

Methylene Blue Binding to DNA with Alternating AT Base Sequence: Minor Groove Binding is Favored over Intercalation

<http://www.jbsdonline.com>

**Remo Rohs^{§,*}
Heinz Sklenar**

Theoretical Biophysics Group
Max Delbrück Center for Molecular Medicine
Robert-Rössle-Str. 10
13092 Berlin, Germany

[§]Present Address:
Department of Structural Biology
Weizmann Institute of Science
Rehovot 76100, Israel

Abstract

The results presented in this paper on methylene blue (MB) binding to DNA with AT alternating base sequence complement the data obtained in two former modeling studies of MB binding to GC alternating DNA. In the light of the large amount of experimental data for both systems, this theoretical study is focused on a detailed energetic analysis and comparison in order to understand their different behavior. Since experimental high-resolution structures of the complexes are not available, the analysis is based on energy minimized structural models of the complexes in different binding modes. For both sequences, four different intercalation structures and two models for MB binding in the minor and major groove have been proposed. Solvent electrostatic effects were included in the energetic analysis by using electrostatic continuum theory, and the dependence of MB binding on salt concentration was investigated by solving the non-linear Poisson-Boltzmann equation. We find that the relative stability of the different complexes is similar for the two sequences, in agreement with the interpretation of spectroscopic data. Subtle differences, however, are seen in energy decompositions and can be attributed to the change from symmetric 5'-YpR-3' intercalation to minor groove binding with increasing salt concentration, which is experimentally observed for the AT sequence at lower salt concentration than for the GC sequence. According to our results, this difference is due to the significantly lower non-electrostatic energy for the minor groove complex with AT alternating DNA, whereas the slightly lower binding energy to this sequence is caused by a higher deformation energy of DNA. The energetic data are in agreement with the conclusions derived from different spectroscopic studies and can also be structurally interpreted on the basis of the modeled complexes. The simple static modeling technique and the neglect of entropy terms and of non-electrostatic solute-solvent interactions, which are assumed to be nearly constant for the compared complexes of MB with DNA, seem to be justified by the results.

Introduction

Methylene blue (MB) is known to undergo binding with nucleic acids. MB-DNA complexation has been studied both experimentally (1) and theoretically (2) because of methylene blue's medical importance as a photosensitizing dye. MB binds to nucleic acids (3) and causes photooxidative damages by generating singlet oxygen through a triplet-triplet energy transfer from the photoexcited dye to molecular oxygen (4). Singlet oxygen luminescence measurements have shown a significant dependence of the singlet oxygen quantum yield on the architecture of the dye-DNA complexes (5). The current experimental knowledge of the MB-DNA complex structures and of the binding behavior of different DNA-target sequences is derived from spectroscopic data, especially from those obtained by linear and circular dichroism (CD) studies (3, 6-10). These results clearly show a different binding behavior of MB to DNA with either AT or GC alternating base sequences (3), but they do not explain the observed differences. Because of the lack of available X-ray crystallography or NMR determined atomic-resolution structures, structural model-

*Phone: +972 8 934 2479
Fax: +972 8 934 4154
Email: mail@remo-rohs.de

ing of MB-DNA complexes and energetic analysis are considered to be useful tools for obtaining more detailed information on MB binding to DNA.

The different conceivable binding modes are intercalation of MB between adjacent base pairs, and insertion of MB either into the minor or the major groove of the DNA helix. In case of MB binding to DNA with alternating GC base sequence, the spectroscopic data clearly indicate intercalation of the planar heterocyclic MB between neighboring base pairs (1, 11-13), whereas in case of binding to DNA with alternating AT base sequence minor groove binding is assumed to be the predominant binding mode (1, 13-15).

MB binding to a DNA decamer with an alternating GC base sequence was investigated in a former modeling study (2) that resulted in six structural models for the different binding modes, where intercalation was found to be energetically favored, in accordance with experimental data. The structural models, derived for the different modes of MB binding to DNA, compose an ensemble of six MB-DNA complex structures. The GC alternating DNA-target sequence contains two intercalation sites, the 5'-CpG-3' and the 5'-GpC-3' base pair steps. The dye can intercalate in either parallel or gauche orientation into both binding pockets. Thus, four structural models for intercalation and two MB-DNA complexes representing minor and major groove binding, respectively, were obtained (2). A detailed energetic analysis and comparison of contributing energy terms has enabled us to estimate the binding energy for each of the different binding modes.

The planar heterocyclic dye is expected to stabilize its binding to DNA through favorable stacking interactions with its adjacent base pairs. However, the energetic analysis pointed out that for binding of the cationic MB to polyanionic DNA, electrostatic contributions discriminate between different binding modes. In particular, the reaction field energy describing solvent electrostatic effects plays a key role in determining the binding energy ranking. In accordance with CD data (1), the estimated binding energy predicts intercalative binding of MB to GC alternating DNA as energetically favored binding mode (2).

The fluorescence quantum yield of MB has been proven to depend not only on the target-DNA base sequence (14) but also on the environment (15). Experimental binding studies have shown a decrease of binding affinity of MB to DNA with increasing ionic strength due to the total charge reduction of the polyanionic target-DNA by the cationic ligand (15-16). We have therefore extended our binding study in a salt-free aqueous solvent (2) by the calculation of electrostatic energy contributions as a function of salt concentration (17). The resulting destabilization of MB-DNA complexes with GC alternating DNA differs slightly in its magnitude for intercalation and groove binding complexes, affecting the stability ranking of the binding modes with a preference of minor groove binding at high salt concentration (17) in accordance with experimental data (1).

Several experimental studies of MB-DNA complexes were focused on comparing the binding behavior of MB to target-DNA molecules with AT or GC alternating base sequences. Spectroscopic results, especially CD data, indicate major differences of MB binding to DNA with AT alternating base sequence in comparison to DNA with GC alternating base sequence (1, 3, 13-15). It was concluded that, in contrast with MB binding to GC alternating DNA, minor groove binding is the preferred binding mode for MB binding to AT alternating DNA, already at low salt concentration. Because of these experimental data, we considered a comparable modeling study of MB binding to AT alternating target-DNA molecules to be worthwhile. This will allow us to compare structural and energetic characteristics of MB binding to the two different RY alternating DNA molecules. For modeling and analyzing MB binding to AT alternating DNA, performed in this study, the same modeling technique and protocols have been used as in case of MB binding to GC alternating DNA (2, 17).

In this context, it was of particular interest if the simple static modeling techniques and, in particular, the approximate 'single point' correction of electrostatic energies, used for taking into account important solvent effects, can faithfully reproduce the experimentally observed sequence-specificity of ligand binding to DNA.

Materials and Methods

Molecular Models

The binding of MB to DNA was studied by using DNA decamers as target molecules, with MB binding in the central part of the decamers to reduce end effects. At first, energy minimizations have been performed in order to locate energy-minimum conformations of the free target-DNA structure with an AT alternating base sequence. Similarly to the GC decamer, two low-energy structures have been found. Both conformers show the characteristic structural features of B-form DNA, but differ by having either O1'-endo sugar puckers at thymine and C2'-endo sugar puckers at adenine covering nucleotides or all C2'-endo sugar puckering. In addition, bifurcated hydrogen bonds in combination with a large Propeller twist are shaped, as reported for AT base pairs by experimental studies (18-19). With the same arguments detailed for the GC decamer (2), the lowest-energy conformer with alternating sugar puckers was selected as the starting point for modeling the complexes that were finally analyzed. It should be stressed that neither the energetic ranking of different binding modes nor major structural features of the complexes are affected by this choice. Atomic monopoles of the target-DNA were taken from a library for standard nucleotides with charge distributions calculated by a re-parameterized Hückel-Del Re method (20).

The structural model and charge distribution of MB were taken over from the former study (2). Here, the MB structure was defined as a rigid body with two-fold symmetry and with internal degrees of freedom, describing rotations of its methylated amino-groups, and the atomic monopoles were adapted to *ab initio* results for electrostatic potentials (2). Starting configurations of MB-DNA complexes with AT alternating DNA have been generated by using helical parameters for positioning MB in the different binding modes. In addition, starting structures obtained by exchange of bases in the lowest-energy complex structures with GC alternating target-DNA were used, but with virtually the same results.

Energy Minimization

Energy minimization has been performed with the JUMNA (Junction Minimization of Nucleic Acids) algorithm (21) using the built-in Flex force field (22). Pairwise electrostatic interactions were damped by a distance dependent sigmoidal dielectric function (23), with a starting value of $\epsilon = 2$ for short distances, slope = 0.25 \AA^{-1} , and an asymptotic dielectric permittivity of $\epsilon = 78$ for distances $> 20 \text{ \AA}$. In addition, reduced effective phosphate charges of $-0.25e$ were used. For complexes with GC alternating DNA, it was shown that, with these electrostatic parameters, the obtained energy minimized structures stay close to the total energy minima after the 'single point' correction of the electrostatic energy based on the continuum solvent treatment. It was verified that the same parameters are valid for complexes with AT alternating DNA.

Continuum Treatment of Solvent Electrostatic Effects

The simple electrostatic damping model, used for energy minimizations, leads to reasonable structures, but fails in prediction of relative stability of alternative structures in aqueous solution (2, 24). We have therefore re-calculated the electrostatic energies of the obtained structures by using a continuum solvent treatment (25-27), whereby the electrostatic energy is calculated as the sum of unscreened Coulomb interactions of the solute charges and the reaction field contributions of the polarized solvent continuum. In the frame of 'single point' corrections, the re-calculat-

ed electrostatic energy is simply added to the unchanged non-electrostatic energy for estimating the relative stability of the modeled MB-DNA complexes.

It has been shown for several DNA and RNA solutes that electrostatic contributions of the aqueous solvent exert the major differential effect on the stability of alternative solute structures, whereas salt effects, non-electrostatic solute-solvent interactions, and entropic terms only slightly affect them (24, 28-29). The neglect of latter interactions allows for simplification of the Poisson-Boltzmann treatment to a boundary value problem defined by the Poisson equation, which is numerically solved by the finite difference algorithm implemented in the DelPhi program (27, 30). In this study, the computational protocol of the numerical procedure was the same as used in the previous study of MB binding to GC alternating DNA (2). With a probe sphere radius of 1.4 Å, five focusing steps with a final grid spacing of < 0.3 Å, and rotational averaging of 16 equally spaced angular molecule orientations, the accuracy of the calculated reaction field energy is in the order of 0.1 kcal/mol.

Continuum Treatment of Salt Effects

The investigation of salt effects on MB binding to AT alternating DNA was a further objective of this study, which is focused on complex stability as a function of the ionic strength. Comparison of MB binding to AT alternating DNA with results of the former study of salt effects on MB binding to GC alternating DNA (17) necessitates the same protocol and parameters for both studies. The continuum treatment of salt effects requires the solution of the non-linear Poisson-Boltzmann equation, performed numerically by the finite difference algorithm implemented in the UHBD program (31). The solvent-solute interface between the two regions of low and high dielectric permittivity is defined by a probe sphere radius of 1.4 Å and is complemented by a region that is accessible for the solvent but not for ions. This ion exclusion layer around the solute molecule is defined, as proposed by Rashin and Honig, by using an ion radius of 1.68 Å (32). The UHBD built-in dielectric boundary smoothing option has been applied as well as a three-step focusing that results in < 0.33 Å of final grid spacing to reduce the grid dependence of the results. The electrostatic free energy calculations are based on solving a volume integral (33) and on assuming that the energy-minimized structures vary only slightly by the addition of salt, as shown for double-stranded DNA and RNA structures (28-29).

Structural Analysis

Ligand binding to DNA induces structural deformations of the target-DNA. These structural changes depend on the ligand-DNA binding mode. Intercalation of MB between two adjacent base pairs requires an unstacking of the base pairs forming the binding pocket, and an unwinding of the DNA helix. MB binding within the DNA grooves causes a change in the groove geometry. The description of distinct structural deformations of the DNA-targets in comparison with free DNA conformations requires a detailed structural analysis, involving helical parameters, sugar puckering modes, backbone torsion angles, and widths and depths of the grooves. The structural analysis was performed with the CURVES algorithm that calculates helical base pair step parameters with reference to a global helical axis system (34-35). CURVES built-in definitions by Stofer and Lavery (36) served for analyzing the groove geometry.

In the former study of MB-DNA binding, we have introduced a helical parameter description of the ligand position and orientation relative to the target-DNA (2) in accordance with the CURVES algorithm (34-35). For this purpose, a ligand fixed axis system, with its origin at the sulfur atom and the z-axis perpendicular to the heterocyclic plane of MB, was defined (2). This axis system allows for the calculation of the helical parameters Xdisplacement, Ydisplacement, Rise, Inclination, Tip, and Twist of MB relative to the helical axis of a chosen flanking base pair (2).

Helical parameters, describing rigid-body translations and rotations of the bases with respect to the helical DNA axis and of the MB ligand with respect to a flanking base pair, have been used to generate starting structures of the different MB-DNA complexes for energy minimizations. The starting structure generation for the different binding modes was oriented at the six representative energy-minimum structures of the formerly presented complexes of MB bound to GC alternating decamers (2). The energy-minimized conformations of the MB-DNA complexes with MB intercalated either into the 5'-TpA-3' or the 5'-ApT-3' base step are shown in Figure 1. This figure shows that the formation of an intercalation pocket requires an approximate Rise doubling between the base pairs, flanking the intercalated dye, to allow for stacking interactions with the ligand. Additional conformational changes at the intercalation site are the local unwinding of the target-DNA helix and transitions of the two α/γ -torsion angle pairs of both DNA strands from the -gauche/+gauche to the trans/trans conformation, termed α/γ -flip. These local structural changes are essential for enabling MB to intrude into the intercalation pocket without any conformational clashes. Structural modeling of the intercalation complexes was complemented by an adiabatic mapping of MB rotation around the helical axis in both possible intercalation pockets of the DNA decamer, whereby for each rotation angle the electrostatic energy was calculated by using the continuum solvent approach. In this way, two different orientations of MB in both intercalation pockets have been identified as local energy-minimum structures.

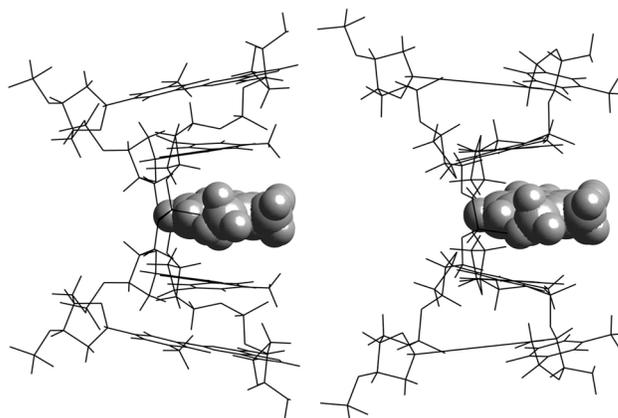


Figure 1: Structural models of the symmetric MB-DNA intercalation complexes ic1-TpA (left) and ic1-ApT (right). The ligand and four central base pairs flanking the intercalation site are shown. Note different Roll angles of adjacent base pairs at the two binding sites.

Figure 2 shows the two MB orientations for intercalation at the 5'-TpA-3' step. In the energetically preferred structure (ic1-TpA), designated as symmetric intercalation complex, the 2-fold symmetry axis of the dye coincides with the dyadic symmetry axis of the target-DNA. In the second structure (ic2-TpA), designated as gauche intercalation complex, the long axis of MB is rotated by approximate 140° around the helical axis. The comparison of the symmetric and gauche intercalation complexes shows a reduced stacking interaction when MB is intercalated in gauche orientation, thus contributing to the higher energetic stability of the symmetric intercalation complex relative to gauche intercalation. Both the two symmetric intercalation complexes (ic1-TpA and ic1-ApT) and the two gauche intercalation complexes (ic2-TpA and ic2-ApT) are structurally quite similar to the corresponding complexes obtained for MB intercalation into GC alternating DNA (2).

Structural similarity is also seen in the two groove binding complexes obtained with AT and GC alternating DNA decamers, respectively. The molecular graphics of the minor groove complex (AT-ming), depicted in Figure 3, shows that minor groove binding deepens and narrows the minor groove. For steric reasons, MB is inserted into the minor groove with its methyl groups facing outside the groove. In

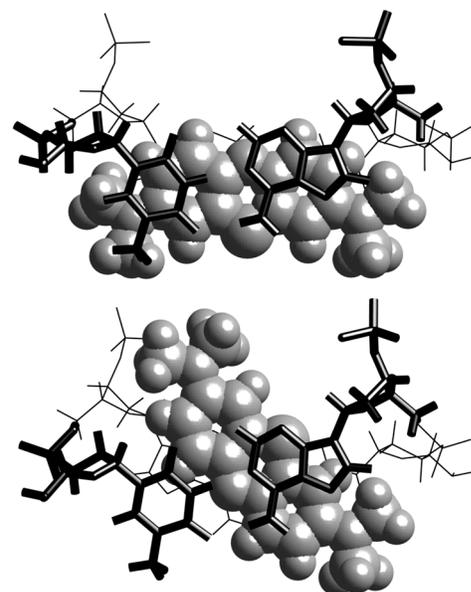


Figure 2: Stacking interactions in the MB-DNA intercalation complexes with symmetric ic1-TpA (upper panel) and gauche ic2-TpA (lower panel) intercalation, shown in a view along the helical axis.

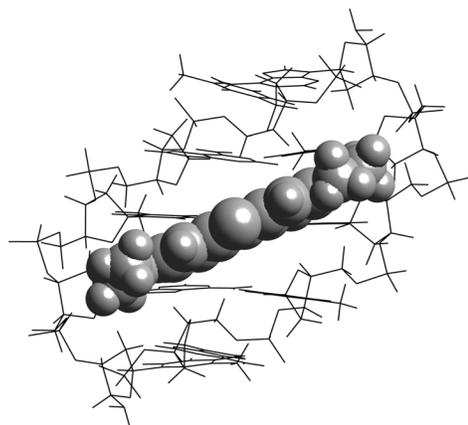


Figure 3: Structural model of the lowest-energy minor groove complex AT-ming. The ligand and five central base pairs located in the binding region are shown.

contrast with minor groove binding, MB binding in the major groove allows for much more movements of MB upon binding. The weak MB binding to AT alternating DNA in the major groove (AT-majg) makes it difficult to locate a global minimum conformation for this binding mode. Accordingly, several local minimum structures with different positions and orientations of MB within the central major groove region, but with very similar energies, have been found by global minimum search for this indisputably weakest binding mode.

Energetic Analysis

The decomposition of the total energy into energy components has enabled us to identify critical components that favor one binding mode against others. The results, given in Table I, show that the non-electrostatic (van der Waals) energy favors minor groove binding of MB to AT alternating DNA over both intercalative and major groove binding. The non-electrostatic energy of the minor groove complex (AT-ming) is 14.9 kcal/mol lower than this energy obtained for the lowest-energy intercalation complex (ic1-TpA). The electrostatic energy, damped by the distance dependent dielectric function lowers this energy difference by only 0.3 kcal/mol. In contrast with this crude approximation, the electrostatic continuum treatment of the aqueous solvent changes the energetic stability ranking of the MB-DNA complexes from 'AT-ming/ic1-TpA/ic1-ApT/AT-majg/ic2-TpA/ic2-ApT' to 'ic1-TpA/AT-ming/ic1-ApT/ic2-TpA/AT-majg/ic2-ApT', i.e., continuum solvent electrostatics increases the stability of intercalation complexes against groove binding. A similar result has been obtained for MB binding to GC alternating DNA (2). For both target-DNA sequences, the higher stability of groove binding, due to a better solvent accessibility of the ligand, is overcompensated by larger Coulomb repulsion of the solute charges. The total energy difference (0.4 kcal/mol) between the symmetric intercalation complex (ic1-TpA) and the minor groove complex (AT-ming) is, however, significantly smaller than for MB binding to GC alternating DNA (2.0 kcal/mol) (2). The changed relative stability of the two binding modes is mainly due to the lower non-electrostatic component obtained for minor groove binding in case of AT alternating DNA.

Table I

Decomposition of the total energy (in kcal/mol) for the six modeled MB-DNA complexes. Symmetric and gauche intercalation at the two different base pair steps are short termed 'ic1' and 'ic2', respectively, and the groove complexes are denoted as 'ming' and 'majg'. The energy components, quoted relative to the lowest-energy complex ic1-TpA, include the non-electrostatic energy ΔE_{Ne} of the Flex force field, the electrostatic energy ΔE_{Ed} obtained by sigmoidal electrostatic damping, the total Flex energy ΔE_{Flex}^{tot} , the Coulomb energy ΔE_{Coul} of unscreened solute charges, the reaction field energy ΔE_{Rf} , the total electrostatic energy ΔE_{cm}^{elec} of the continuum model, and the final total energy ΔE_{cm}^{tot} obtained by adding up ΔE_{Ne} and ΔE_{cm}^{elec} .

Structure	ΔE_{Ne}	ΔE_{Ed}	ΔE_{Flex}^{tot}	ΔE_{Coul}	ΔE_{Rf}	ΔE_{cm}^{elec}	ΔE_{cm}^{tot}
ic1-TpA	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ic2-TpA	3.5	3.4	6.9	30.0	-31.8	-1.8	1.7
ic1-ApT	0.2	3.8	4.0	26.5	-25.4	1.1	1.3
ic2-ApT	7.5	6.6	14.1	29.9	-32.0	-2.1	5.4
AT-ming	-14.9	0.3	-14.6	163.0	-147.7	15.3	0.4
AT-majg	0.2	5.6	5.8	169.6	-165.2	4.4	4.6

Another decomposition of the total energy is shown in Table II. Here, the total energy (denoted by E_{Bdg}) is given relative to the energy of the free structures and decomposed into energy terms describing energy changes upon binding and the respective interaction energy of the deformed structures. For each complex, the sum of these terms can be considered as an estimate of the binding energy, with the same stability ranking as discussed on the basis of the total energy values of Table I. The considerable deformation of the DNA in the intercalation complexes, particularly for gauche intercalation, is reflected by their larger deformation energy compared to groove binding. Due to the rigidity of the MB molecule with only rotatable methyl

groups, the deformation energy of MB is very small. The interaction energy of the deformed conformations (including continuum electrostatics corrections), given in Table II, shows its smallest value for major groove binding and indicates again that this is the energetically least favored binding mode. For comparison of MB binding to GC and AT alternating DNA, the binding energy of corresponding binding modes with GC alternating DNA (2) is included in Table II, indicating again that symmetric intercalation to AT alternating DNA against minor groove binding is energetically less favored than in case of binding to GC alternating DNA. The generally lower binding energy of MB binding to the AT sequence is caused by the higher deformation energy of AT alternating DNA compared with GC alternating DNA (2). The deformation energy can be decomposed into positive non-electrostatic and negative electrostatic components. Both contributions are smaller for the AT alternating target decamer than for the GC target sequence. This decomposition indicates that the higher deformation energy of AT target results from the less favorable electrostatic energy of the deformed AT structures. In structural terms, this effect can be partially attributed to the decreased phosphate-phosphate distances in the deformed AT decamers due to the higher structural flexibility of this sequence. It should be stressed that the higher deformation energy of the AT decamer is observed for all binding modes. The contribution of base pair unstacking to the deformation energy in the intercalation mode is virtually the same for both sequences and has therefore no differential effect. However, it should be emphasized that these results are based on an energetic analysis in a salt-free aqueous solvent.

Table II

Decomposition of estimated binding energy (in kcal/mol) for the six modeled MB-DNA complexes. Notation of structures corresponds to Table I. The deformation energy ΔE^{def} is the difference of the total energy calculated for the deformed structure and the respective free structure. E_{Int} is the interaction energy of the deformed structures in the complex, and E_{Bdg} the resulting estimate of the binding energy. Note that the stability ranking of the complexes is exactly the same as shown in Table I for the total energy. For comparison, respective binding energy values of MB to GC alternating DNA are included.

Structure	ΔE_{DNA}^{def}	ΔE_{MB}^{def}	E_{Int}	E_{Bdg}	Structure	E_{Bdg}
ic1-TpA	8.8	0.1	-22.1	-13.2	ic1-CpG	-14.8
ic2-TpA	10.6	0.0	-22.1	-11.5	ic2-CpG	-12.2
ic1-ApT	6.5	0.1	-18.5	-11.9	ic1-GpC	-14.7
ic2-ApT	10.8	0.1	-18.7	-7.8	ic2-GpC	-12.4
AT-ming	6.3	0.0	-19.1	-12.8	GC-ming	-12.8
AT-majg	3.5	0.2	-12.3	-8.6	GC-majg	-10.3

Salt Effects on Binding

The electrostatic energy in aqueous solvent with monovalent salt was calculated by solving the non-linear Poisson-Boltzmann equation (27) using the UHBD program (31). The results for the different MB-DNA complexes are given in Table III, for ionic strengths between 10^{-4} moles/liter and 2 moles/liter and relative to the electrostatic energy at zero salt concentration. The respective total energy, relative to the energy of free structures at zero salt, is plotted in Figure 4. As expected, adding of salt increases the stability of free DNA due to the polyanionic charges. Because of the reduced total charge by $-e$, the stabilization effect of salt is significantly smaller

Table III

Electrostatic free energy of the modeled complexes as a function of monovalent salt concentration (in moles/liter). The energy values (in kcal/mol) are given relative to the electrostatic energy of the complexes in aqueous solution without added salt. Notation of structures corresponds to Table I.

Structure	0.0	0.0001	0.001	0.01	0.1	0.5	1.0	2.0
ic1-TpA	0.0	-7.6	-20.0	-36.1	-50.9	-58.1	-60.3	-62.1
ic2-TpA	0.0	-7.8	-20.3	-36.4	-51.3	-58.5	-60.8	-62.6
ic1-ApT	0.0	-7.5	-19.8	-35.8	-50.6	-57.6	-59.8	-61.5
ic2-ApT	0.0	-7.5	-19.8	-35.8	-50.6	-57.7	-59.8	-61.6
AT-ming	0.0	-8.3	-21.0	-37.3	-52.3	-59.8	-62.1	-64.0
AT-majg	0.0	-8.2	-20.9	-37.2	-52.2	-59.7	-62.0	-63.9
free (DNA+MB)	0.0	-10.8	-25.4	-43.3	-59.8	-68.0	-70.6	-72.8

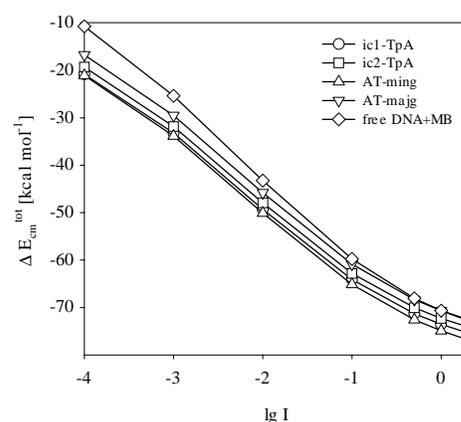


Figure 4: Total energy of the symmetric (ic1-TpA) and gauche (ic2-TpA) intercalation, and minor (AT-ming) and major (AT-majg) groove complexes as a function of the decadic logarithm of ionic strength, in comparison the total energy of the unbound molecules. The energy values are relative to the energy at zero salt concentration.

for the complexes with the cationic MB ligand and, therefore, causes a decrease of the binding energy with increasing salt concentration. This result is in accordance with experimental data (15-16). The same behavior was found for binding to GC alternating DNA, i.e., it is virtually independent of DNA base sequences.

The binding energy of the MB complexes with AT alternating DNA in dependence of salt concentration is quoted in Table IV. In addition, for the sake of comparison with GC alternating DNA (17), the results for the four most stable complexes with both sequences are plotted as functions of salt concentration in Figure 5. Almost independently of both the binding mode and salt concentration, the binding energy obtained for complexes with AT alternating DNA is generally lower than with GC alternating DNA by roughly two kcal/mol, in qualitative agreement with experimental findings (13). In both cases, the binding energy of intercalation complexes as a function of salt concentration decreases with a larger slope than the binding energy of groove complexes. A possible explanation for this difference is the prolongation of the DNA in the intercalation complexes resulting in a lower charge density in the center of the structure and subsequently in a slightly lower stabilization by increasing salt concentration in comparison with the groove binding complexes. We should note that the observed curvature in the salt dependence might be an artifact due to the limits of volume integration that is limited to the volume of the final focusing grid in the PB calculations (the volume outside this box has an effect that depends on the salt concentration). However, for the comparison of different binding modes this does not affect the conclusions of the paper.

Figure 5: Estimated binding energy of MB binding to AT alternating (left panel) and GC alternating (right panel) DNA decamers as a function of the decadic logarithm of ionic strength. Plots are shown for symmetric and gauche intercalation at 5'-YpR-3' sites, and for minor and major groove binding.

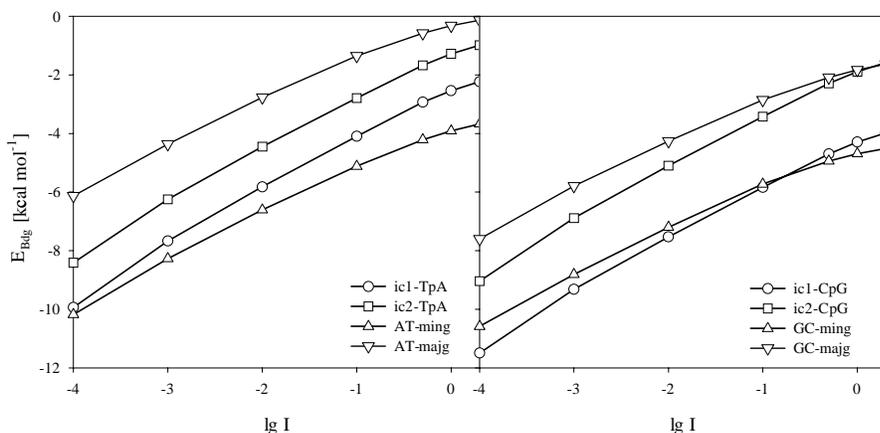


Table IV

Estimated binding energy (in kcal/mol) of the modeled complexes as a function of salt concentration (in moles/liter). Notation of structures corresponds to Table I.

Structure	0.0	0.0001	0.001	0.01	0.1	0.5	1.0	2.0
ic1-TpA	-13.2	-9.9	-7.7	-5.8	-4.1	-2.9	-2.5	-2.2
ic2-TpA	-11.5	-8.4	-6.2	-4.4	-2.8	-1.7	-1.3	-1.0
ic1-ApT	-11.9	-8.8	-6.5	-4.7	-3.0	-1.8	-1.4	-1.0
ic2-ApT	-7.8	-4.7	-2.4	-0.6	1.1	2.2	2.7	3.0
AT-ming	-12.8	-10.2	-8.3	-6.6	-5.1	-4.2	-3.9	-3.7
AT-majg	-8.6	-6.2	-4.4	-2.8	-1.4	-0.6	-0.3	-0.1

In view of the significantly different binding behavior of the two sequences, our data show that the energetic preference of minor groove binding to AT alternating DNA (AT-ming) amounts already to 0.3 kcal/mol at very low salt concentration and reaches 1.5 kcal/mol at 2 moles/liter, whereas a small preference for minor groove binding to GC alternating DNA (GC-ming) is only seen above 0.1 moles/liter (0.5 kcal/mol at 2 moles/liter) (17). The binding energy differences for zero salt and a very small salt concentration of 10^{-4} moles/liter are due to the fact that in the continuum solvent treatment without salt, based on the Poisson equation, the counterions are completely neglected, whereas in the treatment with salt, based on the non-linear Poisson-Boltzmann equation, the counterions are always

present and compensate the phosphate charges of the electrically neutral systems. Using the Poisson-Boltzmann equation is therefore more realistic.

Salt effects on binding energetics are very similar for both sequences. Significant differences, with respect to the binding energy and binding mode preferences, are mainly due to non-electrostatic interactions, as discussed on the basis of salt-free solvent treatment, and are therefore not changed by salt effects. For AT alternating DNA, the energetic preference of minor groove binding is stable at the whole range of tested salt concentrations, as opposed to a change from symmetric intercalation to minor groove binding at salt concentrations > 0.1 moles/liter, in case of GC alternating DNA.

Structural Analysis

Although minor groove binding is the preferred binding mode of MB with AT alternating DNA, it was of interest to analyze the intercalation complexes structurally in some detail, since symmetric intercalation is the competing binding mode, which is favored in MB binding to GC alternating DNA. Helical parameters were calculated by using the CURVES algorithm (34-35). In Figure 6, the helical base pair step parameters Rise, Twist, and Roll and the intra-base pair parameter Propeller are shown for both intercalation modes and both intercalation sites in comparison to the free target-DNA. The Rise values between the base pairs forming the intercalation pocket are 6.7 \AA for the 5'-TpA-3' and 7.2 \AA for the 5'-ApT-3' base pair step, but are virtually the same for symmetric and gauche intercalation. A similar behavior has been observed for MB binding to GC alternating DNA (2). In contrast to the Rise, symmetric and gauche intercalation into RY alternating DNA

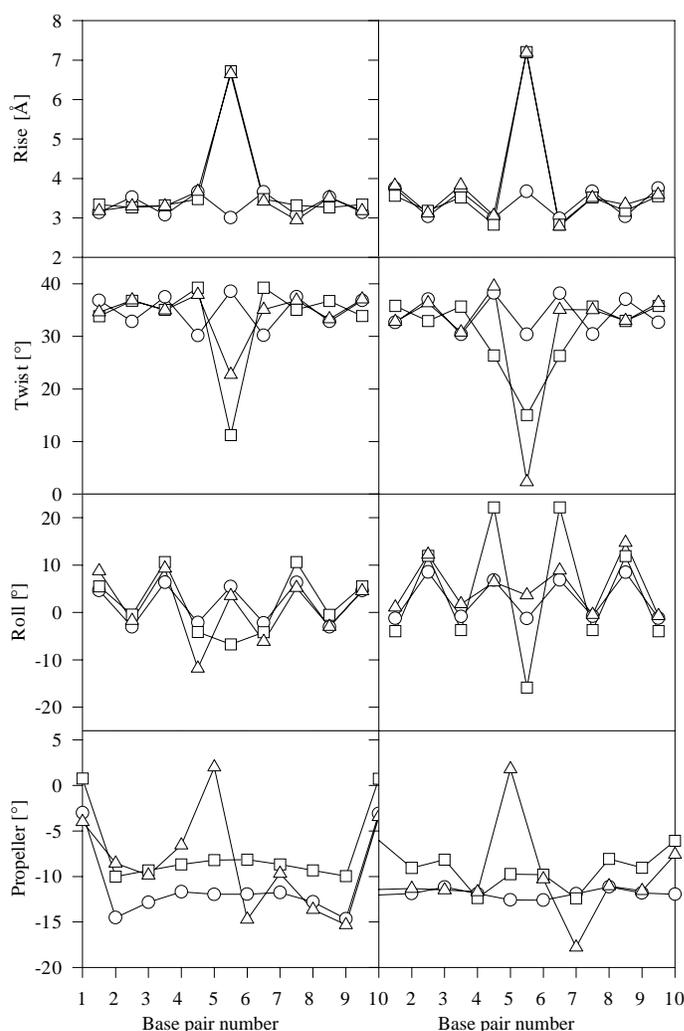


Figure 6: Helical base pair step parameters (Rise, Twist, and Roll), and the intra base pair Propeller twist of DNA in symmetric (squares) and gauche (triangles) intercalation complexes for MB intercalation at 5'-TpA-3' (left panel) and 5'-ApT-3' (right panel) sites. The intercalation sites are located between the central base pairs 5 and 6. For comparison, the parameters of the energy-minimized free DNA decamer (circles) are included.

show a reverse unwinding behavior for the two binding pockets, again in accordance with MB binding to GC alternating DNA (2). The DNA in the ic1-TpA complex is stronger unwound than in the respective ic2-TpA complex (helical Twist of 11.2° compared to 22.7°), whereas the ic1-ApT complex is less unwound than the ic2-ApT complex (helical Twist of 15.0° compared to 2.3°). As already seen in the molecular graphics of the symmetric intercalation complexes ic1-TpA and ic1-ApT in Figure 1, the Roll angle describes most significantly conformational differences of the DNA induced by intercalation into the 5'-TpA-3' and the 5'-ApT-3' pocket. Figure 6 shows moderate Roll variations from 10.7° to -6.7° in the ic1-TpA complex, in contrast to larger variations between 22.2° and -15.9° in the ic1-ApT complex. Interestingly, the Roll changes its sign in the ic1-ApT complex from the central base pair step to both neighboring steps and in the ic1-TpA complex from the central base pair step gradually via the neighboring steps to the next neighboring steps. Thus, the Roll angle is the most specific structural parameter for describing conformational changes of DNA induced by intercalations into 5'-TpA-3' and 5'-ApT-3' base steps. The helical base pair step parameters Rise, Twist, and Roll are much more sensitive to intercalation into either 5'-YpR-3' or 5'-RpY-3' base pair steps than to changing from AT to GC alternating DNA. However, due to the weaker hydrogen bonding of A-T compared to G-C base pairs the intra base pair parameter Propeller twist behaves quite differently for symmetric intercalation into the 5'-ApT-3' and the 5'-GpC-3' base pair step. The more negative Propeller twist in the symmetric intercalation complex with AT alternating DNA is almost constant, in contrast to large changes at the 5'-GpC-3' intercalation site. In accordance with the explanation by weaker hydrogen bonding of A-T base pairs, the Propeller twist changes more efficiently between neighboring base pairs in the gauche intercalation complexes with AT alternating DNA, due to the less favorable stacking pattern of the ligand with its adjacent base pairs.

The conformational differences, described for 5'-YpR-3' and 5'-RpY-3' intercalation sites, are correlated with characteristic sugar puckering modes at the flanking nucleotides, namely 5'-Y(C3'-endo)-p-R(C2'-endo)-3' and 5'-R(C3'-endo)-p-Y(C3'-endo)-3', respectively. This intercalation site-specific sugar puckering is observed for both DNA-target sequences as well as for both intercalation modes. The correlation between intercalation site-specific helical parameters and sugar puckering modes of the ligand flanking nucleotides is in agreement with X-ray crystallography results, which are, however, restricted to the intercalation of an aromatic ligand between only two flanking nucleotide pairs (37). This conformational correlation seems to be a general feature of DNA intercalation complexes.

The position and orientation of the ligand are described by helical parameters, given in Table V, with respect to the helical axis of the 5'-adjacent base pair. The positive Xdisplacement indicates a shift of the sulfur atom, where the ligand fixed coordinate system has its origin, into the major groove, and the Ydisplacement of 0.0 \AA means perfect dyadic symmetry of the intercalation complex. The widths and depths of the DNA grooves in the different MB-DNA complexes are quoted in Table VI in comparison to the free target-DNA. The minor groove deformations of AT intercalation complexes are similar to MB intercalation into GC alternating DNA (2), although the minor grooves are shaped less differently at the 5'-YpR-3' or the 5'-RpY-3' sites, if MB is bound to GC alternating DNA.

In the minor groove complex, the position of MB relative to the target-DNA results in a positive Xdisplacement, because the ligand fixed x -axis (2) points towards the outside of the minor groove, whereas the x -axis points into the major groove in the symmetric intercalation complex. The opposite position of the ligand in the minor groove complex (with its methyl groups facing outside the groove) is also reflected by its helical Twist of 174.1° . The Inclination of 27.3° points out that the orientation of MB's long axis is nearly parallel to both nucleic acid backbones. The reduced Xdisplacement of 8.1 \AA (compared to 9.6 \AA in the minor groove complex

Table V

Helical parameters describing the position and orientation of MB in the different MB-DNA complexes. Translation parameters are given in Å and rotations in degrees. Notation of structures corresponds to Table I. The parameters are defined as ligand-axis parameters with reference to the helical axes system of the 5'-flanking base pair (2), in analogy to base pair-axis parameters used in CURVES (34-35).

Structure	<i>Xdisp</i>	<i>Ydisp</i>	<i>Rise</i>	<i>Inc</i>	<i>Tip</i>	<i>Twist</i>
ic1-TpA	2.1	0.0	3.3	-4.0	0.5	4.9
ic2-TpA	2.5	0.8	2.9	1.4	12.8	141.4
ic1-ApT	1.0	0.0	3.3	2.1	-0.1	6.4
ic2-ApT	2.9	0.6	3.5	2.9	-1.1	144.7
AT-ming	8.1	0.8	0.3	27.3	-5.1	174.1
AT-majg	5.1	1.5	2.6	0.5	-38.4	7.0

with GC alternating DNA) reflects a deeper insertion of MB into the minor groove of AT alternating DNA. As expected and observed in minor groove complexes with other ligands (38-39), minor groove binding of MB induces a minor groove narrowing and deepening. The deeper insertion of the dye into the minor groove of AT rich sequences is favored by the absence of bulky guanine amino groups (38). In comparison with GC rich sequences, this could explain more favorable non-electrostatic interactions, resulting in the smaller binding energy difference for minor groove binding and symmetric intercalation as shown in Table II.

The helical parameters of MB and groove geometry parameters for major groove binding are given for completeness. As shown by the energetic analysis, MB major groove binding is energetically unstable, leading us to abstain from a more detailed discussion of this rather unrealistic binding mode.

Discussion

Based on a detailed energetic analysis of modeled MB complexes with AT alternating DNA, we have found that MB favors minor groove binding to AT alternating DNA, whereas the symmetric intercalation complex is most stable in binding to GC alternating DNA (2). Although the energy differences are small and several energy terms have been approximated or neglected, the change from intercalation to minor groove binding seems to be significant. Neglected entropy terms are important for calculations of free energy, but are assumed to cancel out in calculations of relative energy. By using this approximation, the energetic analysis indicates that the higher preference for minor groove binding is mainly due to more favorable van der Waals interactions of MB in the minor groove. This major result of our study is in accordance with the interpretation of experimental data for complexes of AT rich DNA with other minor groove binding ligands (38).

We will now compare our results with some conclusions derived from spectroscopic studies of similar systems. Absorption spectra of MB bound to DNA with alternating AT base sequences refer to two coexisting binding modes, whereas MB is shown, in absorption and LD studies, to bind always via intercalation to GC alternating DNA (1, 13). The two binding modes, observed for MB binding to AT alternating DNA at low MB-DNA binding ratios, were attributed to intercalative and non-intercalative binding (13). The fluorescence quantum yield of MB in the environment of AT rich sequences reaches the same magnitude as for free MB molecules. In contrast, the fluorescence quantum yield of MB in the environment of GC rich sequences is significantly quenched (14). Furthermore, the lifetime of MB's excited singlet state was proven by a fs-transient absorption study to be strongly affected by GC, but are only slightly affected by AT containing DNA (40). These experimental data provide strong indications for a groove binding mode of MB to DNA with AT sequences, whereas intercalation is the predominant binding mode for GC alternating DNA.

Table VI

Groove widths and depths (in Å) of the target-DNA deformed by interaction with MB in the different binding modes. Notation of structures corresponds to Table I. The widths and depths of the minor and major groove were averaged over the central base pair steps of the decamer. The groove definitions of Stofer and Lavery (36), as built-in in the CURVES program, have been used.

Structure	w_{\min}	d_{\min}	w_{\max}	d_{\max}
free helix	6.3	4.7	12.5	6.5
ic1-TpA	7.4	5.0	19.9	5.3
ic2-TpA	7.8	4.7	16.0	2.3
ic1-ApT	11.7	0.1	18.2	9.1
ic2-ApT	11.3	3.9	18.0	5.0
AT-ming	5.0	5.5	12.8	5.0
AT-majg	7.3	4.3	10.2	7.4

The CD spectra, reported by Tuite and Nordén (1), do not only substantially differ in sign and magnitude for MB binding to either AT or GC alternating DNA, but also allow for the interpretation that a non-intercalative binding mode is stabilized over intercalation with increasing salt concentration (1). The change to non-intercalative binding, when the salt concentration increases, has been proven by laser excitation luminescence measurements (15), in agreement with our calculation of binding energy as a function of salt concentration. The higher sensitivity of MB binding with GC sequences to salt concentration, concluded from spectroscopic data (15), corresponds with the binding energy plots of Figure 5. The cross intersection of the two binding energy plots for symmetric intercalation and minor groove binding, in case of GC alternating DNA, allows for the interpretation that both binding modes coexist at salt concentrations in the range of 0.1- 0.5 moles/liter.

The CD data do not distinguish between minor and major groove binding. Based on LD measurements, however, Tuite and Nordén have calculated an angle of about 57° between MB's long axis and the helical axis of the DNA (1). The 90° complement of this angle roughly agrees with MB's Inclination of 27.3° , given in Table V for minor groove binding. The prediction of minor groove binding, favored over intercalation at the tested salt concentrations, is therefore in accordance with available CD and LD data (1).

The reported decrease of binding energy with increasing ionic strength, independent of the binding mode (6, 16), is also seen in the binding energy plots of Figure 5. This behavior was explained by the reduced total charges of the complexes compared with free DNA. Tuite and Kelly have found higher binding constants for MB binding to GC alternating sequences than for AT alternating DNA-targets (13), in qualitative agreement with our data given in Table II, where estimated binding energy values of both systems are compared. According to our results, a lower binding constant with AT sequences is caused by the larger deformation energy.

Conclusions

Understanding of ligand-DNA interactions in structural and energetic terms is complicated, because large free energy contributions, particularly due to the polyanionic character of DNA and the important role of counterions, add up to small differences that decide on the favorable structure of the complex. In this and the former study on MB binding to DNA, six energy-minimized structural models for the different binding modes have been used for a differential energetic analysis. The total energy includes van der Waals (Lennard-Jones) and angular energy terms of the Flex force field and the electrostatic energy calculated by the Poisson-Boltzmann approach. Entropy terms and solute-solvent interactions have been neglected by assuming that they are very similar for the different structures. A justification for this approximation is given by the results, which are in good agreement with available experimental data for the investigated systems. We can therefore conclude that our simple static modeling technique has provided reasonable structural models for MB binding to GC and AT alternating sequences of DNA. In case of the GC sequence (2), the proposed structural models have served already for discussing experimental results (41-42). In addition, the energetic analysis has given some insight into the different binding behavior of the two sequences.

It should, however, be emphasized that flexible DNA and its ligand complexes are dynamic systems, whereby also a sequence-specific mobility of the ligand in the different binding modes can be expected. Such data are even more difficult to obtain by experimental techniques than static high-resolution structures. The progress in computer simulations of the dynamics of large biomolecular systems gives promising hope that such simulation can also be successfully applied to ligand-DNA binding. The results of a first study of MB binding to DNA, based on an efficient Monte Carlo technique (43), will be presented in a forthcoming paper.

The authors would like to thank first and foremost Prof. Zippora Shakked from the Weizmann Institute of Science for all her helpful advice and valuable discussions and Adi Rolider for proofreading the manuscript. R. Rohs acknowledges funding by the Helmholtz Network for Bioinformatics and the Minerva Foundation (Max Planck Society).

References and Footnotes

1. E. M. Tuite and B. Nordén. *J. Am. Chem. Soc.* 116, 7548-7556 (1994).
2. R. Rohs, H. Sklenar, R. Lavery, and B. Röder. *J. Am. Chem. Soc.* 122, 2860-2866 (2000).
3. E. M. Tuite and J. M. Kelly. *J. Photochem. Photobiol. B* 21, 103-124 (1993).
4. G. Laustriat. *Biochimie* 68, 771-778 (1986).
5. R. Rohs. *Spektroskopische Untersuchungen und theoretische Modellierungen physikalischer Eigenschaften von Farbstoff-Oligonukleotid-Konjugaten*. Diploma Thesis, Humboldt University Berlin (1997).
6. B. Nordén and F. Tjerneld. *Biopolymers* 21, 1713-1734 (1982).
7. J. Wang, M. Hogan, and R. H. Austin. *Proc. Natl. Acad. Sci. USA* 79, 5896-5900 (1982).
8. M. Hogan, J. Wang, R. H. Austin, C. L. Monitto, and S. Hershkowitz. *Proc. Natl. Acad. Sci. USA* 79, 3518-3522 (1982).
9. R. Lyng, T. Hård, and B. Nordén. *Biopolymers* 26, 1327-1345 (1987).
10. D. F. Bradley, N. C. Stellwagen, C. T. O'Konski, and C. M. Paulson. *Biopolymers* 11, 645-652 (1972).
11. W. Müller and D. M. Crothers. *Eur. J. Biochem.* 54, 267-277 (1975).
12. C. OhUigin, D. J. McConnell, J. M. Kelly, and W. J. M. van der Putten. *Nucleic Acids Research* 15, 7411-7427 (1987).
13. E. M. Tuite and J. M. Kelly. *Biopolymers* 35, 419-433 (1995).
14. J. M. Kelly, W. J. M. van der Putten, and D. J. McConnell. *Photochem. Photobiol.* 45, 167-175 (1987).
15. W. J. M. van der Putten and J. M. Kelly. *Photochem. Photobiol.* 49, 145-151 (1989).
16. P. Hagmar, S. Pierrou, P. Nielsen, B. Nordén, and M. Kubista. *J. Biomol. Struct. Dyn.* 9, 667-679 (1992).
17. R. Rohs and H. Sklenar. *Indian J. Biochem. Biophys.* 38, 1-6 (2001).
18. R. E. Dickerson and H. R. Drew. *J. Mol. Biol.* 149, 761-786 (1981).
19. H. C. M. Nelson, J. T. Finch, B. F. Luisi, and A. Klug. *Nature* 330, 221-226 (1987).
20. R. Lavery, K. Zakrzewska, and A. Pullman. *J. Comput. Chem.* 5, 363-373 (1984).
21. R. Lavery, K. Zakrzewska, and H. Sklenar. *Comput. Phys. Commun.* 91, 135-158 (1995).
22. R. Lavery, H. Sklenar, K. Zakrzewska, and B. Pullman. *J. Biomol. Struct. Dyn.* 3, 989-1014 (1986).
23. B. E. Hingerty, R. H. Ritchie, T. L. Ferrell, and J. E. Turner. *Biopolymers* 24, 427-439 (1985).
24. K. Zakrzewska, A. Madami, and R. Lavery. *Chem. Phys.* 204, 263-269 (1996).
25. S. C. Harvey. *Proteins: Structure, Function, and Genetics* 5, 78-92 (1989).
26. M. E. Davis and J. A. McCammon. *Chem. Rev.* 90, 509-521 (1990).
27. B. Honig and A. Nicholls. *Science* 268, 1144-1149 (1995).
28. M. Zacharias and H. Sklenar. *Biophys. J.* 73, 2990-3003 (1997).
29. M. Zacharias and H. Sklenar. *J. Mol. Biol.* 289, 261-275 (1999).
30. M. K. Gilson, K. A. Sharp, and B. Honig. *J. Comput. Chem.* 9, 327-335 (1987).
31. M. E. Davis, J. D. Madura, B. A. Luty, and J. A. McCammon. *Comput. Phys. Commun.* 62, 187-198 (1991).
32. A. A. Rashin and B. Honig. *J. Phys. Chem.* 89, 5588-5593 (1985).
33. K. A. Sharp and B. Honig. *J. Phys. Chem.* 94, 7684-7692 (1990).
34. R. Lavery and H. Sklenar. *J. Biomol. Struct. Dyn.* 6, 63-91 (1988).
35. R. Lavery and H. Sklenar. *J. Biomol. Struct. Dyn.* 6, 655-667 (1989).
36. E. Stofer and R. Lavery. *Biopolymers* 34, 337-346 (1994).
37. H. M. Sobell. *Biological Macromolecules and Assemblies 2, Nucleic Acids & Interactive Proteins*, pp. 171-232, Eds., F. A. Jurnak and A. MacPherson. Wiley, New York (1985).
38. C. A. Hawkins, C. Watson, Y. Yan, B. Gong, and D. E. Wemmer. *Nucleic Acids Res.* 29, 936-942 (2001).
39. M. Bennett, A. Krah, F. Wien, E. Garman, R. Mckenna, M. Sanderson, and S. Neidle. *Proc. Natl. Acad. Sci. USA* 97, 9476-9481 (2000).
40. E. Tuite, J. M. Kelly, G. S. Beddard, and G. S. Reid. *Chem. Phys. Lett.* 226, 517-524 (1994).
41. G. D. Reid, D. J. Whittaker, M. A. Day, C. M. Creeley, E. M. Tuite, J. M. Kelly, and G. S. Beddard. *J. Am. Chem. Soc.* 123, 6953-6954 (2001).
42. G. D. Reid, D. J. Whittaker, M. A. Day, D. A. Turton, V. Kayser, J. M. Kelly, and G. S. Beddard. *J. Am. Chem. Soc.* 124, 5518-5527 (2002).
43. R. Rohs. *Simulation der Strukturbildung und Ligandenbindung von Nukleinsäuren im Raum kollektiver und innerer Variablen*. Ph. D. Thesis, Free University Berlin (2002).

Date Received: September 10, 2003

Communicated by the Editor
Zippora Shakked