Synthesis and X-Ray Crystal Structure of Novel *trans-syn* Thymine Photodimers: Effect of a Polyoxyethylene Spacer Chain on Photodimer Stereochemistry

Bargur P. Gangamani,[a] Cheravakkattu G. Suresh*[b] and Krishna N. Ganesh*[a]

[a] *Division of Organic Chemistry (syn)*
[b] *Division of Biochemical Sciences, National Chemical Laboratory, Pune-411008 India*

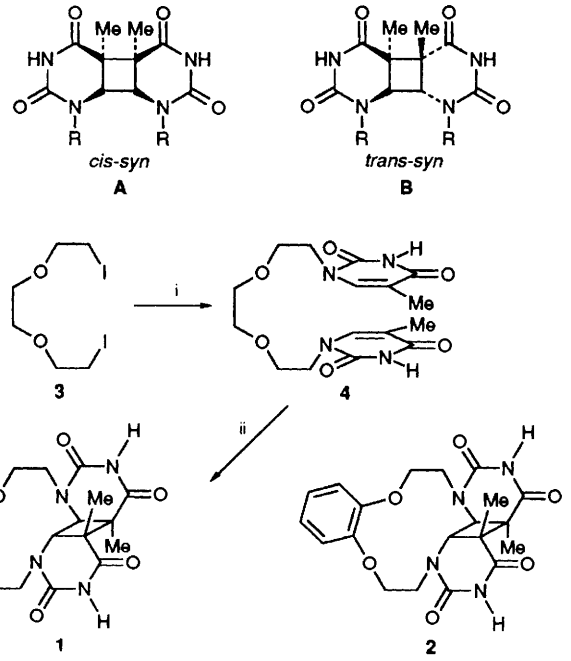
X-ray crystal structures have been determined for the *trans-syn* photodimers **1** and **2** which are the exclusive products upon photodimerization of bis-thymines.

UV light induces adjacent stacked pyrimidine bases in DNA to undergo [2 + 2] cycloaddition reaction to produce cyclobutane linked pyrimidine photodimers.[1] Among the various possible isomers of thymine photodimers, only two are normally formed in DNA: the *cis-syn* T[*c-s*]T (**A**) in both duplex and single stranded (ss) DNA and the *trans-syn* T[*t-s*]T (**B**) formed only in ss DNA, the ratio of the products being 7 : 1 in favour of the *cis-syn* isomer. DNA photolyase, a DNA repair enzyme, reverts pyrimidine photodimers with high specificity towards the *cis-syn* isomers.[2] Attempts to understand and model this reaction have resulted in extensive research on the chemistry of *cis-syn* photodimers from natural dinucleotides[3] and on models possessing two thymine bases linked *via* non-sugar phosphate spacers.[4] In contrast, much less is known about the repair mechanism of *trans-syn* isomers as they are produced in very low amounts. T[*t-s*]T isomers have biological significance as lethal mutagenic lesions since they are not substrates for most DNA photolyases.[5] Recently, it has been demonstrated that *Escherichia coli* DNA photolyase repairs T[*t-s*]T dimers 10^4 fold less effectively than T[*c-s*]T isomers.[6]

In our attempts to investigate the effect of spacer chain on thymine photodimer formation, we chose to synthesize hitherto unknown polyoxyethylene linked bis-thymines. Herein we report interesting results on the exclusive formation of the rare *trans-syn* thymine photodimers **1**, **2** from these model compounds. They have been characterized by X-ray crystallography which reveals interesting structural features. It may be mentioned that only one crystal structure of a *trans-syn* pyrimidine photodimer from orotic acid methyl ester is known.[7]

1,2-bis(2-iodoethoxy)ethane **3** was treated with thymine in hexamethyldisilazane and trimethylsilyl chloride to furnish 1,2-bis(2-thymin-1-ylethoxy)ethane **4** in 65% yield after chromatography (Scheme 1).[8] The bis-thymine **4** was photoirradiated using a Hanovia lamp PC-451 with a Pyrex filter (λ = 290 nm) in acetonitrile–water (7 : 3 v/v) with acetophenone as the sensitizer. This yielded the photodimer **1** as the sole product (HPLC)† in 98% yield. The structure of the photodimer was supported by spectroscopic data†, with an upfield shift of CH_3 and H6 signals in the ^1H NMR spectrum and disappearance of the absorption at λ = 270 nm in the UV spectrum. The photodimer **2** having a rigid aromatic ring as part of the spacer chain was synthesized from 1,2-bis(2-bromoethoxy) catechol by following a similar sequence of reactions.

X-Ray diffraction studies‡ were carried out on the crystals of both **1** and **2**. This has conclusively shown them to be *trans-syn* cyclobutane dimers. The magnitudes of the bond lengths and angles (Figs. 1 and 2) show that the geometry around the cyclobutyl ring is highly strained in order to accommodate the methyl groups attached to C(5) and C(5′). The effect of this could be observed in the lengthening of C(5)–C(5′) single bond and a corresponding widening of the angles opposite to this and excluding this bond in the cyclobutyl ring. Similarly the bond angles C(5′)–C(5)–C(7) and C(5)–C(5′)–C(7′) are also widened. The four-membered cyclobutyl ring is folded with each of the four atoms deviating



Scheme 1 i, HMDS, TMSCl, thymine; ii, hv, acetophenone

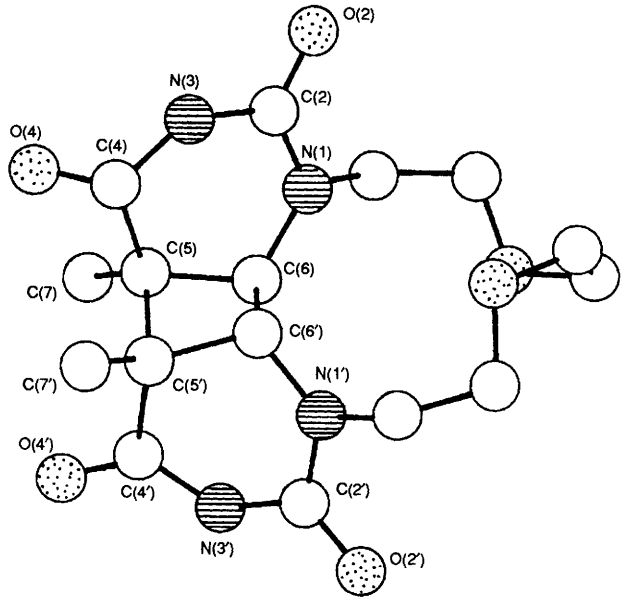


Fig. 1 Structure of compound **1** with hydrogen atoms omitted for clarity. Selected bond lengths (Å), bond angles (°) and interatomic distances (Å) are: N(1)–C(6) 1.466(7), C(4)–C(5) 1.513(13), C(5)–C(6) 1.541(11), C(6)–C(6′) 1.536(11), C(5)–C(5′) 1.598(15), N(1′)–C(6′) 1.441(8), C(4′)–C(5′) 1.542(12), C(5′)–C(6′) 1.543(10), N(1)–C(6)–C(5) 115.7(5), C(4)–C(5)–C(6) 110.5(6), C(6′)–C(6)–C(5) 88.7(6), C(6)–C(6′)–C(5′) 89.9(6), C(6)–C(5)–C(5′) 87.7(6), C(6′)–C(5′)–C(5) 86.5(6), N(1′)–C(6′)–C(5′) 115.5(5), C(4′)–C(5′)–C(6′) 111.3(7), C(5′)–C(5)–C(7) 122.8(6), C(5)–C(5′)–C(7′) 122.9(9), C(6)–C(5)–C(7) 117.1(9), C(6′)–C(5′)–C(7′) 118.8(11), C(7)...C(7′) 3.75(3)

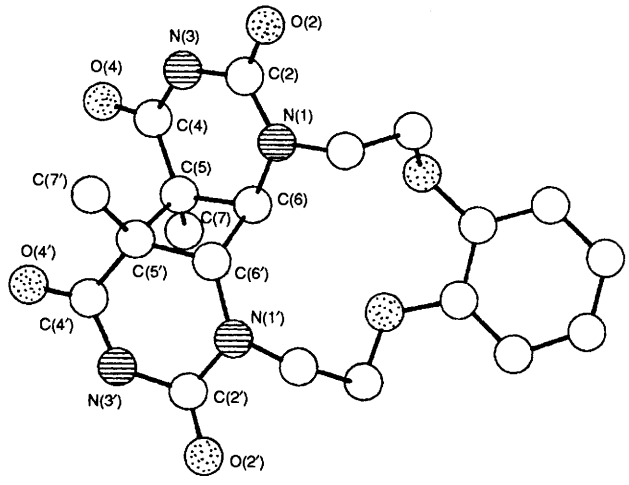


Fig. 2 Structure of compound **2** with hydrogen atoms omitted for clarity. Selected bond lengths (Å), bond angles (°) and interatomic distances (Å) are: N(1)–C(6) 1.46(1), C(4)–C(5) 1.49(1), C(5)–C(6) 1.56(1), C(6)–C(6′) 1.54(1), C(5)-C(5′) 1.59(1), N(1′)–C(6′) 1.46(1), C(4′)–C(5′) 1.49(1), C(5′)–C(6′) 1.56(1), N(1)–C(6)–C(5) 114.1(8), C(4)–C(5)–C(6) 116.9(8), C(6′)–C(6)–C(5) 89.5(7), C(6)–C(6′)–C(5′) 91.3(7), C(6)–C(5)–C(5′) 89.3(7), C(6′)–C(5′)–C(5) 88.1(6), N(1′)–C(6′)–C(5′) 115.0(8), C(4′)–C(5′)–C(6′) 118.8(8), C(5′)–C(5)–C(7) 116.0(8), C(5)–C(5′)–C(7′) 116.3(8), C(6)–C(5)–C(7) 110.4(9), C(6′)–C(5′)–C(7′) 111.0(8), C(7)...C(7′) 3.93(2)

from the plane of the other three by dihedral angles of 20 and 10° for **1** and **2**, respectively. However, this folding is in an opposite sense for **1** and **2** when identical enantiomers are compared. This contrast in the folding of cyclobutyl ring thus increases the separation of the methyl groups at C(5) and C(5′) in **2** while they are brought closer in **1**. This can also be inferred from the C(7)···C(7′) separations. The mutual disposition of these methyl groups thus enhances the strain around the cyclobutyl ring in **1** compared to **2**. Surprisingly the rigid catechol group in the spacer chain reduces the strain around the cyclobutyl ring upon formation of the thymine dimer.

The non-availability so far of structural data on *trans-syn* thymine dimers could either be due to lack of good crystals or result from monomerization upon X-ray irradiation. Here we have reported on the structure of two *trans-syn* dimers linked through linker chains; with these compounds monomerization was not a problem. Further work is in progress to study photodimers with larger spacers and novel structures to aid in the understanding of stereochemical effects during photodimerisation and photoreversion.

B. P. G. thanks UGC-New Delhi for financial support. This work is N. C. L. Communication No. 6017

Received, 18th May 1994; Com. 4/02958E

Footnotes

† All new compounds gave satisfactory spectroscopic and analytical data: Selected ^{1}H NMR (200 MHz, $CDCl_3$) spectroscopic data, Compound **1**: δ 1.24 (s, 2 × CH_3), 3.65 (m, 4 × OCH_2, 2 × NCH_2, 2 × H6), 9.2 (2 × NH). **2**: δ 1.25 (s, 2 × CH_3), 3.92 (m, 2 × OCH_2, 2 × NCH_2, 2 × H6), 6.81 (s, 4 × Ar-H) 9.85 (2 × NH). **4**: δ 1.93 (d, 2 × CH_3, J = 1.2 Hz), 3.61 (s, OCH_2CH_2O), 3.72 (t, 2 × NCH_2, J = 5 Hz), 3.91 (t, 2 × CH_2, J = 5 Hz), 7.13 (d, 2 × H6, J = 1.2 Hz), 8.35 (2 × NH). HPLC, Compound **1**: R_t = 4.37 min, buffer; 17% MeOH, 0.1 mol dm^{-3} $NHEt_3OAc$ flow rate: 2 ml/min^{-1}. **4** R_t = 2.2 min, buffer; 40% MeOH, 0.1 mol dm^{-3} $NHEt_3OAc$ flow rate: 1ml/min^{-1} RP-C-18 column, isocratic run.

‡ *Crystal data*: X-Ray data on both compounds were collected on a CAD-4 P.C. controlled diffractometer using ω–2θ scan mode and Mo-Kα (λ = 0.71069 Å) radiation. Unit-cell parameters were refined from the setting angles of 25 well centred reflections. Both **1** and **2** were colourless plates. Lorentz and polarisation corrections were applied to reflection data but absorpion corrections were not applied. The structures were solved by direct methods (SHELXS-86)[9] and refined by full-matrix least-squares analysis on F (SHELX-76)[10] with non-hydrogen atoms anisotropic and hydrogens isotropic. Hydrogens were either located in the difference map or fixed at ideal geometries. Compound **1**: $C_{16}H_{22}N_4O_6$, M = 366.4, triclinic, space group $P\bar{1}$, a = 7.900(4), b = 10.061(8), c = 12.048(4) Å, α_c = 71.37(5), β = 78.13(3), γ = 72.17(5)°, V = 857.7(9) Å^3 Z = 2, D_c = 1.418 gcm^{-3}, for cell refinement 14.2 ≤ 2θ ≤ 27.8°, F(000) = 388, μ = 0.69 cm^{-1}, T = 295 K, crystal size: 0.1 × 0.3 × 0.6 mm, 2752 measured reflections with 2535 unique, 1482 reflections [$F > 5\sigma(F)$] used in the refinement, 0 ≤ h ≤ 8, −11 ≤ k ≤ 11, −13 ≤ l ≤ 13, $2\theta_{max}$ = 47°, three standard reflections monitored every 3600 s showed a maximum variation of 1.1%, R = 0.0952, R_w = 0.1056 with w = 4.0133/[$\sigma^2(F)$ + 0.002153F^2], highest shift/esd in the final cycle among 235 parameters 0.018, $\Delta\rho_{max,min}$ = 0.38 −0.51 e Å^{-3}.

Compound **2**: $C_{20}H_{22}N_4O_6$, M 414.4, monoclinic space group, $P2_1/n$, a = 10.357(4), b = 14.578(5), c = 12.676(2) Å, β = 97.99(3)°, V = 1895(1)Å^3, Z = 4, D_c = 1.453 g cm^{-3}, for cell refinement 10 ≤ 2θ ≤ 30°, F(000) = 872, μ = 0.68 cm^{-1}, T = 295 K, crystal size: 0.02 × 0.12 × 0.25 mm, 2954 measured reflections with 2629 unique reflections, 1077 reflections [$F > 5\sigma(F)$] used in the refinement 0 ≤ h ≤ 11, ≤ k ≤ 16, −13 ≤ l ≤ 13, $2\theta_{max}$ = 46°, three standard reflections monitored every 5000 s showed only 1% variation. R = 0.0628, R_w = 0.0521, w = 2.6968/[$\theta^2(F)$ + 0.0F^2], highest shift/esd ratio in the final cycle among 358 parameters 0.199, $\Delta\rho_{max,min}$ = 0.32, −0.29 e Å^{-3}. Atomic coordinates, bond lengths and angles and thermal parameters for **1** and **2** have been deposited at the Cambridge Crystallographic Data Centre. See Information for Authors, Issue No. 1.

References

1 G. J. Fisher, and H. E. Johns, in *Photochemistry and Photobiology of Nucleic Acids*, ed. S. Y. Wang, Academic, 1976, vol. 1, pp. 225–294; J. Cadet and P. Vigny, in *Bio-organic Photochemistry*, ed H. Morrison, Wiley interscience, 1990, vol. 1, pp. 1–272.
2 A. Sancar, *Biochemistry*, 1994, **33**, 2; S. T. Kim, and A. Sancar, *Photochem. Photobiol.*, 1993, **57**, 895.
3 A. Sancar, and S.T. Kim, *Biochemistry*, 1991, **30**, 8623.
4 R. F. Hartmann and S. D. Rose, *J. Org. Chem.* 1992, **57**, 2302; S. R. Yeh and D. E. Falvey, *J. Am. Chem. Soc.* 1991, **113**, 8557; M. W. Logue and N. J. Leonard, *J. Am. Chem. Soc.* 1972, **94**, 2893; M. S. Goodman and S. D. Rose, *J. Org. Chem.* 1992, **57**, 3268; R. Austin, S. McMordie, E. Altmann and T. P. Begley, *J. Am. Chem. Soc.* 1993, **115**, 10370.
5 A. Sancar and G. B. Sancar, *Ann. Rev. Biochem.* 1988, **57**, 29; J.-S. Taylor, *Acc. Chem. Res.* 1994, **27**, 76.
6 S. T. Kim, K.Malhotra, C. A. Smith, J.-S. Taylor and A. Sancar, *Biochemistry*, 1993, **32**, 7067.
7 G. I. Birnbaum, *Acta. Crystallogr., Sect. B.*, 1972, **28**, 1248.
8 The residue was adsorbed and loaded on a silica gel (100–200 mesh) column and eluted with increasing gradient 1–10% MeOH–CH_2Cl_2; R_f = 0.35 in 10% MeOH–CH_2Cl_2. See D. T. Browne, in *Synthetic Procedures in Nucleic Acids*, Interscience, 1968, vol. 1, p. 96.
9 G. M. Sheldrick, SHELXS-86, Program for Crystal Structure Determination, University of Göttingen, 1986.
10 G. M. Sheldrick, SHELX-76, Program for Crystal Structure Analysis, University of Cambridge, 1976.