Theoretical Simulations on d(CGCGAATTCGCG)$_2$ with *cis-syn* Thymine–Thymine Cyclobutane Dimer

Shashidhar N. RAO* and Peter A. KOLLMAN†
Searle Research and Development, 4901 Searle Parkway, Skokie, IL 60077, U.S.A.
†School of Pharmacy, University of California, San Francisco, CA 94143, U.S.A.
(Received February 19, 1993)

Synopsis. We present molecular mechanics and molecular dynamics simulations on the dodecanucleotide d(CGCGAATTCGCG)$_2$, incorporating a *cis-syn* thymine–thymine cyclobutane dimer, represented as T[]T, at the TT site using the all atom force field. Although the starting optimized structures were essentially kink-free, kinked models are observed during the molecular dynamics simulations. Most conformational changes are seen in the sugar puckering and phosphodiester conformations of residues in and around the site of photolesion.

Ultraviolet rays damage DNA through photochemical reactions leading to the formation of cyclobutane dimer and 6—4 photoadducts. Of these, the former lesion is dominant in numbers and its frequency of formation has been correlated to the frequency of mutations within a single gene.[1] In the light of their significance, a number of physicochemical studies have attempted to elucidate the structural changes in DNA upon the modification by these two forms of lesion.[2]

Molecular modeling studies on *cis-syn* thymine–thymine cyclobutane dimer (T[]T) incorporated DNA have reported both kinked and kink-free models.[3,4] Here, we present a preliminary report on an extension of our molecular mechanics investigations by describing the results of molecular dynamics simulations on d(CGCGAATTCGCG)$_2$ incorporating a T[]T cyclobutane dimer replacing TT. The principal goal of these investigations is to explore the conformational space of the dodecanucleotide to identify alternative DNA structures that can accommodate cyclobutane dimers. In contrast to the earlier studies,[3] the dynamics simulations show interesting structural alternatives (including kinked forms) as low energy conformers. The conformationally constrained thymine dimers induce significant bending in the dodecanucleotide. Our results provide useful input to further high resolution solid state and solution studies.

Methods

The *cis-syn* cyclobutane dimer was modeled across the neighboring thymine residues in d(CGCGAATTCGCG)$_2$, in the B-form[5] as described earlier.[3] The oligomer was then surrounded by the solvated Na^+ counterions (ε=0.418 kJ mol^{-1} and $R^*=5.0\times10^{-10}$ m) placed along the phosphate bisector using AMBER.[6] The structure was energy optimized with the all atom force field[7] until a root mean square gradient of <0.1 was obtained and equilibrated using molecular dynamics for 5 ps to 300 K, followed by 80 ps of simulation at 300 K with a time step of 0.001 ps. A residue based nonbonded cutoff of 15×10^{-10} m was used and the nonbonded pair list was updated every 20 steps. Structures were stored every 0.2 ps and analyzed for their torsion angles and helix parameters. The IUPAC nomenclature[8] for nucleic acids are used. According to this nomenclature, the nucleic acid backbone torsions about the C4′–C5′, C3′–O3′, P–O3′, P–O5′, O5′–C5′, and C3′–C4′ are referred to as γ, ε, ζ, α, β, and δ, respectively. Throughout this text, the values of these torsions will be generally described in terms of *gauche*$^+$, *trans*, and *gauche*$^+$. They correspond to ranges centered around 60°, 180°, and −60°, respectively. The glycosidic torsions are represented by χ, and their values are usually referred in terms of *anti* (ca. −30° to 120°) and *syn* (ca. 150° to 250°). The nucleotides are referred to as CYT1, GUA2 etc. on the photodimer containing strand, and as CYT1′, GUA2′ etc. on the complementary strand.

Results

Table 1 contains the average dihedral angles of the dodecamer backbone. The C4′–C5′ torsion, γ, remains predominantly *gauche*$^+$ for most of the nucleotides. However, this torsion undergoes a transition *gauche*$^+$→*trans* in GUA4, ADE5, and GUA10, as was also noted in the molecular dynamics simulations of unmodified DNA.[10] In all these three nucleotides, at the beginning of the simulations (<10 ps), the γ values range in the *gauche*$^+$ region. Interestingly, the average γ value for the thymine dimer region is still *gauche*$^+$ though a few *gauche*$^-$ values are observed. The average C3′–O3′ torsion remains *trans* in all the nucleotides, with the CYC region (3′ ends of ADE6 and ADE6′) exhibiting an average tending towards *gauche*$^-$. This observation is in contrast to the corresponding results on the unmodified oligomers where a nontrivial number of *gauche*$^-$ values were seen.[9,10]

The phosphodiester conformation (ζ,α) remains predominantly (*gauche*$^-$, *gauche*$^-$) except for the following: The P–O3′ torsions (ζ) at the 3′ end of THY8 and CYT3 are around *trans*, while the P–5O′ (α) torsion exhibits greater degree of variability. The 3′ ends of GUA4, CYT9, GUA2′, and GUA4′ have average values in the trans region for α. The torsion ζ in THY8, undergoes a transition from *trans* to *gauche*$^-$ at around 25 ps into the molecular dynamics simulation. The corresponding variations in ζ of CYT3 are more significant, as this torsion flip-flops between the *trans* and *gauche*$^-$ regions.

The P–O5′ torsions at the 3′ ends of GUA4 and GUA4′ spend about 10 and 30 ps, respectively in the

Table 1. Averages of Backbone and Glycosidic Torsion Angles in the Molecular Dynamics Simulations of d(CGCGAATTCGCG)$_2$, Incorporating a T[]T Cyclobutane Dimer over a Period of 80 ps

Torsion angles	γ (C4′–C5′)	ε (C3′–O3′)	ζ (O3′–P)	α (P–O5′)	β (O5′–C5′)	χ (pur) (C1′–N9)	χ (pyr) (C1′–N1)
Aver1	85.7	202.5	263.5	265.6	171.8	45.0	60.1
RMSF	20.5	19.8	26.4	41.8	16.1	19.8	20.4
Aver2	72.7		270.5	279.7			53.6
Aver3	171.8		219.3	227.4			131.1

Aver1 is the average of the torsion angles irrespective of the ranges they lie in. RMSF: Root mean square fluctuations. Aver2 is the average of the torsion angles excluding the non-standard ranges. Aver3 is the average of the torsion angles in the non-standard ranges.
In B-DNA, the standard ranges for γ, ε, ζ, α, and β are *gauche*$^+$, *trans*, *gauche*$^-$, *gauche*$^-$, and *trans* respectively. The glycosidic torsion is typically in the *anti* range.

gauche$^-$ region before undergoing transition to the *trans* region. The P–O5′ torsions at the 3′ ends of GUA2′ spends the first 30 ps in the *gauche*$^-$ region, the next 25 ps in the *trans* region and the subsequent 25 ps back in the *gauche*$^-$ region. By contrast, the corresponding torsion in CYT9 spends all its time in the *trans* region. As earlier,[9,10] the C5′–O5′ torsion (β) is *trans* throughout the simulations. The glycosidic torsions are mostly *anti* (χ ranging from $-20°$ to $140°$) except in THY7 where χ varies from $94°$ to $168°$. The average values for this torsion vary from $20°$ to $75°$ in all nucleotides except for THY7 where the average is $131°$.

The furanose sugars demonstrate considerable flexibility, consistent with the observations based on crystallographic, solution and theoretical studies on various DNA fragments reported in the literature.[11] A number of nucleotides have their average sugar puckers spend a nontrivial percentages of the total simulation times (12 to 55%) in the O1′ *endo* region. The sugar in THY8 has the maximum percentage of O1′ *endo* geometry (55%) while 17 of the 24 nucleotides spend more than 15% time in the O1′ *endo*-C1′ *exo* region. As in our earlier reported simulations on the DNA oligomers,[9,10] no sugar puckers are observed in the higher energy O1′ *exo*–C1′*endo* region. ADE5 has its sugar geometry in the C3′ *endo*–C4′ *exo* region for 79% of the simulation time, while five other sugars spend more than 25% of their time in this region. Fifteen nucleotides have their sugars with the standard C2′ *endo* geometry over more than 50% of the simulation time. This distribution of sugar geometries is in contrast to the case of the corresponding simulations on the unmodified oligomers which predominantly exhibit C2′ endo geometry, as found in the starting model.[5]

The dodecamer models with cyclobutane dimer undergo significant kinking and bending, around the site of the dimer (Fig. 1) during the molecular dynamics simulations. These structural variations are consistent with the calculated average helix twists listed in Table 2. The average value is ca. 33.3 ($n\approx10.8$). Two of the helix repeats (26.5° in the ADE6–THY7 step and 22.3° in the THY7–THY8 step) overwind the helix leading effectively to a kink and closely resembling the canonical A-DNA structure. In the regions removed from the cyclobutane dimer, the helix characteristics are predominantly those of B-DNA. The hydrogen bonding between the base pairs in regions other than that containing the T[]T dimer is not significantly altered.

Table 2. Average Helix Twists in All the Dinucleotide Steps during the Molecular Dynamics of d(CGCGAATTCGCG)$_2$, Incorporating a T[]T Cyclobutane Dimer over a Period of 80 ps

Step	Helix repeat (rmsf)
CYT1–GUA2	43.6 (11.4)
GUA2–CYT3	31.7 (4.0)
CYT3–GUA4	38.3 (5.6)
GUA4–ADE5	31.9 (3.8)
ADE5–ADE6	31.6 (4.3)
ADE6–THY7	26.5 (5.3)
THY7–THY8	22.3 (3.6)
THY8–CYT9	35.6 (3.9)
CYT9–GUA10	34.4 (5.0)
GUA10–CYT11	32.5 (5.6)
CYT11–GUA12	38.3 (6.7)
Average	33.3 (5.4)
Number of residues/turn	10.8 (9.3—12.9)

Root mean square fluctuations are indicated in parentheses.

Are the results of our calculations general enough to be applicable to oligonucleotide sequences consistent of non-self-complementary sequences ? Although we have not carried out explicit calculations to investigate the effects of non-self-complementary sequences, the extents of variations in the DNA conformational characteristics are suggestive of possible accommodation of asymmetric sequences within the DNA altered by photolesion. The key conformational parameters, the phosphodiester torsions and sugar geometry do exhibit wide variations in the non-cyclobutane dimer part of the oligonucleotide structures. Such variations are capable of absorbing non-self-complementary sequences of bases (unpublished data).

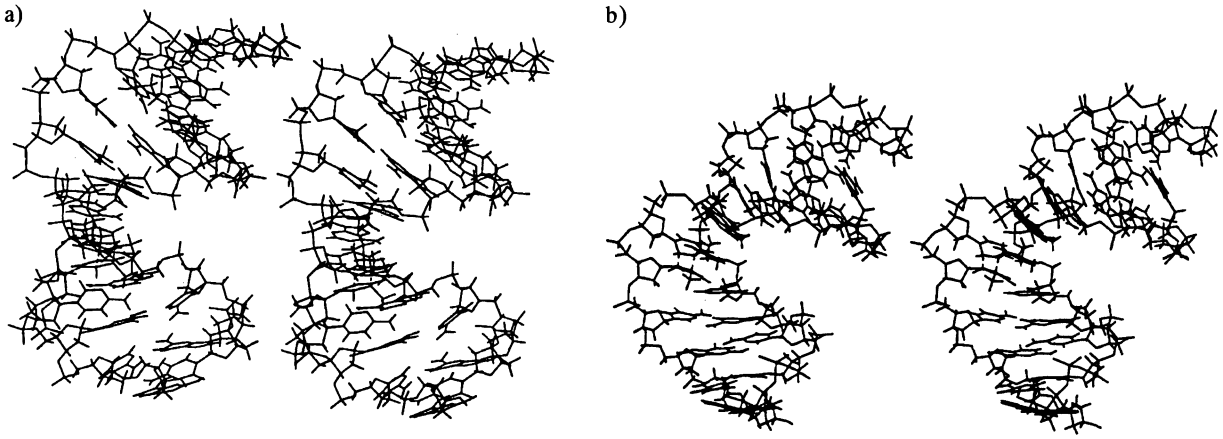

Fig. 1. Stereo pairs of d(CGCGAATTCGCG)$_2$, with T[]T cyclobutane dimer at the end of 40 ps (a) and 80 ps (b) of molecular dynamics.

How do the molecular dynamics models stack up with the biological significance of photolesions and their recognition by the complexes of repair enzymes ? In contrast to one of the previously reported molecular mechanics models,[3] the present studies demonstrate significant structural changes in the overall DNA double helix in addition to the conformational variations in the dimer region. It is tempting to speculate that the repair enzymes do recognize the gross structural changes (kinks and bends, coupled with variations in groove sizes). The conformational changes in the T[]T dimer region, particularly those found in the base orientations relative to the furanose, could be potential basis of recognition of the photolesion by the enzymes involved in the *N*-glycosylase activity.[12]

Conclusions

Molecular dynamics simulations on *cis-syn* thymine–thymine cyclobutane dimer incorporated dodecamer show significant kinks and bends and have helical parameters characteristic of the A-form rather than the B-form. These structural features result principally from the overwinding of the double helix in the region of photolesion. The models discussed here could provide structural insights into the mechanisms of actions of repair enzymes that recognize the cyclobutane dimers. Our results will hopefully stimulate further high resolution X-ray crystallographic and solution studies to characterize the structures of photolesion incorporated nucleic acids.

Financial support from the NCI (grant # CA-25644) is gratefully acknowledged. We thank Drs. Robert Langridge (Director) and T. E. Ferrin (Facility manager) for the usage of the computer graphics facilities at the University of California, San Francisco.

References

1) W. A. Haseltine, *Cell*, **33**, 13 (1983).
2) a) V. Lyamichev, *Nucleic Acids Res.*, **19**, 4491 (1991); b) F. Hruska, D. Wood, K. Ogilvie, and J. Charlton, *Can. J. Chem.*, **53**, 1193 (1975).
3) S. N. Rao, J. W. Keepers, and P. A. Kollman, *Nucleic Acids Res.*, **12**, 4789 (1984).
4) D. A. Pearlman, S. R. Holbrook, D. H. Pirkle, and S. -H. Kim, *Science*, **227**, 1304 (1985).
5) S. Arnott, P. Campbell-Smith, and R. Chandrasekaran, in "CRC Handbook of Biochemistry," ed by G. D. Fasman, CRC Cleveland, Ohio (1976), Vol. II, pp. 411—422.
6) U. C. Singh, P. K. Weiner, J. W. Caldwell, and P. A. Kollman, "AMBER-UCSF (v3.0)," Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143 (1986).
7) S. J. Weiner, P. A. Kollman, D. Nguyen, and D. Case, *J. Comput. Chem.*, **7**, 230 (1986).
8) IUPAC-IUB Joint Commission on Biochemical Nomenclature, *Eur. J. Biochem.*, **131**, 9 (1983).
9) S. N. Rao, U. C. Singh, and P. A. Kollman, *Isr. J. Chem.*, **27**, 189 (1987).
10) S. N. Rao and P. A. Kollman, *Biopolymers*, **29**, 517 (1990).
11) W. Saenger, in "Principles of Nucleic Acid Structure," ed by W. Saenger, Springer-Verlag, New York (1984), pp. 44—199.
12) L. K. Gordon and W. A. Haseltine, *J. Biol. Chem.*, **256**, 6608 (1981).