Acta Cryst. (1992). B48, 714–719

Crystal and Molecular Structure of d(GTCTAGAC)

BY ALDO R. CERVI

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England

BEATRICE LANGLOIS D'ESTAINTOT

Laboratoire de Cristallographie, Université de Bordeaux, Talence, France

AND WILLIAM N. HUNTER*

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England

(*Received* 25 *November* 1991; *accepted* 10 *April* 1992)

Abstract

d(GTCTAGAC), $C_{78}H_{92}N_{30}O_{46}P_7.3H_2O$, M_r for DNA only = 2403, tetragonal, $P4_32_12$, a = 42.56 (11), c = 24.61 (11) Å, V = 44577 Å^3, Z = 8, λ(Cu $K\alpha$) = 1.5418 Å, μ(Cu $K\alpha$) = 9.7 cm^{-1}, T = 295 K, R = 0.211 for 801 unique reflections with $F > \sigma(F)$ between 6 and 2.5 Å resolution. The asymmetric unit consists of a single strand of oligonucleotide and three well-defined solvent molecules. The structure is quasi-isomorphous with d(CTCTAGAG) [Hunter, Langlois d'Estaintot & Kennard (1989). *Biochemistry*, **28**, 2444–2451] and was solved by molecular replacement. Restrained least-squares methods interspersed with computer-graphics manipulation were used to refine the structure. Two strands, related by a crystallographic dyad axis, coil about each other to form a right-handed helix. The mean helix rotation is 32° resulting in 11.3 base pairs per turn. The average rise per base pair is 3.3 Å, most of the furanose rings adopt a C3′-*endo* conformation and the duplex shows a shallow minor groove and a deep major groove. These global features suggest classification of d(GTCTAGAC) as A-form DNA and tend to mask significant variations in conformational parameters at the base-pair level. In particular, the central TpA (=TpA) step displays extensive interstrand purine–purine overlap and an unusual sugar-phosphate backbone conformation.

1. Introduction

The DNA duplex adopts a range of conformations in the crystalline state depending not only on the base sequence (Saenger, 1984) but also on the degree of hydration, the molecular packing and temperature (Shakked, Guerstein-Guzikevich, Eisenstein, Frolow & Rabinovich, 1989). This highlights the ability of DNA to undergo what may be important conformational changes resulting from sometimes modest variations in environment. Conformational variability in DNA offers potential sequence-dependent structural diversity that can be utilized by proteins in site-specific interactions. Considerable effort has been devoted towards studying the principles of nucleotide and nucleic acid structure and the determinants of DNA conformation (Saenger, 1984). In this context we report the structure determination of the octanucleotide d(GTCTAGAC).

* Author to whom correspondence should be addressed.

0108-7681/92/050714-06$06.00

2. Experimental

The octanucleotide of sequence d(GTCTAGAC) was synthesized by phosphoramidite methodology (McBride & Caruthers, 1983) on an Applied Biosystems 381A synthesizer and then purified by ion exchange and reverse-phase high-pressure liquid chromatography. Crystals were grown at 295 K by vapour diffusion from 20 µl droplets sitting in Corning glass depression paltes (McPherson, 1990). Small, well-formed tetragonal bipyramids grew in a matter of weeks from solutions of 6 mg ml^{-1} octamer, 60 mM sodium cacodylate buffer at pH 6.8, 25 mM magnesium chloride, 6 mM spermine tetrachloride and 10% v/v 2-methyl-2,4-pentanediol (MPD) equilibrated against a 50% MPD reservoir. A single colourless crystal (0.2 × 0.2 × 0.1 mm) was sealed in a glass capillary, with a drop of mother liquor to maintain a humid environment, for characterization and use in data collection. Data out to a $2\theta_{max}$ of 46° were collected using a Rigaku AFC-5 diffractometer on an RU200 rotating anode with graphite-monochromated Cu $K\alpha$ radiation, λ = 1.5418 Å, 50 kV, 160 mA settings. The crystal-to-detector distance was set at 400 mm and a continuously evacuated beam path was used to reduce absorption by air.

Lattice parameters were obtained from a least-squares fit of 25 reflections within the range $15 \le \theta \le 30°$. Based on the Laue group 4/*mmm*, the systematic

absences of $00l$: $l = 4n$ and $h00$: $h = 2n$ the space group was identified as $P4_12_12$ (No. 92) or enantiomorph $P4_32_12$ (No. 96). Three standard reflections (210, 220, 221) were recorded every 150 measurements during data collection. A small linear correction factor (0.7%) was applied on the basis of standard monitoring. A total of 1861 reflections were recorded of which 1757 were unique, $R_{int} = 0.12$. Measurements were made in the $\omega/2\theta$ mode, with a scan width of $(1.42 + 0.30\tan\theta)^\circ$ at a speed of $8^\circ\ min^{-1}$ in ω. Reflections with $I < 7\sigma(I)$ were scanned in triplicate to improve the counting statistics. Intensities were corrected for Lorentz–polarization factors and absorption with transmission factors ranging from 0.68 to 1.00 (North, Phillips & Mathews, 1968), and secondary extinction (0.16541×10^{-5}). Data collection and processing was carried out with software provided by Molecular Structure Corporation, Texas, USA.

Although data to a resolution of 2.0 Å were recorded we noted that the number of reliable measurements beyond 2.5 Å was negligible. This is a consequence of using a small crystal for data collection. We did attempt to improve the resolution by subsequently using larger crystals, up to $0.5 \times 0.5 \times 0.3$ mm in one case. However, it was observed that the larger specimens were not as well defined in appearance as the small samples and although the scattering power was improved these crystals displayed a much increased mosaic spread from a value of 0.34° for the average full width at half height from the small crystals to values in excess of 0.8° for the larger specimens. We decided not to use the larger crystals and proceeded with the 2.5 Å data set. This observation concerning the improved perfection, as measured by mosaic spread, of the smaller crystals was also made with d(CTCTAGAG) (W. N. Hunter, unpublished results).

The initial model for the refinement, selected on the basis of a similar unit cell and symmetry, was that of d(CTCTAGAG) (Hunter, Langlois d'Estaintot & Kennard, 1989) in space group $P4_32_12$. Subsequent analysis showed that this choice was correct and the model was a satisfactory starting point. Difference density maps with coefficients $(2F_o - F_c)$ and $(F_o - F_c)$ at 2.7 Å resolution were calculated with phases derived from the initial model and displayed on a Silicon Graphics IRIS 4D/240GTX computer-graphics system using *FRODO* (Jones, 1978) as modified for a UNIX workstation (Jones & Cambilleau, 1989). Sequence changes were introduced and the model coordinates manipulated to fit the electron density. The R factor of this model was 0.481. Refinement continued with the *CCP*4 package (*CCP*4, 1979) and the restrained least-squares method of Hendrickson & Konnert (1981) using *NUCLSQ* (Westhof, Dumas & Moras, 1985). Atomic scattering factors were taken from Cromer & Waber (1974). Negligible restraints were employed on sugar conformations and none on sugar-phosphate torsion angles. Restrained isotropic temperature factors were refined. A search for solvent molecules was initiated as soon as the resolution was extended to the limit of the reliable data, 2.5 Å as discussed above. The refinement was terminated after the location of only three solvent molecules. This number of solvent molecules is very much less than has been located in other oligonucleotide structures (Kennard & Hunter, 1989) and is both a consequence of limited resolution and the very strict selection criteria we imposed on identifying solvent positions.

After 147 cycles of *NUCLSQ* refinement, the final R value was 0.211 for the 801 unique reflections with $F > \sigma(F)$ in the range 6–2.5 Å. This corresponds to 77% of the theoretical number of reflections within this resolution range. The average thermal parameters (Å^2) for phosphate, furanose and base groups are approximately 19, 15 and 10 respectively. The average thermal parameter for the three solvent molecules is 32 Å^2. The correctness of the structure is indicated by an excellent fit of atomic coordinates to electron density. H atoms were not included in the model. The largest discrepancies in the final difference maps were located close to the phosphorus atoms. Statistics from the final cycle of refinement are presented in Table 1* and illustrate that the final model has good geometry with only 12 distances deviating from ideality by greater than 2σ. Diagrams were obtained using *PLUTO* (Motherwell & Clegg, 1978) as modified by Dodson and Evans for the *CCP*4 package (*CCP*4, 1979).

3. Discussion

The asymmetric unit in the crystal structure of d(GTCTAGAC) is a single strand. Two strands coil about each other to form a right-handed double helix with Watson–Crick base pairing (Fig. 1). A crystallographic dyad which traverses the unit-cell diagonal is coincident with the duplex dyad. Nucleotides are labelled G1 to C8 on strand 1, the asymmetric unit, and G9 to C16 on strand 2 in a 5′ to 3′ direction. The crystal packing is similar to that encountered in other A-type structures (Kennard & Hunter, 1989) with the terminal base pairs lying in the minor groove of adjacent symmetry-related duplexes. This assembly is promoted by extensive van der Waals interactions.

* Lists of structure factors, isotropic thermal parameters and atomic coordinates have been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 55119 (18 pp.). Copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England. [CIF reference: LI0119]

Table 1. *Statistics from the final cycle of restrained least-squares refinement of* d(GTCTAGAC)

Restraint groups and parameters	Standard deviations	R.m.s.	$\Sigma w\Delta^2$*
Distance restraints (Å)			
Sugar and base bond lengths	0.011	0.025	
Sugar and base bond angles	0.023	0.035	0.25×10^2
Phosphate bond lengths	0.031	0.035	
Phosphate bond angles	0.044	0.055	
12 distances deviate from ideality by more than 2σ			
Planar restraints (Å)			
Deviation from the plane	0.017	0.030	0.26×10^2
Non-bonded contacts (Å)			
Single torsion	0.154	0.063	
Multiple torsion	0.247	0.063	0.12×10^2
Hydrogen bond		0.063	
Isotropic thermal parameters (Å^2)			
Sugar and base bond	2.6	5.00	
Sugar and base angle	3.2	5.00	
Phosphate bond	5.5	7.50	0.21×10^3
Phosphate angle	4.5	7.50	
Restraint on magnitude of positional shift	0.35		
Unrestrained chiral volumes		0.10	

* Where $w = 1/\sigma^2$ and Δ = deviation from ideal value.

Fig. 2 shows the four unique base-pair steps in the duplex. Step 1 GpT (=ApC) displays extensive overlap between the six-membered rings of the adjacent G1 purine and the T2 pyrimidine residues. The N4 amino group of C16 also shows an overlapping interaction with the six-membered ring of A15. No cross-strand interaction is observed at this step. Steps 2 TpC (=GpA) and 3 CpT (=ApG) are pyrimidine–pyrimidine (=purine–purine) steps. The adjacent purines all interact with the five-membered ring of the other. The overlap between adjacent pyrimidines is generally reduced although the methyl groups of the thymines do appear to interact strongly with the cytosines in the CpT (=ApG) step. Step 3 also shows a small overlap involving atoms of bases on opposing strands, these being residues T4 and G14. This cross-strand interaction occurs on the minor-groove side of the duplex and involves the amino N2 of the guanine and the keto O2 of the thymine.

The central TpA (=TpA) step (step 4) displays a marked degree of cross-strand purine–purine overlap involving the six-membered rings of each adenine. The overlap of adjacent bases is poor. This pyrimidine–purine step looks very similar to the TpA (=TpA) steps observed in d(GGTATACC) (Shakked, Rabinovich, Kennard, Cruse, Salisbury & Viswamitra, 1983) and r(GCG)d(TATACGC) (Wang, Fujii, van Boom, van der Marel, van Boeckel & Rich, 1982) and also to the central CpG (=CpG) steps found in d(CCCCGGGG) (Haran, Shakked, Wang & Rich, 1987), d(GCCCGGGC) (Heinemann, Lauble, Frank & Blocker, 1987), d(GGGCGCCC) (Rabinovich, Haran, Eisenstein & Shakked, 1988), d(GGCCGGCC) (Wang, Fujii, van Boom & Rich, 1982) and d(GTACGTAC) (Courseille, Dautant, Hospital, Langlois d'Estaintot, Precigoux, Molko & Teoule, 1990).

Geometrical parameters relevant to base–base interactions, calculated with the *NEWHELIX* program (Dickerson, 1989) and brought into agreement with the Cambridge Conventions (1989) are presented in Table 2. For comparative purposes values for d(CTCTAGAG) are included. In d(GTCTAGAC), step 3 (C3–G14) displays the lowest roll value of 1° which indicates that the relevant base pairs are nearly parallel to each other. Step 2 displays the highest roll value of 13°. The average value for the slide is −2.9 Å towards strand 1 of the second base pair compared to the first base pair. The highest value is −3.2 Å at the TpA (=TpA) step, which is greater than the −1.7 Å found for the TpA (=TpA)

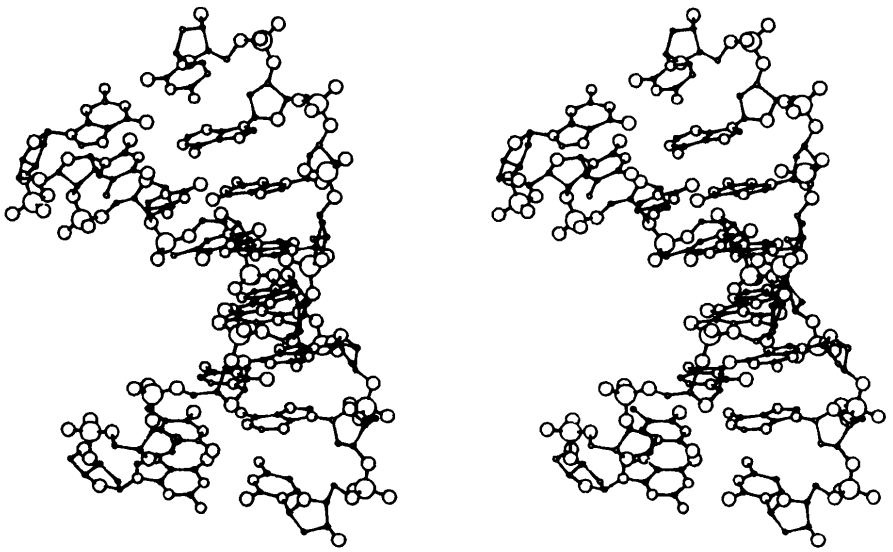


Fig. 1. A stereoview of the d(GTCTAGAC) octamer duplex showing the major (deep) and minor (shallow) grooves. Residue A8 of the asymmetric unit is at the top left-hand corner. The symmetry operation $y, x, -z$ generates the second strand of the duplex. Atoms are represented as spheres of decreasing radii in the order P, O, N, C.

step of d(CTCTAGAG) (Hunter, Langlois d'Estaintot & Kennard, 1989). The twist displays an average value of 32° resulting in 11.3 base pairs per turn of helix. The smallest value, as for d(CTCTAGAG) is observed at the central region and will be discussed later. The rise per base pair shows a range of 2.8–3.7 Å with an average value of 3.3 Å and an average tilt value of 1°. This angle is unusually small for A-DNA (usually 20°) and more closely resembles that of B-DNA (about 0°). The distance of base pairs from the helix axis is 3.5 Å, almost 1 Å less than the value for classical fibre A-DNA, but comparable to later A-form octamer results (Kennard & Hunter, 1989). The propeller twist is larger for the A–T base pairs, with an average value of −13°, than for the C–G base pairs with an average of −9°.

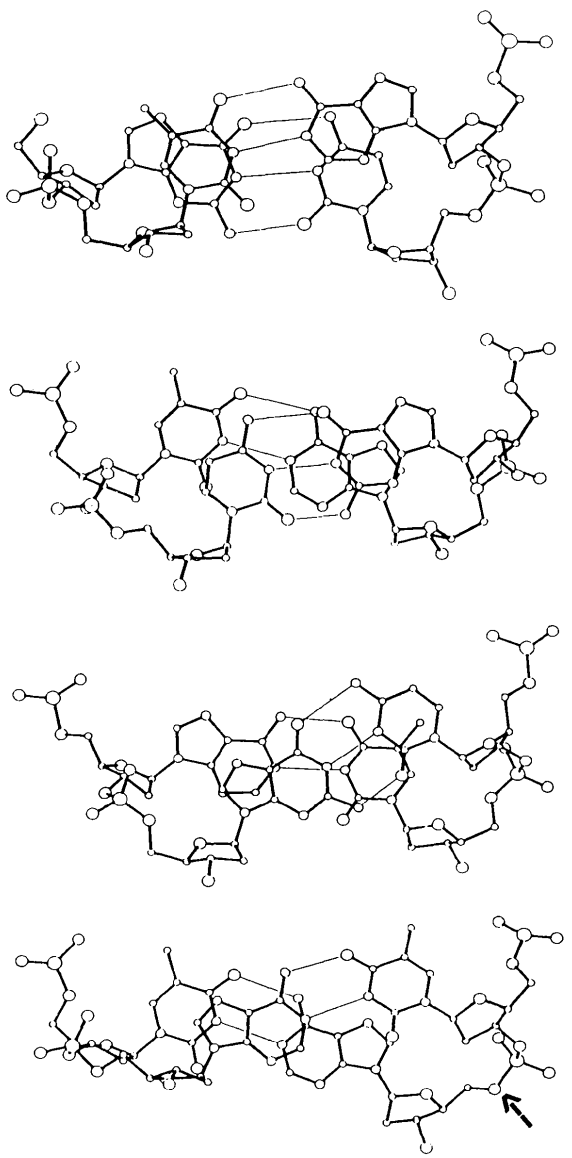


Fig. 2. Views of base-pair steps, top to bottom: G1–C16: T2–A15, T2–A15: C3–G14, C3–G14: T4–A13, T4–A13: A5–T2. Hydrogen bonds are shown as thin lines. Atoms are depicted as spheres of decreasing radii in the order P, O, N, C. The arrow highlights the unusual *trans* conformation observed in step 4.

Table 2. *Geometrical properties of base-pair steps and base pairs*

Parameters for d(GTCTAGAC) were calculated using the *NEWHELIX* program (Dickerson, 1989) brought into agreement with the Cambridge Conventions (1989). The values for d(CTCTAGAG) are in brackets.

Base pair	Step	Roll (°)	Slide (Å)	Twist (°)	Rise (Å)	Base-pair tilt (°)	Propeller twist (°)
G1–C16		9				0	10
(C1–G16)		(7)				(1)	(−9)
	1		2.9	32	3.7		
			(−1.4)	(33)	(3.3)		
T2–A15		13				3	−14
(T2–A15)		(10)				(0)	(−11)
	2		−3.2	34	3.3		
			(−1.6)	(32)	(3.2)		
C3–G14		1				2	−9
(C3–G14)		(5)				(2)	(−4)
	3		3.0	35	2.8		
			(−1.3)	(41)	(2.9)		
T4–A13		10				0	−11
(T4–A13)		(6)				(0)	(−12)
	4		−2.7	25	3.4		
			(−1.7)	(21)	(3.2)		
A5–T12							
(A5–T12)							
Average		8	−2.9	32	3.3	1	−11
		(7)	(−1.5)	(32)	(3.1)	(1)	(−9)

The sugar-phosphate torsion angles and intra-strand phosphorus distances are presented in Table 3. Values of χ are − antiperiplanar except for χ of C8 which is anticlinal; α extends from ± synclinal to ± anticlinal; β values are ± antiperiplanar and γ angles are ± synclinal/ − antiperiplanar/ + anticlinal whilst ε and ζ torsion angles show − anticlinal/ − antiperiplanar and − synclinal/ − anticlinal conformations respectively. Most of the furanose moieties in d(GTCTAGAC) adopt the C3′-*endo* conformation, as evidenced by δ values (range 77–81°, average 79°), the amplitude of pucker (range 41–45°, average 43°) and the pseudorotation phase angle (range 16–30°, average 23°). Residues C3, A5 and C8 however, adopt C1-*exo*, O4′-*endo* and C1′-*exo*/C2′-*endo* conformations respectively. The α and γ angles associated with A5 angles are anomalous in that they are antiperiplanar and produce an unusual all-*trans* conformation about the P—O5′—C5′—C4′ bonds. This behaviour has been observed at the central pyrimidine–purine step in a number of structures, for example in d(CCCCGGGG) (Haran, Shakked, Wang & Rich, 1987) and d(GCCCGGGC) (Heinemann, Lauble, Frank & Blocker, 1987). In these structures there is a marked degree of interstrand overlap of the guanines. In d(GTCTAGAC) similar stacking interactions are observed despite replacement of the C–G base pairs by A–T pairs. This central step has the smallest helical twist (25°)

Table 3. *Torsion angles* (°) *and distances between adjacent phosphorus atoms* (Å) *for* d(GTCTAGAC)

Residue	χ	α	β	γ	δ	ε	ζ	P−P +1
G1	−169			79	80	−122	−93	
T2	−154	−61	−205	62	78	−143	−105	5.9
C3	−150	−302	−174	284	122	−176	−64	6.4
T4	−152	−69	−177	52	81	−164	−88	5.8
A5	−168	−230	−162	196	94	−164	−68	6.9
G6	−170	−117	−164	93	78	−133	−83	5.7
A7	−168	−56	−193	55	77	−169	−71	6.1
C8	−113	−320	−182	322	133			
Average	−156	−244	−180	143	93	−153	−72	

Torsion angles refered to above correspond as follows: $\alpha = {}_{(n-1)}$O3′—P—O5′—C5′, β = P—O5′—C5′—C4′, γ = O5′—C5′—C4′—C3′, δ = C5′—C4′—C3′—O3′, ε = C4′—C3′—O3′—P, ζ = C3′—O3′—P—O5′$_{(n+1)}$, χ = O4′—C1′—N1—C2(pyrimidines) and O4′—C1′—N9—C4(purines). Atoms designated $(n-1)$ and $(n+1)$ belong to adjacent nucleotide units.

possibly due to strong cross-strand stacking interactions. These cross-strand associations force the pliable sugar-phosphate backbone to adopt this unusual conformation which is arrowed in Fig. 2. Similar stacking interactions and backbone conformations at the central TpA (=TpA) step have been observed in previous structures (Jain, Zon & Sundaralingam, 1987; Hunter, Langlois d'Estaintot & Kennard, 1989). The interstrand adjacent phosphorus atom distances (P−P+1) highlight this unusual structural feature. The normal value is about 6 Å in A-form DNA and 7 Å in B-form DNA (Saenger, Hunter & Kennard, 1986). At the central step the distorted backbone results in a value of 6.9 Å.

In d(GTCTAGAC) the minor groove width ranges from 9.4 to 9.9 Å with an average of 9.5 Å. The fibre models for A-DNA and B-DNA have minor groove widths of about 11 and 6 Å respectively (Arnott & Hukins, 1972). The major-groove width for d(GTCTAGAC) is 9 Å. This value is much greater than predicted from fibre diffraction studies of A-DNA where a value of 4 Å is obtained. Rather it resembles that of B-DNA where values of 11 or 12 Å are found.

In summary, we have determined the crystal structure of the self-complementary oligonucleotide d(GTCTAGAC) and compared it with the isomorphous companion d(CTCTAGAG) (Hunter, Langlois d'Estaintot & Kennard, 1989). The right-handed duplex formed in the crystal displays an overall A-type structure. On close inspection it becomes apparent that a number of structural features are similar to those of B-DNA. The dominant feature in the structure at the local level occurs at the central TpA (=TpA) step. Pyrimidine–purine steps are highly bistable (Calladine & Drew, 1984) since the purines can have strong interactions with either the base on the same strand or with the base of the opposite strand. In d(GTCTAGAC) as in d(CTCTAGAG) (Hunter, Langlois d'Estaintot & Kennard, 1989) there is a cross-strand adenine stacking at the central step. This may influence the sugar-phosphate backbone to adopt an anomalous extended conformation.

We thank the Cervi family, the University of Manchester and the Wellcome Trust for financial support, the Science and Engineering Council (UK) for the purchase of equipment. BLdE acknowledges receipt of a short-term EMBO fellowship. We thank Gordon Leonard for comments on the manuscript.

References

ARNOTT, S. & HUKINS, D. W. L. (1972). *Biochem. Biophys. Res. Commun.* **17**, 1504–1509.

CALLADINE, C. R. & DREW, H. R. (1984). *J. Mol. Biol.* **178**, 773–782.

CAMBRIDGE CONVENTIONS (1989). *Nucl. Acids Res.* **5**, 1797–1803.

*CCP*4 (1979). The SERC (UK) Collaborative Computing Project No. 4, Daresbury Laboratory, Daresbury, Warrington WA4 4AD, England.

COURSEILLE, C., DAUTANT, A., HOSPITAL, M., LANGLOIS D'ESTAINTOT, B., PREGIGOUX, G., MOLKO, D. & TEOULE, R. (1990). *Acta Cryst.* A**46**, FC9–FC12.

CROMER, D. T. & WABER, J. T. (1974). *International Tables for X-ray Crystallography*, Vol. IV. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

DICKERSON, R. E. (1989). *NEWHELIX*. Molecular Biology Institute, Univ. of California, Los Angeles, CA 90024, USA.

HARAN, T. E., SHAKKED, Z., WANG, A. H.-J. & RICH, A. (1987). *J. Biomol. Struct. Dyn.* **5**, 199–217.

HEINEMANN, U., LAUBLE, H., FRANK, R. & BLOCKER, H. (1987). *Nucl. Acids Res.* **22**, 9531–9550.

HENDRICKSON, W. A. & KONNERT J. H. (1981). *Biomolecular Structure and Conformation, Function and Evolution*, edited by R. SRINIVARSEN, pp. 43–57. Oxford: Pergamon Press.

HUNTER, W. N., LANGLOIS D'ESTAINTOT, B. & KENNARD, O. (1989). *Biochemistry*, **28**, 2444–2451.

JAIN, S., ZON, G. & SUNDARALINGAM, M. (1987). *J. Mol. Biol.* **197**, 141–145.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

JONES, T. A. & CAMBILLEAU, C. (1989). *Silicon Graphics Geometry Partners Directory*, 89.

KENNARD, O. & HUNTER, W. N. (1989). *Q. Rev. Biophys.* **22**, 327–379.

MCBRIDE, L. J. & CARUTHERS, M. H. (1983). *Tetrahedron Lett.* **24**, 245–248.

MCPHERSON, A. (1990). *Eur. J. Biochem.* **189**, 1–23.

MOTHERWELL, S. & CLEGG, W. (1978). *PLUTO*. Program for plotting molecular and crystal structures. Univ. of Cambridge, England.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

RABINOVICH, D., HARAN, T., EISENSTEIN, M. & SHAKKED Z. (1988). *J. Mol. Biol.* **200**, 151–161.

SAENGER, W. (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.

SAENGER, W., HUNTER, W. N. & KENNARD, O. (1986). *Nature (London)*, **324**, 385–388.

SHAKKED, Z., GUERSTEIN-GUZIKEVICH, G., EISENSTEIN, M., FROLOW, F. & RABINOVICH, D. (1989). *Nature (London)*, **342**, 446–460.

SHAKKED, Z., RABINOVICH, D., KENNARD, O., CRUSE, W. B. T., SALISBURY, S. A. & VISWAMITRA, M. A. (1983). *J. Mol. Biol.* **166**, 183–201.
WANG, A. H.-J., FUJII, S., VAN BOOM, J. H. & RICH, A. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 3968–3972.
WANG, A. H.-J., FUJII, S., VAN BOOM, J. H., VAN DER MAREL, G. A., VAN BOECKEL, S. & RICH, A. (1982). *Nature (London)*, **299**, 601–604.
WESTHOF, E., DUMAS, P. & MORAS, D. (1985). *J. Mol. Biol.* **184**, 119–145.

Acta Cryst. (1992). B**48**, 719–725

Two Closely Related Structures: (−)-(7a*S*)-2,3,7,7a-Tetrahydro-7a-phenylthio-1*H*-indene-1,5(6*H*)-dione and its Racemic Compound

BY CAROLYN PRATT BROCK, STEFAN KWIATKOWSKI* AND DAVID S. WATT

Department of Chemistry, University of Kentucky, Lexington, KY 40506-0055, *USA*

AND ASFAQ SAYED

Enraf–Nonius, 390 *Central Avenue, Bohemia, NY* 11716, *USA*

(*Received* 20 *August* 1991; *accepted* 25 *March* 1992)

Abstract

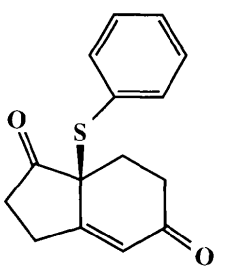
$C_{15}H_{14}O_2S$, $M_r = 258.34$. Racemic compound: $T_{fus} =$ 397–399 K (124–126 °C), monoclinic, $P2_1/c$, $a =$ 7.349 (1), $b = 7.203$ (2), $c = 24.729$ (4) Å, $\beta =$ 91.27 (1)°, $V = 1308.7$ (4) Å^3, $Z = 4$, $D_x =$ 1.311 g cm^{-3}, λ(Mo $K\alpha$) = 0.7107 Å, $\mu = 2.27$ cm^{-1}, $(\sin\theta/\lambda)_{max} = 0.650$ Å^{-1}, $F(000) = 544$, $T =$ 297 (1) K, $R = 0.038$ for 219 variables and 2228 observed reflections [$I > 3\sigma(I)$]. Chiral crystals: T_{fus} = 350–351 K (77–78 °C), orthorhombic, $P2_12_12_1$, a = 7.3342 (8), $b = 7.3587$ (4), $c = 24.613$ (6) Å, $V =$ 1328.4 (4) Å^3, $Z = 4$, $D_x = 1.292$ g cm^{-3}, λ(Cu $K\alpha$) = 1.5406 Å, $\mu = 20.43$ cm^{-1}, $(\sin\theta/\lambda)_{max} =$ 0.497 Å^{-1}, $F(000) = 544$, $T = 295$ (1) K, $R = 0.027$ for 220 variables and 1318 observed reflections [$I >$ $3\sigma(I)$], absolute configuration determined at the 0.001 confidence level so that the molecule is identified as the (*S*)-(−)-enantiomer. The common motif in these very closely related structures is a double layer of molecules; adjacent double layers are related by inversion centers in the racemic compound and by 2_1 screw axes in the chiral crystals. The conglomerate of the chiral crystals is very unstable relative to the racemic compound. The chiral crystals are also 1.5% less dense; the root-mean-square atomic displacements are *ca* 17% larger in the $P2_12_12_1$ than in the $P2_1/c$ crystals. Calculation of an idealized phase diagram suggests that attempts to grow well formed chiral crystals are unlikely to succeed unless the solution contains less than *ca* 1% of the 'wrong' enantiomer.

* On leave from the Warsaw Institute of Technology, Noakowskiego 3, 00-664 Warsaw, Poland.

0108-7681/92/050719-07$06.00

Introduction

Samples of 2,3,7,7a-tetrahydro-7a-phenylthio-1*H*-indene-1,5(6*H*)-dione (hereafter, THIOPH) enriched



(−)-THIOPH

in the (−)-enantiomer (see below) were synthesized as described previously (Kwiatkowski, Syed, Brock & Watt, 1989) by application of the Hajos–Parrish asymmetrically biased aldol condensation to achiral 2-(3-oxobutyl)-2-phenylthio-1,3-cyclopentanedione using (+)-(*R*)-proline as a catalyst. A ^{1}H NMR analysis using a chiral shift reagent showed that the resulting product was obtained in greater than 95% enantiomeric excess. Preliminary estimates suggested that it would be possible to determine the absolute configuration of the product from an X-ray structure determination if Cu radiation were used ($\Delta f'' =$ 0.56 e for the S atom, *i.e.*, 0.4% of the scattering power of the molecule). Some of the product was therefore dissolved in warm methanol, and the solution was allowed to cool and evaporate until a small crop of good-looking crystals had been obtained. One of the resulting lath-like crystals was mounted in