J|A|C|S
COMMUNICATIONS

Published on Web 12/10/2004

Resolving Conflicting Crystallographic and NMR Models for Solution-State DNA with Solution X-ray Diffraction

Xiaobing Zuo and David M. Tiede*

Chemistry Division, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439

Received September 9, 2004; E-mail: tiede@anl.gov

X-ray diffraction and NMR are the most frequently used techniques for DNA structure determination. X-ray fiber diffraction first revealed the gross features (such as pitch, step height, and helix radius) of DNA in the 1950s[1] and is used for DNA that cannot be crystallized. Single-crystal X-ray diffraction provides most accurate and detailed structure and has dramatically increased the knowledge of DNA structures since the late 1970s.[2] However, DNA crystal structures may or may not fairly represent the solution structure because they are subject to crystal packing forces that can have a big impact on both local and global structures.[3,4] NMR has achieved great success in solving solution-state atomic structures for proteins and nucleic acids in the past 20 years.[5] Nuclear Overhauser enhancement (NOE) measurements are the most basic source for 3-D NMR structure determination, generating a large number of internuclear distance constraints, typically within 5 Å. Recently, new NMR techniques, such as residual dipolar coupling[6,7] and chemical shift anisotropy,[8,9] were developed to provide more and longer distance and orientation constraints. Such advances have already made NMR a relatively robust technique for protein structure determination. However, the relatively low density of protons and short contact distances measured make accurate global DNA structure determination by NMR a challenge and significantly dependent on the choice of mean-force interaction potentials in configurational refinement.[4,10] Therefore, a need exists for a complementary solution structure characterization technique capable of evaluating the accuracy of NMR conformational models and methods that led to them.

Herein, we report on synchrotron-based high-angle X-ray solution scattering measured to 2 Å resolution for two synthetic DNA sequences for which there are numerous conflicting X-ray crystal and solution NMR models. Our results demonstrate that high-angle X-ray scattering is capable of discriminating between each of the models through measurement of composite structural parameters for solution-state DNA and provides a direct, independent method for testing structural models and measuring solution-state configurational dispersions.

For monodispersed macromolecules in solution, scattering can be understood to arise from atomic coordinate-based, orientationally averaged molecular diffraction. Small-angle X-ray scattering reflects the overall macromolecular size and shape,[11,12] while at high angle an interference pattern (molecular diffraction) is measured that arises from spatially resolved atomic-pair correlations. High-angle solution X-ray diffraction combined with coordinate-based model analyses has been demonstrated to provide a unique measure of supramolecular conformation and configurational dispersion in liquids.[13−15] A brief description of solution X-ray diffraction simulation is provided in Supporting Information (SI). Because of its repetitive structure, DNA shows a highly structured interference pattern whose peaks correspond to those measured in DNA fiber diffraction.[1] Simulations indicate that the interference pattern in the q range 0.2−1.7 Å^{-1} is dominated by sugar and phosphate backbone

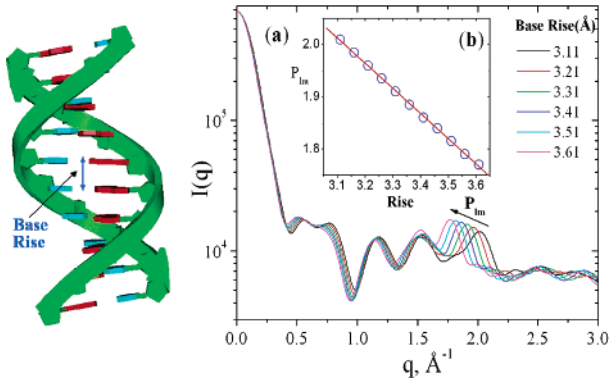

Figure 1. Simulations on solution X-ray scattering using eq S1 (SI) for $d(A)_{10}-d(T)_{10}$ with various base rise otherwise in canonical B-form (a) and relationship between the peak position of $\mathbf{P_{lm}}$ and base rise (b). Scattering angle $q = 4\pi \sin\theta/\lambda$.

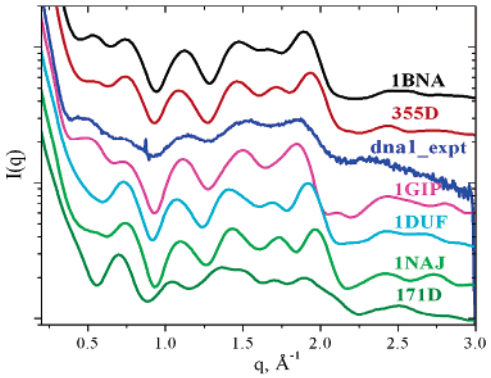

Figure 2. Experimental solution interference pattern (**dna1_expt**) for **dna1** and calculated patterns for its crystal (**1BNA** and **355D**) and NMR (**1GIP**, **1DUF**, **1NAJ**, and **171D**) models using eq S1. ∼1 mM **dna1** in tris-buffer (pH = 7.0), 0.10 M NaCl. $I(q)$ in logarithm scale, and curves are raised to avoid overlapping otherwise with same $I(0)$.

moieties. However, the last major peak ($\mathbf{P_{lm}}$) with $q \approx 1.7-2.1$ Å^{-1} can be shown to arise nearly exclusively from the base pair stacking rise (Figure S1). Figure 1 shows the correlation between $\mathbf{P_{lm}}$ and the magnitude of the base rise modeled for otherwise canonical B-form $d(A)_{10}-d(T)_{10}$. $\mathbf{P_{lm}}$ shifts to low q while increasing the base rise.

The self-complementary sequence $d(CGCGAATTCGCG)_2$ (**dna1**) is known as Drew−Dickerson DNA, which was the first B-form DNA solved by crystallography and has been most intensively investigated using crystallography and NMR. The experimental data for **dna1** and the calculated X-ray interference patterns for crystal (Protein DataBank ID: **1BNA**[16] and **355D**[17]) and NMR (**171D**,[18] **1DUF**,[10] **1GIP**,[4] and **1NAJ**[9]) models are displayed in Figure 2. The experimental interference pattern shows a series of configurationally broadened peaks that can be correlated with interference patterns calculated from model structures. The NMR structure **1GIP**

10.1021/ja044533+ CCC: $30.25

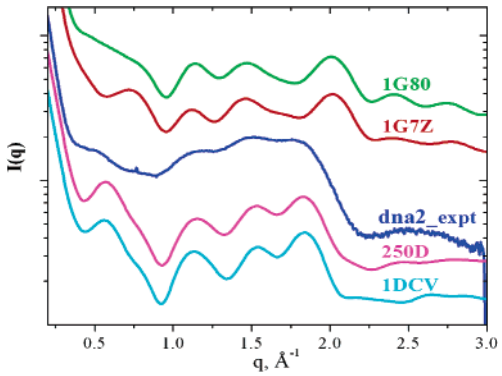

Figure 3. Experimental solution interference pattern (**dna2_expt**) for **dna2** and calculated patterns for its NMR (**1G80**, **1G7Z**) and crystal (**250D**, **1DCV**) models. Same experimental conditions as **dna1**.

was found to have the best peak match with experimental data, having a deviation no greater than 0.01 Å^{-1} for each major peak. For crystal structure **1BNA**, both the first and last major peaks (at 0.4 and 1.8 Å^{-1}, respectively) of calculated interference pattern are shifted by 0.04 Å^{-1} to higher angles, the former due to different backbone geometry and the latter due to the slightly shorter base rise in crystalline compared to solution forms. Similarly, $\mathbf{P_{lm}}$ of the calculated pattern for **355D** has a further shift of 0.04 Å^{-1}, indicating an even shorter base rise in this crystal. Among the NMR models, **1DUF**, **1NAJ**, and **1GIP** match experimental data better than **171D** as judged by the peak positions and shapes of the calculated patterns. In the latter model, interatomic distances were mostly determined by NOE measurements. Extra internuclear distance and orientation restraints were employed in the former three models, which came from residual dipolar coupling and dramatically improved the quality of NMR structures. In model **1NAJ**, more restraints, i.e., ^{31}P−^{1}H dipolar coupling and ^{31}P chemical shift anisotropy, were employed than in **1DUF**. However, both of them miss the first peak and significantly shift $\mathbf{P_{lm}}$ compared to the experimental data. **1NAJ** has limited improvement for peaks in the *q* range of 1.0−1.6 Å^{-1}, but is a worst match for $\mathbf{P_{lm}}$. Kuszewski et al.[4] used the same NMR data as **1DUF** but a different base−base potential interaction taken from high-resolution crystal structures. The resulting best match model **1GIP** demonstrates that the description of nonbonded contacts has a large impact on the accuracy of local and global NMR structures.

Figure 3 shows the interference pattern measured for a shorter self-complementary sequence, d(CGCTAGCG)$_2$ (**dna2**) and calculated patterns for four structural models for this or similar sequences. The experimental peak positions measured for **dna2** show that this sequence, like **dna1**, adopts a conformation in solution resembling that of B-form calf thymus DNA (Figure S3), although additional configurational broadening effects[15] are also evident for the shorter **dna2** sequence. **250D**[19] and **1G7Z**[20] are crystal and NMR models for **dna2**, respectively, while **1G80**[20] is an NMR structure for a similar sequence d(GCGTACGC)$_2$ and **1DCV**[21] is a crystal structure with one extra base pair at each end, otherwise identical to **dna2**. Only the **dna2** part in **1DCV** was used for calculation in Figure 3. Closely analogous interference patterns were calculated for crystal models **250D** and **1DCV**. Except for a shift in the first diffraction peak and significant configurational broadening in the experimental data, the peak position and pattern matches of the crystal models to **dna2** experimental data identify these as excellent models for equilibrium solution-state **dna2**. The shift of the first peak may reflect a small conformational deviation for **dna2** between crystalline and solution states. Interference patterns for NMR models **1G7Z** and **1G80** are similar but deviate significantly from those of crystal models and data. The absence of the first peak in calculated patterns for **dna2** NMR models is due to the large helical radius (∼1 Å larger than that of crystal models, Figure S2), which causes this peak to be lost in base-pair scattering. $\mathbf{P_{lm}}$ for the **dna2** NMR models are shifted by 0.2 Å^{-1} to higher angle compared to experimental data and crystal models, corresponding to a 0.4 Å shorter base rise. As discussed above, the mismatches of the NMR models can arise from a combination of inaccurate measurements of internuclear distances and orientations or improper treatment of nonbonded contacts. Crystallographic models were found to provide a good match to solution X-ray diffraction data for **dna2** but not **dna1**.

In summary, we have demonstrated that the composite structural parameters determined by solution X-ray diffraction provide a direct measure of DNA solution-state conformation capable of discriminating between differing NMR and crystallographic models. The solution diffraction method provides a complementary approach for evaluating accuracies of NMR refinement procedures, for configurational analyses of DNA sequences and other macromolecular assemblies not amenable for crystallographic or NMR structure determination, and for 100-ps time-resolved structural studies.

Acknowledgment. This work was supported by the Office of Basic Energy Sciences, U.S. Department of Energy (contract W-31-109-ENG-38). We thank Drs. R. Zhang and A. Goshe for insightful discussions and the Advanced Photon Source (Argonne National Laboratory) Sector 11-12 staff, especially Drs. S. Seifert and G. Jennings, for their support on synchrotron experiments.

Supporting Information Available: Solvent influence on DNA X-ray scattering (Figure S1), components scattering for **dna2** models (Figure S2), and experimental scattering for calf thymus DNA (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Klug, A. *J. Mol. Biol.* **2004**, *335*, 3−26.
(2) Neidle, S., Ed. *Oxford Handbook of Nucleic Acid Structure*; Oxford University Press: New York, 1999.
(3) (a) Abrescia, N. G. A.; Thompson, A.; Huynh-Dinh, T.; Subirana, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2806−2811. (b) Abrescia, N. G. A.; González, C.; Gouyette, C.; Subirana, J. A. *Biochemistry* **2004**, *43*, 4092−4100.
(4) Kuszewski, J.; Schwieters, C.; Clore, G. M. *J. Am. Chem. Soc.* **2001**, *123*, 3903−3918.
(5) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.
(6) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(7) Lipsitz, R. S.; Tjandra, N. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 387−412.
(8) Wu, Z.; Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **2001**, *123*, 3617−3618.
(9) Wu, Z.; Delaglio, F.; Tjandra, N.; Zhurkin, V. B.; Bax, A. *J. Biomol. NMR* **2003**, *26*, 297−315.
(10) Tjandra, N.; Tate, S.-i.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 6190−6200.
(11) Glatter, O.; Kratky, O. *Small-Angle X-ray Scattering*; Academic Press: New York, 1982.
(12) Feigin, L. A.; Svergun, D. I. *Structure Analysis by Small Angle X-ray and Neutron Scattering*; Plenum Press: New York, 1987.
(13) Tiede, D. M.; Zhang, R.; Seifert, S. *Biochemistry* **2002**, *41*, 6605−6614.
(14) Zhang, R.; Thiyagarajan, P.; Tiede, D. M. *J. Appl. Crystallogr.* **2000**, *33*, 565−568.
(15) Tiede, D. M.; Zhang, R.; Chen, L. X.; Yu, L.; Lindsey, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 14054−14062.
(16) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 2179−2183.
(17) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355.
(18) Schweitzer, B. I.; Mikita, T.; Kellogg, G. W.; Gardner, K. H.; Beardsley, G. P. *Biochemistry* **1994**, *33*, 11460−11475.
(19) Urpi, L.; Tereshko, V.; Malinina, L.; Huynh-Dinh, T.; Subirana, J. A. *Nat. Struct. Biol.* **1996**, *3*, 325−328.
(20) Isaacs, R. J.; Spielmann, H. P. *J. Mol. Biol.* **2001**, *307*, 525−540.
(21) Eichman, B. F.; Vargason, J. M.; Mooers, B. H. M.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3971−3976.

JA044533+