Published in final edited form as:

Chembiochem. 2013 February 11; 14(3): 323–331. doi:10.1002/cbic.201200706.

Molecular Basis For Sequence-Dependent Induced DNA Bending

Dr. Michael Rettig, Prof. Markus W. Germann, Shuo Wang, and Prof. W. David Wilson

Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA

Markus W. Germann: mwg@gsu.edu

Abstract

With a growing understanding of the microstructural variations of DNA it has become apparent that subtle conformational features are essential for specific DNA molecular recognition and function. DNA containing an A-tract has a narrow minor groove and a globally bent conformation but the structural features of alternating AT DNA is less understood. Several studies indicate that alternating AT sequences are polymorphic with different global and local properties than A-tracts. The mobility of alternating AT DNA in gel electrophoresis is extensively reduced upon binding with minor groove binding agents such as netropsin. Although this suggests that such complexes are bent, similarly as A-tract DNA, direct evidence and structural information on AT DNA and the induced conformational change is lacking. We have used NMR and residual dipolar coupling together with restrained molecular dynamics simulations to determine the solution structures of an alternating AT DNA segment, with and without netropsin, in order to evaluate the molecular basis of the binding induced effects. Complex formation causes a significant narrowing of the minor groove and a pronounced change in bending, from a slight bend towards the major groove for the free DNA to a pronounced minor groove bend in the complex. This observation demonstrates that conformational features and the inherent malleability of AT sequences are essential for specific molecular recognition and function. These results take the field of DNA structures into new areas while opening up avenues to target novel DNA sequences

Keywords

DNA structure; ligand binding DNA bending; NMR

Introduction

Local variations in structure have been recently established to be important in specific recognition and function of DNA sequences.^[1,2] Intrinsic and induced bending of DNA, for example, is a common recognition mechanism for conformational control of gene expression through complex formation.^[3–9] “Bendability” of a DNA sequence can influence both protein recognition specificity and affinity for a DNA sequence^[3,7–9]. Such induced bending typically involves AT rich DNA sequences and can involve both direct and

Correspondence to: Markus W. Germann, mwg@gsu.edu.

Accession Numbers

PDB coordinates and NMR restraints have been deposited. Complex: PDB ID (2LWH) and RCSB ID (RCSB102918). Control: PDB ID (2LWG) and RCSB ID (RCSB102917).

Supplementary Data

Supplementary tables 1–7. Proton assignments for free d(GGATATATCC)₂, complexed d(GGATATATCC)₂, notable chemical shift differences, proton assignment of bound netropsin, intermolecular netropsin DNA NOE contacts, NOE distance monitors and RDC restraints of free d(GGATATATCC)₂ complexed d(GGATATATCC)₂. Supplementary figures 1–2, of NOESY base H1 pathway free d(GGATATATCC)₂ and RMSD plots of the trajectory of free and netropsin complexed DNA.

Supporting information for this article is available on the WWW under <http://www.chembiochem.org>

indirect interactions, although DNA modifications can also cause helix bending (for example, [10]). Conversion of a DNA sequence from the native state to a bent conformation in the complex contributes to the overall free energy of binding. Therefore the malleability of DNA impacts recognition specificity and has direct functional benefit. [3,5,7,11] AT sequences of the same composition have two limiting sequence arrangements, either A-tracts, with all A bases on one strand and all T on the other, or alternating A and T bases on both strands. Extensive experimental evidence indicates that these two sequence arrangements have quite different structures and properties. [12–15] A-tract sequence-associated DNA curvature was discovered based on electrophoretic anomalies of sequences from the AT-rich mitochondrial kinetoplast DNA (kDNA) of some parasitic microorganisms. [16–19] Bending of A-tracts is strongly supported by polyacrylamide gel electrophoresis (PAGE) of DNAs of a variety of sequences and lengths, [18–24] by FRET analysis of fluorescently labeled DNA oligomers [5] and by X-ray and NMR structural investigations of oligomer DNAs. [15,25,26]

Studies of alternating AT structural variations have, however, been less extensive than with A-tracts. Several studies have shown that the alternating AT DNA has a more “B-like” structure and physical properties. [12,14,15,24] The alternating AT sequence, for example, migrates more like straight DNA sequences than like curved A-tract DNA in PAGE. [14,24] Hydroxyl radical and DNase I cleavage patterns are quite sensitive to minor groove width and cleavage results with alternating AT DNA sequences are similar to those of B-form DNA sequences but not to those of A-tract DNA. [12] X-ray and NMR structures of DNA duplexes with A-tracts have narrow minor grooves while alternating AT structures have been observed with both narrow and wide minor grooves and a variety of polymorphic structures. [25–30]

As part of an effort to develop new antiparasitic drugs to target the AT-rich mitochondrial, kDNA of kinetoplastid parasites, the structure-dependent global conformational changes induced by a number of minor groove targeting compounds were evaluated for both alternating AT and A-tract DNA sequences. [14,24,27] PAGE results suggest that the compounds typically leave the curvature of intrinsically bent A-tract DNA nearly unchanged. In contrast, significant compound induced mobility decreases that correlate with bending distortions are observed for alternating AT sequences. [14,24] Ligand induced changes in bending are an important new observation since such bending can interfere with gene expression and contribute to antiparasitic biological activity. [28] DNA curvature can be due to a global feature of nucleic acids associated with the summation of small influences on the topology for individual base pair steps or it can be due to more dramatic effects such as local kinks in the helical conformation. A key unanswered question from the PAGE observations on mobility changes is what is the molecular basis of compound induced structural changes in alternating AT DNA? A high resolution structure of the alternating AT sequence and its complex with a minor groove binder is a critical missing piece of current information in our understanding of local DNA microstructures. We have used NMR including residual dipolar coupling (RDC) methods together with restrained MD simulations to determine the structures of the free DNA and complexes. Specific questions to answer are: (i) is the gel mobility reduction actually due to DNA induced bending by minor groove complex formation; (ii) what are the specific changes in the DNA helical structure that cause the conformational change that leads to the gel anomaly; and (iii) is the conformational change a local kink or a more global effect and what is its magnitude? The results presented here answer these questions and take the field of local structural variations of DNA forward in an important new area. We have previously shown that minor groove binding ligands exhibit microscopic rearrangements on the millisecond time scale without dissociating off the DNA into the bulk solution. [30] This rearrangement results in chemical exchange processes and NMR signals that greatly complicate structure determination. It is therefore

necessary to explore solution conditions and minor groove binding ligands where exchange process are minimal. Netropsin exhibits the most favorable dynamic properties and was therefore chosen as a representative ligand to investigate induced bending in alternating AT DNA sequences.

Results and Discussion

Gel Migration Studies

In order to compare PAGE and NMR results, the same DNA sequence was used for both methods (Figure 1). The alternating AT sequence was evaluated in ligation ladder PAGE experiments with and without netropsin (Figure 1D). The migration patterns obtained were compared to A-tract DNA and two mixed sequence straight DNAs as curvature references for bent and straight DNA, respectively. [24] It is clear from Figure 1A and the R_L (mobility observed / mobility predicted for a straight DNA of the same number of base pairs, L_{act}) versus L_{act} plot for the free DNA and netropsin complex in Figure 1C that the free ATATAT sequence migrates slightly slower than both of the mixed DNA sequences whereas the netropsin-DNA complexes are significantly slower and more comparable to the A-tract DNA migration rate. Through the comparison of mobility of the 189 bp AAAAA and 180 bp Net-ATATAT complex (indicated by arrows in Figure 1A), the netropsin binding induced bending in ATATAT is calculated to be $\sim 15^\circ$.

^1H NMR Resonance Assignments and Structure Calculation

Several (AT)₃ sequences were tested for NMR experiments and the self-complementary 10 mer, d(GGATATATCC)₂, showed favorable NMR signal dispersion for NMR structure analysis (data not shown). The imino proton spectra of the 10-mer sequence in the presence of netropsin and the base pair numbering scheme are shown in Figure 2. At ligand to DNA ratios below 1:1 both free and complexed DNA coexist in solution showing that exchange between bound and free DNA is slow on the NMR time scale. Tight binding of netropsin is also manifested in the absence of free DNA at the final 1:1 titration step. As the binding of netropsin lifts the intrinsic symmetry of the self-complementary oligonucleotide the five imino protons seen in the free duplex spectrum are doubled in the final titration step (from top to bottom spectra in Figure 2). However, in addition to the slow exchange between free and bound DNA, netropsin similar to other minor groove binding ligands, exhibits microscopic rearrangements on the millisecond time scale. [30] In order to minimize this process we reduced the salt concentration to 20mM NaCl and temperature to 278K. Under these conditions, exchange cross peaks visible in NOESY and ROESY experiments between symmetry related protons (i.e. T8/T18 methyl protons etc) are greatly reduced (data not shown).

DNA signals were assigned using standard protocols for B-DANN. [10,29,30,39] Briefly, continuous base-to-sugar NOESY networks between neighboring residues were used for the sequential assignment of both the free DNA (Supplementary Material, SM Figure 1) and netropsin complex (Figure 3A). In addition, DQF-COSY spectra were used to support the assignment of intraresidue sugar protons. It is noted that the low temperature together with significant overlap of the sugar protons of the nonequivalent, but chemically identical strands, hindered a pseudorotation analysis. The netropsin resonances were identified following the assigning of the DNA resonances. Overviews of the chemical shifts of both the exchangeable and the non-exchangeable protons of the free DNA as well as the netropsin-DNA complex are given in tables (SM Tables 1–4). The fixed distance between the 2 and 4 positioned pyrrole protons of netropsin also served as an independent verification of the validity of the applied NOE restraints (described below).

Modeling and Structure Calculations

Distance and RDC restraints were determined as described in Methods (refinement statistics are in Table 1). To exclude a bias of the initial geometry on the final models, both A- and B-type DNA models of the drug-DNA complex were built and subjected to simulated annealing using the Born implicit solvent model with NOE and Watson-Crick restraints. The initial mass weighted rmsd values of 5.6 Å between the A- and B-type starting structures were greatly reduced to 1.2 Å at the end of the second simulated annealing run showing that the force field and the applied NOE restraints are sufficient to drive different initial geometries towards a common final structure. After charge neutralization with Na⁺ ions, solvation with TIP3P water molecules and several equilibration steps, the RDC restraints were added as described in Methods. The fully restrained MD simulations were then performed for the two netropsin-DNA complex structures. Only one model was used for the free DNA structure.

Rmsd plots showed that after 2.8 ns (free DNA) and 0.5 ns (DNA-netropsin complex) the solvated and fully restrained systems were equilibrated. After omitting the equilibration periods the two trajectories of the netropsin-DNA system were merged. Rmsd fitting was applied to the average structure of all frames (SM Figure 2) clearly showing that the two netropsin-DNA runs have essentially converged towards common final structures. The following analysis with Curves+ and its utility program Canal was thus based on the full 5 ns and 19 ns of the rMD trajectories of the free DNA and the DNA-netropsin complex, respectively. For visualization purposes, one final representative structure was chosen for each individual system by the method described below.

After averaging of all production run snapshots, the snapshot having the lowest rmsd from the average was selected. This structure is both physically meaningful and most closely represents the structural average of the trajectory. To remove random thermal fluctuations, a 10000 step energy minimization in explicit solvent was performed and the obtained model of the netropsin-DNA complex is shown in Figure 3B. An inspection of the NMR ensembles (Figure 4) shows that the minor groove of the free DNA is significantly larger and also hints that the DNA is more dynamic. The larger backbone variation compared to the netropsin-DNA complex can be due to the nature and number of the constraints or actual differences in dynamics. Free MD simulations support the notion that the minor groove topology of the free DNA is more dynamic (data not shown).

DNA Chemical Shifts Changes in the Complex

The intermolecular NOE contacts clearly reflect the binding of netropsin in the minor groove of the duplex (Figure 3A, SM Table 5) and significant chemical shift changes were observed for some of the DNA resonances upon binding of netropsin (Figure 5A, B, SM Table 3). As netropsin fully covers the central four AT base pairs, chemical shifts changes are especially pronounced for protons in that region. The pronounced upfield shifts for the T6, A7, T16 and A17 H1' sugar protons can be rationalized by their positioning in the shielding region of netropsin's pyrrole rings (Figure 5C). In contrast, the H1' sugar protons of T4, A5, T8, T14 and A15 are not in proximity of the pyrrole groups and their observed chemical shift changes are generally much smaller. Lying within the deshielding region of the pyrrole rings in the floor of the minor groove, significant downfield shifts are observed for the A5, A15 and, to a somewhat lesser extent, for the A17 H2' protons. Significant downfield shifts can also be observed for the T4, T14 and T16 imino protons. Taken together these results show the good correlation between the derived structure and experimental NMR raw data and highlight the high quality of the derived solution structure.

Final Structure Analysis

Both the free DNA and the netropsin-DNA complex share the structural characteristics of B-DNA with South type sugar puckers and glycosyl torsion angles in the anti-range (Figure 4). The free DNA exhibits base stacking features similar to those observed in X-ray structures. As such, the stacking between adenine and thymine bases (AT steps) in the central AT-tract is optimized, whereas stacking at the TA steps are interrupted. Binding of netropsin results in changes in the TATA segment to optimize interactions with the ligand. Compared to free DNA the complex is over wound by 23° (Figures 6 A, B).

The twist changes are not evenly distributed but are localized to the TA steps while the central AT step is essentially unaffected. Binding of netropsin lifts the symmetry of the self-complementary oligonucleotide substrate as is evident from the NMR data (Figure 2,3,5). At the base pair level the most pronounced changes are seen in the roll and twist changes (Figures 6, 7). The average minor groove widths of the free and complexed DNA are shown in Figure 7A. The larger standard deviations of the free DNA illustrate its higher structural flexibility in the rMD runs despite the larger number of restraints. It is clear from the plot that the maximization of close contacts between netropsin and the DNA significantly narrows the minor groove in the duplex center. At the same time, netropsin binding slightly pushes the interacting base pairs toward the major groove and changes the roll angles in the netropsin binding site (Figure 7A).

The bend angles were determined by analyzing the production runs of the free and complexed DNA using Curves+ and Canal.^[38] Plots of these bend angles are shown in Figure 7A. An average value of 8° into the major groove was found for the free DNA while surprisingly the netropsin-DNA complex was bent 18° into the minor groove. Overlays of the final models for free DNA and the netropsin-DNA complex are shown in Figure 7B along with helix axis traces from Curves+. The dramatic increase in curvature of the DNA in its netropsin complex is easily seen. An analysis of the production runs and representative structures reveals that the free DNA is slightly bent towards the major groove. In contrast, the bending of the netropsin-DNA complex is into the minor groove.

Discussion

Induced PAGE mobility changes of alternating AT sequences on binding is a general property of minor groove binding agents.^[14,24] including the 10 mer duplex used in the NMR structural analysis (Figure 1). The compounds affect the intrinsically bent A-tract DNA sequences much less. Given that the free AT DNA is essentially straight, one explanation for the PAGE mobility change is bending of the AT DNA due to binding effects on minor groove width. If free AT sequences have a wider groove than typical A-tracts, as predicted by cleavage studies^[12], then binding of netropsin would narrow the groove in order to create an optimal binding cavity and make van der Waals contacts with the walls of the minor groove. The solution structure of an alternating AT repeat sequence and its complex with a minor groove binder have not been determined by current high resolution 2D NMR methods with RDCs. X-ray results with ATATAT DNA sequences are variable with both wide and narrow grooves.^[15] Previous NMR results for alternating AT sequences are also quite varied with left-handed and right-handed B as well as D-form structures proposed.^[39–43]

In order to test the hypothesis of minor groove narrowing as the basis of induced bending of (AT)_n DNA sequences by minor groove binders, and to determine the solution structure of free (AT)_n for comparison with A-tract sequences, the NMR studies reported here were undertaken. Both the free DNA and its complex with netropsin gave well resolved 2D NMR spectra for structural analysis. A number of other minor groove binding agents were also

evaluated with this sequence but exchange rates limited the quality of structural information for many of them. The structure of the free ATATAT DNA sequence has a much more linear helix axis (Figures 1, 7A, B) than that found in A-tract sequences. [20,22,24] In the X-ray structures of the free DNA the stacking of the bases in A-T steps was much better than in T-A steps so that the helical twist was high at T-A and low at A-T steps. [26] This variation in stacking and twist is also seen in the solution structure. As expected from the gel results in Figure 1, the bend angle for the free DNA (Figures 7A, B) is small and in reasonable agreement with the value predicted from quantitative analysis of PAGE results.

The structure of the netropsin-DNA complex differs from the free DNA in several interesting ways. First, the minor groove width is reduced through the central AT sequences to roughly the width of the netropsin conjugated heterocyclic-amide system (Figure 7A). As can be seen in Figure 3B, this lets the entire netropsin molecule fit snugly between the walls of the minor groove. The netropsin amide protons form H-bonds with A(N3) and T(O2) at the floor of the minor groove and the pyrrole rings pack against the AH2 protons as illustrated in the view in Figure 5C. These interactions are critical for the AT specific binding of netropsin. The narrowing of the minor groove causes an increase in curvature of the helix axis to approximately 18° (Figure 7). This value is essentially the same as the intrinsically bent A-tract DNA and is in excellent agreement with predicted results from PAGE ligation ladder experiments.

The free DNA and the netropsin-DNA models obtained from NMR results clearly show that the changes in PAGE mobility on binding netropsin in the minor groove are due to the summation of local bending at the binding sites. The reason for the bending of (AT)_n sequences by netropsin can now be defined at the molecular level. Netropsin is too small to cover a full turn of the DNA helix and thus, its local effect on the DNA structure is directional. The narrowing of the minor groove requires the backbone to counterbalance the movement towards netropsin. As the backbone cannot increase its length, the walls of the minor groove move toward the interaction site with netropsin to improve the binding energetics and groove narrowing leads to an overall bend into the minor groove. The local helical parameter that best correlates with the induced bending is the roll angle (Figure 7A). The free DNA has some positive roll values in the duplex center, but binding of netropsin reverses this and gives a large negative roll at this position. As a positive roll value is defined as an opening of the minor groove (versus an opening of the major groove for a negative value) between two adjacent base pairs, these combined changes lead to the overall bend towards the minor groove.

In summary, netropsin and several other minor groove binding dications interact very strongly with A-tract sequences but do not significantly change the PAGE mobility. [14,24] The reason for this observation is now quite clear: A-tracts already have a narrow minor groove and can bind minor groove agents without a significant average change in groove width or local helix axis bends in solution. Straight alternating AT sequences on the other hand require the minor groove to narrow with bending of the helix on binding minor groove agents. These results provide new insight into the molecular basis of minor groove recognition.

Experimental Section

Sample Preparation

The self-complementary 10 mer DNA oligonucleotide 5'-GGATATATCC-3' was purchased from Integrated DNA Technologies, Inc. (IDT, Coraville, IA) with HPLC purification. The sample was dialyzed three times against DI water using a Spectra/Por (Spectrum Laboratories, Inc., Rancho Dominguez) dialysis membrane with a cutoff of 1,000 Daltons,

then lyophilized and finally redissolved in BPES buffer (20 mM NaCl, 20 mM phosphate, 0.1 mM EDTA, pH 7.0). The final duplex concentration was about 0.9 mM for NMR experiments. For experiments involving exchangeable protons, a 90% H₂O-10% D₂O (Sigma Aldrich, St. Louis) sample was used. The DNA-compound complex was formed by titrating netropsin (Sigma Aldrich, St. Louis) dissolved in D₂O into the DNA solution until saturation of the oligonucleotide as judged by the imino proton spectral region. For experiments with non-exchangeable protons the DNA sample was lyophilized twice and redissolved in 99.994% D₂O (Isotec, Miamisburg). The same amount of netropsin as for the water sample was added. For the ligation ladders in PAGE, the procedure is as described previously [14, 27].

Phage pf1 was purchased from Asla (Asla Biotech Ltd., Latvia) and prepared as suggested by the manufacturer. The phage was exchanged into the D₂O buffer through a series of ultracentrifugation washing steps and finally added to the existing NMR sample giving a deuterium splitting at 278 K of 10.4 and 8.1 Hz for the free DNA and the netropsin-DNA complex, respectively. The final sample concentration was 1.2 mM for the free DNA and 0.6 mM for the complex.

Gel Electrophoresis and bending calculations

DNA ligation ladders of ATATAT and AAAAA oligomers (Figure 1) were analyzed on 8% native polyacrylamide gels run with 1xTBE at 25°C. AAAAA tract bends the DNA overall structure by 18° per helical turn. [20]. Therefore, phased A5 containing multimers are used as a reference for the calculation of netropsin induced bending angle as described previously. [22,24] Briefly, R_L values for ligated multimers in range of 105 bp L_{actual} 189 bp were averaged over three gel experiments for AAAAA. Data are fitted to $R_L - 1 = (a - b) C_r^2$, where C_r is the relative curvature and equals to 1.0 for AAAAA by definition.

NMR experiments and NOE distance restraints

The NMR data were collected on a Bruker Avance 600 MHz NMR spectrometer equipped with a QXI probe. HSQC experiments of the netropsin-DNA complex were recorded on a Bruker Avance 500 MHz NMR spectrometer equipped with a TXI cryoprobe. The NMR experiments were performed at 278 K using the water resonance as reference (5 ppm). Water suppression for 1D titrations was achieved by applying a 1-1 pulse sequence. Phase sensitive 2D NOE experiments were collected with 2048 × 800 data points in the two dimensions and 64 scans per t_1 increment. A mixing time of 100 ms was used. For samples in H₂O the recycle delay was set to 2.5 s and the spectral width was 11.5 kHz. A WATERGATE 3-9-19 scheme was used for solvent suppression. For NOESY experiments in D₂O the spectral width was reduced to 4.8 kHz. Suppression of the residual water signal was achieved by applying a weak presaturation pulse during mixing time and the 4 s recycle delay. Linear prediction with 25 coefficients in t_1 and zero filling gave a symmetrical 4 K × 4 K matrix. Both dimensions were apodized with shifted (/2) squared sine bell functions.

Phase sensitive DQF-COSY experiments with 4096 × 800 data points in t_2 and t_1 , 64 scans per t_1 increment and a spectral width of 4.8 kHz were additionally recorded for the D₂O sample. The residual HDO signal was suppressed by a weak presaturation pulse during the 3 s recycle delay. Linear prediction with 20 coefficients in t_1 and zero filling gave a symmetrical 4 K × 4 K matrix. Both dimensions were apodized with shifted (/3) squared sine bell functions.

RDC restraints were obtained from high resolution F_2 coupled ¹³C-¹H HSQC experiments. Two sets of experiments were recorded which were individually optimized for the DNA base and sugar regions (SM Table 7). The F_1 dimension was folded and experiments were

typically run with 2048×400 data points in t_2 and t_1 , 120 scans per t_1 increment and a spectral width of 3.9 kHz in the ^1H and 5 kHz in the ^{13}C dimension. The residual HDO signal was suppressed by a weak presaturation pulse during the recycle delay (2–5 s). Linear prediction with 25 coefficients in t_1 and zero filling gave a symmetrical $2\text{ K} \times 2\text{ K}$ matrix. Both dimensions were apodized with shifted ($1/2$) squared sine bell functions.

Distance restraints were calculated from cross-peak volumes of NOESY experiments at 100 ms mixing time. SPARKY [31] was used for integration with the cytosine H5–H6 cross-peak (2.45 Å) used as a standard. These reference cross peaks were within 5% – 10% of each other for the control and complex, indicating that these residues experience the same correlation time. Depending on peak quality error bounds of $\pm 25\%$ and $\pm 30\%$ were assigned to derived distances. No lower limits but upper error bounds of 5.5 Å were applied to strongly overlapped peaks, those affected by the presaturation pulse and to all exchangeable protons. The validity of the distances determined from the NOESY data was examined by monitoring the results obtained for invariant distances in netropsin and distances that are insensitive to structure such as the deoxyribose H2 –H4 distance (SM Table 6). Watson-Crick hydrogen bond restraints [32] were allowed to vary by ± 0.2 Å and have been included to maintain proper base pairing especially during the simulated annealing. The NOE restraints were implemented into AMBER using a well potential. For the RDC restraints, one bond ^{13}C - ^1H couplings were determined from the difference in $^1J_{\text{CH}}$ values from the F2 coupled natural abundance ^{13}C - ^1H HSQC spectra of the aligned and unaligned free DNA and netropsin-DNA samples. In accordance with the NOE restraints, peak quality dependent error bounds of ± 0.75 (1), 1.25 (1.5) and 2 (2) Hz were applied to the free DNA and netropsin-DNA complex.

Molecular Modeling

The self-complementary 10 mer d(GGATATATCC)₂ was built as both A- and B-type DNA using the NUCGEN application included in the AMBER 9 [33] software package and both were used as starting structures for the simulation of the DNA-netropsin complex. The starting geometry for netropsin was derived from geometry optimization in vacuum at the B3LYP/6-31G* level with GAUSSIAN 03. [34] Using this structure and the RESP approach, the partial charges for netropsin were computed at the HF/6-31G* level. The compound was then hand-docked into the minor groove most closely covering the central ATATAT steps of both oligonucleotides. To simulate the DNA as well as solvent and sodium ions, the PARMBSC0 force field [35] was used. Except for the terminal amidinium and guanidinium groups which were simulated by ff99 as described by ref [36], GAFF [37] was used for the netropsin molecule.

To remove bad initial contacts, simulated annealing incorporating a generalized Born implicit solvent model was performed for both the free DNA as well as the DNA-netropsin complexes with the AMBER software package. After an initial 5000 step energy minimization the structures were heated to 800 K and kept at this temperature for 30 ps. Over the next 50 ps the system was gradually cooled down to 100 K followed by 2 ps with the systems temperature set to 0 K. During the first 25 ps of simulation the force constants were linearly increased from $1.5\text{ kcal mol}^{-1}\text{ Å}^{-2}$ and $2.5\text{ kcal mol}^{-1}\text{ Å}^{-2}$ to their final values of $30\text{ kcal mol}^{-1}\text{ Å}^{-2}$ and $50\text{ kcal mol}^{-1}\text{ Å}^{-2}$ for the NMR and Watson-Crick restraints, respectively. The force constants were kept constant for the rest of the simulated annealing. The procedure was completed by a 5000 step energy minimization and performed twice for each structure.

For the simulations in explicit solvent the Watson-Crick restraints were removed from all but the terminal two base pairs. For these pairs, the force constants were brought to $30\text{ kcal mol}^{-1}\text{ Å}^{-2}$ and are in line with the NMR restraints. In addition to sodium ions for charge

neutralization, 4500 preequilibrated TIP3P water molecules in a truncated octahedral box were added to the system. Prior to the production run, the system was equilibrated by an energy minimization with the DNA or netropsin and DNA atom positions fixed followed by a minimization of the whole system. The system was then slowly heated to 300 K with weak positional restraints on the DNA or DNA and drug atoms under constant-volume conditions. The system was then allowed to run for 3 ns in the NPT ensemble. The particle mesh Ewald method was used to evaluate the electrostatic interactions with a direct space sum cutoff of 10 Å. Time steps of 1 and 2 fs were employed for the free DNA and the netropsin-DNA complex, respectively, as the bond length involving hydrogen atoms were kept fixed with the SHAKE algorithm. As RMSD comparisons showed that the system was fully equilibrated after 1.5 ns, a representative structure was chosen out of the last 1.5 ns for the free DNA and DNA-netropsin systems. After energy minimization the RDC restraints were fitted against the fixed DNA and DNA-ligand complex thus obtaining the initial alignment tensor. Consecutive runs of short energy minimizations using the RDC data followed by refitting with fixed atom coordinates were additionally performed to improve the alignment tensor. The final production run in explicit solvent was performed using the optimized alignment tensor as well as all RDC and NOE restraints.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institute of Health [NIAID AI 64200 to W.D.W.].

References

1. Bishop EP, Rohs R, Parker SCJ, West SM, Liu P, Mann RS, Honig B, Tullius TD. ACS Chem Biol. 2011; 6:1314–1320. [PubMed: 21967305]
2. Rohs R, West SM, Sosinsky A, Liu P, Mann RS, Honig B. Nature. 2009; 461:1248–1253. [PubMed: 19865164]
3. Watkins D, Mohan S, Koudelka GB, Williams LD. J Mol Biol. 2010; 396:1145–1164. [PubMed: 20053356]
4. Peters JP, Becker NA, Rueter EM, Bajzer Z, Kahn JD, Maher LJ III. Methods Enzymol. 2011; 488:287–335. [PubMed: 21195233]
5. Dragan AI, Privalov PL. Methods Enzymol. 2008; 450:185–199. [PubMed: 19152861]
6. Petrov AS, Harvey SC. Biophys J. 2008; 95:497–502. [PubMed: 18487310]
7. Stella S, Cascio D, Johnson RC. Genes Dev. 2010; 24:814–826. [PubMed: 20395367]
8. Luijsterburg MS, Noom MC, Wuite GJL, Dame RT. J Struct Biol. 2006; 156:262–272. [PubMed: 16879983]
9. Crothers DM. Proc Natl Acad Sci U S A. 1998; 95:15163–15165. [PubMed: 9860938]
10. Wu Z, Maderia M, Barchi JJ Jr, Marquez VE, Bax A. Proc Natl Acad Sci U S A. 2005; 102:24–28. [PubMed: 15618396]
11. Chen B, Young J, Leng F. Biochemistry. 2010; 49:1590–1595. [PubMed: 20108983]
12. Fox KR. Nucleic Acids Res. 1992; 20:6487–6493. [PubMed: 1336176]
13. Abudaya A, Brown PM, Fox KR. Nucleic Acids Res. 1995; 23:3385–3392. [PubMed: 7567447]
14. Hunt RA, Munde M, Kumar A, Ismail MA, Farahat AA, Arafa RK, Say M, Batista-Parra A, Tevis D, Boykin DW, Wilson WD. Nucleic Acids Res. 2011; 39:4265–4274. [PubMed: 21266485]
15. Yuan H, Quintana J, Dickerson RE. Biochemistry. 1992; 31:8009–8021. [PubMed: 1510987]
16. Marini JC, Levene SD, Crothers DM, Englund PT. Cold Spring Harb Symp Quant Biol. 1983; 1:279–283. [PubMed: 6574847]

17. Kitchin PA, Klein VA, Ryan KA, Gann KL, Rauch CA, Kang DS, Wells RD, Englund PT. *J Biol Chem.* 1986; 261:11302–11309. [PubMed: 3015964]
18. Cons BM, Fox KR. *Biochem Biophys Res Commun.* 1990; 171:1064–1070. [PubMed: 1699520]
19. Wilson WD, Tanious FA, Mathis A, Tevis D, Hall JE, Boykin DW. *Biochimie.* 2008; 90:999–1014. [PubMed: 18343228]
20. Crothers DM, Drak J, David MJL, James ED. *Methods Enzymol.* 1992; 212:46–71. [PubMed: 1518459]
21. Diekmann S, David MJL, James ED. *Methods Enzymol.* 1992; 212:30–46. [PubMed: 1518451]
22. Ross ED, Den RB, Hardwidge PR, Maher LJ III. *Nucleic Acids Res.* 1999; 27:4135–4142. [PubMed: 10518603]
23. Maki AS, Kim TW, Kool ET. *Biochemistry.* 2004; 43:1102–1110. [PubMed: 14744156]
24. Tevis DS, Kumar A, Stephens CE, Boykin DW, Wilson WD. *Nucleic Acids Res.* 2009; 37:5550–5558. [PubMed: 19578063]
25. Hud NV, Feigon J. *Biochemistry.* 2002; 41:9900–9910. [PubMed: 12146955]
26. Fratini AV, Kopka ML, Drew HR, Dickerson RE. *J Biol Chem.* 1982; 257:14686–14707. [PubMed: 7174662]
27. Wang S, Munde M, Wang SM, Wilson WD. *Biochemistry.* 2011; 50:7674–7683. [PubMed: 21800847]
28. Wilson WD, Nguyen B, Tanious FA, Mathis A, Hall JE, Stephens CE, Boykin DW. *Curr Med Chem: Anti-cancer Agents.* 2005; 5:389–408.
29. Wüthrich, K. *NMR of Proteins and Nucleic Acids.* Wiley-Interscience; New York: 1986. p. 203-233.
30. Rettig M, Germann MW, Ismail MA, Batista-Parra A, Munde M, Boykin DW, Wilson WD. *J Phys Chem B.* 2012; 116:5620–5627.
31. Goddard, TD.; Kneller, DG. *SPARKY 3.* University of California; San Francisco: 2008.
32. Saenger, W. *Principles of Nucleic Acid Structure.* Springer-Verlag; New York: 1984.
33. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Merz, KM.; Pearlman, DA.; Crowley, M.; Walker, RC.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, KF.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, DH.; Schafmeister, C.; Ross, WS.; Kollman, PA. *Amber 9.* University of California; San Francisco: 2006.
34. Frisch, MJ.; Trucks, GW.; Schlegel, HB.; Scuseria, GE.; Robb, MA.; Cheeseman, JR.; Montgomery, JA., Jr; Vreven, T.; Kudin, KN.; Burant, JC.; Millam, JM.; Iyengar, SS.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, GA.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J. *Gaussian 03.* Gaussian, Inc; Wallingford: 2003.
35. Perez A, Marchan I, Svozil D, Sponer J, Cheatham TE III, Loughton CA, Orozco M. *Biophys J.* 2007; 92:3817–3829. [PubMed: 17351000]
36. Athri P, Wilson WD. *J Am Chem Soc.* 2009; 131:7618–7625. [PubMed: 19445463]
37. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA. *J Comput Chem.* 2004; 25:1157–1174. [PubMed: 15116359]
38. Lavery R, Moakher M, Maddocks JH, Petkeviciute D, Zakrzewska K. *Nucleic Acids Res.* 2009; 37:5917–5929. [PubMed: 19625494]
39. Sklenar V, Kypr J, Bax A, Zon G, Vorlickova M. *Int J Biol Macromol.* 1989; 11:273–277. [PubMed: 2489092]
40. Suzuki E, Pattabiraman N, Zon G, James TL. *Biochemistry.* 1986; 25:6854–6865. [PubMed: 3801398]
41. Gupta G, Sarma MH, Dhingra MM, Sarma RH, Rajagopalan M, Sasisekharan V. *J Biomol Struct Dyn.* 1983; 1:395–416. [PubMed: 6400880]
42. Kerwood DJ, Zon G, James TL. *Eur J Biochem.* 1991; 197:583–595. [PubMed: 2029892]
43. Assa-Munt N, Kearns DR. *Biochemistry.* 1984; 23:791–796. [PubMed: 6712923]

Figure 1.

A) 8% native PAGE of ATATAT with netropsin in 1:1 and 2:1 molar ratio of compound / binding site. A 21 mer A-tract ligation ladder, AAAAA, was loaded as the curvature reference and two random sequences with either 20 bp or 21 bp were used as the migration markers. The arrows are for 180 bp sequences of the 20 mer marker and the free ATATAT and its netropsin complexes, and for 189 bp ladders for the 21 mer marker and AAAAA sequences; B) Sequences of target DNA ATATAT and AAAAA. C) Relative mobility, R_L , as a function of L_{act} for ATATAT in the absence (open circle) and presence of netropsin in 1:1 molar ratio (solid circle). $R_L = L_{app}/L_{act}$, was calculated for each ligation ladder and plotted versus L_{act} for each sample. L_{act} is the actual molecular weight in base pair of each ligated multimer and L_{app} is the apparent molecular weight in base pair determined according to the distance, relative to the standard, each ladder migrates in the gel. From the linear fit, $R_L - 1$ is obtained as $(-0.024 + 5.0 \times 10^{-5}) C_T^2$. Using this equation, the relative curvature, C_T value, for unbound ATATAT and the Net-ATATAT complex can be determined. Then the C_T values can be converted to bending angles with a 1.0 C_T value equivalent to an 18° bending. Free ATATAT is a straight sequence, $C_T \approx 0$, while the binding of netropsin bends ATATAT by 15° . Note that the ATATAT sequence, as used in NMR, has a 10 base pair repeat while the AAAAA sequence has a 10.5 base pair repeat from a 21 base pair ligation ladder sequence. D) The structure and atom labeling of netropsin.

Figure 2. (Top) Changes in the imino proton spectral region during titration with netropsin at 278 K. Duplex to netropsin ratios are indicated on the right of the spectra and the five imino proton signals of the palindromic decamer DNA are doubled at a netropsin to DNA ratio of 1:1, (Bottom). The imino protons of G2 and T8 provide a convenient monitor to follow the extent of binding. The DNA sequence used and the numbering scheme of the bases is shown.

Figure 3.

A) NOESY spectrum depicting the H6/H8 base - H1' sugar region of the netropsin-DNA complex. The sequential walk through the H6/H8-H1' network is shown in blue and red for the two complementary strands. The selected DNA-netropsin NOE contacts illustrate the binding of netropsin in the minor groove of the DNA. B) Netropsin-DNA complex. Netropsin fits tightly into the minor groove of the AT DNA. The DNA is depicted as a transparent Connolly surface while netropsin is rendered as a space-filling representation.

Figure 4. NMR structure bundles of A) Netropsin-DNA complex and B) free DNA. Snapshots were taken every 10 ps for the last 100 ps of the restrained MD simulation and subsequently minimized. The heavy atom RMSD are 0.53 Å and 0.68 Å for the complex and free DNA respectively.

Figure 5.

Chemical shift changes ($\Delta\delta = \text{complex} - \text{duplex}$) for DNA minor groove protons: A) H1' sugar and; B) adenine H2 and imino protons. C) Close-up view into the minor groove of the netropsin-DNA complex. H1' sugar protons highlighted (+) (A7 H1', T16 H1') show large upfield chemical shift changes whereas the A15 H2 (-) is shifted downfield upon binding of netropsin. These shifts result from the proximity of these atoms to the shielding (+) and deshielding regions (-) of netropsin's pyrrole ring. The H1' sugar protons in grey (T8 H1', A15 H1') are too distant to be affected by the aromatic pyrrole.

Figure 6.

Overall and local twist changes. Top (black) and bottom (grey) base pair of A) the netropsin-DNA Duplex and B) free DNA. The complex is overwound by 23° compared to the free DNA. (C) Twist change for each base pair step. The twist changes are localized to base step 4–5 and 6–7 while the central step, 5–6, is not affected.

Figure 7.

A) Selected parameters of the free DNA (green) and the netropsin-DNA complex (orange) derived from the rMD simulations. Average minor groove widths, X displacements and roll angles including standard deviations are given at the base pair level. The bend angles for all 5000 and 9500 snapshots (bin size 0.5°) are shown as a histogram. The average bend angles are 8° into the major groove for the DNA and 18° into the minor groove for the complex. B) Final energy-minimized structure of the free DNA (green) and the netropsin-DNA complex (orange). The binding of netropsin bends the DNA towards the minor groove as illustrated by the helix axes.

Table 1

NMR statistics for free DNA and the netropsin-DNA complex

	free DNA	netropsin-DNA complex
NMR constraints		
Total NOE	542	460
Intra-residue	310	185
Inter-residue	232	209
Sequential ($ i - j = 1$)	191	176
Nonsequential ($ i - j > 1$)	41	33
DNA–netropsin intermolecular	-	48
Netropsin intramolecular	-	18
Total Watson-Crick	6	6
Total RDCs	46	40
Structure statistics		
Restraint violation energies (kcal mol ⁻¹)	3.3/1.5/4.8	10.6/4.7/15.4
NOEs / RDCs / total No. of distance violations d_v		
$0 \text{ \AA} < d_v < 0.1 \text{ \AA}$	37	49
$0.1 \text{ \AA} < d_v < 0.2 \text{ \AA}$	4	12
$0.2 \text{ \AA} < d_v < 0.3 \text{ \AA}$	0	0
Deviations from idealized geometry		
Bond lengths (Å)	0.025	0.025
Bond angles (°)	2.3	2.4