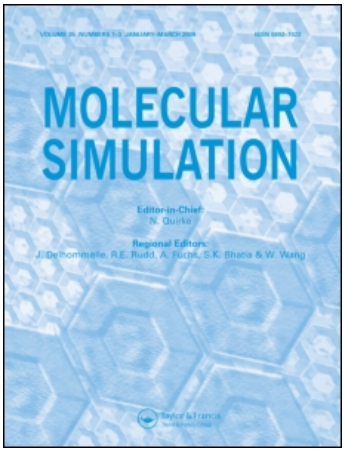


Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Molecular Dynamics Simulations of DNA Triple-Helices. Does The Triplex $d(A)_{10}\cdot d(T)_{10}\cdot d(T)_{10}$ Have A-Form or B-form Geometry?

C. A. Laughton[a]
[a] CRC Biomolecular Structure Unit, Institute of Cancer Research, U.K.

To cite this Article Laughton, C. A.(1995) 'Molecular Dynamics Simulations of DNA Triple-Helices. Does The Triplex $d(A)_{10}\cdot d(T)_{10}\cdot d(T)_{10}$ Have A-Form or B-form Geometry?', Molecular Simulation, 14: 4, 275 — 289
To link to this Article: DOI: 10.1080/08927029508022023
URL: http://dx.doi.org/10.1080/08927029508022023

Molecular Simulation, 1995, Vol. 14, pp. 275–289
Reprints available directly from the publisher
Photocopying permitted by license only

MOLECULAR DYNAMICS SIMULATIONS OF DNA TRIPLE-HELICES. DOES THE TRIPLEX $d(A)_{10}\cdot d(T)_{10}\cdot d(T)_{10}$ HAVE A-FORM OR B-FORM GEOMETRY?

C.A. LAUGHTON

CRC Biomolecular Structure Unit, Institute of Cancer Research, Sutton, Surrey SM2 5NG, U.K.

(Received September 1993, accepted July 1994)

In view of the current paucity of experimental structural data, the molecular modelling of DNA triple helices plays a particularly important role in helping to understand more clearly their structure and the factors contributing to their stability. At the same time, it makes it difficult to assess the "realism" of the simulations. We have used molecular modelling, and particularly molecular dynamics, to study both (polypurine)·(polypyrimidine)·(polypyrimidine) and (polypyrimidine)·(polypurine)·(polypurine) triplexes. The importance of treating solvation and electrostatic interactions in as accurate a manner as possible is clear, resulting in the need for large scale molecular simulations.

Recently, there has been discussion as to whether the (polypurine)·(polypyrimidine)·(polypyrimidine)-type triplexes have conformational features (particularly sugar puckers) resembling A-DNA, as in the original fibre-diffraction-derived model, or more like B-DNA. Reported here are the results of two 160ps molecular dynamics simulations of the triplex $d(A)_{10}.d(T)_{10}.d(T)_{10}$, performed with explicit solvent under periodic boundary conditions, starting from both A- and B-type structures. Both simulations show similar behaviour, with partial unwinding of the helices (reduction in twist and increase in rise), and a preponderance of O1'-endo sugar puckers.

KEY WORDS: DNA triple helix; DNA sugar pucker; molecular dynamics

1 INTRODUCTION

Although it has been known for 36 years that DNA can form triple-stranded structures *in vitro* [1], it is only much more recently that there has been resurgence in interest in them. Firstly there has been the realisation that they may have a role *in vivo* [2], and secondly there has been the realisation that the base-specificity of triplex formation provides a powerful route to sequence-specific DNA recognition and targeting [3, 4, 5]. The first model for the structure of the triple helix was derived from fibre-diffraction data on $d(T)_n.d(A)_n.d(T)_n$ [6, 7]. It is generally compared to the structure of A-type DNA duplexes, in that the sugars have C3'-endo puckers, the bases are inclined to the helix axis and the helix has 12 residues per turn (Figure 1). With the recent upturn in interest in triple helices have come further structural studies, however as yet there has been no report of the determination of the crystal structure of any DNA triplex. Many of these studies have thrown some doubt on the fibre-diffraction-derived model, and in particular on the sugar puckers where a C2'-endo, B-DNA-like conformation has seemed in better accord with the data [8, 9]. As a result, a new model for

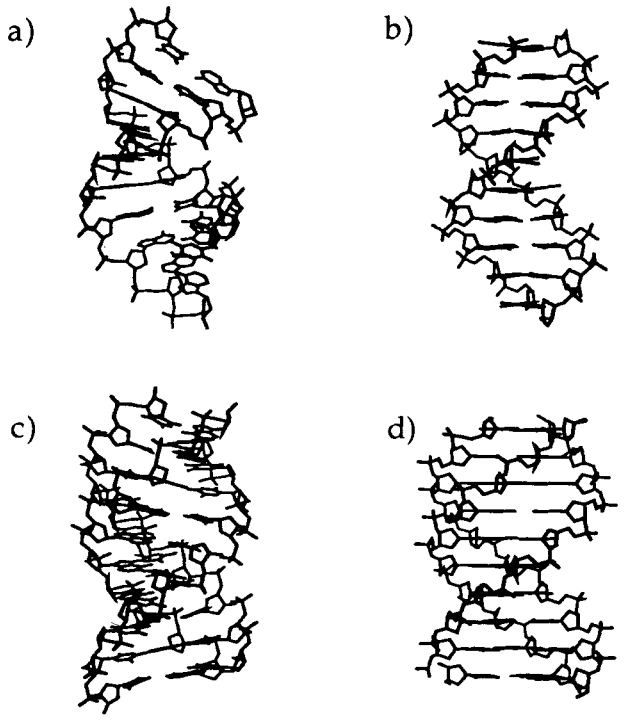


Figure 1 Comparison of the structures of a) an A-form duplex, b) a B-form duplex, c) an A-form triplex, and d) a B-form triplex.

$d(T)_n.d(A)_n.d(T)_n$ has recently been proposed in which all three strands have a C2′-endo sugar pucker and the bases lie perpendicular to the helix axis [10, 11]. In view of these characteristics it has been described as having a B-type geometry.

We have been interested in using molecular modelling, and particularly molecular dynamics, to study the structure and flexibility of DNA triple helices. In order to speed up computations, it is common to perform molecular dynamics simulations without an explicit consideration of the solvent but to approximate its effect on the solute through a modification of the electrostatic interaction term. This modification can take the form of a reduction in the partial charges on the solute atoms [12], and/or the use of distance-dependent dielectric constant [13]. However, we have found that such an approximation is not safe when applied to such highly-charged systems as DNA triple helices. This was exemplified in a study of the (polypyrimidine)·(polypurine)·(polypurine) type of triplex [14], where molecular modelling using a distance-dependent dielectric constant was unable to discriminate on energetic grounds between two models for the structure of this type of triplex, with *syn-* or *anti-* conformations for the third strand purines. In contrast, inclusion of explicit solvent revealed the *anti-* structure to be favoured, in agreement with all the experimental evidence to date.

We have previously performed short (30ps) molecular dynamics simulations of the triplex $d(TC)_5.d(GA)_5.d(C^+T)_5$ and related sequences with explicit consideration of solvent and under periodic boundary conditions [15]. Staring from the A-type

structure, we observed a tendency for sugar puckers to increase to O1′-endo or even C2′-endo values, particularly in the purine strand. Similar results have been obtained by other workers with other systems [16]. We hypothesised that the driving force behind this change in sugar pucker might be a reorientation of the phosphate groups to improve their solvation. Here are presented the results of longer (160ps) simulations on the triplex $d(A)_{10}.d(T)_{10}.d(T)_{10}$, explicitly solvated under periodic boundary conditions, starting from both the A-type structure studied previously, and also from the recently described B-type model with C2′-endo sugar puckers.

METHODS

General

All simulations were performed on a Hewlett Packard HP720 workstation using the AMBER 4.0 suite of programs [17] and the standard all-atom parameter set [18]. Visualization and interactive modelling were performed using the programs Midas-Plus [19] and MD Display [20] running on Silicon Graphics Indigo workstations. Helical parameters were calculated using the programs Newhelix92 [21] and Curves 4.1 [22].

Construction of the System

In order that the two simulations could be readily compared, it was desirable that they should begin from as similar a position as possible, and in particular that the counterion distributions were initially similar and both solutes were solvated by the same number of water molecules. To achieve this, structures of the $d(A)_{10}.d(T)_{10}.d(T)_{10}$ triplexes with both A-type and B-type geometries were generated from the helical coordinates given in the appropriate references [7, 11], and then a triplex with intermediate geometry was generated by coordinate-averaging the two structures. The base labelling scheme is shown in Figure 2. It was this average structure which was subsequently used to build the full simulation system. The system was made electrically neutral by the addition of sodium counterions, placed according to the procedure previously described [15]. Finally the solute was placed in a box of Monte Carlo TIP3P waters so that there was a minimum of 8 Å between the edge of the box and any atom of the solute. The final system contained 984 solute atoms (including 27 counterions) and 3595 water molecules in a box 54.7 Å × 49.5 Å × 48.3 Å.

Minimisation and Dynamics Protocol

All energy minimisations and dynamics were performed using an 8.5 Å non-bonded cutoff and a pair list that was updated every 20 steps.

Minimisations were performed under constant volume conditions to an rms gradient 0.1 kcal/mol/Å. Molecular dynamics were performed under constant pressure condi-

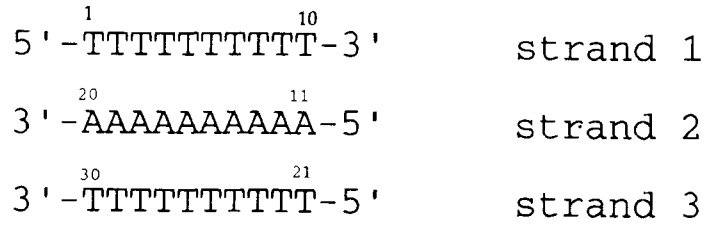


Figure 2 Strand definitions, strand sense and base labelling for the $d(A)_{10}\cdot d(T)_{10}\cdot d(T)_{10}$ triplex.

tions, with coupling to a heat bath and with restraints to maintain the base hydrogen-bonding pattern. The first stage consisted of a minimisation of the whole system, during which the DNA was converted from its intermediate structure to either the A- or B-type geometries by application of position restraints. The second stage consisted of 5 ps of molecular dynamics on the water molecules at 100 K, to help remove voids in the water structure that can result from the system building procedure. The third stage consisted of a re-minimisation of the waters, and the fourth a minimisation of the whole system, without any restraints. At this point the molecular dynamics simulations proper could begin. The protocol consisted of a 10 ps heating phase, in which the temperature was raised linearly to 300 K, followed by 150 ps of constant temperature dynamics, during which snapshots of the structure were saved every 0.2 ps. Each 160 ps simulation required about 272 CPU hours.

Helix Analyses

Helical analyses were performed on structures generated by coordinate-averaging snapshots from successive 10 ps periods, i.e. 0–10 ps, 10–20 ps, etc. The coordinate averaging was done after least-squares fitting over all DNA atoms.

RESULTS AND DISCUSSION

Rms Deviations

The rms deviations of the snapshots for each simulation from the initial A- and B-type models models are shown in Figure 3. For the simulation begun from the A-type model (simulation A, Figure 3a), the rms deviations indicate an initial phase of structural reorganisation of about 60 ps with the rms deviation rising to about 3 Å, followed by a relatively stable phase of about the same length, and finally a phase in which the structure deviates yet further from both initial models to a final rms deviation of about 4 Å. For the simulation begun from the B-type model (simulation B, Figure 3b) a similar initial phase of 60 ps is observed, but the analysis of the rest of the simulation depends rather on which initial structure it is referred to. Relative to its own initial structure (B-type) a short stable phase of about 30–40 ps is followed by a final stage in which the rms deviation rises towards 4 Å, while relative to the A-type structure the rms deviation oscillates between about 3 and 3.5 Å for the remainder of the simulation. The initial

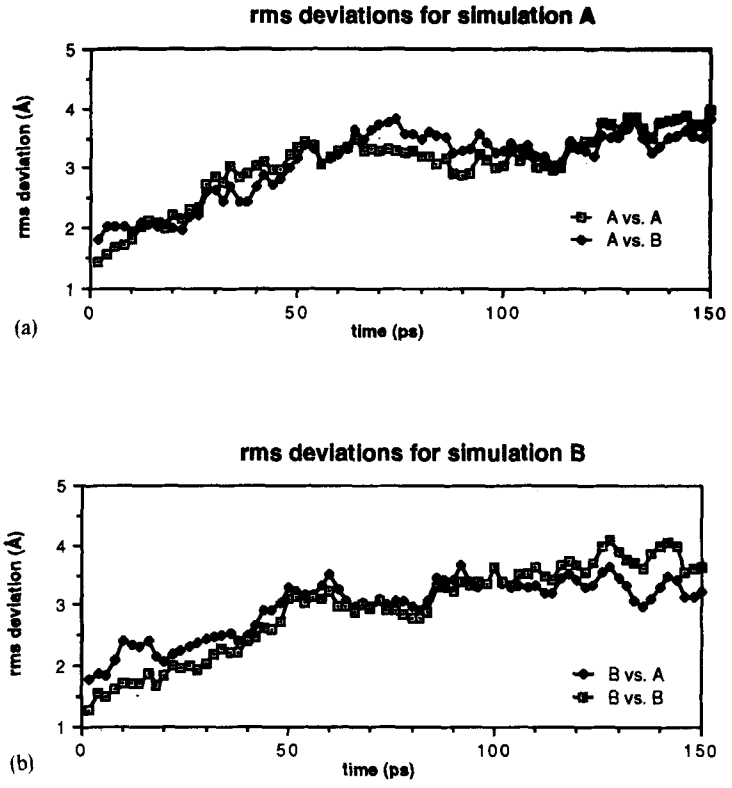


Figure 3 Root-mean-square deviations (measured over all DNA atoms) relative to initial A- and B-form models for a) simulation A (begun from the A-form structure), and b) simulation B (begun from the B-form structure).

conclusion, then, is that neither of the model structures, A-type or B-type, provides a good representation of the conformation of this triplex in solution.

Helical Changes

From a graphical analysis of the trajectories, it was clear that one of the gross features of both simulations was a partial unwinding of the helix. This is demonstrated in Figure 4, where it can be seen that this unwinding was manifested in both a reduction in helical twist and an increase in rise. Most of this stretching and unwinding of the triplex took place during the first 60–70 ps of the simulations, i.e. it correlates withthe initial phase that was characterised by rapidly increasing rms deviations from the starting structures. Similar unwinding has also been observed in MD simulations of hydrated DNA duplexes [23], though not in all cases [24]. The nmr-derived structure of a triplex containing a G.TA triple [8] also shows increased rise, though the change is more modest (to 3.44 Å) than that observed in these simulations, and is not accompanied by any reduction in twist.

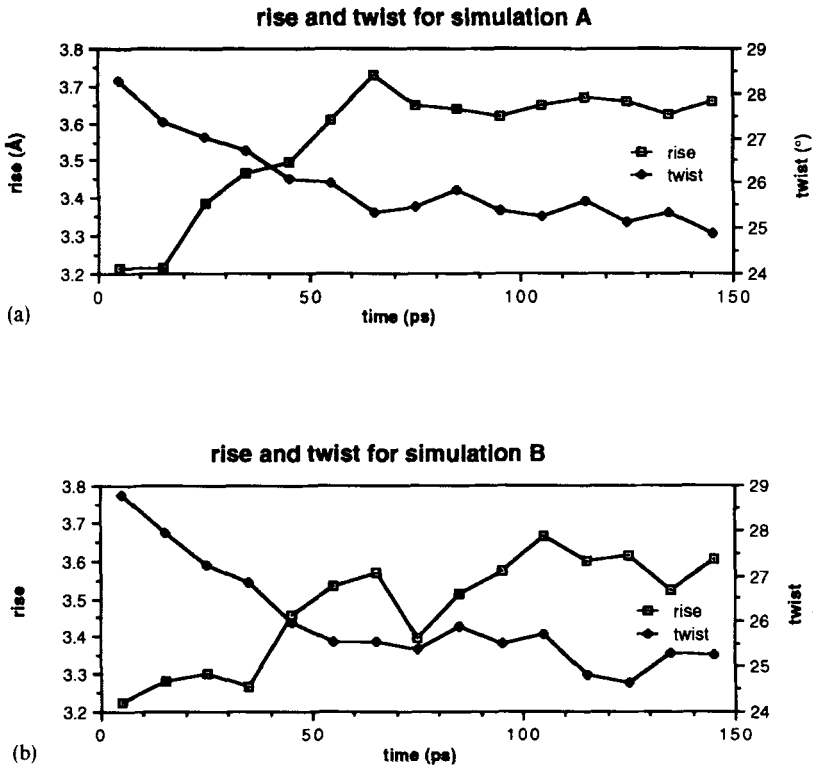


Figure 4 Average helical rise and twist for consecutive 10 ps-averaged triplex structures taken from a) simulation A, and b) simulation B.

DNA Bending

Bending of the triple helices is also observed during the course of the simulation. As a rough measure of this, the angle between the first and last helix axis segments calculated using Curves is plotted in Figure 5. It can be seen that for simulation A bending is at a maximum (36°) 80–90 ps into the simulation, and from then on to the end decreases almost to zero. For simulation B the angle increases unsteadily throughout the simulation, reaching 39° just before the end. The most bent 10 ps- averaged structures from each simulation could be overlaid with an all-DNA-atom rms deviation of just 1.52 Å, indicating very similar bending modes. For comparison, the rms deviation between the two final 10 ps-averaged structures, one of which (from simulation B) is bent and the other not, is 2.60 Å, and the rms deviation between the two initial models is 2.43 Å.

Base Steps with Anomalous Rise

A further major structural alteration was observed in both simulations, namely a major increase in rise at one base-pair step-in effect the spontaneous creation of a potential

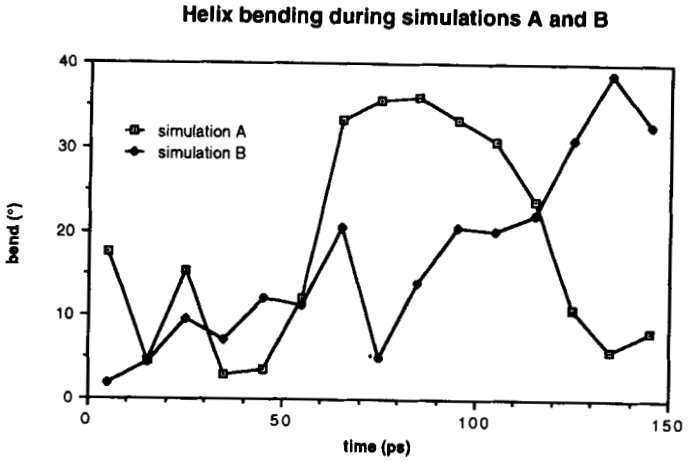


Figure 5 DNA bending (angle between first and last helix axis segments calculated using Curves) for consecutive 10 ps-averaged triplex structures taken from simulations A and B.

intercalation site (Figure 6). For simulation A this took place 40–50 ps into the simulation and occurred between the fifth and sixth base-triples; for simulation B it occurred 100–110 ps into the simulation, between the seventh and eighth triples. In both cases the rise at this step is increased to 7.5–8.0 Å. Examination of molecular

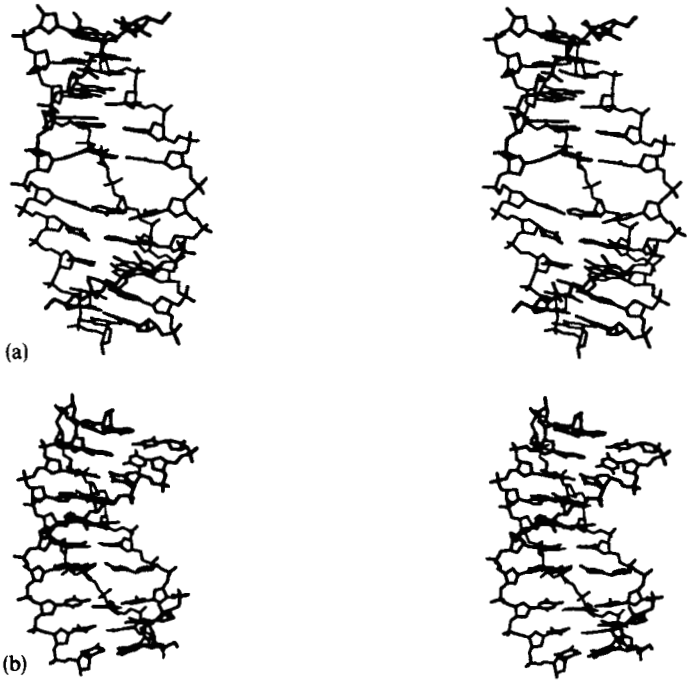
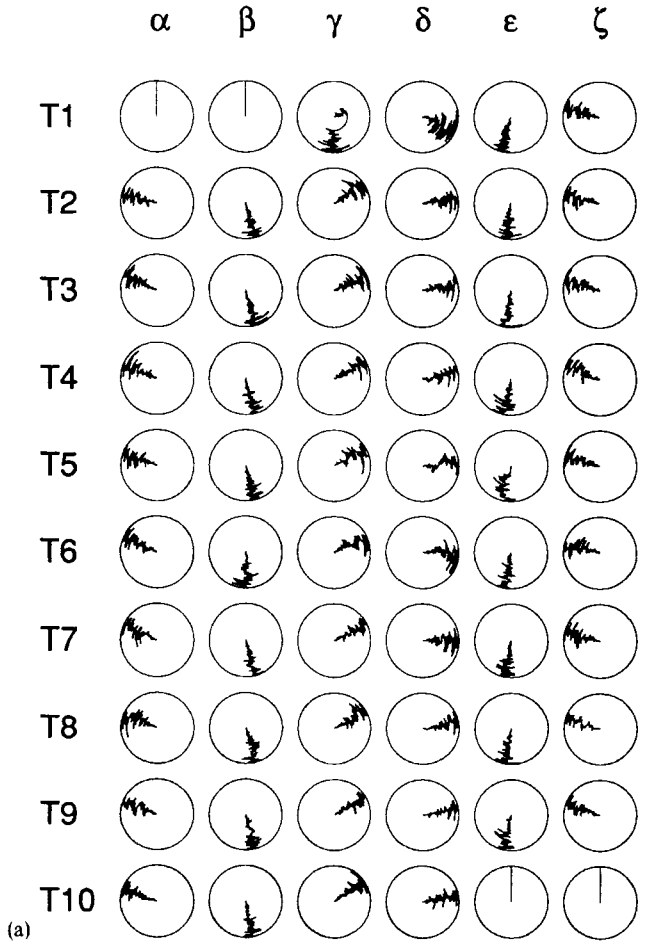


Figure 6 Structures of the $d(A)_{10}\cdot d(T)_{10}\cdot d(T)_{10}$ triplex at the end of a) simulation A, and b) simulation B. For clarity hydrogen atoms and counterions have been omitted.

dynamics snapshots reveals that the space between the base-pair planes remains essentially a void – there is little or no penetration of water molecules.

Backbone Torsion Angles

In Figure 7 and 8 the fluctuations in individual backbone torsion angles over the course of the simulation are presented as "dials" plots. As was found in our previous simulation, a common structural transition involves correlated changes in alpha and gamma angles from *gauche*$^-$, *gauche*$^+$ to *trans*, *trans*. In the case of simulation A, there are no such transitions in strand 1 (the Watson-Crick hydrogen bonded pyrimidine



(a)

Figure 7 "Dial" plots of the backbone torsion angles over the constant temperature phase (150 ps) of simulation A for a) strand 1, b) strand 2, and c) strand 3. For each dial 0 ps is at the centre and 150 ps at the circumference; 0° is at 12 o'clock and positive angles are clockwise.

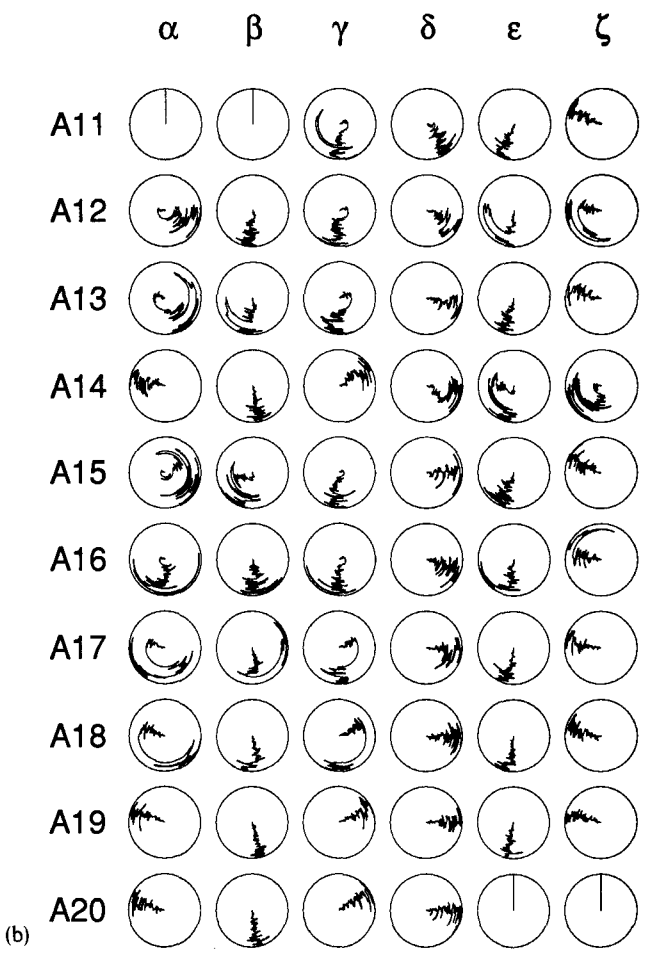
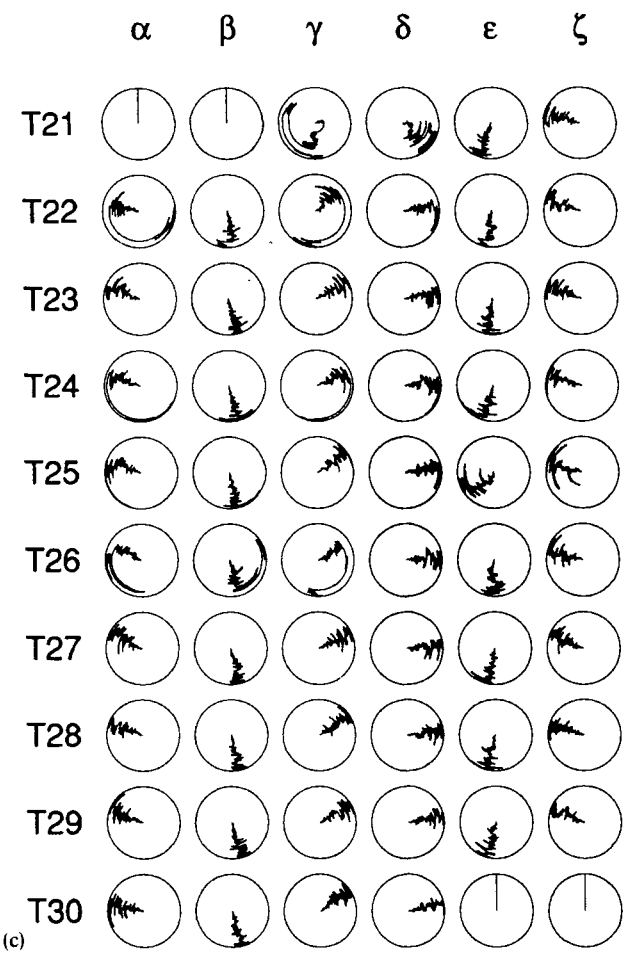


Figure 7 (*Continued*)

strand), but in strand 2 (the purine strand) they occur at 20 ps (A15), 26 ps (A16), 34 ps (A 12), 50 ps(A 13), 80 ps(A 17) and 112 ps(A 18). In strand 3 they occur at 124 ps (T22 and T26) and at 140 ps (T24). For simulation B there are a number of very early transitions; at A17 at the very start of the simulation, and at T2 and T4 after just 4 ps. Further transitions occur in strand 2 at 16 ps (A18), 52 ps (A15), 70 (A14) and 82 ps (A 12), and in strand 3 at 34 ps (T24), 38 ps (T27) and 46 ps (T25). Referring to the rms deviation plots, it is observed that there is no apparent correlation between these transitions and sudden changes in the structure, but it is clear that the transition does lead to increased flexibility in the backbone, as evidenced by the large fluctuations in alpha after the transition. Similar transitions have been noted in duplex DNA simulations [24].



(c)

Figure 7 (*Continued*)

Sugar Puckers

As mentioned above a great deal of the interest in the structure of triplexes has been focused on the question of sugar pucker. In Table 1 are given the distribution of the sugar puckers from the last 100 ps of each simulation. Both simulations give very similar results. For all strands the most commonly observed sugar pucker is O1′-endo, i.e. intermediate between A- and B-form values. C2′-endo sugar puckers are most commonly observed in strand 2 (> 20%) and C3′-endo puckers in strand 1. These results are still not in full accordance with some of the latest nmr studies, which indicated a high proposition of S-type (near C2′-endo) sugar puckers in all three strands [9]. However that analysis was based on a model for sugar pucker behaviour involving a fast, two-state equilibrium between C2′-endo and C3′-endo conformations, whereas

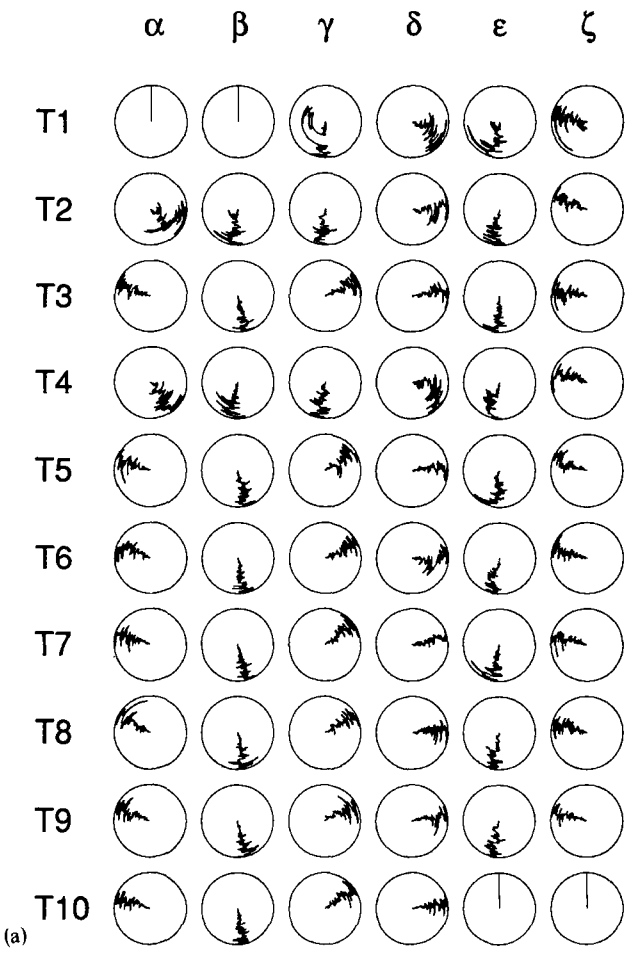


Figure 8 "Dial" plots from simulation B, for a) strand 1, b) strand 2, and c) strand 3.

these simulations indicate the existence, and high population, of the intermediate O1′-endo conformation.

CONCLUSIONS

Despite starting from quite different structures (rms deviation 2.43 Å) both simulations result in very similar trajectories. The motion of the triplexes is characterised by two large-scale components: i) bending of the helix, which is reversible, and ii) stretching and unwinding of the helix, which is irreversible. Overlaying 10 ps-averaged structures from the two simulations at points where they are similarly bent can be done with an

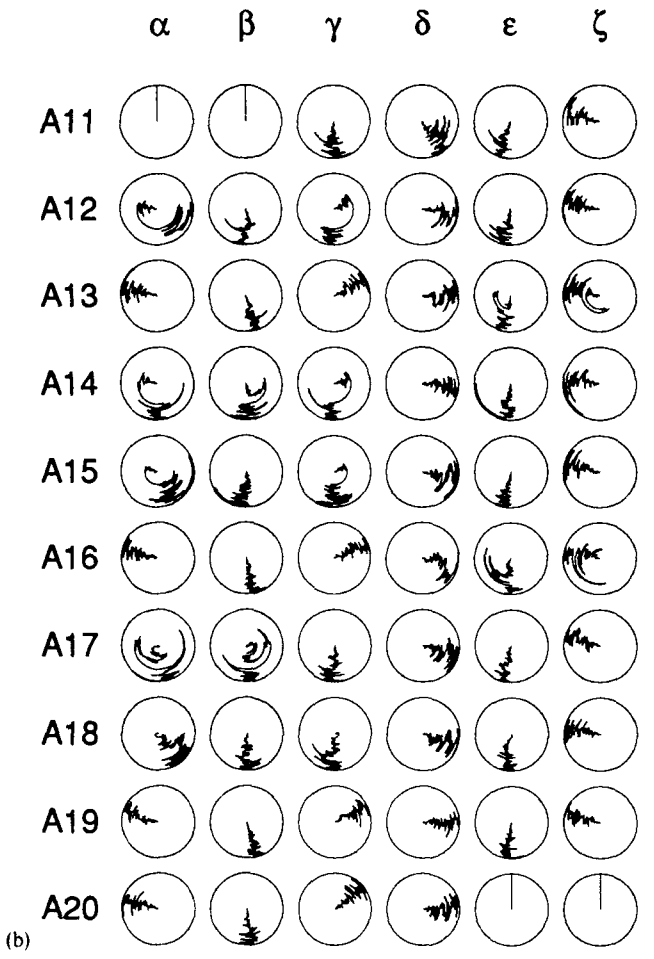


Figure 8 (*Continued*)

all-DNA-atom rms deviation of just 1.52 Å; similarly straight structures can be overlaid with a rms deviation of 1.60 Å. Thus the simulations are tending to converge, but to a structure which is about equidistant from, but not close to, either of the initial models. It would be nice to produce a new model for triplex structure, in terms of a regular, helically-repeating, unit; on the basis of these simulations. However this is hampered by the heterogeneity in the backbone conformations, particularly for the purine strand. Whilst there is some evidence for the occasional occurrence of the *trans*, *trans* alpha/gamma conformation from nmr studies on triplexes [25], and it is occasionally observed in duplex crystal structures (both A- and B-form), there must be some doubt that it is as common as these simulations would lead one to expect. It may be that these transitions are, at least in part, the result of beginning the simulations from conformations which are quite far from the apparent equilibrium geometry. It will be interesting

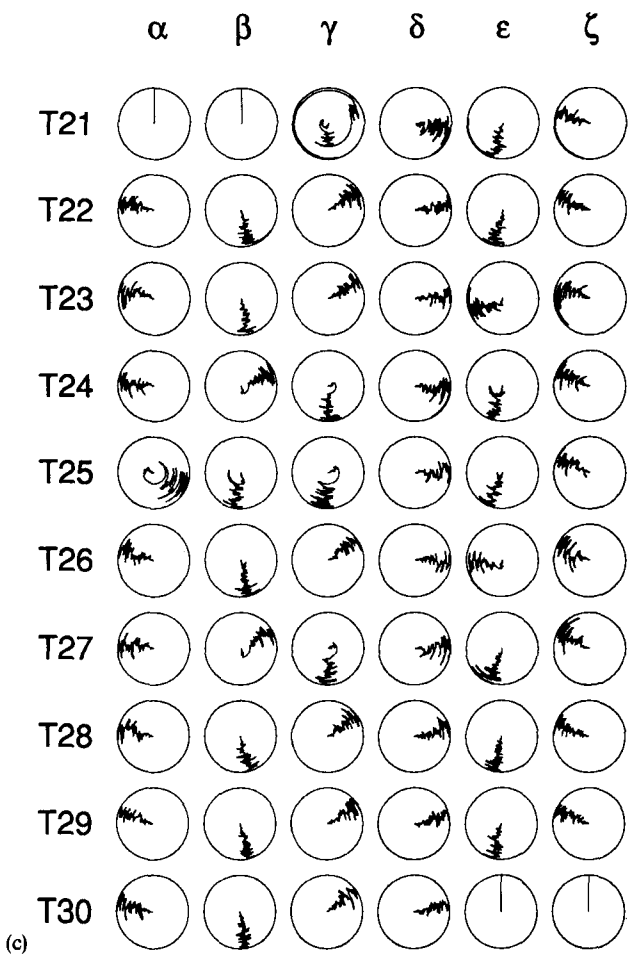


Figure 8 (*Continued*)

Table 1 Relative populations of different sugar-pucker states during the last 100 ps of both simulations, taken over all bases in each strand.

Simulation	*Strand*	$\langle P \rangle$[a] (°)	*C3'-endo* (%)	*O1'-endo* (%)	*C2'-endo* (%)
A	1	77.6	19.2	74.8	5.9
A	2	103.8	7.7	64.8	27.3
A	3	82.5	12.4	79.8	7.8
B	1	83.4	16.4	72.8	10.8
B	2	103.8	4.2	73.8	21.9
B	3	82.1	13.8	80.1	6.0

[a] Average pseudorotation phase angle.

to see if simulations begun from structures with modified values for helical rise, twist and sugar pucker show similar dynamical behaviour.

Acknowledgements

I thank Professor Stephen Neidle for helpful discussions. This work was supported by the Cancer Research Campaign of Great Britain.

References

[1] G. Felsenfeld, D.R. Davies and A. Rich, "Formation of a Three-Stranded Polynucleotide Molecule", *J. Amer. Chem. Soc.*, **79**, 2023 (1957).
[2] M.D. Frank-Kaminetskii, in *DNA Toplogy and its Biological Effects*, N.R. Cozzarelli and J.C. Wang eds. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. 1990, pp 185–215.
[3] C. Helene, "The anti-gene strategy: control of gene expression by triplex-forming-oligonucleotides", *Anticancer Drug Des.* **6**, 569 (1991).
[4] S.A. Strobel and P.B. Dervan, "Single-site enzymatic cleavage of yeast genomic DNA mediated by triple helix formation", *Nature*, **350**, 172 (1991).
[5] F.M. Orson, D.W. Thomas, W.M. McShan, D.J. Kessler and M.E. Hogan, "Oligonucleotide inhibition of IL2R alpha mRNA transcription by promoter region collinear triplex formation in lymphocytes", *Nucl. Acids Res*, **19**, 3435 (1991).
[6] S. Arnott and E. Selsing, "Structures of the polynucleotide complexes poly(dA)·poly(dT) and poly(dT)·poly(dA)·poly(dT)", *J. Mol. Biol.*, **88**, 509 (1974).
[7] S. Arnott, P.J. Bond, E. Selsing and P.J.C. Smith, "Models of triple-stranded polynucleotides with optimised stereochemistry.", *Nucl. Acids Res.*, **11**, 4141 (1976).
[8] I. Radhakrishnan and D.J. Patel, "Solution Conformation of a G·TA Triple in an Intramolecular Pyrimidine·Purine·Pyrimidine DNA Triplex.", *J. Amer. Chem. Soc.*, **144**, 6913 (1992).
[9] R.F. Macaya, P. Schultze and J. Feigon, "Sugar Conformations in Intramolecular DNA Triplexes Determined by Coupling Constants Obtained by Automated Simulation of P. COSY Cross Peaks.", *J. Amer. Chem. Soc.*, **114**, 781 (1992).
[10] F.B. Howard, H.T. Miles, K. Liu, J. Frazier, G. Raghunathan and V. Sasisekharan, "Structure of d(T)n·d(A)n·d(T)n: the DNA triple helix has B-form geometry with C2'-endo sugar pucker", *Biochemistry*, **31**, 10671 (1992).
[11] G. Raghunathan, H.T. Miles and V. Sasisekharan, "Symmetry and molecular structure of a DNA triple helix: d(T)n·d(A)n·d(T)n", *Biochemistry*, **32**, 455 (1993).
[12] B. Tidor, K.K. Irikura, B.R. Brooks and M. Karplus, "Dynamics of DNA Oligomers", *J. Biomol. Struct. Dyn.*, **1**, 231 (1983).
[13] M. Orozco, C.A. Laughton, P. Herzyk and S. Neidle, "Molecular-mechanics modelling of drug-DNA structures; the effects of differing dielectric treatment on helix parameters and comparison with a fully solvated structural model", *J. Biomol Struct Dyn.* **8**, 359 (1990).
[14] C.A. Laughton and S. Neidle, "Prediction of the structure of the $Y^+ \cdot R^- \cdot R^+$-type DNA triple helix by molecular modelling", *Nucl. Acids Res.*, **20**, 6535 (1992).
[15] C.A. Laughton and S. Neidle, "Molecular dynamics simulation of the DNA triplex $d(TC)_5 \cdot d(GA)_5 \cdot d(C^+T)_5$", *J. Mol. Biol.* **223**, 519 (1992).
[16] F.H. Hausheer, U.C. Singh, J.D. Saxe, O.M. Colvin and P.O. Tsu, "Can oligonucleoside methylphosphonates form a stable triplet with a double DNA helix?", *Anticancer Drug Des.* **5**, 159 (1990).
[17] D.A. Pearlman, D.A. Case, J.C. Caldwell, G.L. Seibel, U.C. Singh, P. Weiner and P.A. Kollman (1991) AMBER 4.0, University of California, San Francisco, USA.
[18] S.J. Weiner, P.A. Kollman, D.T. Nguyen and D.A. Case, "An All Atom Force Field for Simulations of Proteins and Nucleic Acids", *J. Comp. Chem.*, **7**, 230 (1986).
[19] T.E. Ferrin, C.C. Huang, L.E. Jarvis and R. Langridge, "The MIDAS display system", *J. Mol. Graphics*, **6**, 13 (1988).
[20] T. Callahan and T.P. Lybrand, MD Display (1991, QCPE program 623). Available from the Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405, USA.
[21] R.E. Dickerson. Newhelix92 (1992), available from the Protein Data Bank, Brookhaven National Laboratory, Upton, NY 11973, USA.
[22] R. Lavery and H. Sklenar, "The Definition of Generalized Helicoidal Parameters and of Axis Curvature for Irregular Nucleic Acids.", *J. Biomol. Struct. Dyn.*, **6**, 63 (1988).

[23] K. Miaskiewicz, R. Osman and H. Weinstein, "Molecular Dynamics Simulation of the Hydrated d(CGCGAATTCGCG)$_2$ Dodecamer", *J. Amer. Chem. Soc.*, **115**, 1526 (1993).
[24] S. Swaminathan, G. Ravishanker and D.L. Beveridge, "Molecular Dynamics of B-DNA Including Water and Counterions: A 140-ps Trajectory for d(CGCGAATTCGCG) Based on the GROMOS Force Field.", *J. Amer. Chem. Soc.*, **113**, 5027 (1991).
[25] M.M. Mooren, D.E. Pulleyblank, S.S. Wijmenga, M.J. Blommers and C.W. Hilbers, "Polypurine/polypyrimidine halfpins form a triple helix structure at low pH", *Nucl. Acids Res.* **18**, 6523 (1990).