KILB, R. W., LIN, C. C. & WILSON, E. B. (1975). *J. Chem. Phys.* **26**, 1695–1703.
KITANO, M. & KUCHITSU, K. (1973). *Bull. Chem. Soc. Jpn*, **46**, 3048–3051.
KITANO, M. & KUCHITSU, K. (1974). *Bull. Chem. Soc. Jpn*, **47**, 67–72.
KOETZLE, T. F. & LEHMANN, M. S. (1976). *The Hydrogen Bond*, edited by P. SCHUSTER, G. ZUNDEL & C. SANDORFY, Vol. II, pp. 457–471. Amsterdam: North Holland.
MCMULLAN, R. K., FOX, R. O. JR & CRAVEN, B. M. (1978). Unpublished.
MEULENAER, J. DE & TOMPA, H. (1965). *Acta Cryst.* **19**, 1014–1018.
MIYAZAWA, T., SHIMANOUCHI, T. & MIZUSHIMA, S. (1956). *J. Chem. Phys.* **24**, 408–418.
OLOVSSON, I. & JÖNSSON, P. G. (1976). *The Hydrogen Bond*, edited by P. SCHUSTER, G. ZUNDEL & C. SANDORFY, Vol. II, pp. 393–456. Amsterdam: North Holland.
OTTERSEN, T. (1975). *Acta Chem. Scand. Ser. A*, **29**, 939–944.
PIMENTEL, C. C. & MCCLELLAN, A. L. (1960). *The Hydrogen Bond*. San Francisco: Freeman.
POPLE, J. A., KRISHNAN, R., SCHLEGEL, H. B. & BINKLEY, J. S. (1979). *Int. J. Quantum Chem. Symp.* No. 13, pp. 225–241.
RADOM, L., LATHAN, W. A., HEHRE, W. J. & POPLE, J. A. (1972). *Aust. J. Chem.* **25**, 1601–1612.
SENTI, F. & HARKER, D. J. (1940). *J. Am. Chem. Soc.* **62**, 2008–2019.
SHULL, D. G. (1972). Unpublished.
STEVENS, E. D. (1978). *Acta Cryst.* B**34**, 544–551.
TAKAGI, K. & OKA, T. (1963). *J. Phys. Soc. Jpn*, **18**, 1174–1180.
TEMPLETON, L. K. & TEMPLETON, D. H. (1973). Am. Crystallogr. Assoc. Meet., Storrs, Connecticut. Abstr. E-10.

Acta Cryst. (1980). B**36**, 2299–2303

The Structure of 4-Thiopseudouridine

BY CHARLES L. BARNES AND STUART W. HAWKINSON

Department of Biochemistry, The University of Tennessee, Knoxville, Tennessee 37916, *USA*

AND PAUL W. WIGLER

University of Tennessee Memorial Research Center, Knoxville, Tennessee 37920, *USA*

(*Received* 21 *December* 1979; *accepted* 8 *April* 1980)

Abstract

$C_9H_{12}N_2O_5S$ is in space group $P2_1$, with a = 15·434 (2), b = 7·4381 (5), c = 14·818 (2) Å, β = 110·45 (1)°, Z = 6, D_c = 1·627, D_o = 1·60 (3) Mg m^{-3}, V = 1593·9 Å^3, M_r = 260·3, λ(Mo $K\alpha_1$) = 0·70926 Å, T = 297 ± 1 K, $F(000)$ = 816. The crystal structure of 4-thiopseudouridine, a thio-substituted derivative of the minor transfer RNA nucleoside pseudouridine, has been determined. The structure was solved by direct methods, but the determination was hampered by a high degree of pseudosymmetry in the crystal. The correct structure was chosen by refinement to an R of 0·043 for the 3972 significant data. The molecules are packed head to head and tail to tail, with the riboses hydrogen bonded and the bases stacked and hydrogen bonded in ribbons. The three independent molecules display no unusual conformational parameters.

Introduction

Pseudouridine is unique among the nucleosides in RNA in having a C–C 'glycosidic' bond, which gives it structural properties different from those of uridine. The synthetic nucleoside 4-thiopseudouridine has been shown by Wigler, Bindslev & Breitman (1974) to inhibit the growth of *Escherichia coli* (strain B5RU) when the cells are grown on pseudouridine. Wigler *et al.* have proposed that a metabolite of 4-thiopseudouridine, presumably the nucleotide, is an inhibitor of the enzyme pseudouridylate synthetase, which catalyzes the key reaction in the only known salvage pathway for pseudouridine 5′-phosphate (Solomon & Breitman, 1971), the conversion of pseudouridine 5′-phosphate to uridine 5′-phosphate. This study was undertaken to determine the conformational features of the 4-thio-substituted pseudouridine for comparison with pseudouridine and with other thio-substituted

0567-7408/80/102299-05$01.00

pyrimidines. Our interest in this project was heightened when preliminary film studies indicated the presence of three independent molecules in the asymmetric unit, for space group $P2_1$, which, according to our own experience and a recent survey (Belsky & Zorkii, 1977), is an uncommon occurrence.

Experimental

4-Thiopseudouridine was synthesized according to Wigler *et al.* (1974) and purified by chromatography (Uziel, Koh & Cohn, 1968). Recrystallization from aqueous ethanol produced well formed yellow monoclinic prisms, elongated along *b*. A specimen 0·35 × 0·2 × 0·1 mm was selected for single-crystal diffraction study. Following initial determination of the space group and approximate cell dimensions by film methods in our laboratory, accurate cell dimensions and intensities were measured by Dr Eric J. Gabe of the National Research Council of Canada, Ottawa, on a modified Picker diffractometer using crystal-monochromatized Mo $K\overline{\alpha}$ radiation in the θ–2θ scanning mode. Of the 5020 independent observations recorded, 3972 were considered observed [$I > 3\sigma(I)$].

The structure determination was approached using *MULTAN* 78 (Main, Hull, Lessinger, Germain, Declercq & Woolfson, 1978). The phase set with the highest combined figure of merit gave a plausible fraction of the structure, which produced more of the structure when used in the Karle (1968) recycling routine. The remaining non-hydrogen atoms were located using weighted Fourier syntheses, though one ribose appeared to be disordered. At this point, the model seemed reasonable, featuring hydrogen-bonded riboses and stacked ribbons of hydrogen-bonded bases. However, refinement of the model, with one whole molecule disordered, converged to an R of 0·14, a result obviously inconsistent with the quality of the data. Examination of a model of the structure led to the observation that the stacked arrangement of the independent molecules results in a set of three reasonable structures differing primarily in the location of the origin along *c*. A shift of the origin of the false solution by $c/3$ gave a set of parameters which eventually refined to an R of 0·043. This structural pseudosymmetry is remarkably similar to that observed by van Zoeren, Oonk & Kroon (1978) in the structure of L-pyroglutamic acid in the space group $P2_12_12_1$, with $Z = 12$.

Anisotropic refinement of the non-hydrogen atoms was performed by the least-squares method with the matrix in six blocks. The H atoms were located by difference Fourier syntheses during the later stages of refinement, and their idealized locations were updated between cycles. Atoms related by the approximate, non-crystallographic, translational symmetry operation $0,0,\frac{1}{3}$ were assigned to the same block during refinement. Each intensity was assigned a variance, $\sigma^2(I)$, based on counting statistics, plus a term,

Table 1. *Positional and thermal parameters* ($\times 10^4$) *for the non-hydrogen atoms*

The isotropic temperature factor is $\exp[-8\pi^2 U(\sin\theta/\lambda)^2]$; the values given are the arithmetic averages of the principle components of the anisotropic temperature factors. Standard deviations in units of the last significant digits are given in parentheses.

	x	*y*	*z*	*U* (Å²)
Molecule *A*				
N(1)	845 (2)	−653 (3)	−523 (2)	342 (14)
C(2)	500 (2)	880 (4)	−274 (2)	279 (14)
O(2)	−54 (2)	879 (3)	155 (2)	382 (12)
N(3)	820 (2)	2438 (3)	−535 (2)	304 (13)
C(4)	1457 (2)	2589 (4)	−994 (2)	278 (15)
S	1749 (1)	4585 (0)	−1266 (1)	548 (5)
C(5)	1821 (2)	917 (4)	−1188 (2)	265 (14)
C(6)	1497 (2)	−626 (4)	−955 (2)	301 (15)
C(1′)	2572 (2)	964 (4)	−1623 (2)	257 (13)
C(2′)	3521 (2)	1412 (4)	−871 (2)	257 (13)
O(2′)	4018 (1)	2504 (3)	−1302 (1)	338 (11)
C(3′)	3955 (2)	−465 (4)	−674 (2)	247 (13)
O(3′)	4924 (1)	−458 (3)	−210 (1)	315 (11)
C(4′)	3614 (2)	−1272 (4)	−1674 (2)	232 (13)
O(1′)	2651 (1)	−777 (3)	−2020 (1)	270 (10)
C(5′)	3723 (2)	−3258 (4)	−1693 (2)	321 (15)
O(5′)	3503 (1)	−3916 (3)	−2650 (2)	411 (12)
Molecule *B*				
N(1)	960 (2)	−483 (4)	2552 (2)	308 (12)
C(2)	623 (2)	1048 (4)	2804 (2)	272 (14)
O(2)	59 (1)	1067 (3)	3216 (2)	413 (12)
N(3)	964 (2)	2611 (3)	2558 (2)	298 (13)
C(4)	1638 (2)	2736 (4)	2140 (2)	244 (13)
S	1926 (1)	4723 (1)	1838 (1)	433 (4)
C(5)	2035 (2)	1062 (4)	2010 (2)	244 (13)
C(6)	1669 (2)	−475 (4)	2209 (2)	280 (14)
C(1′)	2856 (2)	1083 (4)	1689 (2)	248 (13)
C(2′)	3750 (2)	1592 (4)	2515 (2)	263 (13)
O(2′)	4291 (1)	2725 (3)	2147 (2)	342 (11)
C(3′)	4221 (2)	−230 (4)	2777 (2)	271 (13)
O(3′)	5181 (1)	−134 (3)	3321 (1)	342 (11)
C(4′)	3957 (2)	−1132 (4)	1798 (2)	265 (14)
O(1′)	2998 (1)	−685 (3)	1363 (1)	305 (11)
C(5′)	4071 (2)	−3137 (4)	1856 (2)	353 (16)
O(5′)	3810 (2)	−3896 (3)	919 (2)	467 (13)
Molecule *C*				
N(1)	656 (2)	−514 (4)	6308 (2)	322 (13)
C(2)	305 (2)	1030 (4)	6544 (2)	295 (14)
O(2)	−245 (1)	1054 (3)	6975 (2)	394 (12)
N(3)	618 (2)	2581 (3)	6261 (2)	311 (12)
C(4)	1248 (2)	2694 (4)	5792 (2)	265 (14)
S	1572 (1)	4692 (1)	5529 (1)	438 (4)
C(5)	1585 (2)	1019 (4)	5577 (2)	260 (13)
C(6)	1280 (2)	−516 (4)	5848 (2)	304 (15)
C(1′)	2286 (2)	1010 (4)	5074 (2)	257 (13)
C(2′)	3293 (2)	1356 (4)	5751 (2)	287 (13)
O(2′)	3731 (1)	2492 (3)	5276 (2)	393 (12)
C(3′)	3703 (2)	−526 (4)	5807 (2)	268 (13)
O(3′)	4684 (1)	−582 (3)	6156 (2)	364 (12)
C(4′)	3221 (2)	−1168 (4)	4785 (2)	280 (14)
O(1′)	2282 (1)	−714 (3)	4637 (1)	335 (12)
C(5′)	3306 (2)	−3142 (5)	4585 (2)	374 (17)
O(5′)	4231 (2)	−3488 (4)	4644 (2)	496 (14)

$(0{\cdot}04I)^2$, empirically derived during refinement. The final values of R, the weighted R $\{[\sum w(F_o - F_c)^2/\sum wF_o^2]^{1/2}\}$, and σ, the goodness of fit $\{[\sum w(F_o - F_c)^2/(n-p)]^{1/2}$, where $n = 3972$ reflections and $p = 460$ variables$\}$, were 0·043, 0·047 and 1·40, respectively. The average shift on the final round of refinement was less than 20% of the estimated standard deviation. All refinements were performed with the XRAY system (Stewart, Kruger, Ammon, Dickinson & Hall, 1972). The final atomic parameters are given in Table 1.*

Structural details

Drawings of the molecules, including bond distances and angles, are shown in Fig. 1. Differences in pseudouridine ring bond angles in comparison to those observed in other diketo pyrimidine structures are attributable to the substitution of the C(5)–C(1′) bond for the normal glycosidic N(1)–C(1′) bond, and are comparable to those observed in α-pseudouridine (Rohrer & Sundaralingam, 1970). The C–S bond distances are in the short end of the range reported for other thio-substituted pyrimidines: 1·645–1·684 Å in a review of several structures (Saenger & Suck, 1971), 1·669 Å in 1-methyl-4-thiouracil (Hawkinson, 1975), and 1·677 Å in 2-thiouridine (Hawkinson, 1977). This difference may reflect an increased degree of double-bond character in the C(4)–S bond in this structure.

* Lists of structure factors, anisotropic thermal parameters and H atom positional parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 35320 (26 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Torsion angles descriptive of the molecular conformation are given in Table 2. We follow the definition of the angle χ [C(6)–C(5)–C(1′)–O(1′)] established by Rohrer & Sundaralingam (1970) as this allows comparison with values determined for other nucleosides. By this definition, all three molecules are in the *anti* conformation. The sugars are intermediate in conformation between C(3′)-*endo* and C(4′)-*exo*. The C(3′)-*endo* conformation has been commonly observed in pyrimidine nucleosides (see, for example, the review by Sundaralingam, 1975). Indeed, in the structure of uridine (Green, Rosenstein, Shiono, Abraham, Trus & Marsh, 1975) the values of χ for the two independent molecules, both with C(3′)-*endo* riboses, are 18·3 and 24·3°, which overlap with the values found in this structure.

Table 2. *Torsion angles and ribose pseudo-rotation parameters for molecules A, B and C*

		A	*B*	*C*
Torsion angles				
χ	O(1′)–C(1′)–C(5)–C(6) [ribose *anti* to S(4)]	18·4 (3)°	16·8 (3)°	21·4 (3)°
τ_0	C(4′)–O(1′)–C(1′)–C(2′)	−9·3 (2)	−10·8 (2)	−15·9 (2)
τ_1	O(1′)–C(1′)–C(2′)–C(3′)	−19·1 (2)	−16·6 (2)	−14·6 (2)
τ_2	C(1′)–C(2′)–C(3′)–C(4′)	38·9 (2)	36·0 (2)	37·5 (2)
τ_3	C(2′)–C(3′)–C(4′)–O(1′)	−45·7 (2)	−43·6 (2)	−48·4 (2)
τ_4	C(3′)–C(4′)–O(1′)–C(1′)	34·3 (2)	33·9 (2)	40·2 (2)
ψ	O(5′)–C(5′)–C(4′)–C(3′)	−172·5 (2)	179·7 (2)	−70·7 (3)
ψ'	C(5′)–C(4′)–C(3′)–O(3′)	71·0 (3)	73·3 (3)	66·8 (3)
Ribose pseudo-rotation parameters				
φ	Phase angle from ${}_2^3T$	28°	30°	35°
Q	Amplitude of pucker	0·44 Å	0·42 Å	0·46 Å
	Conformational descriptor	${}^3T_4$	${}^3T_4$	${}_4^3T$

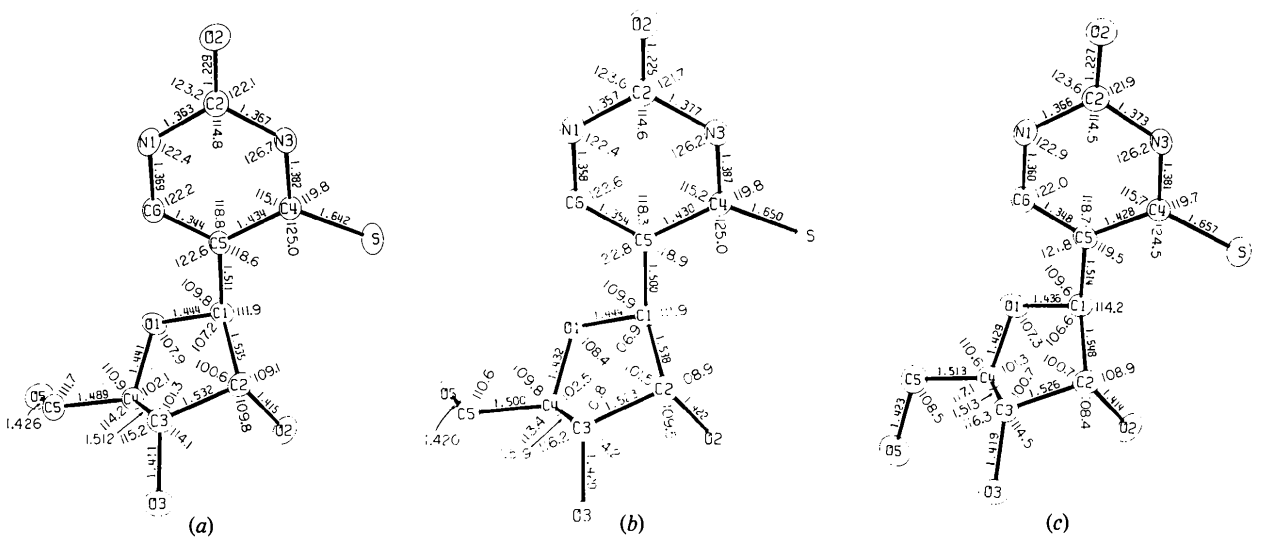

Fig. 1. Bond distances (Å) and angles (°) for (*a*) molecule *A*, (*b*) molecule *B* and (*c*) molecule *C*. The thermal ellipsoids are drawn at the 50% probability level. The average standard deviations in the bond lengths and bond angles are 0·004 Å and 0·3°, respectively. Illustrations were prepared with the aid of the program *ORTEP* (Johnson, 1965).

Nucleosides are flexible in conformation about the C(4′)–C(5′) exocyclic bond of the ribose. All three possible conformations defined by the angle C(3′)–C(4′)–C(5′)–O(5′), *gauche*$^+$ (about 60°), *trans* (about 180°) and *gauche*$^-$ (about 300°), have been observed. In this structure, molecules *A* and *B* are *trans*, while molecule *C* is *gauche*$^-$, and this difference, which constitutes the major conformational dissimilarity among the three molecules, may have implications in the packing scheme discussed below.

Packing details

Fig. 2 is a stereoview of the packing interactions. The head-to-head, tail-to-tail packing of the nucleoside molecules results in alternating layers of hydrogen-bonded riboses and stacked, hydrogen-bonded bases. The hydrogen-bonded bases form two crystallographically distinct ribbons, both parallel to the *b* axis. Fig. 3 shows part of the ribbon composed of screw-related molecules *A*. Fig. 4 shows the second kind of ribbon, composed of alternating molecules *B* and *C*, arrayed along a pseudo-twofold screw axis. The major conformational difference between molecules *B* and *C* is the torsion angle about the C(4′)–C(5′) bond, the molecules being otherwise quite similar.

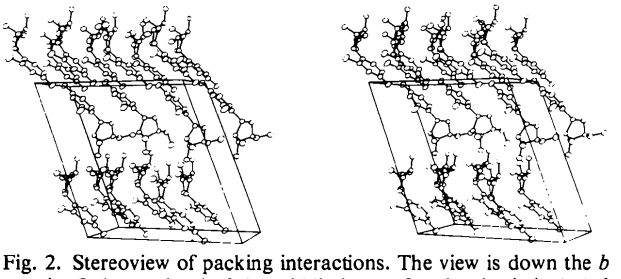

Fig. 2. Stereoview of packing interactions. The view is down the *b* axis. Only one level of the middle layer of molecules is included for clarity.

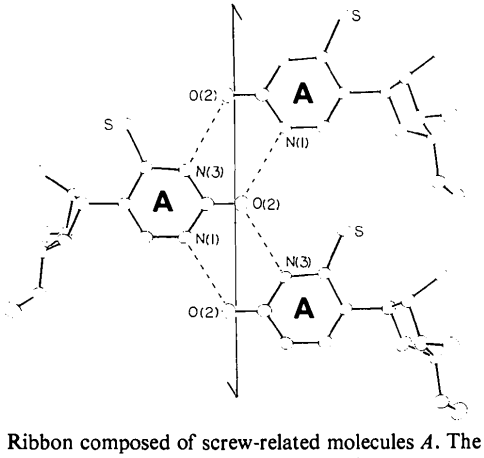

Fig. 3. Ribbon composed of screw-related molecules *A*. The dotted lines indicate hydrogen bonds.

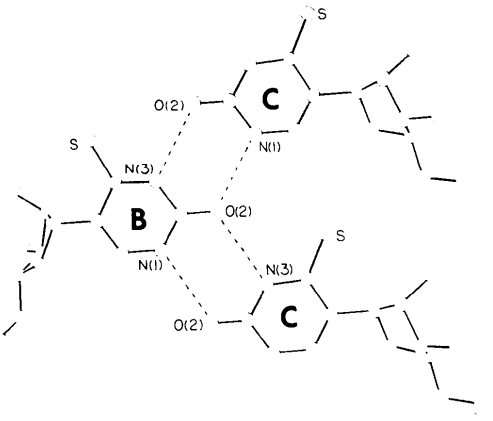

Fig. 4. Ribbon composed of pseudo-twofold-screw-related molecules *B* and *C*. The dotted lines indicate hydrogen bonds.

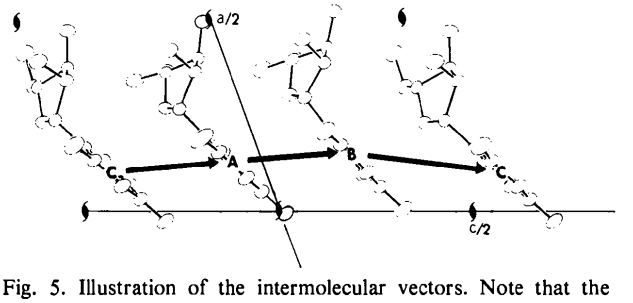

Fig. 5. Illustration of the intermolecular vectors. Note that the pseudo-translational symmetry results in pseudo-twofold screw axes parallel to the crystallographic screw axes.

A careful examination of the figures reveals that the best position for the pseudo-twofold screw axis relating molecules *B* and *C* is not quite on the *c* axis, but that the pseudosymmetry is good enough that relocation of the origin to $0,0,\frac{1}{6}$ or $0,0,\frac{1}{3}$ would result in structures closely resembling the true structure. Such a false solution caused the initial difficulty with the structure determination. The pseudo-translational symmetry relating molecules *A*, *B* and *C* is further illustrated in Fig. 5. Note that the intermolecular vectors $C \rightarrow A$ and $A \rightarrow B$ are very nearly identical, but that $B \rightarrow C$ does not exactly overlap the other two. Were molecule *C* not shifted, with the concomitant ribose conformational change, the result would be a new unit cell $\frac{1}{3}$ of the volume of the observed cell. The packing would then be identical to that found in a number of pyrimidine nucleoside structures.

The hydrogen-bond distances and angles are given in Table 3. The sequence of hydrogen bonds involving ribose hydroxyls is O(2′)H···O(3′)H···O(5′)H···*X*. In many nucleoside structures the O(5′)H···*X* interaction is weak, as in, for example, 2-thiouridine (Hawkinson, 1977), where *X* = S. In the present case, there are three different types of hydrogen bonds O(5′)H···*X*, where *X* is O(2′), O(3′) or S, and all three

Table 3. *Hydrogen bonds and close contacts of the type D—H···A*

a, *b* and *c* after the atom labels indicate molecules *A*, *B* and *C*.

Donor (*D*)	Acceptor (*A*)	In molecule at	*D*—*A* (Å)	H—*A* (Å)	*D*—H—*A* (°)
O(2')*a*	O(3')*a*	$1-x, \frac{1}{2}+y, -z$	2·725 (3)	1·9	170
O(3')*a*	O(5')*b*	$1-x, \frac{1}{2}+y, -z$	2·775 (4)	2·0	157
O(5')*a*	Sc	$x, -1+y, -1+z$	3·408 (3)	2·7	144
O(2')*b*	O(3')*c*	$1-x, \frac{1}{2}+y, 1-z$	2·762 (3)	1·9	164
O(3')*b*	O(5')*a*	$1-x, \frac{1}{2}+y, -z$	2·709 (4)	1·9	157
O(5')*b*	O(2')*b*	$x, -1+y, z$	3·039 (3)	2·3	140
O(2')*c*	O(3')*b*	$1-x, \frac{1}{2}+y, 1-z$	2·796 (3)	1·9	175
O(3')*c*	O(5')*c*	$1-x, \frac{1}{2}+y, 1-z$	2·833 (4)	2·0	161
O(5')*c*	O(3')*b*	$1-x, -\frac{1}{2}+y, 1-z$	3·083 (3)	2·3	160
N(1)*a*	O(2)*a*	$-x, -\frac{1}{2}+y, -z$	2·984 (4)	2·1	176
N(3)*a*	O(2)*a*	$-x, \frac{1}{2}+y, -z$	2·955 (3)	2·1	177
N(1)*b*	O(2)*c*	$-x, -\frac{1}{2}+y, 1-z$	2·981 (4)	2·1	169
N(3)*b*	O(2)*c*	$-x, \frac{1}{2}+y, 1-z$	2·967 (4)	2·1	177
N(1)*c*	O(2)*b*	$-x, -\frac{1}{2}+y, 1-z$	2·954 (4)	2·1	176
N(3)*c*	O(2)*b*	$-x, \frac{1}{2}+y, 1-z$	2·997 (4)	2·2	176

interactions are longer than normal. This change in structure must be responsible for a slight lowering in energy resulting in the crystal structure having three molecules per asymmetric unit rather than one.

Conclusions

It has been noted previously that the substitution of S for O in pyrimidine nucleosides has little effect on the conformational details of the nucleoside (Hawkinson, 1977). As the structure of the β isomer of pseudouridine has not been reported in the literature, a direct comparison with the present structure is not possible, though the 'normal' conformational parameters of 4-thiopseudouridine would seem to support this conclusion. The activity of 4-thiopseudouridine as an inhibitor of pseudouridine metabolism may correlate with the greater length of the C—S bond, the larger van der Waals radius for S, or the decreased ability of S to participate as an acceptor in hydrogen bonding, compared to O, or with an altered conformation of the nucleotide as it is bound to the enzyme active site. In the present structure, only one of the S atoms is involved as an acceptor in a weak interaction.

References

BELSKY, V. K. & ZORKII, P. M. (1977). *Acta Cryst.* A**33**, 1004–1006.

GREEN, E. A., ROSENSTEIN, R. D., SHIONO, R., ABRAHAM, D. J., TRUS, B. L. & MARSH, R. E. (1975). *Acta Cryst.* B**31**, 102–107.

HAWKINSON, S. W. (1975). *Acta Cryst.* B**31**, 2153–2156.

HAWKINSON, S. W. (1977). *Acta Cryst.* B**33**, 80–85.

JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Tennessee.

KARLE, J. (1968). *Acta Cryst.* B**24**, 182–186.

MAIN, P., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J. P. & WOOLFSON, M. M. (1978). *MULTAN 78. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univs. of York, England, and Louvain-la-Neuve, Belgium.

ROHRER, D. C. & SUNDARALINGAM, M. (1970). *J. Am. Chem. Soc.* **92**, 4950–4955.

SAENGER, V. W. & SUCK, D. (1971). *Acta Cryst.* B**27**, 1178–1186.

SOLOMON, L. R. & BREITMAN, T. R. (1971). *Biochem. Biophys. Res. Commun.* **44**, 299–304.

STEWART, J. M., KRUGER, G. J., AMMON, H. L., DICKINSON, C. & HALL, S. R. (1972). The XRAY system – version of June 1972. Tech. Rep. TR-192. Computer Science Center, Univ. of Maryland, College Park, Maryland.

SUNDARALINGAM, M. (1975). *Structure and Conformation of Nucleic Acids and Protein–Nucleic Acid Interactions*, edited by M. SUNDARALINGAM & S. T. RAO, pp. 487–524. Baltimore: Univ. Park Press.

UZIEL, M., KOH, C. K. & COHN, W. E. (1968). *Anal. Biochem.* **25**, 77–98.

WIGLER, P. W., BINDSLEV, B. & BREITMAN, T. R. (1974). *J. Carbohydr. Nucleosides Nucleotides*, **1**, 307–321.

ZOEREN, E. VAN, OONK, H. A. J. & KROON, J. (1978). *Acta Cryst.* B**34**, 1898–1900.