Heteroatom Chemistry
Volume 5, Number 4, 1994

# Modified Components of RNA. The Molecular and Crystal Structure of 5-Carboxymethylaminomethyl-2-thiouridine ($s^2cmnm^5U$)

Michał W. Wieczorek* and Grzegorz D. Bujacz

*Institute of Technical Biochemistry, Technical University of Łódź, Stefanowskiego 4/10, 90-924 Łódź, Poland*

Dorota A. Adamiak

*Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznań, Poland*

Andrzej Małkiewicz and Barbara Nawrot

*Institute of Organic Chemistry, Technical University of Łódź, Żwirki 36, 90-924 Łódź, Poland*

*Received 1 March 1994; revised 20 April 1994*

## ABSTRACT

*The structure of the title compound* **1** *has been determined by X-ray analysis. The following crystal data were found: orthorhombic,* $P2_12_12_1$, $a = 5.243(2)$, $b = 24.864(4)$, *and* $c = 37.083(6)$ Å. *Three molecules of the compound with four molecules of water are present in the independent part of the unit cell. The sugar ring puckering is* $^3E$ *for the molecules A and C and* $^3T_2$ *for the molecule B. The conformation of the hydroxymethylene group is* $g^+$ *and* $g^-$ *for B and A, C, respectively. The torsion angles* $\chi$ *for all molecules are in the range of the anti conformation. In the crystal lattice of* **1** *both the intra-and intermolecular hydrogen bonding systems occur.*

## *INTRODUCTION*

5-Carboxymethylaminomethyl-2-thiouridine **1** ($s^2cmnm^5U$) (Figure 1) has been discovered in several sequences of tRNAs from *B. subtilis* [1], and it is a by-product of the biosynthesis pathway of 5-methylaminomethyl-2-thiouridine ($s^2mnm^5U$) in *E. coli* strains [2]. Both modified nucleosides are located in the first position of tRNAs anticodon ($U^*_{34}$, "wobble position"), and both distinctively influence decoding capacity of the biopolymer [3–5]. The function of $U^*_{34}$ as an identity element, by which tRNAs recognize their cognate aminoacyl tRNA synthetase (RS), has been recently experimentally confirmed [6,7]. Moreover, the status of the post-transcriptional modification of $U_{34}$ in $\text{tRNA}^{\text{Glu}}_{(\text{mnm}^5\text{s}^2\text{U})}$ from *E. coli* and $\text{tRNA}^{\text{Glu}}_{(U^*_{34})}$ from *B. subtilis* seems to be essential also for the mentioned tRNAs recognition/binding with glutamyl tRNA reductase (GluRT), the key step of chlorophyll bio-synthesis via the lately discovered $C_5$ pathway [8].

To correlate structural features of $U^*_{34}$ with its biological function, the conformation and III-D structure dynamics of variously 5-, 2-substituted uridines [9–12], as well as oligoribonucleosides with these units as components [12–14], have been determined by NMR and CD spectroscopy. Recently, striking progress in the cocrystallization of tRNAs with aminoacyl tRNA synthetases has been

---

*To whom correspondence should be addressed.

1042-7163/94/$5.00 + .25

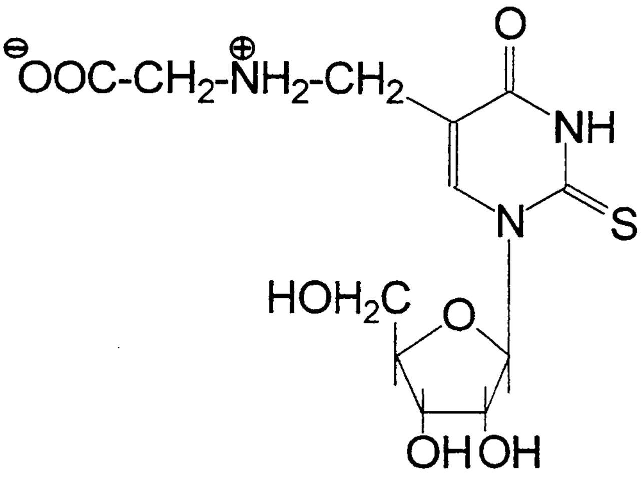

**FIGURE 1** 5-carboxymethylaminomethyl-2-thiouridine (**1**).

**TABLE 1** Crystal Data and Experimental Details

| | |
|---|---|
| Molecular formula | $3C_{12}H_{17}N_3O_7S \cdot 4H_2O$ |
| Formula weight | 347.34 |
| Crystallization solvent | methanol-isopropanol-water |
| Space group | $P2_12_12_1$ |
| $a$ (Å) | 5.243(2) |
| $b$ (Å) | 24.864(4) |
| $c$ (Å) | 37.083(5) |
| $Z$ | 4 |
| $V$ (Å$^3$) | 4834(2) |
| $\mu$ (cm$^{-1}$) | 22.1 |
| $D_c$ (g/cm$^3$) | 1.531 |
| Crystal dimensions (mm) | 0.04, 0.05, 0.3 |
| Maximum $2\theta$ (°) | 130 |
| Radiation, $\lambda$ (Å) | Cu $K_\alpha$, 1.54178 |
| Scan mode | $\omega/2\theta$ |
| Scan width (°) | $0.70 + 0.14 \tan \theta$ |
| $hkl$ ranges: | $h = 0\ 6$ |
| | $k = 0\ 29$ |
| | $l = 0\ 43$ |
| No. of reflections measured: | |
| total | 4789 |
| independent | 3762 |
| $I \geq 2\sigma(I)$ | 2829 |

achieved, and the molecular structures of several such complexes (tRNA-RS) were successfully refined [15].

In this context, comparison of the conformation in the solid state and in the solution of a representative series of wobble uridines should be valuable [21]. In connection with our previous work on the X-ray structure of cmnm$^5$U [16], we present here the molecular and crystal structure of its 2-thioanalog **1**.

## RESULTS AND DISCUSSION

The studied nucleoside crystallizes as a zwitterion. Neither of the oxygen atoms of the carboxyl function is protonated, and the similar lengths of bonds within the carboxyl function indicate delocalization of a negative charge between both oxygen atoms. The nitrogen atom is bound to four atoms and has a positive charge. It is noteworthy that the "decarboxylated" analog of **1**, 5-methylaminomethyl-2-thiouridine (s$^2$mnm$^5$U) [18b], and also 5-aminomethyluridine (nm$^5$U) [19] crystallize as betaines, which are formed by proton transfer from the $N_3H$ imine to the amine function.

Nonhydrogen atom coordinates and equivalent temperature parameters are presented in Table 2. Table 3 includes coordinates of hydrogen atoms. A general view of the refined independent part of the unit cell of compound **1**, with its atom numbering scheme, is shown in Figure 2. The three molecules of **1** with solvent are also presented separately to show intra- and intermolecular hydrogen bonding (Figure 3a–c).

The nucleoside crystallizes in the space group $P2_12_12_1$ with three molecules of the compound and four molecules of water in the asymmetric part of the unit cell. In this context, **1** exhibits a more complex crystal structure than its 2-oxo-analog, which crystallized in space group $P2_1$, with one water molecule in the asymmetric part of the unit cell [16].

The thermal vibrations of the oxygen atoms O1w and O2w are evidently weaker in comparison with analogous data for O3w and O4w, reflecting stronger binding of the first two water molecules with their nucleoside counterparts.

The complex H-bonding system distinctively influences the crystal structure of nucleoside **1**. Table 4 lists selected hydrogen contacts. All hydrogen contacts not greater than 3.00 Å are given in the supplementary material.

Tables 5–7 contain bond lengths, bond angles, and selected torsion angles, respectively. The torsion angles $\chi_{CN}$ (O1′–C1′–N1–C2) are −160(1), −159(1), and −163(1)° for the molecules *A*, *B*, and *C*, respectively and fall in the range of the *anti* conformation. Generally, these values are only slightly different from the respective data obtained for the 2-oxo-analog of **1** [16] and speak for relative freedom of the conformation around the *N*-glycosidic bond [16].

The sugar ring puckering is $^3E$ for the molecules *A* and *C* (for *C*, this is slightly deformed toward $^3T_2$) and $^3T_2$ for the molecule *B*. In terms of pseudorotation [25], the ribofuranose ring conformation is characterized by the parameters $P$ = 19(1), 10(1), 15(1)° and $\tau_m$ = 42(1), 37(1), 42(1) [26] for molecules *A*, *B*, and *C*, respectively. Figures 3a–c present separately the molecules of the asymmetric unit to show the view of the ribofuranose rings with the hydrogen bonding scheme of the respective molecules *A*, *B*, and *C*.

5-Methylaminomethyl-2-thiouridine [18b], as well as methoxy-2-thiouridine [20] and 2-thiouri-

**TABLE 2** Atomic Coordinates ($\times 10^4$) and Equivalent Isotropic Displacement Coefficients ($Å^2 \times 10^3$)

| | *x* | *y* | *z* | *U(eq)* | | *x* | *y* | *z* | *U(eq)* |
|---|---|---|---|---|---|---|---|---|---|
| S2*A* | 4841(9) | 784(1) | 892(1) | 58(1) | S2*B* | 2471(9) | −2267(1) | 1450(1) | 69(1) |
| O1′*A* | 763(15) | −435(3) | 286(2) | 34(2) | O1′*B* | −1277(15) | −3783(3) | 1681(2) | 41(2) |
| O2′*A* | 2047(19) | 648(3) | −69(2) | 49(2) | O2′*B* | 181(19) | −3614(3) | 898(2) | 52(2) |
| O3′*A* | 3986(16) | −97(3) | −552(2) | 40(2) | O3′*B* | 2418(17) | −4597(3) | 1060(2) | 48(2) |
| O5′*A* | −1686(15) | −1271(3) | −103(2) | 36(2) | O5′*B* | 584(23) | −4680(4) | 2056(2) | 78(3) |
| O4*A* | 9631(26) | −776(4) | 1339(3) | 108(4) | O4*B* | 7934(19) | −2788(4) | 2458(2) | 66(3) |
| O14*A* | 12148(17) | −2820(3) | 484(2) | 46(2) | O14*B* | 11014(19) | −4578(4) | 3263(2) | 69(3) |
| O15*A* | 12131(17) | −1932(3) | 432(2) | 49(2) | O15*B* | 10907(17) | −3768(4) | 3021(3) | 60(3) |
| N1*A* | 4259(17) | −228(3) | 669(2) | 27(2) | N1*B* | 2164(19) | −3203(3) | 1808(2) | 34(2) |
| N3*A* | 7234(21) | −99(4) | 1121(2) | 45(2) | N3*B* | 5174(23) | −2594(4) | 2008(2) | 49(3) |
| N11*A* | 7093(19) | −1829(3) | 544(2) | 37(2) | N11*B* | 5819(17) | −3781(4) | 2917(2) | 29(2) |
| C1′*A* | 2316(22) | −11(4) | 403(2) | 26(2) | C1′*B* | 292(26) | −3387(4) | 1529(3) | 42(3) |
| C2′*A* | 3610(24) | 239(4) | 71(2) | 34(3) | C2′*B* | 1712(24) | −3637(4) | 1209(2) | 34(3) |
| C3′*A* | 3538(23) | −232(4) | −180(3) | 30(2) | C3′*B* | 1790(22) | −4233(4) | 1324(3) | 36(3) |
| C4′*A* | 966(21) | −480(4) | −108(3) | 26(2) | C4′*B* | −765(25) | −4300(4) | 1504(3) | 41(3) |
| C5′*A* | 750(25) | −1060(4) | −216(3) | 38(3) | C5′*B* | −1158(30) | −4754(5) | 1767(3) | 60(4) |
| C2*A* | 5389(23) | 134(4) | 897(3) | 36(3) | C2*B* | 3275(27) | −2711(5) | 1763(3) | 49(3) |
| C4*A* | 7977(31) | −623(5) | 1131(3) | 55(4) | C4*B* | 6224(28) | −2927(5) | 2267(3) | 45(3) |
| C5*A* | 6712(22) | −970(4) | 888(3) | 31(2) | C5*B* | 4942(24) | −3450(4) | 2293(3) | 37(3) |
| C6*A* | 4893(22) | −767(4) | 675(3) | 29(2) | C6*B* | 3019(24) | −3562(4) | 2067(2) | 33(3) |
| C10*A* | 7522(28) | −1551(4) | 894(3) | 44(3) | C10*B* | 6091(23) | −3884(5) | 2525(3) | 43(3) |
| C12*A* | 8170(23) | −2374(4) | 539(4) | 44(3) | C12*B* | 7024(25) | −4221(5) | 3135(3) | 42(3) |
| C13*A* | 11083(25) | −2380(4) | 472(3) | 37(3) | C13*B* | 9952(25) | −4185(5) | 3136(3) | 42(3) |
| S2*C* | 12912(9) | −1693(1) | 2705(1) | 65(1) | O1w | 7934(27) | −1613(4) | 2035(3) | 109(4) |
| O1′*C* | 17242(15) | −1686(3) | 3714(2) | 33(2) | O2w | 4620(50) | −5421(6) | 2420(3) | 244(10) |
| O2′*C* | 16200(17) | −626(3) | 3320(2) | 51(2) | O3w | 787(61) | −5419(8) | 2785(7) | 350(16) |
| O3′*C* | 14250(16) | −445(3) | 4014(2) | 41(2) | O4w | 13253(50) | −865(8) | 1876(6) | 293(11) |
| O5′*C* | 19734(16) | −1686(3) | 4396(2) | 40(2) | | | | | |
| O4*C* | 7487(17) | −2948(3) | 3363(2) | 51(2) | | | | | |
| O14*C* | 5888(17) | −3368(3) | 4902(2) | 42(2) | | | | | |
| O15*C* | 5949(14) | −2735(3) | 4483(2) | 42(2) | | | | | |
| N1*C* | 13615(18) | −1995(3) | 3396(2) | 31(2) | | | | | |
| N3*C* | 10337(19) | −2403(4) | 3083(2) | 40(2) | | | | | |
| N11*C* | 10903(18) | −2804(3) | 4347(2) | 33(2) | | | | | |
| C1′*C* | 15605(22) | −1578(4) | 3409(3) | 33(2) | | | | | |
| C2′*C* | 14547(26) | −1018(4) | 3469(3) | 38(3) | | | | | |
| C3′*C* | 14549(21) | −975(4) | 3880(3) | 28(2) | | | | | |
| C4′*C* | 17118(21) | −1235(4) | 3969(3) | 28(2) | | | | | |
| C5′*C* | 17282(24) | −1447(4) | 4346(3) | 35(3) | | | | | |
| C2*C* | 12298(23) | −2049(4) | 3075(3) | 31(3) | | | | | |
| C4*C* | 9520(22) | −2701(5) | 3373(3) | 38(3) | | | | | |
| C5*C* | 11157(21) | −2688(4) | 3687(2) | 24(2) | | | | | |
| C6*C* | 13123(22) | −2332(4) | 3681(3) | 31(2) | | | | | |
| C10*C* | 10511(25) | −3058(4) | 3986(3) | 40(3) | | | | | |
| C12*C* | 9811(23) | −3136(5) | 4647(3) | 46(3) | | | | | |
| C13*C* | 6973(23) | −3075(4) | 4679(3) | 30(2) | | | | | |

Equivalent isotropic *U* is defined as one-third of the trace of the orthogonalized $U_{ij}$ tensor.

**TABLE 3** Hydrogen Atom Coordinates ($\times 10^3$)

| | x | y | z | | x | y | z |
|---|---|---|---|---|---|---|---|
| H2oA | 304(4) | 98(1) | 0(1) | HO1C | 1150(5) | −337(1) | 397(1) |
| H3oA | 263(4) | 2(1) | −66(1) | HO2C | 858(5) | −318(1) | 396(1) |
| H5oA | −190(3) | −164(1) | −18(1) | H21C | 1039(5) | −354(2) | 464(1) |
| H3nA | 810(4) | 15(1) | 129(1) | H22C | 1061(5) | −293(2) | 487(1) |
| H11A | 520(4) | −183(1) | 47(1) | H2oB | −167(4) | −379(1) | 92(1) |
| H12A | 813(4) | −164(1) | 36(1) | H3oB | 96(3) | −464(1) | 90(1) |
| H1′A | 125(4) | 26(1) | 52(1) | H5oB | 122(4) | −495(1) | 220(1) |
| H2′A | 540(5) | 37(1) | 12(1) | H3nB | 571(4) | −225(1) | 199(1) |
| H3′A | 480(5) | −48(1) | −11(1) | H11B | 404(3) | −376(1) | 298(1) |
| H4′A | −47(4) | −23(1) | −18(1) | H12B | 661(3) | −345(1) | 298(1) |
| H5′A | 192(5) | −122(1) | −7(1) | H1′B | −52(5) | −305(1) | 145(1) |
| H5″A | 110(5) | −117(1) | −47(1) | H2′B | 341(5) | −348(1) | 116(1) |
| H6A | 372(4) | −98(1) | 49(1) | H3′B | 322(4) | −426(1) | 152(1) |
| HO1A | 646(6) | −176(1) | 110(1) | H4′B | −217(5) | −435(1) | 129(1) |
| HO2A | 934(6) | −157(1) | 95(1) | H5′B | −283(6) | −478(1) | 186(1) |
| H21A | 755(5) | −258(1) | 77(1) | H5″B | −53(6) | −511(2) | 166(1) |
| H22A | 737(5) | −257(1) | 32(1) | H6B | 201(5) | −392(1) | 208(1) |
| H2oC | 1628(3) | −24(1) | 337(1) | HO1B | 520(5) | −425(1) | 246(1) |
| H3oC | 1481(3) | −22(1) | 389(1) | HO2B | 793(5) | −392(2) | 246(1) |
| H5oC | 2037(3) | −184(1) | 461(1) | H21B | 590(5) | −420(1) | 339(1) |
| H3nC | 946(4) | −243(1) | 284(1) | H22B | 654(5) | −459(1) | 301(1) |
| H11C | 1276(4) | −269(1) | 438(1) | H11w | 879(5) | −147(1) | 181(1) |
| H12C | 974(4) | −250(1) | 434(1) | H12w | 666(5) | −146(1) | 212(1) |
| H1′C | 1662(4) | −158(1) | 318(1) | H21w | 588(10) | −523(2) | 234(1) |
| H2′C | 1268(5) | −103(1) | 338(1) | H22w | 437(10) | −532(1) | 265(1) |
| H3′C | 1319(4) | −120(1) | 398(1) | H31w | −82(12) | −542(2) | 272(2) |
| H4′C | 1865(4) | −94(1) | 393(1) | H32w | 110(12) | −512(2) | 288(2) |
| H5′C | 1614(5) | −177(1) | 434(1) | H41w | 1448(10) | −61(2) | 174(1) |
| H5″C | 1632(5) | −121(1) | 452(1) | H42w | 1431(10) | −118(2) | 197(1) |
| H6C | 1476(4) | −237(1) | 385(1) | | | | |

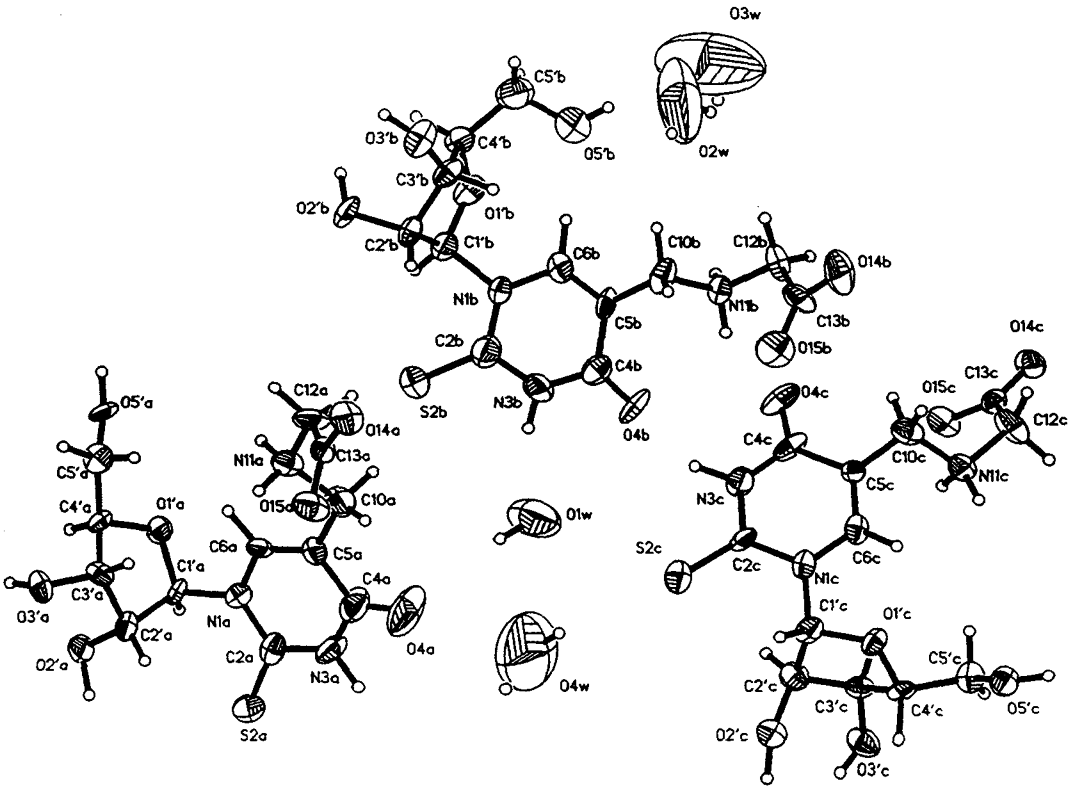

**FIGURE 2** Thermal ellipsoidal view and atom numbering scheme of the asymmetric part of the unit cell of **1**.

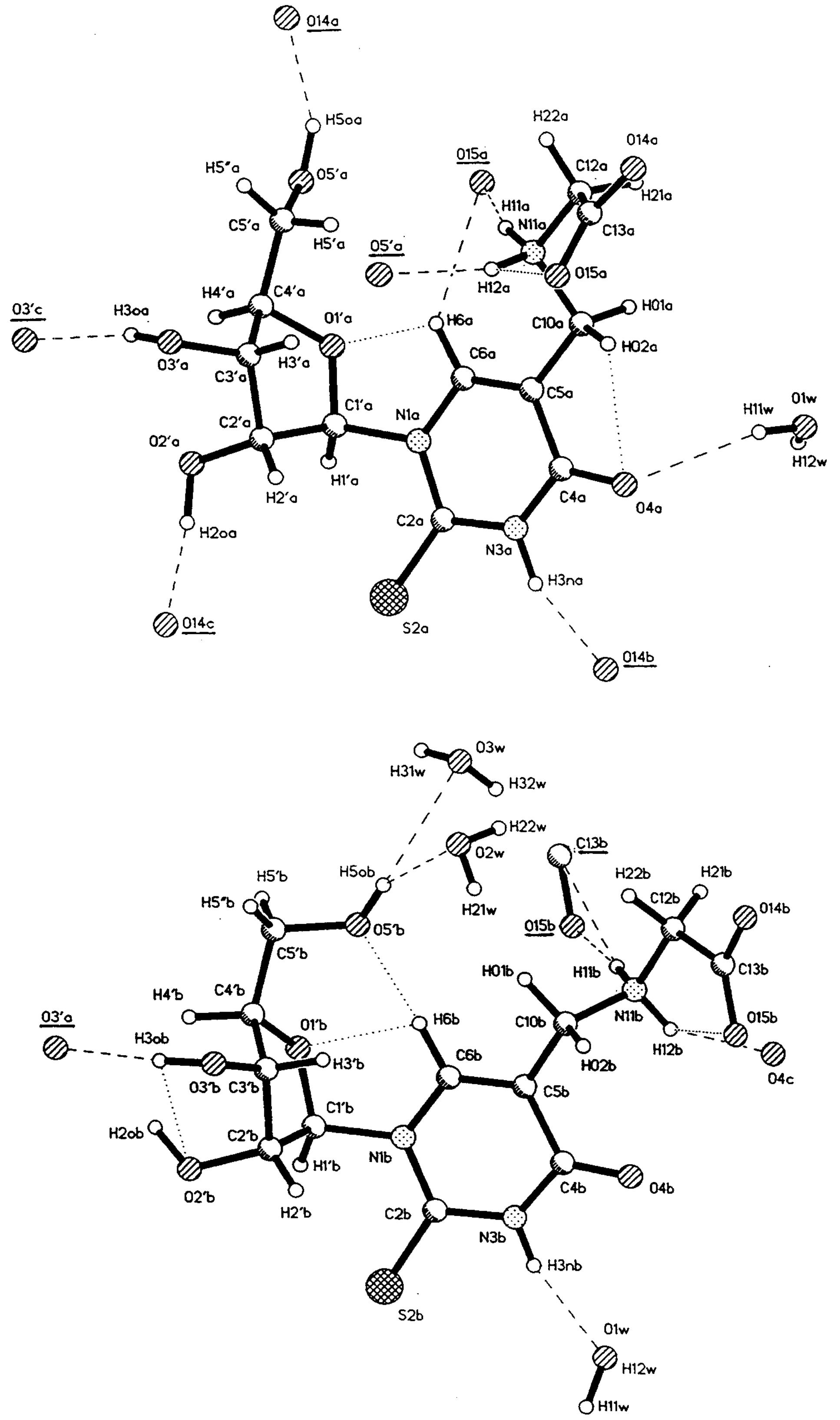

**FIGURE 3** The individual hydrogen bonding schemes of each molecule of the asymmetric unit of **1** with atom numbering; the intramolecular bonds are marked by dotted lines, and the intermolecular ones are denoted by dashed lines. Atoms with underlined names are from symmetrical molecules (Table 4).

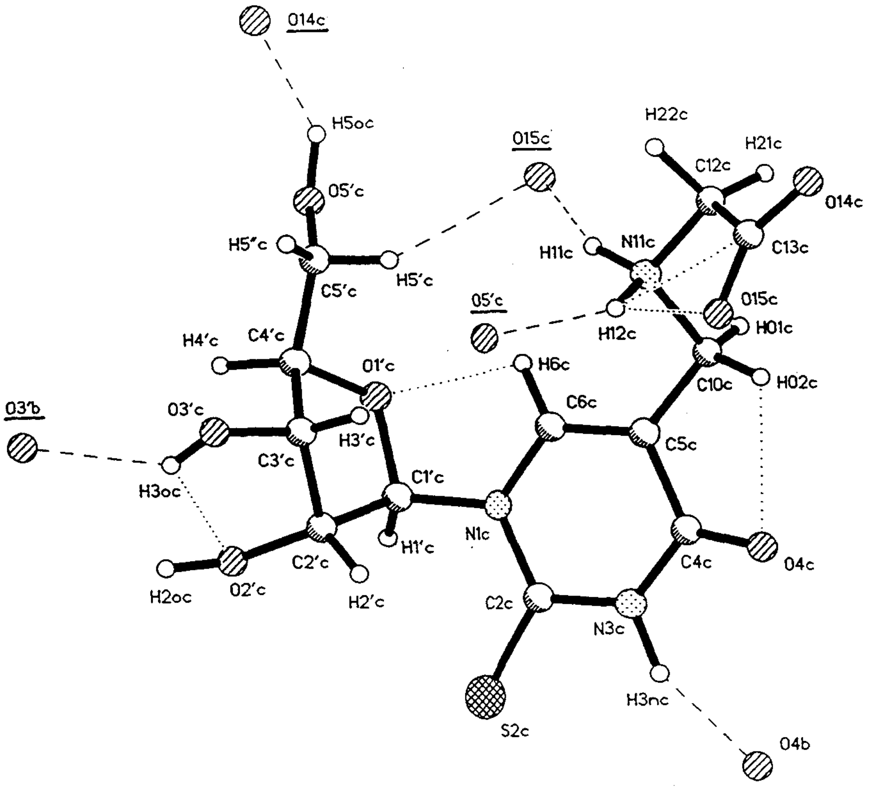

**FIGURE 3** Continued

dine-oxyacetic acid methyl ester [20], the 2-thioanalogs of two wobble position uridines bearing 5-substituents coplanar with the heterobase ring, exhibit also the $^3E$ conformation in the solid state.

One should also notice that all 2-thiouridines hitherto studied exist in solution predominantly as $^3E$ conformers [5,10–12]. The accumulated results strongly suggest that the 2-thio substituent shifts the conformation of uridines toward the biologically important $^3E$ conformation [5,10–12], either in solution or in the solid state. The torsion angles O5′–C5′–C4′–O4′ are 60(1), 61(1), and 65(1)°, and those of O5′–C5′–C4′–C3′ (γ) are 174(1), 61(1), and 179(1)° for molecules *A*, *B*, and *C*, respectively. Thus, the hydroxymethylene group prefers the $g^+$ conformation in molecule *B*, but $g^-$ for *A* and *C* [21]. The torsion angle N11–C10–C5–C6, which describes the spatial relation between the 5-substituent and the heterobase ring, is 27(2) and 38(2)° for molecules *A* and *C*, respectively. The analogous angle for the *B* molecule is equal to −118(1)° and corresponds to that determined for the 2-oxoanalog of **1** [16]. The bond lengths (Table 5) do not differ from those registered for other wobble 2-thiouridines studied so far [18–20].

Three almost flat fragments are present in the molecule of nucleoside **1** (Table 8). The first is formed by the atoms of the heterobase ring (plane 1), the second is formed by four atoms of the sugar moiety (plane 2), and the third is the carboxyl group of the amino acid residue (plane 3). The spatial structure of the nucleoside **1** can be represented by means of the angles between these planes. Least-squares planes, standard deviations of atoms, and dihedral angles between planes are listed in Table 8. Dihedral angles between planes 1 and 2 are 69.6(4) and 70.0(4)°, respectively, for molecules *A* and *C*. The analogous angle for molecule *B* is distinctively higher (81.2(4)°) and indicates an almost perpendicular arrangement of the mentioned fragments.

The "skew" conformation of the 5-substituent is characterized also by three different values of the torsion angle between planes 1 and 3, which are 55.2(5), 39.3(4), and 48.9(4)° for conformers *A*, *B*, and *C*, respectively.

In conclusion, the presence of the three conformers (*A*, *B*, and *C*) in the independent part reflects relative flexibility of the nucleoside **1** structure, although the sugar moiety seems to be more rigid than its other fragments.

The role of post-transcriptional modification of $U_{34}$ can be discussed in terms of an influence of its 2- and 5-substituents on the geometry and dynamics of the tRNA anticodon arm [13], either via interaction of skew 5-substituents with the sugar-

**TABLE 4** Selected Hydrogen Contacts

| *X* | *H* | *Y* | *X–H(Å)* | *H … Y(Å)* | *X … Y(Å)* | *(X–H … Y)(°)* | *Symmetry* |
|---|---|---|---|---|---|---|---|
| O2′*A* | H2o*A* | O14*C* | 1.02(3) | 1.74(2) | 2.75(1) | 168(2) | 1−*X*, 0.5+*Y*, 0.5−*Z* |
| O3′*A* | H3o*A* | O3′*C* | 0.87(3) | 1.88(2) | 2.70(1) | 156(2) | 1.5−*X*, −*Y*, *Z*−0.5 |
| O5′*A* | H5o*A* | O14*A* | 0.95(3) | 1.84(2) | 2.74(1) | 155(2) | *X*−1.5, −0.5−*Y*, −*Z* |
| N3*A* | H3n*A* | O14*B* | 0.98(4) | 1.86(2) | 2.78(1) | 156(2) | 2−*X*, 0.5+*Y*, 0.5−*Z* |
| N11*A* | H11*A* | O15*A* | 1.03(4) | 1.63(3) | 2.65(1) | 167(2) | *X*−1, *Y*, *Z* |
| N11*A* | H12*A* | O5′*A* | 1.00(3) | 1.94(2) | 2.84(1) | 149(2) | 1+*X*, *Y*, *Z* |
| N11*A* | H12*A* | O15*A* | 1.00(3) | 2.24(3) | 2.69(1) | 105(2) | |
| C6*A* | H6*A* | O1′*A* | 1.06(5) | 2.19(3) | 2.73(1) | 109(2) | |
| C6*A* | H6*A* | O15*A* | 1.06(5) | 2.52(3) | 3.36(1) | 136(2) | *X*−1, *Y*, *Z* |
| C10*A* | HO2*A* | O4*A* | 0.98(6) | 2.45(4) | 2.77(2) | 98(2) | |
| O3′*B* | H3o*B* | O3′*A* | 0.97(3) | 1.78(2) | 2.71(1) | 160(2) | *X*−0.5, −0.5−*Y*, −*Z* |
| O3′*B* | H3o*B* | O2′*B* | 0.97(3) | 2.59(2) | 2.78(1) | 91(2) | |
| O5′*B* | H5o*B* | O2w | 0.92(3) | 2.28(6) | 3.11(2) | 149(2) | |
| O5′*B* | H5o*B* | O3w | 0.92(3) | 2.46(6) | 3.27(3) | 146(3) | |
| N3*B* | H3n*B* | O1w | 0.91(4) | 1.97(6) | 2.84(2) | 159(3) | |
| N11*B* | H11*B* | O15*B* | 0.96(3) | 1.65(3) | 2.60(1) | 171(2) | *X*−1, *Y*, *Z* |
| N11*B* | H11*B* | C13*B* | 0.96(3) | 2.47(5) | 3.34(2) | 151(2) | *X*−1, *Y*, *Z* |
| N11*B* | H12*B* | O15*B* | 0.94(4) | 2.39(3) | 2.70(1) | 98(2) | |
| N11*B* | H12*B* | O4*C* | 0.94(4) | 1.97(2) | 2.79(1) | 146(2) | |
| C6*B* | H6*B* | O1′*B* | 1.05(4) | 2.31(3) | 2.72(1) | 102(2) | |
| C6*B* | H6*B* | O5′*B* | 1.05(4) | 2.03(4) | 3.06(2) | 169(3) | |
| O3′*C* | H3o*C* | O3′*B* | 0.78(3) | 2.12(2) | 2.75(1) | 137(2) | 2−*X*, 0.5+*Y*, 0.5−*Z* |
| O3′*C* | H3o*C* | O2′*C* | 0.78(3) | 2.47(2) | 2.81(1) | 108(2) | |
| O5′*C* | H5o*C* | O14*C* | 0.94(3) | 1.91(2) | 2.68(1) | 137(2) | 1.5+*X*, −0.5−*Y*, 1−*Z* |
| N3*C* | H3n*C* | O4*B* | 1.02(3) | 1.86(3) | 2.81(1) | 154(2) | |
| N11*C* | H11*C* | O15*C* | 1.02(3) | 1.72(3) | 2.70(1) | 160(2) | 1+*X*, *Y*, *Z* |
| N11*C* | H12*C* | O5′*C* | 0.98(3) | 2.02(2) | 2.85(1) | 141(2) | *X*−1, *Y*, *Z* |
| N11*C* | H12*C* | O15*C* | 0.98(3) | 2.14(2) | 2.65(1) | 111(2) | |
| C5′*C* | H5′*C* | O15*C* | 1.00(5) | 2.46(2) | 3.32(1) | 143(2) | 1+*X*, *Y*, *Z* |
| C6*C* | H6*C* | O1′*C* | 1.06(5) | 2.20(3) | 2.69(1) | 106(2) | |
| O1w | H11w | O4*A* | 1.00(4) | 2.51(5) | 3.43(2) | 154(3) | |
| O1w | H12w | O4w | 0.83(5) | 2.48(6) | 3.14(3) | 137(3) | *X*−1, *Y*, *Z* |
| O2w | H22w | O3w | 0.88(5) | 1.96(7) | 2.42(4) | 111(4) | |
| O2w | H22w | O4w | 0.88(5) | 2.55(6) | 3.05(3) | 116(4) | 2−*X*, *Y*−0.5, 0.5−*Z* |
| O3w | H31w | O4w | 0.88(9) | 2.26(9) | 2.70(4) | 111(4) | 1−*X*, *Y*−0.5, 0.5−*Z* |
| O3w | H32w | O14*B* | 0.86(8) | 1.94(7) | 2.74(2) | 156(4) | *X*−1, *Y*, *Z* |
| O4w | H42w | O1w | 1.02(8) | 2.20(7) | 3.14(3) | 152(3) | 1+*X*, *Y*, *Z* |

**TABLE 5** Bond Lengths (Å)

| | | *Molecules* | | | | | *Molecules* | | |
|---|---|---|---|---|---|---|---|---|---|
| | | *A* | *B* | *C* | | | *A* | *B* | *C* |
| S2 | C2 | 1.64(1) | 1.66(1) | 1.66(1) | N3 | C2 | 1.40(2) | 1.38(2) | 1.35(2) |
| O1′ | C1′ | 1.40(1) | 1.40(1) | 1.45(1) | N3 | C4 | 1.36(2) | 1.38(2) | 1.37(1) |
| O1′ | C4′ | 1.47(1) | 1.47(1) | 1.47(1) | N11 | C10 | 1.49(1) | 1.48(1) | 1.49(1) |
| O2′ | C2′ | 1.41(1) | 1.41(1) | 1.42(1) | N11 | C12 | 1.47(1) | 1.50(2) | 1.50(1) |
| O3′ | C3′ | 1.44(1) | 1.37(1) | 1.42(1) | C1′ | C2′ | 1.54(1) | 1.53(2) | 1.52(2) |
| O5′ | C5′ | 1.44(2) | 1.42(2) | 1.43(2) | C2′ | C3′ | 1.50(1) | 1.54(1) | 1.53(2) |
| O4 | C4 | 1.22(2) | 1.19(2) | 1.23(2) | C3′ | C4′ | 1.51(2) | 1.51(2) | 1.53(2) |
| O14 | C13 | 1.23(1) | 1.22(2) | 1.24(1) | C4′ | C5′ | 1.50(1) | 1.51(2) | 1.50(2) |
| O15 | C13 | 1.25(1) | 1.23(2) | 1.24(1) | C4 | C5 | 1.41(2) | 1.47(2) | 1.45(1) |
| N1 | C1′ | 1.52(1) | 1.50(2) | 1.47(1) | C5 | C6 | 1.34(2) | 1.34(2) | 1.36(2) |
| N1 | C2 | 1.37(1) | 1.36(2) | 1.38(1) | C5 | C10 | 1.51(1) | 1.51(2) | 1.48(1) |
| N1 | C6 | 1.38(1) | 1.39(1) | 1.37(1) | C12 | C13 | 1.55(2) | 1.54(2) | 1.50(2) |

**TABLE 6** Bond Angles (°)

| | | | *Molecules* | | |
|---|---|---|---|---|---|
| | | | *A* | *B* | *C* |
| C1′ | O1′ | C4′ | 108.9(8) | 109.2(8) | 109.6(8) |
| C1′ | N1 | C2 | 117.3(8) | 118.0(9) | 116.8(8) |
| C1′ | N1 | C6 | 121.1(8) | 119.6(8) | 122.5(8) |
| C2 | N1 | C6 | 121.6(9) | 121.6(9) | 120.6(9) |
| C2 | N3 | C4 | 127.6(10) | 128.2(11) | 127.2(10) |
| C10 | N11 | C12 | 112.4(9) | 111.2(9) | 112.3(8) |
| O1′ | C1′ | N1 | 108.9(8) | 108.8(9) | 108.4(8) |
| O1′ | C1′ | C2′ | 108.2(8) | 108.2(9) | 105.9(9) |
| N1 | C1′ | C2′ | 111.6(8) | 109.9(9) | 113.1(9) |
| O2′ | C2′ | C1′ | 109.3(8) | 110.0(9) | 110.6(9) |
| O2′ | C2′ | C3′ | 108.8(9) | 106.3(8) | 109.9(9) |
| C1′ | C2′ | C3′ | 99.8(8) | 100.9(8) | 102.2(9) |
| O3′ | C3′ | C2′ | 114.2(9) | 116.3(9) | 114.5(9) |
| O3′ | C3′ | C4′ | 114.3(9) | 117.2(9) | 114.5(9) |
| C2′ | C3′ | C4′ | 103.5(9) | 101.8(9) | 100.7(9) |
| O1′ | C4′ | C3′ | 102.1(8) | 105.3(9) | 102.9(8) |
| O1′ | C4′ | C5′ | 109.5(9) | 110.0(9) | 109.3(9) |
| C3′ | C4′ | C5′ | 114.4(9) | 119.5(10) | 113.6(9) |
| O5′ | C5′ | C4′ | 109.8(9) | 107.7(10) | 108.6(9) |
| S2 | C2 | N1 | 124.4(8) | 125.1(9) | 124.2(8) |
| S2 | C2 | N3 | 122.3(8) | 120.3(9) | 120.7(8) |
| N1 | C2 | N3 | 113.2(9) | 114.6(10) | 115.1(9) |
| O4 | C4 | N3 | 121.3(12) | 122.5(12) | 121.0(10) |
| O4 | C4 | C5 | 123.1(13) | 124.2(12) | 123.3(10) |
| N3 | C4 | C5 | 115.7(11) | 113.2(11) | 115.6(10) |
| C4 | C5 | C6 | 118.8(11) | 119.2(10) | 116.8(9) |
| C4 | C5 | C10 | 116.4(10) | 119.3(10) | 117.0(9) |
| C6 | C5 | C10 | 124.9(10) | 120.6(10) | 126.2(9) |
| N1 | C6 | C5 | 123.2(10) | 122.9(10) | 123.5(10) |
| N11 | C10 | C5 | 113.0(9) | 113.4(9) | 112.1(9) |
| N11 | C12 | C13 | 113.0(10) | 112.3(10) | 112.5(9) |
| O14 | C13 | O15 | 126.7(11) | 128.7(12) | 126.4(10) |
| O14 | C13 | C12 | 116.8(10) | 114.2(11) | 116.6(10) |
| O15 | C13 | C12 | 116.4(10) | 117.1(11) | 116.9(10) |

phosphate backbone [18] or the contribution of the 2-thio group to the restriction of the local conformation ($^3E$) [5,10–12] and enhancement of the stacking forces [10–12].

In this context, the presented results of the studies on the conformation and crystal structure of nucleoside **1** do not contradict these statements.

## EXPERIMENTAL

5-Carboxymethylaminomethyl-2-thiouridine **1** was synthesized according to the previously reported procedure [17]. The crystallization of sample **1** was achieved by a slow evaporation of the more volatile solvent from an isopropanol-methanol and water mixture (pH 5). Although this procedure was found as optimal, the nucleoside crystallized in the form of thin needles, and it somewhat conditioned the process of the data collection and methodology of the structure refinement.

The crystal data and experimental details are

**TABLE 7** Torsional Angles Including Hydrogen Atoms of Hydroxyl Groups (°)

| | | | | *Molecules* | | |
|---|---|---|---|---|---|---|
| | | | | *A* | *B* | *C* |
| S2 | C2 | N1 | C1′ | −2(1) | 8(2) | 3(1) |
| S2 | C2 | N1 | C6 | 177(1) | 177(1) | −175(1) |
| S2 | C2 | N3 | C4 | −175(1) | −174(1) | −176(1) |
| O1′ | C1′ | N1 | C2 | −160(1) | −159(1) | −163(1) |
| O1′ | C1′ | N1 | C6 | 22(1) | 32(1) | 15(1) |
| O1′ | C1′ | C2′ | O2′ | 89(1) | 85(1) | 88(1) |
| O1′ | C1′ | C2′ | C3′ | −25(1) | −27(1) | −29(1) |
| O1′ | C4′ | C3′ | O3′ | −165(1) | −163(1) | −163(1) |
| O1′ | C4′ | C3′ | C2′ | −40(1) | −35(1) | −40(1) |
| O1′ | C4′ | C5′ | O5′ | 60(1) | −61(1) | 65(1) |
| O2′ | C2′ | C1′ | N1 | −151(1) | −157(1) | −153(1) |
| O2′ | C2′ | C3′ | O3′ | 50(1) | 51(1) | 48(1) |
| O2′ | C2′ | C3′ | C4′ | −75(1) | −78(1) | −76(1) |
| O3′ | C3′ | C2′ | C1′ | 164(1) | 165(1) | 165(1) |
| O3′ | C3′ | C4′ | C5′ | 77(1) | 73(1) | 79(1) |
| O5′ | C5′ | C4′ | C3′ | 174(1) | 61(1) | 179(1) |
| O4 | C4 | N3 | C2 | 180(1) | 177(1) | 169(1) |
| O4 | C4 | C5 | C6 | 179(1) | 179(1) | −170(1) |
| O4 | C4 | C5 | C10 | 0(2) | −12(2) | 10(2) |
| O14 | C13 | C12 | N11 | 176(1) | −169(1) | 174(1) |
| O15 | C13 | C12 | N11 | 0(2) | 12(2) | −6(1) |
| N1 | C1′ | O1′ | C4′ | −121(1) | −113(1) | −118(1) |
| N1 | C1′ | C2′ | C3′ | 95(1) | 91(1) | 90(1) |
| N1 | C2 | N3 | C4 | 0(2) | 6(2) | 1(2) |
| N1 | C6 | C5 | C4 | 3(2) | 0(2) | 0(2) |
| N1 | C6 | C5 | C10 | −178(1) | −170(1) | −179(1) |
| N3 | C2 | N1 | C1′ | −177(1) | −173(1) | −174(1) |
| N3 | C2 | N1 | C6 | 1(1) | −4(2) | 8(1) |
| N3 | C4 | C5 | C6 | −1(2) | 2(2) | 8(2) |
| N3 | C4 | C5 | C10 | 180(1) | 172(1) | −172(1) |
| N11 | C10 | C5 | C4 | −153(1) | 73(1) | −141(1) |
| N11 | C10 | C5 | C6 | 27(2) | −118(1) | 38(2) |
| C1′ | O1′ | C4′ | C3′ | 25(1) | 18(1) | 23(1) |
| C1′ | O1′ | C4′ | C5′ | 146(1) | 148(1) | 144(1) |
| C1′ | N1 | C6 | C5 | 176(1) | 170(1) | 174(1) |
| C1′ | C2′ | C3′ | C4′ | 39(1) | 37(1) | 42(1) |
| C2′ | C1′ | O1′ | C4′ | 0(1) | 6(1) | 4(1) |
| C2′ | C1′ | N1 | C2 | 81(1) | 83(1) | 80(1) |
| C2′ | C1′ | N1 | C6 | −98(1) | −86(1) | −102(1) |
| C2′ | C3′ | C4′ | C5′ | −158(1) | −159(1) | −158(1) |
| C2 | N1 | C6 | C5 | −3(2) | 1(2) | −8(2) |
| C2 | N3 | C4 | C5 | −1(2) | −6(2) | −9(2) |
| C5 | C10 | N11 | C12 | 173(1) | −179(1) | 170(1) |
| C10 | N11 | C12 | C13 | −80(1) | 75(1) | −79(1) |
| H2o | O2′ | C2′ | C1′ | 108(2) | −55(2) | −162(2) |
| H2o | O2′ | C2′ | H2′ | −18(2) | 177(2) | 75(3) |
| H2o | O2′ | C2′ | C3′ | −144(2) | 53(2) | −50(3) |
| H3o | O3′ | C3′ | C2′ | −81(2) | −75(2) | −31(3) |
| H3o | O3′ | C3′ | H3′ | 158(2) | 168(2) | −154(2) |
| H3o | O3′ | C3′ | C4′ | 38(2) | 46(2) | 85(2) |
| H5o | O5′ | C5′ | C4′ | 178(2) | −155(2) | 178(2) |
| H5o | O5′ | C5′ | H5′ | −73(3) | 81(3) | −74(3) |
| H5o | O5′ | C5′ | H5″ | 44(3) | −38(3) | 41(3) |

**TABLE 8** Least-Squares Planes[a]

| | | *Deviation (Esd) (Å)* | | |
|---|---|---|---|---|
| *Plane* | *Atom* | *Molecule A* | *Molecule B* | *Molecule C* |
| 1 | N1 | 0.007(8) | −0.004(8) | −0.042(8) |
| | C2 | −0.001(11) | 0.029(12) | 0.038(10) |
| | N3 | −0.005(9) | −0.022(10) | 0.025(9) |
| | C4 | 0.005(13) | 0.019(12) | −0.077(11) |
| | C5 | 0.010(11) | 0.003(11) | 0.027(10) |
| | C6 | −0.018(11) | −0.005(10) | 0.035(10) |
| | M1= | 0.6985(30) | 0.6557(33) | 0.6126(34) |
| | M2= | 0.1810(47) | −0.3747(43) | −0.7125(29) |
| | M3= | −0.6923(30) | −0.6555(33) | −0.3422(38) |
| | D= | −0.2676(193) | −0.6636(486) | 3.6390(593) |
| 2 | C1′ | −0.001(10) | −0.046(12) | −0.023(11) |
| | C2′ | 0.001(11) | 0.020(10) | 0.018(12) |
| | C4′ | 0.000(10) | −0.027(12) | −0.014(10) |
| | O1′ | 0.000(7) | 0.018(7) | 0.011(7) |
| | M1= | −0.7667(37) | −0.7497(36) | 0.7822(36) |
| | M2= | 0.6336(43) | 0.2004(76) | 0.4149(58) |
| | M3= | −0.1037(69) | −0.6308(48) | −0.4648(53) |
| | D= | −1.1024(62) | −5.3321(645) | −0.0815(963) |
| 3 | C12 | −0.011(14) | −0.003(11) | −0.001(11) |
| | C13 | 0.023(11) | 0.011(11) | 0.003(10) |
| | O14 | −0.004(7) | −0.002(7) | −0.001(7) |
| | O15 | −0.004(7) | −0.004(10) | −0.001(7) |
| | M1= | −0.1393(74) | 0.0350(65) | −0.1303(61) |
| | M2= | −0.0867(48) | −0.4067(53) | −0.6952(36) |
| | M3= | −0.9865(11) | −0.9129(23) | −0.7069(35) |
| | D= | −2.0461(536) | −6.2126(866) | −7.4300(886) |
| Dihedral angles between planes (°) | | | | |
| Plane 1/Plane 2 | | 69.6(4) | 81.2(4) | 70.0(4) |
| Plane 2/Plane 3 | | 55.2(5) | 39.3(4) | 48.9(4) |
| Plane 1/Plane 3 | | 81.1(6) | 62.1(5) | 86.5(4) |

[a] Equation of the plane: $M1x + M2y + M3z - D = 0$.

shown in Table 1. The intensity data were collected using a CAD4 diffractometer with graphite monochromatized Cu $K_\alpha$ radiation. The lattice constants were refined by a least-squares fit of 25 reflections in the $\theta$ range 9.9–20.1°. A small decline in intensities of three standard reflections (1, −2, −4; 1,1,−4; 1, 3, 2) was observed: 4.5% during 113.3 hours of exposure time. For that reason, the data file was corrected by use of the DECAY program [27].

The structure was solved by use of the SHELXS-86 program [22]; 2829 reflections with $I \geq 2\sigma(I)$ were used to refine it by full-matrix least-squares using $F$'s (SHELXTL program [28]). All hydrogen atoms were found in a difference Fourier map and set as riding with fixed isotropic thermal parameters (all $U = 0.08$ Å$^2$). Anisotropic thermal parameters were applied for all nonhydrogen atoms. The refinement converged to $R = 0.0755$ and $R_w = 0.0707$ with the weight $w = 1/[\sigma^2(F) + 0.001521F^2]$ for 660 refined parameters. The largest residual peak in the final difference Fourier map was 0.54 eÅ$^{-3}$.

Scattering factors were taken from the International Tables for X-ray Crystallography [23]. Hydrogen bonds, torsion angles, and least-squares planes were calculated by the CSU program [24]. Values of $F_{obs}/F_{calc}$, anisotropic displacement coefficients, and full data including hydrogen atoms are deposited at the Cambridge Crystallographic Data Centre [29].

## ACKNOWLEDGMENTS

This work was partially supported by Grant KBN No. PB 0506/P3/93/05 (for A.J.M.). We thank M. Kluge for his assistance during the crystallization tests.

## REFERENCES

[1] M. Sprinzl, T. Hartmann, F. Meissner, I. Moll. T. Vorderwülbecke, *Nucleic Acids Res., Suppl.*, 1987, r53.

[2] T. G. Hagervall, C. G. Edmonts, J. A. McCloskey, G. R. Björk, *J. Biol. Chem.*, *262*, 1987, 8488.

[3] R. Adamiak, P. Górnicki, *Prog. Nucl. Ac. Res. Mol. Biol.*, *32*, 1985, 27.

[4] K. Murao, H. Ishikura, *Nucleic Acids Res., Spec. Publ.*, *5*, 1978, 333.

[5] S. Yokoyama, T. Watanabe, H. Ishikura, Z. Yamaizumi, S. Nishimura, T. Miyazawa, *Proc. Natl. Acad. Sci. USA*, *82*, 1985, 4905.

[6] L. L. Kisselev, *Prog. Nucl. Ac. Res. Mol. Biol.*, *32*, 1985, 237.

[7] D. Söll, *Experientia*, *46*, 1990, 1089.

[8] D. Jahn, E. Verkamp, D. Söll, *T.I.B.S.*, *17*, 1992, 215.

[9] S. Yokoyama, Z. Yamaizumi, S. Nishimura, T. Miyazawa, *Nucleic Acids Res.*, *6*, 1979, 2611.

[10] H. Sierzputowska-Gracz, E. Sochacka, A. Małkiewicz, K. Kuo, Ch. Gehrke, P. F. Agris, *J. Am. Chem. Soc.*, *109*, 1987, 7171.

[11] P. F. Agris, H. Sierzputowska-Gracz, A. Smith, A. Małkiewicz, E. Sochacka, B. Nawrot, *J. Am. Chem. Soc.*, *114*, 1992, 2652.

[12] W. Smith, H. Sierzputowska-Gracz, E. Sochacka, A. Małkiewicz, P. F. Agris, *J. Am. Chem. Soc.*, *114*, 1992, 7989.

[13] V. Dao, D. Guenther, A. Małkiewicz, A. Nawrot, E. Sochacka, A. Kraszewski, K. Everett, P. F. Agris, *Proc. Natl. Acad. Sci.*, *91*, 1994, 2125.

[14] M. Sekine, K. Seio, T. Saton, K. Sakamoto, S. Yokoyama, *Nucleosides Nucleotides*, *12*, 1993, 305.

[15] *15th International tRNA Workshop*, Cap d'Agde, France, May 30–June 4, 1993, in *Programme and Abstracts*, vol. C.II, p. 253 (1993).

[16] Z. Gałdecki, B. Luciak, A. Małkiewicz, B. Nawrot, *Monatsh. Chem.*, *122*, 1991, 487.

[17] A. Małkiewicz, B. Nawrot, E. Sochacka, *Z. Naturforsch.*, *B42*, 1987, 360.

[18] (a) W. Hillen, E. Egert, H. J. Lindner, H. G. Gassen, *Biochemistry*, *17*, 1978, 5314; (b) W. Hillen, E. Egert, H. J. Lindner, H. G. Gassen, *FEBS Lett.*, *94*, 1978, 361.

[19] Z. Gałdecki, D. Adamiak, A. Małkiewicz, E. Sochacka, Technical University of Łódź, 1992, unpublished results.

[20] Z. Gałdecki, D. Adamiak, A. Małkiewicz, B. Nawrot, Technical University of Łódź, 1993, unpublished results.

[21] W. Saenger: *Principles of Nucleic Acid Structure*, Springer-Verlag, Inc., New York, 1984.

[22] G. M. Sheldrick, G. M. Krüger, R. Goddard: SHELXS-86. *Crystallographic Computing 3*, Oxford University Press, 1985, Oxford, United Kingdom, p. 175.

[23] *International Tables for X-ray Crystallography*, Kynoch Press, Birmingham, England, 1974.

[24] I. Vickovic: *Crystal Structure Utility, Programme for the Presentation of Crystal Structure Analysis*, Release May, 1988, University of Zagreb, Zagreb, Yugoslavia.

[25] C. Altona, M. Sundaralingam, *J. Am. Chem. Soc.*, *94*, 1972, 8205.

[26] M. Jaskólski, *Acta Cryst.*, *A40*, 1984, 364.

[27] B. A. Frenz: *SDP—Structure Determination Package*, Enraf-Nonius, Delft, Holland, 1984.

[28] SHELXTL PC, Release 4.1 (May 1990), Siemens Analytical X-Ray Instruments, Inc., Madison, WI.

[29] University Chemical Laboratory, Cambridge Crystallographic Data Centre, Lensfield Road, Cambridge CB2 1EW, United Kingdom.