LEGIBILITY NOTICE

A major purpose of the Technical Information Center is to provide the broadest dissemination possible of information contained in DOE's Research and Development Reports to business, industry, the academic community, and federal, state and local governments.

Although a small portion of this report is not reproducible, it is being made available to expedite the availability of information on the research discussed herein.

LA-UR--90-1604

DE90 011977

TITLE STRUCTURE AND DYNAMICS OF A [1:1] DRUG-DNA COMPLEX:
ANALYSIS OF 2D NMR DATA USING MOLECULAR MECHANICS AND
MOLECULAR DYNAMICS CALCULATIONS

AUTHOR(S) Ramaswamy H. Sarma
Mukti H. Sarma
Kimiko Unmemoto
Goutam Gupta
Angel E. García

SUBMITTED TO Proceedings of the 23rd Jerusalem Symposium on Molecular
Basis of Specificity in Nucleic Acid-Drug Interactions,
B. Pullman, ed. Kluwer Academic Publishers, Dordrech,
The Netherlands

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

By acceptance of this article the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution or to allow others to do so for U.S. Government purposes.

The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

**Structure and Dynamics of a [1:1] Drug-DNA
Complex: Analysis of 2D NMR Data Using
Molecular Mechanics and Molecular Dynamics
Calculations.**

**Ramaswamy H. Sarma¹, Mukti H. Sarma¹, Kimiko
Umemoto¹, Goutam Gupta² and Angel E. Garcia²**

**¹ Institute of Biomolecular Stereodynamics,
Department of Chemistry,
State University of New York at Albany,
Albany, NY 12222.**

**² Theoretical Biology and Biophysics Group,
T-10, MS K710,
Los Alamos National Laboratory,
Los Alamos, NM 87545.**

Abstract

1D/2D NMR studies are reported for a [1:1] complex of d(GA₄T₄C)₂ and Dst2 (an analogue of distamycin A). Full-Matrix NOESY Simulations, Molecular Mechanics and Molecular Dynamics Calculations are performed to analyze the NMR data. Results show that drug-DNA complex formation is driven by static features like H-bonding and steric interactions in the minor-groove of DNA. As a consequence of drug binding, a non-linear oscillatory mode is activated. In this mode the molecule samples equilibrium structural states of different degrees of bending. It is noted that these structures belong to three distinctly different energy wells that satisfy the same NMR data.

Introduction

Based upon physico-chemical studies [1-8], it is often concluded that static features like H-bonding and steric interactions in the minor-groove of DNA account for the sequence specificity of the minor-groove binding drugs. However, it is quite possible that some elements of specificity can be dynamic in nature. In other words, in presence of the drug, the target DNA molecule may exhibit characteristic motions at the site of binding and adjacent to it. This prompted us to analyze the dynamics of *Dst2*-d(GA₄T₄C)₂ [1:1] complex by combining 2D NMR and theoretical methods (*Dst2* is an analogue of distamycin A- see Fig. 1).

Methodology

NMR. The details of NMR methods were previously reported [9]. All NMR experiments on *Dst2*-d(GA₄T₄C)₂ complex in which drug:DNA ratio was kept at 1:1 molar equivalent. Assignment of the exchangeable and non-exchangeable DNA and drug protons in the complex were derived by combining 1D/2D NMR data of the free DNA, the free drug and the drug-DNA complex in H₂O/D₂O [9]. The NOESY data of the *Dst2*-d(GA₄T₄C)₂ [1:1] complex for mixing times τ_m = 50, 100 and 150 msec, were used to carry out the Full Matrix NOESY Simulations and associated R-factor test on several structures generated using constrained least-square method [10]. R-factor is defined as

$$\text{R-Factor} = (1/N) \frac{\sum_{ij}^N |a_{ij}^o - a_{ij}^c|}{\sum_{ij}^N a_{ij}^o} \quad (1)$$

where

N = total number of measurable NOESY (i,j) peaks,

a_{ij}^o = observed NOESY intensity for the (i,j) peak

& a_{ij}^c = calculated NOESY intensity for the (i,j) peak.

For overlapping NOESY cross-peaks, it is hard to estimate the experimental value of the intensity due to a single pair-wise interaction. Therefore, for overlapping NOESY peaks, theoretical NOESY slices were constructed taking into consideration of the contributions of all (i,j) pairs in the region of overlap. Theoretical NOESY slices, so obtained, were then compared with the corresponding observed ones. The details of the methodology employed to treat the overlapping peaks were previously discussed [9,11].

Constrained Energy Minimization. NOESY data resulted in a set of about 30 conformationally similar DNA models that showed similar agreement with the NOESY data. From these DNA models about 60 independent average inter-proton distances were extracted as structural constraints required for agreement with the NMR data of the Dst2-d(GA4T4C)₂ [1:1] complex. A pseudo potential $U = K(d_{\text{const}} - d_{\text{actual}})^2$ was added in the potential function with $K = 10 \text{ Kcal/Mole/\AA}^2$ for distances $< 3\text{\AA}$ and $K = 5 \text{ Kcal/Mole/\AA}^2$ for distances $> 3\text{\AA}$. The Dst2-d(GA4T4C)₂ [1:1] complex was energy minimized using the AMBER force field [12]. All atoms (i.e., heavy atoms and hydrogens) of Dst2 and d(GA4T4C)₂ were included in all calculations. Dst2 was docked at the center of the minor-groove of the target DNA duplex consistent with the NOESY data. The B-DNA model of Arnott and Hukins [13] was used as the starting DNA structure for energy minimization using NOESY constraints.

Constrained Molecular Dynamic Simulation. The NOESY constrained energy minimized structure of the Dst2-d(GA4T4C)₂ [1:1] complex was used as the starting structure for the MD simulation. An 8-psec MD run was used for equilibration. A 52-psec (after equilibration) constrained MD simulation at constant energy (with T close to 300 K) was performed on the complex using the 60 NMR constraints with the associated pseudo-potential as mentioned above. The last 30-psec MD data were used to analyze the variations in the NMR constraints and other structural parameters. 104 structures at regular intervals of

0.5 psec were taken out of the 52-psec MD trajectory. Root-Mean-Square(RMS) deviations among all these structures were computed and presented in the form of a (104 X 104) matrix. Global translations and rotations were appropriately deducted from the RMS matrix such that the matrix reflected true conformational differences among these structures. Each of the 104 structures from the MD trajectory was fully energy minimized (to an average gradient of 0.01 Kcal/Mole/Å) with the NMR constraints. Structural variations among the 104 energy minimized structures were analyzed again by constructing another (104 X 104) RMS matrix.

Results

Dst2 is located at the central 5 basepairs of the Dst2-d(GA4T4C)2 [1:1] complex. The numbering scheme of the DNA duplex and Dst2 is shown in Figure 1. The presence of H-bonded amide protons of Dst2 were located within 9.3-8.6 ppm [9]. The length of Dst2 suggests that the drug can span about 5 consecutive A.T pairs in the minor-groove of a B-DNA duplex. From Fig. 1, it appears that Dst2 can bind to d(GA4T4C)2 in 5 different ways i.e., it can span

A2 - A3 - A4 -A5 -T6
T19-T18-T17-T16-A15 ,

A3 - A4 - A5 - T6 -T7
T18- T17-T16-A15-A14 or their symmetry
equivalents etc.

The following data [9] supported the notion that Dst2 is predominantly bound at the central 5 base-pairs i.e., the segment

A4 - A5 - T6 - T7 - T8
T17-T16-A15-A14-A13

or its symmetry equivalent

A14-A15-T16-T17-T18
T7 - T6 - A5 - A4 - A3

(See Fig. 1).

- (i) Strong NOEs (at $\tau_m = 50, 100$ and 150 msec) are present between HC3/HC11 of Dst2 and H2(A5/A15) of DNA.
- (ii) Weaker NOEs are observed between HC3/HC11 of Dst2 and H1' of T6/T16, T7/T17; these peaks begin to take shape at $\tau_m = 100$ msec or higher.
- (iii) Upon Dst2 binding H2 of A4/A14, A5/A15 are down field shifted and H1' of A5/A15, T6/T16, T7/T17 are high field shifted. Other protons show very little change upon drug binding.
- (iv) There is no resonance doubling of DNA and Dst2 protons i.e., chemical shift values of H8 of A4 and A14 are equivalent and so are H6 of T6 and T16 and similarly for all other protons. This observation eliminates the possibility of any other binding mode present as a major population.

However, NMR data cannot rule out the possibility of minor conformers showing other than the central binding mode [Fig. 1].

A Static Structure Obtained from the Constrained Energy Minimization. Full-Matrix NOESY Simulations were performed on the DNA structures generated by a constrained least-square method and calculated NOESY intensities for proton pairs (i,j) were compared with the corresponding observed ones at $\tau_m = 50, 100$ and 150 msec. In this simulation, an isotropic correlation time τ_c of 3 nsec was used and Dst2 protons were not included in the relaxation matrix (inclusion of Dst2 protons did not significantly alter the NOESY pattern involving the DNA protons). Details of the methodology were previously reported [9,11].

NOESY Simulations resulted in a set of about 30 conformationally similar models that give similar agreement (i.e., a low R-Factor) with the observed NOESY data. From these 30 models, an estimate of average inter-proton distances can be obtained for various pair-wise interactions in the DNA; e.g., for T7 we obtain estimates of distances like

for the inter-nucleotide $H6(T_i) \cdots H6(T_{i-1})$ interactions. However, at the ApT sequence intra- and inter-nucleotide NOEs show similar NOE values. Comparison of Fig. 3A and Fig. 3B shows that the distribution plots for intra- and inter-nucleotide distances involving H6 at the ApT junction are similar in nature which is in agreement with the NOESY data [9]. Table I lists average inter-proton distances (and standard deviations) for the pair-wise interactions described in Fig. 3.

Three Energetically Stable But Distinctly Different Structures Isolated from the MD Trajectory. In summary, it is possible to construct an ensemble of structures from the MD trajectory that agree with the NOESY data. However, the question is : how different are these structures from each other? We address this question in this section.

104 Structures were chosen as representative snap-shots from the 52-psec MD trajectory; two neighboring snap-shots were separated by a time step of 0.5 psec. One RMS (104 X 104) matrix was constructed; each element (i,j) in the matrix gives the RMS deviations between two structures i and j in the MD trajectory. Global translation and rotations were appropriately deducted from each element. Therefore, the RMS matrix reflects true conformational differences among the 104 structures. It is noted that values of the elements along a row or a column of the RMS matrix show a periodic variation i.e., at first the values gradually increase and then reaches a maximum and finally drops to a minimum. The period of flipping from one structure to another is approximately 5 psec.

Structures along the MD trajectory can either belong to the same energy minimum or to different ones. If they belong to the same energy minimum, upon minimization they should show a small RMS deviation ($\leq 0.5 \text{ \AA}$) among each other. If, however, the structures belong to different energy minima, upon minimization RMS deviations among the structures should still be fairly large ($\geq 1.5 \text{ \AA}$). Therefore, NOESY constrained energy minimization was performed on all the 104 structures from the MD trajectory. The (104 X 104) RMS matrix of the 104 minimized also shows a similar

H6(T7)---H2'(T7), H6(T7)---H2''(T6), H6(T7)---H1'(T7), H6(T7)---H1'(T6), H6(T7)---CH3(T8) etc. and similarly for other residues. Estimates of inter-proton distance were extracted for 60 pair-wise interactions by considering 30 NMR models. These 60 distances were taken as structural constraints required for agreement with the NMR data. The B-DNA model of Arnott and Hukins [13], instead of any one of the NMR models, was chosen as the starting model of DNA in the complex. This was done to ensure an unbiased search of the conformational space during minimization. Indeed, energy minimization led to a model of the *Dst2*-d(GA₄T₄C)₂ [1:1] complex that showed agreement with the NOESY data. The energy minimized model have the following features:

- (i) A.T pairs are propellar twisted leading to an array of bifurcated H-bonds in the major-groove of the A/T tract [Fig. 1].
- (ii) The amide protons of *Dst2* form a series of H-bonds in the minor-groove of DNA [Fig. 1]. The location of the drug at the center of the DNA molecule results in the following drug-DNA contacts: H2(A5/A15)---HC3/HC11= 2.4 Å and H1'(T6/T16,T7/T17)---HC3/HC11= 3.2-3.5 Å. This distance profile mimics the observed drug-DNA NOESY pattern [9].

Figure 2 shows the observed and theoretically constructed NOESY slices through H8/H6 protons for τ_m =50 and 100 msec. Theoretical NOESY slices are constructed using the cartesian coordinates of the DNA and drug protons in the energy minimized model of the *Dst2*-d(GA₄T₄C)₂ [1:1] complex. Note that the energy minimized model reproduces the observed NOESY pattern.

An Ensemble of Structures derived from the 52-psec Constrained MD Simulation. The structure obtained after NOESY constrained energy minimization is merely a stable static structure of the *Dst2*-d(GA₄T₄C)₂ [1:1] complex. In reality, NMR data correspond to an ensemble of structures. These structures originate due to the presence of low frequency internal motions in the molecule and there is a continuous interconversion among these structures within

the NMR time scale. In order to obtain an ensemble of structures, we carried out a 52-psec MD simulation (at constant energy with temperature T close to 300 K). The same set of 60 inter-proton distances were kept as structural constraints during the MD simulation and the energy minimized model of the *Dst*2-d(GA₄T₄C)₂ [1:1] complex was taken as the initial structure before temperature equilibration. Hence, this enabled us to monitor the MD trajectory of the structures that agree with the NMR data.

We selected structures along the last 30-psec of the MD trajectory to analyze the NOESY constraints and other structural parameters. 3,000 structures (at time steps of 0.01 psec) were chosen from the trajectory. For the purpose of illustration, we only show here the result of our analysis describing the variations in the intra-nucleotide distances of the type H8/H6(i)---H2'(i) and the inter-nucleotide distances of the type H8/H6(i)---H2''(i-1). These distances are the shortest (hence the most reliable ones) estimated from the Full-Matrix NOESY Simulation [9]. It appears that H8(A_i)---H2'(A_i) for all A's show similar distance distribution profiles and so do H6(T_i)---H2'(T_i) for all T's. Therefore, for a given structure in the trajectory, we computed the distance H8(A_i)---H2'(A_i) averaged over all A's and similarly for H6(T_i)---H2'(T_i). Figure 3A shows the $P(x)$ vs. x (P = % of structures with a given x) for H8(A_i)---H2'(A_i) [dotted line] and H6(T_i)---H2'(T_i) [solid line]. Primary NOESY pattern indicate that the H8(A_i)---H2'(A_i) distances are larger than the H6(T_i)---H2'(T_i) distances. The distribution in Fig. 3A is consistent the observed NOESY pattern. Figure 3B shows $P(x)$ vs. x plots for three types of pairwise interactions i.e., H8(A_i)---H2''(A_{i-1}) averaged over all A's in a given structure of the trajectory [dotted line], H6(T_i)---H2''(T_{i-1}) averaged over all T's except T6/T16 [solid line] and H6(T6/T16)---H2''(A5/A15) [solid-dotted line]. Note that three distance distribution profiles are distinctly different; the same conclusion is also derived from the primary NOESY pattern. The primary NOESY pattern revealed a particularly interesting feature i.e., inside the A/T tract NOE for the intra-nucleotide H6(T_i)---H2'(T_i) interaction is stronger (by a factor of about 2) than the NOE

periodic variation as observed in the previous RMS matrix before minimization i.e., along a row or a column, at first the RMS value gradually increases and then goes to a maximum and finally drops to a minimum. What is more interesting is that upon minimization three structures (designated as I, II and III) are isolated as three well defined minima that show quite large deviations ($\geq 1.5\text{\AA}$) among each other. Thus, in this way we are able to locate at least three distinctly different energy minima in the potential surface. And all these three minimized structures satisfy the NOESY data. Examination of the two (104 X 104) RMS matrices (i.e., before and after minimization) shows that transitions between I and III are very frequent (time period of 5 psec) while transitions between I and II are very rare (happens only once during 52-psec). Thus, emerges a simple picture of dynamics of the *Dst2*-d(GA₄T₄C)₂ [1:1] complex. The structures are located in three distinctly different energy wells and there is a continuous interconversion of the structures within one well and between structures in two different wells (I and III).

Table II lists the base and helix parameters for these structures calculated using the methodology of Soumpasis and Tung [14]. The values of the corresponding parameters for structures I, II and III are different even though it is not immediately obvious how exactly these differences translate into their 3-dimensional structures. However, differences among these structures are clearly visible from Figure 4 in which structures I, II and III are shown along their long-axis. Structures I, II and III clearly show three different states of bending in the DNA decamer duplex. In structure III, the segment

G1- A2 - A3
C20-T19-T18

seems to show a sharp kink with the rest of the molecule. In structure II, this kink is less abrupt while in structure I it is barely visible. Therefore, it appears that upon *Dst2* binding, the region of the target DNA adjacent to the site of binding, can show a characteristic motion. Although, it is visually clear that bending may be a principle component of motion, a quantitative estimate of bending has not yet been done.

Conclusion

In this article we show that an ensemble of structures can be constructed for the $Ds:2\text{-d}(\text{GA}_4\text{T}_4\text{C})_2$ [1:1] complex that agree with the NOESY data. Our studies also reveal that the dynamics of the system can be described by a simple picture in which structures are located in at least three well defined energy wells and there is a continuous interconversion of the structures within one well or between structures in two different wells. It is also apparent that upon drug binding, there is a characteristic motion at the site of binding and adjacent to it.

Acknowledgement

This work was supported by the U.S. Department of Energy, by grants from the National Institute of Health (GM29787) and by a contract from the National Foundation of Cancer Research. The high-field NMR experiments were performed at the Francis Bitter National Magnet Laboratory, MIT. The NMR facility is supported by the grant PR0095 from the Division of Research Resources of the NIH and by the National Science Foundation under contract C-670.

Authors wish to thank Dr. C. -S. Tung for providing us with the base and helix parameters in Table II and for many helpful discussions.

References

1. Zimmer, C. (1975), *Prog. Nucl. Acid Res. Mol. Biol.*, **15**, 285.
2. Dervan, P. B. (1986), *Science*, **232**, 464.
3. Krowicki, Lee, M., Hartley, J. A., Ward, B., Kissinger, K., skorobogaty, A., Dabrowiak, J. C. and Lown, J. W. (1988), in *Structure & Expression, Vol 2: DNA its drug Complexes*, Eds. Sarma, R. H. and Sarma, M. H., Adenine Press, NY, pp 251.
4. Zimmer, C., Luck, G., Burckhardt, G. and Krowicki, K. and Lown, J. W. (1988), in *Structure & Expression, Vol 2: DNA its drug Complexes*, Eds. Sarma, R. H. and Sarma, M. H., Adenine Press, NY, pp 291.
5. Kopka, M. L., Yoon, C., goodsell, D. and Dickerson, R. E. (1985), in *Structure & Motion: Membranes, Nucleic Acids and Proteins*, Eds., Clementi, E., Corongiu, G., Sarma, M. H. and Sarma, R. H., Adenine Press, NY, pp 461.
6. Klevit, R. E., Wemmer, D. E. and Reid, B. R. (1986), *Biochemistry*, **25**, 3296.
7. Umemoto, K., Sarma, M. H., Gupta, G., Luo, J. and Sarma, R. H. (1990), *J. Amer. Chem. Soc.*, in press.
8. Breslauer, K., Farrante, R. and Marky, L. a., Dervan, P. B. and Youngquist, R. S. (1985), in *Structure & Expression, Vol 2: DNA its drug Complexes*, Eds. Sarma, R. H. and Sarma, M. H., Adenine Press, NY, pp 273.
9. Sarma, M. H., Gupta, G., Garcia, A. E., Umemoto, K. and Sarma, R. H. (1990), *Biochemistry*, in press.
10. Gupta, G., Sarma, M. H. and Sarma, R. H. (1988), *Biochemistry*, **27**, 7909.
11. Umemoto, K., Sarma, M. H., Gupta, G. and Sarma, R. H. (1990), *Biochemistry*, in press.
12. Weiner, S. J., Kollman, P. J., Case, D. A., Singh, U. C., ghio, C., Alagona, G., proteta, S. and Weiner, P. (1984), *J. Amer. chem. Soc.*, **106**, 765.
13. Arnott, S. and Hukins, D. W. I., (1972), *J. Mol. Biol.*, **81**, 93.
14. Soumpasis, D. M. and Tung, C. -S. (1988), *J. Biomol. Str. Dyn.*, **6**, 397.

Legends to Figures

Figure 1. Description of bifurcated H-bonds in the major- and minor-groove of the constrained energy minimized structure of the *Dst2*-d(GA₄T₄C)₂ [1:1] complex. (A) Numbering scheme of the DNA duplex in the complex. The array of bifurcated H-bonds along their lengths are shown. These H-bonds are formed between N6(A) and O4(T) across the strand between two neighboring A.T pairs. (B) The array of H-bonds in the minor-groove where N1, N3, N5 and N7 are the proton donors and O2(T) and N3(A) are the acceptors. (C) The numbering scheme of the *Dst2* atoms.

Figure 2. Calculated and observed NOESY (for $\tau_m=50$ and 100 msec) are shown for several H8 and H6 protons through three panels A, B and C.

Figure 3. Analysis of the short inter-proton distances using 3,000 snap-shots taken from the last 30-psec of the MD trajectory. (A) P(x) vs. x plot for the intra-nucleotide distances i.e., H8(A)---H2'(A) [dotted line] and H6(T)---H2'(T) [solid line]. (B) P(x) vs. x plot for inter-nucleotide distances i.e., H8(A_i)---H2"(A_{i-1}) [dotted line], H6(T_i)---H2"(T_{i-1}) except for the ApT sequence [solid line] and H6(T6/T16)---H2"(A5/A15) [solid-dotted line].

Figure 4. Energy minimized structures I, II and III shown from left to right. All three structures show drug-DNA interactions in the minor-groove that are consistent with the NMR data. Although nucleotide conformations of all residues in structures I, II and III belong to an average B1 geometry, there are small but finite differences among the three structures. These small changes are sufficient to produce different bending states as exhibited by structures I, II and III. The RMS values for the three structures are : RMS(I,II)= 3.1 Å, RMS(I,II)= 2.2 Å and RMS(II,III)= 1.9 Å. The total potential energies of the three structures are within about 2 Kcal/Mole of each other (one mole is in terms of one mole of the drug DNA complex).

Table I
Average values and corresponding standard deviations (Å)
of the short inter-proton distances in DNA as obtained from
the 3,000 snap-shots along the last 30-psec of the MD
trajectory.

Proton-Pair	Average Distance, $\bar{x}$	Standard Deviation, σ	Deviation from Gaussian, γ
1. H8(A)---H2'(A)	2.38	0.11	0.14
2. H6(T)---H2'(T)	2.25	0.10	0.25
3. H8(Ai)---H2"(Ai-1)	2.45	0.13	0.17
4. H6(Ti)---H2"(Ti-1) (except for ApT)	2.42	0.12	0.18
5. H6(T6/T16) ---H2"(A5/A15)	2.35	0.11	0.12

$\gamma=(1/\sigma^3)\langle (x - \bar{x})^3 \rangle$ measures the deviation of $P(x)$ from a Gaussian distribution in Fig. 3.

Table II
Base and Helix Parameters of Structures I, II and III Calculated Using
the Methodology of Soumpasis and Tung [14].

Base	Tilt	Helical Twist	Roll	Wedge	Xdis	Ydis	Helical Rise	Prop.	Buckle
I	-	-	-	-	-	-	-	20	-12
G II								14	-14
III								12	-22
A I	5	26	7	8	-1.0	-0.9	3.0	27	4
II	3	31	4	5	-1.2	-1.0	3.0	17	-4
III	2	32	5	6	-1.1	-0.7	3.1	18	-16
A I	0	31	-13	13	-0.9	-0.3	3.1	18	11
II	0	28	-5	5	-1.2	-0.2	3.0	16	1
III	1	32	2	2	-0.8	0.1	3.0	21	-8
A I	-5	35	-13	14	-0.9	0.2	3.1	12	6
II	-4	31	-4	6	-1.1	0.3	3.0	16	0
III	-3	30	3	4	-0.6	0.2	3.0	22	2
A I	-6	25	-5	8	-0.7	0.3	3.2	13	2
II	-6	23	-1	7	-0.4	0.3	3.1	17	-2
III	-5	22	0	5	-0.2	-0.2	3.1	18	0
T I	-9	41	-4	10	-0.1	0.5	3.0	17	3
II	-9	40	-3	9	-0.1	0.6	3.0	20	9
III	-7	41	-3	8	0.0	0.3	3.1	16	8
T I	8	45	-3	8	-0.9	-0.4	2.9	20	3
II	8	39	0	8	-0.8	-0.4	3.0	17	7
III	5	45	-3	9	-0.8	-0.6	2.9	17	7
T I	6	31	-1	6	-0.8	0.0	3.1	19	3
II	6	35	-3	7	-0.8	0.0	3.0	19	3
III	5	31	-3	6	-0.9	0.1	3.1	15	3
T I	4	42	-2	4	-0.8	-0.2	3.0	20	9
II	4	42	-2	5	-0.8	-0.4	3.0	20	9
III	2	42	-5	5	-0.8	0.0	3.0	18	5
C I	-1	40	1	2	-1.2	-0.1	3.1	13	20
II	-1	41	1	1	-1.3	-0.1	3.1	13	20
III	-2	11	1	3	-1.2	0.1	3.1	14	17

I, II and III are the three minimized structures discussed in the text.

Calculated and Observed NOESY Slices for [1:1] *Dst2*-d(GA4T4C)₂ complex

Legends to Figures

Figure 1. Description of bifurcated H-bonds in the major- and minor-groove of the constrained energy minimized structure of the *Dst2*-d(GA₄T₄C)₂ [1:1] complex. (A) Numbering scheme of the DNA duplex in the complex. The array of bifurcated H-bonds along their lengths are shown. These H-bonds are formed between N6(A) and O4(T) across the strand between two neighboring A.T pairs. (B) The array of H-bonds in the minor-groove where N1, N3, N5 and N7 are the proton donors and O2(T) and N3(A) are the acceptors. (C) The numbering scheme of the *Dst2* atoms.

Figure 2. Calculated and observed NOESY (for $\tau_m=50$ and 100 msec) are shown for several H8 and H6 protons.

Figure 3. Analysis of the short inter-proton distances using 3,000 snap-shots taken from the last 30-psec of the MD trajectory. (A) P(x) vs. x plot for the intra-nucleotide distances i.e., H8(A)---H2'(A) [dotted line] and H6(T)---H2'(T) [solid line]. (B) P(x) vs. x plot for inter-nucleotide distances i.e., H8(A_i)---H2''(A_{i-1}) [dotted line], H6(T_i)---H2''(T_{i-1}) except for the ApT sequence [solid line] and H6(T6/T16)---H2''(A5/A15) [solid-dotted line].

Figure 4. Energy minimized structures I, II and III shown from left to right. All three structures show drug-DNA interactions in the minor-groove that are consistent with the NMR data. Although nucleotide conformations of all residues in structures I, II and III belong to an average B1 geometry, there are small but finite differences among the three structures. These small changes are sufficient to produce different bending states as exhibited by structures I, II and III. The RMS values for the three structures are : RMS(I,III)= 3.1 Å, RMS(I,II)= 2.2 Å and RMS(II,III)= 1.9 Å.

Table I

Average values and corresponding standard deviations (Å) of the short inter-proton distances in DNA as obtained from the 3,000 snap-shots along the last 30-psec of the MD trajectory.

Proton-Pair	Average Distance, $\bar{x}$	Standard Deviation, σ	Deviation from Gaussian, γ
1. H8(A)---H2'(A)	2.38	0.11	0.14
2. H6(T)---H2'(T)	2.25	0.10	0.25
3. H8(Ai)---H2''(Ai-1)	2.45	0.13	0.17
4. H6(Ti)---H2''(Ti-1) (except for ApT)	2.42	0.12	0.18
5. H6(T6/T16) ---H2''(A5/A15)	2.35	0.11	0.12

$\gamma = (1/\sigma^3) \langle (x - \bar{x})^3 \rangle$ measures the deviation of $P(x)$ from a Gaussian distribution in Fig. 3.

Table II
Base and Helix Parameters of Structures I, II and III
Calculated Using the Methodology of Soumpasis and Tung
[14].

	Base Tilt	Helical	Roil	Wedge	Xdis	Ydis	Helical	Prop.	Buckle
I	-	-	-	-	-	-	-	20	-12
G II								14	-14
III								12	-22
A I	5	26	7	8	-1.0	-0.9	3.0	27	4
II	3	31	4	5	-1.2	-1.0	3.0	17	-4
III	2	32	5	6	-1.1	-0.7	3.1	18	-16
A I	0	31	-13	13	-0.9	-0.3	3.1	18	11
II	0	28	-5	5	-1.2	-0.2	3.0	16	1
III	1	32	2	2	-0.8	0.1	3.0	21	-8
A I	-5	35	-13	14	-0.9	0.2	3.1	12	6
II	-4	31	-4	6	-1.1	0.3	3.0	16	0
III	-3	30	3	4	-0.6	0.2	3.0	22	2
A I	-6	25	-5	8	-0.7	0.3	3.2	13	-2
II	-6	23	-1	7	-0.4	0.3	3.1	17	-2
III	-5	22	0	5	-0.2	-0.2	3.1	18	0
T I	-9	41	-4	10	-0.1	0.5	3.0	17	3
II	-9	40	-3	9	-0.1	0.6	3.0	20	9
III	-7	41	-3	8	0.0	0.3	3.1	16	8
T I	8	45	-3	8	-0.9	-0.4	2.9	20	3
II	8	39	0	8	-0.8	-0.4	3.0	17	7
III	5	45	-3	9	-0.8	-0.6	2.9	17	7
T I	6	31	-1	6	-0.8	0.0	3.1	19	3
II	6	35	-3	7	-0.8	0.0	3.0	19	3
III	5	31	-3	6	-0.9	0.1	3.1	15	3

Table II(continued)

T I 4	42	-2	4	-0.8	-0.2	3.0	20	9
II 4	42	-2	5	-0.8	-0.4	3.0	20	9
III 2	42	-5	5	-0.8	0.0	3.0	18	5
C I -1	40	1	2	-1.2	-0.1	3.1	13	20
II -1	41	1	1	-1.3	-0.1	3.1	13	20
III -2	41	1	3	-1.2	0.1	3.1	14	17

I, II and III are the three minimized structures discussed in the text.

**Bifurcated H-Bonds
in the Major Groove**

(A)

**Bifurcated H-Bonds
in the Minor Groove**

(B)

(C)

dna dis

UCSF MidasPlus