FULL PAPER

B → A → B Transitions in a Molecular Dynamics Trajectory of Low Salt DNA Solution

Alan N. Real and Robert J. Greenall

Department of Physics, University of York, York, YO10 5DD, United Kingdom; Tel: (+44) (0)1904 432235, Fax: (+44) (0)1904 432214. E-mail: rjg2@york.ac.uk

Abstract A molecular dynamics simulation of the DNA dodecamer d(CGCATATATGCG) has been performed with AMBER 5.0 under low salt conditions. Both B → A and A → B transitions are observed. This may have biological significance for the formation of complexes between DNA and TATA-box binding proteins.

Keywords Molecular dynamics, Simulation, 1DN9, A-DNA, B-DNA, Conformational transitions

Introduction

The conformations adopted by DNA and the transitions between them are related to biological functions. The main conformations - A, B and Z [1] - are dependent on base sequence, cationic environment and water activity. For example, in fibres, transitions between the 'wet' B-form and the 'dry' A-form can be induced by changing the water content.[2] Fibres of DNA containing alternating AT tracts can adopt the D-form,[3] which is not available to DNA molecules with general base sequences, and a transition to the B-form can occur if the humidity is increased.[4] Such alternating sequences are of particular interest since, during transcription, TATA-box binding proteins (TBPs) bind in the minor groove of DNA with consensus sequence TATA T/A A T/A X.[5] X-ray crystallographic studies [5] have shown that TBP binding produces a radical deformation of the DNA

from the B-form (which the DNA would normally be expected to adopt in solution) to a new form, known as TA-DNA,[6] which is similar to A-DNA but with a significantly higher base inclination. Thus it appears likely that the consensus sequence might be susceptible to a B → A transition during, or immediately prior to, TBP binding.

Molecular dynamics (MD) is becoming a powerful tool for probing the structure and dynamics of DNA, particularly with regard to the effects of base sequence and environmental factors. Recent developments in technique, particularly the Particle Mesh Ewald algorithm [7-10] for calculating electrostatic interactions, have improved the realism of MD simulations so that they can now reproduce A and B forms under experimentally observed conditions.[11, 12] Several simulations have revealed conformational transitions, but there is some indication that the direction of the B → A transition is potential-dependent. For example the CHARMM23 potential [13] appears to favor the B → A transition [14-16] although the CHARMM27 potential [17] samples an equilibrium between A and B that can be perturbed by changing the environmental conditions.[18] Con-

Correspondence to: R. J. Greenall

versely, in the case of the Cornell et al. [19] potential with one exception, only the A B transition has been reported under low salt and high humidity conditions,[11, 12, 20] which is consistent with experimental observations.[1] The one exception was a gradual B → A transition in the TATA-box sequence d(GCGTATATAAAACGC). [21] However B → A → B transitions have been observed with this potential [14] when sufficient Na+ and Cl- ions were added to model a 0.45M NaCl solution. Thus at present the Cornell et al. [19] potential appears to favor the B-form under low salt conditions although there is some evidence that this property may be sensitive to the base sequence of the DNA.

We have conducted an MD simulation, using AMBER 5.0 [22] and the Cornell et al. [19] force field, of the dodecamer d(CGCATATATGCG)2. In crystals [23] this sequence adopts a B-DNA structure, but in the AT-tract the rotation per residue and the base-pair roll have different values at ApT and TpA steps consistent with the alternating-B model for poly d(A-T) proposed by Klug et al..[24] During the first 2ns of this trajectory the molecule undergoes a concerted B → A → B transition that we describe in this preliminary report.
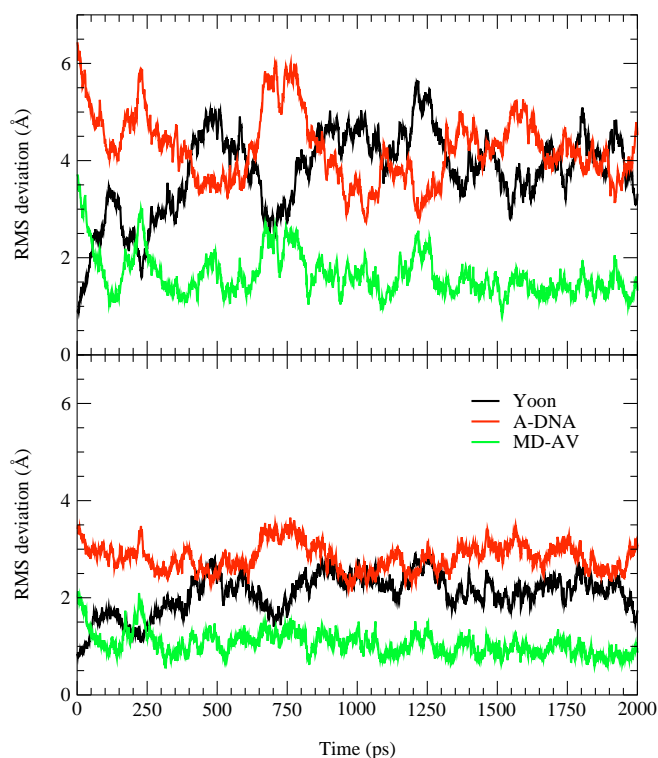
Figure 1 *Root mean square (RMS) deviations of the trajectory with respect to the initial 1DN9 (Yoon) coordinates (black), canonical A-DNA (red) and the MD average (green). The top figure shows RMS deviations calculated for the whole molecule and the lower figure is for the central AT tract*
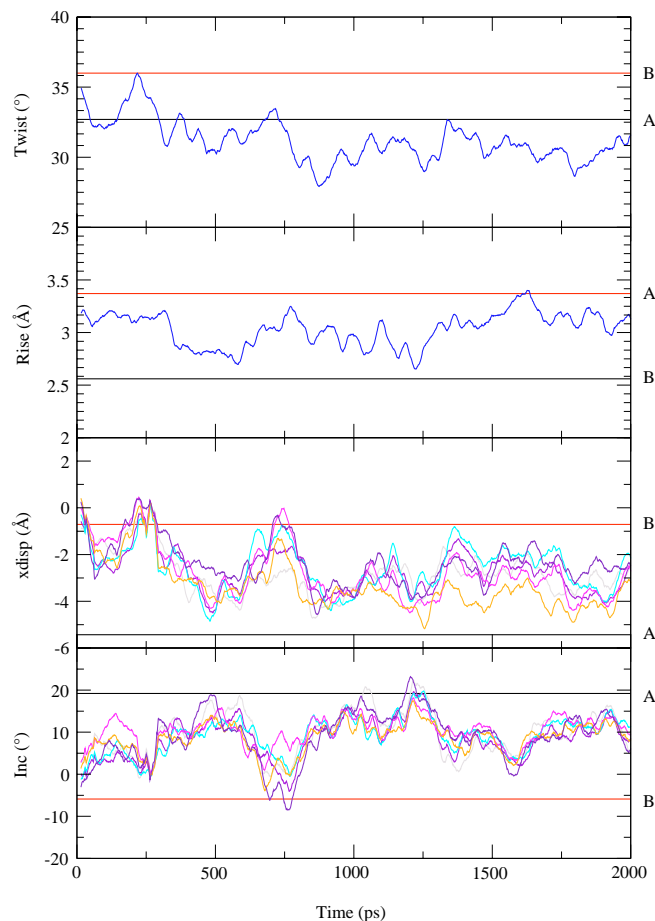
Figure 2 *Evolution with time of twist, rise, xdisp and inc. In the case of twist and rise the average values over the central AT region of the molecule are shown. For xdisp and inc individual values for each nucleotide in the AT tract are shown superimposed. All data has been 'smoothed' for clarity using a running average over 30 ps. Values of these parameters for canonical A- and B-forms are shown by horizontal lines*

Methods

Three simulations were performed with the initial DNA coordinates from the Protein Data Bank entry 1DN9 [23] and canonical A and B fibre structures.[25] 22 sodium ions were placed around the DNA to neutralize the phosphate charges and the whole system was immersed in 4990 Monte-Carlo equilibrated TIP3P waters, resulting in a bath of dimensions ~52×58×69 Å MD was initiated after several rounds of semi-constrained and eventually unconstrained steepest descent minimizations. Harmonic constraints of 25 kcal·mol⁻¹·Å² were placed on the DNA atom and ion positions during the first 500 steps of energy minimization and gradually reduced to zero in 5 kcal·mol⁻¹·Å² and 12.5 kcal·mol⁻¹·Å² steps for the DNA and ions respectively, ending with 500 steps of unconstrained minimization. With the DNA held rigid the system was then heated rapidly from 100 K to 300 K over 1 ps and

held at 300 K for 24 ps followed by a slower reheating from 10 to 300K over 5 ps using similar constraints on the DNA to those employed during the minimization stage. These constraints were reduced to zero by 5 kcal·mol^{-1}·Å^{2}· ps^{-1} over the next 5 ps before initiating a production run. Long range electrostatics were treated via the Particle Mesh Ewald (PME) method. During the constant pressure MD a 2 fs timestep was used with coordinate snapshots taken every 1 ps. Temperature regulation was achieved through Berendsen's coupling method.[26] SHAKE constraints [27] using a geometrical tolerance of 5×10^{-5} Å were imposed on all covalent bonds involving hydrogen atoms.

Results

Figure 1 shows the root mean square deviations (RMSD) of the 1DN9 trajectory with respect to the initial B-DNA coordinates,[23] canonical A-DNA [25] and the MD average. The RMSD from canonical B-DNA [25] was very close to that from the initial coordinates and has been omitted. The RMSD indicates that the DNA changes from the initial B-form to an A-like conformation at ~500 ps, reverts to the B-form at ~750 ps and then returns to the A-form at ~1250 ps. However the RMSD is only a crude indicator of DNA conformation and so further evidence is needed to establish that B $\rightarrow$ A $\rightarrow$ B transitions have occurred.

The helical twist and rise per residue and xdisp (the displacement of a base-pair from the helical axis) and inc (the inclination angle between a base-pair and the helical axis) are shown as a function of time in figure 2. Values of these parameters for A- and B-DNA are also shown. Xdisp and inc are sensitive indicators of A-like or B-like conformations. Clearly they both alternately adopt B-like and A-like values

in synchrony with the B $\rightarrow$ A $\rightarrow$ B transitions observed in RMSD. In addition the behaviors of xdisp and inc are synchronized in each ApT and TpA step. Thus we see concerted B $\rightarrow$ A $\rightarrow$ B transitions throughout the AT-tract rather than uncorrelated transitions at each nucleotide. This is emphasized by viewing the molecule along the helix axis (figure 3) which clearly shows transitions from the B-DNA initial structure (0 ps), with its bases clustered in the center, to A-DNA (500 ps) in which the bases move radially outwards leaving a distinct central hole. The molecule then reverts to B-DNA at 750 ps and back to A-DNA at 1250 ps.

The average rise per residue and base rotation (twist) in the AT tract are shown in the top of figure 2. In fibre diffraction experiments two closely related parameters (the helix pitch and the number of residues per helix turn) are widely used to classify DNA conformations as they can be directly determined from resulting diffraction patterns of regular helices. Model B-DNA has a rise of 3.38 Å and a twist of 36° (10 base pairs per helical turn and a helix pitch of 33.8 Å) and A-DNA has a base-pair rise of 2.56 Å and a twist of 32.7° (11 base pairs per turn and a helix pitch of 28.2 Å). For DNA oligomers the criterion is a little less strict as variations in the base sequence have a significant effect on the regularity of the helix. In the simulation the rise adopts a B-like starting value and reduces towards A-like values around 500 ps before returning to B at 750 ps. A-like values are seen at 1250 ps and B-form values seen at ~1650 ps. For helical twist the picture is somewhat less clear. DNA in solution is known to adopt rotations per residue slightly lower than in the model values (e.g. 10.6 base-pairs per turn for B-DNA of random sequence [28]). Hence the twist reduces from canonical B-DNA early on in the simulation. At 500 ps the twist is low and at 750ps higher values are regained. During the slower B $\rightarrow$ A transition from 800 $\rightarrow$ 1250 ps a low twist is adopted. Low twists have been observed in previous B-DNA simula-

Figure 3 *Schematic plots of the DNA looking along the helical axis at 0, 500, 750 and 1250 ps. Similar projections for canonical A- and B-DNA are also shown*
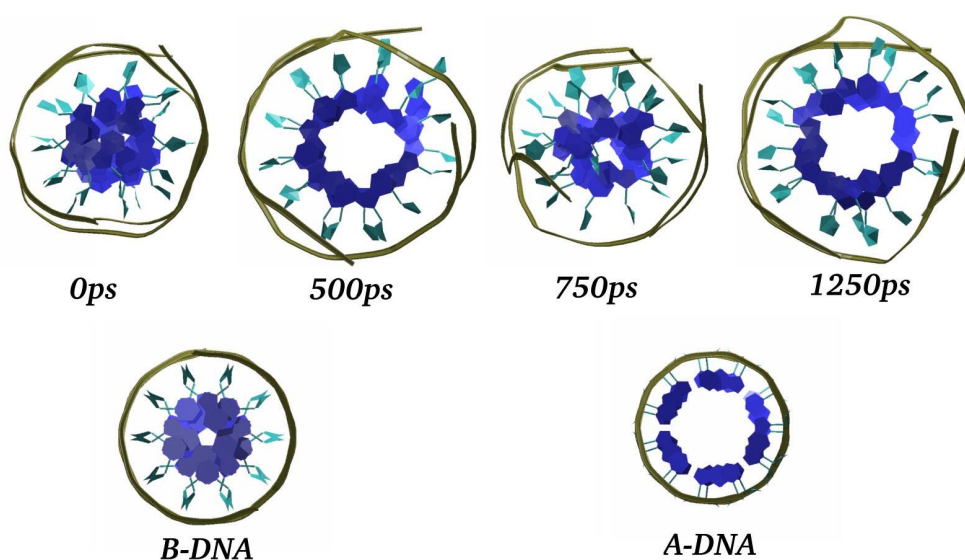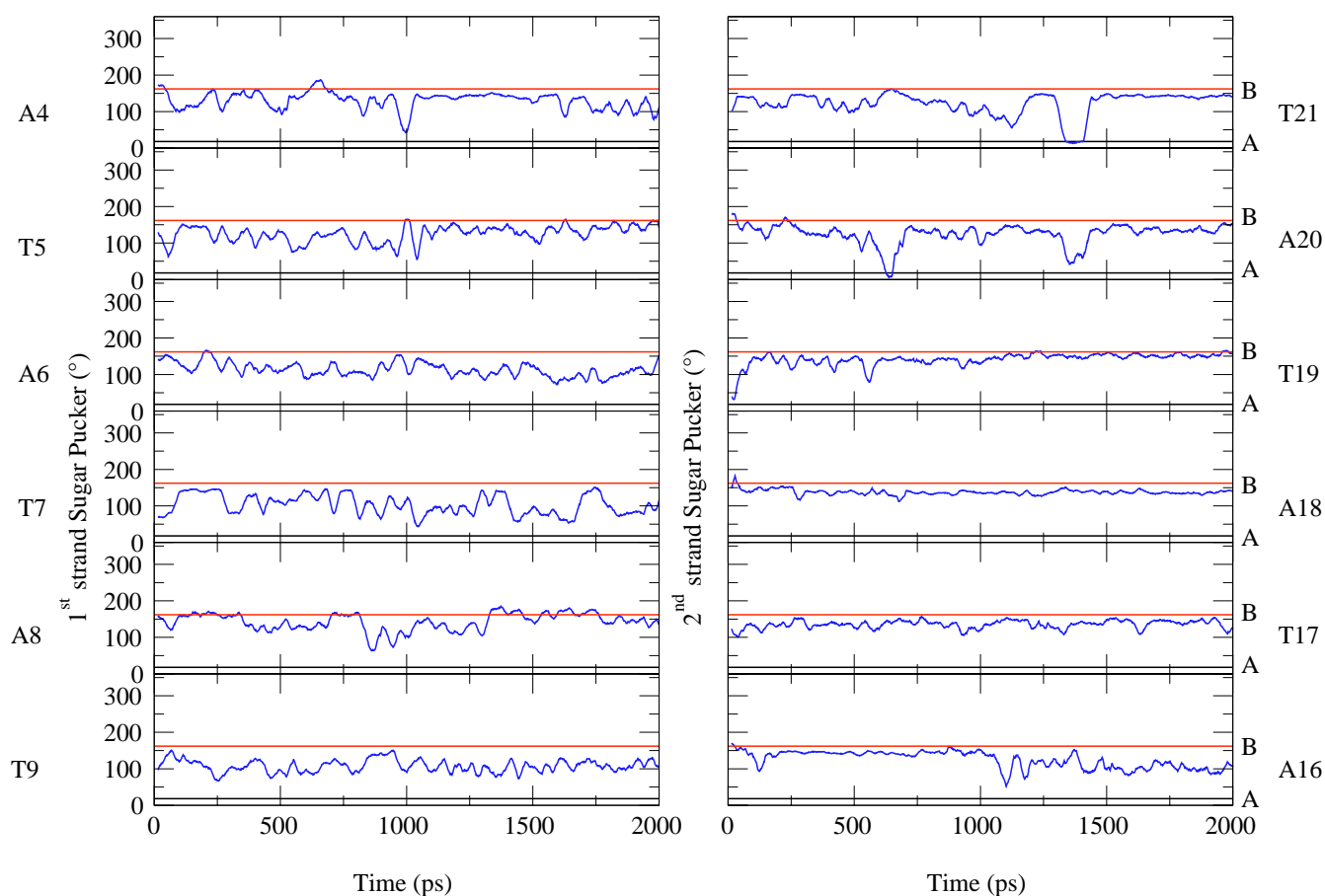
Figure 4 *Evolution with time of the sugar pucker angles, P, in the AT tract. Values for canonical A- (C3'-endo) and B-DNA (C2'-endo) are shown by horizontal lines*

tions with this potential.[7, 11, 29, 30] Hence the B → A → B transitions are seen in the twist if this and the lowered twist due to solvation are taken into account.

Figure 4 shows the sugar pucker pseudorotation, P, adopted in the AT tract during the simulation. Although a certain amount of A-like (C3'-endo) puckering is adopted by some of the sugars during the trajectory, it is not obvious that concerted transitions seen in the base movements are correlated with sugar re-puckering during the simulation. For the majority of the trajectory B-like (C2'-endo) and intermediate (O4'-endo) sugar puckers are adopted.

Similar results, including B → A → B transitions, were obtained with initial canonical A and B coordinates, showing that these observations are not an artefact of the initial conditions. The three simulations converge to similar average properties in agreement with previous work.[11]

Discussion

We believe that this is the first observation of both B → A and A → B transitions in low salt DNA solution with the

Cornell et al. [19] potential. This appears to show that previous concerns [11] of overstabilization of the B-form are unfounded. In addition, it indicates that the occurrence of conformational transitions is dependent on base sequence.

It is noteworthy that the change from B to A values in the bases is not correlated with analogous changes in the sugar puckers. In a recent MD study [30] one dihedral angle (C1'C2'C3'C4') in each sugar was driven from B to A values, which led to a B → A transition throughout the molecule, and it was suggested that sugar puckering is a major determinant of DNA conformation. But our results show that the B → A transition - at least insofar as the bases are concerned - is not driven by sugar repuckering.

This work suggests that DNA containing alternating AT tracts, of the kind found in TBP binding regions, may be readily susceptible to a B → A transition. In this regard it supports another simulation with a similar sequence,[21] although in that case the evidence for a transition, being based only on a gradual change in xdisp and some sugar repuckering, was much weaker than the concerted transitions described here. Such susceptibility would be an attractive property since TBP binding is associated with deformation of the DNA from B- to a TA-form.[5, 6] In addition this base sequence appears

to be finely balanced on a knife-edge between the A- and B-forms; such dynamic fluctuations may provide a recognition motif for TBPs.

We are currently investigating the robustness of the results presented here with regard to extended simulation times, high salt conditions, base sequence and new AMBER parameters.[31]

Supplementary material available Coordinates of the DNA at times shown in figure 3 have been provided in PDB format.

References

1. Saenger, W. Principles of nucleic acid structure. Springer, New York, 1984.
2. Franklin, R. E.; Gosling, R. G. *Acta Cryst.* **1953**, *6*, 673.
3. Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratcliff, R. L. *J. Mol. Biol.* **1980**, *143*, 49-72.
4. Mahendrasingam, A.; Forsyth, V. T.; Hussain, R.; Greenall, R. J.; Pigram, W. J.; Fuller, W. *Science* **1986**, *233*, 195-197.
5. Bewley, C. A.; Gronenborn, A. M.; Clore, G. M. *Annu. Rev. Biophys. Biolmol. Str.* **1998**, *27*, 105-131.
6. Guzikevich-Guerstein, G.; Shakked, Z. *Nature Struct. Biol.* **1996**, *3*, 32-37.
7. Cheatham, T. E., III.; Miller, J. L.; Fox, T.; Darden, T.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193-4194.
8. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
9. York, D. M.; Darden, T. A.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345-8348.
10. York, D. M.; Yang, W.; Lee, H.; Darden, T. A.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001-5002.
11. Cheatham, T. E., III.; Kollman, P. *J. Mol. Biol.* **1996**, *259*, 434-444.
12. Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Phys. Chem. B* **1997**, *102*, 4658-4667.
13. MacKerell, A. D.; Wiorkiewiczkuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946-11975.
14. Yang, L.; Pettitt, B. M. *J. Phys. Chem. B* **1996**, *100*, 2564-2566.
15. Feig, M.; Pettitt, B. M. *J. Phys. Chem. B* **1997**, *101*, 7361-7363.
16. Feig, M.; Pettitt, B. M. *Biophys. J.* **1998**, *75*, 134-149.
17. Floppe, N.; MacKerell, A. D. *J. Comput. Chem.* **2000**, *21*, 86-104.
18. MacKerell, A. D.; Banavali, N. K. *J. Comput. Chem.* **2000**, *21*, 105-120.
19. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M. J.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.
20. Cieplak, P.; Cheatham III, T. E.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 6722-6730.
21. Flatters, D.; Young, M.; Beveridge, D. L.; Lavery, R. *J. Biomol. Str. Dyn.* **1997**, *14*, 757-765.
22. Pearlman, D. A.; Case, D. A.; Cadwell, J. W.; Ross, W. S.; Cheatham, T. E., III.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. Comp. Phys. Comm. **1995**, *91*, 1-41.
23. Yoon, C.; Prive, G. G.; Goodsell, D. S.; Dickerson, R. E. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 6332-6336.
24. Klug, A.; Jack, A.; Viswamitra, M. A.; Kennard, O.; Shakked, Z.; Steitz, T. A. *J. Mol. Biol.* **1979**, *131*, 669-680.
25. Arnott, S.; Hukins, D. W. L. *Biochem. Biophys. Res. Comm.* **1972**, *47*, 1504-1509.
26. Berendsen, H. J. C.; Postma, J. P. M.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684-3690.
27. Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids.* Oxford University: New York, 1987.
28. Rhodes, D.; Klug, A. *Nature* **1980**, *286*, 573-578.
29. Duan, Y.; Wilkosz, P.; Crowley, M.; Rosenberg, J. M. *J. Chem. Soc. Faraday Trans.* **1994**, *90*, 1415-1427.
30. Soliva, R.; Luque, F. J.; Alhambra, C.; Orozco, M. *J. Biomol. Str. Dyn.* **1999**, *17*, 89-99.
31. Cheatham, T. E., III.; Cieplak, P.; Kollman, P. A. *J. Biomol. Str. Dyn.* **1999**, *16*, 845-862.