Influence of Crystal Packing Forces on Molecular Structure in 4-Thiouridine. Comparison of *anti* and *syn* Forms

Bogdan Lesyng * and Wolfram Saenger

Max-Planck-Institut für experimentelle Medizin, Abteilung Chemie, Hermann-Rein-Str. 3, D-3400 Göttingen, Bundesrepublik Deutschland

Z. Naturforsch. **36 c,** 956–960 (1981); received July 6, 1981

4-Thiouridine, Crystal Packing Forces, Molecular Structures

4-thiouridine crystallizes from butyric acid in the form of yellow needles, monoclinic space group $P2_1$ with $a = 10.136$ (3), $b = 4.843$ (2), $c = 11.257$ (3) Å, $\beta = 93.91$ (5)°. The crystal structure was solved on the the basis of 895 X-ray counter data using direct methods and refined to a reliability index of $R = 5.6\%$. 4-thiouridine is in the *anti* conformation, the torsion angle O(1′)-C(1′)-N(1)-C(6) being 26.5°. Pseudorotation parameters $\tau_{max} = 41.1°$ and $P = 15.7°$ refer to C(3′)-*endo* (^{3}E) envelope form of the ribose. The conformation about the C(5′)-O(5′) bond is *gauche, gauche.*

The present structure is compared with the previous one, crystallized from water as sesquihydrate and existing as the *syn* conformer (W. Saenger and K. H. Scheit, J. Mol. Biol. **50,** 153–169 (1970). The influence of crystal packing forces on flexible molecules such as nucleosides is discussed.

Introduction

β-4-Thiouridine (s^4U), the sulfur analog of uridine (U), is a minor constituent of tRNA [1–4]. Owing to its particular physical and chemical properties, it has attracted considerable interest and has been the subject of studies in relation to tRNA and concerning the general behaviour of s^4U as substitute for U in biological systems [5–13].

NMR data have demonstrated that s^4U and its 5′-phosphate occur preferentially in the *anti* form in aqueous solution [14–19], similar as other pyrimidine nucleosides and nucleotides. When crystallized from water as sesquihydrate, however, it adopts the unusual *syn* form [13] which in the pyrimidine series has been only observed in a few rare cases with either the base substituted in 6-position [20, 21] or with the ribose moiety modified [22, 23].

As s^4U · sesquihydrate in the crystalline state displays a unique conformation, it was of interest to find out whether crystallization from another solvent would produce the generally preferred *anti* form and if so, what would be the structural differences of the two conformational isomers.

* Present address: Warsaw University, Institute of Exp. Physics, Department of Biophysics, Warsaw, Poland.

Reprint requests to Dr. W. Saenger.

0341-0382/81/1100-0956 $ 01.00/0

Experimental

s^4U was dissolved in butyric acid in the hope to obtain crystals of a specific s^4U · butyric acid complex as model system for nucleoside protein interactions. The yellow, needleshap crystals grown after slow cooling of such a solution, contained however, only s^4U. Systematic absences of X-ray reflections OkO with k odd indicated space group $P2_1$; relevant crystallographic data were measured on an automated four-circle diffractometer, using Ni-filtered CuK_α radiation, $2\theta/\theta$ scan mode with maximum $2\theta = 120°$. Absorption correction was not applied. Crystallographic data are given in Table I. The crystal structure was solved by direct methods [24] and refined by full matrix least squares. Hydrogen

Table I. Crystal data for 4-thiouridine.

Formula	$C_9H_{12}N_2O_5S$
Space group	$P2_1$, monoclinic
Lattice constants	$a = 10.136$ (3) Å
	$b = 4.843$ (2) Å
	$c = 11.257$ (3) Å
	$\beta = 93.91$ (5)°
Volume of unit cell	551.3 Å^3
Molecular weight	260.25
Density	$D_{calc} = 1.568$ g/cm^3
Z	2
F (000)	272
Number of measured data	895
$2\theta_{max}$	120°
R-factor	5.6%

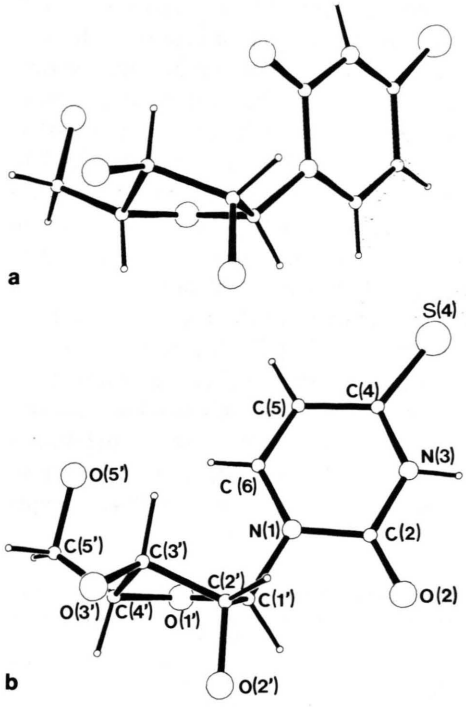

Fig. 1. Schematic views of *syn* (a) and *anti* (b) forms of the s^4U molecule. The plane C(1′)-O(1′)-C(4′) is perpendicular to the plane of the paper.

atoms bonded to C, N and O(5′) were located from difference Fourier maps but those attached to O(2′) and O(3′) did not show up as clearly and were therefore omitted.

Results and Discussion

The molecular conformation of s^4U in the present crystal form is *anti*. In Fig. 1, the two *syn* [13] and *anti* forms are displayed and in Table III, bond angles and distances computed from entries in Table II a, b are compared. As already discussed [13], the main differences are found in bond angles around atoms C(1′) and N(1) because they are directly influenced by the close contacts occurring between O(2) and ribose atoms if the uracil heterocycle is in the *syn* orientation. *I.e.* angles around C(1′) and N(1) in the *anti* form are "normal", those in the *syn* form are widened, compare C(2′)-C(1′)-N(1) and C(1′)-N(1)-C(2), Table III. A noticeable difference one can also observe for the O(3′)-C(3′)-C(2′) angle. Otherwise, bond angles and distances are similar in the two conformational isomers of s^4U.

The torsion angles, (Table IV), however, differ greatly, especially χ_{CN}, defined by O(1′)-C(1′)-N(1)-C(6), is −87.1° in *syn* [13] but 26.5° in the *anti* form. Further, the ribose puckering modes are different, with pseudorotation parameters [25] $\tau_{max} = 41.0°$, $P = 34.2°$ for *syn* corresponding to a C(3′)-*endo*, C(4′)-*exo* ($^3_4$T) twist pucker and for *anti*, $\tau_{max} = 41.1°$, $P = 15.7°$ refers to an ideal C(3′)-*endo* (^{3}E) envelope form. Differences in puckering modes are also clear from Fig. 1 and from the deviations of atoms from the plane defined in Table V. The orientations about the exocyclic C(4′)-C(5′) bonds

Table II. Positional and thermal parameters. Standard deviations in the least significant figures are given in parentheses. Thermal parameters are in the form $T = \exp[-2\pi^2(U_{11}h^2a^{*2} + U_{22}k^2b^{*2} + U_{33}l^2c^{*2} + 2U_{12}hka^*b^* + 2U_{13}hla^*c^* + 2U_{23}klb^*c^*)]$. (a) Nonhydrogen atoms (all values are × 10^4). (b) Hydrogen atoms.
Table II a.

Atom	x/a	y/b	z/c	U11	U22	U33	U23	U13	U12
N (1)	− 353 (4)	5324 (13)	7130 (4)	116 (19)	249 (24)	299 (23)	17 (24)	34 (18)	22 (22)
C (2)	649 (5)	5512 (15)	6344 (5)	173 (24)	353 (30)	221 (24)	− 51 (28)	69 (21)	− 47 (25)
N (3)	1686 (4)	3726 (13)	6582 (4)	162 (21	304 (25)	223 (21)	43 (23)	62 (19)	34 (22)
C (4)	1867 (5)	1965 (16)	7508 (5)	137 (24)	286 (28)	283 (26)	− 10 (27)	4 (21)	16 (25)
C (5)	816 (6)	1911 (17)	8312 (5)	236 (26)	451 (31)	293 (26)	92 (28)	78 (23)	25 (27)
C (6)	− 242 (6)	3581 (17)	8093 (5)	254 (26)	399 (30)	250 (26)	64 (28)	92 (22)	16 (28)
O (2)	597 (4)	7125 (11)	5517 (4)	258 (20)	350 (25)	328 (21)	93 (23)	90 (18)	26 (21)
S (4)	3213 (1)	23 (6)	7658 (1)	238 (8)	412 (10)	373 (9)	61 (10)	50 (7)	114 (9)
C (1′)	− 1505 (6)	7182 (15)	6874 (5)	177 (24)	348 (30)	289 (26)	− 35 (27)	64 (22)	55 (26)
C (2′)	− 2580 (6)	5882 (15)	6061 (5)	167 (24)	392 (30)	192 (24)	− 31 (26)	40 (21)	50 26)
C (3′)	− 3489 (6)	4583 (15)	6966 (5)	175 (23)	262 (29)	305 (26)	− 35 (26)	−18 (21)	− 30 (25)
C (4′)	− 3440 (6)	6809 (16)	7919 (5)	185 (24)	327 (28)	269 (26)	− 7 (27)	86 (22)	93 (25)
O (1′)	− 2061 (4)	7772 (11)	7979 (4)	207 (19)	341 (23)	300 (20)	− 117 (21)	75 (17)	− 7 (20)
O (2′)	− 3227 (4)	8116 (11)	5434 (4)	212 (19)	381 (24)	279 (20)	102 (21)	55 (17)	49 (20)
O (3′)	− 4792 (4)	4242 (12)	6403 (4)	208 (20)	456 (27)	453 (22)	44 (23)	−75 (18)	− 56 (21)
C (5′)	− 3758 (7)	5904 (17)	9155 (6)	308 (27)	532 (33)	367 (28)	95 (29)	168 (24)	164 (29)
O (5′)	− 3028 (5)	3507 (13)	9511 (4)	500 (24)	490 (26)	287 (21)	35 (23)	117 (20)	133 (25)

Table II b.

Atom	x/a	y/b	z/c	U
H (3)	2291	3910	6094	500
H (5)	864	659	9042	500
H (6)	− 882	3446	8552	500
H (1′)	− 1198	9067	6471	500
H (2′)	− 2286	4607	5386	500
H (3′)	− 2912	2749	7360	500
H (4′)	− 4096	8367	7570	500
H (5′)	− 4643	5164	9133	500
H (5′)	− 3614	6958	10072	500
H (5″)	− 2998	3199	10277	500

are *gauche, gauche* in both cases, with relevant torsion angles given in Table IV.

In the crystal structure of uridine, two molecules (denoted U(A) and U(B)) are found in the asymmetric unit [26]. Both of them are in the *anti* conformation with glycosidic torsion angles χ_{CN} and pseudorotation parameters P and τ_{max}:

U(A) $\chi_{CN} = 18.3°$, $\tau_{max} = 40.4°$, $P = 3.7°$;
U(B) $\chi_{CN} = 24.3°$, $\tau_{max} = 42.4°$, $P = 13.8°$.

Comparison with s⁴U shows that *anti*-s⁴U is very similar to U(B).

Table III. Atomic distances (Å) and bond angles (°) in 4-thiouridine.

Atoms	syn [13]		*anti*	
A-B-C	Distance A-B	Angle A-B-C	Distance A-B	Angle A-B-C
N(1)-C(2)-N(3)	1.382	115.3	1.396	114.5
C(2)-N(3)-C(4)	1.381	126.7	1.373	127.7
N(3)-C(4)-C(5)	1.373	114.5	1.349	115.3
C(4)-C(5)-C(6)	1.433	119.8	1.445	119.0
C(5)-C(6)-N(1)	1.346	122.5	1.353	122.2
C(6)-N(1)-C(2)	1.370	120.8	1.372	121.1
O(2)-C(2)-N(1)	1.221	123.4	1.214	122.2
S(4)-C(4)-N(3)	1.662	121.4	1.656	120.4
C(1′)-C(2′)-C(3′)	1.535	102.2	1.512	101.9
C(2′)-C(3′)-C(4′)	1.558	100.5	1.553	100.6
C(3′)-C(4′)-O(1′)	1.534	103.4	1.519	104.0
C(4′)-O(1′)-C(1′)	1.434	109.6	1.470	109.1
O(1′)-C(1′)-C(2′)	1.413	107.6	1.429	107.4
O(2′)-C(2′)-C(1′)	1.416	106.0	1.426	105.8
O(2′)-C(2′)-C(3′)	1.416	109.7	1.426	111.1
O(3′)-C(3′)-C(2′)	1.409	114.8	1.435	109.0
O(3′)-C(3′)-C(4′)	1.409	114.3	1.435	112.2
C(5′)-C(4′)-C(3′)	1.511	118.4	1.514	116.4
C(5′)-C(4′)-O(1′)	1.511	108.7	1.514	108.1
O(5′)-C(5′)-C(4′)	1.439	113.2	1.419	110.9
C(1′)-N(1)-C(2)	1.466	120.6	1.487	115.7
C(1′)-N(1)-C(6)	1.466	118.2	1.487	123.1
O(1′)-C(1′)-N(1)	1.413	108.8	1.429	107.7
C(2′)-C(1′)-N(1)	1.535	119.2	1.512	112.9

The crystal packing patterns of s⁴U in the *syn* [13] and *anti* (Fig. 2) forms are reminiscent of each other because separation into hydrophilic and hydrophobic zones is observed. The former zone is built up of riboses (and water of hydration in *syn*-s⁴U) which are hydrogen bonded to each other and to N(3), S(4), and the latter are formed by heterocycles stacked nearly perpendicular to the stack axis in *syn*-s⁴U but at an angle 44.3° in *anti*-s⁴U, leading to a corrugated sheet structure in that case.

The stacking interactions in thiouracil derivatives have been summarized in [27]. A consistent picture emerged showing (a), that S can be involved in hydrogen bonding as well as O(2) or O(4) and (b), that stacking overlap of adjacent bases is such that S interacts with N(1) or N(3). A similar pattern is also observed in *anti*-s⁴U (Fig. 2) where S(4) accepts

Table IV. Torsion angles in *syn*- and *anti* 4-thiouridine.

	syn [13]	*anti*
C(1′)-C(2′)-C(3′)-C(4′)	33.1	38.9
C(2′)-C(3′)-C(4′)-O(1′)	− 40.6	− 38.4
C(3′)-C(4′)-O(1′)-C(1′)	33.3	23.6
C(4′)-O(1′)-C(1′)-C(2′)	− 11.3	2.0
O(1′)-C(1′)-C(2′)-C(3′)	− 14.9	− 26.0
N(1)-C(1′)-O(1′)-C(4′)	−141.7	−119.9
N(1)-C(1′)-C(2′)-C(3′)	109.4	92.6
N(1)-C(1′)-C(2′)-O(2′)	−135.7	−151.3
O(2′)-C(2′)-C(3′)-O(3′)	43.9	44.7
O(2′)-C(2′)-C(3′)-C(4′)	− 79.1	− 73.4
O(1′)-C(1′)-C(2′)-O(2′)	100.0	90.1
O(3′)-C(3′)-C(4′)-C(5′)	−164.0	−154.1
O(3′)-C(3′)-C(4′)-C(5′)	75.7	87.1
C(1′)-C(2′)-C(3′)-O(3′)	156.2	157.0
C(2′)-C(3′)-C(4′)-C(5′)	−160.8	−157.2
C(5′)-C(4′)-O(1′)-C(1′)	159.9	148.0
C(3′)-C(4′)-C(5′)-O(5′)	65.7	49.1
O(1′)-C(4′)-C(5′)-O(5′)	− 51.7	− 67.5
O(1′)-C(1′)-N(1)-C(6)	− 87.1	26.5

Table V. Atomic distances from the plane C(1′)-O(1′)-C(4′) in *syn* and *anti* 4-thiouridine.

	syn	*anti*
C(1′)	0.0	0.0
C(2′)	0.286	− 0.050
O(2′)	− 0.969	− 1.435
C(3′)	0.819	0.591
O(3′)	0.676	0.155
C(4′)	0.0	0.0
C(5′)	0.491	0.763
O(1′)	0.0	0.0

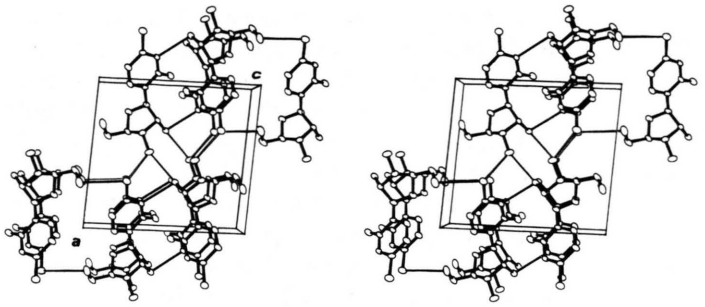

Fig. 2. A stereoscopic view of the crystal structure of *anti*-s^{4}U viewed along b. Hydrogen bonds are indicated by open lines.

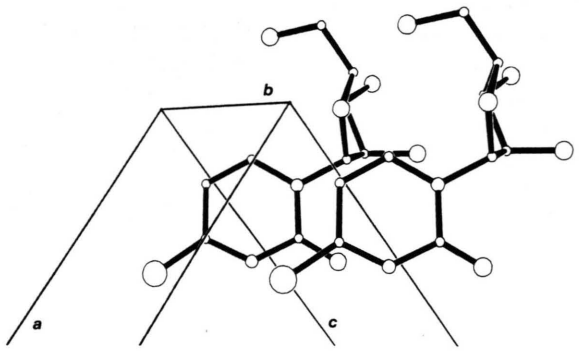

Fig. 3. Overlap projection of the s^{4}U heterocycles and their orientation with respect to the unit cell.

hydrogen bonds from O(3′) and is located close to and over N(3) of the neighbouring heterocycle, as found in the crystal structure of arabino-4-thiouridine [27]. The O(2) oxygen, however, does not accept a hydrogen bond. As indicated in the stacking diagram (Fig. 3), base-base contacts only involve bonds C(2)-O(2) and C(4)-S(4) but heterocycles practically do not overlap. A futher intermolecular interaction is found in the close contact of 2.999 (9) Å between (partially negatively charged) O(2) with (positively charged) C(2) of an adjacent base, and this O(2) atom is not involved in hydrogen bonding. Short intermolecular contacts are summarized in Table VI.

Table VI. Interatomic distances in hydrogen bonds for 4-thiouridine (*anti*).

N(3)-O(2′)	2.855
S(4)-O(3′)	3.263
S(4)-O(5′)	3.289
O(2′)-O(3′)	2.835

Conclusions

The two different crystal structures of *syn* and *anti* s^{4}U, one containing water of hydration and the other not, show that packing forces can substantially influence the three-dimensional structure of a nucleoside. This should be kept in mind if crystallographic data of a flexible molecule are interpreted in structural terms. In the nucleoside series, however, where a vast body of data, both from crystallographic and from spectroscopic studies is available and preferred conformational ranges are known, such rare conformations as *syn*-s^{4}U demonstrate the flexibility of a molecule and help to recognize extreme cases and conformational transitions. If the correlation of sugar pucker and angle distortion with *syn-anti* interchange are concerned, a more complete picture of the structural properties of s^{4}U and other nucleosides can be drawn.

Up to the present time, there are no quantitative theoretical data relating energetical characteristics of crystal packing forces with intrinsic properties of bigger biological molecules. The promising theoretical investigations of amides, carboxylic acids and other simpler model compounds [28] can be extended in future on nucleosides. A consistent application of Force Field Methods and Monte Carlo techniques to isolated molecules, and molecules influenced by crystal forces as well as free in solution, is the urgent task.

s^{4}U constitutes a good experimental basis for further theoretical investigations.

[1] M. N. Lipsett, J. Biol. Chem. **240,** 3975–3978 (1965).
[2] M. N. Lipsett and B. P. Doctor, J. Biol. Chem. **242,** 4072–4077 (1967).
[3] S. Cory, K. A. Marcker, S. K. Dube, and B. F. C. Clark, Nature **220,** 1039–1040 (1968).
[4] S. K. Dube, K. A. Marcker, B. F. C. Clark, and S. Cory, Nature **218,** 232–233 (1968)
[5] B. P. Doctor, B. J. Wayman, S. Cory, P. S. Rudland, and B. F. C. Clark, Eur. J. Biochem. **8,** 93–100 (1969).
[6] A. Favre, A. M. Michelson, and M. J. Yaniv, J. Mol. Biol. **58,** 367–379 (1971).
[7] M. J. Yaniv, A. Chestier, F. Gros, and A. Favre, J. Mol. Biol. **58,** 381–388 (1971).
[8] L. H. Schulman, J. Mol. Biol. **58,** 117–131 (1971).
[9] O. Bergstrom and N. Leonard, J. Am. Chem. Soc. **94,** 6178–6182 (1972).
[10] R. T. Walker and U. L. RajBhandary, J. Biol. Chem. **247,** 4879–4892 (1972).
[11] K. H. Scheit and E. Gaertner, Biochim. Biophys. Acta **182,** 10–16 (1969).
[12] M. Geller, A. Pohorille, and A. Jaworski, Biochim. Biophys. Acta **331,** 1–8 (1973).
[13] W. Saenger and K. H. Scheit, J. Mol. Biol. **50,** 153–169 (1970).
[14] F. E. Hruska, K. K. Ogilvie, A. A. Smith, and H. Wayborn, Canad. J. Chem. **49,** 2449–2452 (1971).
[15] B. J. Blackburn, A. A. Grey, I. C. P. Smith, and F. E. Hruska, Canad. J. Chem. **48,** 2866–2870 (1970).
[16] T. R. Emerson, R. J. Swan, and T. L. V. Ulbricht, Biochem. **6,** 843–850 (1967).
[17] S. I. Chan and J. H. Nelson, J. Am. Chem. Soc. **91,** 168–183 (1969).
[18] P. O. P. Ts'o, N. S. Kondo, M. P. Schweizer, and D. P. Hollis, Biochem. **8,** 997–1029 (1969).
[19] K. H. Scheit and W. Saenger, FEBS Lett. **2,** 305–308 (1969).
[20] D. Suck and W. Saenger, J. Amer. Chem. Soc. **94,** 6520–6526 (1972).
[21] W. Saenger, G. Ritzmann, and W. Pfleiderer, Acta Crystallogr. **B 33,** 2989–2993 (1977).
[22] C. L. Coulter, J. Amer. Chem. Soc. **95,** 570–575 (1973).
[23] D. Suck, W. Saenger, P. Main, G. Germain, and J. P. Declercq, Biochim. Biophys. Acta **361,** 257–265 (1974).
[24] G. Germain, P. Main, and M. M. Woolfson, Acta Cryst. **B 26,** 274–285 (1970).
[25] C. Altona and M. Sundaralingam, J. Amer. Chem. Soc. **94,** 8205–8212 (1972).
[26] E. A. Green, R. D. Rosenstein, R. Shiono, D. J. Abraham, B. L. Trus, and R. E. Marsh, Acta Cryst. **B 31,** 102–107 (1975).
[27] W. Saenger and D. Suck, Eur. J. Biochem. **32,** 473–478 (1973).
[28] P. Dauber and A. T. Hagler, Acc. Chem. Res. **13,** 105–112 (1980).