the CRAY YMP was provided by the Pittsburgh Supercomputer Center.

2 August 1991; accepted 4 November 1991