

The Rare *trans-syn* Thymine Photodimers: Structures of Polyoxyethylene-Linked Bisthymines and the Derived *trans-syn* Thymine Photodimers. Comparison of the Stereochemistry Before and After Photodimerization

C. G. SURESH,^{a*} B. P. GANGAMANI^b AND K. N. GANESH^b

^aDivision of Biochemical Sciences, and ^bDivision of Organic Chemistry (Synthesis), National Chemical Laboratory, Pune-411008, India. E-mail: suresh@ncl.ernet.in

(Received 8 October 1994; accepted 29 August 1995)

Abstract

trans-syn Photodimer of polyoxyethylene bisthymine (1) [1,1'-(ethylenedioxydiethylene)dithymine], *trans-syn*-1,8-(ethylenedioxyethylene)hexahydro-4a,4b-dimethylcyclobuta[1,2-*d*:4,3-*d'*]dipyrimidine-2,4,5,7(3*H*,6*H*)-tetrone (2), C₁₆H₂₂N₄O₆, *M_r* = 366.4, triclinic *P*1̄, *a* = 7.900 (4), *b* = 10.061 (8), *c* = 12.048 (4) Å, α = 71.37 (5)°, β = 78.13 (13)°, γ = 72.17 (5)°, *V* = 857.7 (9) Å³, *Z* = 2, *D_x* = 1.418 g cm⁻³, μ = 0.103 mm⁻¹, *F*(000) = 388, *R* = 0.0953 for 1482 observed reflections; 1,1'-(1,2-phenylenedioxydiethylene)dithymine (3), C₂₀H₂₂N₄O₆, *M_r* = 414.4, triclinic, *P*1̄, *a* = 8.558 (2), *b* = 11.077 (2), *c* = 11.437 (3) Å, α = 108.57 (2)°, β = 94.37 (2)°, γ = 99.77 (2)°, *V* = 1003.1 (4) Å³, *Z* = 2, *D_x* = 1.377 g cm⁻³, μ = 0.096 mm⁻¹, *F*(000) = 436, *R* = 0.0443 for all reflections; *trans-syn* photodimer of (3), 1,8-(1,2-phenylenedioxydiethylene)hexahydro-4a,4b-dimethylcyclobuta[1,2-*d*:4,3-*d'*]dipyrimidine-2,4,5,7(3*H*,6*H*)-tetrone (4), C₂₀H₂₂N₄O₆, *M_r* = 414.4, monoclinic, *P*2₁/n, *a* = 10.357 (4), *b* = 14.578 (5), *c* = 12.676 (2) Å, β = 97.99 (3)°, *V* = 1895 (1) Å³, *Z* = 4, *D_x* = 1.453 g cm⁻³, μ = 0.102 mm⁻¹, *F*(000) = 872, *R* = 0.0628 for 1077 observed reflections. The bisthymines (1) and (3) on photodimerization produce cyclobutane-linked *trans-syn* thymine dimers as the exclusive products. The planar thymines stack one above the other in (1), while they are approximately perpendicular in (3). Planarities of the thymines are altered on photodimerization. The stereochemistry around the cyclobutane ring in photoproducts (2) and (4) is highly strained. The distortions in the bond lengths and angles due to this strain are concentrated mainly on the C5 and C6 atoms, which form the interpyramidal bonds. This strain is comparatively less in the rigid catechol-linked dimer (4). The N3 atom of thymines forms hydrogen bonds in all structures. Photoproducts show closer packing compared with the parent compounds.

1. Introduction

The deleterious effects of UV radiation on the genetic material of living cells are well known (Setlow, 1978).

The most susceptible sites to UV radiation in deoxyribonucleic acid (DNA), the primary carrier of genetic information, are the adjacently stacked pyrimidine bases (Taylor, 1990). Any mutation in DNA due to photodamage, if not repaired correctly, would lead to cell death or cancer. Skin cancer is known to be caused by damage to DNA by sunlight (Taylor, 1994). Patients suffering from the genetic disorder *Xeroderma pigmentosum*, characterized by a defective repair mechanism of DNA photodamage, are reported to be most sensitive to sunlight (Setlow, 1978). Presently, research in the field of the photochemistry of nucleic acids has acquired special status due to a fast depleting ozone layer and the consequent mutagenic and carcinogenic effects of solar UV radiation increasingly felt by the biosphere.

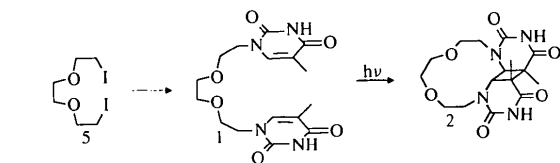
The most significant photoproducts of DNA, both in terms of easy formation and adverse effects, are the cyclobutane dimers of thymine (Fisher & Johns, 1976). The possible cyclobutane photodimers from DNA, based on their stereochemistry, can be classified into *cis-syn* and *trans-syn* isomers (Cadet & Vigny, 1990). The former is an irradiation product of either duplex or single-stranded DNA (ssDNA), whereas the latter is formed only from ssDNA or at the junctions of DNA chains with different conformations (Taylor, 1990). Organisms have developed various repair mechanisms to combat DNA damage (Sancar & Tang, 1993; Sancar & Sancar, 1988). In a most ingenious method, the energy of the damaging radiation, in the near-UV region, is being utilized by a DNA photolyase enzyme system to monomerize the pyrimidine photodimers (Kim & Sancar, 1993). Photolyases were often found to have comparatively higher or even exclusive specificity towards *cis-syn* rather than *trans-syn* thymine dimers (Sancar & Sancar, 1988). As a result, studies have been oriented heavily towards an understanding of the repair mechanism of comparatively easily available *cis-syn* thymine dimers of natural dinucleotides or model compounds. There are few such studies for *trans-syn* thymine photodimers.

Recently, it has been demonstrated that *E. coli* DNA photolyase does repair *trans-syn* thymine dimers, but with 10⁴-fold less efficiency compared with the *cis-syn*

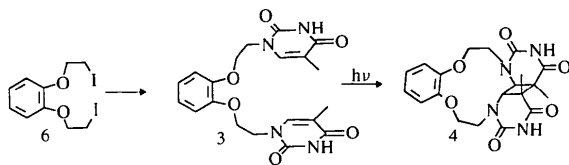
thymine dimer (Kim, Malhotra, Smith, Taylor & Sancar, 1993). Such a subdued mechanism present in the cell for the repair of *trans-syn* dimers deem them to be extremely lethal to life. In our attempts to understand and mimic the stereochemical analogues of cyclobutane thymine dimers, we synthesized model compounds wherein two thymines were linked by a polyoxyethylene chain. These compounds on photodimerization yielded exclusively the *trans-syn* thymine dimers as photoproducts (Gangamani, Suresh & Ganesh, 1994). The stereochemistry of the product compounds has been established conclusively through X-ray crystal structure determination. It may be noted that the only other crystal structure reported before for a *trans-syn* isomer was that of the model compound orotic acid methyl ester (Birnbaum, 1972). Here we report the detailed crystal structures of the photoproduct (2) of polyoxyethylene bithymine (1), catechol polyoxyethylene bithymine (3) and its photoproduct (4). Thus, the crystal structures of two model compounds both before and after photodimerization are available and provide an opportunity to assess structural effects and the stereochemical outcome of photodimerization of thymines. A comparison of the structures of two photodimers, one with the polyoxyethylene spacer chain and the other containing an additional catechol group in the spacer, could delineate the effects of spacer chains on the stereochemistry of *trans-syn* thymine photodimers.

2. Experimental

Compound (1) was prepared by treating 1,2-bis(2-iodoethoxy)ethane (5) with thymine in hexamethyldisilazane and trimethylsilyl chloride. Compound (2), the photoproduct of (1), was obtained by photoirradiating (1) in an acetonitrile–water (7:3 v/v) mixture using a Hanovia lamp PC-451 with a Pyrex filter ($\lambda = 290$ nm) in the presence of acetophenone as the sensitizer (Scheme I). Compound (3) was synthesized by treating 1,2-bis(2-iodoethoxy)benzene (6) with thymine, following a procedure similar to that used for (1). Photodimerization of (3) to produce (4) was achieved using a procedure identical to that described for (2) (Scheme II).



Scheme I.



Scheme II.

Plate crystals of (1) and (3) were readily obtained from methanol solutions of the respective compounds on evaporation. Compound (4) also crystallized from methanol as thin plates under controlled evaporation. Crystals suitable for X-ray analysis of (2) could be prepared only by repeated seeding of the initially grown very thin long crystals in methanol–water solution. The resultant crystals were not of the excellent quality reflected in the slightly high *R* factor for (2). However, we decided to use them for analysis in the absence of better crystals. All crystals were colourless.

X-ray intensity data for all compounds were collected on an Enraf–Nonius CAD-4 PC controlled single crystal diffractometer using zirconium-filtered Mo $K\alpha$ radiation ($\lambda = 0.71068$ Å). Data collection was carried out at 295 K employing ω -2 θ scans with scan width $(0.80 + 0.35 \tan \theta)^\circ$. Data were assimilated in 96 steps with 16 steps on either side being used for background estimation. The intensities of three reflections monitored at intervals of 3600 s for the first three and 5000 s for the last showed no significant variation throughout data collection, ruling out any crystal damage, especially to photoproducts, on incidence of X-rays. Further, the unit-cell estimation before and after data collection showed little difference. Lorentz and polarization corrections were applied to the reflection data during data reduction. However, no absorption correction was applied.

The crystal structures were solved by direct methods and Fourier calculations in *SHELXS86* (Sheldrick, 1985) and refined by full-matrix least-squares in *F* using *SHELXL76* (Sheldrick, 1976). After the input and refinement of non-H atoms, with anisotropic temperature factors, H atoms were either fixed at ideal geometries or located from a difference map phased on refined non-H atom positions. Hydrogens were assigned the isotropic thermal parameters of the atoms to which they were bonded. In all cases except (2) the hydrogen parameters were refined. Nonetheless, no H atoms, although included for structure-factor calculations, were refined in (2) during the last cycles as some tend to assume unacceptable values. Experimental details, including parameters pertaining to data collection and structure refinement, are summarized in Table 1. [An earlier report (Bideau, Courseille & Bravic, 1982) on the structure of (1) has been brought to our notice subsequent to submission of this manuscript. Hence, we include our parameters for (1) only as supplementary material,* although the comparisons here are based on our structure determination.] Computations were carried out on a Magnum RISC computer with a unix operating system.

* Lists of structure factors, atomic coordinates, anisotropic displacement parameters and complete geometry have been deposited with the IUCr (Reference: HU0428). Copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 1. *Experimental details*

	(1)	(2)	(3)	(4)
Crystal data				
Chemical formula	C ₁₆ H ₂₂ N ₄ O ₆	C ₁₆ H ₂₂ N ₄ O ₆	C ₂₀ H ₂₂ N ₄ O ₆	C ₂₀ H ₂₂ N ₄ O ₆
Chemical formula weight	366.4	366.4	414.4	414.4
Cell setting	Triclinic	Triclinic	Triclinic	Monoclinic
Space group	<i>P</i> $\bar{1}$	<i>P</i> $\bar{1}$	<i>P</i> $\bar{1}$	<i>P</i> 2 ₁ / <i>n</i>
<i>a</i> (Å)	8.534 (4)	7.900 (4)	8.558 (2)	10.357 (4)
<i>b</i> (Å)	9.352 (1)	10.061 (8)	11.077 (2)	14.578 (5)
<i>c</i> (Å)	12.371 (2)	12.048 (4)	11.437 (3)	12.676 (2)
α (°)	92.59 (1)	71.37 (5)	108.57 (2)	90.0
β (°)	96.16 (2)	78.13 (3)	94.37 (2)	97.99 (3)
γ (°)	116.05 (2)	72.17 (5)	99.77 (2)	90.0
<i>V</i> (Å ³)	877.3 (5)	857.7 (9)	1003.1 (4)	1895 (1)
<i>Z</i>	2	2	2	4
<i>D_x</i> (Mg m ⁻³)	1.387	1.418	1.377	1.453
Radiation type	Mo <i>K</i> α	Mo <i>K</i> α	Mo <i>K</i> α	Mo <i>K</i> α
Wavelength (Å)	0.71073	0.71073	0.71073	0.71073
No. of reflections for cell parameters	25	25	25	25
θ range (°)	8.65–16.4	7.1–13.9	7.5–19.95	5–15
μ (mm ⁻¹)	0.068	0.103	0.096	0.102
Temperature (K)	295	295	295	295
Crystal form	Block	Block	Block	Block
Crystal size (mm)	0.50 × 0.12 × 0.10	0.60 × 0.30 × 0.10	0.65 × 0.40 × 0.25	0.25 × 0.20 × 0.12
Crystal colour	Colourless	Colourless	Colourless	Colourless
Data collection				
Diffractionmeter	Enraf–Nonius CAD-4	Enraf–Nonius CAD-4	Enraf–Nonius CAD-4	Enraf–Nonius CAD-4
Data collection method	ω -2 θ	ω -2 θ	ω -2 θ	ω -2 θ
No. of measured reflections	2801	2752	3187	2954
No. of independent reflections	2596	2535	2955	2629
No. of observed reflections	1499	1482	2818	1077
Criterion for observed reflections	<i>F</i> > 5σ(<i>F</i>)	<i>F</i> > 5σ(<i>F</i>)	<i>F</i> > 0	<i>F</i> > 5σ(<i>F</i>)
<i>R</i> _{int}	0.0272	0.0455	0.0234	0.0485
$\theta_{\max}$ (°)	23.5	23.5	23.5	23
Range of <i>h</i> , <i>k</i> , <i>l</i>	0 → <i>h</i> → 9 -10 → <i>k</i> → 10 -13 → <i>l</i> → 13	0 → <i>h</i> → 8 -11 → <i>k</i> → 11 -13 → <i>l</i> → 13	0 → <i>h</i> → 9 -12 → <i>k</i> → 12 -12 → <i>l</i> → 12	0 → <i>h</i> → 11 0 → <i>k</i> → 16 -13 → <i>l</i> → 13
No. of standard reflections	3	3	3	3
Frequency of standard reflections (s)	3600	3600	3600	5000
Intensity decay (%)	0.2	0.4	0.2	0.2
Refinement				
Refinement on	<i>F</i>	<i>F</i>	<i>F</i>	<i>F</i>
<i>R</i>	0.0343	0.0953	0.0443	0.0628
<i>wR</i>	0.0385	0.1057	0.0570	0.0520
<i>S</i>	0.596	5.306	1.106	2.669
No. of reflections used in refinement	1499	1482	2818	1077
No. of parameters used	323	235	235	359
H-atom treatment	See text	See text	See text	See text
Weighting scheme	$w = a/[\sigma^2(F_o) + b(F_o)^2]$	$w = a/[\sigma^2(F_o) + b(F_o)^2]$	$w = a/[\sigma^2(F_o) + b(F_o)^2]$	$w = a/[\sigma^2(F_o) + b(F_o)^2]$
<i>a</i>	—	3.9769	1.0	2.7820
<i>b</i>	—	0.0022	0.0035	0.1 × 10 ⁻⁵
(Δ/σ) _{max}	0.12	0.005	0.090	0.016
$\Delta\rho_{\max}$ (e Å ⁻³)	0.15	0.37	0.32	0.32
$\Delta\rho_{\min}$ (e Å ⁻³)	-0.18	-0.50	-0.22	-0.29
Extinction method	None	None	None	None
Source of atomic scattering factors	<i>International Tables for X-ray Crystallography</i> (1974, Vol. IV)	<i>International Tables for X-ray Crystallography</i> (1974, Vol. IV)	<i>International Tables for X-ray Crystallography</i> (1974, Vol. IV)	<i>International Tables for X-ray Crystallography</i> (1974, Vol. IV)

3. Results and discussion

The refined atomic coordinates for non-H atoms are listed in Table 2. Corresponding *ORTEP*II diagrams (Johnson, 1976) for molecular structures along with atom labelling are shown in Fig. 1. Selected interatomic distances and bond angles are tabulated in Table 3. Some

dihedral angles relevant to our discussion appear in Table 4. Hydrogen-bond parameters and information on short atomic contacts for the structures reported here are provided in Table 5. Although for each of the compounds one half of the molecule is identical to the other, it was interesting to note that none depicted this symmetry in their crystal structures. In each structure, the asymmetric

Table 2. Fractional atomic coordinates and equivalent isotropic displacement parameters ($\text{\AA}^2$)

	$U_{eq} = (1/3)\sum_i \sum_j U_{ij} a_i^* a_j^* \mathbf{a}_i \cdot \mathbf{a}_j$			
	<i>x</i>	<i>y</i>	<i>z</i>	U_{eq}
(2)				
O2	0.2540 (8)	0.1050 (8)	−0.0583 (4)	0.091 (3)
O4	0.3630 (17)	−0.3627 (9)	0.1169 (6)	0.0216 (7)
O2'	−0.2563 (7)	−0.0562 (8)	0.5481 (4)	0.088 (3)
O4'	0.2324 (13)	−0.4138 (8)	0.5084 (5)	0.135 (5)
O10	−0.0229 (10)	0.3090 (6)	0.2264 (4)	0.090 (3)
O10'	−0.2807 (7)	0.2182 (7)	0.1725 (4)	0.078 (3)
N1	0.2020 (7)	0.0478 (7)	0.1430 (4)	0.050 (3)
N3	0.2490 (8)	−0.1224 (9)	0.0414 (4)	0.074 (3)
N1'	−0.1522 (7)	−0.0645 (8)	0.3587 (4)	0.056 (3)
N3'	0.0139 (11)	−0.2112 (8)	0.5137 (5)	0.072 (3)
C2	0.2310 (9)	0.0166 (11)	0.0374 (6)	0.055 (3)
C4	0.3004 (15)	−0.2461 (12)	0.1333 (7)	0.116 (5)
C5	0.2781 (12)	−0.2207 (10)	0.2533 (6)	0.077 (4)
C6	0.1673 (8)	−0.0647 (7)	0.2506 (5)	0.041 (3)
C7	0.4620 (14)	−0.2718 (11)	0.3006 (7)	0.105 (5)
C2'	−0.1407 (10)	−0.1076 (11)	0.4776 (6)	0.068 (4)
C4'	0.1262 (16)	−0.3079 (10)	0.4590 (7)	0.086 (5)
C5'	0.1073 (16)	−0.2750 (8)	0.3274 (6)	0.082 (4)
C6'	−0.0034 (10)	−0.1180 (8)	0.2786 (5)	0.051 (3)
C7'	0.0554 (29)	−0.3953 (13)	0.3110 (13)	0.176 (11)
C8	0.2489 (12)	0.1737 (13)	0.1460 (6)	0.078 (5)
C9	0.1079 (18)	0.3110 (11)	0.1242 (7)	0.097 (5)
C11	−0.1880 (21)	0.4123 (10)	0.2026 (8)	0.134 (7)
C11'	−0.2881 (16)	0.3646 (14)	0.1352 (8)	0.114 (6)
C9'	−0.3639 (10)	0.1788 (14)	0.2875 (7)	0.079 (5)
C8'	−0.3328 (10)	0.0184 (15)	0.3210 (6)	0.088 (6)
(3)				
O2	0.4478 (2)	0.2451 (2)	0.5512 (2)	0.0707 (7)
O4	0.0253 (2)	0.0331 (1)	0.6719 (1)	0.0515 (6)
O2'	0.7159 (2)	0.9135 (1)	0.8589 (1)	0.0609 (6)
O4'	0.9132 (2)	0.7111 (1)	0.5096 (1)	0.0491 (5)
O10	0.7203 (1)	0.3175 (1)	0.8418 (1)	0.0423 (5)
O10'	0.6589 (2)	0.5435 (1)	0.9354 (1)	0.0436 (5)
N1	0.4990 (2)	0.1086 (1)	0.6549 (1)	0.0406 (6)
N3	0.2403 (2)	0.1356 (1)	0.6135 (1)	0.0375 (5)
N1'	0.6343 (2)	0.6928 (1)	0.7715 (1)	0.0436 (3)
N3'	0.8150 (2)	0.8058 (2)	0.6848 (1)	0.0419 (6)
C2	0.4009 (2)	0.1694 (2)	0.6036 (2)	0.0415 (7)
C4	0.1716 (2)	0.0516 (2)	0.6697 (2)	0.0387 (6)
C5	0.2818 (2)	−0.0070 (2)	0.7240 (2)	0.0534 (8)
C6	0.4372 (2)	0.0237 (2)	0.7127 (2)	0.0532 (8)
C7	0.2197 (4)	−0.0995 (4)	0.7896 (5)	0.1033 (13)
C2'	0.7216 (2)	0.8109 (2)	0.7783 (2)	0.0446 (7)
C4'	0.8268 (2)	0.6974 (2)	0.5875 (2)	0.0390 (6)
C5'	0.7359 (2)	0.5761 (2)	0.5889 (2)	0.0407 (6)
C6'	0.6451 (2)	0.5792 (2)	0.6794 (2)	0.0428 (7)
C7'	0.7468 (3)	0.4530 (2)	0.4886 (2)	0.0554 (8)
C8	0.6729 (2)	0.1398 (2)	0.6526 (2)	0.0490 (7)
C9	0.7659 (2)	0.1976 (2)	0.7823 (2)	0.0463 (7)
C11	0.7860 (2)	0.3821 (2)	0.9623 (1)	0.0357 (6)
C12	0.8798 (2)	0.3348 (2)	1.0351 (2)	0.0489 (7)
C13	0.9377 (3)	0.4099 (2)	1.1578 (2)	0.0586 (9)
C13'	0.9023 (3)	0.5300 (2)	1.2091 (2)	0.0618 (9)
C12'	0.8085 (3)	0.5778 (2)	1.1374 (2)	0.0519 (8)
C11'	0.7508 (2)	0.5054 (2)	1.0144 (1)	0.0374 (6)
C9'	0.6268 (3)	0.6713 (2)	0.9798 (2)	0.0468 (7)
C8'	0.5348 (3)	0.6895 (2)	0.8712 (2)	0.0517 (8)
(4)				
O2	−0.0485 (8)	0.4779 (5)	0.6291 (5)	0.052 (3)
O4	0.2743 (7)	0.3438 (5)	0.4769 (5)	0.048 (3)
O2'	0.1426 (8)	−0.0508 (5)	0.7579 (5)	0.055 (3)
O4'	0.2106 (8)	0.0893 (5)	0.4527 (5)	0.061 (3)
O10	0.1584 (6)	0.3741 (4)	0.9309 (4)	0.027 (2)
O10'	0.1369 (7)	0.1907 (5)	0.9401 (5)	0.049 (3)
N1	0.0368 (7)	0.3474 (5)	0.7069 (5)	0.024 (3)
N3	0.1078 (9)	0.4020 (7)	0.5534 (7)	0.039 (4)
N1'	0.0678 (8)	0.0965 (9)	0.7378 (6)	0.033 (3)
N3'	0.1681 (10)	0.0234 (7)	0.6042 (7)	0.045 (4)
C2	0.0246 (10)	0.4128 (7)	0.6316 (7)	0.030 (4)

Table 2 (cont.)

	<i>x</i>	<i>y</i>	<i>z</i>	U_{eq}
C4	0.2010 (10)	0.3375 (7)	0.5449 (7)	0.030 (4)
C5	0.1955 (9)	0.2551 (6)	0.6141 (6)	0.024 (3)
C6	0.1170 (10)	0.2648 (7)	0.7103 (7)	0.027 (4)
C7	0.3345 (11)	0.2225 (9)	0.6525 (9)	0.040 (4)
C2'	0.1254 (11)	0.0198 (7)	0.7063 (8)	0.039 (4)
C4'	0.1628 (10)	0.0943 (8)	0.5336 (8)	0.042 (4)
C5'	0.0958 (9)	0.1785 (6)	0.5653 (6)	0.024 (3)
C6'	0.0386 (10)	0.1783 (7)	0.6725 (8)	0.029 (4)
C7'	−0.0063 (12)	0.2083 (10)	0.4733 (8)	0.038 (4)
C8	−0.0365 (10)	0.3601 (8)	0.7966 (7)	0.030 (4)
C9	0.0365 (10)	0.4149 (7)	0.8866 (8)	0.029 (4)
C11	0.1588 (9)	0.3313 (7)	1.0299 (7)	0.032 (4)
C12	0.1720 (9)	0.3862 (8)	1.1210 (8)	0.035 (4)
C13	0.1753 (10)	0.3434 (9)	1.2207 (8)	0.040 (4)
C13'	0.1759 (10)	0.2491 (8)	1.2264 (7)	0.040 (5)
C12'	0.1650 (10)	0.1943 (8)	1.1357 (7)	0.037 (4)
C11'	0.1539 (9)	0.2364 (7)	1.0367 (7)	0.032 (4)
C9'	0.0939 (14)	0.0975 (8)	0.9374 (8)	0.045 (5)
C8'	0.0025 (11)	0.0852 (8)	0.8330 (8)	0.032 (4)

unit was formed by a complete molecule. The thymines were in *keto* form in all compounds and the pyrimidine rings were planar in (1) and (3). The pyrimidine bond lengths and angles were comparable to that of thymine in its crystal structure (Gerdil, 1961). The C4—C5 bond, although single, acquired partial double-bond character due to the neighbouring double bonds and is still the longest intraring bond in the pyrimidine, while the C5—C6 double bond is the shortest. Others have intermediate lengths. Out of the six intraring bond angles, those at positions 1, 3 and 6 were wider than the rest. In the crystal structure of (1) the thymine rings were stacked one above the other in an infinite mode of packing and could perhaps be compared with ssDNA (Fig. 2). The mutual orientation of the thymines connected by the spacer chain was such that the projection of one onto the plane of the other showed a rotation of 53° about an axis perpendicular to the plane. This gave a helical twist to the polyoxyethylene chain. The value of this rotation angle was slightly higher compared with the angle of rotation between adjacent thymines in B-DNA, which was only 36°. The two thymine planes were almost parallel, making an angle of 10°. In (3), where the spacer chain contained a rigid catechol group, the thymines were disposed almost perpendicular to each other (74°), as well as with the planar catechol. Perhaps the presence of this rigid group was a deterrent for the spacer chain in (3) to twist itself in order to bring the thymines in a stacked position.

The photoproducts of (1) and (3) are cyclobutane-linked dimers of thymine and were established to be the *trans-syn* isomers by determination of their crystal structures. Belonging to the centrosymmetric lattice, the crystals contained in them are both enantiomers D and L of the *trans-syn* dimer. The formation of the photodimer was through interpyrimidine bonds between C5 and C6 atoms, respectively, of the two thymine bases. Thus, when the geometries of the compounds were compared before and after photodimerization, the main difference,

and an obvious corollary was in the stereochemistry of C5 and C6. Both atoms have changed from a planar to a tetrahedral configuration. Other features observed on dimerization were the stretching of bonds, the most notable being N1—C6, C4—C5 and C5—C6, along

with their pairs in the twin pyrimidine ring, and the alteration of some bond angles. There was substantial reduction in the bond angles C5—C6—N1 and C4—C5—C6 and their pairs when (1) was converted to (2). However, such an effect was confined to the C5—C6—N1 pair when (3) changed to its photoproduct (4). All these changes could be attributed to the saturation of the C5—C6 bond as a result of photodimerization. Again, the widest intraring pyrimidine angle, that at N3, has further widened on photodimerization; this effect being more pronounced in the case of (4). While the intraring angles of cyclobutane assumed nearly equal values in both (2) and (4), the other bond angles around C5 and C6 atoms have a similar situation only in (4) whereas their magnitudes varied considerably in (2) (Table 3). In both the cases the angle pairs C5—C4—O4 displayed contraction on photodimerization, hinting at a plausible weak interaction between the fourth atom pairs C7 and O4.

On examining the stereochemistry of the cyclobutane rings in (2) and (4) the interpyrimidine bonds C5—C5' and C6—C6' were found to have an asymmetry in their lengths, with the former being longer than the latter. This could well be due to the bulk methyl groups at the C5 position. Similar asymmetry has also been reported in the crystal structures of other photodimers with bulky groups at such positions (Leonard, Golankiewicz, McGredie, Johnson & Paul, 1969; Birnbaum, 1972; Jankun, Maluszynska, Kaluski & Wiesz, 1977; Koziol & Rajchel, 1982). All the intraring cyclobutane angles were around 90°. It is a consequence of the asymmetry in the two

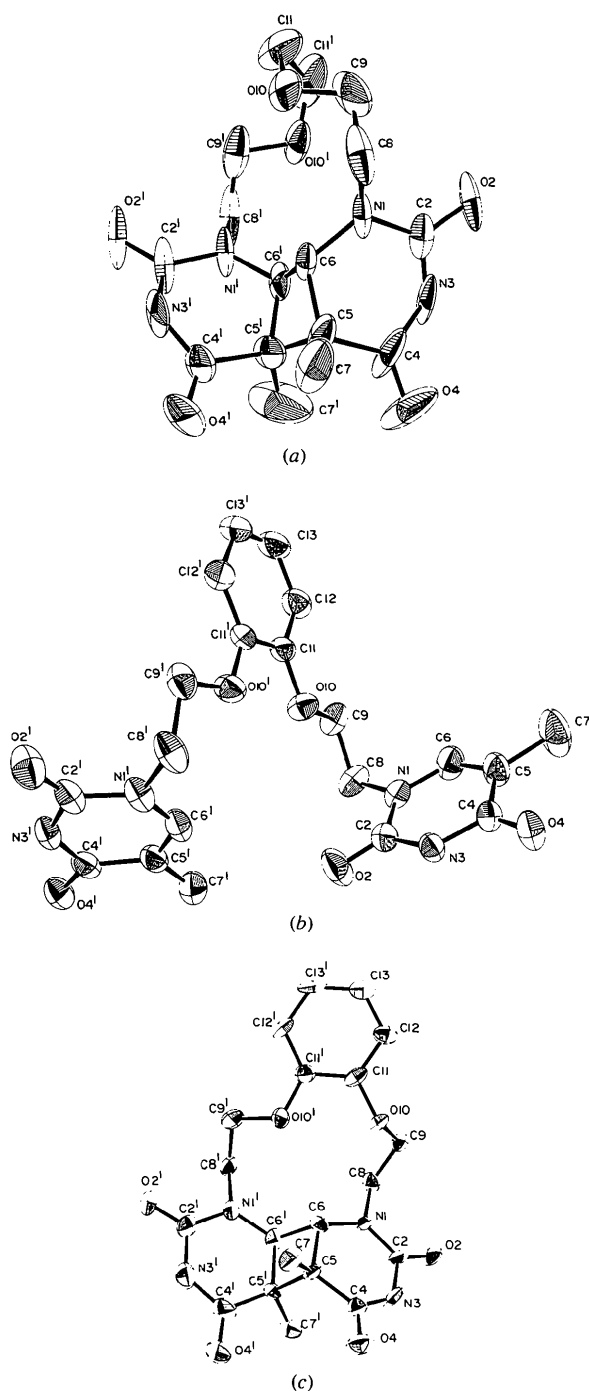


Fig. 1. Ellipsoid plots (50% displacement probability) showing the structures and atom labelling for (a) (2), (b) (3) and (c) (4).

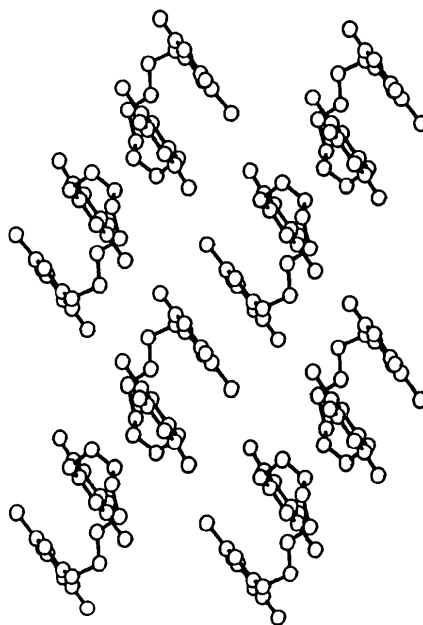


Fig. 2. Packing of the molecules in the crystal structure of (1), showing resemblance to ssDNA. View along a^* .

Table 3. *Selected interatomic distances (Å) and bond angles (°)*

	(2)	(3)	(4)
N1—C2	1.363 (10)	1.368 (2)	1.343 (11)
N1'—C2'	1.372 (8)	1.368 (2)	1.353 (12)
C2—N3	1.349 (15)	1.385 (2)	1.409 (12)
C2'—N3'	1.385 (10)	1.375 (2)	1.426 (12)
N3—C4	1.391 (10)	1.370 (2)	1.363 (12)
N3'—C4'	1.345 (11)	1.378 (2)	1.363 (13)
C4—C5	1.514 (13)	1.441 (3)	1.493 (12)
C4'—C5'	1.542 (12)	1.439 (3)	1.492 (13)
C5—C6	1.539 (11)	1.342 (3)	1.563 (12)
C5'—C6'	1.542 (10)	1.336 (3)	1.556 (12)
C6—N1	1.465 (7)	1.372 (3)	1.461 (11)
C6'—N1'	1.442 (8)	1.382 (2)	1.460 (12)
C2—O2	1.230 (9)	1.211 (2)	1.212 (10)
C2'—O2'	1.222 (9)	1.226 (2)	1.221 (11)
C4—O4	1.188 (15)	1.237 (2)	1.228 (11)
C4'—O4'	1.203 (11)	1.230 (2)	1.202 (11)
C5—C5'	1.600 (15)		1.587 (12)
C6—C6'	1.537 (12)		1.539 (12)
C7...C7'	3.75 (3)		3.93 (2)
C6—N1—C2	118.3 (7)	120.4 (2)	127.0 (8)
C6'—N1'—C2'	120.3 (5)	121.6 (2)	124.9 (8)
N1—C2—N3	116.8 (6)	114.7 (2)	114.5 (9)
N1'—C2'—N3'	115.8 (6)	114.2 (2)	116.3 (9)
C2—N3—C4	127.7 (8)	127.6 (2)	129.9 (9)
C2'—N3'—C4'	128.6 (8)	127.4 (2)	129.4 (10)
N3—C4—C5	115.5 (9)	115.0 (2)	115.0 (10)
N3'—C4'—C5'	115.1 (7)	115.2 (2)	115.2 (9)
C4—C5—C6	110.5 (6)	117.6 (2)	116.9 (8)
C4'—C5'—C6'	111.4 (7)	118.2 (2)	118.8 (8)
C5—C6—N1	115.7 (5)	124.8 (2)	114.0 (8)
C5'—C6'—N1'	115.4 (5)	123.3 (2)	115.0 (8)
N1—C2—O2	123.6 (10)	123.7 (2)	124.8 (9)
N1'—C2'—O2'	122.9 (7)	123.5 (2)	125.2 (10)
N3—C2—O2	119.4 (8)	121.6 (2)	120.7 (9)
N3'—C2'—O2'	121.3 (6)	122.3 (2)	118.5 (10)
N3—C4—O4	121.7 (9)	119.6 (2)	121.2 (9)
N3'—C4'—O4'	123.1 (8)	118.8 (2)	122.0 (10)
C5—C4—O4	122.8 (8)	125.4 (2)	123.4 (10)
C5'—C4'—O4'	121.8 (8)	126.0 (2)	122.8 (10)
C5'—C5—C6	87.8 (6)		89.3 (7)
C5—C5'—C6'	86.4 (6)		88.1 (6)
C5—C6—C6'	88.7 (6)		89.6 (7)
C5'—C6'—C6	90.0 (6)		91.3 (7)

interpyrimidine bond lengths that the angles opposite C5—C5' were slightly wider than the angles included by this bond. Cyclobutyl rings were twisted with intraring dihedral angles of 20 and 10° for (2) and (4), respectively. This twist, for identical enantiomers, is opposite in a sense for (2) and (4), with the effect that the methyl group attached at C5 is brought closer in (2) while placed farther away in (4). This is also reflected in the interatomic distances of methyl carbons in (2) and (4) (Table 3). On the surface this was surprising, because the mutual disposition of the methyl groups in (2) only increased the strain on the cyclobutane ring. A possible explanation for such an anomalous folding of the cyclobutane ring in this structure could be the influence of the spacer chain. What turned out to be really interesting was that a rigid group such as catechol, present in the spacer chain, only eased the strain on the *trans-syn* thymine photodimer stereochemistry. Consolidation of this argument needs more such examples and any biological implications of this phenomena has to be

explored further. The relative twists between the two pyrimidine rings in the photodimer defined as the average deviation of two torsional angles, C4—C5—C5'—C4' and N1—C6—C6'—N1', from the undistorted value of 120°, were 29.3 and 12.8° for (2) and (4), respectively. The twists defined here, taken together with the twists of the cyclobutane rings referred to earlier, were also opposite in a sense for (2) and (4). These relative twists between the pyrimidine rings influence the interatomic distances of the methyl group on one ring to the O4 atom on the other ring. The closeness of these atoms in (2) compared with (4), along with a resultant interaction between them (Table 5), may be compensating to some extent the strain originated from the mutual closeness of the methyl groups in (2). The geometrical strains at the site of photodimerization for nucleotides and model compounds were also reported for several other structures (Karle, 1976). Nevertheless, this was the first time a comprehensive study of the stereochemistry of the photoproducts in comparison to that of their respective parent bisthymines was accomplished through the determination of three-dimensional structures of both the parent compound and its photodimer.

The planar geometry of the pyrimidine rings was altered on photodimerization. This has been most significant in the case of (2). Here the pyrimidine rings assume a boat conformation, with N3 and C6 located at bowsprit positions (Figs. 3a and b). The pyrimidines also assumed the boat conformation in (4). However, the difference is that here at the bowsprit positions the atoms were C2 and C5 in one ring, but N1' and C4' in the other (Figs. 3c and d). Distortion of the planarity of pyrimidines could also be due to saturation at C5—C6 bonds, as has been reported for other cyclobutane dimers of thymines (Cadet, Voituriez, Hruska & Grand, 1985; Hruska, Voituriez, Grand & Cadet, 1986). The spacer backbone chain of bisthymines has also undergone drastic changes in conformation after photodimerization (Table 4). This could, in all probability, turn out to be one aspect which may frustrate attempts to carry out X-ray investigations on the same crystals of bisthymines before and after dimerization, since these large displacements of the spacer atoms could easily rupture the single crystals.

Crystal densities calculated for pre- and post-dimerized compounds showed an apparent increase in close packing for photodimers. This could be attributed more to the shrinking of the molecules on dimerization rather than to any tighter packing, as no indications were available for any additional intermolecular short contacts being propped up in the structures of photoproducts. From the magnitudes of the bond lengths of C2—O2 and C4—O4 the *keto* form of pyrimidines could be assumed. Thus, the only potential proton donor N3, *sans* N1 which is involved in the linkage, in each of the pyrimidine rings was forming a hydrogen bond with either a pyrimidine

oxygen or a spacer chain oxygen of a symmetry-related molecule (Table 5). Few noticeable short carbon–oxygen contacts, including those already discussed, mostly intramolecular, were expected to provide extra stability to individual molecules as well as to the overall packing in their respective crystals, Table 5.

trans-syn Thymine dimers which were the main photoproducts of ssDNA on irradiation of UV light have been identified as more lethal to the cell, because their repair by photolyases is less efficient. Detailed study of the stereochemical interactions of *trans-syn* thymine dimers with the photolyases has been hampered by the sparse yield of this isomer in any particular experiment. In this context we have shown that the *trans-syn* thymine is the exclusive product upon UV irradiation

of hitherto unknown bisthymines linked through novel polyoxyethylene spacer chains. The study reported here would help in the design and exclusive synthesis of the rare *trans-syn* photodimers and provide a tool for understanding and modelling the repair mechanisms of DNA or even in designing DNA photolyases targeted at *trans-syn* thymine photodimers in particular.

The authors wish to express their profound gratitude to Professor John Barnabas for his encouragement and for sparing us the computing facilities. BPG thanks UGC-New Delhi for financial support. KNG acknowledges Jawaharlal Nehru Centre, Bangalore, of which he is a Senior Honorary Fellow. NCL Communication No. 6127.

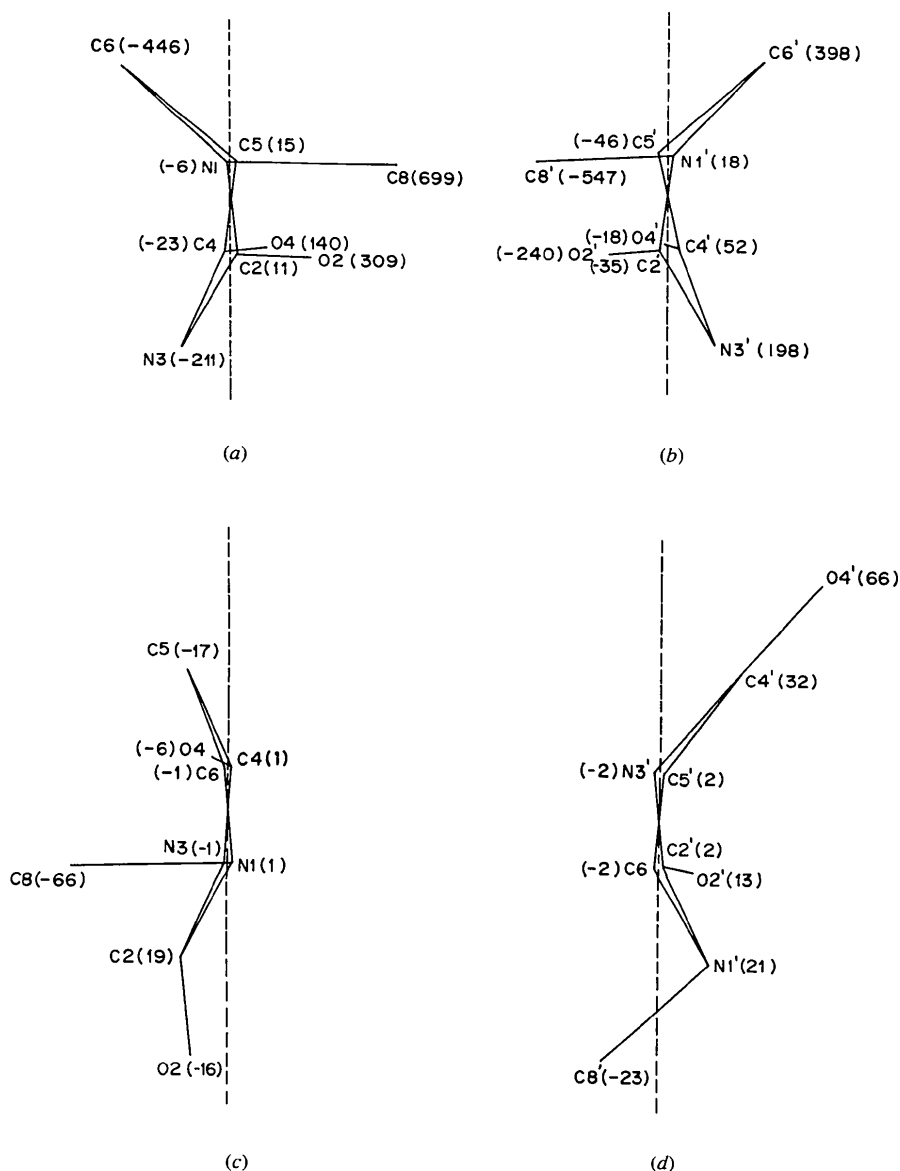


Fig. 3. Projection of pyrimidine atoms (a) and (b) in (2), and (c) and (d) in (4), perpendicular to a mean plane (indicated by dotted lines) calculated with respect to the four atoms having the lowest deviations from a least-squares plane for all atoms in the respective thymine rings. The numbers in parentheses are the deviations of respective atoms from the mean plane in units of 10^{-3} Å.

Table 4. *Pyrimidine and the spacer backbone chain dihedral angles in the structure to show the large shifts in their magnitudes accompanied by conformational changes on dimerization*

	(1)	(2)	(3)	(4)
C6—N1—C2—N3	3.5 (4)	—8.6 (10)	—1.2 (2)	—5.6 (13)
C6'—N1'—C2'—N3'	3.5 (4)	—4.6 (11)	1.7 (3)	4.6 (14)
N1—C2—N3—C4	1.0 (4)	—23.3 (13)	1.2 (3)	—2.5 (15)
N1'—C2'—N3'—C4'	—4.3 (5)	—23.4 (13)	1.0 (3)	2.0 (16)
C2—N3—C4—C5	—5.1 (4)	20.5 (13)	—0.1 (3)	14.8 (15)
C2'—N3'—C4'—C5'	1.9 (5)	16.6 (14)	—3.1 (3)	—3.9 (16)
N3—C4—C5—C6	4.7 (4)	11.3 (11)	—1.0 (3)	—18.5 (12)
N3'—C4'—C5'—C6'	1.2 (4)	14.4 (11)	2.5 (3)	—0.4 (13)
C4—C5—C6—N1	—0.8 (4)	—38.5 (9)	1.0 (3)	12.1 (11)
C4'—C5'—C6'—N1'	—1.8 (4)	—37.6 (10)	0.0 (3)	5.7 (12)
C5—C6—N1—C2	—3.6 (4)	38.9 (9)	0.2 (3)	0.2 (13)
C5'—C6'—N1'—C2'	—0.6 (4)	34.1 (10)	—2.2 (3)	—8.1 (13)
C2—N1—C8—C9	81.2 (3)	89.8 (10)	118.4 (2)	—87.9 (10)
C2'—N1'—C8'—C9'	—96.5 (3)	94.6 (10)	—88.7 (2)	85.9 (11)
N1—C8—C9—O10	64.8 (3)	74.0 (11)	—63.3 (2)	—59.5 (11)
N1'—C8'—C9'—O10'	—65.6 (4)	80.2 (10)	—67.1 (2)	64.6 (11)
C8—C9—O10—C11	168.2 (3)	—161.5 (9)	177.1 (2)	—104.7 (9)
C8'—C9'—O10'—C11'	—84.5 (4)	—175.0 (9)	175.9 (2)	143.2 (9)
C9—O10—C11—C11'	—168.1 (3)	73.8 (11)	174.2 (2)	103.6 (10)
C9'—O10'—C11'—C11	131.8 (3)	64.4 (12)	—176.1 (2)	—162.4 (9)
O10—C11—C11'—O10'	—83.5 (4)	40.7 (13)	—1.5 (2)	—5.5 (13)

Table 5. *Hydrogen-bond geometries*

<i>D</i> —H... <i>A</i>	<i>D</i> ... <i>A</i>	<i>D</i> —H... <i>A</i>
(2)		
N3—N3H...O10 ⁱ	2.966 (9)	17
N3'—N3'H...O10 ⁱⁱ	2.976 (8)	19
C7—C7H1...O4 ⁱⁱⁱ	3.006 (12)	53
C7—C7H2...O4 ⁱⁱⁱ	3.006 (12)	55
C7'—C7'H3...O4 ⁱⁱⁱ	3.023 (21)	40
(3)		
N3—N3H...O4 ^{iv}	2.930 (2)	12 (1)
N3'—N3'H...O4 ^v	2.888 (2)	5 (2)
C6'—C6'H...O10 ^{vi}	3.073 (3)	58 (1)
C8—C8H1...O2 ⁱⁱⁱ	2.775 (3)	52 (1)
C9—C9H1...O4 ^{vi}	3.191 (3)	40 (2)
C8'—C8'H1...O2 ⁱⁱⁱ	2.744 (3)	57 (2)
(4)		
N3—N3H...O2 ⁱⁱ	2.897 (11)	16 (5)
N3'—N3'H...O10 ^{vii}	2.895 (12)	8 (6)
C6—C6H...O10 ⁱⁱⁱ	3.194 (11)	41 (5)
C6—C6H...O10 ⁱⁱⁱ	3.085 (11)	37 (5)

Symmetry codes: (i) $-x, -y, -z$; (ii) $-x, 1 - y, 1 - z$; (iii) x, y, z ; (iv) $1 - x, 1 - y, 1 - z$; (v) $x + 1, y + 1, z$; (vi) $x + 1, y, z$; (vii) $\frac{1}{2} - x, y - \frac{1}{2}, z + \frac{3}{2}$.

References

- Bideau, P. J., Courseille, C. & Bravic, G. (1982). *Acta Cryst.* **B38**, 2075–2077.
- Birnbaum, G. I. (1972). *Acta Cryst.* **B28**, 1248–1254.
- Cadet, J. & Vigny, P. (1990). *Bioorganic Photochemistry*, edited by H. Morrison, Vol. 1, pp. 1–272. New York: John Wiley & Sons.
- Cadet, J., Voituriez, L., Hruska, F. E. & Grand, A. (1985). *Biopolymers*, **24**, 897–903.
- Fisher, G. J. & Johns, H. E. (1976). *Photochemistry and Photobiology of Nucleic Acids*, edited by S. Y. Wang, Vol. 1, pp. 225–294. New York: Academic Press.
- Gangamani, B. P., Suresh, C. G. & Ganesh, K. N. (1994). *Chem. Commn.* pp. 2275–2276.
- Gerdil, R. (1961). *Acta Cryst.* **14**, 333–344.
- Hruska, F. E., Voituriez, L., Grand, A. & Cadet, J. (1986). *Biopolymers*, **25**, 1399–1417.
- Jankun, E. S., Maluszynska, H., Kaluski, Z. & Wiesz, K. G. (1977). *Acta Cryst.* **B33**, 1624–1627.
- Johnson, C. K. (1976). *ORTEP*. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.
- Karle, I. L. (1976). *Photochemistry and Photobiology of Nucleic Acids*, edited by S. Y. Wang, Vol. 1, pp. 483–519. New York: Academic Press.
- Kim, S.-T. & Sancar, A. (1993). *Photochem. Photobiol.* **57**, 895–904.
- Kim, S.-T., Malhotra, K., Smith, C. A., Taylor, J.-S. & Sancar, A. (1993). *Biochemistry*, **32**, 7065–7068.
- Kozioł, A. E. & Rajchel, A. (1982). *Acta Cryst.* **B38**, 999–1001.
- Leonard, N. J., Golankiewicz, K., McGredie, R. S., Johnson, S. M. & Paul, I. C. (1969). *J. Am. Chem. Soc.* **91**, 5855–5862.
- Sancar, A. & Sancar, G. B. (1988). *Ann. Rev. Biochem.* **57**, 29–67.
- Sancar, A. & Tang, M.-S. (1993). *Photochem. Photobiol.* **57**, 905–921.
- Setlow, R. B. (1978). *Nature (London)*, **271**, 713–717.
- Sheldrick, G. M. (1976). *SHELX76. Program for Crystal Structure Determination*. University of Cambridge, England.
- Sheldrick, G. M. (1985). *SHELXS86. Program for the Solution of Crystal Structures*. University of Göttingen, Germany.
- Taylor, J.-S. (1990). *J. Chem. Educ.* **67**, 835–841.
- Taylor, J.-S. (1994). *Acc. Chem. Res.* **27**, 76–82.