

Crystal Structure and Conformation of 2-Thio-5-methylaminomethyluridine Dihydrate

BY ERNST EGERT AND HANS JÖRG LINDNER

Fachgebiet Organische Chemie I, Technische Hochschule Darmstadt, Petersenstrasse 22, D-6100 Darmstadt, Federal Republic of Germany

WOLFGANG HILLEN AND HANS GÜNTER GASSEN

Fachgebiet Biochemie, Technische Hochschule Darmstadt, Petersenstrasse 22, D-6100 Darmstadt, Federal Republic of Germany

AND HELMUT VORBRÜGGEN

Schering AG, Müllerstrasse 170–178, D-1000 Berlin 65, Federal Republic of Germany

(Received 29 July 1978; accepted 6 September 1978)

Abstract

$C_{11}H_{21}N_3O_7S$ is orthorhombic, space group $P2_12_12_1$, with $a = 13.695$ (7), $b = 11.364$ (6), $c = 9.748$ (5) Å, $Z = 4$. The structure, which was refined to $R = 0.063$ for 915 counter reflections, exhibits a zwitterionic base with N(3) deprotonated and N(51) protonated. The orientation of the base is *anti* ($\chi = 15^\circ$); the ribosyl moiety shows a C(3')-*endo* conformation and a *gauche-gauche* arrangement of C(4')–C(5'). These conformational features are stabilized by the characteristic intramolecular interaction C(6)–H...O(5'). One water molecule forms an eight-membered ring with the aid of two hydrogen bonds to O(4) and the amino group. The nucleosides exhibit no base stacking. The packing is mainly determined by the interaction with the water molecules leading to a great variety of hydrogen bonds and to water channels parallel to c .

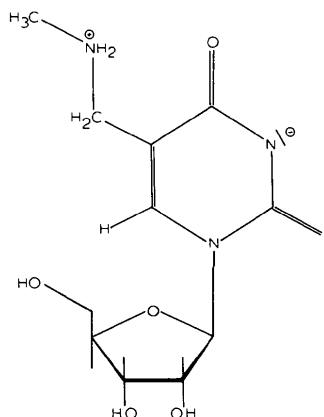
Uridine is seldom found in the wobble position of tRNA's but is replaced by 5-substituted derivatives which may, furthermore, contain S instead of the 2-keto O atom. The structure of Phe-tRNA_{yeast} (Kim *et al.*, 1974; Ladner *et al.*, 1975) shows the wobble nucleoside in a sterically exposed position. The base-pairing properties between codon and anticodon are certainly influenced by the exact conformation of the anticodon region which depends on the interaction of the wobble nucleoside with its neighbours and the anticodon loop backbone.

In our attempt to clarify the function of 5-substitution, we report here the crystal structure of 2-thio-5-methylaminomethyluridine dihydrate ($s'mnm^sU \cdot 2H_2O$). This nucleoside which was first isolated and whose structure was elucidated by Carbon, David & Studier (1968)

occurs in the wobble position of Gln-Lys- and Glu-tRNA_{*E. coli*}, and forms base pairs only with adenine and not with guanine, *i.e.* the wobble recognition is restricted by the 5-substituent.

Experimental

1- β -D-Ribofuranosyl-2-thio-5-methylaminomethyluracil was synthesized by the method of Vorbrüggen & Krolkiewicz (1975). Crystals were grown from oxygen-free water as irregular colourless blocks. Weissenberg photographs showed them to be orthorhombic with the systematic absences $h00$, $0k0$, and $00l$ for odd indices determining the space group as $P2_12_12_1$.



Crystal data are summarized in Table 1. Intensities were collected on a Stoe two-circle diffractometer (Cu $K\alpha$ radiation) equipped with a graphite monochromator; the crystal was orientated along c . 1010 symmetry-independent reflections $hk0$ to $hk7$ with $\theta \leq$

Table 1. *Crystal data*

| | |
|--|--|
| Molecular formula: $C_{11}H_{21}N_3O_7S$ | $V = 1517.1 \text{ \AA}^3$ |
| $M_r = 339$ | $Z = 4$ |
| Space group: $P2_12_12_1$ | $\rho_o = 1.49 \text{ Mg m}^{-3}$ (floatation) |
| $a = 13.695 (7) \text{ \AA}$ | $\rho_c = 1.484$ |
| $b = 11.364 (6)$ | $\mu = 2.14 \text{ mm}^{-1}$ |
| $c = 9.748 (5)$ | |

60° were measured in the θ - 2θ scan mode. The data were corrected for background and for Lorentz and polarization factors, but not for absorption.

Structure determination and refinement

The structure was solved with *SHELX 76* (Sheldrick, 1976) which uses scattering factors from *International Tables for X-ray Crystallography* (1974).

Table 2. *Positional parameters and their e.s.d.'s*

| | x | y | z |
|----------|------------|-------------|------------|
| N(1) | 0.3572 (6) | 0.0030 (8) | 0.6594 (8) |
| C(2) | 0.3653 (7) | -0.1144 (9) | 0.6112 (9) |
| S(2) | 0.3284 (2) | -0.2244 (2) | 0.7201 (3) |
| N(3) | 0.4035 (6) | -0.1358 (7) | 0.4921 (7) |
| C(4) | 0.4285 (7) | -0.0498 (9) | 0.3994 (9) |
| O(4) | 0.4647 (5) | -0.0756 (6) | 0.2829 (7) |
| C(5) | 0.4179 (8) | 0.0721 (9) | 0.4472 (9) |
| C(51) | 0.4371 (8) | 0.1736 (9) | 0.3517 (9) |
| N(51) | 0.5428 (6) | 0.1922 (7) | 0.3183 (8) |
| C(52) | 0.6091 (9) | 0.1969 (9) | 0.4359 (9) |
| C(6) | 0.3806 (7) | 0.0903 (8) | 0.5710 (9) |
| C(1') | 0.3190 (7) | 0.0260 (9) | 0.7951 (9) |
| C(2') | 0.3995 (7) | 0.0140 (9) | 0.9124 (9) |
| C(3') | 0.4397 (8) | 0.1393 (9) | 0.9173 (9) |
| C(4') | 0.3470 (8) | 0.2130 (9) | 0.9044 (9) |
| O(1') | 0.2876 (5) | 0.1463 (6) | 0.8022 (6) |
| O(2') | 0.3537 (5) | -0.0158 (6) | 1.0349 (6) |
| O(3') | 0.4945 (5) | 0.1586 (6) | 1.0302 (7) |
| C(5') | 0.3543 (9) | 0.3364 (9) | 0.8475 (9) |
| O(5') | 0.4037 (6) | 0.3439 (7) | 0.7266 (7) |
| O(A) | 0.8097 (6) | -0.0192 (7) | 0.4229 (7) |
| O(B) | 0.6537 (5) | -0.0210 (7) | 0.2102 (7) |
| H(C51) | 0.405 (6) | 0.163 (7) | 0.228 (8) |
| H'(C51) | 0.412 (6) | 0.255 (7) | 0.399 (6) |
| H(N51) | 0.558 (5) | 0.140 (7) | 0.220 (7) |
| H'(N51) | 0.544 (5) | 0.256 (7) | 0.268 (6) |
| H(C52) | 0.590 (7) | 0.261 (9) | 0.503 (8) |
| H'(C52) | 0.617 (6) | 0.133 (7) | 0.488 (8) |
| H''(C52) | 0.683 (6) | 0.228 (8) | 0.381 (8) |
| H(C6) | 0.379 (6) | 0.171 (6) | 0.598 (6) |
| H(C1') | 0.265 (5) | -0.034 (6) | 0.807 (6) |
| H(C2') | 0.443 (5) | -0.032 (7) | 0.866 (6) |
| H(C3') | 0.473 (5) | 0.149 (6) | 0.794 (7) |
| H(C4') | 0.302 (5) | 0.208 (7) | 0.995 (7) |
| H(O2') | 0.398 (6) | -0.021 (7) | 1.153 (7) |
| H(O3') | 0.517 (6) | 0.205 (7) | 1.042 (7) |
| H(C5') | 0.395 (6) | 0.385 (7) | 0.937 (7) |
| H'(C5') | 0.297 (6) | 0.401 (7) | 0.811 (8) |
| H(O5') | 0.392 (7) | 0.381 (7) | 0.640 (7) |
| H(OA) | 0.824 (6) | 0.007 (7) | 0.518 (7) |
| H'(OA) | 0.820 (6) | -0.088 (8) | 0.394 (8) |
| H(OB) | 0.690 (6) | -0.040 (8) | 0.301 (8) |
| H'(OB) | 0.591 (6) | -0.037 (8) | 0.232 (8) |

Isotropic refinement with unit weights for 915 reflections having $|F| > 3\sigma_F$ reduced R to 0.099, which dropped to 0.085 with inclusion of anisotropic temperature factors. At this stage only eight of the 21 H atoms could be located by a difference synthesis. After weights $w = 1/\sigma(|F|)^2$ were introduced, further refinement with fixed H temperature factors yielded the positions of the thirteen remaining H atoms and finally converged at an R of 0.063. None of the positional parameters of the heavy atoms shifted more than 0.2σ in the last cycle. Positional parameters are listed in Table 2.*

Results and discussion

Structure of the base

The base exists as a zwitterion with N(3) deprotonated and the amino N(51) protonated. Apart from the difference syntheses there are further indications strongly supporting this fact: (i) Compared with 2-

* Lists of structure factors and anisotropic thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 33892 (8 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

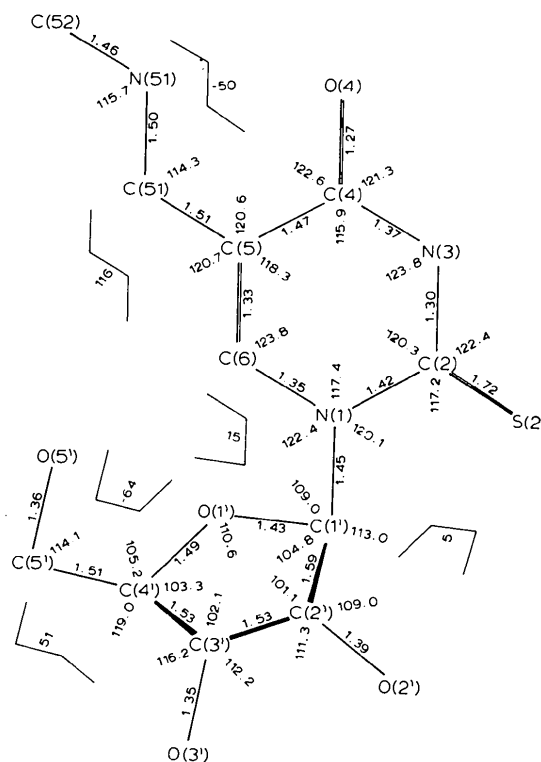


Fig. 1. The bond lengths (\AA), the bond and dihedral angles ($^\circ$) of 2-thio-5-methylaminomethyluridine ($\sigma_{xx} = 0.01 \text{ \AA}$, $\sigma_{xxx} = 0.8^\circ$, $x = C, O, N, S$).

thiouridine (Hawkinson, 1977), C(2)—N(3) and N(3)—C(4) are shorter by 0.06 and 0.02 Å (Fig. 1) whereas the adjacent N(1)—C(2), C(2)—S(2), C(4)—O(4), and C(4)—C(5) are longer (0.05, 0.04, 0.05, and 0.03 Å). These bond-length differences are the result of the delocalization of the negative charge within the π system. (ii) N(3) and O(4) exhibit an unusual hydrogen-bonding scheme (Table 3). N(3) acts as an acceptor in the direction of the expected N(3)—H bond whereas O(4) participates in three hydrogen bonds and has a tetrahedral environment. Furthermore, S(2) forms an unusually strong hydrogen bond due to its partly negative charge.

The pyrimidine ring is slightly non-planar in a shallow chair conformation with a standard deviation of the ring atoms of $\sigma = 0.02$ Å. C(1') and O(4) lie in the plane whereas S(2) (0.08 Å) and C(51) (0.11 Å) show perceptible deviations. C(51)—N(51) is approximately perpendicular to the base plane (Fig. 1), and C(52) is rotated around C(51)—N(51) to form a donating hydrogen bond N(51)—H...O(B) with a water molecule which, for its part, forms an eight-membered ring by the aid of a second hydrogen bond O(B)—H'...O(4) with the same nucleoside molecule (Fig. 2). The substituents of C(51)—N(51)—C(52) show a nearly perfect staggered arrangement with N(51)—C(52) *gauche* to C(5)—C(51).

Conformation of the ribose

The ribose has a C(3')-*endo* conformation (Fig. 2) with C(1'), C(2'), C(4'), and O(1') in the plane ($\sigma = 0.02$ Å) while C(3') deviates by 0.64 Å. The orientation at the glycosidic bond N(1)—C(1') is *anti* with a dihedral angle C(6)—N(1)—C(1')—O(1') of $\chi = 15^\circ$ which is in the range between 0 and 20° typical for 2-thiopyrimidines (Lin & Sundaralingam, 1971; Lin, Sundaralingam & Arora, 1971; Kojić-Prodić, Kvič & Ružić-Toroš, 1976; Kojić-Prodić, Liminga, Šljukić &

Ružić-Toroš, 1974; Hawkinson, 1977) and fits in the common relation between the puckering of the ribose and the orientation of the base (Sundaralingam, 1969; Egert, Lindner, Hillen & Gassen, 1977).

C(1')—O(1') is shorter than C(4')—O(1'), as is normally found in nucleoside structures, but these two lengths are longer than expected. Furthermore, N(1)—C(1') (1.45 Å) is unusually short for a C(3')-*endo* uridine. These deviations might be the result of the negatively charged π system. C(5')—O(5') is *gauche* to C(3')—C(4') and to O(1')—C(4') with dihedral angles of 51 and -64° (Fig. 1). This arrangement allows a fairly strong intramolecular interaction C(6)—H...O(5') (Table 3) often found in nucleosides (Saenger, 1973; Egert, Lindner, Hillen & Gassen, 1978).

Packing of the molecules

As there is no short axis the usual base stacking by translation is not possible in this case. A view down *a* (Fig. 3) illustrates the crystal structure. The packing differs from that of other nucleosides, because there is no separation of base and ribose regions. Although there are many hydrogen bonds in which all possible donor groups participate (Table 3), there is no interaction between the ribosyl moieties. The molecules are connected along *c* (vertical in Fig. 3) by O(2')—H...O(4) and N(51)—H...O(3') and along *b* (horizontal) by N(51)—H'...O(4) and O(3')—H...N(3).

The two water molecules which are arranged in a tetrahedral alignment by two donating and two

Table 3. List of intermolecular hydrogen bonds including the characteristic intramolecular C(6)—H...O(5') interaction

| X—H...Y | X—H (Å) | H...Y (Å) | X...Y (Å) | \angle X—H—Y (°) |
|-----------------|------------|--------------|--------------|-----------------------|
| N(51)—H...O(3') | 1.15 | 2.05 | 2.91 | 129 |
| N(51)—H...O(B) | 1.15 | 2.25 | 3.05 | 124 |
| N(51)—H'...O(4) | 0.87 | 1.99 | 2.82 | 160 |
| O(2')—H...O(4) | 1.30 | 1.68 | 2.94 | 160 |
| O(3')—H...N(3) | 0.63 | 2.13 | 2.73 | 161 |
| O(5')—H...O(A) | 0.96 | 2.03 | 2.78 | 135 |
| O(A)—H...O(B) | 0.99 | 1.91 | 2.88 | 167 |
| O(A)—H'...S(2) | 0.85 | 2.40 | 3.24 | 170 |
| O(B)—H...O(A) | 1.03 | 2.04 | 2.98 | 150 |
| O(B)—H'...O(4) | 0.90 | 1.86 | 2.76 | 177 |
| C(6)—H...O(5') | 0.96 | 2.36 | 3.27 | 161 |

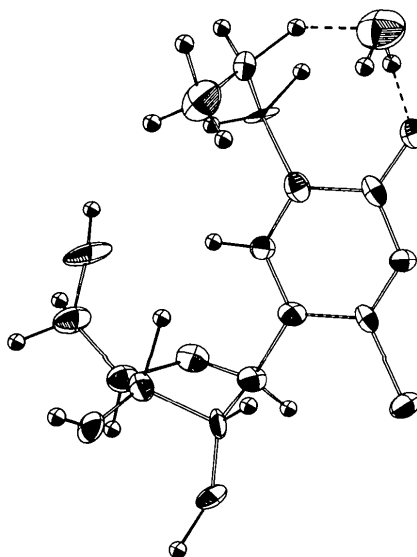


Fig. 2. The structural representation of 2-thio-5-methylaminomethyluridine with the thermal ellipsoids of the heavy atoms including one of the water molecules which forms an eight-membered ring with the aid of two hydrogen bonds (dashed).

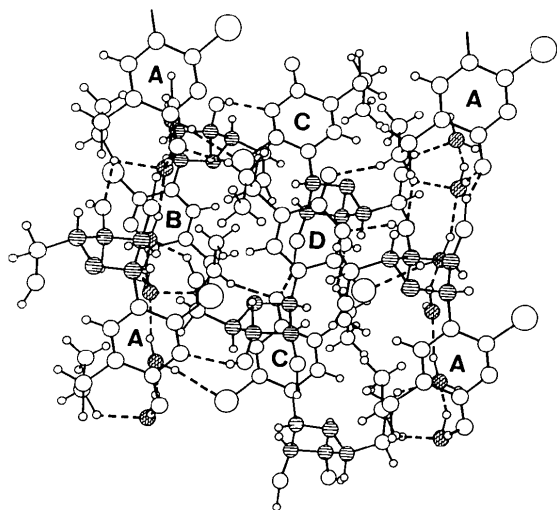


Fig. 3. A view down a with b horizontal and c vertical. The molecules are designated with letters and are interrelated through the following symmetry operations: $A \leftrightarrow B$: screw axis parallel to c ; $A \leftrightarrow C$: screw axis parallel to a ; $A \leftrightarrow D$: screw axis parallel to b . The ribose atoms and the water oxygens are hatched and the hydrogen bonds are dashed.

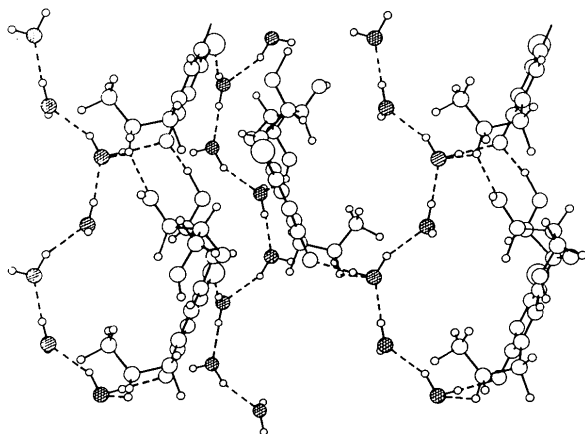


Fig. 4. A view along b with a horizontal and c vertical showing the water channels parallel to c . Only the nucleosides A and B from Fig. 3 are plotted with the water oxygens hatched and the hydrogen bonds dashed.

accepting hydrogen bonds are connected by two hydrogen bonds to each other forming channels of water parallel to c (Fig. 4). The water molecules strengthen the interaction of the nucleoside moieties parallel to b by the hydrogen bonds $O(5')-H \cdots O(A)$

and $O(A)-H \cdots S(2)$. They are very important for the arrangement of the molecules parallel to a , because in this direction the nucleoside moieties are connected only by the water. Fig. 4 also demonstrates the lack of base stacking in the crystal. The intermolecular arrangement is mainly determined by the solvent interaction due to the increased polarity of the base moiety by the 5-substituent. This has to be considered when the structural and functional features of this nucleoside in tRNA's are discussed.

We thank Mrs E. Rönnefeldt and Dr O. E. Beck for their help in preparing the manuscript. This work was supported by the Fonds der Chemischen Industrie and the Deutsche Forschungsgemeinschaft.

References

- CARBON, J., DAVID, H. & STUDIER, M. H. (1968). *Science*, **161**, 1146–1147.
- EGERT, E., LINDNER, H. J., HILLEN, W. & GASSEN, H. G. (1977). *Acta Cryst.* **B33**, 3704–3707.
- EGERT, E., LINDNER, H. J., HILLEN, W. & GASSEN, H. G. (1978). *Acta Cryst.* **B34**, 2204–2208.
- HAWKINSON, S. W. (1977). *Acta Cryst.* **B33**, 80–85.
- International Tables for X-ray Crystallography* (1974). Vol. IV, pp. 99–102. Birmingham: Kynoch Press.
- KIM, S. H., SUDDATH, F. L., QUIGLEY, G. J., MCPHERSON, A., SUSSMAN, S. J., WANG, A. H., SEEMAN, N. C. & RICH, A. (1974). *Science*, **185**, 435–440.
- KOJIĆ-PRODIĆ, B., KVICK, A. & RUŽIĆ-TOROŠ, Ž. (1976). *Acta Cryst.* **B32**, 1090–1095.
- KOJIĆ-PRODIĆ, B., LIMINGA, R., ŠLJUKIĆ, M. & RUŽIĆ-TOROŠ, Ž. (1974). *Acta Cryst.* **B30**, 1550–1555.
- LADNER, J. E., JACK, A., ROBERTUS, J. D., BROWN, R. S., RHODES, D., CLARK, B. F. C. & KLUG, A. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 4414–4418.
- LIN, G. H.-Y. & SUNDARALINGAM, M. (1971). *Acta Cryst.* **B27**, 961–969.
- LIN, G. H.-Y., SUNDARALINGAM, M. & ARORA, S. K. (1971). *J. Am. Chem. Soc.* **93**, 1235–1241.
- SAENGER, W. (1973). *Angew. Chem.* **85**, 680–690.
- SHELDRIK, G. M. (1976). *SHELX 76*. Program for crystal structure determination. Univ. of Cambridge, England.
- SUNDARALINGAM, M. (1969). *Biopolymers*, **7**, 821–860.
- VORBRÜGGEN, H. & KROLIKIEWICZ, K. (1975). *Angew. Chem.* **87**, 251; *Angew. Chem. Int. Ed. Engl.* **14**, 255–256.