Simulations of DNA Coiling around a Synthetic Supramolecular Cylinder That Binds in the DNA Major Groove

Syma Khalid,[a, c] Michael J. Hannon,[a, b] Alison Rodger,[a] and P. Mark Rodger*[a]

Abstract: In this work we present the results of a molecular simulation study of the interaction between a tetracationic bis iron(II) supramolecular cylinder, $[Fe_2(C_{25}H_{20}N_4)_3]^{4+}$, and DNA. This supramolecular cylinder has been shown to bind in the major groove of DNA and to induce dramatic coiling of the DNA. The simulations have been designed to elucidate the interactions that lead the cylinder to target the major groove and that drive the subsequent DNA conformational changes. Three sets of multi-nanosecond simulations have been performed: one of the uncomplexed d(CCCCCTTTTTCC)·d(GGAAAAAGGGGG) dodecamer; one of this DNA complexed with the cylinder molecule; and one of this DNA complexed with a neutralised version of the cylinder. Coiling of the DNA was observed in the DNA–cylinder simulations, giving insight into the molecular level nature of the supramolecular coiling observed experimentally. The cylinder charge was found not to be essential for the DNA coiling, which implies that the DNA response is moderated by the short range interactions that define the molecular shape. Cylinder charge did, however, affect the integrity of the DNA duplex, to the extent that, under some circumstances, the tetracationic cylinder induced defects in the DNA base pairing at locations adjacent to the cylinder binding site.

Keywords: bio-supramolecular chemistry · DNA recognition · helical structures · molecular dynamics · noncovalent interactions

Introduction

Control of gene expression is currently one of the key areas of interest in molecular medicine.[1–4] Gene expression involves the transfer of information encoded within the gene to produce a biologically active protein. However, not all genes are expressed in every cell all the time; bio-regulation of gene expression is effected by proteins that activate or repress transcription by binding to short, specific DNA sequences.[1,5] The ability to turn genes on or off artificially by the action of synthetic analogues of DNA-binding proteins is an important goal that would open up new possibilities for disease control and prevention as well as cure.

Proteins that bind DNA frequently achieve sequence-specific code recognition by binding non-covalently in the major groove of DNA.[6–8] In biological systems, the major groove is the preferred binding site for sequence recognition as this groove shows the greater variation in size and shape with base sequence and also the greater number and variation in pattern of hydrogen-bond donor and acceptor units; it is therefore the ideal target for synthetic molecules designed to recognise and bind to specific DNA sequences. However, relatively little progress had been made in producing synthetic major groove binders because of the size entailed. Most of the compounds synthesised are either minor groove binders,[9] intercalators,[10,11] or metal complexes that span only two to three base pairs and whose interaction with DNA is too limited for sequence selectivity.[12–14] Design of synthetic agents that target the major groove in a non-covalent, sequence-selective fashion remains a challenge of considerable importance in molecular medicine and biology.

[a] Dr. S. Khalid, Prof. M. J. Hannon, Prof. A. Rodger, Prof. P. M. Rodger
Department of Chemistry, University of Warwick
Coventry CV4 7AL (UK)
E-mail: p.m.rodger@warwick.ac.uk

[b] Prof. M. J. Hannon
Current address:
School of Chemistry, University of Birmingham
Birmingham B15 2TT (UK)

[c] Dr. S. Khalid
Current address:
Department of Biochemistry, University of Oxford
South Parks Rd, Oxford, OX1 3QU (UK)

Recently, a major step has been taken toward achieving the goal of designing synthetic agents to target the major groove: we have developed a novel compound that binds strongly in the major groove of DNA and is large enough to span five base pairs.[15,16] The compound is a metallo-supramolecular tetracationic cylinder (**1**) with a triple helical framework. This cylinder is approximately 19 Å in length and 11 Å in diameter. It is too big to bind in the minor groove of DNA, but has just the right shape and size to lie along the major groove.[15]
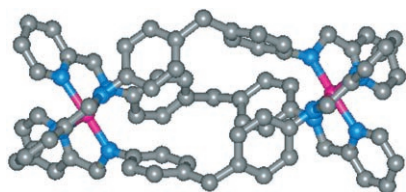

Figure 1. Tetracationic cylinder [Fe$_2$L$_3$]$^{4+}$, where L is *N,N'*-bis(pyridin-2-ylmethylene)-4,4'-diaminodiphenylmethane: Fe^{II} is pink, N is blue, C is grey; H atoms have been omitted for clarity.

Experimental techniques have provided information regarding the binding strength and preferred binding sites of **1** on DNA.[16] Spectroscopic studies indicate that the cylinder binds very strongly to DNA with a binding constant well in excess of 10^7M^{-1} in 20 mM salt. Flow linear dichroism (LD) reveals coiling or bending of the DNA on addition of the cylinder, and tapping mode atomic force microscopy (AFM)[16] reveals this to be dramatic intramolecular coiling of DNA that is unprecedented with a synthetic agent: **1** was found to induce DNA to bend by 40–60° per ligand,[15] with the smaller angles (40°) occurring at the maximum loading of one cylinder per DNA turn. In contrast cobalt ammines—which are known to cause DNA to bend and ultimately to condense into multi-molecular aggregates—generate just 2–5° bend per ligand.[17] The effect of cobalt ammines can largely be explained by the Manning charge condensation theory,[18] but the magnitude of the effect with **1** suggests the bending is of a very different character. The *M*-helical enantiomer of the cylinder coils DNA more aggressively than the *P* enantiomer.[16] NMR data have confirmed a major groove-binding mode for the *M* cylinder.[19] From all the experimental data it is clear that the cylinder binds in the major groove and is able to induce dramatic conformational changes in the DNA. However, gaining molecular level information about the effects and interactions which lead to this remarkable supramolecular event is challenging. Molecular dynamics (MD) simulations can provide information at the molecular level that is complementary to experiment and therefore are an ideal way to get a better understanding of this system.

There is comparatively little modelling literature on the interaction between transition-metal complexes and DNA. Some modelling has been performed on DNA with ruthenium(II)–tris(1,10-phenanthroline)[20] in vacuum, and the re-

sults compared with experimental data.[12] The Λ-ruthenium complex showed a preference for the major groove of DNA but little or no bending of the DNA at low loading and whereas the Δ enantiomer preferred the minor groove at low loadings where it did bend the DNA—but only by a few degrees—and bound to both grooves at higher loadings with no additional bending. Several other papers have used modelling studies to interpret experimental data on DNA/transition-metal complexes,[17,21] but since their emphasis was on understanding the experimental data, these calculations again omitted the solvent. Rigorous molecular dynamics simulations with explicit solvent have been performed for cobalt(III)–hexammine[22] and a nickel(II)–metalloprotein[23] in the presence of DNA. In both cases, the simulations produced good agreement with a range of experimental data; in neither case was there significant bending or coiling of DNA.

In this paper we present results of a molecular dynamics (MD) study of the effect of the supramolecular cylinder, **1**, on DNA. Docking calculations, both manual and using high-temperature MD with constrained or rigid molecules, have been used to identify favourable initial binding sites for the cylinder on a DNA dodecamer, and then multi-nanosecond MD simulations performed with explicit solvent to monitor the DNA response. Previously, MD simulations[19] were performed on this system to interpret the NMR spectra, but these used NOE data as additional force-field restraints. The simulations described in the present paper were designed to probe any changes in DNA structure upon cylinder binding and are not restrained in any way. As a consequence, these new simulations enable a direct analysis of the DNA response—both its nature and the forces that give rise to it.

Computational Methods

System: The DNA dodecamer sequence (CCCCCTTTTTCC)·d(GGAAAAAGGGGG) was chosen for this study. Preliminary simulations were performed with a decamer, d(CCCCCTTTTT)·d(AAAAAGGGGG), which was chosen since it might offer potential for identifying preference of A–T versus G–C tracts within a single series of simulations; this sequence was subsequently extended by adding two C–G pairs to the A–T end so as to minimise any tendency for the DNA ends to fray in the absence of the cylinder; note that no restraints were introduced into the force field to hinder end-fraying. The numbering used in this paper to identify specific bases and base pairs is defined below.

5'	C1	C2	C3	C4	C5	T6	T7	T8	T9	T10	C11	C12	3'
	|	|	|	|	|	|	|	|	|	|	|	|	
3'	G24	G23	G22	G21	G20	A19	A18	A17	A16	A15	G14	G13	5'
base pair no.	1	2	3	4	5	6	7	8	9	10	11	12	

The *M* enantiomer of the cylinder, **1** (denoted C^{4+}), was modelled using the CHARMM22 all-atom force field, but with the FeN$_6$ sub-unit treated as a rigid body with geometry taken from the crystal structure. For comparison, a simulation was also performed in which the *M* cylinder was

made electrically neutral, denoted C^0. Although the C^0 system has no experimental counterpart, a comparison of the C^{4+} and C^0 systems enables an exploration of the relative importance of molecular shape (as defined by the van der Waals interactions) and electrostatic interactions in inducing the DNA response. Analogous comparisons with netropsin[24,25] have proved useful in understanding the influence of charge in minor-groove binding ligands. Many methods could be used to achieve this neutrality, including setting all atomic charges to zero, neutralising the Fe atoms, or adding a neutralising negative charge to the organic part of the cylinder. In this work we have adopted the last option: $-0.03\,e$ was added to the charge on each C and H atom in the cylinder, resulting in a counterbalancing $-4\,e$ charge spread over the surface of the cylinder. While still giving rise to electric field gradients, the negative surface charge should considerably reduce the *electrostatic* attraction of the cylinder to the DNA phosphate backbone, and thus makes the comparison with C^{4+} instructive.

Technical details: Compounds were modelled using the CHARMM force-field series. This contains two main variants for modelling DNA: the older CHARMM22 force field,[26] and a more refined CHARMM27.[27] There have been a number of comparisons of how well these two force fields reproduce the behaviour of DNA,[28] and the general consensus is that the CHARMM27 force field is much better. In particular, it correctly predicts B-DNA to be the stable conformation in low ionic strength solvents at normal temperatures, whereas the CHARMM22 force field tends to cause the DNA to adopt an A-like form. However, it should be noted that most of these comparisons have treated just the duplex in aqueous solution: no comparisons have been reported for DNA interacting with a major-groove-binding ligand, particularly when that ligand carries a substantial positive charge. The experimental data discussed above indicates that C^{4+} perturbs the DNA strongly, bending it well away from the canonical B-form. This is outside the parameter-space that has so far been used to derive and validate the DNA force fields, and so it is of interest to compare the performance of these two force fields in modelling DNA/C^{4+} complexes.

All simulations were performed with DL_POLY.[29] The conversion from CHARMM to DL_POLY force-field formats was achieved by using a purpose-built program that interprets the CHARMM prm and crd files. As in previous work[30] checks on the force field were performed on numerous configurations to ensure the energies and forces calculated with DL_POLY and CHARMM agreed *exactly*. Water was modelled with the TIP3P[31] potential, and kept rigid using the SHAKE algorithm,[32] implemented in DL_POLY with a tolerance of 0.0001. All hydrogen atoms were assigned a mass of 2 u. This gave good energy conservation in constant energy simulations with a 2 fs time step for the CHARMM22 force field. An Ewald sum[33] was used to evaluate electrostatic interactions, with $\boldsymbol{k}_{max}$ = (5,5,6) and $\alpha = 0.12604\,\text{Å}^{-1}$. Simulations were performed with orthorhombic periodic boundaries at constant volume and temperature (NVT) using the Hoover thermostat[33] with a time constant of 0.1 ps and a time step of 2 fs; some 0.25 ns segments of the trajectories were repeated with a 1 fs time step and showed no significant differences. Some difficulties were encountered initiating CHARMM27 simulations with the 2 fs time step, and so this was reduced to 1 fs for all the CHARMM27 simulations reported herein. In all other respects, the CHARMM22 and CHARMM27 simulations followed the same protocols and methodology.

Protocol: The multi-nanosecond simulations were performed using a four-step protocol designed to identify good binding sites while still ensuring a disperse background ion distribution:

1) High temperature MD simulations were performed on the cylinder and ions moving around an immobilised B-DNA dodecamer, with the relative dielectric constant $\varepsilon_r = 80$;
2) low potential energy configurations were identified and optimised with respect to just the cylinder position; ions were kept fixed at their positions from the high temperature snapshot and the DNA also remained fixed;
3) water was added, the relative dielectric constant reset to 1, and the system allowed to relax while restraining the DNA with a series of progressively weaker harmonic tethering potentials;
4) full, unrestrained MD simulations were performed on the aqueous system for 2–5 ns.

This protocol is described more fully hereafter.

Docking (steps 1 and 2): A dodecamer of B-DNA and one cylinder molecule (*M* enantiomer) were embedded in a neutralising atmosphere of Na^+ ions. A further 58 Na^+ and 58 Cl^- ions were added in a 45 Å× 45 Å×60 Å box, with DNA aligned along the *z* (long) axis; this gave [NaCl] = 0.8 M, which is similar to the concentration used in some other MD studies.[34,35] The DNA was immobilized and an NVT MD simulation performed for 1 ns at 900 K; this temperature was found to be high enough to ensure that the ligand sampled the whole surface of the DNA efficiently, but without any significant deformation in its own shape. The temperature was also high enough to ensure that the background atmosphere of sodium and chloride ions behaved like a homogeneous gas, with no aggregation of the ions onto the DNA or cylinder. The conformations with the lowest configurational energy were extracted and energy-minimised with respect to the position of just the ligand. No change was allowed in the position of the DNA atoms, Na^+ ions or Cl^- ions during this optimisation of the cylinder within its DNA/ionic gas environment. The resulting configuration was then used to start fully solvated MD simulations as described below. Variants on this procedure were tried in which the DNA and/or Na^+ and Cl^- ions were harmonically restrained rather than frozen during the cylinder minimisation, but these produced essentially the same results when carried forward into fully solvated MD simulations. Manual docking calculations were also performed, but did not give more favourable docking sites than those identified using this high temperature MD docking protocol.

A number of alternative low energy conformations was examined and all found to involve major groove binding. Several of these were then carried forward into full solvated MD simulations following the protocol given below.[36] All such simulations exhibited a very similar DNA response to that reported herein—particularly with respect to the extent to which they bent the DNA, and the resultant stability of the DNA—and confirmed that the results of our simulations were not sensitive to variations in the initial binding site. We note that these repeat simulations also proved our results to be robust with respect to the initial background ion distribution, since each configuration also involved a very different, essentially random, initial arrangement of the sodium and chloride ions.

Molecular dynamics simulations (steps 3 and 4): Water molecules, taken from an equilibrated liquid water simulation, were added to the optimised configurations (DNA, cylinder, Na^+ and Cl^- positions) identified in the docking calculations. Any water molecule that overlapped with DNA, cylinder, Na^+ or Cl^- was removed, which left a total of 3720 water molecules in the final DNA/ligand system, or 3758 for the uncomplexed DNA. Equilibration followed a similar protocol to that used by other research groups.[23,37] The DNA atoms were tethered to their original positions with a harmonic force constant of 100 kcal mol^{-1} Å^{-2} and an NVT MD simulation performed for 10 ps at 310 K. A further five 10 ps simulations were then performed successively in which the tethering force constant was 50, 25, 10, 5 and 1 kcal mol^{-1} Å^{-2}, respectively. The tethering potentials were then removed completely and multi-nanosecond simulations performed, saving configurations every 1 ps for later analysis.

Up to four repeat simulations of 2–5 ns duration, using slightly different starting configurations, were performed for each ligand/DNA system to validate the reproducibility and timescale of the DNA response. In every case, the response of the DNA was found to be rapid and repeatable, with most changes occurring in the first 0.5 ns, and no noticeable changes occurring between 2 and 5 ns.[36] In this paper we therefore concentrate on the behaviour during the first 2 ns after binding.

Analysis methods: A number of different methods have been used in this work to monitor DNA structure. The time dependence of conformational parameters has been monitored using the Curves algorithm[38] implemented in MDToolchest.[39] This uses seven torsion angles (α–ζ, ϕ, χ) to describe the DNA backbone, while the intra-base-pair geometry is described by six *helicoidal* parameters: three displacements (shear,

stretch and stagger) and three angles (buckle, propeller and opening). This parameter set is now well established as a means of describing DNA conformation;[40,41] a full definition is supplied with the Supporting Information, though it is useful to note here that of the backbone angles, α and ζ refer to torsions about P–O bonds, δ and ϕ refer to the ribose ring, while χ is for the bond that links the backbone to the base.

In understanding the DNA conformation and flexibility, it is also useful to examine some of the parameters that give a more correlated view of the overall behaviour. In particular, it is useful to study the relationship between neighbouring base pairs. In this work these have been monitored using a distance (slide) and three angles (tilt, roll and twist; see Figure 9) as implemented in 3DNA.[42] In this case, the analysis has been applied to the average structure determined from a continuous 50 ps portion of the MD trajectory; this length of trajectory was found to be long enough to smooth out the instantaneous fluctuations in shape, but was still short with respect to the systematic relaxation induced by the presence of the cylinder, **1**. The 3DNA analysis of average structures has also been used to generate normal vector plots,[7,43] which are useful for identifying linear and bent regions within the DNA, and hence to estimate the degree to which the DNA is bent by the cylinder.

Results and Discussion

Response of the DNA to a bound cylinder

As will be discussed below, the simulations with the two DNA force fields (CHARMM22 and CHARMM27) gave broadly similar responses for the DNA/ligand complex, in that the binding site was consistent with the two models and both indicated the DNA to be bent by the cylinder. Some quantitative differences were observed between the two force fields, with CHARMM27 tending to give the more minimal response to the cylinder. Two examples (discussed in more detail below) are worthy note at this stage. The overall bend of the DNA was about 40° with CHARMM22 compared with 20° for CHARMM27, and the CHARMM22 DNA did respond differently to C^{4+} and C^0 whereas the CHARMM27 model was remarkably insensitive to this large change in ligand charge. In the following we focus first on the results for the CHARMM22 model: since the response is greater with this model it is likely to be easier to identify the underlying physics of the DNA/ligand interaction. This is then followed by a more detailed comparison of the force fields to place confidence limits on the interpretation of the simulations.

Charmm22: As described in the protocol, favourable binding sites for the cylinder on the DNA were identified and then explicit solvent molecules added. In all cases, binding was found to occur preferentially in the region of the A–T tract (base pairs 6–10). Using the position from the docking calculation as a starting point, water was added, the system equilibrated while the DNA conformation was restrained to the B-form, and then simulations allowed to proceed without restraints. We re-iterate that up to four such simulations were performed for each DNA/cylinder system.

Once the restraints were removed, the DNA was observed to respond rapidly to the presence of the cylinder. Major conformational changes were observed in the presence of both C^{4+} and C^0. In every trajectory these occurred within the first 0.5 ns, with the precise timing depending on the starting configuration. In all cases, no significant conformational changes were observed in the subsequent 4–5 ns, suggesting that our simulations were probing the final DNA state, or at least identifying a long-lived intermediate in the response process.

It is instructive to begin by comparing before and after images of the DNA conformation in the three different systems: uncomplexed, bound to C^{4+}, and bound to C^0. Although such snapshots provide only limited information about the range of DNA conformations, for the present study they prove to give a useful and visual overview of the more quantitative data presented below. Images of the initial (directly after docking) and relaxed (after 2 ns) configurations for DNA and cylinder are given in Figure 2. Initially, the DNA adopts a nearly linear B-form. The initial binding mode of the cylinder is in the major groove, lying symmetrically between the two strands of the DNA, and spanning base pairs 6–11 (T6–A19 to C11–G14). This is consistent with the published structure obtained from NMR (NOE) data.[15,16,19] This binding geometry was retained throughout the equilibration phases, and was not disrupted until the tethering potentials that restrained the DNA conformation were removed.

After 2 ns of the unrestrained simulations, the *uncomplexed* DNA retained its near linear B-form, albeit with some fraying at the first base pair. In contrast, the final configurations of the DNA/cylinder simulations suggest substantial curvature of the DNA for both charged and uncharged cylinders. As will be discussed below (see Figure 7), these snapshots do depict a bend of about 40° in the helix axis in the presence of the cylinders.

At the end of the simulations the charged cylinder, C^{4+}, remained in the major groove, but some distortion of the groove shape was evident. There was close association between the cylinder and base-pairs 5–9 (closest interatomic distances between the cylinder and each of these nucleic acids are 2.4–3.5 Å); thus the cylinder remained within the A-T tract. There was no evidence of hydrogen bonding between the cylinder and DNA, but this is not surprising since the cylinder has no conventional hydrogen-bond donor groups, and its potential acceptors are obscured by nonpolar carbons. It is particularly interesting to note that some breakdown of the Watson–Crick structure is apparent in the last three or four base pairs of the DNA (Figure 2, middle, bottom), that is, adjacent to the cylinder binding site. As is shown below, this is due to a mispairing between T9 and A15, which then leaves T10 and A16 unpaired.

Importantly, the simulations reproduce both the major groove binding characterized by NMR and the DNA coiling effects as observed in LD and AFM experiments. The implication of the simulations is that, at least for some sequences, this bending might be associated with alterations to the DNA base pairing. Such effects have also been observed in some DNA–protein complexes,[44,45] and can, indeed, be an important feature of the way that proteins process DNA.
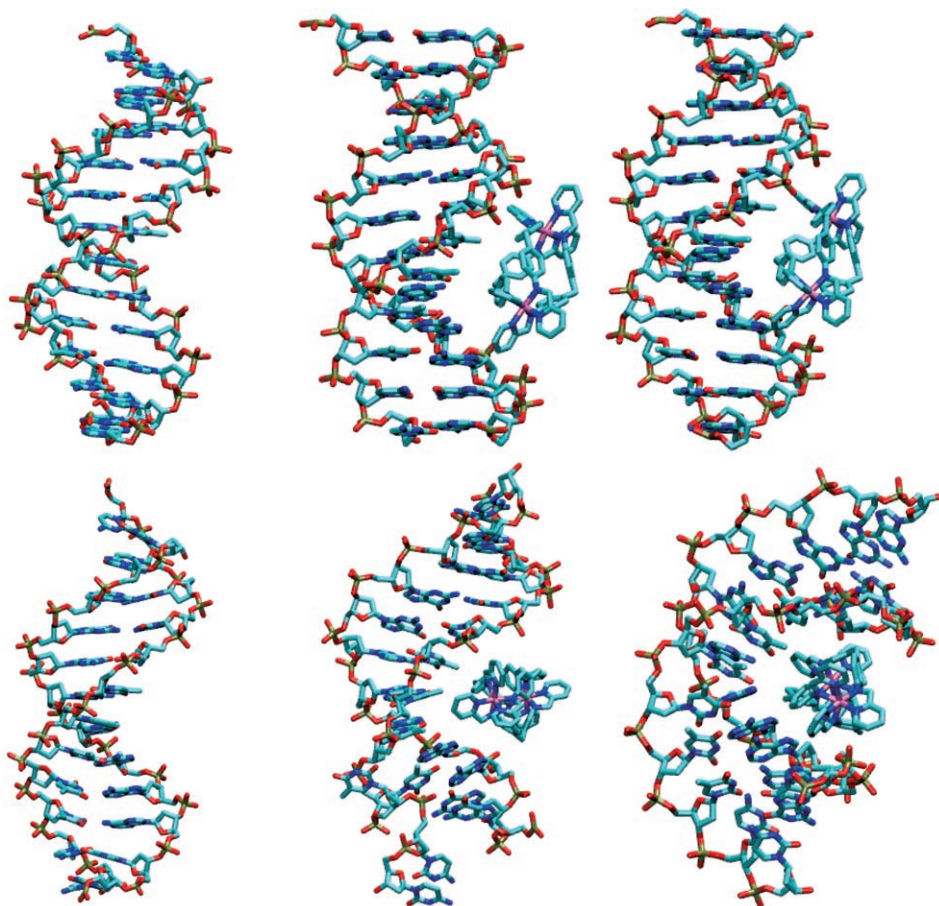
The helicoidal parameters developed by Lavery and Sklenar[38] are a good source of quantitative data on how these ligands affect the integrity of the base pairs within the DNA double helix. These parameters describe how two nucleic acid bases move and rotate relative to each other and are defined such that zero displacement/rotation corresponds to the ideal base pair geometry for B-DNA. Helicoidal parameters have been calculated from all three simulations and are presented in Figure 3. For the uncomplexed DNA, nearly all the base pairs can be seen to be stable, with typically small fluctuations about values of zero (the ideal base-pair geometry). The only exception is the first pair, C1–G24, which, from about 1 ns, showed large amplitude motions characteristic of fraying at the end of the double helix; end-fraying is a not uncommon event for DNA, both in simulations and in vivo. Some large-amplitude fluctuations were also seen for T8–A17 and T9–A16 pair at about 1 ns, but these were transient

Figure 2. Conformations of DNA taken from the beginning (top) and after 2 ns (bottom) of MD simulations using the CHARMM22 force field: uncomplexed DNA (left), DNA/C^{4+} (middle) and DNA/C^0 (right). Conformations after 5 ns were very similar to those at 2 ns.

Moreover, DNA bending is commonly associated with A–T tracts.[46,47]

The behaviour of the neutralised cylinder, C^0, provides a fascinating contrast to that of C^{4+}. Its final configuration shows C^0 to lie symmetrically within the major groove, spanning base pairs 5–10, but with the DNA having bent substantially to create a much deeper pocket for the cylinder than was found with C^{4+}. It might originally be thought that the coiling of the DNA was, in large part, due to the attraction between the large positive charge on C^{4+} and the negatively charged phosphates in the DNA backbone. However, it is clear from Figure 2 (and the quantitative analysis presented below) that the neutral cylinder, C^0, also causes strong coiling in the DNA, but does so without inducing any of the disruption of the Watson–Crick base pairs that was evident with C^{4+}. This contrast suggests that the shorter ranged van der Waals forces play a significant, perhaps even dominant, role in coiling the DNA backbone while the large Coulombic forces generated by the exposed cationic charge of C^{4+} may generate a localised stress that is strong enough to disrupt inter-strand hydrogen bonding in the DNA duplex.

and rapidly returned to stable values; such behaviour is indicative of the flexibility inherent within the DNA duplex rather than of irreversible conformational changes.

Given the amount of curvature evident in Figure 2 (right), the C^0 ligand gives rise to helicoidal parameters that are remarkably similar to those of the uncomplexed DNA. Indeed, the principal difference between these two sets of data is that the transient large amplitude oscillations seen in the uncomplexed DNA are completely absent for DNA in the presence of C^0. As with the uncomplexed DNA, fraying is again found only at the first base pair. Thus, the presence of C^0 actually appears to enhance the stability of the Watson–Crick base pairs, despite the overall curvature of the DNA evident in Figure 2 (right).

For C^{4+} bound to DNA, the helicoidal parameters clearly reveal a deformation of the base-pair stack. There are two distinct zones in the double helix for this system. The first seven base pairs show stable behaviour as seen in the uncomplexed DNA, with just small fluctuations around zero. Similar behaviour is seen for T8–A17, albeit with larger amplitude vibrations for the shear (SHR), buckle (BKL) and open (OPN) parameters. However, the plots for the last four base pairs show extremely large variations that are
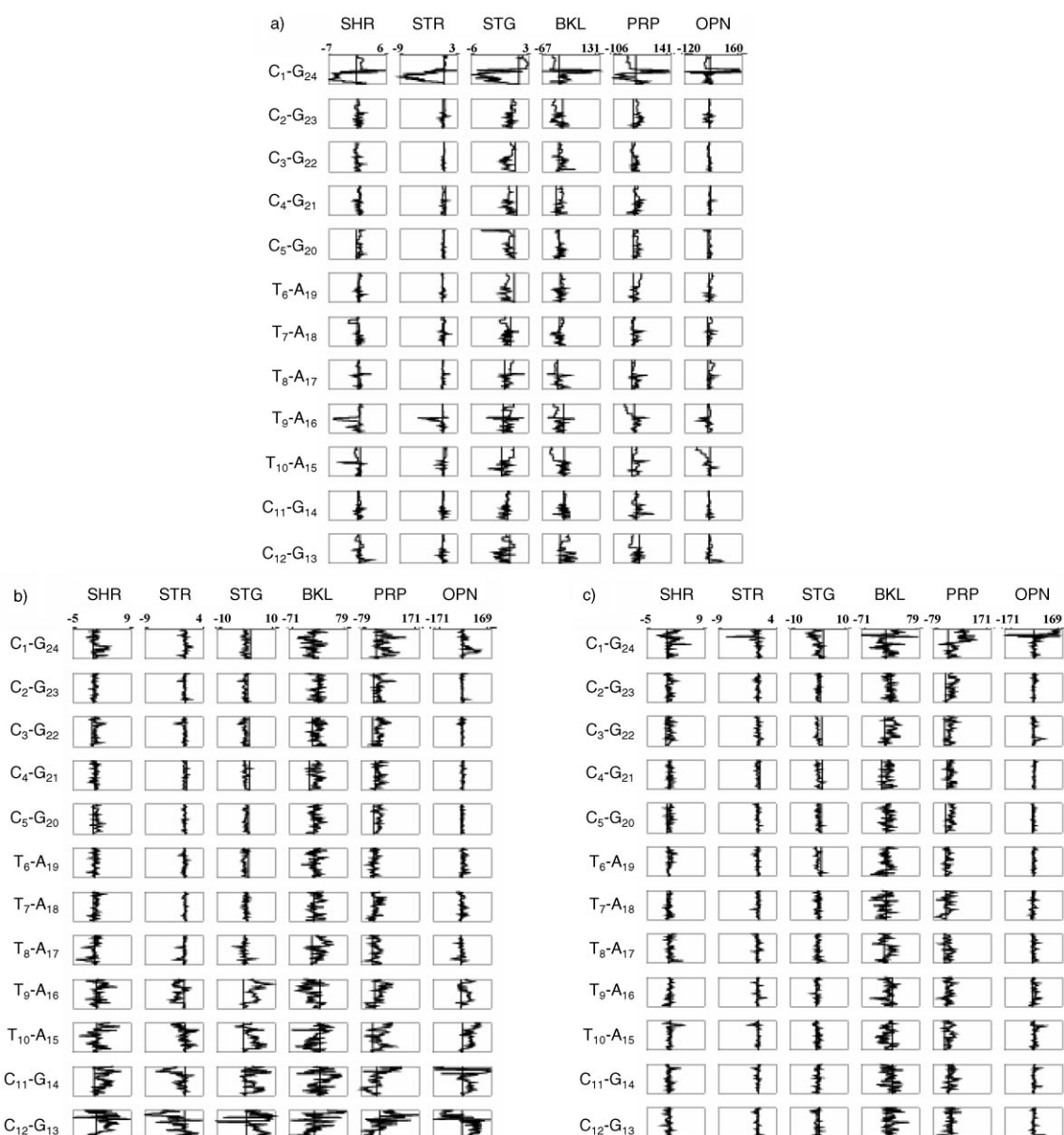
Figure 3. Helicoidal parameters for a) uncomplexed DNA, b) DNA/C^{4+} and c) DNA/C^0. Time is given on the vertical axis (0–2 ns, 0 top) and the helicoidal parameter on the horizontal axis (scale as marked). For a definition of these parameters[38] see Supporting Information.

simply inconsistent with a stable Watson–Crick base pair: shear deformations vary by as much as 8 Å during a simulation, while the base pair "stretches" are, at times, 3–4 Å *less* than their equilibrium value.

Visual inspection of the DNA/C^{4+} trajectory showed that this apparent disruption of the double helix below base-pair T9–A16 was actually a complex rearrangement amongst the base pairs, resulting in the formation of a new pairing between T9 and A15, so that T10 and A16 were left unpaired (see Figure 4). The sequence of events for the formation of this defect was as follows:

1) Some stress became evident in base pairs 10–12 from about 170 ps:

2) T9–A16 broke apart at 390 ps;

3) a new base pair, T9–A15, formed within the following 10 ps.

T10 was then observed to swing out of the helix, kinking the backbone, and subsequently disrupting the last two base pairs (C11–G14 and C12–G13). It is interesting to note that this activity occurred adjacent to, rather than at, the cylinder binding site: the cylinder remained bound to the base pairs numbered 5–8 (i.e., C5–G20 to T8–A17) throughout the simulation.

This change observed in the base-pair stacking is also revealed in the pattern of hydrogen bonding between the nucleic acids at the end of the simulation. Hydrogen-bond
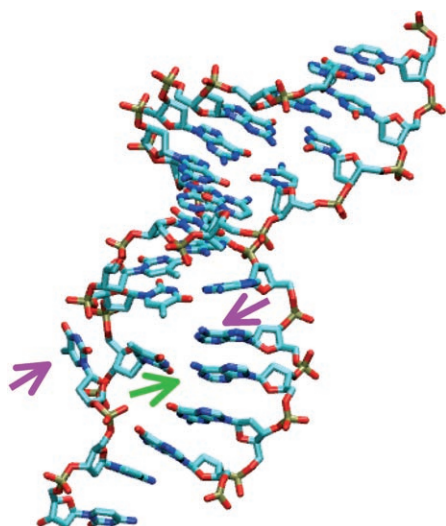
Figure 4. Snapshot of DNA showing the base-pair defect induced by C^{4+}. The green arrow points to the slipped mismatch between T9–A15; the magenta arrows indicate the resulting unmatched bases, T10 and A16.

lengths were obtained from the average DNA structure, as calculated by the 3DNA package, using the final 50 ps of each trajectory. This time window was found to be long enough to smooth out the instantaneous vibrations, but still short enough to avoid artefacts due to long timescale conformational motions of the DNA. The results for the two DNA/cylinder simulations are listed in Table 1 and confirm the formation of a mismatched Watson–Crick base pair T9–A15, with the two hydrogen bonds, in the presence of C^{4+}.

Dial plots of the DNA backbone torsion angles from the three sets of simulations are given in Figure 5. Data for the uncomplexed DNA are consistent with the average values expected for B-DNA, and show no evidence of a transition to one of the other forms. The presence of the cylinder does change the DNA backbone conformation, but the changes are smaller than was seen in the helicoidal parameters. In-

terestingly, while the helicoidal parameters indicated that the base-pair structure was more constrained in DNA/C^0 than in the uncomplexed DNA, the opposite appears to be true for the backbone, with the backbone torsions showing larger fluctuations in the presence of C^0 than without. The changes in base-pair stacking induced by C^{4+} are also seen in the backbone, but the difference between the two cylinders is less obvious than it was with the direct measures of base-pair geometry. The most obvious indication is in the C3′-O-P-O (ζ) and O-P-O-C5′ (α) angles, which show almost random angular variations between base pairs 10–12 in the presence of C^{4+}, suggestive of rotation that is unhindered by base-pair formation. This effect appears to be more pronounced at the end of the first strand (nucleotides T10 to C12) than on the complementary strand (nucleotides G13 to A15).

A schematic representation of the global helix axis after 2 ns is given in Figure 6. We note that quantifying curvature in the DNA helix can be problematical. Any measure must be able to distinguish between the local oscillations in base-pair orientation found in, particularly, the A-form, and a real bend in the overall helix direction. This can be particularly difficult in short strands of DNA, such as the dodecamer studied here, since one cannot use the behaviour of the helix beyond the binding site to confirm the persistence of any bend. To measure the extent of bending we have used the 3DNA definition of the local helix direction. This uses the geometry of any two adjacent base pairs (a step) to define a unit vector, U_i, along the helix axis at that step. Bending of the DNA can then be monitored by defining an angle, $\vartheta_i = \cos^{-1}(U_{ref} \cdot U_i)$, which describes the total bend in the DNA between some reference step (taken to be the first stable step in this work) and the ith base-pair step. Most importantly, this definition gives perfectly aligned U_i ($\vartheta_i = 0$ for every i) in *both* the canonical A and B forms, so that deviations from 0 can be interpreted as bending of the DNA.

ϑ_i values for the, averaged, DNA structures after 2 ns are shown in Figure 7; we reiterate that the DNA response was complete on this timescale, and no further relaxation was observed in simulations continued to 5 ns. Thus, these plots give a good indication of both the global bend induced in the DNA by the ligands, and how it is distributed along the dodecamer. Both C^{4+} and C^0 can be seen to bend the DNA by about 40°, which compares very favourably with the 40–60° per cylinder measured experimentally. In the presence of C^{4+} this bend is strongly localised on steps 5 and 6 (base pairs 5–7), but with the consequent degradation of base pairs 9–12 noted above. In contrast, the
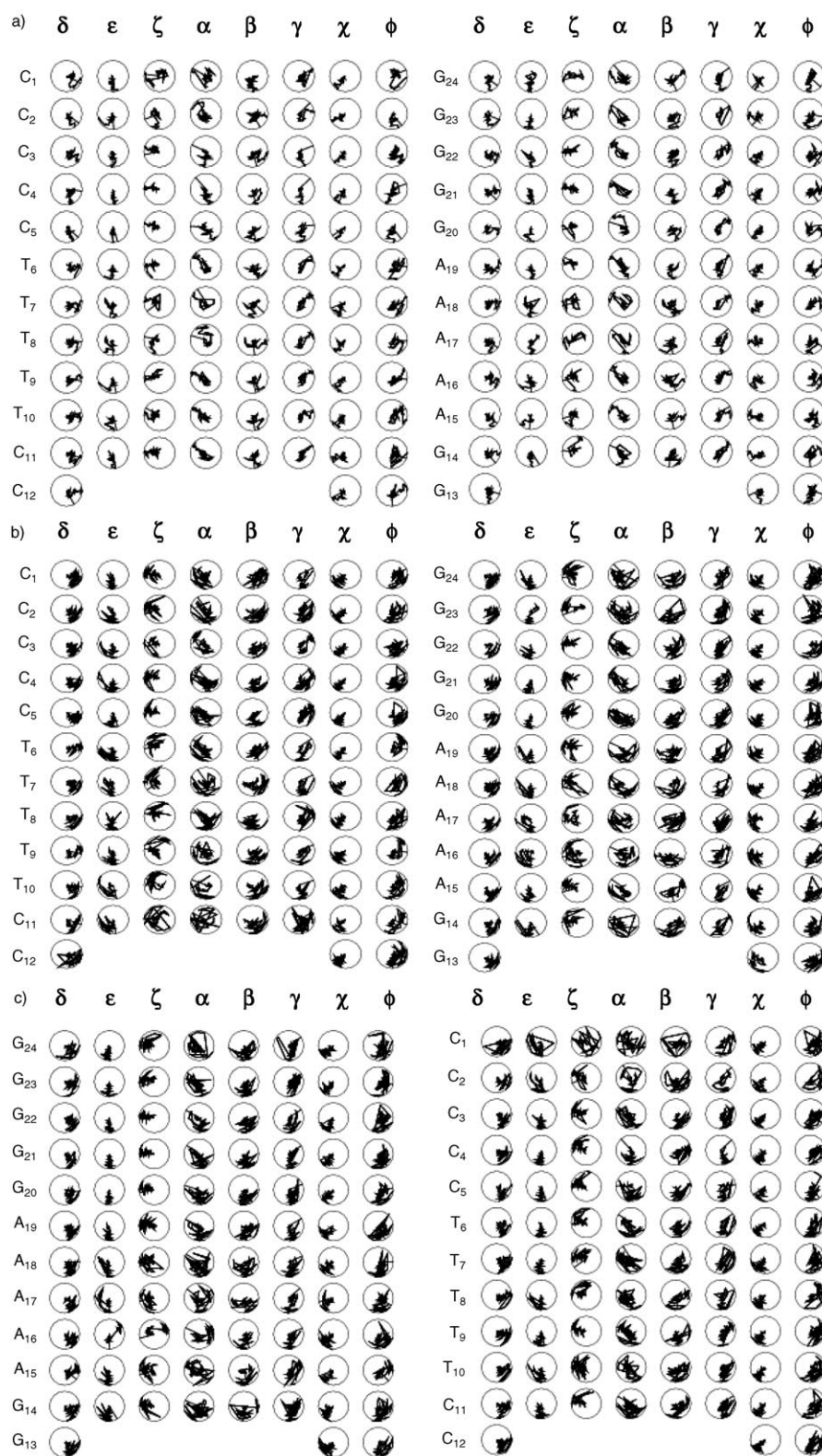
Table 1. Intra-base pair hydrogen bond length, taken from the average DNA structure observed during the final 50 ps of the simulation; blank values indicate no hydrogen bond was found. The mismatch induced by C^{4+} is indicated in bold.

Base pairings	Length [Å]					
	DNA/C^{4+}			DNA/C^0		
	O2–N2	N3–N1	N4–O6	O2–N2	N3–N1	N4–O6
C1–G24	2.75	3.06	3.22			
C2–G23	2.96	3.04	2.97	2.82	3.00	3.02
C3–G22	2.86	3.00	3.01	2.86	3.01	2.97
C4–G21	2.80	3.03	3.10	2.80	3.15	3.35
C5–G20	2.91	3.05	2.99	3.03	2.99	2.84
T6–A19		2.98	3.18		2.99	3.01
T7–A18		3.05	2.91		3.06	2.95
T8–A17		3.03	2.89		3.00	2.87
T9–A16					3.03	3.00
T9–A15		**3.17**	**2.89**			
T10–A15					3.13	2.88
C11–G14				3.00	3.01	2.94
C12–G13				2.85	3.01	3.00

Figure 5. DNA backbone torsion angles for a) uncomplexed DNA, b) DNA/C^{4+} and c) DNA/C^0. Circular polar plot of time (radius, 0–2 ns) against torsion angle (0° top, 90° right).

12° over 10 base pairs, with some evidence of more pronounced curvature over the first few steps.

Another common descriptor for helix curvature is the normal vector plot (NVPs).[43] NVPs are projections of the unit vector normal to the plane of each base pair onto a plane that is normal to the helix axis. Linear segments of the DNA can then be identified as clusters of neighbouring points, while coiling is seen as a steady change of orientation across a number of adjacent base pairs. NVPs calculated at the end of each simulation are presented in Figure 8. The uncomplexed DNA shows a reasonably tight cluster of points. In the presence of the neutralised cylinder (C^0), however, there is a large and systematic variation in orientation across base pairs 3–9, consistent with the strong curvature noted above. For the fully charged cylinder, C^{4+}, there is strong curvature but as was seen with the cumulative bending angle (Figure 7), this effect is localised between base pairs 6 and 7, that is, to the start of the region in which the ligand is bound, and is suggestive of a localised kink in the double helix.

Various parameters describing the geometry of base-pair steps have also been calculated using 3DNA, and the average structure obtained during the final 50 ps portion of the trajectory. In Figure 9 we report values of the slide displacement, and roll, twist and tilt angles,[40,41] for all stable base pairs. There are several important points to note from these plots. In the first place, the step parameters observed in the presence of the cylinders are not readily identified with any of the common DNA conformations. Slide values below −1 Å, as observed along most of the stable duplex in the presence of both C^{4+} and C^0, are often taken to be indica-

40° bend for C^0 is spread over 10 base pairs, which probably explains why the integrity of the duplex is retained in this case. The uncomplexed DNA shows a net bend of only 10–
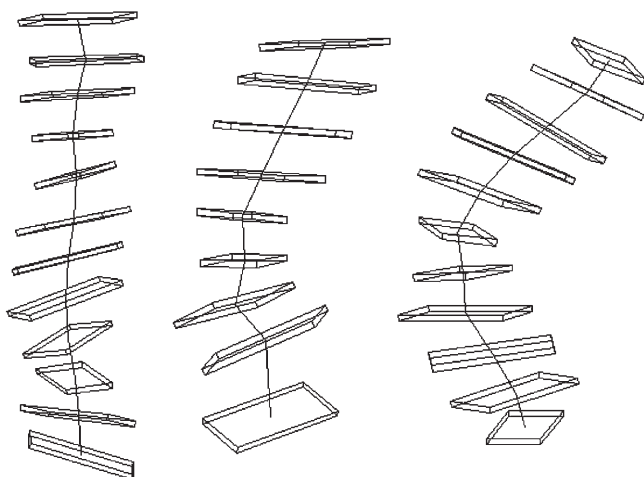
Figure 6. Representation of the average DNA structure in the uncomplexed (left), DNA/C^{4+} (middle) and DNA/C^0 (right) simulations. Averages were calculated from the final 50 ps of each simulation using 3DNA.[42] Only the stable base pairs are depicted.
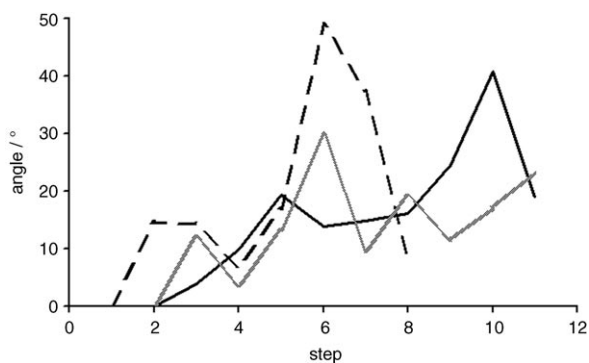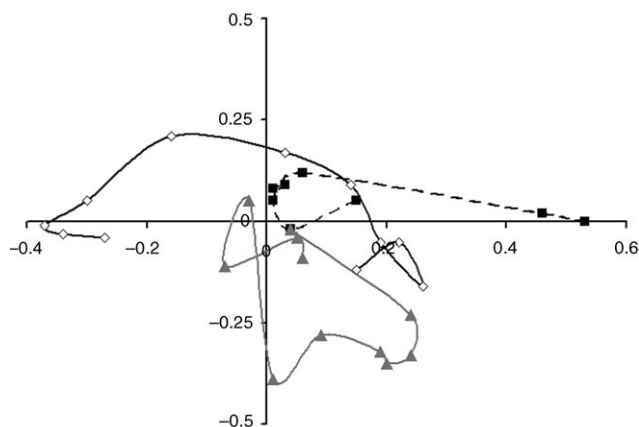

Figure 8. Normal vector plots (NVPs) for the average DNA structure calculated from the final 50 ps of each simulation using the CHARMM22 force field. NVPs are a projection of a unit vector normal to each base pair onto a plane that is normal to the local helical axis, with the points corresponding to successive base pairs being joined by a smooth line. ◇: C^0, ■: C^{4+}, ▲: uncomplexed.

Figure 7. Degree of bend in the DNA, calculated from the local helix direction as defined by 3DNA. The angles are defined as $\vartheta_i = \cos^{-1}(U_{ref} \cdot U_i)$, where U_{ref} is the helix director for the first stable step. The helix bend (left) is identically zero for both the canonical A- and B-forms. ——: C^0, -----: C^{4+}, ——: uncomplexed.

tive of A-DNA. However, both the slide and roll values obtained in this work are really intermediate between the values expected for A- and B-DNA, while, for C^0, the twist is more consistent with B-DNA. On balance, 3DNA assigns the duplex as having the B-form in the presence of the neutral cylinder, and fails to make any assignment in the presence of C^{4+}. It is also interesting to note that both the slide and the roll angles take their most extreme values at step 6, which is both the start of the region to which the cylinders bound and the start of the A-T tract. Finally, we note that the twist angles indicate that neither C^{4+} nor C^0 cause significant winding or unwinding of the DNA—either locally or globally. The average twist for C^0 is 33.4° per step, which is intermediate between the A- and B-forms. A slightly lower value (29.4° per step) is seen in the presence of C^{4+}, but the extent to which this is influenced by the mismatched base pairing is unclear. This behaviour is in contrast to some pro-

teins, such as 434 Cro, which bend the DNA via alternating regions of under- and overwinding.[1]

A concomitant question when DNA bending is observed is to ask how the size of major and minor grooves is affected. Following Hassan and Calladine,[48] we have monitored these by measuring the distance between appropriately displaced phosphate groups, with the P atom representing the position of the phosphate. The width of the major groove can be defined as the distance between the P atom in the nth nucleotide, counting from the 5′ end of one strand, and the P atom of the $n+3$rd nucleotide from the 3′ end of the complementary strand; this numbering assumes 5′ nucleotides. Similarly, for the minor groove one can use the distance between the P atoms in the nth nucleotide from the 5′-end of one strand with the $n-4$th nucleotide from the 3′-end of the complementary strand. The distribution of these inter-phosphate distances was determined by analyzing every configuration saved from the final 0.25 ns of each simulation. To avoid artefacts from end-fraying, those phosphates that were adjacent to the first and last base pair were omitted from the analysis. This left five measurements of the major groove width, and six for the minor groove. It should be noted that, for a dodecamer, this meant that the minor and major groove measurements were made on largely different regions of the DNA, with only two base pairs (T6–A19 and T7–A18) being used to measure both the major and the minor groove widths.

Unlike crystal structure analyses, the DNA shows considerable flexibility during a simulation, and so gives rise to a distribution of groove widths for each pair of phosphate groups. The average groove width and its standard deviation are presented in Figure 10. It is striking that the DNA-coiling induced by C^0 does not in any way affect the width of the major groove—the site into which C^0 binds. An enlargement in the width of the minor groove (ca. 2.5 Å) is seen in the vicinity of the C^0 binding site, but only for four of the
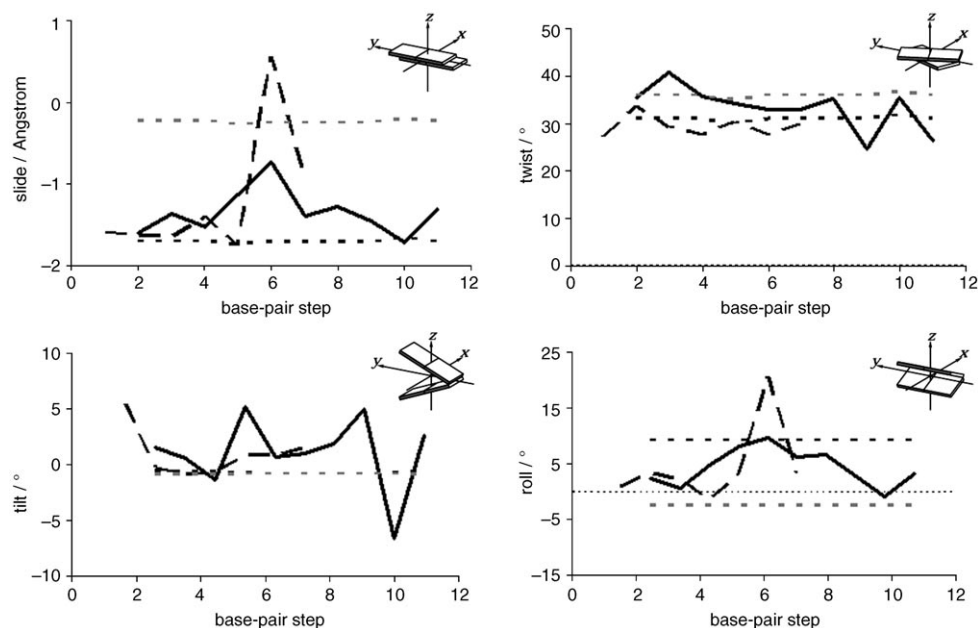
Figure 9. Slide displacement, and twist, tilt and roll angles between adjacent base pairs; a schematic definition of each inter-base pair coordinate is given with the plot. The horizontal axis corresponds to the step between base pairs, so that step 1 compares the first two base pairs (C1–G24 and C2–G23), etc. Only stable base-pairs have been included. Data for the canonical A- (-----) and B-forms (-----) are given as a reference; ——: C^0, ---: C^{4+}.
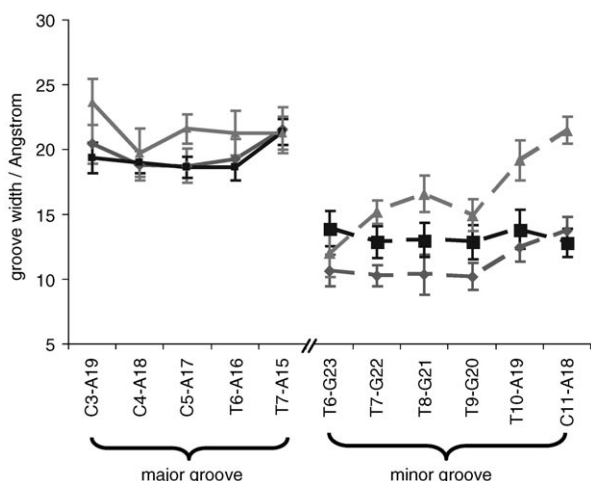

Figure 10. Widths of the major (solid) and minor (dashed) grooves, as determined from specified inter-phosphate distances. Calculations were performed with CHARMM22. "Error bars" denote the width of the distribution of distances ($\pm$ one standard deviation), not the uncertainty in the calculation. No correction has been made for the size of the phosphate group; ♦: uncomplexed, ■: C^0, ▲: C^{4+}.

six inter-phosphate distances found in this region (those involving nucleosides T6 through T9, but not A18 and A19).

In contrast to C^0, C^{4+} has a substantial effect on the widths of both the major and minor groove. The major groove width increases by an average of 1.8 Å compared with both the uncomplexed DNA and the DNA with bound C^0. Even larger increases are seen in the width of the minor

groove. However the largest increases (nearly 8 Å for C11–A18) are in regions where the backbone torsion angles α and ζ have already been shown to be ill defined (Figure 5) and so probably indicate that there is no well defined minor groove in this region when C^{4+} is bound to the DNA.

In summary, the supramolecular cylinder, **1**, is seen to have a strong influence on the conformation of DNA as modelled with the CHARMM22 force field. The interactions within a neutralised analogue of **1** were sufficient to induce a bend of about 40° in the DNA helix axis, which is very similar to the angle seen in experiments with **1**. When the tetracationic charge is also accounted for, **1** was found to strain the base pairing within the duplex, and induced a mispairing of A–T bases adjacent to the ligand binding site. It is important to stress that the induction of base-pair defects was reproducible. Four separate C^{4+} simulations were performed, starting with different low-energy cylinder/DNA binding sites and with different arrangements of the solvent and counterions. In each case a mismatched A–T base pair formed within the first 0.5 ns. It is possible that the defects were introduced as part of the coiling process (i.e., whether the additional strain during coiling made the DNA susceptible to defects in the base-pair sequence) in which case they would not reflect a global free energy minimum. To test this, a further simulation was performed starting from a stable configuration obtained after 2 ns of the DNA/C^0 simulation and changing the atomic charges back to those of C^{4+}; four Na$^+$ ions were also removed to maintain electrical neutrality in the system. An equivalent extended simulation with C^0 was used as a control. Within just 0.25 ns, mispairings in the A–T region again appeared in the presence of C^{4+} (Supporting Information, Figure 4), but no mispairings appeared with

C⁰ during an additional 3 ns. We conclude that strong coiling of DNA by **1**, coupled with a disruption of the duplex base pairing, is a real prediction from the CHARMM22 force field.

Force-field dependence: The two DNA force fields have been compared by measuring the extent to which a stable binding geometry obtained with one force field is maintained by the other. This amounts to determining whether an important minimum on the free energy surface of one force field is reproduced with the other. By starting in a well-defined free-energy minimum, instead of repeating the docking protocol, one avoids the possibility of simulations with different potentials simply converging on different local minima and thereby ends up with an unambiguous comparison of the two force fields. Accordingly, simulations with the CHARMM27 force field were started from the stable conformation obtained after 2 ns in the C⁰/CHARMM22 simulation.

Simulations were again performed with both tetracationic and neutralised cylinders and are denoted $_{27}C^{4+}$ and $_{27}C^0$, respectively. For the $_{27}C^{4+}$ system, four Na^+ ions chosen at random were removed to maintain electrical neutrality. The system was then relaxed in a series of short, restrained simulations, as outlined in Section Computational Methods. All restraints (i.e., tethering potentials) were then removed and the system simulated for an additional 1.5 ns. The 3DNA analysis of bend and step parameters, obtained from the average structure at the end of these simulations, is presented in Figures 10–12, while the distribution of groove widths observed during the final 0.25 ns is depicted in Figure 14. Where clarity permits, results for CHARMM22 are also depicted for comparison.

In general, the CHARMM27 DNA is less strongly influenced by the cylinder than is the CHARMM22 model. For both $_{27}C^{4+}$ and $_{27}C^0$, the DNA uncoils slightly to give an overall bend of only about 20° (Figure 11), while the normal vector plots (Figure 12) remain relatively compact. At the same time the slide parameters (Figure 13) indicate a shift to the B-form that is known to be favoured by CHARMM27 for an aqueous duplex in the absence of a
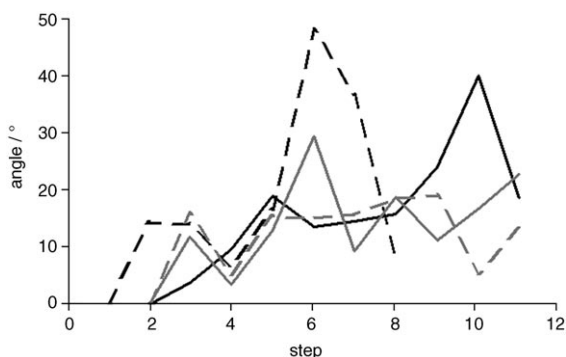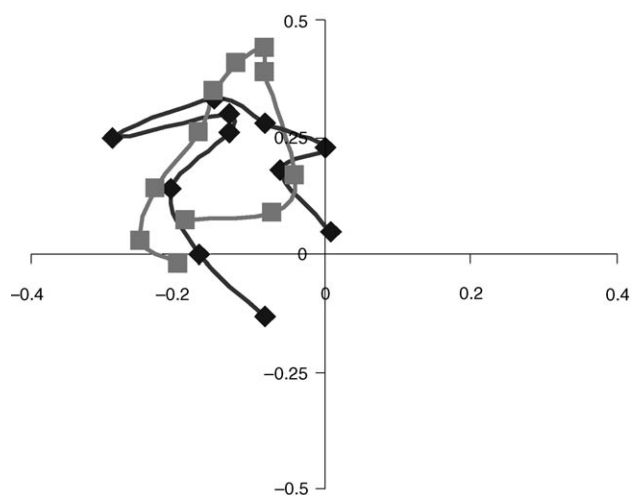

Figure 12. Normal vector plots for the average DNA structure obtained during the final 50 ps of the CHARMM27 simulations. ◆: $_{27}C^0$, ■: $_{27}C^{4+}$.

ligand. No evidence was found that either cylinder disrupted the base pairings at any stage during the simulations, with the average structure from the end of the simulations showing the expected three (G–C) and two (A–T) hydrogen bonds for every base pair; at least one of the hydrogen-bond lengths was less than 3.0 Å in every base pair, while none was more than 3.13 Å. This is a major difference from the analogous CHARMM22 simulation described above, where reinstating the $+4\,e$ cylinder charge on the DNA/C⁰ complex induced base-pair mismatches within just 0.25 ns.

Perhaps most surprisingly, the cylinder charge appears to have very little effect on the CHARMM27 DNA. The normal vector plots, global bend and step parameters (with the exception of the roll angle at the 6th step) all show remarkably similar behaviour for the $_{27}C^{4+}$ and $_{27}C^0$ systems. Given the strongly ionic character of DNA and its flexibility, this apparent invariance to the ligand charge is unexpected. One property that does show differences between the two force fields is the groove widths. Relative to CHARMM22/C⁰, the major groove of CHARMM27-DNA contracts in the presence of C⁴⁺, but expands in the presence of C⁰. We note that CHARMM27 uses a more strongly charged DNA backbone than does CHARMM22, with the PO_4 net charge being $-1.2\,e$, compared with $-1.0\,e$ for CHARMM22. This increased negative charge may lead the DNA backbone to contract more strongly onto the tetracationic cylinder, resulting in a stiffening and straightening of the DNA backbone and thereby inhibiting the subsequent bending of the DNA. A similar response to ligand charge has been seen with netropsin bound in the minor groove,[24] but greater scope for DNA relaxation might have been anticipated when the ligand binds in the major groove.

The comparison between these two force fields highlights several influences on the DNA response. In the first place we note that both force fields predict the neutralised ligand to coil DNA at least as strongly as the tetra cationic ligand. Previous work has shown that regions of low dielectric con-

Figure 11. Global bend of the DNA, calculated as for Figure 6. CHARMM27 results are given as dashed lines, and CHARMM22 as the solid lines; ——: C⁰, -----: C⁴⁺, ——: $_{27}C^0$, -----: $_{27}C^{4+}$.
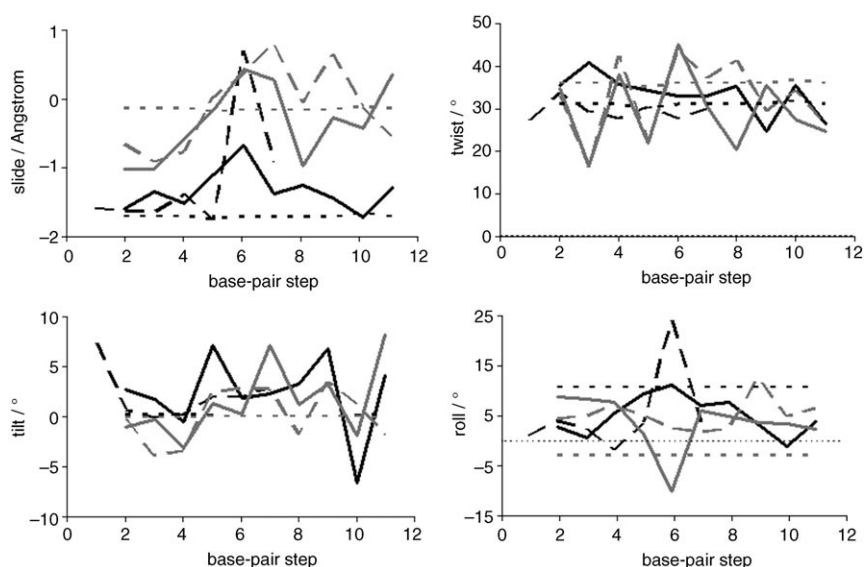
Figure 13. Selected step parameters for stable base pairs, calculated as for Figure 8. Values for the canonical A- and B-forms are given for reference, ——: C^0, -----: C^{4+}, ——: $_{27}C^0$, -----: $_{27}C^{4+}$, •••••: A-form, ······: B-form.
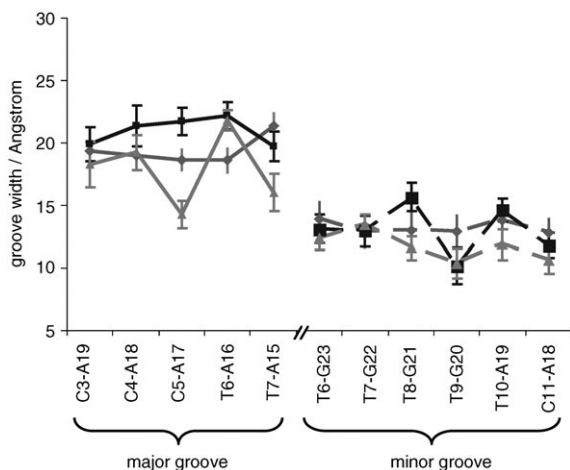

Figure 14. Groove widths calculated during the final 0.25 ns of the CHARMM27 simulations. Solid lines indicate the major groove and dashed lines the minor groove, ◆: C^0, ■: $_{27}C^0$, ▲: $_{27}C^{4+}$.

stant enhance the phosphate repulsions in DNA and cause it to bend *away* from the low dielectric region.[49] Although the neutralised ligand does generate electric field gradients, its enhanced negative surface charge should also accentuate the phosphate repulsions and induce the DNA to bend away from the ligand, or at least coil less tightly about it. The fact that this is not observed, and indeed the coiling may be enhanced for the neutralised ligand, implies that it is the other forces—the van der Waals forces—that drive the coiling.

Ligand charge does have some affect on the DNA response, but its principal manifestation is different in the two force fields. The contrast is instructive. In CHARMM27, the $+4e$ charge attracts the negatively charged DNA backbone, causing the major groove to contract onto the cylinder. With

CHARMM22 its effect is to strain the hydrogen bonding between base pairs in the duplex. The induction of such strain is reasonable given that most force fields describe hydrogen bonding primarily through electrostatic interactions. Orbital interactions, such as charge transfer and π-electron polarisation, do contribute strongly to the hydrogen bonding between base pairs,[50,51] but these are also likely to be strongly perturbed by the proximity of a tetra-cationic ligand. In reality, both the strain to the base pair hydrogen bonds and the contraction of the DNA backbone onto the ligand are likely to be present, and getting the balance right will be an important validation of the force field. In this context, the CHARMM22 result that duplex base-pairing will be disrupted by C^{4+} provides an interesting prediction that awaits the measurement of definitive X-ray or NMR structures.

Sensitivity to cylinder charge: If, as indicated above, the disruption of base pairs in a DNA duplex with CHARMM22 is due to a competition between inter-base pair hydrogen bonding and ligand base electrostatics, then it is of interest to identify how strong the ligand electrostatics need to be in order to effect the base pair mismatches. Analogous simulations to those reported in the last section were performed with CHARMM22 and with a variant of the cylinder in which the overall charge was set to $+3e$ (C^{3+}), again achieved by a constant shift of the atomic charges for all C and non-polar H atoms. Within 0.25 ns this system also showed disruption of the base pairing within the double helix (see Figure 15). Similar calculations with a $+2e$ variant of the cylinder (not shown) gave results that were intermediate between C^{3+} and C^0, with frequent bifurcated pairings in which an A was found to bridge two T bases in the opposite strand, but these did not lead to irreversible breakdown in the base pairing on a nanosecond timescale. We conclude that some disruption of the duplex base pairs is a robust feature of the CHARMM22 model, and is not especially sensitive to the parameterisation of the cylinder charge distribution.

Conclusion

Our molecular dynamics simulations of the supramolecular cylinder, **1**, in explicit solvent show major groove binding of the cylinder on the DNA and coiling of the DNA in re-
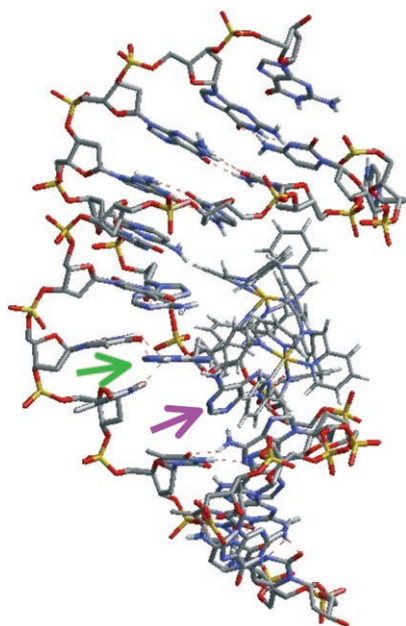
Figure 15. Configuration of DNA 0.25 ns after continuing from the final configuration of the 2 ns DNA/C^0 simulation with a $+3\,e$ cylinder (C^{3+}). The magenta arrow points to an unpaired T, while the green arrow points to two A bases paired to a single T.

sponse to this binding. This is consistent with all available experimental observations. Moreover, the simulations have, as envisaged, yielded molecular level information about the coiling effect. Other highly charged cations, such as spermine and cobalt(III)–hexamine,[17,22,52,53] bend DNA by a few degrees per ligand and cause it to condense into poly-molecular aggregates (often of well-defined shape). Our simulations indicate that the DNA bend induced by **1** is an order of magnitude greater than this: the simulations give 20–40°, depending on the force-field model, which compares favourably with the experimental measurements of 40–60° per ligand. Somewhat surprisingly, the tetracationic charge of the supramolecular cylinder was found *not* to be essential for inducing the coiling, as a neutral analogue was also found to bend the DNA. Instead the molecular shape appears to be the key factor and must give rise to a coil with a pitch which ensures an intramolecular effect observed experimentally with this ligand, rather than the more commonly observed intermolecular aggregation. This highlights the significance of our molecular design which creates, through metallo-supramolecular assembly, a cylindrical agent of very similar shape and dimensions to the various cylindrical motifs employed by proteins for DNA recognition and structure control.

This is *not* to say electrostatic interactions were unimportant. Our simulations identified two competing influences arising from a cationic ligand in the major groove that are likely to be general effects. On the one hand, the large cationic charge attracted the phosphate backbone, causing the major groove to contract around the ligand. On the other hand, such a large positive charge density close to the base

pairs places strain upon the inter-base hydrogen bonding that holds the DNA duplex together. The balance of these two effects did depend on the choice of DNA force field, but the latter effect was sufficiently large that C^{4+} was observed to generate base-pair mismatches in some systems. When these mismatches were observed, they occurred in the A–T tract and adjacent to, rather than at, the site at which the cylinder bound. Any weakening of the base pair hydrogen bonding at or near this site could prime the DNA for transcription or replication processes. This is an unexpected molecular-level feature associated with the experimentally observed cylinder binding and DNA conformational change. Intriguingly, we note that many proteins also bend DNA at locations adjacent to the site of transcription. For example, the CAP (catabolic gene activating protein) binds adjacent to the initial site at which RNA polymerase acts and causes a 50-fold increase in the rate of transcription initiation.[54] This activation is usually ascribed to the provision of an interaction surface for the polymerase, thereby enhancing binding to the promoter site. Our C^{4+}/DNA simulations suggest there may be an alternative mechanism: that the bending induced by CAP binding can help to open the DNA at an adjacent site, and could thus contribute to the increased rate at which transcription is initiated in the presence of promoters. This opens up new potential applications for these DNA-coiling metallo-supramolecular cylinders.

Thus we conclude that the tetracationic supramolecular cylinder, **1**, does have two substantive effects on DNA. In the first place it can induce coiling of the DNA, with potentially very large curvature being introduced in DNA dodecamers. However, at the same time it weakens the hydrogen bonding between base pairs near the binding site and so can tend to introduce defects into the alignment of the base-pair sequences in the two strands of the double helix. In our simulations this occurred consistently within the A–T rich region of the DNA and resulted in misaligned A–T crosslinks between the two strands. It is not clear whether this represents a selectivity for A–T over G–C and simulations with other DNA sequences are in progress to resolve this issue.[1,2]

[1] See, for example, J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, 5th ed., Freeman, New York, **2002**.

[2] J. P. Baak, F. R. Path, M. A. Hermsen, G. Meijer, J. Schmidt, E. A. Janssen, *Eur. J. Cancer* **2003**, *39*, 1199–1215.

[3] C. F. Calkhoven, C. Muller, A. Leutz, *Trends Mol. Med.* **2002**, *8*, 577–583.

[4] L. Rogge, *Arthritis Res. Ther.* **2003**, *5*, 47–50.

[5] B. S. Parekh, G. W. Hatfield, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1173–1177.

[6] R. R. Sinden, *DNA Structure and Function*, Academic Press, London, **1994**.

[7] R. E. Dickerson, *Nucleic Acids Res.* **1998**, *26*, 1906–1926.

[8] C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland, New York, 2nd ed., **1999**.

[9] B. C. Baguley, *Mol. Cell. Biochem.* **1982**, *43*, 167–181.

[10] W. A. Denny, *Anti-Cancer Drug Des.* **1989**, *4*, 241–263.

[11] S. J. Lippard, J. M. Berg, *Principles of Bioinorganic Chemistry*, University Science Books, Mill Valley, CA, **1994**.

[12] D. Z. Coggan, I. S. Haworth, P. J. Bates, A. Robinson, A. Rodger, *Inorg. Chem.* **1999**, *38*, 4486–4497.

[13] B. Onfelt, P. Lincoln, B. Nordén, *J. Am. Chem. Soc.* **1999**, *121*, 10846–10847.

[14] K. E. Erkkila, D. T. Odom, J. K. Barton, *Chem. Rev.* **1999**, *99*, 2777–2795.

[15] M. J. Hannon, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletton, I. Meistermann, C. J. Isaac, K. J. Sanders, A. Rodger, *Angew. Chem.* **2001**, *113*, 903–908; *Angew. Chem. Int. Ed.* **2001**, *40*, 879–884.

[16] I. Meistermann, A. Rodger, V. Moreno, M. J. Prieto, E. Moldrheim, E. Sletton, S. Khalid, P. M. Rodger, J. Peberdy, C. J. Isaac, M. J. Hannon, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5069–5074.

[17] A. Rodger, K. J. Sanders, M. J. Hannon, I. Meistermann, A. Parkinson, D. S. Vidler, I. S. Haworth, *Chirality* **2000**, *12*, 221–236.

[18] G. S. Manning, K. K. Ebralidse A. D. Mirzabekov, A. Rich, *J. Biomol. Struct. Dyn.* **1989**, *6*, 877–889.

[19] E. Moldrheim, M. J. Hannon, I. Meistermann, A. Rodger, E. Sletten, *J. Biol. Inorg. Chem.* **2002**, *7*, 781–790.

[20] I. S. Haworth, A. Elcock, A. Rodger, W. G. Richards, *J. Biomol. Struct. Dyn.* **1991**, *9*, 553–569.

[21] B. P. Hudson, J. K. Barton, *J. Am. Chem. Soc.* **1998**, *120*, 6877–6888.

[22] T. E. Cheatham, P. A. Kollman, *Structure* **1997**, *5*, 1297–1311.

[23] Y. Y. Fang, B. D. Ray, C. A. Claussen, K. B. Lipkowitz, E. C. Long, *J. Am. Chem. Soc.* **2004**, *126*, 5403–5412.

[24] B. Wellenzohn, W. Flader, R. H. Winger, A. Hallbrucker, E. Mayer, K. R. Liedl, *Biopolymers* **2001**, *61*, 276–286.

[25] J. Dolenc, C. Oostenbrink, J. Koller, W. F. van Gunsteren, *Nucleic Acids Res.* **2005**, *33*, 725–733.

[26] A. D. MacKerell, Jr., D. Bashford, M. Bellot, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnik, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

[27] N. Foloppe, A. D. MacKerell, Jr., *J. Comput. Chem.* **2000**, *21* 86–104.

[28] S. Y. Reddy, F. Leclerc, M. Karplus, *Biophys. J.* **2003**, *84*, 1421–1449.

[29] W. Smith, T. R. Forester, *J. Mol. Graphics* **1996**, *14*,136.

[30] a) M. A. San Miguel, R. Marrington, P. M. Rodger, A. Rodger, C. Robinson, *Eur. J. Biochem.* **2003**, *270*, 3345–3352; b) M. A. San Miguel, P. M. Rodger, *Phys. Chem. Chem. Phys.* **2003**, *5*, 575–581;

c) R. Lukac, A. J. Clark, M. A. San Miguel, A. Rodger, P. M. Rodger, *J. Mol. Liq.* **2002**, *101*, 261–272; d) T. Astley, G. G. Birch, M. G. B. Drew, P. M. Rodger, *J. Phys. Chem. A* **1999**, *103*, 5080–5090.

[31] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.

[32] J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comp. Physiol.* **1977**, *23*, 327.

[33] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford (UK), **1987**.

[34] M. Feig, B. M. Pettitt, *Biophys. J.* **1998**, *75*, 134–149.

[35] N. Korolev, A. P. Lyubartsev, A. Laaksonen, L. Nordenskiöld, *Biophys. J.* **2002**, *82*, 2860–2875.

[36] S. Khalid, PhD Thesis, University of Warwick (UK), **2004**.

[37] B. Wellenzohn, R. H. Winger, A. Hallbrucker, E. Mayer, K. R. Liedl, *J. Am. Chem. Soc.* **2000**, *122*, 3927–3931.

[38] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 655–667.

[39] G. Ravishanker, S. Swaminathan, D. L. Beveridge, R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1989**, *6*, 669–699.

[40] R. E. Dickerson, M. Bansal, C. R. Calladine, S. Diekmann, W. N. Hunter, O. Kennard, E. von Kitzing, R. Lavery, H. C. M. Nelson, W. K. Olson, W. Saenger, Z. Shakked, H. Sklenar, D. M. Soumpasis, C. S. Tung, A. H. J. Wang, V. B. Zhurkin, *EMBO. J.* **1989**, *8*, 1–4.

[41] W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinmann, X. J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C. S. Tung, E. Westhof, C. Wolberger, H. M. Berman, *J. Mol. Biol.* **2001**, *313*, 229–237.

[42] X. J. Lu, W. K. Olson, *Nucleic Acids Res.* **2003**, *31*, 5108–5121.

[43] D. L. Beveridge, S. B. Dixit, G. Barreiro, K. M. Thayer, *Biopolymers* **2004**, *73*, 380–403; D. McDonald, P. Lu, *Curr. Opin. Struct. Biol.* **2002**, *12*, 337–343.

[44] R. E. Dickerson, T. K. Chiu, *Biopolymers* **1997**, *44*, 361–403.

[45] R. R. Sinden, *DNA Structure and Function*, Academic Press, New York, **1994**.

[46] A. Barbic, D. P. Zimmer, D. M. Crothers, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2369–2373.

[47] J. Hizver, H. Rozenberg, F. Frowlow, D. Rabinovich, Z. Shakked, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8490–8495.

[48] M. A. Hassan, C. R. Calladine, *J. Mol. Biol.* **1998**, *282*, 331–343.

[49] A. H. Elcock, A. McCammon, *J. Am. Chem. Soc.* **1996**, *118*, 3787–3788.

[50] C. F. Guerra, F. M. Bickelhaupt, *Angew. Chem.* **2002**, *114*, 2194–2197; *Angew. Chem. Int. Ed.* **2002**, *41*, 2092–2095.

[51] C. F. Guerra, F. M. Bickelhaupt, *Angew. Chem.* **1999**, *111*, 3120–3122; *Angew. Chem. Int. Ed.* **1999**, *38*, 2942–2945.

[52] I. S. Haworth, A. Rodger, W. G. Richards, *Proc. R. Soc. London Ser. B* **1991**, *244*, 107–116.

[53] K. Bryson, R. J. Greenall, *J. Biomol. Struct. Dyn.* **2000**, *18*, 393–412.

[54] See reference [1], p. 873.