Exocyclic groups in the minor groove influence the backbone conformation of DNA

Bernd Wellenzohn, Wolfgang Flader, Rudolf H. Winger, Andreas Hallbrucker, Erwin Mayer and Klaus R. Liedl*

Institute of General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innrain 52a, A-6020 Innsbruck, Austria

Received August 31, 2001; Revised and Accepted October 22, 2001

ABSTRACT

Exocyclic groups in the minor groove of DNA modulate the affinity and positioning of nucleic acids to the histone protein. The addition of exocyclic groups decreases the formation of this protein–DNA complex, while their removal increases nucleosome formation. On the other hand, recent theoretical results show a strong correlation between the B/BII phosphate backbone conformation and the hydration of the grooves of the DNA. We performed a simulation of the d(CGCGAATTCCGCG)2 dodecamer and one simulation of the d(CGCAAATTTCGCG)2 dodecamer in order to investigate the influence of the exocyclic amino group of guanine. The removal of the amino group introduces a higher intrinsic flexibility to DNA supporting the suggestion that make the enhanced flexibility responsible for the enlarged histone complexation affinity. This effect is attributed to changes in the destacking interactions of both strands of the DNA. The differences in the hydration of the minor groove could be the explanation of this flexibility. The changed hydration of the minor groove also leads to a different B/BII substate pattern. Due to the fact that the histone preferentially builds contacts with the backbone of the DNA, we propose an influence of these B/BII changes on the nucleosome formation process. Thus, we provide an additional explanation for the enhanced affinity to the histone due to removal of exocyclic groups. In terms of B/BII we are also able to explain how minor groove binding ligands could affect the nucleosome assembly without disrupting the structure of DNA.

INTRODUCTION

Knowledge of the detailed structural behavior of DNA is of extraordinary interest in order to understand sequence-specific complexation. Indirect readout is mediated by the sequence-dependent conformational deformability of the DNA. Bending, unwinding and other recognition tools in the indirect readout have already been extensively investigated (2–16). An analysis (17) of protein–DNA complexes indicates that >50% of all contacts are between the amino acid side chain and the DNA backbone demonstrating the importance of DNA backbone conformations in recognition processes (18).

Experimental and theoretical studies showed that these backbone phosphates occupy either the B or BII conformational substrate (19–25). The B/BII pattern is sequence dependent, thus transferring sequence information from the bases to the chemical degenerate phosphate groups. Hence, we surmise that these conformational substrates of the phosphates influence the sequence-specific recognition of the DNA. The backbone angles ε and ζ are used to define the B and BII states. In the B state the corresponding ε and ζ angles are between 120–210° (trans) and 235–295° (gauche+), respectively, and for the BII state the ε and ζ angles are between 210–300° (gauche−) and 150–210° (trans), respectively (26–28).

A recently performed molecular dynamics simulation of the trp repressor–operator complex shows a strong correlation between the B/BII phosphate substrate and the number of interactions with this phosphate (Wellenzohn et al., submitted for publication). There is also evidence for the B/BII equilibrium inducing a dynamic curvature in the NF-kB binding site playing an active role in DNA–protein recognition (29,30).

The role of water molecules in the mechanism of the BII →B state transitions were investigated by experimental (31) and theoretical (23,32) methods showing strong correlations with hydration patterns in the minor and major groove. Thus, such exocyclic groups should influence the BII equilibrium. Recent experimental studies showed that the exocyclic groups modulate the affinity and positioning of DNA to the histone octamer (33–36). It was suggested that addition of exocyclic groups decrease and removal increase the local deformability of DNA (37) by maximizing and minimizing the steric resistance to bending. Inosine is a purine nucleoside, which differs from guanine only by the removal of the exocyclic amino group. It occurs naturally in the wobble position of some t-RNAs where it appears to pair also with adenosine in addition to cytidine and uridine (38–42). The X-ray structure of the d(CGCAAATTTCGCG)2 (43) dodecamer, which is an inosine analog of the
Drew Dickerson dodecamer d(CGCGAATTCGCG)₂ (44–46), was derived recently. We performed a 3 ns long molecular dynamics simulation of the d(CGCGAATTCGCG)₂ dodecamer and compared it with a 10 ns long reference simulation of the Drew Dickerson dodecamer. The intention of this work was to investigate the influence which the exocyclic amino group in the minor groove exerts on DNA structure and dynamic. As a result of the strong correlation between the $B_I \rightarrow B_{II}$ equilibrium and the hydration pattern of the grooves the removal of the exocyclic amino group should also influence the phosphate conformations. An analysis of the simulation results indicates that the phosphates of the CpI steps have a preference for $B_{II}$ while in the reference simulation the respective CpG steps are mostly in $B_I$. These results underline the importance of the hydration for the DNA conformations supporting theoretical (32) and experimental (31) work carried out previously. The removal of the amino group also enhances the flexibility (37) of the whole DNA, presumably because of altered base stacking interactions (33–36). The d(CGCGAATTCGCG)₂ dodecamer shows no intrinsic curvature supporting the explanations that hold the enhanced flexibility responsible for the increased affinity to the histone octamer (33,47). In the nucleosome core particle the majority of the protein–DNA contacts are between charged amino acid side chains and the sugar phosphate backbone (48,49). Thus, we propose that besides the flexibility the change in the $B_I/B_{II}$ pattern could also contribute to the altered histone binding. Observations at the trp repressor–operator complex in which $B_I \rightarrow B_{II}$ transitions occur synchronous to hydrogen bond breaking or formation (Wellenzohn et al., submitted for publication) and other experimental results (29,30), which attribute an active role in DNA–protein recognition to $B_I \rightarrow B_{II}$ support our suggestion. Additionally, minor groove binding ligands, which normally cause only small distortions in the DNA structure, also inhibit the formation of nucleosomes (50). Such ligands are changing the $B_I \rightarrow B_{II}$ (51,52) equilibrium, this being a possible explanation for this effect.

**METHODS AND COMPUTATIONAL DETAILS**

Molecular dynamics simulations of DNA (53–57) and DNA complexes (24,51,58–60) are providing complementary information to experiments and thus are of great interest in the field of structural biology. The inclusion of long-range interactions via the Ewald Summation in the form of the particle mesh Ewald method leads to stable B-form DNA trajectories (61–63). We performed one 3 ns long molecular dynamics simulation of the d(CGCGAATTCGCG)₂ and one 10 ns long simulation of the d(CGCGAATTCGCG)₂ (Drew Dickerson dodecamer; Fig. 1, top) dodecamer as reference system. The simulation parameters of the reference system are already described elsewhere (32). As a starting structure for the d(CGCGAATTCGCG)₂ dodecamer (Fig. 1, bottom), the crystal structure with the PDB
code 1D77 was used. Each strand of the DNA has 11 PO₄⁻ anions and in order to achieve electroneutrality 22 Na⁺ counterions were added using the program CION of the AMBER (64) package. Subsequently, solvation of the DNA with TIP3P Monte Carlo water boxes requiring a 12 Å solvent shell in all directions resulted in a system with dimensions 67.8695 × 48.7944 × 45.3952 Å³. The respective Γ-value = 176.6 (water/nucleotide). The force field parameters of inosine were selected in analogy to existing parameters in the force field of Cornell et al. (65) with the modifications of Cheatham et al. (25). The charges for inosine were derived by using the RESP standard procedure calculating the ab initio electrostatic potential with GAUSSIAN98 (66) at HF/6-31G* level of theory. As simulation protocol standard protocols (23,67–70) were adapted for our needs. The all atom force field of Cornell et al. (65) with the modifications of Cheatham et al. (25) was used. At the beginning, minimizations were carried out with harmonic restraints on DNA and counterion positions. The restraints were stepwise relaxed and at the end a 500 step minimization without restraints was performed. For equilibration the system was heated from 50 to 300 K during 10 ps under constant volume conditions and harmonic restraints. Subsequently, the restraints were once again relaxed and finally an unrestrained 5 ps constant temperature and pressure equilibration was carried out. Analysis of the resulting trajectories was performed with CARNAL, PTRAJ and different visualization programs (71,72). All calculations were performed on an SGI octane.

Figure 2. The rms-deviation with respect to the X-ray starting structure in Å as function of time. The mean value over the whole simulation is 2.2 Å, which is the normal range for such molecular dynamics simulations. The rms-values show only the normal fluctuation indicating that the system is in equilibrium.

Figure 3. The figure indicates the mean angle in (°) between the successive base pair planes. The most pronounced curvature exhibits at the AT step of the AATT tract. The solid line shows the values for the d(CGCGAATTTCGCG)₂ simulation and the dotted line gives the respective value for the d(CGCAATTTCGCG)₂ dodecamer. Both end-standing base pairs are not shown in this graph because they are too flexible.

Figure 4. The graphs show the B-factors estimated from the simulation of the d(CGCGAATTTCGCG)₂ (black line) dodecamer, the values of the d(CGCGAATTTCGCG)₂ (dotted line) DNA and of the d(CGCGAATTTCGCG)₂–netropsin complex (gray line). The figure on the top specifies the calculated values by each atom and the bottom picture gives the average for each residue. Larger peaks mean more motion.
RESULTS

The total energy of the d(CGCAATTCCGGC)₂ simulation is constant (not shown) during the simulation and the rms-deviation depicted in Figure 2 shows the normal fluctuations but no drift, indicating that the system is in the equilibrium. The mean value of the rms-deviation with respect to the X-ray starting structure is 2.2 Å.

In Figure 1 (bottom) it can be seen that one terminal C·G base pair is broken after \( \sim 500 \) ps (is not responsible for the enhancement in the rms-deviation at 250 ps) and stays open for the rest of the simulation, thus they are not considered in this work. This occurs often in molecular dynamics simulations of DNA with explicit water molecules. The rest of the molecule remains unaffected.

The wrapping of nucleic acids by the histone octamer is a mechanical deformation of the DNA, which is also influenced by the exocyclic groups that penetrate into the minor groove (33–36). Most of the contacts are between the sugar phosphate backbone and the amino acid side chains, thus it is unlikely that direct contacts with the minor grooves are influencing the protein–DNA interaction. Figure 3 shows the angles between the successive base pair planes for d(CGCGAATTCGCG)₂ and d(CGCAATTCCGGC)₂ in order to investigate whether the guanine → inosine transformation leads to an intrinsic curvature.

The most pronounced curvature exhibits the ApT step of the ApApTpT tract. The angles of the base pair planes of the CipA steps are neither significantly larger than in the case of the corresponding CpGpA steps nor do they show a distinctive absolute value. The sizes of the DNA molecules defined as the average distance (in the simulation) between the center of mass of the second base pair (first pair is opened as mentioned above) and the center of mass of the last base pair are also an indicator of the amount of bending. The value for the d(CGCGAATTCGCG)₂ simulation is 33.5 Å and for d(CGCAATTCCGGC)₂ the respective value is 34.3 Å which is <3% difference. These results show that inosine does not enlarge the bending of the dodecamer.

Buttinelly et al. (33) suggest that the removal of exocyclic groups of the bases increases the deformability of DNA by minimizing the steric resistance to bending. This enlarged deformability should be the explanation for the better binding to the histone octamer. The drug-induced inhibition of the inherent DNA flexibility is up to now also the most plausible mechanism for the fact that minor groove binders are able to inhibit the formation of the histone–DNA complex (50). We calculated the B-factors (Fig. 4) as an indication for the positional fluctuations for both simulations and for a recently published d(CGCGAATTCGCG)₂–netropsin complex (24; Wellenzohn et al., submitted for publication). Netropsin is a widely studied (73–83) minor groove-binding ligand, which selectively recognizes AT-rich regions.

Figure 4 clearly shows that the d(CGCAATTCCGGC)₂ dodecamer exhibits the largest B-values and thus agrees with the above-mentioned suggestions. The peaks in the left, middle and right represent the ends of the DNA helices and the periodic peaks (top) represent the backbone phosphates. The
largest fluctuation exhibits the d(CGCIATTCGCCG)₂ dodecamer on the left side which is due to the above-mentioned opening of the base pair. The enlarged flexibility of the rest of the DNA seems not to be an artifact of this base pair opening because the other end of the DNA (middle), which is 12 bp distant from the opened one, also shows a higher positional fluctuation. The influence of the minor groove binding ligand netropsin is by far much less distinct than the guanine → inosine substitution. Netropsin does not substantially affect the flexibility, being in agreement with recent theoretical results which showed that the hydration of the exocyclic groups in both the major and the minor groove are contributing to the $B_I \rightarrow B_{II}$ equilibrium. Thus, the removal of an amino group in the minor groove (= guanine → inosine transformation) should alter the conformational behavior of the phosphates. Figure 7 (top) shows the $B_I \rightarrow B_{II}$ pattern of the d(CGCGAATTCGCG)₂ Drew Dickerson dodecamer and of the d(CGCIATTCGCCG)₂ DNA.

The substitution from guanine to inosine changes the $B_I/B_{II}$ equilibrium substantially. Phosphate number 4 is in the case of a CpG step nearly always in $B_I$ while the respective CpG step has a distinctive preference for the $B_{II}$ state. Phosphate 5 of the CpG step also changes the substate due to the guanine → inosine transformation. Phosphate 16, which also represents a CpG step, shows an enhanced preference for the $B_{II}$ substate, but this phosphate very often interconverts between $B_I$ and $B_{II}$. This may be because phosphate number 17 is most of the time in $B_{II}$ and it is very unlikely that two successive phosphates occupy the $B_{II}$ state at the same time. The results underline that the hydration of the grooves influences the backbone conformation of DNA and we propose again that a conversation between the major and minor groove could be mediated by these $B_I/B_{II}$ substrates.

**SUMMARY AND CONCLUSIONS**

We performed two molecular dynamics simulations; one of the d(CGCGAATTCGCG)₂ Drew Dickerson dodecamer and one of the d(CGCIATTCGCCG)₂ DNA. The transformation from guanine to inosine leads to the removal of an exocyclic amino group in the minor groove of the nucleic acid. Such exocyclic groups have already proven their influence on the binding of the DNA to the histone protein. In the complex, the DNA structure is strongly disrupted and the enlarged flexibility due to the removal of the amino group is used as an explanation for the enlarged complexation affinity. Our simulations show that the substitution of guanine by inosine in the Drew Dickerson dodecamer does not introduce an intrinsic bending but indeed
introduces a higher flexibility to the whole DNA. The calculated B-factors are enhanced not only at the place of the guanine → inosine transformation and we propose that the differences in the hydration of the minor groove changes the stacking interactions that are responsible for this effect. The differences in the hydration also influence the conformation of the phosphate backbone according to Flader et al. (32) The phosphates of the Cpl steps show a high preference for the B II state while the respective CpG step is mainly in B I. As a result of the fact that the histone preferentially contacts the backbone of the DNA, we suggest an influence of the B I/BII backbone substrate pattern on the nucleosome formation process. Thus, the correlation between the B I/BII substrate pattern and the exocyclic amino groups provides an additional explanation for this effect. Altogether the exact conformational behavior of the backbone phosphates could be of extraordinary interest for DNA recognition processes.

Acknowledgement

This work was supported by a grant of the Austrian Science Fund (grant number P13845-TPH).

REFERENCES


73. Lah,J. and Vesnaver,G. (2000) Binding of distamycin a and netropsin to the 12mer DNA duplexes containing mixed AT,GC sequences with at most five or three successive AT base pairs. Biochemistry, 39, 9317–9327.


